1

Issues and Problems in Decision Making under Uncertainty

1.1 Introduction

The future cannot be predicted exactly, but one may learn from past observations. Past decisions can also improve future predictability. This is the context in which decisions are generally made. Herein, we discuss some mathematical issues pertaining to this topic.

1.1.1 Decision Making as Constrained Optimization Problems

Making decisions in a rational way is a problem which can be mathematically formulated as an *optimization* problem. Generally, several conflicting goals must be taken into account simultaneously. A choice must be made about which goals are formulated as constraints to be satisfied at a certain "level" (apart from constraints which are imposed by physical limitations), and which goals are reflected by (and aggregated within) a *cost function*.[1] Duality theory for *constrained optimization* problems provides a way to analyze, afterwards, the sensitivity of the best achievable cost as a function of constraint levels which were fixed a priori, and, possibly, to tune those levels to achieve a better trade-off between conflicting goals.

Problems that involve systems evolving in time enter the realm of *Optimal Control*. In a deterministic setting, Optimal Control has a long history dating back to the fifties with famous names such as Lev Pontryagin [124] and Richard Bellman [15]. The former, with his *Maximum Principle*, was more in the line of a *variational* approach of such problems, whereas the latter introduced the *Dynamic Programming* (DP) technique in connection with the state space approach.

[1] Throughout this book, without loss of generality, optimization problems are formulated as *minimization* problems, hence the objective function to be minimized is called a *cost*.

1.1.2 Facing Uncertainty

In general, when making decisions, one is faced with *uncertainties* which affect the cost function and, generally, the constraints. There are several possible attitudes associated with uncertainties, and consequently, several possible mathematical formulations of decision making problems under uncertainty. Let us mention two main possibilities.

Worst Case Design

The assumption here is that uncertainties lie in particular bounded subsets and, that one must consider the *worst situation* to be faced and try to make it as good as possible. In more mathematical terms, and considering the cost only for the time being (see hereafter for constraints), since one would like to minimize that cost, one must minimize the *maximal* possible value Nature can give to that cost by playing with uncertainties within the assumed bounded subsets. That is, a *min-max* (game like) problem is formulated and a *guaranteed* performance can be evaluated (as long as assumptions on uncertainties hold true).

The treatment of constraints in such an approach should normally follow the same lines of thought (one must fight against the worst possible uncertainty outcomes from the point of view of constraint satisfaction). Sometimes the terminology of *robust* decision making (or control) is used for approaches along those lines [16].

Stochastic Programming or Stochastic Control

Here, uncertainties are viewed as random variables following *a priori* probability laws. We shall call them "primitive" random variables as opposed to other "secondary" random variables involved in the problem and which are derived from the primitive ones by applying functions such as dynamic equations, feedback laws (see hereafter), etc. Then the cost to be minimized is the mathematical expectation of some performance index depending on those random variables and on decisions.

For this mathematical expectation to make sense, the decisions must also become random variables defined on the same underlying probability space. A trivial case is when those decisions are indeed *deterministic*: we shall call them *open-loop* decisions or"controls" later on. But they may also be true random variables because they are produced by applying functions to either primitive or secondary random variables. Here, we enter the domain of *feedback* or *closed-loop* control which plays a prominent part in decision making under uncertainty.

Let us now say a few words about constraint satisfaction. Constraints may be imposed as *almost sure* (a.s.) constraints. This is generally the case of equality or inequality constraints expressing physical laws or limitations. Other constraints may be formulated with mathematical expectations, although it is generally difficult to give a sound practical meaning to this approach. If a.s. requirements may sometimes be either unfeasible or not economically viable, one may appeal to "constraints in probability": the satisfaction of those constraints is required only "sufficiently often", that is, with a certain prescribed probability. We do not pursue this discussion here, as we mostly consider a.s. constraints in this book.

In the title of this section, we have used the words "Stochastic Programming" and "Stochastic Control". Stochastic Control, or rather Stochastic Optimal Control (SOC), is the extension of the theory of Deterministic Optimal Control to the situation when uncertainties are present and modeled by random variables, or stochastic processes since control theory mostly addresses dynamic problems. SOC problems were introduced not long after their deterministic counterparts, and the DP approach has been readily extended (under specific assumptions) to the stochastic framework. "Pontryagin like" or "variational" approaches appeared much later in the literature [25] and we shall come back to explanations for this fact. SOC is used to deal with *dynamic* problems. The notion of *feedback*, as naturally delivered by the DP approach, plays a central part in this area.

Stochastic Programming (SP), which can be traced back to such early contributors as George Dantzig [50], is the extension of Mathematical Programming to the stochastic framework. As such, the initial emphasis is on optimization, possibly in a *static* setting, and numerical resolution methods are based on variational techniques; randomness is generally addressed by appealing to the Monte Carlo technique which, roughly speaking, amounts to representing this uncertainty through the consideration of several "samples" or "scenarios". This is why, historically, the notions of *feedback* and *information* were less present in SP than they were in SOC.

However, the SP community[2] has progressively considered two-stage, and then multi-stage problems. Inevitably, the question of *information structures* popped up in the field, at least to handle the elementary constraint of *nonanticipativeness*: one should not assume that the exact realizations of random variables at and after stage $t+1$ are known when making decisions at stage t; only a probabilistic description of future occurrences can be taken into account.

It is therefore natural that the two communities of SOC and SP tend to merge and borrow ideas from each other. The concepts of information and feedback are more developed in the former, and the variational and Monte Carlo approaches are more widespread in the latter. Getting closer to each other for the two communities should perhaps begin with unifying the terminology: as far as we understand, *recourse* in the SP community is used as a substitute for *feedback*. This book is an attempt to close the gap. The com-

[2] The official web page of the SP community `http://www.stoprog.org/` offers links to several tutorials and examples of applications of SP.

parison between SOC and SP approaches is already addressed by Varaiya and Wets in this interesting paper [148].

1.1.3 The Role of Information in the Presence of Uncertainty

In Deterministic Optimal Control, as mentioned previously, there are two main approaches in connection with Pontryagin's and Bellman's contributions. The former focuses on open-loop controls, whereas the latter provides closed-loop solutions. By open-loop controls, we mean that the decisions are given as a function of *time* only, whereas closed-loop strategies compute the control to be implemented at each time instant as a function of both *time and observations*; the observations may be the state itself.

In fact, there are no discrepancies in the performance achieved by both approaches because, in a deterministic situation, everything is uniquely determined by the decision maker. Therefore, if closed-loop strategies are implemented, one can simulate the closed-loop dynamic system, record the trajectories of state, control and observations variables, substitute those trajectories in the control strategy, and compute an open-loop control history that would generate exactly the same trajectories.

The situation is quite different in an uncertain environment, since trajectories are not predictable in advance (off-line) because they depend on on-line realizations of random variables. Available observations reveal some information about those realizations, at least on *past* realizations (because of *causality*). By using this on-line information, one can do better than simply apply a *blind* open-loop control which has been determined only on the basis of a priori probability laws followed by the random "noises".

This means that the achievable performance is dependent on what we call the *information pattern* or *information structure* of the problem: a decision making problem under uncertainty is not well-posed until the exact amount of information available prior to making every decision has been defined. Open-loop problems are problems in which no actual realization can be observed, and thus, the optimal decisions solely depend on a priori probability laws. In dynamic situations, every decision may depend on certain on-line observations that must be specified. Of course, the optimal decisions also depend on a priori probability laws since, generally, not all random realizations can be observed prior to making decisions, if only because of causality or nonanticipativeness.

