Konrad-Zuse-Zentrum für Informationstechnik Berlin
Heilbronner Str. 10, D-10711 Berlin - Wilmersdorf

Rüdiger Schultz

Strong Convexity in Stochastic Programs with Complete Recourse II: Partially Random Right-Hand Side

Preprint SC 95–21 (September 1995)

Strong Convexity in Stochastic Programs with Complete Recourse II: Partially Random Right-Hand Side*

Rüdiger Schultz

Abstract: We establish a verifiable sufficient condition for strong convexity of the expected recourse as a function of the tender variable in a two-stage stochastic program with linear recourse. Generalizing a former result where all components of the second-stage right-hand side vector were random we treat the case where only a subvector of the right-hand side is random. As prerequisite, a refined analysis of the polyhedral complex of lineality regions of the second-stage value function is carried out. The sufficient condition for strong convexity allows to widen the class of recourse models for which certain quantitative results on stability and asymptotic convergence of optimal solutions are valid.

Key Words: Stochastic programs with complete recourse, strong convexity, polyhedral complex, stability, asymptotic convergence

AMS 1980 subject classification: 90C15, 90C31

* This research is supported by the Deutsche Forschungsgemeinschaft.

1 Introduction

This paper extends the analysis of [18] where strengthened convexity properties of the expected recourse function in a linear stochastic program with complete recourse were established. In contrast to [18], the present paper is directed to models where only parts of the second-stage right-hand side vector are random, a much more realistic assumption in practical applications. Our model reads

$$\min\{g(x) + Q(x) \ : \ x \in C\} \tag{1.1}$$

where

$$Q(x) = \int_\Omega \Phi(z(\omega) - Ax) \ P(d\omega) \tag{1.2}$$

and

$$\Phi(t) = \min\{q^T y \ : \ Wy = t, \ y \in I\!R^{\bar{m}}_+\}. \tag{1.3}$$

Here, $g : I\!R^m \longrightarrow I\!R$ denotes a convex function and $C \subset I\!R^m$ is a nonempty closed convex set. The random vector z acts from a probability space $(\Omega, \mathcal{A}, P)$ to $I\!R^s$ and A, q, W are fixed data with proper dimensions.

It is well known that, under mild assumptions, Q is a convex function on $I\!R^m$ ([6], [26]). Here, we are aiming at verifiable sufficient conditions for improved convexity properties of Q. We will extend the following statement on the function

$$\tilde{Q}(\chi) = \int_\Omega \Phi(z(\omega) - \chi) \ P(d\omega) \tag{1.4}$$

that was established in [18].

Theorem 1.1 *Let $W(I\!R^{\bar{m}}_+) = I\!R^s$ and $\{u \in I\!R^s : W^T u \leq q\}$ have nonempty interior. Suppose further that $\int_\Omega \|z(\omega)\| \ P(d\omega) < \infty$ and that there exist a convex open set $V \subset I\!R^s$, constants $r > 0, \rho > 0$ as well as a density θ of $P \circ z^{-1}$ such that $\theta(t') \geq r$ for all $t' \in I\!R^s$ with dist $(t', V) \leq \rho$. Then $\tilde{Q}$ is strongly convex on V, i.e., there exists some $\kappa > 0$ such that for all $\chi, \chi' \in V$ and all $\lambda \in [0, 1]$*

$$\tilde{Q}(\lambda\chi + (1-\lambda)\chi') \ \leq \ \lambda\tilde{Q}(\chi) + (1-\lambda)\tilde{Q}(\chi') - \kappa\lambda(1-\lambda)\|\chi - \chi'\|^2.$$

Geometrically, this theorem says that $\tilde{Q}$ is locally supported by positive definite quadratic forms instead of merely linear forms (hyperplanes) as would follow from the convexity of $\tilde{Q}$. Since assumptions are in terms of model data, the theorem actually permits verification of this rather implicit property for the quite involved functional $\tilde{Q}$.

A simple but often quite powerful conclusion from the quadratic support property is that strong convexity implies quadratic growth of the objective on neighbourhoods of global minimizers. Conversely, bounds on differences of function values yield bounds on distances of arguments by taking the square root. From the perspective adopted in [1], strong convexity yields a quadratic conditioning of minimizers.

Since Q is constant on translates of the null space of A, strong convexity of Q is possible only for the exceptional case where the null space is $\{0\}$. Therefore, we are heading for strong convexity of the function $\tilde{Q}$ which depends on the tender variable $\chi = Ax$. Stability results in [4], [8], [17],

[20], [21] and convergence rates for simulation based algorithms [22] are based on assumptions that can be verified via the strong convexity of $\tilde{Q}$. This motivates the search for comprehensive classes of recourse models for which strong convexity of $\tilde{Q}$ can be verified. Due to the density assumption, Theorem 1.1 excludes models where components of $z(\omega)$ are deterministic. The latter, however, is met in many applications. The aim of the present paper is to extend our earlier work in [18] to cover that case.

In Section 2, we elaborate consequences of meeting a partially random right-hand side vector z in (1.1) - (1.3). In particular, we study properties of the polyhedral complex of lineality regions of Φ. In Section 3, we adapt the assumptions of Theorem 1.1 to the more general framework and prove our main result, a sufficient condition for strong convexity. In Section 4, we present conclusions and applications of the main result. We show that certain statements now become available for an extended class of recourse problems. These statements concern the structure of the expected recourse function, quantitative stability of optimal solutions and asymptotic properties in the context of estimation via empirical measures. Section 5 contains some final remarks.

2 Recourse Models with Partially Random Right-Hand Side

The two-stage stochastic program (1.1) - (1.3) originates from a random optimization problem

$$\min\{g(x) + q^T y \ : \ Ax + Wy = z(\omega),\ x \in C,\ y \in I\!R^{\bar{m}}_+\} \tag{2.1}$$

where decisions x and y have to be taken before and after the realization of $z(\omega)$, respectively. The idea behind the two-stage (or recourse) approach in stochastic programming is to find an $x \in C$ such that the sum of the direct cost $g(x)$ and the expected recourse cost

$$Q(x) \ = \ E_\omega(\min\{q^T y \ : \ Wy = z(\omega) - Ax,\ y \in I\!R^{\bar{m}}_+\}) \ = \ \int_\Omega \Phi(z(\omega) - Ax)\ P(d\omega)$$

is minimal. For further reading on modeling in stochastic programming we refer to [7] and the references therein.

Now assume that $z(\omega)$ admits a representation $z(\omega) = (z_1(\omega), z_2)$ with $z_1(\omega) \in I\!R^{s_1}$, $z_2 \in I\!R^{s_2}, s_1 + s_2 = s$. Accordingly, we rewrite the random optimization problem (2.1)

$$\min\{g(x) + q^T y \ : \ A_1 x + W_1 y = z_1(\omega),\ A_2 x + W_2 y = z_2,\ x \in C,\ y \in I\!R^{\bar{m}}_+\}$$

from where we derive the expected recourse function

$$Q(x) \ = \ \int_\Omega \Phi(z_1(\omega) - A_1 x, z_2 - A_2 x)\ P(d\omega) \tag{2.2}$$

with

$$\Phi(t_1, t_2) = \min\{q^T y \ : \ W_1 y = t_1,\ W_2 y = t_2,\ y \in I\!R^{\bar{m}}_+\}.$$

In practical applications, the simple constraint $y \in I\!R^{\bar{m}}_+$ in (1.3) often has to be replaced by a general polyhedral constraint on y. Note that this corresponds to setting $A_2 = 0$ in (2.2). The system

$$W_2 = z_2,\ y \in I\!R^{\bar{m}}_+ \tag{2.3}$$

then reflects the polyhedral constraint on y.
We find it convenient to denote the image measure $P \circ z_1^{-1}$ on $I\!R^{s_1}$ by μ. From (2.2) we derive the function

$$\bar{Q}(\chi_1, \chi_2) = \int_{I\!R^{s_1}} \Phi(z_1 - \chi_1, z_2 - \chi_2)\ \mu(dz_1)$$

where $\chi_1 \in I\!R^{s_1}, \chi_2 \in I\!R^{s_2}$. The strong-convexity analysis in the present paper concerns $\bar{Q}$ as a function of χ_1. We fix some $\bar{\chi}_2 \in I\!R^{s_2}$ and denote $\bar{t}_2 = z_2 - \bar{\chi}_2$. The function

$$\tilde{Q}(\chi_1) = \int_{I\!R^{s_1}} \Phi(z_1 - \chi_1, \bar{t}_2)\ \mu(dz_1) \tag{2.4}$$

is our central object of study. For the special case mentioned in (2.3), $\bar{Q}$ does not depend on χ_2, and our investigations concern the full function $\bar{Q}$.
The following basic assumptions are imposed throughout the paper.

(A1) The matrix $W = \binom{W_1}{W_2} \in L(I\!R^{\bar{m}}, I\!R^{s})$ has full rank, and for each $t_1 \in I\!R^{s_1}$ there exists a $y \in I\!R^{\bar{m}}_+$ such that $W_1 y = t_1$ and $W_2 y = \bar{t}_2$.

(A2) There exists a $u = (u_1, u_2) \in I\!R^{s_1} \times I\!R^{s_2}$ such that $W_1^T u_1 + W_2^T u_2 \leq q$.

(A3) It holds $\int_{I\!R^{s_1}} \|z_1\|\ \mu(dz_1) < \infty$.

(A4) The probability measure μ has a density.

