Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

MARC C. STEINBACH

General Information Constraints in Stochastic Programs

ZIB-Report 01-24 (September 2001)

GENERAL INFORMATION CONSTRAINTS IN STOCHASTIC PROGRAMS

MARC C. STEINBACH

ABSTRACT. Scenario tree models of stochastic programs arise naturally under standard nonanticipativity assumptions. We demonstrate how tree-sparse programs cover the general case, with *arbitrary* information constraints. Detailed examples and intuitive interpretations illuminate the basic thoughts behind the abstract but elementary construction.

0. INTRODUCTION

Information constraints—or *nonanticipativity* conditions—play a fundamental role in decision processes under uncertainty: they model precisely when and to what extent the decision maker acquires information on the outcomes of random events. In probabilistic terms nonanticipativity is expressed by measurability requirements with respect to σ-fields that vary over time, typically forming a filtration. These concepts provide the theoretical foundation for any kind of stochastic decision model and are indispensable in situations that involve continuous distributions; see, e.g., [3]. In the narrower area of discrete-time stochastic programming models with finitely many scenarios, the theoretical background is unnecessarily general and inconvenient to work with. Nonanticipativity is either modeled by explicit equality constraints or by scenario tree representations (where it is implicit in the tree topology). Such formulations are well suited for computation; they arise naturally under the standard assumption that the acquired information is always up-to-date; cf. [10, 13, 14]. However, the more abstract probabilistic perspective proves again useful when a scenario tree formulation is ultimately desired but the random data and observed data (the information available to the decision maker) are related in a nontrivial way, as in situations involving delayed observations or foresight. General information constraints in this sense are considered by Rockafellar and Wets in [12] (where nonanticipativity is enforced by explicit projections and a probabilistic perspective is deliberately avoided), and in [11] where information σ-fields model arbitrary nonanticipativity requirements in the context of generalized linear-quadratic control [9]. These models remain on an abstract level in that no concrete representations of the random data and variables are specified.

The present paper constructs natural scenario tree representations for stochastic programs with arbitrary random data and information fields. In the process we discuss the role of standard assumptions within the generalized context and provide a couple of examples to illustrate important cases or subtle details. The presentation employs probabilistic notation and language most of the time (until a direct translation into tree notation is possible), but interpretations are given throughout. Thus, rather than establishing fundamentally new results, our emphasis is on exploring the rich possibilities arising in the general setting and giving explanations in terms of the intuitive tree concept. The problem classes used in the construction are *tree-sparse* programs as developed in [16, 17, 18]; they share important structural properties with the generalized linear-quadratic problems of [11].

The material is organized as follows. In §1 we recall the basic theoretic concepts and establish their relation to scenario trees. Data and information fields are then introduced

2000 *Mathematics Subject Classification.* 90C15, 60G99.

Key words and phrases. information constraints, general nonanticipativity, sigma-fields, filtrations, scenarios, stochastic programs, tree-sparse programs.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant STE 566/1-1.

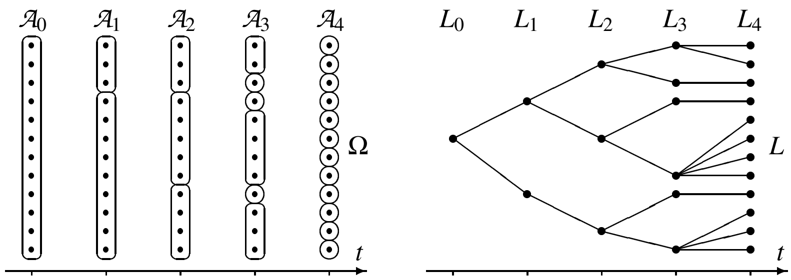

FIGURE 1. A finite filtration (left) and its associated scenario tree (right).

in §2 as the key ingredients in general stochastic decision models, and §3 presents the tree-sparse programs modeling the precise problem classes of interest. The transition from abstract problem formulations to concrete scenario tree representations is developed in §4, followed by a brief comparison to the problem class of Rockafellar and Wets [11] in §5. Some final remarks in §6 conclude the paper.

1. PRELIMINARIES

1.1. **Fields.** The fundamental probabilistic concepts for modeling the evolution of information and stochastic influences in decision processes are σ-fields and filtrations. To introduce the notation we recall basic definitions and facts that may be found in standard textbooks on measure theory and probability, such as Bauer [1, 2], Dudley [6], or Shiryaev [15].

Let Ω denote a set and 2^Ω its power set. Recall that a *σ-field* $\mathcal{F}$ in Ω is a collection of subsets of Ω that is closed under forming complements, countable unions, and countable intersections. The sets $A \in \mathcal{F}$ are called *$\mathcal{F}$-measurable*, the $\subseteq$-minimal elements $\alpha \in \mathcal{F}$ are called *$\mathcal{F}$-atoms*. A *sub-σ-field* $\mathcal{G} \subseteq \mathcal{F}$ is a subset of $\mathcal{F}$ that is also a σ-field; $\mathcal{G} \subseteq \mathcal{F}$ is *coarser* than $\mathcal{F}$ and $\mathcal{F}$ *finer* than $\mathcal{G}$. The σ-field $\sigma(\{A_i\}_{i\in I})$ *generated* by an arbitrary collection of subsets $A_i \subseteq \Omega$ is the unique coarsest σ-field containing every A_i. Likewise $\sigma(\{\mathcal{F}_i\}_{i\in I})$ denotes the unique coarsest σ-field containing every $\mathcal{F}_i$ as a sub-σ-field. The empty set and Ω itself generate the trivial σ-field, $\sigma(\emptyset) = \sigma(\Omega) = \{\emptyset, \Omega\}$.

Given a *measurable space* $(\Omega, \mathcal{F})$ and an ordered set $(I, \preceq)$, a *filtration* $\mathbb{F} = \{\mathcal{F}_t\}_{t\in I}$ in Ω (in $\mathcal{F}$) is a family of σ-fields in Ω (in $\mathcal{F}$) satisfying $\mathcal{F}_\tau \subseteq \mathcal{F}_t$ for all $\tau, t \in I$ with $\tau \preceq t$. For instance, the set of all sub-σ-fields of $\mathcal{F}$ (ordered by inclusion) is a filtration in $\mathcal{F}$.

A *probability measure* on $(\Omega, \mathcal{F})$ is a monotonous, σ-additive function $P\colon \mathcal{F} \to [0,1]$ satisfying $P(\emptyset) = 0$ and $P(\Omega) = 1$. The triple $(\Omega, \mathcal{F}, P)$ is a *probability space*. In this context $A \in \mathcal{F}$ is called an *event*, $\omega \in \Omega$ an *elementary event*, and Ω the *sample space*.

A mapping $f\colon (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ is *measurable* if $f^{-1}(A') \in \mathcal{F}$ for every $A' \in \mathcal{F}'$. A *random variable* Z on a probability space $(\Omega, \mathcal{F}, P)$ is such a measurable mapping. A *stochastic process* (or *random process*) $\xi = (\xi_t)_{t\in I}$ is a family of random variables; ξ is *adapted* to the filtration $\mathbb{F}$ if every ξ_t is $\mathcal{F}_t$-measurable. The *canonical* (*minimal*) filtration associated with ξ consists of the σ-fields $\sigma(\{\xi_\tau\}_{\tau\preceq t})$ generated by the pre-images $\xi_\tau^{-1}(A')$ of $\mathcal{F}'$-measurable sets. It is the coarsest filtration in $\mathcal{F}$ to which ξ is adapted. (Note that measurability with respect to a coarser σ-field is a *stronger* restriction.)

Every *finite* σ-field $\mathcal{F}$ (having finite cardinality $|\mathcal{F}|$) is completely characterized as $\mathcal{F} = \sigma(\mathcal{A}) = 2^{\mathcal{A}}$ where $\mathcal{A}$ denotes the (finite) set of atoms; hence $|\mathcal{F}| = 2^{|\mathcal{A}|}$ and every measurable set $A \in \mathcal{F}$ is a union of $\mathcal{F}$-atoms. Moreover, if $|\Omega|$ is finite, then every σ-field $\mathcal{F}$ corresponds to a partitioning of Ω into mutually disjoint subsets (the $\mathcal{F}$-atoms), and $\mathcal{G} \subseteq \mathcal{F}$ holds if and only if every $\mathcal{G}$-atom is a union of $\mathcal{F}$-atoms. In that case we assume $\mathcal{F} = 2^\Omega$ without loss of generality: if $\mathcal{F} \subset 2^\Omega$ (proper inclusion) then Ω contains irrelevant or unobservable sub-atomic events and should be redefined as the set of $\mathcal{F}$-atoms.

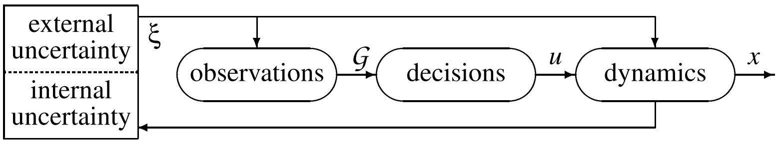

FIGURE 2. Flow of information in a stochastic dynamic decision process.

We then have $\mathcal{A} = \Omega$ and $P = \sum_{\omega\in\Omega} p_\omega \delta_\omega$ for suitable probabilities $p_\omega \geq 0$, $\sum_{\omega\in\Omega} p_\omega = 1$. A filtration (with linearly ordered I) now simply corresponds to a sequence of successively finer partitionings of Ω; see Fig. 1.

