Modeling and Analysis of Interconnected Systems in the Presence of Uncertainty

S. K. Mitter
Department of Electrical Engineering and Computer Science
and Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
U. S. A.

September 12, 2016

Abstract

In this paper we discuss aspects of the problem of modeling interconnected stochastic dynamical systems in the presence of modeling and measurement uncertainties. These ideas may be relevant to the study of financial systems, for example, the banking system, from the point of view of simulation, architecture and stability.

1 Introduction

The modeling of complex engineering systems and the design of their control and information architecture to obtain adaptive behavior in the presence of uncertainty are fundamental problems of Systems Theory. An integral part of this theoretical problem is understanding the modeling of uncertainty both in the physical models as well as in the observation (measurement) process. Indeed, understanding how uncertainty affects economic behavior is the fundamental problem of Economics [1]. The objective of this paper is to describe and suggest that some of these ideas may be relevant to the modeling of financial systems and the analysis of their stability.

The simulation of the financial system requires modeling it as an interconnection of heterogeneous, uncertain, stochastic dynamical systems. These stochastic dynamical systems will typically be modeled as stochastic differential equations or discrete Markov Chains, obtained through a process of tearing and zooming, see Section 2. Also in Section 2, it is suggested that, ideally, these subsystems will be modeled as behavioral systems — behavior being a family of random trajectories. The structure of the interconnected system can be modeled as a graph. In Section 6, we present a theory of interconnection of stochastic systems is presented following J.C. Willems [2]. See also M. Agarwal and S.K. Mitter [3].

The remainder of this paper deals with uncertainty, feedback and robustness. One of the questions in modeling of uncertainty is whether a probabilistic description is always appropriate. This is discussed in Section 5 in the context of De Finetti's ideas on subjective probability. It is pointed out that De FInetti's idea of coherence of assessments is closely related to the fundamental theorem of asset pricing. Section 5.1 is devoted to stochastic control with imperfect models. It is suggested that the ideas here would be worth examining to make stochastic finance more robust by using a worst-case point of view. Issues of feedback and dissipative systems are discussed in Sections 4 and 6. The important fact that real-time pricing may give rise to large volatility due to the creation of a feedback loop between the consumer and the energy market is pointed out. Thus feedback can have both positive and negative effects, a fact not always appreciated. Section 6.1 deals with the important subject of dissipative systems.

If one had the luxury of redesigning the financial system, it should be designed as a layered hierarchical system. This is the subject of Appendix B. One of the fundamental issues here is the design of interfaces so that the layers are functionally separated. The resulting architecture principles may have a role in elucidating regulatory principles for creating a robust stable financial system.

2 Modeling Interconnected Systems: The Behavioral View of J.C. Willems

Complex interconnected systems [4] are usually modeled by tearing, zooming and linking. "Tearing" refers to viewing a system as an interconnection of subsystems, "zooming" refers to modeling the subsystems, and "linking" refers to setting up the relations which describe a mathematical structure underlying the above modeling process (see Figure 1).

Consider a system that interacts with its environment through terminals, and the dynamics of the variables involved in this interaction need to be described. Assume that there are N-terminals and the variables associated with the kth terminal belongs to the space W_k. The internal structure of the system and the parameter values of the elements lead to constraints on the possible time functions

$$W = (W_1, \ldots, W_N) : \mathbb{R} \to W = W_1 \times W_2 \cdots \times W_N \ .$$

The set of possible time trajectories denoted by $B \subset W$ is called the "behavior" of the system. In the event that $W = U \times Y$, with U designated as input and Y designated as output and, in addition, there is a state space X such that B can be described as the set of (u, y), satisfying (1)

$$\left.\begin{aligned} \frac{dx}{dt} &= f(x,u) \\ y &= h(x,u) \end{aligned}\right\} \tag{1}$$

or through differential algebraic equations. In this case, a causality structure has been imposed. But, there are several possible choices of (u, y) and f, h leading to the same behavior.

The structure of an interconnected system can be modeled as a graph with leaves. The vertices of the graph correspond to subsystems with their dynamical description, the edges correspond to connected terminals and the leaves correspond to external terminals that interact with the environment. The dynamics corresponding to this architecture has a hierarchical structure, since the subsystem at a particular vertex can be viewed as an interconnection architecture of a subsystem.

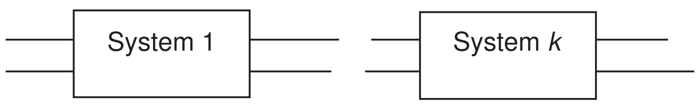


Figure 1: Teasing and Zooming

3 Modeling of Uncertainty: Model Uncertainty vs. Measurement Uncertainty

In the late 1950s and early 1960s, in the Systems and Control field, the input-output view was replaced by the state-space point of view involving differential equations and Markov processes. The Canonical point of view was that there was an underlying probability space $(\Omega, \mathcal{F}, P)$ where the dynamics of the state process evolves according to a controlled stochastic differential equation

$$dx_t = f(x_t)dt + g(u_t)dt + \sigma(x_t)dw_t \ , \tag{2}$$

where w_t is a standard Brownian motion, and the measurement process evolves according to

$$dy_t = h(x_t)dt + dv_t \ , \tag{3}$$

and v_t is also standard Brownian motion, typically independent of w_t.

The problem of control is to choose the controls u_t, adapted to the σ-field $\mathcal{F}_t^y$, to minimize

$$J(u) = \int_0^T c(u_t, x_t)dt + b(x_T) \ . \tag{4}$$

In this model, the existence of a canonical probability space $(\Omega, \mathcal{F}, P)$ is assumed and a full probabilistic description involving the primitive random processes (w_t, v_t) is available to the controller. The question of robustness of these models, in the context of linear systems, that is, the design of controls whose performance (for example, stability) are robust against model uncertainty, was a subject of extensive research in the late 1970s are uncertainty (see, for example, [5]). It was argued that a probabilistic description of model uncertainty was not appropriate since it was unclear how to put a canonical probability measure on the space of models. Indeed, a worst-case view, namely models as elements of bounded sets in H^∞ (Hardy space) would be more appropriate. One of the fundamental questions here is understanding the trade-off between robustness and performance (stability, for example).

Feedback, that is closing the loop by dynamically interconnecting inputs and outputs, plays an important role here. A general view of the function of feedback is to reduce the complexity of uncertain models where complexity is defined as the metric entropy of the H^∞-ball centered round the nominal plant (see [5]). The problem of control for stochastic dynamical systems, given by (2)–(4) with model uncertainty, has been investigated as Convex Programming problems [6].