(For notational convenience, column vectors often will be written rowwise as done with u in (A2).)
The second part of assumption (A1), often referred to as complete recourse, ensures that the second-stage linear program

$$\min\{q^T y\ :\ W_1 y = t_1,\ W_2 y = \bar{t}_2,\ y \in I\!R^{\bar{m}}_+\}$$

has a feasible point for each $t_1 \in I\!R^{s_1}$. Assumption (A2) states the existence of feasible points to

$$\max\{t_1^T u_1 + \bar{t}_2^T u_2\ :\ W_1^T u_1 + W_2^T u_2 \leq q\}$$

which is the dual to the aforementioned linear program. Hence, (A1) and (A2) imply that $\Phi(t_1) \in I\!R$ for each $t_1 \in I\!R^{s_1}$. Moreover, it is well known from parametric linear programming that Φ is piecewise linear and convex on $I\!R^{s_1}$. Since, later on, the polyhedral complex of lineality regions of Φ will be important, we look a bit closer at Φ.
Since W has full rank, there exist vertices of the (nonempty) feasible set

$$M_D = \{(u_1, u_2) \in I\!R^{s_1 + s_2}\ :\ W_1^T u_1 + W_2^T u_2 \leq q\}.$$

By $\tilde{d}_1, \ldots, \tilde{d}_N$ we denote the vertices of M_D that arise as optimal ones when maximizing over M_D with respect to the objective functions

$$t_1^T u_1 + \bar{t}_2^T u_2,\ \ t_1 \in I\!R^{s_1}.$$

Denoting by $\tilde{d}_{i1}, \tilde{d}_{i2}$ $(i = 1, \ldots, N)$ the projections of $\tilde{d}_i$ on $I\!R^{s_1}$ and $I\!R^{s_2}$, respectively, we obtain (cf. [11], [25])

$$\Phi(t_1) = t_1^T \tilde{d}_{i1} + \bar{t}_2^T \tilde{d}_{i2}$$

for all $t_1 \in I\!R^{s_1}$ such that

$$(t_1, \bar{t}_2) \in \mathcal{K}_i = \{v \in I\!R^s \ : \ v^T(u - \tilde{d}_i) \leq 0 \ \text{ for all } u \in M_D\}.$$

The set $\mathcal{K}_i$ is called the outer normal cone to M_D at the vertex $\tilde{d}_i$. Now, the lineality regions $\mathcal{K}_i^*$ of Φ are given by

$$\mathcal{K}_i^* = \pi_1(\mathcal{K}_i \cap \{I\!R^{s_1} \times \{\bar{t}_2\}\}), \ \ i \in \{1, \ldots, N\} \tag{2.5}$$

where π_1 denotes the projection from $I\!R^s$ to $I\!R^{s_1}$.

Since $\Phi(t_1) \in I\!R$ for any $t_1 \in I\!R^{s_1}$, it holds $\cup_{i=1}^N \mathcal{K}_i^* = I\!R^{s_1}$. Moreover, the sets $\mathcal{K}_i^*$ are the full-dimensional elements of a polyhedral complex ([11], [25]) which, in particular, says that together with a polyhedron all its faces are in the complex and that the intersection of two polyhedra is a face of both. If all components of z are random, i.e., $\Phi(t) = \min\{q^T y \ : \ Wy = t, \ y \in I\!R_+^{\bar{m}}\}$, and if $\Phi(t) \in I\!R$ for all $t \in I\!R^s$, then the polyhedral complex of lineality regions of Φ is especially simple. It is the fan of outer normal cones to the dual polyhedron $\{u \in I\!R^s : W^T u \leq q\}$. Hence it solely consists of polyhedral cones with vertex zero. Difficulties in moving from completely random to partially random right-hand sides in linear two-stage stochastic programs have their roots in the more complicated lineality complex of Φ.

Assumptions (A1) - (A3) imply that $\tilde{Q}(\chi_1) \in I\!R$ for any $\chi_1 \in I\!R^{s_1}$. Of course, $\tilde{Q}$ is also convex on $I\!R^{s_1}$. Well known statements in stochastic programming ([2], [6], [26]) guarantee that $\tilde{Q}$ is continuously differentiable at $\chi_1 \in I\!R^{s_1}$ provided that none of the sets $\chi_1 + \text{bd } \mathcal{K}_i^*$ has positive $\mu-$measure. (Here,"+" denotes Minkowski addition and "bd" the boundary.) Therefore, (A1) - (A4) imply that $\tilde{Q}$ is continuously differentiable on $I\!R^{s_1}$. Denoting $d_i = -\tilde{d}_i$ $(i = 1, \ldots, N)$, the gradient reads

$$\tilde{Q}'(\chi_1) = \sum_{i=1}^{N} d_{i1} \mu(\chi_1 + \mathcal{K}_i^*). \tag{2.6}$$

Later on, this representation will be used to verify strong convexity via strong monotonicity of the gradient. Recall that the recession cone $0^+\mathcal{K}_i^*$ of $\mathcal{K}_i^*$ is defined as the cone of unbounded directions of $\mathcal{K}_i^*$, [14]. The following lemmas provide some useful identities.

Lemma 2.1 *For each $i = 1, \ldots, N$, $0^+\mathcal{K}_i^*$ coincides with $\mathcal{K}_{io}^* = \pi_1(\mathcal{K}_i \cap \{I\!R^{s_1} \times \{0\}\})$.*

Proof: With a suitable matrix $B_i = (B_{i1}, B_{i2})$, the normal cone $\mathcal{K}_i$ to M_D at $\tilde{d}_i$ can be written as

$$\mathcal{K}_i = \{t \in I\!R^s : B_i t \leq 0\} = \{(t_1, t_2) \in I\!R^{s_1 + s_2} : B_{i1} t_1 + B_{i2} t_2 \leq 0\}.$$

Hence, by definition

$$\mathcal{K}_{io}^* = \{t_1 \in I\!R^{s_1} : B_{i1} t_1 \leq 0\}.$$

On the other hand,

$$\mathcal{K}_i^* = \{t_1 \in I\!R^{s_1} : B_{i1}t_1 \le -B_{i2}\bar{t}_2\}$$

implying (cf. [14])

$$0^+\mathcal{K}_i^* = \{t_1 \in I\!R^{s_1} : B_{i1}t_1 \le 0\} = \mathcal{K}_{io}^*.$$

■

Lemma 2.2 *For each $i = 1, \ldots, N$, $\mathcal{K}_{io}^*$ coincides with the outer normal cone to $\pi_1 M_D$ at $\tilde{d}_{i1} = \pi_1 \tilde{d}_i$.*

Proof: The following are equivalent

$$t_1 \in \mathcal{K}_{io}^*$$

and

$$(t_1, 0) \in \mathcal{K}_i$$

and

$$t_1^T \tilde{d}_{i1} + 0^T \tilde{d}_{i2} \ge t_1^T u_1 + 0^T u_2 \ \text{ for all } (u_1, u_2) \in M_D$$

and

$$t_1^T \tilde{d}_{i1} \ge t_1^T u_1 \ \text{ for all } u_1 \in \pi_1 M_D.$$

■

Remark 2.3 *Of course, under the projection of a polyhedron, vertices (if any) are not necessarily sent to vertices. Therefore, the following cases can arise in the above lemmas.*

(i) $\tilde{d}_{i1} \in$ *int* $\pi_1 M_D$. *Then* $\mathcal{K}_{io}^* = \{0\}$, *and* $\mathcal{K}_i^*$ *is bounded.*

(ii) $\tilde{d}_{i1} \in$ *bd* $\pi_1 M_D$, *but* $\tilde{d}_{i1}$ *is not a vertex of* $\pi_1 M_D$. *Then the linear span of* $\mathcal{K}_{io}^*$ *has dimension greater than zero and less than* s_1, *and* $\mathcal{K}_i^*$ *is unbounded with a recession cone of dimension less than* s_1.

(iii) $\tilde{d}_{i1}$ *is a vertex of* $\pi_1 M_D$. *Then* $\mathcal{K}_{io}^*$ *has full dimension* s_1 *and so does the recession cone of the (unbounded) set* $\mathcal{K}_i^*$.

At the following example, we will illustrate the above considerations.
In

$$\min\{q^T y \ : \ W_1 y = t_1, \ W_2 y = z_2, \ y \in I\!R_+^{\bar{m}}\}$$

we put

$$\begin{aligned}
q^T &= (21, 21, 21, 21, 7, 7, 3, 3, 1, 0) \in I\!R^{10}, \\
W_1 &= \begin{pmatrix} -3 & -3 & -3 & 3 & -1 & 1 & 0 & 0 & 0 & 0 \\ -5 & 1 & 2 & -2 & -2 & 2 & -1 & 1 & 0 & 0 \end{pmatrix} \in L(I\!R^{10}, I\!R^2), \\
W_2 &= \begin{pmatrix} 12 & 12 & 9 & 9 & 3 & 3 & 0 & 0 & 1 & -1 \end{pmatrix} \in L(I\!R^{10}, I\!R^1), \\
z_2 &= 1.
\end{aligned}$$

Using the code PORTA (Polyhedron Representation Transformation Algorithm), cf. [3], we compute the vertices of the dual polyhedron M_D together with a vertex-inequality incidence table that displays the binding inequalities for each vertex. The gradients of the binding (linear) inequalities, of course, generate the respective outer normal cone. We obtain the following vertices together with their outer normal cones

$$\begin{aligned}
\tilde{d}_1 &= (-7,0,0), & \mathcal{K}_1 &= \text{cone}\,\{(0,0,-1),(-1,-2,3),(-3,2,9),(-3,1,12),(-3,-5,12)\},\\
\tilde{d}_2 &= (-1,-3,0), & \mathcal{K}_2 &= \text{cone}\,\{(0,0,-1),(0,-1,0),(-1,-2,3)\},\\
\tilde{d}_3 &= (5,-3,0), & \mathcal{K}_3 &= \text{cone}\,\{(0,0,-1),(0,-1,0),(3,-2,9)\},\\
\tilde{d}_4 &= (7,0,0), & \mathcal{K}_4 &= \text{cone}\,\{(0,0,-1),(1,2,3),(3,-2,9)\},\\
\tilde{d}_5 &= (1,3,0), & \mathcal{K}_5 &= \text{cone}\,\{(0,0,-1),(1,2,3),(0,1,0)\},\\
\tilde{d}_6 &= (-5,3,0), & \mathcal{K}_6 &= \text{cone}\,\{(0,0,-1),(0,1,0),(-3,2,9)\},\\
\tilde{d}_7 &= (-3,0,1), & \mathcal{K}_7 &= \text{cone}\,\{(0,0,1),(-3,1,12),(-3,-5,12)\},\\
\tilde{d}_8 &= (2,-3,1), & \mathcal{K}_8 &= \text{cone}\,\{(0,0,1),(0,-1,0),(-1,-2,3),(3,-2,9),(-3,-5,12)\},\\
\tilde{d}_9 &= (4,0,1), & \mathcal{K}_9 &= \text{cone}\,\{(0,0,1),(1,2,3),(3,-2,9)\},\\
\tilde{d}_{10} &= (-2,3,1), & \mathcal{K}_{10} &= \text{cone}\,\{(0,0,1),(1,2,3),(0,1,0),(-3,2,9),(-3,1,12)\}.
\end{aligned}$$

The following picture shows a top view at M_D along the third coordinate. In the picture, the number $i \in \{1,\ldots,10\}$ corresponds to the vertex $\tilde{d}_i$. The marked vertices $\tilde{d}_7,\ldots,\tilde{d}_{10}$ belong to the affine hyperplane with third coordinate equal to 1. For the remaining vertices $\tilde{d}_1,\ldots,\tilde{d}_6$, the third coordinate equals 0. Note that, beside conv $\{\tilde{d}_1,\ldots,\tilde{d}_6\}$, there are two further facets of M_D hidden in the picture: conv $\{\tilde{d}_2,\tilde{d}_3,\tilde{d}_8\}$ and conv $\{\tilde{d}_5,\tilde{d}_6,\tilde{d}_{10}\}$.

