# Development of statistical mechanics emulators for system analysis and uncertainty quantification in nuclear engineering

by

Abhinav Gairola

M.Sc., École Centrale de Paris, 2013

---

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

Doctor of Philosophy

Department of Mechanical and Nuclear Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2019

# Abstract

Despite the tremendous growth of computational resources in recent years, the "*curse of dimensionality*" associated with the canonical evolutionary partial differential equations in nuclear reactor engineering and thermal-hydraulics such as the direct solution to non-linear Navier-Stokes equations is still a challenge to tackle. To solve this fundamental problem and to get a reasonable answer in the time-constrained engineering design and optimization process, reduced spatio-temporal complexity models are required. Typically, this is done by linearization and spatio-temporal averaging of parameters in the context of thermal-hydraulic applications. This procedure may lead to an inaccurate depiction of the system behavior–with a number of structural errors in the model. These structural errors can also arise because of the principles on which the model relies. Since it is almost impossible to completely decipher the principles on which the model rests, and a certain degree of distortion is always required to build a system of equations. This approximated view of the reality is the basis of many engineering system simulations toolboxes such as RELAP and TRACE–which can lead to a tremendous amount of uncertainties in the quantities of interests. The gap between the reality and approximated system is typically bridged by uncertainty quantification routes which rely on black-box approaches, such as constructing low-cost regression surrogates or emulators and conducting sensitivity studies with those surrogates.

To tackle these uncertainties, this work takes a unique approach of using the high-resolution datasets to learn the dynamics of uncertainties with a statistical mechanics approach. While doing so, it is inherently assumed that the Navier-Stokes equation is a structurally perfect and correct model. The unique contribution in the general context of nuclear engineering is the constraining of non-linear Langevin equation on the high-resolution datasets via the non-equilibrium statistical mechanics route. This is achieved by computing

the parameters of the model via solving an inverse problem through the utilization of the Kramers-Moyal expansion method. These statistical mechanics equations are used to quantify uncertainties in the scalar dispersion. A similar approach is used to emulate grid load pattern and renewable energy generation, and the emulators are integrated with established reactor system models to design a novel stochastic control strategy.

# Development of statistical mechanics emulators for the system analysis and uncertainty quantification in nuclear engineering

by

Abhinav Gairola

M.Sc., École Centrale de Paris, 2013

---

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

Doctor of Philosophy

Department of Mechanical and Nuclear Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2019

Approved by:

Major Professor
Hitesh Bindra

# Copyright

# Abstract

Despite the tremendous growth of computational resources in recent years, the "*curse of dimensionality*" associated with the canonical evolutionary partial differential equations in nuclear reactor engineering and thermal-hydraulics such as the direct solution to non-linear Navier-Stokes equations is still a challenge to tackle. To solve this fundamental problem and to get a reasonable answer in the time-constrained engineering design and optimization process, reduced spatio-temporal complexity models are required. Typically, this is done by linearization and spatio-temporal averaging of parameters in the context of thermal-hydraulic applications. This procedure may lead to an inaccurate depiction of the system behavior–with a number of structural errors in the model. These structural errors can also arise because of the principles on which the model relies. Since it is almost impossible to completely decipher the principles on which the model rests, and a certain degree of distortion is always required to build a system of equations. This approximated view of the reality is the basis of many engineering system simulations toolboxes such as RELAP and TRACE–which can lead to a tremendous amount of uncertainties in the quantities of interests. The gap between the reality and approximated system is typically bridged by uncertainty quantification routes which rely on black-box approaches, such as constructing low-cost regression surrogates or emulators and conducting sensitivity studies with those surrogates.

To tackle these uncertainties, this work takes a unique approach of using the high-resolution datasets to learn the dynamics of uncertainties with a statistical mechanics approach. While doing so, it is inherently assumed that the Navier-Stokes equation is a structurally perfect and correct model. The unique contribution in the general context of nuclear engineering is the constraining of non-linear Langevin equation on the high-resolution datasets via the non-equilibrium statistical mechanics route. This is achieved by computing

the parameters of the model via solving an inverse problem through the utilization of the Kramers-Moyal expansion method. These statistical mechanics equations are used to quantify uncertainties in the scalar dispersion. A similar approach is used to emulate grid load pattern and renewable energy generation, and the emulators are integrated with established reactor system models to design a novel stochastic control strategy.

# Table of Contents

List of Figures . . . . . xii

List of Tables . . . . . xix

Acknowledgements . . . . . xx

1 Introduction . . . . . 1

2 Uncertainity quantification and Sensitivity analysis - Using Best Estimate System tools . . . . . 6

2.1 Motivation . . . . . 7

2.2 An Overview of Safety Analysis . . . . . 9

2.2.1 Methodology . . . . . 10

2.3 Model description . . . . . 11

2.3.1 Fuel Composition and Uncertainties . . . . . 11

2.3.2 RELAP model description . . . . . 11

2.4 Results and Discussion . . . . . 14

2.4.1 Thermal-hydraulic analysis using RELAP model . . . . . 14

2.4.2 Uncertainties in RELAP thermal-hydraulic model . . . . . 16

2.4.3 Effect of uncertainties in multiple parameters . . . . . 18

2.5 Conclusion . . . . . 20

3 Statistical modeling of turbulence . . . . . 22

3.1 Origin of Uncertainties . . . . . 22

3.1.1 State of the art in engineering applications . . . . . 23

3.1.2 Limitations of the system codes . . . 23
3.1.3 Physics based approach . . . 25
3.2 Stochastic flow equation and Fokker-Planck representation . . . 26
3.2.1 From Newton's law to local stochastic model . . . 26
3.2.2 Fokker-Planck equation . . . 27
3.2.3 Itô's Lemma . . . 28
3.2.4 Advection-diffusion equation–a type of Fokker Planck equivation . . . 29
3.3 Kolmogorov's & Taylor's hypothesis-local isotropy, randomness of energy cascade & stationarity . . . 31
3.4 Surrogate modeling . . . 33
3.4.1 Non-linear Langevin equation . . . 33
3.4.2 Parameters of non-linear Langevin equation . . . 34
3.4.3 The Kramers-Moyal expansion . . . 34
3.4.4 Data learning: Estimating moments . . . 37
3.5 Wall-bounded turbulent flow problem: Channel flow . . . 39
3.5.1 Description of domain and flow conditions . . . 39
3.5.2 Near wall structures . . . 39
3.6 Scalar transport with mixed convection: Liquid metal channel flow . . . 45
3.6.1 Results & discussion: Liquid metal channel flow . . . 47
3.6.2 Velocity data and surrogate . . . 47
3.6.3 Temperature data and surrogate . . . 49
3.7 Conclusions . . . 51

4 Uncertainty quantification of the contaminant dispersal in the environment . . . 52
4.1 Motivation . . . 53
4.2 Problem Description . . . 55
4.3 Numerical and analytical methods . . . 59
4.3.1 Solution to Convection-Diffusion equation . . . 61

4.3.2 SDE simulations . . . 61
4.4 Results and Discussion . . . 63
4.4.1 First passage time . . . 66
4.4.2 The definition of first passage time . . . 67
4.4.3 Numerically computing the first passage time . . . 68
4.5 Radionuclide dispersal: quantifying the effect of radioactive decay . . . 69
4.5.1 Colloid facilitated radioactivity transport . . . 69
4.5.2 Decaying radioactive particle–Bernoulli trials . . . 70
4.5.3 SDE driven by colored noise . . . 71
4.5.4 OU-process for colored noise . . . 73
4.5.5 Colored noise generation by generalized Langevin Equation . . . 74
4.6 Conclusions . . . 76

5 Statistical modeling approach to load following nuclear power . . . 78
5.1 Motivation . . . 79
5.1.1 Impact on Nuclear Power . . . 80
5.2 Load following NPPs . . . 81
5.3 Reduced Order model-Nuclear Reactor Dynamics . . . 83
5.3.1 Temperature feedback . . . 85
5.3.2 Energy balance . . . 86
5.3.3 Control System . . . 87
5.4 Estimating power generation from wind energy systems . . . 90
5.4.1 Statistical model of wind speed . . . 90
5.4.2 Load curve for NPP . . . 92
5.5 Numerical Results . . . 94
5.5.1 Integrated Micro-grid - Expected Outcomes . . . 95
5.6 Summary . . . 95

6 Conclusion and future work . . . . . 98

Bibliography . . . . . 101

A Data-Filtering . . . . . 111

A.1 High pass filtering–via differencing . . . . . 111

A.2 Summation or Moving average–Inverse of differencing . . . . . 113

A.3 Amplitude . . . . . 114

B Fluctuation dissipation relation . . . . . 116

C Generalized Langevin equation: Memory kernel computation . . . . . 118

# List of Figures

1.1 Black box approach: Here $f(x)$ is an input to output mapping for a computational tool ($f : x \rightarrow y$). Inputs are being sampled from a prescribed distribution and the computational tool is run multiple times. . . . . 2

1.2 Computation of average and fluctuating part of the signal via the solution of the inverse problem depicted in the picture. . . . . 4

2.1 Single Channel Model. . . . . 13

2.2 Axial Peaking factor profile for B-ring fuel elements with associated uncertainty band. . . . . 13

2.3 Axial distribution of Coolant temperature next to B-ring and F-ring fuel elements. . . . . 15

2.4 Axial distribution of DNBR in B-ring ($\alpha_{pmax}$ is the maximum power peaking factor). . . . . 16

2.5 Effect of uncertainties in axial peaking factors: Differential spread (99% confidence band) of coolant temperature adjacent to B-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet coolant temperature around mean value. . . . . 16

2.6 Effect of uncertainties in axial peaking factors: Differential spread (99% confidence band) of coolant temperature adjacent to F-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet temperature around mean value. . . . . 17

2.7 Effect of uncertainties in form friction factor: Differential spread (99% confidence band) of coolant temperature adjacent to B-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet temperature around mean value. . . . . 18
2.8 Joint sampling of multiple parameters: Example of jointly sampled (In Purple) friction factors and PPFs for any axial node with the uncertainties. . . . . 19
2.9 Effect of uncertainties in multiple parameters: Differential spread (99% confidence band) of coolant temperature adjacent to B-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet temperature around mean value. . . . . 20

3.1 One-dimensional RELAP model can be enhanced by coupling it with a realistic-looking velocity signal generated by a 'Langevin' surrogate model. . . . . 24
3.2 Richardson cascade, $n$ is the generation number. . . . . 32
3.3 Turbulence is random in the inertial range of scales and can be approximated via a Gaussian process. . . . . 32
3.4 Delta function can be approximated by the limiting Gaussian distribution. The derivative of delta function can then be computed. . . . . 35
3.5 A Markov chain represented by the equation 3.22 . . . . . 35
3.6 Joint probability density–a vertical cut in it reveals the conditional probability. 37
3.7 Shifted signal. . . . . 38
3.8 Fitted and extrapolated $D^2$ coefficient. . . . . 38
3.9 Domain description of the channel flow problem (taken from John Hopkins Turbulence Database (https://www.jhu.edu/)). . . . . 40
3.10 Streamwise velocity at $\mathbf{z} = \mathbf{0}$ plane (data taken from John Hopkins Turbulence Database). . . . . 40

3.11 Wall effect: near the wall time series show less randomness and more organized behavior. However, turbulence introduces a meandering (or a zig zag) effect in them. . . . . 41
3.12 Diffusion coefficient ($\mathbf{D^2}$) from center to wall. . . . . 41
3.13 Mean diffusion coefficient vs wall normal distance. . . . . 41
3.14 Pre-multiplied energy spectra vs wavelength of Langevin and DNS singal. Here $\lambda = \frac{\mathbf{2\pi}}{\mathbf{k}}$ where $\mathbf{k}$ is the wave number defined as: $\frac{\mathbf{2\pi f}}{\mathbf{U}}$ and $\mathbf{U}$ is the mean velocity. . . . . 42
3.15 Contour map of premultiplied energy spectra DNS signal (left), Langevin signal (right). Here $\mathbf{y}^+ = \frac{\mathbf{y}}{\delta_\nu}$ where $\delta_\nu$ is the viscous length scale defined as: $\frac{\nu}{\mathbf{U}_\tau}$. 42
3.16 Contour map of premultiplied energy spectra–Langevin signal. Wall normal distance is replaced by the fourth-order fitted polynomial ($\mathbf{f}(\mathbf{y_4^+})$) of the same– which represents mean diffusion coefficient ($\mathbf{D^2}$). . . . . 43
3.17 Convolution of Langevin model with Gaussian kernel–smoothing out the signal output. This kernel is responsible for introducing an artificial viscosity. The length of filter is proportional to the Taylor microscale. . . . . 44
3.18 Taylor microscale computed via osculating parabola method. . . . . 44
3.19 Comparison of the Langevin model with memory effects and the DNS-data. The signal of the convolved Langevin model shows remarkable similarity with the true DNS signal. . . . . 44
3.20 Comparison of the Langevin model with memory effects and the DNS-data. The correlation function of the model are now remarkably similar to the true DNS signal. . . . . 44
3.21 Domain description of the differentially heated liquid metal channel flow problem (data courtesy: computational fluid dynamics group at Paul Scherrer Institute (https://www.psi.ch/en)). . . . . 46

3.22 Streamwise velocity (right) and temperature (left) plot at the $\mathbf{z = 0}$ plane (data courtesy: computational fluid dynamics group at Paul Scherrer Institute). 46
3.23 Fluctuations of probed velocity time series. . . . . 47
3.24 One dimensional spectra of probed velocity time series. . . . . 47
3.25 Quadratic diffusion. . . . . 47
3.26 Linear drift with minor third-order corrections. . . . . 47
3.27 Generated surrogate from the time series captured in DNS data–one series shifted in time from the other. . . . . 48
3.28 PDFs of the velocity signal. In red is the DNS data and in black is the velocity surrogate. The green band around the pdf's represent the pdf of multiple trajectories. . . . . 48
3.29 KL divergence tells how well the probability distribution 'n' approximates the distribution'm.' If the KL divergence ($\mathbf{D_{KL}(m||n)} = \int \mathbf{m(y)}\ \mathbf{ln}\frac{\mathbf{m(y)}}{\mathbf{n(y)}}\mathbf{dy}$, '$\mathbf{m}$' is the target pdf and '$\mathbf{n}$' is the computed pdf) is zero then the computed pdf is exactly similar to the target pdf. The green band around 100 mark signifies no serious gain of information after running 100 trajectories. . . . . 49
3.30 Generated surrogate from the time series captured in the DNS data–one series shifted in time from the other. . . . . 49
3.31 PDFs of the temperature signal. In red is the DNS data and in black is the velocity surrogate. The green band around the pdf's represent the pdf of multiple trajectories. . . . . 50
3.32 Combined spectrum of velocity and temperature data. . . . . 50

4.1 Lagrangian points (shown in blue) situated on the Eulerian trajectories $x(t)$ are perturbed by a random forcing sampled from a 'Gaussian' distribution. . 53
4.2 Transport of contaminants, here the colors represents ■ the top surface, ■ the soil just below it, ■ water table, ■ rocks and minerals, while the arrows in ■ represents the torturous path of the released particles. . . . . 55

4.3 Analytical solution at 3 different times after the initial particle release(at time period 2, 4 & 8 in arbitrary units). . . . . 62
4.4 Mean Square Displacement (MSD) of particles computed from the multiple realizations of Eqn. 4.19 after tracking for a period of 10 (in arbitrary units) in a media with $D_e = 0.1$ and $\bar{u} = 0.01$. . . . . 63
4.5 Plot showing 100 realizations of concentration profiles with mean concentration profile shown in red. . . . . 64
4.6 Concentration profiles with mean and uncertainty bands with 99% confidence interval for 5000 trajectories (left) and 100,000 trajectories (right). . . . . 65
4.7 Different ensembles of simulations with 1000 particles and different number of total realizations. . . . . 66
4.8 3-Dimensional volume $\Omega$ with absorbing boundaries $d\Omega$ and the realization of Langevin equation. On this domain, the first passage time is defined. . . . . 67
4.9 Absorbing boundary condition: truncated trajcetories of Itô's sde. . . . . 68
4.10 Distribution of hitting time: The distribution of hitting time leads to a better emergency preparedness by building an uncertainty band around the mean hitting time. . . . . 69
4.11 Colloid particle shown here in ■ color attached to it are the radioactive species shown in ■ color . . . . . 70
4.12 Concentration computed for decaying colloidal particle. . . . . 71
4.13 Trajectories of particles with Bernoulli trial. Here the survival probability is set to 0.994. . . . . 71
4.14 Computed concentrations from trajectories. . . . . 72
4.15 A particle in smooth (less uncertain) and rough flows (more uncertain)–which characterizes the degree of temporal correlation. The smoother the trajectory is the higher the temporal correlation is. Correlated noise can be generated by using an OU process (see equation 4.32). . . . . 72

4.16 OU process for Γ. . . . . 73
4.17 Correlation of simulated OU process. . . . . 74
4.18 Analytical correlation function and the correlation of simulated gLe trajectories. 76
4.19 Simulated time trajectory of a generalized Langevin model. . . . . 76

5.1 A typical curve of wind and solar output, along with net (L-S-W) demand. L:Load, S:Solar, W:Wind (NERC-2009[1]). . . . . 81
5.2 Emsland NPP in Germany showing 140 MW/min ramping rate[2]. . . . . 83
5.3 Synergistic existence of nuclear and wind power in a microgrid. **(10 small wind turbines and one nuclear power plant–feeding a small locality)** 84
5.4 Classical output feedback control (AP: Actual Power, DP: Demand Power, SG: Stochastic Generation). . . . . 88
5.5 Comparison of autocorrelation and spectra of the OU process at various values of '$\gamma$': at higher values of '$\gamma$' the process relaxes faster while the spectra of the process will become flatter leading to a random white noise signal. . . . . 92
5.6 Generated wind speed time series and the probability distribution. . . . . 93
5.7 Response of the reactor dynamics model ($x$, $y$, $z_F$, $z_C$ from left to right). Time scale is in 'seconds'. . . . . 94
5.8 Response of the reduced order model owing to stochastic & deterministic generation. . . . . 96

6.1 Lagrangian correlation function of the streamwise velocity of turbulent channel flow. The correlation function becomes less flat at shorter time scales as one goes away from the wall. The Markovian models have exponentially decaying correlation which cannot capture the short-time behavior of the turbulent signal. . . . . 99

A.1 Comparison of the gain of difference and moving average filter (see equation A.15)–it will let go the higher and lower frequencies oscillations respectively. This also means the differencing and moving average or convolution integral are complementary filters to each other. . . . . 112
A.2 Two point moving average filter and differencing. . . . . 113
A.3 Fifty point moving average filter and differencing. . . . . 114
A.4 Convolution Integral done by moving the window 'u(k)' along the axis 'k', a window is shown in the zoomed plot and the Fourier transform is performed to show the response of the box filter. . . . . 114
A.5 Box-car filter response–covers the lower end of the spectrum–depending on the length of the window the frequency response of the filter changes. . . . . 115

# List of Tables

2.1 Single channel model parameters. . . . . 14

2.2 Core-wide parameters. . . . . 14

3.1 Non-dimensional parameters used in the channel flow simulation. . . . . 39

3.2 Non-dimensional parameters used in the differentially heated channel flow simulation. . . . . 45

5.1 Non-dimensional variables. . . . . 86

5.2 Parameter values used in this reduced order model[3]. . . . . 88

5.3 Nomenclature table for the reduced order model. . . . . 89

# Acknowledgments

First and foremost, I am thankful to my Ph.D. supervisor Dr. Hitesh Bindra at Kansas State University for invaluable guidance, puzzeling questions and support. The streamlined supervision, constructive reviews and resourcefulness of Dr. Bindra was immensely helpful, right from the commencement of this work to the present day. I would also like to thank: all the fellow graduate students in the nuclear engineering and physics department. All the valuable discussions with them always steered me in the right direction through the course of this work. Without the continuous support of these people, completion of this work would have been highly uncertain. For my dissertation, I also thank the jury members:- Dr. Mohammad H. Hosni, Dr. Lado Samushia, Dr. Sanjoy Das and Dr. Walter McNeil for their invaluable time and effort to review my work. I would also like to thank Dr. Bojan Ničeno of Paul Scherrer Instute in Villigen, Switzerland for his valuable advice and guidance. Last but not the least, I am thankful to my friends: Dr. Ajit Rai, Chirayu Batra, Shashank Pathak and Zayed Ahmed whose friendship always brought happiness to me. I also thank my wife Roma for being immensely supportive. I am thankful to my family for their unconditional support, love and all the positive vibes they sent me all the way from India. Especially, I thank my parents for their undying love towards me and always inspiring me. For all the above-mentioned people, words can not do enough justice to summarize their contribution and support. Without their guidance, encouragement, criticism, and concerns, realizing this work would have been impossible.

# Chapter 1

# Introduction

The state of a physical system is studied by constructing mathematical models, which often take the form of differential equations–partial and ordinary. Specific examples can be quoted from a variety of different fields, for example, Lorenz model which was invented to study oceanic convection, Euler and Navier Stokes equations for fluid flow, Advection-Diffusion equations (ADE) for temperature, concentration, etc. These equations are the workhorse of many modern computational environments which solves them by applying (or not applying) certain assumptions on it in a specific physical domain. Like any mathematical model, these equations were derived with certain known assumptions embedded in them. Therefore, it can be said that they are an idealized representation of the system under investigation. The embedded idealization originates from the *a priori* truths or *first principles* on which the model rests–the full deduction of which is an insurmountable task. In other words, such models are distorted but a useful view of the reality[4;5] Hence, it is quite possible that any calculation performed via them with (or without) any assumptions have a significant amount of uncertainty embedded in it. To tackle this, engineers and scientists often rely on quantifying the uncertainties in the performed calculations via some statistical measure. This can be done by:- a) repeated execution of the imperfect model by varying the certain parameters of it through phenomenologically derived distributions, b) which is followed by analyzing the variances produced in the outputs.

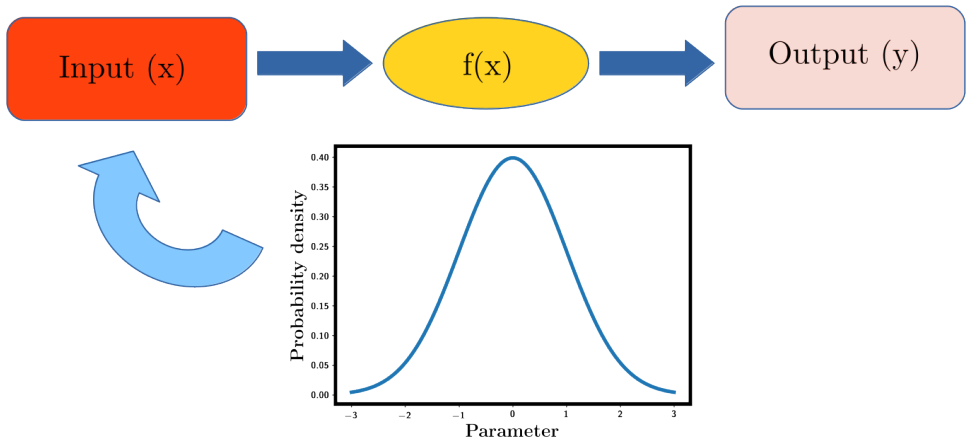



**Figure 1.1**: *Black box approach: Here $f(x)$ is an input to output mapping for a computational tool ($f : x \rightarrow y$). Inputs are being sampled from a prescribed distribution and the computational tool is run multiple times.*

At the scales of nuclear energy systems, it is often required to solve different mathematical models which feed on to each other to perform a multiphysics calculation.For the obvious reason of saving time, cost and computational resources, this is done by building a one-dimensional approximation of the three-dimensional physical phenomena (an imperfect model) occurring inside the nuclear system– covering the heat transfer, fluid flow, and neutron transport. Traditionally these thermal-hydraulics and neutronics calculations were performed on codes like RELAP, TRACE, MCNP, and PARCS which has the best available one-dimensional physical models. Even then owing to the approximations made these codes have a significant degree of parametric uncertainties.

Therefore it becomes essential to rely on an uncertainty quantification (UQ) technique. The traditional UQ techniques typically rely on a black-box approach. In the black-box approach, the inputs were mapped to the output space via a complex mathematical function or a surrogate of the original model–which then is perturbed by various possible sampling techniques (random or more sophisticated sampling approaches such as stratified and Latin hypercube). This is done to separate out the systematic (Epistemic) uncertainties from the stochastic one (Aleatory). This all is performed by relying on the state of the art software tools such as Dakota[6] and Raven[7]. Both of which are design and optimization tools being

developed at Sandia national laboratory and Idaho national laboratory respectively. These toolkits enable the development of powerful surrogates based on data-driven approaches, for example, but not limited to principal component analysis; nearest neighbors approach, spline regressors and support vector machines. While Dakota is geared towards black box uncertainty quantification techniques with almost no room for intrinsic variability of the parameters (dynamic evolution of the uncertain parameters), the current version of RAVEN has started looking into the dynamic nature of the key system quantities from the similar route mentioned in the current work. These surrogates based on the black box approach lack a deep physical interpretation.

This is true because the black box route relies on the input distributions based on expert judgments (see figure 1.1) than an actual physical understanding. While this can provide a tractable way of uncertainty quantification, they tend to fell short on detailed physics and are not informed by it. For example, the key engineering quantities of interests often depend on the prediction of velocity and scalar fields. The transport of these two quantities is governed by the respective diffusion coefficients which vary from one time scale to another (for e.g., scalar quantities (like temperature, concentration, etc.) advected by turbulent diffusion falls under the Batchelor's regime at short time scales[8] and in Richardson's regime at larger time scales[9]). Lack of this physical understanding will always create a knowledge gap in the models based on the black box route with externally perturbed parameters. Owing to this time scale dependence of the key parameters of the model, it can be said that the model parameters tend to evolve and in reality, are non-linear in nature. Thus, it can be said that the surrogates provided by the black box route tend to neglect the origin of the two types of uncertainties (Epistemic & Aleatory) i.e., the chaotic nature of the micro-physics. In other words, they tend to the neglect the stochastic (or Aleatoric) nature of parameter. Therefore, the variance produced by them may put an unnecessary bias on the system performance itself. Since the two types of uncertainties originates from the chaotic motion of the building blocks of the system. Both of which get coupled to each other at the small scale. Hence, it is important to study the behavior of the system from the microscopic standpoint and then to propagate the uncertainties to the macro level. This is done by studying the specific random

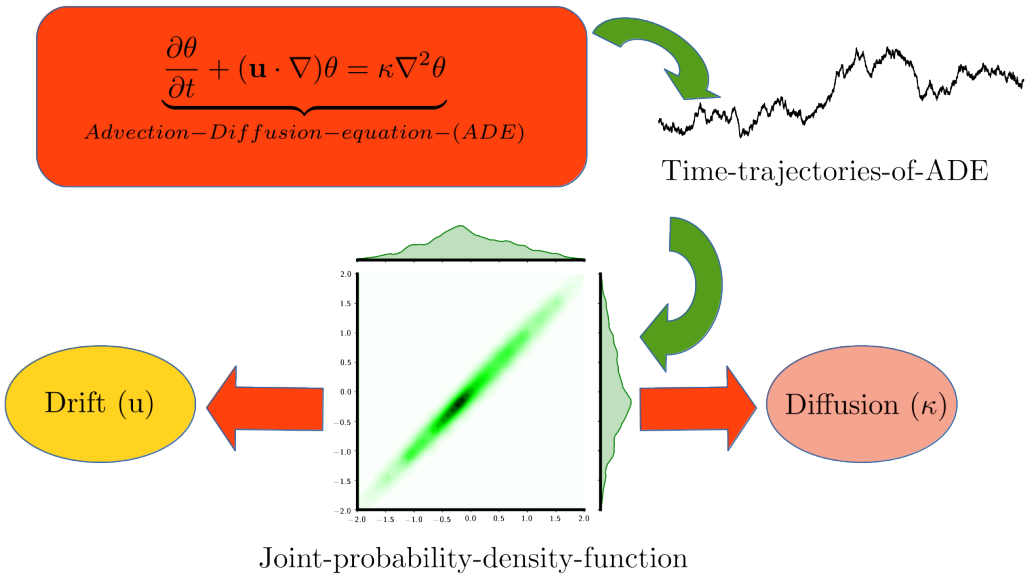



**Figure 1.2**: *Computation of average and fluctuating part of the signal via the solution of the inverse problem depicted in the picture.*

process driving the system in the second part of the current work. Where time signals were analyzed via resorting to an unsupervised learning methodology introduced in the context of non-equilibrium statistical mechanics[10]. In its very essence, the method projects the collected data into a conditionally averaged and a fluctuating part (see figure 1.2). These two parameters have the interpretation of drift and diffusion of the random signal, respectively.

These physics-driven surrogates can be directly embedded in the legacy safety analysis codes. The utility of these codes can be understood from the fact that the high-resolution DNS simulations are still beyond the reach of current computational capabilities which render them of little significance to the nuclear community. On the one hand, averaged out differential equation has a considerable amount of uncertainties while the resolution of all the trajectories of the dynamical system is still beyond the reach. Hence, a unique low complexity (mesoscopic) model of the velocity signal constrained directly on the high-resolution DNS data or experiments may fasten the nuclear safety analysis process. Since the surrogate is unsupervised, it tends to find the governing parameters with well known physical interpretations directly from the collected data.

The noise driving the statistical emulator is completely memoryless in nature. However,

a real system is driven by a random noise only at a certain spatiotemporal scale. For the specific case of turbulent flows (or the chaotic motion of the fluid) the process tends to evolve through three distinct scales–a) energy-containing regime, b) inertial sub-range & c) the viscous dissipative sub-range. It is memoryless only in the inertial sub-range while it shows memory in the viscous dissipative regime. Therefore, a delaying or the memory term can be introduced into the system via a convolution operation.

Rest of the thesis details the concepts briefly presented here with relevant examples. The next chapter begins with the introduction of black box uncertainty quantification before delving into the specifics of perturbing the input variables of the single-channel RELAP model of the Kansas State University TRIGA reactor for epistemic uncertainty quantification purpose. The second chapter demonstrates a methodology to directly obtain a physics-based surrogate by constraining a non-linear Langevin equation on the collected data. The latter two chapters show the application of simple linear Langevin equation for spatial uncertainty quantification and generating uncertain power profiles for synergistic existence of nuclear and wind power producers in a small micro-grid.

# Chapter 2

# Uncertainity quantification and Sensitivity analysis - Using Best Estimate System tools

Much of the design and engineering of the system relies on large-scale simulation for progress because the phenomena of interest typically incorporate multiple physical processes interacting over a wide range of spatio-temporal scales. This is true even for the study of the basic processes, e.g., the flow of fluid between two plates, which requires simulations that can challenge massively parallel computer architectures. Therefore approximated physical models of the complex physical process are utilized in the realm of engineering and scientific design, optimization, and control. Modern nuclear safety analysis codes are not an exception to this. The nuclear industry relies on various hydraulics (RELAP, TRACE & GOTHIC) and neutron transport codes which have approximated and imperfect models embedded in them. They are imperfect like every other model, but they include the best available physics. They are thus providing the analyst with most updated modeling capabilities. Despite being very capable on the physics and code architecture front, still, the uncertainties impose a limit on the results of those hydraulic and neutronics calculations. One identifiable source of uncertainty in these models can trace its roots to the parameters used. These parameters

may have many uncertainties because of the various approximations made during their measurement and calculation phase.Therefore for a credible computational driven safety analysis of the modern nuclear reactor and for the license extension of the old fleet, it is essential to quantify the uncertainties associated with the nuclear codes and calculations.

To this end, a case study is presented which tries to answer the uncertainties associated with the fuel burnup and form friction factor in an existing KSU TRIGA reactor code. The epistemic uncertainties were propagated through the reactor code by treating it as a "black box". This is advantageous from the standpoint that one does not have to analytically propagate the uncertainties through the individual arithmetic operations involved in the simulations. To perform which a code user has to be extremely skillful in applying the rules of advanced mathematical techniques (interval arithmetic) to account for the range of floating-point values taken by the input parameter and the corresponding range of values generated by the code[11;12].