4 Uncertainty and Feedback

In Control Systems, feedback represents an interconnection structure where the outputs of a dynamical system are connected to the inputs of a dynamical system through another dynamical system. The objective is to change the behavior of the original dynamical system to a more desirable behavior through the feedback interconnection. In particular, feedback is used to stabilize an unstable system. However, the feedback control needs to be designed correctly. If not designed correctly, the feedback may have a destabilizing effect. These ideas are illustrated here in the context of real-time pricing in smart grids [7, 8].

The large-scale integration of renewable resources into the existing power grid for economic and environmental considerations is a fundamental problem in so-called Smart Grids. Dynamic pricing and demand response are often considered to be mechanisms for neutralizing the adverse effects of the essential non-stationary and uncertain nature of renewable generation. One mechanism is to allow the consumers to adjust their consumption in response to a signal (price mechanism) that affects the wholesale market conditions, possibly real-time prices. However, this real-time, or near real-time, coupling between supply and demand poses significant challenges for the stability and robustness of power systems.

The direct linking of price-sensitive consumers to the wholesale electricity markets creates a closed-loop, dynamical system. In the absence of a stabilizing control law, direct feedback may lead to increased volatility, decreased robustness to disturbances and systemic risk. There are additional sources of dynamic behavior, namely, time delay between market clearing and consumption decision (information asymmetry between consumers and system operator) leading to a requirement of demand and price prediction. The other source of dynamical behavior is the dynamics of consumer behavior.

The modeling philosophy to be adopted here is one where models are created at different levels of abstraction, and the architecture of the information and control system is a layered, hierarchical, architecture. In the Appendix, an example is given of a layered architecture rising from the internet architecture [10].

In this current exposition, an abstracted model of consumer behavior is considered, for example, where the "state" of the individual consumer is given by the location of marginal prices. The electricity market model has three participants: (1) suppliers, (2) consumers, and (3) an Independent System Operator (ISO). The suppliers and consumers are price-taking, profit-maximizing agents. The ISO is in charge of clearing the market, that is matching supply and demand, subject to network constraints with the objective of maximizing the social welfare. The solution to the resulting optimization problems (concave, maximizing) and information exchange of prices, results in dynamic models of supply-demand under real-time pricing.

In [8], a theoretical framework to study the effects of real-time pricing on the stability and volatility of power systems has been studied. Exposing retail consumers to real-time whole market prices creates a closed-loop feedback control system. It was pointed out in [8] that the feedback control system needs to be properly designed to limit volatility and to provide adequate robustness of margins against external disturbances. For recent work on this subject, see [9].

5 Stochastic Finance [11]

The standard setting in mathematical finance is probabilistic where it is assumed that there is a probability P describing the uncertainties present in the system. The absence of arbitrage implies the existence of an equivalent martingale measure P^* which is interpreted as a consistent price system that represents the market's belief. Since P and P^* are equivalent, and P^* is a martingale measure, one may move between P and P^* via the Girsanov Theorem [12]. The probability measure P captures typical patterns of behavior observed in the past. There is a stationarity assumption here which is difficult to justify. In De Finetti's work [13], the existence of such an objective probability P is questioned. However, an expectation $\int H dP^*$ of the discounted future cash flow H, generated by the bond, is assigned each day in the financial market, either directly through the market price of the bond or by the prices of instruments such as credit default swaps which provide insurance agains a default of the bond. Thus the probability measure P^* reflects the aggregate of a large number of bets made on the market. Therefore even if, as De Finetti argues, an objective probability does not exist, one can take bets on a given bet of certain odds. In De Finetti's terminology these are assessments, that is, functions from the space of events in the interval $[0, 1]$. If these assessments satisfy a "coherence" condition, then it can be shown that "coherence" is equivalent to the existence of a probability measure. For a detailed discussion, see the Appendix A. See also [14]. A detailed exposition of

De Finetti's ideas in the context of expert assessments is presented in the doctoral dissertation of Peter Jones at MIT [15]. A detailed relationship between coherence and arbitrage is discussed there.

Taking a dynamical view, the market's predication of future developments at a given time, t, is given by a conditional distribution

$$P_t^*(\,\cdot\,|\mathcal{F}_t) \;\text{ defined on } \hat{\mathcal{F}}_t$$

where $\mathcal{F}_t$ is the information available at time t and $\hat{\mathcal{F}}_t$ is the σ-field generated by the pay-offs of traded contingent claims with maturities $T > t$. An understanding of the dynamics of P_t^* would involve the microstructure of financial markets, that is, the dynamic behavior of agents with heterogeneous preferences and expectations.

5.1 Stochastic Control with Imperfect Models

Recent work on stochastic control with imperfect models [6] is relevant to estimating risk or, more generally, an appropriate performance index.

Consider the problem of estimating risk in the context of an underlying state process that is a continuous time diffusion. Take the "worst case" approach, that is, estimating the minimal performance ($\approx$ negative of maximal risk in risk estimation) over the allowed class of models.

The following two basic forms of uncertainties can be identified:

- The first is the *modeling uncertainty* wherein the drift and diffusion coefficients of the diffusion are not exactly known. This is modeled by introducing a hypothetical control process. This changes it to a controlled diffusion model. The performance is then minimized over the allowed control processes. With finance applications in mind, this problem with additional constraints, can be thought of as a "constrained" control problem [16].
- The second form of uncertainty has to do with *unmodeled dynamics*. Here we assume that there may be certain state variables about whose dynamics something is known, but these are not observed. On the other hand, certain other state variables are observed and are modeled *separately* by Markov diffusions, but there is uncertainty about their dependence structure. This scenario is labeled as "uncertainty in dynamics." This formulation is explicitly motivated by problems arising in credit risk (see, e.g., [17, ch. 9]), where stochastic dynamic models for two or more processes are separately available, but their dependence structure is unknown.

Combining both of these sources of uncertainty, the problem is cast as an abstract linear program over suitably defined occupation measures. This is an infinite-dimensional linear program. The dual programming problem is defined over the space of functions and it can be shown that there is no duality gap.

5.1.1 Model Uncertainty

Let $x(\cdot)$ be a d-dimensional diffusion process

$$\begin{aligned} dx(t) &= [x(t)]dt + \sigma^*[x(t)]dW(t) \\ x(0) &= x_0\ ,\ t \geq 0 \end{aligned} \tag{5}$$

satisfying standard non-degeneracy conditions.

$$\begin{aligned} dx(t) &= m[x(t), u'(t)]dt + \sigma[x(t), u''(t)]dW_t \\ \text{Law of } &\rightarrow x(0) = \sigma_0 \end{aligned} \tag{6}$$

where u' and u'' are U-valued processes of the form $u'(t) = v'[x(t), t]$, and $u''(t) = v''[x(t), t]$. here U is a compact, metric space. u' and u'' are referred to as Markov controls, which are non-anticipative. Eq. 6 captures both models of uncertainty.