A random variable Z is $\mathcal{G}$-measurable if and only if Z is constant with respect to every $\mathcal{G}$-atom $\beta \in \mathcal{B}$. We write $Z = (Z_\beta)_{\beta\in\mathcal{B}} = (Z_\omega)_{\omega\in\Omega}$ to refer to its *realizations*; likewise $\xi = (\xi_\beta)_{\beta\in\mathcal{B}} = (\xi_{t\beta})_{t\in I,\beta\in\mathcal{B}}$ etc. for a random process. If Z takes values in a vector space, the *conditional expectation* $\mathbf{E}^{\mathcal{G}}\{Z\}$ is the $\mathcal{G}$-measurable random vector $\bar{Z}$ whose realizations are the convex combinations $\bar{Z}_\beta = \sum_{\omega\in\beta} p_\beta^{-1} p_\omega Z_\omega$ with $p_\beta = \sum_{\omega\in\beta} p_\omega$. The mapping $\mathbf{E}^{\mathcal{G}}$ defines a projection, and the usual expectation is given as the special case $\mathbf{E}^{\{\emptyset,\Omega\}} = \mathbf{E}$.

1.2. **Trees.** Abstractly a *tree* is a connected, acyclic graph. We consider only *rooted* trees where the root is predecessor of all other nodes. Let V denote the set of nodes (or *vertices*), $L_t \subseteq V$ the level set of nodes at depth t, and $L \supseteq L_T$ the set of leaves (where T is the depth of the tree). Further let $0 \in L_0$ denote the root, $j \in L_t$ the "current" node, $i \equiv \pi(j)$ its unique predecessor (if $t > 0$), $S(j)$ its set of successors, and $\Pi(j) = \{0,\dots,i,j\}$ the unique path from the root to j. Expressing father-son relations by π and S is more convenient than using the standard graph-theoretic concept of an arc set. Other authors use similar tree representations [10, 13, 14] or simply two indices (time and scenario) for every node.

In the probabilistic context, *scenario trees* (also known as *event trees* or *decision trees*) are often used as more convenient representations of finite filtrations when $\mathcal{F}_0 = \{\emptyset,\Omega\}$. Given such a filtration $\{\mathcal{F}_t\}_{t\in I}$, the tree's level sets are formally defined as the sets of atoms, $L_t := \mathcal{A}_t$, and the successors of $\alpha \in L_t$ are $S(\alpha) := \{\beta \in L_{t+1} : \beta \subseteq \alpha\}$. A path $\Pi(\alpha)$ (or sometimes α itself) is called a *stage-t scenario* for $\alpha \in L_t$, and a *scenario* if α is a leaf. Clearly $\Pi(\alpha)$ corresponds to a decreasing sequence of atoms $\Omega = \alpha_0 \supseteq \alpha_1 \supseteq \cdots \supseteq \alpha_t = \alpha$. (Thus each node is an atom; we just disregard that it is a subset of Ω and write j for α.) The scenario probabilities are $p_j > 0$, $j \in L$. All other nodes also have probabilities p_j satisfying $p_j = \sum_{k\in S(j)} p_k$. Hence $\sum_{j\in L_t} p_j = 1$ holds for all t, and $p_0 = 1$. The tree thus defined has the special properties $L = L_T$ and $\pi(j) \in L_{t-1}$ for $j \in L_t$, $t > 0$; see Fig. 1.

Conversely, given a scenario tree with the properties just mentioned, the $\mathcal{F}_T$-atoms are defined as the singleton sets $\alpha_j := \{j\}$, $j \in L_T$, and the $\mathcal{F}_t$-atoms for $t < T$ are recursively defined as the unions $\alpha_j := \bigcup_{k\in S(j)} \alpha_k$, $j \in L_t$, yielding eventually $\alpha_0 = L$. Thus we have $\Omega = L = L_T$, $\mathcal{F} = \mathcal{F}_T = 2^\Omega$, and $P(\alpha_j) = p_j$.

2. STOCHASTIC DECISION PROCESSES

Consider a sequential decision problem in discrete time over a finite *planning horizon* $I = \{0,1,\dots,T\}$ where the *states* x_t of a dynamic process are influenced via *controls* u_t, the decisions to be made. The goal is to minimize some cost subject to constraints under random influences ξ_t representing, for instance, external disturbances or process-inherent uncertainty. Decisions react to observations of "past" random events (*recourse*) while taking into account the known distribution of as yet unobserved "future" events, in a sense to be made precise below. The flow of information is illustrated in Fig. 2.

2.1. **Data and information fields.** Suppose that $\xi = (\xi_t)_{t\in I}$ is a stochastic process on some finite probability space $(\Omega,\mathcal{F},P)$. The realizations of ξ_t generate σ-fields $\mathcal{D}_t \subseteq \mathcal{F}$. We call ξ the *data process* and $\mathcal{D}_t$ the *data fields*. For simplicity of exposition suppose further that ξ_t contains all random data elements that influence the process at time t, such as matrix and vector elements in a linear stochastic program.[1] Finally let $\mathcal{D}_t^* := \sigma(\{\mathcal{D}_\tau\}_{\tau \le t})$ denote the σ-field generated by all random events up to time t, and $\mathbb{F}_d := \{\mathcal{D}_t^*\}_{t \in I}$ the associated minimal filtration in $\mathcal{F}$, which we call the *data filtration*.

The central issue of the current paper relates to the question when and to what extent the decision maker obtains information on the realizations of random events. Following Rockafellar and Wets [11], we model the information available at time t as an *information field* $\mathcal{G}_t \subseteq \mathcal{F}$ with respect to which u_t must be measurable, $u_t = \mathbf{E}^{\mathcal{G}_t}\{u_t\}$. This restriction is the general *nonanticipativity condition*. Abstractly it means that decision u_t relies on the *observation* of a $\mathcal{G}_t$-atom or, in other words, that the information available at time t permits the distinction of $\mathcal{G}_t$-atoms. Practically observations are usually made by measuring the values of certain data elements, state components, or functions thereof.

The standard setting in the stochastic programming literature assumes *full information* in the sense that the observed data are precisely the random data $\{\xi_\tau\}_{\tau \le t}$ up to the current time, and hence $\mathcal{G}_t = \mathcal{D}_t^*$ for all $t \in I$. Thus the decision process u is required to be $\mathbb{F}_d$-adapted, which implies the same property for the resulting state process x. Here we consider instead the general case of arbitrary and independent data and information fields $\mathcal{D}_t$ and $\mathcal{G}_t$. As will be seen, this covers situations where ξ_t is observed before or after time t (or perhaps never) and where the time of observation may be scenario-dependent, or even where previously acquired information may get lost.

2.2. **Dynamics.** The dynamic aspect of a discrete-time process is modeled by state transition equations. Decisions may take effect instantaneously (such as financial transactions) or continously over an extended period of time (such as heating in a chemical process). With extended action dynamic equations take the general form

$$x_t = \theta_t(x_{t-1}, u_{t-1}). \tag{1}$$

If decisions take effect instantaneously one has different states x_t^- and x_t^+ immediately before and after decision u_t, yielding separate transition equations

$$x_t^- = \vartheta_t^-(x_{t-1}^+), \qquad x_t^+ = \vartheta_t^+(x_t^-, u_t). \tag{2}$$

Depending on whether x_t is chosen to mean x_t^- or x_t^+, those combine to one of the forms

$$x_t = \theta_t^-(x_{t-1}, u_{t-1}) := \vartheta_t^-(\vartheta_{t-1}^+(x_{t-1}, u_{t-1})) \qquad (x_t \equiv x_t^-), \tag{3}$$

$$x_t = \theta_t^+(x_{t-1}, u_t) \quad := \vartheta_t^+(\vartheta_t^-(x_{t-1}), u_t) \qquad (x_t \equiv x_t^+). \tag{4}$$

The choice is a matter of taste unless one alternative offers an advantage in numerical computations. Conceptually we prefer the latter form (which will be used throughout).

Uncertainty may enter all state transitions above, that is, $\theta_t, \vartheta_t^\pm, \theta_t^\pm$ are elements of ξ_t with (discrete) values in suitable function spaces. The precise meaning of the time index on u_t, ξ_t is therefore that these quantities enter the process at time t, either through dynamics, objective, or constraints. This is what we mean by saying that a decision u_t is "made" at time t or that an event ξ_t "happens" at time t.

3. Stochastic Programs

In the following we focus on *stochastic programs*, where an *expected* cost is to be minimized. We assume that the objective is convex and smooth and that the constraints are linear (polyhedral). Two closely related problem classes are considered; they correspond to the *tree-sparse programs* of [18] but are stated in probabilistic notation with abstract feasible sets, like the extended linear-quadratic control problems of Rockafellar and Wets [11].

[1] In practice one usually prefers *factor models* where actual random parameters ξ_t have just a few components and all problem data are expressed as functions of ξ_t. The dependence is typically affine to preserve convexity.

