



Foundations of Predictive Computational Science

CSE 397 / EM 397: Special Topics in
Computational Science

Unique Number 66248

Fall Semester, 2016

TTH 9:30–11:00 AM

Room CMA 5.190 Communications Building

J. Tinsley Oden

A special course on contemporary topics in computational sciences will be offered to survey several developing areas at the foundation of predictive computational science and at the forefront of research in the related areas. Topics include:

- Logical Probability Theory: The Cox-Jaynes theory of probability and its relation to Kolmogorov.
- Bayesian methods for model selection, calibration, validation, and uncertainty quantification.
- Information Theoretical Methods: Akaike Information Criteria and other methods for model selection.
- Model Inadequacy: OPAL (the Occam Plausibility Algorithm), Gaussian Processes, Bayesian networks, Deep Learning.
- Continuum Mixture theory and Phase-Field Methods, the general approaches for modeling interacting heterogeneous media.
- Introduction to tumor growth modeling: multiscale models of living tissue, cellular behavior, sub-cell models of signalling pathways and genomics.
- Applications in complex systems: molecular dynamics, multi-species biological systems and other systems.

Contents

1 Predictive Computational Science: Computer Predictions in the Presence of Uncertainties	9
1.1 Introduction	9
1.2 Mathematical Models	13
1.3 Sources of Uncertainty in Predictions	15
2 Probability Theory - A Review of the Classical Frequentist Theory of Kolmogorov	19
2.1 Introductory Comments	19
2.2 Measure and σ -algebras	20
2.3 The Bochner Integral	22
2.4 The Radon-Nikodym Theorem	24
2.5 Summary of Basic Properties of a Probability in the Kolmogorov-Measure Theory Setting	27
2.5.1 Conditional Probability	28
2.6 Probability Properties in Terms of Probability Densities	30
2.7 Examples of Probability Density Functions	33
3 The Logical Theory of Probability of R.T. Cox and E.T. Jaynes	37
3.1 Introduction	37
3.2 Plausible Reasoning	38

3.3	Logic and Philosophical Theories of Probability	40
3.4	Van Horn's Guide to Cox's Theorem	43
3.5	The Terenin-Draper Account of Cox's Theorem and Cox-Jaynes Probability	50
3.6	Summary Comments	52
4	Statistical Calibration and Validation of Parametric Models.	55
4.1	Introduction	55
4.2	Data Uncertainty: Experimental Noise	57
4.3	Model Inadequacy	58
4.4	Frequentist and Fisherian Approaches	59
4.5	Principles of Bayesian Model Calibration and Validation	63
4.6	Confidence Intervals versus Credible Intervals	67
4.7	Examples: Optimization, Least-Square Misfit, and Bayesian Learning	67
4.7.1	Least-Square Misfit	67
4.7.2	Bayesian Learning	68
4.8	Gaussian Processes: Tools to Build Models of Noise and Model Inadequacies	69
5	Information Theory and the Principle of Maximum Entropy	77
5.1	Introduction	77
5.2	An Elementary Look at Information	77
5.3	A First Look at Information Entropy	79
5.4	The Maximum Entropy Principle	80
5.5	Mutual Information	82
5.6	Shannon's Theorem	84
5.7	Example: Entropy of the Uniform Distribution	85
5.8	Maximum Entropy Priors: The General Solution for Maximum Entropy with Linear Constraints	87

5.9	The Tan Bui Thanh Lemma: Entropy as a path to Bayes	91
6	Monte Carlo Methods: MCMC and Sensitivity Analysis	93
6.1	Introduction	93
6.2	Markov Chains	95
6.3	Markov Chain Monte Carlo	96
6.4	Parameter Sensitivity Analysis	97
6.4.1	Morris's Method of Elementary Effects	98
6.4.2	Variance-Based Sensitivities	100
6.4.3	Scatter Plots	101
6.4.4	Design of Validation Experiments	102
6.4.5	An Algorithm for Sensitivity Indices	103
6.4.6	An Example: Coarse-Grained Models of Atomistic Systems	104
7	Model Selection and OPAL, the Occam Plausibility ALgorithm	111
7.1	Introduction	111
7.2	Bayesian Model Plausibilities	112
7.2.1	Frequentist Approaches: The Akaike Information Criterion	113
7.3	The OPAL Algorithm: Adaptive Model Selection and Model Validation	115
8	Introduction to Cell and Cancer Biology	121
8.1	Introduction	121
8.2	DNA: Deoxyribonucleic Acid	121
8.3	The Cell	130
8.4	RNA: Ribose Nucleic Acid	130
8.5	Transcription	132
8.6	Protein Synthesis Steps	136
8.7	Signal Transduction	138

8.8	The Cell Cycle	148
8.9	Cancer	152
9	Phenomenological Models of Tumor Growth	157
9.1	Introduction	157
9.2	A Brief Review of Continuum Mechanics	158
9.3	A Continuum Theory of Mixtures	165
9.4	Mixture Theory Models of Avascular Tumor Growth	175
9.5	An Application of OPAL to the Selection and Validation of Models of Tumor Growth	184
10	Chemical Kinetics Models of Signal Transduction	193
10.1	Introduction	193
10.2	An Overview of Chemical Kinetics	195
10.3	Complex Reactions	199
10.4	An Application to a Cell-Signaling Pathway	202
11	Phase-Field Models	209
11.1	Introduction	209
11.2	Structure of Phase-field Models	209
11.3	The Allen-Cahn Equation	211
11.4	The Cahn-Hilliard Equation	212
11.5	Phase-Field Models of Tumor Growth	213

Chapter 1

Predictive Computational Science: Computer Predictions in the Presence of Uncertainties

1.1 Introduction

Predictive computational science is the scientific discipline concerned with assessing the predictability of mathematical and computational models of events that occur in our physical universe in the presence of uncertainties in all of the factors that determine the reliability and scientific significance of the prediction. Prediction is a fundamental mission of basic science. It is also the essence of engineering, which is concerned with the application of science in the prediction at the performance of engineered systems under design conditions: the overarching goal of engineering for centuries. It is also at the core of medical science, which seeks to predict the outcome of therapies and medical procedures.

What is a “science-based prediction”? Indeed, what is science? The word **science** is derived from the Latin, **scientia**, meaning knowledge. Thus, the Oxford dictionary defines science as the activity concerned with the systematic acquisition of knowledge: the enterprise that builds and organizes knowledge in forming testable explanations and predictions about events that take place or can take place, in the universe. Engineering and medicine are the activities concerned with the application of scientific and practical knowledge for the benefit or needs of mankind.

Exactly how knowledge is acquired is a subject debated for millennia. To Socrates and Plato (ca. 370 BC), humankind is endowed with pre-existing knowledge, knowledge that is already present in our minds and is accessible to our intuitive powers. Aristotle had a quite different view; in his Organon describing his six works on logics, he asserts that the acquisition of knowledge comes through sensation and perception. **Sensation** referring to the use of one's five senses to experience reality through empirical events and thus interpretation by the consciousness of the subject acquiring knowledge. To this he adds **perception**, the recognition that the event has a sensation. The term **observation** refers to the acquisition of information through one's senses, or through aids to one's senses made possible by using instruments, all to either observe natural events or events orchestrated in experimental setups. Aristotle was the father of logic – the science of reasoning, and he laid down the rules of deductive logic as the foundations of algebra and much of mathematics.

Deductive reasoning, or deductive logic, began with the two-sided logic advanced by Aristotle. It is embodied in a process of reasoning beginning with one or more general statements (axioms or principles) and following the rules of logic to reach logical conclusions. It thus provides the formalisms for mathematics and mathematical proof. It (deduction) is sometimes called "top-down" logic. It is infallible and not influenced by empirical observations. On the other hand, scientific knowledge is based on **inductive logic**, the process of reasoning by generalizing or extrapolating from initial information or hypotheses. Inductive logic is "bottom up" logic, an open system including domains of epistemic uncertainty, thus allowing a conclusion to be false.

Thus, from antiquity there came the notions that knowledge can arise from observations of natural events or events manufactured in experiments, and through logical explanations of the cause of events and hypotheses concerning the effects of certain actions on events taking place in the future. But not all philosophers accepted these ideas.

Philosopher David Hume (1711-1776) launched the scientific philosophy of **skepticism** whereby he asserted that knowledge can only be obtained through the study of empirical observations and data. He thus rejected inductive logic as a means to gain knowledge: if inductive logic can allow a possibly false conclusion, how can it be a source of knowledge? Fortunately, the great theoretical scientists of the eighteenth, nineteenth, and twentieth centuries (Maxwell, Einstein, Dirac, Boltzmann, ...) ignored this point of view and went on to vastly increase our knowledge of the universe through miraculous advances in theoretical science. The paradox between experimental evidence and scientific theory was more or less put to rest by the philosopher Karl Popper (1902-

1994) who introduced the **principle of falsification**. According to this principle, any scientific claim, any hypothesis put forth as a possible scientific theory, cannot qualify as a scientific theory unless it is possible that it be falsified, that is, it can possibly be disproved when new conflicting information is acquired. Quite opposite to Hume, famous astronomer Arthur Eddington (1882-1944) announced that "one should never accept any experimental data that has not been confirmed by theory" – an affront to Hume's philosophy and an acknowledgment of the great advances in theoretical physics made by his associate Paul Dirac, who discovered anti-matter through purely theoretical arguments four years before it was observed experimentally. To the great probabilist and physicist E.T. Jaynes (1922-1998), all knowledge is inductive; inductive logic must precede the performance of meaningful experiments or the interpretation of observations of natural events.

Today it is generally recognized that knowledge is acquired through observation and theory, the two classical pillars of science. These are the foundations of the scientific method, the body of techniques for investigating phenomena and acquiring new knowledge based on **observation** (empirical or measurable evidence) and **theory** (the foundation, testing, and modification of hypotheses based on inductive logic laid down to explain observed or perceived event).

Throughout this work, theory is presumed to be characterized by mathematical models or models derived from surrogate models that arise from simplifying and subjective assumptions. When possible, such models are generally based on long-standing scientific principles complemented by empirical laws, all expressed in mathematical constructions.

Modern scientific philosophy often recognizes a third pillar of science: **computational science**, a new means of acquiring knowledge through computer simulations where experiments are impossible or inadequate for the purpose at hand. There are countless examples in which computer simulations based on models derived from accepted physical laws and empirical relationships, provide a path to knowledge.

These three cornerstones of predictive science; the collection and interpretation of observational and experimental data, the creation and use of theory and models, and the use of computer simulations, are by no means independent; they are intrinsically intertwined and inseparably connected. Inductive logic precludes the design of experiments; theory and models underlie the design and construction of instruments and measuring devices, and data and mathematical abstractions of theoretical models form the foundation of computer simulations. All are fundamental to science-based prediction.

Today, prediction in modern science and engineering stands at a threshold brought on by enormous advances in computational methods and machines over the last two or three decades. On one side of the threshold, we have growing expectations of quantitative predictions of physical realities produced by computer simulations of events at the center of a long list of technological and scientific grand challenges: climate change, effects of natural hazards, the behavior of biological and biomedical systems, design of materials, aging civil infrastructures, and many more, all relying heavily on computer predictions that form the basis for critical, often life-and-death decisions. On the other side of this threshold is the realization that every step of the prediction process faces sometimes overpowering uncertainties: uncertainties in observational data, in model selection, in model parameters, and in the targets of prediction; the so-called *quantities of interest*.

The importance, indeed the urgency, to extend and enrich the processes of prediction to take such uncertainties into account has gradually taken hold in scientific and mathematical research over the last two decades, and particularly over the last decade, and a large and growing literature is devoted to the subject (as representative, see [2, 3, 11–13, 78, 79, 81, 92, 109]). Predictive computational science embraces methods for recognizing and handling randomness in observational data, model selection, model calibration, model validation, and verification of computational models in the presence of uncertainties, and includes the developments of methods for solving the forward prediction problem for quantities of interest, inverse problems in parameter identification, and quantifying the uncertainty in the predictions. The well-worn acronym "VV & UQ" (Validation, Verification, and Uncertainty Quantification) is often thrown out to label the mission of computational predictive science.

This exposition is intended to present a broad view of the foundations of predictive computational science at a level accessible to researchers and practitioners interested in model reliability and how it is determined when uncertainties in all of the components of prediction exist. The presentation is necessarily based on theories of probability, statistics, information theory, with meaningful applications drawn from cell biology, continuum physics, and molecular dynamics, methods of mathematical statistics and probability and includes discussions of powerful methods for model selection and for quantifying uncertainties in predictions of quantities of interest.

1.2 Mathematical Models

At the heart of predictive computational science is the idea of a **mathematical model** of a class of physical events of interest, a mathematical abstraction of reality generally based on a combination of scientific principles and empirical laws. Mathematical models stand as surrogates or as representatives of theories – the inductive hypotheses put forth to approximate perceived realities and to provide the essential component in prediction of future events.

It is convenient (indeed, essential for our purposes) to consider mathematical models as representing an abstract mathematical problem, such as the following: Given a vector $\boldsymbol{\theta}$ and a domain S in which the problem is posed, find a function $u(\boldsymbol{\theta}, S) \in V$ such that,

$$\mathcal{A}(\boldsymbol{\theta}, S; u(\boldsymbol{\theta}, S)) = 0 \quad (1.1)$$

where \mathcal{A} is a collection of operators (algebraic, differential, partial-differential, integral, ...), including possibly constraints, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^k$ is a vector of model parameters taken from a parameter space Θ , S is the scenario in which the problem is posed, and $u(\boldsymbol{\theta}, S)$ is the solution of Equation 1.1 which belongs to a function space V for specific $\boldsymbol{\theta}$ and S .

The notion of a scenario is a especial idea that refers to the features of the mathematical problem which can be exactly specified, independent of parameters, such as the domain of the solution and certain boundary-and initial conditions, the idea being that the same model may be used in several different scenarios. For example, an elasticity model can be defined for several scenarios, each consisting of collections of a material bodies of different shape and boundary conditions. We use the term *scenario* interchangeably to refer to both the domains in which the mathematical model (1.1) is solved and the actual physical environments in which experimental data are collected, or in which the target quantities to be predicted reside. Later we will argue that it makes sense to consider a hierarchy of scenarios, beginning with component or unit calibration tests in scenarios S_c to initially adjust model parameters, then moving to more complex subsystem validation scenarios S_v to test the accuracy with which model can predict observables (to see if the model is “valid”), and then the prediction scenario S_p , involving the full physical system and model, in which the forward problem (1.1) is solved to predict the *Quantities of Interest* (QoIs). A revealing example of various scenarios emerges in models of biomedical systems in which *in vitro* data are selected in lab experiments (e.g. cell proliferation in a petri dish supplied with concentration of nutrients, or *in vitro* experiments with laboratory animals) and the models are there expected to be validated using *in vivo* data collected through biopsies or emerging imaging techniques. These

are the calibration scenarios S_c (the *in vitro* experiments) and the validations scenarios S_v (the *in vivo* experiments). The model may then be called upon to make an out-of-data prediction in the prediction scenario S_p . We discuss these ideas in more detail in later lectures.

The model parameters $\boldsymbol{\theta}$ are generally random vectors, and, thus, the solutions $u(\boldsymbol{\theta}, S)$ are random or stochastic functions. We elaborate on this fact in the next section. Equation (1.1), therefore, describes an infinite class of models parametrized by $\boldsymbol{\theta}$. Equation (1.1) is called the *forward problem* as it describes a projection of data "forward" into a prediction $u(\boldsymbol{\theta}, S)$, which is the solution of the forward problem.

Two fundamental issues come to light at this point:

1. There must exist solutions $u(\boldsymbol{\theta}, S)$, to the forward problem for every $\boldsymbol{\theta} \in \Theta$. This fact leads to conditions on the space Θ of parameters. The solution will be an element of an appropriate space V of generally random functions.
2. The goal of constructing and solving the model (1.1) is to compute specific quantities of interest (QoIs), generally described as functionals Q on the space V containing the solution. This calculation is to be made in the full prediction scenario, $S = S_p$, the full system in which the model extrapolates all relevant information to predict an outcome beyond those events available through the analysis of observational data:

$$Q : V \rightarrow \mathbb{R} : Q(u(\boldsymbol{\theta}, S_p)) = \bar{Q}(\boldsymbol{\theta}), \quad (1.2)$$

\bar{Q} being also a random function.

As we will emphasize later, the determination of the validity of the model (1.1) as a means of predicting \bar{Q} of (1.2) must, by definition, involve determining the accuracy with which model predictions agree with experimental data- the observations. These data must be collected in some accessible experimental setting which corresponds to a model *validation scenario* S_v , in which experiments are designed that can, hopefully, represent how well the model can predict the QoI. Different scenarios S_c appear when we use experimental data to calibrate a model, that is to adjust its parameters so that predicted responses for experimental measurements.

1.3 Sources of Uncertainty in Predictions

So what is the anatomy of scientific prediction within the context laid down thus far, and how do we cope with the multitude of uncertainties in each step on the predictions process? We wish to predict a specific reality in a possibly very complex scenario S_p using a computational model which is based on a mathematical model of a real situation from which the target quantity of interest (the QoI) can be specified. We must therefore select a model that we hope can deliver a sufficiently, accurate estimate of the QoI and for which we have some confidence in its predictability. The predictability of our model rests on knowledge gained through observation – and our access to observational data – and our inductive hypotheses embodied in our selections of a model. The model, in general will be characterized by a set of parameters which themselves are generally unknown, and which at best involve uncertainties. Then we must discretize the mathematical model to produce a computational model, a corruption necessary to render the model into a form that can be processed by digital computers.

Collecting all these sources of uncertainty, and following, in part, the classifications plan of Terenin and Draper (2015), we list the following sources of uncertainty in predictive computational science and in making prediction of realities:

1. **\mathcal{P} -uncertainty.** This is so-called *systemic uncertainty* – uncertainty about the appropriate systems of reasoning or mathematical framework in which to study and quantify uncertainty in a logically-consistent way (e.g. do we frame our predictive process in the language of probability theory (if so, which one?)), possibility theory, fuzzy sets, Dempster-Shafer theory, interval analysis ... which logical system produces a self-consistent way to deal with uncertainty in prediction?
2. **\mathcal{Y} -uncertainty.** This embodies uncertainty in observational data. All data recorded from physical experiments can be polluted by experimental noise. How do we cope with uncertainties in data?
3. **\mathcal{M} -uncertainty.** One of the most elusive and important sources of uncertainty is the selection of the mathematical model itself: \mathcal{M} -uncertainty. We agree with the famous quote of E.P.Box [1919-2013]: "Essentially all models are wrong. Some are useful". One of the most important missions of model selection and validation processes is to identify such useful models.
4. **$\boldsymbol{\theta}$ -uncertainty.** Model parameter $\boldsymbol{\theta}$ uncertainty, the uncertainty in parameter values or in parametric families of distributions – a key source of uncertainty in predictive models.

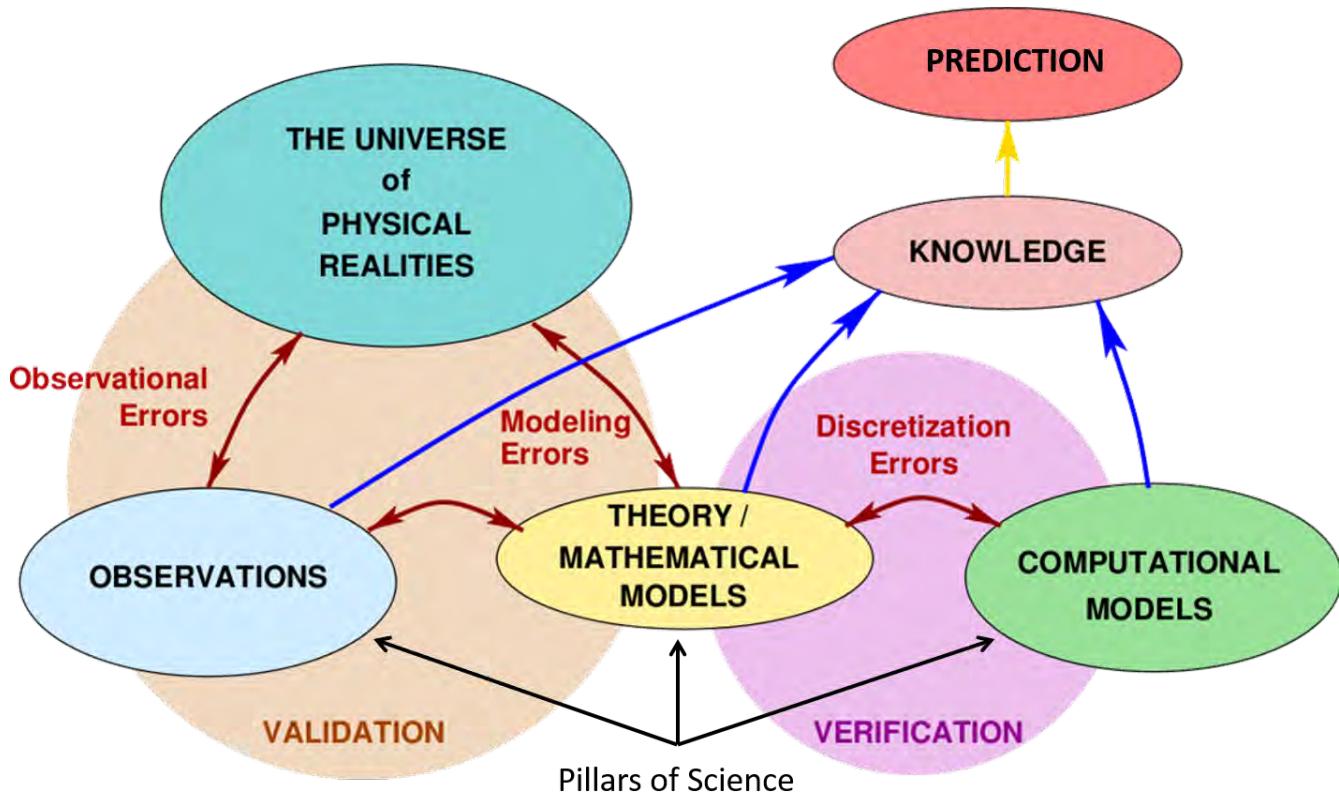


Figure 1.1: The Imperfect Paths to Knowledge and Prediction. A schematic of the imperfect paths to knowledge that involve uncertainties in all phases of the process for gaining knowledge and making predictions. [From Oden, Moser, and Ghattas, 2010 [80]]

5. ***h*-uncertainty.** This is uncertainty brought about by the discretization process that produces the computational model – “*h*” representing an appropriate discretization parameter. This uncertainty can sometimes be estimated *a posteriori* through error estimation or bounded through *a priori* estimates or at least studied in some quantitative sense through various verification processes.

The various paths to knowledge, made imperfect by the uncertainties discussed , are represented symbolically in Figure 1.1.

In the remainder of these notes, we will systematically develop the machinery to study each of these sources of uncertainty (except possibly *h*-uncertainty, which is covered in depth elsewhere and which is a bit out of the scope of this exposition). The next two chapters focus on \mathcal{P} -uncertainty – which we resolve by adopting forms of probability theory. Parameter (θ) and model (\mathcal{M}) uncertainties are taken up in a chapter on statistical calibration, validation, and on information theory. We describe algorithms for systematically addressing \mathcal{Y} and θ uncertainties. Then we develop the

foundations of model prediction, parameter sensitivity analysis, and discuss an introduction to information theory and entropy methods. Then we consider several application areas where these tools are applied to various problems in engineering, science, and medicine.

Chapter 2

Probability Theory - A Review of the Classical Frequentist Theory of Kolmogorov

2.1 Introductory Comments

Some scholars cite the beginning of a systematic study of the idea of probability with correspondence between de Fermat and Pascal in 1654. (e.g. Davis [1]). The theory was expanded and formalized in the writings of Bayes in 1763 and Laplace in 1820 and onward, and advanced in the works of Cournot in 1843 and Venn in 1820. However, to most of today's circles in mathematics and statistics, a mathematically rigorous theory of probability did not appear until the work of Kolmogorov, beginning in 1933, which laid the mathematical foundations for the so-called *frequentist* view of probability. This view is embedded in the language of measure theory, measurable functions, and a view of the probability of an event as directly related to the frequency the event will occur over infinitely many samples.

Other quite different notions and theorems, not as well known or established as Kolmogorov, began to emerge in the 1930's-early 1940's in the work of deFinetti [1931-1933], on Bayesian inference based on *belief* and *betting odds*, the so-called *subjective probability theorem*, and then came the powerful *logical probability* of R.T. Cox in 1946, also based on Bayesian inference. Cox's theory was embraced by E.T. Jaynes and expanded and enriched in his treatise, *Probability Theory: The Logic of Science* published in 2003. We refer to the latter as the *Cox-Jaynes theory of probability*.

In the sections that follow in these notes we will review the Kolmogorov theory. The Cox-Jaynes theory is taken up in Chapter 3.

2.2 Measure and σ -algebras

Let A be an arbitrary non-empty set. We wish to construct a mathematical abstraction of the idea of size or volume of subsets of A , which we will call a *measure* on A : a real number assigned to subsets of A . While it is natural to attempt to construct such a measure for every subset of A , one can find counterexamples in which this is impossible: [the classical Vitali theorem confirms, using the Axiom of Choice, that subsets of the real line are not (Lebesgue) measurable. Another set-theoretic problem Kolmogorov encountered in using as a basis for measure spaces is that the countable, or uncountable, infinite power set 2^A is inappropriate for the assignment of measured probabilities is embodied in the *Von-Neumann Paradox* in which one can partition in unit square in \mathbb{R}^2 into a finite member of subsets, subject these subsets to an affine area-measuring map, and end up with two unit squares of area 2 (c.f., Terenin and Draper 2015 [110])]. The introduction of a σ -algebra, closed under countable union, intersection, and complementation set operations, avoids this paradox. Thus, to define a meaningful notion of measure, we need to identify a smaller collection of special subsets of A . The class \mathcal{A} of subsets of A that qualifies, has two basic properties: 1) the class is closed under complementation and 2) countable unions of such sets belong to the class:

- 1) $\emptyset \in \mathcal{A}$
 - 2) $B \in \mathcal{A} \Rightarrow B' = A/B = \text{complement of } B \in \mathcal{A}$
 - 3) $A_1, A_2, \dots, A_n, \dots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$
- (2.1)

A class of subsets of A with properties (2.1) is called a *sigma-algebra* or σ -*algebra* on A and the member sets making up a σ -algebra are called *measurable sets*. The pair (A, \mathcal{A}) where A is a set and \mathcal{A} is a σ -algebra on A is called a *measurable space*.

Example. Let $A = \{\alpha, \beta, \gamma, \delta\}$. The class of subsets, $\mathcal{A} = \{\emptyset, \{\alpha, \beta\}, \{\gamma, \delta\}, \{\alpha, \beta, \gamma, \delta\}\}$ is a σ -algebra on A , as complements of every set and unions belong to \mathcal{A} .

A *measurable function* is defined in the following way: if (A, \mathcal{A}) and (B, \mathcal{B}) are measurable spaces, a function $f : A \rightarrow B$ is measurable iff $\forall B \in \mathcal{B}, f^{-1}(B) \in \mathcal{A}$; i.e. the inverse image of every measurable set is a measurable set.

Generated σ -algebra. Let \mathcal{C} be any arbitrary class of subsets of a set A . There exists a unique smallest σ -algebra that contains every set in \mathcal{C} . This σ -algebra is called the *σ -algebra generated by \mathcal{C}* .

Example. Let $A = \{\alpha, \beta, \gamma, \delta\}$, $C = \{\{\alpha, \beta\}\}$. We construct sets of complements and unions of sets in \mathcal{C} to produce a σ -algebra. For instance,

$$(\mathcal{A}|\mathcal{C})_1 = \{\{\alpha, \beta, \gamma, \delta\}, \{\gamma, \delta\}, \emptyset, \{\alpha, \beta\}\} ,$$

$$(\mathcal{A}|\mathcal{C})_2 = \{\{\alpha, \beta, \gamma, \delta\}, \{\alpha\}, \{\beta, \gamma, \delta\}, \{\beta\}, \{\alpha, \beta, \delta\}, \{\alpha, \beta\}, \emptyset\} ,$$

etc. The smallest of these is $(\mathcal{A}|\mathcal{C})_1$, which is therefore the σ -algebra generated by \mathcal{C} .

Borel σ -algebras, A common class of measurable sets occur in Borel σ -algebras. If (X, \mathcal{X}) is a topological space (a set X and a class of subsets \mathcal{X} defining a topology on X), then the sets in \mathcal{X} are called open sets with respect to \mathcal{X} . (They could also be defined as closed sets). A set in (X, \mathcal{X}) that can be formed from open sets in \mathcal{X} through relative complementation, countable unions, and countable intersection is a *Borel set*. The collection of all Borel sets is a *Borel σ -algebra* on X . It is the smallest σ -algebra containing all open sets.

Measure. A function μ from a measurable space (A, \mathcal{A}) into $[-\infty, +\infty]$ is a *measure* iff

$$i) \mu(A) \geq 0 \quad \forall A \in \mathcal{A}$$

$$ii) \mu(\emptyset) = 0$$

$$iii) \mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i) \quad \text{for every countable collection of pairwise disjoint sets } \{A_i\}_{i \in I}$$

Measure Space. If (A, \mathcal{A}) is a measurable space and μ is a measure, the triple (A, \mathcal{A}, μ) is called a measure space.

Probability Space. A measure space $(A, \mathcal{A}, \mathbb{P})$ with measure $\mathbb{P}(A) = 1$ is called a *probability space* and \mathbb{P} is called a *probability measure*: $(\mathbb{P} : \mathcal{A} \rightarrow [0, 1])$.

σ -finite Measure. If (A, \mathcal{A}, μ) is a measurable space and $\mu(A)$ is *finite*, the space is said to be finite. It is σ -finite if it can be partitioned into a countable union of sets with finite measure.

2.3 The Bochner Integral

- Let (A, \mathcal{A}, μ) be a measure space.
- Let $\{A_i\}_{i \in I}$ be a set of disjoint sets in \mathcal{A}

- Let

$$\chi_{A_i}(a) = \begin{cases} 1 & \text{if } a \in A_i \\ 0 & \text{if } \text{otherwise} \end{cases} \quad (2.2)$$

$(\chi_{A_i} = \text{the characteristic function of } A_i)$.

- Let U be a Banach space with norm $\|.\|_U$
- Let $\{u_i\}_{i \in I}$ be distinct elements of U

The finite sum

$$S(a) = \sum_{i \in I} u_i \chi_{A_i}(a) \quad (2.3)$$

is called *simple function*.

- If $\mu(A_i)$ is finite when even $u_i \neq 0$, then $S(a)$ is said to be *integrable* and its integral in U is defined as follows.
- Let $S_n(a) = \sum_{i=1}^n u_i \chi_{A_i}(a)$, then we can compute the sum in U :

$$\sum_{i=1}^n u_i \mu(A_i) = \int_A S_n(a) d\mu \quad (2.4)$$

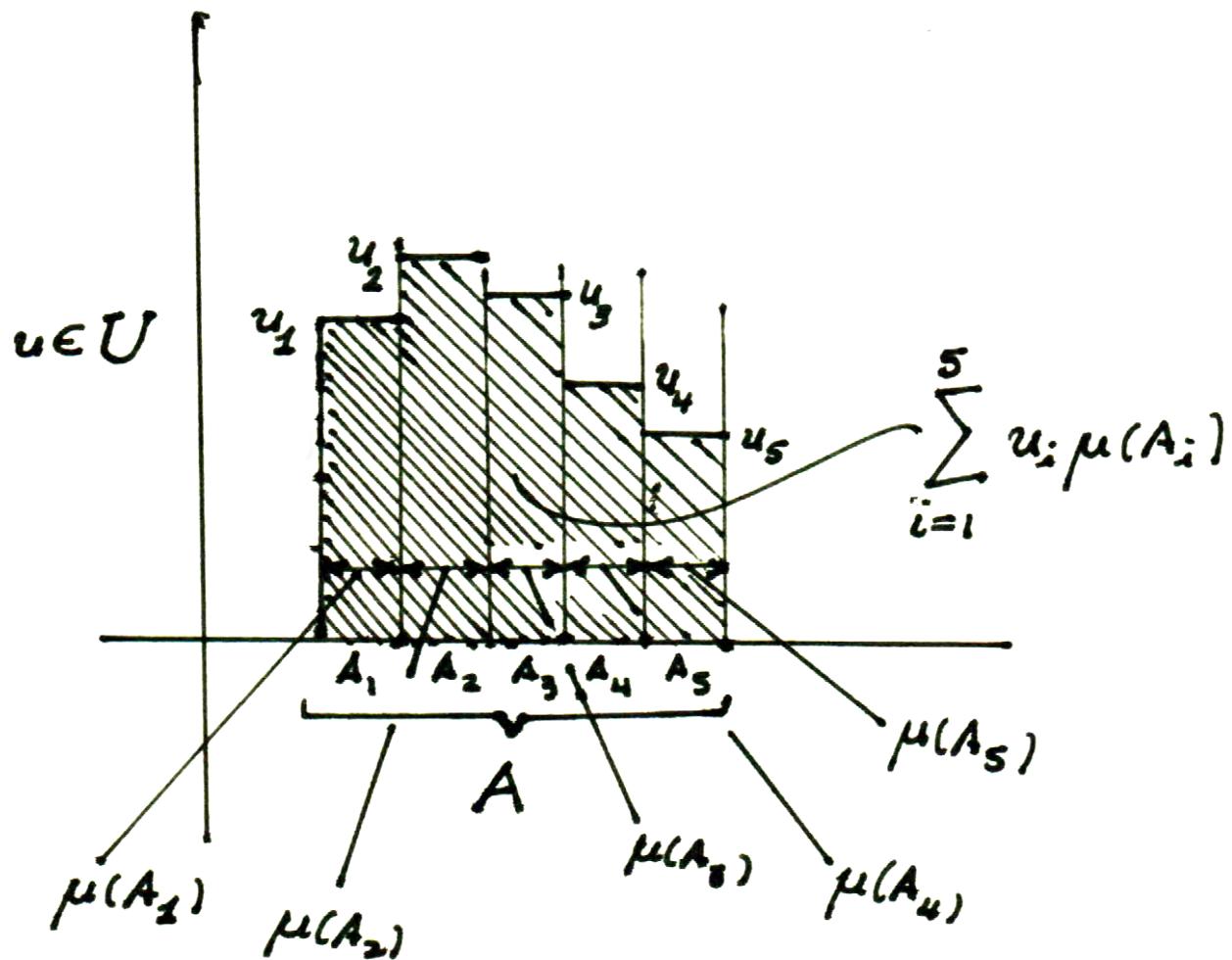


Figure 2.1: Graphical description of a single function representing an approximation of the Bochner integral of measurable function on a set A .

A measurable function $f : A \rightarrow U$ is *Bochner integrable* if there exists a sequence of integrable simple functions S_n such that,

$$\lim_{n \rightarrow \infty} \int_A ||f - S_n||_U d\mu = 0 \quad (2.5)$$

where the integral is the usual Lebesgue integral. Then the Bochner integral is,

$$\lim_{n \rightarrow \infty} \int_A S_n d\mu = \int_A f d\mu \quad (2.6)$$

2.4 The Radon-Nikodym Theorem

A function f is *absolutely continuous* with respect to a measure μ iff f is μ -integrable on a set A , the integral over subsets $B \subset A$, and for any $\epsilon > 0$ there exists a $\delta > 0$ such that $|\int_B f d\mu| < \epsilon$ whenever $\mu(B) < \delta$. In the case of real-valued function $f : [a, b] \rightarrow \mathbb{R}$, f is absolutely continuous with respect to Lebesgue measure if $\exists g \in L^1(a, b)$ such that

$$f(x) = f(a) + \int_a^x g(t) dt \quad (2.7)$$

Thus, $g = df/dx$ almost everywhere.

Theorem. *Let (A, \mathcal{A}) be a measurable space and ν be a σ -finite measure on (A, \mathcal{A}) that is absolutely continuous with respect to a σ -finite measure μ on (A, \mathcal{A}) . Then there is a measurable function $f : A \rightarrow [0, \infty)$ such that*

$$\nu(A) = \int_A f d\mu \quad (2.8)$$

In the general case covered by the theorem, f is written

$$f = d\nu/d\mu \quad (2.9)$$

and is called the *Radon-Nikodym derivative*.

Clearly,

$$\nu(A) = \int_A d\nu = \int_A \frac{d\nu}{d\mu} d\mu = \int_A f d\mu \quad (2.10)$$

A remarkable fact is that the Radon-Nikodym theorem fails to hold, in general, for Bochner integrals defined on all Banach spaces. Those Banach spaces for which it does hold are said to have the Radon-Nikodym property- which can be defined more precisely in the language of vector-valued measures. The property does hold for reflexive spaces, including therefore all Hilbert spaces, but does not hold for $L^1(\Omega)$, $L^\infty(\Omega)$, where Ω is an open bounded domain in \mathbb{R}^n .

The integral $\int_A f d\mu$ is an element of U . Thus, $\|\int_A f d\mu\|_U$ makes sense. Indeed, if f is integrable, $\|\int_A f d\mu\|_U \leq \int_A \|f\|_U d\mu$. For $B \in \mathcal{A}$, $\int_B f d\mu$ is a *vector-valued* (U -valued) measure on A , which is absolutely continuous with respect to μ .

Random Variables

Let \mathbb{P} be a probability measure on a measurable space (Ω, \mathcal{F}) . The sets in the σ -algebra \mathcal{F} are called events and $\mathbb{P}(A)$ ($A \in \mathcal{F}$) is the probability of the event A .

Random Variable. Let (U, \mathcal{U}) be another measurable space. A measurable function $X : \Omega \rightarrow U$ is a U -valued *random variable*, the term random coming from the fact that \mathbb{P} is a probability measure. Its values in U are regarded as not known with certainty but only to within some probability. Since X is measurable, the pre-image (or inverse image) of sets B in \mathcal{U} are sets in $\mathcal{F} : X^{-1}(B) \in \mathcal{F}$. If U is a topological space, the σ -algebra on U can be taken to be the Borel σ -algebra (that generated by the open sets in U). When U is the real line, X is called a real random variable or simply a random variable. Then $X : \Omega \rightarrow \mathbb{R}$ is a random variable if

$$\{\omega \in \Omega : X(\omega) \leq r\} \in \mathcal{F} \quad \forall r \in \mathbb{R}. \quad (2.11)$$

Here $X^{-1}((-\infty, r]) = \{\omega \in \Omega : X(\omega) \leq r\}$, arising from the fact that $\{(-\infty, r] : r \in \mathbb{R}\}$ generates a Borel σ -algebra on the real numbers.

Cumulative Distribution Function. The notion at a cumulative distribution function associated with a random variable X provides a means for characterizing X while hiding the particular probability space used to define X , recording only probabilities of various values of X . Consider $X : \Omega \rightarrow \mathbb{R}$. The event $\{\omega \in \Omega : X(\omega) \leq x\}$ is the probability that X is smaller than x ($x \in \mathbb{R}$), written $\mathbb{P}(X \leq x)$. Running these ranges over all $x \in \mathbb{R}$ gives the cumulative distribution function (CDF) for X :

$$F_X(x) = \mathbb{P}(X \leq x) \quad (2.12)$$

The CDF thus characterizes X without specifically displaying Ω . (The image measure space can be written $(\mathbb{R}, \mathcal{B}, dF_X)$).

Functions of a Random Variable. Let $X : \omega \rightarrow \mathbb{R}$ be a real random variable and let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a Borel measurable function. Since the composition of a measurable function is also measurable, $g(X) = Y$ is also a random variable on Ω . The CDF at Y is

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq y). \quad (2.13)$$

If g^{-1} exists and is increasing,

$$F_Y(y) = \mathbb{P}(X \leq g^{-1}(y)) = F_X(g^{-1}(y)). \quad (2.14)$$

The random variable X maps the measure \mathbb{P} to a measure df_X on \mathbb{R} , in the sense that

$$\mathbb{P}(a \leq X \leq b) = \int_a^b dF_X = F_X(b) - F_X(a). \quad (2.15)$$

In particular, the CDF has the properties:

- F_X is right continuous on \mathbb{R} ,
- $\lim_{x \rightarrow \infty} F_X(x) = 1$.

Probability Density. Let $X : \Omega \rightarrow U$ be a U -valued random variable and (U, \mathcal{U}, μ) a measure space. The variable X has a probability distribution measure $Xo\mathbb{P}$ on (U, \mathcal{U}) . The density of X with respect to μ is the Radon-Nikodym derivative

$$\rho = \frac{dXo\mathbb{P}}{d\mu}, \quad (2.16)$$

by which we mean ρ is a measurable function such that

$$\int_A \rho d\mu = \int_{X^{-1}(A)} d\mathbb{P} = \mathbb{P}(X \in A), \quad (2.17)$$

$\forall A \in \mathcal{U}$. When $(U, \mathcal{U}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (\mathcal{B} being the usual Borel σ -algebra), the CDF is

$$F_X(x) = \int_{-\infty}^x \pi(s) ds. \quad (2.18)$$

Or, if π is continuous at x ,

$$\pi(x) = \frac{dF_X(x)}{dx}. \quad (2.19)$$

The probability that values of X fall with the interval $[x, x + dx]$ can be interpreted as $\pi(x)$.

2.5 Summary of Basic Properties of a Probability in the Kolmogorov-Measure Theory Setting

To sum up some basic probability properties, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. The map $\mathbb{P} : \mathcal{F} \rightarrow \mathbb{R}^+$ (also written $\mathbb{P} : \Omega \rightarrow \mathbb{R}^+$) is a probability on \mathcal{F} iff,

1. $\mathbb{P}(A) \geq 0$ (the probability of the “event” A is non-negative), $\forall A \in \mathcal{F}$,
2. \forall pair of disjoint sets A and B in \mathcal{F} , $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$,
3. $\mathbb{P}(\Omega) = 1$.

From these elementary axioms we can derive the following properties of a probability:

Theorem. *If axiom 1 – 3 hold, then:*

- i) $\mathbb{P}(A') = 1 - \mathbb{P}(A)$ ($A' = A^{(c)}$ = complement of A),
- ii) $\mathbb{P}(\emptyset) = 0$.
- iii) $\mathbb{P}(A) \leq \mathbb{P}(B)$ if $A \subset B$.
- iv) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Proof. Property *i*) follows from the simple calculation,

$$\mathbb{P}(\Omega) = \mathbb{P}(A \cup A') = \mathbb{P}(A) + \mathbb{P}(A') = 1. \quad (2.20)$$

Property *ii*), similarly, comes from the fact that

$$\mathbb{P}(\Omega) = 1 = \mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) = 1 + \mathbb{P}(\emptyset). \quad (2.21)$$

Identity *(iii)* use the fact that

$$\mathbb{P}(B) = \mathbb{P}((B - A) \cup A) = \mathbb{P}(A) + \mathbb{P}(B - A) \geq \mathbb{P}(A). \quad (2.22)$$

To prove condition *(iv)*, we note that

$$\mathbb{P}(A) = \mathbb{P}(A - B) \cup (A \cap B)) = \mathbb{P}(A - B) + \mathbb{P}(A \cap B). \quad (2.23)$$

Thus,

$$\begin{aligned}
 \mathbb{P}(A \cup B) &= \mathbb{P}((A - B) \cup (B - A) \cup (A \cap B)) \\
 &= \mathbb{P}(A - B) + \mathbb{P}(B - A) - \mathbb{P}(A \cap B) \\
 &= \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B),
 \end{aligned} \tag{2.24}$$

and, therefore,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \tag{2.25}$$

□

2.5.1 Conditional Probability

What is the probability of an event A given that an event B has occurred? This is called the **conditional probability** and is denoted

$$\mathbb{P}(A|B) = \text{the conditional probability.} \tag{2.26}$$

$\mathbb{P}(A|B)$ is the probability of A according to a new probability measure on the same sample space such that the outcomes not in B have probability zero, and that is consistent with the probability on the original space.

Theorem

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{2.27}$$

Proof: The “proof” of this relation is really just an observation of elementary properties of sets. Suppose A and B are contained in the sample set Ω . Write $\mathbb{P}(A) = A/\Omega$ to represent the percentage of the probability of A of $\mathbb{P}(\Omega)$, which is unity. Think of A/Ω as the area of a set A in a Venn diagram with A in a unit circle Ω . Then $\mathbb{P}(B) = B/\Omega$ and $\mathbb{P}(A \cap B) = AB/\Omega$. Then $\mathbb{P}(A|B) = AB/B = (AB/\Omega)/B/\Omega = \mathbb{P}(A \cap B)/\mathbb{P}(B)$.

Another way to prove Equation (2.27) is as follows: let ω_i be a sample from Ω that falls in the set B . Let $\tilde{\Omega} = \{\omega_i \in \Omega : \omega_i \in B \subset B \forall i\}$. Define $\tilde{\mathcal{F}} = \{C \cap B \forall C \in \mathcal{F}\}$. Then $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is a new probability space. If $\tilde{\mathbb{P}}(\tilde{\Omega}) = 1$, we must have $\tilde{\mathbb{P}} = \mathbb{P}/\mathbb{P}(B)$. Then $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$. □

The usual probability of an event A can be regarded as a conditional probability, conditional on

all possible events:

$$\mathbb{P}(A) = \mathbb{P}(A|\Omega) = \mathbb{P}(A \cap \Omega)/\mathbb{P}(\Omega) = \mathbb{P}(A)/1. \quad (2.28)$$

Bayes' Rule (Bayes' Theorem)

From (2.27), we immediately have:

$$\mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A), \quad (2.29)$$

therefore,

$$\boxed{\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}}. \quad (2.30)$$

This fundamentally important relation is called Bayes rule (or Bayes law or Bayes theorem). It is one of the most important relationships in all of probability and mathematical statistics. This interpretation evokes deep philosophical and basic revelations that will be taken up in later lectures—particularly in Section 3.4 to follow.

Joint Probability. Given two events $A, B \in \mathcal{F}$, the probability that A and B will occur jointly, i.e., at the same time, is called the **joint probability** and is written $\mathbb{P}(A, B)$. If the events are **independent**, then

$$\mathbb{P}(A, B) = \mathbb{P}(A)\mathbb{P}(B) \quad (2.31)$$

In general, the joint and conditional probabilities are related as follows:

$$\boxed{\mathbb{P}(A, B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(B)}. \quad (2.32)$$

So $\mathbb{P}(A, B) = \mathbb{P}(A \cap B)$.

2.6 Probability Properties in Terms of Probability Densities

Throughout these lectures we will confine our attention to cases in which every real-valued random variable X has a **probability distribution** G_X a **probability density function** (pdf) π (or sometimes π_X for specificity), such that

$$G_X(A) = \mathbb{P}(X^{-1}(A)), = \mathbb{P}(X(a) \in A), = \int_A \pi(x) dx. \quad (2.33)$$

Thus, $\pi(x)$ (with $dx = d\mu(X)$), is the Radon-Nikodym derivative of the measure μ_X with respect to Lebesgue measure on Borel sets in: \mathbb{R} . We write $dG_X/dx = \pi$. If x is a realization of X , the CFD of which is $G_X(x) = \mathbb{P}(X)(\omega \leq x, \omega \in \Omega)$ and we write $G(x) == G_X(x) = \int_{-\infty}^x \pi(x) dx$. In particular,

$$\mathbb{P}(a \leq x \leq b) = \int_a^b \pi(x) dx , \quad (2.34)$$

If the random variable takes on a countably infinite or finite numbers of discrete values, x_1, x_2, \dots then the probability of $X = x_i$ is determined by

$$\pi_i = \pi(x_i) = \mathbb{P}(\{\omega \in \Omega; X(\omega) = x_i\}) = \mathbb{P}(X = x_i), \quad i = 1, 2, \dots \quad (2.35)$$

The distribution function is then,

$$G(x) = \sum_{I(x_i \leq x)} \pi(x_i). \quad (2.36)$$

We will now lay out several key probability properties and relations that form the basic tools of modern statistics and applied probability theory that are key to subsequent developments. We begin with the *expectation* or *first moment* of X :

$$\bar{x} = \mathbb{E}[X] = \int_{\mathbb{R}} x \pi(x) dx \quad (2.37)$$

This is the mean of X , and the moment about the origin of the curve $\pi(x)$:

$$\bar{x} \text{ (denoted sometimes } \mu) = \frac{\int_{\mathbb{R}} x \pi(x) dx}{\int_{\mathbb{R}} \pi(x) dx} . \quad (2.38)$$

The second moment about the mean is the **variance**,

$$V(X) = \mathbb{E} \{(X - [X])^2\} = \int_{\mathbb{R}} (x - \bar{x})^2 \pi(x) \, dx. \quad (2.39)$$

The k th moment is obviously,

$$\mathbb{E} \{(X - \bar{x})^k\} = \int_{\mathbb{R}} (x - \bar{x})^k \pi(x) \, dx. \quad (2.40)$$

Other properties are listed as follows.

The **joint probability density** (of two random variable X and Y) is $\pi(x, y)$:

$$\mathbb{P} \{X \in A, Y \in B\} = \mathbb{P} \{X^{-1}(A) \cap Y^{-1}(B)\} = \iint_{A \times B} \pi(x, y) \, dx \, dy. \quad (2.41)$$

The variable X and Y are **independent** if,

$$\pi(x, y) = \pi(x)\pi(y). \quad (2.42)$$

The **covariance** of X and Y is the mixed central moment,

$$cov(X, Y) = \mathbb{E} \{(X - \bar{x})(Y - \bar{y})\} = \mathbb{E} \{XY\} - \mathbb{E}[X]\mathbb{E}[Y]. \quad (2.43)$$

and the **correlation coefficient** of X and Y :

$$corr(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y}, \quad (2.44)$$

where σ_X and σ_Y are the **standard deviations**,

$$\sigma_X = \sqrt{V(X)} \quad , \quad \sigma_Y = \sqrt{V(Y)}. \quad (2.45)$$

If $cov(X, Y) = 0$, the variables are **uncorrelated**. If there are independent, they are uncorrelated (since $\mathbb{E} \{(X - \bar{x})(Y - \bar{y})\} = \mathbb{E} \{(X - \bar{x})\} \mathbb{E} \{(Y - \bar{y})\} = 0$). They are **orthogonal** if $\mathbb{E} \{XY\} = 0$. Then $\mathbb{E} \{X^2\} + \mathbb{E} \{Y^2\} = \mathbb{E} \{(X + Y)^2\}$.

In correspondence with (2.27), the **conditional probability density** $\pi(x|y)$ of X given Y is the probability of X assuming Y is equal to its realization y ; i.e., in the limit when Y assumes values

in an interval shrinking to a single point y . We have

$$\pi(x|y) = \frac{\pi(x,y)}{\pi(y)}, \quad \pi(y) \neq 0. \quad (2.46)$$

Again we have the symmetry

$$\pi(x|y)\pi(y) = \pi(y|x)\pi(x) = \pi(x,y). \quad (2.47)$$

This leads to **Bayes rule in terms of probability densities**:

$$\pi(x|y) = \frac{\pi(y|x)\pi(x)}{\pi(y)}. \quad (2.48)$$

As we shall dwell on later, the pdf $\pi(x)$ on the right-hand side of 2.48 is called the **prior** probability; $\pi(y|x)$ is the **likelihood** and $\pi(x|y)$ is the **posterior** pdf. The denominator is a normalization factor: $\pi(y) = \int_{\mathbb{R}} \pi(y|x)\pi(x) dx$, because we must have $\int_{\mathbb{R}} \pi(y|x) dx = 1$ for any y . The denominator $\pi(y)$ is also called the **evidence** pdf.

These definitions are easily extended to n dimensions. For example, for vector-valued random variables $\mathbf{X} = \{X_1, X_2, \dots, X_n\}^T$, each component being a real-valued random variable. Then $\pi(\mathbf{x}) = \pi(x_1, x_2, \dots, x_n)$, the joint pdf, and $\bar{x}_i = \int_{\mathbb{R}^n} x_i \pi(\mathbf{x}) \pi(\mathbf{x}) d\mathbf{x}_1 d\mathbf{x}_2 \dots d\mathbf{x}_n$, $1 \leq i \leq n$.

The **covariance matrix** is the $n \times n$ matrix,

$$\text{cov}(X) = \int_{\mathbb{R}^n} (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T \pi(\mathbf{x}) d\mathbf{x}, \quad (2.49)$$

(with i, j components $\int_{\mathbb{R}^n} (x_i - \bar{x}_i)(x_j - \bar{x}_j)^T \pi(\mathbf{x}) d\mathbf{x}$, $d\mathbf{x} = dx_1 dx_2 \dots dx_n$).

For illustration purpose, we list some standard and common example of pdfs as follows.

2.7 Examples of Probability Density Functions

We write these in the form of a likelihood: $\pi(y|\theta)$, $Y \in y \equiv$ a realization of a random variable y (often representing observational data) and θ , a vector of parameters.

1. Gaussian (or Normal) Distribution.

$$\pi(y|(\mu, \sigma)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right), \quad -\infty < y < \infty. \quad (2.50)$$

with $\mu = \bar{y}$ as the mean, and σ^2 as the variance.

2. Cauchy Distribution.

$$\pi(y|(\mu, \tau^2)) = \frac{1}{\pi} \frac{\tau}{(y-\mu)^2 + \tau^2}. \quad (2.51)$$

3. Laplace Distribution.

$$\pi(y|(\mu, \tau)) = \frac{1}{2\tau} \exp\left(-\frac{|y-\mu|}{\tau}\right). \quad (2.52)$$

4. Poisson Distribution.

$$\pi(y|\lambda) = \frac{\lambda^y}{y!} e^{-\lambda}, \quad y = 0, 1, 2, 3, \dots \quad \& \quad 0 < \lambda < \infty, \quad (2.53)$$

$$\mathbb{E}[y] = \lambda \quad , \quad \mathbb{V}[y] = \lambda. \quad (2.54)$$

5. Pearson's Family of Distribution.

$$\pi(y|(\mu, \tau^2, b)) = \frac{\Gamma(b)\tau^{2b-1}}{\Gamma(b-\frac{1}{2})\Gamma(\frac{1}{2})} \frac{1}{[(y-\mu)^2 + \tau^2]^b}. \quad (2.55)$$

6. Weibull Distribution.

$$\pi(y|(\lambda, k)) = \begin{cases} \frac{k}{\lambda} \left(\frac{y}{\lambda}\right)^{k-1} \exp\left(-\frac{y}{\lambda}\right)^k, & y > 0 \\ 0, & y \leq 0. \end{cases} \quad (2.56)$$

7. Exponential Distribution.

$$\pi(y|\lambda) = \frac{1}{\lambda} \exp\left(-\frac{y}{\lambda}\right). \quad (2.57)$$

8. Rayleigh Distribution.

$$\pi(y|\sigma) = \frac{2}{2\sigma} \left(\frac{y}{\sqrt{2\sigma}} \right) \exp \left(-\frac{y^2}{2\sigma^2} \right). \quad (2.58)$$

Many other examples could be cited. Most often, we do not know a pdf or its parameters and we must determine it by solving a stochastic problem.

We note that when X is Gaussian, we write for shorthand,

$$X \sim \mathcal{N}(\mu, \sigma^2), \quad (2.59)$$

in which \mathcal{N} stands for “Normal” distribution, μ is the mean, and σ^2 is the variance. A Gaussian \mathbf{X} in \mathbb{R}^n is,

$$\pi(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp \left(-\frac{1}{2} (\mathbf{x} - \bar{\mathbf{x}})^T \Sigma^{-1} (\mathbf{x} - \bar{\mathbf{x}}) \right), \quad (2.60)$$

where Σ is the covariance matrix.

In closing this section, we record some fundamental classical results in probability theory for future reference. These include the central limit theorem and the law of large numbers.

The Central Limit Theorem

We note that:

Convergence in probability \Rightarrow convergence in distribution \Rightarrow convergence almost surely (a.s.).

Let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ be n real-valued random variables that are,

1. independent [i.e., $\pi(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) = \prod_{i=1}^n \pi(x_i)$], and
2. identically distributed [i.e., drawn from the same distribution],

each with expectation μ and variance σ^2 . Then the distribution of

$$Z_n = \frac{1}{\sqrt{n}\sigma} (\mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_n - n\mu) \quad (2.61)$$

converges to the distribution of a Gaussian random variable,

$$\lim_{n \rightarrow \infty} \mathbb{P}(Z_n \leq x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left(-\frac{x^2}{2\sigma^2} \right) \mathrm{d}x. \quad (2.62)$$

Equivalently, if

$$Y_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad (\text{average or mean}), \quad (2.63)$$

then, for large n ,

$$Y_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right). \quad (2.64)$$

The law of large numbers

Assume $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are i.i.d. random variables with finite mean μ . Then,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n \right) = \mu \quad (2.65)$$

almost surely (with probability 1) the realization $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ converges to the mean μ .

In these definition, *i.i.d.* stands for independent and identically distributed sample.

We close with comments on different notions of convergence in probability spaces. In particular, if X is random map from $(\Omega, \mathcal{F}, \mu)$ to a metric space (S, d) , a sequence X_n of random variables,

1. converges in probability, $X_n \xrightarrow{P} X$, in a metric d if $\forall \epsilon > 0$, $\mu(\omega \in \Omega : d(X_n(w), X(w)) > \epsilon) \rightarrow 0$ as $n \rightarrow \infty$, or,
2. converges in distribution, $X_n \xrightarrow{d} X$, $\mathbb{E}f(X_n) \rightarrow \mathbb{E}f(X)$ as $n \rightarrow \infty$, for every bounded continuous function f , and,
3. converges almost surely, $X_n \rightarrow X$, $\mu - a.s.$, if $\mu(\omega \in \Omega : \lim X_n(w) = X(w)) = 1$.

Chapter 3

The Logical Theory of Probability of R.T. Cox and E.T. Jaynes

3.1 Introduction

An alternative and quite different mathematical framework than Kolmogorov's for developing a theory of probability was begun by R.T. Cox in 1946 [31], expanded on in subsequent accounts in 1961 [32] and 1979 [33], and enriched and extended by E.T. Jaynes in his treatise, *Probability Theory: the Logic of Science*, published posthumously in 2003 [60]. A significant literature has accumulated on this so-called subjective or logical theory of probability based on Bayesian inference, including works on the Bayesian approach to scientific reasoning by Howson and Urbach [55], counter examples to Cox's original examples of Halpern [49], “patches” (corrections) to the theorem by Halpern [48], and alternate proofs and interpretations by Arnborg and Sjodin [10], Knuth and Skilling [70], Zimmerrnan and Cremers [26], and Dupre and Tipler [38]. A “guide to Cox's Theorem” was provided by a comprehensive account of Van Horn [116], and a recent paper by Terenin and Draper [110] is devoted to “rigorizing and extending the Cox-Jaynes derivation of probability”. We provide in this chapter brief overview of Cox-Jaynes theory, drawing from Jaynes [60], Van Horn [116] and Terenin and Draper [110].

3.2 Plausible Reasoning

The notion of logical or CJ (Cox-Jaynes) probability begins with the concept of *plausible reasoning*. The idea is that in science, and indeed, in every day life, we rarely have enough information to permit deductive reasoning. Instead, we make plausible conclusions based on past experience. Certain reasoning processes assess the plausibility of a proposition accepting that it may not be a certainty.

Thus, we regard the logical approach to scientific reasoning as one way of assessing the plausibility of a proposition (as opposed to the frequency of an event) that can be true or false or that can be assigned a degree of plausibility. The appropriate mathematical setting for such considerations is that furnished by a *Boolean algebra*.

Propositional Boolean Algebra. A propositional Boolean algebra C consists of the six-tuple,

$$C \sim \{\mathcal{A}, \wedge, \vee, \neg, \mathbf{0}, \mathbf{1}\} \quad (3.1)$$

wherein

- \mathcal{A} = a finite or (countably or uncountably) infinite set of true propositions A, B, \dots
- Two binary operations, \wedge and \vee , and a unitary operation \neg on the propositions in \mathcal{A} :
 $A \wedge B$ = the *conjunction* of A and B , meaning that both A and B are true. $A \vee B$ = the *disjunction* of A and B , meaning that at least one of A and B is true (“ A or B ”)
 $\neg A = A \neg$ = the negation of A , meaning A is false
- Two identity elements $\mathbf{0}$ and $\mathbf{1}$

all such that,

- 1) C is *closed* under countable applications of \wedge and \vee and
- 2) For all propositions $(A, B, C) \in \mathcal{A}$, the following five Boolean axioms hold:

Associative Laws

$$A \wedge (B \wedge C) = (A \wedge B) \wedge C; \quad A \vee (B \vee C) = (A \vee B) \vee C$$

Commutative Laws

$$A \wedge B = B \wedge A \quad ; \quad A \vee B = B \vee A$$

Identity Elements

$$A \wedge \mathbf{1} = A \quad ; \quad A \vee \mathbf{0} = A$$

Complementation or Negation

$$A \wedge (\neg A) = \mathbf{0} \quad ; \quad A \vee (\neg A) = \mathbf{1}$$

Distributive Laws

$$A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$$

$$A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C).$$

The conjunction of two propositions, $A \wedge B$, is also written AB as their “product” and the disjunction, $A \vee B$, written as their sum $A + B$, all in analogy with the (“two-valued”) algebra of numbers. In this context, the Boolean algebra of propositions is said to follow the *rules of Aristotelian logic*.

The plausibility of a proposition A given that a proposition B is true is characterized by a real-valued function $pl(\cdot|.)$ on pairs of propositions:

$$pl : \mathcal{A} \times \mathcal{A}/\mathbf{0} \rightarrow \mathbb{R}^+. \tag{3.2}$$

Thus, $pl(A|B)$ is the degree of plausibility of proposition A given that B is true. If B is true, A becomes plausible, while evidence does not guarantee that A is true, but by observation of one or more of its consequences gives confidence that A is plausible. Also, if, given B , A is more plausible than C , we have $pl(A|B) > pl(C|B)$. Other fundamental rules will follow.

Jaynes Example: From Jaynes [60]: suppose we state the proposition $A \sim$ “it will start raining at 10:00 a.m. at the latest”, and $B \sim$ “the sky will become cloudy before 10:00 a.m”. Plausible reasoning not only establishes whether a proposition is more or less plausible, it evaluates the degree of plausibility in some way. For example the darkness of the clouds may suggest the degree

of plausibility of rain. □

3.3 Logic and Philosophical Theories of Probability

Logic is the science of reasoning, an intellectual discipline with origins dating millennia, that provides the cornerstones to mathematics, science, and all forms of human knowledge.

As noted in Chapter 1, deductive logic or logical deduction is a form of reasoning from one or more premises or axioms so as to deduce a logically certain conclusion. Deductive logic is said to be “top-down”, infallible and irrevocable when its conclusions are reached by following strict laws of reasoning. Mathematics is a discipline founded on deductive logic – one lays down axioms and definitions and then deduces their consequences following rules of logic.

Inductive logic, or “bottom-up” logic, represents a system of reasoning in which logical hypotheses or explanations are put forth to explain observed events or phenomena. Inductive arguments are, therefore, fallible – they can prove to be false when new information is acquired.

While deductive logic is absolute and infallible, it is independent of observation or experiment and, therefore, provides no way to test the validity of the axioms on which it is based. Inductive logic, on the other hand, can never be used to absolutely prove a scientific theory.

According to Jaynes [60], all knowledge is inductive. In deductive reasoning, if A and B are elements or propositions, one writes $A \Rightarrow B$ to mean if A is true, B is logically deduced from A . According to Jaynes, in almost all situations, “we do not have the right kind of information to allow this kind of deductive reasoning. In the formal logic of plausible reasoning $A \Rightarrow B$ means only that A and B have the same “truth value”.

If A and B are true propositions, a proposition C which is a function f of A and B , ($C = f(A, B)$) can be either true or false. Since A and B take on values [TT, TF, FT, FF], there are 16 logical functions: $f(T, T) = T, f(T, F) = F, \dots$ etc. A logical function of n propositions as 2^n function values. One can therefore construct truth tables analogous to the elementary logic of operations on real numbers, to cover possible truth values of such logic functions.

Plausible reasoning not only establishes whether a proposition is more or less plausible, it evaluates the degree of plausibility in some way (e.g. the darkness of the clouds may suggest the plausibility of rain). Examples of the Boolean algebra of propositions are:

A, B, C, \dots = propositions = unambiguous, simple, definitive, statements that are either true or false (two-valued, Aristotelian logic)

As noted earlier, conjunction “cojoins” two propositions into a “logical product”:

$$A \wedge B = AB \quad (\text{both } A \text{ and } B \text{ are true})$$

Disjunction, called the “logical sum”, “disjoins” two propositions,

$$A \vee B = A + B \quad (\text{at least one of } A \text{ and } B \text{ are true})$$

Negation describes “denial”,

$$\neg A \equiv A \quad \text{is false}$$

We understand equality and identity of proposition as

$$\begin{aligned} A &= B && (A \text{ and } B \text{ are equal in truth (or “truth value”)}) \\ A &\equiv B && (A \text{ and } B \text{ are equal by definition}) \end{aligned}$$

Some properties:

$A \vee \neg A =$ a necessary (or logical) truth (A is either true or false),

$A \wedge \neg A =$ a necessary (or logical) falsehood (A is true and false),

$\neg A \wedge B = AB$ is false,

$\neg A \wedge \neg B =$ both A and B are false,

if $C = A \wedge B$, then $\neg C = \neg A \vee \neg B$,

if $D = A \vee B$, then $\neg D = \neg A \wedge \neg B$.

These properties are recognized as analogous to elementary set properties with A, B, C, D sets, $AB = A \cup B$, $A + B = A \cap B$, $\bar{A} = A'$, etc. $A \Rightarrow B$ (A implies B) does not assert that either A or

B is true, it means $A \wedge \bar{B}$ (A and (B is false)) is false or $\bar{A} \vee B$ (either A is false or B is true) is true.

Technically, propositional statement of Boolean algebras in symbolic logic follow rules similar to ordinary algebra, with values **0** or **1**, and no operations other than conjunction (multiplication) and disjunction (addition). To understand true logical meaning of some functional (called a *logic function*) relationship among proposition, we use the rules just stated to construct truth tables similar to those encountered in constructing the foundations of algebra of numbers. In two-valued logic, any proposition has a truth value of 0 or 1, with a true proposition being assigned 1 and a false proposition 0. Two propositions are equal if they have the same truth value.