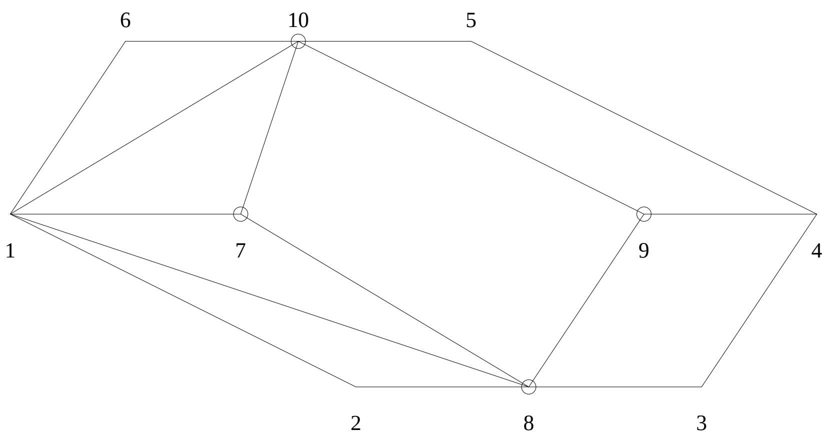

Figure 2.1: Top view at M_D

Computing the cones $\mathcal{K}_{io}^*$ yields

$$\begin{array}{ll}
\mathcal{K}_{1o}^* = \text{cone}\,\{(-3,2),(-1,-2)\}, & \mathcal{K}_{2o}^* = \text{cone}\,\{(-1,-2),(0,-1)\}, \\
\mathcal{K}_{3o}^* = \text{cone}\,\{(0,-1),(3,-2)\}, & \mathcal{K}_{4o}^* = \text{cone}\,\{(3,-2),(1,2)\}, \\
\mathcal{K}_{5o}^* = \text{cone}\,\{(1,2),(0,1)\}, & \mathcal{K}_{6o}^* = \text{cone}\,\{(0,1),(-3,2)\}, \\
\mathcal{K}_{7o}^* = \{0\}, & \mathcal{K}_{8o}^* = \text{cone}\,\{(0,-1)\}, \\
\mathcal{K}_{9o}^* = \{0\}, & \mathcal{K}_{10o}^* = \text{cone}\,\{(0,1)\}.
\end{array}$$

The next picture shows the projection $\pi_1 M_D$ and its fan of outer normal cones. Numbers correspond to the projections $\tilde{d}_{i1}$ and the normal cones $\mathcal{K}_{io}^*$ $(i = 1,\ldots,6)$, respectively. We only marked projections $\tilde{d}_{i1}$ leading to vertices of $\pi_1 M_D$. The projections $\tilde{d}_{81}, \tilde{d}_{101}$ are contained in the relative interior of the edges conv $\{\tilde{d}_{21}, \tilde{d}_{31}\}$ and conv $\{\tilde{d}_{51}, \tilde{d}_{61}\}$, respectively. In terms of normal cones, this means that $\mathcal{K}_{8o}^*, \mathcal{K}_{10o}^*$ are the common facets of $\mathcal{K}_{2o}^*, \mathcal{K}_{3o}^*$ and $\mathcal{K}_{5o}^*, \mathcal{K}_{6o}^*$, respectively. The points $\tilde{d}_{71}, \tilde{d}_{91}$ belong to the interior of $\pi_1 M_D$. Therefore, $\mathcal{K}_{7o}^* = \mathcal{K}_{9o}^* = \{0\}$.

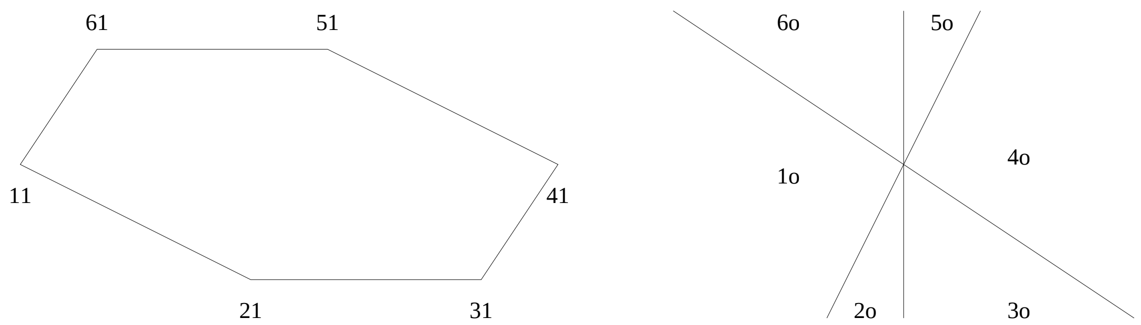

Figure 2.2: Projection $\pi_1 M_D$ and normal fan of $\pi_1 M_D$

Finally, the lineality complex of Φ looks as follows. It arises by intersecting the normal fan of M_D with the affine hyperplane of all points in $I\!R^3$ whose third coordinate equals 1 (recall that $z_2 = 1$ and see (2.5)). Numbers $i \in \{1, \ldots, 10\}$ correspond to lineality regions $\mathcal{K}_i^*$. Note that the normal cones $\mathcal{K}_{io}^*$ reappear as recession cones in the various members of the complex.
In Section 3, the unbounded facets of the full-dimensional members of the complex will be essential for adapting the density assumption from Theorem 1.1. In the example, these are the facets of $\mathcal{K}_2^*, \mathcal{K}_3^*, \mathcal{K}_5^*$ and $\mathcal{K}_6^*$.

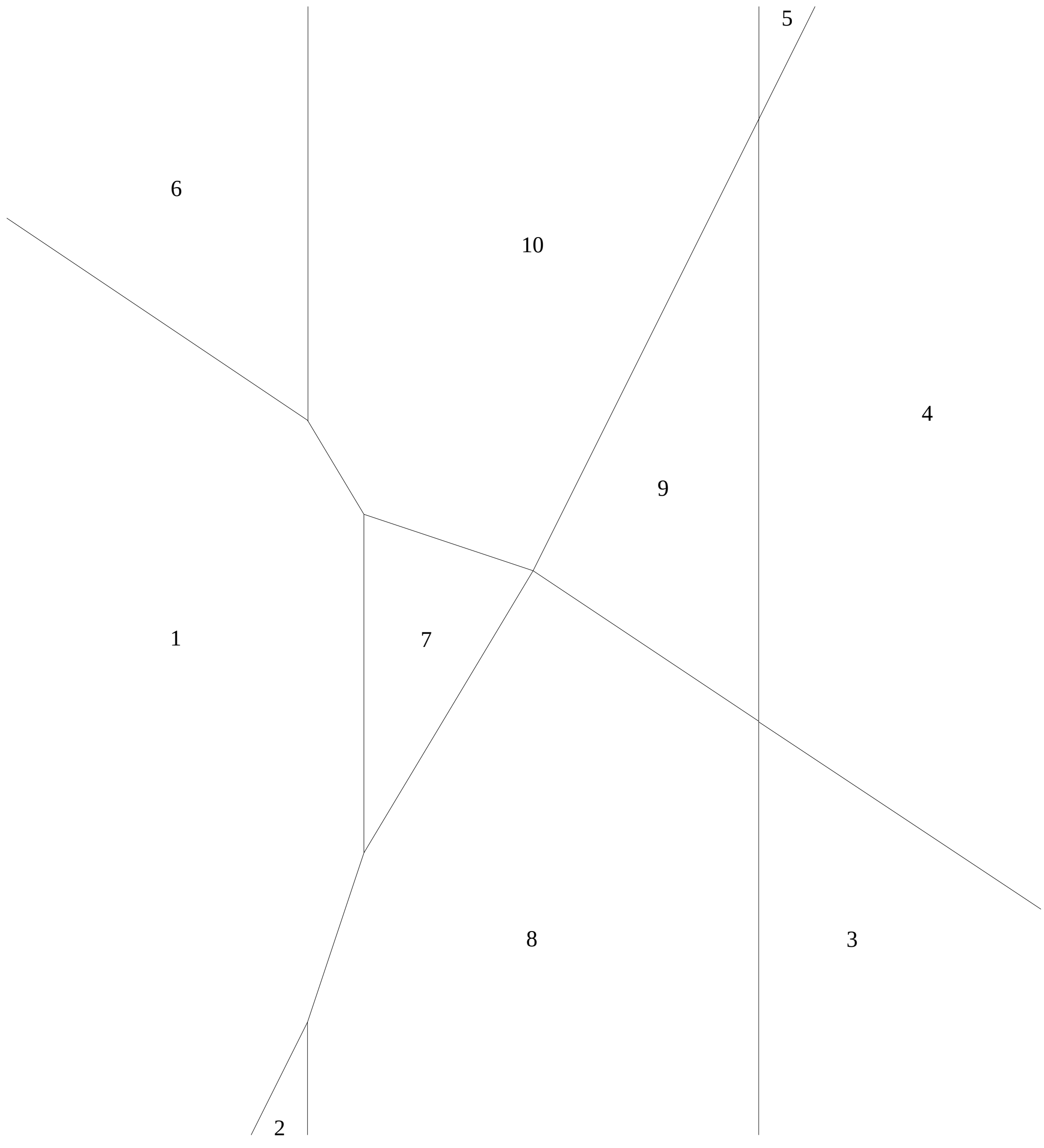

Figure 2.3: Lineality complex of Φ

3 Strong Convexity

Now we turn our attention to the announced sufficient conditions for the strong convexity of the function $\tilde{Q}$ defined in (2.4). We begin with introducing strengthened versions of the assumptions (A2) and (A4) that will be needed for the main result.

Recall that there are finitely many lineality regions $\mathcal{K}_i^*$ whose union makes up the whole of $\mathbb{R}^{s_1}$. Therefore, the recession cones fulfil $\cup_{i=1}^N 0^+\mathcal{K}_i^* = \mathbb{R}^{s_1}$, and Lemmas 2.1 and 2.2 imply that $\pi_1 M_D$ is compact. Note that, under (A1), (A2), M_D itself is not necessarily compact. We strengthen (A2) to

(A2)* The interior int $\pi_1 M_D$ of $\pi_1 M_D$ is nonempty.