Let $Z = [0, T] \times \mathbb{R}^d$. Associated with (6) is the occupation measure $\nu \in \mathcal{P}(Z \times U)$ defined by

$$\int_{Z \times U} f d\nu = \frac{1}{T} E\Big[\int_0^T f[t, x(t), u(t)]dt\Big]\ . \tag{7}$$

Let $\mathcal{L}$ be the controlled extended generator of (6) with domain $\mathcal{D}(\mathcal{L})$. Let $\mathcal{M}$ denote the set of occupation measures.

Proposition 5.1

$$\mathcal{M} = \Big\{\nu \in \mathcal{P}(Z \times U)\ :\ T\int_{Z \times U} dg d\nu = -\int_{\mathbb{R}^d} g(0, \cdot)d\varphi_0\ :\ \forall\ g \in \mathcal{D}(\mathcal{L})\Big\}.\ \blacksquare \tag{8}$$

The goal is to estimate performance metrics of the form

$$\int_{Z \times U} k d\nu = \frac{1}{T} E\Big[\int_0^T k[t, x(t), u(t)]dt\Big]\ , \tag{9}$$

with constraints $\displaystyle\int_{Z \times U} k_i d\nu = c_i,\ 1 \leq i \leq m.$

It turns out that both the uncertain model situation, as well as the unmodeled dynamics situation, can be studied through an optimization problem of the following "representative" form:

$$\min_{\nu \in m} \int_{Z \times U} k d\nu$$

where $\mathcal{M}$ is given by (8) with the further restraints given by (9). Note that $\mathcal{M}$ is a convex compact set and the constraints given by (9) are linear. This is a convex programming problem. The dual problem is given by

$$\sup_{g \in \mathcal{D}(\mathcal{L})} \int g(0)d\varphi(0) - \sum_{i=1}^m \lambda_i c_i\ :\ k(x, u) + \sum_{i=1}^m \lambda_i k_i(x, u) + \mathcal{L}g(x, u) \geq 0$$

$$(x, u) \in Z \times U \ .$$

One can prove that there is no duality gap under a natural constraint qualification. Approximation schemes for the dual problem are available. This problem is analogous to dual problems arising in Optimal Transportation.

This formulation is a worst-case formulation, but there is an underlying probability measure describing the diffusion process. However, there are situations in economics and finance where it would be difficult to assert the existence of this probability measure. This would correspond to Knightian Uncertainty.

6 Interconnections and Stability

In the previous sections, recent ideas suggested from Systems Theory that could be relevant to modeling the Financial system and its control and information architecture are discussed. The ideas of Layering and Layering Architecture, discussed in Appendix B, may be relevant to defining structural modifications, as, for example, in inserting regulatory interfaces to make the current system more robust against disturbances. Fundamental to answering these questions is understanding the mathematical structure of interconnecting stochastic systems. The discussion of this question follows J.C. Willems [2].

A stochastic system is simply a probability space $(\Omega, \mathcal{F}, P)$, where Ω is a complete separable metric space, $\mathcal{F}$, the Borel σ-field, and $\tilde{\mathcal{F}}$ as sub σ-field of $\mathcal{F}$. For example, if $(x_t)_{t \geq 0}$ is a diffusion process, given

$$dx_t = m(x_t)dt + \sigma(x_t)dw_t \ , \tag{10}$$

defined on a probability space $(\Omega, \mathcal{F}, P)$, then the stochastic system defined by (10) is $(\Omega, \mathcal{F}_t^x, P)$, where $\mathcal{F}_t^x$ is the sub-σ-field generated by (x_s), $0 \leq s \leq t$.

For the remainder of this discussion, assume there is a measurable space $(\Omega, \mathcal{F})$, with Ω countable. Two σ-algebras, $\mathcal{F}_1$ and $\mathcal{F}_2$ on Ω are said to be complementary if all non-empty sets $F_1, F_1' \in \mathcal{F}_1$ and $F_2, F_2' \in \mathcal{F}_2$

$$[F_1 \cap F_2 = F_1' \cap F_2'] \Rightarrow [F_1 = F_1' \text{ and } F_2 = F_2'] \ .$$

Consider two stochastic systems,

$$\Sigma_1 = (\Omega, \mathcal{F}_1, P_1) \ \text{ and } \ \Sigma_2 = (\Omega, \mathcal{F}_2, P_2) \ .$$

They are said to be complementary if $F_1, F_1' \in \mathcal{F}_1$ and if $F_2, F_2' \in \mathcal{F}_2$,

$$[F_1 \cap F_2 = F_1' \cap F_2'] \Rightarrow P_1(F_1)P_2(F_2) = P_1(F_1') = P_2(F_2') \ .$$

Complementarity of σ-algebras imply complementarity of stochastic systems.

Let $\Sigma_1 = (\Omega, \mathcal{F}_1, P_1)$ and $\Sigma_2 = (\Omega, \mathcal{F}_2, P_2)$ be complementary stochastic systems. Then the intersection of Σ_1 and Σ_2 is defined as a system $(\Omega, \mathcal{F}, P)$, where $\mathcal{F}$ is the σ-field generated by $\mathcal{F}_1 \cup \mathcal{F}_2$ and P is defined on $[F_1 \cap F_2 | F_1 \in \mathcal{F}_1 \text{ and } F_2 \in \mathcal{F}_2]$ by

$$P(F_1 \cap F_2) = P_1(F_1) P_2(F_2) \ .$$

P is then extended to all of $\mathcal{F}$.

As an example, a probabilistic version of the supply-demand curve can be constructed using the above ideas. These ideas can be extended to the interconnection

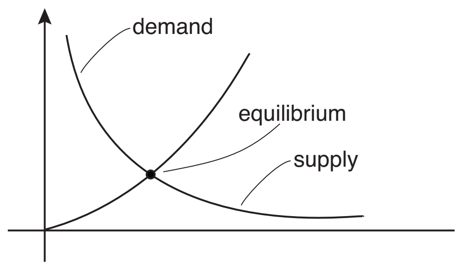


Figure 2: Supply demand curve.

of stochastic systems given by

$$\begin{aligned} dx_1(t) &= m_1[x_1(t)]dt + dw_1(t) \\ dx_2(t) &= m_2[x_2(t)]dt + dw_2(t) \end{aligned}$$

by working with σ-fields $\mathcal{F}_t^{x_1}$ and $\mathcal{F}_t^{x_2}$ on $(\Omega, \mathcal{F}, P)$. The stochastic system $(\Omega, \mathcal{F}, P)$ is said to be deterministic if $\tilde{\mathcal{F}} = \{\phi, F, F^{\text{complement}}, \Omega\}$, $P(F) = 1$. $\tilde{\mathcal{F}}$ is called the behavior of the system.