For example, consider the compound propositions $A = \neg(B \wedge C)$ (the negation of $B \wedge C$) and $D = \neg B \vee \neg C$. The truth value of these two logic functions is established by a truth table of the following form:

a)	b)	c)	d)
B	C	$B \wedge C$	$\neg(B \wedge C)$
T	T	T	F
T	F	F	T
F	T	F	T
F	F	F	T

Column *d*) describe the truth value of $A = \neg(B \wedge C)$. Now look at $A' = \neg B \vee \neg C$ and the truth table:

e)	f)	g)	h)
$B \vee C$	$\neg B \vee C$	$\neg B \wedge \neg C$	$\neg B \vee \neg C$
T	F	F	F
T	F	F	T
T	F	F	T
F	T	T	T

Thus the proposition $B' = \neg B \vee \neg C$ (column *h*) has the same truth value as A' , (column *d*), so A' and B' are *logically equal*, they have the same truth value. This is one of many examples of propositional logic that can be derived from straightforward applications of elementary Aristotelian

logic. Other examples can be found in the book of Jaynes [60], in the book on *Mathematics and Plausible Reasoning* by G. Polya [86]; a summary can be found in the text of Gregory [45].

3.4 Van Horn's Guide to Cox's Theorem

It is informative to follow a few arguments given by Van Horn in his guide to Cox's Theorem [116]. There are two central notions:

- 1) *the state of information* X summarizes the information one has on what is called the “atomic proposition Λ ”, those that cannot be decomposed into other propositions. These atomic propositions form the *basis* for the state of information X .
- 2) the *plausibility of a proposition* $A \in \mathcal{A}$, given the state of information X , is a number $pl(A|X)$. The state of information obtained from A by adding the additional information that X is true; thus written (AX) or $(A|X)$.

We may now lay down several assumptions. In the notation inspired by Van Horn; there are five “rules **Ri**”:

R1. The plausibility $pl(A|X)$ of A given X is a single non-negative real number. Moreover there exists a number T (the plausibility of a known true propositions) such that $pl(A|X) \leq T$. We say that X is *consistent* if there is no proposition A for which both $pl(A|X)$ and $pl(\neg A|X) = T$ (are true).

R2. Plausibilities are consistent with so-called propositional calculus, in the sense that the following hold:

- 1) if $A \Leftrightarrow A'$, meaning that “ $A \Leftrightarrow B$ ” is a *tautology*, (meaning a fundamental truth in formal logic), then $pl(A|X) = pl(A'|X)$.
- 2) if A is a tautology, then $pl(A|X) = T$.
- 3) $pl(A|B, C, X) = pl(A|(B \wedge C), X)$.

4) if X is consistent and $pl(\neg A|X) < T$, then A, X is also consistent.

R3. There is a non-increasing function S such that $pl(\neg A|X) = S(pl(A|X))$ for all A and consistent X .

R4. (Universality) There exists a nonempty set Y_o that is a dense subset of $(S(T), T)$. Thus, for real numbers $a, b, c \in Y_o$ there is a consistent X and three atomic propositions A_1, A_2, A_3 such that $a = pl(A_1|X)$, $b = pl(A_2|A_1, X)$ and $c = pl(A_3|A_2, A_1, X)$.

Rules **R1-R4** are set up to establish more or less commonsense properties of properties of the plausibility of a proposition. **R1** establishes the reasonable notion that real numbers can be used to represent degrees of plausibility, and that larger numbers represent higher degrees of plausibility, T being the real numbers, representing the plausibility of a known true propositions. The idea that for a given state of information X , two plausibilities A and B are equally plausible or one is more plausible than the other, is an assumption of *universal comparability*. The idea is not without some controversy. Some (e.g. Shafer [102]) propose that a “two-dimensional” logic may be called for, in which both the “degree of belief” of a proposition as well as the belief in its negation is justified. Jaynes [60] and Van Horn [116], among others, provide strong arguments in favor of **R1**.

For **R2**, we simply demand that if two propositions are equivalent, we can use them interchangeably, and that the plausibility of any proposition known to be a tautology should be assigned a value reflecting this knowledge. Condition 3) in **R2** is obvious, and in 4), if we cannot say with certainty that the negation $\neg A$ of A is true, there remains the fact that A may be true and, then, cannot contradict assertion on $\neg A$.

In **R3** (Negation), we demand that A and $\neg A$ are opposite answers to the same question: if A is known to be true, then $\neg A$ is false $pl(A|X)$ and $pl(\neg A|X)$ cannot vary independently in two-dimensional Aristotelian logic.

In **R4**, the goal is to lay down rules that enable plausible reasoning to fall under a universal system of logic applicable to any problem domain in which meaningful propositions can be formulated, allowing the range of the set of allowed plausibilities to be dense in $(S(T), T)$. Van Horn [116] comments that Halpern [48] argues that **R4** is problematic as it cannot be satisfied on finite domains, an assertion Van Horn discounts. We return to this point later.

Conjunction: *The Product Rule of Plausibilities.* We now come to the centerpiece of Cox's theorem, the conjunction or product rule for propositions: given consistent X and the conjunction $A \wedge B = AB$ of two propositions conditioned on X , what can be said about the relationship of $pl(A \wedge B|X)$ with the plausibilities $pl(A|B, X)$ and $pl(B|X)$? A systematic list of arguments leads to the following:

R5. Given $S(T)$ and a tautology T , there exists a continuous function $F : [S(T), T]^2 \rightarrow [S(T), T]$, strictly increasing on $(S(T), T]$, such that

$$pl(A \wedge B|X) = F[pl(A|B, X), pl(B|X)] \quad (3.3)$$

for any A, B and consistent X .

The establishment of (3.3) follows some surprisingly intuitive arguments. They begin with writing down all the combinations of A, B , and X that $pl(A \wedge B|X)$ could be a function of (omitting “ pl ” temporarily, for simplicity), they are $(A|X)$, $(B|X)$, $(A|B, X)$ and $(B|A, X)$ and all combinations of these numbers, 15 in all. Because $A \wedge B = B \wedge A$, only 9 of these survive. Van Horn [116], following Tribus [111], rules out as many of these as possible:

- 1) $(A \wedge B|X) = F[(A|X)]$
- 2) $(A \wedge B|X) = F[(A|BX)]$
- 3) $(A \wedge B|X) = F[(A|X), (A|BX)]$
- 4) $(A \wedge B|X) = F[(A|BX), (B|AX)]$
- 5) $(A \wedge B|X) = F[(A|X), (B|X)]$

Remarkably, all five of these candidates lead to a contradiction and can, therefore, be discarded. For example, in 1), if A is a tautology (a fundamental truth), and B is a “atomic” proposition, then $A \wedge B$ is equivalent to B and $(A|X) = T$, so $S(T) = (B|X)$. But X can always be chosen to make $(B|X)$ any value in Y_0 . So we have a contradiction. 1) cannot, therefore hold. Similarly, in 3), for instance, if A is a tautology and B is any atomic proposition, then $S(T) = (B|X)$, a contradiction. Similar arguments can be made for all these candidates leaving only four possibilities out of the original 15:

- I) $(A \wedge B|X) = F[(A|BX), (B|X)],$

- II) $(A \wedge B|X) = F[(A|B, X), (B|X), (A|X)],$
- III) $(A \wedge B|X) = F[(A|B, X), (B|X), (B|A, X)],$
- IV) $(A \wedge B|X) = F[(A|B, X), (B|X), (B|A, X), (A|X)].$

Van Horn refers to the arguments of Tribus [111] who claims that III) and IV) can be ruled out because “they cannot avoid dealing with plausibility expressions that involve some inconsistent state of information $X = Y$, and hence an unavoidable ambiguity”. Also, rule II) is eliminated for certain similar reasons, leaving only I). But Van Horn asserts that “there is really no compelling reason for choosing any one of the four remaining candidates”. He leans toward I) as the other candidates merely add additional conditions to those used in I) and because Jaynes [60] provides an intuitive argument in support of I).

Assuming I) holds, next consider the plausibility $A \wedge B \wedge C|X)$, X being consistent. Applying rule I), with “ B ” now replaced by $(B \wedge C)$, gives

$$(A \wedge (B \wedge C)|X) = F[(A|(B \wedge C), X), ((B \wedge C)|X)] = F[(A|B, C, X), F[(B|C), X], (C|X)] \quad (3.4)$$

Now use the equivalent grouping of the propositions on the left-hand side:

$$((A \wedge B) \wedge C|X) = F[((A \wedge B)|C, X), (C|X)] = F[F[(A|B, CX), (B|C, X)](C|X)] \quad (3.5)$$

But $A \wedge (B \wedge C) = (A \wedge B) \wedge C$, so the functions derived in (3.4) and (3.5) must coincide. Setting

$$x = (A|B, C, X),$$

$$y = (B|C, X), \text{ and}$$

$$z = (C|X),$$

where $x, y, z \in [0, T]$, we are looking for a function F that for numbers x, y, z is such that,

$$F(x, F(y, z)) = F(F(x, y), z) \quad . \quad (3.6)$$

Since F is continuous and is dense, this relation holds for all $x, y, z \in [0, T]$. This property of denseness is fundamental to Cox’s theorem. The fact that “ \wedge ” is associative ensures that F is associative, which limits choices for F to be “isomorphic to multiplication”. A little inventory:

if we can find any function F satisfying (3.6), then, by rule I), we will have determined how the “product” $A \wedge B$, conditional on X , depends on $(A|BX)$ and $(B|X)$. Thus, we establish (3.3).

Equation (3.6) is called the *associativity equation*; it is an example of a *functional equation*. Such equations are studied in the book of Aczel [1], where the following theorem can be found.

Aczel’s Theorem. *Let $f : (a, b]^2 \rightarrow (a, b]$ be a continuous, strictly increasing function in both arguments a and b , with $a, b \in \mathbb{R}$ and $a < b$, and suppose that for all $x, y, z \in (a, b]$,*

$$f(x, f(y, z)) = f(f(x, y), z).$$

Then there exist a continuous, strictly increasing function g such that

$$g(f(x, y)) = g(x) + g(y)$$

for all $x, y \in (a, b]$. □

Now set $f = F$; $a = S(T)$, $b = T$ and define a real-valued, continuous, increasing, and non-negative function $w = w(x)$ on $(S(T), T]$ such that

$$\ln w(x) = g(x), \quad x \in [S(T), T].$$

with $w(S(T))$ the limit as $x \rightarrow S(T)$. We then must have by continuity

$$w(F(x, y)) = w(x)w(y),$$

for $x, y \in [S(T), T]$. Thus, for every $A, B \in \mathcal{A}$ and consistent X , we have

$$w((A \wedge B|X)) = w((A|B, X))w((B|X)). \quad (3.7)$$

The function w thus suggests that conditioned probabilities obey a “product rule” similar to probabilities (e.g. $\mathbb{P}(A \cup B) = \mathbb{P}(A|B)\mathbb{P}(B)$, now conditioned on information X).

A final step for the product rule calls upon the fact that,

$$w(S(T)) = 0, \quad w(T) = 1, \quad \text{and} \quad 0 < w(x) < 1 \quad \text{for } S(T) < x < T.$$

A proof of this assertion is given in [116]. Briefly, it involves the observation that for any proposition A , consistent X , and tautology D ,

$$w((A|X)) = w((D \wedge A|X)) = w((D|A, X))w((A|X)) = w(T)w((A|X)),$$

so dividing by $w((A|X))$ gives $w(T) = 1$ and, therefore, $w(x) < 1$ for $x < T$. The proof that $w(S(T)) = 0$ follows from contradiction: assume otherwise and consider $A_1, A_2 \in \mathcal{A}$ and $x \in Y_o$. If $w(S(T)) = y > 0$, then $0 < z < 1$ and $z < \sqrt{z} < 1$, $w^{-1}(z) < x < w^{-1}(\sqrt{z})$. If $(A_1|X) = (A_1, A_2|X) = x$, $w((A_1 \wedge A_2|X)) = w^2(x) < z = w(S(T))$, so $(A_1 \wedge A_2|X) < S(T)$, a contradiction.

So we can add to 3.7, the condition that the monotone continuous function satisfying 3.7 must also satisfy,

$$0 \leq w(x) \leq 1, \quad (3.8)$$

with $w(x) = 0$ when x is impossible and $w(x) = 1$ when x is certain.

Disjunction: *The Sum Rule for Plausibilities.* The aim of the sum or disjunction rule for plausibilities defined on pairs of propositions is to establish a relationship involving the plausibility of a proposition A , conditioned on consistent X , and its negation $\neg A$ conditioned on X . Full details can be found in Van Horn's guide [116] and in the paper of Paris [84]. It suffices here to point out that a similar thread of logic as that in the conjunction relations lead again to a functional equation of the form,

$$yS\left(\frac{x}{y}\right) = S(x)S(S(y)/S(x)) \quad (3.9)$$

for all $0 < x \leq y < 1$, where

$$yS\left(\frac{x}{y}\right) = p(\neg A \wedge B|X)$$

$$S(y) = p((A \vee \neg B) \wedge (\neg A \vee \neg B)|X)$$

$$S(x) = p(\neg A \vee \neg B|X),$$

$p(x) = w(x)^m$, $m = -\log(2)/\log w(\alpha)$, and α is a fixed point of a map such that $S(A|X) = (\neg A|X)$ for all A and consistent X . From these relations, it possible to show that,

$$w^m(\neg A|X) + w^m(A|X) = 1 , \quad (3.10)$$

for positive m . This is known as the *sum rule*. We note that the product rule 3.7 can be equally-well be written,

$$w^m((A \wedge B|X)) = w^m((A|BX))w^m((B|X)),$$

for positive m . For whatever m we choose, we can set,

$$pl(x) = w^m(x).$$

Collecting all of the results laid down thus far, we arrive at the following version of Cox's theorem:

Cox's Theorem as given by Van Horn.

Given that rules R1 - R5 hold, there exists a continuous, strictly increasing function $pl : \mathcal{A} \times \mathcal{A} \setminus \mathbf{0} \rightarrow \mathbb{R}^+$ such that for every A and $B \in \mathcal{A}$ and consistent X ,

- 1) $pl(A|X) = 0$ iff A is known to be false, given the information X .
- 2) $pl(A|X) = 1$ iff A is known to be true given the information X .
- 3) $0 \leq pl(A|X) \leq 1$.
- 4) $pl(A \wedge B|X) = pl(A|X) pl(B|A, X)$
- 5) $pl(\neg A|X) = 1 - pl(A|X)$.

Because of the commutative property, $A \wedge B = B \wedge A$, assertion 4 in Cox's theorem implies that,

$$pl(A|X)pl(B|AX) = pl(B|X)pl(A|BX), \quad (3.11)$$

which is recognized as the Bayes rule, provided we interpret the plausibilities as probabilities.

The five properties laid down in Cox's theorem are formally the rules of probability theory, stated in terms of plausibilities. As pointed out by Van Horn, “any system of plausible reasoning that is not isomorphic to probability theory must necessarily violate one of the requirements” (in the theorem). So “set-theoretic probability then becomes a model theory for our logic”. We may thus

interpret states of information to be probability distributions, with AX understood to be “the probability distribution obtained from X by conditioning the set of values for which A is true”. Henceforth, we refer to the Cox-Jaynes plausibilities as probabilities in the spirit of those covered in the previous chapter.

Property 4) sets the Cox-Jaynes theory apart, in a sense, from frequentist theory, featuring conditional probability and Bayes’ rule as central property of the theory. The existence of the Kolmogorov set - theoretic framework of probability, which involves the same rules, proves that these rules are not contradictory, the primary difference being that in Cox theory, conditional probability now is the fundamental measure. In the Kolmogorov theory described earlier, all probabilities are conditioned on the choice of the sample set Ω , as indicated in equation (2.28) in section 2.5.1.

3.5 The Terenin-Draper Account of Cox’s Theorem and Cox-Jaynes Probability

Terenin and Draper [110] provide a detailed account of the structure and different philosophical issues underlying major probability theories, particularly Kolmogorov’s set-theoretic theory based on frequentist definition of probability and the Cox-Jaynes theory based on the notions of plausible reasoning. Some key aspects of their arguments are summarized in the following paragraphs:

- 1) In Jaynes, *Logic of Science*, he insists that the framework of his theory is established for only *finite sets of propositions* (generally meaning it can be applied to statistical models with only a finite numbers of parameters), with “extensions to countably infinite sets permitted only when this is the result of a well-defined and well-behaved limiting process from a finite set”. Terenin and Draper show this position is “overly cautious”. The Van Horn proof also does not assume the sets of propositions are necessarily finite.
- 2) Terenin and Draper claim that Van Horn’s notion of *state of information* is defined at a non-rigorous level, making it hard to evaluate. They go on to “rigorize” the statement and proof of Cox’s theorem.
- 3) They invoke *Stone’s Representation Theorem* [108], which they interpret to establish that

“every Cox-Jaynes propositional Boolean algebra $(\mathcal{A}, \wedge; \vee, \neg, 0, \mathbf{1})$ is isomorphic to a Kolmogorov set-theoretic space $(\mathcal{F}, \cup, \cap, (\cdot)^c, \emptyset, \Omega)$.

4) They invoke the conditions of *sequential continuity* for non-empty (or true) B :

a) if $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$ such that $A_i \nearrow A$, then $\mathbb{P}(A_i|B) \nearrow \mathbb{A}|B$

b) if $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$ such that $A_i \searrow A$, then $\mathbb{P}(A_i|B) \searrow \mathbb{A}|B$

5) They provide a extension of the σ -algebra to resolve the requirement of denseness in the range of the probability function \mathbb{P} ; specifically, for any subset M of \mathbb{R} , they define $S = [\min M, \max M]$ and Ω_E (= the extension of the sample set Ω) = $\Omega \times S$, and \mathcal{F}_E (= the extension of the σ -algebra \mathcal{F}) = $\mathcal{F} \otimes \mathcal{B}(S)$, $\mathcal{B}(S)$ being the Borel σ -algebra on S . One then can define the extended probability function $\mathbb{P}_E : \mathcal{F}_E \times \mathcal{F}_E \setminus \emptyset \rightarrow S$ that satisfies all of the axioms for probabilities (plausibility) including the product rule and the sum rule. They justify that this density property “defeats” the counterexample of Halpern [49]. Further details can be found in [110].

Example Application

We close this chapter by showing that several of the elementary properties of frequentist probability can be deduced from the Cox-Jaynes theory with a few straightforward arguments. We again call upon Jaynes [60] for these arguments. The first case we mention is the notion of exclusivity.

Given a set of propositions; say A_1, A_2, A_3 , Given B , what is the probability (plausibility) that *at least one* is true? that is, $pl(A_1 + A_2 + A_3|B)$? By the sum rule,

$$pl(A_1 + A_2 + A_3|B) = pl(A_1|B) + pl(A_2|B) + pl(A_3|B) - pl((A_1 + A_2)A_3|B). \quad (3.12)$$

But $(A_1 + A_2)A_3 = A_1A_3 + A_2A_3$. Applying the sum rule again, we get,

$$\begin{aligned} pl(A_1 + A_2 + A_3|B) &= pl(A_1|B) + pl(A_2|B) + pl(A_3|B) \\ &\quad - pl(A_1A_2|B) - pl(A_1A_3|B) - pl(A_2A_3|B) \\ &\quad + pl(A_1A_2A_3|B). \end{aligned} \quad (3.13)$$

Let the A_i be *mutually exclusive* (evidence B implies that no two can be true simultaneously).

Then,

$$pl(A_i A_j | B) = pl(A_i | B) \delta_{ij}. \quad (3.14)$$

Then the *last four terms vanish* (e.g., if only B implies A_1, A_2 is true, $pl(A_2 A_3 | B) = 0 = pl(A_1 A_3 | B)$ and $pl(A_1 A_2 | B) + pl(A_1 A_2 A_3 | B) = 0$). Then for three mutually exclusive propositions,

$$pl(A_1 + A_2 + A_3 | B) = \sum_{i=1}^3 pl(A_i | B), \quad (3.15)$$

or, for n such propositions,

$$pl(A_1 + A_2 + \dots + A_m | B) = \sum_{i=1}^m pl(A_i | B) \quad , \quad 1 \leq m \leq n. \quad (3.16)$$

If in addition to being mutually exclusive, the propositions are also *exhaustive*, meaning that B stipulates that one and only one can be true (Jaynes [60], p.38), then for $m = n$, $pl(A_1 + A_2 + \dots + A_m | B) = 1$ and equation (3.16) implies that,

$$\sum_{i=1}^n pl(A_i | B) = 1. \quad (3.17)$$

Now let $\{A_1, A_2, \dots, A_n\}$ be a collection of n mutually exclusive and exhaustive propositions and B evidence on these propositions. A simple argument can be made that if B is indifferent between all of the propositions, then $pl(A_i | B) = pl(A_j | B)$ for $1 \leq i, j \leq n$. Then from Equation 3.17,

$$pl(A_i | B) = \frac{1}{n} \quad 1 \leq i \leq n. \quad (3.18)$$

Thus, we arrive at a classical result directly analogous to the frequentist approach.

3.6 Summary Comments

Among conclusions that can be drawn from the developments outlined in this chapter are the following:

- The Cox-Jaynes theory of logical probability, based on the notions of plausible reasoning and on plausibilities defined on propositional Boolean algebras, is a self-consistent theory, equivalent to the Kolmogorov theory, but structured in a framework in which conditional probability is the primitive concept. In this sense, there is no \mathcal{P} -uncertainty if this theory of probability is employed to deal with uncertainties.
- “Every natural extension of Aristotelian logic with uncertainties is Bayesian” (c.f. [60]).

Chapter 4

Statistical Calibration and Validation of Parametric Models.

4.1 Introduction

Recall that in Chapter 1 we introduced in equation (1.1) the notion of an abstract mathematical model defining a forward problem with a solution $u = u(\boldsymbol{\theta}, S)$ depending on parameters $\boldsymbol{\theta} \in \Theta$ and set in a scenario S for a problem $\mathcal{A}(\boldsymbol{\theta}, S; u(\boldsymbol{\theta}, S)) = 0$. Now we return to this setting and consider the question of predictability of such models in the presence of uncertainty-addressing the principal sources of uncertainty described in Section 1.3: \mathcal{P} , \mathcal{M} , \mathcal{Y} , $\boldsymbol{\theta}$, and h uncertainty.

We dispense with \mathcal{P} uncertainty immediately. We argued in Chapter 3 within the framework of Cox-Jaynes probability theory (and equivalently, Kolmogorov theory), that there is no \mathcal{P} -uncertainty and that every natural extension of Aristotelian logic in the presence of uncertainties is Bayesian [110]. We will therefore employ tools and also concepts from the equivalent Kolmogorov theory when convenient.

This leaves us with \mathcal{M} , \mathcal{Y} , $\boldsymbol{\theta}$, and h uncertainties. In [110], it is stated that “all uncertainty in probabilistic models is (therefore) the result of model specification: either via the parameters ($\boldsymbol{\theta}$ -uncertainty), or via the structure of the model itself (\mathcal{M} -uncertainty)”. We add to this uncertainty in the data (\mathcal{Y} -uncertainty) and, since we are interested in computational models obtained through discretization of mathematical models, h-uncertainty as well.

This latter area, h -uncertainties, is in the realm of *model verification* and is either an issue of code verification, requiring such things as documentation, version control, benchmarks, etc., or solution verification, which is a matter of mathematically-based *a posteriori*, error estimation. These topics fall outside the scope of this exposition and will not be considered further.

That leaves us with \mathcal{M} , \mathcal{Y} , and $\boldsymbol{\theta}$ uncertainties. \mathcal{M} -uncertainty is a deep and challenging subject involving the study of model selection, model inadequacy, and the complicated interaction of uncertainties in the model with those in data and parameters. We take up this important subject after first looking at \mathcal{Y} and $\boldsymbol{\theta}$ uncertainty, later in the chapter. First, we address \mathcal{Y} and $\boldsymbol{\theta}$, beginning with their appearance in Bayes' rule, where one finds naturally the push and pull of the two classical pillars of science: observation – with uncertain data $\mathbf{y} \in Y$, and theory, represented by models with uncertain parameters $\boldsymbol{\theta} \in \Theta$.

Thus, from Bayes' rule, we convert to probability densities π , and from the rule $\mathcal{P}(A|B) = \mathcal{P}(B|A)\mathcal{P}(A)/\mathcal{P}(B)$, we let B represent the observational data \mathbf{y} for scenario S and A represent the model parameters $\boldsymbol{\theta}$, leading to

$$\boxed{\pi(\boldsymbol{\theta}|\mathbf{y}, S) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}, S)\pi(\boldsymbol{\theta}|S)}{\pi(\mathbf{y}|S)}}, \quad (4.1)$$

where $\pi(\boldsymbol{\theta}|S)$ is the *prior* density of parameters $\boldsymbol{\theta} \in \Theta$ for science S , $\pi(\mathbf{y}|\boldsymbol{\theta}, S)$ is the *likelihood*, $\pi(\boldsymbol{\theta}|\mathbf{y}, S)$ is the *posterior* density, and $\pi(\mathbf{y}|S)$ is the *evidence*. Since $\int_{\Theta} \pi(\boldsymbol{\theta}|\mathbf{y}, S) d\boldsymbol{\theta} = 1$ for all \mathbf{y} in the space \mathcal{Y} of observational data and for all scenarios S (S_c , the calibration scenarios, S_v the validation scenarios, and S_p the prediction scenario), then the evidence must be the marginalization of the numerator,

$$\pi(\mathbf{y}|S) = \int_{\Theta} \pi(\mathbf{y}|\boldsymbol{\theta}, S)\pi(\boldsymbol{\theta}|S) d\boldsymbol{\theta}. \quad (4.2)$$

Here $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n) \in \mathcal{Y}_s \subset \mathbb{R}^n$, and it is understood that (4.2) holds for every scenario S , the data \mathbf{y} being different in each.

4.2 Data Uncertainty: Experimental Noise

Bayes' rule (4.1) embraces all of the pillars of science drawn upon to acquire (or increase) one's knowledge: observations, characterized by observational data $\mathbf{y} \in \mathcal{Y}$; theory, mathematical representation of hypotheses founded on inductive logic put forth to explain observed events – as represented by mathematical models, and computations needed to evaluate the resulting distributions. To the statistician, models are represented by the likelihood, $\pi(\mathbf{y}|\boldsymbol{\theta}, S)$, $\boldsymbol{\theta}$ representing specific features of the model and $\pi(\boldsymbol{\theta}|S)$, the probability distribution for data given $\boldsymbol{\theta}$, for any scenario S ; the outcome, $\pi(\boldsymbol{\theta}|\mathbf{y}, S)$ is the result of the information in the data and the predictability of the model. To the scientist or engineer, the model is reached after a process of using the solution of the forward problem to evaluate the observations available to the models. These predictions of \mathbf{y} of the models are quantities $\mathbf{d}(\boldsymbol{\theta}, S)$ which rarely coincide with the recorded observations \mathbf{y} .

Let us consider the problem of using model (4.1) for a given scenario S to predict the value of some physical entity. The target prediction could be, for example, the temperature at a point in a turbine engine, the volume fraction of cancer cells in a sample of human tissue, the maximum von-Mises stress in a machine part, or the time-to-failure of a steel shaft in an aircraft engine. Such quantities are rarely deterministic. A fundamental assumption in classical mathematical statistics is that the physical entity we wish to predict with model (4.1) in any particular scenario S is determined by a stochastic law (the “true” structure of the entity) characterized by a probability distribution \mathbf{G} with density \mathbf{g} . The structure of \mathbf{g} (or \mathbf{G}) is available to us only through samples taken from a set of experiments, the experimental observations \mathbf{y} , which are generally corrupted by experimental noise $\boldsymbol{\epsilon}$ (or experimental errors), since no measuring device is perfect. The observations \mathbf{y} are realizations of a random variable Y .

But the observational data we record is not the truth; rather it is some function \mathbf{f} of the true density \mathbf{g} and of unknown experimental errors $\boldsymbol{\epsilon}$, often called the *experimental noise*:

$$\mathbf{y} = \mathbf{f}(\mathbf{g}, \boldsymbol{\epsilon}) \tag{4.3}$$

(or the so-called stochastic extension $Y = f(\mathbf{G}, E)$).

To go further, we must assume (or construct) an appropriate form of \mathbf{f} and propose a probability distribution $p_{\text{noise}}(\boldsymbol{\epsilon})$ for $\boldsymbol{\epsilon}$. These constitute a *noise model*. The construction of an appropriate noise model for observational data is often a challenging endeavor requiring detailed information

on the measurement processes and instrumentation. This may often mean we must construct a new parametric model of \mathbf{g} , knowing very little about the “truth”. But there are approaches, some discussed later in this chapter, of constructing approximations of \mathbf{g} that may improve in accuracy as the new data is acquired. These approximations involve new parameters, not the parameters $\boldsymbol{\theta}$ of the model of the forward problem – sometimes called hyper-parameters. Methods for estimating these parameters and enhancing our approximation of \mathbf{g} are discussed later (e.g., section 4.8).

In the majority of statistics applications, the most common noise model is the additive model,

$$\mathbf{g} + \boldsymbol{\epsilon} = \mathbf{y}, \quad (4.4)$$

showing an additive correction to the truth is what is observed; the noise distribution $p_{\text{noise}}(\boldsymbol{\epsilon})$ is often assumed to be Gaussian. The extra parameters in the definition of $p_{\text{noise}}()$ may be added to the model parameters to produce the total parameters vector $\boldsymbol{\theta}$. If we have n samples of \mathbf{g} , (4.4) is replaced by: $\mathbf{g}_i + \boldsymbol{\epsilon}_i = \mathbf{y}_i$, $i = 1, 2, \dots, n$.

4.3 Model Inadequacy

Suppose we have solved the forward problem and that we use the solution $u(\boldsymbol{\theta}, S)$ to compute the model prediction of \mathbf{g}_i . Let $\mathbf{d}_i(\boldsymbol{\theta}, S)$ denote the model prediction of \mathbf{g}_i for parameters $\boldsymbol{\theta}$ and scenario S . Since the model* is imperfect, there is modeling error $\boldsymbol{\gamma}_i(\boldsymbol{\theta}, S)$, also assumed for convenience to be additive:

$$\mathbf{g}_i - \mathbf{d}_i(\boldsymbol{\theta}, S) = \boldsymbol{\gamma}_i(\boldsymbol{\theta}, S), \quad 1 \leq i \leq n. \quad (4.5)$$

$\boldsymbol{\gamma}_i(\boldsymbol{\theta}, S)$ is referred to as *model inadequacy* or *model form error* or *model bias*.

In analogy with (4.4), relation (4.5) can be thought of as part of a model for model error, to possibly be complemented with the introduction of an assumed probability distribution $p_{\text{model}}(\boldsymbol{\theta}, S)$ for $\boldsymbol{\gamma}_i(\boldsymbol{\theta}, S)$ (see, e.g. [15, 65, 66]).

It is unfortunate, and possibly a surprise to some, that *one can never compute the reality* \mathbf{g}_i ! The best we can know about \mathbf{g}_i is through the imperfect data \mathbf{y}_i and imperfect model prediction $\mathbf{d}_i(\boldsymbol{\theta}, S)$. Eliminating \mathbf{g}_i from (4.4) and (4.5) gives

*We reproduce here, the discussions in OBF [79]

$$\boldsymbol{\varepsilon}_i + \boldsymbol{\gamma}_i(\boldsymbol{\theta}, S) = \mathbf{y}_i - \mathbf{d}_i(\boldsymbol{\theta}, S), \quad 1 \leq i \leq n, \quad \mathbf{y}_i \in \mathcal{Y}_S, \quad \boldsymbol{\theta} \in \Theta. \quad (4.6)$$

That is, the total error, experimental noise and model error, is the difference between the observational data \mathbf{y}_i and model prediction $\mathbf{d}_i(\boldsymbol{\theta}, S)$ for parameters $\boldsymbol{\theta} \in \Theta$ and scenario S . We remark that in much of the literature does not recognize these two distinctly different forms of error.

A common theme in classical statistics, particularly frequentist statistics which we will discuss briefly in sections 4.4 and 4.5, is that the “true” distribution is contained in the probability distributions $p_{\text{noise}}(\boldsymbol{\epsilon})$ and $p_{\text{model}}(\boldsymbol{\theta}, S)$ accessible by the parametric model that defines the prediction $\mathbf{d}(\boldsymbol{\theta}, S)$. This is equivalent to assuming that there exists a (magic) parameter $\boldsymbol{\theta}^*$ such that $\boldsymbol{\gamma}(\boldsymbol{\theta}^*, S) = 0$, $1 \leq i \leq n$; i.e., the exact prediction can be derived by the model. Such models are said to be “correct” or “well-specified”; otherwise, they are misspecified. It is no surprise that most models are indeed *misspecified* – no parameters exist that produce the truth. Many fundamental questions arise about such misspecified model such as how close they can approximate the truth and if, while inadequate and misspecified, are they still sufficiently accurate to be used to make important predictions. This fundamental point brings us to recall once again the famous statement of George E. P. Box [21, 22]: “essentially all models are wrong, but some are useful”. Finding such useful models is the mission of model validation processes.

4.4 Frequentist and Fisherian Approaches

Statistics can be defined as the discipline concerned with the classification study, and interpretation of data in accordance with probability theory. Depending on the probability theory selected, general approaches to statistics in contemporary statistical thinking are: 1) the *frequentist approach* and 2) the *Bayesian approach* (and there are several variants of each of these).

Broadly speaking, the frequentist approach to statistics is generally thought of as one based on the Kolmogorov interpretation of probability in which any given event or experiment is considered as one of an infinite sequence of possible repetitions of the same experiment, each producing independent results. The probability of an event is thus, the frequency with which the event is expected to be observed over infinite samples. To the frequentist, the underlying model parameters remain constant during the repeatable process. It, therefore, makes sense to seek the value of the parameters that best describe observations or samples. For this later reason, a fundamental frequentist approach, referred to as the Fisher approach [41] or the maximum likelihood approach, or the “Fisherian” approach, is put forth for determining the parameter that enables the model to

deliver the best characterization of the true target distribution.

The maximum likelihood approach focuses on determining the true unknown probability distribution (here, the pdf or probability density function) \mathbf{g} that generated the observational data $\mathbf{y} \in \mathcal{Y}$, for some scenario S , through samples $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ of \mathbf{g} . It is often postulated that \mathbf{g} belongs to a family \mathcal{P}_{Θ} of probability densities $\pi(\mathbf{y}, \boldsymbol{\theta})$,

$$\mathbf{g} \in \mathcal{P}_{\Theta} = \{\pi(\mathbf{y}, \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\} \quad (4.7)$$

generally characterized by a conditional probability density $\pi(\mathbf{y}|\boldsymbol{\theta})$ (or $\pi(\mathbf{y}|\boldsymbol{\theta}, S)$) for observational data \mathbf{y} (and scenario S). The likelihood is such that, with the parameter $\boldsymbol{\theta}$ given, the model gives the probability of those outcomes \mathbf{y} .

According to the great statistician R. A. Fisher [41], “The likelihood that any parameter should have any assigned value is proportional to the probability that, if this were true, the totality of all observations should be that observed.” Fisher pointed out that likelihood and probability are of an “entirely different nature” [43]. So if $\boldsymbol{\theta}$ is assigned a value, the probability of \mathbf{y} being observed is a conditional probability, $\pi(\mathbf{y}|\boldsymbol{\theta})$,

Suppose that the samples \mathbf{y}_i are independent and identically distributed (“i.i.d.”); then the joint distribution of the sample is $\prod_{i=1}^n \pi(\mathbf{y}_i|\boldsymbol{\theta})$, viewed as a function of $\boldsymbol{\theta}$ only. This results in what is called the *likelihood function*:

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^n \pi(\mathbf{y}_i|\boldsymbol{\theta}) = \pi(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n|\boldsymbol{\theta}) \quad (4.8)$$

the latter equality holding for independent samples. It is customary to work with the more manageable log-likelihood function,

$$l_n(\boldsymbol{\theta}) = \sum_{i=1}^n \log \pi(\mathbf{y}_i|\boldsymbol{\theta}) = \log \pi(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n|\boldsymbol{\theta}). \quad (4.9)$$

One common approach to defining a likelihood function begins with the relation (4.6) of total error (noise and model inadequacy) and the difference between observations and forward problem prediction $\mathbf{d}_i(\boldsymbol{\theta}, S)$. One introduces arguments that the probability density p_{μ} of the total error p_{noise} plus p_{model} (or more often, just the experimental noise) is known in terms of additional

parameters μ . Then the likelihood is given by,

$$p_\mu(\mathbf{y} - \mathbf{d}(\boldsymbol{\theta}, S)) = \pi(\mathbf{y} | \boldsymbol{\theta}_a) \quad (4.10)$$

where $\boldsymbol{\theta}_a$ is the parameter vector augmented with the parameters μ defining p_μ and, for simplicity, we omit writing the likelihood as also conditional on the scenario S , the parameter μ , being the hyper-parameter.

We remark that the assumption of additive noise embodied in (4.4) tacitly assumes that the noise and the parameters are independent. If this is not the case, we must determine a conditional density of the noise, $p_{\text{noise}}(\boldsymbol{\varepsilon} | \boldsymbol{\theta})$, and then $\pi(\mathbf{y} | \boldsymbol{\theta}) = p_{\text{noise}}(\mathbf{y}_i - \mathbf{d}_i(\boldsymbol{\theta}, S) | \boldsymbol{\theta})$. The “noise model” $p_{\text{noise}}(\boldsymbol{\varepsilon})$ need not be based on the assumption of additive noise, but can be a more complex relation of the type $\mathbf{y} = \mathbf{f}(\mathbf{d}(\boldsymbol{\theta}), \boldsymbol{\varepsilon})$, \mathbf{f} being referred to as the noise model. See, e.g., [64].

As noted earlier, the goal is often to find the parameter $\hat{\boldsymbol{\theta}} \in \Theta$ that maximizes the likelihood of a model. The *maximum likelihood estimator*, or MLE, is thus defined by

$$\hat{\boldsymbol{\theta}}(\mathbf{y}) = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}} [\log \pi(\mathbf{y} | \boldsymbol{\theta})], \quad \mathbf{y} \in \mathcal{Y}. \quad (4.11)$$

We remark, as will be discussed more fully later, that the sequence of MLEs, under specific conditions on the likelihood, converges (in distribution) to a normal distribution with a covariance matrix equal to the inverse of the Fisher information matrix, $\mathbf{I}_1(\boldsymbol{\theta})$,

$$\begin{aligned} \mathbf{I}_1(\boldsymbol{\theta}) &= \int_{\mathcal{Y}} \pi(\mathbf{y} | \boldsymbol{\theta}) \frac{\partial \log \pi(\mathbf{y} | \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \log \pi(\mathbf{y} | \boldsymbol{\theta})^T}{\partial \boldsymbol{\theta}} d\mathbf{y}, \\ &= \text{var}_{\boldsymbol{\theta}} \left[\frac{\partial \log \pi(\mathbf{y} | \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right], \\ &= -\mathbb{E}_{\boldsymbol{\theta}} \left[\frac{\partial^2 \log \pi(\mathbf{y} | \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right], \end{aligned} \quad (4.12)$$

where $\text{var}(\cdot)$ is the variance and $\mathbb{E}_{\boldsymbol{\theta}}$ is the expected value with respect to $\pi(\mathbf{y} | \boldsymbol{\theta})$. We write this as the limit,

$$\frac{1}{\sqrt{n}} (\hat{\boldsymbol{\theta}}^n - \boldsymbol{\theta}) \rightarrow d \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_1^{-1}(\boldsymbol{\theta})) \quad (4.13)$$

where $\hat{\boldsymbol{\theta}}^n$ is the MLE for n i.i.d. observations and we write $\hat{\boldsymbol{\theta}}^n \rightarrow \hat{\boldsymbol{\theta}}$ and $\mathcal{N}(\mu, \Sigma^2)$ denotes the normal distribution with mean μ and standard deviation Σ .

Of special significance is parameter $\boldsymbol{\theta}^*$ weighted by the truth \mathbf{g} as the expected value, defined by,

$$\boldsymbol{\theta}^*(\mathbf{y}) = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}} \mathbb{E}_g [\log \pi(\mathbf{y}|\boldsymbol{\theta})]. \quad (4.14)$$

This parameter arises when we consider information content in various densities \mathbb{E}_g being the expected value with respect to \mathbf{g} .

The *Kullback-Leibler divergence* between two probability distributions with densities p and q is defined by

$$D_{KL}(p \parallel q) = \int p(\mathbf{x}) \log \frac{p(\mathbf{x})}{q(\mathbf{x})} d\mathbf{x}, \quad (4.15)$$

which can be expanded to give,

$$D_{KL}(p \parallel q) = -H(p) + H(p, q), \quad (4.16)$$

where $H(p)$ is the Shannon information entropy,

$$H(p) = - \int p(\mathbf{x}) \log p(\mathbf{x}) d\mathbf{x}, \quad (4.17)$$

and $H(p, q)$ is the cross entropy,

$$H(p, q) = - \int p(\mathbf{x}) \log q(\mathbf{x}) d\mathbf{x}. \quad (4.18)$$

We discuss information theory, information entropy and the maximum entropy principle in the next chapter.

If \mathbf{g} is the “true” distribution, as discussed earlier, the $D_{KL}(\mathbf{g} \parallel \pi(\cdot, \boldsymbol{\theta}))$ is a measure of the information lost by assuming the likelihood $\pi(\cdot|\boldsymbol{\theta})$ is the true distribution,

$$D_{KL}(\mathbf{g} \parallel \pi(\cdot, \boldsymbol{\theta})) = \int \mathbf{g}(\mathbf{y}) \log \frac{\mathbf{g}(\mathbf{y})}{\pi(\mathbf{y}|\boldsymbol{\theta})} d\mathbf{y} \quad (4.19)$$

$$= -H(\mathbf{g}) + H(\mathbf{g}, \pi(\cdot|\boldsymbol{\theta})). \quad (4.20)$$

The parameter $\boldsymbol{\theta}^*$ of (4.14) is defined as minimum of the expected value of log-likelihood weighted

by the truth \mathbf{g} :

$$\begin{aligned}\boldsymbol{\theta}^* &= \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmin}} D_{KL}(\mathbf{g} \| \pi(\cdot, \boldsymbol{\theta})) \\ &= \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmin}} H(\mathbf{g}, \pi(\cdot, \boldsymbol{\theta})) \\ &= \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}} \mathbb{E}_{\mathbf{g}}[\log \pi(\cdot, \boldsymbol{\theta})].\end{aligned}\tag{4.21}$$

If the model is well specified, then $\mathbf{g}(\mathbf{y}) = \pi(\mathbf{y} | \boldsymbol{\theta}_0)$ and $\boldsymbol{\theta}^* = \hat{\boldsymbol{\theta}}$.

In the language introduced in the previous section, if (4.7) holds, the model is well specified, but if there does not exist a parameter $\boldsymbol{\theta}$ such that \mathbf{g} is in \mathcal{P}_{Θ} , the model is misspecified. For a misspecified model, the best one can hope for is to determine the parameter $\boldsymbol{\theta}^*$ in (4.21) that minimizes the D_{KL} “distance” from the model predictions to the truth \mathbf{g} , but the truth is unknown.

One fundamental question is whether or not the maximum likelihood estimates $\hat{\boldsymbol{\theta}}^n$ for n i.i.d data samples do indeed converge in some sense to $\boldsymbol{\theta}^*$ as $n \rightarrow \infty$? It can be shown that under special assumptions on the likelihood functions [67, 68], the $\hat{\boldsymbol{\theta}}^n$ do indeed converge to $\boldsymbol{\theta}^*$ in the sense of distributions and $\frac{1}{\sqrt{n}}(\hat{\boldsymbol{\theta}}^n - \hat{\boldsymbol{\theta}}) \rightarrow \mathcal{N}(0, V(\boldsymbol{\theta}^*))$ where V is called the sandwich matrix [42, 67, 68] depending on the $L_n(\boldsymbol{\theta})$ and, approximately, $V(\boldsymbol{\theta}^*) = V(\hat{\boldsymbol{\theta}})$.

4.5 Principles of Bayesian Model Calibration and Validation

Let us now depart from the frequentist world and venture into the world of Bayesian inference and logical probability. Bayes’ rule (4.1) provides a powerful framework to bring together all of the pillars of scientific knowledge: beginning with a specifications of prior knowledge of probability of parameters (the prior), and laying down the probability of observational data for given parameters provided by a model (the likelihood), we then compute the posterior distribution to upgrade one’s knowledge of the parameters of the system. Once upgraded, use the best information of the parameters in hand to solve the forward problem (e.g. (1.1)), to obtain an informed prediction of a physical event (or quantity) of interest.

The goal is to determine the validity of a given model (or a set of models) as a means for predicting an event, given only certain observational data and incomplete prior information on model param-

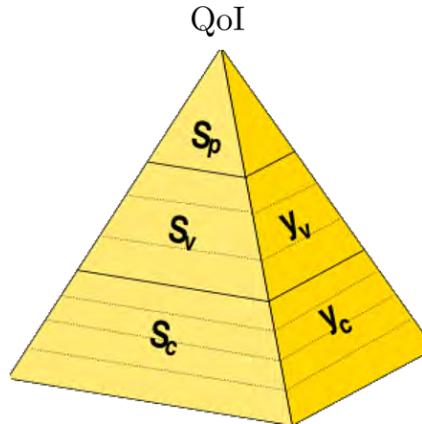


Figure 4.1: The prediction “pyramid”. A symbolic representation of the hierarchy of scenarios and observational data, beginning with unit calibration scenario: S_c delivering data y_c , then validation scenario: S_v with data y_v , and finally the full prediction scenario S_p with the target QoI. The “pyramid” is actually a tetrahedron with the hidden face representing the never seen reality we attempt to predict.

eters. It is informative to consider a hierarchy of scenarios, each with different observational data, beginning with calibration scenario, moving to validation scenario, and then to the full prediction scenario to predict the QoI, as depicted symbolically in the prediction pyramid shown in Figure 4.1.

We propose the following rules and principles for the model validation process.

1. **Quantities of Interest:** It is necessary (and fundamental) to specify precisely the intended purpose of the model and the specific *quantities of interest* (QoIs) that we want to predict. A given mathematical model may be acceptable for one QoI and not another. It is unreasonable (indeed, illogical) to expect a model to be able to predict all features of the behavior of a complex physical system. The target prediction will generally be a probability distribution (or system of probability distributions) describing an event in a complex scenario we call S_p , the prediction scenario, which rarely pertains to observable events (since the S_p may involve predictions of systems that do not yet exist, as in engineering design).
2. **Calibration:** we subscribe to the philosophy of E. T. Jaynes [59], who said :

“ *The essence of ‘honesty’ and ‘objectivity’ demand that we take into account all of the evidence we have and not just some arbitrary chosen subset of it*”

We can apply this idea to model calibration, the tuning of model parameters so that the model predictions agree with experiments on simple component tests. These calibration tests are performed in calibrations scenarios S_c .

3. Validation: Model validation is generally defined as a process designed to estimate the accuracy with which a model can predict QoIs based on the accuracy it can predict actual experimental observations. The validation experiments must involve experimental scenarios, denoted S_v , designed to depict the influence of QoIs on subsystems of the full prediction scenarios S_p . The use of the word “accuracy” here is of fundamental importance. To determine model validity one must choose a metric to measure the distance between model predictions in S_v and experimental observations, manifested in validation data \mathbf{y}_v . In addition, one must specify the tolerance that this measure of error must meet if the model is deemed to be “valid” (or technically “not valid”, since additional information could always falsify a model originally thought to meet tolerances).

In summary, we follow these principles and rules:

- 1) We identify QoIs, Q , in the full model scenario S_p and seek appropriate measures of distributions or densities $\pi(Q)$ and their uncertainty content.
- 2) Before the validation experiments are performed, we calculate initial pdfs defining model parameters in calibration experiments performed on component units of the systems in experimental scenarios S_c .
- 3) We construct validation experiments set up in scenario S_v involving subsystem models of S_p and we identify (subjectively) an appropriate metric d and tolerance γ_{tol} so that the validity of the model for prediction can be declared (or rejected) if $d(\text{prediction}, \text{experimental observation}) \leq \gamma_{\text{tol}}$ (or $> \gamma_{\text{tol}}$).

It is clear that a critical and fundamentally important step in the above validation process is the design of validation experiments that adequately inform the model of the QoIs, the selection of a validation metric, and the (often subjective) step of selecting tolerances.

The process is implemented through the following Bayesian-statistical inverse calculations:

- a) Choose a prior $\pi(\boldsymbol{\theta})$ for calibration scenario S_c ,
- b) Collect calibration data \mathbf{y}_c and update (initialize) the parameters via Bayes’s rule,

$$\pi(\boldsymbol{\theta}|\mathbf{y}_c, S_c) = \frac{\pi(\mathbf{y}_c|\boldsymbol{\theta}, S_c)\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}_c|S_c)}. \quad (4.22)$$

- c) Use the calibration posterior as a prior for the validation experiments involving validation data \mathbf{y}_v :

$$\pi(\boldsymbol{\theta} | \mathbf{y}_v, \mathbf{y}_c, S_v, S_c) = \frac{\pi(\mathbf{y}_v | \boldsymbol{\theta}, \mathbf{y}_c, S_c, S_v) \pi(\boldsymbol{\theta} | \mathbf{y}_c, S_c)}{\pi(\mathbf{y}_v | \mathbf{y}_c, S_v)}. \quad (4.23)$$

- d) With the parameters determined by the validation tests, solve the validation forward problem ($\mathcal{A}(\boldsymbol{\theta}, S_v; u(\boldsymbol{\theta}, S_v)) = 0$) and determine the accuracy with which the model predicts validation observables:

$$d(\mathbf{d}(\boldsymbol{\theta}, S_v), \mathbf{y}_v) \leq \gamma_{\text{tol}} \quad . \quad (4.24)$$

Here $d(\cdot, \cdot)$ is the metric chosen to measure the accuracy of the validation predictions (we remark that other measures of differences between pdfs can be used that are not true “metrics” such as the $D_{KL}(\cdot \| \cdot)$ -divergence, which is not symmetric).

- e) If model satisfies the validity criterion, solve the full forward problem, $\mathcal{A}(\boldsymbol{\theta}, S_p; u(\boldsymbol{\theta}, S_v)) = 0$, on S_p and compute the QoI, $\pi(Q)$. Determine uncertainty inherent in $\pi(Q)$.

How do we quantify uncertainty in the prediction $\pi(Q)$? There are many ways to answer this question. The calculation of the mean and/or the variance of $\pi(Q)$ or higher moments provide quantification of uncertainties in $\pi(Q)$. The uncertainty content given by the information entropy $H(\pi(Q)) = - \int \pi(Q(\mathbf{y})) \log \pi(Q(\mathbf{y})) d\mathbf{y}$ is also a measure. We discuss this in more detail in Chapter 5.

Several major issues emerge:

- How do we compute the priors?
- In general, not all parameters influence the QoI; which parameters are relevant and which have negligible influence on the target predictions?
- What rules do we follow in designing the validation tests?
- How do we choose validation metrics and tolerances?
- How do we solve the forward problem?
- How do we solve the statistical inverse problems (4.22) and (4.23)?
- Most importantly, for given data \mathbf{y} in various scenarios, how do we choose the best models among sets of possible models?

We take up these questions in the chapters to follow.

4.6 Confidence Intervals versus Credible Intervals

Another difference in philosophy and interpretations of frequentist and Bayesian approaches is in the notion of *confidence interval* in frequentist methods and *credible interval* in Bayesian.

In the former, if $\boldsymbol{\theta}$ represents the true unknown parameter, and we are $p\%$ confident it lies between bounds $\boldsymbol{\alpha}(\mathbf{y})$ and $\boldsymbol{\beta}(\mathbf{y})$,

$$\boldsymbol{\alpha}(\mathbf{y}) \leq \boldsymbol{\theta} \leq \boldsymbol{\beta}(\mathbf{y}),$$

then $[\boldsymbol{\alpha}(\mathbf{y}_i), \boldsymbol{\beta}(\mathbf{y}_i)]$, $i = 1, 2, \dots, n$ are called *confidence intervals*, where i indicates the number of experiments performed. This means that the true unknown value of the parameter will surely fall within the confidence interval in $p\%$ of n repetitions of the experiment, if the experiment were repeated n times.

In the corresponding Bayesian approach, we have the posterior distribution $\pi(\boldsymbol{\theta}|\mathbf{y})$, $p\%$ of which may fall between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$,

$$\pi(\boldsymbol{\alpha}(\mathbf{y}) \leq \boldsymbol{\theta} \leq \boldsymbol{\beta}(\mathbf{y})) = \frac{p}{100}.$$

In this case, the interval is called the *credible interval*. The difference in these two concepts often arises in studies of the asymptotic behavior of estimates as $n \rightarrow \infty$.

4.7 Examples: Optimization, Least-Square Misfit, and Bayesian Learning

We describe in this section some rather standard but fundamentally important examples of calculating posteriors of statistical inverse problem (for the parameters $\boldsymbol{\theta}$) involving minimum-least-square formulations, along with some very important observations on asymptotic behavior of Bayesian methods in cases in which the number n of data samples becomes arbitrary large.

4.7.1 Least-Square Misfit

A standard example that permits a comparison of certain aspects of the frequentist and bayesian approaches involves the case in which the likelihood is given by:

$$\pi(\mathbf{y}_s|\boldsymbol{\theta}, s) \propto \exp \left\{ -\frac{1}{2} \|\mathbf{y}_s - \mathbf{d}(\boldsymbol{\theta}, S)\|_{\Sigma}^2 \right\} \quad (4.25)$$

where $\|\cdot\|_{\Sigma}^2 = (\cdot)^T \Sigma^{-1} (\cdot)$, $\Sigma = [\Sigma_{ij}]$ being the $n \times n$ covariance matrix. In some cases, Σ is simply $\sigma^2 \mathbf{I}$ ($\Sigma_{ij} = \sigma^2 \delta_{ij}$), σ^2 being the variance, which can be added to $\boldsymbol{\theta}$ as an additional parameter. If $\mathbf{d}(\boldsymbol{\theta}, S)$ is linear in $\boldsymbol{\theta}$, the likelihood in (4.25) is Gaussian; otherwise, it is non-Gaussian.

Let $\mathbf{d}(\boldsymbol{\theta}, S)$ be linear: $\mathbf{d}(\boldsymbol{\theta}, S) = \mathbf{G}\boldsymbol{\theta}$, \mathbf{G} being a $p \times p$ matrix. Then the MLE is

$$\boldsymbol{\theta}^* = (\mathbf{G}^T \sigma^{-1} \mathbf{G})^{-1} \mathbf{G}^T \sigma^{-1} \mathbf{y}. \quad (4.26)$$

If the prior $\pi(\boldsymbol{\theta})$ is also Gaussian, the posterior becomes

$$\pi(\boldsymbol{\theta} | \mathbf{y}, S) \propto \exp \left\{ -\frac{1}{2} \mathcal{S}(\boldsymbol{\theta}) \right\}, \quad (4.27)$$

where $\mathcal{S}(\boldsymbol{\theta})$ is the misfit function,

$$\mathcal{S}(\boldsymbol{\theta}) = \|\mathbf{y}_i - \mathbf{d}_i(\boldsymbol{\theta}, S)\|_{\Sigma}^2 + \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|_{\Sigma_{\text{prior}}}^2, \quad (4.28)$$

Σ_{prior} being the covariance matrix of the prior density and $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|_{\Sigma_{\text{prior}}}^2 = (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \Sigma_{\text{prior}}^{-1} (\boldsymbol{\theta} - \boldsymbol{\theta}_0)$, $\boldsymbol{\theta}_0$ being a reference parameter vector. The prior is seen to provide a Tikhonov-type regularization of the misfit $\mathcal{S}(\boldsymbol{\theta})$. The vector,

$$\boldsymbol{\theta}^{**} = \arg \max_{\boldsymbol{\theta}} \pi(\boldsymbol{\theta} | \mathbf{y}, S) \quad (4.29)$$

is called the *Maximum A Posterior estimate* (MAP) for scenario S . We note that for uniform priors, MAP = MLE.

4.7.2 Bayesian Learning

One of the most interesting aspects of the Bayesian approach is so-called Bayesian learning. Suppose a single observation \mathbf{y}_1 is recorded. Beginning with a prior $\pi(\boldsymbol{\theta})$, we update $\boldsymbol{\theta}$ throughout the calculation of the posterior:

$$\pi(\boldsymbol{\theta} | \mathbf{y}_1) \propto \pi(\mathbf{y}_1 | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \quad (4.30)$$

When new additional data \mathbf{y}_2 is attained, the posterior becomes the prior for another posterior update:

$$\pi(\boldsymbol{\theta} | \mathbf{y}_1, \mathbf{y}_2) \propto \pi(\mathbf{y}_2 | \boldsymbol{\theta}, \mathbf{y}_1) \pi(\boldsymbol{\theta} | \mathbf{y}_1). \quad (4.31)$$

Continuing in this way, we “learn” more about $\boldsymbol{\theta}$ with each new observation,

$$\begin{aligned}\pi(\boldsymbol{\theta}|\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n) &\propto \pi(\mathbf{y}_2|\boldsymbol{\theta}, \mathbf{y}_1)\pi(\boldsymbol{\theta}|\mathbf{y}_1) \times \pi(\boldsymbol{\theta}|\mathbf{y}_{n-1}) \\ &= \left(\prod_{i=1}^n \pi(\mathbf{y}_i|\boldsymbol{\theta}) \right) \pi(\boldsymbol{\theta})\end{aligned}\quad (4.32)$$

Given the sequence $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$, we can use the posterior in (4.32) as a prior, but if we consider the samples to be i.i.d., then the same prior must be retained.

A fundamental question arises is what is the limiting posterior distribution attained as $n \rightarrow \infty$? The answer to this question is captured by Bernstein–von Mises theorem for the case in which the model is correctly specified and the likelihood is sufficiently smooth [19, 118]). This is also called the *Bayesian Central Limit Theorem*. Suppose that $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ are i.i.d. samples from a probability density $g(\mathbf{y}|\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$. Suppose $\pi(\boldsymbol{\theta})$ is a prior density and denote by $\pi(\mathbf{y}^n|\boldsymbol{\theta})$ the likelihood, $\pi(\mathbf{y}^n|\boldsymbol{\theta}) = \prod_{i=1}^n \pi(y_i|\boldsymbol{\theta})$. Let both $\pi(\boldsymbol{\theta})$ and $\pi(\mathbf{y}^n|\boldsymbol{\theta})$ be twice differentiable near $\hat{\boldsymbol{\theta}}$, the MLE, which is assumed to exist. Then, for large n and commonly satisfied assumptions, the posterior density

$$\pi(\boldsymbol{\theta}, \mathbf{y}^n) = \frac{\pi(\mathbf{y}^n|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathbf{y}^n)}, \quad (4.33)$$

converges, in distribution, as $n \rightarrow \infty$, to the normal distribution

$$\mathcal{N}(\hat{\boldsymbol{\theta}}, \hat{\mathbf{I}}(\hat{\boldsymbol{\theta}})^{-1}), \quad (4.34)$$

where $\hat{\mathbf{I}}(\hat{\boldsymbol{\theta}})$ is the Fisher information matrix of (4.12).

4.8 Gaussian Processes: Tools to Build Models of Noise and Model Inadequacies

At some steps in the processes of formulating the Bayesian inference problem we encountered the need to contemplate new parametric models, outside models relevant to the principal forward problem, to approximate such entities as experimental noise or model inadequacy. The idea of Gaussian processes provides a popular and often powerful approach for constructing such models. In this section we present a brief introduction to this subject.

Stochastic or Random Process. A *stochastic* or *random process* is a collection of random variables $X = \{X(t); t \in I\}$ defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, I being an index set and the values of X lying in some “state space” Y . In general, one selects a finite number n of *trial points* $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in I^n$ and the process generates a probability distribution $\{X(\mathbf{x}_1), X(\mathbf{x}_2), \dots, X(\mathbf{x}_n)\}$ for which we compute probabilities $\mathbb{P}[X(\mathbf{x}_1), X(\mathbf{x}_2), \dots, X(\mathbf{x}_n)]$. Thus, a stochastic process can be characterized by a function $f(X)$ of a random variable X and, for any set of sample or trial points $\{\mathbf{x}_i\}_{i=1}^n$, a corresponding set of probabilities

$$\{\mathbb{P}(X(\mathbf{x}_1)), \mathbb{P}(X(\mathbf{x}_2)), \dots, \mathbb{P}(X(\mathbf{x}_n))\}.$$

Gaussian Process. A *Gaussian process* (or GP) is a random process in which any finite number of random variables have a joint probability distribution that is Gaussian.

So here we think of a real Gaussian process as being defined by random variables represented by the value of a function $\mathbf{f}(\mathbf{x})$ at the location of a point \mathbf{x} (e.g. $\mathbf{x} \in \mathbb{R}^m$, or $\mathbf{x} = t$, a particular time value). Since every Gaussian distribution is completely specified by its mean and covariance, a Gaussian process $\mathbf{f}(\mathbf{x})$ is defined by its *mean function* $\mathbf{m}(\mathbf{x})$ and its *covariance function* $\mathbf{k}(\mathbf{x}, \mathbf{x}')$:

$$\begin{cases} \mathbf{m}(\mathbf{x}) = \mathbb{E}[\mathbf{f}(\mathbf{x})] \\ \mathbf{k}(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(\mathbf{f}(\mathbf{x}) - \mathbf{m}(\mathbf{x}))(\mathbf{f}(\mathbf{x}') - \mathbf{m}(\mathbf{x}'))^T]. \end{cases} \quad (4.35)$$

It is customary to use the following notation for such a process (Rasmussen and Williams, 2006 [91]).

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}(\mathbf{m}(\mathbf{x}), \mathbf{k}(\mathbf{x}, \mathbf{x}')). \quad (4.36)$$

The great utility of Gaussian processes resides in their so called *marginalization property*: since all distributions at any sample point are, by definition, Gaussian, any enlargement of sample points taking the process outside the initial sample wetset must also be Gaussian, and any such larger sample will not change the distribution of the smaller sample set. This subtle property, also called the *consistency property*, becomes clear in the basic Gaussian extension described below.

The Consistency or Marginalization Property of GPs. The fundamental consistency or marginalization property of GPs is embodied in the following theorem.

The GPs Theorem. Let \mathbf{y}_1 and \mathbf{y}_2 be two random variables with joint Gaussian probability distribution

$$(\mathbf{y}_1, \mathbf{y}_2) \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix}\right). \quad (4.37)$$

Then

$$\mathbf{y}_1 \sim \mathcal{N}(\mathbf{a}, \mathbf{A}) \quad (4.38)$$

and

$$(\mathbf{y}_1 | \mathbf{y}_2) \sim \mathcal{N}(\mathbf{a} - \mathbf{B}\mathbf{C}^{-1}(\mathbf{y}_2 - \mathbf{b}), \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T). \quad (4.39)$$

proof: It is understood that \mathbf{y}_1 is associated with a density $\pi(\mathbf{y}_1)$ equal to the marginalization $\int \pi(\mathbf{y}_1, \mathbf{y}_2) d\mathbf{y}_2$. The conditional probability (4.39) follows the evaluation of,

$$\pi(\mathbf{y}_1 | \mathbf{y}_2) = \frac{\pi(\mathbf{y}_1, \mathbf{y}_2)}{\pi(\mathbf{y}_2)} \quad (4.40)$$

being Gaussian, making use of the Gaussian distribution equation: if $\mathbf{f} = \{f_1, f_2, \dots, f_n\}^T$, and $\mathbf{f} \sim \mathcal{N}(\mathbf{m}, \Sigma)$, then,

$$\pi(\mathbf{f}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{f} - \mathbf{m})^T \Sigma^{-1} (\mathbf{f} - \mathbf{m}) \right\}. \quad (4.41)$$

□

The joint density is,

$$\pi(\mathbf{y}_1, \mathbf{y}_2) = \frac{1}{\sqrt{2\pi|\Sigma|}} \exp \left(-\frac{1}{2} \Gamma(y_1, y_2) \right)$$

where,

$$\begin{aligned} \Gamma(y_1, y_2) &= \{(y_1 - a)^T (y_2 - b)^T\} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix}^{-1} \begin{bmatrix} y_1 - a \\ y_2 - b \end{bmatrix} \\ &= (y_1 - a)^T \mathbf{A}' (y_1 - a) + 2(y_1 - a)^T \mathbf{B}' (y_2 - b) + (y_2 - b)^T \mathbf{C}' (y_2 - b) \end{aligned}$$

where

$$\begin{aligned} \mathbf{A}' &= (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T)^{-1} \\ \mathbf{B}' &= -\mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B})^{-1} = \tilde{\mathbf{B}}^T \\ \mathbf{C}' &= (\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B})^{-1} \end{aligned}$$

Thus, with some algebra, one has,

$$\Gamma(y_1, y_2) = \mathcal{N}(b, \mathbf{C})\mathcal{N}(\mu, \Sigma')$$

where,

$$\begin{aligned}\mu &= a - \mathbf{B}\mathbf{C}^{-1}(y_2 - b) \\ \Sigma' &= \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T.\end{aligned}$$

Then,

$$\begin{aligned}\pi(y_1|y_2) &= \pi(y_1, y_2)/\pi(y_2) \\ &= \mathcal{N}(b, \mathbf{C})\mathcal{N}(\mu, \Sigma')/\mathcal{N}(b, \mathbf{C}) \\ &= \mathcal{N}(\mu, \Sigma') \\ &= \mathcal{N}(a - \mathbf{B}\mathbf{C}^{-1}(y_2 - b), \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T),\end{aligned}$$

which was asserted. \square

The Gaussian Process Response Surface. We use the GP framework to construct *response surfaces*. The idea is this: Let $\mathbf{Y}(\mathbf{x})$ denote a random field for which realizations

$$\{\mathbf{y}(x_1), \mathbf{y}(x_2), \dots, \mathbf{y}(x_m)\}$$

are given, and $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ being *training points* ($\mathbf{x}_i \in \mathcal{R}^n$) (or ‘input points’ or ‘control points’). Then the *training set* is

$$\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}(\mathbf{x}_i)), i = 1, 2, \dots, m\}. \quad (4.42)$$

We represent \mathbf{y} by a GP:

$$\mathbf{y} \sim \mathcal{GP}(\psi^T(\mathbf{x})\boldsymbol{\beta}, \Sigma), \quad (4.43)$$

where

$$\psi^T(\mathbf{x})\boldsymbol{\beta} = \sum_{i=1}^q \psi_i(\mathbf{x})\beta_i = \mathbf{m}(\mathbf{x}) = \text{the mean function}, \quad (4.44)$$

(also called the “trend”), $\psi_i(\mathbf{x})$ are appropriate basis functions, $\boldsymbol{\beta}$ a vector of coefficients, and Σ is

the *covariance matrix*, which is any symmetric, positive semi-definite matrix. Commonly, we take

$$\begin{aligned}\Sigma_{ij} &= k(x_i, x_j) \\ &= \lambda_0 \exp \left\{ -\frac{1}{2} \sum_{k=1}^d \frac{(x_i^k - x_j^k)^2}{l_k} \right\} + \lambda_1 + \lambda_2 \delta_{ij}\end{aligned}\quad (4.45)$$

Here λ_0 , l_h , λ_1 and λ_2 are parameters, with

λ_0 = a measure of precision of variance (an amplitude of a covariance envelope),

l_k = characterization of flatness of covariance,

λ_1 = a bias “latent value”, and,

λ_2 = a regularization.

