6

Discretization Methodology for Problems with Static Information Structure (SIS)

In this chapter, we consider problems formulated as in (1.3), which we repeat here for convenience

$$\min_{\boldsymbol{U} \preceq \mathcal{G}} \mathbb{E}\big(j(\boldsymbol{U}, \boldsymbol{W})\big) \tag{6.1a}$$

where $\mathcal{G}$ is a σ-field, or

$$\min_{\boldsymbol{U} \preceq \boldsymbol{Y}} \mathbb{E}\big(j(\boldsymbol{U}, \boldsymbol{W})\big) \ , \tag{6.1b}$$

where $\boldsymbol{Y}$ is a random variable (called observation). Both $\mathcal{G}$ and $\boldsymbol{Y}$ are *static*, that is, they do not depend on the control $\boldsymbol{U}$ (in §1.2.2, we used the acronym SIS for this situation). Recall that problems with DIS (see again §1.2.2), but no dual effect, are also amenable to this formulation (such situations are considered in §10.3).

We are mainly interested in devising systematic approaches to the discretization of such problems in order to solve them numerically with the help of a computer. Essentially, in the discretized problem, any random variable, be it part of the data as $\boldsymbol{W}$, or of the unknowns as $\boldsymbol{U}$, is represented by a finite set of values (e.g. $\{w^i\}_{i=1,\dots,N}$), and its associated probability law is represented by a sum of atomic measures (Dirac measures δ_{w^i} located at w^i) with positive weights p^i summing up to 1. Consequently, in the discrete problem, expectations reduce to finite sums, and optimization is w.r.t. a finite set $\{u^i\}$ of variables.

Before we can address this main topic, the next section briefly discusses the theory of *quantization* which is essentially a tool to derive approximate, but finite, representations of random variables, and which provides a framework in which to discuss the quality of those approximations.

6.1 Quantization

When trying to solve stochastic optimization problems numerically, one may have to manipulate approximate, but *finite*, representations of random variables. The quantization technique reduces the amount of information necessary to represent a random variable while trying to preserve as much as possible of the original random variable. It has its origin in Communication Theory [67], in which random signals must be sent through a channel with limited bandwidth. By reducing the amount of information necessary to describe, and thus transmit, the signal, one hopes to increase the flow of signals sent through the channel. At the same time, the signals should be distorted as little as possible. There is clearly a trade-off here.

In this text, we do not address this trade-off directly; we rather assume that the amount of information retained to represent a random variable is given,[1] and we try to minimize the distortion in the representation of the random variable under this constraint.

Indeed, we first start with set-theoretic notions that are limited to algebraic aspects of quantization. Then, we move to the more quantitative notion of *optimal* quantization, where the set over which quantization is considered must be a *normed vector space*.

6.1.1 Set-Theoretic Quantization

A random variable $\boldsymbol{W}$ is a measurable mapping from a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to a measurable space $(\mathbb{W}, \mathcal{W})$. In this subsection, the probability law plays no role, but it is used in the next subsection.

Consider a *projection* $Q : \mathbb{W} \to \mathbb{W}$, that is, a measurable mapping such that $Q \circ Q = Q$. Assume, moreover, that its image $\operatorname{im} Q$ has a finite cardinality. That is, it contains a finite number N of distinct values. We call Q a *quantization* and $Q \circ \boldsymbol{W}$ a *quantized* approximation of $\boldsymbol{W}$.

We may consider Q as factorized into two mappings,

$$Q = d \circ e \ ,$$

where

1. $e : \mathbb{W} \to \{1, \ldots, N\}$ is called the *encoding*;
2. $d : \{1, \ldots, N\} \to \operatorname{im} Q \subset \mathbb{W}$ is a *bijection*, which is called the *decoding*.

In a communication context, instead of sending values $w \in \mathbb{W}$ over the communication channel, only the *code* $i = e(w) \in \{1, \ldots, N\}$ is sent; at the other end of the channel, the message i is *decoded* by using $d(i) \in \operatorname{im} Q$; for this reason, $\operatorname{im} Q \subset \mathbb{W}$ is called the *codebook*: this is a collection of N values in $\mathbb{W}$.

[1] This is likely to determine the complexity of the discretized optimization problem we seek to formulate.

Of course, $\Omega/(Q \circ \boldsymbol{W})$ is a partition[2] of Ω with N elements. Observe that this partition which defines the information carried by the quantized random variable is independent of the decoding d (as long as this is a bijection), it only depends on the encoding e. Otherwise stated, $\Omega/(Q \circ \boldsymbol{W}) = \Omega/(e \circ \boldsymbol{W})$.

Similarly,

$$\mathbb{W}/Q = \mathbb{W}/e \,, \tag{6.2}$$

and this defines a partition of $\mathbb{W}$, the N elements of which are called *quantization cells*. All elements in cell i are represented by the same *representative* $d(i)$ in the codebook. In summary, while the encoding e defines the cells (and thus the information carried by the quantized variable), the decoding d defines the representative in each cell (called the *centroid* of the cell), which has an importance as long as the values w are physical quantities (for example, consumption of energy, prices, etc.). The whole situation is illustrated by Figure 6.1.

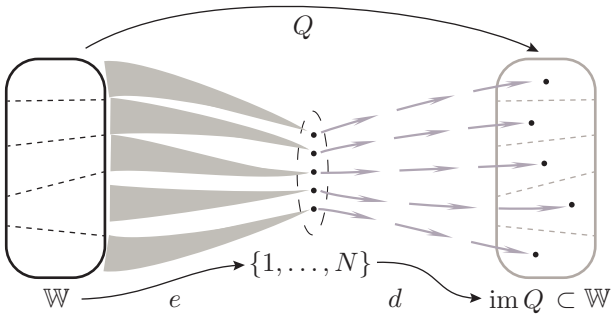

Fig. 6.1. Quantization Q, encoding e, decoding d

6.1.2 Optimal Quantization In Normed Vector Spaces

We assume now that $\mathbb{W}$ is a normed vector space (the norm is denoted $\|\cdot\|$). Given a value of N, the idea is to choose the least distorted quantized variable. Distortion may be defined by the L^2-distance between the original and the quantized random variable, which is equal to the square root of the *Mean Quadratic Error* (MQE):

$$\mathrm{MQE} := \mathbb{E}\big(\big\| \boldsymbol{W} - Q \circ \boldsymbol{W} \big\|^2 \big) \,. \tag{6.3}$$

Optimizing the quantization amounts to reducing this distortion measure to the minimum. This task can be split up into two parts: choosing the best encoding e, or equivalently defining the best partition $\mathbb{W}/e$; and specifying

[2] The notation $\Omega/(Q \circ \boldsymbol{W})$ was introduced in Definition 3.30.

the best decoding d, which consists of choosing the best representative (centroid) w^i in each cell C^i of this partition. The next lemma provides two necessary conditions related to those two optimal choices.

Lemma 6.1. *An optimal quantization must satisfy the following two conditions:*

1. *given the centroids* $\{w^i\}_{i=1,\dots,N}$*, the cells* C^i *must be such that,* $\mathbb{P}_{\boldsymbol{W}}$*-almost surely, if* $w \in C^i$*, then* $\|w - w^i\| \leq \|w - w^j\|, \forall j \neq i$*;*
2. *given the cells* C^i*, the centroid* w^i *is equal to* $\mathbb{E}\big(\boldsymbol{W} \mid \boldsymbol{W}^{-1}(C^i)\big)$*.*

Proof. The MQE can be written as follows

$$\mathbb{E}\big(\|\boldsymbol{W} - Q \circ \boldsymbol{W}\|^2\big) = \mathbb{E}\Big(\sum_{i=1}^{N} \mathbb{E}\Big(\|\boldsymbol{W} - w^i\|^2 \;\Big|\; \boldsymbol{W}^{-1}(C^i)\Big)\Big) .$$

Suppose the first condition of the lemma is not satisfied. Then, since N is finite, it means that there exist a subset of $\mathbb{W}$ with *positive* probability for $\mathbb{P}_{\boldsymbol{W}}$ and two indices i and j such that every w in that subset belongs to C^i whereas $\|w - w^j\| < \|w - w^i\|$. Then, by changing the definition of cells so that this whole subset is moved from C^i to C^j, the above performance index is improved, which contradicts optimality.

Now, considering the i-th conditional expectation in the right-hand side above, it is well known that $w^i = \mathbb{E}\big(\boldsymbol{W} \mid \boldsymbol{W}^{-1}(C^i)\big)$ is the value which minimizes that expression (see Definition B.5). □

The first condition in this lemma defines what is known as a *Voronoi diagram* or *tessellation* (see Figure 6.2). This condition may be called *"the nearest neighbor"* condition in that any w should be represented by its nearest neighbor in the codebook.

Using (6.2), observe that

$$\mathbb{E}\big(Q(\boldsymbol{W})\big) = \sum_{i=1}^{N} \mathbb{P}_{\boldsymbol{W}}(C^i) w^i = \mathbb{E}\Big(\mathbb{E}\big(\boldsymbol{W} \mid \boldsymbol{W}^{-1}(\mathbb{W}/e)\big)\Big) = \mathbb{E}(\boldsymbol{W}) , \quad (6.4)$$

that is, an optimal quantized variable is respectful of the first moment of the original random variable. However, the next lemma shows that the second order moment or the variance is underestimated when replacing the original variable by its optimal quantized version.