Because of these considerations, one must keep in mind that solving stochastic optimization problems, especially in dynamic situations when on-line observations are made available, is not just a matter of optimization, of dealing with conventional constraints, or even of computing or evaluating mathematical expectations (which is generally a difficult task by itself); it is also the question of properly handling specific constraints that we shall call *informational constraints*. Indeed, as this book illustrates, there are essentially two ways of dealing with such constraints. That used by the DP approach is a *functional* way: decisions are searched for as *functions* of observations (feedback laws). But another way, which is more adapted to variational approaches in stochastic optimization, may also be considered: all variables of the problem, including decisions, are considered as random variables or stochastic processes; then the dependency of decisions upon observations must go through notions of *measurability* as used by Measure Theory. We shall call this alternative approach an *algebraic* handling of informational constraints (this terminology stems from the fact that information may be mathematically captured by σ-algebras, also called σ- fields, another important notion introduced by Measure Theory). A difficult aspect of numerical resolution schemes is precisely the practical translation of those measurability or algebraic constraints into the numerical problem.

An even more difficult aspect of *dynamic* information patterns is that future information may be affected by past decisions. Such situations are called situations with *dual effect*, a terminology which tries to convey the idea that present decisions have two, very often conflicting, effects or objectives: directly contributing to optimizing the cost function on the one hand, modifying the informational constraints to which future decisions are subject, on the other. Problems with dual effect are generally among the most difficult decision making problems (see again [148] about this topic).

1.2 Problem Formulations and Information Structures

In this section, two formulations of stochastic optimization problems are proposed: they pertain to the two schools of SOC and SP alluded to above. The important issue of *information structures* is also discussed.

1.2.1 Stochastic Optimal Control (SOC)

General Formulation

We consider the following formulation of a stochastic optimal control (SOC) problem in discrete time: for every time instant t, $\boldsymbol{X}_t$ ("state"[3]), $\boldsymbol{U}_t$ (control) and $\boldsymbol{W}_t$ (noise) are all random variables over a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. They are related to each other by the *dynamics*

$$\boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) \tag{1.1a}$$

which is satisfied $\mathbb{P}$-almost surely for $t = 0, \ldots, T-1$. Here, to keep things simple, T, the *time horizon*, should be a given deterministic integer value, but it may be a random variable in more general formulations. The variable $\boldsymbol{X}_0$ is a given random variable. It is convenient to view $\boldsymbol{X}_0$ as a given function of some other random variable called $\boldsymbol{W}_0$, in such a way that all primitive random

[3] Those quotes around the word *state* become clearer when discussing the *Markovian case* by the end of this subsection.

variables are denoted $\boldsymbol{W}_s$, $s = 0, \ldots, T$, whereas $\boldsymbol{W}$ denotes the corresponding stochastic process $\{\boldsymbol{W}_s\}_{s=0,\ldots,T}$. The purpose is to minimize a cost function

$$\mathbb{E}\Big(\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + K(\boldsymbol{X}_T) \Big) \tag{1.1b}$$

in which K is the *final* cost whereas L_t is called the *instantaneous* cost. The symbol $\mathbb{E}(\cdot)$ denotes *expectation* w.r.t. $\mathbb{P}$ (assuming of course that the functions involved are measurable and integrable). The minimization is achieved by choosing the control variable $\boldsymbol{U}_t$ at each time instant t, but as previously mentioned, this is done after some *on-line* information has been collected (in addition to the *off-line* information composed of the model — dynamics and cost — and the a priori distribution of $\{\boldsymbol{W}_s\}_{s=0,\ldots,T}$). This on-line information is supposed to be at least *causal* or *nonanticipative*, that is, the largest possible amount of information available at time instant t is equivalent to the observation of the realizations of the random variables $\boldsymbol{W}_s$ for $s = 0, \ldots, t$ (but not beyond t). In the language of Probability Theory, this amounts to saying that $\boldsymbol{U}_t$, as a random variable, is *measurable* w.r.t. the σ-field generated by $\{\boldsymbol{W}_s\}_{s=0,\ldots,t}$ which is denoted $\mathcal{F}_t$:

$$\mathcal{F}_t = \sigma\big(\{\boldsymbol{W}_s\}_{s=0,\ldots,t}\big) \tag{1.1c}$$

(the reader may refer to Appendix B for all those standard notions). Of course, this σ-field increases as time passes, that is, $\mathcal{F}_t \subset \mathcal{F}_{t+1}$: it is then called a *filtration*.

Remark 1.1. Observe that in the right-hand side of (1.1a), $\boldsymbol{U}_t$ must be chosen *before* $\boldsymbol{W}_{t+1}$ is observed: this is called the *decision-hazard* framework, as opposed to the *hazard-decision* framework in which the decision maker plays after "nature" at each time stage. This is why we put $\boldsymbol{W}_{t+1}$ rather than $\boldsymbol{W}_t$ in the right-hand side of (1.1a). ♢

It may be that $\boldsymbol{U}_t$ is constrained to be measurable w.r.t. some σ-field $\mathcal{G}_t$ smaller than $\mathcal{F}_t$:

$$\boldsymbol{U}_t \text{ is } \mathcal{G}_t\text{-measurable, } \mathcal{G}_t \subset \mathcal{F}_t, \quad t = 0, \ldots, T-1\,. \tag{1.1d}$$

Unlike $\mathcal{F}_t$, the σ-field $\mathcal{G}_t$ is not necessarily increasing with t (see hereafter).

Information Structure

Very often, $\mathcal{G}_t$ itself is a σ-field generated by some random variable $\boldsymbol{Y}_t$ called *observation*. Actually, $\boldsymbol{Y}_t$ should be considered as the collection of *all* observations available at t. That is, if $\boldsymbol{Z}_t$ denotes a new observation made available at t, but if the decision maker has *perfect memory* of all observations made so far, then $\boldsymbol{Y}_t = \{\boldsymbol{Z}_s\}_{s=0,\ldots,t}$. In this case, as for $\mathcal{F}_t$, the σ-field $\mathcal{G}_t$ is increasing with t, but this is not necessarily always true.

The σ-fields $\mathcal{F}_t$, generated by $\{\boldsymbol{W}_s\}_{s=0,\ldots,t}$, are of course only dependent upon the data of the problem, and this is also the case of the $\mathcal{G}_t$ if the observations $\boldsymbol{Y}_t$ are solely dependent on the primitive random variables $\boldsymbol{W}_s$. But if the observations depend also on the controls $\boldsymbol{U}_s$ (for example, if $\boldsymbol{Z}_t$ is a function of the "state" $\boldsymbol{X}_t$, possibly a function corrupted by noise), it is likely that the σ-field $\mathcal{G}_t$ depends on controls too, and therefore, the measurability constraint (1.1d) is an implicit constraint in that control is subject to constraints depending on controls! Fortunately, thanks to causality, this implicit character is only apparent, that is, the constraint on $\boldsymbol{U}_t$ depends on controls $\boldsymbol{U}_s$ with s strictly less than t.

Nevertheless, this is generally a source of huge complexity in SOC problems which is known under the name of the *dual effect* of control. This terminology tries to convey the fact that when making decisions at every time instant s, the decision maker has to take care of the following double effect: on the one hand, his decision affects cost (directly, at the same time instant, and in the future time instants, through the "state" variables); but, on the other hand, it makes the next decisions $\boldsymbol{U}_t$, $t > s$ more or less constrained through (1.1d).

Example 1.2. Let us give an example of this double or dual effect in the real life: the decision of investing in research in any industrial activity. On the one hand, investing in research costs money. On the other hand, an improved knowledge of the field of activity may help save money in the future by allowing better decisions to be made. This example shows that this future effect is very often contradictory with immediate cost considerations and thus the matter of a trade-off to be achieved. △

We now return to our general discussion of information structure in SOC problems. Even if the observations $\boldsymbol{Y}_t$ depend on past controls, it may happen than the σ-fields $\mathcal{G}_t$ they generate *do not* depend on those controls. This tricky phenomenon is discussed in Chapter 10. Apart from this rather exceptional situation, there are other circumstances when things turn out to be less complex than it may have seemed a priori.

