Stable Bounds on the Duality Gap of Finite Sum Minimization Problems

Anonymous Author(s)
Affiliation
Address
email

Abstract

The Shapley-Folkman theorem shows that Minkowski averages of uniformly bounded sets tend to be convex when the number of terms in the sum becomes much larger than the ambient dimension. In optimization, Aubin and Ekeland [1976] show that this produces an a priori bound on the duality gap of separable nonconvex optimization problems involving finite sums. This bound is highly conservative and depends on unstable quantities, and we relax it in several directions to show that non convexity can have a much milder impact on finite sum minimization problems such as empirical risk minimization and multi-task classification. As a byproduct, we show a new version of Maurey's classical approximate Carathéodory lemma where we sample a significant fraction of the coefficients, without replacement.

1 Introduction

We focus on separable optimization problems written

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} f_i(x_i) \\ \text{subject to} & Ax \leq b, \\ & x_i \in Y_i, \quad i = 1, \ldots, n, \end{array} \tag{P}$$

in the variables $x_i \in \mathbb{R}^{d_i}$ with $d = \sum_{i=1}^{n} d_i$, where the functions f_i are lower semicontinuous (but not necessarily convex), the sets $Y_i \subset \mathbf{dom}\, f_i$ are compact, and $A \in \mathbb{R}^{m \times d}$ and $b \in \mathbb{R}^m$. Aubin and Ekeland [1976] showed that the duality gap of problem (P) vanishes when the number of terms n grows towards infinity while the dimension m remains bounded, provided the nonconvexity of the functions f_i is uniformly bounded. The result in [Aubin and Ekeland, 1976] hinges on the fact that the epigraph of problem (P) can be written as a Minkowski sum of n sets in dimension $m+1$. In this setting, the Shapley-Folkman theorem shows that if $V_i \subset \mathbb{R}^m$, $i = 1, \ldots, n$ are arbitrary subsets of $\mathbb{R}^m$ and

$$x \in \mathbf{Co}\left(\sum_{i=1}^{n} V_i\right) \qquad \text{then} \qquad x \in \sum_{[1,n]\setminus\mathcal{S}} V_i \;+\; \sum_{\mathcal{S}} \mathbf{Co}(V_i)$$

for some $|\mathcal{S}| \leq m+1$. If the sets V_i are uniformly bounded, n grows and m remains bounded, the term $\sum_{\mathcal{S}} \mathbf{Co}(V_i)$ becomes negligible and the Minkowski sum $\sum_i V_i$ is increasingly close to its convex hull. In fact, several measures of nonconvexity decrease monotonically towards zero when n grows in this setting, with [Fradelizi et al., 2017] showing for instance that the Hausdorff distance

$$d_H\left(\sum_i V_i, \mathbf{Co}\left(\sum_i V_i\right)\right) \to 0.$$

We illustrate this phenomenon graphically in Figure 1, where we show the Minkowski mean of n unit $\ell_{1/2}$ balls for $n = 1, 2, 10, \infty$ in dimension 2, and the average of five arbitrary point sets (defined

Submitted to 32nd Conference on Neural Information Processing Systems (NIPS 2018). Do not distribute.

from digits here). In both cases, Minkowski averages are nearly convex for relatively small values of n.

The Shapley-Folkman theorem was derived by Shapley & Folkman in private communications and first published by [Starr, 1969]. It was used by Aubin and Ekeland [1976] to derive *a priori* bounds on the duality gap. The continuous limit of this result is known as the Liapunov convexity theorem and shows that the range of non-atomic, vector valued measures is convex [Aumann and Perles, 1965, Berliocchi and Lasry, 1973]. The results of Aubin and Ekeland [1976] were extended in [Ekeland and Temam, 1999] to generic separable constrained problems, and also by [Lauer et al., 1982, Bertsekas, 2014] to more precise yet less explicit nonconvexity measures, who describe applications to large-scale unit commitment problems. Extreme points of the set of solutions of a convex relaxation to problem (P) are used to produce good approximations and Udell and Boyd [2016] describe a randomized purification procedure to find such points with probability one.

The Shapley-Folkman theorem is a direct consequence of the conic version of Carathéodory's theorem, with the number of terms in the conic representation of optimal points controlling the duality gap bound. Our first contribution seeks to reduce this number by allowing a small approximation error in the conic representation. This essentially trades off approximation error with duality gap. In general, these approximations are handled by Maurey's classical approximate Carathéodory lemma [Pisier, 1981]. Here however we need to sample a very high fraction of the coefficients, hence we produce a *high sampling ratio* version of the approximate Carathéodory lemma using results by [Serfling, 1974, Bardenet et al., 2015, Schneider, 2016] on sampling sums without replacement.