Lemma 6.2. *For an optimal quantization* Q*, with* MQE *defined by* (6.3)*, one has that:*

$$\mathbb{E}(\boldsymbol{W}) = \mathbb{E}\big(Q(\boldsymbol{W})\big) , \quad (6.5a)$$

$$\mathrm{Var}\,\big(\boldsymbol{W}\big) = \mathrm{Var}\,\big(Q(\boldsymbol{W})\big) + \mathrm{MQE} . \quad (6.5b)$$

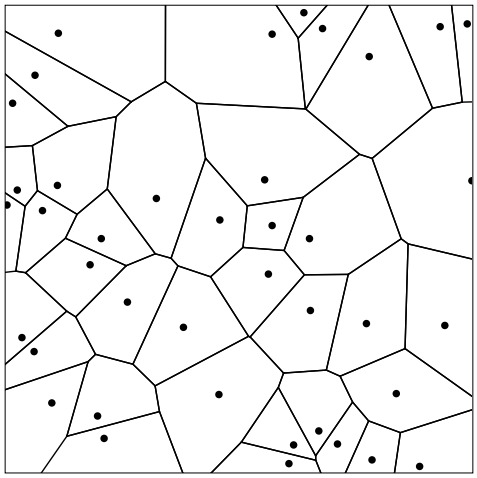

Fig. 6.2. Voronoi tessellation

Proof. The former claim is a repetition of (6.4). We concentrate on the latter.

$$\begin{aligned}\mathrm{Var}\,(\boldsymbol{W}) &= \mathbb{E}\big(\,\|\boldsymbol{W}-\mathbb{E}(\boldsymbol{W})\|^2\,\big)\\ &= \mathbb{E}\big(\,\|\boldsymbol{W}-Q(\boldsymbol{W})+Q(\boldsymbol{W})-\mathbb{E}(\boldsymbol{W})\|^2\,\big)\\ &= \mathbb{E}\big(\,\|\boldsymbol{W}-Q(\boldsymbol{W})\|^2\,\big)+\mathbb{E}\big(\,\|Q(\boldsymbol{W})-\mathbb{E}(\boldsymbol{W})\|^2\,\big)\\ &\qquad +2\mathbb{E}\Big(\sum_{i=1}^{N}\mathbb{E}(\langle \boldsymbol{W}-w^i\,,w^i-\mathbb{E}(\boldsymbol{W})\rangle \mid \boldsymbol{W}^{-1}(C^i))\Big)\,.\end{aligned}$$

The first term is precisely the MQE; the second term is the variance of $Q(\boldsymbol{W})$ thanks to (6.5a); the third term is zero since, in the scalar product, the second factor is constant over C^i whereas, the first factor has a zero conditional expectation given the value of w^i (see Lemma 6.1). □

The two necessary optimality conditions in Lemma 6.1 may be used to build up an iterative (Lloyd's) algorithm to find an optimal quantization of a given random variable. This algorithm proceeds by alternating the following two stages:

1. given a collection of N centroids w^i, the Voronoi cells are (re)drawn using the first conditon in Lemma 6.1; this amounts to defining all the half spaces delimited by the medians (hyperplanes) of all segments (w^i, w^j) (it involves manipulating affine inequalities);
2. given the cells, the centroids are (re)defined by using the second optimality condition in the lemma (it involves computing integrals over the cells).

Unfortunately, such an algorithm sometimes tends to stop on local minima which are not true minima (see [67]).

6.2 A Systematic Approach to Discretization

We come to the main topic of this chapter, namely the reformulation of problems as (6.1) as finite dimensional problems in order to solve them numerically. The language of quantization briefly developed previously is used throughout this section, and the contribution by Pennanen [109] is "translated" in this language in order to make easier its comparison with other solutions proposed hereafter.

6.2.1 The Problematics of Discretization

As we have seen in §1.4, the resolution of the discrete problem does not directly provide an acceptable answer in that what is expected as a "solution" to (6.1) is a random variable over the *original* probability space, satisfying, in addition, the measurability conditions imposed in the formulation (6.1). Therefore, some "reconstruction" is needed after the discrete problem has been solved, and examples of this reconstruction were given in §1.4 (see (1.15) and Figure 1.3).

Notice that, even if the formulation of the discrete problem seemingly only requires the consideration of random variables assuming finitely many distinct values, these random variables must be embedded into the original space by defining "cells" around the atomic support of the discrete probability distribution defined by the weights $\{p^i\}_{i=1,\dots,N}$. Then, a reasonable requirement for convergence of the reconstructed solution towards the optimal original problem solution is that each p^i gets arbitrarily close to the true probability mass of the cell around atom i as the number N of cells goes to infinity.

However, as shown by the discussion around examples in §1.4, it is not enough to handle the approximation of mathematical expectations in a sound manner: some caution is also required to properly represent the essential contraints of *information structure* in the discrete problem. Here again, cells around atoms, or as otherwise stated, partitions of the original spaces, play a part. This section concentrates on this particular issue.

Pennanen [109] was probably the first author to envisage this topic in a systematic way. Regarding the approximation techniques for expectations, he considered not only usual Monte Carlo sampling but also Quasi-Monte Carlo and other sophisticated quadrature techniques. In order to be able to give an asymptotic epi-convergence theorem as the number of samples goes to infinity, he imposed a rather strong condition which tightly coordinates the samples used to approximate expectations with the way informational constraints[3] are translated in the discretized problem. This condition naturally leads to the construction of scenario trees that were introduced in the end of Chapter 1 (see Figure 1.5) and that are very popular in the SP community. However, as we shall see by the end of this chapter, the technique of scenario trees is hindered by the poor convergence rate of the underlying Monte Carlo approximation

[3] Actually, Pennanen considers *non anticipativity* constraints only.

technique (see §6.3.2), and therefore it is important to show that other more flexible conditions can also be used that still enable a convergence proof to be given.

In this chapter, the focus is on the links between the sampling technique and the translation of informational constraints in the discrete problem; convergence issues are discussed in Chapter 8. We also consider more general informational constraints beyond non anticipativity constraints. Finally, we use the language of lattice of partition fields introduced in Chapter 3, which provides a formalism making clearer the comparison of Pennanen's work with alternative methods we propose in §6.2.3.

Remark 6.3. Although the issue of convergence is deferred to Chapter 8, the following observation may be kept in mind as a safeguard against unreasonable discretization schemes. Remember that, according to §3.5.2, Problem (6.1) can be reformulated as

$$\mathbb{E}\Big(\min_{u}\mathbb{E}\big(j(u,\boldsymbol{W})\;\big|\;\mathcal{G}\big)\Big)\quad\text{or}\quad\mathbb{E}\Big(\min_{u}\mathbb{E}\big(j(u,\boldsymbol{W})\;\big|\;\boldsymbol{Y}\big)\Big)\,. \tag{6.6}$$

Therefore, in any discretized version derived for the original problem, one should try to identify an expression which serves as an approximation of the conditional expectation of the cost and check that this approximation is sound enough. An application of this observation is encountered, for example, in Remark 6.13. ◊

6.2.2 The Approach Inspired by Pennanen's Work

In [109], sequential stochastic optimization problems are considered with non anticipativity constraints: in the framework (6.1b), this amounts to considering that $\boldsymbol{Y} = h(\boldsymbol{W})$ with $\boldsymbol{Y} = \{\boldsymbol{Y}_t\}_{t=1,2,\dots}$, $\boldsymbol{W} = \{\boldsymbol{W}_t\}_{t=1,2,\dots}$ and $\boldsymbol{Y}_t = h_t(\boldsymbol{W}) = (\boldsymbol{W}_1,\dots,\boldsymbol{W}_t)$.

In what follows, we forget about the time index t which plays no particular part as long as $\boldsymbol{Y}$ does not depend on $\boldsymbol{U}$ (the SIS assumption) and we just retain that $\boldsymbol{Y} = h(\boldsymbol{W})$ where h is a function (generally non injective) from $\mathbb{Y}$ to $\mathbb{W}$.

Remark 6.4. In (6.1b), the cost function j depends on the two random variables $\boldsymbol{U}$ and $\boldsymbol{W}$. The observation $\boldsymbol{Y}$ is another random variable which may, or may not, be in relation with $\boldsymbol{W}$. In the present framework, since we assume that $\boldsymbol{Y} = h(\boldsymbol{W})$, we may consider all involved random variables (including $\boldsymbol{U}$) as measurable mappings from $\mathbb{W}$ to another space, that is, we may consider that $\Omega = \mathbb{W}$. Then $\boldsymbol{W}$ is the identity mapping from $\mathbb{W}$ to itself and $\boldsymbol{Y}$ is identical to the mapping h. This is what is assumed hereafter unless otherwise explicitly stated. ◊

First Stage

The formulation of a discretized problem involves the definition of finite sets to approximate the spaces $\mathbb{Y}$ and $\mathbb{W}$. This is formalized by defining *quantizations* as defined in §6.1. We first consider $\boldsymbol{W}_N$, a quantized version of $\boldsymbol{W}$. More precisely, we consider the following applications and properties:

- $q_{\mathbb{W}} : \mathbb{W} \to \mathbb{W}_N$ where $\mathbb{W}_N$ is a finite subset of cardinality N of $\mathbb{W}$; we require that $q_{\mathbb{W}}(w) = w$ whenever $w \in \mathbb{W}_N$; [4]
- $\iota_{\mathbb{W}}$, the canonical injection of $\mathbb{W}_N$ into $\mathbb{W}$; then,
  - $q_{\mathbb{W}} \circ \iota_{\mathbb{W}}$ is the identity $\mathrm{I}_{\mathbb{W}_N}$ in $\mathbb{W}_N$;
  - $Q_{\mathbb{W}} = \iota_{\mathbb{W}} \circ q_{\mathbb{W}}$ is then a projection in $\mathbb{W}$: this is indeed a quantization as defined in §6.1, and $\boldsymbol{W}_N = Q_{\mathbb{W}} \circ \boldsymbol{W} = Q_{\mathbb{W}}$ (according to Remark 6.4) is the quantized random variable we were looking for;
  - $\mathbb{W}/q_{\mathbb{W}} = \mathbb{W}/Q_{\mathbb{W}}$ is a partition of $\mathbb{W}$ into N cells.

Next, in order to obtain a quantized version of $\boldsymbol{Y}$, Pennanen proceeds in the following way. He first considers the discrete random variable $\boldsymbol{Y}_N$ defined as $h \circ \boldsymbol{W}_N$. Notice that since h is not injective in general, it may happen that the set of values $\mathbb{Y}_N = h \circ Q_{\mathbb{W}}(\mathbb{W})$ has a cardinality smaller than N (despite the subscript N in this notation). Then, consider:

- $\iota_{\mathbb{Y}}$, the canonical injection of $\mathbb{Y}_N$ into $\mathbb{Y}$;
- $h_N : \mathbb{W}_N \to \mathbb{Y}_N$ such that $\iota_{\mathbb{Y}} \circ h_N = h \circ \iota_{\mathbb{W}}$.