The *outgoing control* variant has dynamic equations in the form (1), (3), reading

$$\min_{x,u} \quad \sum_{t=0}^{T} \mathbf{E}\{\varphi_t(x_t,u_t)\} \tag{5}$$

$$\text{s.t.} \quad x_t = G_t x_{t-1} + E_t u_{t-1} + h_t \qquad \forall t = 0,\dots,T, \tag{6}$$

$$(x_t,u_t) \in (X_t \times U_t) \cap Y_t \qquad \forall t = 0,\dots,T, \tag{7}$$

$$\mathbf{E}^{\mathcal{G}_t}\{u_t\} - u_t = 0 \qquad \forall t = 0,\dots,T, \tag{8}$$

$$\sum_{t=0}^{T} \mathbf{E}\{F_t x_t + D_t u_t + e_t\} = 0, \tag{9}$$

whereas the *incoming control* variant has dynamic equations in the form (4), reading

$$\min_{u,x} \quad \sum_{t=0}^{T} \mathbf{E}\{\varphi_{t-1,t}(x_{t-1},u_t) + \varphi_t(x_t)\} \tag{10}$$

$$\text{s.t.} \quad x_t = G_t x_{t-1} + E_t u_t + h_t \qquad \forall t = 0,\dots,T, \tag{11}$$

$$(x_{t-1},u_t) \in Y_{t-1,t} \qquad \forall t = 0,\dots,T, \tag{12}$$

$$(u_t,x_t) \in U_t \times X_t \qquad \forall t = 0,\dots,T, \tag{13}$$

$$\mathbf{E}^{\mathcal{G}_t}\{u_t\} - u_t = 0 \qquad \forall t = 0,\dots,T, \tag{14}$$

$$\sum_{t=0}^{T} \mathbf{E}\{D_t u_t + F_t x_t + e_t\} = 0. \tag{15}$$

Here all quantities are understood as random variables on the probability space $(\Omega,\mathcal{F},P)$. The data elements are $\xi_t = (\varphi_t, G_t, E_t, h_t, X_t, U_t, Y_t, F_t, D_t, e_t)$ in the outgoing control case and $\xi_t = (\varphi_{t-1,t}, \varphi_t, G_t, E_t, h_t, Y_{t-1,t}, U_t, X_t, F_t, D_t, e_t)$ in the incoming control case, with $X_t, U_t, Y_t, Y_{t-1,t}$ denoting polyhedral sets in appropriate dimensions.[2]

Directly corresponding scenario tree formulations are obtained if the decision maker has full current information, $\mathcal{G}_t = \mathcal{D}_t^*$. Random variables Z_t are then simply replaced by their realizations Z_j, $j \in L_t$, where the scenario tree is constructed from the data filtration $\mathbb{F}_d$. Thus, recalling the notational convention $i \equiv \pi(j)$, the outgoing control problem becomes

$$\min_{x,u} \quad \sum_{j\in V} p_j \varphi_j(x_j,u_j) \equiv \sum_{t=0}^{T} \sum_{j\in L_t} p_j \varphi_j(x_j,u_j) \tag{16}$$

$$\text{s.t.} \quad x_j = G_j x_i + E_j u_i + h_j \qquad \forall j \in V, \tag{17}$$

$$(x_j,u_j) \in (X_j \times U_j) \cap Y_j \qquad \forall j \in V, \tag{18}$$

$$\sum_{j\in V} p_j [F_j x_j + D_j u_j + e_j] = 0, \tag{19}$$

and the incoming control problem becomes

$$\min_{u,x} \quad \sum_{j\in V} p_j [\varphi_{ij}(x_i,u_j) + \varphi_j(x_j)] \tag{20}$$

$$\text{s.t.} \quad x_j = G_j x_i + E_j u_j + h_j \qquad \forall j \in V, \tag{21}$$

$$(x_i,u_j) \in Y_{ij} \qquad \forall j \in V, \tag{22}$$

$$(u_j,x_j) \in U_j \times X_j \qquad \forall j \in V, \tag{23}$$

$$\sum_{j\in V} p_j [D_j u_j + F_j x_j + e_j] = 0. \tag{24}$$

These are precisely the tree-sparse programs studied in [18], except that the feasible sets are kept in the more convenient abstract notation. Appropriate representations by hierarchies

[2] Since the probability space is finite, there is no conceptual difficulty with set-valued or "function-valued" random variables X_t, φ_t etc., and no integrability assumptions or "P-almost surely" specifications are required.

of full-rank linear equations and inequalities are actually a key feature of the tree-sparse modeling and solution framework. (This explains the presence of "unnecessary" sets X_t, U_t in (7) and U_t in (13): they represent certain inequalities in those hierarchies.) The underlying natural classification of constraints will again play a significant role below.

4. Scenario Tree Formulations

We are now ready to construct scenario tree formulations of the discrete-time problems (5)–(9) and (10)–(15) with arbitrary information constraints. Recall which items are given:

(a) a finite probability space $(\Omega, \mathcal{F}, P)$;
(b) a data process $(\xi_t)_{t\in I}$ on $(\Omega, \mathcal{F}, P)$;
(c) a sequence of information fields $\mathcal{G}_t \subseteq \mathcal{F}$, $t \in I$.

The reformulation proceeds in two steps. First, a filtration representing the finest relevant σ-fields is determined to construct the scenario tree. Second, decisions that must be measurable with respect to coarser σ-fields are mapped onto the tree. Artificial states or extra nodes may have to be introduced and control variables may have to be reinterpreted in the process so that information constraints are properly incorporated into the tree topology. This provides the transition from the abstract viewpoint of [11] to a practical formulation.

It is also possible to model information constraints explicitly within our problem classes, as *global* equality constraints. This might even be advantageous in certain unusual situations, but in general it contravenes the very idea of a scenario tree formulation.

4.1. Assumptions.

To get familiar with the probabilistic setting adopted here we discuss some elementary assumptions that can be made without restriction of generality—even for arbitrary (nonconvex, nonsmooth) problem functions and feasible sets.

Formal proofs are mostly omitted in the sequel. Almost all considerations rely directly on the following well-known result (which is trivial in the finite case).

Theorem 1. *Suppose that Z is a random vector on* $(\Omega, \mathcal{F}, P)$ *and that* $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}$. *Then* $\mathbf{E}^{\mathcal{F}_1}\{\mathbf{E}^{\mathcal{F}_2}\{Z\}\} = \mathbf{E}^{\mathcal{F}_2}\{\mathbf{E}^{\mathcal{F}_1}\{Z\}\} = \mathbf{E}^{\mathcal{F}_1}\{Z\}$. *In particular,* $\mathbf{E}\{\mathbf{E}^{\mathcal{F}_2}\{Z\}\} = \mathbf{E}\{Z\}$. □

4.1.1. *Global Constraints Vectors.* Observe first that the randomness of e_t is irrelevant: only the expectation $\bar{e} := \sum_{t=0}^{T} \mathbf{E}\{e_t\} \equiv \mathbf{E}\{\sum_{t=0}^{T} e_t\}$ enters the problem. That is, two problems differing only in the realizations of e_t are equivalent whenever the values of $\bar{e}$ agree. Thus one may drop e_t from ξ_t, setting for instance $e_0 := \bar{e}$ and $e_t := 0$ for $t > 0$.

4.1.2. *Probability Space.* A common assumption under standard information constraints is $\mathcal{F} = \mathcal{F}_T$ where $\mathcal{F}_t := \mathcal{D}_t^*$ (data filtration $\mathbb{F}_d$). Given a set of data scenarios $(\xi_\omega)_{\omega\in\Omega}$ with associated probabilities $(p_\omega)_{\omega\in\Omega}$, this is achieved by the natural definitions $\mathcal{F} := 2^\Omega$ and $P := \sum_{\omega\in\Omega} p_\omega \delta_\omega$. Under general information constraints the σ-fields $\mathcal{D}_t$ and $\mathcal{G}_t$ may be totally unrelated—even $\mathcal{G}_t \not\subseteq \mathcal{D}_T^*$ is possible. In this case Ω must be larger than the set of data scenarios, that is, a probability space must be specified *a priori*. The assumption $\mathcal{F} = \mathcal{F}_T$ now holds with no loss of generality for the refined filtration $\mathcal{F}_t := \sigma(\{\mathcal{D}_\tau, \mathcal{G}_\tau\}_{\tau\le t})$: every feasible solution is automatically $\mathcal{F}_T$-measurable. Thus $\mathcal{F}$ is unnecessarily fine if $\mathcal{F} \supset \mathcal{F}_T$: there exist identical scenarios (having the same data elements ξ_t *and* belonging to the same $\mathcal{G}_t$-atoms throughout), and $(\Omega, \mathcal{F}, P)$ should be replaced with $(\mathcal{A}_T, \mathcal{F}_T, P|\mathcal{F}_T)$.

Example. Consider a simple two-stage problem with $\Omega = L = \{1,2,3,4\}$, $V = \{0\} \cup \Omega$:

$$\min_{u,x} \sum_{j\in V} p_j x_j^2 \qquad \text{s.t.} \qquad x_0 = u_0, \quad \begin{Bmatrix} x_1 = x_0 - u_1, & x_3 = x_0 + u_{23} \\ x_2 = x_0 - u_{23}, & x_4 = x_0 + u_4 \end{Bmatrix}.$$

The σ-fields are $\mathcal{D}_0 = \mathcal{G}_0 = \{\emptyset, \Omega\}$, $\mathcal{D}_1 = \sigma(\{1,2\},\{3,4\})$, and $\mathcal{G}_1 = \sigma(\{1\},\{2,3\},\{4\})$; hence $\mathcal{D}_1 = \mathcal{D}_T^* \neq \mathcal{F}_1 = \mathcal{F} = 2^\Omega$. If we have $\Omega' = \{1,\dots,5\}$ initially with $x_5 = x_0 + u_{23}$, then the identical scenarios $\omega' \in \{3,5\}$ are replaced with the single (aggregated) scenario $\omega = 3$ where $p_3 = p_3' + p_5'$, yielding the coarser model just discussed.