An alternative way to describe stochastic systems is in terms of a stochastic kernel. Let X and Y be complete separable metric spaces. Then a stochastic kernel is a map, $x \mapsto \mathcal{P}(Y, \mathcal{B}_Y)$. Then the calculus of defining joint stochastic kernels from two stochastic kernels and the disintegration of a joint stochastic kernel into the marginal and conditional stochastic kernels allows one to interconnect stochastic systems.

6.1 Dissipative Systems

The concept of a dissipative system [18] will now be described. The concept involves

(i) A memoryless function of the input and output $s(u, y)$, called the "supply rate."

(ii) A non-negative memoryless function of the state $V(x)$.

(iii) An inequality that involves the system trajectories, the supply rate, and the storage function called the "dissipation inequality"

$$V[x(t_1)] - V[x(t_0)] \leq \int_{t_0}^{t_1} s[u(t), y(t)]dt \ , \tag{11}$$

$\forall\, t_0 \le t_1$, and all trajectories $u(\,\cdot\,), y(\,\cdot\,), x(\,\cdot\,)$ satisfying the dynamical equations

$$(\Sigma) \qquad \begin{aligned} \frac{dx}{dt} &= f[x(t), u(t)] \\ y(t) &= h[x(t), u(t)] \end{aligned} \tag{12}$$

A system (Σ) is said to be dissipative if there exists a non-negative storage function $V(x)$ such that the dissipation inequality (11) is satisfied.

A system (Σ) is said to be lossless if the inequality (11) is satisfied as equality. The concept of a dissipative system can be defined when a behavioral description of a dynamical system is considered. The idea of a storage function is intimately related to that of a Lyapunov function and constitutes a generalization of a Lyapunov function to open systems, in this case, input-output systems. Thus, dissipative systems are stable in an appropriate sense. It can be shown that the interconnection of two dissipative systems through a lossless system is dissipative (see, J.C. Willems, loc. cit).

Thus, to check stability of an interconnected system, it needs to be "torn" into sub-systems and interconnections with checks of both the dissipativeness of sub-systems by finding storage functions satisfying the dissipation inequality, and the dissipativeness of the interconnections. A similar theory can be developed for interconnected behavioral systems (cf: J.C. Willems and K. Takaba [19]).

The ideas of dissipativeness and storage functions can be extended to a stochastic setting where the idea of a deterministic non-negative storage function V is replaced by V being a super-martingale (cf: V.S. Borkar and S.K. Mitter [20]).

Conclusions

In this paper we have suggested that ideas of modeling of interconnected stochastic dynamical systems and their control, resulting in a robust and stable system, are relevant to understanding issues of robustness and stability of the financial system. There are several subjects which have not been discussed in this paper. Replacing Brownian motion with Levy processes (power law) would make Stochastic Finance more robust. Models for systemic risk, where extreme volatility is at the system level, which may occur through interactions even when the volatility is low at the subsystem level, should be studied. The theory of Large Deviation also has a natural role to play here. I believe that an interdisciplinary research program centered around these ideas would be beneficial to fields such as Stochastic Finance and Macroeconomics.

Appendix: Modeling of Uncertainty

An exposition of De Finetti's ideas of subjective probability is presented in the appendix.

A On De Finetti Coherence and Kolmogorov Probability [10]

A.1 Finite Probability Spaces

Let Ω be an arbitrary set and $\mathcal{A} = \{A_i\}_{i=1}^N$ a finite collection of nonempty subsets of Ω.

Definition *A probability assessment on* $(\Omega, \mathcal{A})$ *is a function* $\tilde{P}$ *mapping each set* A *in* $\mathcal{A}$ *to a number* $\tilde{P}(A) \in [0,1]$. *We denote a probability assessment by* $(\Omega, \mathcal{A}, \tilde{P})$.

Definition *A probability assessment* $(\Omega, \mathcal{A}, \tilde{P})$ *is said to be* coherent *if for all* $[c_1, \cdots, c_N] \in \mathcal{R}^N$,

$$\max_{\omega \in \Omega} \sum_{i=1}^{N} c_i [I_{A_i}(\omega) - \tilde{P}(A_i)] \geq 0. \tag{13}$$

This can be given the following convex analytic interpretation: The vector $[\tilde{P}(A_1), \cdots, \tilde{P}(A_N)]$ is in the closed convex hull of the finite set

$$B = \{[I_{A_1}(\omega), \cdots, I_{A_N}(\omega)] : \omega \in \Omega\}.$$

See for example [21].

Therefore, probabilities $p(e)$ can be assigned to each element $e = (e_1, \ldots, e_N)$ of the finite set B such that

$$[\tilde{P}(A_1), \ldots, \tilde{P}(A_N)] = \sum_{e \in B} e p(e).$$

It is easy to see that the following collection of subsets of Ω

$$\left\{ \bigcap_i I_{A_i}^{-1}(e_i); e = (e_1, \ldots, e_N) \in B \right\}.$$

form a partition that generate the same σ-field as the collection $\mathcal{A}$. The probabilities on the set B can be thought of as probabilities of these partitions and therefore define a probability measure on the σ-field generated by the collection $\mathcal{A}$. Therefore we have the following theorem.