Compared with the corresponding assumption in Theorem 1.1, we no longer need a nonempty interior for the whole of M_D but only for a projection related to the multipliers belonging to the rows with random entries in the right-hand side.

As to strengthening (A4), we consider the unbounded facets (i.e., $(s_1-1)-$dimensional faces)

$$\mathcal{F}_{ij}^* = \mathcal{K}_i^* \cap \mathcal{K}_j^*$$

of the full-dimensional members of the lineality complex of Φ. Given $\rho > 0$ and some convex open set $V \subset \mathbb{R}_1^s$ we denote $V_\rho = \{\tau \in \mathbb{R}^{s_1} : \text{dist } (\tau, V) \leq \rho\}$ where "dist" denotes the usual point-to-set distance. Then, (A4) is strengthened to

(A4)* There exist a convex open set $V \subset \mathbb{R}^{s_1}$, constants $r > 0, \rho > 0$, points $e_{ij}^* \in \mathcal{F}_{ij}^*$ and a density θ of μ such that $\theta(t') \geq r$ for all $t' \in \cup_{(i,j)}\{e_{ij}^* + V_\rho\}$.

To relate (A4)* to the corresponding assumption in Theorem 1.1, confirm that there it was claimed $\theta(t') \geq r$ for all $t' \in 0 + V_\rho$. Moreover, in that situation, the lineality complex of Φ is a fan of polyhedral cones all with vertex in 0. Hence, the point 0 belongs to all the unbounded facets of the full-dimensional members of the lineality complex of Φ, and the choice $e_{ij}^* = 0$ is possible for all pairs (i,j) in question.

The following two lemmas will be needed in the proof of our main result. Recall that $\mathcal{K}_{io}^*$ $(i = 1, \ldots, N)$ are outer normal cones to $\pi_1 M_D$ (Lemma 2.2) and let $I_o \subset \{i = 1, \ldots, N\}$ denote the set of all i such that $\mathcal{K}_{io}^*$ has (full) dimension s_1. By Remark 2.3, $\tilde{d}_{i1}$ with $i \in I_o$ are the vertices of $\pi_1 M_D$.

Lemma 3.1 *Assume (A1) and (A2)*. Then it holds*

$$\alpha^* := \min_{j \in I_o} \inf_{\substack{\tilde{v} \in \mathcal{K}_{jo}^* \\ \|\tilde{v}\|=1}} \max_{\substack{i \in I_o \\ d_{i1}, d_{j1} \text{ adjacent}}} (d_{i1} - d_{j1})^T \tilde{v} > 0.$$

Proof: Since $\mathcal{K}_{jo}^*$ is the outer normal cone to $\pi_1 M_D$ at $\tilde{d}_{j1} = -d_{j1}$ and $\tilde{v} \in \mathcal{K}_{jo}^*$, it holds

$$d_{i1}^T \tilde{v} \geq d_{j1}^T \tilde{v} \quad \text{for all} \quad i \in I_o.$$

Hence, $\alpha^* \geq 0$. Assume that $\alpha^* = 0$. Then, for some $j \in I_o$

$$\inf_{\substack{\tilde{v} \in \mathcal{K}_{jo}^* \\ \|\tilde{v}\|=1}} \max_{\substack{i \in I_o \\ d_{i1}, d_{j1} \text{ adjacent}}} (d_{i1} - d_{j1})^T \tilde{v} = 0.$$

Then, for any $n \in I\!N \setminus \{0\}$, there exists a $v_n \in \mathcal{K}^*_{jo}, \|v_n\| = 1$ such that $(d_{i1} - d_{j1})^T v_n \leq 1/n$ for all $i \in I_o$ such that d_{i1} and d_{j1} are adjacent. Passing to the limit yields an accumulation point $\bar{v} \in \mathcal{K}^*_{jo}, \|\bar{v}\| = 1$ of $\{v_n\}$ such that $(d_{i1} - d_{j1})^T \bar{v} = 0$ for all $i \in I_o$ such that d_{i1} and d_{j1} are adjacent. Since $\bar{v} \neq 0$, this implies int $\pi_1 M_D = \emptyset$ contradicting (A2)*. ■

Lemma 3.2 *If there is no vertex $\tilde{d}_l$ of M_D such that $\pi_1 \tilde{d}_l = \tilde{d}_{l1}$ belongs to the relative interior of the edge $[\tilde{d}_{i1}, \tilde{d}_{j1}]$ connecting the vertices $\tilde{d}_{i1}$ and $\tilde{d}_{j1}$ $(i, j \in I_o)$ in $\pi_1 M_D$, then there exists a joint unbounded facet $\mathcal{F}^*_{ij}$ of $\mathcal{K}^*_i$ and $\mathcal{K}^*_j$.*

Proof: Since there is an edge connecting $\tilde{d}_{i1}$ and $\tilde{d}_{j1}$ in $\pi_1 M_D$, the set $\mathcal{F}^o_{ij} = \mathcal{K}^*_{io} \cap \mathcal{K}^*_{jo}$ is a joint facet of $\mathcal{K}^*_{io}$ and $\mathcal{K}^*_{jo}$. If there is no joint unbounded facet of $\mathcal{K}^*_i$ and $\mathcal{K}^*_j$ then $\mathcal{F}^o_{ij}$ arises as recession cone of an $(s_1 - 1)-$dimensional face in a further element of the lineality complex, say $\mathcal{K}^*_k$. Since $\mathcal{K}^*_{io}$ and $\mathcal{K}^*_{jo}$ are the only s_1-dimensional outer normal cones to $\pi_1 M_D$ with facet $\mathcal{F}^o_{ij}$, the set $\mathcal{K}^*_{ko}$ must have exactly dimension $s_1 - 1$. Then, $\pi_1 \tilde{d}_k = \tilde{d}_{k1}$ must belong to the relative interior of the edge $[\tilde{d}_{i1}, \tilde{d}_{j1}]$, a contradiction. ■

Theorem 3.3 *Suppose (A1), (A2)*, (A3) and (A4)*. Then $\tilde{Q}$ is strongly convex on V.*

Proof: The proof will be given via strong monotonicity of the gradient based on the equivalence: $\tilde{Q}$ is strongly convex (with constant $\kappa > 0$) on some convex subset $V \subset I\!R^{s_1}$ if and only if

$$(\tilde{Q}'(\chi_1) - \tilde{Q}'(\chi_1'))^T (\chi_1 - \chi_1') \geq 2\kappa \|\chi_1 - \chi_1'\|^2$$

for all $\chi_1, \chi_1' \in V$ (cf. e.g. [12]).
Let $\chi_1 \in V$ and $v \in I\!R^{s_1}$ such that $\chi_1 + v \in V$. By (2.6),

$$(\tilde{Q}'(\chi_1 + v) - \tilde{Q}'(\chi_1))^T v = \sum_{i=1}^{N} f_i(\chi_1 + v) d_{i1}^T v - \sum_{i=1}^{N} f_i(\chi_1) d_{i1}^T v \tag{3.1}$$

with $f_i(\xi) = \mu(\xi + \mathcal{K}^*_i)$ for all $i = 1, \ldots, N,\ \xi \in I\!R^{s_1}$.
Assumption (A4) and $\cup_{i=1}^N \mathcal{K}^*_i = I\!R^{s_1}$ imply

$$\sum_{i=1}^{N} f_i(\chi_1 + v) = \sum_{i=1}^{N} f_i(\chi_1) = 1.$$

Therefore, the sums in (3.1) can be understood as expectations of two discrete probability measures on $I\!R$ with mass points in $d_{i1}^T v$ and probabilities $f_i(\chi_1 + v)$ and $f_i(\chi_1), i = 1, \ldots, N$, respectively. Denoting the accompanying distribution functions by F_v and $F_{o,v}$, respectively, we obtain

$$(\tilde{Q}'(\chi_1 + v) - \tilde{Q}'(\chi_1))^T v = \int_{I\!R} \tau dF_v(\tau) - \int_{I\!R} \tau dF_{o,v}(\tau) = \int_{I\!R} (F_{o,v}(\tau) - F_v(\tau)) d\tau$$

where we have used integration by parts of Riemann-Stieltjes integrals.
Next, we show that

$$F_{o,v}(\tau) - F_v(\tau) \geq 0 \qquad \text{for all} \quad \tau \in I\!R. \tag{3.2}$$

Let $I_v(\tau) = \{i \in \{1,\dots,N\} : d_{i1}^T v \le \tau\}$. By (A4),

$$F_{o,v}(\tau) = \mu\Big(\bigcup_{i\in I_v(\tau)} \{\chi_1 + \mathcal{K}_i^*\}\Big) \qquad \text{and} \qquad F_v(\tau) = \mu\Big(\bigcup_{i\in I_v(\tau)} \{\chi_1 + v + \mathcal{K}_i^*\}\Big).$$

Therefore, (3.2) is shown when having verified

$$\bigcup_{i\in I_v(\tau)} \{\chi_1 + v + \mathcal{K}_i^*\} \subset \bigcup_{i\in I_v(\tau)} \{\chi_1 + \mathcal{K}_i^*\}.$$

Assume on the contrary that there exist $i_o \in I_v(\tau)$ and $w_{i_o} \in \chi_1 + v + \mathcal{K}_{i_o}^*$ such that

$$w_{i_o} \notin \bigcup_{i\in I_v(\tau)} \{\chi_1 + \mathcal{K}_i^*\}$$

Since $\cup_{\{1,\dots,N\}}\{\chi_1 + \mathcal{K}_i^*\} = \mathbb{R}^{s_1}$, there exists $i_1 \in \{1,\dots,N\} \setminus I_v(\tau)$ with

$$w_{i_o} \in \{\chi_1 + \mathcal{K}_{i_1}^*\}. \tag{3.3}$$

From $w_{i_o} \in \chi_1 + v + \mathcal{K}_{i_o}^*$ it follows $w_{i_o} - \chi_1 - v \in \mathcal{K}_{i_o}^*$ implying (cf. (2.5))

$$(w_{i_o} - \chi_1 - v, \bar{t}_2) \in \mathcal{K}_{i_o}.$$

Now recall that $\mathcal{K}_{i_o}$ is the outer normal cone to M_D at $\tilde{d}_{i_o} = -d_{i_o}$. Therefore, $\tilde{d}_{i_o}$ is an optimal vertex when maximizing over M_D the linear function whose gradient is $(w_{i_o} - \chi_1 - v, \bar{t}_2)$. Thus, we obtain

$$(w_{i_o} - \chi_1 - v)^T d_{i_o1} + \bar{t}_2^T d_{i_o2} \le (w_{i_1} - \chi_1 - v)^T d_{i_11} + \bar{t}_2^T d_{i_12}. \tag{3.4}$$