Remark 6.5. This h_N can be obtained as

$$h_N = \iota_{\mathbb{Y}}^{-1} \circ h \circ \iota_{\mathbb{W}} \,, \tag{6.7}$$

where $\iota_{\mathbb{Y}}^{-1}$ is any mapping such that $\iota_{\mathbb{Y}} \circ \iota_{\mathbb{Y}}^{-1} = \mathrm{I}_{\mathbb{Y}_N}$, the identity over $\mathbb{Y}_N$. This $\iota_{\mathbb{Y}}^{-1}$ is not uniquely defined: any mapping from $\mathbb{Y}$ to $\mathbb{Y}_N$ can play the role of $\iota_{\mathbb{Y}}^{-1}$, as long as its restriction to $\mathbb{Y}_N$ behaves as the identity. Nevertheless, h_N is well defined since precisely, in (6.7), only the restriction of $\iota_{\mathbb{Y}}^{-1}$ to $\mathbb{Y}_N$ is involved. ◊

Therefore, $\boldsymbol{Y}_N$ is well defined:

$$\boldsymbol{Y}_N = h(\boldsymbol{W}_N) = h \circ Q_{\mathbb{W}} = \iota_{\mathbb{Y}} \circ h_N \circ q_{\mathbb{W}} \,. \tag{6.8}$$

The expression $h \circ Q_{\mathbb{W}}$ is also well defined once $Q_{\mathbb{W}}$ has been chosen (while h is given). The situation is summarized in Figure 6.3.

Obviously, $\boldsymbol{Y}_N \preceq \boldsymbol{W}_N$. However, $\boldsymbol{Y}_N$ is not necessarily a quantized version of $\boldsymbol{Y}$ and, in particular, one cannot claim that $\boldsymbol{Y}_N \preceq \boldsymbol{Y}$. Here is a counterexample.

[4] The finite set $\mathbb{W}_N$ generally results from some sampling of the noise $\boldsymbol{W}$ or alternative quadrature methods. But, as recognized by Pennanen himself, in order to define a consistent discretization scheme, it is not enough to introduce the finite set $\mathbb{W}_N$, but it is also necessary to define how the whole original set $\mathbb{W}$ is mapped onto that finite set: this is why $q_{\mathbb{W}}$ must be defined.

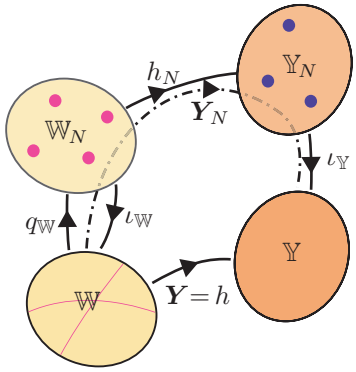

Fig. 6.3. Discretization according to Pennanen, stage 1

Example 6.6. Consider

$$\mathbb{W} = [-2,2] \times [-2,2]\,, \quad q_{\mathbb{W}}(w_1,w_2) = \begin{pmatrix} \operatorname{sign}(w_1) \\ \operatorname{sign}(w_2) \end{pmatrix}\,, \quad \mathbb{W}_N = \{(\pm 1, \pm 1)\}\,,$$
$$h(w_1,w_2) = w_1 + w_2\,, \quad \mathbb{Y}_N = \{-2,0,2\}\,.$$

Observe that $h(-1,2) = h(1/2,1/2) = 1$ whereas $\boldsymbol{Y}_N(-1,2) = \operatorname{sign}(-1) + \operatorname{sign}(2) = 0$ is different from $\boldsymbol{Y}_N(1/2,1/2) = \operatorname{sign}(1/2) + \operatorname{sign}(1/2) = 2$. According to Proposition 3.38 (item 2), $\boldsymbol{Y}_N$ is not measurable with respect to $\boldsymbol{Y}$. The two dots with coordinates $(-1,2)$ and $(1/2,1/2)$ are represented in Figure 6.4. This figure also displays three partitions of $\mathbb{W} = [-2,2] \times [-2,2]$. The first partition corresponds to $\mathbb{W}/h$, and it has infinitely many elements (only a few are represented in the figure). The second partition corresponds to $\mathbb{W}/Q_{\mathbb{W}}$ and it has four elements (the values of $Q_{\mathbb{W}}$ belonging to $\mathbb{W}_N$ are indicated). The third partition, namely $\mathbb{W}/(h \circ Q_{\mathbb{W}})$, has only three elements corresponding to the three elements of $\mathbb{Y}_N$ indicated in the figure. △

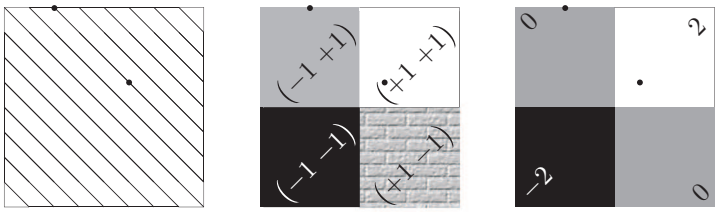

Fig. 6.4. Representations of the partitions corresponding to $\mathbb{W}/h$, $\mathbb{W}/Q_{\mathbb{W}}$ and $\mathbb{W}/(h \circ Q_{\mathbb{W}})$ in Example 6.6

Second Stage

The discretized problem may be considered to "live" on a discrete probability space:

- Ω_N is $\mathbb{W}_N$;
- the associated σ-algebra is the complete partition field of this set;
- the discrete probability law $\mathbb{P}_N$ is the original probability law $\mathbb{P}$ transported from $\mathbb{W}$ to $\mathbb{W}_N$ by $q_{\mathbb{W}}$.

The decision variable in this discretized problem is denoted $\boldsymbol{U}_N$: this is a mapping from $\mathbb{W}_N$ to $\mathbb{U}$, and this mapping can assume *at most* N distinct values; these values result from a numerical optimization. However, the expected answer is a random variable, that is, a mapping from $\mathbb{W}$ to $\mathbb{U}$. As it is natural, $\boldsymbol{U}_N$ is extended to the whole $\mathbb{W}$ by considering $\boldsymbol{U} = \boldsymbol{U}_N \circ q_{\mathbb{W}}$, which amounts to building up a piecewise constant function using the cells of $\mathbb{W}/q_{\mathbb{W}}$ (but again, although there are N cells, there might be less than N distinct values of the control).

The searched solution $\boldsymbol{U}$ should also satisfy some measurability requirements in order to reflect the information structure of the original problem. As a straightforward translation of the informational constraint in (6.1b), in the discrete problem T. Pennanen requires that

$$\boldsymbol{U}_N \preceq h_N \,. \tag{6.9}$$

Lemma 6.7. *The condition* (6.9) *implies that* $\boldsymbol{U} \preceq \boldsymbol{Y}_N$.

Proof. Indeed, $\boldsymbol{U}_N \preceq h_N$ implies that $\boldsymbol{U}_N \circ q_{\mathbb{W}} \preceq h_N \circ q_{\mathbb{W}}$ (see (3.41)), and $h_N \circ q_{\mathbb{W}} \equiv \iota_{\mathbb{Y}} \circ h_N \circ q_{\mathbb{W}}$ since $\iota_{\mathbb{Y}}$ is injective (see Proposition 3.41). The latter is just $\boldsymbol{Y}_N$ (see (6.8)). □

Condition (6.9) implies that there exists a mapping $\gamma_N : \mathbb{Y}_N \to \mathbb{U}$ such that $\boldsymbol{U}_N = \gamma_N \circ h_N$. However, since $\boldsymbol{Y}_N$ is not necessarily measurable w.r.t. $\boldsymbol{Y}$ (as shown by Example 6.6), the proposed $\boldsymbol{U}$ is not necessarily measurable w.r.t. $\boldsymbol{Y}$ either, and, in this case, it would not be an admissible solution for problem (6.1b). Therefore, Pennanen finally requires the following additional condition:

$$\boldsymbol{Y}_N \preceq \boldsymbol{Y} \tag{6.10a}$$

which is equivalent (see Proposition 3.46) to

$$\exists Q_{\mathbb{Y}} : \operatorname{im} \mathbb{Y} \to \mathbb{Y} \quad \text{such that} \quad \boldsymbol{Y}_N = Q_{\mathbb{Y}} \circ \boldsymbol{Y} = Q_{\mathbb{Y}} \circ h \,. \tag{6.10b}$$

With this condition, one then has that $\boldsymbol{U} \preceq \boldsymbol{Y}_N \preceq \boldsymbol{Y}$, that is, $\boldsymbol{U}$ is now an admissible solution for (6.1b).