Theorem A.1 *Consider a probability assessment* $(\Omega, \mathcal{A}, \tilde{P})$. *Let* $\mathcal{F}$ *be the finite algebra generated by the collection* $\mathcal{A}$. *Then there exists a probability measure* $\mathbf{P}$ *on* $(\Omega, \mathcal{F})$ *such that* $\mathbf{P}(A) = \tilde{P}(A)$ *for all* A *in* $\mathcal{A}$ *if and only if the probability assessment* $(\Omega, \mathcal{A}, \tilde{P})$ *is coherent.*

A.2 Finitely Additive Extensions

We first extend the above definitions as follows:

Definition For an arbitrary collection $\mathcal{A}$ **of nonempty subsets of** Ω**, a probability assessment on** $(\Omega, \mathcal{A})$ **is defined exactly as in Definition 1, whereas a probability assignment** $(\Omega, \mathcal{A}, \tilde{P})$ **will be said to be coherent if (1) holds for all finite subcollections** $\{A_1, \cdots, A_N\} \subset \mathcal{A}$ **and** $[c_1, \cdots, c_N] \in \mathcal{R}^N$**,** $N \geq 1$**.**

We first consider a countable $\mathcal{A}$, enumerated as $\{A_1, A_2, \cdots\}$. Let $(\Omega, \mathcal{A}, \tilde{P})$ be a coherent probability assessment. We shall denote by $\sigma(\mathcal{C})$ the σ-algebra generated by a family $\mathcal{C}$ of sets. Let $\mathcal{A}_n = \{A_1, \cdots, A_n\}$ for $n \geq 1$. By Theorem 2.1, the set $\mathcal{P}_n$ of probabilities compatible with $\tilde{P}$ restricted to $\mathcal{A}_n$ is nonempty for each n. Identifying $\mathcal{P}_n$ with a subset of the simplex of probability vectors in $\mathcal{R}^{|\sigma(\mathcal{A}_n)|}$, one easily verifies that it is convex compact. For $m > n$, let $\Pi_{m,n}(P)$ for $P \in \mathcal{P}_m$ denote the element of $\mathcal{P}_n$ obtained by restricting P to $\sigma(\mathcal{A}_n)$.

Lemma A.1

(i) *For* $k > m > n$, $\Pi_{m,n} \circ \Pi_{k,m} = \Pi_{k,n}$.

(ii) *For* $m > n$, $\Pi_{m,n}(\mathcal{P}_m) \subset \mathcal{P}_n$ *and is compact nonempty.*

(iii) *For* $k > m > n$, $\Pi_{k,n}(\mathcal{P}_k) \subset \Pi_{m,n}(\mathcal{P}_m)$.

(iv) $\mathcal{P}_n^* \triangleq \bigcap_{m>n} \Pi_{m,n}(\mathcal{P}_m) \subset \mathcal{P}_n$ *is compact nonempty.*

(v) $\Pi_{m,n}(\mathcal{P}_m^*) = \mathcal{P}_n^*$ *for* $m > n$.

Proof. $(i) - (iii)$ are easily verified. (iv) follows from the finite intersection property of families of compact sets. (v) follows from the definition of $\mathcal{P}_n^*$. □

Pick $\mu_n^0 \in \mathcal{P}_n^*$ for $n \geq 1$ and for $m > n$, let $\mu_n^{m-n} = \Pi_{m,n}(\mu_n^0) \in \mathcal{P}_n^*$. Let $\{\mu_1^{n(k)}\}$ denote a subsequence of $\{\mu_1^n\}$ in $\mathcal{P}_1^*$ converging to some $\mu_1^* \in \mathcal{P}_1^*$. Let $\{\mu_2^{n(k(m))}\}$ denote a subsequence of $\{\mu_2^{n(k)}\}$ converging to some $\mu_2^* \in \mathcal{P}_2^*$. Proceeding thus and using a diagonal argument, we can pick $\tilde{\mu}_n \in \mathcal{P}_n^*, n \geq 1$, such that $\Pi_{m,n}(\tilde{\mu}_m) \to \mu_n^*$ as $n < m \to \infty$. Clearly, $\Pi_{m,n}(\mu_m^*) = \mu_n^*$. We have:

Theorem A.2 *For an arbitrary family $\mathcal{A}$ of nonempty subsets of Ω with $\mathcal{F}$ = the algebra it generates, a probability assessment $(\Omega, \mathcal{A}, \tilde{P})$ is coherent if and only if there exists a finitely additive probability $\mathbf{P}$ on $(\Omega, \mathcal{F})$ that agrees with $\tilde{P}$ on $\mathcal{A}$.*

Proof. The '*only if*' part follows as in Theorem 2.1. For the '*if*' part, consider first a countable $\mathcal{A}$. In the above notation, set $\mathbf{P} = \mu_n^*$ on $\mathcal{A}_n$ for $n \geq 1$. This consistently defines a finitely additive probability on $\mathcal{F} = \bigcup_n \sigma(\mathcal{A}_n)$. For arbitrary $\mathcal{A}$, let $(\mathcal{A}_\alpha, \mu_\alpha), \alpha \in \mathcal{I}$, denote a nested family of countable subsets $\{\mathcal{A}_\alpha\}$ of $\mathcal{A}$ equipped with finitely additive probabilities $\{\mu_\alpha\}$ on the corresponding algebras $\{\mathcal{F}_\alpha\}$ such that the following consistency condition holds: $\mathcal{F}_{\alpha_1} \subset \mathcal{F}_{\alpha_2}$ implies μ_{α_2} restricts to μ_{α_1} on $\mathcal{F}_{\alpha_1}$. Then $\mathcal{F} = \bigcup_\alpha \mathcal{F}_\alpha$ is the algebra generated by $\bar{\mathcal{A}} = \bigcup_\alpha \mathcal{A}_\alpha$ and for $A \in \bar{\mathcal{A}}$, $\mu(A) \stackrel{\Delta}{=} \mu_\alpha(A)$ for any α such that $\mathcal{A}_\alpha$ contains A, defines a finitely additive probability on $(\Omega, \mathcal{F})$ in a consistent way. Consider the family of pairs $(\hat{\mathcal{A}}, \hat{\mu})$, where $\hat{\mathcal{A}} \subset \mathcal{A}$ and $\hat{\mu}$ is a finitely additive probability on the algebra generated by $\hat{\mathcal{A}}$. Define a partial order on this family by setting $(\mathcal{B}, \nu) < (\mathcal{D}, \eta)$ if $\mathcal{B} \subset \mathcal{D}$ and η restricts to ν on $\mathcal{B}$. By the foregoing, this family is nonempty. Also, every ordered chain w.r.t. this partial order has a least upper bound: for any ordered family $\{(\mathcal{A}_\alpha, \mu_\alpha), \alpha \in \mathcal{I}\}$, $\bar{\mathcal{A}}, \mu$ defined as above would provide a least upper bound. Thus by Zorn's lemma, there exists a maximal element $(\mathcal{A}^*, \mu^*)$. We are done if $\mathcal{A}^* = \mathcal{A}$. Suppose not. Take $A \in \mathcal{A} - \mathcal{A}^*$. Then the algebra generated by $\mathcal{A}^* \cup \{A\}$ is given by $\tilde{\mathcal{G}} \stackrel{def}{=} \{(A \cap B_1) \cup (A^c \cap B_2) : B_1, B_2 \in \mathcal{G}\}$, where $\mathcal{G}$ is the algebra generated by $\mathcal{A}$. Extend μ^* to a finitely additive probability $\tilde{\mu}$ on $\tilde{\mathcal{G}}$ by setting

$$\tilde{\mu}((A \cap B_1) \cup (A^c \cap B_2)) \stackrel{def}{=} \tilde{\mathcal{P}}(A)\mu^*(B_1) + (1 - \tilde{\mathcal{P}}(A))\mu^*(B_2)$$

for $B_1, B_2 \in \mathcal{G}$. That this does indeed define a finitely additive probability on $\tilde{\mathcal{G}}$ is easily verified. Then $(\mathcal{A} \cup \{A\}, \tilde{\mu})$ contradicts the maximality of $(\mathcal{A}^*, \mu^*)$. It follows that $\mathcal{A}^* = \mathcal{A}$. This completes the proof. □