By $i_1 \notin I_v(\tau)$ and $i_o \in I_v(\tau)$,

$$d_{i_11}^T v > \tau \ge d_{i_o1}^T v,$$

and (3.4) implies

$$\begin{aligned}(w_{i_o} - \chi_1)^T d_{i_o1} + \bar{t}_2^T d_{i_o2} &\le (w_{i_o} - \chi_1)^T d_{i_11} + \bar{t}_2^T d_{i_12} + d_{i_o1}^T v - d_{i_11}^T v \\ &< (w_{i_o} - \chi_1)^T d_{i_11} + \bar{t}_2^T d_{i_12}.\end{aligned}$$

On the other hand, we have by (3.3)

$$(w_{i_o} - \chi_1)^T d_{i_11} + \bar{t}_2^T d_{i_12} \le (w_{i_o} - \chi_1)^T d_{i_o1} + \bar{t}_2^T d_{i_o2}$$

yielding a contradiction. Thus, (3.2) is shown.
The further strategy for the proof is deriving a lower bound for

$$(\tilde{Q}'(\chi_1 + v) - \tilde{Q}'(\chi_1))^T v = \int_{\mathbb{R}} (F_{o,v}(\tau) - F_v(\tau))d\tau \tag{3.5}$$

by properly restricting the domain of integration and minorizing the integrand on the right. Recall the sets $\mathcal{K}_{io}^* \subset \mathbb{R}^{s_1}$ arising in Lemmas 2.1, 2.2 and that I_o is the set of all $i \in \{1,\dots,N\}$ such that $\mathcal{K}_{io}^*$ has (full) dimension s_1. Since $\cup_{i\in I_o}\mathcal{K}_{io}^* = \mathbb{R}^{s_1}$, there exists a $j = j(v) \in I_o$ such that $v \in \mathcal{K}_{jo}^*$.

By Lemma 3.1, there exists $i_* = i_*(v) \in I_o$ such that d_{j1} and d_{i_*1} are adjacent in $-\pi_1 M_D$ and

$$d_{j1}^T v + \alpha^* \|v\| < d_{i_*1}^T v. \tag{3.6}$$

Now let $l = l(v) \in \{1, \ldots, N\}$ be such that $\pi_1 d_l = d_{l1}$ belongs to the half-open line segment $[d_{j1(v)}, d_{i_*1(v)})$ and is nearest to $d_{i_*1(v)}$ of all vertices with this property. Then, there exists $\beta(v)$ with $0 \leq \beta(v) < 1$ such that

$$d_{l1(v)} = d_{j1(v)} + \beta(v)(d_{i_*1(v)} - d_{j1(v)}).$$

Since for arbitrary $v' \in I\!\!R^{s_1}$ there are only finitely many possible edges $[d_{j1(v')}, d_{i_*1(v')}]$ we have

$$\beta = \max_{v' \in I\!\!R^{s_1}} \beta(v') < 1.$$

Now, the following holds

$$d_{i_*1(v)}^T v - d_{l1(v)}^T v = (1 - \beta(v))(d_{i_*1(v)}^T v - d_{j1(v)}^T v) > (1 - \beta)\alpha^* \|v\|. \tag{3.7}$$

Fix some $\alpha > 0$ such that $\alpha < (1 - \beta)\alpha^*$. With $j(v), i_*(v)$ and $l(v)$ as above, we estimate in (3.5)

$$(\tilde{Q}'(\chi_1 + v) - \tilde{Q}'(\chi_1))^T v \geq \int_{d_{l1}^T v}^{d_{l1}^T v + \alpha \|v\|} (F_{o,v}(\tau) - F_v(\tau)) d\tau. \tag{3.8}$$

Consider $\mathcal{F}_{i_*j}^o = \mathcal{K}_{i_*o}^* \cap \mathcal{K}_{jo}^*$ which is a joint facet of $\mathcal{K}_{i_*o}^*$ and $\mathcal{K}_{jo}^*$. Since $\mathcal{K}_{i_*o}^* = 0^+ \mathcal{K}_{i_*}^*$ (Lemma 2.2), there exists a point $e_{i_*j}^*$ arising in assumption (A4)* such that $e_{i_*j}^* + \mathcal{F}_{i_*j}^o$ is contained in a facet of $\mathcal{K}_{i_*}^*$.
Let $\tau \in I\!\!R$ be such that

$$d_{l1}^T v \leq \tau \leq d_{l1}^T v + \alpha \|v\|. \tag{3.9}$$

Then, it holds

$$\bigcup_{0<\lambda<1} \{\chi_1 + \lambda v + e_{i_*j}^* + \mathcal{F}_{i_*j}^o\} \subset \bigcup_{i \in I_v(\tau)} \{\chi_1 + \mathcal{K}_i^*\} \setminus \bigcup_{i \in I_v(\tau)} \{\chi_1 + v + \mathcal{K}_i^*\}. \tag{3.10}$$

To establish this inclusion we first show that

$$\bigcup_{0<\lambda<1} \{\chi_1 + \lambda v + e_{i_*j}^* + \mathcal{F}_{i_*j}^o\} \cap \bigcup_{i \in I_v(\tau)} \{\chi_1 + v + \mathcal{K}_i^*\} = \emptyset. \tag{3.11}$$

Assume on the contrary that there exist $\lambda \in (0, 1), \tau \in I\!\!R$ fulfilling (3.9), $i \in I_v(\tau)$, $u_i \in \mathcal{K}_i^*$ and $w_{i_*j} \in \mathcal{F}_{i_*j}^o$ such that

$$\chi_1 + \lambda v + e_{i_*j}^* + w_{i_*j} \ = \ \chi_1 + v + u_i.$$

Then

$$u_i \ = \ -(1 - \lambda)v + e_{i_*j}^* + w_{i_*j} \ \in \ \mathcal{K}_i^*$$

implying

$$(-(1 - \lambda)v + e_{i_*j}^* + w_{i_*j}, \bar{t}_2) \ \in \ \mathcal{K}_i.$$

Therefore,

$$(-(1-\lambda)v + e^*_{i_*j} + w_{i_*j})^T d_{i1} + \bar{t}_2^T d_{i2} \;\le\; (-(1-\lambda)v + e^*_{i_*j} + w_{i_*j})^T d_{i_*1} + \bar{t}_2^T d_{i_*2}.$$

This implies

$$(e^*_{i_*j} + w_{i_*j})^T d_{i1} + \bar{t}_2^T d_{i2} - (e^*_{i_*j} + w_{i_*j})^T d_{i_*1} - \bar{t}_2^T d_{i_*2} \;\le\; (1-\lambda)(d_{i1}^T v - d_{i_*1}^T v) \;<\; 0 \tag{3.12}$$

where the last inequality follows from $i \in I_v(\tau)$ and $d_{i_*1}^T v > \tau$ (cf. (3.7) and (3.9)).
On the other hand,

$$e^*_{i_*j} + w_{i_*j} \;\in\; e^*_{i_*j} + \mathcal{F}^o_{i_*j} \;\subset\; \mathcal{K}^*_{i_*}$$

and, hence,

$$(e^*_{i_*j} + w_{i_*j}, \bar{t}_2) \;\in\; \mathcal{K}_{i_*}$$

implying

$$(e^*_{i_*j} + w_{i_*j})^T d_{i_*1} + \bar{t}_2^T d_{i_*2} \;\le\; (e^*_{i_*j} + w_{i_*j})^T d_{i1} + \bar{t}_2^T d_{i2}$$

in contradiction to (3.12). This verifies (3.11).
Let us now show that

$$\bigcup_{0<\lambda<1} \{\chi_1 + \lambda v + e^*_{i_*j} + \mathcal{F}^o_{i_*j}\} \;\subset\; \bigcup_{i \in I_v(\tau)} \{\chi_1 + \mathcal{K}^*_i\}. \tag{3.13}$$

Assume on the contrary that there exist $\lambda \in (0,1), \tau \in I\!R$ fulfilling (3.9), $w_{i_*j} \in \mathcal{F}^o_{i_*j}$, $k \in \{1,\ldots,N\} \setminus I_v(\tau)$ such that

$$\chi_1 + \lambda v + e^*_{i_*j} + w_{i_*j} \;\in\; \chi_1 + \mathcal{K}^*_k.$$

Since $e^*_{i_*j} + \mathcal{F}^o_{i_*j}$ is contained in a facet of $\mathcal{K}^*_{i_*}$, there exists a full-dimensional polyhedron $\mathcal{K}^*_{l_*}$ such that $e^*_{i_*j} + \mathcal{F}^o_{i_*j}$ is contained in a joint facet of $\mathcal{K}^*_{l_*}$ and $\mathcal{K}^*_{i_*}$.
Now, it holds

$$(\lambda v + e^*_{i_*j} + w_{i_*j})^T d_{k1} + \bar{t}_2^T d_{k2} \;\le\; (\lambda v + e^*_{i_*j} + w_{i_*j})^T d_{l_*1} + \bar{t}_2^T d_{l_*2}$$

from where it follows

$$(e^*_{i_*j} + w_{i_*j})^T d_{k1} + \bar{t}_2^T d_{k2} - (e^*_{i_*j} + w_{i_*j})^T d_{l_*1} - \bar{t}_2^T d_{l_*2} \;\le\; \lambda(d_{l_*1}^T v - d_{k1}^T v). \tag{3.14}$$

Two cases may arise: Either $l_* = j$ or $l_* \neq j$.
If $l_* = j$ then

$$d_{l_*1}^T v - d_{k1}^T v \;=\; d_{j1}^T v - d_{k1}^T v \;<\; 0,$$

since $j \in I_v(\tau)$ and $k \notin I_v(\tau)$ which yields $d_{j1}^T v \le \tau < d_{k1}^T v$.
If $l_* \neq j$, then there exists no joint unbounded facet of $\mathcal{K}^*_{i_*}$ and $\mathcal{K}^*_j$, and Lemma 3.2 applies. Hence, d_{l_*1} belongs to the interior of the edge $[d_{j1}, d_{i_*1}]$ in $-\pi_1 M_D$. By construction of $l(v)$ it follows

$$d_{l_*1}^T v \;\le\; d_{l1}^T v \;\le\; \tau \;<\; d_{k1}^T v$$

and again

$$d_{l_*1}^T v - d_{k1}^T v \ < \ 0.$$

Now, (3.14) yields

$$(e_{i_*j}^* + w_{i_*j})^T d_{k1} + \bar{t}_2^T d_{k2} \ < \ (e_{i_*j}^* + w_{i_*j})^T d_{l_*1} + \bar{t}_2^T d_{l_*2}. \tag{3.15}$$