By adjusting values of the parameters, we can conceivably manipulate the shape and flatness of the covariance between sample points in \mathcal{D} .

Given the GP (4.43), the *predicted output* $\mathbf{y}^* = \mathbf{y}(\mathbf{x}^*)$ at a point outside the training set \mathcal{D} , $(\mathbf{x}^*, \mathbf{y}^*) \notin \mathcal{D}$, is, according to (4.39),

$$(\mathbf{y}^* | \mathbf{y}) = \mathcal{GP}(\mathbb{E}((\mathbf{y}^* | \mathbf{y})), \boldsymbol{\Sigma}^*), \quad (4.46)$$

where,

$$\begin{cases} \mathbb{E}(\mathbf{y}^* | \mathbf{y}) = \boldsymbol{\psi}^T(\mathbf{x})\boldsymbol{\beta} - \mathbf{k}(\mathbf{x}^*, \mathbf{x})\mathbf{K}^{-1}(\mathbf{y} - \mathbf{m}) \\ \boldsymbol{\Sigma}^* = \mathbf{k}(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{k}(\mathbf{x}^*, \mathbf{x})\mathbf{K}^{-1}\mathbf{k}(\mathbf{x}, \mathbf{x}^*) \end{cases} \quad (4.47)$$

and,

$$\mathbf{K} = \mathbf{k}(\mathbf{x}, \mathbf{x}). \quad (4.48)$$

Clearly, $\mathbf{m}(\mathbf{x}) = \boldsymbol{\psi}^T(\mathbf{x})\boldsymbol{\beta}$.

The new parameters ω defining the GP are called hyperparameters, e.g.

$$\boldsymbol{\omega} = (\boldsymbol{\beta}, \lambda_0, \{l_k\}_{k=1}^d, \lambda_1, \lambda_2). \quad (4.49)$$

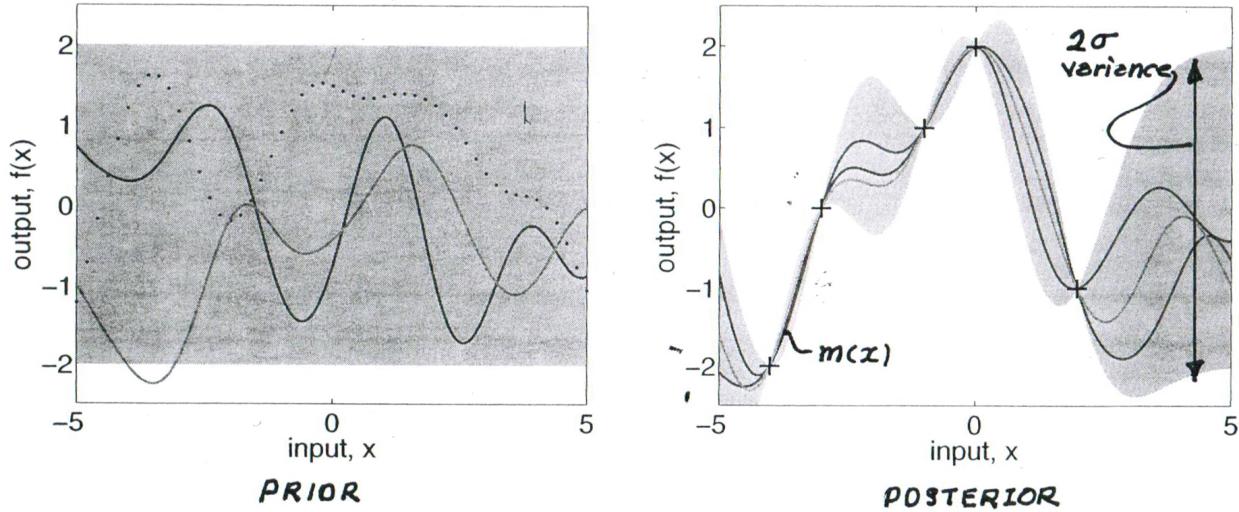


Figure 4.2: Illustration of a 1-dimensional data set fit by a GP with noise-free data points, $y(x)$ anchoring the curve at input points x and the variance defined through chains of hyperparameters.

Figure 4.2 shows a conceptualization of a GP representation over a one-dimensional input set. On the left, we see a collection of prior pdfs. The use of five input or training points *anchor* the output function $\mathbf{y}(x)$ at points $\mathbf{y}(\mathbf{x}_i)$ on the mean curve, $\mathbf{m}(\mathbf{x})$, as shown. The shape of the covariance between these points is fashioned by adjusting the covariance parameters. The factorization rule can be used to update function values $\pi(\mathbf{y}(\mathbf{x}_1), \mathbf{y}(\mathbf{x}_2), \dots, \mathbf{y}(\mathbf{x}_m)) = \prod_i^m \pi(\mathbf{y}(\mathbf{x}_1))\pi(\mathbf{y}(\mathbf{x}_2)), \dots, \pi(\mathbf{y}(\mathbf{x}_m))$.

Model Inadequacy and Validation using Gaussian Processes.

We close this section with an example of the use of GPs to represent model inadequacy and to also account of experimental noise, following Bayarri, Berger, Paulo, Sachs, and Cafeo (2007) [15].

Let $y^R(\mathbf{x})$ denote the truth or reality of a physical entity at point $\mathbf{x} \in \mathbb{R}^n$ (denoted \mathbf{g} earlier), and let $\mathbf{y}^M(\mathbf{x}, \boldsymbol{\theta})$ denote the model prediction of this field at \mathbf{x} (denoted $d(\boldsymbol{\theta})$ earlier). Then the model inadequacy γ ($= \gamma_i(\boldsymbol{\theta}, S)$ of 4.5) is

$$\gamma(\boldsymbol{\theta}, \mathbf{x}) = \mathbf{y}^R(\mathbf{x}) - \mathbf{y}^M(\mathbf{x}, \boldsymbol{\theta}). \quad (4.50)$$

Here $\boldsymbol{\theta}$ is the parameter vector of the forward prediction problem. Let $\mathbf{y}^F(\mathbf{x}_i)$ denote the “field data” at \mathbf{x}_i in a set of test points $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ (denoted \mathbf{y}_i earlier). Thus, for a additive model,

$$\mathbf{y}^F(\mathbf{x}_i) = \mathbf{y}^R(\mathbf{x}_i) + \epsilon_i^F \quad (4.51)$$

where ϵ_i^F is experimental noise. To complete the noise model, we assume ϵ_i^F is i.i.d. and

$$\epsilon_i^F \sim \mathcal{N}(0, 1/\lambda^F). \quad (4.52)$$

Thus,

$$\mathbf{y}^F(\mathbf{x}_i) - \mathbf{y}^M(\mathbf{x}_i, \boldsymbol{\theta}) = \gamma(\boldsymbol{\theta}, \mathbf{x}_i) + \epsilon_i^F \quad (4.53)$$

We next introduce priors for the key pdfs:

$$\begin{cases} \pi(\boldsymbol{\theta}), \pi(\lambda^F), \text{ and} \\ \pi(\gamma) \sim \mathcal{GP}(\psi^T(\mathbf{x})\boldsymbol{\beta}, \Sigma_\gamma) \end{cases} \quad (4.54)$$

with hyperparameters $\boldsymbol{\omega}_j = (\beta_\gamma, \lambda_0^\gamma, \{l_k^\gamma\}_{k=1}^{d_\gamma}, \lambda_1^\gamma, \lambda_2^\gamma)$. The posterior update is,

$$\pi(\boldsymbol{\theta}, \lambda^F, \gamma | \mathbf{y}^F) \propto \mathbf{f}(\mathbf{y}^F | \boldsymbol{\theta}, \lambda^F, \gamma) \pi(\theta, \lambda^F, \gamma), \quad (4.55)$$

with \mathbf{f} being a multivariate normal distribution. The model output is,

$$\mathbf{y}^M \sim \mathcal{GP}(\psi^T(\mathbf{x})\boldsymbol{\beta}, \Sigma_\gamma). \quad (4.56)$$

Computing the Hyperparameters.

One popular method to estimate the hyperparameters $\boldsymbol{\omega}$ is to use the MLE:

$$\boldsymbol{\omega}_{MLE} = \arg \max_{\boldsymbol{\omega}} \pi(\mathbf{y} | \mathbf{x}, \boldsymbol{\omega}), \quad (4.57)$$

where,

$$\pi(\mathbf{y} | \mathbf{x}, \boldsymbol{\omega}) \propto \prod_k \exp \left\{ -\frac{1}{2} \left(\mathbf{y}^F - \mathbf{y}^m(\boldsymbol{\omega}, \mathbf{x}) \right)^2 / \sigma^2 \right\}, \quad (4.58)$$

with $\mathbf{y}^F = \mathbf{y}^M + \boldsymbol{\epsilon}$, $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$. A common approach is to use the marginal likelihood,

$$\begin{aligned} \frac{\partial}{\partial \omega_j} \log \pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\omega}) &= 0 \\ &= \frac{1}{2} \mathbf{y}^T \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \omega_j} \boldsymbol{\Sigma}^{-1} \mathbf{y} - \frac{1}{2} \text{trace} (\boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{K}}{\partial \omega_j}). \end{aligned} \quad (4.59)$$

The full Bayesian uses the posterior,

$$\pi(\boldsymbol{\theta}, \boldsymbol{\omega} | \mathbf{y}^F) \propto \pi(\mathbf{y}^F | \boldsymbol{\theta}, \boldsymbol{\omega}) \pi(\boldsymbol{\theta}, \boldsymbol{\omega}). \quad (4.60)$$

In implementing this GP updates, Bayarri et al. (2007) [15] describe an MCMC process starting with the posterior,

$$\begin{aligned} \pi(\boldsymbol{\theta}, \lambda^F, \boldsymbol{\gamma}(\boldsymbol{\omega}_\gamma) | \mathbf{y}^F) &\propto f(\mathbf{y}^F | \boldsymbol{\theta}, \lambda^F, \boldsymbol{\gamma}(\boldsymbol{\omega})) \pi(\boldsymbol{\theta}, \lambda^F, \boldsymbol{\gamma}) \\ &\approx \mathbf{y}^F - \mathbf{y}^M(\mathbf{x}_i, \boldsymbol{\theta}) - \boldsymbol{\epsilon}_i^F, \end{aligned} \quad (4.61)$$

where the outputs $\{\boldsymbol{\theta}^{(i)}, \lambda^{F(i)}, \mathbf{y}^M(\mathbf{x}, \boldsymbol{\theta}^{(i)}), \boldsymbol{\gamma}^{(i)}(\mathbf{x})\}$ are sampled and all quantities of interest can be estimated from these samples.

The prediction of reality $\mathbf{y}^F(\mathbf{x})$ at a new set of training points \mathcal{D}_{NEW}^F are drawn from the predictive posterior distribution, $\pi(\mathbf{y}_{NEW}^R | \mathbf{y}^F, \mathbf{y}^M)$ where $\mathbf{y}^R = \mathbf{y}^M + \boldsymbol{\gamma}$, determined from,

$$\mathbf{y}_{NEW}^R \sim \frac{1}{N} \sum_{k=1}^N (\mathbf{y}_{NEW}^{M(k)} + \boldsymbol{\gamma}_{NEW}^{(k)}), \quad (4.62)$$

where $\mathbf{y}_{NEW}^{M(k)}$ and $\boldsymbol{\gamma}_{NEW}^{(k)}$ are drawn from the joint posterior. Then, \mathbf{y}_{NEW}^R is the “new” prediction of the reality. Variance and tolerance bounds (credible intervals) can be computed approximately from these calculations.

Chapter 5

Information Theory and the Principle of Maximum Entropy

5.1 Introduction

Apart from Bayesian approaches, but very much complementary to them, are the methods and concepts underlying information theory and the concept of information entropy. Introduced in the landmark papers of Claude Shannon in 1948 [104], information theory provides a rational approach to quantifying and understanding the information content in a probability distribution (or a digital signal). When enriched by the observations and works of E.T. Jaynes [60], well-defined measures of uncertainty in information or in probability emerge that are fundamentally important in many applications, including the rational construction of priors and in Bayesian approaches to inverse problems, such as image reconstruction and pattern recognition. We partake on a brief journey into the subject in the present chapter. More detailed accounts of this subject can be found in the original paper of Shannon [104] and the books of Jaynes [60], Gregory [45], Klir [69], Cover and Thomas [30], and the notes of Carter [27].

5.2 An Elementary Look at Information

Let us first address the problem of defining a meaningful measure of the information content in a probability density p . As a measure of information, some common-sense conventions and properties might be the following:

We consider a random experiment with n sampled outputs: with probabilities p_1, p_2, \dots, p_n , respectively and with each $p_i > 0$. We wish to construct a measure of the information obtained by observing a particular p_i . Such an information measure should also be related to the uncertainty in p_i which we also want to quantify. We propose that the information I has the following common-sense properties:

$$I : (0, 1] \rightarrow [0, \infty) \quad (I(p) \geq 0)$$

$$I(p_i = 1) = 0 \text{ (there is no information in a certain probability)}$$

$I(p)$ should behave properly for joint, independent outcomes. In particular, if the outcomes of n experiments is denoted p_i ($\in \mathbb{N}^n$) and the outcomes of another m independent experiments is q_j ($\in \mathbb{N}^m$) then the joint outcome should be $p_i \cdot q_j \forall i \in \mathbb{N}^n$ and $j \in \mathbb{N}^m$. The “anticipatory” information (c.f., Klir [69]; Carter [27]) of the joint outcome will such be the sum of the individual outcomes:

$$I(p_i \cdot q_j) = I(p_i) + I(q_j) \quad (5.1)$$

Equation (5.1) is referred to by some authors (e.g. Klir [69]) as a *Cauchy equation*, the solution of which is a class of functions I defined for all $\alpha \in (0, 1]$ by,

$$I(a) = c \log_b a$$

where c is an arbitrary constant, b is a distinct constant $\neq 1$. Since $I(\cdot)$ must be decreasing on $(0, 1]$ and $\log(\cdot)$ is increasing, c must be negative. By scaling so that $I(\frac{1}{2}) = 1$ and choosing $c = -1$, we get

$$I(a) = -\log_b a, \quad (5.2)$$

or, for probability p_i of outcome i ,

$$I(p_i) = -\log_b p_i = \log_b \left(\frac{1}{p_i} \right) \quad (5.3)$$

We remark that the property (5.2) and hence (5.3), follows from (5.1), or more generally from properties of continuous real-valued functions $g : (0, \infty) \rightarrow \mathbb{R}$ with the property that $g(x_1 x_2) = g(x_1) + g(x_2)$. Indeed, from this one shows that $g(x^n) = n g(x)$ and, by continuity, $g(x^\alpha) = \alpha g(x)$, $\alpha > 0$. Setting $x = a$ reveals that $g(a^\alpha) = \alpha g(a)$ or $g(a)^{-1} g(a^\alpha) = \alpha$. This shows that $g(a)^{-1} g(x)$ is a left inverse of a^x . A straightforward argument shows that $g(0)$ is also injective and, therefore,

that $g(a)^{-1}g(x) =$ the inverse of $a^x = \log_a x$, so that

$$g(x) = g(a) \log_a x .$$

Setting $I(\cdot) = g(\cdot)$, a , $g(a) = 1$, and $a = b$ yields (5.2).

Example. If we take $b = 2$, the units of information are called *bits*. Carter [27] notes the example of flipping a fair coin to give heads H or tails T with probability $p = 1/2$. The information is $I(p) = -\log_2(1/2) = 1$ bit. For n flips $I(p^n) = -\log((1/2)^n) = n \log(2) = n$ bits. \square

5.3 A First Look at Information Entropy

The above properties will prove useful in determining the connection between information content in a pdf and its “entropy”. For a discrete probability distribution $p = \{p_1, p_2, \dots, p_n\}$, we will define the *information entropy* $H(p)$ as the real number,

$$H(p) = - \sum_{i=1}^n p_i \log p_i \tag{5.4}$$

where the logarithm can be selected for any base b . We generally take $b = 2$. Exactly why (5.4) is a meaningful definition of information entropy will be taken up in section 5.5 where we present Shannon’s theorem and arguments as to why this property is a fundamental measure of uncertainty content in a probability distribution.

Consider a situation in which we are given a set of n quantities $\{a_1, a_2, \dots, a_n\}$ with corresponding probabilities $\{p_1, p_2, \dots, p_n\}$, each a_i sampled independently of the others. It is natural to ask: what is the average amount of information we get from each a_i ?

From entity a_i , we get information,

$$I(p_i) = -\log(p_i) = \log(1/p_i) .$$

In a run of N observations, we see around Np_i occurrences of a_i . Thus, in the N independent

observations, we get total information of,

$$I = \sum_{i=1}^n N p_i \log\left(\frac{1}{p_i}\right).$$

Then, the expectation (mean) value of the information is I/N , or,

$$\mathbb{E}[I(p)] = \langle I(p) \rangle = - \sum_{i=1}^n p_i \log(p_i) = H(p),$$

We have shown that the *entropy of a probability distribution is the expectation value of the information of the distribution:*

$$\mathbb{E}[I(p)] = \langle I(p) \rangle = H(p) .$$

(5.5)

5.4 The Maximum Entropy Principle

Gibb's Inequality. A simple inequality attributed to J. Williard Gibbs has to do with the rate of increase of $\ln(x)$, $x \in \mathbb{R}$, compared to the function $y(x) = x - 1$. Since $(\ln(x))' = 1/x$, the log function is tangent to $y(x)$ at $x = 1$. Since $\ln(x)$ is concave for $x > 0$, we have

$$\ln(x) \leq x - 1 \quad (5.6)$$

with equality holding only when $x = 1$. This is the *Gibb's inequality*. \square

A classical application of Gibb's inequality arises when we consider the relative entropy of two probability distributions say $p = \{p_1, p_2, \dots, p_n\}$ and $q = \{q_1, q_2, \dots, q_n\}$; $p_i, q_i \geq 0$, and $\sum_i p_i = \sum_i q_i = 1$. Their Kullback-Leibler divergence (or *relative entropy*) is,

$$D_{\text{KL}}(p||q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}$$

$$\begin{aligned}
-D_{\text{KL}}(p||q) &= - \sum_{i=1}^n p_i \log \frac{p_i}{q_i} \\
&= \leq \sum_{i=1}^n p_i \left(1 - \frac{q_i}{p_i} \right) \\
&= \sum_{i=1}^n (q_i - p_i) \\
&= 0,
\end{aligned} \tag{5.7}$$

from Gibb's inequality. Therefore,

$$D_{\text{KL}}(p||q) = -H(p) + H(p, q) \geq 0 . \tag{5.8}$$

where $H(p, q)$ is the cross entropy, $H(p, q) = -\sum_{i=1}^n p_i \log q_i$. Next, given a probability distribution $\{p_1, p_2, \dots, p_n\}$, we compute the difference,

$$\begin{aligned}
H(p) - \log(n) &= - \sum_{i=1}^n p_i \log p_i - \log(n) \\
&= \sum_{i=1}^n p_i \log\left(\frac{1}{p_i}\right) - \log(n) \sum_{i=1}^n p_i \\
&= \sum_{i=1}^n p_i \left(\log\left(\frac{1}{p_i}\right) + \log\left(\frac{1}{n}\right) \right) \\
&= \sum_{i=1}^n p_i \log\left(\frac{1/n}{p_i}\right)
\end{aligned} \tag{5.9}$$

But, from the Gibbs inequality (5.6), this latter sum is less than or equal to zero, the equality holding only when $p_i = 1/n$. Thus,

$$0 \leq H(p) \leq \log(n) \tag{5.10}$$

The entropy $H(p)$ is zero only when one of the p_i 's is one and all others are zero (then the average information $\langle I(p) \rangle = 0$) and $H(p) = \log(n)$, its maximum, only when all of the probabilities have the same value, $1/n$; i.e. all events are equally likely.

We are thus led to a fundamental properties of the entropy of a probability; *The Maximum Entropy Principle*:

Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ be a set of points in \mathbb{R}^m and $p_i = p(\mathbf{x}_i)$ denote the probability of an event at \mathbf{x}_i . Let $\mathcal{P}_n = \{(p_1, p_2, \dots, p_n) | p_i \in [0, 1], \sum_i p_i = 1\}$ be the set of all probability distributions on X with n components. Then of all $p \in \mathcal{P}_n$, the one that maximizes the entropy defines the information content $\mathbb{E}[I(p)]$ and quantifies the uncertainty in the information.

5.5 Mutual Information

In Shannon's theory of information, the following situation is considered: a "source" supplies a stream of data $\{a_1, a_2, \dots, a_n\} = A$ which are enclosed and sent through a channel, possibly then disturbed by noise, and then received by a receiver which is supposed to decode and derive information from the sequence. The channel is (ignoring noise) capable of carrying data $\{c_1, c_2, \dots, c_r\}$, so encoding is accomplished by a function,

$$f : A \rightarrow C^* , \quad C^* = \text{set of finite strings of data from } C.$$

We denote,

$$l_i = |f(a_i)| = \text{length of strings in } C^* \text{ encoding data symbol } a_i$$

Given a probability measure of the source, let $P(a_i)$ denote the prior estimate that the unit a_i will be sent next through the channel. Its information is then $I(P(a_i)) = -\log(P(a_i))$. But suppose we observe a signal b_j , so that our estimate of the probability is revised to $P(a_i|b_j)$, with new information $I(P(a_i|b_j)) = -\log P(a_i|b_j)$. The *mutual information* is defined as the change in information (due to new data) and is denoted $I(a_i; b_j)$:

$$I(a_i; b_j) = I(P(a_i)) - I(P(a_i|b_j)) = \log\left(\frac{1}{P(a_i)}\right) - \log\left(\frac{1}{P(a_i|b_j)}\right) = \log\left(\frac{P(a_i|b_j)}{P(a_i)}\right) \quad (5.11)$$

The following properties follow easily:

$$\left. \begin{aligned} I(a_i; b_j) &= I(b_j; a_i) \\ I(a_i; b_j) &= \log(P(a_i|b_j)) + I(a_i) \\ I(a_i; b_j) &\leq I(a_i) \end{aligned} \right\} \quad (5.12)$$

If a_i and b_j are independent, $P(a_i|b_j) = P(a_i)P(b_j)$ and $I(a_i; b_j) = 0$.

Generalization. Now we allow noise in the channel:

$$A = \{a_1, a_2, \dots, a_n\} = \text{input stream}$$

$$B = \{b_1, b_2, \dots, b_m\} = \text{output stream}$$

$$\text{Channel Characterization } P(a_i|b_j)$$

The average mutual information over all symbols is,

$$\left. \begin{aligned} I(A; b_j) &= \sum_i P(a_i|b_j)I(a_i; b_j) \\ I(a_i; B) &= \sum_j P(a_i|b_j)I(a_i; b_j) \end{aligned} \right\} \quad (5.13)$$

and,

$$\begin{aligned} I(A, B) &= \sum_i P(a_i)I(a_i; B) \\ &= I(B, A) \\ &= \sum_i \sum_j P(a_i|b_j) \log\left(\frac{P(a_i|b_j)}{P(a_i)P(b_j)}\right) \end{aligned} \quad (5.14)$$

Denoting the total entropies and conditional entropies as,

$$H(A) = \sum_{i=1}^n P(a_i) \log(1/P(a_i)) \quad (5.15)$$

$$H(B) = \sum_{k=1}^m P(b_k) \log(1/P(b_k)) \quad (5.16)$$

$$H(A|B) = \sum_{i=1}^n \sum_{j=1}^m P(a_i|b_j) \log(1/P(a_i|b_j)) \quad (5.17)$$

$$H(A, B) = \sum_{i=1}^n \sum_{j=1}^m P(a_i, b_j) \log(1/P(a_i, b_j)) \quad (5.18)$$

Then,

$$H(A, B) = H(A) + H(B|A) \quad (5.19)$$

$$H(A, B) = H(B) + H(A|B) \quad (5.20)$$

Moreover,

$$\begin{aligned}
 I(A; B) &= H(A) + H(B) - H(A, B) \\
 &= H(A) - H(A|B) \\
 &= H(B) - H(B|A)
 \end{aligned} \tag{5.21}$$

5.6 Shannon's Theorem

It is often asserted that the information entropy $H(p)$ is the only meaningful measure of uncertainty and information in probability theory. More rigorous developments that support this assertion can be found in the original paper of Shannon [104], in the treatise of Jaynes [60], the descriptive exposition of Klir [69], and other works. Without providing full details, we summarize the basic axioms underlying the theorem following [69].

Axiom 1. When a zero component is added to a probability distribution, the quantity of uncertainty should not change: $H(p_1, p_2, \dots, p_n) = H(p_1, p_2, \dots, p_n, 0)$.

Axiom 2. $H(\cdot)$ is a continuous function of all of its arguments, $H(p) = H(p_1, p_2, \dots, p_n)$.

Axiom 3. Uncertainty is invariant under permutation π of the probability components: $H(p_1, p_2, \dots, p_n) = H(\pi(p_1, p_2, \dots, p_n))$ $\pi(p)$ being a permutation of p .

Axiom 4. The maximum uncertainty is obtained when all probability components are equal to $1/n$:

$$H(p_1, p_2, \dots, p_n) \leq H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right).$$

Axiom 5. The uncertainty of any joint probability must not exceed the sum of the uncertainties of the corresponding marginal distributions ([[69]; p.73]): i.e. given a joint probability distribution p_{ij} , $1 \leq i \leq n$, $1 \leq j \leq m$,

$$\begin{aligned}
 H(p_{11}, p_{12}, \dots, p_{1m}, p_{21}, p_{22}, \dots, p_{2m}, \dots, p_{n1}, p_{n2}, \dots, p_{nm}) &\leq \\
 H\left(\sum_{j=1}^m p_{1j}, \sum_{j=1}^m p_{2j}, \dots, \sum_{j=1}^m p_{mj}\right) + H\left(\sum_{i=1}^n p_{i1}, \sum_{i=1}^n p_{i2}, \dots, \sum_{i=1}^n p_{im}\right).
 \end{aligned}$$

Axiom 6. We have for joint probability $p_i q_j$, $1 \leq i \leq n$, $1 \leq j \leq m$,

$$\begin{aligned} H(p_1 q_1, p_1 q_2, \dots, p_1 q_m, p_2 q_1, p_2 q_2, \dots, p_2 q_m, \dots, p_n q_1, p_n q_2, \dots, p_n q_m) = \\ H(p_1, p_2, \dots, p_n) + H(q_1, q_2, \dots, q_m) . \end{aligned}$$

Thus, the uncertainty of a joint probability should equal the sum of uncertainties of the corresponding marginal distributions.

Axiom 7. The uncertainties of probabilities with components equal to $1/n$ (as in **Axiom 4**) should increase with increasing n .

Axiom 8. A probability distribution must have the *branching property*: to $A = \{x_1, x_2, \dots, x_s\}$ and $B = \{x_{s+1}, x_{s+2}, \dots, x_n\}$, then,

$$H(p_1, p_2, \dots, p_n) = H(p_A, p_B) + p_A H\left(\frac{p_1}{p_A}, \frac{p_2}{p_A}, \dots, \frac{p_s}{p_A}\right) + p_B H\left(\frac{p_{s+1}}{p_B}, \frac{p_{s+2}}{p_B}, \dots, \frac{p_n}{p_B}\right) .$$

Axiom 9. $H(\frac{1}{2}, \frac{1}{2}) = 1$ if the units of H are to be bits.

Shannon's theorem is then:

Theorem 5.2. *The only functional H on \mathcal{P}^n that satisfies the axioms listed above is the Shannon entropy, $H(p) = -\sum_{i=1}^n p_i \log p_i$.* \square

Details of the proof can be found in Klir [69] or, in slightly different forms of the axioms, in Jaynes [60] and, of course, Shannon [104].

5.7 Example: Entropy of the Uniform Distribution

Let U be a *uniform random variable*, that is, U takes on some constant values in a bounded domain $D \subset \mathbb{R}^n$, with volume (or measure) $|D| < \infty$ (see Figure 5.1). The corresponding density is,

$$\pi(x) = \frac{1}{|D|}, \quad a \leq x \leq b ,$$

meaning that for any x in $[a, b]$, the same probability, $1/(b - a)$, occurs. The entropy of U is:

$$H(U) = \log(|D|) .$$

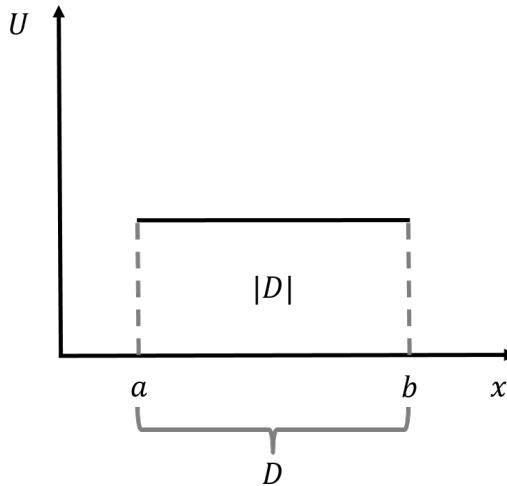


Figure 5.1: A m -dimensional example of uniform distribution over a domain D .

The question arises, how uncertain is the random variable U ? The answer is that the entropy of any other random variable X (call it $H(X)$ for now) will always be smaller than $H(U)$:

$$H(X) \leq H(U) .$$

The uniform distribution (or, for instance, the uniform prior $\pi_{\text{prior}}(x) = \text{const.}$) is thus the distribution of maximum uncertainty (or, as some say, maximum ignorance).

Considerably more information is contained in the case in which the mean μ and the variance σ^2 of a distribution are known. Then, the correct distribution $\pi(x)$ must maximize the entropy subject to appropriate constraints; i.e.,

$$\begin{aligned} \pi(x) &= \arg \max_{\pi} H(x) \\ &= \arg \max_{\pi} \left\{ \int \pi(x) \log \pi(x) d x \right\} \end{aligned}$$

subject to the constraints,

$$\begin{aligned} \int \pi(x) d x &= 1 , \\ \int x \pi(x) d x &= \mu , \\ \int (x - \mu)^2 \pi(x) d x &= \sigma^2 . \end{aligned}$$

Constructing a Lagrangian functional Υ defined by the entropy and Lagrange multiplier times these constraints, we easily solve the constrained maximization problem and establish that the solution to the maximization problem is,

$$\pi(x) = \exp \left\{ -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\} ,$$

a Gaussian, as probably expected.

This process of computing prior distributions by maximizing the entropy subject to constraints was introduced by Jaynes [58] in 1968 and used by him to generate priors for very general situations in which some a priori information is known about the variable x (generally, in our applications $x = \theta$, the parameters). We examine Jaynes' approach more thoroughly in the next section.

5.8 Maximum Entropy Priors: The General Solution for Maximum Entropy with Linear Constraints

An important application of the maximum entropy principle for constructing prior distributions when some statistical information is known, was developed by Jaynes [58] and described later in his book *Probability Theory: The Logic of Science* [60]. This section reviews the basic ideas.

Consider a real-valued variable x about which we have n propositions $\{A_1, A_2, \dots, A_n\}$. We have prior information I about x that takes on values $\{x_1, x_2, \dots, x_n\}$ corresponding to the respective propositions. The probabilities p_i that we assign to the possibilities x_i form the set $\{p_1, p_2, \dots, p_n\}$.

Next, suppose we have a set of m functions of x , $f_k(x)$, $1 \leq k \leq m < n$. We want to prescribe the expectations of the f_k relative to the probability distribution $\{p_1, p_2, \dots, p_n\}$;

$$\langle f_k(x) \rangle = \sum_{i=1}^n p_i f_k(x_i) = F_k \quad (5.22)$$

where F_k are given numbers. *Thus, given the F_k , and given the x_i . we wish to find the probabilities p_i which has maximum entropy, subject to the relevant constraints, including $\sum_i p_i = 1$.* This leads to the problem of finding critical points of the functional,

$$\Upsilon(p, \lambda_0, \lambda_1, \dots, \lambda_m) = H(p) - \lambda'_0(1 - \sum_{i=1}^n p_i) - \sum_{k=1}^m \lambda_k(F_k - \sum_i^n p_i f_k) \quad (5.23)$$

where $\lambda'_0, \lambda_1, \dots, \lambda_m$ are Lagrange multipliers and $f_k = f_k(x_i)$. We have,

$$\sum_{i=1}^n \left[\frac{\partial H(p)}{\partial p_i} \delta p_i - \lambda'_0 \delta p_i + \sum_{k=1}^m \lambda_k f_k \delta p_i \right] = 0$$

$\forall \delta p_i$. Since $H(p) = -\sum_i p_i \log p_i$, $\partial H/\partial p_i = -\log p_i - 1$,

$$\log p_i - 1 - \lambda'_0 + \sum_k \lambda_k f_k = 0,$$

or,

$$p_i^{max} = p_i = \exp(\lambda_0 - \sum_{k=1}^m \lambda_k f_k) \quad (5.24)$$

where $\lambda_0 = \lambda'_0 + 1$. Returning to the constraints:

$$1 = \sum_{i=1}^n p_i = \exp \left[-\lambda_0 \sum_i \exp(-\sum_{k=1}^m \lambda_k f_k) \right]$$

At this point, we introduce the *partition function*

$$Z(\lambda_1, \lambda_2, \dots, \lambda_m) = \sum_i^n \exp(-\sum_{k=1}^m \lambda_k f_k) , \quad (5.25)$$

so,

$$\lambda_0 = \log Z . \quad (5.26)$$

Returning to (5.19), we have,

$$F_k = \sum_i^n \exp[\log Z - \sum_{k=1}^m \lambda_k f_k(x_i)]$$

which we recognize as,

$$F_k = -\frac{\partial}{\partial \lambda_k} (\log Z(\lambda_1, \lambda_2, \dots, \lambda_m)) \quad (5.27)$$

which determines the λ_k as functions of the given F_k . The maximum entropy is,

$$H_{max} = -\sum_{i=1}^n p_i^{max} \log p_i^{max}$$

where p_i^{max} is given by (5.24). Thus,

$$H_{max} = \lambda_0 + \sum_{k=1}^m \lambda_k F_k. \quad (5.28)$$

We can now assert that H attains its maximum if and only if $p_i = p_i^{max}$,

$$p_i = \frac{1}{Z} \exp\left[-\sum_{k=1}^m \lambda_k f_k(x_i)\right]. \quad (5.29)$$

One significance of problem (5.22) is this: if we know certain properties of a distribution, such as its mean, moments, or other features F_k , then we construct the prior which fits these constraints and which maximizes the entropy.

Examples

The canonical examples of application of this process are two cases:

1. The mean of the distribution μ is known:

$$\sum_i p_i x_i = \langle x \rangle = \mu,$$

then, maximizing $H(p)$ subject to the constraints $\sum_i p_i = 1$ and $\sum_i p_i x_i = \mu$ leads to the exponential density,

$$\pi(x) = \frac{1}{\mu} \exp\left\{-\frac{x}{\mu}\right\}.$$

2. If, in addition, the variance is known,

$$\sum_i p_i(x_i - \langle x \rangle)^2 = \sigma^2,$$

then, as indicated earlier, the maximum of $H(p)$ is attained at the Gaussian density,

$$\pi(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}.$$

Covariance and Other Properties. Several other properties of the prior can be derived from the maximum entropy theory and the solution (5.24) (or (5.29)). Details can be found in Jaynes [60]. Among these, we mention the relation:

$$\frac{\partial \langle f_k \rangle}{\partial \lambda_j} = \frac{\partial \langle f_j \rangle}{\partial \lambda_k} = \frac{\partial^2 \log Z(\lambda_1, \lambda_2, \dots, \lambda_m)}{\partial \lambda_j \partial \lambda_k} \quad (5.30)$$

Given the canonical distribution (5.29), one may ask how the mean values $\langle f_k(x) \rangle$ are correlated with each other? The departures from the mean determine the *covariance*,

$$C_{jk} = \langle (f_j - \langle f_j \rangle)(f_k - \langle f_k \rangle) \rangle = \langle f_j f_k \rangle - \langle f_j \rangle \langle f_k \rangle \quad (5.31)$$

From (5.30), it can be seen that,

$$C_{jk} = -\frac{\partial \langle f_j \rangle}{\partial \lambda_k} = -\frac{\partial \langle f_k \rangle}{\partial \lambda_j} \quad (5.32)$$

Thus, *the (negative of) the second derivatives of the log of the partition function determine the covariance of f_j and f_k in our distribution.*

The Continuous Case. In the case the probability density is a continuous function $p = p(x)$, we face the problem that $-\int p \log(p) dx$ is not invariant under a change of variable $x \rightarrow y(x)$, and hence is not an acceptable definition of entropy. To correct this, we introduce (following Jaynes [60] p. 375) an invariant measure $m(x)$ defined as the limit of discrete points x_i , $i = 1, 2, \dots, n$, in, say an interval $[a, b]$:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \quad (N = \text{no. of points } m(a, b) \stackrel{\text{def}}{=} \int_a^b m(x) dx), \quad (5.33)$$

or

$$\frac{1}{m(x_i)} = \lim_{n \rightarrow \infty} [n(x_{i+1} - x_i)].$$

Then one can argue that the continuous information measure is

$$H(p) = - \int p(x|I) \log \frac{p(x,I)}{m(x)} dx , \quad (5.34)$$

with I the prior information. In many accounts on entropy, the role of $m(x)$ is ignored and one simply sets $m(x) = 1$ in (5.34), as was done in earlier sections of this chapter.

The “continuous version” of problem (5.19) is, given F_k , $k = 1, 2, \dots, m$ and the functions f_k , find the probability distribution $p(x)$ such that

$$\int p(x|I) f_k(x) dx = F_k, \quad 1 \leq k \leq m . \quad (5.35)$$

In this case, the solution to the maximum entropy problem is

$$p(x, I) = Z^{-1}(\lambda_1, \lambda_2, \dots, \lambda_m) m(x) \exp[\lambda_1 f_1(x) + \lambda_2 f_2(x) + \dots + \lambda_m f_m(x)] . \quad (5.36)$$

5.9 The Tan Bui Thanh Lemma: Entropy as a path to Bayes

A interesting observation on the connection between maximum entropy and Bayes’ theorem was made by Tan Bui Thanh [24]. Let us look at two optimization problems that are designed to produce the optimal choice of a posterior density π_{post} , given data \mathbf{y} and a likelihood function $\pi_{\text{like}} = \pi(\mathbf{y}|\mathbf{x})$.

First, let us choose π_{post} to minimize the relative entropy (the D_{KL} divergence) between a density π and a given prior, π_{prior} . Then,

$$\begin{aligned} \pi_{\text{post}} &= \arg \min_{\pi} \int \pi(\mathbf{x}) \log \frac{\pi(\mathbf{x})}{\pi_{\text{prior}}(\mathbf{x})} d\mathbf{x} \\ &= \arg \min_{\pi} D_{\text{KL}}(\pi || \pi_{\text{prior}}) , \end{aligned} \quad (5.37)$$

subject to the constraints,

$$\int \pi(\mathbf{x}) d\mathbf{x} = 1 \quad \text{and} \quad \pi(\mathbf{x}) \geq 0 . \quad (5.38)$$

Second, let us choose, for given data \mathbf{y} and likelihood $\pi_{\text{like}}(\mathbf{y}|\mathbf{x})$, a posterior π_{post} that minimizes the mean-square error:

$$\begin{aligned} \pi_{\text{post}} &= \arg \min_{\pi} \int \pi(\mathbf{x}) \|\mathbf{y} - \mathbf{d}(\mathbf{x})\|_{\Sigma}^2 d\mathbf{x} \\ &= \arg \min_{\pi} \left\{ - \int \pi(\mathbf{x}) \log \pi_{\text{like}}(\mathbf{y}|\mathbf{x}) d\mathbf{x} \right\} \end{aligned} \quad (5.39)$$

To attempt to satisfy requirement of (5.37) and (5.39), one consider a “compromise” functional and optimization problem,

$$\pi_{\text{post}} = \arg \min_{\pi} \left\{ D_{\text{KL}}(\mathbf{x} \parallel \pi_{\text{prior}}) - \int \pi(\mathbf{x}) \log \pi_{\text{like}}(\mathbf{y}|\mathbf{x}) d\mathbf{x} \right\} ,$$

subject to the constraints (5.38).

Remarkably, appending Lagrange multipliers to the constraint, and solving the constrained minimization problem leads to the following solution:

$$\pi_{\text{post}}(\mathbf{x}|d) = \frac{\pi_{\text{like}}(\mathbf{d}|\mathbf{x}) \pi_{\text{prior}}(\mathbf{x})}{\int \pi_{\text{like}}(\mathbf{d}|\mathbf{x}) \pi_{\text{prior}}(\mathbf{x}) d\mathbf{x}} , \quad (5.40)$$

which, remarkably, is Bayes’ rule. Thus, Bayes’ rule is optimal for computing the posterior under the condition laid down above.

Chapter 6

Monte Carlo Methods: MCMC and Sensitivity Analysis

6.1 Introduction

Some basic issues are, as yet, unaddressed in our study of Bayesian inference and statistical calibration and validation. First and foremost, how do we numerically evaluate the influence of the posterior distributions arising from the Bayes rule on key quantities of interest? This issue of numerically evaluating Bayes' rule was a principal contributing factor to the long dormant period in which the importance of Bayesian approaches was largely ignored by the statistics community. Secondly, there is the question of what model parameters appreciably affect the quantities of interest in a computer simulation? Specifically, in almost every application of interest, the quantities of interest (QoIs) which are the target centerpiece of the prediction, are not influenced by the choice of some, often many, values of the model parameters. This begs the question: of all the parameters in a given model, and given a QoI, which parameters can be eliminated, because variations in them have negligible effect on the values of the QoIs?

Remarkably, both of these questions can be addressed with great efficiency using *Monte Carlo* methodologies. Monte Carlo methods represent a broad class of computational algorithms that rely on repeated random samples of a function to obtain numerical representations of values of the function over its domain.

The term ‘Monte Carlo’ was used by the earlier inventors and users of the idea, in analogy with

gambling with random bets at the Monte-Carlo casino in Monaco. It was introduced by atomic bomb scientists, Nicholes Metropolis and Stanislaw Ulam in 1949, Metropolis et al. in 1953 [75], and with further improvements by Wilfred Hastings in 1970 [52]. The basic idea involves solving deterministic problems using a probabilistic analogy.

The most common Monte-Carlo algorithm is that associated with integration of a function $f(x)$ over a domain Ω , $\int_{\Omega} f(x)dx$. There the idea is to introduce a set X of randomly selected points $X = \{x_1, x_2, \dots, x_n\}$ in Ω and approximate the integral with the sum,

$$\int_{\Omega} f(x)dx \approx |\Omega| \frac{1}{n} \sum_{i=1}^n f(x_i) , \quad (6.1)$$

$|\Omega|$ being the volume of Ω . In actual applications, it is sufficient to employ a pseudo-random set of sample points. The error is $O(1/\sqrt{n})$:

$$\text{Var}\left(\int_{\Omega} f(x)dx - |\Omega| \frac{1}{n} \sum_{i=1}^n f(x_i)\right) = \frac{1}{n} \left\{ \sum_{i=1}^n f^2(x_i) - \left[\frac{1}{n} \sum_{i=1}^n f(x_i) \right]^2 \right\} \quad (6.2)$$

In the applications of Bayes' rule of interest here, the goal is to use Monte-Carlo sampling of a distribution to trace out its variation over its domain, as opposed to integrating a posterior over that domain. Thus, we have, for example, the posterior probability density $\pi_{\text{post}}(\boldsymbol{\theta}|\mathbf{y}, S)$ – denoted $\pi_{\text{post}}(\boldsymbol{\theta})$ for simplicity – and we generate a sequence of samples from $\pi_{\text{post}}(\boldsymbol{\theta})$ by randomly selecting parameter value $\boldsymbol{\theta}_i$ in the domain, $\text{dom } \pi_{\text{post}}$. The challenge is to develop a systematic strategy to sample all (or a sufficient number) of values to fully represent the structure of π_{post} . This, we accomplish by following a random walk over the contours of the surface which is defined by a Markov chain, described in Section 6.2. An example of such an approximation of the surface is given in Figure 6.1.

As to sensitivity analysis, the idea is to find out how the QoI, or an appropriate surrogate value Y , varies when specific parameters of a model take on random incremental variations. Ideally, we want to know how large are the variations in the QoI due to variations in the parameters. The parameter sensitivities are thus related to the numbers,

$$\left| \frac{\partial Q(u(\boldsymbol{\theta}, S))}{\partial \boldsymbol{\theta}_i} \right| , \quad 1 \leq i \leq n , \quad \boldsymbol{\theta}_i \in \boldsymbol{\theta} \in \Theta . \quad (6.3)$$

Unfortunately, such sensitivity calculations are often very computationally intensive and do not provide a useful measure of parameter sensitivity. Instead, we are led to surrogate measures of

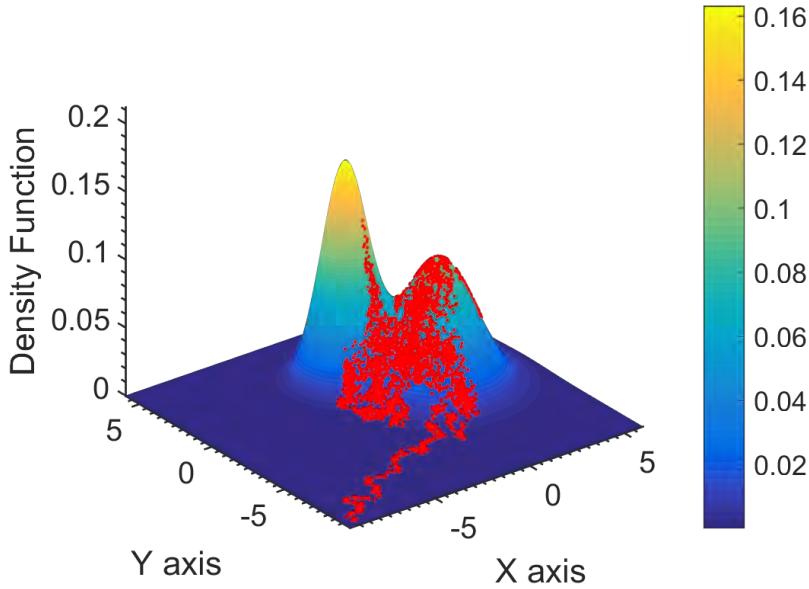


Figure 6.1: The structure of a posterior distribution $\pi_{\text{post}}(\boldsymbol{\theta})$ depending on the parameters $\boldsymbol{\theta} = (x, y)$ constructed by computing samples along a random walk, starting at $\boldsymbol{\theta} = (-9, -9)$.

output dependency on the parameter, say $Y = Y(\boldsymbol{\theta})$, and to numerical methods which compute the sensitivity of $Y(\boldsymbol{\theta})$ to perturbations in model parameter $\boldsymbol{\theta}$. We describe such sensitivity measures and calculations advocated by Saltelli [97] and others later in this chapter. As mentioned earlier, these sensitivity calculations employ Monte-Carlo methodologies.

In this chapter we provide an introduction to Monte-Carlo sampling methods underlying the famous and powerful MCMC methods (Markov-Chain Monte Carlo Methods) and later variance-based Monte-Carlo sampling methods of sensitivity analysis for model sensitivities.

6.2 Markov Chains

We must build a representation of the structure of the surface of probability density such as $\pi_{\text{post}}(\boldsymbol{\theta})$ by using a Monte Carlo sampling of a process that traces out a path connecting the random sample points \boldsymbol{x} in a way that can cover the domain of the function as the number of samples is increased. Such a sampling can be performed by a *Markov chain*, which is a sequence of random variables generated by sampling along paths in the function domain called a chain. The Markov Chain is a random process that traverses from one “state” to another in such a way that they create a “chain” through a state space in which the predicted distribution on the next state in a chain of probabilities depends only on the current state and not on the events or their probabilities that

produces them. Such a “memoryless-ness” property is called a *Markov property*. In applications of interest here, we pick an initial point in, say, \mathbb{R}^n , and create a *random walk*, a path through the space, guided by a set of decision rules.

In a more formal setting, a Markov chain can be considered to be a sequence of random variables X, Y, Z accessed by a process in a certain order, $X \rightarrow Y \rightarrow Z$ such that the conditional probability of Z depends only on Y and is independent of X . The joint probability (mass) can be thus written,

$$\begin{aligned}\pi(x, y, z) &= \pi(z|y, x)\pi(y|x)\pi(x) \\ &= \pi(z|y)\pi(y|x)\pi(x) \\ &= \pi(z|x, y)\pi(x|y) \\ \Rightarrow \pi(x, z|y) &= \pi(x|y)\pi(z|y) .\end{aligned}\tag{6.4}$$

The memoryless property shows that no processing of information about step Y can increase the information at Y about information at X (written in terms of mutual information as $I(X, Y) \geq I(X, Z)$). See [30].

6.3 Markov Chain Monte Carlo

The challenge in applying Monte Carlo sampling is to limit the number of samples in regions of the domain of the integrand where the integrand is very small. A process for tracing out the random walk along a chain that selects directions in which the highest probability of capturing meaningful values of the integrand is the goal of the Markov Chain Monte Carlo (MCMC) strategy.

Suppose that we want to map out and explore the structure of a posterior density $\pi_{\text{post}}(\boldsymbol{\theta})$ ($= \pi(\boldsymbol{\theta}|\mathbf{y}, S)$), through a series of trial points $\boldsymbol{\theta}^k$ on a Markov chain starting at an initial value $\boldsymbol{\theta}^0$ of the parameters $\boldsymbol{\theta}$, and designed to have the Markovian property; that each new sample $\boldsymbol{\theta}^{(k+1)}$ does not depend on the earlier samples, beyond $\boldsymbol{\theta}^{(k-1)}$. The new value $\boldsymbol{\theta}^{(k+1)}$ depends on the previous sample $\boldsymbol{\theta}^k$ through a *transition probability* $p(\boldsymbol{\theta}^{(k+1)}, \boldsymbol{\theta}^{(k)})$ that is designed to make the path and samples generate all of the posterior.

A first step is to pick a *proposal (conditional) distribution* $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{(k)})$ that gives the value $\boldsymbol{\theta}^*$ of the $k + 1$ step, given the k th value $\boldsymbol{\theta}^{(k)}$. The density q can be of any form, such as multivariate Gaussian with mean equal to the current sample $\boldsymbol{\theta}^{(k)}$. The density then decreases with distance away from $\boldsymbol{\theta}^{(k)}$, the current sample.

Secondly, we must decide if we accept the candidate sample $\boldsymbol{\theta}^*$. We do this on the basis of the *Metropolis ratio*,

$$r = \frac{\pi_{\text{post}}(\boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta}^{(k)})} \times \frac{q(\boldsymbol{\theta}^k | \boldsymbol{\theta}^{(*)})}{q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^{(k)})}. \quad (6.5)$$

If the proposal distribution is symmetric, then the second factor is 1. If not, and $r \geq 1$, we set $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^*$. If $r < 1$, we accept $\boldsymbol{\theta}^*$ with probability r , this latter step is done with by sampling a random variable U from the uniform distribution $U(0, 1)$. If $U \leq r$, we set $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^*$; otherwise, $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)}$.

This second step is captured by the *acceptance probability* $\alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(k)})$, defined by,

$$\alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(k)}) = \min(1, r) = \min\left(1, \frac{\pi_{\text{post}}(\boldsymbol{\theta}^*) q(\boldsymbol{\theta}^{(k)} | \boldsymbol{\theta}^*)}{\pi_{\text{post}}(\boldsymbol{\theta}^k) q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^{(k)})}\right) \quad (6.6)$$

Thus,

$$\boldsymbol{\theta}^{(k+1)} = \begin{cases} \boldsymbol{\theta}^* & \text{with probability } \min\{1, \alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(k)})\} \\ \boldsymbol{\theta}^{(k)} & \text{with probability } 1 - \min\{1, \alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(k)})\} \end{cases} \quad (6.7)$$

This version of MCMC is called the *Metropolis-Hastings algorithm*. It is used as a basis for many popular statistical sampling programs, and, in particular, the code QUESO [87] of Prudencio and Schultz [89] is employed in calculations in later chapters.

6.4 Parameter Sensitivity Analysis

As noted earlier, it is frequently the case that QoIs or experimental observations are not sensitive to changes in certain model parameters. Since computational cost and complexity may strongly depend on the number of parameters appearing in a model, any rational process that leads to the reduction in the number of relevant parameters is regarded as very valuable. This fact has led to the development of several effective methods for parameter sensitivity analysis, represented in the work of Sobol in 1990, 1993, 2007 [105–107], Homma and Saltelli (1996) [54], and the book of Saltelli [97], as well as Saltelli et al. (2009, 2010) [95, 96, 98], and Saltelli and Sobol (1995) [99]. It must be appreciated that parameter sensitivity depends on the particular QoI; parameter changes may strongly affect some QoI-values but not affect at all other choices of QoIs.

The setting for parameter sensitivity analysis is this: for a given parametric model involving a set

of parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, \dots, \theta_m\}$, select an *output function* $Y(\boldsymbol{\theta})$ that represent a principal output of interest. The output function is thus a surrogate for the QoI, and indeed, when feasible, one can select $Y(\boldsymbol{\theta})$ to be the QoI: $Y(\boldsymbol{\theta}) = \tilde{Q}(\boldsymbol{\theta}) = Q(u(\boldsymbol{\theta}, S_p))$, S_p being the prediction scenario. The goal is to measure the sensitivity of $Y(\boldsymbol{\theta})$ to changes in components of $\boldsymbol{\theta}$.

6.4.1 Morris's Method of Elementary Effects

A popular method of parameter sensitivity analysis is that of Morris [77] as the method of *elementary effects*. The principal ideas behind the method are listed as follows,

1. We identify an output function $Y(\boldsymbol{\theta}) = Y([\theta_1, \theta_2, \dots, \theta_n])$ which is an output model of the QoI and is a function of the n-dimensional vector of parameters $\boldsymbol{\theta}$.
2. We rescale each parameter θ_i to the unit interval $[0, 1]$.
3. We consider an n-dimensional hypercube, $\Omega^n = [0, 1]^n$ which we partition into p^n cells, p being an integer number > 2 , such that each parameter domain $[0, 1]$ is partitioned into bins $\theta_i \sim [0, 1] \sim [0, \frac{1}{p-1}, \frac{1}{p-2}, \dots, 1]$. These bins $[\frac{1}{p-i}, \frac{1}{p-i-1}]$ contain sample points from which Monte Carlo samples of θ_i will be selected. The bins are called the *levels* of the cube.
4. We introduce a sampling distribution π_i for each θ_i , π_i possibly denoting a prior density, or simply a uniform distribution assigned to all θ_i , $i = 1, 2, \dots, n$.
5. We next choose a small number δ , for instance $\delta = p/2(p - 1)$, $p \geq 2$, and we compute the elementary effect EE_i as the difference approximations,

$$EE_i = \frac{Y(\theta_1, \theta_2, \dots, \theta_i + \delta, \theta_{i+1}, \dots, \theta_n) - Y(\boldsymbol{\theta})}{\delta}, \quad 1 \leq i \leq n. \quad (6.8)$$

The EE_i 's thus represent the change in the output due to a δ -perturbation in each parameter θ_i .

6. We now compute EE_i^j of the output for random (Monte Carlo) samples taken from each bin. If the parameter distribution is not uniform, we sample quantiles in $[0, 1]$ deriving input values from the inverse of the cumulative distribution function.
7. Now we must traverse through the hypercube doing sample paths, called *trajectories*, consisting of paths of $n + 1$ orthogonal steps through Ω^n defined as follows:
 - (a) Select a random starting point $\boldsymbol{\theta}^{(1)}$

- (b) If $\boldsymbol{\theta}_i$ denotes a set of unit coordinate vectors on the axes of the hypercube, select a new vector $\boldsymbol{\theta}^{(2)}$ differing from $\boldsymbol{\theta}^{(1)}$ in its i^{th} component by δ :

$$\boldsymbol{\theta}^{(2)} = \boldsymbol{\theta}^{(1)} \pm \mathbf{e}_i \delta,$$

with + or - taken such that $\boldsymbol{\theta}^{(2)} \in \Omega$.

- (c) Continuing in this manner,

$$\boldsymbol{\theta}^{(3)} = \boldsymbol{\theta}^{(2)} \pm \mathbf{e}_j \delta$$

we finally get,

$$\boldsymbol{\theta}^{(n+1)} = \boldsymbol{\theta}^{(n)} \pm \mathbf{e}_k \delta.$$

This defines a trajectory r through Ω^n .

8. Next, we compute the elementary effects sensitivities and variances over r trajectories,

$$\mu_i = \frac{1}{r} \sum_{j=1}^r EE_i^j , \quad (6.9)$$

$$\sigma_i^2 = \frac{1}{r-1} \sum_{j=1}^r (EE_i^j - \mu_i)^2 , \quad (6.10)$$

with μ_i being a mean sensitivity and σ_i^2 a variance at the i^{th} parameter sensitivity (over r trajectories). A common indication of the sensitivity assigned to parameter i is,

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^r |EE_i^j| . \quad (6.11)$$

By laying out the numbers $\mu_1^*, \mu_2^*, \dots, \mu_n^*$, we are able to estimate the relative sensitivity of the output to changes in the parameters. When a measure μ_k^* is regarded as very “small” according to some preset tolerance, that parameter can be eliminated or set equal to a deterministic constant, this latter choice being sometimes made when sensitivities of other parameters are correlated with those of θ_i . We note that the sensitivity measures μ_i , σ_i^2 and μ_i^* depend on the number r of trajectories. Presumably, these measures should converge to constants as $r \rightarrow \infty$. Some examples of such calculations are given in Chapter 10.

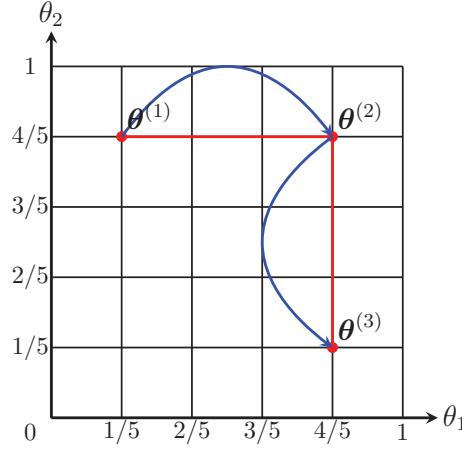


Figure 6.2: An illustration of a 2-dimensional hypercube ($n = 2$) with $p = 6$ levels, $\delta = 3/5$, and a trajectory r along $n + 1 = 3$ sampling points: $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(3)}$.

6.4.2 Variance-Based Sensitivities

A popular measure of parameter sensitivity in the variance-based sensitivity advocated in several papers and books (e.g. Cukier et al (1973) [35], Sobol (1993) [106], Saltelli et al (2000) [97]).

As in the case of the elementary effects measures, we construct a n -dimensional unit hypercube $\Omega^n = \{\boldsymbol{\theta}, 0 \leq x_i \leq 1, i = 1, 2, \dots, n\}$ and generate random (Monte Carlo) samples of the uncertain inputs $\theta_1, \theta_2, \dots, \theta_n$ from some prescribed distribution (generally uniform or Gaussian). Let $\boldsymbol{\theta}_{\sim i}$ denote the parameters vector which excludes the i th component of $\boldsymbol{\theta}$,

$$\boldsymbol{\theta}_{\sim i} = (\theta_1, \theta_2, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n)^T$$

We denote by $V_i(y|\boldsymbol{\theta}_{\sim i})$ the variance of the Monte Carlo sample of Y evaluates at $\boldsymbol{\theta}_{\sim i}$ and by $\mathbb{E}_{\boldsymbol{\theta}_{\sim i}}[V_i(y|\boldsymbol{\theta}_{\sim i})]$ the average over all such variance. Then the *variance-based sensitivity* of θ_i is defined by

$$S_{Ti} = \frac{\mathbb{E}_{\boldsymbol{\theta}_{\sim i}}[V_i(y|\boldsymbol{\theta}_{\sim i})]}{V(Y)} \quad (6.12)$$

where $V(Y)$ is the variance of the output function.

Generally, it is convenient to regard the output as being defined by a square-integrable function $f(\boldsymbol{\theta})$ defined on the n -dimensional hypercube Ω^n . Then, to derive (6.12) or alternative sensitivity

measure, one begins with the so-called Hoeffding decomposition of f [105, 106] and the conditional expectations $\mathbb{E}(Y|\theta_i)$ to derive the following decomposition of the output variance:

$$V(Y) = \sum_i V_i + \sum_j \sum_{j>i} V_{ij} + \cdots + V_{12\cdots m}, \quad (6.13)$$

where

$$\begin{aligned} V_i &= V(f_i(\theta_i)) = V_{\theta_i}(\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i)), \\ V_{ij} &= V(f_{ij}(\theta_i, \theta_j)), \\ &= V_{\theta_i \theta_j}(\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i, \theta_j)) - V_{\theta_i}(\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i)) - V_{\theta_j}(\mathbb{E}_{\theta_{\sim j}}(Y|\theta_j)). \end{aligned} \quad (6.14)$$

Here $V_{\theta_i}(\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i))$ and $\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i, \theta_j)$ are first-order effects, $V_{\theta_i \theta_j}(\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i, \theta_j))$ are joint effects of the pair (θ_i, θ_j) on V , etc. Complete details can be found in [54]; see also the appendix in [39].

For so-called non-additive models, Y depends on the interaction between different parameters θ_i and θ_j ($i \neq j$). Then we divide (6.14) by $V(Y)$ forming sensitivity indices $S_i, S_{ij}, \dots, S_{1,2,\dots,m}$ yielding

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \cdots + S_{1,2,\dots,m} = 1. \quad (6.15)$$

The *total effective index* can be derived from this property and (6.12); i.e.,

$$\begin{aligned} S_{T_i} &= \frac{\mathbb{E}_{\theta_{\sim i}}(V_{\theta_i}(Y|\theta_i))}{V(Y)} \\ &= 1 - \frac{V_{\theta_i}(\mathbb{E}_{\theta_{\sim i}}(Y|\theta_i))}{V(Y)}. \end{aligned} \quad (6.16)$$

The total effect sensitivity S_{T_i} measures the total contributions to the output variation in parameter θ_i . Straightforward algorithms based on Monte Carlo sampling for computing such indices are discussed in [94–96, 98]; see a slightly amended version in [39].

The process of using such sensitivity measures to eliminate irrelevant parameters involves computing S_{T_i} 's for each parameter $\theta_i \in \boldsymbol{\theta}$, and removing from the list of parameters those for which the index is sufficiently small (e.g. for which $S_{T_i}/\max_j S_{T_j} < 0.1$).

6.4.3 Scatter Plots

A generally convenient and often inexpensive qualitative way to visualize the sensitivity of outputs to parameter variations is through scatter plots, which involve clouds of points showing an output

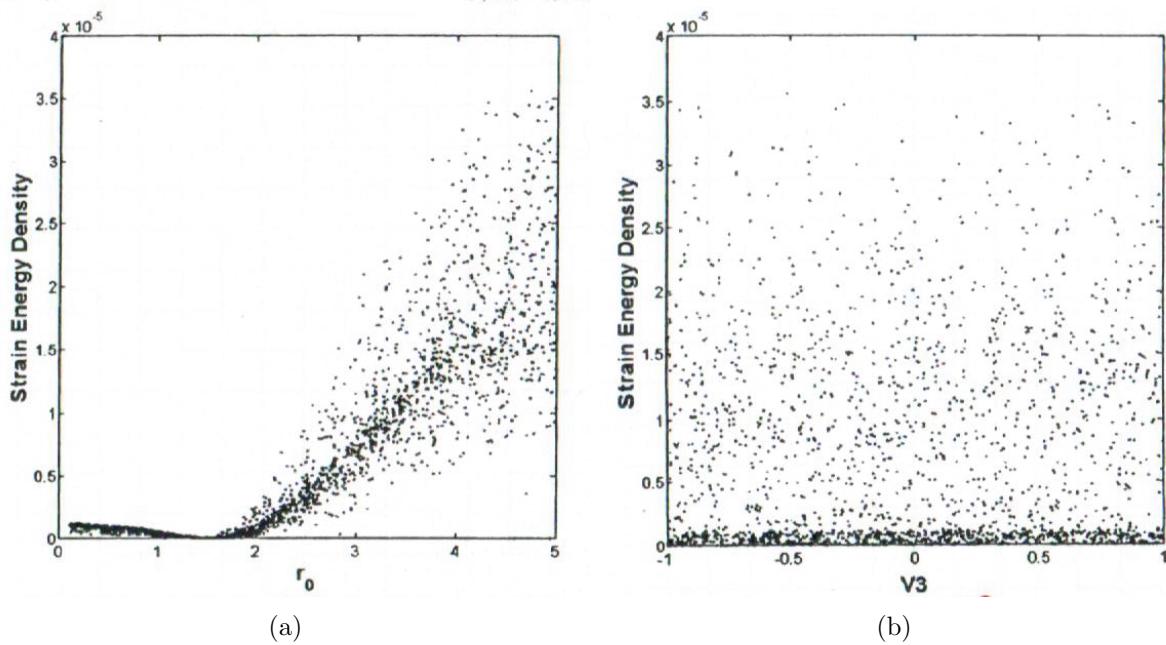


Figure 6.3: Parameters may have a significant influence on the output. The uniform distribution in b) suggests an insensitivity of the outputs to the parameters.