With this motivation of performing the reactor safety analysis together with the quantification of uncertainties, a brief overview of the same is presented in the next section together with a motivating example of applying the methodology to the Kansas State University research reactor.

## 2.1 Motivation

Safety analysis of nuclear reactors (both research and power)[13;14] is performed by constructing models in best-estimate computer codes that use the best available physics[15], techniques, and practices to simulate steady-state, accident, and anticipated transient conditions. However, the use of best-estimate codes and models does not preclude the existence of errors in the approximated system models. The sources of these potential errors are diverse and can range from round-off errors, system modeling approximations, input data uncertainties[16], and user throughput. Thus, it is advisable to assess the multi-physics simulation results produced with best-estimate codes and models by using appropriate uncertainty quantification[17].

Such uncertainty quantification has become especially important in recent years as research reactors that were designed and constructed decades ago have undergone modifications for increased licensed powers or to convert from high- to low-enriched fuel in support of safeguard efforts.

As a specific example, the Kansas State University (KSU) TRIGA Mark II research reactor recently made a request to the U.S. Nuclear Regulatory Commission to permit loading of four, 12% U (by mass; 20% enriched) ZrH-U fuel elements. Since its first operation in 1962, the KSU reactor has only used 8.5% U fuel, with a change from aluminum to stainless-steel cladding in 1973, at which point the majority of KSU's fuel in use today was first loaded. The larger U mass of the new elements will help balance the reactivity loss due to the rather large burnup of those original elements and allow the reactor to be operated at its licensed power of 1.25 $MW_{th}$.

However, to use these elements requires a better understanding of the potential impact on safety their use would entail. The current methodology used by KSU reactor staff to model the core physics is rather simple and limited essentially to a steady-state neutronic analysis based on MCNP[18]. A significant approximation made in the model is to compute element-wise burnup using the core-averaged power. In other words, the spatial dependence of the neutron flux has not generally been used to update fuel compositions. Moreover, the majority of fuel loaded in 1973 already had significant burnup (more, in fact, than has been incurred in the KSU core since that time). By using the documented burnup of these elements (quantified by the mass of $^{235}$U depleted) and log-book records of power over time, the staff computes the estimated amounts of $^{235}$U and $^{238}$U for entry into the MCNP model, neglecting all other actinides and fission products. This model is then used to compute local powers without further, explicit thermal-hydraulic analysis. Hence, although the method used is adequate for normal operation of the current core, to understand the potential safety impact of loading 12% U elements requires a more detailed understanding of the current core composition and local power-peaking factors, their associated uncertainties, and how these uncertainties ultimately propagate to a detailed thermal-hydraulic, safety model.

To this end, a detailed depletion sequence was generated in which explicit, element-wise

compositions were tracked based on core-averaged, ring-averaged, and element-averaged powers. Because the documented burnups of the originally-loaded elements were computed on a ring-averaged power basis, a comparison of the present-day, ring-averaged and element-averaged compositions provided a first estimate for the potential uncertainty in those documented burnups. Specifically, the deviations between the ring- and element-averaged burnup were fitted to a normal distribution, from which 1000 possible core loadings were sampled and subsequently depleted, leading to 1000 different, present-day compositions and element-wise power profiles for a core loaded with the 12% U fuel. Moreover, to demonstrate that the new loading pattern will not lead the reactor to unwanted situations and breach the defined safety limits, it is crucial to demonstrate the safety through a qualified computer code. Hence, a representative, single-channel, thermal-hydraulic model of the reactor was developed in RELAP5/MOD 3.3[19], which is a best-estimate computer code designed primarily for the safety analysis of Light Water Reactors(LWRs) that has been used previously for safety analysis in TRIGA reactors[20]. By using the sampled, present-day core compositions, the uncertainty of the original fuel burnups was propagated directly through the safety-analysis model.

The rest of this chapter is organized as follows. Presented in Section 2.2 is a brief overview of safety analysis, with particular attention paid to uncertainties and methodologies. In Section 2.3, the neutronic and thermal-hydraulic models used to analyze the KSU TRIGA reactor are described, while the major results of the analysis and subsequent conclusions are presented in Sections 2.4 and 2.5, respectively.

## 2.2 An Overview of Safety Analysis

Rules for systematic evaluation of the accident-like situations in a nuclear power plant were established in the year 1974 when the U.S. Nuclear Regulatory Commission (NRC) published rules for LOCA analysis in 10 CFR 50.46 and Appendix K[21]. The rules prepared by the NRC, which are based on conservative assumptions and rely on minimizing possible damage or consequences pertaining to some fixed criteria, were adopted due to insufficient knowledge

of the physics governing the systems. To address this knowledge gap, the NRC launched an extensive R&D program that led to the development of large-scale test facilities where vital; experimental data could be collected. The result of this effort was essentially a new method, CSAU (Code, Scaling, Applicability and Uncertainty)[22], which led to best-estimate codes and models quantified with an uncertainty band for the safety analyses of the nuclear power plants. These best-estimate models provided operators the significant benefit of being able to increase the power output of their plants. Previously, the whole industry was heavily penalized as there was always a large gap between the pessimistic calculations and the best estimate calculations and the available margin of safety.[17]

### 2.2.1 Methodology

It was mentioned above that several uncertainties may enter and propagate through simulation models due to variance in the epistemic variables.

Understanding the effect of uncertainties in the burnup estimates on thermal-hydraulic behavior of reactor can be critically important for safety assessment, especially when any refueling operations or mechanical design changes are intended. In this work, 1000 cases with uncertain input conditions for estimating burnup were sampled to obtain variations in the neutron flux profile using MCNP. The axial flux profiles were used as input for axial volumetric heat generation in the thermal-hydraulic model i.e. RELAP input file. However, in RELAP itself there are varying degrees of uncertainties arising from thermal-hydraulic model parameters such as the form friction loss coefficient. To compare the effect of uncertainties in the burnup on the reactor safety calculations, the independent sensitivity analysis of the RELAP model was conducted with variation in a parameter which is completely independent of thermal energy models such as friction parameter. This independent analysis allows the comparative sensitivity analysis of the RELAP only models and the RELAP+MCNP models. In a final study, uncertainties in axial peaking factors and friction factors were jointly sampled to model the system behavior.

The safety analysis in these studies was mainly quantified based on the estimates for

axial coolant temperature profile, and the critical heat flux (CHF).

The coolant outlet temperature and the minimum margin between predicted and critical heat flux in the core were used as classifiers to assess the safety of the system. The quantitative spread about the mean coolant temperatures evaluated for various cases with different input parameters were used to present sensitivity analysis results. More details on the methodology for burnup estimation and thermal-hydraulic modeling are provided in the following section.

## 2.3 Model description

### 2.3.1 Fuel Composition and Uncertainties

Since 1973 till the year 2016 KSU reactor core has been arranged in 29 different configurations with 100 different fuel elements making up the core. The fuel in the current configuration has been sourced from three different vendors (General atomics, Northrop corporation and CERCA). Since the burnup of the fuel sourced from the General atomics (GA) group is reported only in the integral form with no quantification of the uncertainties in the reported equations which might have created a significant error in the calculated numbers. In addition to this in the General atomics methodology, a ring averaged power peaking factor was used. While the fuel in KSU's Triga reactor core is not arranged in a ring like fashion, hence, the methodology prescribed by the GA group might have caused a major source of error in the burnup calculations of the current core configuration of the KSU's Triga reactor. To understand this, a series of Monte-Carlo computation were performed starting from the as-loaded and, ultimately, the present-day fuel compositions in the KSU reactor. Details of such Monte-Carlo calculations can be gleaned from the Master thesis of S.M Alshogeathri[23].

### 2.3.2 RELAP model description

Reactor Excursion and Leak Analysis Program (RELAP) solves the fundamental equations of heat and mass transfer together with numerous empirically derived closure laws; thus the

usage of it can be extrapolated to other industrial processes too.

Specifically, RELAP uses one-dimensional, one-fluid or two-fluid hydrodynamic models for simulating single-phase or two-phase flows, respectively[19]. The governing hydrodynamic equations are averaged in volume and time. Transverse scalar transport phenomena between multiregions like heat transfer are formulated in terms of the bulk properties by utilizing various empirically-determined coefficients such as the heat transfer coefficient.

A model simulating the natural circulation phenomena in the hottest channel of a specific ring (B-ring or F-ring) of the reactor facility was constructed in RELAP. The model consists of four significant components to simulate natural convection: a) coolant source and sink, b) cold leg, c) hot leg, and d) fuel element. The coolant sink and source were modeled using the time dependent volume component of the RELAP[20]. The source and sink were initialized to a pressure of $1.43 \times 10^5$Pa, which is the sum of the atmospheric pressure plus height of the water above the cold and hot channel. The coolant temperature was set to 49°C (322 K), which is the maximum temperature of the coolant in the reactor by design. The cold leg in this model is a representative of a downcomer for the coolant, while the hot channel is the flow region adjacent to the fuel element. The fuel element was modeled using the cylindrical heat structure component of the RELAP. The two legs of the model, i.e., the cold and hot legs, were connected using the single junction component of RELAP.

The RELAP model was built using the parameters listed in Tables 5.3 and 5.2, and the diagram shown in Figure 2.1. That diagram shows two time-dependent volumes that enforce the pressure boundary conditions and two pipes that simulate the cold and hot channel connected via a single junction component of RELAP. Heat is added to the fluid by incorporating the heat structure component (simulating a fuel element) of RELAP with an appropriate axial power profile (Figure 2.2) and power level. For example, the power level for the "B" ring is at 22 kW, corresponding to a ring-average peaking factor of 1.49 (with addition of the 12% fuel).

As shown in the Figure 2.2, the axial distribution of peaking factors with the uncertainty bands, explained in the following section, were used to compute power levels for individual simulations. This power level was applied to the heat structure within the single channel.

Similarly, power levels were computed for F-ring with similar uncertainty bands.

Owing to the smallest flow area, the dimensions of the average flow area were based on the "B" ring. Due to this reason and lack of cross flow in this model, this single channel model can be considered as conservative.

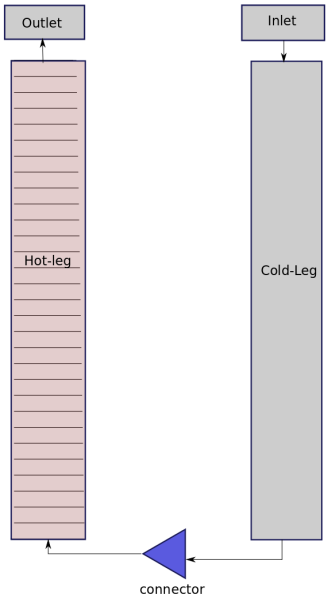



**Figure 2.1**: *Single Channel Model.*

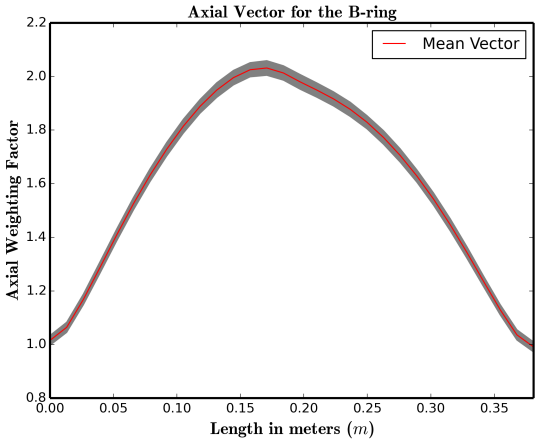



**Figure 2.2**: *Axial Peaking factor profile for B-ring fuel elements with associated uncertainty band.*

**Table 2.1**: *Single channel model parameters.*

| | |
|---|---|
| Flow area | $2.5091 \cdot 10^{-4}$ ($m^2$) |
| Hydraulic Diameter | 0.01668 (m) |
| Heated Diameter | 0.0184 (m) |

**Table 2.2**: *Core-wide parameters.*

| Time dependent volume | |
|---|---|
| Pressure (Pa) | $1.43 \cdot 10^5$ |
| Temperature (K) | 322.15 |
| Cold & Hot leg height | |
| Height (m) = 0.711 | |

## 2.4 Results and Discussion

The total burnups were sampled using assumed, ring-averaged total burnup values $B_r$ for the current core and the distribution described in Section 2.3.1. In total, 1000 core compositions were defined in which the total burnup of each fuel element was sampled independently. These cases were run with 30 inactive cycles and 250 active cycles with 100,000 histories per generation in the MCNP model of the reactor core. These simulations were used to generate the axial peaking factors for fuel rods located in each core ring. The power-peaking factors for each ring were found to be normally distributed. A example axial power profile the for B-ring is shown in Figure 2.2 with a 95% confidence band resulting from the 1000 simulations. The average standard deviation of the power at individual axial nodes is 1.13%. These axial peaking factors were used in RELAP5/MOD 3.3 to understand thermal response and safety characteristics.

### 2.4.1 Thermal-hydraulic analysis using RELAP model

Before simulating the actual cases for reactor safety analyses, RELAP model was tested for grid convergence and was verified with the help of higher fidelity 3-D computational fluid dynamics code. As the RELAP model is a 1-D, single-channel model, the discretization in this 1-D or axial direction can significantly impact the results. The heated region of the

channel was divided into 15, 30 and 60 axial nodes for grid convergence tests. The maximum change observed in local temperatures was less than 0.1 K after changing the number of axial nodes from 30 to 60. Based upon this analysis, it was concluded that 30 nodes were sufficient to simulate the heated section of this 1-D channel. The commercial multiphysics software COMSOL[24] was used to construct a 3-D thermal-hydraulic model for verification of RELAP models. The two models show very good agreement on the axial node temperatures for both the B and F ring channels. More details on these verification tests can be found in previous work[25].

The verified model was then used with various axial peaking power profiles or localized volumetric heat generation rates assigned to corresponding axial nodes. As described before, RELAP model was constructed as a single channel model which implies the axial power profile of a particular ring (such as, B-ring) only was used as input data for each simulation case. The mean coolant temperature for B-ring and F-ring are plotted in Figure 2.3. Departure from Nucleate Boiling Ratio (DNBR), the ratio of the critical heat flux to the actual local heat flux, was computed for the B-ring in the same model(Figure 2.4). These results show that the maximum local heat flux in the B-ring channel is at least a factor of three below the critical heat flux.

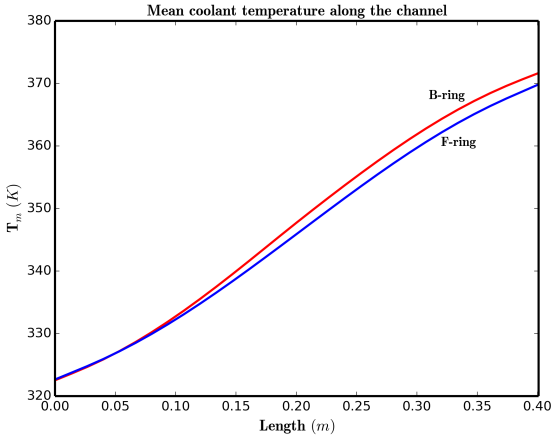



**Figure 2.3**: *Axial distribution of Coolant temperature next to B-ring and F-ring fuel elements.*

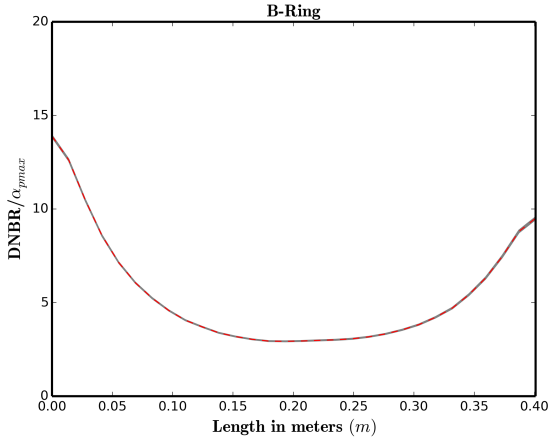



**Figure 2.4**: *Axial distribution of DNBR in B-ring ($\alpha_{pmax}$ is the maximum power peaking factor).*

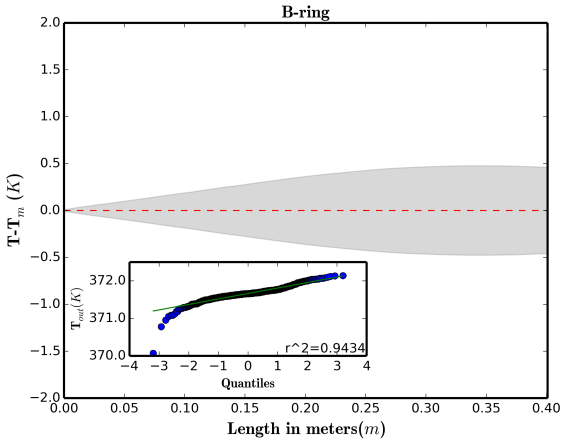



**Figure 2.5**: *Effect of uncertainties in axial peaking factors: Differential spread (99% confidence band) of coolant temperature adjacent to B-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet coolant temperature around mean value.*

### 2.4.2 Uncertainties in RELAP thermal-hydraulic model

Although single-channel RELAP models are widely used and accepted for thermal-hydraulic studies and safety analysis, there can be uncertainties associated with model input parameters, such as heat transfer coefficient[26] and friction coefficients. A particularly important uncertainty comes from the modeling of friction factors and can be significant because of conversion from a 3-D physical domain to 1-D single channel models. Accurately modeling

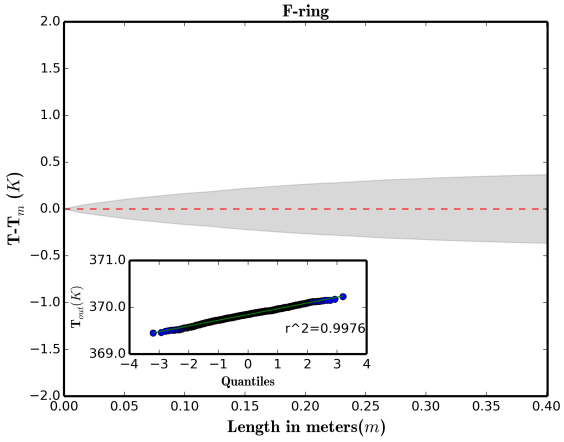



**Figure 2.6**: *Effect of uncertainties in axial peaking factors: Differential spread (99% confidence band) of coolant temperature adjacent to F-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet temperature around mean value.*

form loss coefficient for actual reactor features, such as bottom and top grid plates, can be quite challenging as the associated flow undergoes changes in the direction and flow area simultaneously. In RELAP models, this form friction factor is modeled at the junction connecting cold leg and hot leg, which is representative of the physical geometry connecting pool water with the core via grid plates. Therefore, to understand the effects of uncertainties within thermal-hydraulic model and their relative magnitude on the coolant temperature prediction, RELAP simulations were performed with different input values of form-friction factors based upon the uncertainties reported in the literature with similar geometry of grid plates and flow-rates[27–30]. A normal distribution of form friction factors ($f \sim \mathcal{N}(1.71, 0.1)$) was obtained to generate 1000 input files, and these 1000 cases were simulated to obtain axial profiles of coolant temperature. The axial power profile for these studies was kept constant using the mean values at individual nodes from the power factor calculations discussed in previous subsection. Detailed statistical analysis of these simulation results is presented in the Figure 2.7.

A comparative analysis of the plots in Figure 2.5 and Figure 2.7 shows that ∼1% average change in axial peaking factors leads to 2 K difference in predicted coolant outlet temperatures. Whereas ∼10% change in friction factor leads to only 0.7 K difference in the predicted

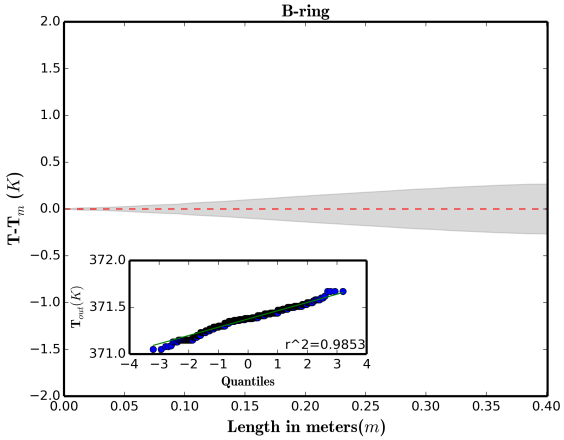



**Figure 2.7**: *Effect of uncertainties in form friction factor: Differential spread (99% confidence band) of coolant temperature adjacent to B-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet temperature around mean value.*

values of coolant outlet temperatures.

Based on these simulation results, it can be concluded that outlet coolant temperature predictions are more sensitive to uncertainties in axial peaking factors as compared to uncertainties in form friction factors.

### 2.4.3 Effect of uncertainties in multiple parameters

In more realistic scenarios, the uncertainties in the axial peaking factors or form friction factor can occur simultaneously. To resolve the coupled effect of uncertainties in two or more model parameters on the predicted behavior, all of these parameters should be varied simultaneously. In the previous two studies, 1000 cases with different axial profiles and 1000 cases with different form friction factors were simulated. The combined sensitivity analysis of simultaneous variation in different parameters requires multivariate sampling of these different parameters with associated uncertainties. A classical approach to this multidimensional sampling is adopted to select a set of 1000 parametric inputs using joint sampling from multivariate normal distribution as all of the individual parameters – axial peaking factors and friction factors are normally distributed. This multivariate normal distribution can be expressed as Equation. 2.1.

$$\mathbf{x} \sim \mathcal{N}_k(\mu, \Sigma) \tag{2.1}$$

where, $\mathbf{x} = [x_1, x_2, x_3, ...x_k]$ is a $k$-th dimensional parametric input set obtained from multivariate normal distribution, $\mu = [\mu_1, \mu_2, \mu_3, ...\mu_i, ..., \mu_k]$, $\mu_i$ is the mean value of an $i$-th parametric input and $\Sigma$ is the covariance matrix explicitly described in Equation. 2.2.

$$\Sigma = \mathbf{Cov}[x_i, x_j] \tag{2.2}$$

In the actual analysis, 31 parameters (1 friction factor and 30 power peaking factors for B-ring) were jointly sampled using this covariance matrix. An example distribution of jointly sampled two parameters, or two-dimensional parametric input, is shown in the Figure 2.8.

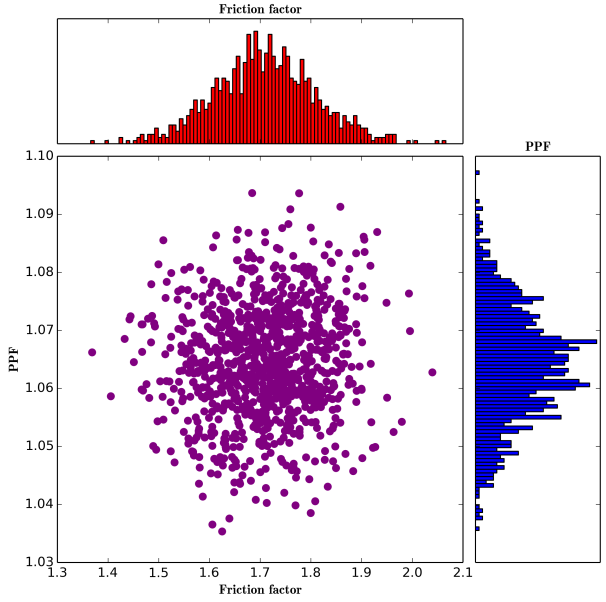



**Figure 2.8**: *Joint sampling of multiple parameters: Example of jointly sampled (In Purple) friction factors and PPFs for any axial node with the uncertainties.*

With the jointly sampled set of axial peaking factors and friction factor, 1000 cases are simulated using RELAP code. The combined effect of these uncertainties on coolant temperature is shown in Figure 2.9. The sensitivity is assessed with the help of maximum change in the prediction of coolant outlet temperature, which in this case is 1.5 K.

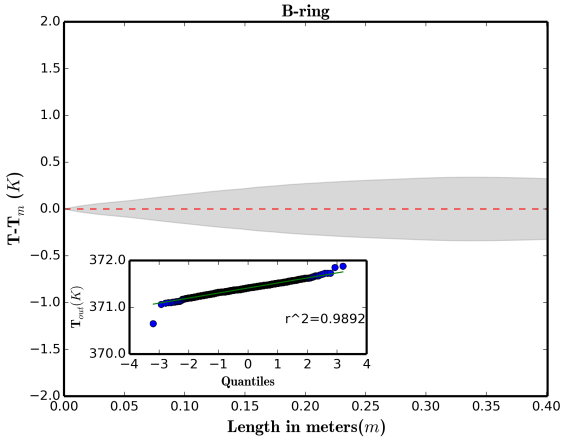



**Figure 2.9**: *Effect of uncertainties in multiple parameters: Differential spread (99% confidence band) of coolant temperature adjacent to B-ring fuel element about the mean value as a function of axial location; Sub-plot shows normal distribution of outlet temperature around mean value.*

## 2.5 Conclusion

Burn-up estimates and the associated uncertainties were computed to obtain a more accurate estimates on power peaking factors in KSU TRIGA reactor. The sampled burn-up values with associated uncertainties were used to generate 1000 MCNP simulation cases to obtain peaking factors with uncertainty bands. Axial peaking factors were used to compute local axial power levels or volumetric heat generation rate in RELAP thermal-hydraulic model. The propagation of uncertainties in burn-up estimates through MCNP and RELAP codes was analyzed. The maximum deviation in the prediction of coolant outlet temperatures, one of the critical output parameters for safety analysis was found to be 2 K with a confidence interval of 99%. RELAP output, coolant temperatures, were found to be more sensitive to uncertainties in axial peaking factors as compared to uncertainties in hydraulic model parameters, e.g, form friction factors.

The effect of uncertainties in multiple parameters on model behavior was then studied using jointly sampled distribution of axial peaking factors and friction factors. The simulation results with jointly sampled parameters show that predicted deviations remain lower than the case where only axial peaking factors were varied. This chapter shows a methodology to

perform sensitivity analysis and uncertainty propagation in multiple codes, and can be used as an alternative method for the reactor safety analysis.

The method presented in this chapter is based on perturbing the existing legacy safety analysis code in a non-intrusive way. In this way, it is beneficial for quantifying the uncertainties in engineering system in a blind manner. However, for a deeper and more physical understanding of the uncertainties itself, it is advisable to invoke the more scientific and rigorous theoretical grounds. The next chapter is an attempt to do the same using the existing tools of non-equilibrium statistical mechanics.

# Chapter 3

# Statistical modeling of turbulence

## 3.1 Origin of Uncertainties

The last chapter presented a case study where the uncertainties attributed to the burnup and friction factor were treated via a black-box approach[31]. Parametric uncertainties were studied by perturbing the computational model. Two best estimate codes were used for uncertainty propagation and sensitivity analysis. The goal of this approach from an uncertainty quantification point of view is to bring the epistemic uncertainties towards aleatoric uncertainties without invoking the rules of advance mathematical methods (e.g., interval arithmetic[12]). However, the two uncertainties can trace back their root to the very nature of the system itself–which is a blend of random and organized behavior. While the organized behavior can be studied by some simpler models, statistical tutiouools best study the random part of the solution. Statistical mechanical point of view ascertains that the various uncertainties originate from the chaotic/ random nature of the governing micro-physics of the system. Hence, it can be said that both the aleatoric and epistemic uncertainties originate at the micro-level. In other words, they are the ramification of the microscopic behavior of the building blocks of the system. To this end, it is essential to understand and quantify the impact of the movement of the micro-physics on global behavior. This motion can be studied by building local stochastic models of the phenomena under consideration.

### 3.1.1 State of the art in engineering applications

The canonical evolutionary partial differential equations governing the specific physical process, e.g. fluid flow (Navier-Stokes and Euler equations) and advection of scalar (advection-diffusion equation) are some examples of high dimensional systems with infinite dimensionality. The solution of such equations takes considerable computational resources. Therefore, to resolve these issues and to save valuable computational time, often Spatio-temporal averaging is performed. This averaging in effect removes either all or partially the associated small scale structures (meso and micro) and leads to a fast running computational tool. It is most effectively done by Computational fluid dynamics (CFD) simulation techniques like LES (Large Eddy Simulation), RANS (Reynolds-Averaged Navier-Stokes) and URANS (Unsteady Reynolds-Averaged Navier-Stokes). It is shown that only high-fidelity approximated CFD techniques like large-eddy simulations (LES) have been able to capture the physics of the fluctuating field with accuracy. Often at times, it is required to quantify and characterize the uncertainties creeping in the system via a repeated execution of the model. However, the substantial computational costs associated with the high-resolution techniques prohibits the same. On the other hand, Reynolds-Averaged Navier-Stokes (RANS) and other similar models are computationally less intensive but are limited in their capability to capture complex and chaotic behavior of the flow accurately. Thus, it can be said that any approximation introduces a degree of uncertainty in the predicted quantity of interest. Hence, it is mandatory to explore other physical models. Typically lower fidelity 0-D or 1-D system codes are used for safety analysis, which is fast, but the various physical parameters embedded in theses codes have a significant degree of uncertainties.

### 3.1.2 Limitations of the system codes

The system-level codes used in reactor safety analysis have highly approximated physics. A 0-D approximation models the 3-D flow patterns in the reactor plena in the system codes. These approximations can significantly underestimate the important physical attributes, for example, the form friction factor in the junction connecting the hot and cold junction in the

RELAP model shown in figure 3.1. These gross approximations can be improved considerably using CFD codes- leading to question of how would 3-D models inform 1-D or 0-D models. Homogenized system parameter through a realization is bound to produce errors as these 3-D flows significantly evolve in time– therefore produce physically inaccurate results due to the '*unfolding of flow*' over time.

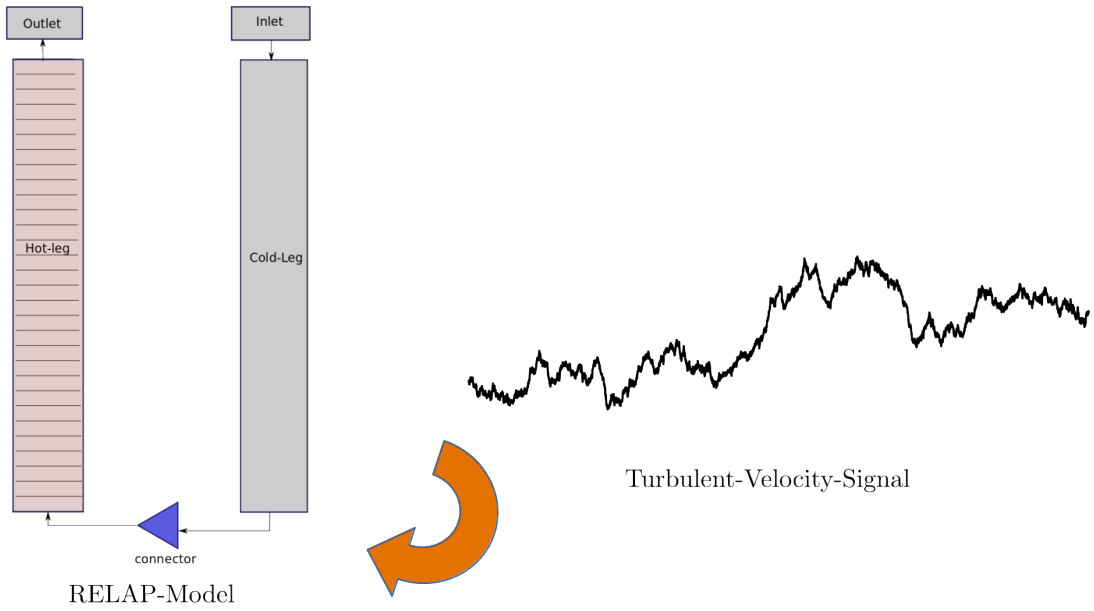



**Figure 3.1**: *One-dimensional RELAP model can be enhanced by coupling it with a realistic-looking velocity signal generated by a 'Langevin' surrogate model.*

This inaccurate depiction of a valuable property which is attributed to the turbulent flow in the system codes can lead to an erroneous prediction for scalar transport. Scalar mixing or the enhancement of the transport of energy from the point of its origin by the advecting scalar ('Temperature') is an important variable. An advection-diffusion equation can describe the dynamics of the advected scalar ('$\theta$'). The system codes handle the solution of this equation by an effective diffusivity type of parameter. While this provides a tractable way of solving the governing equation, this falls short in the light of the fact that the turbulent diffusion coefficient varies with turbulent fluctuations and time[9;32–37]. Hence, the effective representation of diffusivity is bound to lose a significant amount of information. Since turbulent diffusion is a phenomenon introduced by the rapid temporal fluctuations

of the velocity, any attempt to treat it as a nonstochastic/nonchaotic process will lead to erroneous results. It is this very important behavior–which affects many other parameters has not been captured by the system analysis codes which do not allow them to accurately depict uncertainty bands for scalar transport with multiple realizations.