On the other hand,

$$e_{i_*j}^* + w_{i_*j} \ \in \ \mathcal{K}_{l_*}^*$$

implying

$$(e_{i_*j}^* + w_{i_*j}, \bar{t}_2) \ \in \ \mathcal{K}_{l_*}$$

and, therefore,

$$(e_{i_*j}^* + w_{i_*j})^T d_{l_*1} + \bar{t}_2^T d_{l_*2} \ \leq \ (e_{i_*j}^* + w_{i_*j})^T d_{k1} + \bar{t}_2^T d_{k2}.$$

This contradicts (3.15) and finally verifies (3.10).
By (A4) and (3.10) we obtain for all $\tau \in I\!R$ with $d_{l1}^T v \leq \tau \leq d_{l1}^T v + \alpha\|v\|$

$$\begin{aligned} F_{o,v}(\tau) - F_v(\tau) \ &= \ \mu\Big(\bigcup_{i \in I_v(\tau)} \{\chi_1 + \mathcal{K}_i^*\} \setminus \bigcup_{i \in I_v(\tau)} \{\chi_1 + v + \mathcal{K}_i^*\}\Big) \\ &\geq \ \mu\Big(\bigcup_{0 \leq \lambda \leq 1} \{\chi_1 + \lambda v + e_{i_*j}^* + \mathcal{F}_{i_*j}^o\}\Big). \end{aligned}$$

The last expression does not depend on τ. Hence, we can continue the estimate in (3.8) as follows

$$(\tilde{Q}'(\chi_1 + v) - \tilde{Q}'(\chi_1))^T v \ \geq \ \alpha \cdot \|v\| \cdot \mu\Big(\bigcup_{0 \leq \lambda \leq 1} \{\chi_1 + \lambda v + e_{i_*j}^* + \mathcal{F}_{i_*j}^o\}\Big). \tag{3.16}$$

The cylindric set

$$\bigcup_{0 \leq \lambda \leq 1} \{\chi_1 + \lambda v + e_{i_*j}^* + \mathcal{F}_{i_*j}^o\}$$

is located between the affine hyperplanes

$$\chi_1 + e_{i_*j}^* + \text{span}\ \mathcal{F}_{i_*j}^o \qquad \text{and} \qquad \chi_1 + v + e_{i_*j}^* + \text{span}\ \mathcal{F}_{i_*j}^o \tag{3.17}$$

whose distance can be minorized as follows:
Since $\mathcal{F}_{i_*j}^o = \mathcal{K}_{i_*o}^* \cap \mathcal{K}_{jo}^*$ and $\mathcal{K}_{i_*o}^*, \mathcal{K}_{jo}^*$ are outer normal cones to $\pi_1 M_D$ at $-d_{i_*1}$ and $-d_{j1}$, respectively, it holds

$$w_{i_*j}^T d_{i_*1} \ = \ w_{i_*j}^T d_{j1} \qquad \text{for all} \ \ w_{i_*j} \in \mathcal{F}_{i_*j}^o$$

implying

$$(\text{span}\ \mathcal{F}_{i_*j}^o)^\perp \ = \ \text{span}\ (d_{i_*1} - d_{j1})$$

where the superscript "$\perp$" denotes the orthogonal complement.
The Hausdorff distance of the hyperplanes in (3.17) thus computes as

$$\frac{(d_{i_*1} - d_{j1})^T v}{\|d_{i_*1} - d_{j1}\|}.$$

With the positive constant

$$\delta = \max\{\|d_{i_1 1} - d_{i_2 1}\| \ : \ i_1, i_2 \in I_o,\ d_{i_1 1}, d_{i_2 1} \text{ adjacent}\}$$

we obtain (cf. also Lemma 3.1)

$$\frac{(d_{i_* 1} - d_{j1})^T v}{\|d_{i_* 1} - d_{j1}\|} \ \geq \ \frac{\alpha^* \|v\|}{\delta} \ > \ \frac{\alpha \|v\|}{\delta}.$$

This minorizes the "height of the cylinder"

$$\bigcup_{0 \leq \lambda \leq 1} \{\chi_1 + \lambda v + e^*_{i_* j} + \mathcal{F}^o_{i_* j}\}.$$

Now consider

$$\mathcal{F}^{o,\rho}_{i_* j} \ = \ \{\tilde{w} \in \mathcal{F}^o_{i_* j} \ : \ \|\tilde{w}\| \leq \rho\}.$$

Assumption (A4)* guarantees that there is a density θ of μ which, on the set

$$\bigcup_{0 \leq \lambda \leq 1} \{\chi_1 + \lambda v + e^*_{i_* j} + \mathcal{F}^{o,\rho}_{i_* j}\},$$

is greater than or equal to some $r > 0$.
Therefore,

$$\mu\Big(\bigcup_{0 \leq \lambda \leq 1} \{\chi_1 + \lambda v + e^*_{i_* j} + \mathcal{F}^o_{i_* j}\}\Big) \ \geq \ \mu\Big(\bigcup_{0 \leq \lambda \leq 1} \{\chi_1 + \lambda v + e^*_{i_* j} + \mathcal{F}^{o,\rho}_{i_* j}\}\Big) \ \geq r \cdot \delta^{-1} \cdot \alpha \cdot \|v\| \cdot \ell_{s_1 - 1}(\mathcal{F}^{o,\rho}_{i_* j})$$

where ℓ_{s_1-1} denotes the $(s_1 - 1)-$ dimensional Lebesgue measure.
Since there are only finitely many facets $\mathcal{F}^o_{i_1 i_2}$, $(i_1, i_2 \in I_o)$, the constant

$$\ell_{min} = \min\{\ell_{s_1-1}(\mathcal{F}^{o,\rho}_{i_1 i_2}) \ : \ i_1, i_2 \in I_o,\ d_{i_1 1}, d_{i_2 1} \text{ adjacent}\}$$

is positive.
Together with (3.16), we finally obtain

$$(\tilde{Q}'(\chi_1 + v) - \tilde{Q}'(\chi_1))^T v \ \geq \ r \cdot \ell_{min} \cdot \delta^{-1} \cdot \alpha^2 \cdot \|v\|^2$$

for all $\chi_1 \in V$ and $v \in \mathbb{R}^{s_1}$ such that $\chi_1 + v \in V$. The constants ℓ_{min}, δ and α are independent on χ_1 and v. Hence, $\tilde{Q}$ is strongly convex on V with constant $\kappa = \frac{1}{2} r \cdot \ell_{min} \cdot \delta^{-1} \cdot \alpha^2$. ∎

Remark 3.4 *Note that the representation we obtained for κ, in a certain sense, confirms the empirical observation in stochastic programming that expected recourse functions are usually "quite flat" in a region surrounding the optimal solution. Indeed $\alpha, \ell_{min}, \delta^{-1}$ are the bigger the "rounder" $\pi_1 M_D$ and r is the bigger the more concentrated the measure μ.*

Compared with the density assumption in Theorem 1.1, the verification of (A4)* is a much more difficult task. Of course, (A4)* holds for distributions like the multivariate normal who have densities that are strictly positive on arbitrary compact sets.

A possible choice for the points e^*_{ij} in (A4)* are *vertices* of unbounded facets in the lineality complex of Φ. (For the example in Section 2, for instance, it is sufficient to take as e^*_{ij} the vertices of $\mathcal{K}^*_2, \mathcal{K}^*_3, \mathcal{K}^*_5$ and $\mathcal{K}^*_6$.) Of course, in general such vertices are not known explicitly. Then, it is sufficient to guarantee strict positivity of θ at some set $B + V_\rho$ where B is a compact set containing all the vertices in the lineality complex of Φ. In what follows, we sketch a respective estimation procedure.
The outer normal cone $\mathcal{K}_i$ to M_D at $\tilde{d}_i$ admits a representation

$$\mathcal{K}_i \;=\; \{u \in I\!R^s \;:\; u = W(i)v, v \geq 0\}$$

where $W(i) \in L(I\!R^{m_i}, I\!R^s)$ is formed by the columns of W leading to constraints of $M_D = \{u \in I\!R^s : W^T u \leq q\}$ that are active in $\tilde{d}_i$. Then it holds

$$\begin{aligned}\mathcal{K}^*_i \;&=\; \{u_1 \in I\!R^{s_1} \;:\; u_1 = W(i)_1 v,\; \bar{t}_2 = W(i)_2 v,\; v \geq 0\} \\ &=\; W(i)_1(\{v \in I\!R^{m_i} \;:\; W(i)_2 v = \bar{t}_2,\; v \geq 0\}).\end{aligned}$$

hence, the vertices of $\mathcal{K}^*_i$ are among the $W(i)_1-$images of the vertices $v(i)_j$ $(j = 1, \ldots, J_i)$ of $\{v \in I\!R^{m_i} \;:\; W(i)_2 v = \bar{t}_2,\; v \geq 0\}$. With some basis submatrix $W(i,j)_2$ of $W(i)_2$, the vertex $v(i)_j$ can be written as

$$v(i)_j \;=\; (W(i,j)_2^{-1}\bar{t}_2, 0),$$

and we obtain

$$W(i)_1 v(i)_j = W(i,j)_1 W(i,j)_2^{-1}\bar{t}_2$$

where $W(i,j)_1$ is given by the columns in $W(i)_1$ corresponding to the basis variables in $v(i)_j$.
An upper bound for the vertices in the lineality complex of Φ is then given by

$$\max_{i=1,\ldots,N} \;\max_{j=1,\ldots,J_i} \|W(i,j)_1\| \cdot \|W(i,j)_2^{-1}\| \cdot \|\bar{t}_2\|. \tag{3.18}$$

Of course, $W(i,j)_2$ is a non-singular square submatrix of W_2 and $W(i,j)_1$ is formed by the corresponding columns in W_1. The above bound, hence, is relaxed by

$$\max_{k=1,\ldots,K} \|W(k)_1\| \cdot \|W(k)_2^{-1}\| \cdot \|\bar{t}_2\| \tag{3.19}$$

where $W(1)_2, \ldots, W(K)_2$ are *all* the invertible submatrices of W_2 and $W(1)_1, \ldots, W(K)_1$ the corresponding submatrices of W_1.
Both (3.18) and (3.19), of course, are practicable for problems of moderate size only. The code PORTA ([3]), used for the computations in the example at the end of Section 2, can be employed to calculate the matrices $W(i)$ and the vertices $v(i)_j$.
For the example in the previous section, the bound (3.19) is easily computed by considering the non-zero elements in W_2 together with the corresponding columns of W_1. Using the $\|\cdot\|_\infty-$norm, the bound computes as 2/3. It is attained by the vertices of $\mathcal{K}^*_2, \mathcal{K}^*_5$.