$Y(\boldsymbol{\theta})$ plotted against randomly chosen values of each parameter in a model. Examples are shown in Figure 6.3. The idea is to take random (Monte-Carlo) samples of a uniform distribution of a parameter θ_i and plot the corresponding $Y(\theta_i)$ -values over a two-dimensional domain. A parameter is considered to have a possibly important influence on $Y(\theta)$ if the cloud exhibits a distinct pattern. A uniform cloud, or a cloud close to zero, suggests that the sensitivity of the output to changes in that particular parameter is small.

6.4.4 Design of Validation Experiments

The calculation of parameter sensitivities can provide an important tool in assessing the effectiveness of the design of the validation experimental scenario S_v . When feasible, one computes parameter sensitivities for the output functions $Y(\boldsymbol{\theta})$ set equal to the QoI, $Q(\boldsymbol{\theta})$, in the full prediction scenario S_p . One then recomputes the sensitivity for a proposed validation scenario with $Y(\boldsymbol{\theta})$ set equal to the validation observables. If the validation sensitivities are not in good agreement with those for the full prediction scenario, then the validation experiments are ineffective and cannot determine parameters that inform the prediction of the QoI. Such comparisons of S_p and S_v sensitivities provides a powerful indication of the effectiveness of the design of the validation experiments. In the same vein, comparing S_c and S_p sensitivities can also provide insight in

designing calibration experiments.

6.4.5 An Algorithm for Sensitivity Indices

Returning to (6.12), a Monte Carlo scheme to compute the indices with a reduced computational cost of evaluating multi-dimensional integrals is presented here, due to Faghihi in Farrell et al. [39] and inspired by [97]. The principal sets are:

1. Generate two $N \times k$ independent sampling matrices, \mathbf{A} and \mathbf{B} , where each row is a sample point in the hyperspace of k dimensions. N is called a base sample with respect to the probability distributions of the input variables.
2. Define a matrix \mathbf{D}_i formed by all columns of \mathbf{A} except the i^{th} column, which is from \mathbf{B} .
3. Compute the model outputs for all the matrices (i.e., $N \times 1$ vectors of model output):

$$\mathcal{Y}_A = f(\mathbf{A}) \quad ; \quad \mathcal{Y}_B = f(\mathbf{B}) \quad ; \quad \mathcal{Y}_{D_i} = f(\mathbf{D}_i)$$

4. The total-effect of indices can be estimated by,

$$S_{T_i} \approx 1 - \frac{\frac{1}{N} \sum_{j=1}^N \mathcal{Y}_A^{(j)} \mathcal{Y}_{D_i}^{(j)} - \left(\frac{1}{N} \sum_{j=1}^N \mathcal{Y}_A^{(j)} \right)^2}{\frac{1}{N} \sum_{j=1}^N \mathcal{Y}_A^{(j)} \mathcal{Y}_A^{(j)} - \left(\frac{1}{N} \sum_{j=1}^N \mathcal{Y}_A^{(j)} \right)^2} \quad (6.17)$$

Relation (6.17) can be obtained by substituting the Monte Carlo estimators,

$$V_{\theta_{\sim i}} [\mathbb{E}_{\theta_i}(Y|\theta_{\sim i})] \approx \frac{1}{N} \sum_{j=1}^N \mathcal{Y}_A^{(j)} \mathcal{Y}_{D_i}^{(j)} - \mathbb{E}^2(Y), \quad (6.18)$$

along with,

$$V_Y = \mathbb{E}(Y^2) - \mathbb{E}^2(Y) = \frac{1}{N} \sum_{j=1}^N \mathcal{Y}_A^{(j)} \mathcal{Y}_A^{(j)} - \left(\frac{1}{N} \sum_{j=1}^N \mathcal{Y}_A^{(j)} \right)^2, \quad (6.19)$$

into (6.12).

6.4.6 An Example: Coarse-Grained Models of Atomistic Systems

An example of the use of variance-based sensitivity analyses is found in comparison of parameter sensitivities for coarse-grained molecular models of atomistic systems in thermodynamic equilibrium.

We know that all matter in our universe is made up of essentially indivisible objects called *atoms* which make up the material universe in which we live. Atoms are characterized according to their number of electrons, atomic mass, number of neutrons, etc., and are tabulated in the well-known Periodic Table of modern chemistry and physics; there being some 118 distinct atomic structures currently known. In classical models of atomic systems discussed here, atoms are considered to be point masses in motion due to the action of forces, and their locations with respect to a fixed origin is given by the generalized coordinate vectors \mathbf{r} , $i = 1, 2, \dots, n$, with n being the number of atoms. We denote by \mathbf{p}_i , the momentum vector associated with i . The motion of such an atomistic system is governed by Hamilton's equations,

$$\dot{\mathbf{r}}_i = +\frac{\partial H}{\partial \mathbf{p}_i} \quad , \quad \dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i} \quad , \quad (6.20)$$

with $1 \leq i \leq n$, and the Hamiltonian H ,

$$H(\mathbf{r}, \mathbf{p}) = \sum_i \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{2m_i} + V(\mathbf{r}) \quad (6.21)$$

where m_i is the mass of particle i and $V(\mathbf{r}) = V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ is the force potential. A typical model for the potential $V(\cdot)$ (employed, for example, in the Molecular Dynamics code LAMMPS [85]) is of the form,

$$V(\mathbf{r}) = V_{bond}(\mathbf{r}) + V_{angle}(\mathbf{r}) + V_{dihedral}(\mathbf{r}) + V_{non-local}(\mathbf{r}) + V_{coulomb}(\mathbf{r}) \quad , \quad (6.22)$$

where

$$V_{bond}(\mathbf{r}) = \sum_{i=1}^{N_b} \frac{1}{2} k_{ri} (\mathbf{r}_i - \mathbf{r}_{0i})^2 \quad , \quad (6.23)$$

$$V_{angle}(\mathbf{r}) = \sum_{i=1}^{N_a} \frac{1}{2} k_{\theta i} (\theta_i - \theta_{0i})^2 \quad , \quad (6.24)$$

$$V_{dihedral}(\mathbf{r}) = \sum_{i=1}^{N_d} \sum_{n=1}^4 \frac{V_n i}{2} [1 + (-1)^{n-1} \cos(n\phi_i)] \quad , \quad (6.25)$$

$$V_{non-local}(\mathbf{r}) = \sum_{i=1}^{N_{nl}} \sum_{j>i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^\alpha - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^\beta \right], \quad r_{ij} \leq r_c, \quad (6.26)$$

$$V_{coulomb}(\mathbf{r}) = \sum_{i=1}^{N_q-1} \sum_{j>i}^{N_q} 4\epsilon_0 \left[\frac{q_i q_j}{r_{ij}} \right]. \quad (6.27)$$

Here V_{bond} represents, in general, covalent bonds between atoms with bonding constants k_{ri} , V_{angle} is the potential energy due to change in angles between successive bonding lengths between atoms, the dihedral angle ϕ_i between consecutive links, $V_{non-local}$ denotes Lennard-Jones type potentials, generally with $\alpha = 6$, $\beta = 12$ or perhaps $\alpha = 6$, $\beta = 9$, and $V_{coulomb}$ is the coulomb potential for discrete charges q_i and q_j at coordinates \mathbf{r}_i and \mathbf{r}_j , respectively, with $r_{ij} = \|\mathbf{r}_j - \mathbf{r}_i\|$. Here in (6.23) and (6.24), \mathbf{r}_{0i} and θ_{0i} are the original coordinate parameters assumed to be given in the initial configuration of the system, ϕ_i in (6.25) is a torsional degree of freedom, and $\|\cdot\|$ is the Euclidean norm.

In most molecular dynamics (MD) applications, the momentum vectors are equated to the Newtonian momentum, $\mathbf{p}_i = m_i \dot{\mathbf{r}}_i$, and (6.20) reduces to (generally huge) system of nonlinear second-order differential equations describing a Newtonian model for the AA system. Because of the enormous size and complexity of such systems, it is standard practice to construct *coarse-grained* (*CG*) approximations of the system, obtained by aggregating groups of atoms into N super atoms or molecules, thereby dramatically reducing the size of the model. This results in a large class \mathcal{M} of models with the atom coordinates replaced by molecular coordinates \mathbf{R}_I , $I = 1, 2, \dots, N \ll n$, and with a force potential of the same form, but with constants,

$$\boldsymbol{\theta} = \{K_{Ri}, R_{0i}, K_{\theta i}, V_{Ni}, \epsilon_{ii}, \sigma_{ij}\} , \quad 1 \leq i \leq N , \quad (6.28)$$

wherein we ignore coulomb-charge effects.

A major problem with this very classical and often used process of coarse graining is that a given AA configuration, which we assume is the *truth* to simplify the analysis, can be segregated into many possible CG configurations, as indicated in the map $G : AA \rightarrow CG$ illustrated in Figure 6.4, and for each choice of a CG model, then are numerous choices of force potential, each choice with different sets of parameters.

In Farrell et al [39], this problem was taken up for specific applications involving modeling the stretching of a polyethylene chain, in constant volume and temperature ($T = 100K$), consisting of 12 “beads” (molecules) with two carbon atoms per bead. The model involved 12 parameters, each assigned a uniform distribution:

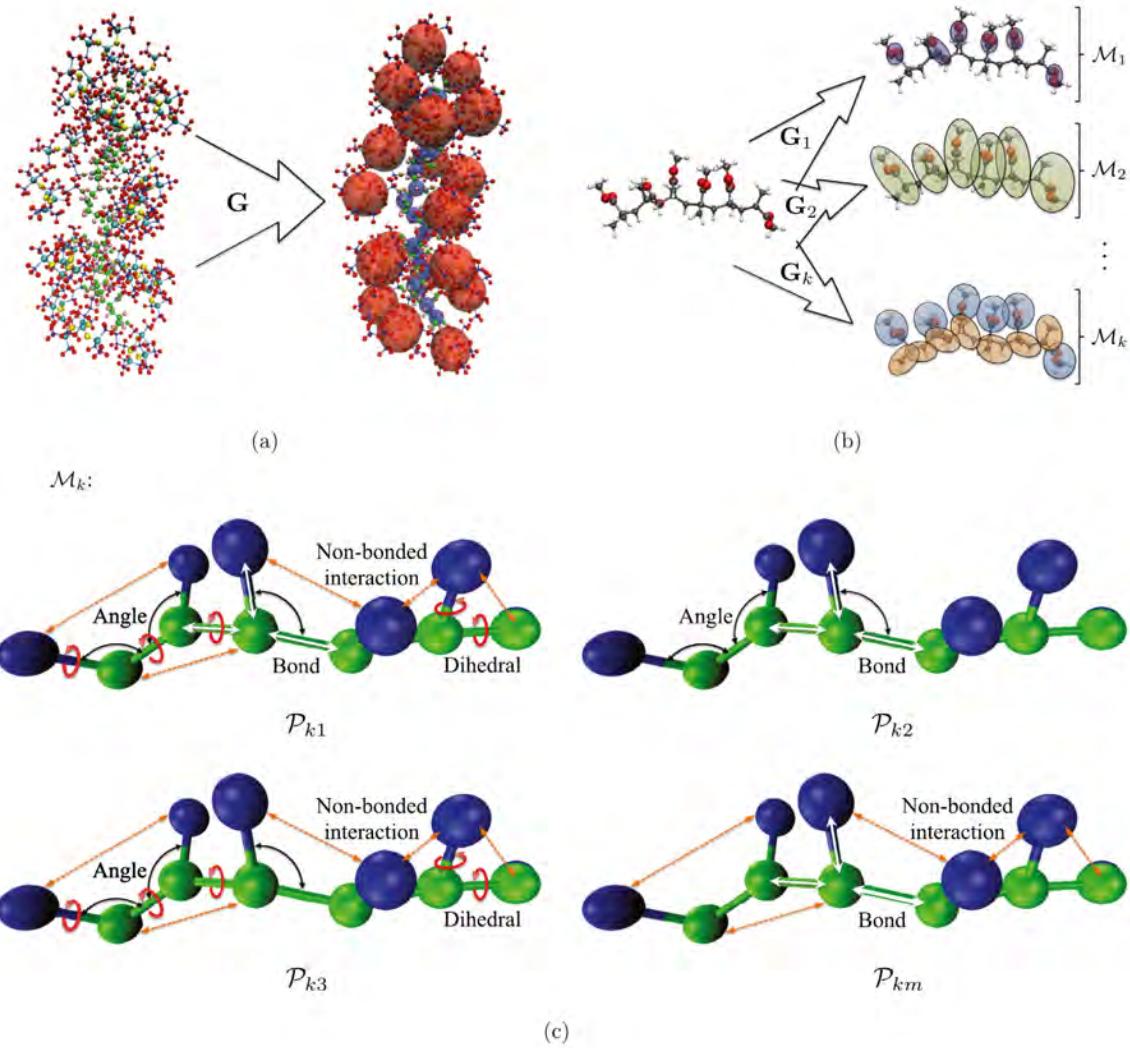


Figure 6.4: (a) Coarse-graining of an all-atom system; (b) different choices of AA-to-GC map G ; (c) different choices of intermolecular force fields V , for each map G . (From Farrell et al. (2015) [39])

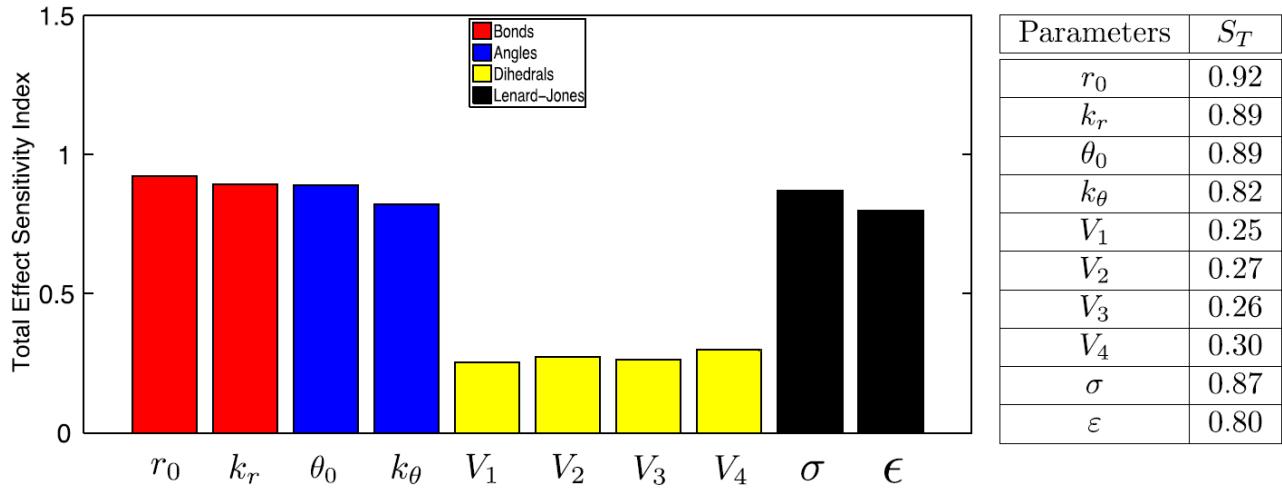


Figure 6.5: Bar charts illustrating the results of the sensitivity analysis using total sensitivity indices for the deformation of the polyethylene chain using 20,000 Monte Carlo samples of the parameter space. (From Farrell et al. (2015) [39])

$$R_0 \sim \mathcal{U}(0.1, 5) , \quad K_R \sim \mathcal{U}(0.1, 90)$$

$$\theta_0 \sim \mathcal{U}(80, 100) , \quad K_\theta \sim \mathcal{U}(0, 15)$$

$$V_1 \sim \mathcal{U}(-1, 1) , \quad V_2 \sim \mathcal{U}(-1, 1)$$

$$V_3 \sim \mathcal{U}(-1, 1) , \quad V_4 \sim \mathcal{U}(-1, 1)$$

$$\sigma \sim \mathcal{U}(0.05, 7) , \quad \epsilon \sim \mathcal{U}(0.01, 5)$$

The goal is to estimate the total sensitivity indices for each parameter with respect to the output function,

$$Y(\boldsymbol{\theta}) = \langle V(\boldsymbol{\theta}, G_1) \rangle - \langle V(\boldsymbol{\theta}, G_2) \rangle ,$$

where G_0 denotes the initial configuration of the chain and G is the map defining the configurations attained by applying a force to one end of the chain while keeping the other end constrained.

We use (6.12) to compute the total sensitivities of each of these 12 parameters. Results are presented in Figure 6.5 which is reproduced from that in [39]. Scatter plots of some parameters are given in Figure 6.6. From these results, we observe that the sensitivities of the dihedral parameters are much smaller than those of other parameters, and we conjecture that by omitting them altogether, the values of the output $Y(\theta)$ will not appreciably change. To confirm the

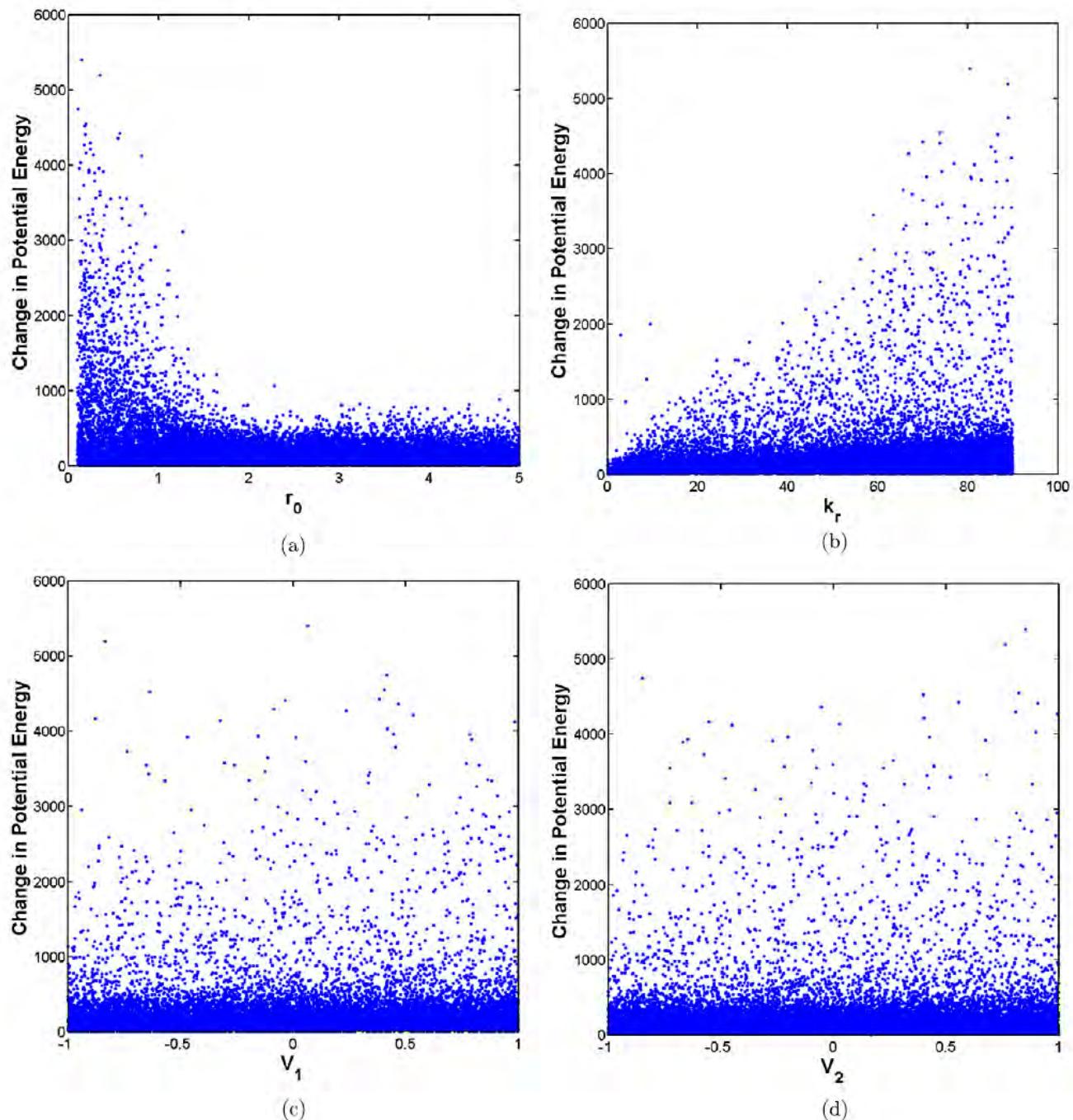


Figure 6.6: Scatterplots illustrating a quantitative sensitivity analysis for the polyethylene chain using 40,000 Monte Carlo samples of the parameter space. The observable is the change in potential energy as the chain is deformed. (From Farrell et al. (2015) [39])

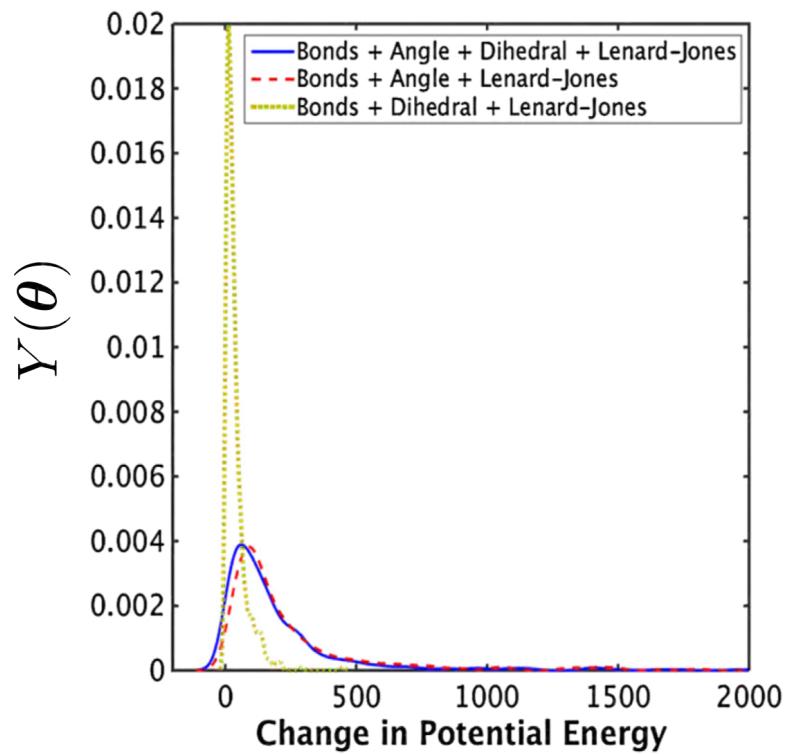


Figure 6.7: Comparison between the kernel density estimates of the change in potential energy resulting from the deformation of a polyethylene chain for the cases when dihedral and angle parameters are excluded from calculations. Results show 500 Monte Carlo samples of the parameter space. (From Farrell et al. (2015) [39])

conjecture, we compute the same output functions except that the dihedral parameters are omitted. Figure 6.7 shows the results of these different models: 1) the full model with covalent bonds, angles, dihedrals, and Lennard-Jones terms; 2) the model without dihedrals (i.e., only with bonds, angles, and LJ term; and 3) the model with bonds, dihedral, and LJ terms. We observe that by eliminating the parameters with lowest sensitivities (the dihedrals), the distributions of the output is essentially unchanged.

Chapter 7

Model Selection and OPAL, the Occam Plausibility ALgorithm

7.1 Introduction

At this point, we have considered Bayesian and Entropy-Information theoretic approaches to address several of the major sources of uncertainty, including \mathcal{P} -uncertainty in Chapters 2 and 3 and \mathcal{Y} and θ uncertainties in Chapters 4 and 5. What remains is one of the most important problems in predictive science: \mathcal{M} -uncertainty, the uncertainty in the selection of the centerpiece of prediction, the model used to make the prediction. For many years, the community of scientist and engineers were taught that model selection was the province of experienced practitioners who had earned the privilege of selecting mathematical models based on divine-given judgment, discounting that the models selected often delivered poor predictions of the observed events. Or more commonly, very classical models of mechanics and physics were morphed and distorted and ultimately used to model events far outside the original domains of the textbook models. In the meantime, the statistics world was concerned about model selection and methods to guide selection, but most of the time only in a limited sense of calculating likelihoods for regression calculations or low-dimensional approximations.

In more recent times, the theory and methodology of model selection has taken strides forward with the advent of modern computing capabilities and modern algorithms for implementing methods based on Bayesian inference and information theory. In this chapter, we describe two classes of methods for model selection, one based on Bayesian ideas involving the notion of posterior model plausibilities, and the other based on frequentist notions, information theory, and relative

entropy. According to Chow (1981) [28], the Bayesian approach is motivated by the work of Jeffreys (1939,1961,1998) [61–63] and revived in more recent work of Beck and co-workers (Beck and Yuen, 2004; Beck, 2010; Beck and Taflanidis, 2013 [16–18]) and others (Farrell et al., 2015; Prudencio and Schulz, 2012 [39, 89]) while the frequentist follows the work of Akaike (1974, 1976, 1977, 1998 [4–7]) and its various generalizations (Hurvich and Tsai, 1989; Lebreton et al., 1992; Schwartz, 1978 [56, 72, 101]). A readable account of these latter approaches is found in Burnham and Anderson (2002) [25] and in Konishi and Kitagawa (2008) [71].

In the sections that follow, we introduce the theory of posterior model plausibilities and we outline an introduction to the information-theoretic and frequentist methods of Akaike and others. We then describe a general framework for model selection and model validation embodied in OPAL – the Occam Plausibility ALgorithm – introduced and implemented in [35] and [79], which brings together most of the tools and concepts presented thus far: model calibration, validation, parameter sensitivity, model plausibility, model inadequacy, and prediction of the QoIs.

7.2 Bayesian Model Plausibilities

Let us suppose that we have in hand a set of m parametric models $\mathcal{P}_i(\boldsymbol{\theta}_i)$, $1 \leq i \leq m$, each with its own set of parameters $\boldsymbol{\theta}_i \in \Theta_i$, describing a manifestation of the forward problem (1.1) for a scenario S . Suppose further that we have acquired a set of observational data $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\} \in \mathcal{Y}_S$ and that for each model and parameter set and a prior $\pi(\boldsymbol{\theta}_j|S)$ is specified. The next step is to simply write down Bayes' rule (4.1), this time acknowledging that we actually have additional conditional information, namely that the model $\mathcal{P}_j(\boldsymbol{\theta}_j)$ is known and that it belongs to the set \mathcal{M} :

$$\pi(\boldsymbol{\theta}_j|\mathbf{y}, \mathcal{P}_j, \mathcal{M}, S) = \frac{\pi(\mathbf{y}|\boldsymbol{\theta}_j, \mathcal{P}_j, \mathcal{M}, S)}{\pi(\mathbf{y}|\mathcal{P}_j, \mathcal{M}, S)} \pi(\boldsymbol{\theta}_j|\mathcal{P}_j, \mathcal{M}, S), \quad 1 \leq j \leq m. \quad (7.1)$$

Here we have merely rewritten (4.1) acknowledging additional prior knowledge (and, thus, we have applied Jaynes' principle of taking into account “ all of the evidence we have and not just some arbitrary chosen subset of it.”)

The key term of interest in (7.1) is surprisingly the *evidence*, the denominator on the right-hand-side of (7.1). More than just a normalization factor, it is the marginalization of the numerator,

$$\pi(\mathbf{y}|\mathcal{P}_j, \mathcal{M}, S) = \int_{\Theta_j} \pi(\mathbf{y}|\boldsymbol{\theta}_j, \mathcal{P}_j, \mathcal{M}, S) \pi(\boldsymbol{\theta}_j|\mathcal{P}_j, \mathcal{M}, S) d\boldsymbol{\theta}_j, \quad 1 \leq j \leq m. \quad (7.2)$$

Examining the left-hand-side of (7.2), we observe that it may be interpreted as a new likelihood function for a discrete version of the Bayes rule over the set of m models \mathcal{M} [18, 28, 88]. We thus compute its posterior, for each model j , denoted ρ_j :

$$\rho_j = \pi(\mathcal{P}_j | \mathbf{y}, \mathcal{M}, S) = \frac{\pi(\mathbf{y} | \mathcal{P}_j, \mathcal{M}, S) \pi(\mathcal{P}_j | \mathcal{M}, S)}{\pi(\mathbf{y} | \mathcal{M}, S)}, \quad 1 \leq j \leq m. \quad (7.3)$$

The prior model plausibilities, $\pi(\mathcal{P}_j | \mathcal{M}, S)$, in the case of m equally plausible models, may be simply set equal to $1/m$. The denominator, $\pi(\mathbf{y} | \mathcal{M}, S)$ is treated as a normalizing factor.

The numbers ρ_j are the *posterior plausibilities* of the set \mathcal{M} of mathematical models and the densities $\pi(\mathcal{P}_j | \mathcal{M}, S)$ are the prior model densities. Choosing the denominator $\pi(\mathbf{y} | \mathcal{M}, S)$ so as to normalize the set of discrete probability components, we have

$$\sum_{j=1}^m \rho_j = 1. \quad (7.4)$$

The model (or models) in \mathcal{M} with plausibility (plausibilities) ρ_k ($\rho_k \geq \rho_j, 1 \leq j \leq m$) is deemed the most plausible in the set \mathcal{M} for the given data \mathbf{y} in scenario S .

7.2.1 Frequentist Approaches: The Akaike Information Criterion

Methods for model selection based on likelihood theory typical of frequentist arguments can be derived using the notion of information lost when a given model deviates from the “true” distribution of the QoI or other observables. These are the Akaike methods or their generalizations, initiated in the 1974 paper of [4] and expanded in subsequent work [5–7]. The idea behind these approaches begins with the notion of a true distribution $g = g(\mathbf{y}), \mathbf{y} \in \mathcal{Y}_S$, defining the target prediction, and a set \mathcal{M} of models $\{\mathcal{P}_1(\boldsymbol{\theta}_1), \mathcal{P}_2(\boldsymbol{\theta}_2), \dots, \mathcal{P}_m(\boldsymbol{\theta}_m)\}$ each defined by their respective likelihood probabilities $\pi(\mathbf{y} | \boldsymbol{\theta}_j, S), 1 \leq j \leq m$. The information lost $I_j(g, \pi)$ when the model \mathcal{P}_j is used to approximate the truth is given by the D_{KL} -divergence,

$$\begin{aligned} I_j(g, \pi) &= D_{KL}(g \parallel \pi(\cdot | \boldsymbol{\theta}_j, S)) \\ &= \int_{\mathcal{Y}_S} g(\mathbf{y}) \log \frac{g(\mathbf{y})}{\pi(\mathbf{y} | \boldsymbol{\theta}_j, S)} d\mathbf{y} \\ &\quad \left(\text{or } \sum_i^n g_i \log g_i / \pi_i \text{ for the discrete case} \right), \quad 1 \leq j \leq m. \end{aligned} \quad (7.5)$$

For each model \mathcal{P}_j , this loss is minimized by choosing the parameters $\boldsymbol{\theta}_j$ to be the MLE (recall

(4.11) $\hat{\boldsymbol{\theta}}(\mathbf{y}) = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmax}} \log \pi(\mathbf{y}|\boldsymbol{\theta}, S)$). Choosing such an MLE in (7.5) and averaging over the samples \mathbf{x} and \mathbf{y} in \mathcal{Y}_S , gives the number

$$\begin{aligned} A_j &= \mathbb{E}_{\mathbf{x}} \mathbb{E}_{\mathbf{y}} \left[\log(\pi(\mathbf{x}|\hat{\boldsymbol{\theta}}_j(\mathbf{y}), S)) \right] \\ &= \int_{\mathcal{Y}_S} \int_{\mathcal{Y}_S} g(\mathbf{x}) g(\mathbf{y}) \log \pi(\mathbf{x}|\hat{\boldsymbol{\theta}}_j(\mathbf{y}), S) d\mathbf{x} d\mathbf{y}, \quad 1 \leq j \leq m. \end{aligned} \quad (7.6)$$

The numbers A_j represent the information lost when model \mathcal{P}_j (namely $\pi(\mathbf{x}|\hat{\boldsymbol{\theta}}_j(\mathbf{y}), S)$) is used to approximate the truth g , \mathbf{x} and \mathbf{y} in (7.6) being independent random samples taken from the same distribution.

The idea of using the A_j , or computable approximations of them, as measures of model goodness was proposed in [4]. To reduce the calculation of measures such A_j to a practical means for model comparison, several simplifying but generally acceptable assumptions can be introduced. These include the assumption that the log-likelihood is twice continuously differentiable with respect to the parameters, that a truncated Taylor expansion of the log-likelihood can be used to approximate its variation over Θ , that a large enough sample $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ is taken that asymptotic properties valid as $n \rightarrow \infty$ hold, and that the sequence of MLEs $\hat{\boldsymbol{\theta}}_j(\mathbf{n})$ taken for sample size n converge to the D_{KL} -minimizer $\boldsymbol{\theta}_j^*$. Under such assumptions, the number A_j can be replaced by the computable *Akaike Information Criteria*,

$$AIC_j(\mathbf{y}) = -2 \log \left(\pi(\mathbf{y}|\hat{\boldsymbol{\theta}}_j, S) \right) + 2k_j, \quad 1 \leq j \leq m, \quad \mathbf{y} \in \mathcal{Y}_S \quad (7.7)$$

where k_j is the number of parameters in model \mathcal{P}_j , the best model in \mathcal{M} being that with the smallest $AIC_j(\mathbf{y})$.

Akaike Weights

In applications of the AIC, the actual AIC_j values are of little use; what one wishes to determine are the relative merits of various models as measured by the relative criteria,

$$\Delta_j = AIC_j - AIC_{\min} \quad , \quad 1 \leq j \leq m . \quad (7.8)$$

The best model among the set of all models has a relative value $\Delta^* = \Delta_{\min} = \min\{\Delta_j\}_{j=1}^m$. It is useful to transform the AIC relative values Δ_j into a model probability by setting them proportional to (minus one-half of) a log-likelihood: $\log \pi_{\text{like}_j} \propto (-\Delta_j/2)$. Then we can define the

model probabilities by the Akaike weights:

$$W_j = \frac{\exp\left(-\frac{1}{2}\Delta_j\right)}{\sum_{k=1}^m \exp\left(-\frac{1}{2}\Delta_k\right)}. \quad (7.9)$$

Clearly,

$$\sum_{k=1}^m W_k = 1. \quad (7.10)$$

The weight W_j is clearly the weight of evidence in favor of \mathcal{P} , and the highest corresponds to the best model.

Generalizations of the AIC have been proposed by several investigators. For example, the “second-order” AIC for small sample size has been proposed by [56],

$$\begin{aligned} AIC_{cj}(\mathbf{y}) &= -2 \log (\pi(\mathbf{y}|\boldsymbol{\theta}_j, S)) \\ &+ 2k_j + \frac{2k_j(1+k_j)}{n-k_j-1} \\ &= AIC_j(\mathbf{y}) + \frac{2k_j(1+k_j)}{n-k_j+1}, \end{aligned} \quad (7.11)$$

n being the sample size. Still another measure has been proposed by [72] for case in which data is “overdispersed.” Then, instead of $AIC_j(\mathbf{y})$ of (7.7), one uses

$$QAIC_j(\mathbf{y}) = -2 \log \left[\pi(\mathbf{y}|\hat{\boldsymbol{\theta}}, S)/\hat{C} \right] + 2k_j \quad (7.12)$$

where \hat{C} is the ratio of the chi-squared statistic χ^2 and its number of degrees of freedom. A small-sample version of this estimator, $QAIC_{cj}(\mathbf{y})$, analogous to (7.11), can be obtained by replacing $AIC_j(\mathbf{y})$ in (7.11) with $QAIC_j(\mathbf{y})$. See [25] for additional details.

7.3 The OPAL Algorithm: Adaptive Model Selection and Model Validation

The Occam Plausibility Algorithm (OPAL), introduced in [39] and described in [79] (from which the present account is extracted), provides a systematic adaptive approach to statistical model calibration and validation that brings together several of the methodologies discussed earlier while

also providing an approach to address and resolve the elusive problem of model inadequacy. The name is understandably derived from Occam's Razor, the “principle” that one should use the “simplest” theory or model among competing models that leads to the same prediction. In OPAL, the simplest model in a class of possible models is defined as that with the fewest parameters, but other measures of simplicity or complexity could be used.

The algorithm involves the following steps.

1. **Initialization.** A set \mathcal{M} of possible models is identified for predicting the QoI:

$$\mathcal{M} = \{\mathcal{P}_1(\boldsymbol{\theta}_1), \mathcal{P}_2(\boldsymbol{\theta}_2), \dots, \mathcal{P}_m(\boldsymbol{\theta}_m)\}, \quad (7.13)$$

each model having its own parameter space ($\boldsymbol{\theta}_k \in \Theta_k$). The models in \mathcal{M} may be closely related, some sharing the same parameters.

2. **Sensitivity Analysis.** Parameter sensitivities in each model class $\mathcal{P}(\boldsymbol{\theta}_j)$ are computed using techniques such as those described in Chapter 6 for the prediction scenario S_p . Using an appropriate tolerance, only those model classes are retained that contain parameters to which the output is sensitive, as indicated by the relative value of the sensitivity measures; models containing parameters to which the output is insensitive are eliminated, resulting in a reduced set $\bar{\mathcal{M}}$ of model classes,

$$\bar{\mathcal{M}} = \{\bar{\mathcal{P}}_1(\boldsymbol{\theta}_1), \bar{\mathcal{P}}_2(\boldsymbol{\theta}_2), \dots, \bar{\mathcal{P}}_l(\boldsymbol{\theta}_l)\}, \quad l \leq m. \quad (7.14)$$

3. **Occam Categories.** The models in $\bar{\mathcal{M}}$ are next partitioned according to the number of parameters in each. These are referred to as the *Occam categories*, with the models containing the smallest number of parameters designated category 1, those in the next higher number category 2, and so forth. We therefore produce a collection of subsets, $\bar{\mathcal{M}}_1, \bar{\mathcal{M}}_2, \dots, \bar{\mathcal{M}}_N$, with models in set $\bar{\mathcal{M}}_k$ containing exactly k parameters, $1 \leq k \leq N$. By our interpretation of Occam's Razor, models of category k are simpler or less complex than those of category $l > k$.
4. **Calibration.** All of the models in the category 1 set $\bar{\mathcal{M}}_1$ are next calibrated using Bayes rule (4.1) for a specified calibration scenario S_c resulting in calibrated Category 1 models,

$$\bar{\mathcal{M}}_1^* \sim \{\mathcal{P}_1^*(\boldsymbol{\theta}_1^*), \mathcal{P}_2^*(\boldsymbol{\theta}_2^*), \dots, \mathcal{P}_k^*(\boldsymbol{\theta}_k^*)\}, \quad k < l. \quad (7.15)$$

5. **Plausibility Calculation.** We next calculate Bayesian posterior plausibilities ρ_j^* , $1 \leq j \leq k$, of the Category 1 model in (7.15) (recall (7.3)). The most plausible model (or models) is

identified as model \mathcal{P}_j^* such that

$$\rho_j^* \geq \rho_i^*, \quad 1 \leq i \leq k. \quad (7.16)$$

6. **Validation Step.** With the most plausible and simplest model now identified, we move to a validation scenario S_v , update the parameters using Bayes' rule (4.1), and solve the forward problem in S_v to determine the model predictions of the S_v experimental observations. We emphasize that the design of the validation experiments is of paramount importance in that the accuracy with which the model agrees with validation observational data should indicate the accuracy with which the QoI is predicted in S_p .

Let $q_v(\mathbf{y})$ be the probability density of the validation observational data $\mathbf{y} = \mathbf{y}_v$ and $\pi_v(\mathbf{y}, \boldsymbol{\theta}^{**})$ be the model-generated density approximating q_v obtained using the validation-updated parameters, denoted $\boldsymbol{\theta}^{**}$. Many different metrics can be used at this point to measure the accuracy with which $\pi_v(\mathbf{y}, \boldsymbol{\theta}^{**})$ approximates q_v . Also, for a given metric, an accuracy tolerance $\gamma = \gamma_{\text{tol}}$ must be specified to assess if the distance between predictions and measurements is small enough to deem the model valid. A natural way to compare predictions and data in the presence of uncertainties is to employ the D_{KL} -divergence, in which case the model is said to be valid if

$$D_{KL}(q_v \parallel \pi_v(\mathbf{y}, \boldsymbol{\theta}^{**})) \leq \gamma_{\text{tol}}, \quad (7.17)$$

with $\mathbf{y} = \mathbf{y}_v$ and,

$$D_{KL}(q_v \parallel \pi_v(\mathbf{y}, \boldsymbol{\theta}^{**})) = \int_{S_v} q_v(\mathbf{y}) \log \frac{q_v(\mathbf{y})}{\pi_v(\mathbf{y}, \boldsymbol{\theta}^{**})} d\mathbf{y}. \quad (7.18)$$

It is sometimes more meaningful to use the normalized measure, D_{KL}/H , H being the Shannon entropy, so that a valid model is such that $D_{KL}/H \leq \gamma_{\text{tol}}$. We emphasize that other measures (“metrics”) could be used.

7. **Iteration or Solution.** If the model is valid, e.g., if (7.17) holds, we use the validated model in scenario S_p with parameter $\boldsymbol{\theta}^{**}$ and solve the forward problem for the QoI. If no model in this category is valid, we return to step 1, now setting $\mathcal{M} = \bar{\mathcal{M}}_2$, the models of Category 2, and repeat steps 2 through 5 until a valid model is determined. The process is terminated when one model is identified that passes the validation criterion such as (7.17). If no models in the original set \mathcal{M} are valid, the modeler must enlarge and enrich the set of models until(if possible) a set containing a valid model can be identified.

Several remarks are in order:

Remark 1 A flow chart indicating the path followed in implementing OPAL is given in Figure 7.1.

Remark 2 In each calculation in the calibration step, Step 4, a prior pdf must be identified. These can often be determined using the maximum entropy based methods discussed in chapter 5.

Remark 3 Several tolerances must be set in implementing OPAL. These include a tolerance to judge which parameters can be ignored on the basis of the level of sensitivity of the QoI exhibits to their changes. Also, the tolerance γ_{tol} defining a valid model for given measures of accuracy of the validation prediction must be specified.

Remark 4 As a cautionary note, even though a sensitivity measure of a parameter may be small, the posterior values may be ‘correlated’ with others so that its influence on the QoI may be negligible. Then we may assign a constant, deterministic value to that parameter and then retain the corresponding model in the set \mathcal{M} .

Remark 5 The plausibilities are computed for the calibration scenario as the calculation of plausibilities for validation scenarios is rarely feasible owing to their computational complexity. In some cases, however, it is conceivable that validation data could be used to determine model plausibilities. Then model validity may be clearly connected to model plausibility for data \mathbf{y}_v in S_v .

Remark 6 A frequentist version of OPAL can easily be envisioned in which Bayesian updates are replaced with MLE calculations and Bayesian plausibilities are replaced by model selection methods based on such measures as the Akaike Information Criterion.

Remark 7 As noted earlier, the use of the number of parameters of a model as a measure of its simplicity in the sense of Occam’s razor, is only one of several approaches to delineate Occam categories. Other measures, such as computational complexity, operation counts of calibration and validation algorithms, number of model dependent variables, etc., could be used as a basis for partitioning \mathcal{M} into categories.

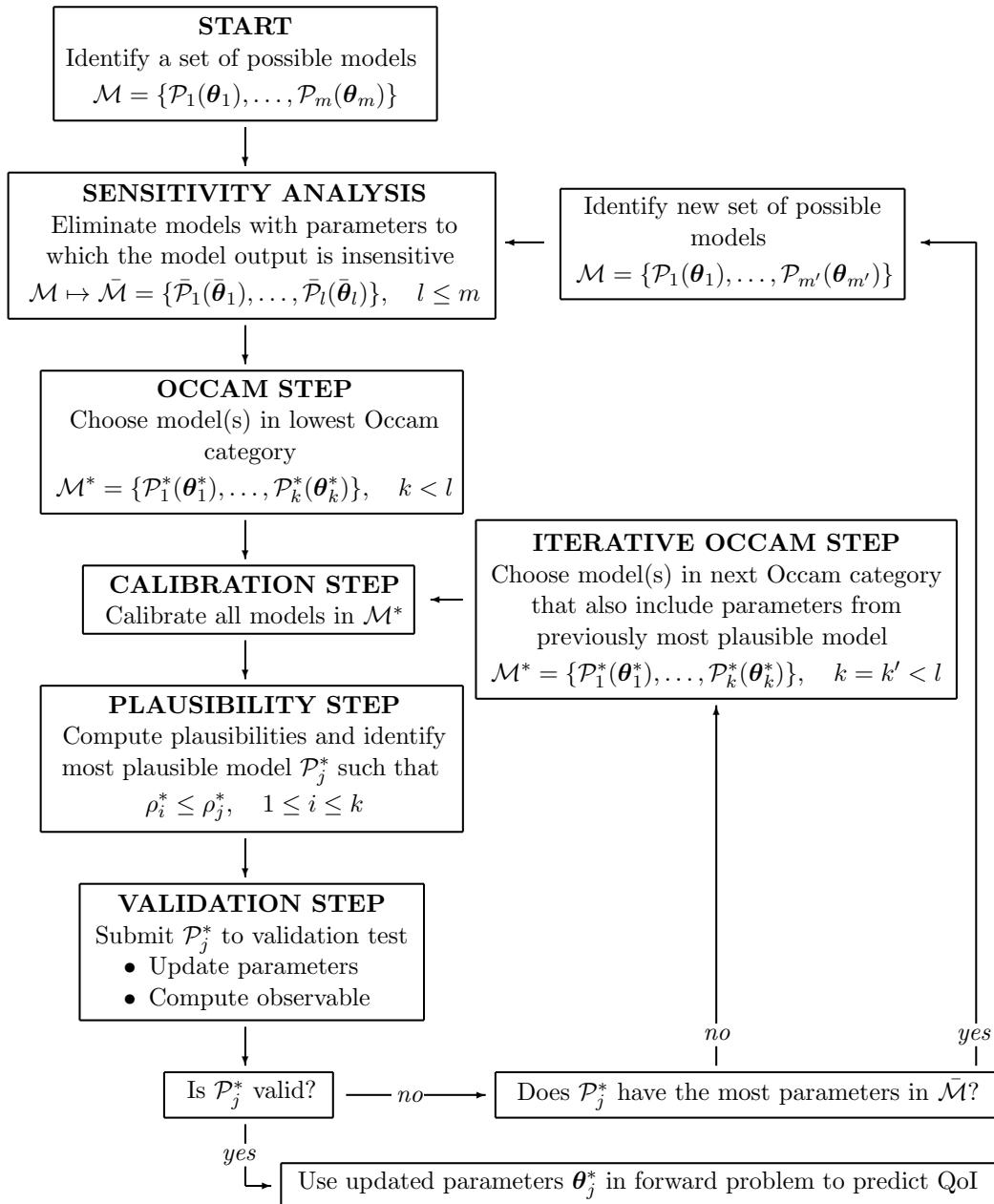


Figure 7.1: Flowchart for OPAL, the Occam-Plausibility Algorithm [39, 79].

Chapter 8

Introduction to Cell and Cancer Biology

8.1 Introduction

One of the most challenging and important areas of predictive science is predictive medicine, particularly the prediction of the emergence and growth of cancer in living organisms and the prediction of the outcomes of various therapies. In Chapters 9 and 10 ahead of us, we will describe applications of the theory and methods discussed in the chapters preceding this one to the prediction of the growth or decline of cancer in living organisms, including methods and philosophy underlying the development of multiscale models of tumor growth.

To set the stage for these developments, it is important to review basic and even elementary facts about cell biology, mainly to record basic definitions, to review the ideas of the cell cycle, protein synthesis, sub-cellular signaling pathways and the cellular-level events that are thought to lead to cancer. This is the goal of this chapter. We begin with an introductory discussion of DNA, and travel through the remarkable phenomena of protein synthesis and cell cycles, to an introductory discussion of the emergence and growth of cancer.

8.2 DNA: Deoxyribonucleic Acid

Within the nuclei of the cells of living organisms there resides DNA, or *deoxyribose nucleic acid*, a long molecule consisting of two single-stranded molecule chains intertwined to form a double helix. The DNA molecule provides the genetic code that governs the growth, function, reproduction, and development of almost all living organisms. To understand the structure and the fundamental role of DNA, a review of some aspects of elementary chemistry is warranted.

The basic building block of a DNA strand is the so-called *deoxyribose sugar unit*, a molecule consisting of five carbon atoms and one oxygen atom, covalently bonded as shown Figure 8.1. A covalent bond is a generally strong intermolecular bond in which one or more pairs of electrons are shared by atoms to complete their outer shell. The sugar units are called *nucleotides*; they are attached to a sequence of phosphorous-based molecules, called the *phosphodiester linkage* or *phosphate molecules*, consisting of a phosphorous atom and four oxygen atoms; that forms the backbone of the strand, beginning with a number 3 carbon in the sugar unit and terminating with a number 5 carbon (or vice versa). This molecule is shown in Figure 8.2.

Chromosomes

Each DNA molecule in a human cell, when stretched out, would be $\sim 2 \times 10^{-19}$ meters wide and 2 meters long, about 200,000 times longer than the typical size length-scale of a cell. Yet, in each human cell, the DNA is packed into a single tiny structure called a *chromosome*, containing millions of base pairs. The DNA is itself packed around proteins called *histones* which are twisted into a packing structure, ultimately forming the chromosomes. Histones may be thought of as spools around which DNA strands are twisted to allow a secondary packing structure within the primary structure that forms the chromosome. These are depicted in Figure 8.3.

Every human has 46 chromosomes, 23 from each parent, arranged in 22 pairs of a matching chromosomes and a pair of sex chromosomes determining sex: the X and Y chromosomes; XX for female, XY for male.

The collection of all genes in the 46 chromosomes is called the human *genome*.

The phosphate backbone connects the sugar units into repeated molecules called *nucleotides*, shown in Figure 8.4, which expose the radicals to inner connections to so-called *base-acid molecules*. For each DNA molecule, two such long strands are intertwined into a helix connected by four of these base acids denoted *T, A, C, G*: *T* for *thymine*, *A* for *adenine*, *C* for *cytosine*, and *G* for *guanine*. The chemical compositions of these molecules is shown in Figure 8.5 and the helical strands are shown in Figure 8.6 and 8.7. The base pairing is strongest between *A* and *T*, and *C* and *G*; i.e. *A* complements *T*, and *C* complements *G*. The *A* and *G* molecules form receptors for the *T* and *C* molecules, the former are called *pyrimides* and the latter are called *purines*. Thus, *A* can only pair with *T* and *G* can only pair with *C* as shown symbolically in Figure 8.8.

The result of these pairings is the two nucleotide chains, oriented in opposite directions, as shown in Figure 8.6, and famously called the **DNA double helix**. The chemical composition of the helix is illustrated in Figure 8.7. The base acids are connected on the interior of the helix (*A* to *T*, *C* to

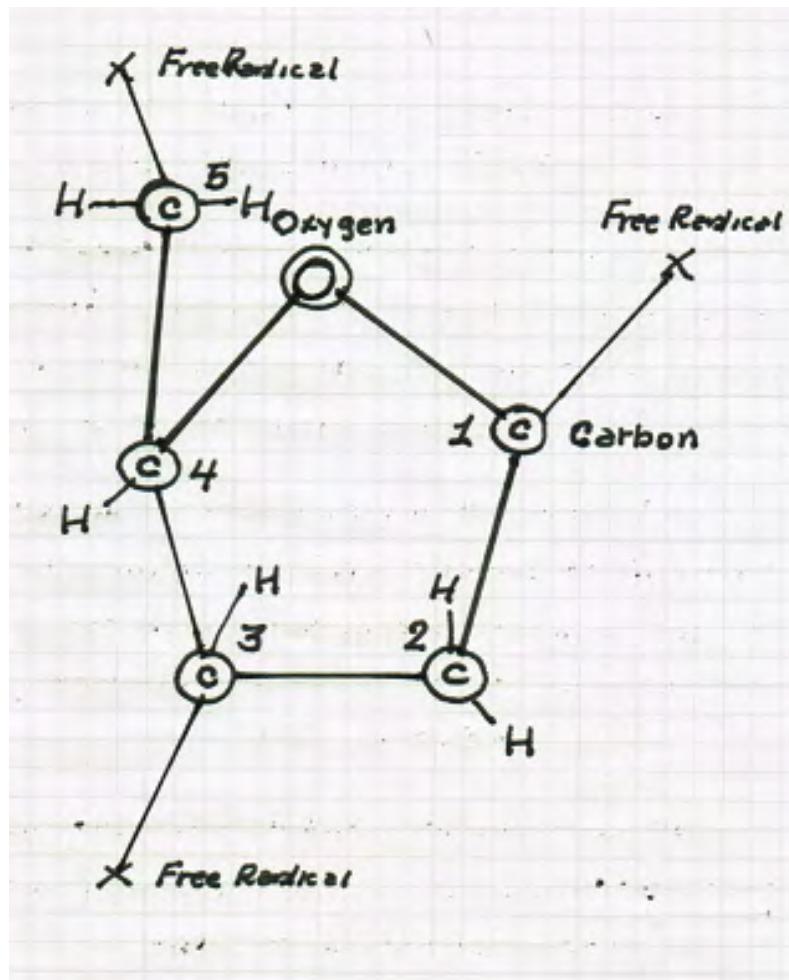


Figure 8.1: Deoxyribose Sugar Unit. A molecule consisting of an oxygen atom, five carbon atoms numbered 1-5, side hydrogen atoms, with the carbons connected through covalent bonds. Three free radicals are indicated by the symbol X. Oxygen, with atomic number 8, has 6 electrons in its outer orbital: Carbon, with atomic number 6, has 4 outer-orbital electrons. The 3- and 5 free radicals connect the unit to the phosphodiester linkage, or backbone. The inside free radical provides a site for connecting the unit to the base nucleic acids (A,G,T,C).

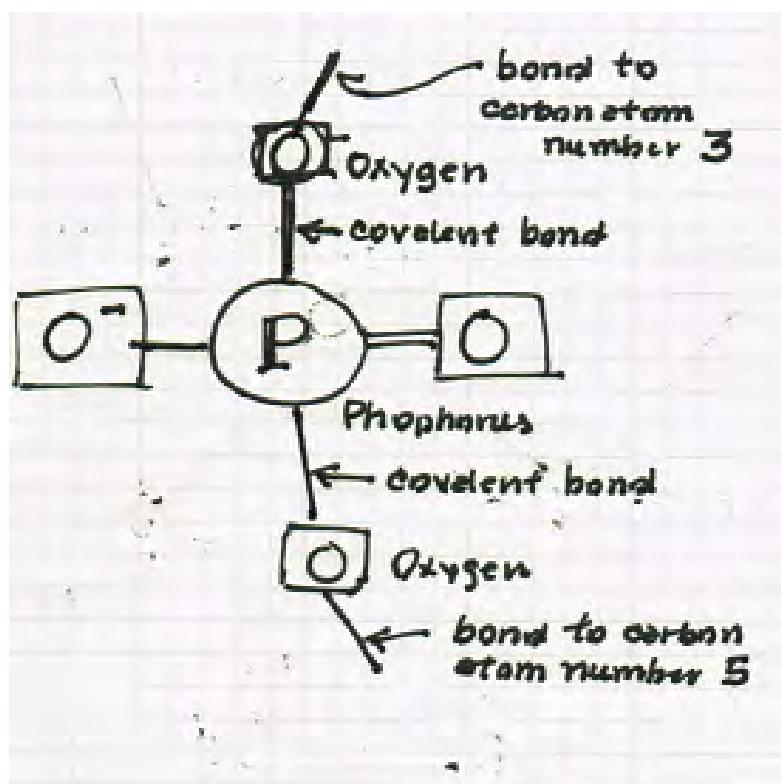


Figure 8.2: The Phosphodiester Linkage. The phosphodiester linkage or backbone of a DNA strand, consists of a phosphorus atom P covalently bonded to oxygen atoms, which are bonded to carbon atoms in the sugar unit. Phosphorus, with atomic number 15, is highly reactive, with 5 electrons in its outer orbital.

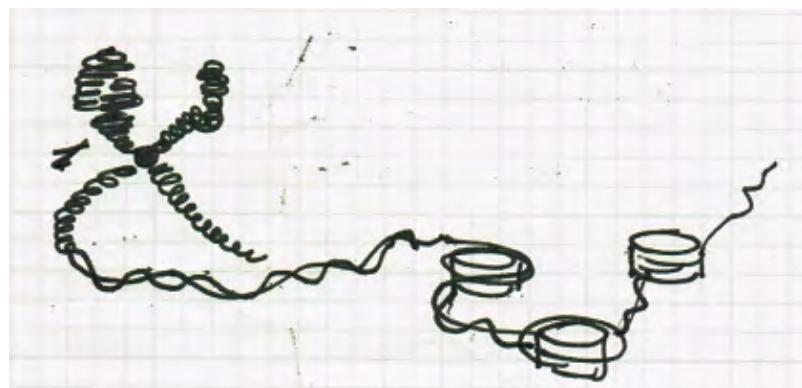


Figure 8.3: A chromosome consisting of tightly packed DNA strands, with center called a centromere, spread out to reveal a secondary packing structure in which strands are wrapped around protein “spools” called histones.

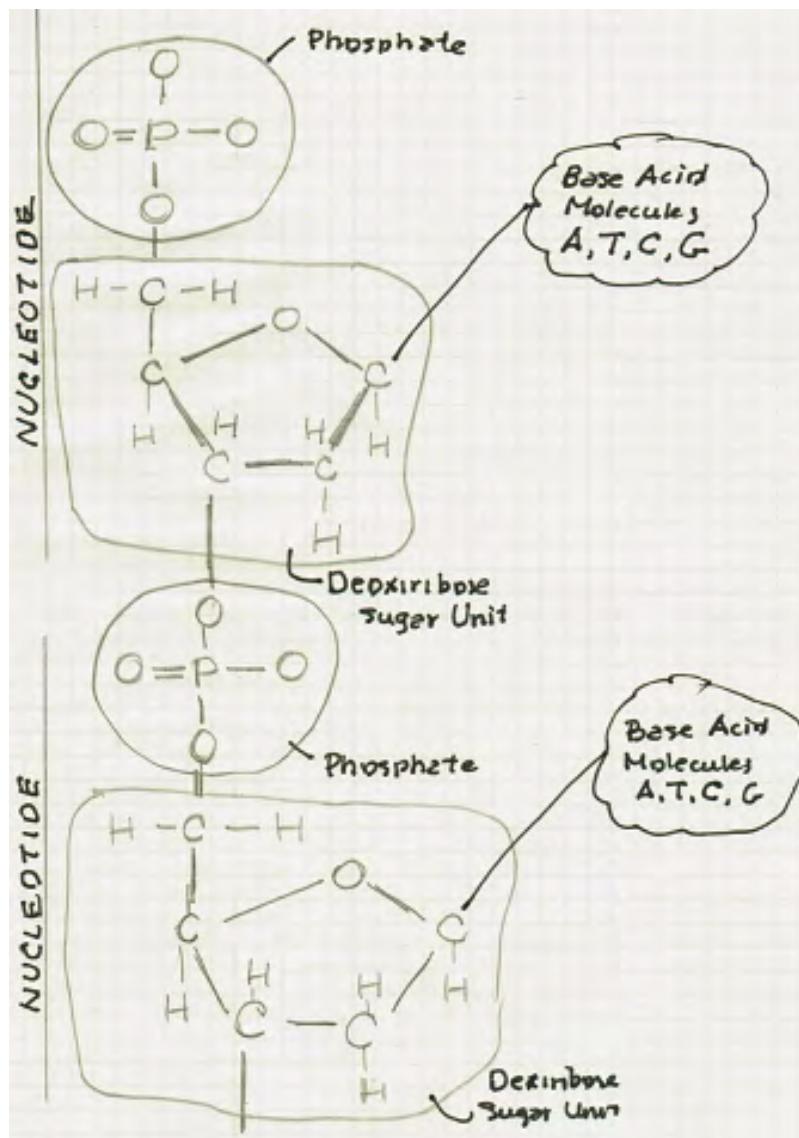


Figure 8.4: Segment of strand of DNA.

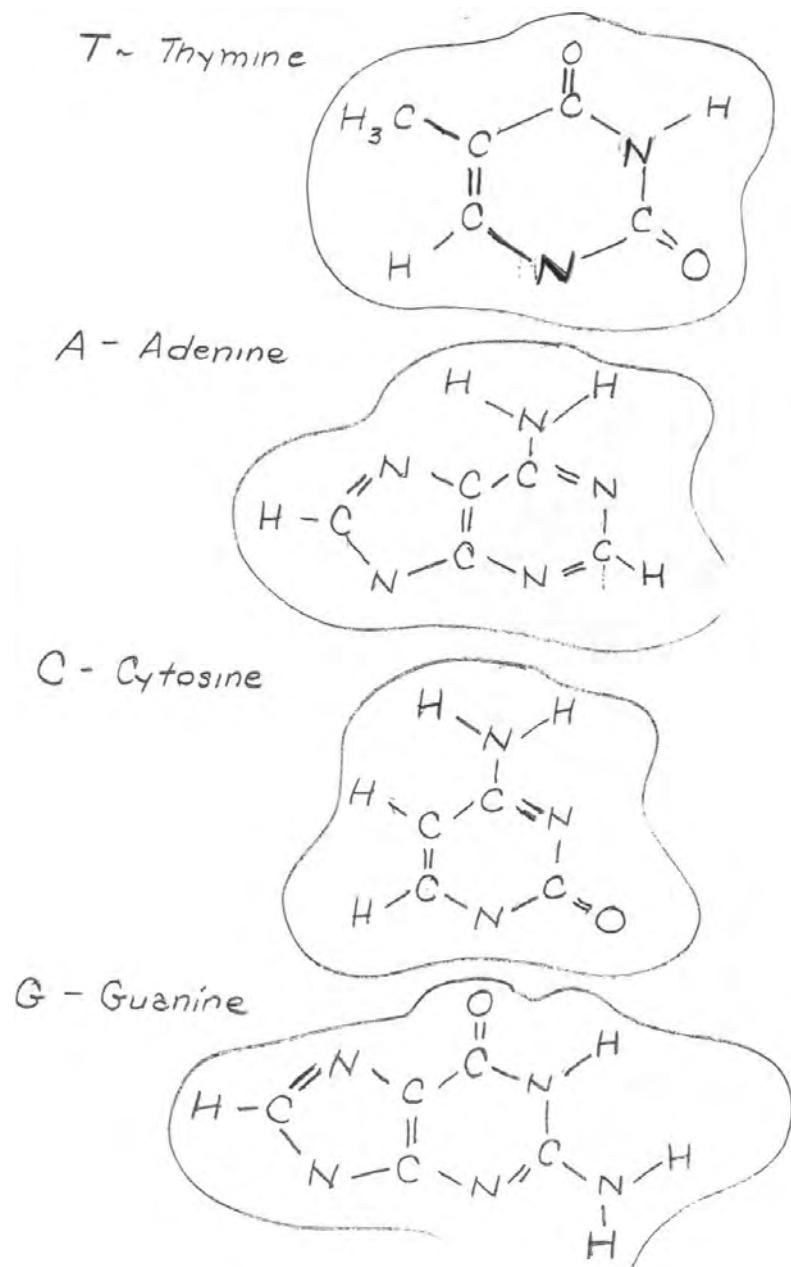


Figure 8.5: The four base nucleotides.

G, T to A, G to C) by hydrogen bonds, relatively weak molecular bonds illustrated in Figure 8.9.

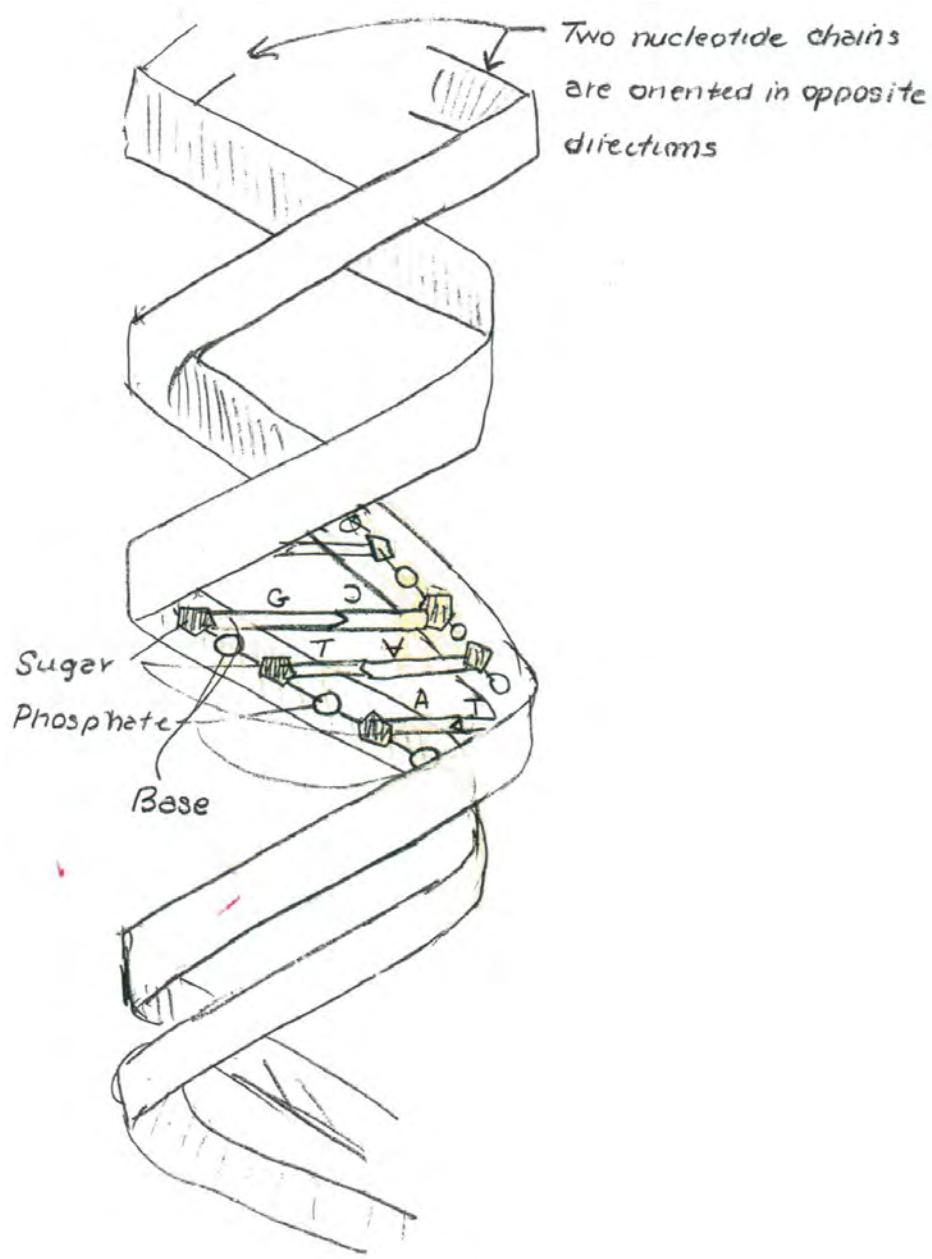


Figure 8.6: The DNA double helix.