The most classical such case is the *Markovian case*. Suppose the stochastic process $\boldsymbol{W}$ is a "white noise", that is, the random variables $\{\boldsymbol{W}_s\}_{s=0,\ldots,T}$, are all mutually independent. Then, $\boldsymbol{X}_t$ truly deserves the name of the *state* variable at time t (this is why, until now, we put the word "state" between quotes — see Footnote 3). Indeed, because of this assumption of white noise, the past realizations of the noise process $\boldsymbol{W}$ provide no additional information about the likelihood of future realizations. Hence, remembering $\boldsymbol{X}_t$ is sufficient information to keep to predict the future evolution of the system after t. That is, $\boldsymbol{X}_t$ "summarizes" the past and additional observations are therefore useless. The *Markovian case* is defined as the situation when $\boldsymbol{W}$ is a white noise stochastic process and $\mathcal{G}_t$ is generated at each time t by the variable $\boldsymbol{X}_t$. Otherwise stated, the available observation $\boldsymbol{Y}_t$ at time t is simply $\boldsymbol{X}_t$. This is a *perfect (noiseless) and full size* observation of the state vector. If the observation is *partial* (a non injective function of $\boldsymbol{X}_t$) and/or a *noisy* such function, then the Markovian situation is broken.

In the Markovian case, $\mathcal{G}_t$ *does depend*, in general, upon past controls $\boldsymbol{U}_s, s < t$, but we *would not do better* with $\mathcal{F}_t$ replacing $\mathcal{G}_t$. This is why the Markovian case, although potentially falling into the most difficult category of problems with a dual effect, is not so complex as more general problems in this category. The Markovian feature is exploited by the Dynamic Programming (DP) approach (see §4.4) which is conceptually simple, but quickly becomes numerically difficult, and, indeed, impossible when the dimension of the state vector $\boldsymbol{X}_t$ becomes large.

1.2.2 Stochastic Programming (SP)

Formulation

Here we consider another formulation of stochastic optimization problems which ignores "intermediate" variables (such as the "state" $\boldsymbol{X}$ in the previous SOC formulation) and which concentrates on the essential items, namely, the

control or **decision** $\boldsymbol{U}$: a random variable over a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in a measurable space $(\mathbb{U}, \mathcal{U})$;
noise $\boldsymbol{W}$: another random variable with values in a measurable space $(\mathbb{W}, \mathcal{W})$;
cost function : a measurable mapping $j : \mathbb{U} \times \mathbb{W} \to \mathbb{R}$;
σ**-fields** : $\mathcal{F}$ denotes the σ-field generated by $\boldsymbol{W}$ whereas $\mathcal{G}$ denotes the one w.r.t. which $\boldsymbol{U}$ is constrained to be measurable; generally, $\mathcal{G}$ is generated by an
observation $\boldsymbol{Y}$: another random variable with values in a measurable space $(\mathbb{Y}, \mathcal{Y})$; in this case, we use the notation

$$\boldsymbol{U} \preceq \boldsymbol{Y} \tag{1.2}$$

to mean that $\boldsymbol{U}$ is measurable w.r.t. (the σ-field generated by) $\boldsymbol{Y}$. As we see in Chapter 3, this relation between random variables corresponds to an order relation. We also use this notation in constraints as $\boldsymbol{U} \preceq \mathcal{G}$ to mean that the random variable $\boldsymbol{U}$ is measurable w.r.t. the σ-field $\mathcal{G}$.

With these ingredients at hand, the problem under consideration is set as follows:

$$\min_{\boldsymbol{U} \preceq \mathcal{G}} \mathbb{E}\big(j(\boldsymbol{U}, \boldsymbol{W})\big) \quad \text{or} \quad \min_{\boldsymbol{U} \preceq \boldsymbol{Y}} \mathbb{E}\big(j(\boldsymbol{U}, \boldsymbol{W})\big) \ . \tag{1.3}$$

Without going into detailed technical assumptions, we assume that expectations do exist, and that infima are reached (hence the use of the min symbol).

Typology of Information Structures

According to the nature of $\mathcal{G}$ or $\boldsymbol{Y}$, we distinguish the following three cases.

Open-loop optimization: this is the case when $\mathcal{G}$ is the trivial σ-field $\{\emptyset, \Omega\}$, or equivalently, $\boldsymbol{Y}$ is any deterministic variable (that is, a constant map over Ω). In this case, an optimal decision is based solely on the a priori (off-line) knowledge of the model, and not on any on-line observation. Therefore, the decision itself is a deterministic variable $u \in \mathbb{U}$ which must minimize a cost function $J(u)$ defined as an expectation of $j(u, \boldsymbol{W})$. The numerical resolution of such problems is considered in Chapter 2.

Static Information Structure (SIS): this is the case when $\mathcal{G}$ or $\boldsymbol{Y}$ are non trivial but *fixed*, that is, a priori given, independently of the decision $\boldsymbol{U}$. The terminology "static" does not imply that no dynamics such as (1.1a) are involved in the problem formulation. It just expresses that the σ-field $\mathcal{G}$ constraining the decision is a priori given at the problem formulation stage. If time t is involved, one must rewrite the measurability constraint as prescribed at each time stage t as "$\boldsymbol{U}_t$ is $\mathcal{G}_t$-measurable" as in (1.1d), and this *does* leave room for information made available on-line as time evolves. "Static" just says that this on-line information cannot be manipulated by past controls.

Remark 1.3. When the collection $\{\boldsymbol{U}_s\}_{s=0,\dots,T-1}$ of random variables is interpreted as a random vector over the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, then its measurability is characterized by the σ-field $\sigma(\{\boldsymbol{U}_s\}_{s=0,\dots,T-1})$ on $(\Omega, \mathcal{A})$. However, with this interpretation, the collection of constraints (1.1d) cannot in general be reduced to a single "vector" constraint $\boldsymbol{U} \preceq \mathcal{G}$ where $\boldsymbol{U}$ would be the "vector" $\{\boldsymbol{U}_s\}_{s=0,\dots,T-1}$ and $\mathcal{G}$ a σ-field on $(\Omega, \mathcal{A})$, like $\sigma(\{\boldsymbol{U}_s\}_{s=0,\dots,T-1})$ is. For example, over a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with $T = 2$, $\mathcal{G}_0 = \{\emptyset, \Omega\}$ and $\mathcal{G}_1 = \mathcal{A}$, consider a random variable $\boldsymbol{U}_1$ such that $\sigma(\boldsymbol{U}_1) = \mathcal{A}$. Writing $\boldsymbol{U} \preceq \mathcal{G}$ implies that $\mathcal{G}$ would be the σ-field $\mathcal{A}$, which does not translate that $\boldsymbol{U}_0$ must be a *constant* (deterministic) variable as implied by $\boldsymbol{U}_0 \preceq \mathcal{G}_0$. ♢

Remark 1.4. If $\mathcal{G}$ is generated by an observation $\boldsymbol{Y}$, either $\boldsymbol{Y}$ does not depend on $\boldsymbol{U}$, or the σ-field it generates is fixed despite $\boldsymbol{Y}$ does depend on $\boldsymbol{U}$ (as already mentioned, this may also happen in some special situations addressed in Chapter 10). One may also wonder whether $\boldsymbol{Y}$ has any relation with $\boldsymbol{W}$, for example, whether $\boldsymbol{Y}$ is given as a function $h(\boldsymbol{W})$, in which case $\mathcal{G}$ would be a sub-σ-field of $\mathcal{F}$, the σ-field generated by $\boldsymbol{W}$. For example, in the SOC problem (1.1), $\boldsymbol{Y}_t$ may be the complete or partial observation of past noises $\boldsymbol{W}_s, s = 0, \dots, t$, so that $\mathcal{G}_t \subseteq \mathcal{F}_t \subset \mathcal{F}_T$. Nevertheless, the fact that $\boldsymbol{Y}$ does or does not have a connection with $\boldsymbol{W}$ is not fundamental. Indeed, by manipulating notation, one can consider that this connection does exist. As a matter of fact, one can redefine the noise variable as the couple $\boldsymbol{W}' = (\boldsymbol{W}, \boldsymbol{Y})$ so that $\boldsymbol{Y}$ *is* a function of $\boldsymbol{W}'$. That the cost function j does not depend on the "full" $\boldsymbol{W}'$ does not matter. ♢

Dynamic Information Structure (DIS): this is the situation when $\mathcal{G}$ or $\boldsymbol{Y}$ depends on $\boldsymbol{U}$, which yields a seemingly implicit measurability constraint. Actually, it is difficult to imagine such problems without explicitly introducing several stages at which decisions must be taken based on observations which may depend on decisions at other stages.