**Uncertainty quantification in Nuclear engineering**

Typically to reduce this information gap introduced by various coarse-graining approximation discussed earlier, sampling-based perturbative forward uncertainty quantification methods have been used in the past[31]. Recently various approaches have been adopted to construct reduced models which mimic the physics and aid safety calculations.While Wang et al.[38] has used a Gaussian process regression with a specified correlator of the spatial noise, Huang et al.[39] has relied on a simple non-linear mapping of input and output vectors, on the other hand, Wu et al.[40] has relied on a polynomial chaos expansion (PCE) based emulator of the process under consideration. These models can typically be simulated by a prescribed distribution of the input parameters for forward uncertainty propagation or in another approach the variance of the input parameter can be computed via an inverse uncertainty quantification methodology[38;41]. These methods may reduce the computational cost and produce a variance around the quantities of interests. Since the input variables are mapped to the output variables via a complex non-linear function, hence it is difficult to get a clear physical picture of the system via the application of this approach.

### 3.1.3 Physics based approach

On the contrary, a physics-based approach to model the critical system attributes governing the system behavior(such as velocity, temperature, pressure, etc.) is to invoke the relevant ideas from statistical mechanics–which in the recent years has penetrated a number of areas of science & engineering[42;43]. The goal of such physical understanding is to characterize the stochastic process associated with turbulent flows and mixing. In this portrayal of the physical system the coarser behavior is attributed to the averages of the relevant quantities

such as flow and scalar field. In its usual context statistical mechanics tries to solve the problem of 'disorganized complexity'–while the idea of averaged out behavior reveals the 'organized simplicity.' The interaction of these extreme ends comes in the middle of the two, i.e., in the so-called 'organized complexity'[42;44] or the mesoscopic regime. It is in the mesoscopic regime where the rich set of non-linear, non-deterministic, and not wholly random interactions occur[42;44]. Hence, it can be said that the microscopic phenomena effect the macroscopic behavior of the system. However, owing to the very short time scales associated with microscopic phenomena, it is often prohibitive to resolve them computationally because of the reasons described earlier. Therefore, a much better physics-based approach relies on coupling a mesoscopic model with the coarse macroscopic system code. This approach can conserve most of the physics of the quantity of interest at a local stochastic level. To this end, understanding from non-equilibrium statistical mechanics is utilized for building a mesoscopic 'Markovian' description of the dynamics of the small scale velocities.

## 3.2 Stochastic flow equation and Fokker-Planck representation

This section introduces the notion of Fokker-Planck representation and its statistical counterpart or the local stochastic model for key 'Lagrangian variables.' Further, the equivalence between the advection-diffusion is established which paves the ground for stochastic flow equation for temperature fluctuations.

### 3.2.1 From Newton's law to local stochastic model

A low complexity statistical model (or a local stochastic model) of a large dynamical system can be constructed by studying the local behavior of the random/ chaotic trajectories of it. This can be studied by resorting to simple statistical models of the process. These models can be constructed by invoking the laws of classical Newtonian mechanics. The idea can be made concrete by considering the case of decaying velocity signal which can be studied by

setting a simple equation of the following type (see equation 3.1):

$$\frac{dv}{dt} = -\gamma v \tag{3.1}$$

Here, $\gamma$ is the decay rate of the velocity signal. Equation 3.1 is a linear first order differential equation having a solution:

$$v(t) = e^{-\gamma t} v(0) \tag{3.2}$$

From equation 3.2 it can be said that the velocity will come to rest at a long time. However, at equilibrium, it should have a finite value of $\infty > \langle v^2 \rangle_{eq} > 0$. This is because of the statistical mechanical nature of the system, according to which the actual velocity cannot go to zero–or else the system will completely come to a 'dead state.' Therefore, the modified equation of motion will be:

$$\frac{dv}{dt} = -\gamma v + C\zeta \tag{3.3}$$

Equation 3.3 is simply Newton's second law driven by a randomly varying force term. It is also known as the Langevin equation after its inventor Paul Langevin.

### 3.2.2 Fokker-Planck equation

Corresponding to a 'Langevin' or an 'Itô' stochastic differential equation a Fokker-Planck equation can always be written which describes the dynamics of the probability current which entails all the microstates of the system. This is possible as by taking conditional moments of the Itô's lemma (see references[45],[46])–which is the formula for taking the derivative of a function of an 'Itô process' (see eq. 3.13) it is possible to recover the Fokker-Planck equation.

### 3.2.3 Itô's Lemma

For a smooth function '$f$' of the stochastic process '$v$' (which is the velocity here) and time '$t$' the two dimensional Taylor expansion can be written as:

$$f(v+\delta v, t+\delta t) = f(v,t) + \delta v f_v(v,t) + \delta t f_t(v,t) + \frac{1}{2}(f_{vv}(v,t)\delta v^2 + 2f_{vt}(v,t)\delta v \delta t + f_{tt}(v,t)\delta t^2) \tag{3.4}$$

Here, terms like $f_v$, $f_t$, $f_{vv}$ etc are the partial derivatives of the function '$f$'. For the limiting case $\lim_{\delta v \to 0} \delta v = dv$. Inserting the stochastic version of Newton's second law (or Langevin corrected) will read as:

$$dv^2 = (-\gamma v dt + C\zeta dt)^2 \tag{3.5}$$

$$\underbrace{dW = C\zeta dt}_{\text{wiener process}} \approx \mathcal{N}(0, dt) \approx C \underbrace{\mathcal{N}(0,1)}_{\epsilon} \sqrt{dt} \tag{3.6}$$

$$dv^2 = (\epsilon C)^2 dt + \mathcal{O}(dt^2) \tag{3.7}$$

Hence, it is important to retain the term related to $dv^2$ in the Taylor expansion because of the peculiar behavior of the Wiener process. This will results into the Itô's lemma for a function of a stochastic process driven by random forcing.

$$\underbrace{df = \frac{\partial f}{\partial v} dv + \frac{\partial f}{\partial t} dt + \frac{C^2}{2} \frac{\partial^2 f}{\partial v^2} dt}_{\text{Itô's lemma}} \tag{3.8}$$

Inserting the stochastic process for $dv$ in Itô's lemma and taking the conditional expectation (see section 4.3) will average out the noise. Since, the function $f$ has compact support therefore integrating once and computing the expected values:

$$\mathbb{E}(\int df = \int \frac{\partial f}{\partial v}(-\gamma v dt + C\zeta dt) + \frac{\partial f}{\partial t} dt + \frac{C^2}{2} \frac{\partial^2 f}{\partial v^2} dt) \tag{3.9}$$

$$\mathbb{E}(\zeta dt) = \mathbb{E}(dW) = 0 \tag{3.10}$$

Results in the Fokker-Planck equation:

$$\frac{\partial p}{\partial t} = \frac{\partial p \gamma v}{\partial v} + \frac{C^2}{2}\frac{\partial^2 p}{\partial v^2} \tag{3.11}$$

If '$\gamma v$' is a constant quantity then it can be pulled out of the derivative:

$$\frac{\partial p}{\partial t} = \underbrace{\gamma v}_{\mu}\frac{\partial p}{\partial v} + \underbrace{\frac{C^2}{2}}_{\sigma}\frac{\partial^2 p}{\partial v^2} \tag{3.12}$$

The Fokker-Planck equation can be written both for the evolution of point density and for the case in which the probability '$p$' entails all the trajectories of the fluid packets. For the latter case a vectored Langevin equation can be written:

$$\dot{\mathbf{v}} = \underbrace{\mu}_{\text{Drift}} + \underbrace{\sqrt{2\sigma}}_{\text{Diffusion}}\ \zeta \tag{3.13}$$

$$\tag{3.14}$$

In the Fokker Planck equation '$\mu$' and '$\sigma$' are the drift and diffusive parameters of the probability current.

If the trajectories described by equation 3.13 in all the directions are uncoupled with each other, then an isotropic field is generated, and a relatively simpler isotropic picture can be presented. However, the coupling between the trajectories entails the anisotropic nature of the field. The diffusion tensor governs the degree of coupling between the trajectories.

### 3.2.4 Advection-diffusion equation–a type of Fokker Planck equivation

The advection-diffusion equation describes the dynamics of the scalar field such as temperature (see eq. 3.15 ), which is a second-order partial differential equation since the temperature is defined in terms of '*entropy*' where entropy is a measure of randomness of the system ac-

counting the number of microstates occupied by the system (which increases rapidly with the energy)[47]. Therefore, the advection-diffusion equation, in other words, represents the dynamics of the probability density function of the occupied microstates. Hence in its most generic form, the advection-diffusion equation can be interpreted as a Fokker-Planck equation.

$$\underbrace{\frac{\partial \theta}{\partial t} + (\mathbf{u} \cdot \nabla)\theta = \kappa \nabla^2 \theta}_{\text{Advection-Diffusion-equation}} \qquad \equiv \underbrace{\frac{\partial p}{\partial t} + (\mu \cdot \nabla)p = \frac{1}{2}\sigma^2 \nabla^2 p}_{\text{Fokker-Planck-equation}} \tag{3.15}$$

Corresponding to a Fokker-Planck equation a Langevin equation can always be written for the fluctuations of the temperature. This equivalence of Fokker-Planck and advection diffusion allows one to write a Langevin equation for the temperature fluctuations or for the time trajectories of the temperature fluctuations.

The above section has introduced two simple statistical models and presented the idea as to how a large dynamical system reduces to these simple local Lagrangian equations. In the subsequent sections, the models (for the velocity and temperature fluctuations) presented in this section were projected on the high-resolution datasets of channel geometries. At first, a relatively simple channel flow data was used subsequently followed by a more complex mixed-convection flow data on which both the velocity and temperature fluctuations were constrained. Since the local trajectories of the fluid variables are statistical with multiple and complicated interactions happening in them. Therefore, at first certain aspects and justifications of building the point models of the fluctuations are need to be considered. Therefore, the basics of isotropy, homogeneity, and the randomness of the flow field are discussed next.

## 3.3 Kolmogorov's & Taylor's hypothesis-local isotropy, randomness of energy cascade & stationarity

In the year 1941 Kolmogorov used the picture presented by Richardson or the so-called Richardson cascade to present a phenomenological picture of turbulence. According to Richardson's cascade–the eddies of size '$l$' will shrink according to a certain law in a successive generation. That is a simple series of the sort can be written:

$$\underbrace{l = l_0 r^n}_{\text{scaling-of-eddies}} \tag{3.16}$$

$$n = 0, 1, 2, \ldots \tag{3.17}$$

.

The exact choice of '$r$' in this series is somewhat arbitrary between $0 < r < 1$ however, according to Frisch[32] this is set to 0.5. Therefore, one can easily see that it represents a simple scaling law for eddies of size (or scale) '$l$.' Kolmogorov in his theory tied the size or the scales of these eddies with the velocity increments (or the fluctuations) which is justified by considering two points on a curve for which the relative velocity (or velocity increments) will only make sense. In his theory, he presented the argument that all the symmetries of the flow will get destroyed at sufficiently high Reynolds number and away from the boundaries. This means that at '*certain scale*' the flow has no knowledge of the production mechanism producing it or the integral scale (or the relevant length scale of the domain.). They are locally isotropic and universal (i.e., the same phenomena can be observed in a star and faucét). Here, the small scale isotropy is the property of velocity increments (or the relative velocity or the velocity tied to a particular size of eddy). Local isotropy means that the flow is invariant from the rotational perspective.

In the limit of high Reynolds number the properties of the small scale velocity (or the velocity increments) are only defined by the local scale ('$l$') and the mean energy dissipation rate ('$\epsilon$') which means that the viscosity will become irrelevant and will only come into

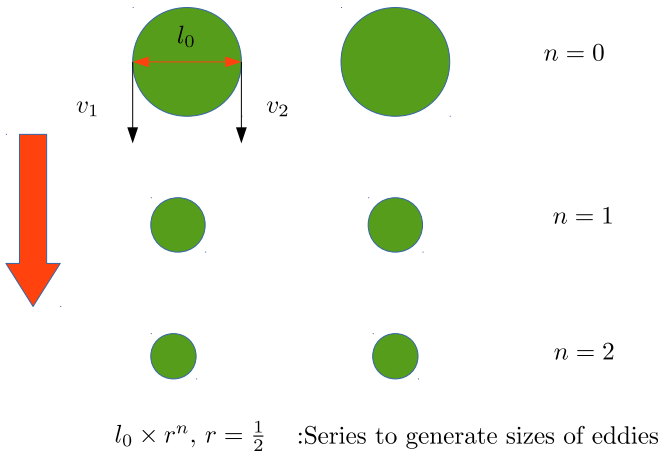



**Figure 3.2**: *Richardson cascade, n is the generation number.*

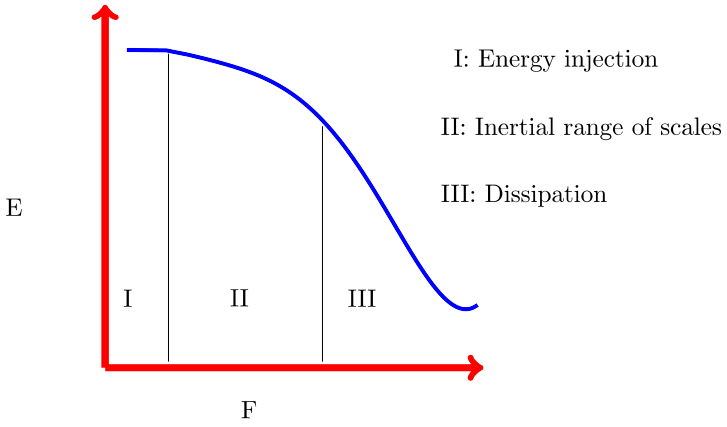



**Figure 3.3**: *Turbulence is random in the inertial range of scales and can be approximated via a Gaussian process.*

play at smaller Reynolds number. It implies that at certain scale this complex process of turbulence is random in nature or memoryless–and it only happens in the so-called inertial range of scales. This idea was successfully exploited in the study of Friedrich & Peinke[48] who demonstrated that it is possible to fit a 'Gaussian' model for the inertial range of scales.

The natural question arises as to how good a time-domain analysis can capture the spatial structures. This can be answered by resorting to the frozen turbulence hypothesis of Taylor.

Under this hypothesis, spatial measurements of fluctuations are possible from a single probe measurement[49] (applicable if the mean flow velocity is much larger than the fluctuating velocity). Owing to this, the spatial correlations can be approximated via the temporal correlations. This is apparent from the the expression for autocorrelation function measured typically by moving a probe at a certain speed 'V' through the flow[49]. In the limit when the probe is moving at a very high speed, then with the temporal measurements it is possible to extract the spatial correlations or alternatively from the single probe measurements if the mean flow is at a much higher speed than the fluctuations. Hence, the measurement at single point can reveal all the spatial information.

Under these two hypotheses, it is possible to build simple statistical models for small increments in the inertial range as they are locally isotropic, stationary and memoryless.

## 3.4 Surrogate modeling

To generate a surrogate model, the coefficients of the Langevin equation need to be computed from the experimentally measured data or high fidelity simulation results. This is done by estimating the conditional probabilities and estimating the various limiting moments of the data as described in the following subsection.

### 3.4.1 Non-linear Langevin equation

Previously, several statistical models for turbulence were proposed by the researchers (see references[9;50]). Based on an experiment by Warhaft[50;51] on a line source in grid turbulence Pope et al.[50] proposed a simple linear 'Markov' model for scalar turbulence. An alternative route to Pope's model of turbulent diffusion is attempted here. A very generic non-linear statistical surrogate for the velocity is built by resorting to the 'Kramers-Moyal expansion' proposed in the statistical mechanics literature[10;48]. The non-linearity in the model comes from the higher order dependence of the coefficients of Langevin equation on the variable under consideration. This means the parameters of the model evolves in time. Very generi-

cally a non-linear Langevin equation (with multiplicative noise) can be written for a generic variable '$y$' which can take any specific form–velocity, temperature fluctuations etc.

$$\underbrace{\dot{y} = \mu(y,t) + \sigma(y,t)\zeta}_{\text{Non-linear-Langevin-equation}} \tag{3.18}$$

### 3.4.2 Parameters of non-linear Langevin equation

The fundamental question which could naturally arise:– is to how to compute the parameters of the Langevin model? For a linear Langevin equation, this is a straightforward task[10] however, for a non-linear Langevin equation, it can be a challenging issue. For a non-linear Langevin equation, the parameters of the model can be computed by first computing the spatio-temporal distribution of the variable governed by the Langevin equation. Once the distribution function of the stochastic variable is known, the expectation values, which are also the parameters of the non-linear Langevin equation can be computed. The relationship between the parameters or the expectation values and the distribution function of the Langevin variable can be found by resorting to the Kramers-Moyal expansion method. If the Kramers-Moyal expansion stops at the second term, then the Fokker-Planck equation results. Therefore, the Kramers-Moyal expansion not only leads to a Fokker-Planck equation but provides a tractable way of computing the parameters of the non-linear Langevin equation. This expansion can be derived using the method presented in the following subsection.

### 3.4.3 The Kramers-Moyal expansion

The Kramers-Moyal expansion for the transition of probability current can be written by following the analysis presented in Risken[10]. Starting from the definition of conditional probability:

$$P(y, t+\tau|y', t) = \int \delta(y'' - y) P(y'', t+\tau|, y', t) dy'' \tag{3.19}$$

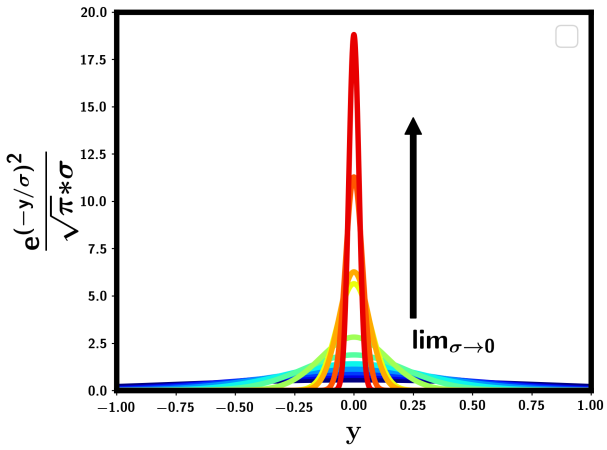



**Figure 3.4**: *Delta function can be approximated by the limiting Gaussian distribution. The derivative of delta function can then be computed.*

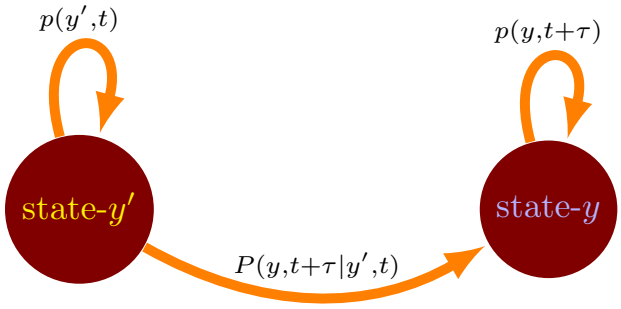



**Figure 3.5**: *A Markov chain represented by the equation 3.22*

where the '$\delta(y''-y)$' is the probability of being in state '$y$'. This particular integral is also known as the shifting property of the dirac delta function. It ensures that the sum on the right handside of the equation 3.19 reduces to the left hand side which means that all the intermediate states were avoided. Now expanding the '$\delta(y''-y)$' function in Taylor series and inserting in equation 3.19 results in,

$$P(y,t+\tau|y',t) = \int (\sum_{n=0}^{\infty} \frac{(y''-y')^n}{n!} \frac{\partial^n}{\partial y^n} \delta(y'-y)) P(y'',t+\tau|y',t) dy'' \tag{3.20}$$

$$P(y,t+\tau|y',t) = \left( (1 + \sum_{n=1}^{\infty} \frac{\partial^n}{\partial y^n} \int \underbrace{\frac{1}{n!}(y''-y')^n P(y'',t+\tau|y',t)dy''}_{\text{Moments:-}M^n(y',t,\tau)}) \delta(y'-y) \right) \tag{3.21}$$

Inserting the resulting equation 3.21 in equation 3.22:

$$p(y,t+\tau) = \int \underbrace{P(y,t+\tau|y',t)}_{\text{state transition probability}} \quad p(y',t)dy' \tag{3.22}$$

$$p(y,t+\tau) = \int \left( (1 + \sum_{n=1}^{\infty} \frac{\partial^n}{\partial y^n} \int M^n(y',t,\tau))\delta(y'-y) \right) p(y',t)dy' \tag{3.23}$$

$$\tag{3.24}$$

The delta function will simplify the integrals in the above equations

$$p(y,t+\tau) - p(y,t) = \sum_{n=1}^{\infty} \frac{\partial^n}{\partial y^n} M^n(y,t,\tau)p(y,t) \tag{3.25}$$

Dividing by $\tau$ and setting limit to 0

$$\lim_{\tau\to 0} \frac{p(y,t+\tau) - p(y,t)}{\tau} = \lim_{\tau\to 0} \frac{1}{\tau} \sum_{n=1}^{\infty} \frac{\partial^n}{\partial y^n} M^n(y,t,\tau)p(y,t) \tag{3.26}$$

$$\lim_{\tau\to 0} \frac{1}{\tau} M^n(y,t,\tau) = D^n(y,t) \tag{3.27}$$

$$\underbrace{\frac{\partial p}{\partial t} = \sum_{n=1}^{\infty} \frac{\partial^n}{\partial y^n} D^n(y,t)p(y,t)}_{\text{The Kramers-Moyal expansion}} \tag{3.28}$$

$$\underbrace{\frac{\partial p}{\partial t} = \frac{\partial}{\partial y} D^1(y,t)p(y,t) + \frac{\partial^2}{\partial y^2} D^2(y,t)p(y,t)}_{\text{The Fokker-Planck equation}} \tag{3.29}$$

$$D^1(y,t) = \mu(y,t) \tag{3.30}$$

$$D^2(y,t) = \frac{1}{2}\sigma^2(y,t) \tag{3.31}$$

If the Kramers-Moyal expansion stops at $2^{nd}$ term then the Fokker-Planck equation results (see equation 3.29) and the Langevin equation can be written (see equation 3.18). Using the

relations 3.30 and 3.31 the equivalent Langevin equation is (see equation 3.32):

$$\dot{y} = D^1 + \sqrt{2D^2}\zeta \tag{3.32}$$

This analysis also reveals that the first two limiting values of the moment are the drift and diffusion parameters of the Fokker-Planck equation and/or Langevin equation (see equation 3.31).

### 3.4.4 Data learning: Estimating moments

From equation 3.21 it is quite clear that one of the key ingredients of this method is the conditional probability ($P(y, t+\tau)|y', t)$) which is required to compute the conditional moments of the data. This can be computed by drawing a vertical (or horizontal) cut in the joint probability distribution function (see figure 3.6). The joint probability can be computed

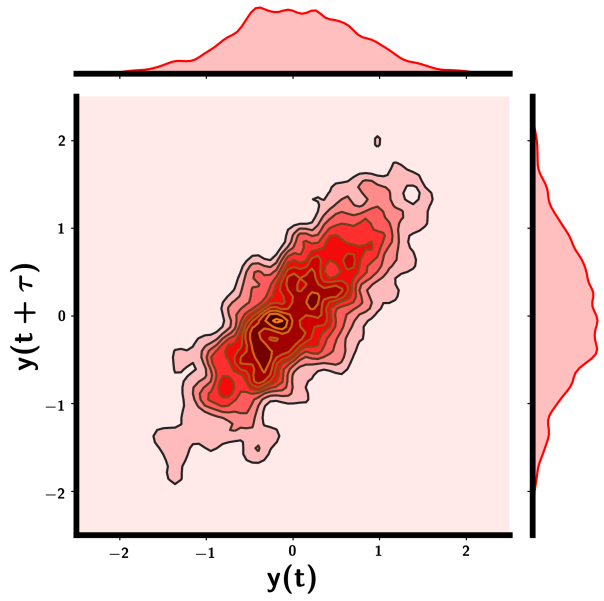



**Figure 3.6**: *Joint probability density–a vertical cut in it reveals the conditional probability.*

via a shifted copy of the signal itself (see figure 3.7). This shift in the signal is generated via binning the data and counting the next state achieved by the signal for each fixed state for the subsequent higher $\tau$ (before this pre-processing of data was done: see appendix for

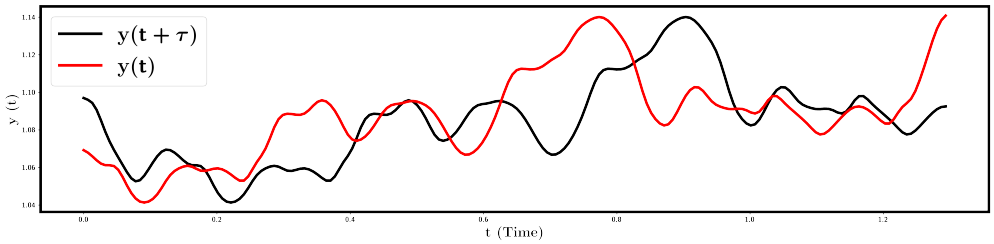



**Figure 3.7**: *Shifted signal.*

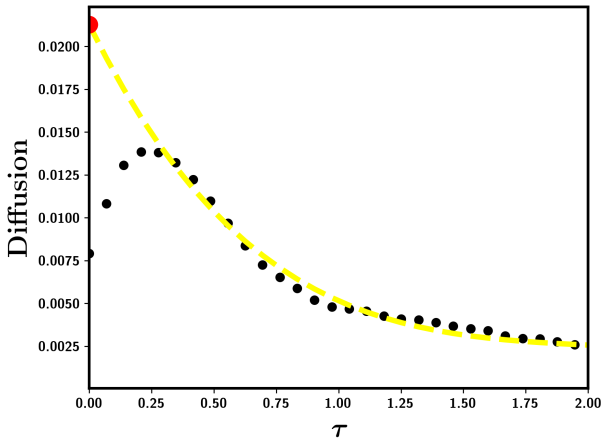



**Figure 3.8**: *Fitted and extrapolated $D^2$ coefficient.*

details A). The signal binning method reveals the full joint probability and the conditional probabilities too. Once the conditional probability is revealed, then the conditional moments were evaluated and the limiting value of the moments were computed by extrapolating the moments to $\tau \to 0$. These conditional moments were evaluated for all the binning points of the series for different lags ($\tau$). Then a simple regression of computed moments versus the binning points reveals the dependence of drift and diffusion coefficients on the signal itself (see figure 3.8). A very important step in the estimation of moments is to ensure that the 'Markovian' properties are always met, the jump has to be big enough. In the case of turbulence Peinke et al. found in a variety of experiments[48;52;53] that Taylor microscale is the minimum length above which the turbulence shows memoryless behavior since at that point inertial length scales starts and the cascade is known to show memoryless behavior.

## 3.5 Wall-bounded turbulent flow problem: Channel flow

For this first case, data was extracted from a publicly available turbulence database of the John Hopkins University[54;55]. This data was generated by DNS of Navier-Stokes equations in a channel domain, the description of which is presented in the following subsection.

### 3.5.1 Description of domain and flow conditions

The size of the channel domain is $8\pi h, 2h, 3\pi h$ ($h = 1$ is half channel height) discretized in $2048 \times 512 \times 1536$ nodes. The Navier Stokes equation was integrated over this domain (see figure 3.9) with 4 periodic and 2 no slip boundary condition via a Fourier pseudospectral code PoongBack using the non-dimensional parameters listed in the table 3.1. The velocity signal was probed at the center of the channel i.e., at location $x = 4\pi h, z = 1.5\pi h$ in the wall-normal direction i.e., at various positions starting from near the wall to the center of the channel.

**Table 3.1**: *Non-dimensional parameters used in the channel flow simulation.*

| $\mathrm{Re}_b$ | $\mathrm{Re}_\tau$ | $\mathrm{U}_\tau$ | $\mathrm{U}_b$ | $\nu$ |
|---|---|---|---|---|
| $3.998\times10^{4}$ | $9.9935\times10^{2}$ | $4.9968\times10^{-2}$ | 0.99994 | $5\times10^{-5}$ |

50 different time series were trained in those locations. Drift and diffusion parameters were extracted from the various probed signals. The diffusion coefficient was plotted vs the fluctuation scales. This analysis indicates that the diffusive coefficient keep on decreasing as one goes away from the wall. This means that the diffusivity is related to the Reynolds stress term which shows similar characteristics (see figure 3.13, 3.12).