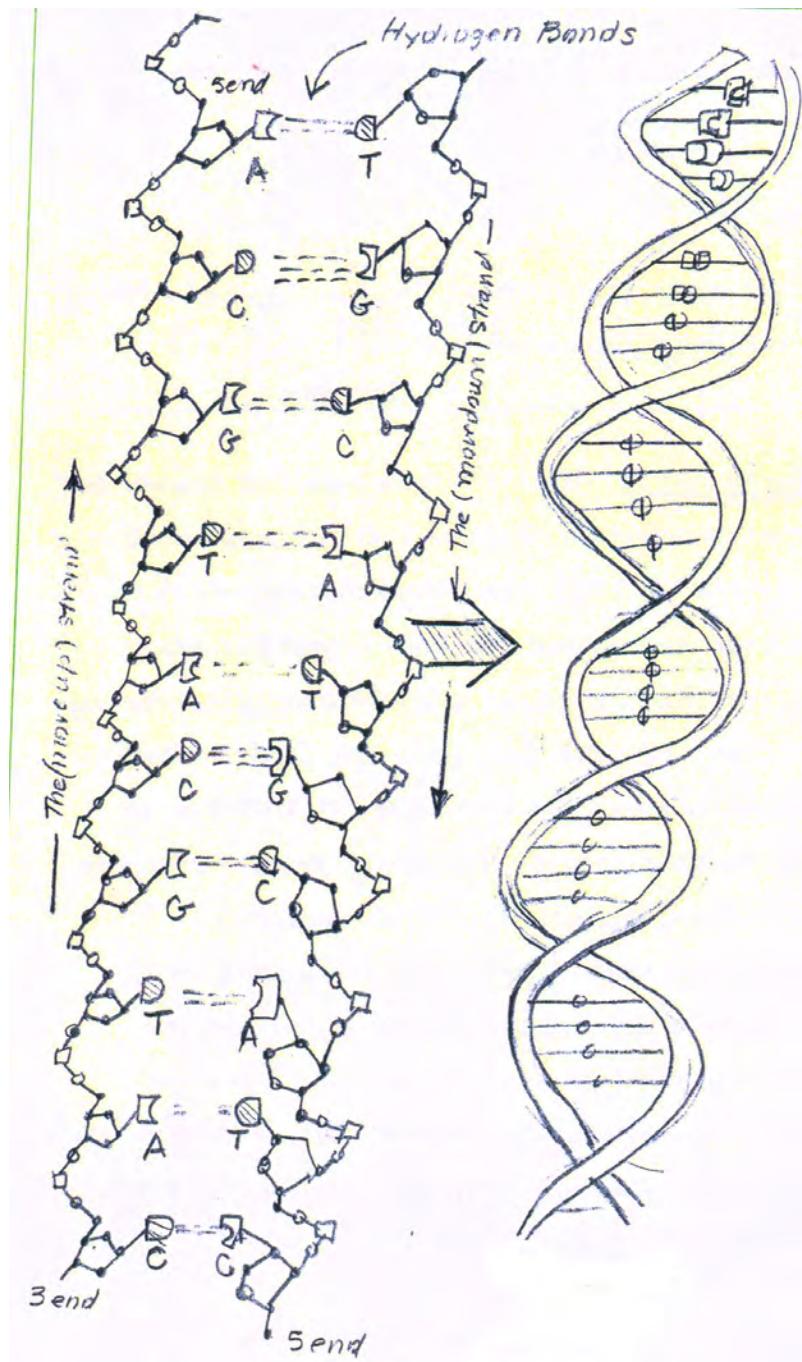


Figure 8.7: Summary of the chemical composition of the double helix.

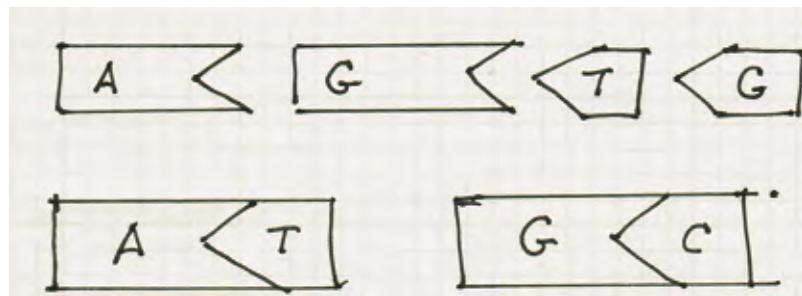


Figure 8.8: Acid base pair and hydrogen bonds inside the double helix.

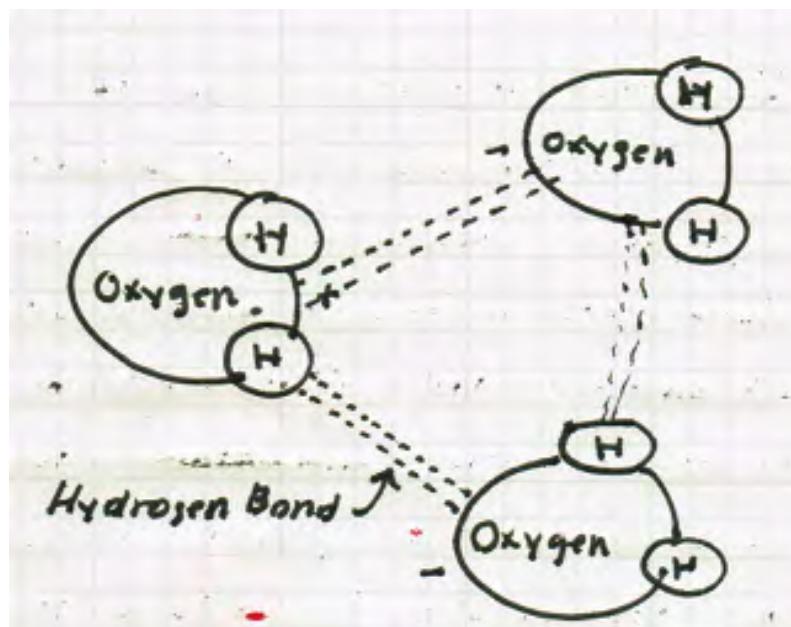


Figure 8.9: Hydrogen bonds. In covalently bonded molecules, such as H_2O , electrons shared by covalent bonds congregate away from the hydrogen atoms leaving a small positive charge away from the hydrogen atom and a small negative charge around the donor (e.g. oxygen or nitrogen). The negatively charged protein of one molecule is attracted to the positively charged end, creating a hydrogen bond.

8.3 The Cell

The *cell* is the basic biological unit of all known living organisms; the smallest unit of life that can replicate itself independently. The cell consists of a gel-like substance, called the *cytoplasm* (as the material of which it is comprised, is called *cytosol*) all enclosed in a thin bilayer cell membrane, with internal structures called *organelles*. The cytoplasm is often colorless and about 80% water. Inside the cell is the *cell nucleus*, which is an organelle also enclosed in a membrane, called *nuclear membrane*. A cell with a well-defined nucleus that is separated from the cytoplasm by its membrane is called a *eukaryote*. The chromosomes reside in the cell nucleus (See figure 8.10).

Through a remarkable process, to be now described, the information imbedded in these base-pair sequences is translated into the creation of new molecules, called *proteins*, through a multi-step process. The sequence is thus said to provide a *code* for manufacturing certain specific proteins that perform fundamental functions in the process of life in organisms.

The segment of the DNA that provides code for a single protein is called a *gene*. The human genome contains 50,000 to 100,000 genes dispersed through the 46 chromosomes. This transcription of base-pair information requires that the DNA unravel itself to release the coded information and then heal itself to replicate again the parent DNA molecule.

Replication of DNA

For the DNA molecule to replicate, three steps are required:

1. *Unwinding*: Special *enzymes* break the hydrogen bonds and the DNA strands unwind, that is, bases become unpaired, a process called *cleavage* of the DNA. (Figure 8.13)
2. *Pairing*: Free nucleotides pair with free basis on separated strands.
3. *Joining*: Complementary bases pair with covalent bonds linking them together: *polymerization* (aided by enzyme polymerase).

8.4 RNA: Ribose Nucleic Acid

In the nucleus of a cell, another molecule emerges during the replication process that, like DNA, consists of a long chain of nucleotides linked by sugar phosphate bonds. These are key differences between these molecular chains and DNA:

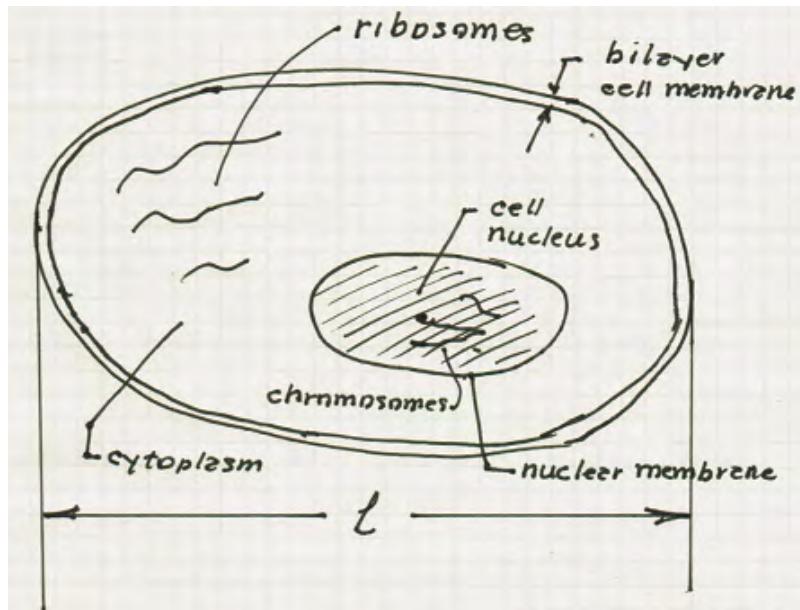


Figure 8.10: A sketch of the eukaryotic cell and its major components. The length l is, on average, around 10 microns, or 10×10^{-6} meters, which is 10,000 nanometers. There are about $10^{13} - 10^{14}$ cells in a human, and around 350 different types of cells.

1. The sugar molecule contains one more oxygen atom than the deoxyribose molecule of DNA, and is simply called *ribose* (Figure 8.11).
2. Because the sugar molecule differs from deoxyribose, the base nucleotide is different; the *T*-base Thymine of DNA is replaced by a new base *U* for *Uracil*.
3. The chain consists of only one strand instead of two.
4. These chains are much shorter than DNA.

These chains are called RNA molecules – the **Ribose Nucleotide Acid** (Figure 8.12).

The coding region of a gene is transcribed into a copy of the DNA sequence called *RNA*, a molecule similar to a single strand of DNA. The information “enclosed” in the DNA molecule enables the cell to *synthesize proteins*.

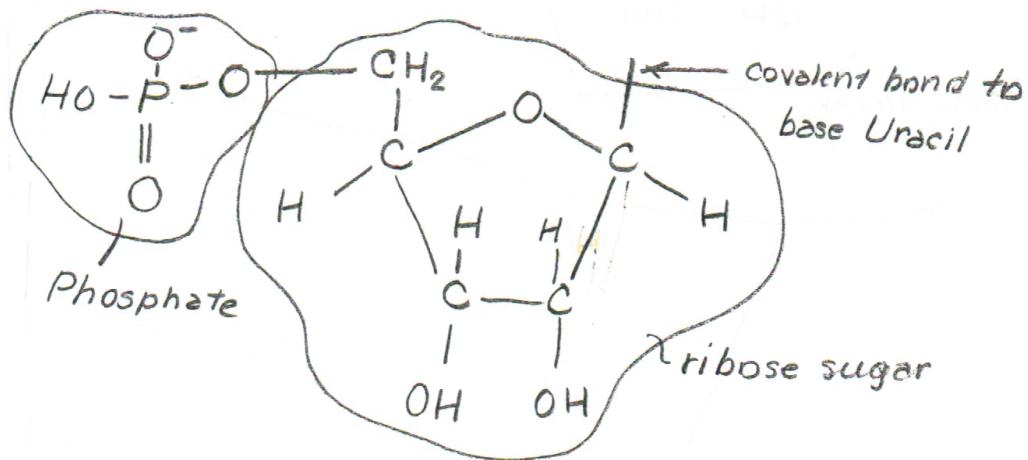


Figure 8.11: Ribose, a single strand with a phosphate backbone and a ribose sugar element.

8.5 Transcription

As the DNA unravels during replication, the information it carries decoded and translated into the single-strand RNA molecule. This process is called *transcription*. The basic steps are as follows:

1. An enzyme *RNA polymerase* joins the ribonucleotides together to form the RNA strand. Only a part of the DNA is transcribed to an RNA strand (unlike replication, where all of the DNA is copied).

The “length” of DNA transcribed into an RNA molecule is defined as a **gene**.
2. One gene can be transcribed to one RNA strand and another gene can be transcribed to another strand, and a single gene may be transcribed thousands of times.
3. The newly made RNA moves out of the nucleus and into the cytoplasm of the cell. This RNA is then called a *messenger RNA* or mRNA.
4. A set of three adjacent bases in an mRNA molecule is called a *codon*. Since there are 4 bases, there exist $4^3 = 64$ ways to combine 4 bases into a 3 letter codon. In the protein synthesis stage of the cell cycle, each of these 64 codons forms one of 20 acid molecules, called *amino acids*, plus 3 so-called *stop codons*. Several different codons may specify the same amino acid. The 64 resulting amino acids and their corresponding three-base codons are displaced in Table 8.1.

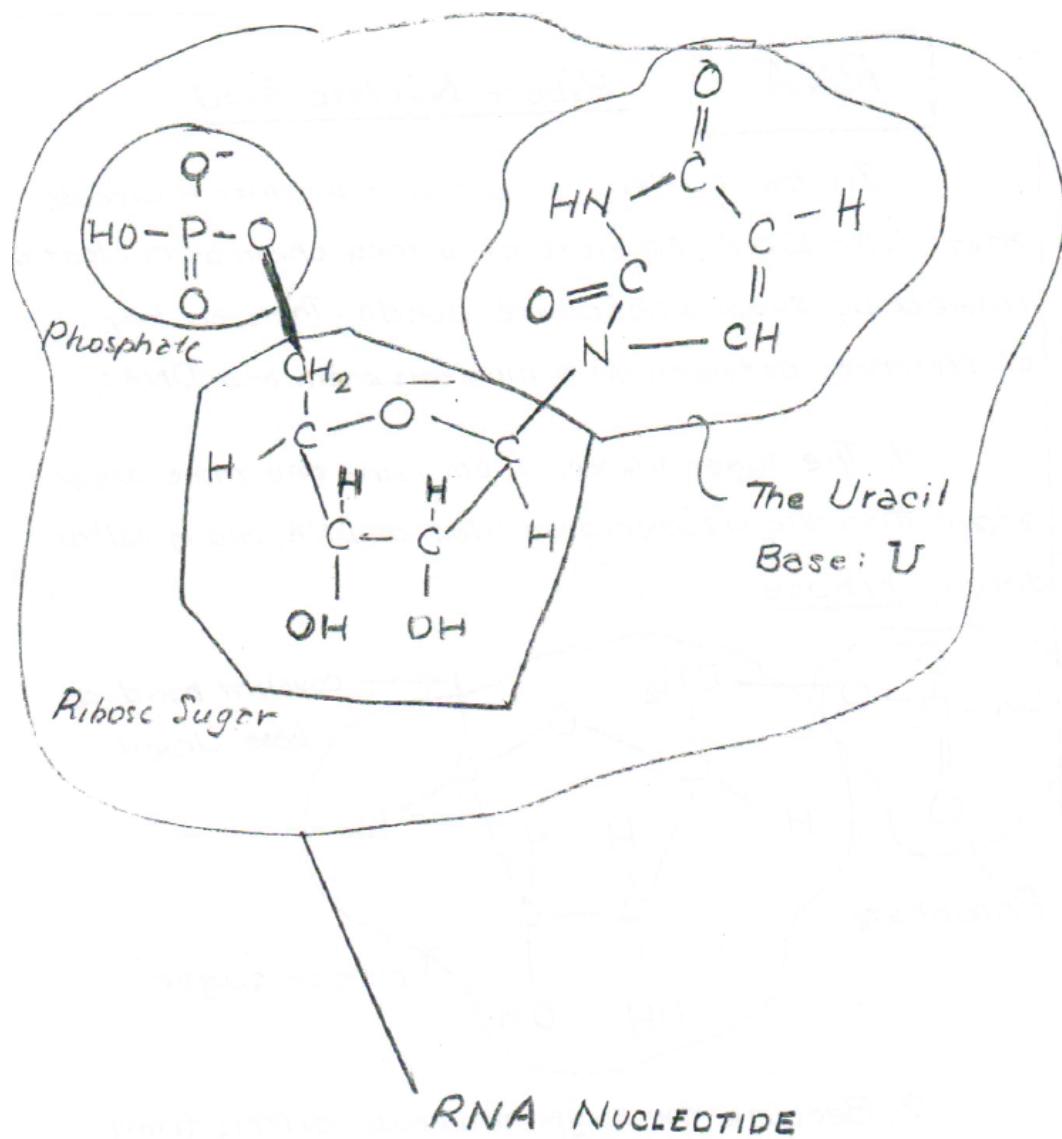


Figure 8.12: The RNA nucleotide molecule, consisting of the base Uracil *U*, and the ribose sugar unit.

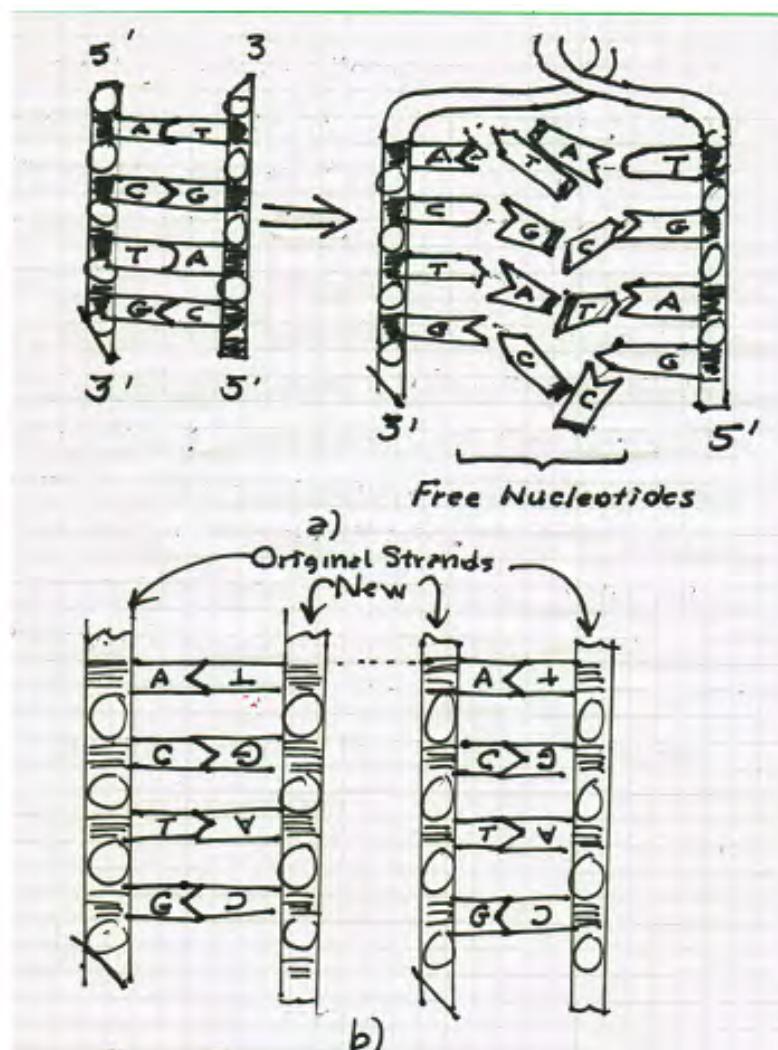


Figure 8.13: The process of DNA replication: a) unwinding b) pairing, and c) polymerization, producing an exact replica of the DNA molecule.

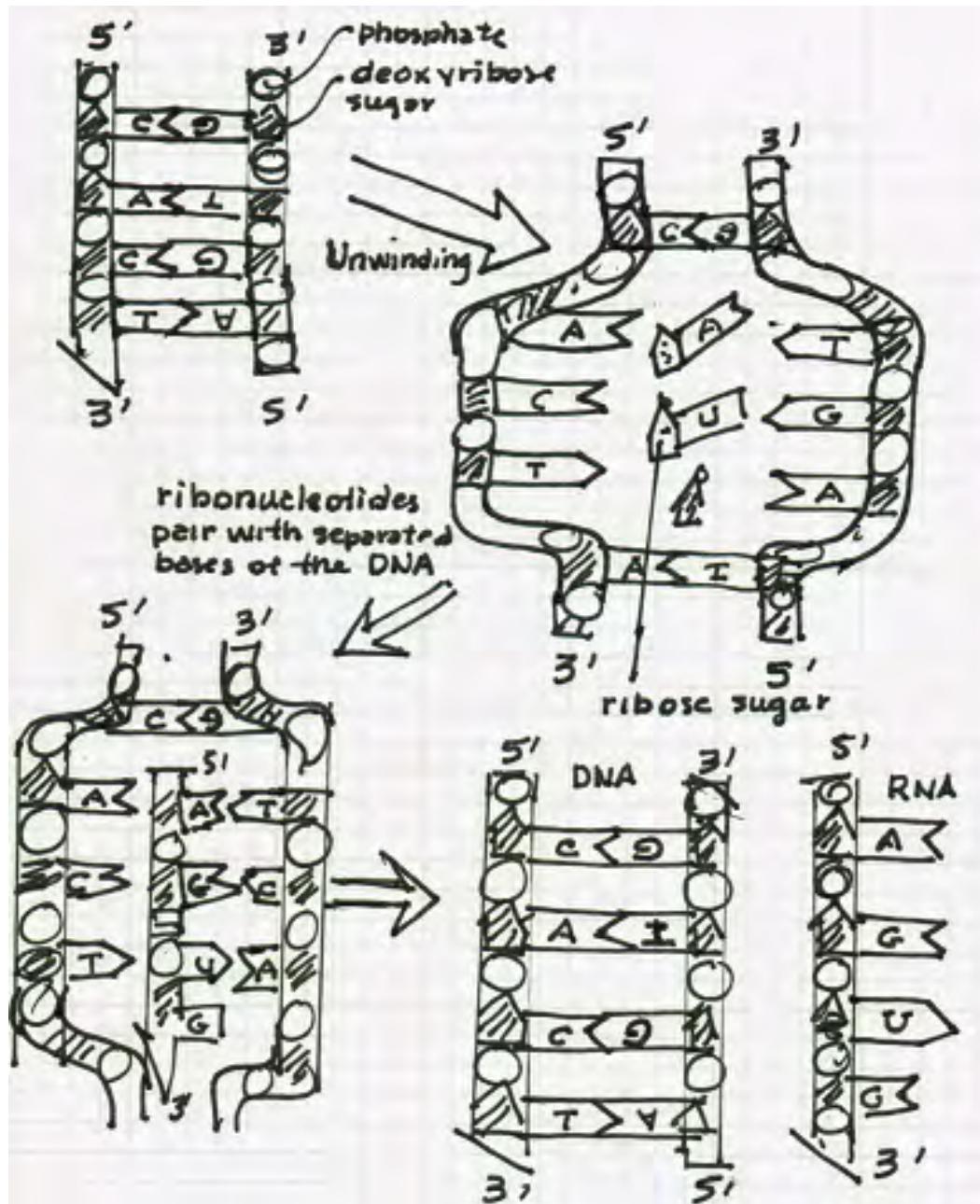


Figure 8.14: DNA Transcription: During unwinding of the DNA helix, so-called ribonucleotides pair with separated DNA bases creating a new row of bases linked to a single RNA strand, leaving as a result, the perfectly preserved DNA and a new RNA molecule.

Table 8.1: The Standard RNA Genetic Code

Codon	Amino Acid	Codon	Amino Acid	Codon	Amino Acid	Codon	Amino Acid
UUU	PHE	CUU	LEU	AUU	ILE	GUU	VAL
UUC	PHE	CUC	LEU	AUC	ILE	GUC	VAL
UUA	LEU	CUA	LEU	AUA	ILE	GUA	VAL
UUG	LEU	CUG	LEU	AUG	MET/ START	GUG	VAL
UCU	SER	CCU	PRO	ACU	THR	GCU	ALA
UCC	SER	CCC	PRO	ACC	THR	GCC	ALA
UCA	SER	CCA	PRO	ACA	THR	GCA	ALA
UCG	SER	CCG	PRO	ACG	THR	GCG	ALA
UAU	TYR	CAU	HIS	AAU	ASN	GAU	ASP
UAC	TYR	CAC	HIS	AAC	ASN	GAC	ASP
UAA	STOP	CAA	GLN	AAA	LYS	GAA	GLU
UAG	STOP	CAG	GLN	AAG	LYS	GAG	GLU
UGU	CYS	CGU	ARG	AGU	SER	GGU	GLY
UGC	CYS	CGC	ARG	AGC	SER	GGC	GLY
UGA	STOP	CGA	ARG	AGA	ARG	GGA	GLY
UGG	TRP	CGG	ARG	AGG	ARG	GGG	GLY

According to Table 8.1, only *one codon* codes for each of the following amino acids: MET (START amino acid), TRP, *two codons* code for each of the amino acids: ASN, ASP, CYS, GLN, GLU, HIS, LYS, PHE, TYR, *three codons* code for amino acid ILE, and three are STOP (nonsense) codons, *four codons* code for each of the amino acids: ALA, GLY, PRO, THR, VAL. There is no amino acid with **five codons**.*Six codons* code for each of the amino acids: ARG, LEU, SER.

8.6 Protein Synthesis Steps

The cell is essentially a protein factory in which information encoded in the DNA is transcribed to the RNA molecules and transmitted to the messenger RNA (the mRNA) which carries the coded information into the cytoplasm of the cell. This is facilitated by a conglomerate of molecules called *ribosome* RNA, which consists of approximately four types of molecules wrapped up by approximately 75 proteins to form a bead-like structure. Each Ribosome consists of a large and small unit separated by a groove through which an mRNA is threaded. The Ribosome slides down the mRNA, and as it moves it joins the amino acid molecules one-by-one to form a polypeptide chain (a cluster of ribosomes may move down an mRNA making multiple copies of the same polypeptide).

The entire synthesis process is facilitated by what is called a *transfer RNAs* (tRNA), single strands of RNA that translate the mRNA message into amino acids and then transfer those amino acids into a growing protein chain (Figure 8.17). Generally each tRNA is around 75 bases, loops back on itself, and each tRNA corresponds to a different amino acid and has a different base sequence.

The sequence of steps in the protein manufacturing process is as follows:

1. An initiator tRNA molecule charged with the (start) amino acid *Methionine* (Met) binds to the small ribosome unit and the mRNA and the small and large ribosome units come together. The anti-codon of the tRNA binds to the *AUG* start codon on the mRNA (Figure 8.19).
2. A second tRNA molecule forms base pairs between its anti-codon and the next codon, thus aligning the first and second amino acids .
3. The bond between the first tRNA is broken and, being uncharged is ejected from the message and is recycled, and the ribosome moves down the mRNA one codon (Figure 8.20).
4. The tRNA from the second amino acid is ejected. At the end of the message, a stop codon is reached (such as *UAG*), a codon for which there is no tRNA, and a termination factor binds in place of tRNA and stops the process (Figure 8.21).
5. The ribosomes separate, release the mRNA, and the polypeptide chain is separated from the last tRNA molecule (Figure 8.22).

Regulatory Sequence and Cell Differentiation

In addition to sequences that code protein construction, there are genes that code *regulatory sequences* and that control the rate and frequency of transcription. Thus, not all genes are “expressed”. There are 50,000-100,000 genes, $10^{13} - 10^{14}$ cells, and ~ 350 different types of cells. *Cell differentiation* means that cells are specialized in structure to perform some specific function. The traits of different types of cells are passed on to their progeny.

In what is called an *epigenetic change* within a genome, the gene expression is modified without changing the information content of DNA. The *stem cell* is a specialized cell with a role of renewing lost cells or repairing damaged cells while maintaining their own population. The role of stem cells is proliferation (unlike other cells that rarely divide). The evolution of stem cell differentiation is illustrated in Figure 8.15

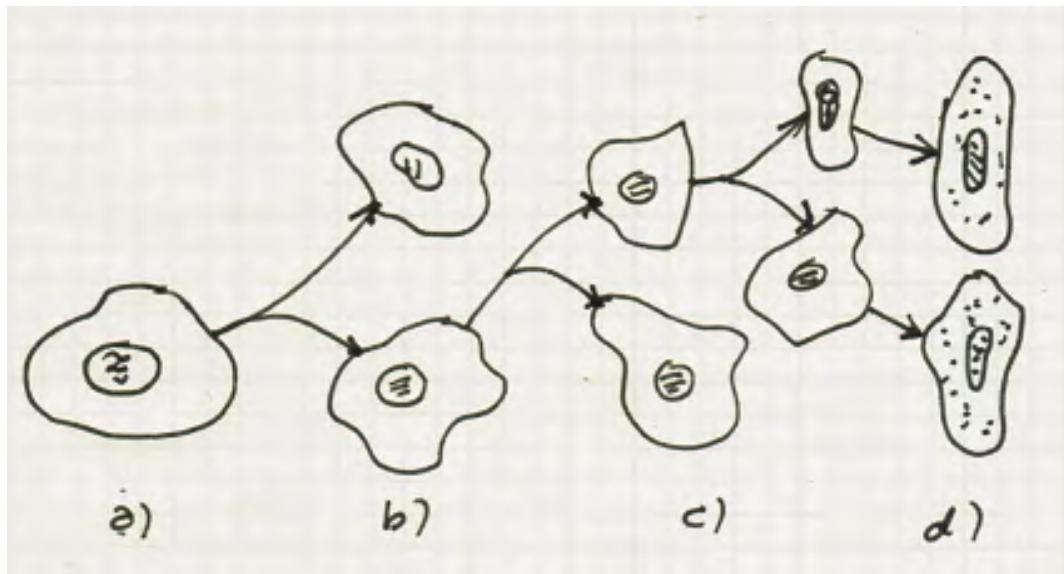


Figure 8.15: Evolution of stem cell differentiation: a) a stem cell in an initial state; b) cell division (mitosis) in which a partially differentiated stem cell is created; c) cell generations; and d) a process of macrophage in which terminally differentiated cells are created.

8.7 Signal Transduction

Individual cells are emerged in a heterogeneous array of constituents that exist outside the cell membrane, collectively called the extracellular matrix or extracellular region. The entire life and behavior of the cell – cell cycle, cell death (called *apoptosis*), gene regulation, protein synthesis, ... – is orchestrated by sequences of molecular signals that penetrate the membrane and propagate within the cell via chemical reactions of protein complexes. This transmission of molecular signals is called *signal transduction*, and the sequence of protein molecules reacting to create these signals constitute a *signaling pathway*.

The cell membrane contains special structures on its extracellular surface or inside the membrane surface called *receptors* which accept extracellular stimuli and which trigger the biochemical chain of events inside the cell. This chain of reactions is called a *signaling cascade*. The response to the signaling cascade may be an alteration of the cell's shape, metabolism, gene expression, or cell cycle and division. The signaling propagates through complex networks of molecules and can be amplified or terminated at any time, and can generate responses from hundreds to millions of molecules.

Signal transduction involves binding of extracellular signaling molecules called *ligands* to receptors that create inner-cellular events. All of these events are manifestations of changes in the structure of particular cell proteins. The binding site may also hold ligands next to molecules to facilitate chemical or catalyzing a chemical reaction between the two molecules. Enzymes are common protein catalysts.

The extracellular receptors are proteins that span across the cell membrane, one part within and the rest outside of the membrane. The transduction occurs as a result of a ligand binding to the outside region of the receptor that produces a change in the conformation of the interior portion of the receptor. The intercellular proteins that react to a ligand-receptor interaction demonstrate an enzymatic response, generally adhering to the laws of chemical kinetics, the enzymes being covalently bonded to the receptor. Chemically, the enzymes include so-called tyrosine kinase and phosphatases. Some create second level messengers that enable such events as releasing calcium into the cytoplasm and which activate enzymes and so-called *adaptor proteins*, the function of which is to link protein-binding partners together (and so to “adapt”), thereby creating larger signaling complexes. The protein kinases, an enzyme that modifies other proteins by adding phosphate groups to them, and *G*-protein, also an enzyme, are recruited in the signal transduction to supply short-lived active complexes in response to an activating signal, such as a growth factor (generally a protein that stimulates cell growth, proliferations, or healing) binding to its receptor.

Within the cell, intercellular receptors are in the form of soluble proteins exist in localized regions. To initiate transduction, the ligand diffuses through the nuclear membrane, into the cell nucleus, to possibly alter gene expression. These so-called nuclear receptors attach to the DNA at sites in what is referred to as the promoter region of genes, and can induce (promote) gene expression. Typical ligands for nuclear receptors are non-polar signaling molecules called *hormones*, that can be directly secreted into the blood stream and carried to the organs and tissues of the body. The effects of such hormones may be realized in the body after a long period of time. The transfer and covalent attachment of a phosphate group to a target protein is called **phosphorylation** (a common means for activating or inhibiting the function of a protein), and the enzymes that catalyze phosphorylation are called **kinases**, which play a regulatory role, turning target proteins functionally on or off.

Cellular responses to extracellular stimulation such as gene activation and alterations of metabolism, require signal transduction. Gene activation promotes other effects within the cell, such as increased access to glucose from the blood stream.

Signal transduction pathways are regarded the key to biological processes within mammals, and many diseases can be attributed to disturbances of such pathways, which cell growth also depends on. For example, the epidermal growth factor (EGF) is a protein that binds with the epidermal growth factor receptor (EGFR), which causes the two monomers to react to form a macromolecular complex and a phosphorylation of the EGFR, activating the intercellular pathway. A schematic of the transduction environment is given in Figure 8.16.

Among major signaling pathways, one can list the ERK pathway (that couples intercellular responses to binding growth factors, often promoting cell division), the CAMP pathway (that works by activating kinase A and a protein PKA), and the IP_3/DAG pathway: PLC that cleaves the molecules phospholipid phosphatidylinositol 4; 5 biophosphate producing diacyl glycerol (DAG) that can lead to altered cellular behavior.

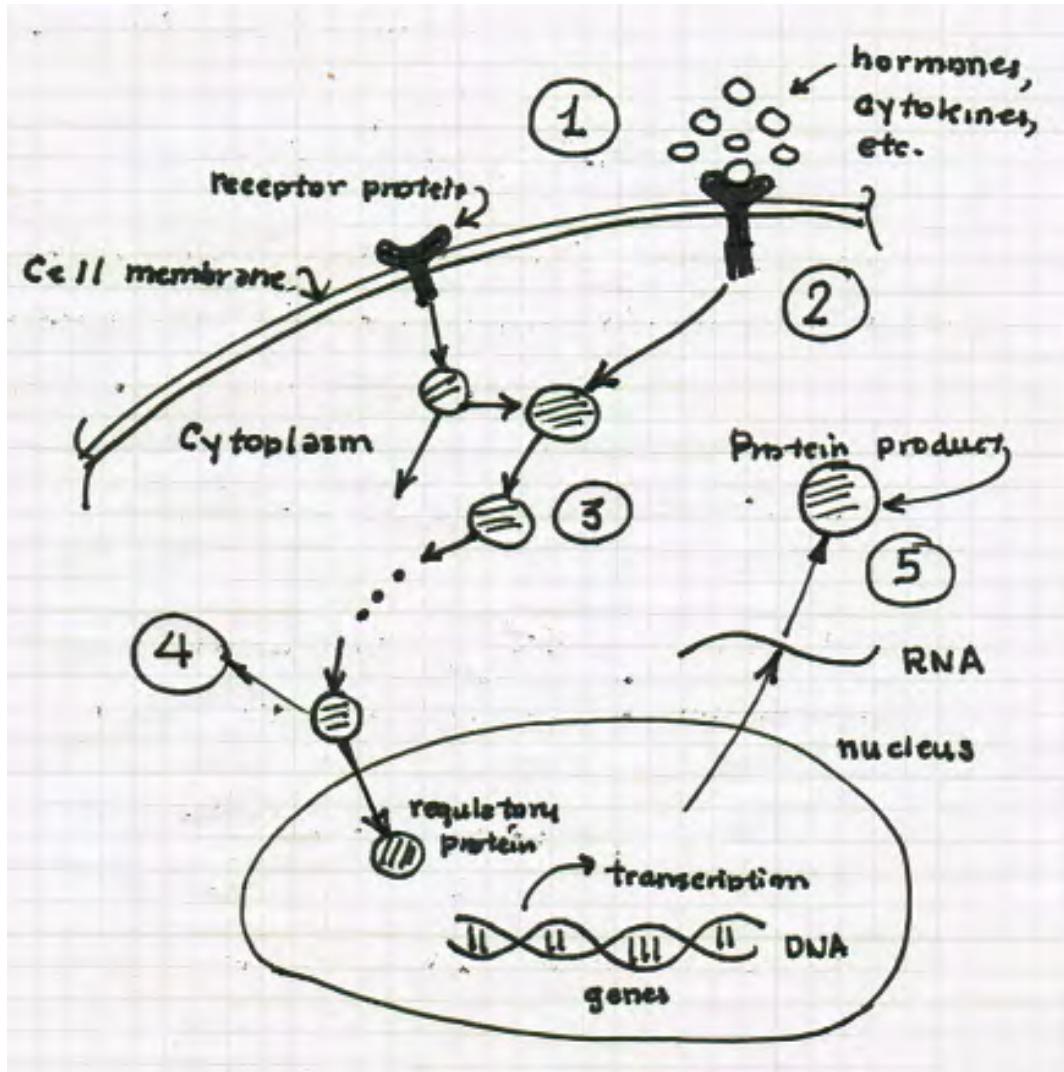


Figure 8.16: Subcellular signaling pathway: 1) Receptor accepts message; 2) message is delivered to protein path; 3) signaling pathways; 4) cell responses; 5) cell behavior changes. [c.f., 93].

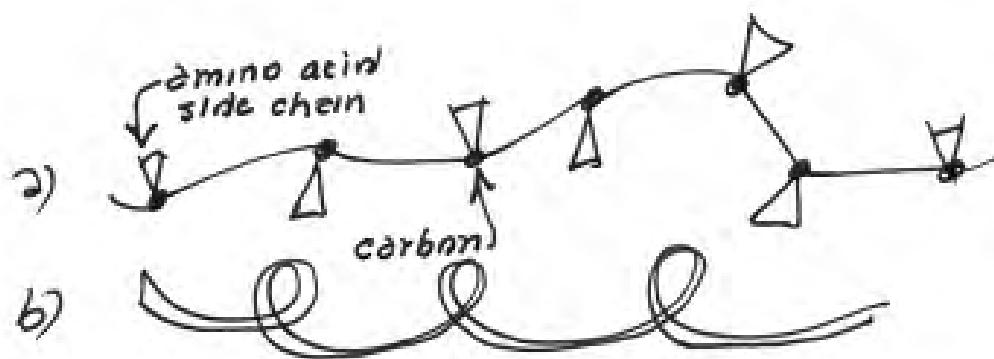


Figure 8.17: a) A protein is a chain of amino acid molecules. The protein manufactured in the protein synthesis process is thus a chain of atoms with amino acid molecules as side chains bonded to a carbon backbone chain. The interaction of the side chains causes the protein to fold into a unique convoluted 3D shape, which is often critical to the function of the protein. b) A protein chain depicted as a continuous ribbon. It may possess binding sites, in which subsets of amino acids allow subsets of target molecules to interact with and *bind* the protein.

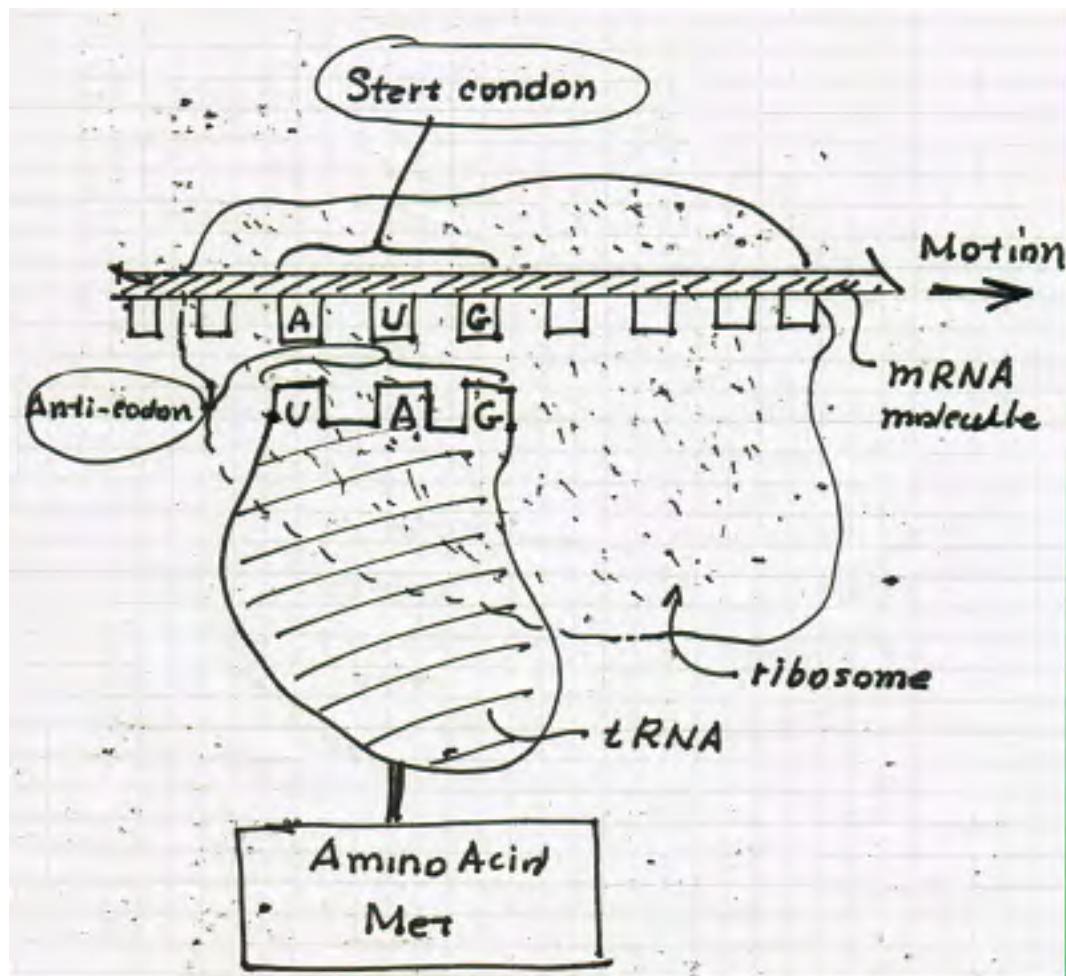


Figure 8.18: Protein Synthesis. – Stage (A). The beginning of the protein synthesis process: construction of a protein chain begins as an initiator “tRNA” molecule charged with a (starting) amino acid Methionine (denoted Met) binds the ribosome and the mRNA.

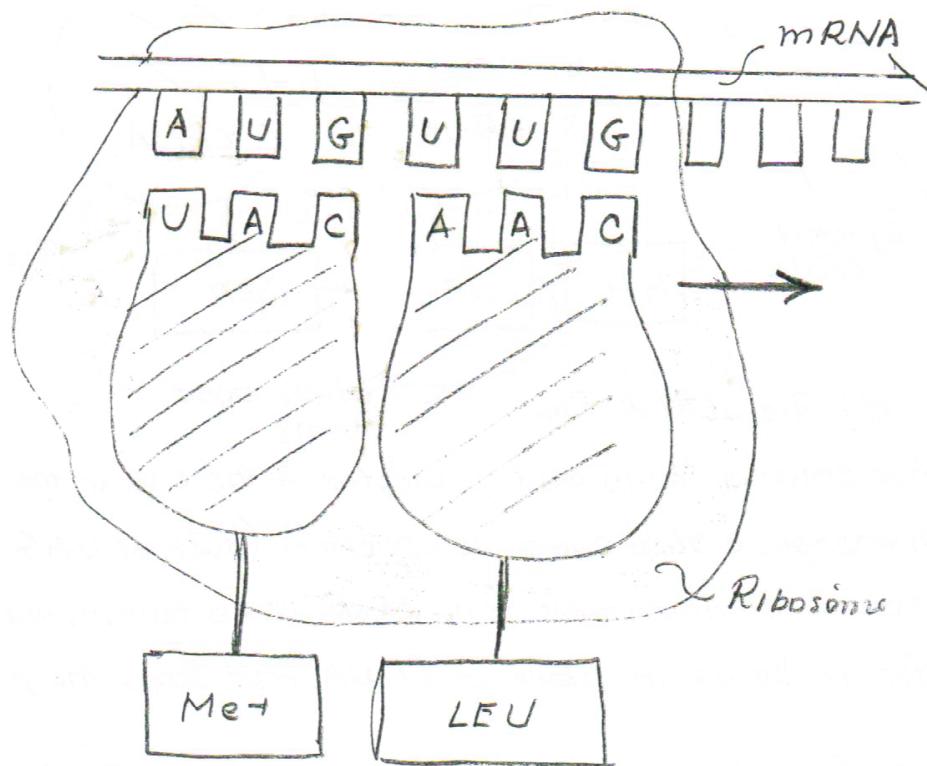


Figure 8.19: Protein Synthesis. – Stage (B). Anti-codon bends with start codon on mRNA, and first and second amino acids are aligned.

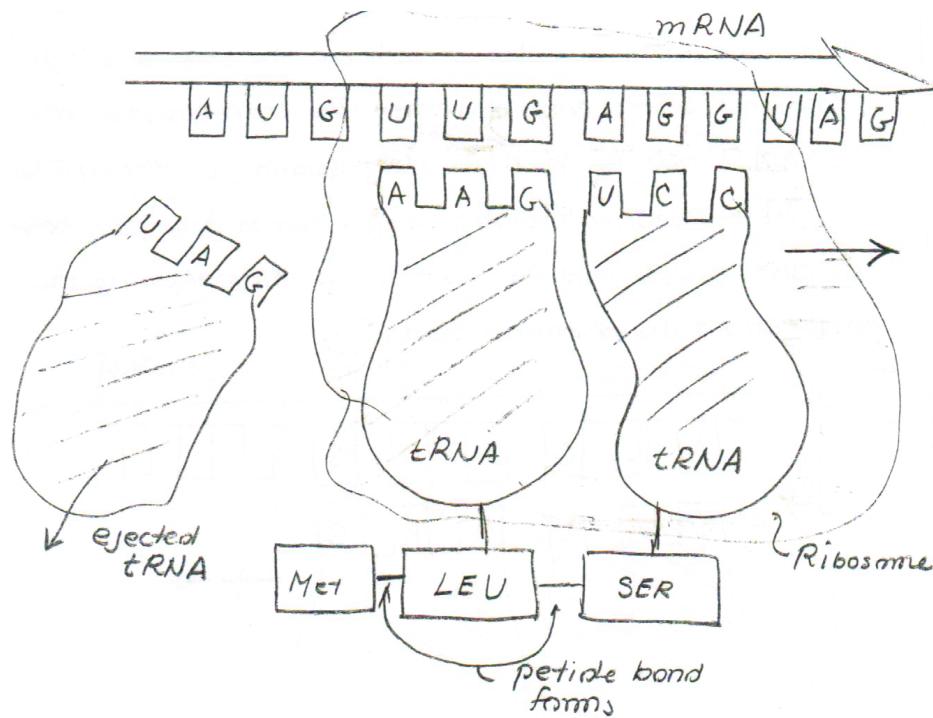


Figure 8.20: Protein Synthesis. – Stage (C). Bond with first tRNA is broken and it is ejected and recycled and ribosome moves down mRNA.

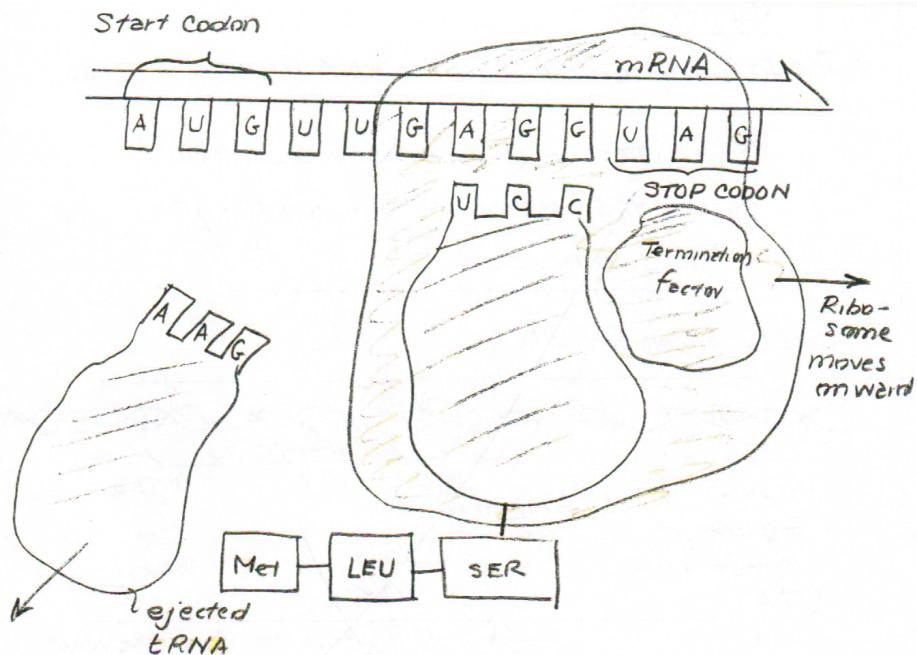


Figure 8.21: Protein Synthesis. – Stage (E). Stop codon is reached and process is terminated.

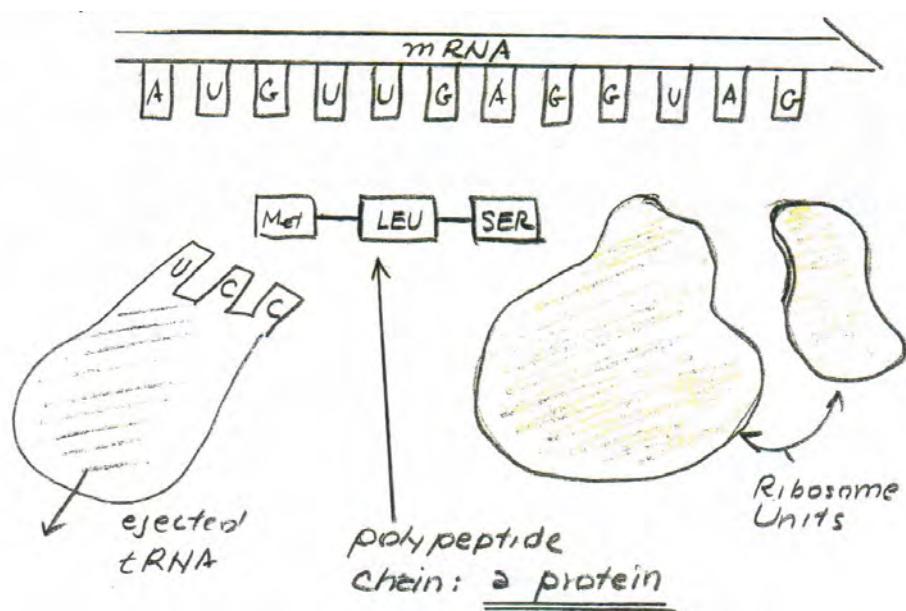


Figure 8.22: Protein Synthesis. – Stage (F). Ribosome separates and releases mRNA: the protein polypeptide is complete and is separated from the last tRNA molecule.

SUMMARY

- The information contained in RNA is translated by *Ribosome*: a large macromolecule that guides the assembly of amino acids into a protein product. Now the mRNA template is converted into an actual protein through a process called *translation*.
- Translation takes place at the Ribosome. The ribosome moves along the mRNA molecule adding an amino acid to the chain until it reaches a *stop codon*, at which point the protein production is finished, and the ribosome detaches.
- Each codon (a word consisting of the three RNA bases) corresponds to a specific *amino acid* (, for example, *GCC* corresponds to the amino acid *Alanine*). There are around 20 common amino acids.
- A molecule called *transfer RNA*: tRNA, is the link between a codon and an amino acid.
- The ribosome steps along the mRNA molecule adding an amino acid to the chain at each step to the growing protein chain. The process stops when it reaches a *stop codon*.
- A single gene can be transcribed many times, each time leading to an identical protein.
- Numerous ribosomes can bind a single mRNA, producing many copies of the same protein.
- **Summary of RNA translation process:**
 1. RNA instigates an unwinding of the DNA double helix, which is done by RNA polymerase II.
 2. RNA transcribes a gene sequence and is encoded as mRNA.
 3. mRNA leaves the cell nucleus (and DNA rewinds).
 4. Ribosome (a macromolecule) attaches to the mRNA and facilitates pairing of mRNA codon with anti-codon of tRNA molecule.
 5. tRNA links codon with an amino acid molecule. This reaction is catalyzed by an enzyme *aminoacyl tRNA synthetase*
 6. Ribosome steps along mRNA adding an amino acid molecule to the chain at each step, stopping at a termination sequence. The protein is thus manufactured.

8.8 The Cell Cycle

The process of cell division begins when growth stimulating factors instigate signal-transduction cascades. The cell volume doubles, all cellular substructures are reproduced, ~ 6 billion nucleic-acid bases are replicated, and 46 chromosomes are synthesized.

The Five Phases of Cell Cycle: In the five stages of the cycle of eukaryotic cells are presented symbolically in Figure 8.23.

G_0 -Phase. The *resting state* – no progress through cycle \Leftarrow terminally differentiated cells permanently arrested in the state.

G_1 -Phase. This is the *Growth Phase* in which new ribosomes, proteins, mitochondria, endoplasmic reticulum, and other constituent are built in preparation for DNA synthesis and eventual cell division.

S-Phase. This is the *Synthesis Phase* in which DNA is duplicated. The histone proteins around which DNA is wrapped are also synthesized.

G_2 -Phase. This is the second *growth phase* in which the cell prepares to divide. Proteins that are needed to promote mitosis are manufactured.

Mitosis. This is the phase of cell division. The division of the nucleus is mitosis, and is performed in 5 steps. The division of the cytoplasm is called *cytokinesis*.

The *external growth signals* that propagate through the transduction pathways, signals the cell to begin a new reproductive cycle. During the *S*-phase (the synthesis phase) genes that produce proteins in the DNA replication are activated. This includes the so-called cyclin molecules, a family of proteins of general types, labeled *A*, *B*, *C*, *D*, *E*, *F*, *G*, and *H* cyclins, and all but *D* are synthesized during the special discrete period in the cell cycle, and are degraded at a second discrete point. Cyclin concentration varies over time, and new complexes are formed that, in turn, activate new cyclin-CDK complexes. The cyclins control the progression through the G_1 phase and the transitions to the *S*-phase, and cyclin-*A* – *Ch2* controls progress through the *S* phase. A cyclin-dependent kinase must couple with a cyclin molecule before it can catalyze phosphorylation.

The six phases of mitosis are illustrated in Figure 8.24. In order of their development, they are

1. Prophase – two sets of chromosomes reside in the cell nucleus; two centrosomes emerge connected by micro tubes called the mitotic spindle.
2. Prometaphase – the nuclear membrane disintegrates, the centrosomes migrate to opposite

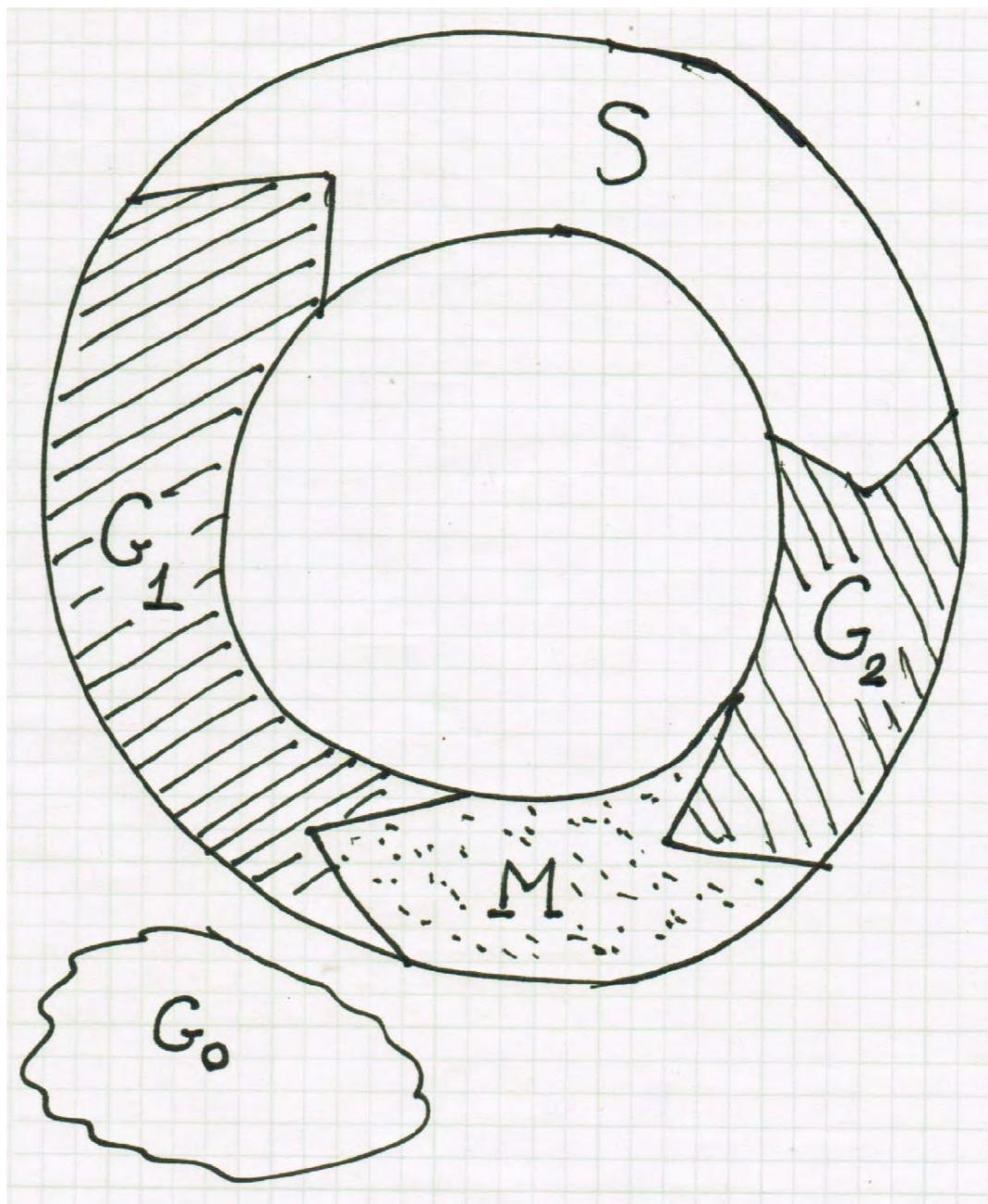


Figure 8.23: The five phases of cell cycle. G_0 , the resting phase in which no progress is made through the cycle; the G_1 -phase (gap 1), in which cyclin D and E form active cyclin and other complexes; S -phase, the synthesis phase in which all DNA is replicated including two identical copies of chromosomes; G_2 -phase in which cell continuous growth and DNA synthesis is checked; and the M -phase, mitosis, in which cell divides into two “daughter” cells and cell chromosomes condense and duplicated chromosomes separate.

ends of the cell and the spindle micro tubes bind each chromosomes at their center (the centromere).

3. Metaphase – Identical chromosomes line up joined at the centromere.
4. Anaphase – Bonds joining the chromosomes breaks.
5. Telophase – New nuclear membranes form around each set of separated chromosomes.
6. G_1 -phase – The cell is cleaved into two cells and chromosomes condense into the cell nuclei.

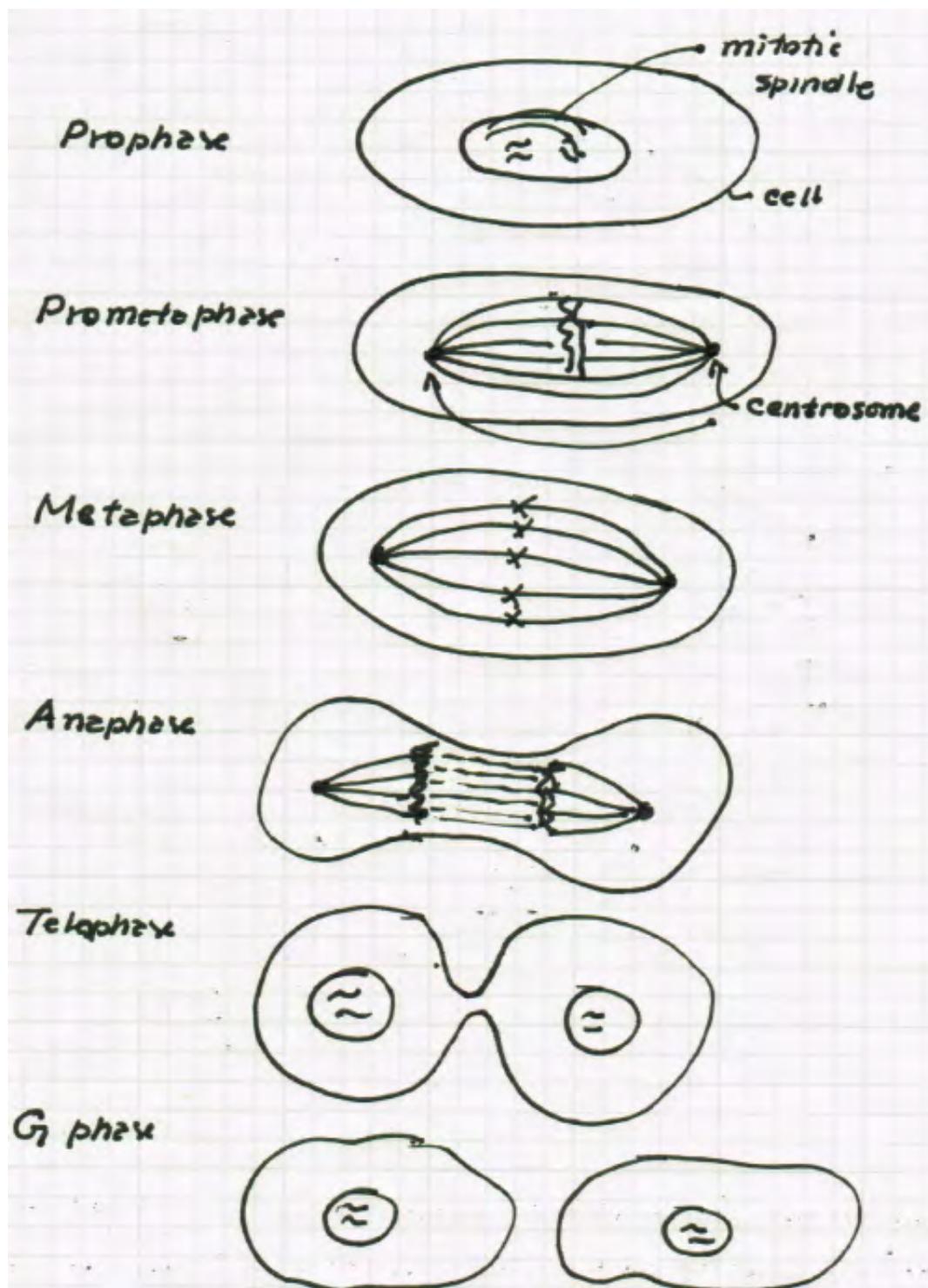


Figure 8.24: Mitosis – The M-phase of cell division.

8.9 Cancer

Cancer is a gross distortion of cell behavior due to abnormalities in the production and functioning of proteins, and due to mutations of genes. Any autonomously growing mass of cells is called a *neoplasm*. A neoplasm is classified as cancer if it is *malignant*, meaning it gains the capacity to invade surrounding tissues, which generally involves destroying proteins that hold the adjacent cells together. The term *tissue homeostasis* are used to describe the balance between cell proliferation and cell death (apoptosis) that preserves the shape, volume, architecture, and functionality of tissue in mammals. Thus, neoplasms constitute a disruption of homeostasis and the occurrence of abnormal growth.