We then use this result to produce an approximate version of the duality gap bound in [Aubin and Ekeland, 1976] which allows a direct tradeoff between the impact of nonconvexity and the approximation error. This approximate formulation also has the benefit of writing the gap bound in terms of stable quantities, thus better revealing the link between problem structure and duality gap.

Nonconvex separable problems involving finite sums such as (P) occur naturally in machine learning, signal processing and statistics. The most direct examples being perhaps empirical risk minimization, sparse recovery and multi-task learning. In this later setting, our bounds show that when the number of tasks grows and the tasks are only loosely coupled (e.g. the separable ℓ_2 constraint [Ciliberto et al., 2017]), nonconvex multi-task problems have asymptotically vanishing duality gap. A stream of recent results have shown that finite sum optimization problems have particularly good computational complexity (see [Roux et al., 2012, Johnson and Zhang, 2013, Defazio et al., 2014] and more recently [Allen-Zhu and Yuan, 2016, Reddi et al., 2016] in the nonconvex case), our results show that *they also have intrinsically low duality gap in some settings.*

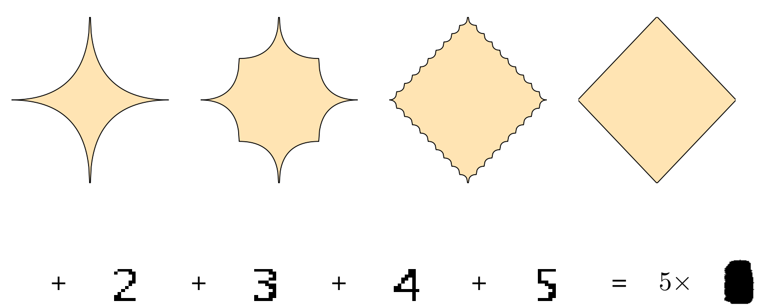

Figure 1: *Top:* The $\ell_{1/2}$ ball, Minkowsi average of two and ten balls, and convex hull. *Bottom:* Minkowsi average of five first digits (obtained by sampling).

2 Convex Relaxation and Bounds on the Duality Gap

We first recall and adapt some key results from [Aubin and Ekeland, 1976, Ekeland and Temam, 1999] producing *a priori* bounds on the duality gap, using an epigraph formulation of problem (P).

2.1 Convex Envelope and Convex Relaxations

Assuming that f is not identically $+\infty$ and is minorized by an affine function, we write $f^*(y) \triangleq \inf_{x \in \mathbf{dom}\, f}\{y^\top x - f(x)\}$ the conjugate of f, and $f^{**}(y)$ its biconjugate. The biconjugate of f (aka the convex envelope of f) is the pointwise supremum of all affine functions majorized by f (see *e.g.*, [Rockafellar, 1970, Th. 12.1] or [Hiriart-Urruty and Lemaréchal, 1993, Th. X.1.3.5]), a corollary then shows that $\mathbf{epi}(f^{**}) = \overline{\mathbf{Co}(\mathbf{epi}(f))}$. For simplicity, we write $S^{**} = \overline{\mathbf{Co}(S)}$ for any set S in what follows. We will make the following technical assumptions on the functions f_i.

Assumption 2.1 *The functions $f_i : \mathbb{R}^{d_i} \to \mathbb{R}$ are proper, 1-coercive, lower semicontinuous and there exists an affine function minorizing them.*

Note that coercivity trivially holds if $\mathbf{dom}(f_i)$ is compact (since f is $+\infty$ outside). When Assumption 2.1 holds, $\mathbf{epi}(f^{**})$, f_i^{**} and hence $\sum_{i=1}^n f_i^{**}(x_i)$ are closed [Hiriart-Urruty and Lemaréchal, 1993, Lem. X.1.5.3]. Finally, as in *e.g.*, [Ekeland and Temam, 1999], we define the lack of convexity of a function as follows.