### 3.5.2 Near wall structures

Wall bounded turbulent flow are known to have large scale structures near the wall region. This can be seen via the dependence of wavelength and Reynolds stress term. Since the

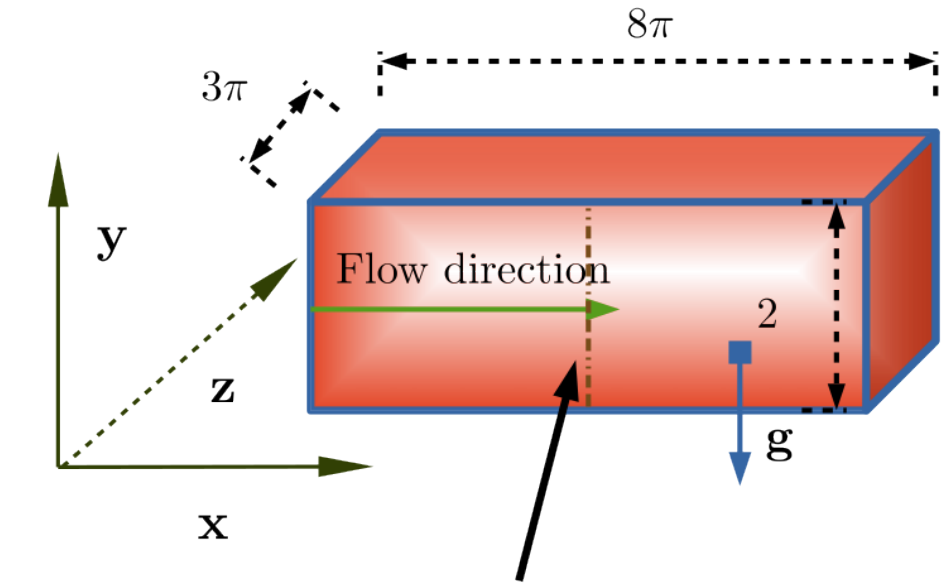



**Figure 3.9**: *Domain description of the channel flow problem (taken from John Hopkins Turbulence Database (https://www.jhu.edu/)).*

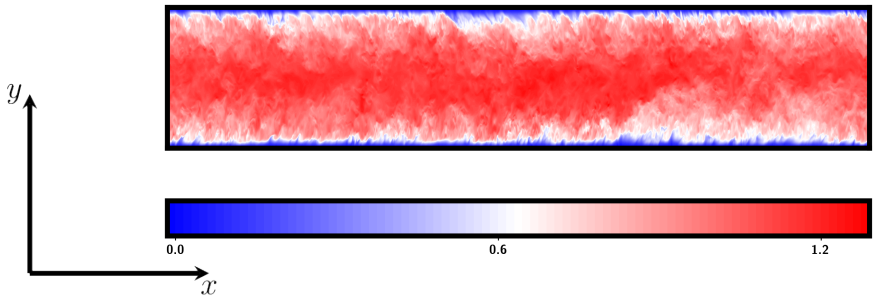



**Figure 3.10**: *Streamwise velocity at* $\mathbf{z = 0}$ *plane (data taken from John Hopkins Turbulence Database).*

diffusion has a similar behavior as that of the Reynolds stress term, then it will be interesting to see how the energy and wavelength translate in relation to the second Kramers-Moyal coefficient. Since the presence of longer wavelength near the wall region indicates the presence

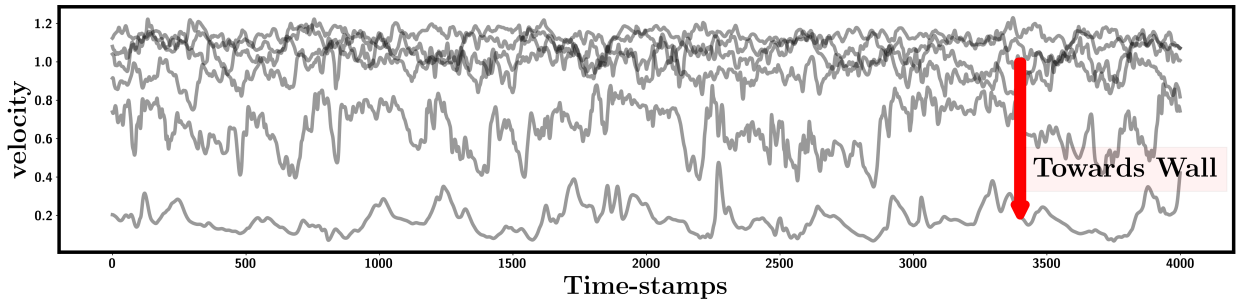



**Figure 3.11**: *Wall effect: near the wall time series show less randomness and more organized behavior. However, turbulence introduces a meandering (or a zig zag) effect in them.*

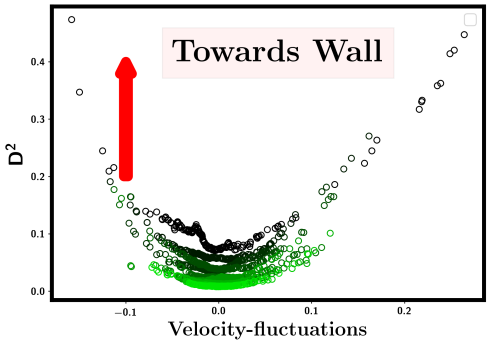



**Figure 3.12**: *Diffusion coefficient ($\mathbf{D^2}$) from center to wall.*

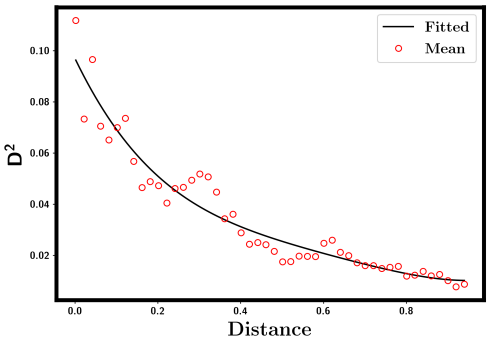



**Figure 3.13**: *Mean diffusion coefficient vs wall normal distance.*

of large scale structures which can be most conveniently analyzed by plotting a contour map of premultiplied and normalized energy spectra obtained at each wall-normal position. It will be interesting to see how the convolved Langevin model performs near the wall region in terms of capturing these large scale structures since in our understanding it tends to perform well away from the wall and not near it. Nevertheless, the spectral trend in the premultiplied spectra is well produced. However, the high energy peak occurs at a lower wavelength (see figures 3.14,3.15). The possible reason for that is the model is not able to work with the slow-moving part of the velocity signal, and the decay is much more rapid in Langevin model with an appropriate kick introduced by the random nature of the forcing. The peak itself is slightly more pronounced than the DNS data–possible reason can be the totally random kicks at short time scales which predicts a slightly higher amplitude of velocity.

The contour maps of premultiplied energy spectra show the same trend that the DNS

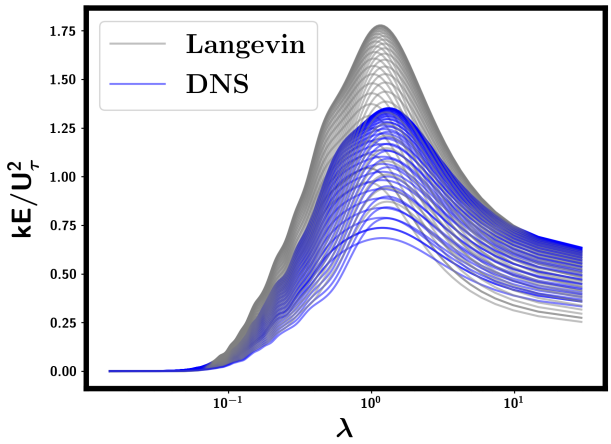



**Figure 3.14**: *Pre-multiplied energy spectra vs wavelength of Langevin and DNS singal. Here* $\lambda = \frac{2\pi}{\mathbf{k}}$ *where* $\mathbf{k}$ *is the wave number defined as:* $\frac{2\pi \mathbf{f}}{\mathbf{U}}$ *and* $\mathbf{U}$ *is the mean velocity.*

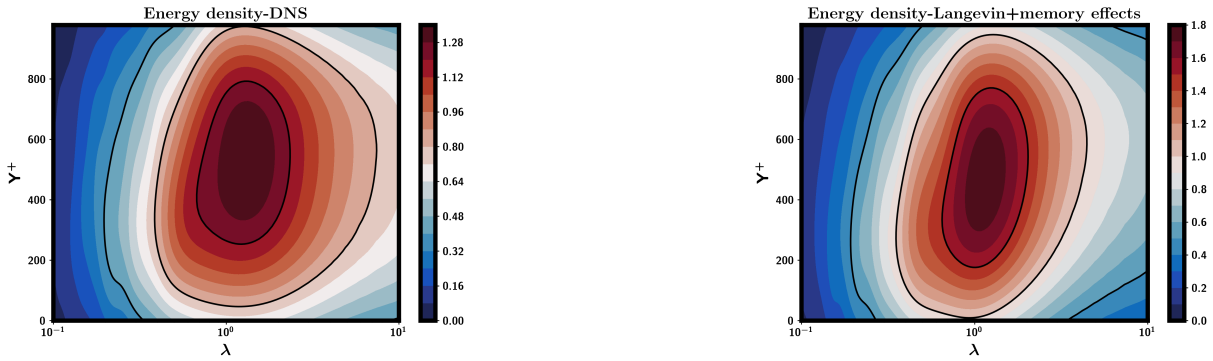



**Figure 3.15**: *Contour map of premultiplied energy spectra DNS signal (left), Langevin signal (right). Here* $\mathbf{y}^+ = \frac{\mathbf{y}}{\delta_\nu}$ *where* $\delta_\nu$ *is the viscous length scale defined as:*$\frac{\nu}{\mathbf{U}_\tau}$.

signal is of higher wavelength near the wall while the Langevin model has a lower wavelength. Thus, the trained Langevin model is not able to reproduce the near-wall behavior correctly. Nevertheless, this model has performed reasonably well in a variety of locations with the very correct description of the statistics away from the wall. This shows that the model can capture the correct physics of the turbulent system. Finally, the relationship between the premultiplied spectra and the diffusion coefficient was plotted, and it indicates that the diffusion varies with the energy in a quite linear fashion on a log-log plot and peaks at the location where the energy is maximum. The peaking of energy and the diffusion coefficient near the wall region is also indicative of a near-wall large scale turbulent structure.

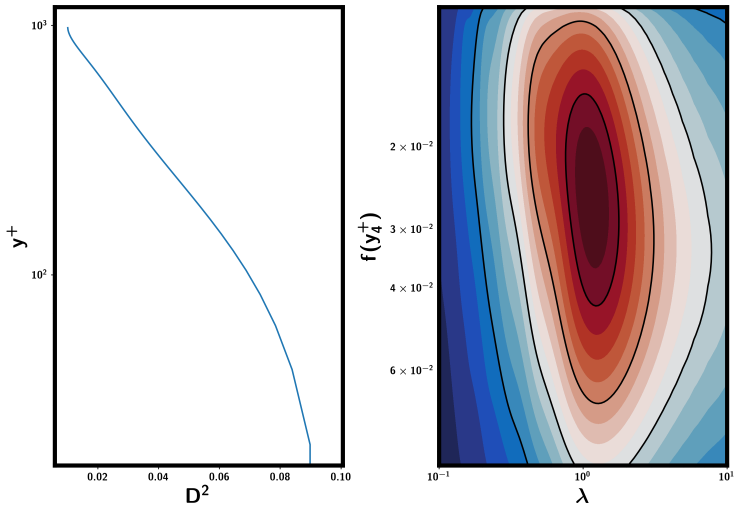



**Figure 3.16**: *Contour map of premultiplied energy spectra–Langevin signal. Wall normal distance is replaced by the fourth-order fitted polynomial (*$\mathbf{f}(\mathbf{y_4^+})$*) of the same–which represents mean diffusion coefficient (*$\mathbf{D^2}$*).*

Before making and analyzing the contour maps presented above the statistical parlance between the true DNS signal and the trained Langevin model-generated time series was maximized by passing the signal through the Gaussian convolution filter. Since, the Fourier transform of a Gaussian (see equations 3.33, 3.34) is also a Gaussian with the variance of two Gaussian related by the

$$g(x) = \frac{1}{\sqrt{2\sigma}} e^{-(\frac{x^2}{2\sigma^2})} \tag{3.33}$$

$$H(f) = e^{-\frac{f^2}{2\sigma_f^2}} \tag{3.34}$$

$$\sigma\sigma_f = \frac{1}{2\pi} \tag{3.35}$$

expression 3.35. Where $\sigma_f$ is the cutoff frequency in Fourier space. This cutoff frequency is set equal to the frequency corresponding to the Taylor microscale ($\lambda_\tau$). The Taylor microscale was conveniently computed via the osculating parabola method (see figure 3.18).

In the next section, a more complicated heated channel flow case was analyzed. Since the statistics of the cascade are universal in nature ( attributed to Kolmogorov[32]) therefore the model should replicate the statistics of the turbulent signal obtained in the heated channel flow.

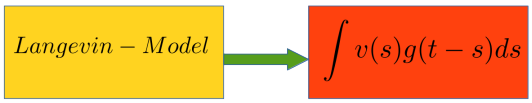



**Figure 3.17**: *Convolution of Langevin model with Gaussian kernel–smoothing out the signal output. This kernel is responsible for introducing an artificial viscosity. The length of filter is proportional to the Taylor microscale.*

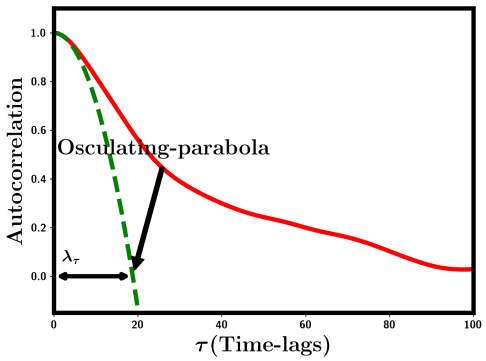



**Figure 3.18**: *Taylor microscale computed via osculating parabola method.*

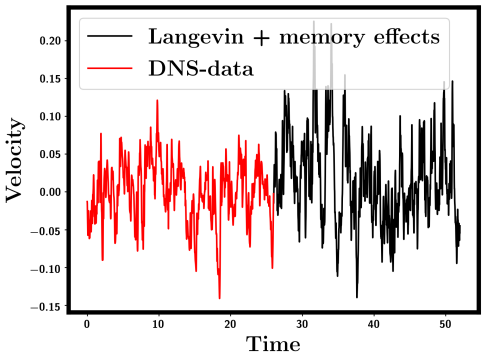



**Figure 3.19**: *Comparison of the Langevin model with memory effects and the DNS-data. The signal of the convolved Langevin model shows remarkable similarity with the true DNS signal.*

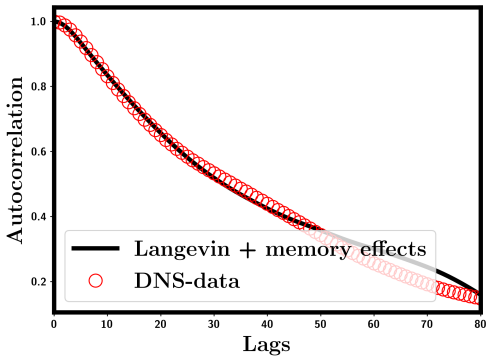



**Figure 3.20**: *Comparison of the Langevin model with memory effects and the DNS-data. The correlation function of the model are now remarkably similar to the true DNS signal.*

## 3.6 Scalar transport with mixed convection: Liquid metal channel flow

The DNS data was generated by using Incompact-3d[56] by the computational fluid dynamics group at PSI, Switzerland by considering the flow configuration in a horizontal channel discretized in $512 \times 257 \times 512$ grid points (see figure 3.21 and 3.22) with the parameters listed in the table 3.2. This flow configuration is used to study the mixed convection flow regime for liquid lead with Prandtl number 0.025. The top and the bottom walls of the channel are heated uniformly at a certain temperature. The cold wall is on the top with a lower temperature ($T_c$), and the hot wall is on the bottom and has a relatively higher temperature ($T_h$). The buoyancy force is induced throughout the flow field due to the temperature difference $\Delta T_{hc} = (T_h - T_c)$ between the top wall and the bottom wall (the acceleration acts downward along y-direction because of gravity). The streamwise, wall-normal and spanwise coordinates are represented by $x$, $y$ and $z$ respectively. Liquid lead is driven by an imposed pressure gradient flowing through the channel. Hence, this flow can provide a good starting point for building a statistical surrogate for the mixing phenomena.

In this flow configuration, the velocity and temperature signals were probed at the $50^{th}$ grid point near the bottom hot plate (see figure 3.21). This point was selected in the computational domain since the effect of mixing is more pronounced near the bottom hot plate. The following figures 3.23, 3.24 shows the probed turbulent velocity signal together with the spectrum of it. The spectrum of the signal is important to look at as any true emulator should be able to follow similar trends in the Fourier space. Since in this work, the scales above the Taylor microscales were separated thus, the applicability of statistical surrogate is till the end of turbulent energy cascading region.

**Table 3.2**: *Non-dimensional parameters used in the differentially heated channel flow simulation.*

| $\mathrm{Re}_b$ | $\mathrm{Re}_\tau$ | $\mathrm{U}_\tau$ | $\mathrm{U}_b$ | $\nu$ |
|---|---|---|---|---|
| $4.667\times10^3$ | $1.96\times10^2$ | 0.0576 | 0.68573 | 0.0002941 |

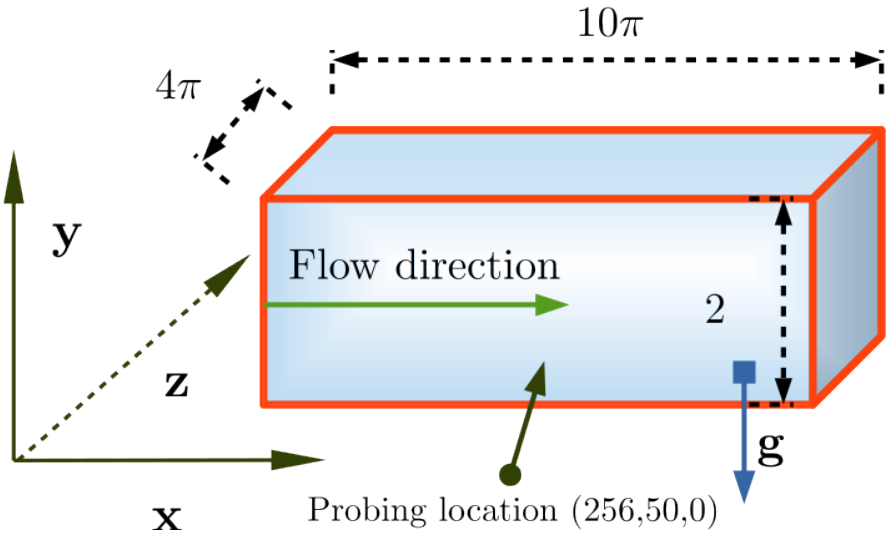



**Figure 3.21**: *Domain description of the differentially heated liquid metal channel flow problem (data courtesy: computational fluid dynamics group at Paul Scherrer Institute (https://www.psi.ch/en)).*

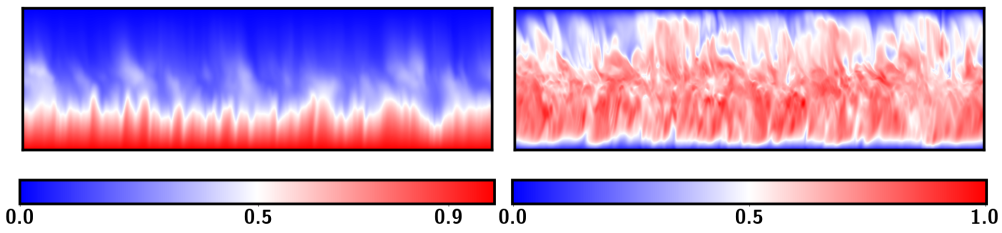



**Figure 3.22**: *Streamwise velocity (right) and temperature (left) plot at the* $\mathbf{z} = \mathbf{0}$ *plane (data courtesy: computational fluid dynamics group at Paul Scherrer Institute).*

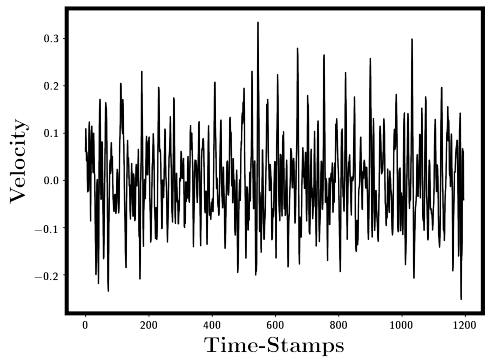



**Figure 3.23**: *Fluctuations of probed velocity time series.*

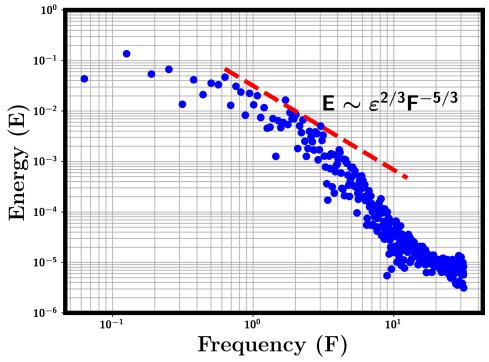



**Figure 3.24**: *One dimensional spectra of probed velocity time series.*

### 3.6.1 Results & discussion: Liquid metal channel flow

The following section presents the essential results extracted from the probed velocity and temperature signals at the location described in the figure 3.21.

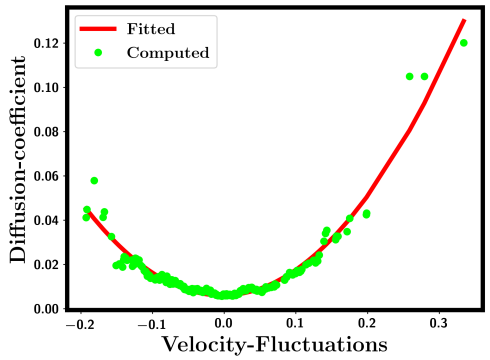



**Figure 3.25**: *Quadratic diffusion.*

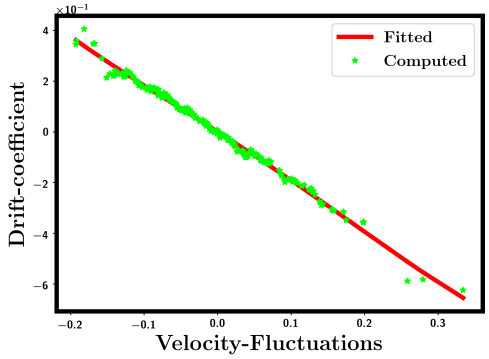



**Figure 3.26**: *Linear drift with minor third-order corrections.*

### 3.6.2 Velocity data and surrogate

The Kramers-Moyal analysis reveals that in order to fit the data to Langevin model for the velocity fluctuations it is essential to compute the various limiting values of the moments of the data which were computed as mentioned in the section 3.4.4. The computed first moment tends to follow a linear trend with the velocity while a quadratic trend was there in the second moment (see figures 3.25 and 3.26). The higher-order moments of the data,

especially the fourth moment, is vanishingly small compared to the second one. This then ensures that the Kramers-Moyal expansion stops at the second term and the Fokker-Planck equation or the so-called advection-diffusion equation results.

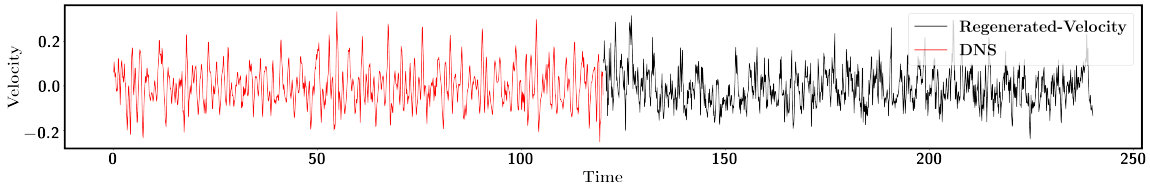



**Figure 3.27**: *Generated surrogate from the time series captured in DNS data–one series shifted in time from the other.*

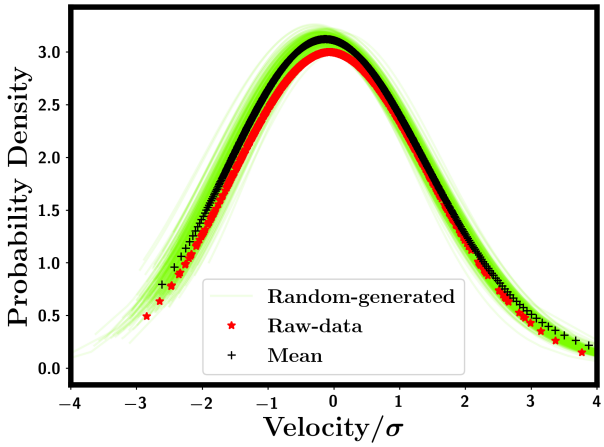



**Figure 3.28**: *PDFs of the velocity signal. In red is the DNS data and in black is the velocity surrogate. The green band around the pdf's represent the pdf of multiple trajectories.*

The fitted coefficients can then be fed back to the Langevin equation for velocity, and a surrogate data can be generated. It can be seen from the above results that the fitted model for the velocity was able to capture the essentials of the velocity statistics quite well. The remarkable accuracy of the goodness of fit of the two pdf's can be appreciated by the very low value of the computed KL-divergence measure (see figure 3.29) between the two pdfs. This model can hence be treated as a point surrogate of the velocity signal[57–60].

It is to be stressed that the Langevin equation presented here is more generic than the otherwise linear 'Markov' model and may reveal more complex features of the turbulent signal.

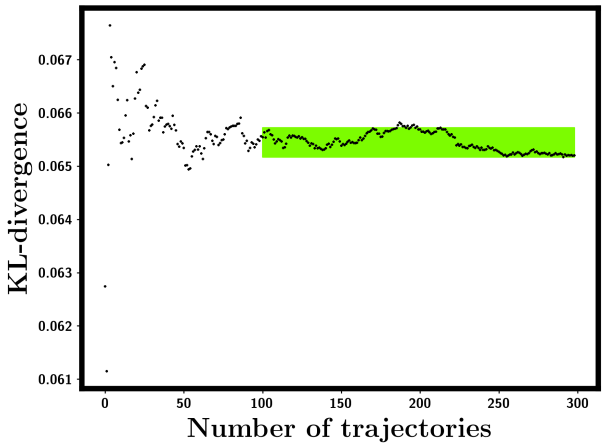



**Figure 3.29**: *KL divergence tells how well the probability distribution 'n' approximates the distribution'm.' If the KL divergence (*$\mathbf{D_{KL}(m||n) = \int m(y)\ ln\frac{m(y)}{n(y)}dy}$*, 'm' is the target pdf and 'n' is the computed pdf) is zero then the computed pdf is exactly similar to the target pdf. The green band around 100 mark signifies no serious gain of information after running 100 trajectories.*

### 3.6.3 Temperature data and surrogate

While velocity is a very important variable for building a generic statistical scalar mixing model. Nevertheless, the ultimate goal of any statistical mixing modeling approach is to understand as to how the temperature is getting advected from the point of its origin. Thus, it is important to understand the scalar itself and to point out the important differences between the behavior of the scalar and temperature time series. Few natural questions arise– does the scalar time series obeys the 'Markov' process. From a visual and spectral inspection point, the scalar time series appears to have the response which is confined to the slower frequencies than the velocity time series–thus it can be better described by a Markov process at the same time scale. It can be seen from the temperature time series that the

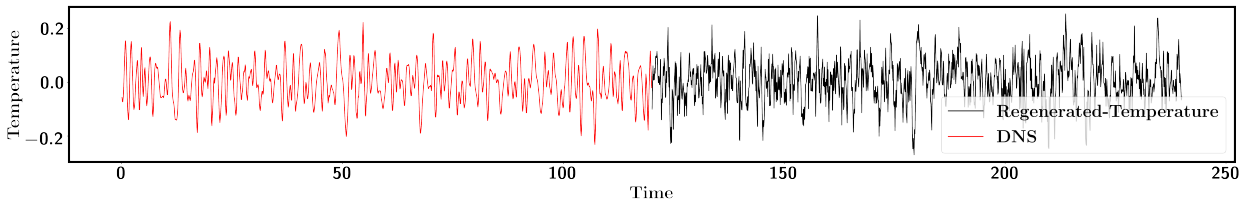



**Figure 3.30**: *Generated surrogate from the time series captured in the DNS data–one series shifted in time from the other.*

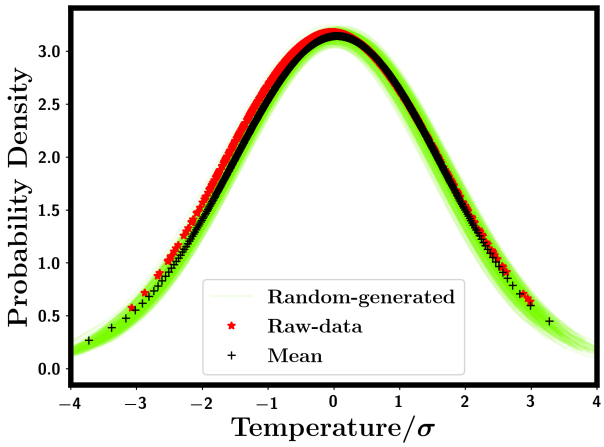



**Figure 3.31**: *PDFs of the temperature signal. In red is the DNS data and in black is the velocity surrogate. The green band around the pdf's represent the pdf of multiple trajectories.*

pdf of the fitted and the raw data matches quite well. Thus, at the same time scale '$\tau$', temperature time series is more 'Markovian' than the velocity time series. In other words temperature fluctuations are more sluggish as compared to the velocity fluctuations.

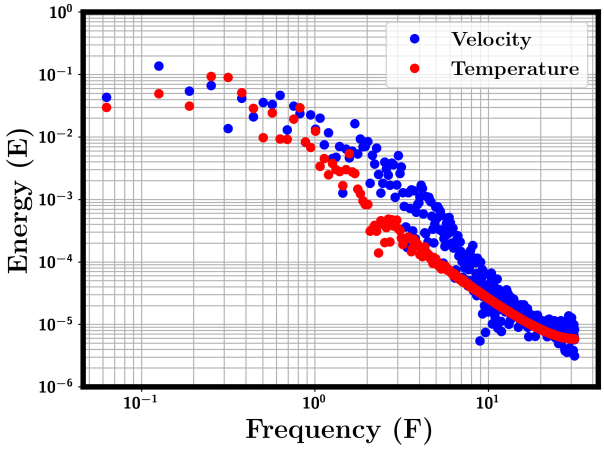



**Figure 3.32**: *Combined spectrum of velocity and temperature data.*

This point can be seen by drawing a combined spectrum (see figure 3.32) of the probed velocity and temperature time series. The spectrum clearly shows that the velocity time series has slightly high-frequency activities when compared to its temperature counterpart. This is due to the low Prandtl number (liquid Lead), which causes the reduction in the amplitude of the temperature fluctuations resulting in the distribution of the energy within

larger length scales[61].

## 3.7 Conclusions

A Langevin equation-based statistical surrogate was developed by probing the velocity and temperature signals in the DNS dataset of the channel flow and differentially heated liquid metal channel flow. The drift and diffusion coefficients of the Langevin equation were obtained with the help of Kramers-Moyal expansion, capturing the markovian processes. The statistical emulator was able to capture the essential details of the turbulent signal and shows good statistical agreement with the raw DNS fluctuations. Some interesting differences between the velocity and temperature time series at the same time scales show that the match for the temperature fluctuation is much better. This can be attributed to the low Prandtl number of the liquid Lead simulated here. The resulting emulator can be coupled to a system scale model and or coupled to the statistical equation for tracking the position ($x$) of the particle which can subsequently be used to estimate the scalar field fluctuations or mixing extent under turbulent flow conditions. These type of statistical emulators can significantly improve the accurate assessment of uncertainties and their propagation. Future work will need improvement of these emulators to capture the full spectrum i.e., also include non-markovian processes which can be achieved by developing emulators based on Generalized Langevin Equation[62].

# Chapter 4

# Uncertainty quantification of the contaminant dispersal in the environment

In this chapter, an application case study with linear 'Markov' model is presented. The data is not real, but the considered problem is genuinely practical where the impact of radioactive particle (or a passive scalar) release in the environment is considered by a linear 'Markov' model with fixed parameters. This is true for the case where a homogeneous velocity field exists in the whole domain i.e., at each point on the trajectory of the fluid a constant '$x$' component of the velocity exists. While the uncertainty at each location is captured by the random forcing term (see figure 4.1)[63]. This is the case presented in the first section of this current chapter where it was assumed that the forcing at each location is random in time–thus describing the location uncertainties by a random white noise term. While this is true for the cases where the correlation of the uncertain term is short-lived. However, this is not the case in scenarios where a random field like turbulent flows drives the particle, and the nature of the correlated (colored) noise may change the computed uncertainties. To this end, a motivating example is presented in the next section in the context of hazardous particle release into a fluid stream.

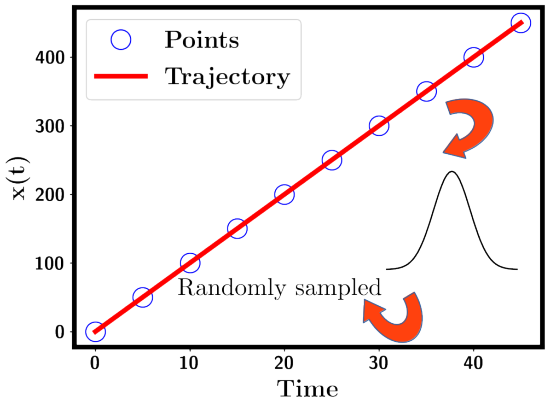



**Figure 4.1**: *Lagrangian points (shown in blue) situated on the Eulerian trajectories $x(t)$ are perturbed by a random forcing sampled from a 'Gaussian' distribution.*

## 4.1 Motivation

Natural calamities generate external events which can severely compromise the safety of the industrial installations which in-turn can result in the release of hazardous particles in the environment, contained otherwise within the various safety barriers or industrial confinements.