A tumor requires nutrients, generally provided by an enhanced supply of blood, which also removes waste (e.g., dead cells). Among disruptions of the DNA that lead to cancer are alterations of the base acids in the DNA, in which bonds between the *A*, *T*, *C*, *G* bases are broken, or additional bonds appears, or when a base molecule is separated from the sugar unit leaving an unpaired base, or a single strand of the DNA breaks, compromising the phosphodiester backbone. Such events can be created by radiation, or chemical species that break chemical bonds.

Various types of cancer – and there are perhaps over 100 types – are often categorized according to the types of cells they invade. For example, *carcinomas* are cancers of epithelial cells, *sarcomas* describe cancers of connective tissues between cell types, and *leukemia* denote cancers of the blood and lymphatic system. An illustration of a transition of the tumor into cancer is given in Figure 8.25.

In a highly cited pair of papers, Hanahan and Weinberg [50, 51] proposed eight common traits shared by all, or at least the majority of types of cancer, and referred to them as *The Hallmarks of Cancer*. These are displayed in Figure 8.26. The hallmarks are:

1. *Sustaining Proliferative Signaling*, in which the ability for cells to divide even without the presence of growth factor is gained.
2. *Deregulating Cellular Energetics*, in which the ability to produce cellular glucose, and thus cellular energy production, is reprogrammed.
3. *Resisting Cell Death*, in which a resistance to signals promoting apoptosis is acquired.
4. *Inducing Angiogenesis*, in which the cell takes on the ability to induce growth of new blood vessels to supply nutrients and oxygen to cancerous cells.

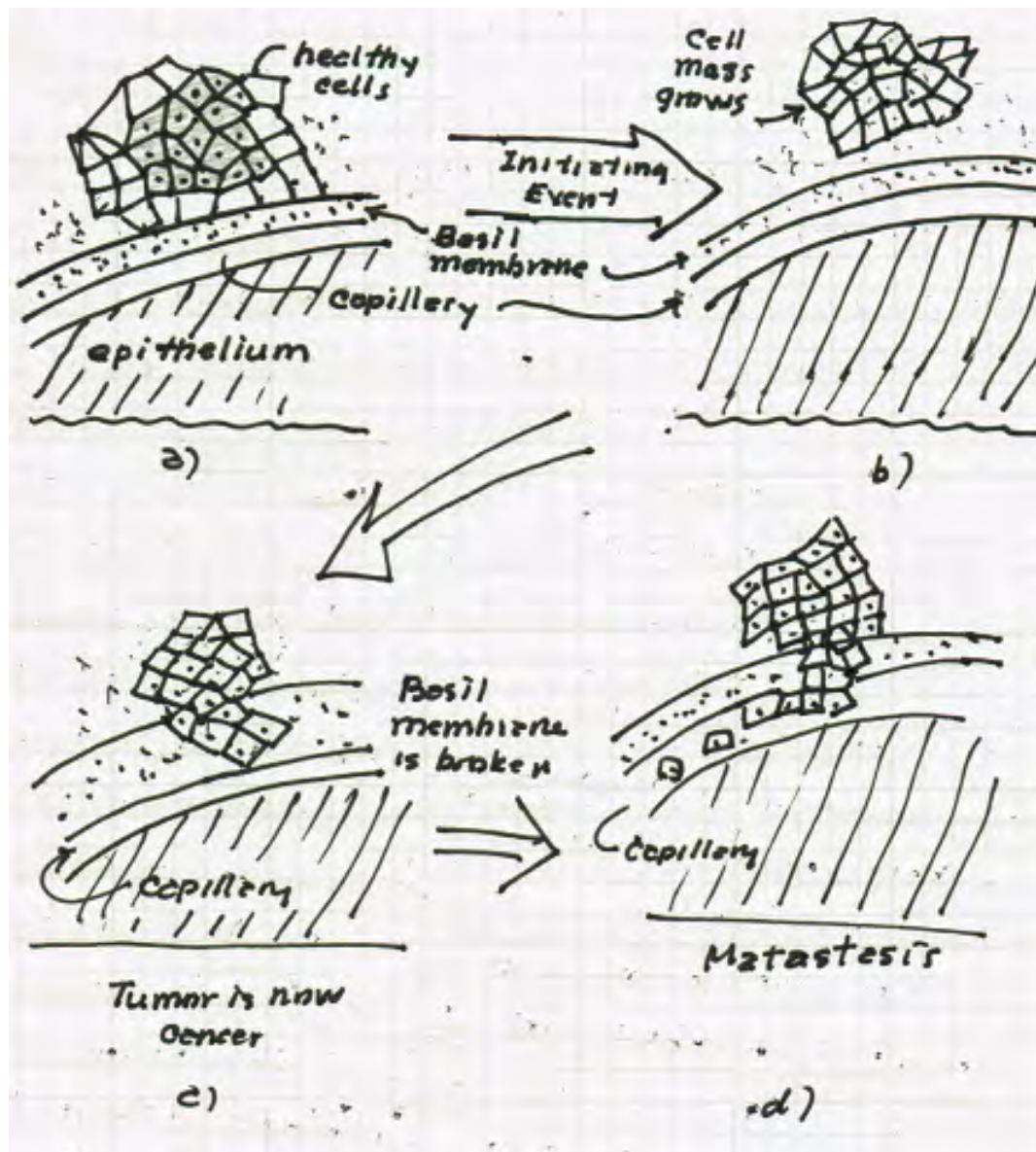


Figure 8.25: Illustration of the growth of a mass of cells into cancer. a) A mass of healthy and stem cells is b) subjected to an initializing event – such as radiation – that damages the DNA and pathways, causing mutations and uncontrolled growth. The baseline membrane is penetrated, allowing tumor cells to invade capillaries and be transmitted to other locations in the body.

5. *Evading Growth Suppressors*, by which the cancer cells gain the ability to ignore anti-proliferation signals that maintain cellular quiescence.
6. *Avoiding Immune Destruction*, refers to the ability of cells to evade attacks and destruction by the immune system.
7. *Enabling Replicative Immortality*, in which the ability to ignore replicative senescence (old age of cells) is avoided.
8. *Activating Invasion and Metastasis*, which involves acquiring the ability to spread to other parts of the body in secondary forms.

These hallmarks form the basis for inductive hypotheses underlying key terms in mathematical models of tumor growth in living tissues. These are discussed in Chapter 10.



Figure 8.26: A display of the eight hallmarks of cancer proposed by Hanahan and Weinberg [50, 51]. C.f., [74].

Chapter 9

Phenomenological Models of Tumor Growth

9.1 Introduction

At this point in this exposition, we reach a threshold in predictive computational science in which we attempt to explore problems of a true frontier of this subject, a threshold in which all of the mathematical, computational, and scientific theory discussed thus far are brought to bear on one of the most elusive and challenging models to predict the emergence, growth, or decline of cancer in living organisms.

The first steps involve, as always, inductive logic, the development of hypothesis in the form of mathematical theories designed to explain observed phenomena. We are thus faced with a question of overriding importance: what phenomena relative to tumor growth do we observe, and what scientific principles, prior information, and judgment do we call on to develop meaningful mathematical models of the physical and biological events observed?

This is not as easy question. The events leading to cancer (recall the Hallmarks of cancer listed in the previous chapter) are manifested at several different spatial and temporal scales. On the phenomenological level (i.e. at scales we observe with our human senses), we encounter living tissue, a highly heterogeneous, hugely complicated mass of interacting constituents that include masses of living cells in various stages of the cell cycle, along with blood vessels, other tissue-level constituents, and nutrients, all evolving under the changing environmental conditions to which

the organism is subjected. Below this scale are events at the cell level, involving the motion and contact of individual cells, cell mitosis, hypoxia and apoptosis. At still smaller scales within individual cells there occur the critical processes of signal transduction – protein signaling along signaling pathways, all orchestrating the behavior of the cell, its DNA, and its role in homeostasis, cell proliferation, and necrosis. Thus, the prediction of tumor growth in living organisms involves multiscale modeling, in which computational models of events at several scales must be developed and in which the influence of events at one scale on those at other scales must be captured. An illustration of these multiscale systems is shown in Figure 9.1.

In the present chapter we focus on phenomenological models of tissue behavior which attempt to capture the behavior of a multitude interacting living species of different chemical and mechanical composition. It is asserted that an appropriate mathematical framework for depicting such behavior is the theory of mixtures, a generalization of continuum mechanics in which the thermo-mechanical behavior of multiple interacting species is described. Thus, in developing this class of models, we begin with time-proven principles of mechanics and we supplement these with mathematical characterization of the events known to influence and promote cancer, as represented, for example, in the Hallmarks described earlier. Ultimately, we must also supplement these components of the models with empirical relationship drawn from experimental observations. The successful implementation of the models-development process must simultaneously account for events at the cellular scale and the signaling-pathway scales. These modeling aspects are taken up in subsequent chapters.

Following this introduction, we present a brief overview of the foundations of continuum mechanics as a prelude to continuum mixture theory, taken up in Section 9.3. Example phase-field models of tumor growth are presented in Section 9.4.

9.2 A Brief Review of Continuum Mechanics

Continuum mechanics can be viewed as branch of mathematical physics that employs phenomenological models to describe the behavior of materials and material bodies. As such, it provides the foundations of fluid mechanics and solid mechanics. The fundamental hypothesis of continuum mechanics is that matter is not discrete. The matter making up a material body is continuously distributed in one-to-one correspondence with points in a subset of \mathbb{R}^3 . Material bodies are thus “continuous media”.

Continuum mechanics models the physical universe as a collection of deformable bodies, including

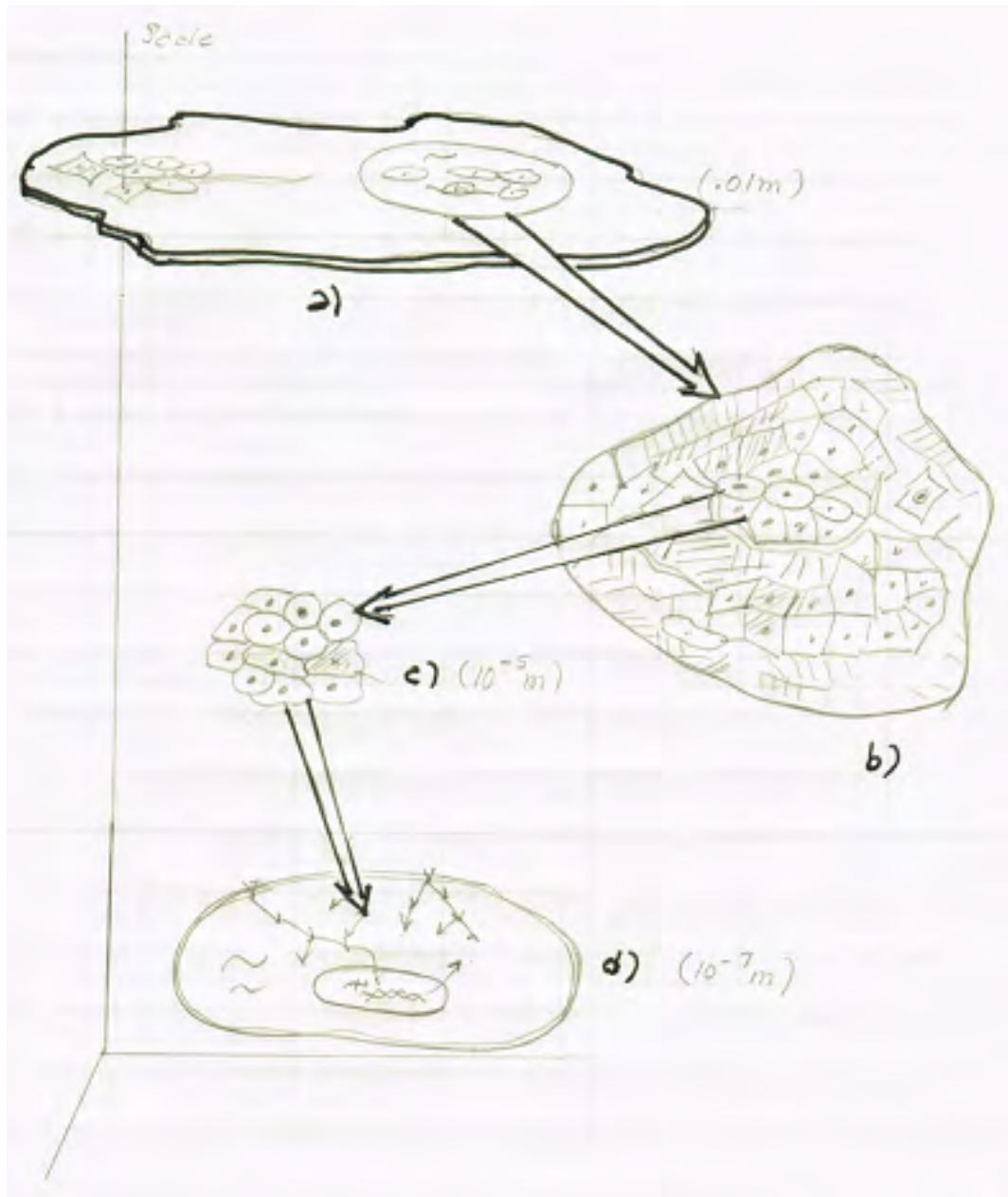


Figure 9.1: Multiple scales in the physics and biology of tumor growth: a) and b) the tissue scale(\sim centimeters); c) cellular scale (\sim 10 microns), and d) sub-cellular signaling scale (\sim 10–100 nanometers).

fluid bodies, that occupy regions in three-dimensional space, and the subsets of space occupied by a body that are called its *configurations*. One configuration, generally one in which the geometry and physical state are known, is selected as the reference configuration, and other configurations are characterized by comparing them with the reference configuration.

Given material body \mathcal{B} , the material points in the reference configuration are in one-to-one correspondence with a subset Ω_0 of \mathbb{R}^3 , with boundary $\partial\Omega_0$. The material points that constitute the body are each assigned a vector \mathbf{X} in the reference configurations. These \mathbf{X} , or its components (x_1, x_2, x_3) in a Cartesian coordinate system in Ω_0 represent a label assigned to each material point of the body.

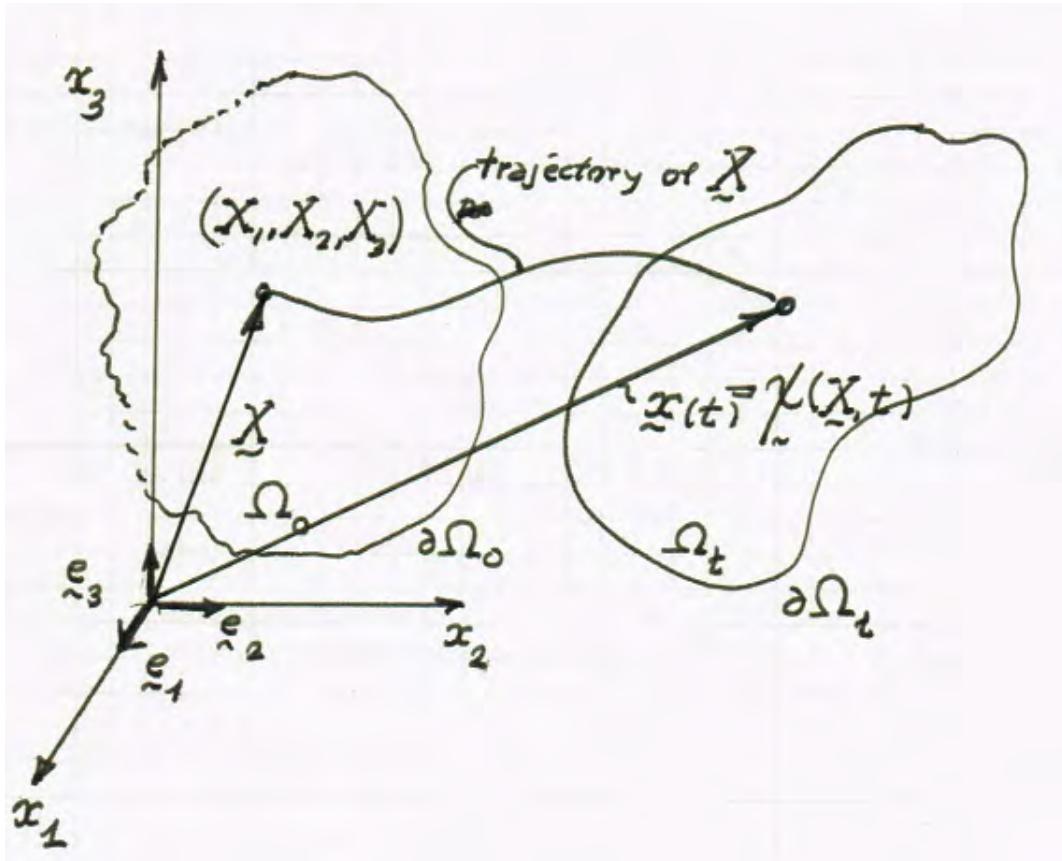


Figure 9.2: Motion of a continuous body from its reference configuration to its current configuration.

The motions of the body carries it through space over a time period $[0, t]$ so that at time t the body occupies a configurations $\Omega_t \subset \mathbb{R}^3$. Thus, material points \mathbf{X} in $\bar{\Omega}_0$ (the classes of Ω_0) are

mapped into positions \mathbf{x} in $\bar{\Omega}_t$ by a smooth vector-valued map $\boldsymbol{\chi}$,

$$\mathbf{x} = \boldsymbol{\chi}(\mathbf{X}, t) \quad (9.1)$$

$\boldsymbol{\chi}$ is called the motion of the body. The displacement of points \mathbf{X} is then

$$\mathbf{u} = \boldsymbol{\chi}(\mathbf{X}) - \mathbf{X} \quad (9.2)$$

(omitting time t for simplicity). The tensor $\mathbf{F} = \mathbf{F}(\mathbf{X})$ defined by

$$\mathbf{F} = \boldsymbol{\nabla} \boldsymbol{\chi} = \frac{\partial X_i}{\partial X_j} e_i \otimes e_j , \quad (9.3)$$

is called the deformation gradient, e_i being unit orthonormal basis vectors along coordinate lines in the reference configurations.

Material line elements in the reference configuration and the current configuration of the body are, respectively,

$$dS_0^2 = d\mathbf{X}^T d\mathbf{X} \quad \text{and} \quad dS^2 = d\mathbf{x}^T d\mathbf{x} ,$$

and since $d\mathbf{x} = \mathbf{F}d\mathbf{X}$,

$$dS^2 - dS_0^2 = d\mathbf{X}^T (\mathbf{C} - \mathbf{I}) d\mathbf{X} = 2d\mathbf{X}^T \mathbf{E} d\mathbf{X} ,$$

where \mathbf{C} is the right Cauchy-Green deformation tensor,

$$\mathbf{C} = \mathbf{F}^T \mathbf{F} , \quad (9.4)$$

and \mathbf{E} is the Green-St. Venant strain tensor,

$$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I}) . \quad (9.5)$$

The velocity and acceleration fields in the body are

$$\dot{\mathbf{x}} = \frac{\partial \boldsymbol{\chi}(\mathbf{X}, t)}{\partial t} \quad \text{and} \quad \ddot{\mathbf{x}} = \frac{\partial^2 \boldsymbol{\chi}(\mathbf{X}, t)}{\partial t^2} , \mathbf{X} \in \bar{\Omega}_0 , t \geq 0 , \quad (9.6)$$

and the *spatial* description of velocity is

$$\mathbf{v} = \mathbf{v}(\mathbf{x}, t) = \dot{\mathbf{x}}(\boldsymbol{\chi}^{-1}(\mathbf{x}, t), t) . \quad (9.7)$$

This leads to the two classical interpretation of representations of motion (and general behavior)

of continuum bodies:

1. the *material* (or *Lagrangian*) descriptions of motion in which functions are defined on *material points* \mathbf{X} of the body, and
2. the *spatial* (or Eulerian) descriptions of motion in which functions are defined on spatial *places* \mathbf{x} in \mathbb{R}^3 .

Thus, given a scalar field $\psi = \psi(X, t)$, its time rate of change in the Lagrangian descriptions is

$$\frac{d\psi(\mathbf{X}, t)}{dt} = \frac{\partial\psi(\mathbf{X}, t)}{\partial t} + \frac{\partial\psi}{\partial\mathbf{X}} \cdot \frac{\partial\mathbf{X}}{\partial t} = \frac{\partial\psi(\mathbf{X}, t)}{\partial t}, \quad (9.8)$$

because, being only a label of a material point, $\partial\mathbf{X}/\partial t = \mathbf{0}$. In the Eulerian case,

$$\frac{d\psi(\mathbf{x}, t)}{dt} = \frac{\partial\psi(\mathbf{x}, t)}{\partial t} + \frac{\partial\psi(\mathbf{x}, t)}{\partial\mathbf{x}} \cdot \frac{\partial\mathbf{x}}{\partial t}, \quad (9.9)$$

but, $\partial\mathbf{x}/\partial t = \dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, t)$, the velocity at \mathbf{x} at time t , so

$$\frac{d\psi(\mathbf{x}, t)}{dt} = \frac{\partial\psi(\mathbf{x}, t)}{\partial t} + \mathbf{v}(\mathbf{x}, t) \cdot \frac{\partial\psi(\mathbf{x}, t)}{\partial\mathbf{x}}. \quad (9.10)$$

We denote $\partial/\partial\mathbf{X} = \nabla = \text{GRAD}$, $(\partial/\partial\mathbf{X}) \cdot \phi = \nabla \cdot \phi = \text{DIV } \phi$, $\partial/\partial\mathbf{x} = \text{grad}$, and $\partial/\partial\mathbf{x} \cdot \mathbf{v} = \text{div } \mathbf{v}$. Thus,

$$\frac{d\psi}{dt} = \frac{\partial\psi}{\partial t} + \mathbf{v} \cdot \text{grad } \psi. \quad (9.11)$$

The velocity gradient is

$$\mathbf{L}(\mathbf{x}, t) = \text{grad } \mathbf{v}(\mathbf{x}, t) = \mathbf{D}(\mathbf{x}, t) + \mathbf{W}(\mathbf{x}, t) \quad (9.12)$$

where

$$\mathbf{D} = \frac{1}{2}(\mathbf{L} + \mathbf{L}^T) \quad \text{and} \quad \mathbf{W} = \frac{1}{2}(\mathbf{L} - \mathbf{L}^T) \quad (9.13)$$

\mathbf{D} is the *rate-of-deformations* tensor and \mathbf{W} is the *spin* tensor.

The Balance Laws of Continuum Mechanics. The thermo-mechanical behavior of continuous bodies is governed by the following five fundamental principles of mechanics and thermodynamics.

I) *The Principle of Conservation of Mass.* The total mass $\mathcal{M}(B)$ of a body B in a closed system

is constant during any motion χ of B . If

$$\mathcal{M}(B) = \int_{\Omega_t} \rho dx , \quad (9.14)$$

where $\rho = \rho(\mathbf{x}, t)$ is the mass density and $dx = dx_1 dx_2 dx_3$ is a volume element, then

$$\frac{d}{dt} \mathcal{M}(B) = 0 . \quad (9.15)$$

Under suitable assumptions on the smoothness of the mass density, this implies, in the spatial description of motion,

$$\frac{\partial \rho}{\partial t} + \operatorname{div} (\mathbf{v}\rho) = 0 , \quad (9.16)$$

$\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ being the velocity.

II) The Principle of Balance of Linear Momentum. The time-rate-of change of the total linear momentum, $I(B)$, of a material body, equals the net force acting on the body,

$$\frac{dI(\mathbf{B})}{dt} = \mathbf{F} , \quad (9.17)$$

where

$$I(\mathbf{B}) = \int_{\Omega_t} \rho \mathbf{v} dx \quad (9.18)$$

and

$$\mathbf{F} = \int_{\Omega_t} \mathbf{b} dx + \int_{\partial\Omega_t} \boldsymbol{\sigma}(\mathbf{n}) dA . \quad (9.19)$$

Here \mathbf{b} is the body force per unit volume and $\boldsymbol{\sigma}(\mathbf{n})$ is the stress vector on surface area element dA with orientation given by the unit exterior normal \mathbf{n} . if momentum is conserved, then the Cauchy stress principle holds:

$$\boldsymbol{\sigma}(\mathbf{n}, \mathbf{x}, t) = \mathbf{T}(\mathbf{x}, t) \mathbf{n} \quad (9.20)$$

where $\mathbf{T}(\mathbf{x}, t)$ is the (second-order) *Cauchy stress tensor*. The local Eulerian form of the principle of balance of linear momentum is thus,

$$\operatorname{div} \mathbf{T} + \mathbf{b} = \rho \frac{d\mathbf{v}}{dt} . \quad (9.21)$$

III) Principle of Balance of Angular Momentum. The time-rate-of change of the total angular momentum $H(\mathbf{B})$ of a body \mathbf{B} about a point \mathbf{O} (the origin of \mathbf{x}), equals the net moment \mathbf{M} about O :

$$\frac{dH(\mathbf{B})}{dt} = \mathbf{M} . \quad (9.22)$$

If (9.21) holds, then locally, this implies that the stress tensor is symmetric:

$$\boxed{\mathbf{T} = \mathbf{T}^T} . \quad (9.23)$$

IV) The Principle of Conservation of Energy. The time-rate-of change of the total energy E of a body \mathbf{B} is equal to the sum of the mechanical power P and the heating per unit time, \dot{Q} :

$$\frac{dE}{dt} = P + \dot{Q} . \quad (9.24)$$

The total energy is the sum of the kinetic energy K ,

$$K = \frac{1}{2} \int_{\Omega_t} \rho \mathbf{v} \cdot \mathbf{v} dx \quad (9.25)$$

and the internal energy U ,

$$U = \int_{\Omega_t} \rho e dx , \quad (9.26)$$

where e is the internal energy per unit mass. This power is given by

$$P = \int_{\Omega_t} \mathbf{b} \cdot \mathbf{v} dx + \int_{\partial\Omega_t} \boldsymbol{\sigma}(\mathbf{n}) \cdot \mathbf{v} dA \quad (9.27)$$

and the heating is given by

$$\dot{Q} = \int_{\partial\Omega_t} -\mathbf{q} \cdot \mathbf{n} dA + \int_{\Omega_t} r dx , \quad (9.28)$$

where \mathbf{q} is the heat flux, \mathbf{n} is unit exterior normal to the boundary, and r is the heat supply per unit volume generated by internal source.

Locally, (9.24)-(9.28) lead to

$$\boxed{\rho \frac{de}{dt} = \mathbf{T} : \mathbf{D} - \operatorname{div} \mathbf{q} + r} . \quad (9.29)$$

V) The Second Law of Thermodynamics. The time-rate-of change of the total entropy $S(B)$ of

body \mathbf{B} minus the entropy of heating, per unit temperature is always greater than or equal to zero:

$$S(B) = \int_{\Omega_t} \rho \eta \, dx , \quad (9.30)$$

$$\frac{dS}{dt} - \frac{1}{\theta} \dot{Q} \geq 0 . \quad (9.31)$$

Here η is the entropy per unit mass and $\theta = \theta(\mathbf{x}, t)$ is the absolute temperature. Locally

$$\boxed{\rho \frac{d\eta}{dt} + \operatorname{div} \frac{\mathbf{q}}{\theta} - \frac{r}{\theta} \geq 0} \quad (9.32)$$

Relation (9.32) is also called the Clausius-Duhem inequality.

Material (Lagrangian) descriptions of the local laws (9.16), (9.21), (9.26), (9.29), and (9.32) are, respectively,

$$\left. \begin{aligned} \rho_0 &= \rho \det \mathbf{F} \\ \operatorname{DIV} \mathbf{S} + \mathbf{b}_0 &= \rho_0 \frac{\partial^2 \mathbf{u}}{\partial t^2} \\ \mathbf{S} &= \mathbf{S}^T \\ \rho_0 \frac{\partial e_0}{\partial t} &= \mathbf{S} : \dot{\mathbf{E}} - \operatorname{DIV} \mathbf{q}_0 + r_0 \\ \rho_0 \frac{\partial \eta_0}{\partial t} + \operatorname{DIV} \frac{\mathbf{q}_0}{\theta} - \frac{r_0}{\theta} &\geq 0 \end{aligned} \right\} \quad (9.33)$$

where the subscript “0” denotes quantities measured in the reference configuration, \mathbf{E} is the Green-St. Venant tensor (9.5) $\mathbf{S} = \det \mathbf{F} \mathbf{F}^{-1} \mathbf{T} \mathbf{F}^{-T}$ is the second Piola-Kirchhoff stress tensor.

9.3 A Continuum Theory of Mixtures

The continuum theory of mixtures provides a unified approach to modeling the evolution of complex, heterogeneous masses of interacting media. Continuum mixture theory dates back to early work of Fick [40], Darcy [37], and work of Truesdell [112], Truesdell and Neepeen [114], and others. Comprehensive accounts of the theory and a fuller review of relevant literature can be found in the review article of Bowen [20] and the monograph of Rajagopal and Tao [90]. The version outlined here follows Oden et al (2010) [82].

The fundamental idea underlying mixture theory is that a material body \mathbf{B} can be composed N constituent species $\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_N$ that occupy the same region Ω of physical space at the same time. As in Section 9.2, we begin by establishing a fixed reference configuration associated with a region Ω in \mathbb{R}^3 . We denote material particles in the d th constituent by the material bodies \mathbf{X}_α and we define the spatial position of each material point at time t by the motions,

$$\mathbf{x} = \boldsymbol{\chi}_\alpha(\mathbf{X}_\alpha, t) , \quad (9.34)$$

$\alpha = 1, 2, \dots, N$. The corresponding deformation gradient for constituent α is

$$\mathbf{F}_\alpha = \text{GRAD } \boldsymbol{\chi}_\alpha(\mathbf{X}_\alpha, t) . \quad (9.35)$$

At each spatial position, \mathbf{x} , N constituent species exists, and each constituent is assigned a mass density $\hat{\rho}_\alpha$ where represent the mass per unit volume of the α th constituent. The mass density of the parent media (the total mixture) is at \mathbf{x} at time t is

$$\rho(\mathbf{x}, t) = \sum_N^{\alpha=1} \hat{\rho}_\alpha(\mathbf{x}, t) . \quad (9.36)$$

It is convenient to describe the distributions of the various species at a point by introducing the idea of their volume fractions. Consider a differential volume dv in $\bar{\Omega}_t \times [0, \tau]$ containing (\mathbf{x}, t) and let dv_α be the portion of dv occupied by constituent α . Then

$$\phi_\alpha(\mathbf{x}, t) = dv_\alpha/dv \quad (9.37)$$

is the volume fraction of the α th constituent at (\mathbf{x}, t) . The assumption that the volume $\phi_\alpha dv = dv_\alpha$ fill up the total volume is called *saturation condition* and is written

$$\sum_{\alpha=1}^N \phi_\alpha(\mathbf{x}, t) = 1 . \quad (9.38)$$

Moreover, each constituent has its own velocity,

$$\mathbf{v}_\alpha = \frac{\partial \boldsymbol{\chi}_\alpha}{\partial t}(X_\alpha, t) = \frac{\partial \boldsymbol{\chi}_\alpha}{\partial t}(\boldsymbol{\chi}_\alpha^{-1}(\mathbf{X}_\alpha, t), t) = \mathbf{v}_\alpha(\mathbf{x}, t) , \quad (9.39)$$

while the parent full mixture velocity is the mass-averaged velocity

$$\mathbf{v}_\alpha = \rho^{-1} \sum_{\alpha=1}^N \rho_\alpha \phi_\alpha \mathbf{v}_\alpha . \quad (9.40)$$

The motion of the mixture relative to that of the α th constituent is called the diffusion velocity \mathbf{u}_α and is defined by

$$\mathbf{u}_\alpha = \mathbf{v}_\alpha - \mathbf{v} . \quad (9.41)$$

A simple calculations shows that $\sum_\alpha^N \rho_\alpha \phi_\alpha \mathbf{u}_\alpha = \mathbf{0}$.

For the mixture two material-time derivations are relevant. For a differentiable function $\phi = \phi(x, t)$ we denote

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi \quad \text{and} \quad \frac{d^\alpha \phi}{dt} = \frac{\partial \phi}{\partial t} + \mathbf{v}_\alpha \cdot \nabla \phi \quad (9.42)$$

$\nabla = \partial/\partial \mathbf{x}$ being the spatial gradient.

The Balance Laws for Mixtures.

Each constituent is required to satisfy its own balance laws, which differ from those of Section 9.2 due to the presence of interaction terms representing the exchange of mass, momentum, and energy between constituents. For a mixture occupying an open region Ω in \mathbb{R}^3 over a time interval $(0, \tau)$, the volume fractions and other independent field variables must satisfy the following local forms of the balance laws for all α , $1 \leq \alpha \leq N$, all $\mathbf{x} \in \Omega$, and $t \in (0, \tau)$.

I. Conservation of Mass for a Mixture

$$\frac{\partial \rho_\alpha \varphi_\alpha}{\partial t} + \nabla \cdot (\rho_\alpha \varphi_\alpha \mathbf{v}_\alpha) = S_\alpha - \nabla \cdot \mathbf{J}_\alpha , \quad (9.43)$$

where S_α is the mass supplied to constituent α by other constituents and \mathbf{J}_α is the mass flux due to changes in the chemical potential defined in terms of gradients in concentrations and changes in nutrient concentrations.

II. Balance of Linear Momentum for a Mixture

$$\rho_\alpha \varphi_\alpha \frac{d^\alpha \mathbf{v}_\alpha}{dt} = \nabla \cdot \mathbf{T}_\alpha + \rho_\alpha \varphi_\alpha \mathbf{b}_\alpha + \hat{\mathbf{p}}_\alpha \quad (9.44)$$

where \mathbf{T}_α is the *partial Cauchy stress tensor*, \mathbf{b}_α is the body force per unit mass, and $\hat{\mathbf{p}}_\alpha$ is the momentum supplied by other constituents to the α^{th} constituent.

III. Balance of Angular Momentum for a Mixture

$$\mathbf{M}_\alpha = \mathbf{T}_\alpha - \mathbf{T}_\alpha^T \quad (9.45)$$

with \mathbf{M}_α the intrinsic moment of momentum vector for constituent α . The partial stress tensor, \mathbf{T}_α , is thus, in general, asymmetric.

IV. Conservation of Energy for a Mixture

$$\rho_\alpha \varphi_\alpha \frac{d^\alpha e_\alpha}{dt} = \text{tr } \mathbf{T}_\alpha^T \mathbf{L}_\alpha - \nabla \cdot \mathbf{q}_\alpha + \rho_\alpha \varphi_\alpha r_\alpha + \hat{\varepsilon}_\alpha + \Upsilon_\alpha. \quad (9.46)$$

Here e_α is the internal energy per unit mass, \mathbf{L}_α is the velocity gradient, $\mathbf{L}_\alpha = \nabla \mathbf{v}_\alpha$, \mathbf{q}_α is the heat flux vector, and r_α is the heat supplied per unit mass per unit time. Comparing (9.46) with (9.29) we find two additional terms: $\hat{\varepsilon}_\alpha$ is the energy supplied to constituent α by other constituents, and

$$\Upsilon_\alpha = \sum_{\beta=1}^N \nabla \cdot \left(\boldsymbol{\sigma}_{\alpha\beta} \frac{d^\alpha \varphi_\beta}{dt} \right) + \sum_{\beta=1}^L \zeta_{\alpha\beta} \frac{d^\alpha m_\beta}{dt}, \quad (9.47)$$

is contribution to the energy from surface effects accruing at the interfaces of constituents.

The quantity $\boldsymbol{\sigma}_{\alpha\beta}$ is a generalized surface traction that is conjugate to time-changes in species volume fractions on constituent interfaces (see Oden [82] for a history of this energy contribution). This term represents the contribution to the change in energy due to actions of surface tensions or adhesion between cell concentrations due to time-rates-of-change of each mass concentration on the surface of the full mixture, and results from a surface power defined by a surface integral,

$$\int_{\partial\omega} \left(\sum_{\beta} \boldsymbol{\sigma}_{\alpha\beta} \frac{d^\alpha \varphi_\beta}{dt} \right) \cdot \mathbf{n} \, d s ,$$

\mathbf{n} being a unit outward normal to the boundary $\partial\omega$ of ω , an arbitrary subdomain of Ω .

In (9.46), m_β denotes a concentration of a nutrient species so that $m_\beta \varphi_\alpha$ defines the reaction between various nutrients in the mixture (such as oxygen) and the constituent φ_α , and $\zeta_{\alpha\beta}$ denotes the chemical or biological forces conjugate to changes in nutrient concentrations. This term in the energy balance provides a means to account for the diffusion of chemical or biological constituents due to chemo- or bio-taxis (see [34]).

V. The Second Law of Thermodynamics for a Mixture

$$\sum_{\alpha=1}^N \left\{ \rho_\alpha \varphi_\alpha \frac{d^\alpha \eta_\alpha}{dt} - \frac{1}{\theta_\alpha} \rho_\alpha \varphi_\alpha r_\alpha + \Gamma_\alpha \eta_\alpha + \nabla \cdot (\mathbf{H}_\alpha - \rho_\alpha \varphi_\alpha \eta_\alpha \mathbf{u}_\alpha) \right\} \geq 0. \quad (9.48)$$

Here η_α is the entropy per unit mass, Γ_α is the total mass supplied,

$$\Gamma_\alpha = S_\alpha - \nabla \cdot \mathbf{J}_\alpha, \quad (9.49)$$

\mathbf{H}_α is the entropy flux in the α^{th} constituent, and θ_α is its absolute temperature. Introducing the Helmholtz free energy ψ_α for constituent α , defined by

$$\psi_\alpha = e_\alpha - \theta_\alpha \eta_\alpha, \quad (9.50)$$

yields, after some algebra,

$$\sum_{\alpha=1}^N \left\{ \frac{1}{\theta_\alpha} \left[-\rho_\alpha \varphi_\alpha \frac{d^\alpha \psi_\alpha}{dt} - \rho_\alpha \varphi_\alpha \eta_\alpha \frac{d^\alpha \theta_\alpha}{dt} + \text{tr } \mathbf{T}_\alpha^T \mathbf{L}_\alpha - \nabla \cdot \mathbf{q}_\alpha \right. \right. \\ \left. \left. + \hat{\varepsilon}_\alpha + \sum_{\beta=1}^N \nabla \cdot \left(\boldsymbol{\sigma}_{\alpha\beta} \frac{d^\alpha \varphi_\beta}{dt} \right) + \sum_{\beta=1}^L \zeta_{\alpha\beta} \frac{d^\alpha m_\beta}{dt} \right] \right. \\ \left. + \Gamma_\alpha \eta_\alpha + \nabla \cdot (\mathbf{H}_\alpha - \rho_\alpha \varphi_\alpha \eta_\alpha \mathbf{u}_\alpha) \right\} \geq 0. \quad (9.51)$$

These principles describe the balance laws and the second law of thermodynamics, for a constituent α in a mixture of N constituents. This system is closed by the addition of appropriate constitutive equations, which put constraints on the physical processes that can be performed on the mixture. There are other constraints imposed by the requirement that the above axioms for the constituents must be consistent with those for the mixture as a whole, as will be made clear below.

The continuum balance laws for the parent mixture are defined by the use of system of laws described earlier (e.g., (9.16), (9.21), (9.26), (9.29), and (9.32)), except that new surface energy terms may appear for the mixture:

$$\left. \begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) &= 0; \\ \rho \frac{d\mathbf{v}}{dt} &= \nabla \cdot \mathbf{T} + \rho \mathbf{b}; \\ \rho \frac{de}{dt} &= \text{tr } \mathbf{T} \mathbf{L} - \nabla \cdot \mathbf{q} + \rho r + \sum_{\alpha=1}^N (\Upsilon_\alpha + \rho_\alpha \varphi_\alpha \mathbf{b}_\alpha \cdot \mathbf{u}_\alpha); \\ \rho \frac{d\eta}{dt} - \sum_{\alpha=1}^N \left(\frac{1}{\theta_\alpha} \rho_\alpha \phi_\alpha r_\alpha - \nabla \cdot \mathbf{H}_\alpha \right) &\geq 0; \end{aligned} \right\} \begin{array}{l} (\mathbf{T} = \mathbf{T}^T) \\ (\psi = e - \theta \eta) \end{array}, \quad (9.52)$$

Note that the body-force term in (9.52)₃ vanishes when there is no diffusion or when $\mathbf{b}_1 = \mathbf{b}_2 = \dots = \mathbf{b}_N = \mathbf{b}$, as pointed out by Bowen [20]. If we sum the constituent balance laws over all N constituents, the sums are compatible with (9.52) if

$$\left. \begin{aligned} \mathbf{T} &= \sum_{\alpha=1}^N (\mathbf{T}_\alpha - \rho_\alpha \varphi_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha); & \mathbf{M} &= \sum_{\alpha=1}^N \mathbf{M}_\alpha = 0; \\ e &= \frac{1}{\rho} \sum_{\alpha=1}^N \left(\rho_\alpha \varphi_\alpha e_\alpha + \frac{1}{2} \rho_\alpha \varphi_\alpha \mathbf{u}_\alpha \cdot \mathbf{u}_\alpha \right); & \mathbf{b} &= \frac{1}{\rho} \sum_{\alpha=1}^N \rho_\alpha \varphi_\alpha \mathbf{b}_\alpha; \\ \eta &= \frac{1}{\rho} \sum_{\alpha=1}^N \rho_\alpha \varphi_\alpha \eta_\alpha; & r &= \frac{1}{\rho} \sum_{\alpha=1}^N \rho_\alpha \varphi_\alpha r_\alpha; \\ \mathbf{q} &= \sum_{\alpha=1}^N \left(\mathbf{q}_\alpha - \mathbf{T}_\alpha^T \mathbf{u}_\alpha + \rho_\alpha \varphi_\alpha e_\alpha \mathbf{u}_\alpha + \frac{1}{2} \rho_\alpha \varphi_\alpha \mathbf{u}_\alpha (\mathbf{u}_\alpha \cdot \mathbf{u}_\alpha) \right) \end{aligned} \right\} \quad (9.53)$$

and, very importantly, the following conditions hold:

$$\left. \begin{aligned} \sum_{\alpha=1}^N \Gamma_\alpha &= \sum_{\alpha=1}^N (S_\alpha - \nabla \cdot \mathbf{J}_\alpha) = 0; \\ \sum_{\alpha=1}^N (\Gamma_\alpha \mathbf{u}_\alpha + \hat{\mathbf{p}}_\alpha) &= 0; \\ \sum_{\alpha=1}^N \mathbf{M}_\alpha &= 0; \\ \sum_{\alpha=1}^N \left[\hat{\varepsilon}_\alpha + \hat{\mathbf{p}}_\alpha \cdot \mathbf{u}_\alpha + \Gamma_\alpha \left(e_\alpha + \frac{1}{2} \mathbf{u}_\alpha \cdot \mathbf{u}_\alpha \right) \right] &= 0. \end{aligned} \right\} \quad (9.54)$$

The mass, momentum, and energy supplied to a constituent for any internal energy e_α and diffusive velocity \mathbf{u}_α must satisfy (9.54). We observe that by taking (9.54)₁ and (9.54)₂ into account, the fourth sum in (9.54) can be written,

$$\sum_{\alpha=1}^N \left[\hat{\varepsilon}_\alpha + \hat{\mathbf{p}}_\alpha \cdot \mathbf{v}_\alpha + \Gamma_\alpha \left(e_\alpha + \frac{1}{2} \mathbf{v}_\alpha \cdot \mathbf{v}_\alpha \right) \right] = 0 \quad (9.55)$$

We henceforth assume that all N constituents experience the same temperature $\theta = \theta(\mathbf{x}, t)$ at a

point \mathbf{x} in the mixture at time t : $\theta = \theta_1 = \theta_2 = \dots = \theta_N$.

An alternate form of (9.51) is obtained if we introduce the Helmholtz free energy per unit volume Ψ_α instead of ψ_α ,

$$\Psi_\alpha = \rho_\alpha \varphi_\alpha \psi_\alpha. \quad (9.56)$$

Thus one can show that,

$$\rho_\alpha \varphi_\alpha \frac{d^\alpha \psi_\alpha}{dt} = \frac{d\Psi_\alpha}{dt} - \Gamma_\alpha \psi_\alpha + \Psi_\alpha \operatorname{tr} \mathbf{L}_\alpha. \quad (9.57)$$

Returning to (9.51), we shall take for the entropy flux:

$$\mathbf{H}_\alpha = \frac{\mathbf{q}_\alpha}{\theta} + \rho_\alpha \varphi_\alpha \eta_\alpha \mathbf{u}_\alpha + \tilde{\mathbf{J}}_\alpha. \quad (9.58)$$

Here \mathbf{q}_α/θ is the standard entropy flux due to the influx of heat at an absolute temperature θ , $\rho_\alpha \varphi_\alpha \eta_\alpha \mathbf{u}_\alpha$ is the entropy flux due to the relative motion of the mixture to constituent α , and $\tilde{\mathbf{J}}_\alpha$ is the entropy flux due to chemical reactions, adhesion, and surface gradients in concentrations of volume fraction and will be defined more precisely below. Eliminating the energy supply $\hat{\varepsilon}_\alpha$ using (9.55), the entropy (Clausius-Duhem) inequality (9.51) assumes the form

$$\begin{aligned} & \sum_{\alpha=1}^N \left\{ \frac{1}{\theta} \left[-\frac{d^\alpha \Psi_\alpha}{dt} - \rho_\alpha \varphi_\alpha \eta_\alpha \frac{d^\alpha \theta}{dt} + \operatorname{tr} (\mathbf{T}_\alpha^T - \Psi_\alpha \mathbf{I}) \mathbf{L}_\alpha + \sum_{\beta=1}^L \zeta_{\alpha\beta} \frac{d^\alpha m_\beta}{dt} \right. \right. \\ & \quad \left. \left. + \sum_{\beta=1}^N \nabla \cdot \left(\boldsymbol{\sigma}_{\alpha\beta} \frac{d^\alpha \varphi_\beta}{dt} \right) - \left(\hat{\mathbf{p}}_\alpha + \frac{1}{2} \Gamma_\alpha \mathbf{v}_\alpha \right) \cdot \mathbf{v}_\alpha \right] \right. \\ & \quad \left. - \frac{1}{\theta^2} \mathbf{q}_\alpha \cdot \mathbf{g} + \nabla \cdot \tilde{\mathbf{J}}_\alpha \right\} \geq 0. \end{aligned} \quad (9.59)$$

General Forms of Constitutive Equations. The system of basic balance laws is closed by introducing constitutive equations for the dependent variables: mass, flux, free energy, Cauchy stress, internal energy or entropy, heat flux, and the momentum. Since we have ignored electromagnetic effects, we confine our attention hereafter to nonpolar materials, so that $\mathbf{M}_\alpha = 0$ and the partial stress tensors \mathbf{T}_α are symmetric. Throughout, we insist that the constitutive equations obey the restrictions imposed by the classical axiom of material frame indifference (as in Ref. [114]). Symbolically, we wish to supply frame indifferent constitutive equations for the thermomechanical fields

$$(\Psi_\alpha, \mathbf{T}_\alpha, \eta_\alpha, \mathbf{q}_\alpha, \hat{\mathbf{p}}_\alpha), \quad 1 \leq \alpha \leq N \quad (9.60)$$

in terms of an array Λ_α of independent state variables, with equations for other quantities, such as $\sigma_{\alpha\beta}$, $\zeta_{\alpha\beta}$, \mathbf{j}_α , $\hat{\mathbf{J}}_\alpha$. We assume that the mixture consists of M solid constituents and $N - M$ fluid components, each of which may be heterogeneous, but which are pointwise isotropic. The M solid constituents are assumed to be heterogeneous isotropic hyperelastic materials. The partial stresses for the fluid phases are assumed to consist of the sum of an equilibrium stress \mathbf{T}_α^e , which is characterized as that of a simple fluid (see Truesdell and Noll [113]) and a non-equilibrium thermoviscous stress \mathbf{T}_α^v , the form of which must be consistent with the principle of material frame indifference and the entropy inequality. We take

$$\Lambda_\alpha = \begin{cases} (\mathbf{X}_\alpha, \theta, \mathbf{g}, \mathbf{C}_\alpha, \boldsymbol{\varphi}, \nabla \boldsymbol{\varphi}, \hat{\mathbf{m}}_\alpha) & \text{for } \alpha \leq M \\ (\mathbf{x}_\alpha, \theta, \mathbf{g}, \mathbf{F}_\alpha, \boldsymbol{\varphi}, \nabla \boldsymbol{\varphi}, \hat{\mathbf{m}}_\alpha) & \text{for } M < \alpha \leq N \end{cases} \quad (9.61)$$

where

$$\left. \begin{array}{l} \mathbf{g} = \nabla \theta \\ \mathbf{C}_\alpha = \mathbf{F}_\alpha^T \mathbf{F}_\alpha \\ \boldsymbol{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_N) \\ \nabla \boldsymbol{\varphi} = (\nabla \varphi_1, \nabla \varphi_2, \dots, \nabla \varphi_N) \\ \hat{\mathbf{m}}_\alpha = (m_1 \varphi_\alpha, m_2 \varphi_\alpha, \dots, m_L \varphi_\alpha), \end{array} \right\} \quad (9.62)$$

\mathbf{C}_α being the right Cauchy-Green deformation tensor.

For a simple fluid, \mathbf{T}_α^e and \mathbf{q}_α can depend on \mathbf{F}_α only through $\det \mathbf{F}_\alpha$ or through the mass density $\hat{\rho}_\alpha$ (see, e.g. Batra[14]). Under these conventions, the constitutive equations assume the form,

$$\left. \begin{array}{ll} \Psi_\alpha = \Psi_\alpha(\Lambda_\alpha), & 1 \leq \alpha \leq N \\ \mathbf{T}_\alpha = \mathbf{T}_\alpha^e + \mathbf{T}_\alpha^v, & 1 \leq \alpha \leq N \\ \mathbf{T}_\alpha^e = \mathbf{T}_\alpha^{e*}(\Lambda_\alpha), & 1 \leq \alpha \leq N \\ \mathbf{T}_\alpha^v = \mathbf{T}_\alpha^{v*}(\Lambda_\alpha), & M < \alpha \leq N \end{array} \right\}. \quad (9.63)$$

It follows that the total time-rate change of the free energy is

$$\begin{aligned} \frac{d^\alpha \Psi_\alpha}{dt} &= \frac{\partial \Psi_\alpha}{\partial \theta} \dot{\theta} + \frac{\partial \Psi_\alpha}{\partial \mathbf{g}} \cdot \dot{\mathbf{g}} + \frac{\partial \Psi_\alpha}{\partial (\mathbf{C}_\alpha, \mathbf{F}_\alpha)} : \mathbf{D}_\alpha \\ &\quad + \sum_{\beta=1}^N (\mu_{\alpha\beta} \dot{\varphi}_\beta + \nabla \cdot \boldsymbol{\sigma}_{\alpha\beta} \dot{\varphi}_\beta) + \sum_{\gamma=1}^L \frac{\partial \Psi_\alpha}{\partial (m_\gamma \varphi_\alpha)} \varphi_\alpha \dot{m}_\gamma \\ &\quad + \sum_{\beta=1}^N \frac{\partial \Psi_\alpha}{\partial \nabla \varphi_\beta} \otimes (\nabla \dot{\varphi}_\beta - \nabla \mathbf{v}_\alpha \cdot \nabla \varphi_\beta) \end{aligned} \quad (9.64)$$

with

$$\frac{\partial \Psi_\alpha}{\partial(\mathbf{C}_\alpha, \mathbf{F}_\alpha)} : \mathbf{D}_\alpha = \begin{cases} \frac{\partial \Psi_\alpha}{\partial \mathbf{C}_\alpha} : \dot{\mathbf{C}}_\alpha = 2\mathbf{F}_\alpha^T \frac{\partial \Psi_\alpha}{\partial \mathbf{C}_\alpha} \mathbf{F}_\alpha : \mathbf{D}_\alpha, & \alpha \leq M \\ \frac{\partial \Psi_\alpha}{\partial \mathbf{F}_\alpha} : \dot{\mathbf{F}}_\alpha = \frac{\partial \Psi_\alpha}{\partial \mathbf{F}_\alpha} \mathbf{F}_\alpha^T : \mathbf{D}_\alpha, & M < \alpha \leq N \end{cases}, \quad (9.65)$$

and $\mu_{\alpha\beta}$ is the chemical potential

$$\mu_{\alpha\beta} = \frac{\partial \Psi_\alpha}{\partial \varphi_\beta} + \delta_{\alpha\beta} \sum_{\gamma=1}^L \frac{\partial \Psi_\alpha}{\partial(m_\gamma \varphi_\alpha)} m_\gamma - \nabla \cdot \boldsymbol{\sigma}_{\alpha\beta}. \quad (9.66)$$

Implementation of the coleman-Noll Principle alters considerably the algebra (see [82]). It can be shown that the Clausius-Duhem inequality assumes the form,

$$\begin{aligned} \sum_{\alpha=1}^N \left\{ \frac{1}{\theta} \left[- \left(\frac{\partial \Psi_\alpha}{\partial \theta} + \rho_\alpha \varphi_\alpha \eta_\alpha \right) \dot{\theta} - \frac{\partial \Psi_\alpha}{\partial \mathbf{g}} \cdot \dot{\mathbf{g}} + \text{tr} \left[\mathbf{T}_\alpha - \left(\Psi_\alpha - \sum_{\beta=1}^N \varphi_\alpha \mu_{\beta\alpha} \right) \mathbf{I} \right. \right. \right. \right. \\ \left. \left. \left. \left. - \frac{\partial \Psi_\alpha}{\partial(\mathbf{C}_\alpha, \mathbf{F}_\alpha)} + \sum_{\beta=1}^N \frac{\partial \Psi_\alpha}{\partial \nabla \varphi_\beta} \otimes \nabla \varphi_\beta \right] \mathbf{L}_\alpha + \sum_{\beta=1}^N \left(\boldsymbol{\sigma}_{\alpha\beta} - \frac{\partial \Psi_\alpha}{\partial \nabla \varphi_\beta} \right) \nabla \dot{\varphi}_\beta \right. \right. \\ \left. \left. \left. + \sum_{\beta=1}^N \left(\zeta_{\alpha\beta} - \frac{\partial \Psi_\alpha}{\partial(m_\beta \varphi_\alpha)} \varphi_\alpha \right) \dot{m}_\beta - \left(\hat{\mathbf{p}}_\alpha + \frac{1}{2} \Gamma_\alpha \mathbf{v}_\alpha - \sum_{\beta=1}^N \left[\frac{1}{\rho_\beta} \mu_{\alpha\beta} \nabla(\rho_\beta \varphi_\beta) \right. \right. \right. \right. \\ \left. \left. \left. \left. - \frac{1}{\rho_\alpha} \mu_{\beta\alpha} \nabla(\rho_\alpha \varphi_\alpha) \right] \right] \cdot \mathbf{v}_\alpha \right] + R_\alpha - \frac{1}{\theta^2} \mathbf{g} \cdot \mathbf{q}_\alpha - \sum_{\beta=1}^N \nabla \left(\frac{\mu_{\alpha\beta}}{\rho_\beta} \right) \cdot \tilde{\mathbf{J}}_\beta \right\} \geq 0, \end{aligned} \quad (9.67)$$

where $R_\alpha = - \sum_{\beta=1}^N \frac{\mu_{\alpha\beta}}{\rho_\beta} \left[S_\beta - \varphi_\beta \frac{d^\alpha \rho_\beta}{dt} \right]$ and we have taken $\tilde{\mathbf{J}}_\alpha = - \sum_{\beta=1}^N \frac{\mu_{\alpha\beta}}{\rho_\beta} \mathbf{J}_\beta$.

The classical method of Coleman and Noll [29] now can be implemented, whereby the rates in (9.67) are arbitrarily varied and, in order for the inequality to hold, it is sufficient that

$$\left. \begin{aligned} \frac{\partial \Psi_\alpha}{\partial \theta} &= -\rho_\alpha \varphi_\alpha \eta_\alpha \\ \frac{\partial \Psi_\alpha}{\partial \mathbf{g}} &= \mathbf{0} \\ \boldsymbol{\sigma}_{\alpha\beta} &= \frac{\partial \Psi_\alpha}{\partial \nabla \varphi_\beta} \\ \zeta_{\alpha\beta} &= \frac{\partial \Psi_\alpha}{\partial(m_\beta \varphi_\alpha)} \varphi_\alpha \\ \mathbf{T}_\alpha^e &= \left(\Psi_\alpha - \sum_{\beta=1}^N \mu_{\beta\alpha} \varphi_\alpha \right) \mathbf{I} + \frac{\partial \Psi_\alpha}{\partial(\mathbf{C}_\alpha, \mathbf{F}_\alpha)} - \sum_{\beta=1}^N \frac{\partial \Psi_\alpha}{\partial \nabla \varphi_\beta} \otimes \nabla \varphi_\beta \end{aligned} \right\} \quad (9.68)$$

and

$$\begin{aligned} \sum_{\alpha=1}^N \left\{ \mathbf{T}_\alpha^{v*}(\boldsymbol{\Lambda}_\alpha) : \mathbf{D}_\alpha + R_\alpha - \sum_{\beta=1}^N \nabla \left(\frac{\mu_{\beta\alpha}}{\rho_\alpha} \right) \cdot \mathbf{J}_\alpha - \frac{1}{\theta} \left[\hat{\mathbf{p}}_\alpha + \frac{1}{2} \Gamma_\alpha \mathbf{v}_\alpha \right. \right. \\ \left. \left. - \sum_{\beta=1}^N \left(\frac{1}{\rho_\beta} \mu_{\alpha\beta} \nabla(\rho_\beta \varphi_\beta) - \frac{1}{\rho_\alpha} \mu_{\beta\alpha} \nabla(\rho_\alpha \varphi_\alpha) \right) \right] \cdot \mathbf{v}_\alpha \right\} \geq 0. \end{aligned} \quad (9.69)$$

It is sufficient to take

$$\hat{\mathbf{p}}_\alpha = -\frac{1}{2} \Gamma_\alpha \mathbf{v}_\alpha + \sum_{\beta=1}^N \left(\frac{1}{\rho_\beta} \mu_{\alpha\beta} \nabla(\rho_\beta \varphi_\beta) - \frac{1}{\rho_\alpha} \mu_{\beta\alpha} \nabla(\rho_\alpha \varphi_\alpha) \right) - \lambda_\alpha \mathbf{v}_\alpha \quad (9.70)$$

with $\lambda_\alpha > 0$, for the bracketed term in (9.69) to be non-negative. For other choices, see Bowen[20] [p. 39], Cristini et al [34], and Araujo and MaElwain [9].

The last term on the right-hand side of inequality (9.69) will be non-negative if we take the mass flux

$$\mathbf{J}_\alpha = - \sum_{\beta=1}^N M_{\beta\alpha}(\boldsymbol{\varphi}, \mathbf{m}_\alpha) \nabla(\mu_{\beta\alpha}/\rho_\alpha), \quad (9.71)$$

where $M_{\beta\alpha}$ is a symmetric positive-semi definite matrix, possibly dependent on the volume fractions $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_N)$ and the taxis factors $m_1 \varphi_\alpha, m_2 \varphi_\alpha, \dots, m_L \varphi_\alpha$, called the mobility of the mixture. The last term on the left-hand side of (9.69) is then always non-negative.

Partial stress and heat flux. Recalling the notation introduced in (9.62), it is observed that the mechanical part of the equilibrium partial stress $\hat{\mathbf{T}}_\alpha^e$ (that in response to deformation and flow) can be written as

$$\left. \begin{aligned} \hat{\mathbf{T}}_\alpha^e &= 2 \mathbf{F}_\alpha^T \frac{\partial \Psi_\alpha}{\partial \mathbf{C}_\alpha} \mathbf{F}_\alpha = \frac{1}{\det \mathbf{F}_\alpha} \mathbf{F}_\alpha \mathbf{S}_\alpha \mathbf{F}_\alpha^T, & \alpha \leq M \\ \hat{\mathbf{T}}_\alpha^e &= \frac{\partial \Psi_\alpha}{\partial \mathbf{F}_\alpha} \mathbf{F}_\alpha^T = \frac{\partial \Psi_\alpha}{\partial \hat{\rho}_\alpha} \frac{\partial \hat{\rho}_\alpha}{\partial \mathbf{F}_\alpha^T} = -\pi_\alpha, & M < \alpha \leq N \end{aligned} \right\}, \quad (9.72)$$

where $\rho_{0\alpha} \varphi_{0\alpha} = \hat{\rho}_{0\alpha}$ is the mass density of the α^{th} constituent in the reference configuration, \mathbf{S}_α is the partial second Piola-Kirchhoff stress tensor,

$$\mathbf{S}_\alpha = \det \mathbf{F}_\alpha \mathbf{F}_\alpha^{-1} \hat{\mathbf{T}}_\alpha \mathbf{F}_\alpha^{-T} = 2 \frac{\partial W_\alpha}{\partial \mathbf{C}_\alpha}, \quad \alpha \leq M \quad (9.73)$$

with $W_\alpha = \hat{\rho}_{0\alpha} \psi_\alpha$, and π_α is the thermodynamic pressure,

$$\pi_\alpha(\theta, \hat{\rho}_\alpha, \boldsymbol{\varphi}, \mathbf{m}_\alpha) = -\hat{\rho}_\alpha^2 \frac{\partial \psi_\alpha}{\partial \hat{\rho}_\alpha}, \quad M < \alpha \leq N.$$

9.4 Mixture Theory Models of Avascular Tumor Growth

The continuum theory of mixtures laid down in preceding sections of this chapter provides a general framework for developing phenomenological models of the behavior of heterogeneous masses of interacting media. It is therefore natural to explore its use in developing mathematical abstractions of the evolution of the complex microenvironment of tumors embedded in a matrix of interacting cellular structures.

An abstraction of such an environment is depicted in Figure 9.3, where we observe a colony of tumor cells in various stages of growth or decline, healthy cells, and other relevant constituents. The structured framework supporting this micro-cellular-environment is called the *stroma*, and it is made up of collagen together with a collection of extracellular molecules called the *extra cellular matrix* (ECM). Cells called *fibroblasts* synthesize the collagen and ECM and make up the main connective tissue cells present in the body. Also present are *macrophages*, large white blood cells that engulf and digest cellular debris. *Endothelial cells* form the lining of blood vessels and lymphatic vessels.

The use of mixture theory to model such systems requires that we first attempt to identify the key constituents in an N-species representation of the microenvironment that prevails and contributes to the growth of tumors in living tissue. We recall that cancer manifests itself in a series of stages in which genetic mutations of cells causes overexpression of oncogenes, the genes causing proliferation of tumor cells, or underexpression of mechanisms that control or suppress cell growth. Thereby, these genes alter homeostasis and lead to the creation of colony of aberrant cells. This mutation process may take years to progress, or can be accelerated by other factors in the microenvironment. Oncogenes are activated by sub-cell growth signals that promote cell proliferation or halt the cell cycle.

So at the tissue-scale, tumor cells that depend on oxygen and other nutrients to survive, may enter the so-called *hypoxic state* between a state of proliferation, in which the possibility of mitosis occurs, and a state of unplanned cell death, called *necrosis*. In the necrotic state, water enters the cell, increasing pressure until it bursts, and the cell undergoes calcification going into a dormant state of microcalcification. This cell death is different from apoptosis, programmed cell death, that occurs when the cells are no longer needed to preserve the latent architecture of the tissue or when they pose a threat to the organism. Thus, a colony of tumor cells can contain proliferative cells which can grow in number due to mitosis, hypoxic cells which are in essentially a suspended state that depends on the local nutrient and oxygen supply, and the necrotic cells.

The general tumor growth scenario is this: mass of such tumor cells establishes itself in its host

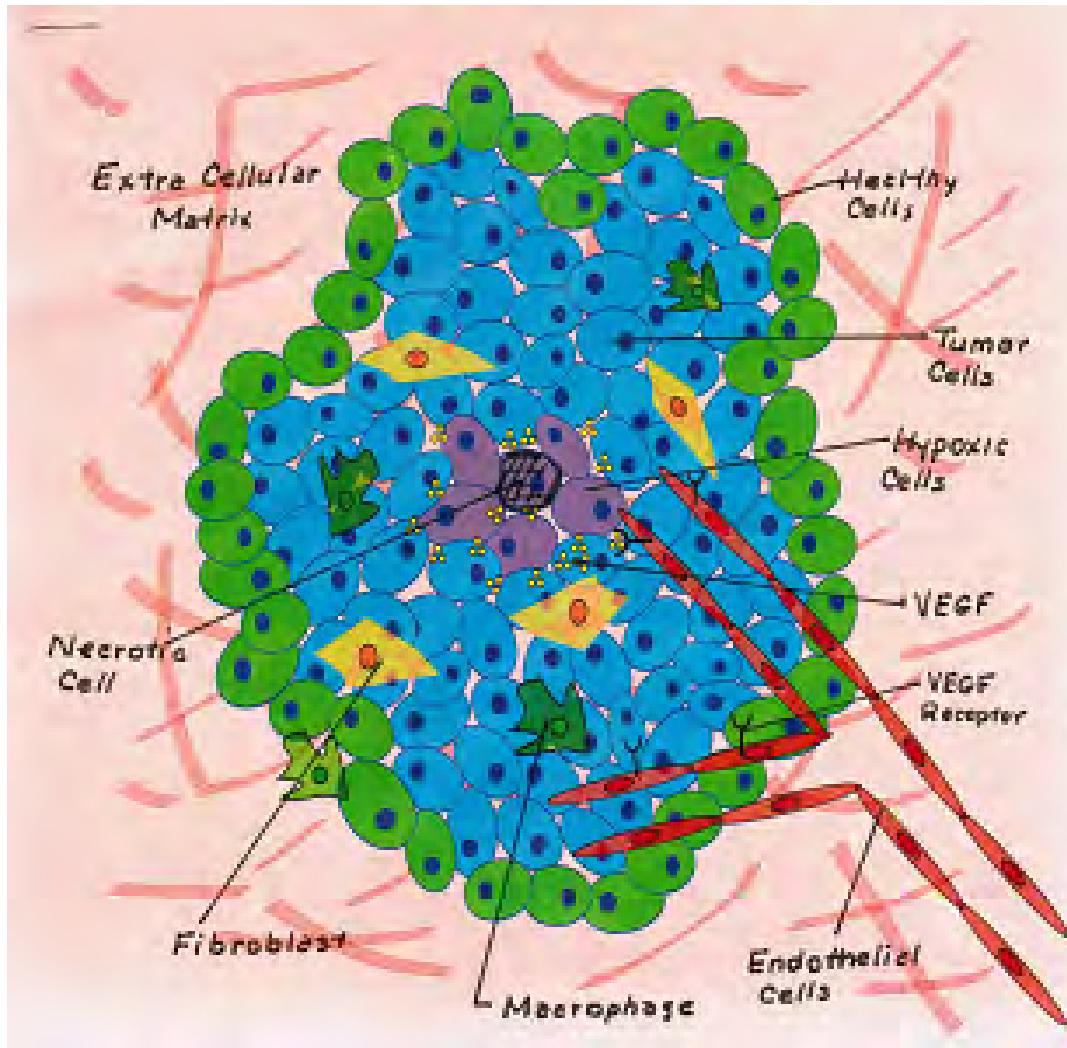


Figure 9.3: An schematic figure of tumor environment and different involved in modeling tumor growth.

tissue surrounded by its microenvironment, beginning with no vascular structure of its own. The survival of this collection of cells depends on its neighboring cell structure, the stroma for a supply of nutrient and growth factors to stimulate its growth. Without sufficient nutrients, an inner core of hypoxic cells forms and eventually some die, creating a necrotic core of dead cells. The outer cells, those with sufficient nutrients and growth factor concentrations, proliferate and cause an outward flux of cells that expands the tumor, causing mechanical compression of surrounding tissue and the base membrane that supports the tumor mass. The proliferating cells also absorb interstitial fluid to promote mitosis and growth of the mass of the tumor. The growing tumor interacts with the ECM, the surrounding environment, which is degraded and modified by the secretion of special enzymes. The ECM, in turn, release growth factors that refuel tumor growth.