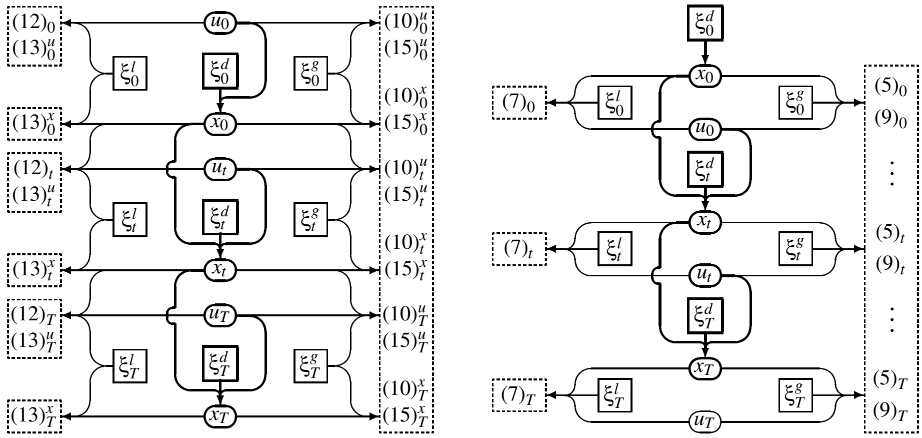

FIGURE 3. Propagation of stochastic information. Left: incoming control; right: outgoing control. $(n)^x_t$ refers to variable x in Eq. (n) at stage t.

4.1.3. *Positive Probabilities.* Another standard assumption is $P > 0$. This means that Ω does not contain scenarios which the decision maker can rule out as impossible ($p_\omega = 0$) based on the information available at $t = 0$. Apart from increasing the solution effort, such scenarios might actually yield a *wrong* problem: they can obviously not contribute to the objective or global constraints but are likely to generate a worse optimum by imposing unjustified local constraints.

Example. Consider a simple tree with root and three leaves (scenarios) having probabilities $p_0 = 1$, $p_1, p_2 > 0$, $p_3 = 0$, with uncertainty occurring only in the constraints:

$$\min_{x,u} \sum_{j\in V} p_j x_j \quad \text{s.t.} \quad x_0 = h_0, \quad x_1 = x_2 = x_3 = x_0 + u_0, \quad x_1 \geq 0, \quad x_2 \geq 1, \quad x_3 \geq 2.$$

The unique solution is $u_0 = 2 - h_0$, $x_1 = x_2 = x_3 = 2$, having cost $h_0 + 2$. Without scenario 3 one obtains the better solution $u_0 = 1 - h_0$, $x_1 = x_2 = 1$, having cost $h_0 + 1$.

4.1.4. *Initial Information Field.* The last assumption is $\mathcal{G}_0 = \{\emptyset, \Omega\}$. This simply means that the *here-and-now* decision u_0 is deterministic: it cannot depend on random events since any event observed up to $t = 0$ is *certain* by construction of the decision model. If optimization takes place before all those events are observed, then every outcome β requires a separate problem on the space $(\beta, \mathcal{F} \cap \beta, \frac{1}{p_\beta} P|\mathcal{F} \cap \beta)$ of conditional probabilities.

4.2. **Dynamic, global, and local data.** A more detailed analysis of the information structure, specifically of the interplay between dynamic and stochastic aspects, will show that the refined filtration of §4.1.2 is in general still unnecessarily fine. To recognize this partition random elements into *dynamic*, *global*, and *local* parts $\xi_t = (\xi^d_t, \xi^g_t, \xi^l_t)$ according to the natural constraints classification in [18]. Having dropped e_t, the components are

$$\xi^d_t = (G_t, E_t, h_t), \qquad \xi^g_t = (\varphi_{t-1,t}, \varphi_t, D_t, F_t), \qquad \xi^l_t = Y_{t-1,t} \times U_t \times X_t \tag{25}$$

in the incoming control problem, and

$$\xi^d_t = (G_t, E_t, h_t), \qquad \xi^g_t = (\varphi_t, F_t, D_t), \qquad \xi^l_t = (X_t \times U_t) \cap Y_t \tag{26}$$

in the outgoing control problem, where ξ^l_t is the local feasible set. The respective σ-fields generated by $\xi^d_t, \xi^g_t, \xi^l_t$ are denoted $\mathcal{D}^d_t, \mathcal{D}^g_t, \mathcal{D}^l_t$, hence $\mathcal{D}_t = \sigma(\mathcal{D}^d_t, \mathcal{D}^g_t, \mathcal{D}^l_t)$.

Probabilistic information (uncertainty) enters the problem directly via data elements ξ_t and indirectly via decisions u_t. Whereas ξ^g_t and ξ^l_t are only relevant in the current stage, information from ξ^d_t and u_t is propagated forward in time by the states. This is illustrated in Fig. 3, which suggests to study the filtration generated by states and controls together.

We therefore introduce state fields $\mathcal{S}_t$ (representing the measurability properties of x_t) and define $\mathcal{F}_t = \sigma(\mathcal{S}_t, \mathcal{G}_t)$.

Consider first the outgoing control case. Since x_t depends on the preceding state x_{t-1}, preceding decision u_{t-1} (if $t > 0$), and on the data element ξ_t^d (but *not* ξ_t^g, ξ_t^l), it must be measurable with respect to $\mathcal{D}_t^d$ and $\mathcal{F}_{t-1}$. This yields the inductive definitions

$$\mathcal{S}_0 := \sigma(\mathcal{D}_0^d) \equiv \mathcal{D}_0^d, \qquad \mathcal{F}_0 := \sigma(\mathcal{S}_0, \mathcal{G}_0), \tag{27}$$

$$\mathcal{S}_t := \sigma(\mathcal{D}_t^d, \mathcal{F}_{t-1}), \qquad \mathcal{F}_t := \sigma(\mathcal{S}_t, \mathcal{G}_t), \quad t = 1, \dots, T. \tag{28}$$

The incoming control case is analogous except that x_t depends on the *current* decision u_t. Obviously this implies $\mathcal{F}_t \equiv \mathcal{S}_t$ so that the appropriate definitions are

$$\mathcal{F}_0 \equiv \mathcal{S}_0 := \sigma(\mathcal{D}_0^d, \mathcal{G}_0) \equiv \mathcal{D}_0^d, \qquad \mathcal{F}_t \equiv \mathcal{S}_t := \sigma(\mathcal{D}_t^d, \mathcal{G}_t, \mathcal{F}_{t-1}), \quad t = 1, \dots, T. \tag{29}$$

It is now easily verified that ξ_t^g, ξ_t^l can be assumed to be $\mathcal{F}_t$-measurable, implying that the filtration just constructed is in fact sufficiently fine for the desired scenario tree formulation. Technically one has to replace the *global* data elements by their conditional expectations with respect to $\mathcal{F}_t$, and the *local* data elements (feasible sets) belonging to the same $\mathcal{F}_t$-atom by their intersection.

Theorem 2. *Define $\bar{\xi}_t^g := \mathbf{E}^{\mathcal{F}_t}\{\xi_t^g\}$. For every $\mathcal{F}_t$-atom α let $\bar{\xi}_{t\alpha}^l := \bigcap_{\omega\in\alpha} \xi_{t\omega}^l$, that is, take intersections of feasible sets. Replacing ξ_t^g, ξ_t^l with $\bar{\xi}_t^g, \bar{\xi}_t^l$ in problem* (10)–(15) *then yields an equivalent problem. An analogous statement holds for the outgoing control case.*

Proof. Let (u,x) be optimal for (10)–(15). Since x_{t-1}, u_t, x_t are all $\mathcal{F}_t$-measurable, we have

$$\begin{aligned}\mathbf{E}\{\varphi_{t-1,t}(x_{t-1}, u_t) + \varphi_t(x_t)\} &= \mathbf{E}\big\{\mathbf{E}^{\mathcal{F}_t}\{\varphi_{t-1,t}(x_{t-1}, u_t) + \varphi_t(x_t)\}\big\} \\ &= \mathbf{E}\big\{\bar{\varphi}_{t-1,t}(x_{t-1}, u_t) + \bar{\varphi}_t(x_t)\big\}.\end{aligned}$$

Thus the objectives are identical. A similar argument applies to the global constraints. Next, by $\mathcal{F}_t$-measurability of u_t,

$$\begin{aligned}u_t \in U_t &\iff \forall \alpha \in \mathcal{A}_t : \forall \omega \in \alpha : u_{t\alpha} \in U_{t\omega} \\ &\iff \forall \alpha \in \mathcal{A}_t : u_{t\alpha} \in \bigcap_{\omega\in\alpha} U_{t\omega} = \bar{U}_{t\alpha} \iff u_t \in \bar{U}_t.\end{aligned}$$

Equivalence of the remaining local constraints is proved similarly, giving $(x_{t-1}, u_t) \in \bar{Y}_{t-1,t}$ and $(u_t, x_t) \in \bar{U}_t \times \bar{X}_t$. Analogous arguments apply to the outgoing control case. ☐

Remark. The proof shows that some of the data elements can actually be assumed to be measurable with respect to even coarser σ-fields. With incoming control these σ-fields are $\mathcal{G}_t$ for U_t, D_t, then $\sigma(\mathcal{F}_{t-1}, \mathcal{G}_t)$ for $\varphi_{t-1,t}, Y_{t-1,t}$, and finally $\mathcal{F}_t$ for φ_t, X_t, F_t. With outgoing control they are $\mathcal{G}_t$ for U_t, D_t and $\mathcal{S}_t$ for φ_t, X_t, Y_t, F_t.