For example, an externally initiated event i.e., a powerful tsunami wave triggered a significant collapse of the engineered safety features of the Fukushima Dai-ichi nuclear power station followed by a series of operational issues leading to the failure of the safety barriers or containments. Due to these breaches, radioactivity was released into air, groundwater and onto a wide area of the Pacific Ocean causing a global calamity[64;65]

Water and air can transport any harmful contaminant released into the environment by convective and diffusive mechanisms. The transport of these effluents can be as complex as the motion of air and water in the atmosphere and hydrosphere. Atmospheric winds play a critical role in the dispersal of environmental contaminants upon instantaneous release, and they can cause environmental impact to vast distances. Ocean currents and rivers are no different, which can lead to widespread dispersal of pollutants. Similarly, groundwater is another intricate part of hydrosphere from where the environmental contaminants can spread and interact with biosphere.

The accurate prediction of the dispersal of contaminants through the environment require a good understanding of the governing principles and effective mathematical techniques. The classical scalar or species transport models can be broadly classified in Eulerian and Lagrangian frameworks. Most of the primitive modeling of the contaminant transport is based on population balance approach, which falls under the broad class of Eulerian framework after continuum approximations. This approach is popularly known as the advection-diffusion equation, which is an example of a Fokker-Planck equation. One approach of obtaining closure for these models is by using Navier-Stokes for momentum transport and coupling the velocity field with the advection-diffusion equation. However, direct numerical solutions to these coupled fundamental equations are often challenging or sometimes even impossible particularly when system domain is as complex as groundwater. Moreover, even if there is a way to solve these coupled equations using large-scale processes it will still need a large number of realizations due to the sensitivity of scalar dispersion to flow field, complex flow path possibilities, and local interactions. Based on several experimental and numerical studies on fluid flow in the porous media or packed beds, it can be concluded that for sizeable overall flow areas with uniformity in flow injection, the local flow rates and local velocity magnitudes are almost constant. The models to evaluate dispersal in groundwater systems take into account several experimentally determined parameters, for instance diffusion coefficient, deposition coefficient, reaction rate, retardation coefficient, soil porosity, etc. The ad-hoc advection-diffusion type models which include a simplified approach to treat momentum transport, are useful for practical purposes but are strongly dependent upon two important factors–a) meticulous modeling with large parametric space, b) parametric variance. In order to accurately assess the effect of these parameters and their associated uncertainties, it is required to perform the simulations of the Eulerian models with parametric sweep i.e., performing multiple realizations, which is a computationally-intensive process. The interpretation of these simulation results will further require understanding sensitivity analysis. Another drawback of these continuum-based models is that they were derived based on gross approximations of uniform mass or species distribution in a control volume or flow stream. This assumption particularly breaks down easily where there is a very small number of ef-

fluent particles such as colloidal particles in the groundwater flow stream. The homogeneity assumptions can lead to significant uncertainties in the model development stage itself. The alternative approach of treating a finite number of particles in a volumetric domain or fluid stream is via Lagrangian framework where the trajectories of every particle can be tracked. This Lagrangian or particle-based approach is able to quantify the uncertainties by coupling the statistical quantities to the appropriate macroscopic variables[42;43].

Current work uses direct simulation of Brownian particles for predicting the dispersal of contaminants upon release in the environment (homogeneous) with an artificially imposed advective and diffusion coefficients, which may be deduced rigorously (experimentally or numerically via molecular dynamics simulation)[66]. The resulting solutions to the Brownian particle simulations are then used to construct a concentration profile of contaminants with associated uncertainties. These results are compared with the solutions of convection-diffusion or Fokker Planck equation.

## 4.2 Problem Description

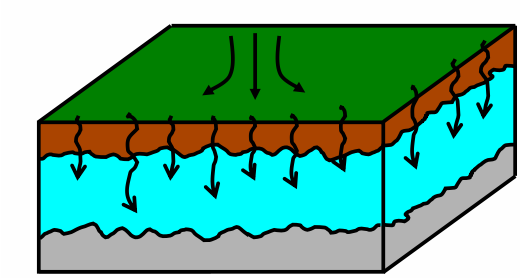


**Figure 4.2**: *Transport of contaminants, here the colors represents* ■ *the top surface,* ■ *the soil just below it,* ■ *water table,* ■ *rocks and minerals, while the arrows in* ■ *represents the torturous path of the released particles.*

To illustrate the aforementioned limitation of modeling capabilities which can result into the uncertainties as a result of continuum level modeling approaches, a simplified representative groundwater media is selected as a very large or infinite domain.

Owing to a sudden release of the contaminant at a surface, the pollutants can convect

and diffuse freely in this semi-infinite groundwater media if there are no additional local kinetics. Transport of the released particles can be studied by the well-known convection-diffusion equation in the mathematical equivalent of the actual geographical domain i.e., in an infinite spatial domain.

$$\frac{\partial C}{\partial t} = \nabla \cdot (D \nabla C) - \mathbf{v} \cdot \nabla \mathbf{C} \tag{4.1}$$

This scalar transport equation can be rewritten in Einsteinian convention as,

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x_i}(u_i C - D\frac{\partial C}{\partial x_i}) \tag{4.2}$$

Defining temperature and velocity as the sum of time-smoothed functions and fluctuating fields,

$$C = \overline{C} + C'; u_i = \overline{u}_i + u_i', \tag{4.3}$$

substituting in the Eqn. 4.2 and time-smoothing leads to

$$\frac{\partial \overline{C}}{\partial t} = -\frac{\partial}{\partial x_i}(\overline{u}_i \overline{C} + \overline{u_i' C'} - D\frac{\partial}{\partial x_i}\overline{C}). \tag{4.4}$$

Typically a new variable, $D_\tau$ (eddy mass diffusivity) is defined.

$$\overline{u_i' C'} = D_\tau \frac{\partial \overline{C}}{\partial x_i}, \tag{4.5}$$

which leads to a simplified time-smoothed transport equation

$$\frac{\partial \overline{C}}{\partial t} = -\frac{\partial}{\partial x_i}(\overline{u_i}\overline{C} - (D_\tau + D)\frac{\partial \overline{C}}{\partial x_i}). \tag{4.6}$$

However, considering the real scenario where large surface area is exposed to an environmental contaminant which can convect or diffuse in the sub-surface region, a 1-D infinite domain model is adequate. Therefore, for an infinite 1-D domain pertaining to a medium with constant drift and diffusion parameters, equation 4.1 will reduce to equation 4.7, with

dominant velocity in x-direction i.e., normal to the surface of contamination. The effective diffusivity $D_e$ is the sum of eddy mass diffusivity $D_\tau$ and molecular diffusivity $D$.

$$\left\{ D_e \frac{\partial^2 \overline{C}}{\partial x^2} - \overline{u} \frac{\partial \overline{C}}{\partial x} = \frac{\partial \overline{C}}{\partial t} \right\} \tag{4.7}$$

$$\left\{ \overline{C}(x,0) = \delta(x) \right\} \tag{4.8}$$

In the above equation, instantaneous fallout at a specific location ($x = 0$) can be thought of as '$\delta$' function. Upon release, the particles spread like a plume in the semi-infinite domain. The above formulation of the problem is a typical approach adopted in previous studies. This methodology completely neglects the uncertainties arising of various reasons and therefore, is only suitable for the case when it is desirable to track the bulk motion of the particles. The described method only leaves the scope of adding uncertainties in an ad-hoc manner in the deterministic Fokker Planck equation[67]. However, when a small population of solid particles are the contaminants, their transport and dispersal in the environment can have a high degree of variations, and the eulerian approach described above may lead to erroneous conclusions. The motion of a few solid particles being transported in complex fluid streams such as in a groundwater system can be highly stochastic due to complex flow path possibilities, Brownian motion and other physical/ chemical interactions with the surrounding media. Therefore, to study the movement of a small population of particles, it is desirable to construct a stochastic counterpart of the convection-diffusion equation. Lagrangian equation of motion in 1-D (x-coordinate) can be written as

$$\frac{dx}{dt} = \bar{u} + u'(t) \tag{4.9}$$

The order of magnitude of the fluctuating velocities can be assumed to be such that the retardation force generated on the particles is comparable to the collision forces exerted by the surrounding fluid molecules. Thus, the fluctuating forces and the resulting acceleration can be expressed as a Langevin equation.

$$m\frac{du'(t)}{dt} = -\zeta u'(t) + F_B(t) \tag{4.10}$$

If the particles are not accelerated, this equation can be simplified to

$$u'(t) = \frac{F_B(t)}{\zeta} \tag{4.11}$$

This fluctuating force $F_B(t)$ resulting from collisions between solid particle and fluid molecules has two components– deterministic and stochastic (noise). '$F_B$'[68]. The fluctuating force term has the following properties :

$$\langle F_B(t) \rangle = 0 \tag{4.12}$$

$$\langle F_B(t) F_B(t') \rangle = 2kT\zeta\delta(t - t') \tag{4.13}$$

Here '$\langle\rangle$' represents the time averaging. The equation 5.5 shows that the strength of fluctuation forces can be expressed in terms of $kT\zeta$ and delta function represents no correlation between forces at different time intervals $t$ and $t'$. Substituting the expression for $u'$ from Eqn. 4.11 into the Eqn. 4.9, results into

$$dX_t = \bar{u}dt + \frac{F_B(t)}{\zeta}dt \tag{4.14}$$

Stochastic part of the fluctuating force can be modeled as different noise type depending upon the physics of the problem. In this study, white or Gaussian noise is used to represent distribution of fluctuating forces. This is characterized as Brownian motion process or Wiener process, mathematically this can be written as:

$$W_t = \int_0^t \eta(\tau)d\tau \tag{4.15}$$

where, $W_t$ is the Wiener process with the following properties:

$$\mathbb{E}[W_t] = 0 \tag{4.16}$$

,

$$\mathbb{E}[W_t^2] = t \tag{4.17}$$

$$W_{t+s} - W_t \sim \mathcal{N}(0, s) \tag{4.18}$$

Defining the strength and stochastic component of the fluctuating force in Eqn. 4.14 leads to a stochastic differential equation

$$dX_t = \bar{u}dt + \sqrt{2D_e}dW_t \tag{4.19}$$

where, diffusion coefficient $D_e = \frac{kT}{\zeta}$.

Thus, state evolution of the particles in a diffusing media with drift parameter may be delineated by an equation similar to 'Langevin' equation or more specifically by an Itô stochastic differential equation (SDE) 4.19. The SDE so developed take care of the drift ($\bar{u}$) and diffusion ($D_e$) of the particles in addition to the innovations ($dW_t$) associated with the positions of particles.

## 4.3 Numerical and analytical methods

Itô SDE describes the trajectory of the individual particle; however, the ultimate question in addressing the uncertainties in the spatio-temporal evolution would require simulating the motion of a large number of particles and then construct their density distribution at different time snapshots. In other words, this direct numerical solution is expected to conserve the physics predicted by Convection-Diffusion (or Fokker-Planck) equation and also inform uncertainties. However, it is important to show that both approaches have equivalence under some approximations.

To find that, one needs to construct the Fokker-Planck equation, which governs the evolution of the density function. In this section equation governing the evolution of un-

derlying probability density function is recovered, and the equivalence of Fokker-Planck and convection-diffusion equation (CDE) is established.

In order to find that, it is important to resort to Itô's lemma[69], which can be written via the Taylor series expansion of the function of the stochastic process $X_t$. While doing so the second-order term in $x$ should be retained, since $dx^2 \propto dt$ (can be followed from the discretized version of Eqns. 4.19 & 4.17) and cannot be ignored.

Therefore for some function '$V$' of the stochastic process '$X_t$' such that $V = 0$ for $t \notin (0, T)$, Itô's lemma can be stated as: :

$$\begin{aligned} dV = (\frac{\partial V}{\partial t} + \bar{u}\frac{\partial V}{\partial x} + D_e\frac{\partial^2 V}{\partial x^2})dt \\ +\sqrt{2D_e}\frac{\partial V}{\partial x}dW_t \end{aligned} \tag{4.20}$$

Integrating Eqn. 4.20 over $t \in [0, T]$ results in:

$$\begin{aligned} V(X_T, T) - V(X_0, 0) = \int_0^T [(\frac{\partial V}{\partial t} \\ +\bar{u}\frac{\partial V}{\partial x} + D_e\frac{\partial^2 V}{\partial x^2})dt + \sqrt{2D_e}\frac{\partial V}{\partial x}dW_t] \end{aligned} \tag{4.21}$$

Taking the conditional expectation of the above equation:

$$\begin{aligned} \mathbb{E}[V(X_T, T) - V(X_0, 0)] = \int_R \Big[\int_0^T \Big[(\frac{\partial V}{\partial t} + \bar{u}\frac{\partial V}{\partial x} \\ +D_e\frac{\partial^2 V}{\partial x^2})\Big]dt\Big]C(x, t|x', t')dx \end{aligned} \tag{4.22}$$

Here, $C(x, t|x', t')$ is the concentration jump or change in the probability density from $(x, t)$ to $(x', t')$. Because $\mathbb{E}[dW_t] = 0$

On integrating by parts this equation can be further simplified to:

$$\begin{aligned} \mathbb{E}[V(X_T, T) - V(X_0, 0)] = \int_R \int_0^T V(x, t)\Big[\frac{-\partial C}{\partial t} - \bar{u}\frac{\partial C}{\partial x} \\ +D_e\frac{\partial^2 C}{\partial x^2}\Big]dtdx \end{aligned} \tag{4.23}$$

As $V(X_t, t) = 0$ for $t \notin (0, T)$[70] therefore:

$$D_e \frac{\partial^2 C}{\partial x^2} = \frac{\partial C}{\partial t} + \bar{u} \frac{\partial C}{\partial x} \tag{4.24}$$

Following explicit solutions of these equations, SDE and CDE are used to illustrate the equivalence and demonstrate the merit of direct numerical solution of SDEs for individual particle trajectories in order to quantify particle dispersion effects.

### 4.3.1 Solution to Convection-Diffusion equation

As mentioned earlier, contaminant transport is dependent upon the flow field, which requires the solution of momentum transport equations. However, due to simplifications and formulating eddy diffusivity along with the uniform mean flow approximation, it can be considered as a standalone convection-diffusion equation with a constant uniform mean flow velocity $\bar{u}$. The convection-diffusion equation with the stated boundary conditions can be solved analytically, which results in the solution

$$C(x, t) = \frac{1}{\sqrt{4\pi D_e t}} e^{-\frac{(x-\bar{u}t)^2}{4D_e t}}, \tag{4.25}$$

which is the closed form expression for concentration field and provides a mean spatio-temporal evolution of the scalar contaminant at different depths in the groundwater. In other words, this solution corresponds to the mean concentration profile in $(x, t)$ for uniformly moving fluid front in the groundwater. These analytical results are plotted for tabulated parameters in figure 4.3.

### 4.3.2 SDE simulations

SDE formulation is solved numerically after discretization of Eqn. 4.19. The developed scheme should preserve two critical physical attributes of the homogenized geometrical features of transport in groundwater networks: 1) memoryless or the absence of long-time correlations 2) purely diffusive nature of homogeneous media attributed to the exclusion of

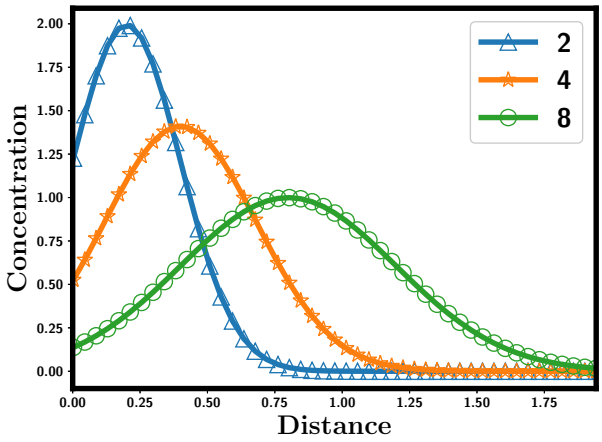



**Figure 4.3**: *Analytical solution at 3 different times after the initial particle release(at time period 2, 4 & 8 in arbitrary units).*

hanging ends, bottlenecks and backbends[71]. The first physical feature is incorporated by sampling a unique data point from a normal distribution at each time step. Second feature is preserved by the term which mimics the condition of diffusion in a homogeneous Euclidean space i.e. by the incorporation of the term which is proportional to the mean square displacement (MSD) of the particle, mathematically it implies: $\langle x^2 \rangle = 2D_e\Delta t$ in the simulation scheme, or the MSD grows linearly with time, which is the case in a purely diffusive media[72]. This also implies that the increments of Wiener process can be written as $dW_t = t^{0.5}\mathcal{N}(0,\,1)$, where $\mathcal{N}(0,\,1)$ is the normally distributed random number.

The numerical solution procedure of the Itô SDE is also known as stochastic simulation algorithm[73], which incorporates the above mentioned features. The algorithm can be stated explicitly as:

a) Generate a normally distributed (with zero mean and unit variance) random number $\mathcal{N}(0,1)$

b) Compute the position of the particle at time $t+\Delta t$ by using the equation: $x(t+\Delta t) = x(t) + \sqrt{2D_e\Delta t}\mathcal{N}(0,1) + \bar{u}\Delta t$. Then continue with step (a) for time $t+\Delta t$.[73].

## 4.4 Results and Discussion

After solving the SDEs, the individual particle trajectories of displacement versus time are plotted. Each trajectory corresponds to an individual particle and is unique due to random components (see figure 4.4).

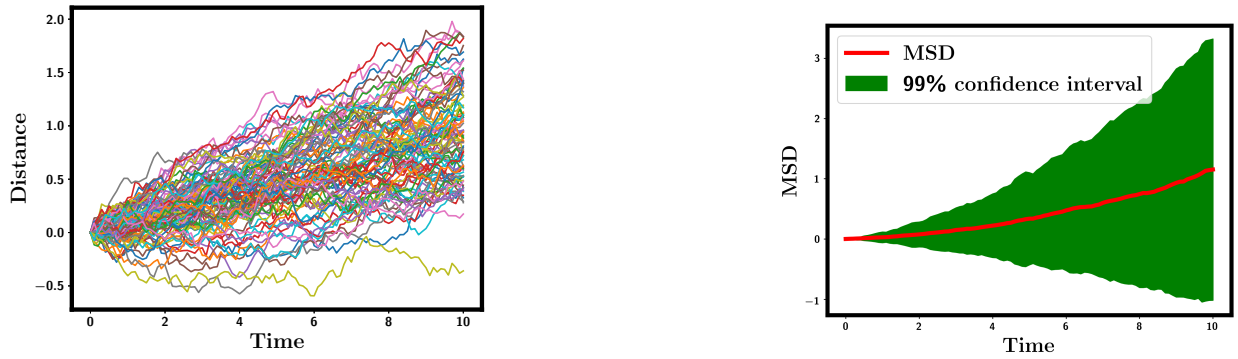



**Figure 4.4**: *Mean Square Displacement (MSD) of particles computed from the multiple realizations of Eqn. 4.19 after tracking for a period of 10 (in arbitrary units) in a media with* $D_e = 0.1$ *and* $\bar{u} = 0.01$.

Different trajectories of 5000 – 100,000 particles were tracked in a media with tabulated parameters for a period of 10 (in arbitrary units). Particle concentration at three different time steps were recovered by binning the data as per the Rice rule. However, this conventional rule may or may not introduce spurious noise or over smoothing, therefore to find an optimal solution is a difficult problem. This is true especially for a limited number of particles and the concentration profile is theoretically accurate only for an infinite number of particles. Therefore in a real application with a limited number of particles injected, kernel-based approaches, for instance, kernel density estimators (KDE) may lead to a better solution[74]. The KDE estimators are used for reconstructing the target density function which are given by $\hat{f}_h(x) = n^{-1} \sum_{i=1}^{n} K_h(x - X_i)$, where '$h$' is bandwidth and '$K$' is a symmetric kernel like 'Gaussian', 'Triangular' etc. The choice of bandwidth selection via the minimization of integrated least square error is a critical criterion in determining the true target density function and the problem is similar in nature to the bin size selection where a small value of bandwidth will introduce spurious oscillations and make the density estimate look 'wiggly', whereas a big value of bandwidth will lead to an estimate which is too smooth in the sense

that it is too biased and may not reveal structural features, for example, the presence of 'bimodal' behavior[75].

$$n_h = 2n^{\frac{1}{3}} \tag{4.26}$$

$$\tilde{C} = \frac{C(x,k)}{\int_{-\infty}^{\infty} C(x,k)dx} \tag{4.27}$$

$$\bar{C} = \frac{\sum_i^N C_i}{N} \tag{4.28}$$

$$\sigma = \sqrt{\frac{\sum_i^N (\bar{C} - C_i)^2}{N}} \tag{4.29}$$

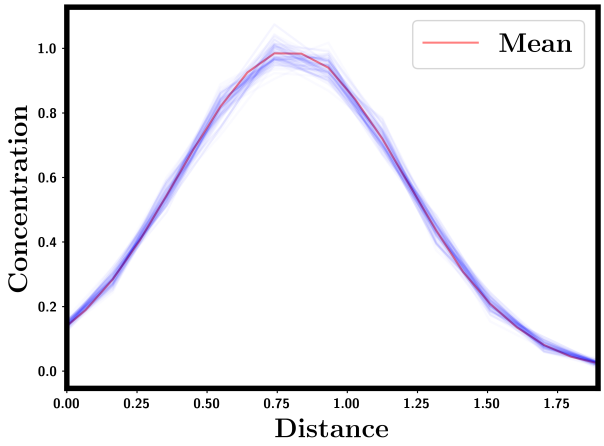



**Figure 4.5**: *Plot showing 100 realizations of concentration profiles with mean concentration profile shown in red.*

The numerically constructed concentration profiles at fixed time steps were normalized using equation 4.27 in which 'k' is a fixed time instant. In which the 'spaghetti' like blue curves (using 100 particles) is the reconstructed concentrations from the different trajectories of the SDE (4.19) simulations together with the mean.

A similar process was repeated for different cases, and data was used to construct the mean and uncertainty (confidence level 99%) bands over the concentration profiles. Simu-

lation results i.e., constructed concentration profiles and variance for two cases with 5000 particles and 100000 particles are shown in Fig. 4.6.

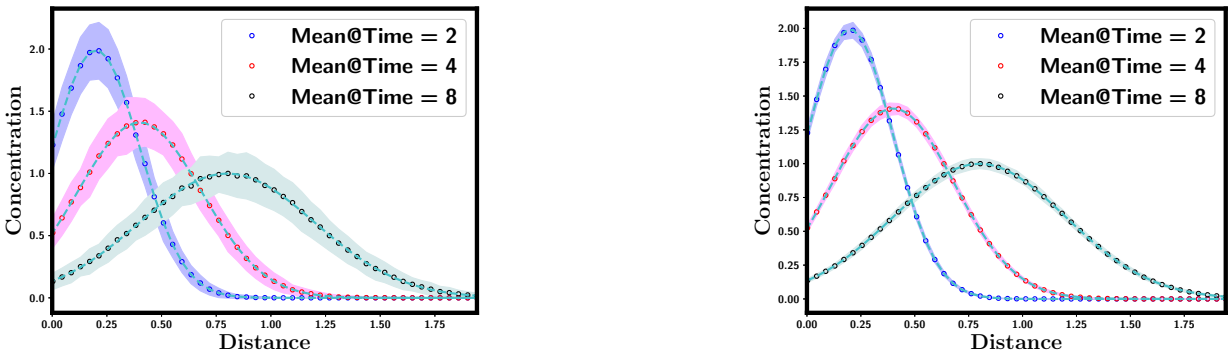
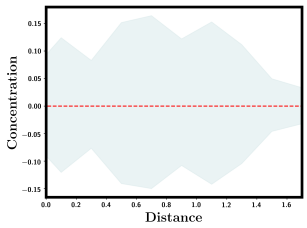


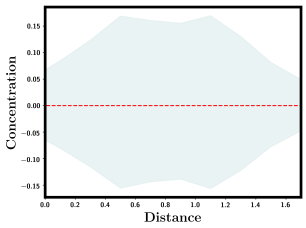

**Figure 4.6**: *Concentration profiles with mean and uncertainty bands with 99% confidence interval for 5000 trajectories (left) and 100,000 trajectories (right).*

It can be seen from the simulation results that as the number of particles were increased the standard deviation associated with concentration profile narrows down and the solution approaches a mean value as shown in Fig.4.6, which confirms that limiting ($\lim_{n\to\infty}\sigma$) case of the equation 4.29 would lead to true homogeneous state of eulerian transport depicted by convection-diffusion equation or the Fokker-Planck equation. This can also be viewed as Borel's law of large number which states that the mean of a random experiment will approach the probability of the outcome, which in this case is the solution of the Fokker-Planck equation. Mean concentration profiles in both cases match well with the analytical predictions. The merit of the SDE based approach is illustrated by capturing the uncertainty of contaminant dispersal for a small number of particles. Particularly for suspended solids or colloids, the SDE approach is preferable and more realistic.

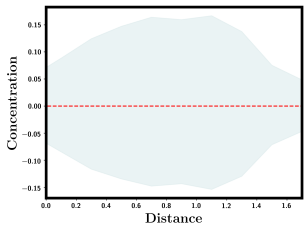
Due to the small number of particles, different ensembles with different realizations of 1000 particles are performed, and the standard deviation is calculated, as shown in Fig. 4.7 with 99% confidence interval. This result signifies that increasing the number of realizations for the same number of particles will smoothen and approximates the mean and standard deviation better i.e., mean profile and 'jittery' behavior of the standard deviation will reduce or in other words the change in variance got converged after 1000 runs of 1000 trajectories and no significant change is observed after that.



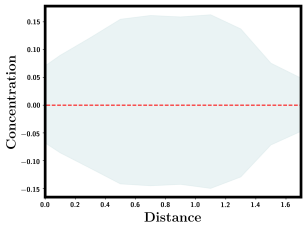

10 realizations




100 realizations




1000 realizations




100,000 realizations

**Figure 4.7**: *Different ensembles of simulations with 1000 particles and different number of total realizations.*

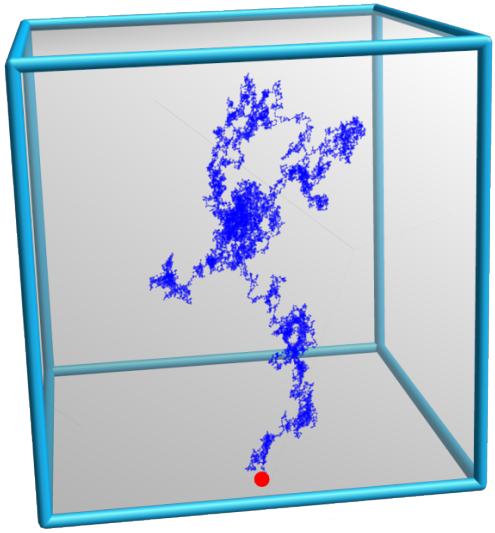
The deviations from the mean behavior are observed both in the length and time scales; this demonstrates the elegance of stochastic calculus which can predict the true state evolution of the particle–an effect not captured in convection-diffusion type equations.

### 4.4.1 First passage time

The uncertainty around the computed concentrations provides a way to accurately quantify the variances around the mean concentrations tracked via a Fokker-Planck equation. However, for emergency preparedness, decision making, and impact quantification after a major nuclear breach, it is imperative to understand at what instant of time the released activity will reach the impact point. This particular time period is the time taken by the particle starting at some specified location to reach (the point of breach or release) another specified location (the point of impact).

### 4.4.2 The definition of first passage time

Mathematically the definition of this particular measure can be found by considering the motion of a set of point $\mathbf{x}$ following the Langevin equation in a volume $\Omega$ bounded by an absorbing boundary $d\Omega$. Because of noise repeated execution of the Langevin equation with the particles starting at the same point, $\mathbf{x_0}$ follow different trajectories. This leads to different hitting times. Focusing only on the points that have not left the volume $\Omega$ and calling their



$\Omega$: Volume
d$\Omega$: Boundaries
$\mathbf{x}$: Trajectories
$\mathbf{x_0}$: Origin (shown in red)

**Figure 4.8**: *3-Dimensional volume $\Omega$ with absorbing boundaries $d\Omega$ and the realization of Langevin equation. On this domain, the first passage time is defined.*

distribution by $p(\mathbf{x}, t)$. Integrating this over all the particles which are inside the volume $\Omega$:

$$S(t, \mathbf{x_0}) = \int_{\Omega} p(\mathbf{x}, t) dx \tag{4.30}$$

This function represents the information that the number of all the starting points that are still inside the volume $\Omega$ at the time '$t$' or the probability that the particle is still in $\Omega$ at the time '$t$.' These points are inside the volume at the time '$t$,' but they might reach the

absorbing boundary $d\Omega$ at a time '$t + dt$.' Therefore, the change:

$$\underbrace{-\frac{dS}{dt} = \chi(t)}_{\text{First passage time or hitting time distribution}} \tag{4.31}$$

Since time at which the particle will hit the boundary is a random variable, therefore, the mean value of it is the mean hitting time[68].

### 4.4.3 Numerically computing the first passage time

Analytical calculation of first passage time can be challenging; therefore, such calculations can be most easily performed by running an Itô's stochastic differential equation– which tracks the distance traveled by the released particle. Here, the absorbing boundary is a representation of the point of impact (see figure 4.9). As soon as any trajectory reaches the point of impact, the particle was removed from the domain and the time is recorded. From a conservative calculation point of view the minimum time period from all the truncated series is important–i.e., the time at which the first trajectory reaches the impact point. However, distribution of time period (see figure 4.10) leads to a better and informed emergency preparedness by providing not only the minimum time period but the mean and appropriate variance around the meantime period.

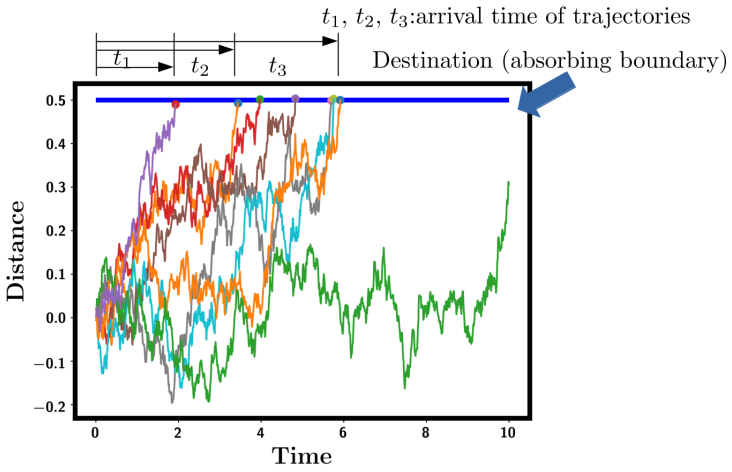



**Figure 4.9**: *Absorbing boundary condition: truncated trajcetories of Itô's sde.*

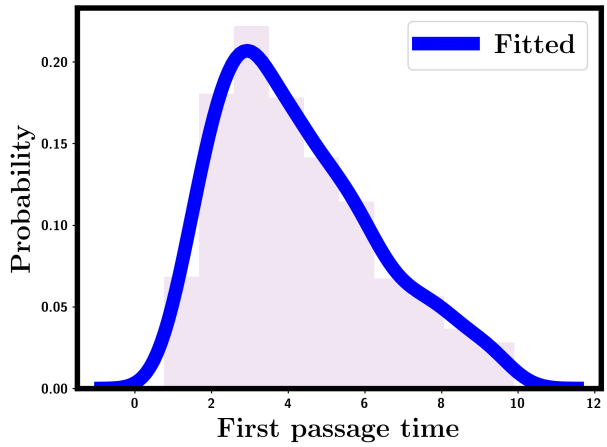



**Figure 4.10**: *Distribution of hitting time: The distribution of hitting time leads to a better emergency preparedness by building an uncertainty band around the mean hitting time.*

## 4.5 Radionuclide dispersal: quantifying the effect of radioactive decay

The absence of decay term in the above method limits its applicability to the case concerning the decay of the radioactive particles. In the following sections, two methods of dealing with the same are presented. Two specific cases were considered a) decay of radioactivity attached to a colloidal particle as it continues its sojourn. b) decay of a single ion or a colloid particle which itself is radioactive i.e., the radioactivity comes from within, and the state of it can change randomly at any instant of time. This random change of state in the second scenario is modeled via a Bernoulli trial.