Angiogenesis. If insufficient nutrients are available to promote proliferation, hypoxia occurs which, remarkably results in the release of proteins collectively called *vascular endothelial growth factors* (VEGF), which are among a group of proteins labeled *tumor angiogenesis factors* (TAFs). These TAFs diffuse outward from the nutrient-deprived hypoxic cells and invade regions outside the tumor, eventually reaching nearby blood vessels, coming into contact with endothelial cells that form the walls of the blood vessels. When these endothelial cells detect a gradient in the TAF they secret special enzymes such as *matrix metalloproteinase* (MMPs) that degrade the ECM and the base membrane allowing the endothelial cells to migrate away from the parent blood vessel toward the TAF source through *sprouts-fingers* of blood vessels that gradually growth toward the hypoxic cell mass in micro-vascular tubes that, on contact with the hypoxic cells, supply the needed nutrients for cell proliferation. The hypoxic cells that are supplied these nutrients, are thus transformed to proliferative cells. These new vessels form a network around the tumor to supply nutrients and growth factors. With this supply of nutrients, the tumor growth accelerates, eventually invading parent blood vessels and entering the blood stream and spreading to other parts of the body. This process is called metastasis and is the onset of cancer.

Model Construction. To begin to construct models that capture all these events, it is clear that the principle of conservation of mass of each of the latent species is of fundamental importance. Referring to (9.43), these conservation laws take the form,

$$\frac{\partial \rho_\alpha \phi_\alpha}{\partial t} + \nabla \cdot (\rho_\alpha \phi_\alpha \mathbf{v}_\alpha) = S_\alpha - \nabla \cdot \mathbf{J}_\alpha , \quad (9.74)$$

$1 \leq \alpha \leq N$, where ρ_α is the mass per unit volume of constituents α , ϕ_α is the corresponding volume fraction, \mathbf{v}_α is the velocity of constituent α , S_α is the mass supplied by other constituents to constituent α , and \mathbf{J}_α is the mass flux. To (9.74) we add the equations of balance of linear and

angular momentum,

$$\operatorname{div} \mathbf{T}_\alpha + \rho_\alpha \phi_\alpha \mathbf{b}_\alpha + \hat{\mathbf{p}}_\alpha = \rho_\alpha \frac{d^\alpha \mathbf{v}_\alpha}{dt} , \quad \mathbf{T}_\alpha - \mathbf{T}_\alpha^T = \mathbf{M}_\alpha , \quad (9.75)$$

\mathbf{I}_α being the partial Cauchy stress, \mathbf{b}_α the body factor per unit mass, $\hat{\mathbf{p}}_\alpha$ the momentum supplied by neighboring constituents, and \mathbf{M}_α the intrinsic moment at points within constituent α . The equations describing the conservation of energy and the Clausius-Duhem inequality could be added to these equations.

Next we must select basic dependent variables that are suggested by the conservation laws (9.74) and (9.75) and that characterize the system described up to this point. Following the development in Lima, Oden, and Almeida [73], we identify the principal volume fraction fields as follows:

$$\alpha = 1 , \quad \phi_1 = \phi_T = \text{volume fraction of tumor cells} \quad (9.76)$$

The tumor cells are divided into three categories: proliferative cells ϕ_P , hypoxic cells ϕ_H , and necrotic cells ϕ_N , so that

$$\phi_T = \phi_P + \phi_H + \phi_N . \quad (9.77)$$

It is important to track the evolution of proliferative, hypoxic, and necrotic cells independently, so we choose ϕ_P , ϕ_H , ϕ_N as unknown dependent variables and calculate ϕ_T using (9.77).

In a closed system, the numbers of tumor cells and healthy cells ϕ_C , excluding endothelial cells is essentially constant: $\phi_T + \phi_C = A = \text{constant}$. We set $\phi_2 = \phi_C = A - \phi_T$.

Nutrients, including oxygen and glucose, are dissolved in the aqueous liquid surrounding the cellular structure, which is essentially water, making up much of the inner-cellular material environment. The cellular water can be rich in nutrients, or poor in nutrients, depending on whether the cells have or have not absorbed nutrients to promote proliferation or to combat hypoxia. Thus, another constituent $\alpha = 3 \sim \phi_\sigma$ can be assigned to nutrient rich extracellular water, and $\alpha = 4 \phi_{\sigma 0}$ can be assigned to its complement nutrient poor extracellular water, so that $\phi_\sigma + \phi_{\sigma 0} = B$, a constant representative of the total water volume fraction.

Similar complementary relations can be postulated for protein-rich TAF, with $\phi_5 = \phi_{TAF}$, and protein poor TAF, $\phi_6 = \phi_{TAF 0}$, with $\phi_{TAF} = \phi_{TAF 0} = C = \text{constant}$, and for proliferative endothelial cell volume fraction $\phi_T = \phi_e$, and not-activated endothelial cell volume fractions $\phi_8 =$

ϕ_{e0} , with $\phi_e + \phi_{e0} = 0 = \text{constant}$. The saturation condition (9.38) requires that

$$\underbrace{\phi_T + \phi_C}_{A} + \underbrace{\phi_\sigma + \phi_\sigma 0}_{B} + \underbrace{\phi_{TAF} + \phi_{TAF0}}_{C} + \underbrace{\phi_e + \phi_e 0}_{D} = 1 , \quad (9.78)$$

for eight constituents. We next introduce simplifying assumptions. We will assume the material is non-polar, meaning the intrinsic moments $\mathbf{M}_\alpha = 0$, so that the partial stress tensor \mathbf{T}_α are symmetric. Since all constituents can be assumed to have a mass density close to that of water. We take $\rho_\alpha = \tilde{\rho} = \text{constant} = 1$, $1 \leq \alpha \leq 8$. All thermal effects are ignored (e.g. $\mathbf{q}_\alpha = \mathbf{0}$, $\boldsymbol{\theta} = \text{const.}$, etc.) and the body forces \mathbf{b}_α are assumed to be negligible. The mass flux \mathbf{J}_α is assumed to be of the form (9.71) depends on the chemical potential $\mu_{\alpha\beta}$. With these assumptions in force, (9.74) leads to the following systems of mass balance equations:

$$\left. \begin{aligned} \frac{\partial \phi_P}{\partial t} + \nabla \cdot (\phi_P \mathbf{v}_P) &= -\nabla \cdot \mathbf{J}_P + S_P, \\ \frac{\partial \phi_H}{\partial t} + \nabla \cdot (\phi_H \mathbf{v}_H) &= -\nabla \cdot \mathbf{J}_H + S_H, \\ \frac{\partial \phi_N}{\partial t} + \nabla \cdot (\phi_N \mathbf{v}_N) &= -\nabla \cdot \mathbf{J}_N + S_N, \\ \frac{\partial \phi_\sigma}{\partial t} + \nabla \cdot (\phi_\sigma \mathbf{v}_\sigma) &= -\nabla \cdot \mathbf{J}_\sigma + S_\sigma, \\ \frac{\partial \phi_{TAF}}{\partial t} + \nabla \cdot (\phi_{TAF} \mathbf{v}_{TAF}) &= -\nabla \cdot \mathbf{J}_{TAF} + S_{TAF}, \\ \frac{\partial \phi_e}{\partial t} + \nabla \cdot (\phi_e \mathbf{v}_e) &= -\nabla \cdot \mathbf{J}_e + S_e, \end{aligned} \right\} \quad (9.79)$$

The relevant mass fluxes are:

$$\left. \begin{aligned} \mathbf{J}_P &= -M_P \nabla D_{\phi_P} E, \\ \mathbf{J}_H &= -M_H \nabla D_{\phi_H} E, \\ \mathbf{J}_N &= -M_N \nabla D_{\phi_N} E, \\ \mathbf{J}_\sigma &= -M_\sigma \nabla D_{\phi_\sigma} E, \\ \mathbf{J}_{TAF} &= -M_{TAF} \nabla D_{\phi_{TAF}} E, \\ \mathbf{J}_e &= -M_e \nabla D_{\phi_e} E, \end{aligned} \right\} \quad (9.80)$$

$M_P(\cdot), M_H(\cdot), \dots, M_e(\cdot)$ are mass-flux mobilities, where $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_8)$ and μ_P, \dots, μ_e are corresponding chemical potentials.

The chemical potentials for multi-species mixtures represent the energy absorbed or released during a chemical reaction or a phase change, and generally is given by the derivative of the energy with respect to the amount of the species, all other species in the mixture remaining constant. In the theory described in previous sections, higher-order terms appear due to surface energy effects represented by gradients of the volume fractions. The correct characterization of the chemical potentials is then as variational or Gateaux derivatives of the free energy. The total energy is given by,

$$E = \int_{\Omega} dx \left(\Psi_T(\boldsymbol{\phi}) + \Psi_\sigma(\boldsymbol{\phi}) + \Psi_{TAF}(\boldsymbol{\phi}) + \Psi_e(\boldsymbol{\phi}) \right), \quad (9.81)$$

where,

$$\left. \begin{aligned} \Psi_T(\boldsymbol{\phi}) &= f(\boldsymbol{\phi}_T) + \frac{\epsilon_T^2}{2} |\nabla \phi_T|^2 + W(\mathbf{C}_T, \phi_T), \\ \Psi_\sigma(\boldsymbol{\phi}) &= \frac{1}{2} \delta_\sigma^{-1} \phi_\sigma^2 - \chi_0 \phi_\sigma \phi_T, \\ \Psi_{TAF}(\boldsymbol{\phi}) &= \frac{1}{2} \delta_{TAF}^{-1} \phi_{TAF}^2, \\ \Psi_e(\boldsymbol{\phi}) &= f(\boldsymbol{\phi}_e) + \frac{\epsilon_e^2}{2} |\nabla \phi_e|^2 + W(\mathbf{C}_e, \phi_e). \end{aligned} \right\} \quad (9.82)$$

Here $f(\cdot)$ is the double-well potential,

$$f(\phi) = \gamma_\phi \phi^2 (1 - \phi)^2 \quad (9.83)$$

and $W(\mathbf{C}_\alpha, \phi_\alpha)$, ($\alpha = T, e$) are the stored energy functions supplied due to mechanical deformation of tumor and endothelial cells and γ_ϕ is an energy scale parameter associated with ϕ . The chemical potentials are then,

$$\left. \begin{aligned} \mu_T &= f'(\phi_T) - \epsilon_T^2 \Delta \phi_T - \chi_0 \phi_T + \partial W(\mathbf{C}, \phi_T) , \\ \mu_H &= \mu_T , \\ \mu_N &= \mu_T , \\ \mu_\sigma &= \delta_\sigma^{-1} \phi_\sigma - \chi_0 \phi_T \phi_T , \\ \mu_{TAF} &= f'(\phi_e) - \epsilon_\sigma^2 \Delta \phi_e - \partial W(\mathbf{C}, \phi_e) / \partial \phi_e . \end{aligned} \right\} \quad (9.84)$$

The source terms S_P , S_H , S_N , S_σ , S_{TAF} , and S_e in (9.84) are to be assigned functional forms that depict and are consistent with the Hallmarks of Cancer given in Chapter 8 (recall Figure 8.26). A key feature is the threshold at which various biological events are initiated or terminated in the organism dictated by sub-cellular phenomena, protein-signaling, and genetics. For example, σ_{PH} may be introduced as a threshold value of the amount of nutrient ϕ_σ in which cancer cells are transformed from a proliferative state (ϕ_P) to an hypoxic state (ϕ_H), suggesting a decline in nutrients needed sufficient to transfer hypoxic cells to a proliferative state with higher probability of proliferation. This would be captured by terms of the form,

$$-\lambda_{PH} \mathcal{H}(\sigma_{PH} - \phi_\sigma) + \lambda_{HP} \mathcal{H}(\phi_\sigma - \phi_{HP}) ,$$

λ_{PH} and λ_{HP} being material parameters, H being the Heaviside step function,

$$\mathcal{H}(x) = \begin{cases} 1 & x > 0 , \\ 0 & x \leq 0 . \end{cases}$$

The proliferative cells volume fraction ϕ_P can grow when assuming nutrient at a constant rate of cell mitosis of λ_P and when hypoxic cells return to a proliferative state when nutrient levels exceed the threshold value σ_{HP} . They may also decay due to programmed death, apoptosis, at a rate λ_A . The hypoxic cells can irreversibly become part of the necrotic core when ϕ_σ drops below a necrotic

threshold σ_{HN} at a constant rate λ_{HN} . The necrotic core can never decrease and are included in a process of calcification. Following Lima et al [73] we observe that the TAF-rich extracellular water volume fraction is produced by hypoxic cells at a rate λ_{TAF} and the growth levels off when TAF-rich extracellular water volume fraction approaches its maximum value. It is also uptaken by proliferative endothelial cells with a volume-action coefficient α_{TAF} during angiogenesis. Both the proliferation of endothelial stalk cells and the TAF-rich uptaken by endothelial cells are regulated by the discrete angiogenesis model so that they are activated whenever TAF-rich volume fraction is higher than the activation threshold ϕ_{TAF_c} . The generation of new endothelial cells occurs at the proliferation rate $\alpha_P(\phi_{TAF})$ that depends on ϕ_{TAF} . When $\phi_{TAF} < \phi_{ref}$, this rate is given by $\alpha_P\phi_{TAF}$. Otherwise, there is a saturation in the proliferation so that this rate reaches its highest value $\alpha_P\phi_{ref}$. Finally, the proliferative endothelial cells provide an increase of the nutrient amount at a rate α_R . The growth levels off in regions where the nutrient approaches its maximum volume fraction. With these assumptions, the source terms take the form

$$\left. \begin{aligned} S_P &= \lambda_P\phi_\sigma\phi_P - \lambda_A\phi_P - \lambda_{PH}\mathcal{H}(\sigma_{PH} - \phi_\sigma)\phi_P \\ &\quad + \lambda_{HP}\mathcal{H}(\phi_\sigma - \sigma_{HP})\phi_H, \\ S_H &= -\lambda_A\phi_H + \lambda_{PH}\mathcal{H}(\sigma_{PH} - \phi_\sigma)\phi_P - \lambda_{HP}\mathcal{H}(\phi_\sigma - \sigma_{HP})\phi_H \\ &\quad - \lambda_{HN}\mathcal{H}(\sigma_{HN} - \phi_\sigma)\phi_H, \\ S_N &= \lambda_{HN}\mathcal{H}(\sigma_{HN} - \phi_\sigma)\phi_H, \\ S_\sigma &= -\lambda_P\phi_\sigma\phi_P - \lambda_{P_h}\phi_\sigma\phi_H + \lambda_A(\phi_P + \phi_H) + \alpha_R(1 - \phi_\sigma)\phi_e, \\ S_{TAF} &= \lambda_{TAF}(1 - \phi_{TAF})\phi_H - \alpha_{TAF}\phi_{TAF}\phi_e\mathcal{H}(\phi_{TAF} - \phi_{TAF_c}), \\ S_e &= \alpha_P(\phi_{TAF})\phi_e\mathcal{H}(\phi_{TAF} - \phi_{TAF_c}), \end{aligned} \right\} \quad (9.85)$$

Assuming that the velocities of the constituents are negligible, the mass fluxes for each constituent

are given by (9.80), and considering $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$), the system (9.79) can be rewritten as

$$\left. \begin{aligned} \frac{\partial \phi_T}{\partial t} &= \nabla \cdot \mathcal{M}(\phi_T, \phi_H, \phi_N) \nabla \mu + S_T, \\ \mu &= \Psi'(\phi_T) - \epsilon_T^2 \Delta \phi_T - \chi_0 \phi_\sigma, \\ \frac{\partial \phi_H}{\partial t} &= \nabla \cdot \bar{M}_H \phi_H^2 \nabla \mu + S_H, \\ \frac{\partial \phi_N}{\partial t} &= \nabla \cdot \bar{M}_N \phi_N^2 \nabla \mu + S_N, \\ \frac{\partial \phi_\sigma}{\partial t} &= \nabla \cdot M_\sigma \delta_\sigma^{-1} \nabla \phi_\sigma - \nabla \cdot M_\sigma \chi_0 \nabla \phi_T + S_\sigma, \\ \frac{\partial \phi_{TAF}}{\partial t} &= \nabla \cdot M_{TAF} \delta_{TAF}^{-1} \nabla \phi_{TAF} + S_{TAF}, \\ \frac{\partial \phi_e}{\partial t} &= \nabla \cdot \bar{M}_e \phi_e^2 \nabla \mu_e + S_e, \\ \mu_e &= \Psi'(\phi_e) - \epsilon_e^2 \Delta \phi_e, \end{aligned} \right\} \text{in } \Omega \times (0, T), \quad (9.86)$$

where $S_T = \lambda_P \phi_\sigma (\phi_T - \phi_H - \phi_N) - \lambda_A (\phi_T - \phi_N)$. Here we take,

$$\left. \begin{aligned} \mathcal{M}(\phi_T, \phi_H, \phi_N) &= \bar{M}_P (\phi_T - \phi_H - \phi_N)^2 + \bar{M}_H \phi_H^2 + \bar{M}_N \phi_N^2, \\ M_H &= \bar{M}_T \phi_H^2, \\ M_N &= \bar{M}_N \phi_N^2, \\ M_e &= \bar{M}_e \phi_e^2, \end{aligned} \right\} \quad (9.87)$$

where \bar{M}_P , \bar{M}_H , \bar{M}_N , and \bar{M}_e are positive constants (or time-independent densities). Among typical boundary conditions on the boundary $\partial\Omega$ over a time interval $(0, \tau)$, we list the following:

$$\left. \begin{aligned} \nabla \phi_T \cdot \mathbf{n} &= \nabla \mu_T \cdot \mathbf{n} = \nabla \phi_H \cdot \mathbf{n} = \nabla \phi_N \cdot \mathbf{n} = \nabla \mu_\sigma \cdot \mathbf{n} = 0, \\ \nabla \phi_{TAF} \cdot \mathbf{n} &= \nabla \phi_e \cdot \mathbf{n} = \nabla \mu_e \cdot \mathbf{n} = \nabla \phi_\sigma \cdot \mathbf{n} = 0, \end{aligned} \right\} \quad (9.88)$$

on $\partial\Omega \times (0, \tau)$, \mathbf{n} being a unit exterior normal to $\partial\Omega$. Numerical implementations of these models

and the result several numerical experiments are reported in [73].

9.5 An Application of OPAL to the Selection and Validation of Models of Tumor Growth

To demonstrate the methods of predictive computational science presented in Chapter 1-8, we now draw from the study of Lima et al [74] examples of application of OPAL to representative classes of models of tumor growth. These involve reaction diffusion models: for tumor cell volume fraction ϕ_T , generalized diffusion models characterized by system of the form 9.86, and mechanical effects based on simple models of elastic deformation. In this case, we take for the energy E of 9.81,

$$E = \int_{\Omega} (\Psi(\phi_T, \nabla \phi_T) + W(\boldsymbol{\epsilon}(\mathbf{u}), \phi_T)) \, d\mathbf{x} \quad (9.89)$$

where $W(\boldsymbol{\epsilon}(\mathbf{u}), \phi_T)$ is the strain energy density,

$$W = \frac{1}{2} \boldsymbol{\epsilon} : \mathbf{C}(\phi_T) \boldsymbol{\epsilon} + \boldsymbol{\epsilon} : \mathcal{T}(\phi_T) \quad (9.90)$$

$\boldsymbol{\epsilon}$ being the strain tensor for small displacement gradients $\nabla \mathbf{u}(\boldsymbol{\epsilon}(\mathbf{u})) = ((\nabla \mathbf{u} + \nabla \mathbf{u}^T)/2) \cdot \mathbf{C}$ the fourth-order elastic compliance tensor, and $\mathcal{T}(\phi_T)$ is a symmetric compositional stress tensor dependent on the tumor volume fraction ϕ_T . We take

$$\mathbf{C}(\phi_T) = \mathbf{C}^h + g(\phi_T)(\mathbf{C}^t - \mathbf{C}^h), \quad (9.91)$$

\mathbf{C}^h being the compliance tensor for healthy cells, \mathbf{C}^t that for tumor cells, and $g(\phi_T)$ a smooth interpolation function with property, $g(\phi_T = 1) = 1$, $g(\phi_T = 0) = 0$. The chemical potential $\mu (= \mu_{\alpha\beta} = \mu_\alpha)$ and the Cauchy stress are

$$\left. \begin{aligned} \mu &= \frac{\partial \Psi}{\partial \phi_T} - \nabla \cdot \frac{\partial \Psi}{\partial \nabla \phi_T} + \frac{\partial W(\phi_T, \boldsymbol{\epsilon}(\mathbf{u}))}{\partial \phi_T}, \\ \mathbf{T} &= \frac{\partial W}{\partial \boldsymbol{\epsilon}} - \frac{\partial \Psi}{\partial \nabla \phi_T} \otimes \nabla \phi_T . \end{aligned} \right\} \quad (9.92)$$

Owing to the fact that the diffusion mechanisms in the tissue take place over a timescale much larger than that associated with inertia effects, we consider a quasi-static deformation governed by the momentum balance,

$$\frac{1}{2} \nabla \cdot [\mathbf{C}(\phi_T)(\nabla \mathbf{u} + \nabla \mathbf{u}^T)] + \lambda \nabla \phi_T = \mathbf{0} , \quad (9.93)$$

where we have set $\mathcal{T}(\phi_T) = \lambda\phi_T\mathbf{I}$, with \mathbf{I} being the identity tensor. As a first approximation, we assume the method is isotropic and homogeneous, and we impose Hooke's law, giving

$$\nabla \cdot G\nabla \mathbf{u} + \nabla \frac{G}{1 - 2\nu} (\nabla \cdot \mathbf{u}) - \lambda \nabla \phi_T = 0 , \quad (9.94)$$

G being the shear modulus of the material.

With these conventions, we have introduced a very simple model of mechanical deformations of the tumor cells, leaving the principal sources of tumor evolution to the conservation of mass, which is dependent on the form of the free energy per unit volume, $\Psi(\phi_T, \nabla\phi_T)$. Three classes of models are considered:

I. Reaction Diffusion Models ($\mu = c\phi_T$)

$$\frac{\partial\phi_T}{\partial t} = \nabla \cdot M_T^* \nabla \phi_T + \lambda_T^{grow} \phi_T (1 - \phi_T), \quad (9.95)$$

II. Reaction Diffusion Models with Mechanical Coupling

$$\begin{cases} \frac{\partial\phi_T}{\partial t} = \nabla \cdot M_T(\phi_T) \nabla \mu + \lambda_T^{grow} \phi_T (1 - \phi_T); \\ \mu = c\phi_T - \lambda \nabla \cdot \mathbf{u}. \end{cases} \quad (9.96)$$

III. Phase-Field Models

$$\begin{cases} \frac{\partial\phi_T}{\partial t} = \nabla \cdot M_{PF}(\phi_T) \nabla \mu + \lambda_T^{grow} \phi_T (1 - \phi_T); \\ \mu = 2\bar{E}\phi_T (1 - 3\phi_T + 2\phi_T^2) - \epsilon_T^2 \Delta \phi_T - \lambda \nabla \cdot \mathbf{u}. \end{cases} \quad (9.97)$$

To account for the effects of mechanical stress on tumor growth, we choose a mobility coefficient that is damped by a stress measure Σ according to

$$M_T \sim \bar{M}_T \exp(-\gamma\Sigma) , \quad M_T^* \sim M_T^* \exp(-\gamma\Sigma) \quad (9.98)$$

and

$$\lambda_T^{grow} = \lambda_T^{grow} \exp(-\gamma^{grow}\Sigma) \quad (9.99)$$

where Σ is taken to be the Von Mises stress and γ, γ^{grow} are constants.

For illustration purpose, we consider a set of 13 possible models, listed in Table 9.1, in the following categories: RD (reaction diffusion), MD (reaction diffusion with mechanical deformation) and PF (phase field). The parameters in each model are listed in the table. Experimental data were obtained from [57] on growth of glioma (brain) cancer in a murine model (laboratory rates) acquired using diffusing weighted magnetic resonance imaging (DW-MMRI), beginning 10 days after implantation and collected on day 10, 12, 14, 15, 16, 18, and 20. Detailed can be found in [57, 74]. The observed progression of the tumor mass over 20 days, shown at a 20-slice selected so that tumor is observed at all measurement samples, is reproduced in Figure 9.4.

Table 9.1: Initial set of possible models.

Model	Variables			Parameters								# Params	Occam	Category			
	ϕ_T	μ	\mathbf{u}	M_T	M_T^*	c	λ_T^{grow}	\bar{E}_T	ϵ_T	E	ν	λ	γ	γ^g			
RD01	✓					✓		✓							2	1	
PF01	✓	✓		✓			✓	✓	✓						4	2	
RD02	✓		✓		✓		✓			✓	✓	✓	✓		6	3	
RD03	✓		✓		✓		✓			✓	✓	✓	✓	✓	6	3	
MD01	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓		6	3	
RD04	✓		✓		✓		✓			✓	✓	✓	✓	✓	7	4	
MD02	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	7	4	
MD03	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	7	4	
PF02	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	7	4	
MD04	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	8	5	
PF03	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	8	5	
PF04	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	8	5	
PF05	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	9	6	

Sensitivity Analysis

The elementary effects method (Morris method) described in Section 6.4 , is employed for a sensitivity analysis using the voxel domain shown in Figure 9.5 using 50 trajectories for each model. Table 9.3 shows the results for all the models in \mathcal{M} . Clearly, the area of the tumor is more sensitive to the rate of tumor growth. While the influence of some of the parameters is very small (on the order of 10^{-7}), we elect to retain all the models for the next OPAL process.

Model Calibration and Validation using MCMC

A set of uniform distributions was used as priors for each parameter and Occam Category 1 is first evaluated (Tables 9.2 and 9.4. Model RD01 is a reaction-diffusion model without mechanical deformation, so only parameters M_T^* and λ_T^{grow} need calibration. In Figure 9.6, the prior and

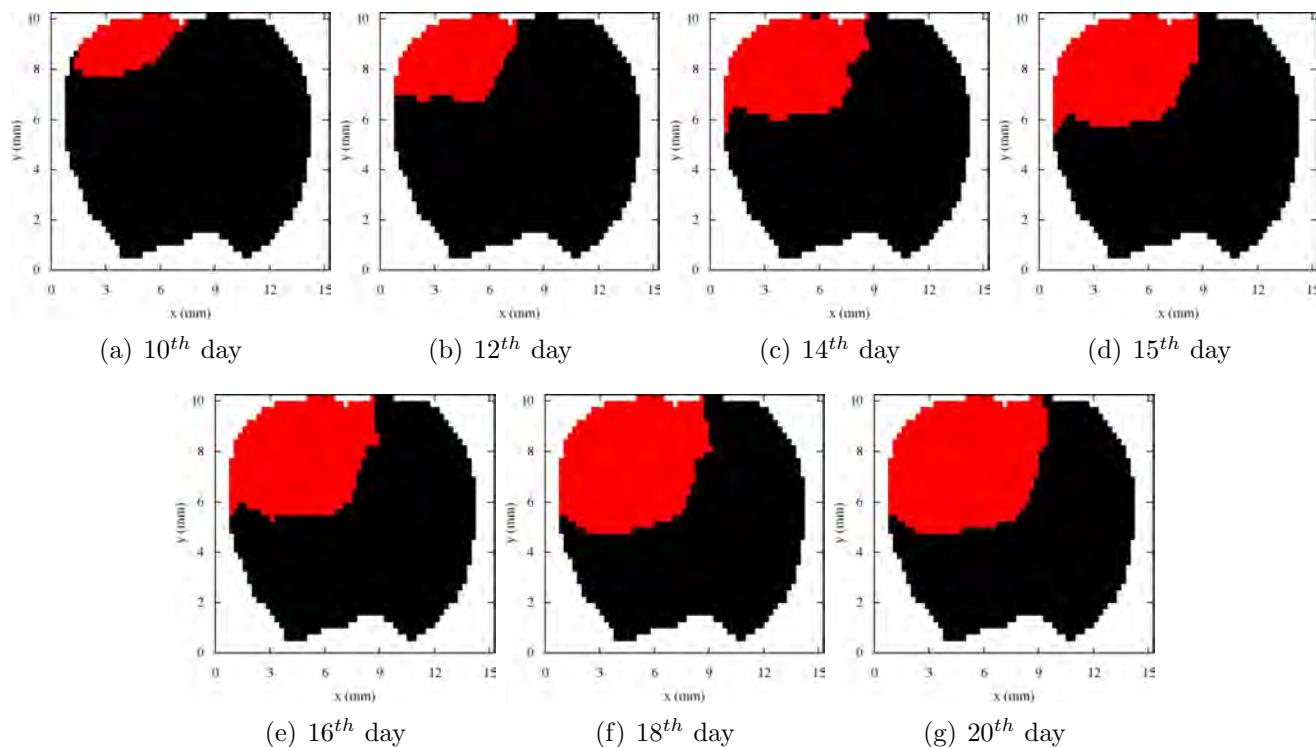


Figure 9.4: Observational data extracted from MRI images of progressive tumor growth in a murine experiment.^[16]

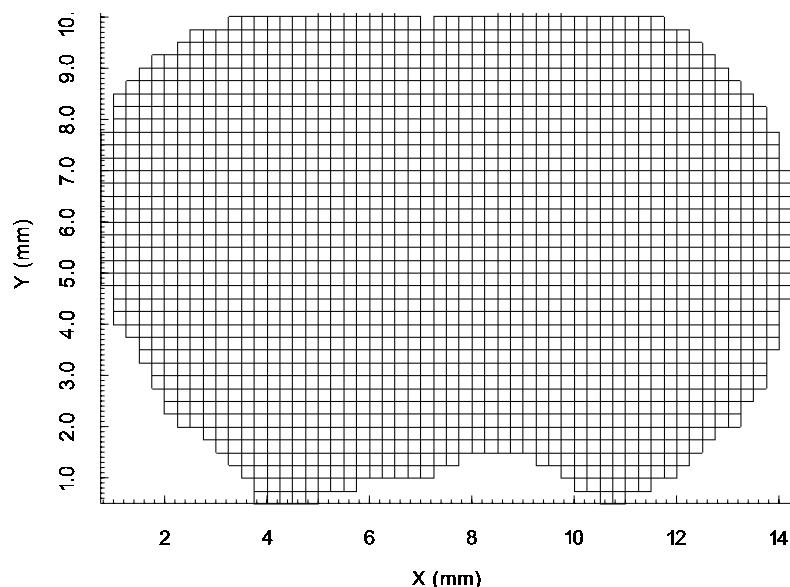


Figure 9.5: Voxels from the MRI that captures the domain of the brain (From [74]).

Table 9.2: Uniform priors for model parameters (From [74]).

Parameter	Prior
M_T	$\mathcal{U}(0.2, 1.0)$
M_T^*	$\mathcal{U}(0.05, 0.35)$
c	$\mathcal{U}(0.1, 0.55)$
λ_T^{grow}	$\mathcal{U}(0.2, 1.6)$
\bar{E}_T	$\mathcal{U}(0.15, 0.35)$
ϵ_T	$\mathcal{U}(0.5, 0.7)$
E	$\mathcal{U}(1.2, 2.8)$
ν	$\mathcal{U}(0.2, 0.49)$
λ	$\mathcal{U}(0.001, 0.004)$
γ	$\mathcal{U}(80, 320)$
γ^g	$\mathcal{U}(80, 320)$

Table 9.3: Sensitivity analysis for models \mathcal{P}_1 to \mathcal{P}_{13} (From [74]).

Model	Sensitivity of the Parameters (μ^*)										
	M_T	M_T^*	c	λ_T^{grow}	\bar{E}_T	ϵ_T	E	ν	λ	γ	γ^g
RD01		19.49		34.74							
PF01	2.59			32.05	0.08	2.63					
RD02		22.06		32.13			3.24×10^{-5}	0.55	1.28	1.41	
RD03		21.01		30.53			2.63×10^{-5}	0.68	2.44		2.65
MD01	20.71		15.70	33.86			2.97×10^{-5}	7.89×10^{-5}	9.54×10^{-5}		
RD04		16.63		29.94			3.02×10^{-5}	1.37	3.72	1.21	2.55
MD02	14.88		15.88	31.75			5.40×10^{-5}	0.74	1.57	1.37	
MD03	14.05		14.93	29.91			4.23×10^{-5}	0.91	2.79		2.48
PF02	2.77			32.02	0.08	3.21	9.00×10^{-7}	2.40×10^{-6}	5.70×10^{-6}		
MD04	12.03		16.32	28.61			2.64×10^{-5}	1.31	3.23	0.95	1.81
PF03	2.89			32.21	0.07	3.60	9.00×10^{-7}	0.16	0.46	0.39	
PF04	2.61			27.74	0.07	3.17	9.00×10^{-7}	0.65	3.20		3.02
PF05	2.46			25.86	0.06	2.86	1.80×10^{-6}	0.44	3.24	0.34	2.39

posterior pdf to each parameter is presented.

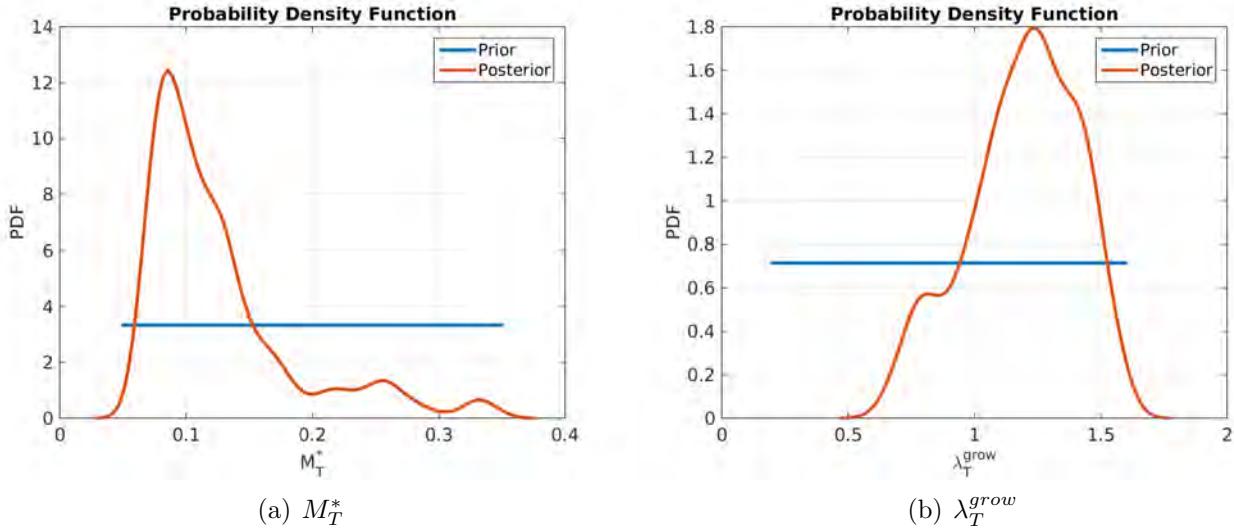


Figure 9.6: Prior ($\pi_{prior}^C(\theta_1|\mathcal{P}_1, \mathcal{M})$) and posterior ($\pi_{post}^C(\theta_1|\mathbf{y}_C, \mathcal{P}_1, \mathcal{M})$) for model RD01 at the calibration step.

In Figure 9.7, the prior and posterior for the validation step are exhibited. The forward problem is solved using Monte Carlo sampling and the pdf for the area of the tumor is computed. The error between the *in vivo* data and the model simulation is given by the Kullback-Leibler divergence divided by the variance of the data. In order to be declared not invalid, we chose $D_{KL}/\sigma_t^2 = D_{KL}(\sigma_t^2) \leq \gamma_{tol} = 0.7$, where σ_t^2 is the variance at time t . The tolerance is a subjective threshold picked to generally correspond to relatively close agreement (within 10-15%) of the tumor mass (area) observed and that predicted by the model. For model RD01, $D_{KL}(\sigma_{16}^2) = 0.92$ and $D_{KL}(\sigma_{18}^2) = 0.70$, so, the model is not valid.

The Occam Category 2 model is also not valid. In Category 3, we have three models with 6 parameters each. The calibration step we compute the plausibility of each model. The most plausible model, in this case model MD01, is selected for the validation step. Even though model MD01 is the most plausible, it is not a valid model. The simulation is continued until a not invalid model is selected.

Table 9.4 contains the results for all the simulations performed. For this specific data set, model PF04 from Category 5 satisfies the prescribed tolerance, with $D_{KL}(\sigma_{16}^2) = 0.70$ and $D_{KL}(\sigma_{18}^2) = 0.66$. In Figure 9.8 we compare the mean results for the tumor area at times 16 and 18 days with their respective *in vivo* data. It is possible to conclude that model PF04, which is a phase-field model that includes the mechanical deformation feedback at the tumor growth rate (λ_T^{grow}), is not invalid. Having concluded that model PF04 is **not** invalid, it is not necessary to consider more complex models, such as PF05.

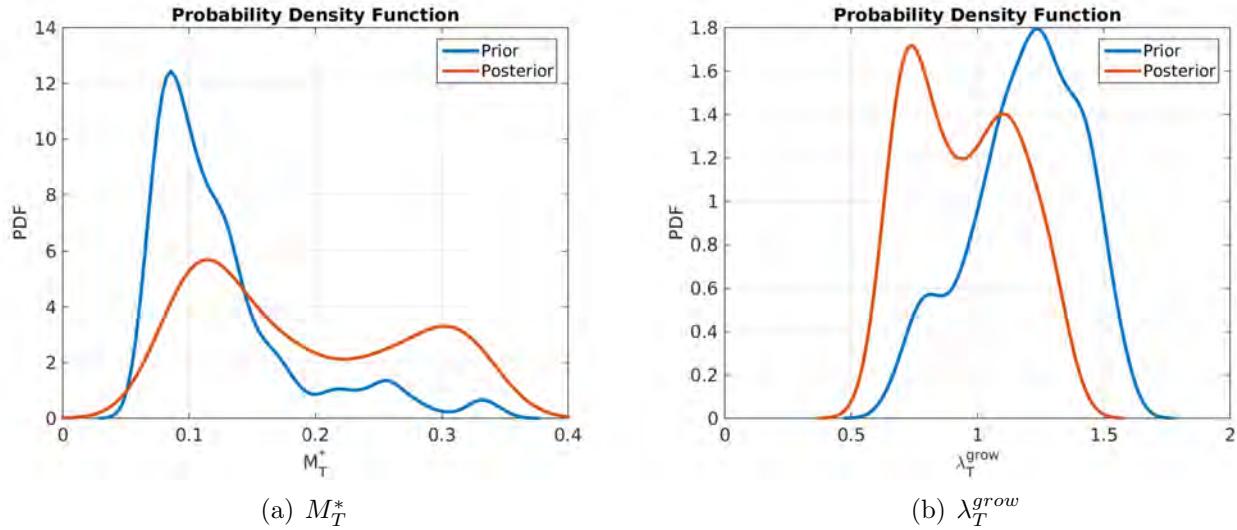


Figure 9.7: Prior ($\pi_{prior}^V(\boldsymbol{\theta}_1 | \mathcal{P}_1, \mathcal{M})$) and posterior ($\pi_{post}^V(\boldsymbol{\theta}_1 | \mathbf{y}_V, \mathcal{P}_1, \mathcal{M})$) for model RD01 at the validation step.

Table 9.4: Results and plausibilities (Reproduced from [74]).

Model	Occam Category	Plausibility	$D_{KL}(\sigma_{16}^2)$	$D_{KL}(\sigma_{18}^2)$
RD01	1	n/a	0.92	0.70
PF01	2	n/a	0.65	0.74
RD02	3	0.22		
RD03	3	0.14		
MD01	3	0.64	0.93	0.72
RD04	4	0.02		
MD02	4	0.06		
MD03	4	0.02		
PF02	4	0.90	0.68	0.76
MD04	5	0.02		
PF03	5	0.34		
PF04	5	0.64	0.70	0.66
PF05	6	n/a		

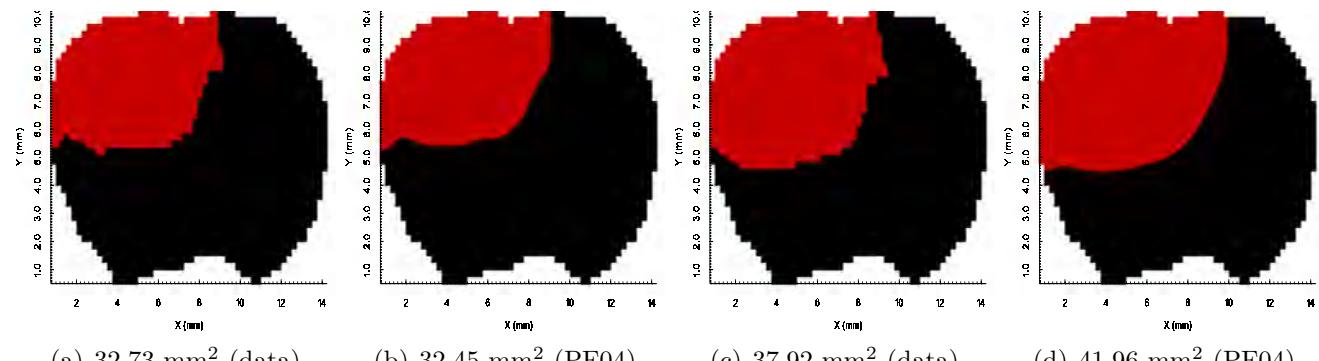


Figure 9.8: Tumor area evolution at slice 12 at $t = 16$ (a and b) and $t = 18$ (c and d).

Chapter 10

Chemical Kinetics Models of Signal Transduction

10.1 Introduction

As described in Chapter 8, signal transduction involves the transmission of molecular events, called protein or molecular signals, through a sequence of chemical reactions along so-called signaling pathways in a cell that trigger key biochemical events in the cell. Transduction occurs when extracellular molecules react with receptor proteins on the surface of the cell or within the cell that initiates the signaling cascade, a biochemical chain of events that leads to cell responses, such as cell differentiation, apoptosis, proliferation through mitosis, gene expression, etc.

It is natural to think of such signaling pathways as a network of nodes representing concentrations of protein molecules or molecule complexes connected by paths that transmit signals representing a probable chemical reaction. Examples of signaling pathways are illustrated in Figure 10.1.

The study of the rates at which chemical reactions occur and to characterize the steps in a chemical reaction process and the sequence in which they occur is precisely the province of *chemical kinetics*, a subject that provides a natural framework for developing discrete network models of signal transduction. The formalism of chemical kinetics is *stoichiometry*, that deals with recipes for balancing reactants and products, as well as their relative proportions needed in a reaction. In this chapter, an introduction to chemical kinetics and enzymatic reactions is given following textbook accounts on biochemistry [e.g., 117] and the report of Shahmoradi and Oden [103].

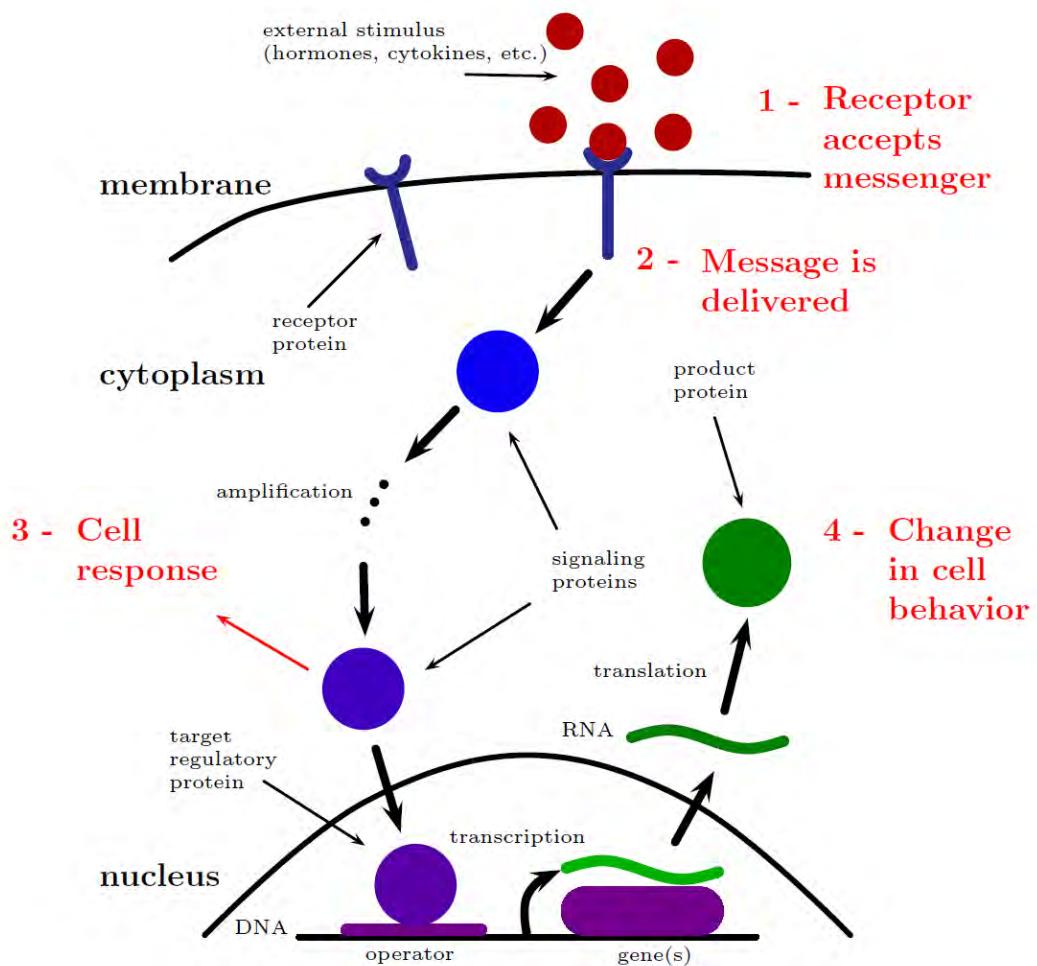


Figure 10.1: Symbolic representation of pathways in a cell, connecting molecules that transmit signals to trigger various cell responses (from [83, p. 748]).

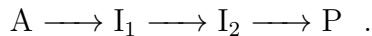
10.2 An Overview of Chemical Kinetics

A chemical reaction is a process in which one or more substances, the *reactants*, are converted into one or more different substances, called the *products*. This conversion is generally realized because the reaction rearranges the constituent atoms of the reactants to create different atomic structures in the products.

The stoichiometry of reactants and products is based on the fundamental principle of conservation of mass: atoms cannot be created or destroyed in the course of a chemical reaction. The simplest case entails a stoichiometric reaction written as,



where A represents the reactants and P the products. There may also be intermediates I_1 and I_2 in the reaction, and we then write,



A bimolecular reaction will involve two molecules, A reacting with B, and we write,



Here A and B denote “species” of atoms; their mass concentration in a solution, such as cytoplasm, is denoted by $[A]$ and $[B]$. These concentrations are generally measured in moles, which, it is recalled, is a unit M of measure in SI (the International System of Units) of the amount of the substrate that contains N_A elementary entities (atoms or molecules, etc.), N_A being the Avogadro’s number: $6.02214057 \times 10^{23}/\text{mol}$. The number N_A of molecules per mole of a substrate is the mean relative molecular mass expressed in grams.

The rate at which A and B react is proportional to the number of random collisions of atoms of A with those of B. If we imagine a single molecule of species A in a concentrations of B in solutions, then the collision note of N molecules of A with molecules B is proportional to the product of N and the total molecular concentration $[B]$ of B. Thus, if R denotes the collision rate of A-atoms with B, and if V is the volume, R is proportional to $N[B]/V$ and N/V is proportional to $[A]$. We write *

$$R = k(T)[A][B] , \quad (10.2)$$

*This argument on reaction rates is found in Shahmoradi and Oden (2016) [103] and Wentworth (2000) [120].

where $k(T)$ is the *rate of reaction*, which depends upon the temperature T , and

$$k(T)[A][B] = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = \frac{d[P]}{dt} . \quad (10.3)$$

This idea can be generalized to a more general elementary reaction involving many reactions; e.g.,



The coefficients a, b, \dots, z being stoichiometric coefficients. The rate of this process must be proportional to the frequency with the reacting molecules combine, which must depend on the products of the concentrations of reactants. Thus,

$$R = k[A]^a[B]^b \cdots [Z]^z \quad (10.5)$$

The sum of the exponents, $(a + b + \dots + z)$, is the *order of the reaction*, and k is the reaction rate constant.

In many cases, it is possible to experimentally measure the molecular mass $[A]$ or $[P]$ in, say, as a function of time, so that the measured μ_{exp} rate of a reaction is,

$$\mu_{exp} = -\frac{d[A]}{dt} = \frac{d[P]}{dt} \quad (10.6)$$

Thus, for a first-order reaction,

$$\mu_{exp} = -\frac{d[A]}{dt} = k[A] , \quad (10.7)$$

and for the reaction $A + B \longrightarrow P$,

$$\mu_{exp} = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = k[A][B] . \quad (10.8)$$

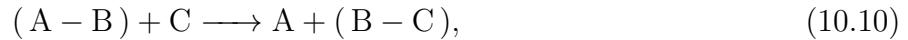
Thus, in a second-order reaction, if we now denote $\mu_{exp} = \nu$ =the instantaneous velocity of reaction, $2A \longrightarrow P$ results in,

$$-\frac{d[A]}{dt} = 2k[A]^2. \quad (10.9)$$

For first-order reactions, the units of k are $[sec^{-1}]$, while for second-order reactions they are $M^{-1}[sec^{-1}]$, M standing for mass.

Transition States.

Following [117, p. 475], consider a bimolecular elementary reaction,



where atom C approaches the diatomic molecule (A-B) in which the (A-B) covalent bond is broken while a new bond (B-C) is being formed. In between these events, an unstable complex, denoted $A \cdot B \cdot C$ exists called a *transition state*. This intermediate state is said to be an *activated complex* and the reaction may decompose back into the reactant, $(A - B) + C$, or progress forward to the product, $A + (B - C)$, depending on the kinetic energy of the system. At the transition state, there is an equal probability that the reaction will occur or that it will decompose back into its components. The general situation for a bimolecular reaction can be written,



where Z is the activated (transition) complex. If k is the usual rate constant,

$$\frac{d[P]}{dt} = k[A][B] = k'[Z]. \quad (10.12)$$

It is customary in transition-state theory to assume that Z is in equilibrium with reactants, so that

$$[Z] = \mathcal{K}[A][B], \quad (10.13)$$

\mathcal{K} being the equilibrium constant. It is argued in [117, p. 476] that the equilibrium constant is related to the change ΔG in the Gibbs free energy of the reactants less than that of the products, as depicted in Figure 10.2, and that,

$$RT \ln \mathcal{K} = -\Delta G, \quad (10.14)$$

with R being the gas constant and T the absolute temperature. Thus,

$$\frac{d[P]}{dt} = k' e^{-\Delta G/RT} [A][B]. \quad (10.15)$$

Thus, the larger the free energy difference of the transition state and the reactants, the slower the reaction occurs.

An appeal to straightforward thermodynamic arguments [117, p. 476] leads to conclusion that the rate constant k for the elementary reaction is,

$$k = \frac{k_B T}{\hbar} e^{-\Delta G/RT}, \quad (10.16)$$

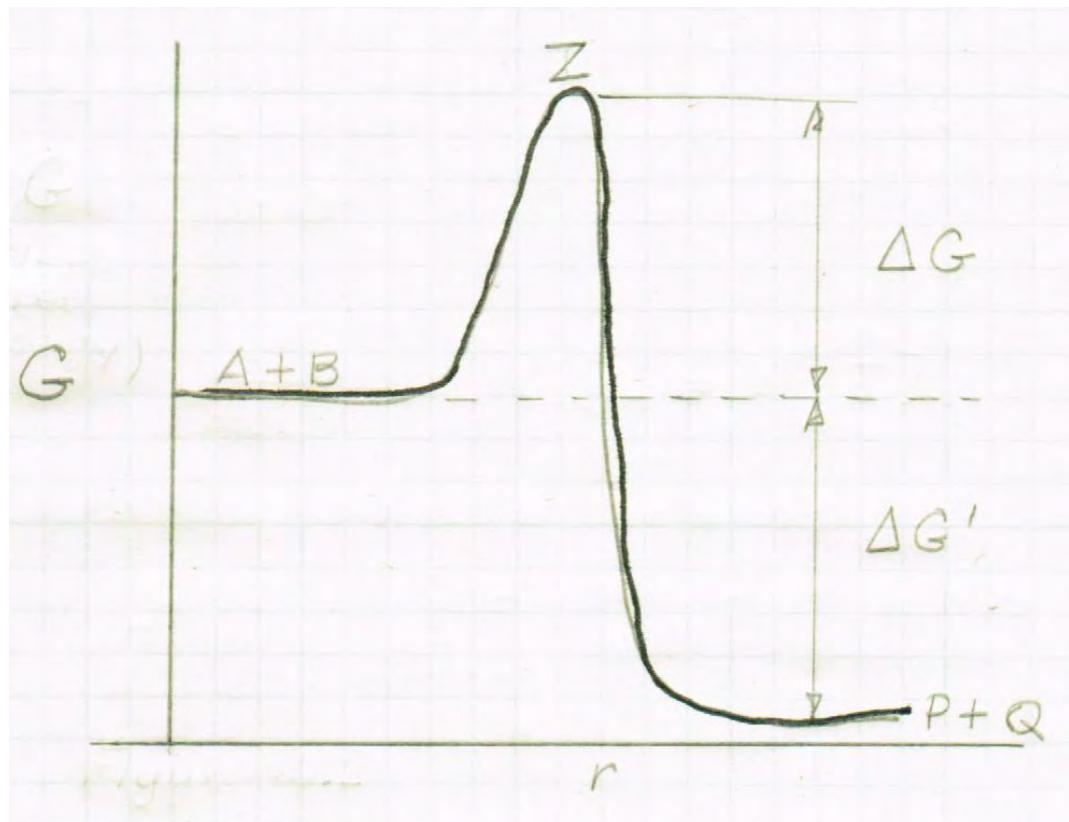


Figure 10.2: Gibbs free energy vs. a reaction coordinate r , showing the drop in free energy ΔG from the activated state Z to that of the reactants (after [117]).

k_B being the Boltzmann constant and \hbar is the Planck constant. Thus, the rate of reaction decreases as the change in Gibbs energy ΔG increases, and the reaction rate increases as the temperature rises.

10.3 Complex Reactions

As noted earlier, most reactions involve multi-step processes in which key intermediates may be obscured by the apparent simplicity of their stoichiometric equations. Shahmoradi and Oden [103] gives the example of ozone decomposing into oxygen:



The full equations for this reaction involve a series of intermediate reactions given by,



The first two equations can be written as,



with k_{-1} describing the rate of the reverse reaction. Collectively, the full reaction can be characterized by the system

$$-\frac{d[\text{O}_3]}{dt} = k_1[\text{O}_3] - k_{-1}[\text{O}_2][\text{O}] + k_2[\text{O}][\text{O}_3], \quad (10.22)$$

$$\frac{d[\text{O}_2]}{dt} = k_1[\text{O}_3] - k_{-1}[\text{O}_2][\text{O}] + 2k_2[\text{O}][\text{O}_3], \quad (10.23)$$

$$\frac{d[\text{O}]}{dt} = k_1[\text{O}_3] - k_{-1}[\text{O}_2][\text{O}] - k_2[\text{O}][\text{O}_3]. \quad (10.24)$$

This system completely defines the rate of change of the concentrations of O_3 , O_2 , and O .

Equilibrium and Steady-State Approximations

Simplified forms of the rate equations such as (10.22), (10.23), (10.24) can be derived in certain situations involving *reaction equilibrium*, in which all reactions are fast, thereby reaching an equilibrated state quickly, and are reversible, in which every forward reaction is equilibrated by its reverse reaction. In the case of reaction equilibrium of the three species of example characterized by (10.22), (10.23), (10.24), the rate of concentration changes of all three species must vanish:

$$-\frac{d[O_3]}{dt} = \frac{d[O_2]}{dt} = \frac{d[O]}{dt} = 0. \quad (10.25)$$

A simple algebraic calculation then yields as a reaction equilibrium condition,

$$k_1[O_3] - k_{-1}[O_2][O] = 0, \Rightarrow \frac{k_1}{K_{-1}} = \frac{[O_2][O]}{[O_3]} \equiv K_1 \quad (10.26)$$

in which K_1 is called the **equilibrium rate constant** of the reaction.

An alternate simplifying approximation pertains to that of a *steady-state* approximation of a reaction process. In steady-state approximation, some but not necessarily all reactants are assumed to be time independent. For example, in the system of equations (10.22), (10.23), (10.24), in the ozone decomposition example embodied by this system of rate equations, the fact that a single oxygen atom O is highly reactive suggests that it reaches a low concentration quickly and becomes approximately constant over time intervals of interest. Then the steady-state approximation assumes the form,

$$\begin{aligned} \frac{d[O]}{dt} &= k_1[O_3] - k_{-1}[O_2][O] - k_2[O][O_3], \\ &= 0 \end{aligned} \quad (10.27)$$

Michaelis-Menten Kinetics

A well-known and highly cited example of a steady-state approximation that arises in chemical kinetics is the *Michaelis-Menten mechanism*, which pertains to enzyme-substrate reactions. We use the following common notations and symbols:

E = enzyme

S = substrate

ES = enzyme-substrate complex

P = products

The enzyme E catalyzes the substrate S to produce the enzyme-substrate complex ES, which can undergo a reverse reaction to recombine the reactant E+S or it can result in products P+E. The non-covalently bonded enzyme-substrate complex ES is known as the *Michaelis complex* [117].



The velocity rate of $d[P]/dt$ is $k_2[ES]$, and the production of ES is the difference between the rates of elementary reactions producing [ES] and those resulting in its decrease in concentration. The entire mechanism is governed by the system,

$$\frac{d[S]}{dt} = -k_1[E][S] + k_{-1}[ES], \quad (10.29)$$

$$\frac{d[E]}{dt} = -k_1[E][S] + k_{-1}[ES] + k_2[ES], \quad (10.30)$$

$$\frac{d[ES]}{dt} = +k_1[E][S] - k_{-1}[ES] - k_2[ES], \quad (10.31)$$

$$\frac{d[P]}{dt} = +k_2[ES], \quad (10.32)$$

The assumption of reaction equilibrium, introduce by Michaelis and Menten in 1913 [76] asserts that $d[S]/dt = 0$ and leads to the condition,

$$\frac{k_{-1}}{k_1} = \frac{[E][S]}{[ES]}. \quad (10.33)$$

The constant $\frac{k_{-1}}{k_1} \equiv K_d$ is the *dissociation constant* of the first step in the enzymatic reaction. If also $d[E]/dt = 0$, then,

$$1 = \frac{K_m[ES]}{[E][S]}, \quad K_m = \frac{k_{-1} + k_2}{k_1}, \quad (10.34)$$

K_m being referred to as the *Michaelis constant*. According to [103], the quantities [ES] and [E] are not directly measurable, but the total enzyme concentration,

$$[E]_T = [E] + [ES], \quad (10.35)$$

is measurable. Setting $d[ES]/dt = 0$ as a steady-state condition, using (10.35) and rearranging terms yields, after some algebra,

$$[ES] = \frac{[E][S]}{K_m + [S]}. \quad (10.36)$$

Returning to (10.29), (10.30), (10.31), (10.32), the initial velocity of reaction, expressed in terms of experimentally measurable quantities is,

$$v_0 = \frac{d[P]}{dt} \Big|_{t^*} = k_2[ES] = \frac{k_2[E]_T[S]}{K_m + [S]}, \quad (10.37)$$

t^* being the time to steady-state condition. Denoting by v_{max} the maximal velocity of reaction is (which as pointed out in [117]) occurs when the enzyme is saturated; i.e., when it is entirely ES),

$$v_{max} = k_2[E]_T. \quad (10.38)$$

Thus the initial velocity of reaction is,

$$v_0 = \frac{v_{max}[S]}{K_m[S]}. \quad (10.39)$$

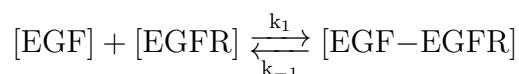
This result is called the Michaelis-Menten equation. It provides the velocity of reaction at the time a steady-state is achieved, and is referred to as the basic equation of enzyme kinetics.

10.4 An Application to a Cell-Signaling Pathway

Basic approaches and methodologies of stoichiometry laid down in the preceding sections provide the framework for developing effective mathematical and computational models of cell signaling pathways and the production of protein complexes essential in key cell functions.

As examples of representative models, we can consider the chemical kinetic model presented by Wang et al. (2007) [119] of the non-small cell lung cancer (NSCLC) - epidermal growth factor receptor (EGFR) signaling pathway shown in Figure 10.3. There one observes a cascade of reactions beginning with binding of the epidermal growth factor (EGF) or transforming growth factor (TGF_α) to the extracellular environment to produce proteins that affect various phenotypic events, such as proliferation, or inhibition of apoptosis. The cascade begins with reactants EGF and EGFR at the cell receptor producing EGF-EGFR, shown within the elliptical domain of the diagram.

This initial reaction is represented by,



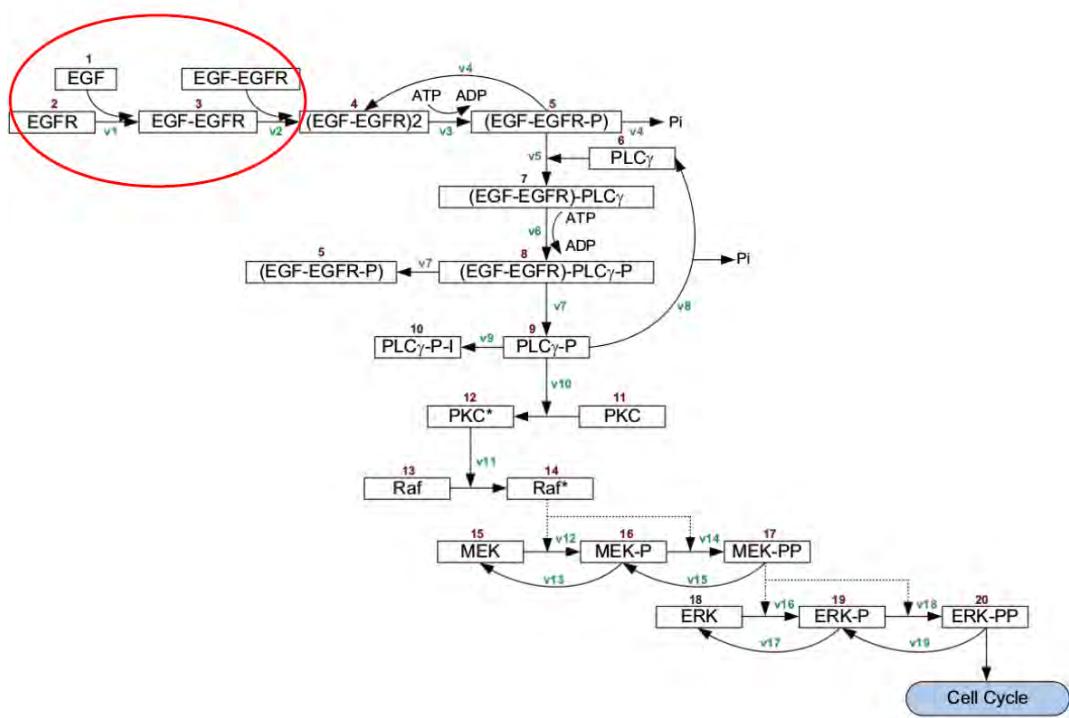


Figure 10.3: Schematic of a network describing a cascade of reactants and products defining an EGFR signaling pathway (reproduced from Wang et al. [119]). The initiation of the cascade begins with the reactants EGF and EGFR producing EGF-EGFR, shown within the initiation ellipse in the figure. Reaction numbers are denoted by v_i .

and is governed by the system,

$$\begin{aligned}\frac{d[EGF]}{dt} &= k_{-1}[EGF - EGFR] - k_1[EGF][EGFR], \\ \frac{d[EGFR]}{dt} &= k_{-1}[EGF - EGFR] - k_1[EGF][EGFR], \\ \frac{d[EGF - EGFR]}{dt} &= k_1[EGF][EGFR] - k_{-1}[EGF - EGFR].\end{aligned}$$

Other protein complexes involved in the network are:

$\text{PLC}_\gamma \sim \text{phospholipase C}_\gamma$,

$\text{ERK} \sim \text{extracellular signal-regulated kinase}$,

$\text{RAF} \sim \text{RAF protein kinase}$,

$\text{MEK} \sim \text{mitogen-activated protein kinase}$.