Those stages may be a priori ordered, and the order may be a total order. This is the case of SOC problems (1.1); but other examples are considered hereafter in which those stages are not directly interpreted as "time instants" but rather as "agents" acting one after the other. As soon as such a total order of stages can be defined a priori, the notion of *causality* (who is "upstream" and who is "downstream") is natural and helps untangling the implicit character of the measurability constraint. Nevertheless, the difficulty of such problems with DIS still remains sometimes tremendous as it is shown with help of an example in §1.3.3.

More general problems may arise in which the order of stages or agent actions is only partial, and the situation may be even worse if this order itself depend on outcomes of the decisions and/or of hazard. At least in the case of a fixed but partial order, it turns out that two notions are paramount for the level of difficulty of the problem resolution:

- Who influences the available observations of whom?
- Who knows more than whom?

We shall not pursue the discussion of this difficult topic here. It is more thoroughly examined in Chapter 9. The forthcoming examples help us scratch the surface.

1.3 Examples

This section introduces a few simple examples in order to illustrate the impact of information structures on the formulation of stochastic optimization problems. The stress is more on this aspect than on being fussy about mathematical details (in particular, we assume that all expectations make sense without going into more precise assumptions).

1.3.1 A Basic Example in Static Information

Consider two given scalar random variables, $\boldsymbol{W}$ and $\boldsymbol{Y}$, plus the decision $\boldsymbol{U}$, and finally the following problem of type (1.3):

$$\min_{\boldsymbol{U} \preceq \boldsymbol{Y}} \mathbb{E}\big((\boldsymbol{W} - \boldsymbol{U})^2\big) . \tag{1.4}$$

It is well known that the solution of this problem, which consists in finding the best approximation of $\boldsymbol{W}$ which is $\boldsymbol{Y}$-measurable (that is, the projection of $\boldsymbol{W}$ onto the subspace of $\boldsymbol{Y}$-measurable random variables), is given by

$U^\sharp = \mathbb{E}(W \mid Y)$, that is, the conditional expectation of W knowing Y (see §3.5.3 and Definition B.5).

Generally speaking, as we see it later on in §3.5.2 and §8.3.5, Problem (1.3) can be reformulated as follows:

$$\mathbb{E}\Big(\min_{u \in \mathbb{U}} \mathbb{E}\big(j(u, W) \mid Y\big)\Big) . \tag{1.5}$$

In this form, since the conditional expectation subject to minimization is indeed a Y-measurable random variable, it should be understood that the minimization operates parametrically for every realization driven by ω and this yields an arg min also parametrized by ω, that is, in fact, a random variable which is also Y-measurable. When using this new formulation for Problem (1.4), the solution is readily derived (Hint: expand the square in the cost function and observe that Y-measurable random variables "get out" of the inner conditional expectation).

1.3.2 The Communication Channel

Description of the Problem

This is the story of two agents trying to communicate through a noisy channel. This story is depicted in Figure 1.1. The first agent (called the "encoder")

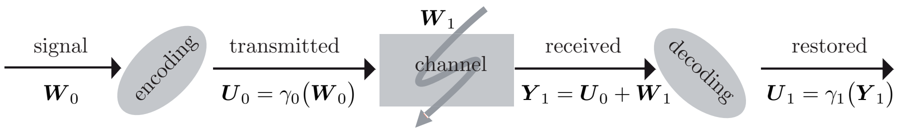

Fig. 1.1. Communication through a noisy channel

gets a "message", here simply a random variable W_0 supposed to be centered ($\mathbb{E}(W_0) = 0$), and he wants to communicate it to the other agent. We may consider that the encoder's observation Y_0 is precisely this W_0. He knows that the channel adds a noise, say a centered random variable W_1, to the message he sends, and so he must choose which "best" message to send. He has to "encode" the original signal Y_0 into another variable U_0 (what he decides to send through the channel), but the other agent (the "decoder") receives a noisy message $U_0 + W_1$. Finally, the decoder has to make his decision U_1 about what was the original message W_0, based on his observation, namely $Y_1 = U_0 + W_1$, the message he received. That is, he has to "decode", in an "optimal" manner, the signal Y_1 which is his observation.

This game is *cooperative* in that the encoder and the decoder try to help each other so as to reduce the error of communication as much as possible (a problem in "team theory" [104], which deals with decision problems involving several agents or decision makers with a common objective function but possibly different observations). Mathematically, this can be expressed by saying that they seek to minimize the expected square error $\mathbb{E}\big((\boldsymbol{U}_1 - \boldsymbol{W}_0)^2\big)$. However, without any other limitation or penalty, such a problem turns out to be rather trivial. For example, if the encoder sends an amplified signal $\boldsymbol{U}_0 = k\boldsymbol{Y}_0$ where k is an arbitrarily large constant, then the noise $\boldsymbol{W}_1$ added by the channel is negligible in front of this very large signal, and the decoder can then decode it by dividing it by the same constant k. For the game to be interesting and realistic, one must put a penalty on the "power" $\mathbb{E}(\boldsymbol{U}_0^2)$ sent over the channel, either with help of a constraint limiting this power to a maximum level, or by introducing an additional term proportional to this power into the cost. To stay closer to the generic formulation (1.3), we choose the latter option. Finally, the problem under consideration is the following:

$$\min_{\boldsymbol{U}_0,\boldsymbol{U}_1} \mathbb{E}\big(\alpha \boldsymbol{U}_0^2 + (\boldsymbol{U}_1 - \boldsymbol{W}_0)^2\big) \tag{1.6a}$$

$$\text{s.t.} \quad \boldsymbol{U}_0 \preceq \boldsymbol{Y}_0\,, \quad \boldsymbol{U}_1 \preceq \boldsymbol{Y}_1\,. \tag{1.6b}$$

The positive parameter α is the unit cost for the power transmitted over the channel. The measurability constraints (1.6b) reflect what each agent knows before making his decision.

Discussion

There are a few remarks to make at this point:

- there is no time index t explicitly involved in this formulation, but still there is a natural order of the agents: the encoder acts first in that his action has an influence on what the decoder observes;
- there is no inclusion (in either direction) between the information available to the encoder and to the decoder although, as just highlighted, the decoder is "downstream" the encoder; if we interpret agents as time stages, it means that, at the second time stage, not all the information available at the first time stage has been retained, a fact referred to as "no perfect memory".

The fact that the encoder can influence what the decoder observes, whereas the decoder does not know as much as the encoder knows, is a source of tremendous difficulties. We are actually here in the heart of what we called "dual effect" earlier: the encoder, when making his decision, should not spend too much money according to the cost function (in particular, he should limit the power send over the channel) but, at the same time, he should be aware of the fact that his encoding impacts the information revealed to the decoder. To make this consideration more concrete, we discuss it further in a simplified setting in the next paragraph.

At this stage, let us say what is known about the resolution of Problem (1.6) [154, 84].