As a third modification, the noise driving the Itô's sde was modified, and some physical reasoning for the same is also presented.

### 4.5.1 Colloid facilitated radioactivity transport

Radioactive elements, for example, cesium attaches to the colloids via ion exchange mechanism in the contaminated subsurface regions[76] (see figure 4.11). This mechanism greatly enhances their transport as now a large amount of radioactivity can be transported to longer distances–leading to a potential greater health hazard. Therefore as a first case trajectories

of the motion of the colloid particle were considered via the Itô stochastic differential equation. Since there is a continuous decay of atoms attached to the colloid–the decay of which can be modeled by the Poisson process. Thus, the decay of radioactivity can then be incorporated by multiplying the computed concentration of the colloid via the term having a continuous decay of radioactivity $N = N_0 e^{-\lambda t}$ ($\lambda$ =decay constant). With this modification, the following results were obtained (see figure 4.12) which clearly shows that the decay of radioactivity attached to the colloidal particle will suppress the computed radioactivity. Without this decay term, the computed radioactivity will be simply the concentration of colloid multiplied by the number density of the radioactive atoms attached to it.

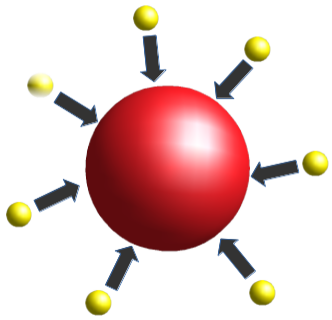


**Figure 4.11**: *Colloid particle shown here in* ■ *color attached to it are the radioactive species shown in* ■ *color*

### 4.5.2 Decaying radioactive particle–Bernoulli trials

In the second case radioactive particles are tracked in the media which may randomly decay as they transit through the computational domain. This decay can be simply achieved by running a Bernoulli trial as the trajectory of a particle is tracked in time. If the Bernoulli trial results in a '0' output, the trajectory will get truncated, and the radioactive particle was no more tracked in time. It can be seen that a number of trajectories stopped randomly–this particular factor depend on the value of the survival probability. If the survival probability is high, almost all the trajectories will survive. However, for the case of $P_s = 0.994$ the

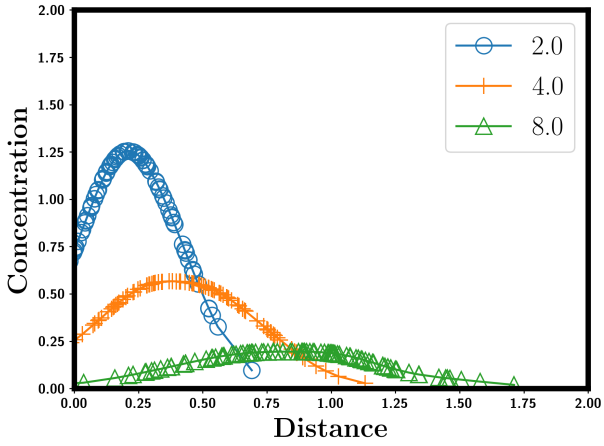



**Figure 4.12**: *Concentration computed for decaying colloidal particle.*

following results were obtained (see figure 4.13).

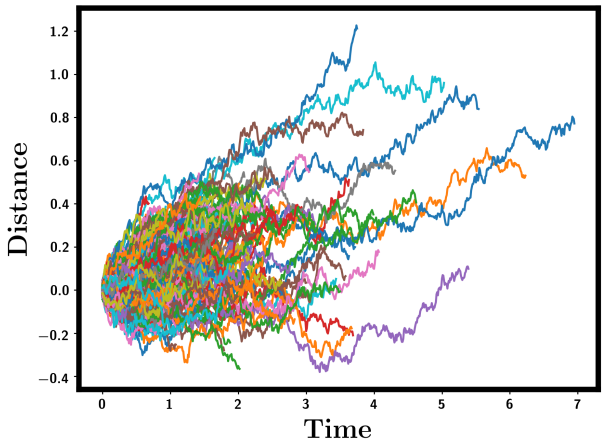



**Figure 4.13**: *Trajectories of particles with Bernoulli trial. Here the survival probability is set to 0.994.*

### 4.5.3 SDE driven by colored noise

#### Motivation

As a third modification it is imperative to divert the attention on the noise driving the system. Since noise is the biggest source of uncertainty in the stochastic system (see figure 4.15). Therefore, it is imperative to make a correct approximation of it. The white noise forcing used in the previous simulation has limited applicability since it is only valid at large time

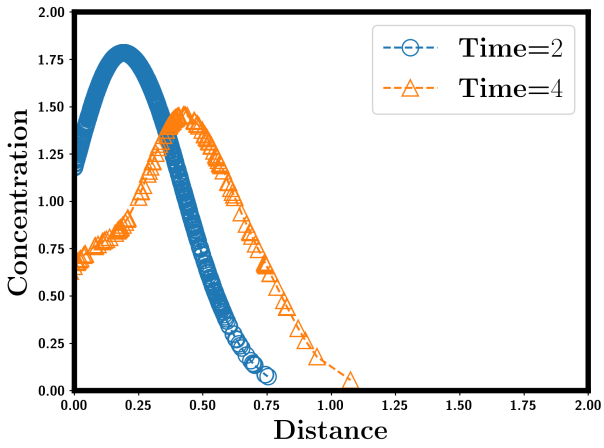



**Figure 4.14**: *Computed concentrations from trajectories.*

scales, including the associated drift-diffusion equation[68]. While for any natural system the spatio-temporal correlations do not decorrelate at such fast pace. This is also true for the particles driven by a rapidly decorrelating velocity field for which the spatio-temporal correlation is not white in time or space. To resolve this, it is imperative to replace the forcing in the Langevin equation by a colored noise generated by some other method. This type of noise can be generated conveniently by another stochastic differential equation, which then can be inserted into the Langevin model. This section of this chapter discusses two such methods to generate the colored noise.

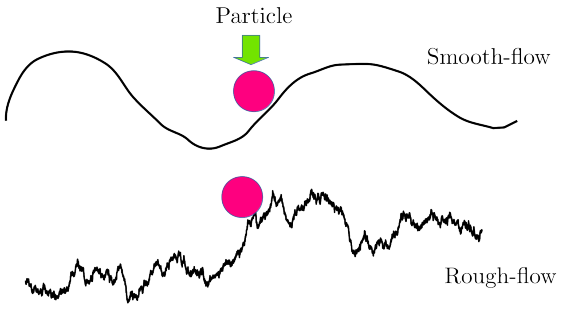



**Figure 4.15**: *A particle in smooth (less uncertain) and rough flows (more uncertain)–which characterizes the degree of temporal correlation. The smoother the trajectory is the higher the temporal correlation is. Correlated noise can be generated by using an OU process (see equation 4.32).*

### 4.5.4 OU-process for colored noise

An OU-process is a corrected version of Newton's second law to account for the small-time fluctuations observed in the course of the trajectories of a particle in a fluid stream. This idea of an OU process can be extended to the situations where the particle is not a real particle at all, but instead some collective property of a macroscopic system, for example, the trajectories of imaginary packets of the fluid stream or the random noise driving the particles[68]. For that particular case an OU-process can be written as:

$$\frac{d\Gamma}{dt} = -\gamma\Gamma + C^{0.5}\phi \tag{4.32}$$

where '$\gamma$' is the relaxation parameter which governs how fast the process ($\Gamma$) reaches its equilibrium value. While a high value of '$\gamma$' guarantees that the process is extremely quick in reaching its equilibrium state, a lower value of it ensures a slow decay. The third and fourth important parameter in the model is '$\phi$' & '$C$' which are the random forcing term and the amplitude multiplier respectively. The term '$\phi$' is random in the sense that it has no correlation with its past values.

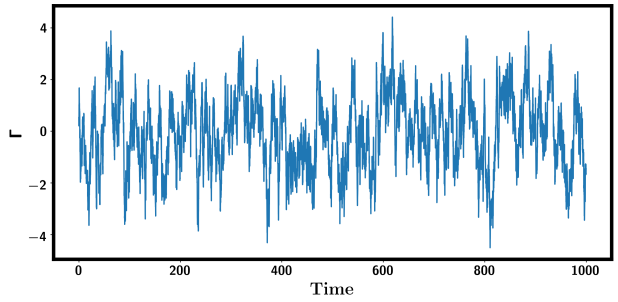



**Figure 4.16**: *OU process for* $\Gamma$.

$$C = 2\langle\Gamma(0)^2\rangle\gamma \tag{4.33}$$

Using the relationship mentioned in equation 5.21 (see appendix B) the parameters of the model were set and a realization of it was generated by using the Euler-Maruyama method. The only user defined parameter here is the $\gamma$ which can be set either manually or can be

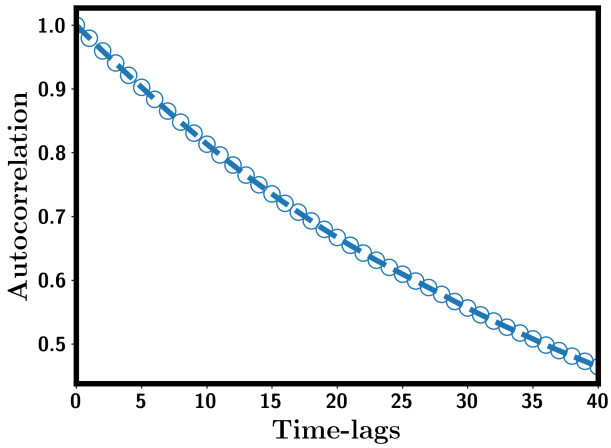



**Figure 4.17**: *Correlation of simulated OU process.*

computed directly from the data. For this particular case the $\gamma$ is set to a value of 0.2.

The generated noise $\Gamma$ can now be inserted into the Itô sde, and the trajectories can be integrated by an approximate method (see equation 4.34):

$$x(t) \approx x(0) + (\bar{u} + \Gamma(0))\Delta t \tag{4.34}$$

Since the noise is not uncorrelated in time hence, the system described above is non-Markovian in nature.

### 4.5.5 Colored noise generation by generalized Langevin Equation

It may happen that the noise driving the particles is non-Markovian itself. This is possible if the particles are driven by a turbulent flow for which the correlations may not look like a decaying exponential (see figure 4.17). For that particular case, it is possible to mimic a turbulent forcing using a generalized Langevin equation[77]. The generalized Langevin equation is similar to the Langevin model with the addition of a convolution integral in it. It can be written as:

$$\frac{d\Gamma}{dt} = -\int_0^t K(t-\tau)\Gamma(\tau)d\tau + F(t) \tag{4.35}$$

Where '$K$' is the memory kernel and '$F(t)$' is the forcing, the correlation of which is not prescribed. In the generalized Langevin equation, it turns out that the memory term '$K$' is an important parameter the form of which cannot be described beforehand. However, it can be computed by resorting to the second fluctuation-dissipation theorem (see equation 4.36) [68].

$$K(t) = \frac{\langle F(0)F(t)\rangle}{\langle \Gamma(0)^2\rangle} \tag{4.36}$$

Assuming that the force $F(t)$ is exponential correlated i.e.:

$$\langle F(0)F(t)\rangle = e^{-\alpha t} \tag{4.37}$$

Inserting equation 4.37 in 4.36 it is possible to get a functional form of the memory Kernel. With this memory kernel in hand it is very easy to compute the covariance structure $(Cov(s))$of the signal $\Gamma$ from the following equation written in the Laplace domain:

$$Cov(s) = \frac{\langle \Gamma(0)^2\rangle}{s + K(s)} \tag{4.38}$$

After inserting the memory kernel in the equation above and some manipulation it is possible to recover the covariance structure of $\Gamma$ in the time domain (for details of Laplace transform see appendix C):

$$Cov(t) = \langle \Gamma(0)^2\rangle (e^{-\alpha t/2}(cos(\omega t) + \frac{\alpha}{2\omega} sin(\omega t))) \tag{4.39}$$

$$\omega = \sqrt{\frac{1}{\langle \Gamma(0)^2\rangle} - \frac{\alpha^2}{4}} \tag{4.40}$$

With a simple Fourier inversion technique it is possible to generate the trajectories of the generalized Langevin model.

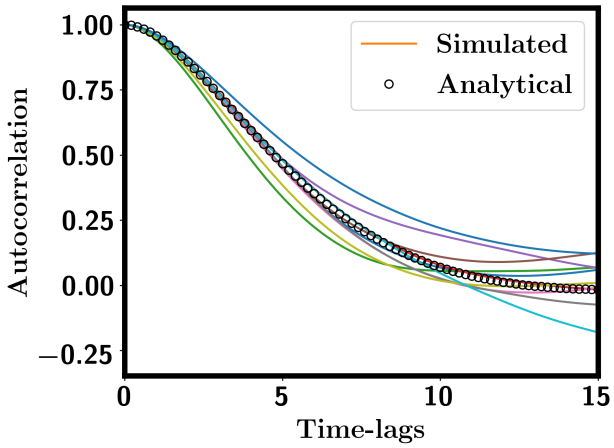



**Figure 4.18**: *Analytical correlation function and the correlation of simulated gLe trajectories.*

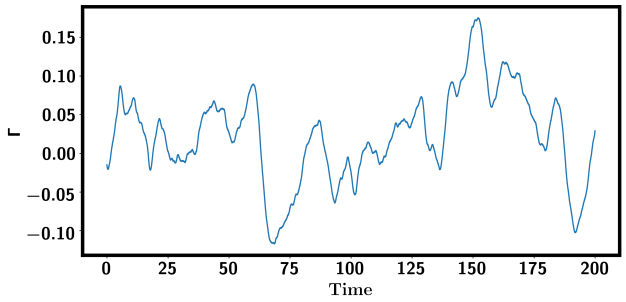



**Figure 4.19**: *Simulated time trajectory of a generalized Langevin model.*

## 4.6 Conclusions

The computational tools presented here evaluate the uncertainties related to the contaminant transport in the homogeneously modeled groundwater system. The limitations of the conventionally used eulerian approach are highlighted, and lagrangian framework of particle motion in 1-D groundwater is analyzed. Langevin equation model for simulating dynamics and kinematics of the particles is modified to account for velocity fluctuations and is simplified to Itô's stochastic differential equation with drift and diffusion components. It is shown with the help of Taylor series expansion and using Itô's lemma that by ignoring second-order terms, Itô's SDE results into Fokker-Planck or Convection-diffusion model. The numerical solutions to SDE are used to construct concentration profiles. As expected, each realization of stochastic simulations should result in a unique concentration profile and especially for

a small number of particles the variations between different realizations is larger. Using simulation, statistics, mean concentration profiles, and their uncertainty bands are plotted and compared with the analytical solution. There is a good agreement between ensemble-averaged concentration profiles and the analytical predictions for all cases. The thickness of uncertainty bands is more thick for The approach presented here studies the evolution of variances by two standpoints a) with the increase in a number of particles b) with an increase in the number of realizations. Rather than coupling the uncertainties externally in a deterministic continuum model, this uncertainty quantification approach is more rigorous and realistic. The discussed approach appears to be the most pragmatic way of treating the various uncertainties which are coupled at a particle level and have involved uncertain and unknown feedback mechanisms owing to the unresolved physics and modeling deficiencies/ errors. The considered approach may go off the target in the presence of dangling ends which may slow down the diffusion process; this can be attributed to the only diffusive nature of the equation attributed to the Wiener process. However, this deficiency can be overcome by driving the stochastic differential equation simply by a colored noise leading to a non-Markovian Langevin equation.

# Chapter 5

# Statistical modeling approach to load following nuclear power

In the last chapters, it was demonstrated how to constrain the stochastic differential equation directly on the data. The trained model can find extensive applications in quantifying the uncertainties attributed to the gross approximations made in the system code formulation. A very important class of problem which has marred the nuclear industry is related to the operation of nuclear power plants in the age of renewable energy. It is attributed to the fact that most of the government subsidies are geared to cater to the renewable industry. This shift in policy has created an indomitable challenge from the grid operation point of view. Since in the recent past owing to this shift in policy a global phenomenon of negative spot prices of electricity is on the rise. Due to the subsidy given to the highly flexible renewable energy producers they can now put their energy into the electrical grid even at negative spot prices. Since the conventional generators like nuclear power plants are not allowed to ramp up and down the power, owing to which these generators are finding it hard to survive. To this end, a motivating example is presented in the next section followed by the modeling of wind speed with a 'Langevin' model and the coupling of the same with a reduced complexity model of the nuclear power plant. The main goal is to demonstrate the thermal-hydraulic response of the nuclear power plant and to show the synergistic existence of it in a small

micro-grid.

## 5.1 Motivation

Most of the fossil fueled plants can supply peak-loads by adding more fuel and, thus, can generate far more revenue during those peak hours. On the other hand, use of NPPs for peak load following is quite complex due to technical constraints associated with reactor behavior. These technical challenges include the adequate handling of reactivity swings caused by time-varying fuel and moderator temperatures, a higher fuel-failure probability due to thermal-structural cycling, and xenon poisoning.

Although there are presently some reactors around the world that are operating with flexible load-following capabilities, such operation is restricted to slowly-varying powers, 2-3 times a day, and only up to 80% of the fuel cycle. In order to evaluate the safe operation under transient conditions generated by load following maneuver, coupled thermal-hydraulics neutronics reduced order is used. A hypothetical model of microgrid is studied where a wind energy generation and nuclear power plant supply total electricity. Thermal response of the reactor shows stable behavior with small fraction of the grid load shared by wind energy system.

The reduced order reactor model considered here is a 0 dimensional. Hence, it is not able to capture spatial variations and fluctuations in the variables, but it can be used as a model to identify any major impact on the overall reactor safety.

The cost of electricity from solar photovoltaic (PV) systems has dropped 15- to 20-fold over the last 20 years, and the LCOE for a PV system currently is about $70-130 per MWhr, including support structures, power conditioning, and land. However, these PV systems can not be used to supply peak load for most of the locations due to time lag in peak electricity generation and peak demand. Moreover, an increased penetration of solar generation on the grid can have unintended consequences as it can increase the load-generation unbalance when weather conditions abruptly change. Similarly, on-shore wind generation is another method of electricity production which can be considered as robust and economical but not

reliable. If the grid is solely depended on solar PV cells or wind energy for a large fraction of the electricity supply, the back-up power required on the grid could equal or even exceed 50% of power consumed because the back-up power would have to be at least equal to the amount of power that is actually used from PV cells/wind turbines. A stable grid, therefore, requires excess capacity which comes with additional back-up power but it remains idle for almost 50% of the time. The reliability of these energy systems in supplying peak load and base load can be substantially improved if there existed economical and durable methods of storing electricity, But the means of energy storage are currently not economical for grid scale deployment and therefore grid acts like a storage with the help of peak demand prices of electricity. Due to these reasons, the cost of electricity to consumers goes up when PV units or wind energy without storage supply energy to the grid. Therefore, under all possible conditions the fractional share of renewables has remained or will remain up to 10-15% in the near future.

In an unregulated electrical grid, wholesale electricity prices are generally driven by the supply and demand which depends on, time of day, weather conditions, seasonal factors, and consumption behaviors. Thus prices fall with low demand and increase during the high demand period. In the current market scenario, government policies lead to much higher revenues for producing electricity from wind and solar energy when compared to the more traditional sources of generation. The typical solar and wind energy production around the U.S. and across the world does not contribute significantly to reduce the demand load throughout a 24 hour period as shown in figure 5.1.

### 5.1.1 Impact on Nuclear Power

The government policies are designed to increase the solar and wind energy penetration in the electrical grid and conventional energy generation is expected to be flexible enough to fill in the deficits to the grid load demand. Due to this, some of the conventional power providers are even getting marred by negative electricity prices, where they have to pay the grid operators to offload the power they generate. Nuclear power plants (NPPs) are

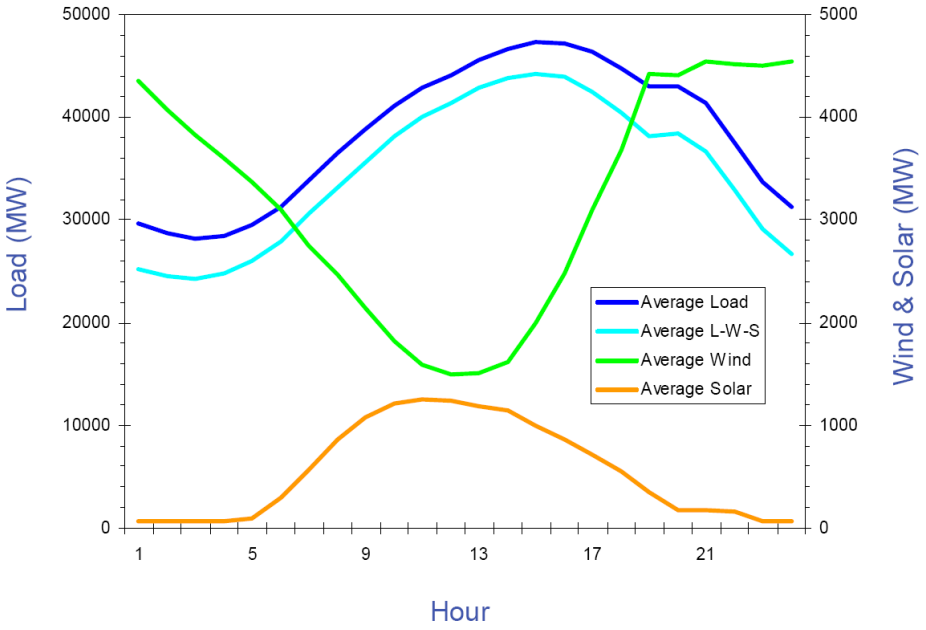



**Figure 5.1**: *A typical curve of wind and solar output, along with net (L-S-W) demand. L:Load, S:Solar, W:Wind (NERC-2009[1]).*

often subjected to negative pricing if they dispatch more energy on the grid than the net demand[78;79].

In the year 2016, Bloomberg New Energy Finance published a report-"Reactors in the Red: Financial Health of the US Nuclear Fleet," which shows that 55 % of America's nuclear plants are losing money and are at serious risk of being replaced by other sources of energy. Owing to which some of the conventional power producers may have to pay the consumers to offload the power from the grid. Lack of grid scale energy storage system suggests that the only options available for NPPs to overcome these economic challenges are to follow the grid load demand and save themselves from potential collapse.

## 5.2 Load following NPPs

NPPs were designed and generally operate as base load providers. Stringent safety criteria derived from conservative considerations imposed by the nuclear regulators makes it difficult to operate NPPs in load following mode. Due to these reasons, load following, adopted so far

in very few nuclear plants, is limited to some specific ramp rates[79] or for only limited number of power changes in a day. From fuel response to stability of NPP systems, numerous aspects must be carefully studied before the load following envelope can be expanded. In United States or other parts of the world, the stringent safety criteria derived from conservative considerations–from the fuel performance point of view doesn't allow power to be ramped up and down in a short period of duration. For example – California's sole remaining nuclear power plant, the Diablo Canyon, does not ramp down during the middle of the day, despite frequent and recurring situations of negative prices, thus losing valuable $s and making it uncompetitive when compared to renewables. From the clad-pellet interaction point of view a ramping rate of 50 MW/min (5 % per minute) load increase is derived for some European reactors[2].

Due to these reasons, it is important to examine reactor response under a scenario in which the electrical grid abruptly demands an increase in power for a finite period of time before the demand drops back to its original level. As energy production from renewable sources can be associated with sudden fluctuations, it is also critical to examine the response of NPPs to provide back-up fluctuating loads. For illustrating this, a hypothetical microgrid is modeled which is connected to, and is supplied electricity from a wind energy generation system and a light-water cooled nuclear power plant. Time-series data for a typical wind energy generation system is used as an input parameter for estimating the demand power for the nuclear power plant. As wind energy generation is dependent upon instantaneous wind speed for any geographical location, the discrete time series information are then used to compute nuclear system response. This work models the nuclear-coupled thermal hydraulic response of the reactor when its power level is governed by the grid demand. These transients were simulated using a simple reduced order model of a pressurized water reactor described in the next section.

The scope of this chapter is to provide a framework to assess possible thermal dynamics of the reactor under different load demand scenarios. In this study, there are two example test scenarios considered to show the implementation of this framework–a) scenario where there is a step change in the grid power demand for a finite amount of time, and b) considering

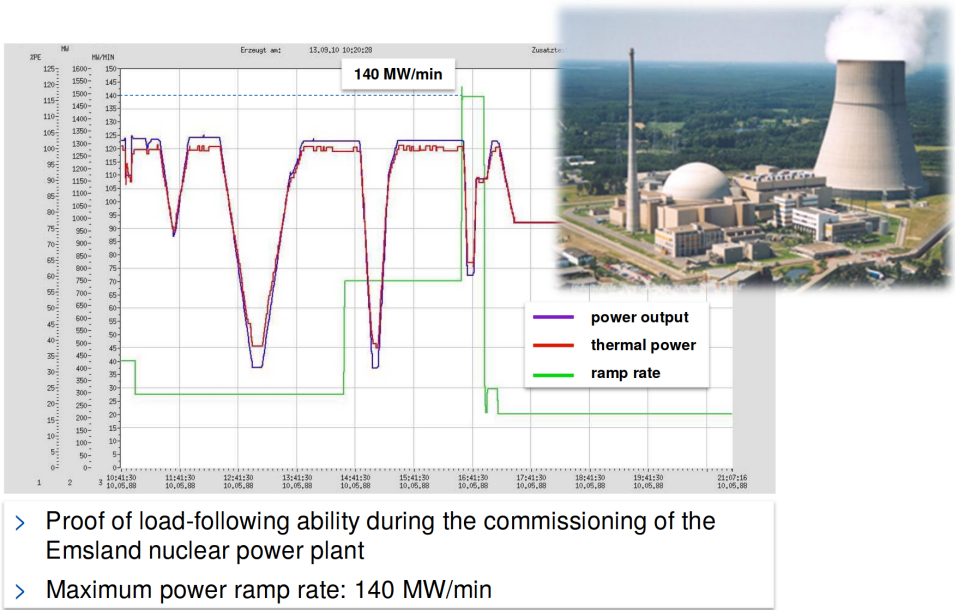



**Figure 5.2**: *Emsland NPP in Germany showing 140 MW/min ramping rate*[2].

a small microgrid where a nuclear power plant and stochastic wind energy generators exists together leading to sudden power demand fluctuations on nuclear plant[80].

## 5.3 Reduced Order model-Nuclear Reactor Dynamics

When load demand and in-turn actual reactor power is changed, it has a direct effect on the temperature, pressure, and enthalpy content of the main reactor systems. These thermodynamic changes impact the neutron balance due to changes in the neutron cross-sections of the material constituents. It is indeed a computationally challenging task to capture the detailed 3D thermal-hydraulics and neutronics of the reactor core, and due to the involvement of different systems, the detailed computational assessment is impossible with current resources. Therefore, a reduced-order model is devised to understand the overall reactor system response to obtain the temporal evolution of important variables in the lumped fuel or coolant volumes. Point reactor kinetics equations (Eqns. 5.1 & 5.2) with effective delayed neutron group are used to model the core neutron population and change in the power levels. The change in thermal power level will impact thermal transport in reactor fuel and coolant

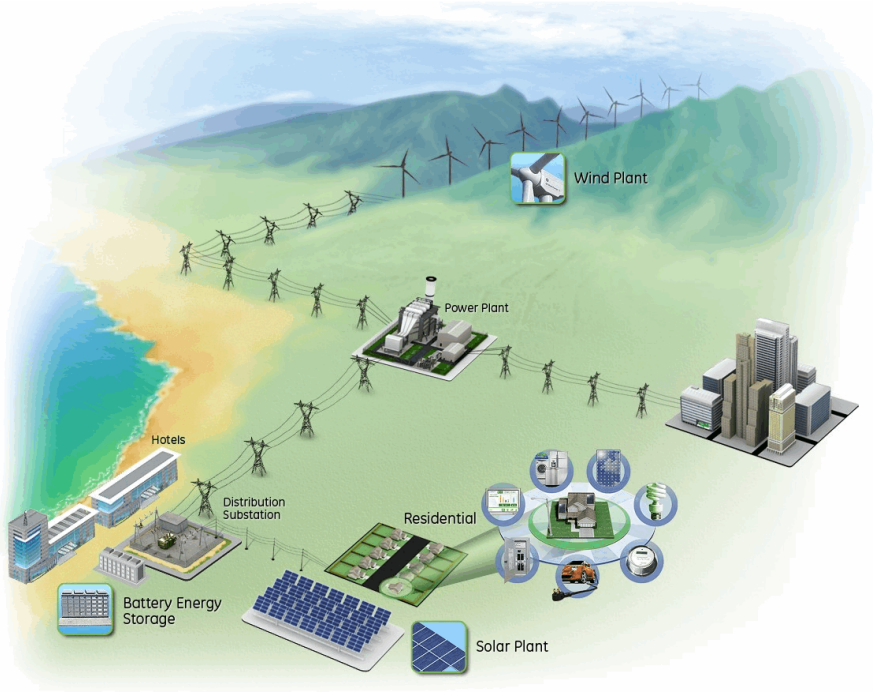



**Figure 5.3**: *Synergistic existence of nuclear and wind power in a microgrid.* ***(10 small wind turbines and one nuclear power plant–feeding a small locality)***

systems. A lumped parameter thermal energy transport model is used to assess thermal response. The reactor dynamics and control is captured with the feedback reactivity term that couples the point reactor kinetics and thermal system equations. A reactor will be stable when the reactivity reduces with increase in power level or temperature, but it is still critical to understand if the reactor will exhibit a limit cycle due to sustained feedback under all circumstances and what is the expected change in the temperature fluctuations.

$$\frac{dn}{dt} = \frac{(\rho - \beta)n}{\Lambda} + \lambda c \tag{5.1}$$

$$\frac{dc}{dt} = \frac{\beta n}{\Lambda} - \lambda c \tag{5.2}$$

$$\tag{5.3}$$

The reactivity feedbacks can be generically defined as,

$$\rho = \rho_0 + \rho_c(t) + \rho_f(u_1, u_2, .., u_k) \tag{5.4}$$

$$\frac{du_k}{dt} = f_k(n, u_1, u_2, .., u_k; w_1, w_2, .., w_m) \tag{5.5}$$

### 5.3.1 Temperature feedback

Fuel temperature change effects the reactivity primarily due to doppler broadening and fuel temperature coefficient for reactivity. The resulting impact of the doppler broadening is to increase the absorption cross-section (e.g. of U-238) resulting in a decrease in reactivity. The microscopic fission cross-section will decrease with the temperature leading to an additional reduction of the reactivity. Due to these effects, the fuel temperature coefficient for reactivity can be modeled as

$$\alpha_T^F = -ln[\frac{1}{p}]\frac{\beta}{2\sqrt{T_F}} \tag{5.6}$$

The coolant or moderator temperature change effects the coolant density which in turn effects the moderating capacity. Lower moderation of neutrons imply higher resonance absorption, therefore leading to significant reduction in the moderating capacity. Although reduction in coolant density also causes reduction in neutron absorption within the moderator or coolant, the increase in temperature effectively causes negative reactivity in PWRs and BWRs. In BWRs, this effect is due to increase in void fraction and in PWRs it is due to reduction in density of pressurized water. Negative moderator or coolant temperature coefficient is less significant than negative fuel temperature coefficient as changes happen in the fuel immediately. Nevertheless, the overall effect of temperature changes in the fuel and coolant or moderator can be described as Eqn. 5.7.

$$\rho = \rho_c(t) + \alpha_T^C \Delta T_C + \alpha_T^F \Delta T_F \tag{5.7}$$

### 5.3.2 Energy balance

In the set of equations mentioned above, equation 5.4 expresses the total reactivity in terms of the various contributions. Where $\rho_0$ is the reactivity under cold conditions with all the control rods withdrawn, $\rho_C(t)$ is the reactivity change owing to various reactivity control mechanisms, and $\rho_F$ is the reactivity introduced into the system by the non-nuclear feedback, for example, temperature changes[3]. The parameter values used for this study are listed in table 5.2. These temperature feedbacks can be captured by the differential equations governing the heat balance in the fuel and the coolant regions respectively[3].

$$m_F C_{pF} \frac{dT_F}{dt} = a_F n - h(T_F - T_C) \tag{5.8}$$

$$m_C C_{pC} \frac{dT_C}{dt} = h(T_F - T_C) - 2W_C C_{pC}(T_C - T_{Cin}) \tag{5.9}$$

Where $T_C$ and $T_{Cin}$ in equation 5.9 are the coolant average temperature and coolant inlet temperature respectively.