Put together, conditions (6.8) and (6.10b) imply that

$$h \circ Q_{\mathbb{W}} = Q_{\mathbb{Y}} \circ h \,. \tag{6.11}$$

Lemma 6.8. *Equation* (6.11) *implies that* $Q_{\mathbb{Y}}$ *(with domain* $\operatorname{im} h$*) is a projection; hence,* $\boldsymbol{Y}_N = Q_{\mathbb{Y}} \circ \boldsymbol{Y}$ *is a quantization of* $\boldsymbol{Y}$*. Therefore,* $Q_{\mathbb{Y}}$ *can be factorized as* $\iota_{\mathbb{Y}} \circ q_{\mathbb{Y}}$ *where* $q_{\mathbb{Y}} : \mathbb{Y} \to \mathbb{Y}_N$ *and* $\iota_{\mathbb{Y}}$ *is again the canonical injection of* $\mathbb{Y}_N$ *into* $\mathbb{Y}$*. Then,* (6.11) *is equivalent to*

$$h_N = q_{\mathbb{Y}} \circ h \circ \iota_{\mathbb{W}} \,. \tag{6.12}$$

Proof. We must show that $Q_{\mathbb{Y}} \circ Q_{\mathbb{Y}} \circ h = Q_{\mathbb{Y}} \circ h$. Indeed, with (6.11) used repeatedly, and the fact that $Q_{\mathbb{W}}$ itself is a projection, one has that

$$Q_{\mathbb{Y}} \circ Q_{\mathbb{Y}} \circ h = Q_{\mathbb{Y}} \circ h \circ Q_{\mathbb{W}} = h \circ Q_{\mathbb{W}} \circ Q_{\mathbb{W}} = h \circ Q_{\mathbb{W}} = Q_{\mathbb{Y}} \circ h \,.$$

Now, with reference to the right-hand side of (6.8), (6.11) can be written as

$$\iota_{\mathbb{Y}} \circ h_N \circ q_{\mathbb{W}} = \iota_{\mathbb{Y}} \circ q_{\mathbb{Y}} \circ h \,, \tag{6.13}$$

which is equivalent to

$$h_N \circ q_{\mathbb{W}} = q_{\mathbb{Y}} \circ h \,,$$

since $\iota_{\mathbb{Y}}$ is injective, and this is again equivalent to (6.12) when composing both sides of the above equation with $\iota_{\mathbb{W}}$ (which is injective) to the right hand and remembering that $q_{\mathbb{W}} \circ \iota_{\mathbb{W}}$ is nothing but the identity in $\mathbb{W}_N$. □

The situation is summarized as follows, and it is illustrated by Figure 6.5. Given the noise $\boldsymbol{W}$ and the observation $\boldsymbol{Y} = h(\boldsymbol{W})$,

- a quantized noise $\boldsymbol{W}_N = Q_{\mathbb{W}} \circ \boldsymbol{W}$ is first defined;
- a discrete random variable $\boldsymbol{Y}_N = h \circ \boldsymbol{W}_N$ is next introduced;
- this discrete random variable is given the status of a quantized observation by imposing Condition (6.10);
- the last two steps finally result in Condition (6.11).

Discussion

Equation (6.11) says that the quantized observation must be the observation of the quantized noise. This condition is intuitively appealing. However, it is unclear how one can ensure it in a systematic construction of a discretization scheme in this general setting. In the particular case of *non anticipativity* constraints, Pennanen [110] proposed a procedure that we briefly discuss in §6.3. What makes things rather locked in general is the initial requirement (6.8) that the observation function h_N in the discrete model should be intimately related to the original observation function h. This is precisely the condition we relax later on.

Equation (6.12) should be compared with Equation (6.7): this shows that the choice of the mapping $\iota_{\mathbb{Y}}^{-1}$ in the latter equation is nothing but the choice

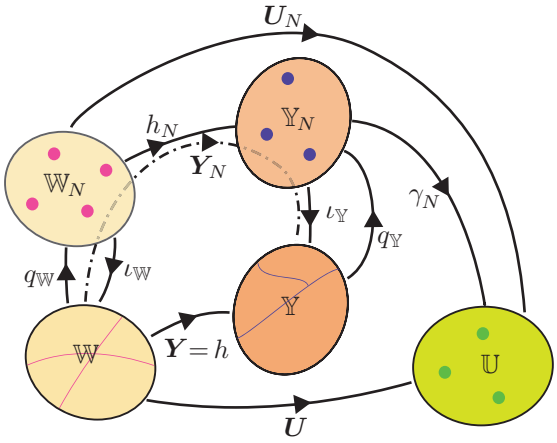

Fig. 6.5. Discretization according to Pennanen, stage 2

of the quantization map $q_{\mathbb{Y}}$ which defines the cells in $\mathbb{Y}$ (as long as the set of centroids $\mathbb{Y}_N$ is already defined).

Observe that (6.8) implies that $\boldsymbol{Y}_N \preceq \boldsymbol{W}_N$ whereas, by (6.10a), $\boldsymbol{Y}_N \preceq \boldsymbol{Y}$, hence $\boldsymbol{Y}_N \preceq \boldsymbol{W}_N \wedge \boldsymbol{Y}$ (see Chapter 3). Then, consider the following example.

Example 6.9. To stay close to the sequential situation considered by Pennanen and still maintain simplicity, we consider (1.8) again. All sample trajectories (w_0^i, w_1^i) of the noise $\boldsymbol{W} = (\boldsymbol{W}_0, \boldsymbol{W}_1)$ are represented as dots in the square $\mathbb{W} = [-1, 1] \times [-1, 1]$ with coordinates (w_0^i, w_1^i).

A quantization based on such a sampling may be obtained by drawing the Voronoi tessellation corresponding to this set of dots (Figure 6.2 illustrates the partition $\Omega/\boldsymbol{W}_N$). On the other hand, $h(w_0, w_1) = w_0 \in \mathbb{Y} = [-1, 1]$. The partition $\Omega/\boldsymbol{Y}$ corresponds to a decomposition of the square into all vertical segments it contains. According to the way the greatest lower bound of partitions is obtained (see §3.3.1), it is realized that $\Omega/(\boldsymbol{W}_N \wedge \boldsymbol{Y})$ is likely to consist of the whole square as the single element, and this remains true even when N goes to infinity. Otherwise stated, $\boldsymbol{W}_N \wedge \boldsymbol{Y}$ remains stuck to the class of bottom elements in the lattice of functions over Ω, namely the class of *constant* functions. Since the "solution" $\boldsymbol{U}$ produced is constrained to be measurable w.r.t. $\boldsymbol{Y}_N \preceq \boldsymbol{W}_N \wedge \boldsymbol{Y}$ (see Lemma 6.7), it cannot be better than the solution in the class of *open-loop* controls. △

This example shows that some necessary conditions derived from Pennanen's conditions (6.8)–(6.10) are not sufficient to ensure convergence of the discrete problem solution towards that of the original problem as N goes to infinity (convergence that Pennanen could prove however). Therefore, Pennanen's conditions are strong enough to avoid the pitfall described in the previous example. The main practical difficulty is that the quantized obser-

vation $\boldsymbol{Y}_N$ is not mastered directly (that is, a priori and directly derived from $\boldsymbol{Y}$ by quantization, which would make things a lot easier to design).

Let us explain why condition (6.8), in fact, reflects a stochastic tree structure as depicted by Figure 1.5. To show this, one must imagine that in our previous model, $\boldsymbol{W}$ represents a stochastic process $\{\boldsymbol{W}_s\}_{s=1,2}$, whereas $\boldsymbol{Y}$ represents the same stochastic process truncated at the first stage, that is, $\boldsymbol{Y} = \boldsymbol{W}_1$.[5] The finite set $\mathbb{W}_N$ is represented by N nodes: each such node carries a *pair* of values (w_1^i, w_2^i) for $i = 1, \ldots, N$. The finite set $\mathbb{Y}_N$ corresponds to the discrete representation of the truncated process $\boldsymbol{W}_1$. It is also represented by a finite set of nodes corresponding to the distinct values found in $\mathbb{Y}_N$: the cardinality of this set is M, which is less than or equal to N; each node carries a value y^j for $j = 1, \ldots, M$. Now, condition (6.8) says two things:

1. the set of N nodes at the second stage is partitioned into M disjoint subsets, each subset being in relation with a node at the first stage: this is the translation of $\boldsymbol{Y}_N \preceq \boldsymbol{W}_N$; otherwise stated, there exists a mapping $\mathfrak{f}$ from $\{1, \ldots, N\}$ to $\{1, \ldots, M\}$ which defines the *preceding* node of each leaf in the tree;
2. moreover, $y^{\mathfrak{f}(i)} = w_1^i$ for $i = 1, \ldots, N$ according to (6.8); as a consequence, it is not necessary to attach a pair of values (w_1^i, w_2^i) to leaf i but attaching w_2^i only is enough since w_1^i can already be read on the preceding node $\mathfrak{f}(i)$ of leaf i as the value $y^{\mathfrak{f}(i)}$.

It should be noticed that while the former item above involves only the encoding parts of the quantizations $Q_{\mathbb{W}}$ and $Q_{\mathbb{Y}}$ (see §6.1.1) which determine the *topology* of the tree, the latter also involves the decoding parts of those quantizations, that is the *numerical values of samples* attached to nodes.

Equation (6.11) claims that $h \circ Q_{\mathbb{W}} \preceq h$. In a representation such as Figures 1.2 or 1.3, in which realizations of $\boldsymbol{W}$ are represented as dots in a square, whereas $\boldsymbol{Y} = h(\boldsymbol{W})$ is the corresponding abscissas of those dots, the previous measurability condition says that if two dots are aligned vertically ($h(w) = h(w')$), then their quantized representations are also aligned vertically ($h \circ Q_{\mathbb{W}}(w) = h \circ Q_{\mathbb{W}}(w')$). This only leaves room for quantizations of $\boldsymbol{W}$ which look like that of the left-hand side of Figure 6.6, with groups of samples aligned vertically and corresponding cells also lined up vertically. The right-hand side of the figure depicts the corresponding stochastic tree. Figure 6.7 shows the Voronoi tessellation that would correspond to the same sample set $\mathbb{W}_N$, but this is not permitted in Pennanen's approach.

[5] The following explanation can then be easily extended by considering $\boldsymbol{W} = \{\boldsymbol{W}_s\}_{s=1,\ldots,T}$ and $\boldsymbol{Y} = \{\boldsymbol{W}_s\}_{s=1,\ldots,t}$ for any intermediate $t < T$. The truncation operator (which retains only the *prefix* of the process up to t) stands for the observation function h of the general theory.