- The exact solution is yet unknown in the general case (see hereafter).
- There are particular cases when the solution is known, namely when the dimension of the message to be transmitted is exactly the same as the dimension of the encoded message, that is, when $\dim \mathbb{W}_0 = \dim \mathbb{U}_0$ (with certain additional assumptions, in particular Gaussian noises). Then, the encoder simply sends the original message ($\boldsymbol{U}_0 = \boldsymbol{W}_0$) and the decoder computes the conditional expectation $\boldsymbol{U}_1 = \mathbb{E}(\boldsymbol{W}_0 \mid \boldsymbol{Y}_1)$, which is a linear function of the observation $\boldsymbol{Y}_1$ when assuming that all primitive random variables are Gaussian. But what is important to notice is that the solution is proved to be optimal not because it satisfies some optimality condition (that, at present, nobody knows how to write), but because it achieves the lower bound of the expected square error provided by the Information Theory of Claude Shannon [11].
- When $\dim \mathbb{W}_0 < \dim \mathbb{U}_0$ (redundancy in coding) or $\dim \mathbb{W}_0 > \dim \mathbb{U}_0$ (compression in coding), the exact solution is not known yet, but it is known to be a *nonlinear* function of observations. Indeed, on the one hand, the best linear feedback strategy satisfying (1.6b) can easily be obtained, and, on the other hand, clever nonlinear feedback strategies have been proposed which outperform the best linear strategy (although they are not claimed to be optimal). This appearance of nonlinear strategies in a Linear-Quadratic-Gaussian (LQG) stochastic optimization problem is an illustration of what is known under the name of *signaling*: by using tricky nonlinear strategies, the encoder tries to provide to the decoder as much information about his observation as possible (here, the message to communicate) at the cheapest cost, using the system "dynamics" itself as the medium of this information transmission. Note that these signaling strategies would be impossible if the encoder could not influence the decoder's observation. In addition, it would be useless if the decoder knew at least as much information as the encoder knows (this would be the case of "perfect memory" in SOC problem (1.1)).

How Signaling Works?

We try to give the feeling of how signaling works, assuming that the encoder uses only linear strategies. Thus, let $\boldsymbol{U}_0 = k\boldsymbol{W}_0$. Of course, the decoder knows k because the strategy is elaborated (off line) jointly by the two decision makers. On line, the decoder observes the value of $\boldsymbol{Y}_1 = k\boldsymbol{W}_0 + \boldsymbol{W}_1$ from which he must guess the value realized by $\boldsymbol{W}_0$.

The primitive random variables of the problem are the couple $(\boldsymbol{W}_0, \boldsymbol{W}_1)$. For the purpose of graphical representation, we assume that this couple lies in the square $[0,1] \times [0,1]$. Figure 1.2 represents this square and the parallel lines corresponding to equations $w_1 = -kw_0 + y_1$ (with slope $-k$ and value-at-zero y_1). Therefore, after the realized value of $\boldsymbol{Y}_1$ has been observed, the decoder knows on which particular line the true realization of the noises is located. Given that his purpose is to determine the realization of $\boldsymbol{W}_0$, it is graphically

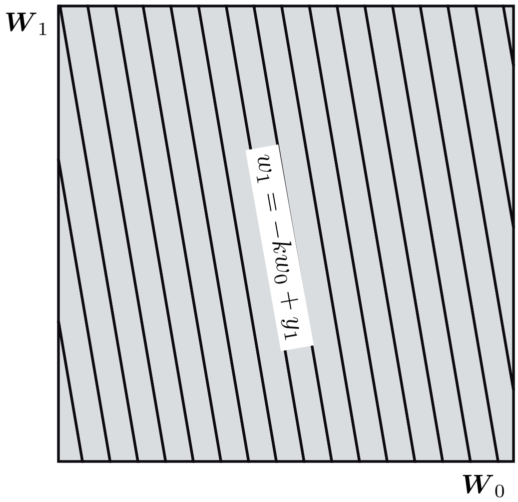

Fig. 1.2. Partition generated by $\boldsymbol{Y}_1$

intuitive that the uncertainty about this value decreases as $|k|$ (that is, the slope, be it negative or positive) increases. In terms of Communication Theory, this means that the ratio signal/noise improves as $|k|$ increases. This shows how the encoder can make the problem of the decoder more or less tractable by choosing his own strategy. But remember that large values of $|k|$, and hence of $\mathbb{E}\big((\boldsymbol{U}_0)^2\big)$, cause a large cost (see (1.6a)).

1.3.3 Witsenhausen's Celebrated Counterexample

The following problem was proposed by Hans Witsenhausen in 1968 [155] as evidence that LQG problems may lead to nonlinear feedback solutions whenever the information structure is not "classical" (say, here, when it does not reflect *perfect memory*). This information feature is similar to that of the previous problem (§1.3.2) and several other features are similar (linear dynamics, dimensions, etc.). The main difference lies in the fact that Witsenhausen's problem belongs to the SOC class (1.1); therefore its cost function is additive in time as (1.1b), whereas (1.6a) is not so because of the cross-product $\boldsymbol{U}_1\boldsymbol{W}_0$. The statement of this problem is as follows:

$$\min_{\boldsymbol{U}_0,\boldsymbol{U}_1} \mathbb{E}\big(k^2\boldsymbol{U}_0^2 + \boldsymbol{X}_2^2\big) \tag{1.7a}$$

$$\text{s.t.} \quad \boldsymbol{U}_0 \preceq \boldsymbol{Y}_0 , \qquad \boldsymbol{U}_1 \preceq \boldsymbol{Y}_1 , \tag{1.7b}$$

$$\boldsymbol{X}_1 = \boldsymbol{X}_0 + \boldsymbol{U}_0 , \tag{1.7c}$$

$$\boldsymbol{X}_2 = \boldsymbol{X}_1 - \boldsymbol{U}_1 , \tag{1.7d}$$

$$\boldsymbol{Y}_0 = \boldsymbol{X}_0 , \tag{1.7e}$$

$$\boldsymbol{Y}_1 = \boldsymbol{X}_1 + \boldsymbol{W} . \tag{1.7f}$$

We have kept Witsenhausen's original notation, but to enhance the parallelism with the previous problem, we could have changed k^2 into α and $\boldsymbol{X}_0$ (resp. $\boldsymbol{W}$) into $\boldsymbol{W}_0$ (resp. $\boldsymbol{W}_1$).

This problem is discussed at length later on in this book (see §4.2), so we just mention it here as another celebrated, yet simple, example of all the difficulties encountered when the assumption of *perfect memory* is dropped (here again, the observation $\boldsymbol{Y}_1$ is not "richer" than $\boldsymbol{Y}_0$). Bansal and Basar [11] discuss the fact that Problem (1.6) (sometimes) admits linear feedback solutions whereas Problem (1.7) has a nonlinear solution. See also a review of this problem by Y. C. Ho [80] and references therein.

1.4 Discretization Issues

So far, several formulations of stochastic optimization problems have been considered, and the role and importance of their information structure have been discussed. Those problems involve random variables and measurability or informational constraints, and they are *infinite-dimensional* problems for which closed-form solutions are scarcely obtainable. Therefore, a numerical resolution goes through some discretization process to make them amenable to a finite-dimensional approximation. However, due to the particular nature of informational constraints, this discretization process requires special care.

1.4.1 Problems with Static Information Structure (SIS)

Most problems with DIS are presently out of reach from the numerical point of view, sometimes even at the early stage of writing down optimality conditions. An exception is provided by problems which are amenable to a Markovian formulation with a very moderate state space dimension. This book mainly concentrates on problems with SIS (nevertheless, problems with no dual effect are also in principle amenable to a SIS formulation).

Accordingly, we may consider problems under the SOC formulation (1.1) or under the more compact SP formulation (1.3).

The subclass of open-loop problems are simpler in that their solution is deterministic (the solution is an element of the control space $\mathbb{U}$ and not an application from Ω to $\mathbb{U}$). However, the cost function involves computing an expectation, a task that cannot generally be achieved analytically. Thus, one must appeal to some sort of *Monte Carlo sampling* one way or another. Chapter 2 considers different ways of exploiting this idea and combining it with numerical optimization itself.

The more general SP or SOC problems with SIS involve the same issue of computing expectations, if not even *conditional expectations*, but their solution, unlike in open-loop problems, are random variables. In addition, this solution is subject to informational or measurability constraints. Such constraints must be reflected, one way or another, in a discretized version of the problem, since, in general, some discretization technique must be used to come up with a numerical problem that can be solved with a computer. It turns out that this twofold aspect of discretization, namely,

- Monte Carlo like sampling for estimating expectations or conditional expectations;
- finite dimensional representation of random variables with mutual measurability constraints;

is a rather subtle issue that must be handled very carefully for, otherwise, a completely irrelevant discrete problem may result. An example is given hereafter.