4 Conclusions and Applications

In this section, we present some consequences of Theorem 3.3. First, we derive a sufficient condition for the strict convexity of $\tilde{Q}$. Recall that $\tilde{Q}$ is strictly convex on some convex set $V \subset I\!R^{s_1}$ if

$$\tilde{Q}(\lambda\chi_1 + (1-\lambda)\chi_1') \;<\; \lambda\tilde{Q}(\chi_1) + (1-\lambda)\tilde{Q}(\chi_1')$$

for all $\chi_1, \chi_1' \in V$, $\chi_1 \neq \chi_1'$ and all $\lambda \in I\!R$ with $0 < \lambda < 1$.

Corollary 4.1 *Assume (A1), (A2)*, (A3), (A4). Suppose that for each unbounded facet $\mathcal{F}_{ij}^*$ of each full-dimensional member in the lineality complex of Φ, there exists a point $e_{ij}^* \in \mathcal{F}_{ij}^*$ such that, for some convex open set $V \subset I\!R^{s_1}$, $\cup_{(i,j)}\{e_{ij}^* + V\} \;\subset\; supp\,\mu$, where $supp\,\mu$ is the smallest closed set in $I\!R^{s_1}$ with $\mu-$measure 1. Then $\tilde{Q}$ is strictly convex on V.*

Proof: Repeat the proof of Theorem 3.3 until (3.16). The above additonal assumption on μ then implies that

$$\mu\Big(\bigcup_{0\le\lambda\le 1}\{\chi_1 + \lambda v + e_{i_*j}^* + \mathcal{F}_{i_*j}^o\}\Big) \;>\; 0$$

for all $\chi_1 \in V$ and all $v \in I\!R^{s_1}$ with $\chi_1 + v \in V$.
Therefore,

$$(\tilde{Q}'(\chi_1 + v) - \tilde{Q}'(\chi_1))^T v \;>\; 0$$

for the mentioned χ_1 and v, and $\tilde{Q}'$ is strictly monotone on V. Hence (e.g. [12]), $\tilde{Q}$ is strictly convex on V. ■

Inequality (3.16), in principal, allows verification of further strengthened versions of convexity arising in the literature. Such notions are characterized by strictly monotone terms that replace the quadratic form in the definition of strong convexity. In view of (3.16), the key for verifying strengthened convexity lies in minorizing $\|v\|\cdot\mu(\bigcup_{0\le\lambda\le 1}\{\chi_1+\lambda v+e_{i_*j}^*+\mathcal{F}_{i_*j}^o\})$ by the mentioned strictly monotone terms.

Theorem 3.3 turns out an essential prerequisite for verifying strong convexity in recourse models (1.1) - (1.3) where, beside the vector $z(\omega)$ also components of the technology matrix A are random. Let us assume that the second-stage problem has the form (cf. (2.3))

$$\Phi(t) = \min\{q^T y \;:\; W_1 y = t,\; W_2 y = z_2,\; y \in I\!R_+^{\bar{m}}\} \tag{4.1}$$

and that Q is given by

$$Q(x) = \int_\Omega \Phi(z(\omega) - A(\omega)x)\,P(d\omega) = \int_{I\!R^{(m+1)s}} \Phi(z - Ax)\,\mu(d(z,A)).$$

We represent $A(\omega) \in L(I\!R^m, I\!R^s)$ as $A(\omega) = (A_o, A_1(\omega))$ where $A_o \in L(I\!R^k, I\!R^s)$, $A_1(\omega) \in L(I\!R^{m-k}, I\!R^s), 0 \le k \le m$, and all components in A_o are non-random (constant $P-$almost surely).
The following theorem from [18] provides a sufficient condition for the strong convexity of Q in terms of strong convexity of

$$\tilde{Q}_A(\chi) = \int_{I\!R^s} \Phi(z - \chi)\mu_1^2(A, dz) \tag{4.2}$$

and conditions on A_o and $A_1(\omega)$. In (4.2), $\mu_1^2(A,.)$ denotes the regular conditional distribution of z given A ([5]). By μ_2, we will denote the marginal distribution of μ with respect to A.

Theorem 4.2 *Assume (A1), (A2), (A3). Let $V \subset \mathbb{R}^m$ be convex and suppose*

(i) for μ_2-almost all $A \in \mathbb{R}^{ms}$, $\tilde{Q}_A$ is strongly convex on $\tilde{V} := A(V)$ with some modulus $\kappa(A)$, and there exists some $\kappa > 0$ such that $\kappa(A) \geq \kappa$ for μ_2-almost all A;

(ii) $\int_{\mathbb{R}^{ms}} \|A\|^2 \, \mu_2(dA) < \infty$;

(iii) $k \leq s$, i.e., in $A_1(\omega)$ there are at least $m - s$ columns;

(iv) A_o has full rank;

(v) the matrix $\int_{\mathbb{R}^{ms}} A_1^T A_1 \, \mu_2(dA) - (\int_{\mathbb{R}^{ms}} A_1 \, \mu_2(dA))^T (\int_{\mathbb{R}^{ms}} A_1 \, \mu_2(dA))$ is positive definite.

Then Q is strongly convex on V.

Theorem 3.3 extends the class of problems for which verification of the crucial assumption (i) is possible. Further issues related to Theorem 4.2 are discussed in [18].

As mentioned in the introduction, strong convexity is closely related to quadratic growth which, for instance, arises as a central assumption in quantitative stability results for optimal solution points of perturbed stochastic programs. By Theorem 3.3, stability results are now available for larger problem classes. In what follows, we illustrate this at some examples. Again we confine ourselves to the special case arising in (2.3), i.e., the second stage in (1.1) - (1.3) has the form

$$\Phi(t) = \min\{q^T y \; : \; W_1 y = t, \; W_2 y = z_2, \; y \in \mathbb{R}_+^{\bar{m}}\}.$$

In [21], Shapiro introduces the following growth condition:
With φ, Ψ denoting the optimal value and optimal solution set to (1.1) - (1.3), let there exist an open convex set $U \supset \Psi$ and a constant $\alpha > 0$ such that

$$g(x) + Q(x) \quad \geq \quad \varphi + \alpha \cdot \text{dist}\,(x, \Psi)^2 \quad \text{for all} \;\; x \in C \cap U. \tag{4.3}$$

Using similar techniques as in the proof of Theorem 2.7 in [16] it can be shown that this condition is fulfilled if Ψ is nonempty, bounded, g is convex quadratic, C is a polyhedron and $\tilde{Q}(\chi) = \int_{\mathbb{R}^s} \Phi(z - \chi)\mu(dz)$ is strongly convex on some convex open set $V \supset A(U)$. In this respect, (4.3) is closely related to our Theorem 3.3.
Results based on (4.3) include a stability estimate for perturbed solutions in stochastic programs with complete recourse ([21], Theorem 2.1), an asymptotic expansion of optimal solutions to recourse models with estimated underlying probability measure ([20], Theorem 2.1) and asymptotic expansions of the stochastic error in simulation based optimization methods ([22], Section 4).
In [17], (1.1) - (1.3) is considered as a parametric program with respect to the underlying probability measure μ that varies in the space $\mathcal{P}(\mathbb{R}^s)$ of all Borel probability measures on $\mathbb{R}^s$.

The optimal value $\varphi = \varphi(\mu)$ and the optimal solution set $\Psi = \Psi(\mu)$ then are (multi)functions of μ. The central result in [17] asserts Lipschitz continuity of the multifunction Ψ with respect to a certain distance of probability measures. Strong convexity of $\tilde{Q}$ in the unperturbed problem is a crucial assumption in this respect. In [17], all components of the second-stage right-hand side vector z are random. For partially random right-hand side, the following counterpart to Corollary 2.5 in [17] is valid. (As usual, "d_H" will denote the Hausdorff distance of sets.)

Proposition 4.3 *Let $\Psi(\mu)$ be nonempty, bounded, g be convex quadratic, C be a polyhedron and assume that $\tilde{Q}$ is strongly convex on some convex open set $V \subset I\!R^s$ such that $A(\Psi(\mu)) \subset V$. Then there exists a constant $L > 0$ such that*

$$d_H(\Psi(\mu), \Psi(\nu)) \quad \leq \quad L \cdot \sum_{i=1}^{N} \sup_{\tau \in I\!R^s} |\mu(\tau + \mathcal{K}_i^*) - \nu(\tau + \mathcal{K}_i^*)| \tag{4.4}$$

whenever $\nu \in \mathcal{P}(I\!R^s)$ is chosen such that $\int_{I\!R^s} \|z\| \; \nu(dz) < \infty$ and that the right-hand side is sufficiently small.