A.3 Countably Additive Extensions

As is well known, not all finitely additive probabilities on σ-algebras lead to countably additive extensions. Thus to make a claim akin to the above for countably additive probability measures, we need to impose additional conditions, *stated in terms of our initial collection $\mathcal{A}$ of events.* We give such a condition below. For a set $A \in \mathcal{A}$, let A^i denote A if $i = 0$ and A^c if $i = 1$. The condition is:

(†) If $A_n \in \mathcal{A}, n \geq 1$, satisfies $\bigcap_n A_n^{i(n)} = \phi$ for some choice of $i(n) \in \{0, 1\}, n \geq 1$, then $\bigcap_{n=1}^N A_n^{i(n)} = \phi$ for some $1 \leq N < \infty$.

Remark.

(1) This condition is necessary. Consider, for example, a countable $\mathcal{A} = \{A_1, A_2, \cdots\}$ and define $\mathcal{A}_n = \{A_1, A_2, \cdots, A_n\}$ for $n \geq 1$. Let $\mathcal{F}$ denote

the Boolean algebra generated by $\mathcal{A}$*. Suppose that for some choice of* $i(n) \in \{0,1\}, n \geq 1$, $\bigcap_n A_n^{i(n)} = \phi$*, but* $\bigcap_{n=1}^N A_n^{i(n)} \neq \phi$ *for all finite* $N \geq 1$*. Define a probability* μ_n *on* $(\Omega, \sigma(\mathcal{A}_n))$ *by setting* $\mu_n(A) = 1$ *if* $A \in \sigma(\mathcal{A}_n)$ *contains* $\bigcap_{m=1}^n A_m^{i(m)}$ *and zero otherwise. Then the finitely additive probability* μ *defined on* $(\Omega, \mathcal{F})$ *by* $\mu(A) = \mu_n(A)$ *for* $A \in \sigma(\mathcal{A}_n)$ *is well-defined and corresponds to a coherent probability assignment by Theorem 2. However,*

$$\lim_{N\to\infty} \mu(\bigcap_{n=1}^{N} A_n^{i(n)}) = 1 \neq 0 = \mu(\phi).$$

Thus μ *does not extend to a countably additive probability on* $\sigma(\mathcal{A})$*. (This example is adapted from [24], pp. 141-142.)*

(2) *As an example of a situation where (†) is satisfied, consider the case when each* $A \in \mathcal{A}$ *intersects at most finitely many other sets in* $\mathcal{A}$*. Then* $\bigcap_{n=1}^N A_n^{i(n)} \neq \phi$ *for* $1 \leq N < \infty$ *would perforce imply that for large* N, $\bigcap_{n=1}^N A_n^{i(n)}$ *equals the intersection of a fixed finite subcollection of sets from* $\mathcal{A}$*, whence (†) follows.*

Theorem A.3 *If a probability assignment* $(\Omega, \mathcal{A}, \tilde{P})$ *is coherent and* $\mathcal{A}$ *satisfies (†), then there exists a countably additive probability* $\boldsymbol{P}$ *on* $(\Omega, \sigma(\mathcal{A}))$ *that agrees with* $\tilde{P}$ *on* $\mathcal{A}$*.*

We shall need two preliminary lemmas.

Lemma A.2 *If* $\mathcal{B} = \{A_1, A_2, \cdots\} \subset \mathcal{A}$ *is a countable subfamily, then the atoms of* $\sigma(\mathcal{B})$ *are precisely the nonempty sets of the form* $\bigcap_n A_n^{i(n)}$, $i(m) \in \{0,1\}, n \geq 1$*.*

Proof. Consider the collection of sets A with the property: Given any set C of the above form, either $C \subset A$ or $C \subset A^c$. It is easy to see that this is a sigma field that contains $\mathcal{B}$, and therefore contains $\sigma(\mathcal{B})$. Also, the latter contains sets of the form $\bigcap_n A_n^{i(n)}$, $\{i(n)\}$ as above. The claim follows. □

Lemma A.3 $\sigma(\mathcal{A}) = \bigcup \sigma(\mathcal{B})$ *where the union is over all countable* $\mathcal{B} \subset \mathcal{A}$*.*

Proof. The r.h.s. is clearly contained in the l.h.s. The claim follows on noting that the r.h.s. is also a σ-field. □

Proof of Theorem 4.1: Let $\mathcal{B} = \{A_1, A_2, \cdots\} \subset \mathcal{A}$ and $\mathcal{A}_n = \{A_1, A_2, \cdots, A_n\}, n \geq 1$. Then $\sigma(\mathcal{A}_n), n \geq 1$, is an increasing family of (finite) σ-fields and $\sigma(\mathcal{B})$ is the smallest σ-field containing $\sigma(\mathcal{A}_n), n \geq 1$. Let μ be the finitely additive probability measure guarranteed by Theorem 3.1. Then by (†), Lemma 4.1 above and Theorem 4.1, pp. 141-143 of [24], it extends to a unique countably additive probability measure on $\sigma(\mathcal{B})$. Since $\mathcal{B}$ was an arbitrary countable subset of $\mathcal{A}$, the claim follows in view of Lemma 4.2 above. □

B Layering and Layered Architectures

[1]Layering, or layered architecture, is a form of hierarchical modularity that is central to data network design. The concept of modularity (although perhaps not the name) is as old as engineering. In what follows, the word *module* is used to refer either to a device or to a process within some computer system. What is important is that the module performs a given function in support of the overall function of the system. Such a function is often called the *service* provided by the module. The designers of a module will be intensely aware of the internal details and operation of that module. Someone who uses that module as a component in a larger system, however, will treat the module as a "black box." That is, the user will be uninterested in the internal workings of the module and will be concerned only with the inputs, the outputs, and, most important, the functional relation of outputs to inputs (*i.e.*, the service provided). Thus, a black box is a module viewed in terms of its input-output description. It can be used with other black boxes to construct a more complex module, which again will be viewed at higher levels as a bigger black box.