As already mentioned at the end of §1.1.3, there are two different ways of translating informational constraints: one called *functional* (essentially, some random variables are represented as *functions* of other random variables), and the other one called *algebraic* (some random variables must be *measurable* with respect to other random variables). This translates into different numerical requirements, but in any case the interaction of the informational constraint representation with the Monte Carlo sampling in order to come up with a meaningful discrete problem is a tricky point as illustrated now by an example.

1.4.2 Working out an Example

The Problem

Consider two independent random variables $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$, each with a uniform probability distribution over $[-1, 1]$ (zero mean, variance $1/3$). The unique decision variable $\boldsymbol{U}$ may only use the observation of $\boldsymbol{W}_0$ (which we view as the initial state $\boldsymbol{X}_0$). The final state $\boldsymbol{X}_1$ is equal to $\boldsymbol{W}_0 + \boldsymbol{U} + \boldsymbol{W}_1$. The goal is to minimize $\mathbb{E}(\varepsilon \boldsymbol{U}^2 + \boldsymbol{X}_1^2)$, where ε is a given "small" positive number ("cheap control"). The statement is thus

$$\min_{\boldsymbol{U} \preceq \boldsymbol{W}_0} \mathbb{E}\big(\varepsilon \boldsymbol{U}^2 + (\boldsymbol{W}_0 + \boldsymbol{U} + \boldsymbol{W}_1)^2\big) . \tag{1.8}$$

Exact Solution

We have that

$$\mathbb{E}\big(\varepsilon \boldsymbol{U}^2 + (\boldsymbol{W}_0 + \boldsymbol{U} + \boldsymbol{W}_1)^2\big) = \mathbb{E}\big(\boldsymbol{W}_0^2 + \boldsymbol{W}_1^2 + (1+\varepsilon)\boldsymbol{U}^2 + 2\boldsymbol{U}\boldsymbol{W}_0 + 2\boldsymbol{U}\boldsymbol{W}_1 + 2\boldsymbol{W}_0\boldsymbol{W}_1\big) .$$

The last two terms on the right-hand side yield zero in expectation since $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$ are centered independent random variables and since $\boldsymbol{U}$ is measurable with respect to $\boldsymbol{W}_0$. The first two terms yield twice the variance $1/3$ of the noises. Therefore, we remain with the problem of minimizing

$$\frac{2}{3} + \mathbb{E}\big((1+\varepsilon)\boldsymbol{U}^2 + 2\boldsymbol{U}\boldsymbol{W}_0\big) \tag{1.9}$$

by choosing $\boldsymbol{U}$ as a measurable function of $\boldsymbol{W}_0$. Using (1.5), one can prove that the solution is given by the feedback rule

$$\boldsymbol{U} = -\frac{\boldsymbol{W}_0}{1+\varepsilon} ,$$

and the corresponding optimal cost is readily calculated to be

$$\frac{1}{3}\,\frac{1+2\varepsilon}{1+\varepsilon} \approx \frac{1}{3} . \tag{1.10}$$

Monte Carlo Discretization

We now proceed to some discretization of this problem. To that purpose, we first consider N noise trajectories $(w_0^i, w_1^i), i = 1, \ldots, N$, which are N sample realizations of a two-dimensional vector $(\boldsymbol{W}_0, \boldsymbol{W}_1)$ with uniform probability distribution over $[-1, 1]^2$. Those samples serve to approximate the cost expectation by a usual Monte Carlo averaging.[4]

However, in this process, we must also consider N corresponding realizations $\{u^i\}_{i=1,\ldots,N}$ of the random decision variable $\boldsymbol{U}$. But, we must keep in mind that this random variable should be measurable with respect to the first component $\boldsymbol{W}_0$ of the previous vector.

To that purpose, we impose the constraint

$$\forall i, j,\;\; u^i = u^j \quad \text{whenever} \quad w_0^i = w_0^j , \tag{1.11}$$

which prevents $\boldsymbol{U}$ from taking different values whenever $\boldsymbol{W}_0$ assumes the same value in any two sample trajectories. For each sample i, the cost is

$$\varepsilon(u^i)^2 + (w_0^i + u^i + w_1^i)^2 = (\varepsilon+1)(u^i)^2 + 2(w_0^i + w_1^i)u^i + (w_0^i + w_1^i)^2 . \tag{1.12}$$

This expression must be minimized in u^i for every $i = 1, \ldots, N$, under the constraint (1.11). Indeed, if the N sample trajectories are produced by a random drawing with the uniform probability distribution over the square $[-1, 1]^2$, then, with probability 1, w_0^i is different from w_0^j for any couple (i, j) with $i \neq j$. Therefore, with probability 1, the constraint (1.11) is not binding, that is, (1.12) can be minimized for each value of i independently. This yields the optimal value

$$u^i = -\frac{w_0^i + w_1^i}{1+\varepsilon} \tag{1.13}$$

and the corresponding contribution to the cost $\varepsilon(w_0^i + w_1^i)^2/(1+\varepsilon)$. This is of order ε, and so is the average over N samples

[4] What we call here "*N samples* or *sample realizations*" may be referred elsewhere in this book as a *N-sample*, whereas N is referred to as the *number of samples* or as the *size* of the N-sample.

$$\frac{1}{N(1+\varepsilon)} \sum_{i=1}^{N} \varepsilon (w_0^i + w_1^i)^2 \tag{1.14}$$

even when N goes to infinity. This is far from the actual optimal cost given by (1.10).

What Is the Real Value of this "Solution"?

However, any admissible solution (any $\boldsymbol{U}$ such that $\boldsymbol{U} \preceq \boldsymbol{W}_0$) cannot achieve a cost better than the optimal cost (1.10). The value (1.14) is just a "fake" cost estimation. The resolution of the discretized problem derived from the previous Monte Carlo procedure yielded an optimal value u^i (see (1.13)) associated with each sample noise trajectory represented by a point (w_0^i, w_1^i) in the square $[-1, 1]^2$. Hence, before trying to evaluate the cost associated with this "solution", we must first derive from it an *admissible* solution for the original problem, that is, a random variable $\boldsymbol{U}$ over $\Omega = [-1, 1]^2$, but with constant value along every vertical line of this square (since the abscissa represents the first component $\boldsymbol{W}_0$ of the 2-dimensional noise $(\boldsymbol{W}_0, \boldsymbol{W}_1)$).

A natural choice is as follows:

- we first renumber the N sample points so that the first component w_0^i is increasing with i;
- then, we divide the square into N vertical strips by drawing vertical lines in the middle of segments $[w_0^i, w_0^{i+1}]$ (see Figure 1.3), that is, the i-th strip is $[a^{i-1}, a^i] \times [-1, 1]$ with $a^i = (w_0^i + w_0^{i+1})/2$ for $i = 2, \ldots, N-1$, $a^0 = -1$, and $a^N = 1$;[5]
- then, we define the solution $\boldsymbol{U}$ as the function of (w_0, w_1) which is piecewise constant over the square divided into those N strips, using of course the optimal value u^i given by (1.13) in strip i; that is, we consider

$$\boldsymbol{U}(w) = \sum_{i=1}^{N} u^i \mathbf{1}_{[a^{i-1}, a^i] \times [-1,1]}(w) , \tag{1.15}$$

  where w ranges in the square $[-1, 1]^2$ and $\mathbf{1}_A(\cdot)$ is the indicator function which takes the value 1 in A and 0 elsewhere.

Since this is an admissible solution for the original (continuous) problem, the corresponding cost value $\mathbb{E}(\varepsilon \boldsymbol{U}^2 + \boldsymbol{X}_1^2)$ can be evaluated. Here, the expectation is over the argument w considered as a random variable over the square with uniform distribution.