The model thus described can be written in terms of non-dimensional parameters using the relations given in table 5.1. In which $x$, $y$, $z_F$, $z_C$ are the normalized values of the neutron density, precursor density, fuel temperature and coolant temperature respectively. It is to be noted that the variables $n_e$, $C_e$, $T_{F,e}$, $T_{C,e}$ mentioned in table 5.1 are the equilibrium values of neutron density, precursor density, fuel temperature and coolant temperature respectively.

**Table 5.1**: *Non-dimensional variables.*

| | |
|---|---|
| $x$ | $\frac{n-n_e}{n_e}$ |
| $y$ | $\frac{C-C_e}{C_e}$ |
| $z_F$ | $\frac{T_F-T_{F,e}}{T_{F,e}}$ |
| $z_C$ | $\frac{T_C-T_{C,e}}{T_{C,e}}$ |

With these substitutions, the normalized lumped parameter reactor dynamics model

becomes,

$$\frac{dx}{dt} = \frac{-\beta x}{\Lambda} + \frac{1}{\Lambda}(\beta y) + \frac{\rho}{\Lambda} + \frac{\rho x}{\Lambda} \tag{5.10}$$

$$\frac{dy}{dt} = \lambda(x - y) \tag{5.11}$$

$$\frac{dz_F}{dt} = \frac{a_F n_e x}{m_F c_{pF} T_{F,e}} - \frac{h z_F}{m_F c_{pF}} + \frac{h T_{C,e} z_C}{m_F c_{pF} T_{F,e}} \tag{5.12}$$

$$\frac{dz_C}{dt} = \frac{h T_{F,e} z_F}{m_C C_{pC} T_{C,e}} - \frac{(2 C_{pC} W_{C,e} + h) z_C}{m_C C_{pC}} + \frac{2 W_{C,e} T_{Cin,e} u}{m_C T_{C,e}} - \frac{2 W_{C,e}(T_{C,e} - T_{Cin,e}) w}{m_C T_{C,e}} \tag{5.13}$$

$$\rho = \rho_c(t) + \alpha_T^C T_{C,e} z_C + \alpha_T^F T_{F,e} z_F \tag{5.14}$$

The variables $u$ and $w$ in equation 5.13 are the dimensionless coolant inlet temperature and the dimensionless coolant flow rate, respectively, which are defined in equations 5.15and 5.16 in terms of $T_{Cin,e}$ and $W_{C,e}$ which are the equilibrium coolant inlet temperature and equilibrium coolant flow rate.

$$u = \frac{T_{Cin} - T_{Cin,e}}{T_{Cin,e}} \tag{5.15}$$

$$w = \frac{W_C - W_{C,e}}{W_{C,e}} \tag{5.16}$$

### 5.3.3 Control System

To simulate the scenario in which the demand side is dictating the system output, it is necessary to introduce an appropriate controller in the system so that the system may adapt to the changing demand. A proportional controller is used in this analysis, and excess reactivity is introduced via a classical output feedback control. In other words, a unity feedback control loop shown in figure 5.4 is implemented. Based on the mismatch between the power being produced and power demand, externally controlled reactivity is varied in such a way so as to minimize the mismatch.

**Table 5.2**: *Parameter values used in this reduced order model*[3].

| | |
|---|---|
| $\beta$ | 0.0075 |
| $\Lambda$ | 0.001s |
| $\lambda$ | 0.1 |
| $C_{pF}$ | 200 $J/kg-K$ |
| $C_{pC}$ | 4000 $J/kg-K$ |
| $m_F$ | 40000 $kg$ |
| $m_C$ | 7000 $kg$ |
| $W_C$ | 8000 $kg/s$ |
| $a_F$ | $7\times10^6$ $J/m^3$ |
| $\alpha_T^F$ | $-0.00001$ $1/K$ |
| $\alpha_T^C$ | -0.00005 $1/K$ |
| $h$ | $4\times10^6$ $W/m^2-K$ |

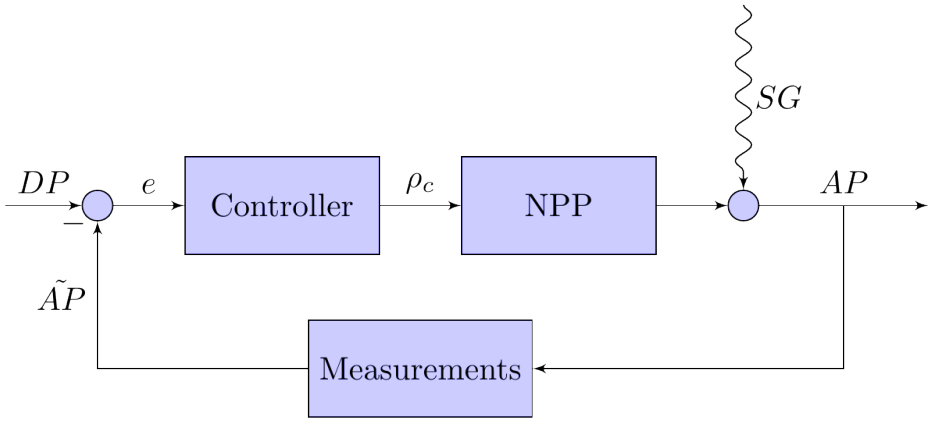



**Figure 5.4**: *Classical output feedback control (AP: Actual Power, DP: Demand Power, SG: Stochastic Generation).*

The reduced-order model equations, along with the proportional control action were numerically solved using the Runge-Kutta-Fehlberg (RK45) method[81;82]. Time step for these calculations was 0.001 sec, and the total simulation time was 150 seconds. The two scenarios are considered to understand the reactor behavior– a) In the first scenario, the demand power level is increased by 13% for 2 minutes and then brought back to the same level. This scenario approximately depicts the controlled load following reactor operations for some European nuclear power plants described earlier in the introduction. b) The second

scenario illustrates the sudden changes in demand power as the wind energy plant and nuclear power plant complement each other to supply the electrical load for a microgrid. The wind energy generation is highly stochastic; therefore, a stochastic and realistic model is developed based on wind speed time-series data to compute the demand power estimates. The methodology to construct energy generation from wind power is described next, which is used to obtain a controller signal. At this juncture, it is worth to mention that a similar model with time-dependent parameters was proposed by Bindra et al.[83] in the context of stability analysis of the boiling water reactors. In his model, various controlling parameters of the model followed an Itô's stochastic differential equation.

**Table 5.3**: *Nomenclature table for the reduced order model.*

| | |
|---|---|
| $n$ | neutron concentration or density |
| $c$ | delayed neutron precursor concentration |
| $\beta$ | delayed neutron fraction |
| $\Lambda$ | average neutron life time |
| $\lambda$ | decay constant |
| $C_{pF}$ | specific heat of fuel |
| $C_{pC}$ | specific heat of coolant |
| $m_F$ | mass of fuel |
| $m_C$ | mass of coolant |
| $W_C$ | coolant mass flow rate |
| $a_F$ | Neutrons to thermal energy factor |
| $\alpha_T^F$ | Doppler feedback coefficient |
| $\alpha_T^C$ | moderator feedback coefficient |
| $h$ | heat transfer coefficient between fuel and coolant |
| $\zeta_{air}$ | Air density |
| $\Gamma_{turbine}$ | Blade cross-sectional area |
| $P(t)$ | Energy flux |
| $P_{elec}(t)$ | Electrical power |
| $B_p$ | Betz factor |
| $\gamma$ | Relaxation parameter |
| $\phi$ | Random forcing |
| $^{17}$O | 12 |

## 5.4 Estimating power generation from wind energy systems

For a given time-dependent wind speed data, the instantaneous electrical energy production rate at the outlet of the wind turbine can be calculated with the following approach. Considering '10' small wind turbines, each of blade radius 12 meters, the kinetic energy flux of the volumetric air can be written as:

$$E(t) = \frac{1}{2}\zeta_{air}(\|u'(t)\|)^2 \tag{5.17}$$

and the corresponding energy flux can be written as

$$P(t) = E(t) \times u'(t) \tag{5.18}$$

It is important to understand that Energy flux is the total power available in the air for extraction–which of course is limited by the design factors of the wind turbine. To convert the net available power to the net electrical power Eqn. 5.18 should be multiplied by the cross-sectional area of the wind turbine blades ($\Gamma_{turbine}$) and by the efficiency factor or the power coefficient ($B_p$) of the wind turbine. This factor ($B_p$) is also known as 'Betz' factor, which has an upper limit of 59 %. Assuming a 'Betz' factor of 0.3 or a 30% efficient turbine:

$$P_{elec}(t) = P(t) \times 0.3 \times \Gamma_{turbine} \tag{5.19}$$

### 5.4.1 Statistical model of wind speed

To calculate the kinetic energy of the wind (Eqn. 5.17) the raw velocity data, which can be fitted into a statistical model. From the perspective of the economic dispatch of electrical power and the operation of the power systems, it is required to predict the wind speed on a shorter time scales as compared to reported data[84;85]. Typically for that reason a stochastic differential equation-based model is used which assumes that the wind time series with some

known statistical parameters is totally random and there are no memory effects involved (see equation 5.20).

$$\frac{du'}{dt} = -\gamma u' + C^{0.5}\phi \tag{5.20}$$

The above model is known as Ornstein–Uhlenbeck (OU) (or a Langevin model and or an Itô's sde) equation which is a Gauss-Markov process. The Markovian nature of the model is attributed to the fact that the current value of the variable $u'$ only depends on the value immediately behind it in time. In this model, '$\gamma$' is the relaxation parameter which governs how fast the process reaches its equilibrium value. While a high value of '$\gamma$' guarantees that the process is extremely quick in reaching its equilibrium state, a lower value of it ensures a slow decay. The third and fourth important parameter in the model is '$\phi$' & '$C$' which are the random forcing term and the amplitude multiplier respectively. The term '$\phi$' is random in the sense that it has no correlation with its past values.

The process described in Eqn. 5.20 is said to be stationary in a sense that its properties are unaffected by change of time origin, that is the the joint distribution of the '$n$' observations made at any times $(u'_{t_1}, u'_{t_2}, ..., u'_{t_n})$ is the same as that associated with '$n$' observations made at time times $(u'_{t_{1+k}}, u'_{t_{2+k}}, ..., u'_{t_{n+k}})$. In other words, the joint distribution of variables will not get affected by shifting all the times of observation forward or backward by an integer amount '$k$'.

A simple OU process with a constant value of '$\gamma$' is used such that the process relaxes not too fast or not too slow, representing a more realistic wind speed scenario. The variance of the force '$\phi$' is set as per the fluctuation-dissipation theorem. This theorem relates the strength of the fluctuating force term with the deterministic part of the solution, i.e. '$\gamma$.'

$$C = 2\langle u'^2 \rangle \gamma \tag{5.21}$$

In other words, fluctuation-dissipation theorem expresses a balance between the friction, which tends to drive the system to a dead state, and noise, which tends to keep the system alive[68].

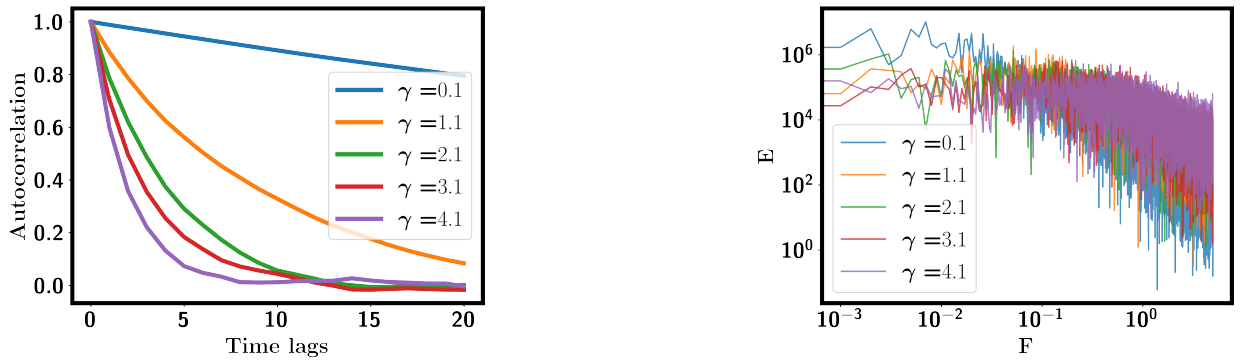



**Figure 5.5**: *Comparison of autocorrelation and spectra of the OU process at various values of 'γ': at higher values of 'γ' the process relaxes faster while the spectra of the process will become flatter leading to a random white noise signal.*

### 5.4.2 Load curve for NPP

The value of relaxation parameter, i.e. 'γ' is set to 0.25, which is a somewhat arbitrary choice. However, as mentioned above a high value of gamma brings the process more towards a state where each realization of it is completely decoupled from the previous one (see figure 5.5). Therefore a relatively higher lower value of it is chosen. While the value of force multiplier is computed using the fluctuation-dissipation relationship (see equation 5.21).

Once all the values of the OU process were calculated, a standard Monte-Carlo program was used to obtain model realizations. These realizations of the synthetically generated time series are normally distributed but the real wind time series is Weibull distributed. Thus, the probability distribution function of the generated series should be transformed. This can be achieved by the following process,

$$f(\tilde{u}) = p(\tilde{y})\frac{d\tilde{y}}{du} \tag{5.22}$$

$$\int_a^b f(\tilde{u})d\tilde{u} = \tilde{y}, \forall \tilde{y} \in \text{uniform distribution} \tag{5.23}$$

$$\text{if } f(\tilde{u}) \text{ is weibull distributed, then } \tilde{y} = 1 - e^{(-\frac{\tilde{u}}{\lambda})^p} \tag{5.24}$$

$$\text{if } \tilde{y} \in \mathcal{N} \text{ and } \tilde{u} \text{ is uniform, } \tilde{u} = 0.5(1 + (\frac{\tilde{y} - \mu}{\sigma\sqrt{2}})) \tag{5.25}$$

Using the above two results one can get: $\tilde{u} = \lambda(-\log((1 - \frac{\tilde{y}-\mu}{\sigma\sqrt{2}})))^{\frac{1}{p}}$ [86;87], to convert a normally distributed variable $\tilde{y}$ to a weibull distributed variable.

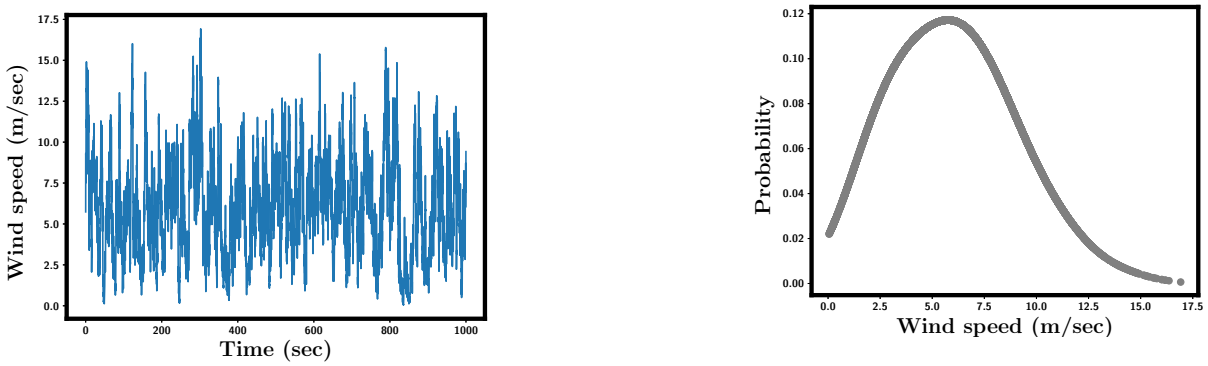



**Figure 5.6**: *Generated wind speed time series and the probability distribution.*

The power output was predicted from the wind speed distributions and then used in the microgrid model to obtain instantaneous total demand power (DP) from the reactor using total power (TP) and wind power output.

$$DP = TP - p(t) \tag{5.26}$$

The predicted instantaneous demand power is used to predict reactor behavior using a reduced-order model and controller system described in the previous section.

## 5.5 Numerical Results

First example test for this load-following response involves the power demand increase 13% for a fixed time period of 120 seconds and then followed by a reduction in power demand to the original level. Numerical simulation results for this study are shown in Fig. 5.7, and the response of the system is observed by the temporal change in normalized variables '$x$', '$y$', '$z_F$' and '$z_C$', which are defined in table 5.1.

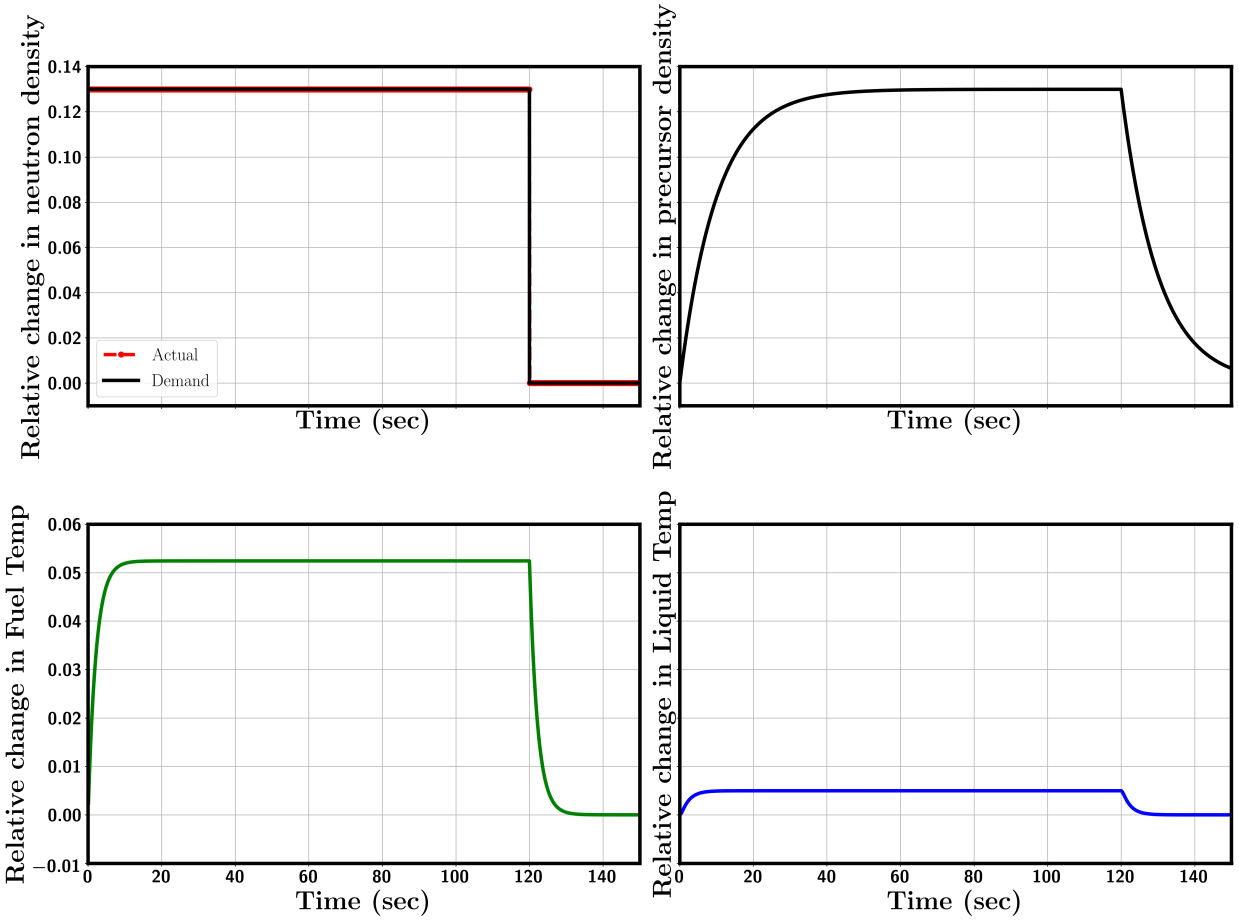



**Figure 5.7**: *Response of the reactor dynamics model ($x$, $y$, $z_F$, $z_C$ from left to right). Time scale is in 'seconds'.*

For a 13% change in the demand power, the fuel temperature jumps from the equilibrium value of 921.88 $K$ to 967.97$K$ over a period of 120 seconds i.e., a relative change of 46.09$K$ while the coolant temperature jumps from 571.88$K$ to 574.73$K$. Thus the coolant experi-

ences a relatively small change in the temperature as compared to the fuel. This is because fluid is flowing at a high mass flow rate and has an additional convection term as compared to solid stationary fuel. Even with a small change in coolant temperature, it is essential to examine its impact on the subcooled nucleate boiling. As this is a lumped parameter model, it is only able to predict an average change in the fuel or coolant temperature, but for more accurate analysis at least 1D models are required to capture the effects of axial or radial peaking.

### 5.5.1 Integrated Micro-grid - Expected Outcomes

In the second scenario, the predicted wind energy production rate (maximum power 13% of the total grid load) and the resulting demand power is used as an input for the model. Due to stochasticity in wind speed described earlier, the resulting demand power input is also stochastic. The reduced-order model simulations show the stochastic effects appear in the resulting solutions for fluid and fuel temperatures. One significantly important aspect to note is that maximum and minimum temperatures of the system remain within limits presented in the results from the first case study. However, the rate of temperature change is also a critical parameter to estimate as thermal fluctuations can lead to thermal fatigue or failures. The maximum rate of fuel temperature change is 22 $K/sec$, and average temperature fluctuations in fuel are around 10-15 $K/sec$. Maximum strain amplitude during these thermal fluctuations is expected to be 0.01%. If the frequency of these fluctuations is much less than the natural frequency of the PWR fuel cladding i.e., 50 Hz, these thermal effects are not expected to cause any major concern to the fuel integrity.

## 5.6 Summary

A reduced-order model of the pressurized water reactor for the simulation of the load following operations is assessed and tested with a proportional controller in the classical feedback control loop. In the first example, the model was used to simulate the variations in the fuel

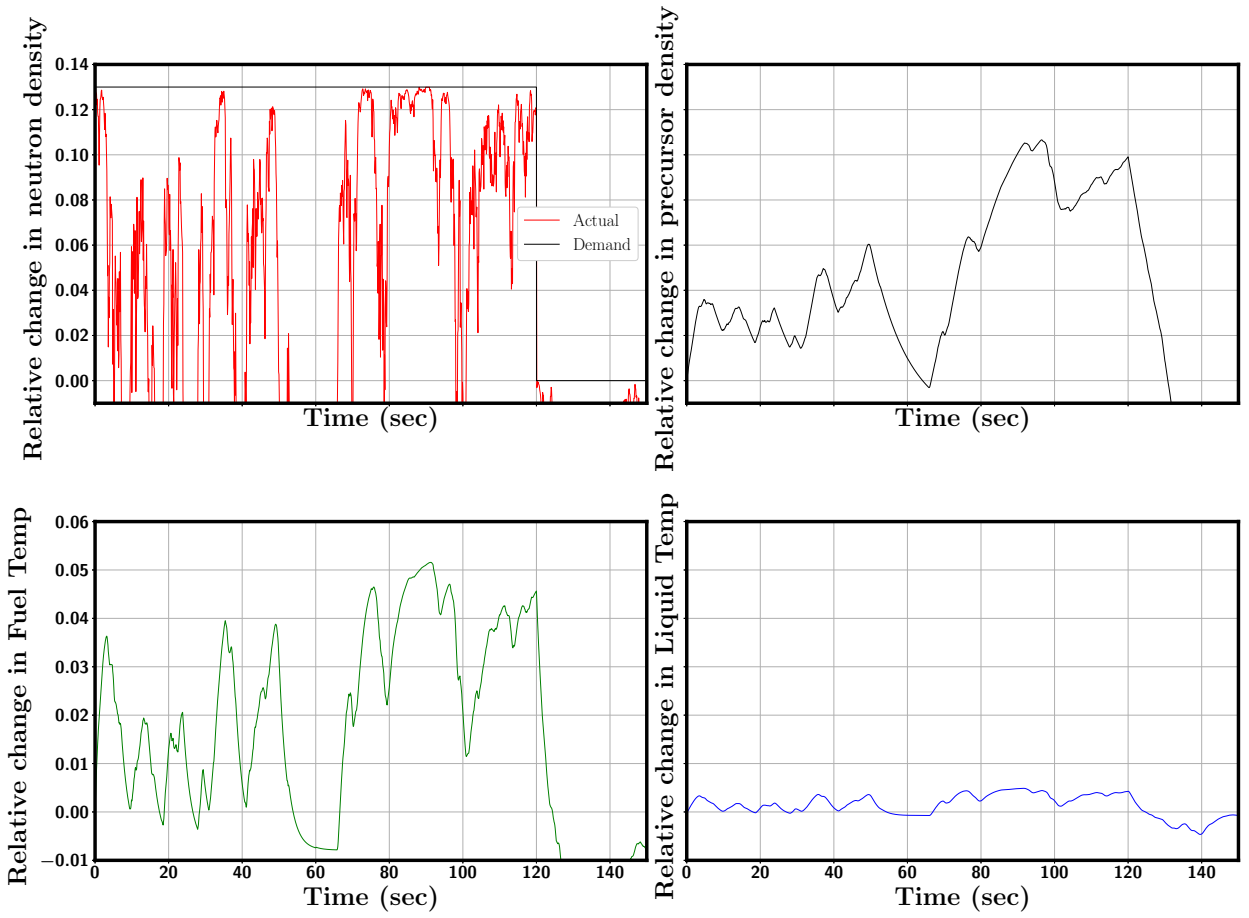



**Figure 5.8**: *Response of the reduced order model owing to stochastic & deterministic generation.*

and coolant temperature following a step change in demand power governed by the grid. The response of the reactor system to this 120-second transient shows that the system is capable of achieving new steady state quickly. Change in coolant temperature is anticipated to be much smaller than the change in the fuel temperature.

Wind and Nuclear energy-powered microgrid are studied with this reduced-order model. The realistic wind speed distribution is used to construct an autoregressive model and predict instantaneous power output from a wind energy system. Deficit power in the microgrid is supplied by a nuclear reactor; thus stochastic demand power input is used to predict reactor thermal behavior. Thermal response of fuel and the coolant shows that temperature limits do not exceed the previous case with a maximum power of wind energy system within 13%

of the total grid load. Under these conditions, thermal fluctuations are monitored closely to assess thermal fatigue problems. The maximum strain amplitude is within the failure limits of the nuclear fuel cladding, but it is vital to ensure that the frequency of thermal cycles remains below the natural frequency of the cladding. Although the model results show stable reactor behavior for controlled load-following or a small fraction of highly variable load demand for more accurate analysis particularly related to thermal fluctuations, it is crucial to improve the fidelity of the models.

In future it is advisable to construct a reduced-order model by more rigorous techniques such as- Proper Orthogonal Decomposition (POD) and or Singular Value Decomposition (SVD)–("Law of Parsimony")[88]. The wind speed model considered here is a point estimate of the wind speed data while a more realistic estimate can be more appropriately given by the space time-correlated models. For a large wind generation component, geographical assessment should be considered. Geographical smoothing / coupled behavior of two geographically isolated renewable power plants is inherently intertwined to the grid-scale power fluctuations–which may not (geographically smoothened) or may effect (coupled) the conventional power generators like nuclear.[89].

# Chapter 6

# Conclusion and future work

This thesis is an evolutionary stride towards applying the understanding from non-equilibrium statistical mechanics. The current work applies it to build statistical surrogates to the problems pertaining to the nuclear engineering community. The application of Langevin surrogate was demonstrated on the wall-bounded flow in two different scenarios. The usage of linear Langevin model was demonstrated in two problems important to the community i.e., the radioactive particle transport and realistic looking wind power signal generation for a synergistic existence of the nuclear power plant in a small microgrid. The problem of black-box uncertainty quantification i.e., an externally perturbed deterministic model was tackled by considering the intrinsic aleatoric behavior of the system.

In this thesis, during the Markovian embedding phase, moments of the data were computed via the Kramers-Moyal expansion method. This leads to a stationary non-linear Langevin equation with fluctuation dependent parameters. This non-linear model was found to be in good agreement with the momentum and scalar transport data obtained from Direct Numerical Simulation (DNS) of Navier-Stokes and energy transport equations.

The statistical mechanics emulators were applied and demonstrated for uncertainty quantification and system controls in the field of nuclear engineering. The spatio-temporal uncertainties were quantified for the scalar dispersion in the geosphere by developing stochastic models of different levels of complexity. Particularly, the effect of different noise processes,

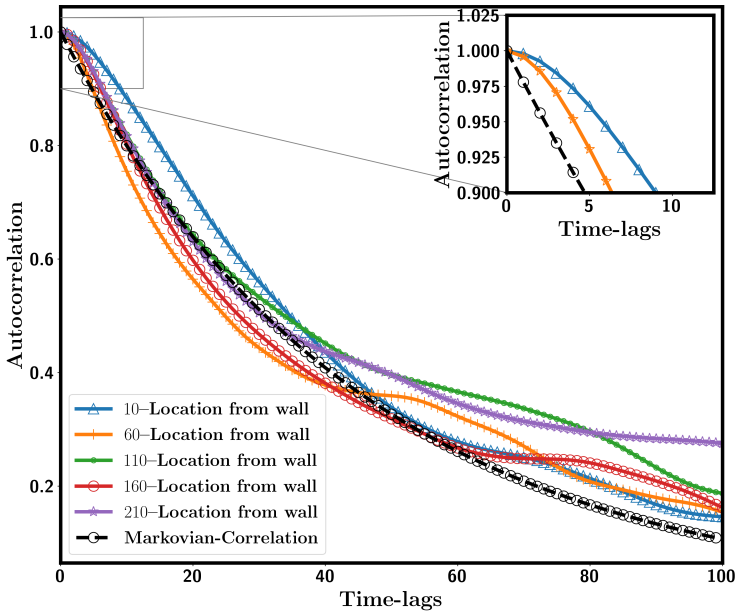



**Figure 6.1**: *Lagrangian correlation function of the streamwise velocity of turbulent channel flow. The correlation function becomes less flat at shorter time scales as one goes away from the wall. The Markovian models have exponentially decaying correlation which cannot capture the short-time behavior of the turbulent signal.*

colloidal transport, and radioactive decay were investigated. The effective surrogate for the renewable generation and for the grid demand was obtained and was integrated with the well-established reactor models to obtain a novel stochastic control methodology.