This approach to design leads naturally to a hierarchy of modules in which a module appears as a black box at one layer of the hierarchy, but appears as a system of lower-layer black boxes at the next lower layer of the hierarchy (see Fig. 3). At the overall system level, (*i.e.*, at the highest layer of the hierarchy), one sees a small collection of top-layer modules, each viewed as black boxes providing some clear-cut service. At the next layer down, each top-layer module is viewed as a subsystem of lower-layer black boxes, and so forth, down to the lowest layer of the hierarchy. As shown in Fig. 3, each layer might contain not only black boxes made up of lower-layer modules but also simple modules that do not require division into yet simpler modules.

As an example of this hierarchical viewpoint, a computer system could be viewed as a set of processor modules, a set of memory modules, and a bus module. A processor module could, in turn, be viewed as a control unit. Similarly, the arithmetic unit could be broken into adders, accumulators, and so on.

In most cases, a user of a black box does not need to know the detailed response of outputs to inputs. For example, precisely when an output changes in response to an input is not important as long as the output has changed by the time it is to be used. Thus, modules (*i.e.*, black boxes) can be specified in terms of tolerances rather than exact descriptions. This leads to standardized modules, which leads, in turn, to the possibility of using many identical, previously designated (*i.e.*, off-the-shelf) modules in the same system. In addition, such standardized modules can easily be replace with new, functionally equivalent modules that are cheaper or more reliable.

All of these advantages of modularity (*i.e.*, simplicity of design; understandability; and standard, interchangeable, widely available modules) provide the motivation for

[1]Taken entirely from Section 1.3, pp. 17–18, from *Data Networks*, 2nd ed., by D. Bertsekas and R. Gallager, Prentice-Hall, Inc., 1992.

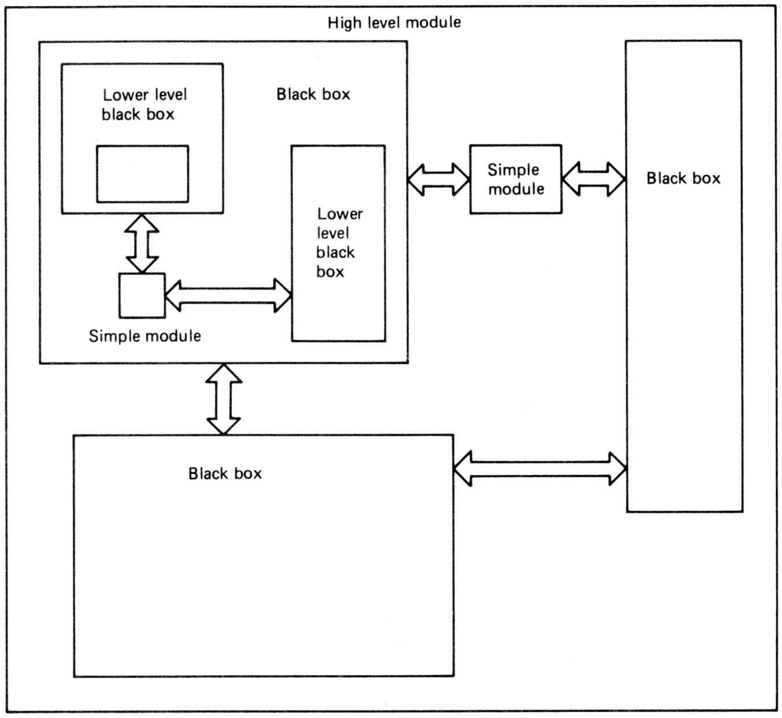


Figure 3: Hierarchy of nested black boxes. Each black box (except that at the lowest level) contains black boxes at a lower level, plus perhaps other modules.

a layered architecture in data networks. A layered architecture can be regarded as a hierarchy of nested modules or black boxes, as described above. Each given layer in the hierarchy regards the next lower layer as one or more black boxes which provide a specified service to the given higher level.

What is unusual about the layered architecture for data networks is that the black boxes at various layers are in fact distributed black boxes. The bottom layer of the hierarchy consists of the physical communication links, and at each higher layer, each black box consists of a lower-layer black box communication system plus a set of simple modules, one at each end of the lower-layer communication system. The simple modules associatied with a black box at a given layer are called *peer processes or peer modules* (see Fig. 4).

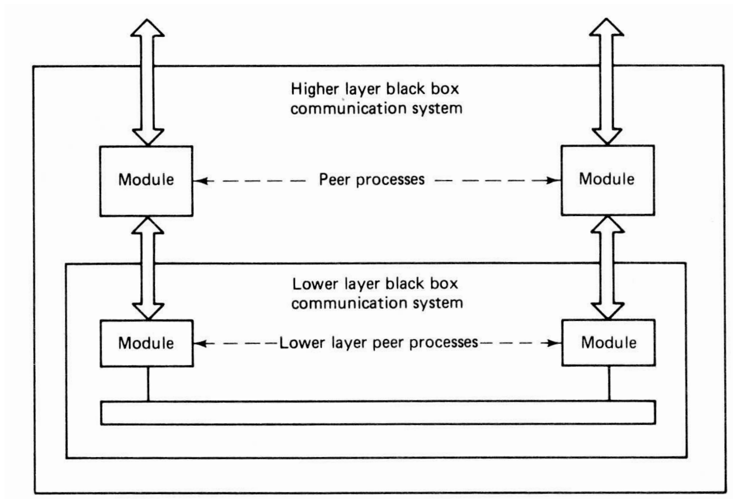


Figure 4: Peer processes with a black box communication system. The peer processes communicate through a lower-layer black box communication system that itself contains lower-layer peer processes.