4.3. **Examples.** Before we proceed with the tree construction let us investigate the two main phenomena that may occur under general information constraints: delayed decisions (possibly caused by delayed observations, by the computing time spent on optimizations, or by the time required to implement decisions), and foresight (the possibility to observe random events before they have an effect on the dynamic process). Here we make the simplifying assumption $\mathcal{D}_t^d = \mathcal{D}_t^g = \mathcal{D}_t^l$ to obviate a distinction of the data fields. Full technical details of the scenario tree formulations will be provided.

4.3.1. *Information Delay.* A common situation occurs when decisions are based on full information up to the *previous* stage, $\mathcal{G}_t = \mathcal{D}_{t-1}^*$. More generally decisions may be delayed by δ stages, $\mathcal{G}_t = \mathcal{D}_{t-\delta}^*$.[3] Information fields are then clearly coarser than the data fields, the realization of ξ_0 is unknown at $t = 0$, and the construction in §4.2 yields $\mathcal{F}_t = \mathcal{D}_t^*$ where $\mathcal{F}_0 \neq \{\emptyset, \Omega\}$ in general. Since $u_{t+\delta}$ is $\mathcal{F}_t$-measurable for $t \in \{0, \dots, T-\delta\}$, it is natural to

[3]For $t - \delta < 0$ we have $\mathcal{D}_{t-\delta}^* = \{\emptyset, \Omega\}$ and the actual delay is $\min(\delta, t+1)$.

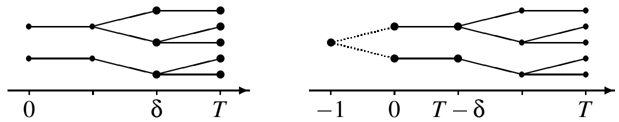

FIGURE 4. Information delay δ stages. Left: data scenarios (no observations available in small nodes: $t < \delta$). Right: scenario tree (no control variables in small nodes: $t > T - \delta$; fictitious stage -1 prepended).

place its realizations $(u'_j)_{j \in L_t}$ on level t (rather than the realizations of u_t). Controls u'_j in the final stages $t > T - \delta$ are then empty and a fictitious stage -1 is required whose control u'_{-1} represents the deterministic decisions $u_0, \dots, u_{\delta-1}$. Decision u_{t-1} (outgoing) or u_t (incoming) is passed forward to stage t through artificial states x'_j. No random data ξ_{-1} exist, so stage -1 is deterministic. The scenario tree rooted in $\mathcal{F}_{-1} = \{\emptyset, \Omega\}$ is thus obtained from the *extended data filtration* $\mathbb{F} := \{\mathcal{F}_t\}_{t=-1}^T \equiv \{\mathcal{D}_t^*\}_{t=-1}^T$; cf. Fig. 4.

The precise reformulation for outgoing control (5)–(9) with delay $\delta \geq 1$ reads

$$\min_{(x,x'),u'} \quad \sum_{t=0}^{T} \sum_{j \in L_t} p_j \varphi_j(x_j, P_j x'_j) \tag{30}$$

$$\text{s.t.} \quad x_j = G_j x_i + E_j P_i x'_i + h_j \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T, \tag{31}$$

$$x'_j = u'_{-1} \qquad \forall j \in L_0, \tag{32}$$

$$x'_j = Q_i x'_i + R_i u'_i \qquad \forall j \in L_t, \quad \forall t = 1, \dots, T, \tag{33}$$

$$(x_j, P_j x'_j) \in (X_j \times U_j) \cap Y_j \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T, \tag{34}$$

$$\sum_{t=0}^{T} \sum_{j \in L_t} p_j [F_j x_j + D_j P_j x'_j + e_j] = 0. \tag{35}$$

Here $P_j x'_j$ replaces u_j, and x_{-1}, x'_{-1} are empty so that stage -1 contributes only u'_i in (33). The controls u'_j are formally unconstrained, and x'_j passes forward the previous δ decisions. Proper components are selected by (deterministic) matrices P_j, Q_j, R_j where

$$\left[\begin{array}{c|c} P_j & \\ \hline Q_j & R_j \end{array}\right] = \left[\begin{array}{cc|c} I & 0 & \\ \hline 0 & I & 0 \\ 0 & 0 & I \end{array}\right] \in \mathbb{R}^{(n_t^u + n_{t+1:t+\delta-1}^u + n_{t+\delta}^u) \times (n_t^u + n_{t+1:t+\delta-1}^u + n_{t+\delta}^u)}, \quad j \in L_t. \tag{36}$$

Here $n_t^u := \dim(u_t)$ and $n_{t:t+\delta}^u := n_t^u + \cdots + n_{t+\delta}^u$. If $\delta = 1$ then $n_{t+1:t+\delta-1}^u$ clearly vanishes, yielding $P_j = I$, $Q_j = 0$, $R_j = I$, and obvious simplifications in (30)–(35). (Likewise P_{-1} and the unused matrices Q_j, R_j for $j \in L_T$ are always empty.)

The reformulation for incoming control (10)–(15) with delay $\delta \geq 1$ is quite similar,

$$\min_{u',(x,x')} \quad \sum_{t=0}^{T} \sum_{j \in L_t} p_j [\varphi_{ij}(x_i, P_j x'_i) + \varphi_j(x_j)] \tag{37}$$

$$\text{s.t.} \quad x_j = G_j x_i + E_j P_j x'_i + h_j \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T, \tag{38}$$

$$x'_{-1} = u'_{-1}, \tag{39}$$

$$x'_j = Q_j x'_i + R_j u'_j \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T-1, \tag{40}$$

$$(x_i, P_j x'_i) \in Y_{ij} \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T, \tag{41}$$

$$(P_j x'_i, x_j) \in U_j \times X_j \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T, \tag{42}$$

$$\sum_{t=0}^{T} \sum_{j \in L_t} p_j [D_j P_j x'_i + F_j x_j + e_j] = 0. \tag{43}$$

Now $P_j x_i$ replaces u_j and x_{-1}, x'_T are empty (time shift), so stage -1 contributes to several conditions. Note in particular that $u_j = P_j x'_i$ appears in the local constraints for all $j \in L_t$.

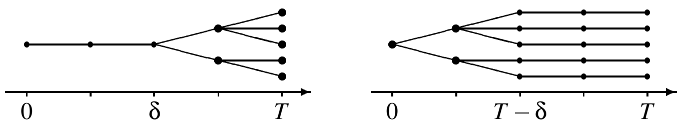

FIGURE 5. Foresight δ stages. Left: data scenarios (deterministic for $t \le \delta$). Right: scenario tree (no further branches for $t \ge T-\delta$).

That is, constraints (41), (42) can be written $(x_i, P_j x_i') \in \bigcap_{j\in S(i)} Y_{ij}$ and $P_j x_i' \in \bigcap_{j\in S(i)} U_j$, respectively.[4] This feasibility issue illustrates a principal difficulty with information delays: decision u_j' (made at time t) must satisfy all constraint realizations at time $t+\delta$.

For $\delta = 1$ we have again $P_j = I$, $Q_j = 0$, and $R_j = I$, implying $x_j' = u_j'$ throughout. This makes artificial states obsolete and we may switch to an outgoing control formulation: the controls are simply shifted back to the previous stage, $u_t = (u_j')_{j\in L_{t-1}}$.

4.3.2. *Foresight.* The opposite case occurs when the decision maker can observe random events δ stages *before* they actually influence the process, $\mathcal{G}_t = \mathcal{D}^*_{t+\delta}$.[5] This implies that $\xi_0, \dots, \xi_\delta$ are deterministic (their realizations are known at $t=0$) and hence $\mathcal{D}_t = \{\emptyset, \Omega\}$ for $t \le \delta$. Information fields are now finer than the data fields and the construction in §4.2 yields $\mathcal{F}_t = \mathcal{D}^*_{t+\delta}$. Thus the scenario tree is constructed from the *information filtration* $\mathbb{F} := \{\mathcal{F}_t\}_{t\in I} \equiv \{\mathcal{G}_t\}_{t\in I}$. With this tree one obtains a straightforward reformulation as usual, except that duplicate data elements ξ_j will generally exist since $\mathcal{D}^*_t \subseteq \mathcal{G}_t$; see Fig. 5.

If the original problem has outgoing control one may alternatively switch to an incoming control reformulation where the states and certain data elements are shifted back one stage:

$$\min_{u,x'} \quad \sum_{t=0}^{T} \sum_{j\in L_t} p_j \varphi_j(x_i', u_j) \tag{44}$$

$$\text{s.t.} \quad x_j' = G_j' x_i' + E_j' u_j + h_j' \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T-1, \tag{45}$$

$$(x_i', u_j) \in Y_j \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T, \tag{46}$$

$$u_j \in U_j \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T, \tag{47}$$

$$x_j' \in X_j' \qquad \forall j \in L_t, \quad \forall t = 0, \dots, T-1, \tag{48}$$

$$\sum_{t=0}^{T-1} \sum_{j\in L_t} p_j F_j' x_j' + \sum_{t=0}^{T} \sum_{j\in L_t} p_j [D_j u_j + e_j] = 0. \tag{49}$$

Here x_i' and $(G_i, E_i, h_i, X_i, F_i)'$ are the respective realizations of $\mathcal{F}_{t-1}$-measurable random variables x_t and $(G_t, E_t, h_t, X_t, F_t)$ for $t > 0$. The initial state $x_0 = x_{-1}'$ does not actually appear in the problem: its fixed value h_0 is directly substituted in (44)–(46).

