Preface

Scope and Purpose of this Book

This monograph arose from my significant confusion about how to understand nervous systems as a physicist. Trying to provide a principled framework for addressing this question led me to a formulation of what we call data assimilation that has applications well beyond my initial inquiry. The point of view presented in this monograph is to view as a communications and dynamical systems problem the general challenge of transferring information in observed data to a physical (or biophysical or geophysical or ...) model of the system producing those data; this is **`data assimilation`**. The model, as the underlying processes, will be nonlinear in some important aspects. This problem is fundamental to how we meld experiments and observations to models, so it is hardly a new issue.

In a sense the problem is quite easy to state: over a period of time $[0,T]$ one makes observations of some properties of a physical system at some discrete times in that interval. Now we step back to think about this system and make a model of its dynamics based on past experience or intuition or whatever. This model can be expressed in differential equations or discrete time rules taking the state from one time to another.

This model could be based on physical principles and basic force laws. We would likely call this a "bottoms-up" approach to modeling the nonlinear dynamics of the processes it seeks to represent. Usually the quantities entering the model are transfer coefficients, viscosity, rate constants telling how one constituent interacts with another, thermal conductivities, etc., and, as such, usually have a distinct physical interpretation. The model could approach the problem from a "top-down" viewpoint where broad phenomenological interactions among state variables are represented by parameters with or without a more basic physical interpretation.

The methods we develop in this book do not help one create a model, except to provide information to the model about the estimated value of its parameters and a systematic method for completing the model through that estimation and to validate (or invalidate) the model through using it to predict further experiments and observations. The methods here do not provide guidance or instruction on how to improve models that are shown to be invalid. That is still the art of the scientist. We find here a principled path to testing the consistency of a given model with given data, and that path is broadly general, useful in many application areas, and has the potential to be a tool of great use in the understanding of complex systems.

The model typically has a lot of state variables, not all of which we are able to observe, and it has some physical parameters we may not know. We want to transfer information from the measurements to that model to allow us to estimate the fixed parameters and to estimate the **unobserved** state variables in the observation interval $[0, T]$. If we can do a good job of this, then using the estimated parameters and using the estimated full model state (observed and unobserved state variables) at T, we can use the model to predict for $t > T$. To test this prediction, we require further data for $t > T$.

The prediction can go wrong for a variety of reasons: (1) The data is very noisy and the interference by the noise masks the behavior of the physical system one is trying to describe. (2) The model is wrong because it operates in a noisy environment and that masks the dynamics we want to uncover or it simply lacks dynamical elements in the differential equation or discrete time map. (3) The method used to extract information from the data and pass it along with the model is flawed; one has an insufficient data assimilation procedure. (4) The model is incorrect.

Once one has noisy data and model errors in the mix, the overall task becomes a statistical problem. There is a probability distribution for the state of the model, conditioned on the observations. One must start at $t = 0$ with an initial condition for the distribution of states. We may have some knowledge of this or may not. The idea of solving initial value problems for probabilistic quantities is totally ingrained in our physics education—think of the Schrödinger equation which produces complex probability amplitudes or the Fokker–Planck equation for real probability distributions—so this is not news. Then we must propagate this distribution function using the dynamical rules of our model to the time $0 \leq t_{\text{meas}-1} \leq T$ where we make our first measurement. At $t_{\text{meas}-1}$ we require a rule letting us know how information in the measurement at that time influences the state distribution function at that time. Then we need to propagate that distribution to the next measurement time $t_{\text{meas}-2} > t_{\text{meas}-1}$ using the dynamics and apply our information transfer rule and so forth until we reach the end of our observation window in time at T.

This set of tasks says "path integral" to contemporary physicists who have seen this question raised in quantum mechanical and statistical physics contexts over and over again. Indeed, precisely this question cast in quantum mechanical language has been a core topic in quantum theory since the 1980s (Caves 1986). The starting point for the classical version of the data assimilation problem is a path integral giving the integral representation of the solution to the data assimilation problem. The path integral is exact, as an exact statement of the information transfer at each measurement comes from an identity on conditional probabilities. The exact path integral is more or less useless for application to any specific question, so approximations and, always, numerical evaluations are required.

This book is about all of that with extensive examples from nonlinear circuits, fluid dynamics, toy geophysical models, and neurobiological simulations and experiments. The general principles are discussed both after some examples and before other examples. Indeed, we start out with a kind of standard least-squares "fitting" of a model to data and only later reveal it to be a saddle point (Laplace 1774; Debye 1909) approximation to the path integral. Since we have the integral representation to the full statistical data assimilation problem, we can formulate methods for evaluating the corrections to the saddle-point approximation, and we can formulate methods for just directly evaluating the integrals involved. The latter leads us to Monte Carlo methods and substantial optimism that contemporary parallel processing techniques will permit large problems to be solved.

The path integral is an integral representation of the linear partial differential equation for the conditional probability distribution. As such it gives the opportunity for a global view of the solution to the underlying stochastic physical problem and permits going beyond the local view of other data assimilation methodologies.

That might even permit one to investigate the stimulating questions about nervous systems as well as many other problems in complex systems.