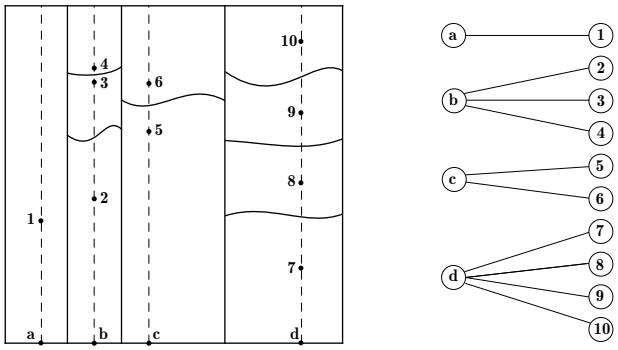

Fig. 6.6. Noise quantization that leads to a stochastic tree

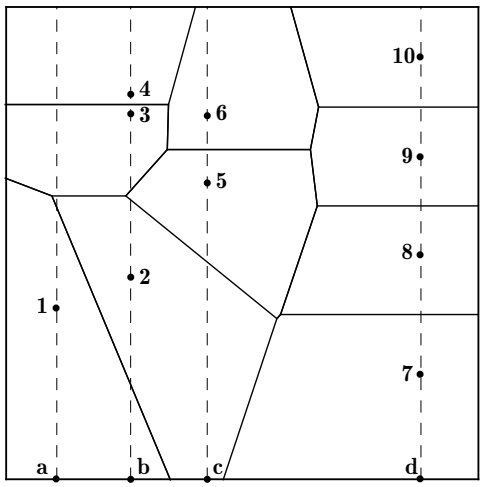

Fig. 6.7. Voronoi tessellation corresponding to the samples of Figure 6.6

6.2.3 A Constructive Proposal

In this subsection, the formulation (6.1b) is considered anew, but $\boldsymbol{Y}$ may or may not be a function h of $\boldsymbol{W}$. In the latter case, a possible choice of Ω is $\mathbb{W} \times \mathbb{Y}$; in the former case, one can again choose $\Omega = \mathbb{W}$.

Remark 6.10. The following observation was already mentioned in Remark 1.4. There is no fundamental difference between the situation when $\boldsymbol{Y}$ is a function of $\boldsymbol{W}$ and the situation when it is not. Indeed, in the latter case, one can redefine the exogeneous noise as the pair $(\boldsymbol{Y}, \boldsymbol{W})$ (this is the new $\boldsymbol{W}$) and then, h is just the linear operator which extracts the first component of this vector (whereas the cost function depends only on the second component of this new $\boldsymbol{W}$).

It is more fundamental to realize that an *optimal* quantization of a pair of random variables $(\boldsymbol{Y}, \boldsymbol{W})$ is generally *not* the Cartesian product of the two

optimal quantizations of $\boldsymbol{W}$ and $\boldsymbol{Y}$ obtained separately (using their marginal probability laws), *even if they are independent random variables.*[6] ◊

First Version

The main departure from Pennanen's approach is that, now, irrespective of the fact that $\boldsymbol{Y}$ is, or is not, a function h of $\boldsymbol{W}$, these two random variables are quantized independently. Therefore, we introduce:

- $q_{\mathbb{W}} : \mathbb{W} \to \mathbb{W}_N$ where $\mathbb{W}_N$ is a finite subset of cardinality N of $\mathbb{W}$; we require that $q_{\mathbb{W}}(w) = w$ whenever $w \in \mathbb{W}_N$;
- $\iota_{\mathbb{W}}$, the canonical injection of $\mathbb{W}_N$ into $\mathbb{W}$; then,
  - $q_{\mathbb{W}} \circ \iota_{\mathbb{W}}$ is the identity $\mathrm{I}_{\mathbb{W}_N}$ in $\mathbb{W}_N$;
  - $Q_{\mathbb{W}} = \iota_{\mathbb{W}} \circ q_{\mathbb{W}}$ is a quantization and $Q_{\mathbb{W}} \circ \boldsymbol{W}$ is the quantized noise $\boldsymbol{W}_N$;
  - $\mathbb{W}/q_{\mathbb{W}} = \mathbb{W}/Q_{\mathbb{W}}$ is a partition of $\mathbb{W}$ into N cells;
- $q_{\mathbb{Y}} : \mathbb{Y} \to \mathbb{Y}_M$ where $\mathbb{Y}_M$ is a finite subset of cardinality M of $\mathbb{Y}$; we require that $q_{\mathbb{Y}}(y) = y$ whenever $y \in \mathbb{Y}_M$;
- $\iota_{\mathbb{Y}}$, the canonical injection of $\mathbb{Y}_M$ into $\mathbb{Y}$; then,
  - $q_{\mathbb{Y}} \circ \iota_{\mathbb{Y}}$ is the identity $\mathrm{I}_{\mathbb{Y}_M}$ in $\mathbb{Y}_M$;
  - $Q_{\mathbb{Y}} = \iota_{\mathbb{Y}} \circ q_{\mathbb{Y}}$ is a quantization and $Q_{\mathbb{Y}} \circ \boldsymbol{Y}$ is the quantized observation $\boldsymbol{Y}_M$;
  - $\mathbb{Y}/q_{\mathbb{Y}} = \mathbb{Y}/Q_{\mathbb{Y}}$ is a partition of $\mathbb{Y}$ into M cells.

Then, in the discretized problem, the decision variable $\boldsymbol{U}$ is subject to the constraint $\boldsymbol{U} \preceq \boldsymbol{Y}_M$, that is, there exists a feedback $\gamma_M : \mathbb{Y}_M \to \mathbb{U}$ such that $\boldsymbol{U} = \gamma_M(\boldsymbol{Y}_M) = \gamma_M \circ q_{\mathbb{Y}} \circ \boldsymbol{Y}$. Of course, this constraint automatically produces an admissible solution for the original problem. The situation is illustrated by Figure 6.8 (compare with Figure 6.5).

At this stage, since there is no connection between $\boldsymbol{W}_N$ and $\boldsymbol{Y}_M$ (even if there is one between $\boldsymbol{W}$ and $\boldsymbol{Y}$), the appropriate Ω to consider in the discretized problem is $\mathbb{Y}_M \times \mathbb{W}_N$ — this finite set has a maximum of MN elements — with the probability law transported from the original Ω to $\mathbb{Y} \times \mathbb{W}$ by the mapping $(q_{\mathbb{Y}}, q_{\mathbb{W}})$.

To make things more concrete, we consider an example.

Example 6.11. We remain in the context of Example 6.9, with a two-dimensional $\boldsymbol{W}$ and with $\boldsymbol{Y}$ being the first coordinate $\boldsymbol{W}_1$; we use the same representation as in Figure 6.6. The left-hand part of Figure 6.9 represents the quantization of $\boldsymbol{W}$ (with $N = 8$). The middle part of that figure represents the quantization of $\boldsymbol{Y}$ on the x-axis (with $M = 5$). Since $\boldsymbol{W}$ and $\boldsymbol{Y}$ are not

[6] Considering two *scalar* random variables with *uniform* distributions over bounded intervals, it can be checked that for the same number of cells, a pavement of a large surface in the plane with hexagons is more efficient in terms of the criterion (6.3) than a pavement with squares. The former cannot obviously be obtained as the Cartesian product of two one-dimensional quantizations.

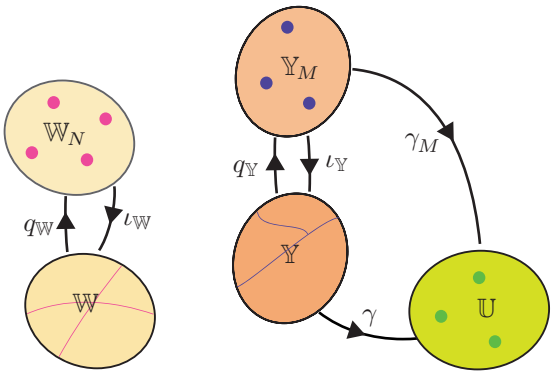

Fig. 6.8. Independent quantization of $\boldsymbol{W}$ and $\boldsymbol{Y}$

independent variables here, all the MN combinations of w^i with y^k are not possible, that is, the probability law transported from $\mathbb{W} \times \mathbb{Y}$ to $\mathbb{W}_N \times \mathbb{Y}_M$ by $(q_{\mathbb{W}}, q_{\mathbb{Y}})$ has only 21 non-zero atoms (out of 40): these are the probability masses of the cells depicted in the right-hand side of Figure 6.9. An alternative representation is that of Figure 6.10 which depicts all possible pairs of

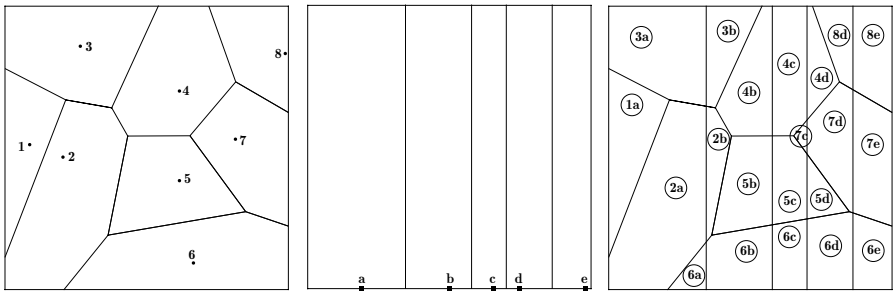

Fig. 6.9. Independent quantizations of $\boldsymbol{W}$ and $\boldsymbol{Y}$: an example

realizations of $(\boldsymbol{Y}_M, \boldsymbol{W}_N)$ in the discrete model: contrary to the situation of Figure 6.6, there is no longer a tree structure involved now. The formulation of the discretized problem is as follows:

$$\min_{\{u^k\}} \sum_{k \in \{\mathbf{a},\ldots,\mathbf{e}\}} \sum_{i=1}^{8} p^{ik}\, j(u^k, w^i) \tag{6.14}$$

in which p^{ik} is the probability weight of the cell ik ($i = 1, \ldots, 8$ and $k \in \{\mathbf{a}, \ldots, \mathbf{e}\}$) in the right-hand side of Figure 6.9. Again, only 21 of those p^{ik} are not zero, but any approximation of the probability masses of the cells that would converge asymptotically to the true values as N and M go to infinity would also be acceptable. △

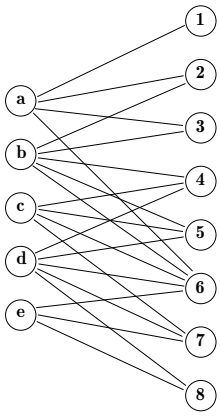

Fig. 6.10. Possible pairs of observations and noises

Remark 6.12. Contrary to the scheme inspired by Pennanen's work described in §6.2.2, it should be clear that only the *encoding* part of $Q_{\mathbb{Y}}$ matters here. That is, only the cells on the horizontal axis in the middle part of Figure 6.9 are important, not the precise values taken by y^k for $k \in \{\mathbf{a}, \ldots, \mathbf{e}\}$. ◊