4.4. **Mapping decisions onto trees.** Returning to the filtration constructed in §4.2, we observe first that the initial σ-field $\mathcal{F}_0$ can be assumed to be trivial without restriction of generality. Recall that $\mathcal{F}_0 = \sigma(\mathcal{D}_0^d, \mathcal{G}_0) = \mathcal{D}_0^d$. If $\mathcal{D}_0^d \ne \{\emptyset, \Omega\}$ then ξ_0^d is uncertain at $t=0$: the problem has an initial delay of information. As in §4.3.1 we prepend a fictitious stage $t=-1$ which contains u_0 and possibly further deterministic decisions, but nothing else. A time shift $t \to t+1$ now yields the desired property. Thus a scenario tree can always be constructed either from the filtration $\{\mathcal{F}_t\}_{t\in I}$ or from the extended filtration $\{\mathcal{F}_t\}_{t=0}^{T+1}$. The data process ξ, decision process u, and state process x are all adapted to that filtration.

The states and data elements are naturally represented in terms of their realizations as $x_t = (x_{t\alpha})_{\alpha\in\mathcal{A}_t}$ and $\xi_t = (\xi_{t\alpha})_{\alpha\in\mathcal{A}_t}$ (or $(x_j)_{j\in L_t}$ and $(\xi_j)_{j\in L_t}$ in tree notation), where $\mathcal{A}_t$ is

[4] Observe that $P_j x_i$ depends on i only (not on j) since P_j is deterministic; cf. (36).

[5] For $t+\delta \ge T$ we have $\mathcal{D}^*_{t+\delta} = \mathcal{D}^*_T = \mathcal{F}$ and the actual foresight is $\min(\delta, T-t)$.

TABLE 1. Random data elements in the general example problem.

ω	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
H_0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E_0	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
b_0	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
F_0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2
H_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
b_1	17	17	17	17	17	17	17	17	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	17	17	17	17	17	17	17	17
F_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H_2	0	0	0	0	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	15	15	15	15	15	15	15	15
E_2	−1	−1	−1	−1	1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1	1	1	1	1	1	1	1	1	−1	−1	−1	−1
b_2	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	11	11	11	11
F_2	6	6	6	6	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−3	−2	−2	−2	−2	−2	−2	−2	−2
H_3	5	5	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
E_3	1	1	1	1	1	1	−2	−2	−2	−2	1	1	1	1	1	1	1	1	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2	−2
b_3	7	7	7	7	7	7	7	7	7	7	7	7	7	7	11	11	11	11	11	11	11	11	11	11	11	11	7	7	7	7	7	7
F_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	−9	−9
H_4	3	3	3	4	4	3	3	3	3	4	4	3	3	4	4	3	3	4	4	3	3	4	4	3	3	4	4	3	3	3	3	4
E_4	−2	−2	−2	−2	3	3	3	3	3	3	3	3	3	3	3	3	−2	−2	−2	−2	−2	−2	3	3	3	3	3	3	3	3	3	3
b_4	42	24	24	24	42	42	42	24	24	24	24	24	24	24	24	42	42	42	42	42	42	42	42	42	42	42	42	42	42	24	24	24
F_4	2	2	2	−1	−1	2	2	2	2	−1	−1	2	2	−1	−1	2	2	−1	−1	2	2	−1	−1	2	2	−1	−1	2	2	2	2	−1

the set of $\mathcal{F}_t$-atoms. Both representations may have duplicate realizations, but this does not cause any difficulties since neither x nor ξ are subject to measurability restrictions.

Decisions u_t must be $\mathcal{G}_t$-measurable so that the direct representation as $u_t = (u_{t\alpha})_{\alpha\in\mathcal{A}_t}$ is perfect if $\mathcal{G}_t = \mathcal{F}_t$, which is the case if the realizations of $\xi_0,\dots,\xi_t$ are known at time t. Otherwise $\mathcal{G}_t$ is strictly coarser than $\mathcal{F}_t$ and there exists at least one $\mathcal{G}_t$-atom $\beta \in \mathcal{B}_t$ which is a union of two or more atoms $\alpha_\nu \in \mathcal{A}_t$. In this case the representation is too fine: it requires explicit equality constraints to ensure that all realizations $(u_{t\alpha})_{\alpha\in\beta}$ are indeed identical. Hence the question arises in which node one should place $u_{t\beta}$. This is now easily decided. Since $\{\mathcal{F}_t\}$ is a filtration with $\mathcal{F}_0 = \{\emptyset,\Omega\}$, a unique largest $\delta(\beta) \in \{1,\dots,t\}$ exists such that β is contained in an atom $\alpha \in \mathcal{A}_{t-\delta(\beta)}$. We interpret $\delta(\beta)$ as the *local delay of information* in β.[6] In tree notation the set $\{\alpha_\nu\}$ (hence β) simply corresponds to a subset $B \subseteq L_t$ and α corresponds to the root i of the unique smallest subtree containing all nodes $j \in B$. As in §4.3.1 we place $u_{t\beta}$ in node $i \in L_{t-\delta(\beta)}$ and pass decisions forward as artificial states on all the paths from i to $j \in B$. This completes the problem reformulation.

Instead of inventing a general notation to write down complete technical details we provide an illustrative example problem. The reader should thus be able to apply the general construction to any particular problem instance that he or she encounters.

4.5. **General case example.** We now study an artificial example which is constructed in such a way that it exhibits all aspects discussed in the previous sections. (This results in a perhaps surprisingly complex information structure.) Consider an outgoing control problem with $x_{t\omega}, u_{t\omega} \in \mathbb{R}$:

$$\min_{x,u} \sum_{t=0}^{T} \mathbf{E}\{H_t x_t^2\} \qquad \text{s.t.} \qquad x_t = x_{t-1} + E_t u_{t-1}, \quad u_t \in [0,b_t], \quad \sum_{t=0}^{T} \mathbf{E}\{F_t x_t\} = \bar{e}.$$

Random data elements $\xi_t = \{H_t, E_t, b_t, F_t\}$ for a 32-scenario instance with $T = 4$ periods (five stages) are listed in Table 1. Some data are deterministic here, such as G_t, h_t, D_t and the lower control bound (zero). Using abbreviated notation $\overline{n} := \{n\}$ and $\overline{m{:}n} := \{m,\dots,n\}$

[6]Conversely, if $\beta \in \mathcal{A}_t$ then there exists a unique smallest $\delta(\beta) \le 0$ such that $\beta \in \mathcal{A}_{t-\delta(\beta)}$. Thus β is an atom of $\mathcal{F}_t,\dots,\mathcal{F}_{t-\delta(\beta)}$, implying that no further branches occur up to stage $t-\delta(\beta)$ in the partial scenario corresponding to $\beta \in \mathcal{A}_t$. We interpret $-\delta(\beta)$ as the *local foresight* in β (or $\delta(\beta)$ as the local delay of information).

for the atoms, information fields are given as

$$\mathcal{G}_0 = \sigma(\overline{1{:}32}) = \{\emptyset, \Omega\},$$
$$\mathcal{G}_1 = \sigma(\overline{1{:}16}, \overline{17{:}18}, \overline{19{:}20}, \overline{21{:}22}, \overline{23{:}24}, \overline{25{:}32}),$$
$$\mathcal{G}_2 = \sigma(\overline{1{:}16}, \overline{17{:}18}, \overline{19{:}20}, \overline{21}, \overline{22}, \overline{23}, \overline{24}, \overline{25{:}26}, \overline{27{:}28}, \overline{29{:}32}),$$
$$\mathcal{G}_3 = \sigma(\overline{1{:}2}, \overline{3{:}4}, \overline{5{:}6}, \overline{7{:}8}, \overline{9{:}10}, \overline{11{:}14}, \overline{15{:}16}, \overline{17{:}20}, \overline{21}, \overline{22}, \overline{23}, \overline{24}, \overline{25{:}26, 31{:}32}, \overline{27{:}30}),$$
$$\mathcal{G}_4 = \sigma(\overline{1{:}6}, \overline{7{:}12}, \overline{13{:}20}, \overline{21}, \overline{22}, \overline{23}, \overline{24}, \overline{25, 27, 29, 31}, \overline{26, 28, 30, 32}).$$

(Note that information get lost in stages 3 and 4.) The dynamic data fields are obtained as

$$\mathcal{D}_0^d = \sigma(\overline{1{:}16}, \overline{17{:}32}), \qquad \mathcal{D}_3^d = \sigma(\overline{1{:}6, 11{:}18}, \overline{7{:}10, 19{:}32}),$$
$$\mathcal{D}_1^d = \sigma(\overline{1{:}32}) = \{\emptyset, \Omega\}, \qquad \mathcal{D}_4^d = \sigma(\overline{1{:}4, 17{:}22}, \overline{5{:}16, 23{:}32}).$$
$$\mathcal{D}_2^d = \sigma(\overline{1{:}4, 13{:}20, 29{:}32}, \overline{5{:}12, 21{:}28}),$$