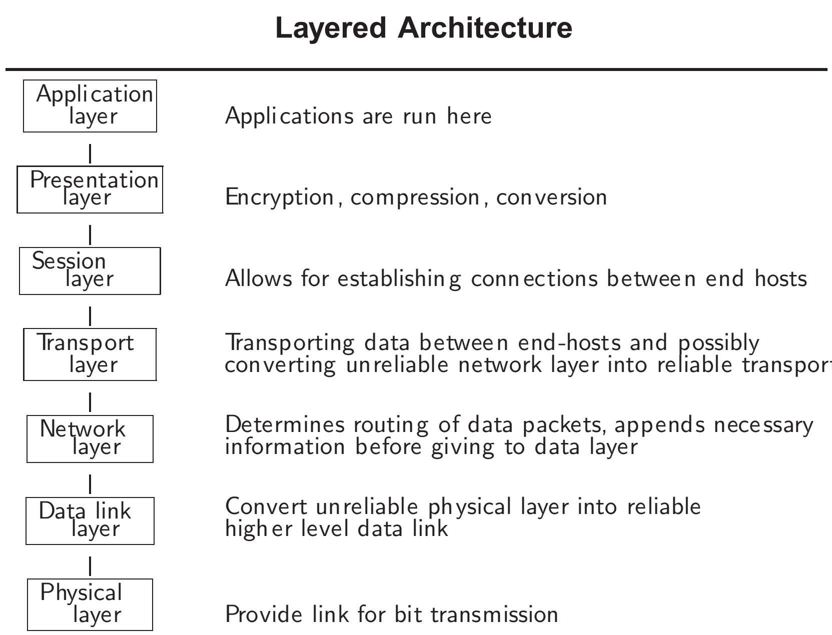
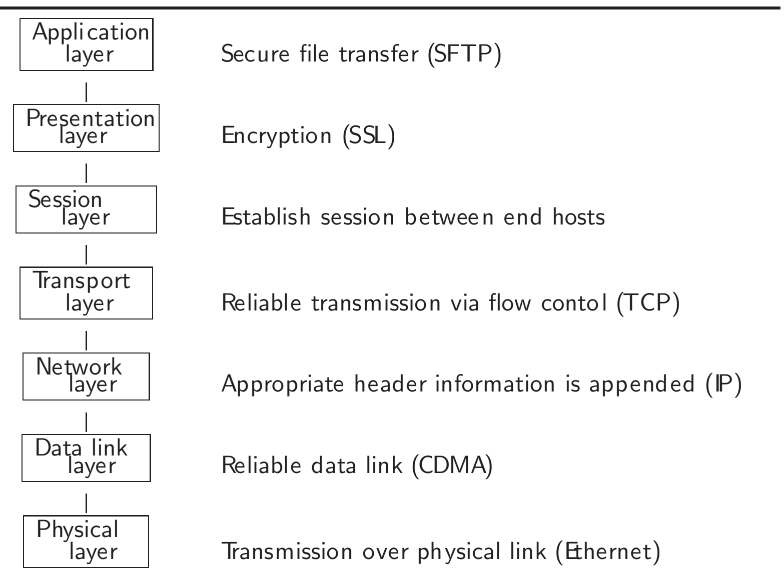


Figure 5: Seven Layer Network Architecture

Layered Architecture: An Example



References

[1] Keynes, John Maynard (1936) *The General Theory of Employment, Interest and Money* (Macmillan, London).

[2] Willems, J.C. (2013) "Open Stochastic Systems," *IEEE Trans. on Auto Control* Vol. 58, No. 2, pp. 406–421.

[3] Agarwal, M. and Mitter, S.K., "Architecture for Communication with a Fidelity Criterion," presented at the IFAC Congress, 2011.

[4] Willems, J.C. (2007) "The behavioral approach to open and interconnected systems," *Control Systems Magazine*, Vol. 27, pp. 46–99.

[5] Mitter, S.K. (May 1998) "The Legacy of George Zames," *IEEE Trans. on Automatic Control*, Vol. 43, No. 5.

[6] Basu, A. and Borkar, V.S., (2008) "Stochastic Control with Imperfect Models," *SIAM J. on Control and Optimization*, pp. 1274–1300.

[7] Kassakian, J.G. and Schmalenscu, R. (2011) "MIT Study on The Future of the Electric Grid."

[8] Roozbehani, M, Dahleh, M.A., and Mitter, S.K (2012) "Volatility of Power Grids under Real-Time Pricing," *IEEE Trans. on Power Systems*, Vol. 27, No. 4, pp. 1926–1939.

[9] Singh, R., Kumar, P.R., Xie, L. (2016) "Decentralized Control via Dynamic Stochastic Prices: The Independent System Operator Problem," to appear.

[10] Borkar, V.S., Konda, V.R. and Mitter, S.K. (2004) "On De Finetti Coherence and Kolmogorov Probability," *Stat. Prob. Lett.* Vol. 66, pp. 417–421.

[11] Föllmer, H. and Schied, A. (2013) "Probabilistic aspects of finance," *Bernoulli* Vol. 19, No, 4, pp. 1306–1326.

[12] Karatzas, I. and Shreve, S. (1998) *Brownian Motion and Stochastic Calculus* (Graduate Texts in Mathematics, Volume 113), 2nd edition (Springer).

[13] De Finetti, B., Machi, A., Smith, A. (1990) *Theory of Probability: A Critical Treatment* (John Wiley and Sons, New York).

[14] Karatzas, I. and Shreve, S. (1998) *Methods of Mathematical Finance* (Applications of Mathematics, Volume 39) (Springer).

[15] Jones, P. (June 2011) "Coherent Approximation of Distributed Expert Assessments," MIT PhD Thesis.

[16] Borkar, V.S. and Ghosh, M.K. (1990) "Controlled diffusions with constraints," *J. Math. Anal. Appl.*, 152, pp. 88–108.

[17] Lando, D. (2005) *Credit Risk Modeling* (Princeton University Press, Princeton, NJ).

[18] Willems, J.C. (1972) "Dissipative Dynamical Systems—Part I: General Theory," *Arch. Rational Mech. Anal.* **45**, pp. 321–351.

[19] Willems, J.C. and Takaba, K. (2007) "Dissipativity and Stability of Interconnections," *Intl. J. Robust Nonlin.* **17**, pp. 563–586.

[20] Borkar, V.S. and Mitter, S.K. (2003) "A Note on Stochastic Dissipativeness," Directions in Mathematical Systems Theory and Optimizations, Lecture Notes in Control and Information Science **286**, pp. 41–49.

[21] Rockafellar, R.T. (1970) *Convex Analysis* (Princeton University Press, Princeton, NJ).

[22] Fishburn, P.C. (1986), "The axioms of subjective probability with discussion," *Statistical Science* **1** (3), 335-358.

[23] Kyborg, H., Smokler, H. (1980) *Studies in Subjective Probability* (Robert E. Krieger Publ. Co., Huntington, NY).

[24] Parthasarathy, K.R. (1967), *Probability Measures on Metric Spaces* (Academic Press, New York-London).

[25] Diaconis, P., Zabell, S.L. (1982), "Updating subjective probability," *Journal of the American Statistical Association* **77**, 822-830.