Proof: As in the proof of Corollary 2.5 in [17] we obtain as a consequence of Theorem 2.4 in [17] that

$$d_H(\Psi(\mu), \Psi(\nu)) \quad \leq \quad L \cdot \sup\{\|(\tilde{Q}_\nu - \tilde{Q}_\mu)'(Ax)\| : \; x \in U \setminus E\}$$

provided that $\int_{I\!R^s} \|z\| \; \nu(dz) < \infty$ and that the right-hand side is sufficiently small. Here, $L > 0$ is some constant, $U \subset I\!R^m$ the closure of an open set containing $\Psi(\mu)$ for which $A(U) \subset V$ and E the set of those $x \in I\!R^m$ such that $\tilde{Q}_\nu - \tilde{Q}_\mu$ is not differentiable at Ax. The dependence of $\tilde{Q}$ on the underlying probabbility measure is indicated by subscripts.
Now apply the gradient formula (2.6). It differs from the gradient formula in [17] by the more complicated lineality regions of Φ but this has no impact on the validity of the argument in the proof of Corollary 2.5 in [17]. Therefore, we finally end up with the sets $\mathcal{K}_i^*$ instead of lower left orthants that were relevant in [17]. ■

As an application of the above proposition we consider estimation via empirical measures. Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of independent $I\!R^s$-valued random variables on the probability space $(\Omega, \mathcal{A}, P)$ with joint distribution μ. The empirical measures $\mu_n(\omega)$ ($\omega \in \Omega$, $n \in I\!N$) are defined by

$$\mu_n(\omega) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i(\omega)}$$

where $\delta_{\xi_i(\omega)}$ denotes the measure with unit mass at $\xi_i(\omega)$ (cf. e.g. [5], [13], [23]).
We are interested in aymptotic properties of $d_H(\psi(\mu), \psi(\mu_n(\cdot)))$ as n tends to infinity. Theorem 2K in [15] implies both $\mathcal{A}-$measurability of the mappings

$$\omega \;\mapsto\; d_H(\psi(\mu), \psi(\mu_n(\omega))), \quad n \in I\!N,$$

and

$$\omega \;\mapsto\; \sup_{\tau \in I\!R^s} |\mu(\tau + \mathcal{K}_i^*) - \mu_n(\omega)(\tau + \mathcal{K}_i^*)|, \quad n \in I\!N, \; i = 1, \ldots, N,$$

see also the discussion in [19], Section 5.

Let $\mathcal{B}$ denote the family of polyhedra in $I\!R^s$ whose facets are parallel to facets arising among $\mathcal{K}_1^*, \ldots, \mathcal{K}_N^*$. Up to location in parallel hyperplanes, then only finitely many different facets occur in $\mathcal{B}$. Therefore, $\mathcal{B}$ is a Vapnik-Červonenkis class of Borel sets in $I\!R^s$ (e.g. [13], [23]). To be a Vapnik-Červonenkis class means, for some family $\mathcal{B}_o$ of Borel sets in $I\!R^s$, that, for some cardinality $k \in I\!N$ and for any finite set $E \subset I\!R^s$ with exactly k elements, there exists a subset of E which is not representable as $E \cap B$ with $B \in \mathcal{B}_o$. For details on Vapnik-Červonenkis classes we refer to the literature ([13], [23]). As prerequisite for the subsequent statement we need the following law of iterated logarithm (log log law) for $\mathcal{B}$.

Lemma 4.4 *For $P-$almost all $\omega \in \Omega$ it holds*

$$\limsup_{n\to\infty} \Big(\frac{n}{2\log\log n}\Big)^{1/2} \cdot \sup\{|\mu(B) - \mu_n(\omega)(B)| : B \in \mathcal{B}\} \leq \frac{1}{2}. \tag{4.5}$$

Proof: Since $\omega \mapsto \sup\{|\mu(B) - \mu_n(\omega)(B)| : B \in \mathcal{B}\}$ is measurable (Theorem 2K in [15]) and $\mathcal{B}$ is a Vapnik-Červonenkis class, the assertion turns out a special case of the log log law established in [10]. ∎

Proposition 4.5 *Adopt the setting of Proposition 4.3 and let μ be estimated via empirical measures $\mu_n(\omega)$ ($\omega \in \Omega, n \in I\!N$). Then it holds*

$$\limsup_{n\to\infty} \Big(\frac{2n}{\log\log n}\Big)^{\frac{1}{2}} \cdot d_H(\psi(\mu), \psi(\mu_n(\omega))) \leq L \cdot N \quad P - \textit{almost surely.}$$

Proof: By (4.5), the estimate in (4.4) is valid with $\nu = \mu_n(\omega)$ provided that $n = n(\omega) \in I\!N$ is sufficiently large. Therefore,

$$\begin{aligned}
&\limsup_{n\to\infty} \Big(\frac{2n}{\log\log n}\Big)^{1/2} \cdot \sup\{|\mu_n(\omega)(B) - \mu(B)| : B \in \mathcal{B}\} \\
&\leq \; 2 \cdot \limsup_{n\to\infty} \Big(\frac{n}{2\log\log n}\Big)^{1/2} \cdot L \cdot \sum_{i=1}^{N} \sup\{|\mu_n(\omega)(B) - \mu(B)| : B \in \mathcal{B}\} \\
&\leq \; 2 \cdot L \cdot N \cdot \frac{1}{2} = L \cdot N \qquad \text{for } P - \text{almost all } \omega \in \Omega.
\end{aligned}$$

∎

In [17] Proposition 3.1, the same law of iterated logarithm was obtained for recourse models with fully-random right-hand side. There, a more specific log log law for the uniform distance of distribution functions was applicable. In the end, this rests on the simpler lineality complex of Φ which leads to an estimate like (4.4) with lower left orthants instead of polyhedra $\mathcal{K}_i^*$, cf. [17], Corollary 2.5.

Further possible conclusions from Proposition 4.3 concern large-deviation estimates as, for instance, estimates for

$$P(\{\omega : \; d_H(\psi(\mu), \psi(\mu_n(\omega))) \geq \varepsilon\}) \;\; \text{with } \varepsilon > 0.$$

Here, estimates for $P(\{\omega : \sup\{|\mu(B) - \mu_n(\omega)(B)| : B \in \mathcal{B}\} \geq \varepsilon\})$ that are available from the literature ([23], p. 829) can be employed. This leads to an estimate in the spirit of [17], Proposition 3.2, for an extended class of problems. Further related work on large deviations in stochastic programming is contained in [8], [9], [24]. In [8], the authors impose a general conditioning assumption for the solution set of the unperturbed problem (cf. [1]) of which the quadratic conditioning implied by strong convexity is a special case.

5 Final Remarks

We have seen that the curvature behaviour of the expected recourse function in a two-stage stochastic program with linear recourse is essentially determined by an interplay of the probability density and the polyhedral complex of lineality regions of the second-stage value function. The latter becomes more involved when passing from models with completely random right-hand side to those where non-random entries occur among the components of the second-stage right-hand side vector. Nevertheless, it is possible to verify strong convexity of the expected recourse as a function of the tender variable for these models too. Beside standard assumptions, an interior point condition now has to hold for a suitable *projection* of the second-stage dual feasible set and uniform positivity of the density must be present at more involved regions determined by the lineality complex of the second-stage value function. Existing vertex enumeration codes can be used to support verification of the mentioned properties.
Strong convexity of the expected recourse as function of the tender variable is essential for quantitative statements on the stability and the asymptotic behaviour of optimal solutions to linear recourse models. Therefore, our main result widens the class of problems for which these statements can be shown to hold.

Acknowledgement:

I wish to thank Wolfgang Maibauer (Konrad-Zuse-Zentrum für Informationstechnik Berlin) for assistance in preparing the figures in Section 2.

References

[1] Attouch, H.; Wets, R.J-B: Quantitative stability of variational systems: II - a framework for nonlinear conditioning, SIAM Journal on Optimization 3 (1993), 359-381.

[2] Bertsekas, D.: Stochastic optimization problems with nondifferentiable cost functionals, Journal of Optimization Theory and Applications 12 (1973), 218-231.

[3] Christof, T.: PORTA - a polyhedron representation transformation algorithm, available via http://www.iwr.uni-heidelberg.de/iwr/comopt/soft/PORTA/readme.html or from the ZIB electronic library ELIB via elib@zib-berlin.de .

[4] Dentcheva, D.; Römisch, W.; Schultz, R.: Strong convexity and directional derivatives of marginal values in two-stage stochastic programming, in: K. Marti, P. Kall

(Eds.) "Stochastic Programming - Numerical Techniques and Engineering Applications", Springer-Verlag, Lecture Notes in Economics and Mathematical Systems 423 (1995), 8-21.

[5] Dudley, R.M.: Real Analysis and Probability, Wadsworth & Brooks/Cole, Pacific Grove, California 1989.

[6] Kall, P.: Stochastic Linear Programming, Springer Verlag, Berlin 1976.

[7] Kall, P.; Wallace, S.W.: Stochastic Programming, John Wiley, Chichester 1994.

[8] Kaniovski, Y.M.; King, A.J.; Wets, R.J-B: Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems, Annals of Operations Research 56 (1995), 189-208.

[9] Kaňková, V.: An approximative solution of a stochastic optimization problem, Transactions of the Eighth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Academia, Prague 1978, 349-353.

[10] Kuelbs, J.; Dudley, R.M.: Log log laws for empirical measures, Annals of Probability 8 (1980), 405-418.

[11] Nožička, F.; Guddat, J.; Hollatz, H.; Bank, B.: Theorie der linearen parametrischen Optimierung, Akademie Verlag, Berlin 1974.

[12] Ortega, J.M.; Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables, Academic Press, New York 1970.

[13] Pollard, D.: Convergence of Stochastic Processes, Springer-Verlag, New York 1984.

[14] Rockafellar, R.T.: Convex Analysis, Princeton University Press, Princeton 1970.

[15] Rockafellar, R.T.: Integral functionals, normal integrands and measurable selections, Nonlinear Operators and the Calculus of Variations (G.P. Gossez et al. Eds.), Lecture Notes in Mathematics, Vol 543, Springer-Verlag, New York 1976, 157-207.

[16] Römisch, W.; Schultz, R.: Stability analysis for stochastic programs, Annals of Operations Research 30 (1991), 241-266.

[17] Römisch, W.; Schultz, R.: Lipschitz stability for stochastic programs with complete recourse, SIAM Journal on Optimization (to appear).

[18] Schultz, R.: Strong Convexity in Stochastic Programs with Complete Recourse, Journal of Computational and Applied Mathematics 56 (1994), 3-22.

[19] Schultz, R.: Rates of convergence in stochastic programs with complete integer recourse, SIAM Journal on Optimization (to appear).

[20] Shapiro, A.: Asymptotic behavior of optimal solutions in stochastic programming, Mathematics of Operations Research 18 (1993), 829-845.

[21] Shapiro, A.: Quantitative stability in stochastic programming, Mathematical Programming 67 (1994), 99-108.

[22] Shapiro, A.: Simulation based optimization - convergence analysis and statistical inference, manuscript, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta 1995.

[23] Shorack, G.R.; Wellner, J.A.: Empirical Processes with Applications to Statistics, Wiley, New York 1986.

[24] Vogel, S.: Stability results for stochastic programming problems, Optimization 19 (1988), 269-288.

[25] Walkup, D.W.; Wets, R.J-B: Lifting projections of convex polyhedra, Pacific Journal of Mathematics 28 (1969), 465-475.

[26] Wets, R.J-B: Stochastic programs with fixed recourse: the equivalent deterministic program, SIAM Review 16 (1974), 309-339.