Observing that $\mathcal{D}_0$ has two atoms (initial information delay) it is easily seen that the data filtration $\mathbb{F}_d = \{\mathcal{D}_t^*\}$ corresponds to two regular binary trees with 16 scenarios each. It is also seen that the probability space with $\mathcal{F} = \mathcal{D}_T^* = 2^\Omega$ is unnecessarily fine: the construction according to §4.2 (with trivial stage -1 prepended) yields the 22-scenario filtration

$$\mathcal{F}_{-1} = \{\emptyset, \Omega\}, \qquad \mathcal{F}_0 = \mathcal{D}_0^d, \qquad \mathcal{F}_1 = \mathcal{G}_1,$$
$$\mathcal{F}_2 = \sigma(\overline{1{:}4}, \overline{5{:}12}, \overline{13{:}16}, \overline{17{:}18}, \overline{19{:}20}, \overline{21}, \overline{22}, \overline{23}, \overline{24}, \overline{25{:}26}, \overline{27{:}28}, \overline{29{:}32}),$$
$$\mathcal{F}_3 = \sigma(\overline{1{:}2}, \overline{3{:}4}, \overline{5{:}6}, \overline{7{:}8}, \overline{9{:}10}, \overline{11{:}12}, \overline{13{:}14}, \overline{15{:}16}, \overline{17{:}18}, \overline{19{:}20}, \overline{21}, \overline{22}, \overline{23}, \overline{24}, \overline{25{:}26}, \overline{27{:}28}, \overline{29{:}30}, \overline{31{:}32}),$$
$$\mathcal{F}_4 = \sigma(\overline{1{:}2}, \overline{3{:}4}, \overline{5{:}6}, \overline{7{:}8}, \overline{9{:}10}, \overline{11{:}12}, \overline{13{:}14}, \overline{15{:}16}, \overline{17{:}18}, \overline{19{:}20}, \overline{21}, \overline{22}, \overline{23}, \overline{24}, \overline{25}, \overline{26}, \overline{27}, \overline{28}, \overline{29}, \overline{30}, \overline{31}, \overline{32}).$$

On the other hand, this filtration is finer than the one obtained from dynamic data fields alone; the information fields are indeed relevant. The associated 6-stage scenario tree has $1 + 2 + 6 + 12 + 18 + 22 = 61$ nodes whose data and variables are given in Table 2, where $p_\omega = 1/32$ for all $\omega \in \Omega$. Here H_j and F_j are the realizations of appropriate conditional expectations whereas b_j is always the minimum of the relevant realizations. The delay of information ranges from -3 (foresight three stages) up to 5: decision $u_{4,13} = \cdots = u_{4,20}$ must be made before any observations become available.

5. Generalized Linear-Quadratic Control

A comparison of the tree-sparse programs with the generalized linear-quadratic control problems of Rockafellar and Wets [11] has been provided in [18] under standard information constraints $\mathcal{G}_t = \mathcal{D}_t^*$. Here we consider again the smooth quadratic case (obtained with $V := \mathbb{R}^{l_1} \times \mathbb{R}_+^{l_2}$ and $Q := 0$ in the notation of [11]; see also [9, Example 3.2]), but under general information constraints. The stochastic problem $(\mathcal{P}_{\text{sto}})$ [11, §4] then reads

$$\min_{u,x} \quad \sum_{t=0}^{T} \mathbf{E}\Big\{\frac{1}{2}u_t^* P_t u_t + p_t^* u_t - c_{t+1}^* x_t\Big\} \tag{50}$$
$$\text{s.t.} \quad x_t = A_t x_{t-1} + B_t u_t + b_t \qquad \forall t = 0, \dots, T, \tag{51}$$
$$u_t \in U_t \qquad \forall t = 0, \dots, T, \tag{52}$$
$$\mathbf{E}^{\mathcal{G}_t}\{u_t\} = u_t \qquad \forall t = 0, \dots, T, \tag{53}$$
$$\mathbf{E}^{\mathcal{G}_t}\{C_t^1 x_{t-1}\} + D_t^1 u_t = q_t^1 \qquad \forall t = 1, \dots, T+1, \tag{54}$$
$$\mathbf{E}^{\mathcal{G}_t}\{C_t^2 x_{t-1}\} + D_t^2 u_t \geq q_t^2 \qquad \forall t = 1, \dots, T+1. \tag{55}$$

(Here A_0, x_{-1}, D_{T+1}, and u_{T+1} are empty.) The data elements P_t, p_t, U_t, D_t, q_t are assumed to be $\mathcal{G}_t$-measurable in [11] whereas no restrictions are imposed on A_t, B_t, b_t, C_t, c_t. Clearly

TABLE 2. Mapping of data and variables to scenario tree in the general example. Nodes j given by t and $\mathcal{F}_t$-atom α; duplicate states indicated by $\mathcal{S}_t$-atoms under x_j; delayed decisions listed under x'_j and u'_j, where $u_{3,25::32} := u_{3,25:26,31:32}$, $u_{4,25::31} := u_{4,25,27,29,31}$, $u_{4,26::32} := u_{4,26,28,30,32}$; local delay of information given under $\delta(j)$; '—' indicates empty items.

$j = t,\alpha$	H_j	E_j	b_j	F_j	x_j	x'_j	u'_j	$\delta(j)$
−1, 1:32	—	—	—	—	—	—	$u_{0,1:32}, u_{4,13:20}$	0
0,1:16	2	−1	3	1	0,1:16	$u_{0,1:32}, u_{4,13:20}$	—	1
0,17:32	1	1	3	−2	0,17:32	$u_{0,1:32}, u_{4,13:20}$	$u_{3,17:20}$	1
1,1:16	0	1	4	0	1,1:16	$u_{4,13:20}$	$u_j, u_{2,1:16}, u_{3,11:14}, u_{4,1:6}, u_{4,7:12}$	0
1,17:18	0	1	4	0	1,17:32	$u_{3,17:20}, u_{4,13:20}$	u_j	−3
1,19:20	0	1	4	0	1,17:32	$u_{3,17:20}, u_{4,13:20}$	u_j	−3
1,21:22	0	1	4	0	1,17:32	—	u_j	0
1,23:24	0	1	4	0	1,17:32	—	u_j	0
1,25:32	0	1	17	0	1,17:32	—	$u_j, u_{3,25::32}, u_{3,27:30}, u_{4,25::31}, u_{4,26::32}$	0
2,1:4	0	−1	47	6	2,1:4	$u_{2,1:16}, u_{4,1:6}$	—	1
2,5:12	8	1	47	−3	2,5:12	$u_{2,1:16}, u_{3,11:14}, u_{4,1:6}, u_{4,7:12}$	—	1
2,13:16	8	−1	47	−3	2,13:16	$u_{2,1:16}, u_{3,11:14}, u_{4,13:20}$	—	1
2,17:18	8	−1	47	−3	2,17:18	$u_{3,17:20}, u_{4,13:20}$	u_j	−2
2,19:20	8	−1	47	−3	2,19:20	$u_{3,17:20}, u_{4,13:20}$	u_j	−2
2,21	8	1	47	−3	2,21:22	—	u_j	−2
2,22	8	1	47	−3	2,21:22	—	u_j	−2
2,23	8	1	47	−3	2,23:24	—	u_j	−2
2,24	8	1	47	−3	2,23:24	—	u_j	−2
2,25:26	15	1	47	−2	2,25:28	$u_{3,25::32}, u_{4,25::31}, u_{4,26::32}$	u_j	−1
2,27:28	15	1	47	−2	2,25:28	$u_{3,27:30}, u_{4,25::31}, u_{4,26::32}$	u_j	−1
2,29:32	15	−1	11	−2	2,29:32	$u_{3,25::32}, u_{3,27:30}, u_{4,25::31}, u_{4,26::32}$	u_j	0
3,1:2	5	1	7	0	3,1:4	$u_{4,1:6}$	u_j	0
3,3:4	9	1	7	0	3,1:4	$u_{4,1:6}$	u_j	0
3,5:6	9	1	7	0	3,5:6	$u_{4,1:6}$	u_j	0
3,7:8	9	−2	7	0	3,7:10	$u_{4,7:12}$	u_j	0
3,9:10	9	−2	7	0	3,7:10	$u_{4,7:12}$	u_j	0
3,11:12	9	1	7	0	3,11:12	$u_{3,11:14}, u_{4,7:12}$	—	2
3,13:14	9	1	7	0	3,13:16	$u_{3,11:14}, u_{4,13:20}$	—	2
3,15:16	9	1	11	0	3,13:16	$u_{4,13:20}$	u_j	0
3,17:18	9	1	11	0	3,17:18	$u_{3,17:20}, u_{4,13:20}$	—	3
3,19:20	9	−2	11	0	3,19:20	$u_{3,17:20}, u_{4,13:20}$	—	3
3,21	9	−2	11	0	3,21	—	u_j	−1
3,22	9	−2	11	0	3,22	—	u_j	−1
3,23	9	−2	11	0	3,23	—	u_j	−1
3,24	9	−2	11	0	3,24	—	u_j	−1
3,25:26	9	−2	11	0	3,25:26	$u_{3,25::32}, u_{4,25::31}, u_{4,26::32}$	—	2
3,27:28	9	−2	7	0	3,27:28	$u_{3,27:30}, u_{4,25::31}, u_{4,26::32}$	—	2
3,29:30	9	−2	7	0	3,29:32	$u_{3,27:30}, u_{4,25::31}, u_{4,26::32}$	—	2
3,31:32	9	−2	7	−9	3,29:32	$u_{3,25::32}, u_{4,25::31}, u_{4,26::32}$	—	2
4,1:2	3	−2	24	2	4,1:2	$u_{4,1:6}$	—	3
4,3:4	3.5	−2	24	0.5	4,3:4	$u_{4,1:6}$	—	3
4,5:6	3.5	3	42	0.5	4,5:6	$u_{4,1:6}$	—	3
4,7:8	3	3	24	2	4,7:8	$u_{4,7:12}$	—	3
4,9:10	3.5	3	24	0.5	4,9:10	$u_{4,7:12}$	—	3
4,11:12	3.5	3	24	0.5	4,11:12	$u_{4,7:12}$	—	3
4,13:14	3.5	3	24	0.5	4,13:14	$u_{4,13:20}$	—	5
4,15:16	3.5	3	24	0.5	4,15:16	$u_{4,13:20}$	—	5
4,17:18	3.5	−2	42	0.5	4,17:18	$u_{4,13:20}$	—	5
4,19:20	3.5	−2	42	0.5	4,19:20	$u_{4,13:20}$	—	5
4,21	3	−2	42	2	4,21	—	u_j	0
4,22	4	−2	42	−1	4,22	—	u_j	0
4,23	4	3	42	−1	4,23	—	u_j	0
4,24	3	3	42	2	4,24	—	u_j	0
4,25	3	3	42	2	4,25:26	$u_{4,25::31}$	—	3
4,26	4	3	42	−1	4,25:26	$u_{4,26::32}$	—	3
4,27	4	3	42	−1	4,27:28	$u_{4,25::31}$	—	3
4,28	3	3	42	2	4,27:28	$u_{4,26::32}$	—	3
4,29	3	3	42	2	4,29:30	$u_{4,25::31}$	—	3
4,30	3	3	24	2	4,29:30	$u_{4,26::32}$	—	3
4,31	3	3	24	2	4,30:31	$u_{4,25::31}$	—	3
4,32	4	3	24	−1	4,30:31	$u_{4,26::32}$	—	3