According to (1.9), this expected cost is easily evaluated analytically as

[5] Later on in this book (see §6.1), we discuss the concept of *Voronoi cells*: here we are defining the N Voronoi cells of the segment $[-1, 1]$ which are based on the "centroids" w_0^i.

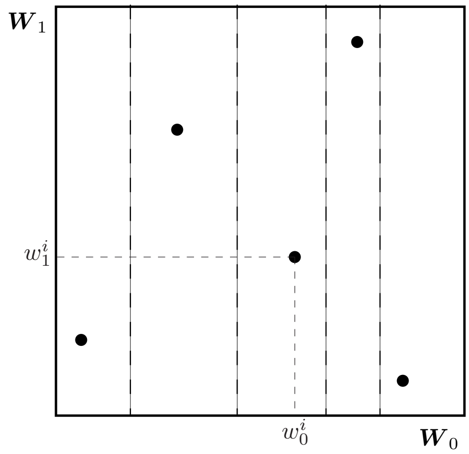

Fig. 1.3. Building an admissible solution for problem (1.8)

$$\frac{2}{3} + \sum_{i=1}^{N} \left((1+\varepsilon)(u^i)^2 \int_{a^{i-1}}^{a^i} \frac{1}{2}\,\mathrm{d}w_0 + 2u^i \int_{a^{i-1}}^{a^i} \frac{w_0}{2}\,\mathrm{d}w_0 \right)$$
$$= \frac{2}{3} + \sum_{i=1}^{N} \left((1+\varepsilon)(u^i)^2 \frac{a^i - a^{i-1}}{2} + u^i \frac{(a^i)^2 - (a^{i-1})^2}{2} \right) . \quad (1.16)$$

Although this is an "expected" cost, it is still a random variable since u^i and a^i are functions of the w_0^i's which result from random drawings (u^i also depends upon the w_1^j's). Indeed, (1.16) should be considered as an *estimation* of the optimal cost resulting from the (random) estimation (1.15) of the true solution.

In order to assess the value of this estimate, and first of all of its possible bias (not to speak of its variance), we must compute the expectation of (1.16) when considering that the w_0^i's are realizations of N independent random variables $\boldsymbol{W}_0^i$, each uniformly distributed over $[-1, 1]$. This calculation is not straightforward. The expression of the a^i's as functions of the w_0^i's is meaningful as long as the w_0^i's have been reordered into an increasing sequence. Therefore, although those N random numbers are the result of independent drawings, the calculation of expectations is made somewhat tricky by this reordering. We therefore skip it here. But, we have used a simple computer program using a pseudo-random number generator to evaluate the mean and standard deviation of this estimated cost as functions of the number N of used samples (for each value of N, the program uses 1,000 series of N drawings in order to evaluate those statistics). Figure 1.4 shows the results: the averaged cost $\pm$ the standard deviation are depicted as functions of N (here ε is taken equal to $1/100$).

By observing Figure 1.4, as N goes to infinity, the expected value of (1.16) goes to $2/3$. Remember that the true optimal cost (see (1.10)) was close

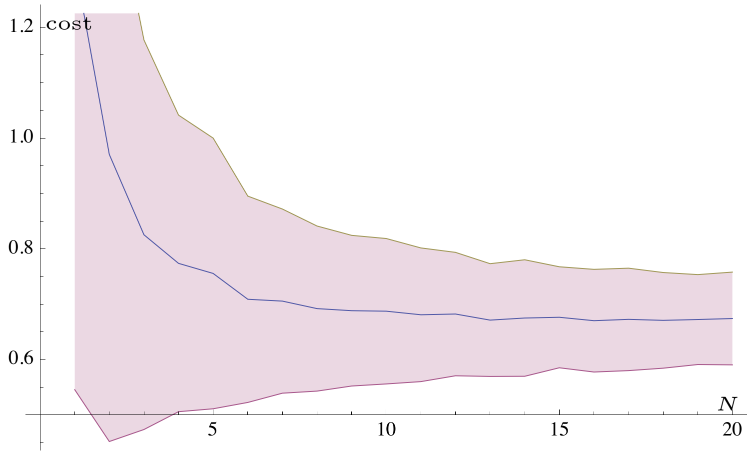

Fig. 1.4. Cost provided by the naive Monte Carlo method as a function of the number N of samples

to 1/3! Moreover, it is readily checked that the optimal *open-loop* solution, that is the optimal $\boldsymbol{U}$ which is measurable w.r.t. the trivial σ-field $\{\emptyset, \Omega\}$, is equal to 0 and that the corresponding cost is also 2/3. Hence the solution we have produced with our naive Monte-Carlo approach (and especially the naive way (1.11) of handling the information structure of the problem) is not better than the open-loop solution!

How to Improve the Monte Carlo Approach? The Idea of Scenario Trees

Reviewing the previous procedure to provide an estimate of the solution of the original problem, one realizes that a crucial step, after the somewhat classical one of Monte Carlo sampling, is to translate the informational constraint $\boldsymbol{U} \preceq \boldsymbol{W}_0$ into the discretized version of the problem. The constraint (1.11) is rather ineffective, and it leads to the fact that the optimal value u^i (see (1.13)) found for sample i is "anticipative": u^i depends on w_1^i, which should not be the case. This explains why the apparent cost (that evaluated by averaging over the N samples) is very optimistic (of order ε whereas the true optimal cost is 1/3).

On the other hand, when one is required to propose an *admissible* solution for the continuous problem, (namely (1.15) which satisfies the measurability constraint), this avoids the drawback of anticipativity, but then we have seen that the corresponding cost is as bad as that of the open-loop solution.

The question is thus: how to make another constraint translating the informational constraint in the discretized problem more effective than (1.11)? An obvious answer is that, in our collection of sample trajectories used in the discrete optimization problem, there should *really* be *distinct* samples with the *same* value of component w_0. This can be viewed as the origin of the idea of *"scenario trees"*. Here "scenario" is another terminology for "sample" and "tree"[6] refers to the shape depicted in Figure 1.5. In this figure, one must

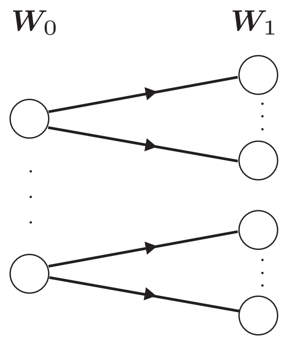

Fig. 1.5. A scenario tree on two stages

imagine that a certain sample value w_0^j is attached to each node j of the first stage in the tree and that sample values w_1^k are likewise attached to nodes k at the second stage. Therefore, since distinct scenarios correspond to distinct "leaves" of the tree (they are still numbered with i ranging from 1 to N), the tree shape implies that several scenarios (couples (w_0^i, w_1^i)) share common values w_0^i. For ease of notation, we assume that all nodes of the first level (numbered with $j = 1, \dots, N_0$) have the same number N_1 of "sons" (successors at the second stage, numbered with $k = 1, \dots, N_1$ for each j). Hence $N = N_0 \times N_1$.

Admittedly, if the scenarios are produced randomly (according to the joint uniform probability law of $(\boldsymbol{W}_0, \boldsymbol{W}_1)$ over the square $[-1, 1] \times [-1, 1]$), or if they have been recorded from real life observations, there is a probability *zero* that a tree shape pops up spontaneously, for any arbitrary large, but finite, N. The question of how a scenario tree can be derived from real recorded data is considered in Chapter 6. The situation is easier if one knows the underlying probability law. In our example, since $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$ are known to be independent (the white noise case), any element in a set of N_0 samples of $\boldsymbol{W}_0$ can be combined with the same, or N_0 distinct, sets of N_1 samples of $\boldsymbol{W}_1$ to produce such a tree. Even if $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$ were not independent, one could first generate N_0 samples of $\boldsymbol{W}_0$ using the marginal probability law of this variable, and then, using each sample w_{0j} and the *conditional* probability

[6] Actually, in Figure 1.5, a "forest", that is, a collection of trees, rather than a "tree", is depicted since there are several "root nodes" which are the nodes at the first level. But we keep on speaking of "trees" to match the traditional terminology of "scenario tree".

law of $\boldsymbol{W}_1$ knowing that $\boldsymbol{W}_0$ assumes the value w_0^j, one could generate N_1 associated samples w_1^k of $\boldsymbol{W}_1$ ("sons" of that w_0^j).