Remark 6.13. In application of Remark 6.3 to (6.14), for $k \in \{\mathbf{a}, \ldots, \mathbf{e}\}$, the approximation of $\mathbb{E}(j(\boldsymbol{U}, \boldsymbol{W}) \mid \boldsymbol{Y} = y^k)$ in the discrete problem is given by the expression

$$\frac{1}{\sum_{i \in I(k)} p^{ik}} \Big(\sum_{i \in I(k)} p^{ik}\, j(u^k, w^i) \Big) , \tag{6.15}$$

where $I(k)$ is the subset of $\{1, \ldots, 8\}$ such that $p^{ik} \neq 0$. Therefore, not only each subset $I(k)$ must be non empty, but its cardinality should asymptotically go to infinity. That is, each vertical strip in Figure 6.9 should intersect asymptotically an infinite number of Voronoi cells: generically, this should be the case when the y^k's and w^i's are sampled independently and when their numbers go to infinity. But this is not the case in the situation illustrated by Figure 1.3 where the cardinality of each $I(k)$ remains equal to 1 asymptotically, even when the number of samples went to infinity (see also §8.5.4 for a related discussion). ◊

Second Version

In Example 6.11, there exists a mapping h such that $\boldsymbol{Y} = h(\boldsymbol{W})$ (namely $\boldsymbol{Y} = \boldsymbol{W}_1$), but there is none which relates the quantized observation $\boldsymbol{Y}_M$ to the quantized noise. Whenever $\boldsymbol{Y} = h(\boldsymbol{W})$, there is a way to recover such a mapping at the price of redefining the quantized noise.

As shown in Example 6.11 (see also the right-hand side of Figure 6.9), as long as $\boldsymbol{U} \preceq \boldsymbol{Y}_M$, the minimal partition of Ω generating a partition field

with respect to which the random variable $(\boldsymbol{U}, \boldsymbol{W}_N)$, hence also $j(\boldsymbol{U}, \boldsymbol{W}_N)$, becomes measurable is that associated with $\boldsymbol{W}_N \vee \boldsymbol{Y}_M = Q_{\mathbb{W}} \circ \boldsymbol{W} \vee Q_{\mathbb{Y}} \circ \boldsymbol{Y}$. Under the assumption that $\boldsymbol{Y} = h(\boldsymbol{W})$, this may be considered a partition of $\mathbb{W}$ (that is, $\boldsymbol{W}_N \preceq \boldsymbol{W}$ and $\boldsymbol{Y}_M \preceq \boldsymbol{Y} \preceq \boldsymbol{W}$; according to Proposition 3.7 and Figure 3.2, the partition on $\mathbb{W}$ is obtained by superposing the previous partition defined by $Q_{\mathbb{W}}$ and that brought back from $\mathbb{Y}$ to $\mathbb{W}$ by the generally multi-valued mapping h^{-1} — see the space $\mathbb{W}$ in the lower left-hand side corner of Figure 6.11).

So, we introduce a new quantized noise $Q'_{\mathbb{W}}(\boldsymbol{W})$, such that $Q'_{\mathbb{W}}(\boldsymbol{W}) \equiv Q_{\mathbb{W}} \circ \boldsymbol{W} \vee Q_{\mathbb{Y}} \circ \boldsymbol{Y}$ (the symbol $\equiv$ is to be taken in the sense of Proposition 3.41). Since the encoding part is already defined, in order to complete the definition of $Q'_{\mathbb{W}}$, it remains to define the decoding part, which amounts to choosing centroids in the cells depicted in the right-hand side of Figure 6.9. The new quantized noise $Q'_{\mathbb{W}}(\boldsymbol{W})$ is denoted $\boldsymbol{W}_{NM}$, but $N \times M$ is just an upper bound of the cardinality of the new discrete noise set. Notice that

- $\boldsymbol{W}_N \preceq \boldsymbol{W}_{NM} \equiv \boldsymbol{W}_N \vee \boldsymbol{Y}_M$ (the new quantized noise is "finer" that the previous one);
- $\boldsymbol{Y}_M = Q_{\mathbb{Y}} \circ \boldsymbol{Y} \preceq Q_{\mathbb{W}} \circ \boldsymbol{W} \vee Q_{\mathbb{Y}} \circ \boldsymbol{Y} \equiv Q'_{\mathbb{W}} \circ \boldsymbol{W}$, that is, $\boldsymbol{Y}_M \preceq \boldsymbol{W}_{NM}$, hence, by Proposition 3.46, there exists $h_{NM} : \mathbb{Y}_M \to \mathbb{W}_{NM}$ such that $Y_M = h_{NM}(\boldsymbol{W}_{NM})$.

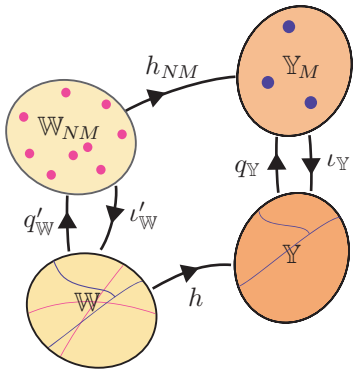

Fig. 6.11. A refined noise quantization

Therefore, we are able to express the quantized observation $\boldsymbol{Y}_M$ as a function h_{NM} of this finer quantized noise $\boldsymbol{W}_{NM}$. In Figure 6.11, $Q'_{\mathbb{W}}$ is the composition $\iota'_{\mathbb{W}} \circ q'_{\mathbb{W}}$, where $q'_{\mathbb{W}} : \mathbb{W} \to \mathbb{W}_{NM}$ is such that $q'_{\mathbb{W}} \equiv Q'_{\mathbb{W}}$ and $\iota'_{\mathbb{W}}$ is the canonical injection of $\mathbb{W}_{NM}$ into $\mathbb{W}$. By definition of h_{NM}, we have that

$$Q_{\mathbb{Y}} \circ h = \iota_{\mathbb{Y}} \circ h_{NM} \circ q'_{\mathbb{W}} \,. \tag{6.16}$$

This is similar to (6.13) and, as in Lemma 6.8, it can be proved that (6.16) is equivalent to

$$h_{NM} = q_{\mathbb{Y}} \circ h \circ \iota'_{\mathbb{W}} \ . \tag{6.17}$$

Discussion

The connection between the original observation function h and that of the discrete model h_{NM} is weaker than it was in Pennanen's approach: essentially, we have nothing similar to (6.8) here.

Returning to Example 6.11, in order to completely define $Q'_{\mathbb{W}}$, we must draw a dot in each cell of the right-hand side part of Figure 6.9. Those dots represent 2-dimensional vectors w^{ik} (namely, the dots with coordinates (w_1^{ik}, w_2^{ik})) with $i \in \{1, \ldots, 8\}$ and $k \in \{\mathbf{a}, \ldots, \mathbf{e}\}$, but not all pairs out of this cartesian product are present (only 21 out of 40). This collection of vectors $\{w^{ik}\}$ is the set $\mathbb{W}_{NM}$. According to (6.17), $h_{NM} : \mathbb{W}_{NM} \to \mathbb{Y}_M$ is such that $h_{NM}(w^{ik}) = y^k$ with $k \in \{\mathbf{a}, \ldots, \mathbf{e}\}$. But obviously, the precise values of those y^k play no particular role: what matters is the partition of $\mathbb{Y}$ generated by $\boldsymbol{Y}_M$ (the encoding part), the codebook $\{y^k\}_{k \in \{\mathbf{a}, \ldots, \mathbf{e}\}}$ is not relevant (the only constraint being that each dot belongs to its corresponding cell).

Therefore, since the codebooks of quantized noises and observations are somewhat flexible, we may use this flexibility to try to get closer to Pennanen's scheme. Graphically, we may try to move the dots y^k within their cells in $\mathbb{Y}$, and simultaneously choose dots $w^1, \ldots, w^{21}$ within the cells $1\mathbf{a}, \ldots, 8\mathbf{e}$ in the right-hand side of Figure 6.9 so that each dot representing a quantized observation be vertically aligned with a subset of the dots representing quantized noises as shown in Figure 6.6. Observe that this is not necessarily possible: if we restrict our attention to the vertical strip labelled **a**, there is *no* vertical line in this strip that crosses simultaneously the cells 1**a** and 6**a**.

Mathematically, the issue is that of choosing $Q_{\mathbb{Y}}$ and $Q_{\mathbb{W}}$ (in which only the encoding parts are important), so that it becomes possible, a posteriori, to choose the decoding parts of $Q_{\mathbb{Y}}$ and of $Q'_{\mathbb{W}}$ (with the constraint that $Q'_{\mathbb{W}} \equiv Q_{\mathbb{W}} \vee Q_{\mathbb{Y}} \circ h$) in such a way that (compare to (6.8))

$$h \circ Q'_{\mathbb{W}} = \iota_{\mathbb{Y}} \circ h_{NM} \circ q'_{\mathbb{W}} \tag{6.18a}$$

$$= Q_{\mathbb{Y}} \circ h \ , \tag{6.18b}$$

the latter equation using (6.16).

At this moment, we do have that

$$Q_{\mathbb{Y}} \circ h \circ Q'_{\mathbb{W}} = Q_{\mathbb{Y}} \circ h \ , \tag{6.19}$$

which is a weaker property than (6.18): indeed, by composing (6.16) with $Q'_{\mathbb{W}} = \iota'_{\mathbb{W}} \circ q'_{\mathbb{W}}$ to the right, (6.19) is derived. But we do not know of a constructive method to ensure (6.18) itself.

In the next section, we briefly sketch the procedure proposed by Pennanen [110] in the particular case when informational constraints reduce to non anticipativity constraints. This procedure can be related to a special Monte

Carlo technique for approximating the expectation of a function of several *independent* random variables. However, we are going to show that this special scheme has a rather bad rate of convergence when compared with the usual Monte Carlo scheme. This is why it is important to be able to get rid of the scenario tree structure as a way to translate informational constraints in the discrete problem.