we have incoming control with $\xi_t^d = (A_t, B_t, b_t)$, $\xi_t^g = (P_t, p_t, c_{t+1})$, and $\xi_t^l = (U_t, C_t, D_t, q_t)$. Thus $\mathcal{G}_t$-measurability of the specified data is not really a restriction: it holds without loss of generality by the direct analogue of Theorem 2 (and likewise $\mathcal{F}_t$-measurability of c_{t+1}).[7] However, constraints (54), (55) deserve closer inspection: the conditional expectations $\mathbf{E}^{\mathcal{G}_t}\{C_t^i x_{t-1}\}$ enforce $\mathcal{G}_t$-measurability of the terms $C_t^i x_{t-1}$ (which may then be assumed of D_t^i, q_t^i as well). Without conditional expectations these constraints would represent polyhedra $Y_{t,t-1}$ that could only be assumed to be measurable with respect to $\sigma(\mathcal{F}_{t-1}, \mathcal{G}_t) \subseteq \mathcal{F}_t$. This would allow slightly more general feasible sets, like in the tree-sparse programs.

6. Conclusions

The previous investigations demonstrate that tree-sparse programs provide natural scenario tree formulations of stochastic programs with general information constraints. Given a problem with arbitrary finite data and information fields, we have shown how the *dynamic* data fields together with the information fields generate a filtration that yields the direct scenario tree reformulation, both for incoming and outgoing control models. Using the relations between atoms of the various σ-fields, we have also interpreted nonstandard information constraints, either as a local delay of information or as local foresight. A local delay δ at time t means that reliable information is only available on random events that have influenced the process until time $t-\delta$ in the past. In other words, the current decision could have been made δ periods earlier if full information had been available. Conversely, a local foresight of δ at time t means that reliable information is already available on all random events that will influence the process until time $t+\delta$ in the future. In the extreme case of an entirely deterministic process (i.e., with foresight $\delta = T$), all decisions can thus be optimized *off line*—at the very beginning. This is also true for the other extreme case (delay $\delta = T$) where no new information is acquired during the entire process.

Further we have shown that outgoing control problems with uniform foresight $\delta \geq 1$ can be recast as incoming control problems, and incoming control problems with uniform delay $\delta = 1$ can be recast as outgoing control problems. Considering the feasibility issues associated with information delays, one may assume that models with delay $\delta > 1$ are probably not reasonable in most practical situations (the time discretization is "too fine"). Thus outgoing control naturally models delays while incoming control naturally models immediate reactions or foresight. This is the reason for considering two versions of tree-sparse programs: they cover almost all practically relevant situations directly. (Actually there is a third variant with dynamic equations in implicit form; cf. [17, 18]. This variant is similar to the standard formulation of stochastic programs [4].)

Of course, the trees constructed in this paper may be further modified to increase numerical efficiency. In the general example problem, for instance, several nodes toward the end of the horizon have empty controls. A smaller problem can be obtained by clustering these nodes and eliminating the associated states.

To add concreteness and to allow finer distinctions of the measurability properties of data elements, we have studied *convex* stochastic programs in §4. However, none of the considerations rely on this. The constructions carry over directly to the general case where the dynamics (6), (11) involve arbitrary transition mappings θ_t and where global constraints (9), (15) have the general form $\sum_t \mathbf{E}\{\gamma_t(x_t, u_t)\} = 0$, with no restrictions whatsoever imposed on the problem functions or feasible sets. For instance, everything applies directly to stochastic integer programs (see [13] and references therein), and the ideas extend to other classes of stochastic optimization models (see, e.g., [5, 7, 8] for a general overview). In any case, the generally applicable concept of distinguishing dynamic, global, and local problem data is the key to a concise problem formulation based on the full understanding of the relation between information structure and algebraic structure.

---

[7]The $\mathcal{G}_t$-measurable equality constraints read $\mathbf{E}^{\mathcal{G}_t}\{C_t^1 x_{t-1}\}_\beta + D_{t\beta}^1 u_{t\beta} = q_{t\beta}^1$ for all $\beta \in \mathcal{B}_t$ (where $D_{t\omega}^1 = D_{t\beta}^1$ and $q_{t\omega}^1 = q_{t\beta}^1$ for all $\omega \in \beta$); similarly the inequalities.

REFERENCES

[1] H. BAUER, *Maß- und Integrationstheorie. (Measure and integration theory)*, de Gruyter Lehrbuch, de Gruyter, Berlin, 2nd, revised ed., 1992. (In German.).

[2] ———, *Probability Theory*, de Gruyter Studies in Mathematics 23, de Gruyter, Berlin, 1996. (Translated from the German by Robert B. Burckel.).

[3] D. P. BERTSEKAS AND S. E. SHREVE, *Stochastic Optimal Control: The Discrete-Time Case*, Academic Press, New York, 1978.

[4] J. R. BIRGE, M. A. H. DEMPSTER, H. I. GASSMANN, E. A. GUNN, A. J. KING, AND S. W. WALLACE, *A standard input format for multiperiod stochastic linear programs*, COAL newsletter, (1987), pp. 1–19.

[5] J. R. BIRGE AND F. LOUVEAUX, *Introduction to Stochastic Programming*, Springer-Verlag, New York, 1997.

[6] R. M. DUDLEY, *Real Analysis and Probability*, Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989.

[7] P. KALL AND S. W. WALLACE, *Stochastic Programming*, Wiley, New York, 1994.

[8] A. PRÉKOPA, *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.

[9] R. T. ROCKAFELLAR, *Linear-quadratic programming and optimal control*, SIAM J. Control Optim., 25 (1987), pp. 781–814.

[10] ———, *Duality and optimality in multistage stochastic programming*, Ann. Oper. Res., 85 (1999), pp. 1–19.

[11] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time*, SIAM J. Control Optim., 28 (1990), pp. 810–822.

[12] ———, *Scenarios and policy aggregation in optimization under uncertainty*, Math. Oper. Res., 16 (1991), pp. 119–147.

[13] W. RÖMISCH AND R. SCHULTZ, *Multistage stochastic integer programs: An introduction*, in Online Optimization of Large Scale Systems, M. Grötschel, S. O. Krumke, and J. Rambau, eds., Springer-Verlag, 2001, pp. 581–600.

[14] A. RUSZCZYŃSKI, *Decomposition methods in stochastic programming*, Math. Programming, 79 (1997), pp. 333–353.

[15] A. N. SHIRYAEV, *Probability*, Graduate Texts in Mathematics, Springer-Verlag, 1984.

[16] M. C. STEINBACH, *Recursive direct algorithms for multistage stochastic programs in financial engineering*, in Operations Research Proceedings 1998, P. Kall and H.-J. Lüthi, eds., New York, 1999, Springer-Verlag, pp. 241–250.

[17] ———, *Hierarchical sparsity in multistage convex stochastic programs*, in Stochastic Optimization: Algorithms and Applications, S. P. Uryasev and P. M. Pardalos, eds., Kluwer Academic Publishers, 2001, Dordrecht, The Netherlands, pp. 385–410.

[18] ———, *Tree-sparse convex programs*, Technical Report ZR-01-08, ZIB, 2001. Submitted for publication.

MARC C. STEINBACH, KONRAD-ZUSE-ZENTRUM FÜR INFORMATIONSTECHNIK BERLIN, DEPARTMENT OPTIMIZATION, TAKUSTR. 7, 14195 BERLIN-DAHLEM, GERMANY

E-mail address: steinbach@zib.de

URL: http://www.zib.de/steinbach