It is not our purpose now to discuss the production of "good" scenario trees. We just assume that such a scenario tree has been obtained, and that it reflects good statistical properties w.r.t. the underlying probability law of the noises when N_0 and N_1 go to infinity, in a sense that we leave to the reader's intuition at this stage. Our purpose is to revisit the resolution of the discretized problem formulated with this scenario tree and to examine its asymptotic behavior when the number of samples becomes very large. To fix notations, we consider scenarios $\{(w_0^j, w_1^{jk})\}_{j=1,\dots,N_0}^{k=1,\dots,N_1}$ and we introduce the following additional symbols:

$$\overline{w}_1^j = \frac{1}{N_1}\sum_{k=1}^{N_1} w_1^{jk} \ , \quad (\overline{\sigma}_1^j)^2 = \frac{1}{N_1}\sum_{k=1}^{N_1}(w_1^{jk})^2 \ . \tag{1.17}$$

Notice that $\overline{w}_1^j$ can be interpreted as an estimate of the *conditional expectation* of $\boldsymbol{W}_1$ knowing that $\boldsymbol{W}_0 = w_0^j$. Likewise, $(\overline{\sigma}_1^j)^2$ can be interpreted as an estimate of the *conditional* second order moment.

To each node of the first level of the tree is attached a control variable u^j. The cost of the discretized problem is

$$\frac{1}{N_0}\sum_{j=1}^{N_0}\Big(\varepsilon(u^j)^2 + \frac{1}{N_1}\sum_{k=1}^{N_1}(u^j + w_0^j + w_1^{jk})^2\Big) \ .$$

The arg min is

$$u^j = -\frac{w_0^j + \overline{w}_1^j}{1+\varepsilon} \ , \ j = 1,\dots,N_0 \ , \tag{1.18}$$

to be compared with (1.13). This yields the optimal cost

$$\frac{1}{N_0(1+\varepsilon)}\sum_{j=1}^{N_0}\big(\varepsilon(w_0^j)^2 + 2\varepsilon w_0^j\overline{w}_1^j - (\overline{w}_1^j)^2 + (1+\varepsilon)(\overline{\sigma}_1^j)^2\big) \ , \tag{1.19}$$

to be compared with (1.14) and (1.10). If we assume that the estimates (1.17) converge towards their right values (respectively, 0 and 1/3) as N_1 goes to infinity, then (1.19) gets close to

$$\frac{1}{N_0(1+\varepsilon)}\sum_{j=1}^{N_0}\Big(\varepsilon(w_0^j)^2 + \frac{1+\varepsilon}{3}\Big) \ .$$

Now, the expression $(1/N_0)\sum_{j=1}^{N_0}(w_0^j)^2$ can also be viewed as an estimate of the second order moment of $\boldsymbol{W}_0$ and, if we assume that it converges to the true value 1/3 when N_0 goes to infinity, then we recover, in the limit, the true optimal cost (1.10). Therefore, unlike with the previous naive Monte Carlo method (see (1.14)), here the optimal cost obtained in the discrete problem appears to converge to the right value.

As seen earlier (see (1.16)), it is also interesting to evaluate the real cost associated with an admissible solution derived from the collection of "optimal" values (1.18) by plugging those values into the formula (1.15) (with N replaced by N_0). Again, we have appealed to a computer program using 1,000 experiments, each consisting in:

- drawing N_0 values w_0^j at random;
- associated with each of those values, drawing a set of N_1 values w_1^{jk} at random;
- computing the $\overline{w}_1^j$'s (see (1.17)), the u^j's (see (1.18)) and forming the admissible solution (1.15) (N replaced by N_0) with those values after reordering the indices j so that w_0^j is increasing with j;
- evaluating the true cost $\mathbb{E}(\varepsilon \boldsymbol{U}^2 + \boldsymbol{X}_1^2)$ by analytic integration w.r.t. the couple $w = (w_0, w_1)$ with uniform probability distribution over the square $[-1, 1]^2$.

Remember that this integral w.r.t. argument w appearing in (1.15) is done for random values u^j depending on the random drawings w_0^j and w_1^{jk}. The 1,000 experiments are used to evaluate the mean and standard deviation of the random cost so obtained. In those experiments, we took $N_0 = N_1$, that is, $N_0 = \sqrt{N}$.

Figure 1.6 depicts the mean $\pm$ the standard deviation of the cost as a function of $N_0 = \sqrt{N}$ (still with $\varepsilon = 1/100$). The limit as N goes to infinity seems to be the correct value of the optimal cost given by (1.10), namely 0.3366, but the convergence appears to be asymptotically very slow, a fact on which we comment further in Chapter 6.

Observe that in the comparison with Figure 1.4, while the abscissa does represent the number of pieces uses to approximate the random variable $\boldsymbol{U}(\cdot)$ in both plots, in Figure 1.4, this abscissa represents also the number of samples used to achieve the Monte Carlo approximation whereas in Figure 1.6, this number of samples is the square of the number of pieces.

By the way, an interesting question is how to choose N_0 and N_1, for a given N with $N = N_0 \times N_1$, so as to get the minimum standard deviation of the cost estimate (or of the estimate of the true solution $\boldsymbol{U}$). This is a question that can be generalized to the question of choosing the best tree topology in a multi-stage problem (here the problem was 2-stage), given the number N of leaves of the tree.

1.5 Conclusion

When moving from *deterministic* to *stochastic* optimization, one must handle the evaluation of mathematical expectations, which typically involves the use of Monte Carlo sampling. However, when considering *dynamic* stochastic

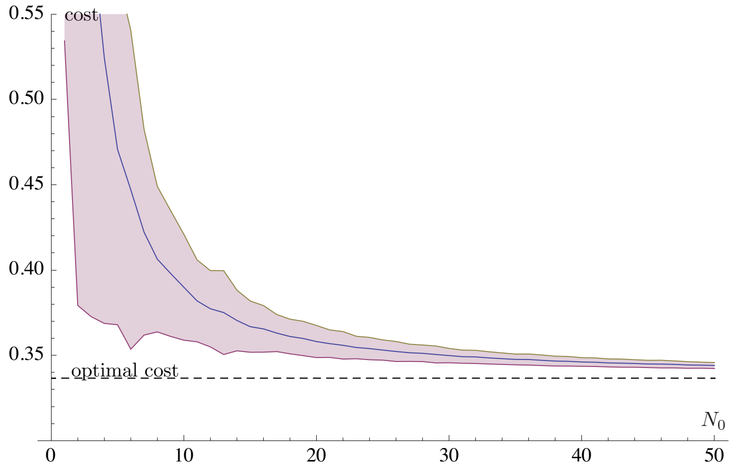

Fig. 1.6. Cost provided by the use of a stochastic tree as a function of the number N_0 of pieces of the piecewise constant $\boldsymbol{U}(\cdot)$ (N_0^2 scenarios)

optimization, another important aspect of the formulation is the specification of the *information structure*, which amounts to defining what one knows each time a decision has to be made.

In this introductory chapter, we described various information structures and the difficulties — which are sometimes tremendous even for seemingly rather simple problems (see Witsenhausen's counterexample at §1.3.3) — that may result from those informational constraints.

Even if we restrict ourselves to problems with SIS (see §1.2.2), obtaining a sound discretized version of the problem with a consistent formulation of the informational constraint is not as trivial a task as we tried to illustrate it in §1.4.

In the rest of this book, the most complex phenomena of DIS and the associated *dual effect* are discussed (see Chapters 4 and 10). However, the attempt to give systematic methodologies to obtain sound discrete versions of stochastic optimization problems is restricted to problems with SIS (Chapter 6).