6.3 A Handicap of the Scenario Tree Approach

In this section, we explain Pennanen's technique [110] to sample noise processes (in the simplest case of *two-stage white noise* processes — see Assumption 5.9) in order to obtain scenario trees,[7] and we make the connection of this technique to a particular way of numerically estimating the expectation of a function of two independent random variables. We then show that this particular (unbiased) estimation technique is not efficient, in terms of the variance of its error, w.r.t. the classical Monte Carlo estimation technique.

6.3.1 How to Sample Noises to Get Scenario Trees

As illustrated by Figure 6.6, scenario trees for two-stage stochastic processes are related to the fact that dots in a two-dimensional space, which represent sample trajectories, are grouped in vertical clusters. That is, there are several trajectories which share common first-stage values. As discussed earlier, the probability that this occurs naturally for continuous-valued stochastic processes is zero when trajectories are generated by pseudo-random Monte Carlo sampling using the probability law of the process. Therefore, the scenario tree structure can be obtained

- either by manipulating a bunch of Monte Carlo samples (or scenarios recorded in the real life) in order to force the tree structure,[8] — but, then, the original sample set must be altered in a way which is not necessarily respectful of the underlying probability law;
- or by making use of special sampling procedures (assuming the underlying probability distribution is known). The latter option is proposed by Pennanen in [110]. We now give a sketch of this idea.

If the stochastic process $\boldsymbol{W}$ is a "white noise", that is, the random variables $\boldsymbol{W}_t$, $\boldsymbol{W}_{t+1}$, ..., are all independent, then the procedure amounts to

- drawing N_0 sample values $w_0^i, i = 1, \ldots, N_0$, of $\boldsymbol{W}_0$ according to the probability distribution of this random variable;

[7] Other references dealing with scenario tree generation will be mentioned at §7.4.1.
[8] Optimal quantization (see §6.1.2) may be used to that purpose and many authors proposed various techniques to build up such trees — see e.g. [113, 60, 70, 12].

- for each such w_0^i, obtaining N_1 sample values $w_1^{ij}, j = 1, \ldots, N_1$, by Monte Carlo sampling according to the distribution of $\boldsymbol{W}_1$, and associating them with that value w_0^i to form two-stage sample trajectories (w_0^i, w_1^{ij}) (thus, there are $N_0 \times N_1$ such trajectories);
- repeating this process over the whole time horizon.

Observe that the clusters of N_1 values w_1^{ij} associated with the w_0^i's may be all identical (in which case, the notation w_1^{ij} can be reduced to w_1^j) or different. We examine later on what is the impact of either choice.

If the stochastic process $\boldsymbol{W}$ is not a white noise, Pennanen assumes that it can be modelled by a recurrent dynamic equation driven by a white noise. Then, the above procedure is used for the driving white noise, and the sample noise trajectories are then obtained by propagating these trajectories through the dynamic equation.

In the rest of this section, to keep things simple, we limit ourselves to the discussion of white noise processes.

6.3.2 Variance Analysis

As discussed throughout this chapter, the discretization of stochastic optimization problems with SIS involves some sort of noise sampling as well as the sound translation of informational constraints in the discrete problem formulation. So far in this chapter, we have given attention to the latter aspect. But it should be clear that the quality of approximation of the mathematical expectations (and *conditional* mathematical expectations, as underlined in Remark 6.3) involved in the problem is also important. If the cost function is badly approximated, one cannot expect that a good approximation of the optimal solution can be derived from the discrete problem solution, whatever care is exercised about the other aspects (in particular, informational constraints) of the problem.

In this subsection, we concentrate on this aspect of the approximation: more precisely, we consider any real-valued function f of two scalar variables, such that the mathematical expectation $\mathbb{E}\big(f(\boldsymbol{X}, \boldsymbol{Y})\big)$, where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are *independent* random variables, makes sense. In relation with the previous discussion, f should be thought of as the cost function. Respectively, $\boldsymbol{X}$ and $\boldsymbol{Y}$, should be interpreted as $\boldsymbol{W}_0$ and $\boldsymbol{W}_1$, the two first stages of a white noise stochastic process. No decision variable appears here since we forget about optimization to pay attention to the quality of approximation of the expectation. The discussion is limited to two-time stages only, but the generalization to several time stages should be straightforward.

In a standard Monte Carlo procedure, N sample values (x^i, y^i) of the pair of random variables $(\boldsymbol{X}, \boldsymbol{Y})$ are generated according to their joint probability distribution, or they have been recorded from real data. An unbiased estimate of $\mathbb{E}\big(f(\boldsymbol{X}, \boldsymbol{Y})\big)$ (which is denoted simply $\mathbb{E}f$ for short) is provided by the arithmetic mean

$$\frac{1}{N}\sum_{i=1}^{N} f(x^i, y^i)\ . \tag{6.20}$$

It is well known that the variance of this estimate is an $O(1/N)$ when N is the number of samples.[9]

In Pennanen's procedure described in §6.3.1, N_x samples are generated for $\boldsymbol{X}$. With each such sample value x^i, a group of samples y^j with $j \in J(i)$, is associated: these samples are also generated by Monte Carlo sampling according to the probability distribution of $\boldsymbol{Y}$. To make things simpler (but this is not essential), we assume that all such index sets $J(i)$ have the same cardinality N_y. Moreover, as suggested earlier, there are two options to consider:

option (a): N_x sample groups of cardinality N_y are generated independently;
option (b): the same group $\{y^j\}_{j\in J}$ is associated with all samples x^i.

Pictorially, those options are illustrated by Figure 6.12 (with $N_x = N_y = 3$). In both cases, overall, $N_x \times N_y$ samples are used to produce the following

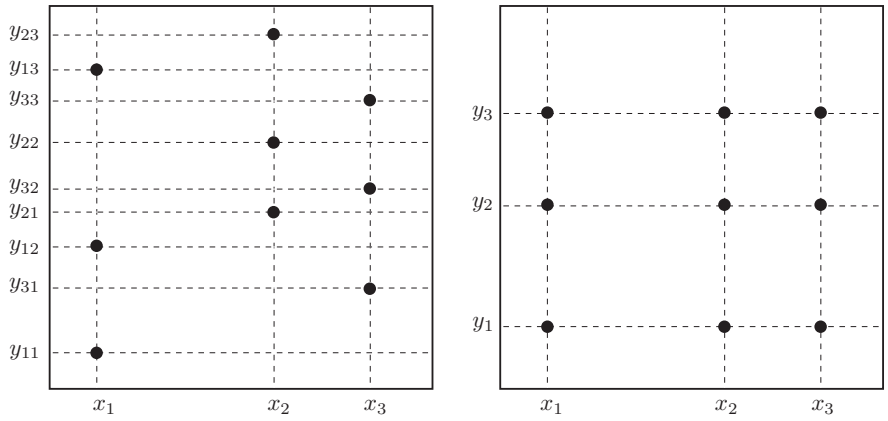

Fig. 6.12. Two ways of sampling to get scenario trees (option (a) left-hand side and option (b) right-hand side)

estimate of $\mathbb{E}f$:

$$\frac{1}{N_x \times N_y}\sum_{i=1}^{N_x}\sum_{j\in J(i)} f(x^i, y^j)\ , \tag{6.21}$$

where $J(i)$ is indeed independent of i in option (b) (that is, $J(i) = \{1, \ldots, N_y\}$ for all i), whereas, in option (a), the $J(i)$'s should be viewed as disjoint subsets of indices to translate the fact that the N_x groups $\{y^j\}_{j\in J(i)}$ have been sampled independently.

[9] For the notation O, see footnote 3 in Chapter 2.

Clearly, the estimate (6.21) of $\mathbb{E}f$ is also unbiased, and if we want to compare it with (6.20) from the point of view of its variance, we must assume that $N = N_x \times N_y$. We first study option (b). Recall that, in this option, all subsets $J(i)$ involved in (6.21) coincide with $\{1, \ldots, N_y\}$.

Proposition 6.14. *Given the value of $N = N_x \times N_y$, with option (b), the variance of estimate (6.21) is minimal when $N_y = N_x$ and this variance is of order $\mathrm{O}(1/N_x)$; therefore the variance is of order $\mathrm{O}(1/\sqrt{N})$.*[10]

Proof. Let $J(i) = \{1, \ldots, N_y\}$ for all i. The variance of the estimate (6.21) is

$$\sigma^2 = \mathbb{E}\left(\left(\frac{1}{N_x N_y}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\big(f(\boldsymbol{X}^i, \boldsymbol{Y}^j) - \mathbb{E}f\big)\right) \times \left(\frac{1}{N_x N_y}\sum_{k=1}^{N_x}\sum_{l=1}^{N_y}\big(f(\boldsymbol{X}^k, \boldsymbol{Y}^l) - \mathbb{E}f\big)\right)\right). \quad (6.22)$$

In this expression, the outer expectation is with respect to the probability distributions of independent random variables $\{\boldsymbol{X}^i\}_{i=1,\ldots,N_x}$, $\{\boldsymbol{X}^k\}_{k=1,\ldots,N_x}$, $\{\boldsymbol{Y}^j\}_{j=1,\ldots,N_y}$ and $\{\boldsymbol{Y}^l\}_{l=1,\ldots,N_y}$ (replicating $\boldsymbol{X}$ and $\boldsymbol{Y}$), the realizations of which are the samples x^i, x^k, y^j and y^l used in the estimate.

If the expression (6.21) is expanded, this yields $N_x^2 N_y^2$ products of random variables (with zero mean) of the type

$$\frac{1}{N_x^2 N_y^2}\mathbb{E}\Big(\big(f(\boldsymbol{X}^i, \boldsymbol{Y}^j) - \mathbb{E}f\big)\big(f(\boldsymbol{X}^k, \boldsymbol{Y}^l) - \mathbb{E}f\big)\Big). \quad (6.23)$$

We split up this set of products into four subsets:

1. the subset for which $i = k$ and $j = l$, of cardinality $N_x N_y$; for this subset, all products of the type (6.23) are squares and their sum contributes to σ^2 (in (6.22)) for a term of order $\mathrm{O}(1/N_x N_y)$;
2. the subset for which $i \neq k$ and $j \neq l$: the cardinality of this subset is $N_x(N_x - 1)N_y(N_y - 1)$; since this subset contains only products of random variables of the type (6.23) which have zero mean and are mutually independent, it contributes for 0 to (6.22);
3. the subset for which $i = k$ but $j \neq l$, of cardinality $N_x N_y(N_y - 1)$ that we study later on;
4. symmetrically, the subset for which $i \neq k$ but $j = l$ of cardinality $N_x(N_x - 1)N_y$.