The quantity of interest in this example is the concentration of ERK-PP extracellular signal-regulated kinase (PP standing for doubly-phosphorylated), which governs the cell cycle.

The network is modeled by a system of 20 ordinary differential equations (ODEs). Using the notation in [119], we denote by X_i , the i th molecular pathway reactant component. The system of ODEs is generated following the chemical kinetics rules described earlier. Tables 10.1 and 10.2 are reproduced from [119] and give the molecular type for each reactant X_i , the corresponding ODE, and the kinetic parameters chosen in [119] for the 19 reactions, numbered V_1, V_2, \dots, V_{19} .

An interesting property of this system is its response to random input concentrations of EGF in the spirit of our study of uncertainties. Figure 10.4 contains results of calculations performed by Ernesto Lima in which a Gaussian distribution of EGF is input, and the ODE system is integrated numerically in time for 150 seconds using a standard fourth-order Runge-Kutta algorithm. The system parameters were taken to be the deterministic values given in Table 10.2. The output QoI which in this case is the ERK-PP at the output node of the network, was computed using a Monte Carlo method and is shown in Figure ???. The output distribution is clearly non-Gaussian. A plot of computed ERK-PP versus ERF is given in Figure 10.5, where a clear nonlinear relation between input and output is seen.

It is interesting that this same cascade can be modeled with several different pathway networks. Figure 10.6, for example, is a EGFR signaling pathway with 100 components representing the

Table 10.1: Kinetic equations and initial concentrations (reproduced from [119]).

Reactant	Molecular Variable	Initial Value [nM]	ODE
X_1	EGF	to be varied	$d(X_1)/dt = -v_1$
X_2	EGFR	80	$d(X_2)/dt = -v_1$
X_3	EGF-EGFR	0	$d(X_3)/dt = v_1 - 2v_2$
X_4	(EGF-EGFR)2	0	$d(X_4)/dt = v_2 + v_4 - v_3$
X_5	EGF-EGFR-P	0	$d(X_5)/dt = v_3 + v_7 - v_4 - v_5$
X_6	PLC $_{\gamma}$	10	$d(X_6)/dt = v_8 - v_5$
X_7	EGF-EGFR-PLC $_{\gamma}$	0	$d(X_7)/dt = v_5 - v_6$
X_8	EGF-EGFR-PLC $_{\gamma}$ -P	0	$d(X_8)/dt = v_6 - v_7$
X_9	PLC $_{\gamma}$ -P	0	$d(X_9)/dt = v_7 - v_8 - v_9 - v_{10}$
X_{10}	PLC $_{\gamma}$ -P-I	0	$d(X_{10})/dt = v_9$
X_{11}	PKC	10	$d(X_{11})/dt = -v_{10}$
X_{12}	PKC*	0	$d(X_{12})/dt = v_{10} - v_{11}$
X_{13}	RAF	100	$d(X_{13})/dt = -v_{11}$
X_{14}	RAF*	0	$d(X_{14})/dt = v_{11} - v_{12} - v_{14}$
X_{15}	MEK	120	$d(X_{15})/dt = v_{13} - v_{12}$
X_{16}	MEK-P	0	$d(X_{16})/dt = v_{12} + v_{15} - v_{13} - v_{14}$
X_{17}	MEK-PP	0	$d(X_{17})/dt = v_{14} - v_{15} - v_{16} - v_{18}$
X_{18}	ERK	100	$d(X_{18})/dt = v_{17} - v_{16}$
X_{19}	ERK-P	0	$d(X_{19})/dt = v_{16} + v_{19} - v_{17} - v_{18}$
X_{20}	ERK-PP	0	$d(X_{20})/dt = v_{18} - v_{19}$

Table 10.2: Kinetic parameters (reproduced from [119]).

Reaction number	Equation	Kinetic parameter	
v_1	$k_1 X_1 X_2 - k_{-1} X_3$	$k_1 = 0.003$	$k_{-1} = 0.06$
v_2	$k_2 X_3 X_3 - k_{-2} X_4$	$k_2 = 0.01$	$k_{-2} = 0.1$
v_3	$k_3 X_4 - k_{-3} X_5$	$k_3 = 1$	$k_{-3} = 0.01$
v_4	$V_4 X_5 / (K_4 + X_5)$	$V_4 = 450$	$K_4 = 50$
v_5	$k_5 X_5 X_6 - k_{-5} X_7$	$k_5 = 0.06$	$k_{-5} = 0.2$
v_6	$k_6 X_7 - k_{-6} X_8$	$k_6 = 1$	$k_{-6} = 0.05$
v_7	$k_7 X_8 - k_{-7} X_5 X_9$	$k_7 = 0.3$	$k_{-7} = 0.006$
v_8	$V_8 X_9 / (K_8 + X_9)$	$V_8 = 1$	$K_8 = 100$
v_9	$k_9 X_9 - k_{-9} X_{10}$	$k_9 = 1$	$k_{-9} = 0.03$
v_{10}	$k_{10} X_9 X_{11} - k_{-10} X_{12}$	$k_{10} = 0.214$	$k_{-10} = 5.25$
v_{11}	$V_{11} X_{12} X_{13} / (K_{11} + X_{13})$	$V_{11} = 4$	$K_{11} = 64$
v_{12}	$V_{12} X_{14} X_{15} / [K_{12}(1 + X_{16}/K_{14}) + X_{15}]$	$V_{12} = 3.5$	$K_{12} = 317$
v_{13}	$V_{13} X_{16} / [K_{13}(1 + X_{17}/K_{15}) + X_{16}]$	$V_{13} = 0.058$	$K_{13} = 2200$
v_{14}	$V_{14} X_{14} X_{16} / [K_{14}(1 + X_{15}/K_{12}) + X_{16}]$	$V_{14} = 2.9$	$K_{14} = 317$
v_{15}	$V_{15} X_{17} / [K_{15}(1 + X_{16}/K_{13}) + X_{17}]$	$V_{15} = 0.058$	$K_{15} = 60$
v_{16}	$V_{16} X_{17} X_{18} / [K_{16}(1 + X_{19}/K_{18}) + X_{18}]$	$V_{16} = 9.5$	$K_{16} = 1.46 \times 10^5$
v_{17}	$V_{17} X_{19} / [K_{17}(1 + X_{20}/K_{19}) + X_{19}]$	$V_{17} = 0.3$	$K_{17} = 160$
v_{18}	$V_{18} X_{17} X_{19} / [K_{18}(1 + X_{18}/K_{16}) + X_{19}]$	$V_{18} = 16$	$K_{18} = 1.46 \times 10^5$
v_{19}	$V_{19} X_{20} / [K_{19}(1 + X_{19}/K_{17}) + X_{20}]$	$V_{19} = 0.27$	$K_{19} = 60$

same cascade given in Figure 10.3, but with many more components, advocated in Schoeberl et al. (2002) [100].

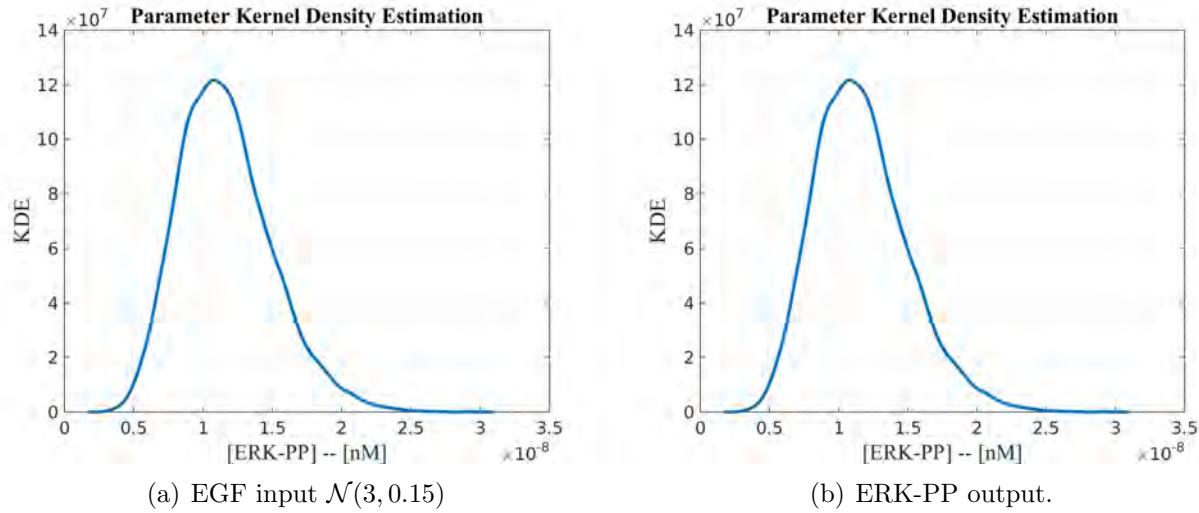


Figure 10.4: a) Input EGF distribution supplied to the pathway network shown in Figure 10.3, with parameters given in Table 10.2 with 100,000 samples of a Gaussian input. b) The computed output ERK-PP at t=150 seconds.

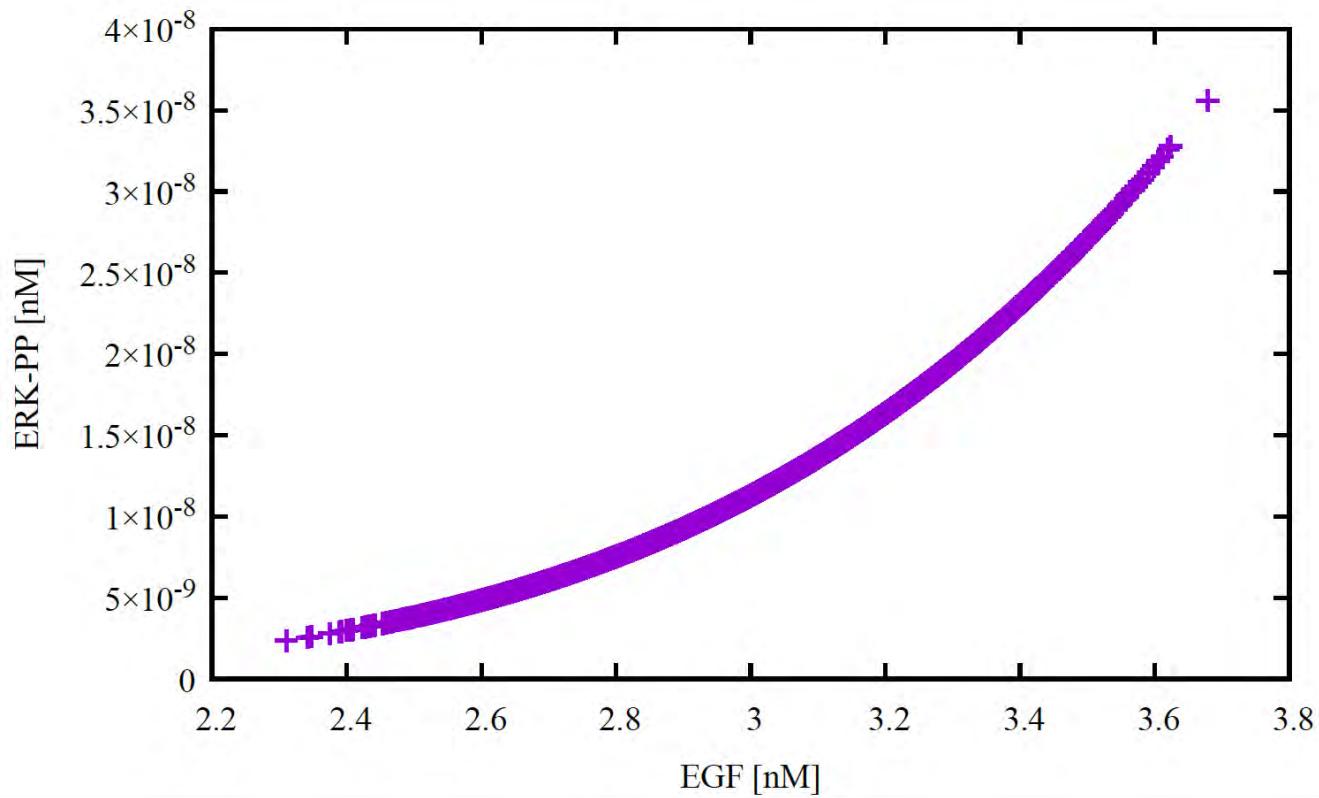


Figure 10.5: Computed variations in output at $t=150$ seconds to EGF input.

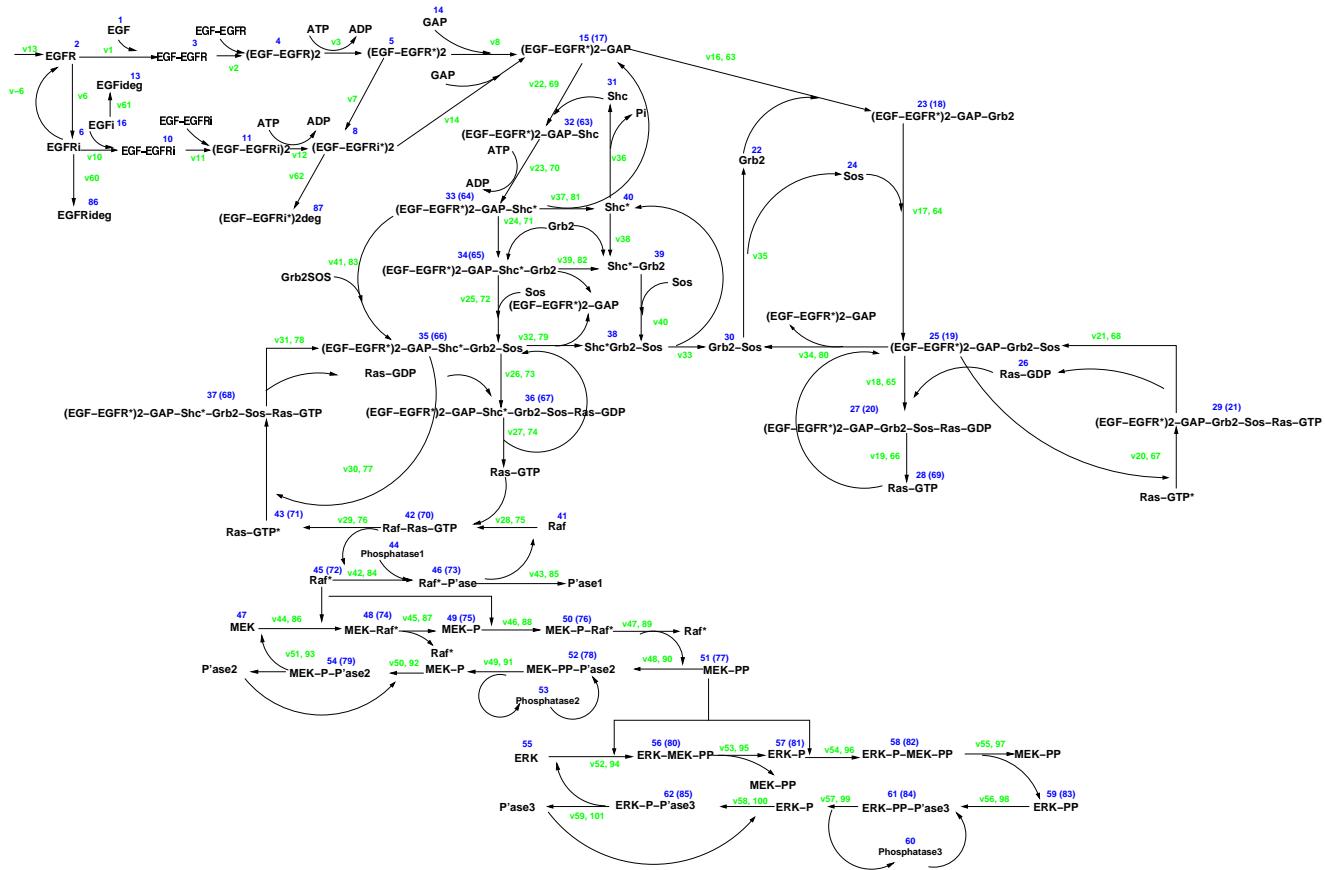


Figure 10.6: A more complex version of the signaling pathway in Figure 10.3 (reproduced from [100]).

Chapter 11

Phase-Field Models

11.1 Introduction

The phase-field models, among the classes of models derived for simulating tumor growth at the tissue scale, belong to a rich class of models that feature moving boundaries between two or more constituents, called “phases”. The solutions of phase-field models exhibit internal layers or boundaries that represent interfaces between phases which have a small but finite width which is controlled by a length-scale parameter. The moving boundary problem is thus characterized by a smooth moving interface which can be regarded as a smoothing or diffusing of the sharp-interface models – and for this reason, phase-field models are also referred to as *diffuse-interface* models. A comprehensive overview of phase-field modeling can be found in the encyclopedia article of Gomez [44] and van der Zee [115], and fundamental ideas behind a large class of such models are discussed in Gurtin [46].

11.2 Structure of Phase-field Models

A fundamental property of thermodynamically consistent phase-field models is that the Helmholtz free energy Ψ depends upon both the values and gradients of the principal field being modeled, the phase field ϕ :

$$\Psi = \widehat{\Psi}(\phi, \nabla\phi) \quad (11.1)$$

The corresponding energy functional E is called the *Ginsberg-Landau* energy:

$$E = \int_{\Omega} \widehat{\Psi}(\phi, \nabla\phi) \, d\boldsymbol{x} \quad (11.2)$$

Ω being the phase-field domain ($\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$). $\mathrm{d}x$ a volume element. The principle of material frame indifference mandates that $\widehat{\Psi}$ depends on $\nabla\phi$ only through $|\nabla\phi|$. In general, $\widehat{\Psi}(0, 0)$ is taken to be of the form

$$\widehat{\Psi}(\phi, \nabla\phi) = W(\phi) + \frac{\epsilon^2}{2} |\nabla\phi|^2 \quad (11.3)$$

where $W(\phi)$ is the phase-field potential. In general, the term $\epsilon^2|\nabla\phi|^2/2$ represent the surface energy, of the surface separating phase, ϵ (or $\sqrt{\epsilon}$) being the transition width of the diffused interface. The phase field energy collapses to the usual Helmholtz free energy as ϵ approaches zero (see Ambrosio and Tortelli [8], Braides [23], Dal Masso [36] and Gomez and van der Zee [44]). The potential W is generally taken to be a characterization of a double-well potential with two local minima, making possible the coexistence of different phases. Typical choices are:

$$\left. \begin{aligned} W &= W_1(\phi) = \frac{1}{4}(1 - \phi)^2, \\ W &= W_2(\phi) = \frac{1}{2} \left[(1 + \phi) \log \left(\frac{1}{2}(1 + \phi) \right) + (1 - \phi) \log \left(\frac{1}{2}(1 - \phi) \right) + a(1 - \phi^2) \right], \\ W &= W_3(\phi) = \begin{cases} \frac{1}{4}(1 + \phi)^2 & , \phi < -1 \\ \frac{1}{4}(1 - \phi^2)^2 & , -1 \leq \phi \leq 1 \\ \frac{1}{4}(\phi - 1)^2 & , \phi > 1 \end{cases}, \end{aligned} \right\} \quad (11.4)$$

a being a positive constant. With $\Psi(0, 0)$ given by (11.3), the Gateaux (vibrational derivative) of the Ginsberg-Landau energy is

$$\langle DE, \delta\phi \rangle = \lim_{\theta \rightarrow 0} \theta^{-1} (E(\phi + \theta\delta\phi) - E(\phi)) = \langle \mu, \delta\phi \rangle \quad (11.5)$$

where μ is the chemical potential, and is given by

$$\mu = DE = W'(\phi) - \epsilon^2 \Delta\phi, \quad (11.6)$$

Δ being the Laplacian.

Free-Energy Dissipation. For $\omega \subset \Omega$, we have, from the conservation of energy principle and the second law of thermodynamics,

$$\frac{d}{dt} \int_{\omega} \widehat{\Psi}(\phi, \nabla\phi) \mathrm{d}x = P_{\omega} - D_{\omega}, \quad (11.7)$$

where P_ω is the mechanical power and D_ω is the total dissipation, which must be non-negative for every physically meaningful process on the material body occupying ω :

$$D_\omega \geq 0 . \quad (11.8)$$

For an arbitrary subdomain $\omega \subset \Omega$,

$$\frac{d}{dt} \int_{\omega} \psi \, dx = \int_{\omega} \mu \frac{\partial \phi}{\partial t} \, dx + \int_{\partial \omega} \frac{\partial}{\partial \nabla \phi} \widehat{\Psi} \cdot \mathbf{n} \frac{\partial \phi}{\partial t} \, da \quad (11.9)$$

where the chemical potential is again the variational derivative,

$$\mu = \frac{\partial \widehat{\Psi}}{\partial \phi} - \nabla \cdot \left(\frac{\partial \widehat{\Psi}}{\partial \nabla \phi} \right) . \quad (11.10)$$

We now present several applications of these relations toward the derivation of classical models by phase transitions and non-classical models of tumor growth.

11.3 The Allen-Cahn Equation

Let ϕ denote a differentiable scalar-valued field on $\Omega \times (0, \tau]$ representing volume fraction of a constituent or a concentration of a species. Then, as we have seen earlier, the conservation of mass of this quantity is characterized by an evolution equation of the form

$$\frac{\partial \phi}{\partial t} + S(\phi, \nabla \phi, \mu) = 0 \quad (11.11)$$

where S is to be designed to not only represent mass supplied or removed but also satisfy the dissipation inequality (11.8). To find restrictions on the form of S , we return to (11.9) and write

$$\frac{d}{dt} \int_{\omega} \widehat{\Psi} \, dx = \int_{\omega} -\mu S \, dx + \int_{\partial \omega} \frac{\partial \widehat{\Psi}}{\partial \nabla \phi} \cdot \mathbf{n} \frac{\partial \phi}{\partial t} \, da , \quad (11.12)$$

we set (recall (11.7)),

$$P_\omega = \int_{\partial \omega} \mathbf{n} \cdot \left(\frac{\partial \widehat{\Psi}}{\partial \nabla \phi} \right) \frac{\partial \phi}{\partial t} \, da = \int_{\omega} -\nabla \cdot \frac{\partial \widehat{\Psi}}{\partial \nabla \phi} \cdot \frac{\partial \phi}{\partial t} \, dx , \quad (11.13)$$

$$D_\omega = \int_{\omega} \mu S \, dx . \quad (11.14)$$

The dissipation inequality is satisfied if we take $S = m(\phi)\mu$, with $m(\phi) \geq 0 \forall \phi$. If $\widehat{\Psi}$ is given by (11.3), then μ is given by (11.6) and

$$\frac{\partial \phi}{\partial t} = -m(\phi)(W'(\phi) - \epsilon^2 \Delta \phi) . \quad (11.15)$$

This is the *Allen-Cahn equation*, presented by Allen and Cahn in 1979. Gertin [47] refers to it as the *Ginzburg-Landau equation*.

11.4 The Cahn-Hilliard Equation

A similar argument leads to the *Cahn-Hilliard equation*. Here the mass balance is of the form

$$\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{H} = 0 \quad (11.16)$$

where \mathbf{H} is a mass flux of the form,

$$\mathbf{H} = \mathcal{H}(\phi, \nabla \phi, \mu, \nabla \mu) , \quad (11.17)$$

Now,

$$\frac{d}{dt} \int_{\omega} \widehat{\Psi} dx = \int_{\omega} \mathbf{H} \cdot \nabla \mu dx + \int_{\partial \omega} \left(\frac{\partial \widehat{\Psi}}{\partial \nabla \phi} \cdot \frac{\partial \phi}{\partial t} - \mu \mathbf{H} \right) \cdot \mathbf{n} da \quad (11.18)$$

If the first term on the right-hand-side is identified as D_{ω} , the it is sufficient to take as \mathbf{H} a function of the form

$$\mathbf{H} = -m(\phi) \nabla \mu , \quad (11.19)$$

with $m(\phi) \geq 0$. Then (11.16) assumes the form of the coupled pair

$$\left. \begin{aligned} \frac{\partial \phi}{\partial t} &= \nabla \cdot (m(\phi) \nabla \mu) \\ \mu &= W'(\phi) - \epsilon^2 \Delta \phi \end{aligned} \right\} \quad (11.20)$$

or, as a single equation

$$\frac{\partial \phi}{\partial t} = \nabla \cdot (m(\phi) \nabla (W'(\phi) - \epsilon^2 \Delta \phi)) . \quad (11.21)$$

This relation is the *Cahn-Hilliard equation*. It is a fourth-order nonlinear partial differential equation in the space variable \mathbf{x} and first order in time. Its similarity to several of the phase-field models of tumor growth derived earlier is apparent.

11.5 Phase-Field Models of Tumor Growth

For illustration purpose, we consider a reduced version of the phase-field tumor growth models introduced in Chapter 9 that lead to systems analogous to the Cahn-Hilliard equations and whose solutions exhibit interesting patterns of phase evolution. These examples are based on the analysis of the four species tumor growth models studied by Hawkins-Daruud, van der Zee, and Oden [53]. First, some simplified notations are in order. We set

$$\phi_1 = \phi_T = u = \text{tumor cell volume fraction},$$

$$\phi_2 = \phi_c = h = \text{healthy cell volume fraction},$$

$$\phi_3 = \phi_\sigma = n = \text{nutrient-rich extracellular water volume fraction},$$

$$\phi_4 = \phi_{\sigma_0} = w = \text{nutrient-poor extracellular water volume fraction}.$$

Then,

$$u + h + n + w = 1 , \quad (11.22)$$

everywhere in $\bar{\Omega} \times [0, \tau)$. We set $u + h = c$, $n + w = 1 - c$ and then rescale u , h and n , w so that u and h and n and w take on values between 0 and 1.

We assume that the nutrient-rich extracellular water is governed by diffusion and that neither n nor w are represented by a double-well-potential in the energy. We include in the energy functional, a term which attains a minimum when full interaction between the tumor cells and the nutrient rich extracellular water when both $u = 1$ and $n = 1$, thereby driving the tumor concentrations toward the oxygen supply. The total free energy is then assumed to be often form

$$\begin{aligned} E &= \int_{\omega} \widehat{\Psi}(u, n) \, dx \\ &= \int_{\omega} \left(W(u) + \frac{\epsilon^2}{2} |\nabla u|^2 + \chi(u, n) + \frac{\delta^{-1}}{2} n^2 \right) \, dx . \end{aligned} \quad (11.23)$$

Under assumptions leading to equations (9.86), we arrive at the system

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \nabla \cdot \mathcal{M}_u \nabla \mu_u + S_u, \\ \frac{\partial h}{\partial t} &= \nabla \cdot \mathcal{M}_h \nabla \mu_h + S_h, \\ \frac{\partial n}{\partial t} &= \nabla \cdot \mathcal{M}_n \nabla \mu_n + S_n, \\ \frac{\partial w}{\partial t} &= \nabla \cdot \mathcal{M}_w \nabla \mu_w + S_w. \end{aligned} \right\} \quad (11.24)$$

(Note. The mass densities of the constituents are assumed to be constant and the constituent velocities are set equal to zero). According to the mixture theory laid down in Chapter 9, the mixture equations (11.24) are compatible with the global balance laws of

$$\left. \begin{aligned} S_u + S_h + \nabla \cdot \mathcal{M}_u \nabla \mu_u + \nabla \cdot (\mathcal{M}_h \nabla \mu_h) &= 0 \\ S_n + S_w + \nabla \cdot \mathcal{M}_n \nabla \mu_n + \nabla \cdot (\mathcal{M}_w \nabla \mu_w) &= 0 \end{aligned} \right\} \quad (11.25)$$

and

$$\sum_i \mu_i S_i \leq 0, \quad (11.26)$$

where

$$\left. \begin{aligned} \mu_u &= W'(u) - \epsilon^2 \Delta u - D_u \chi(u, n), \\ \mu_h &= -W'(c-h) - \epsilon^2 \Delta(c-h) - D_u \chi((c-h), n), \\ \mu_n &= D_n \chi(u, n) + \delta^{-1} n, \\ \mu_w &= -D_n \chi(u, (1-c)-w) + \delta^{-1} ((1-c)-w). \end{aligned} \right\} \quad (11.27)$$

We easily verify that $\sum_i \mu_i = 0$. In the case of equal mobilities: $M_u = M_h$ and $M_n = M_w$, we arrive at the conditions,

$$\left. \begin{aligned} \nabla \left(M_u \nabla \mu_u \right) + \nabla \cdot \left(M_h \nabla \mu_h \right) &= 0, \\ \nabla \cdot \left(M_n \nabla \mu_n \right) + \nabla \cdot \left(M_w \nabla \mu_w \right) &= 0, \end{aligned} \right\} \quad (11.28)$$

which reduces to

$$S_u + S_h = 0 \quad \text{and} \quad S_n + S_w = 0 . \quad (11.29)$$

Assuming for simplicity, that cell mobility for tumor and healthy cells is the same ($M_u = M_h$) and that $M_n = M_w$, we see that thus $\partial u / \partial t = -\partial h / \partial t$ and $\partial n / \partial t = -\partial w / \partial t$. In these cases, the system (11.24) reduces to only two equations,

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(M_u \nabla (W'(u) - \epsilon^2 \Delta u + D_u \chi(u, n)) \right) + S_u , \quad (11.30)$$

and

$$\frac{\partial n}{\partial t} = \nabla \cdot \left(M_n \nabla (D_n \chi(u, n) + \delta^{-1} n) \right) + S_n \quad (11.31)$$

subject to constraints (11.25) and (11.26).

To proceed, we make additional simplifying assumptions. Motivated by the success of using linear constitutive laws for certain chemical reactions, we take [c.f., 53],

$$\left. \begin{aligned} S_u &= P(u)(\mu_n - \mu_u) , \\ S_n &= -S_u , \end{aligned} \right\} \quad (11.32)$$

Where $P(u) = \delta P_\circ u$, $u \geq 0$, $P(u) = 0$ elsewhere P_\circ being a non-negative threshold constant, and δ is a small constant. Following [53], we take

$$\left. \begin{aligned} S_u &= P_\circ u n + \delta P_\circ u (D_n \chi(u, n) - \mu_u) \\ W(u) &= \gamma u^2 (1 - u)^2 \quad (\gamma = \text{const.}) \\ \chi(u, n) &= -\chi_\circ u n \end{aligned} \right\} \quad (11.33)$$

χ_\circ being a constant ≥ 0 .

Given the governing equations (11.30) and (11.31), with relations (11.32) and (11.33) in force, we can not implement numerical methods, described fully in [53], to solve the resulting system for appropriate boundary conditions.

Example Solutions. The following example is found in [53]. Taking for simplicity, a quasi-slide nutrient concentration ($\partial n / \partial t = 0$), $\Omega = (0, 25.6)^2$. The momenta are taken to be $M_u = \widehat{M} u^2$,

$M_n = \delta \hat{D}$, leading to the system

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \nabla \cdot (\hat{M} u^2 \nabla (W'(u) - \epsilon^2 \Delta u - \chi_\circ n)) + S_u \\ 0 &= \nabla \cdot (\delta \hat{D}(\delta^{-1} n - \chi_\circ u)) + S_n \end{aligned} \right\} \quad (11.34)$$

with $\nabla u \cdot \mathbf{n} = 0$, $n = 1$ on $\partial\Omega \times (0, \tau]$, $u(\mathbf{x}, 0) = u_\circ$, $\mathbf{x} \in \Omega$. We set $\hat{D} = 1$, $P_\circ = 0, 1$ and take as an initial condition, the situation in which the tumor resides in the elliptical domain $\{(x, y) : (x - 12.8)^2/2.1 + (y - 12.8)^2/1.9 \leq 1\}$. For this example, one takes $\delta = 0.01$, $\gamma = 0.045$, $\epsilon = 0.005$, and $\chi_\circ = 0.05$.

Results of a numerical solution of (11.34) given in [53] are represented in Figure 11.1.

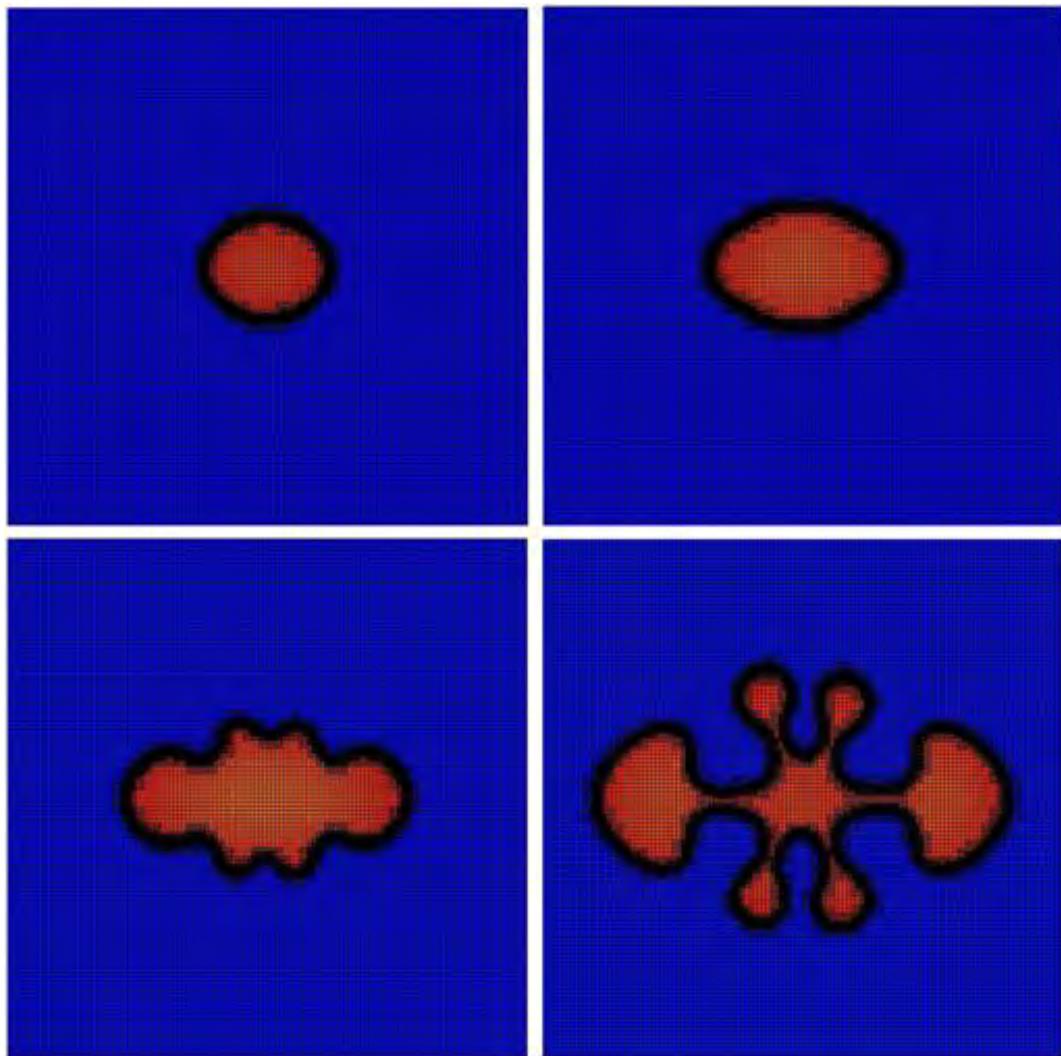


Figure 11.1: Adaptive meshing: illustration of the mesh adaptivity for the simulation shown in Figure 11.2. The patch-recovery method is used for the adaptive meshing (from [53]).

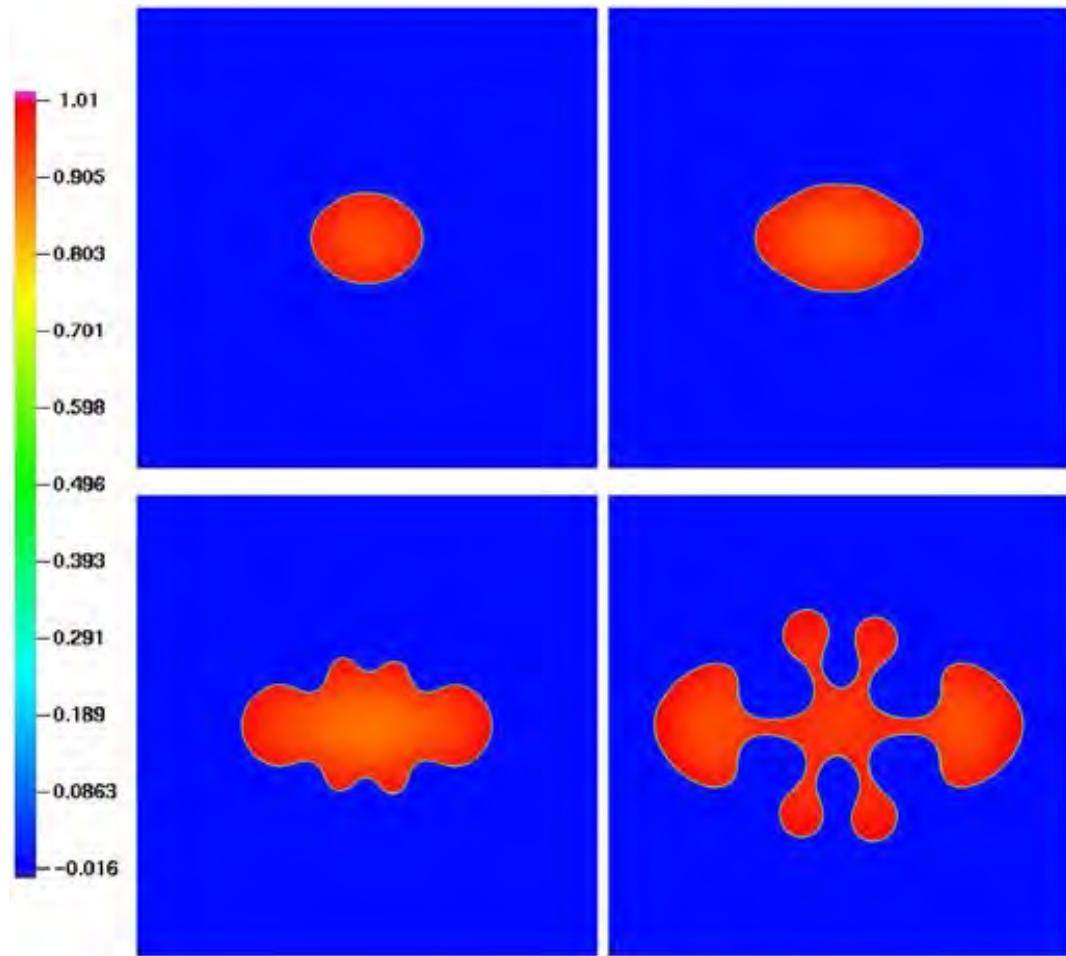


Figure 11.2: Example simulation: snapshots are shown at $t=20, 40, 60$, and 80 of a simulation with $\Gamma = 0.045$, $\epsilon = 0.005$, $\chi_0 = 0.05$, $\delta = 0.01$, $P_0 = 0.1$, $\widehat{M} = 200$, and $\widehat{D} = 1$ (from [53]).

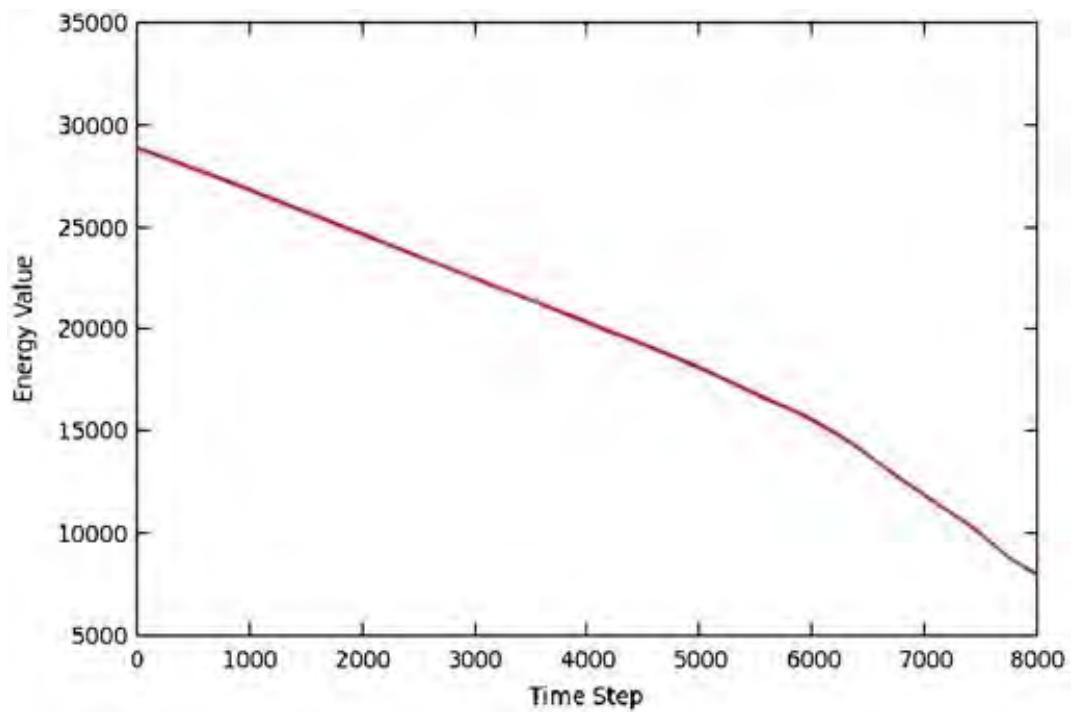


Figure 11.3: Simulation energy profile: shown here is the evolution of the energy value for the simulation shown in Figure 11.2 (from [53]).

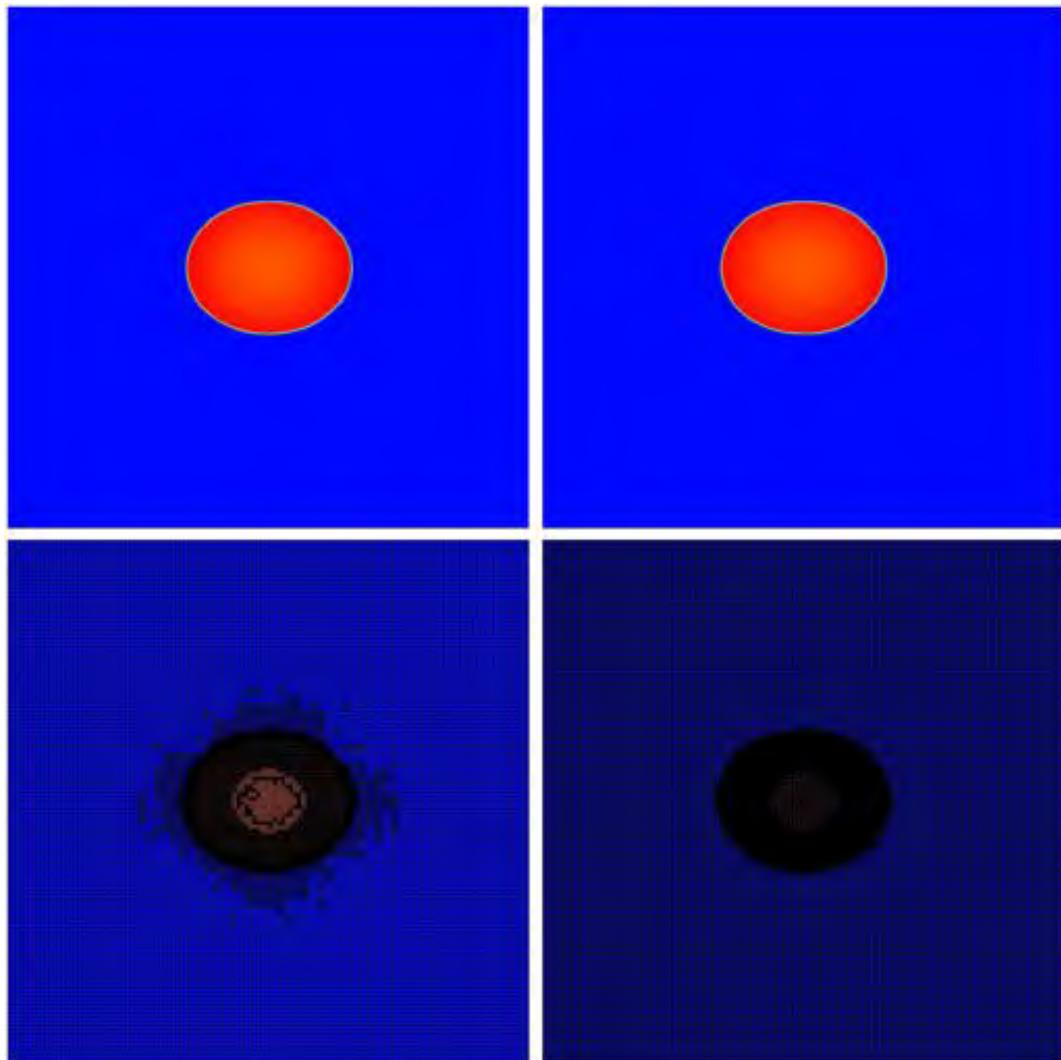


Figure 11.4: Evidence for spatial accuracy: on the top row, the solution at $t = 20$ is shown for the same set of parameters and initial conditions as in Figure 11.2, but with one mesh at a finer resolution than the other. The corresponding final meshes are shown below the solutions (from [53]).

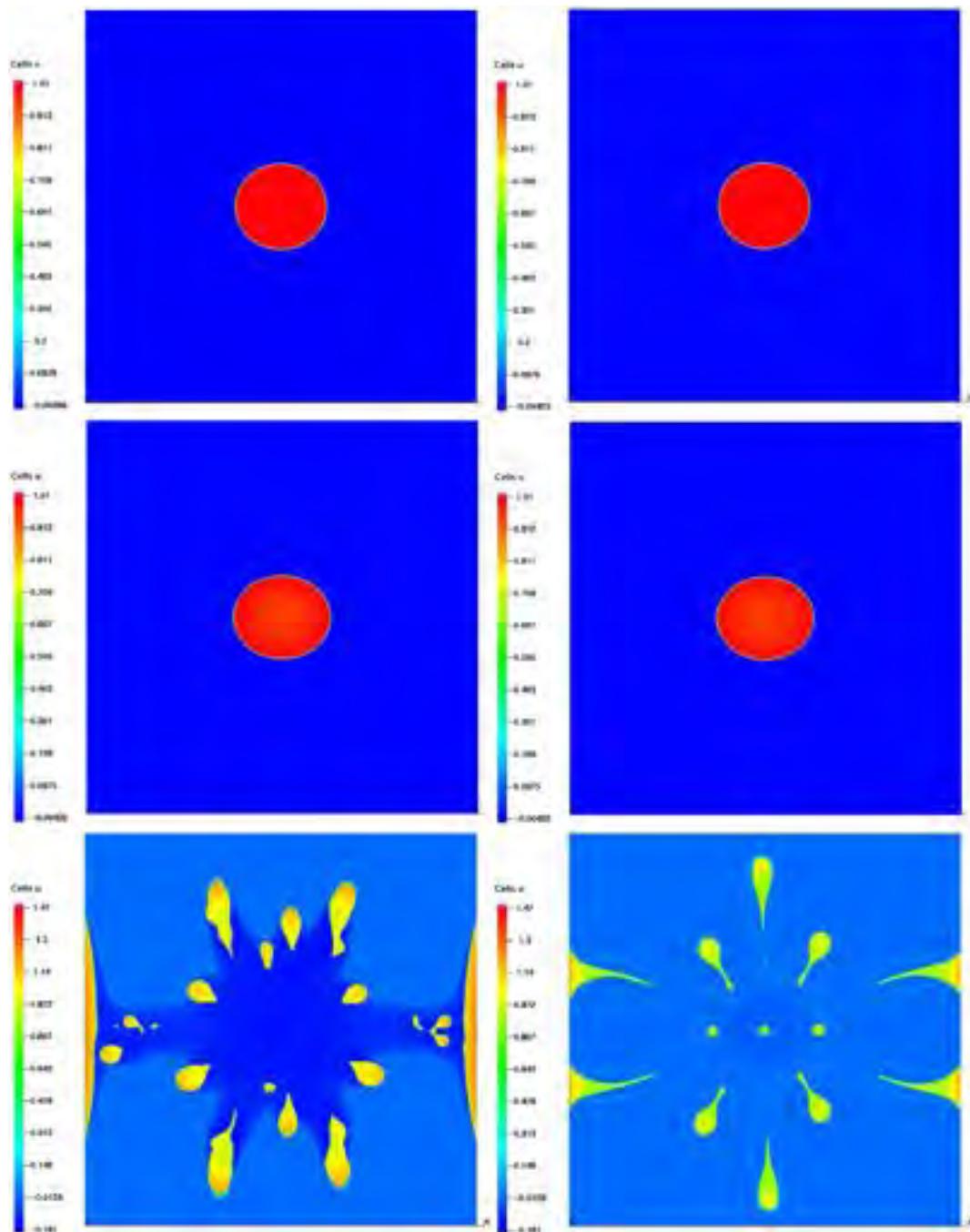


Figure 11.5: Effects of parameter χ_0 : illustrated here are the effects of different values of χ_0 when $\Gamma = 0.045$ and $\epsilon = 0.005$ are held constant. In the first row, $\chi_0 = 0.005$; in the second row, $\chi_0 = 0.05$; and in the third row, $\chi_0 = 0.5$. In the first column, $\delta = 0.1$; and in the second column, $\delta = 0.01$ (from [53]).

Bibliography

- [1] Aczel, J. (1966). *Lectures on Functional Equations and Their Applications*. Academic Press.
- [2] AIAA Standards (1998). Guide for the verification and validation of computational fluid dynamics simulations (AIAA G-077-1998(2002)).
- [3] Ainsworth, M. and Oden, J. T. (2000). *A Posteriori Error Estimation in Finite Element Analysis*. John Wiley & Sons, N.Y.
- [4] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.
- [5] Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion. *System Identification: Advances and Case Studies*, pages 27–96.
- [6] Akaike, H. (1977). *On entropy maximization principle*. P.R. Krishnaiah (ed.), Applications of statistics. North-Holland, Amsterdam, The Netherlands.
- [7] Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In Parzen, E., Tanabe, K., and Kitagawa, G., editors, *Selected Papers of Hirotugu Akaike*, Springer Series in Statistics, pages 199–213. Springer New York.
- [8] Ambrosio, L. and Tortorelli, V. M. (1990). Approximation of functional depending on jumps by elliptic functional via t-convergence. *Communications on Pure and Applied Mathematics*, 43(8):999–1036.
- [9] Araujo, R. P. and McElwain, D. L. S. (2005). A mixture theory for the genesis of residual stresses in growing tissues ii: solutions to the biphasic equations for a multicell spheroid. *SIAM J. Appl. Math.*, 66:447–467.
- [10] Arnborg, S. and Sjodin, G. (2001). On the foundations of bayesianism. In *AIP Conference Proceedings*, volume 568, pages 61–71. AIP Publishing.

- [11] ASME Committee PTC-60 V&V 10 (2006). Guide for verification and validation in computational solid mechanics.
- [12] Babuška, I. and Oden, J. T. (2004). Verification and validation in computational engineering and science: Part I, basic concepts. *Computer Methods in Applied Mechanics and Engineering*, 193(1):4047–4068.
- [13] Babuška, I., Tempone, R., and Nobile, F. (2008). A systematic approach to model validation based on Bayesian updates and prediction-related rejection criteria. *Computer Methods in Applied Mechanics and Engineering*, 197:2517–2539.
- [14] Batra, R. C. (2005). *Elements of Continuum Mechanics*. AIAA.
- [15] Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. M., Cavendish, C., C.-H., L., and Tu, J. (2007). A framework for validation of computer models. *Technometrics*, 49(2):138–154.
- [16] Beck, J. L. (2010). Bayesian system identification based on probability logic. *Structural Control and Health Monitoring*, 17(7):825–847.
- [17] Beck, J. L. and Taflanidis, A. A. (2013). Prior and posterior robust stochastic predictions for dynamical systems using probability logic. *International Journal for Uncertainty Quantification*, 3(4):271–288.
- [18] Beck, J. L. and Yuen, K.-V. (2004). Model selection using response measurements: Bayesian probabilistic approach. *Journal of Engineering Mechanics*, 130(2):192–203.
- [19] Bernstein, S. N. (1927). *Theory of Probability*. Moscow.
- [20] Bowen, R. M. (1976). Theory of mixtures. In Eringen, A. D., editor, *Continuum Physics*, volume 3. Academic Press.
- [21] Box, G. and Draper, N. (1987). *Empirical model-building and response surfaces*. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley.
- [22] Box, G. E. (1979). Robustness in the strategy of scientific model building. *Robustness in statistics*, 1:201–236.
- [23] Braides, A. (2002). *Gamma-convergence for Beginners*, volume 22. Clarendon Press.
- [24] Bui-Thanh, T. (2016). Private communication.
- [25] Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer-Verlag, New York.

- [26] Calude, C. S., Rozenberg, G., and Salomaa, A. (2011). *Rainbow of Computer Science: Essays Dedicated to Hermann Maurer on the Occasion of His 70th Birthday*. Springer.
- [27] Carter, T. and Fe, S. An introduction to information theory and entropy.
- [28] Chow, G. C. (1981). A comparison of the information and posterior probability criteria for model selection. *Journal of Econometrics*, 16(1):21 – 33.
- [29] Coleman, B. and Noll, W. (1963). The thermodynamics of elastic materials with heat conduction and viscosity. *Arch. Ration. Mech. Anal.*, 13:167–178.
- [30] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. Wiley-Interscience, Hoboken, 2nd edition.
- [31] Cox, R. T. (1946). Probability, Frequency and Reasonable Expectation. *American Journal of Physics*, 14(1):1–13.
- [32] Cox, R. T. (1961). *Algebra of Probable Inference*. Johns Hopkins University Press.
- [33] Cox, R. T. (1979). Of inference and inquiry, an essay in inductive logic. *The maximum entropy formalism*, pages 119–167.
- [34] Cristini, V., Li, X., Lowengrub, J., and Wise, S. (2009). Nonlinear simulations of solid tumor growth using a mixture model: invasion and branching. *J. Math. Biol.*, 58:723–763.
- [35] Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., and Schaibly, J. H. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i theory. *The Journal of Chemical Physics*, 59(8):3873–3878.
- [36] Dal Maso, G. (2012). *An introduction to Γ -convergence*, volume 8. Springer Science & Business Media.
- [37] Darcy, H. (1856). *Les Fontaines Publiques de la Ville de Dijon*. Dalmont, Paris.
- [38] Dupre, M. J. and Tipler, F. J. (2006). The Cox Theorem: Unknowns And Plausible Value. *arXiv: math/0611795*.
- [39] Farrell, K., Oden, J. T., and Faghihi, D. (2015). A bayesian framework for adaptive selection, calibration, and validation of coarse-grained models of atomistic systems. *Journal of Computational Physics*, 295(0):189 – 208.
- [40] Fick, A. (1855). *Poggendorff's Annal. Physik*, 94:59.

- [41] Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222:309–368.
- [42] Geyer, C. J. (2003). 5601 notes: The sandwich estimator. *University of Minnesota School of Statistics*.
- [43] Gillies, D. (2000). *Philosophical Theories of Probability*. Routledge, London and New York.
- [44] Gomez, H. and van der Zee, K. G. (2015). Computational phase-field modeling.
- [45] Gregory, P. (2005). *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support*. Cambridge University Press.
- [46] Gurin, M. and McFadden, G. (1992). *On the evolution of Phase boundaries*. Springer.
- [47] Gurin, M. E. (1989). On phase transitions with bulk, interfacial, and boundary energy. In *Analysis and Continuum Mechanics*, pages 429–450. Springer.
- [48] Halpern, J. Y. (1999). Cox’s Theorem Revisited. *arXiv: cs/9911012*.
- [49] Halpern, J. Y. (2011). A Counter Example to Theorems of Cox and Fine. *arXiv: 1105.5450 [cs]*.
- [50] Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *cell*, 100(1):57–70.
- [51] Hanahan, D. and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *cell*, 144(5):646–674.
- [52] Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109.
- [53] Hawkins-Daarud, A., van der Zee, K. G., and Tinsley Oden, J. (2012). Numerical simulation of a thermodynamically consistent four-species tumor growth model. *International journal for numerical methods in biomedical engineering*, 28(1):3–24.
- [54] Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52(1):1 – 17.
- [55] Howson, C. and Urbach, P. (2006). *Scientific Reasoning: The Bayesian Approach*. Open Court Publishing.
- [56] Hurvich, C. M. and Tsai, C. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307.

- [57] II, D. A. H., Weis, J. A., Barnes, S. L., Miga, M. I., Rericha, E. C., Quaranta, V., and Yankeelov, T. E. (2015). Predicting in vivo glioma growth with the reaction diffusion equation constrained by quantitative magnetic resonance imaging data. *Physical Biology*, 12(4).
- [58] Jaynes, E. T. (1968). Prior probabilities. *Systems Science and Cybernetics, IEEE Transactions on*, 4(3):227–241.
- [59] Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge.
- [60] Jaynes, E. T. and Bretthorst, G. L. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.
- [61] Jeffreys, H. (1939). The times of p, s and sks, and the velocities of p and s. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, 4(7):498–533.
- [62] Jeffreys, H. (1961). Small corrections in the theory of surface waves. *Geophysical Journal International*, 6(1):115–117.
- [63] Jeffreys, H. (1998). *The Theory of Probability*. Oxford University Press, 3 edition.
- [64] Kaipio, J. and Somersalo, E. (2004). *Statistical and Computational Inverse Problems*. Springer, Springer-Verlag New York.
- [65] Kennedy, M. C. and O'Hagan, A. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13.
- [66] Kennedy, M. C. and O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3).
- [67] Kleijn, B. J. K. and van der Vaart, A. (2002). The asymptotics of misspecified bayesian statistics. In *Proceedings of the 24th European Meeting of Statisticians*, (eds. T. Mikosch and M. Janzura),.
- [68] Kleijn, B. J. K. and van der Vaart, A. (2012). The Bernstein-Von-Mises theorem under misspecification. *Electronic Journal of Statistics*, 6:354–381.
- [69] Klir, G. J. (2005). *Uncertainty and information: foundations of generalized information theory*. John Wiley & Sons.
- [70] Knuth, K. H. and Skilling, J. (2012). Foundations of Inference. *Axioms*, 1(1):38–73.

- [71] Konishi, S. and Kitagawa, G. (2008). *Information criteria and statistical modeling*. Springer Science & Business Media.
- [72] Lebreton, J.-D., Burnham, K. P., Clobert, J., and Anderson, D. R. (1992). Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. *Ecological Monographs*, 62(1):67–118.
- [73] Lima, E., Oden, J., and Almeida, R. (2014). A hybrid ten-species phase-field model of tumor growth. *Mathematical Models and Methods in Applied Sciences*, 24(13):2569–2599.
- [74] Lima, E., Oden, J., Hormuth, D., Yankeelov, T., and Almeida, R. (2016). Selection, calibration, and validation of models of tumor growth. *Mathematical Models and Methods in Applied Sciences*, pages 1–28.
- [75] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.
- [76] Michaelis, L. and Menten, M. L. (1913). Die kinetik der invertinwirkung. *Biochem. z*, 49(333–369):352.
- [77] Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174.
- [78] Oberkampf, W. L. and Roy, C. J. (2010). *Verification and Validation in Scientific Computing*. Cambridge University Press; 1 editio.
- [79] Oden, J., Babuska, I., and Faghihi, D. (2016). Predictive computational science: Computer predictions in the presence of uncertainty. *Encyclopedia of Computational Mechanics*.
- [80] Oden, J., Moser, R., and Ghattas, O. (2010a). Computer predictions with quantified uncertainty, part i. *SIAM News*, 43(9):1–3.
- [81] Oden, J. T. (2014). Predictive computational science. *IACM Expressions*, 35:2–4.
- [82] Oden, J. T., Hawkins, A., and Prudhomme, S. (2010b). General diffuse-interface theories and an approach to predictive tumor growth modeling. *Mathematical Models and Methods of Applied Science*, 20(3):1–41.
- [83] Oden, J. T., Lima, E. A., Almeida, R. C., Feng, Y., Rylander, M. N., Fuentes, D., Faghihi, D., Rahman, M. M., DeWitt, M., Gadde, M., et al. (2015). Toward predictive multiscale modeling of vascular tumor growth. *Archives of Computational Methods in Engineering*, pages 1–45.

- [84] Paris, J. B. (1994). *The Uncertain Reasoner's Companion: A Mathematical Perspective*. Cambridge University Press.
- [85] Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19.
- [86] Polya, G. (1990). *Mathematics and Plausible Reasoning: Patterns of plausible inference*, volume 2. Princeton University Press.
- [87] Prudencio, E. and Schulz, K. (2012). The parallel C++ statistical library queso: Quantification of uncertainty for estimation, simulation and optimization. In Alexander, M., D'Ambra, P., Belloum, A., Bosilca, G., Cannataro, M., Danelutto, M., Martino, B., Gerndt, M., Jeannot, E., Namyst, R., Roman, J., Scott, S., Traff, J., VallÃ©e, G., and Weidendorfer, J., editors, *Euro-Par 2011: Parallel Processing Workshops*, volume 7155 of *Lecture Notes in Computer Science*, pages 398–407. Springer Berlin Heidelberg.
- [88] Prudencio, E. E., Bauman, P. T., Faghihi, D., Ravi-Chandar, K., and Oden, J. T. (2015). A computational framework for dynamic data-driven material damage control, based on bayesian inference and model selection. *International Journal for Numerical Methods in Engineering*, 102(3-4):379–403.
- [89] Prudencio, E. E. and Schulz, K. W. (2011). The parallel c++ statistical library ‘queso’: quantification of uncertainty for estimation, simulation and optimization. In *European Conference on Parallel Processing*, pages 398–407. Springer.
- [90] Rajagopal, K. R. and Tao, L. (1995). *Mechanics of Mixtures*. World Scientific.
- [91] Rasmussen, C. E. (2006). Gaussian processes for machine learning.
- [92] Roache, P. J. (1998). *Verification and Validation in Computational Science and Engineering*. Hermosa Press, Albuquerque, N.M.
- [93] Rybinski, M. (2008). *Analysis of mathematical models of signalling pathways*. PhD thesis, Master thesis, Uniwersytet Warszawski.
- [94] Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280 – 297.
- [95] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259 – 270.

- [96] Saltelli, A., Chan, K., and Scott, E. (2009). *Sensitivity Analysis*. Number no. 2008 in Wiley paperback series. Wiley.
- [97] Saltelli, A., Chan, K., Scott, E. M., et al. (2000). *Sensitivity analysis*, volume 1. Wiley New York.
- [98] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. Wiley.
- [99] Saltelli, A. and Sobol', I. M. (1995). About the use of rank transformation in sensitivity analysis of model output. *Reliability Engineering and System Safety*, 50(3):225 – 239.
- [100] Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D., and Müller, G. (2002). Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nature biotechnology*, 20(4):370–375.
- [101] Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464.
- [102] Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press.
- [103] Shahmoradi, A. and Oden, J. T. (2017). Computational modeling of cell. *ICES Report, in preparation*.
- [104] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656.
- [105] Sobol', I. M. (1990). Sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 2:112–118.
- [106] Sobol', I. M. (1993). Sensitivity analysis for non-linear mathematical models. *Mathematical Modeling and Computational Experiment*, 1:407–414.
- [107] Sobol', I. M. (2007). Global sensitivity analysis indices for the investigation of nonlinear mathematical models,. *Matematicheskoe Modelirovanie*, 19:23–24.
- [108] Stone, M. H. (1936). The Theory of Representation for Boolean Algebras. *Transactions of the American Mathematical Society*, 40(1):37–111.
- [109] Szabo, B. and Babuška, I. (2011). *Introduction to Finite Element Analysis : Formulation, Verification and Validation*. Wiley.
- [110] Terenin, A. and Draper, D. (2015). Rigorizing and Extending the Cox-Jaynes Derivation of Probability: Implications for Statistical Practice. *ArXiv e-prints*.

- [111] Tribus, M. (2016). *Rational Descriptions, Decisions and Designs: Pergamon Unified Engineering Series*. Elsevier.
- [112] Truesdell, C. (1962). Mechanical basis for diffusion. *J. Chem. Phys.*, 37:2336–2344.
- [113] Truesdell, C. and Noll, W. (1965). The non-linear field theories of mechanics. In Flügge, S., editor, *Handbuch der Physik*, volume III. Springer-Verlag, Berlin.
- [114] Truesdell, C. and Toupin, R. (1960). The classical field theories. In Flugge, S., editor, *Handbuch der Physik*, volume III/I. Springer-Verlag.
- [115] van der Zee, K. G., Tinsley Oden, J., Prudhomme, S., and Hawkins-Daarud, A. (2011). Goal-oriented error estimation for cahn–hilliard models of binary phase transition. *Numerical Methods for Partial Differential Equations*, 27(1):160–196.
- [116] Van Horn, K. S. (2003). Constructing a logic of plausible inference: a guide to Cox’s theorem. *International Journal of Approximate Reasoning*, 34(1):3–24.
- [117] Voet, D., Voet, J. G., and Pratt, C. W. (2016). Fundamentals of biochemistry: life at the molecular level.
- [118] von Mises, R. (1931). *Wahrscheinlichkeitsrechnung*. Springer.
- [119] Wang, Z., Zhang, L., Sagotsky, J., and Deisboeck, T. S. (2007). Simulating non-small cell lung cancer with a multiscale agent-based model. *Theoretical Biology and Medical Modelling*, 4(1):1.
- [120] Wentworth, W. E. (2000). *Physical Chemistry: A Short Course*. Blackwell Publishing.