The major goal of the presented approach is to drive the physical dynamics of the system via a more accurate representation of the uncertain part of the solution i.e., the chaotic motion of the unresolved part of the physics of the system. The data-driven approach demonstrated in this thesis embed the 'Markovian' approximation to the unresolved or uncertain part of the dynamics. While these are a better representation of the reality than simply perturbing the linearized dynamics via a white in time forcing. They still lack one key feature found in the relevant system of interests i.e., the lack of fully random behavior at shorter time scales. This point can be made clear by looking at the flatness (see figure 6.1) of the velocity correlation function of the streamwise velocity of turbulent channel flow[54] at shorter time scales (see inset in the figure 6.1) and comparing the same with the Markovian model correlations. For the case of systems commonly found in nuclear engineering appli-

cations, these highly correlated structures dictate the fluctuations in the signal at shorter time scales–which translates to the presence of ballistic trajectories on shorter time scales than the completely random one. Therefore this separation of time scales leads to two sets of variables i.e., a) the 'slow variables' which decays slower than the pure Markovian assumption[68;90] b) fast decaying Markovian variables. In this thesis, this correction to the Markovian assumption was rather treated by passing the Langevin model output through a Gaussian filter to artificially introduce this flatness of correlation function. This phenomena commonly found in natural systems can be captured more rigorously by invoking a more involved theoretical framework, Mori-Zwanzig projection formalism[68;77;91–93]. This formalism introduces a memory effect in an otherwise completely randomly perturbed representation of the signal. The rationale behind such a modeling approach is that those neglected short time scale fluctuations ('slow: or slow decaying'[68;90]) are not completely decoupled from the large time scale ('fast: or fast decaying'[68;90]) fluctuations modeled in the Markovian system. In fact, they are intrinsically coupled to those modeled fluctuations in the Langevin scheme. This approach then directly leads to two tightly coupled system of equations leading to the introduction of memory term, which has much similarity with the convolution integral.

The moments computed via the Kramers-Moyal expansion method rely heavily on the computed conditional probabilities and the extrapolation methodologies. Therefore these fitted Kramers-Moyal parameters shall be dealt more carefully, and a proper sensitivity analysis on the same should be performed to quantify the errors creeping in during the fitting procedure.

Being evolutionary in nature this thesis paves a way to the rigorous understanding of the natural systems in the nuclear engineering community via the application of understanding from non-equilibrium statistical mechanics i.e., the Kramers-Moyal expansion method and the fluctuation-dissipation theorem. This can be applied to a variety of problems relevant to the community (e.g., radioactive particle transport, turbulent flows inside a nuclear reactor, reactor noise analysis, synergistic existence of stochastic generators with nuclear power, etc.) where there is a clear need of physics-driven low complexity data-driven surrogates.

# Bibliography


[1] NERC Intermittent and Variable Generation Task Force. Accommodating high levels of variable generation. *North American Electric Reliability Corp.(NERC)*, 2009.

[2] C Bruynooghe, A Eriksson, and G Fulli. Operation and maintenance (o&m) costs. compatibility with wind power variability. 2010.

[3] Henryk Anglart. *Nuclear reactor dynamics and stability*. KTH Royal Institute of Technology, 2011.

[4] Roman Frigg and Stephan Hartmann. Models in science. 2006.

[5] Erica Lucy Thompson. *Modelling North Atlantic storms in a changing climate*. PhD thesis, Imperial College London, 2013.

[6] Brian M Adams, Mohamed S Ebeida, Michael S Eldred, John D Jakeman, Laura P Swiler, J Adam Stephens, Dena M Vigil, Timothy M Wildey, William J Bohnhoff, Keith R Dalbey, et al. Consortium for advanced simulation of lwrs. 2014.

[7] Andrea Alfonsi, Cristian Rabiti, Diego Mandelli, Joshua Cogliati, Robert Kinoshita, and Antonio Naviglio. Raven and dynamic probabilistic risk assessment: Software overview. Technical report, Idaho National Lab.(INL), Idaho Falls, ID (United States), 2014.

[8] GK Batchelor. Diffusion in a field of homogeneous turbulence: Ii. the relative motion of particles. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 48, pages 345–362. Cambridge University Press, 1952.

[9] PA Durbin. A stochastic model of two-particle dispersion and concentration fluctuations in homogeneous turbulence. *Journal of Fluid Mechanics*, 100(2):279–302, 1980.

[10] Hannes Risken. Fokker-planck equation. In *The Fokker-Planck Equation*, pages 63–95. Springer, 1996.

[11] Timothy Hickey, Qun Ju, and Maarten H Van Emden. Interval arithmetic: From principles to implementation. *Journal of the ACM (JACM)*, 48(5):1038–1068, 2001.

[12] Alan C Calder, Melissa M Hoffman, Donald E Willcox, Maximilian P Katz, F Douglas Swesty, and Scott Ferson. Quantification of incertitude in black box simulation codes. In *Journal of Physics: Conference Series*, volume 1031, page 012016. IOP Publishing, 2018.

[13] IAEA. International Atomic Energy Agency(IAEA) Report, Safety Analysis for Research Reactors. 2008.

[14] IAEA. International Atomic Energy Agency(IAEA) Report, Best Estimate Safety Analysis for Nuclear Power Plants: Uncertainty Evaluation. 2008.

[15] Augusto Hernández-Solís. *Uncertainty and sensitivity analysis applied to LWR neutronic and thermal-hydraulic calculations*. PhD thesis, 2012.

[16] Ting Zhu. Sensitivity and uncertainty analysis of (Th, Pu) O2 benchmark experiments in ZED-2 using TSUNAMI. 2011.

[17] Francesco DAuria, Anis Bousbia-Salah, Alessandro Petruzzi, and Alessandro Del Nevo. State of the art in using best estimate calculation tools in nuclear technology. *Nuclear Engineering and Technology*, 38(1):11, 2006.

[18] Judith F Briesmeister et al. **MCNP™**-A general Monte Carlo N-particle transport code. *Los Alamos National Laboratory*, 2000.

[19] Information systems laboratories Inc. Relap5/mod3.3 patch 04 code manual volume i: Code structure, system models, and solution methods, 2010.

[20] Wade R Marcum. Thermal hydraulic analysis of the Oregon State TRIGA Reactor using RELAP5-3D, 2008.

[21] US Nuclear Regulatory Commission et al. Acceptance criteria for emergency core cooling systems for light-water nuclear reactors, 1978.

[22] BE Boyack, I Catton, RB Duffey, KR Katsma, GS Lellouche, S Levy, GE Wilson, and N Zuber. Quantifying reactor safety margins part 1: an overview of the code scaling, applicability, and uncertainty evaluation methodology. *Nuclear Engineering and Design*, 119(1):1–15, 1990.

[23] Saqr Mofleh Alshogeathri. Application of the reactivity method on ksu triga fuel. Master's thesis, Kansas State University, 2017.

[24] COMSOL Multiphysics. Comsol multiphysics user guide (version 5.2). *COMSOL, AB*, pages 39–40, 2014.

[25] A. Gairola, D. Gould, J. Guether, J. Roberts, and H. Bindra. Thermal hydraulics safety analyses of triga reactor upon fuel reconfiguration. *Transactions of American Nuclear Society*, 2016.

[26] Amir Zacarias Mesquita, Luiz Carlos Duarte Ladeira, Daniel Artur Pinheiro Palma, and Maritza Rodríguez Gual. Uncertainty assessment of the experimental thermal-hydraulic parameters for cdtn triga research reactor. *International Journal of Nuclear Energy Science and Technology*, 10(3):201–216, 2016.

[27] SV Karmare and AN Tikekar. Heat transfer and friction factor correlation for artificially roughened duct with metal grit ribs. *International Journal of Heat and Mass Transfer*, 50(21):4342–4351, 2007.

[28] Tae-Hyun Chun and Dong-Seok Oh. A pressure drop model for spacer grids with and without flow mixing vanes. *Journal of nuclear science and technology*, 35(7):508–510, 1998.

[29] P Sivashanmugam and S Suresh. Experimental studies on heat transfer and friction factor characteristics of laminar flow through a circular tube fitted with regularly spaced

helical screw-tape inserts. *Experimental Thermal and Fluid Science*, 31(4):301–308, 2007.

[30] Jin-Sheng Leu, Ying-Hao Wu, and Jiin-Yuh Jang. Heat transfer and fluid flow analysis in plate-fin and tube heat exchangers with a pair of block shape vortex generators. *International Journal of Heat and Mass Transfer*, 47(19):4327–4338, 2004.

[31] A Gairola, S Alshogeathri, Y Cheng, D Laramore, H Bindra, and J Roberts. Estimating burnup uncertainties in ksu triga mark ii reactor and their impact on reactor safety. *Progress in Nuclear Energy*, 99:177–184, 2017.

[32] Uriel Frisch. *Turbulence: the legacy of AN Kolmogorov*. Cambridge university press, 1995.

[33] Katepalli R Sreenivasan. Turbulent mixing: A perspective. *Proceedings of the National Academy of Sciences*, page 201800463, 2018.

[34] Geoffrey I Taylor. Diffusion by continuous movements. *Proceedings of the london mathematical society*, 2(1):196–212, 1922.

[35] DR Osborne, JC Vassilicos, K Sung, and JD Haigh. Fundamentals of pair diffusion in kinematic simulations of turbulence. *Physical Review E*, 74(3):036309, 2006.

[36] Marie-Caroline Jullien, Jérôme Paret, and Patrick Tabeling. Richardson pair dispersion in two-dimensional turbulence. *Physical review letters*, 82(14):2872, 1999.

[37] Philip JW Roberts and Donald R Webster. *Turbulent diffusion*. ASCE Press, Reston, Virginia, 2002.

[38] Chen Wang, Xu Wu, and Tomasz Kozlowski. Surrogate-based inverse uncertainty quantification of trace physical model parameters using steady-state psbt void fraction data. *Proc. 17th Int. Topl. Mtg. Nuclear Reactor Thermal Hydraulics (NURETH-17)*, pages 3–8, 2017.

[39] Dongli Huang, Hany Abdel-Khalik, Cristian Rabiti, and Frederick Gleicher. Dimensionality reducibility for multi-physics reduced order modeling. *Annals of Nuclear Energy*, 110:526–540, 2017.

[40] Xu Wu and Tomasz Kozlowski. Inverse uncertainty quantification of reactor simulations under the bayesian framework using surrogate models constructed by polynomial chaos expansion. *Nuclear Engineering and Design*, 313:29–52, 2017.

[41] Yang Liu and Nam Dinh. Validation and uncertainty quantification for wall boiling closure relations in multiphase-cfd solver. *Nuclear Science and Engineering*, 193(1-2): 81–99, 2019.

[42] George J Klir and Baozung Yuan. *Fuzzy sets and fuzzy logic: theory and applications*, volume 574. Prentice Hall PTR New Jersey, 1995.

[43] JM Booker and TJ Ross. An evolution of uncertainty assessment and quantification. *Scientia Iranica*, 18(3):669–676, 2011.

[44] Warren Weaver. Science and complexity. In *Facets of systems science*, pages 449–456. Springer, 1991.

[45] John C Hull. *Options futures and other derivatives*. Pearson Education India, 2003.

[46] Lawrence C Evans. *An introduction to stochastic differential equations*, volume 82. American Mathematical Soc., 2012.

[47] Frederick Reif. *Fundamentals of statistical and thermal physics*. Waveland Press, 2009.

[48] Rudolf Friedrich and Joachim Peinke. Description of a turbulent cascade by a fokker-planck equation. *Physical Review Letters*, 78(5):863, 1997.

[49] Stephen B Pope. *Turbulent flows*. Cambridge university press, 2000.

[50] Stephen B Pope. Simple models of turbulent flows. *Physics of Fluids*, 23(1):011301, 2011.

[51] Z Warhaft. The interference of thermal fields from line sources in grid turbulence. *Journal of Fluid Mechanics*, 144:363–387, 1984.

[52] Christoph Renner, Joachim Peinke, and Rudolf Friedrich. Experimental indications for markov properties of small-scale turbulence. *Journal of Fluid Mechanics*, 433:383–409, 2001.

[53] Ch Renner, J Peinke, R Friedrich, et al. The markov–einstein coherence lengtha new meaning for the taylor length in turbulence. *Physics Letters A*, 359(5):335–338, 2006.

[54] Lee M. Malaya N. Moser R.D. Eyink G. Meneveau C. Graham, J. Turbulent channel flow data set. *The JHU Turbulence Databases (JHTDB)*.

[55] J Graham, K Kanov, XIA Yang, M Lee, N Malaya, CC Lalescu, R Burns, G Eyink, A Szalay, RD Moser, et al. A web services accessible database of turbulent channel flow and its use for testing a new integral wall model for les. *Journal of Turbulence*, 17(2): 181–215, 2016.

[56] Sylvain Laizet and Eric Lamballais. High-order compact schemes for incompressible flows: A simple and efficient method with quasi-spectral accuracy. *Journal of Computational Physics*, 228(16):5989–6015, 2009.

[57] A Gairola and H Bindra. Statistical emulator models for scalar mixing in reactor safety analysis. *18th International Topical Meeting on Nuclear Reactor Thermal Hydraulics - NURETH*, 2019.

[58] A Gairola, W Guo, B Niceno, and H Bindra. Reduced order models based on markovian and non-markovian frameworks. *Bulleting of American Physical Society*, 2018.

[59] D Gould, A Gairola, and H Bindra. Analyzing air-ingress currents in gas-cooled nuclear reactors using kramers moyal expansion. *Bulleting of American Physical Society*, 2018.

[60] A Gairola and H Bindra. Matryoshka inspired statistical surrogate for turbulent mixing.

*Symposium on Advanced Sensors and Modeling Techniques for Nuclear Reactor Safety*, 2018.

[61] G Grötzbach. Challenges in low-prandtl number heat transfer simulation and modelling. *Nuclear engineering and design*, 264:41–55, 2013.

[62] Wonjung Lee. Generalized langevin equation and the linear regression model with memory. *Physical Review E*, 98(2):022137, 2018.

[63] A Gairola and H Bindra. Uncertainty quantification and sensitivity analysis for the environmental dispersion of contaminants. *VERIFICATION AND VALIDATION SYMPOSIUM*, 2018.

[64] Yukio Masumoto, Yasumasa Miyazawa, Daisuke Tsumune, Takaki Tsubono, Takuya Kobayashi, Hideyuki Kawamura, Claude Estournel, Patrick Marsaleix, Lyon Lanerolle, Avichal Mehra, et al. Oceanic dispersion simulations of 137cs released from the fukushima daiichi nuclear power plant. *Elements*, 8(3):207–212, 2012.

[65] Jessica Kratchman and Chuck Norton. Fukushima water contamination- impacts on the u.s. west coast. *US, NRC*, 2015.

[66] EM Kirova and GE Norman. Viscosity calculations at molecular dynamics simulations. In *Journal of Physics: Conference Series*, volume 653, page 012106. IOP Publishing, 2015.

[67] Mesut Cayar and M Levent Kavvas. Ensemble average and ensemble variance behavior of unsteady, one-dimensional groundwater flow in unconfined, heterogeneous aquifers: an exact second-order model. *Stochastic Environmental Research and Risk Assessment*, 23(7):947–956, 2009.

[68] Robert Zwanzig. *Nonequilibrium statistical mechanics*. Oxford University Press, 2001.

[69] Daniel Zwillinger. *Handbook of differential equations*, volume 1. Gulf Professional Publishing, 1998.

[70] Martin Burger and Marie-Christin Tacke. Mathematische modelle fur phasenseparation bei aktiven teilchen. 2015.

[71] BA Boughton, JM Delaurentis, and WE Dunn. A stochastic model of particle dispersion in the atmosphere. *Boundary-Layer Meteorology*, 40(1-2):147–163, 1987.

[72] Shlomo Havlin and Daniel Ben-Avraham. Diffusion in disordered media. *Advances in Physics*, 36(6):695–798, 1987.

[73] Radek Erban, Jonathan Chapman, and Philip Maini. A practical guide to stochastic simulations of reaction-diffusion processes. *arXiv preprint arXiv:0704.1908*, 2007.

[74] X Sanchez-Vila and Daniel Fernàndez-Garcia. Debates stochastic subsurface hydrology from theory to practice: Why stochastic modeling has not yet permeated into practitioners? *Water Resources Research*, 52(12):9246–9258, 2016.

[75] Berwin A Turlach et al. *Bandwidth selection in kernel density estimation: A review*. Université catholique de Louvain Louvain-la-Neuve, 1993.

[76] Gang Chen, Markus Flury, James B Harsh, and Peter C Lichtner. Colloid-facilitated transport of cesium in variably saturated hanford sediments. *Environmental science & technology*, 39(10):3435–3442, 2005.

[77] Burkhard MO Heppe. Generalized langevin equation for relative turbulent dispersion. *Journal of Fluid Mechanics*, 357:167–198, 1998.

[78] Frank Huntowski, Aaron Patterson, and Michael Schnitzer. Negative electricity prices and the production tax credit. *The North Bridge Group (September)*, 2012.

[79] Alexey Lokhov. Technical and economic aspects of load following with nuclear power plants. *NEA, OECD, Paris, France*, 2011.

[80] Abhinav Gairola and Hitesh Bindra. Nuclear-wind powered microgrid: Reduced-order modeling of lwr response. In *Storage and Hybridization of Nuclear Energy*, pages 31–48. Elsevier, 2019.

[81] Erwin Fehlberg. Low-order classical runge-kutta formulas with stepsize control and their application to some heat transfer problems. 1969.

[82] Gerhard Wanner and Ernst Hairer. Solving ordinary differential equations ii. *Stiff and Differential-Algebraic Problems*, 1991.

[83] Hitesh Bindra and Rizwan-uddin. Effects of modeling assumptions on the stability domain of bwrs. *Annals of Nuclear Energy*, 67:13–20, 2014.

[84] Xinxin Zhu and Marc G Genton. Short-term wind speed forecasting for power system operations. *International Statistical Review*, 80(1):2–23, 2012.

[85] Pierre Pinson et al. Wind energy: Forecasting challenges for its operational management. *Statistical Science*, 28(4):564–585, 2013.

[86] Daniel Villanueva, Andrés Feijóo, and José Pazos. Multivariate weibull distribution for wind speed and wind power behavior assessment. *Resources*, 2(3):370–384, 2013.

[87] Daniel Villanueva, José Luis Pazos, and Andrés Feijoo. Probabilistic load flow including wind power generation. *IEEE Transactions on Power Systems*, 26(3):1659–1667, 2011.

[88] Athanasios C Antoulas. *Approximation of large-scale dynamical systems*. SIAM, 2005.

[89] Mahesh M Bandi. Spectrum of wind power fluctuations. *Physical Review Letters*, 118 (2):028301, 2017.

[90] George DJ Phillies. *Elementary lectures in statistical mechanics*. Springer Science & Business Media, 2012.

[91] Kyozi Kawasaki. Simple derivations of generalized linear and nonlinear langevin equations. *Journal of Physics A: Mathematical, Nuclear and General*, 6(9):1289, 1973.

[92] Hazime Mori. Transport, collective motion, and brownian motion. *Progress of theoretical physics*, 33(3):423–455, 1965.

[93] Alexandre Chorin and Panagiotis Stinis. Problem reduction, renormalization, and memory. *Communications in Applied Mathematics and Computational Science*, 1(1):1–27, 2007.

[94] The $\delta$-function convolution.impulse response transfer function. https://www.robots.ox.ac.uk/~dwm/Courses/2TF_2011/2TF-N3.pdf. Accessed: 2019-05-28.

[95] The running-sum lowpass filter. https://ccrma.stanford.edu/~jos/sasp//Running_Sum_Lowpass_Filter.html. Accessed: 2019-05-28.

# Appendix A

# Data-Filtering

## A.1 High pass filtering–via differencing

Higher frequency oscillations can be directly separated by applying a high pass filter, which can be defined easily via differencing the data series. Following shows that it is equivalent of applying a high pass filter. Defining the velocity increments

$$L(v(t) = v(t-1) \tag{A.1}$$

$$v(t) - Lv(t) = v(t) - v(t-1) \tag{A.2}$$

$$(1-L)v(t) = v(t) - v(t-1) \tag{A.3}$$

$$\tag{A.4}$$

Fourier transform of the following equation

$$(1-L)v(t) = v(t) - v(t-1) \tag{A.5}$$

$$\mathscr{F}((1-L)v(t)) = \mathscr{F}(v(t) - v(t-1)) \tag{A.6}$$

$$(\mathscr{F}(v(t-1)) = e^{-j\omega}v(j\omega)) \tag{A.7}$$

Gain is defined as:

$$G(j\omega) = \frac{Fourier - output}{Fourier - input} \tag{A.8}$$

$$\frac{\mathscr{F}((1-L)v(t))}{\mathscr{F}v(t)} = \frac{\mathscr{F}(v(t)) - e^{-j\omega}\mathscr{F}v(t)}{\mathscr{F}v(t)} \tag{A.9}$$

$$G(j\omega) = 1 - e^{-j\omega} \tag{A.10}$$

$$|G(j\omega)|^2 = |1 - e^{-j\omega}|^2 \tag{A.11}$$

$$|G(j\omega)|^2 = ((1 - e^{-j\omega}) \times (1 - e^{j\omega})) \tag{A.12}$$

$$|G(j\omega)|^{=} 4 \times sin^2\omega/2 \tag{A.13}$$

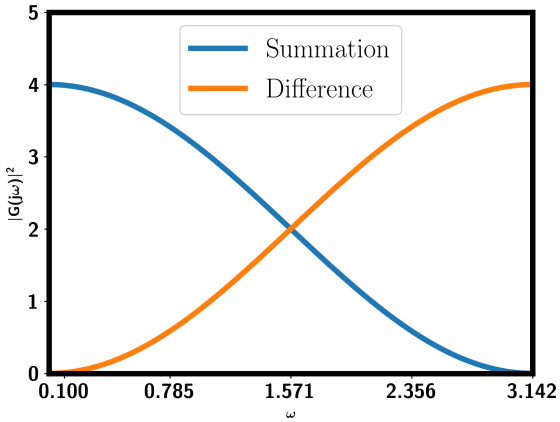



**Figure A.1**: *Comparison of the gain of difference and moving average filter (see equation A.15)–it will let go the higher and lower frequencies oscillations respectively. This also means the differencing and moving average or convolution integral are complementary filters to each other.*

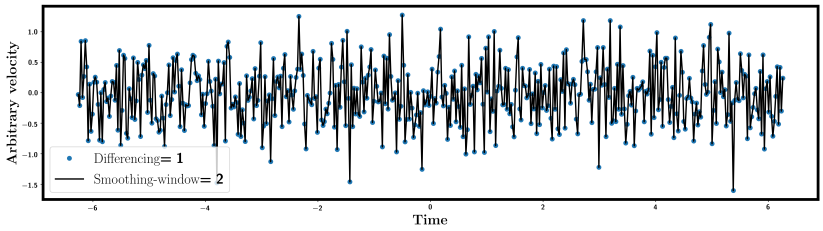



**Figure A.2**: *Two point moving average filter and differencing.*

## A.2 Summation or Moving average–Inverse of differencing

The inverse of differencing is the summation or the moving average operator

$$(1-L)^{-1} = \underbrace{(1+L+L^2+L^3+...)}_{Binomial-expansion} \tag{A.14}$$

$$(1+L+L^2+L^3..)*u(t) = u(t)+u(t-1)+u(t-2)+... = \int u(\tau)\delta(t-\tau)d\tau \text{ [94]} \tag{A.15}$$

$$\mathscr{F}(1+L+L^2+L^3..)*u(t) = u(\omega)(1+e^{j\omega}+e^{2j\omega}+...) \tag{A.16}$$

Therefore the squared magnitude of the gain of the filter truncated to the first term (i.e. two point moving average formula) can be written as:

$$|G(j\omega)|^2 = (1+e^{j\omega})*(1+e^{-j\omega}) \tag{A.17}$$

Increasing the number of points in the summation will give a good time response. However, the frequency response will become poor.[95]. This all is conveniently defined in terms of a sliding window or a convolution operation. The number of points inside in each box governs the time and frequency response of the process.

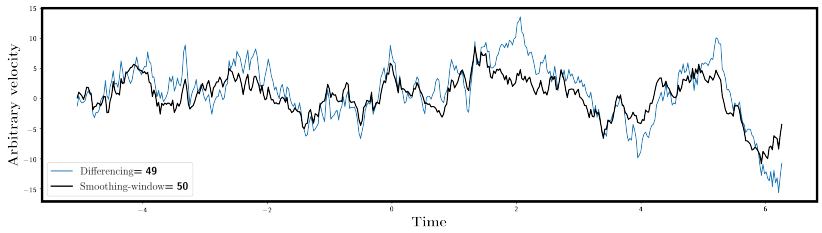



**Figure A.3**: *Fifty point moving average filter and differencing.*

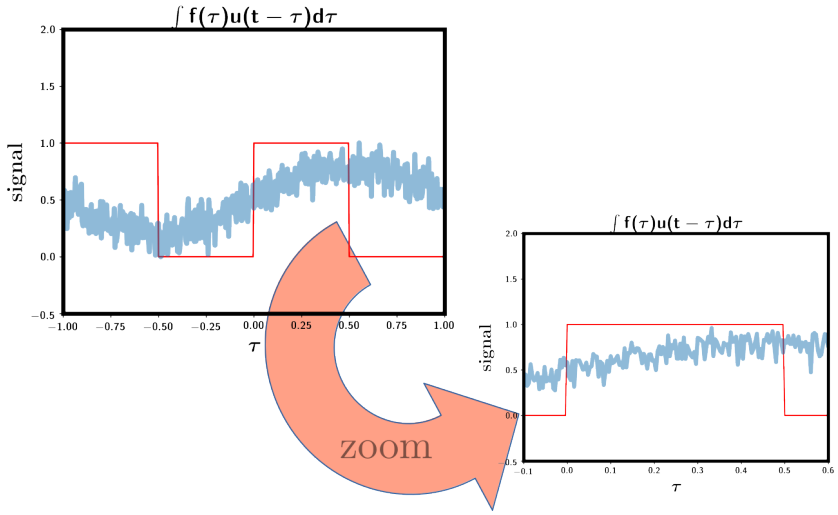



**Figure A.4**: *Convolution Integral done by moving the window 'u(k)' along the axis 'k', a window is shown in the zoomed plot and the Fourier transform is performed to show the response of the box filter.*

## A.3 Amplitude

Time smoothing and removing the fluctuations will decrease the amplitude of the extracted signal. The reason for that can be understood from this proof:

$$(1-L)^{-1}v(\omega) = \frac{1}{2}v(\omega)(1+e^{j\omega}) \tag{A.18}$$

$$\epsilon(\omega) = v(\omega) - \frac{1}{2}v(\omega)(1+e^{j\omega}) \tag{A.19}$$

$$\epsilon(t) = \frac{1}{2}(v(t) - v(t-1)) \tag{A.20}$$

which is exactly half of the fluctuations obtained by differencing.

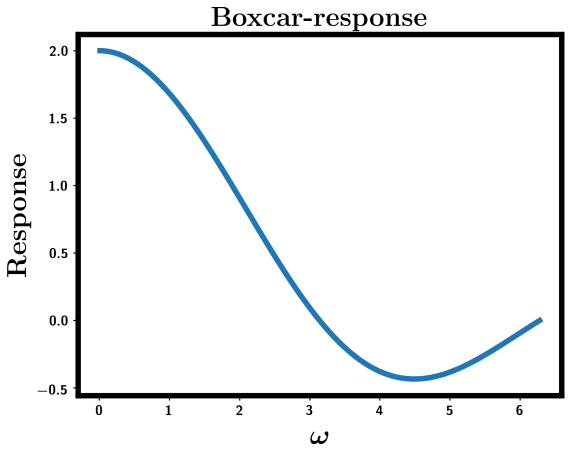



**Figure A.5**: *Box-car filter response–covers the lower end of the spectrum–depending on the length of the window the frequency response of the filter changes.*

# Appendix B

# Fluctuation dissipation relation

For the relaxing dynamics, it is possible to write an equation of sort:

$$\frac{dx}{dt} = -\gamma x(t) + C\phi(t) \tag{B.1}$$

Here, $\phi$ is the stochastic forcing with the following properties:

$$\langle \phi(t) \rangle = 0 \tag{B.2}$$

$$\langle \phi(t)\phi(t') \rangle = \delta(t - t') \tag{B.3}$$

Equation B.1 can be solved by multiplying with the integrating factor $e^{\gamma t}$

$$x(t) = x(0)e^{-\gamma t} + \int_0^t C\phi(t')e^{-\gamma(t-t')}dt' \tag{B.4}$$

Since $x(t)$ is stochastic; therefore, the general quantity of interest, which measures its deviation from the mean is the variance of the process. This can be computed by squaring the above equation B.4. This process will generate three terms one deterministic and two stochastic. One stochastic term is first order in noise; therefore, ensemble averaging will

remove that term[68]. What is important to look at the term which is second order in noise:

$$\int_0^t dt' e^{-\gamma(t-t')} C\phi(t') \int_0^t dt'' e^{-\gamma(t-t'')} C\phi(t'') \tag{B.5}$$

Equation B.5 can be simplified by pulling the noise term inside the integral on the right side and doing the ensemble averaging.

$$\int_0^t dt' e^{-\gamma(t-t')} \int_0^t dt'' e^{-\gamma(t-t'')} C^2 \langle \phi(t'')\phi(t') \rangle \tag{B.6}$$

$$\int_0^t dt' e^{-\gamma(t-t')} \int_0^t dt'' e^{-\gamma(t-t'')} C^2 \delta(t'-t'') \rangle \tag{B.7}$$

Therefore,

$$\langle x(t)^2 \rangle = x(0)^2 e^{-2\gamma t} + \int_0^t e^{-2\gamma(t-t')} C^2 dt' \tag{B.8}$$

$$\langle x(t)^2 \rangle = x(0)^2 e^{-2\gamma t} + \frac{C^2}{2\gamma}(1 - e^{-2\gamma t}) \tag{B.9}$$

At long time the exponential drops:

$$\langle x(t)^2 \rangle = \frac{C^2}{2\gamma} \tag{B.10}$$

$$C = \sqrt{2\gamma \langle x(t)^2 \rangle} \tag{B.11}$$

Which is the desired fluctuation dissipation relationship.

# Appendix C

# Generalized Langevin equation: Memory kernel computation

Starting from the gLe:

$$\frac{d\Gamma}{dt} = -\int_0^T K(t-\tau)\Gamma(\tau)d\tau + F(t) \tag{C.1}$$

Similar to OU process (see appendix B) it is possible to find a relationship for the correlation function of $\Gamma$.

$$\frac{d\langle\Gamma(0)\Gamma(t)\rangle}{dt} = -\int_0^T K(t-\tau)\langle\Gamma(0)\Gamma(\tau)\rangle d\tau \tag{C.2}$$

From non-Markovian fluctuation dissipation theorem a relationship can be found between the memory kernel and the random force ($F(t)$)[68].

$$K(t) = \frac{\langle F(0)F(t)\rangle}{\langle\Gamma(0)^2\rangle} \tag{C.3}$$

Setting the force $F(t)$ to be exponentially correlated:

$$K(t) = \frac{e^{-\alpha t}}{\langle\Gamma(0)^2\rangle} \tag{C.4}$$

Writing everything in Laplace domain:

$$sCov(s) - Cov(0) = -K(s)Cov(s) \tag{C.5}$$

$$Cov(s) = \frac{Cov(0)}{s + K(s)} \tag{C.6}$$

Here, $Cov(s)$ is the Laplace transform of the correlation function $\langle \Gamma(0)\Gamma(t) \rangle$.

$$K(s) = \frac{1}{\langle \Gamma(0)^2 \rangle (s + \alpha)} \tag{C.7}$$

Inserting the Laplace transform of memory kernel in equation C.6.

$$\frac{Cov(s)}{Cov(0)^2} = \frac{(s + \alpha)}{s(s + \alpha) + \frac{1}{Cov(0)^2}} \tag{C.8}$$

$$\frac{Cov(s)}{Cov(0)^2} = \frac{s + \frac{\alpha}{2}}{(s + \frac{\alpha}{2})^2 + \omega^2} + \frac{\frac{\alpha}{2}}{(s + \frac{\alpha}{2})^2 + \omega^2} \tag{C.9}$$

Inverting the Laplace transform results in:

$$\frac{Cov(t)}{Cov(0)^2} = e^{-t\alpha/2} cos(\omega t) + e^{-t\alpha/2} \frac{\alpha}{2\omega} sin(\omega t) \tag{C.10}$$

Where $\omega$ is defined as:

$$\omega = \sqrt{\frac{1}{Cov(0)^2} - \frac{\alpha^2}{4}} \tag{C.11}$$