The path integral also points our attention to the fundamental quantity in data assimilation, namely the paths of a stochastic system through its state space as they are influenced by observations. Formulating the questions one wants to answer as a path integral and focusing on performing the integrals that answer those questions bypass the efforts in other methods of data assimilation to estimate auxiliary quantities from which one might be able to extract the answers of interest. Other books give excellent instruction on those other methods, for example (Evensen 2009).

The examples presented here emphasize the importance of using the dynamical model in a series of what we call "twin experiments" to explore the requirements on measurements to perform the desired transfer of information. Twin experiments, a phrase borrowed from the geophysical literature, generate data with a given model, then uses that model with "unknown" parameters to test data assimilation methods. These twin experiments also provide a testing ground for the methods one selects for performing the path integrals at hand. Further, they are very useful for estimating the number of required measurements and identifying which measurements one needs to carry out a data assimilation task. In this, they are very helpful in designing experiments and observations.

The other theme in our formulation of statistical data assimilation is that of potential instability in the communications channel between the data as a transmitter and the model as a receiver. This is a feature of the nonlinearity of the models we use to formulate the dynamics of the underlying physical processes. It is not a feature of linear models, but it appears and has been recognized in excellent monographs (Evensen 2009). An emphasis on its importance connects the need to regularize the instabilities with the goal of using the model for predicting behavior of the physical systems of interest.

The emphasis is also on the use of prediction as the testing ground for the quality and consistency of the data, the model, and the assimilation methods. A good "fit" to a data set by a plausible model of the underlying dynamics can be rather misleading. Stopping there and evaluating the outcome of the model representation of unobserved states is rather subjective and avoids the scrutiny a model must face in providing accurate predictions.

Returning to the nervous system questions that prompted the discussions given here, one can make any model one wishes to give a quantitative dynamical framework, and the methods in this book can be used to evaluate that model. The viewpoint on constructing quantitative, predictive models one finds here is "bottoms-up" (Rabinovich et al. 2006) starting with biophysical models and detailed experiments at the neuron level, then building the nervous system networks from that as a basis, along with further experiments, of course.

One may choose how to build a model, in this scientific arena or others, and that is an essential ingredient in using statistical data assimilation tools. The tools are indifferent to how the model is made or interpreted. It provides a path for testing those models, completed by the estimation of any unknown parameters within them.

Returning to the problem which stimulated this inquiry, the path that was initiated by the challenge of understanding functional nervous systems grew well beyond that. I made an informal survey of fields in which the data assimilation methods developed in this book are of importance and ended my survey at ten or so distinct areas ranging from toxicology and genetics to numerical weather prediction and predicting how coastal flows drive river-borne pollution dispersion. The material in this monograph might well have implications for a such a diverse set of applications, but we address only two. One is the motivating question of how one may understand from a biophysical viewpoint how functional nervous systems are constructed and operate. The other, touched upon in the chapter on twin experiments, encompasses meteorological models of the interacting atmosphere and ocean, though what is presented is only an initial study of the core ingredient for those, namely, shallow water flows. Our hope, of course, is that readers will find the material here a stepping off point for further numerical development of methods and interesting applications across many disciplines.

While perhaps a sidebar to the flow of this book, I note that I am not a fan of "punitive pedagogy" seen as the mode of presenting each idea once, and only once, regardless of its importance. The reader will find, therefore, repetition without regret presented in the hope that the pedagogy is more elevated in value.

Acknowledgments

The questions that led to the path integral formulation of data assimilation were discussed extensively with Dan Margoliash, Misha Rabinovich, Ulrich Parlitz, and Al Selverston in the context of understanding the biophysics of neurons. The confusion I noted earlier was not theirs, but mine. The ability to make the smallest steps toward understanding the issues was done in collaboration with my UCSD colleague Philip Gill. Kody J. H. Law at the University of Warwick was kind enough to read the book in draft form and make numerous valuable suggestions for its improvement. To all of them, many thanks.

While puzzling about all this I had the opportunity to take a sabbatical from UC, San Diego, in 2009, and that provided environments where I could investigate issues without the usual distractions of home. In the laboratory of Richard Hahnloser at the University of Zurich/ETH I labored extensively with what I came to see as the saddle-point approximation to the overall answer, and at the Bernstein Center for Computational Neuroscience at the Ludwig Maximilians University in Munich hosted by Andreas Herz, I was able to formulate the first steps in understanding the questions as path integral based. The sabbatical concluded with a stay in the Margoliash lab at the University of Chicago where we began to formulate experiments that could be used to employ the path integral methods for investigating the biophysical properties of nerve cells. On returning to UCSD, I had the pleasure of many conversations with colleagues at Scripps Institution of Oceanography on these matters, especially with Bruce Cornuelle.

The research was supported by the Office of Naval Research, the Collaborative Research in Computational Neuroscience (CRCNS) jointly organized by the National Science Foundation and the National Institutes of Health, and finally by grants from the US Department of Energy, Office of Science, Applied Mathematics Program. The generous sabbatical program of the University of California allowed me to initiate and to complete the research leading to this manuscript.

This work is dedicated to my family who gave me the special context for intensely focusing on this research. I acknowledge them again here, and they can now find out what I think I was doing those many hours and months.

http://www.springer.com/978-1-4614-7217-9

Predicting the Future
Completing Models of Observed Complex Systems
Abarbanel, H.
2013, XVI, 238 p. 97 illus., 91 illus. in color., Hardcover
ISBN: 978-1-4614-7217-9