It can be checked that the sum of cardinalities of the four subsets is equal to $N_x^2 N_y^2$, as it must be.

[10] The authors are indebted to Prof. Benjamin Jourdain for preliminary results in this direction.

It remains to study the contribution of products in the third and fourth subsets. Those products involve either the same $\boldsymbol{X}^i$ but different $\boldsymbol{Y}^j$ and $\boldsymbol{Y}^l$, or symmetrically, the same $\boldsymbol{Y}^j$, but different $\boldsymbol{X}^i$ and $\boldsymbol{X}^k$. We only study the former subset since conclusions also apply to the latter by symmetry. We first prove that all terms of the type (6.23) such that $\boldsymbol{X}^i$ is the same but $\boldsymbol{Y}^j$ is independent of $\boldsymbol{Y}^l$ have nonnegative expectations. Indeed, with the short-hand notation

$$\mathbb{E}^{\boldsymbol{X}} f := \mathbb{E}\big(f(\boldsymbol{X},\boldsymbol{Y}) \;\big|\; \boldsymbol{X}\big)\,,$$

one has that

$$\mathbb{E}\Big(\big(f(\boldsymbol{X}^i,\boldsymbol{Y}^j)-\mathbb{E}f\big)\big(f(\boldsymbol{X}^i,\boldsymbol{Y}^l)-\mathbb{E}f\big)\Big) = \mathbb{E}\Bigg(\mathbb{E}^{\boldsymbol{X}^i}\Big(\big(\underbrace{f(\boldsymbol{X}^i,\boldsymbol{Y}^j)-\mathbb{E}^{\boldsymbol{X}^i}f}_{\boldsymbol{B}^j}+\underbrace{\mathbb{E}^{\boldsymbol{X}^i}f-\mathbb{E}f}_{\boldsymbol{C}}\big) \times \big(\underbrace{f(\boldsymbol{X}^i,\boldsymbol{Y}^l)-\mathbb{E}^{\boldsymbol{X}^i}f}_{\boldsymbol{B}^l}+\underbrace{\mathbb{E}^{\boldsymbol{X}^i}f-\mathbb{E}f}_{\boldsymbol{C}}\big)\Big)\Bigg)\,.$$

The independence of $\boldsymbol{Y}^j$ and $\boldsymbol{Y}^l$, and therefore of $\boldsymbol{B}^j$ and $\boldsymbol{B}^l$, and the fact that the latter variables have zero conditional means, imply that $\mathbb{E}^{\boldsymbol{X}^i}\big(\boldsymbol{B}^j\boldsymbol{B}^l\big)=0$. Moreover, since $\boldsymbol{C}$ is $\boldsymbol{X}^i$-measurable,

$$\mathbb{E}^{\boldsymbol{X}^i}(\boldsymbol{B}^j\boldsymbol{C}) = \mathbb{E}^{\boldsymbol{X}^i}(\boldsymbol{B}^j)\times\boldsymbol{C} = 0\,.$$

The same applies to $\mathbb{E}^{\boldsymbol{X}^i}(\boldsymbol{C}\boldsymbol{B}^l)$. The only nonzero term is thus the nonnegative term $\mathbb{E}(\boldsymbol{C}^2)$, which is the variance of $\mathbb{E}^{\boldsymbol{X}^i}f$ (generically of order $\mathrm{O}(1)$).

Finally, the terms in the third and fourth subsets above contribute all together for a nonnegative term of order $\mathrm{O}\Big(\big((N_x-1)+(N_y-1)\big)/N_xN_y\Big) \sim \mathrm{O}(1/N_x+1/N_y)$. This contribution is added to that of the first subset which was $\mathrm{O}(1/N_xN_y)$. For the comparison of (6.20) and (6.21), we assume that $N=N_xN_y$: for N given, the variance of the estimate (6.21) is minimal when $N_x=N_y=\sqrt{N}$, and this variance is of order $\mathrm{O}(1/\sqrt{N})$, to be compared with $\mathrm{O}(1/N)$ of the standard Monte Carlo estimate (6.20). □

It is easy to figure out how this result extends to the case of T stages instead of 2: the variance of the "tree" estimate (6.21) is of order $\mathrm{O}(1/\sqrt[T]{N})$ instead of $\mathrm{O}(1/N)$ for (6.20). Needless to say, this quickly becomes a dramatic loss of quality of the tree estimate as T keeps growing.

Consider now option (a). The calculations in the proof of Proposition 6.14 must be adapted in the following way. First, (6.21) is now valid with subsets $J(i)$ which must be considered as disjoint subsets (that is, the corresponding subsets of random variables $\{\boldsymbol{Y}^j\}_{j\in J(i)}$ are independent). Consequently, (6.22) must be replaced by

$$\sigma^2 = \mathbb{E}\left(\left(\frac{1}{N_x N_y}\sum_{i=1}^{N_x}\sum_{j\in J(i)}\big(f(\boldsymbol{X}^i,\boldsymbol{Y}^j)-\mathbb{E}f\big)\right)\right.$$
$$\left.\times\left(\frac{1}{N_x N_y}\sum_{k=1}^{N_x}\sum_{l\in J(k)}\big(f(\boldsymbol{X}^k,\boldsymbol{Y}^l)-\mathbb{E}f\big)\right)\right). \quad (6.24)$$

Among the four subsets considered in the previous proof, the first and the third ones have the same cardinalities as earlier and contribute to σ^2 for the same amounts as in that proof, namely, $\mathrm{O}(1/N_xN_y)$ and $\mathrm{O}\big((N_y-1)/N_xN_y\big)$ respectively. Since $i \neq k$ implies that $j \neq l$ (because $J(i) \cap J(k)$ should be considered as empty), then the fourth subset (which earlier contained $N_x(N_x-1)N_y$ elements) is now empty whereas the cardinality of the second subset increases to $N_x(N_x-1)N_y^2$ (that is, the elements of the fourth subset which previously contributed for nonnegative terms are transferred to the second subset whose terms contribute for 0). Finally, with option (a), the variance of (6.21) is of order $\mathrm{O}(1/N_xN_y) + \mathrm{O}\big((N_y-1)/N_xN_y\big)$.

Under the constraint that $N = N_xN_y$ (to make the comparison with (6.20) possible), we reach the conclusion that the best (minimal) variance is obtained when $N_y = 1$ and $N_x = N$. Notice that this is no longer a tree, but indeed N independent scenarios, that is (6.20) and (6.21) actually coincide. This just says that, from only the point of view of minimizing the variance of the estimate of the cost function over the whole time horizon, the structure of N independent scenarios is far better than a tree structure. But of course, one should again recall Remark (6.3): if $N_y = 1$, the conditional expectation of the cost (knowing $\boldsymbol{W}_0$) which serves as the objective function in the minimization problem when choosing $\boldsymbol{U}_0$ is approximated with help of a *single* sample of $\boldsymbol{W}_1$, which is very bad. To avoid this, we should have put a lower bound on N_y in order to bound the variance of this conditional expectation, and in this simple case, it is clear that, with the constraint $N = N_xN_y$, the best trade-off between the variance of the estimate of the expected cost over the two-stage horizon and the variance of the estimate of the conditional expectation restrained to the second stage only is again to take $N_x = N_y = \sqrt{N}$.

The conclusion of this rough variance analysis is that the tree structure, which is one way to represent informational (or simply, non anticipativity) constraints, is not very efficient from the point of view of the variance of estimates it provides. Therefore, we should avoid the tree structure and find another way to translate the informational constraints in the discrete problem. The methodology presented in §6.2.3 suggested that this is indeed possible. A more concrete technique in the context of stochastic optimal control problems is presented in Chapter 7.

6.4 Conclusion

In this chapter, we have proposed a methodology to derive approximate finite-dimensional versions of the generic stochastic optimization problem (6.1) which involves *static* informational constraints (the so-called SIS — see §1.4). This discretization stage should not only limit the computations to finite-dimensional objects (probability measures, decision variables, etc.) but it should also translate the original informational constraints in a way which makes it possible to build up a *feasible* solution for the original problem after the discrete finite-dimensional problem has been solved. In addition, one expects that the performance of this feasible solution approaches the true optimal performance when the dimension of the approximate optimization problem goes to infinity. This convergence issue has not been explicitly considered in this chapter and is deferred to Chapter 8.

A popular technique to formulate discrete optimization problems taking care at least of non anticipativity constraints (the minimal form of informational constraints in multi-stage problems) is the so-called scenario tree technique. Pennanen [109] proposed a complete study of this technique, including the reconstruction of a feasible solution for the original problem and its asymptotic convergence to the optimal solution. We have given a description of his approach using the language of quantization presented in the first part of this chapter. We then have proposed other approaches which attempt to relax some of the constraints imposed by Pennanen's approach, and more generally by the scenario tree technique. Indeed, as explained in the end of the chapter, the convergence speed of this technique is seriously handicapped by the variance of the expectation estimates it produces, which is typically in $O(\sqrt[T]{N})$ when using N sample trajectories for a problem with T time stages. This result is also mentioned by Shapiro [138] who uses large deviation techniques to establish it.

The argument above provides motivation to eliminate the scenario tree structure and to find alternative ways to account for informational constraints. The approach described in this chapter is a first proposal in this new direction, and it must be confirmed by the convergence study presented in Chapter 8. Chapter 7 presents another approach, however, that is not directly comparable with the one introduced here because they proceed along two different paths. In this chapter, we have formulated a discretized optimization problem which attempts to mimic the original problem and we have derived the proposed solution from the resolution of this problem. In the next chapter, we consider the optimality conditions of the infinite-dimensional original problem (established in Chapter 5), and we propose discretization schemes in order to approximately solve those conditions.