Definition 2.2 *Let $f : \mathbb{R}^d \to \mathbb{R}$ we let $\rho(f) \triangleq \sup_{x \in \mathbf{dom}(f)}\{f(x) - f^{**}(x)\}$.*

Many other quantities measure lack of convexity, see *e.g.*, [Aubin and Ekeland, 1976, Bertsekas, 2014] for further examples. In particular, the nonconvexity measure $\rho(f)$ can be further refined, using the fact that

$$\rho(f) = \sup_{\substack{x_i \in \mathbf{dom}(f) \\ \alpha \in \mathbb{R}^{d+1}}} \left\{ f\left(\sum_{i=1}^{d+1} \alpha_i x_i\right) - \sum_{i=1}^{d+1} \alpha_i f(x_i) : \mathbf{1}^T \alpha = 1, \alpha \geq 0 \right\}$$

when f satisfies Assumption 2.1 (see [Hiriart-Urruty and Lemaréchal, 1993, Th. X.1.5.4]). In this setting, Bi and Tang [2016] define the k^{th}-nonconvexity measure as

$$\rho_k(f) \triangleq \sup_{\substack{x_i \in \mathbf{dom}(f) \\ \alpha \in \mathbb{R}^{d+1}}} \left\{ f\left(\sum_{i=1}^{d+1} \alpha_i x_i\right) - \sum_{i=1}^{d+1} \alpha_i f(x_i) : \mathbf{1}^T \alpha = 1, \mathbf{Card}(\alpha) \leq k, \alpha \geq 0 \right\} \tag{1}$$

which restricts the number of nonzero coefficients in the formulation of $\rho(f)$. Note that $\rho_1(f) = 0$.

In the supplementary material, we show that the dual of problem (P) maximizes a linear form over the convex hull of a Minkowski sum of n epigraphs. We also show that this dual matches the dual of a convex relaxation of (P), formed using the convex envelopes of the functions $f_i(x)$. In what follows, we will assume without loss of generality that $Y_i = \mathbb{R}^{d_i}$, replacing f_i by $f_i(x) + \mathbf{1}_{Y_i}(x)$. We use the biconjugate to produce a convex relaxation of problem (P) written

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n f_i^{**}(x_i) \\ \text{subject to} & Ax \leq b \end{array} \tag{CoP}$$

in the variables $x_i \in \mathbb{R}^{d_i}$.

2.2 Bounds on the Duality Gap

We now recall results by [Aubin and Ekeland, 1976, Ekeland and Temam, 1999] bounding the duality gap in (P) using the lack of convexity of the functions f_i. In the formulation below, the dual is more explicit than in [Ekeland and Temam, 1999] because the constraints are affine here.

Proposition 2.3 *Suppose the functions f_i in* (P) *satisfy Assumption 2.1. There is a point $x^\star \in \mathbb{R}^d$ at which the primal optimal value of* (CoP) *is attained, such that*

$$\underbrace{\sum_{i=1}^n f_i^{**}(x_i^\star)}_{CoP} \;\leq\; \underbrace{\sum_{i=1}^n f_i(\hat{x}_i^\star)}_{P} \;\leq\; \underbrace{\sum_{i=1}^n f_i^{**}(x_i^\star)}_{CoP} \;+\; \underbrace{\sum_{i \in \mathcal{S}} \rho(f_i)}_{\text{gap}} \tag{2}$$

with $\hat{x}^\star$ is an optimal point of (P)*, and*

$$\mathcal{S} \triangleq \{i : \; (f_i^{**}(x_i^\star), A_i x_i^\star) \notin \mathbf{Ext}(\mathcal{F}_i)\}$$

where $\mathcal{F}_i \subset \mathbb{R}^{m+1}$ is defined as

$$\mathcal{F}_i = \left\{ (f_i^{**}(x_i), A_i x_i) : \ x_i \in \mathbb{R}^{d_i} \right\}$$

writing $A_i \in \mathbb{R}^{m \times d_i}$ the i^{th} block of A.

This last result bounds *a priori* the duality gap in problem (P) by $\sum_{i \in \mathcal{S}} \rho(f_i)$, where $\mathcal{S} \subset [1, n]$. The dual problem in (D) shows that the optimal solution maximizes an affine form over the closed convex hull of the epigraph of the primal (P) and is thus attained at an extreme point of that epigraph. Separability means this epigraph is the Minkowski sum of the closed convex hulls of the epigraphs of the n subproblems, while $|\mathcal{S}|$ counts the number of terms in this sum for which the optimum is attained at an extreme point of these subproblems. The Shapley-Folkman theorem together with the results of the next sections will produce upper bounds on the size of $\mathcal{S}$ and show that it is typically much smaller than n.

3 The Shapley-Folkman Theorem

Carathéodory's theorem is the key ingredient in proving the Shapley-Folkman theorem and is recalled in the supplementary material. The Shapley-Folkman theorem below was derived by Shapley & Folkman in private communications and first published by [Starr, 1969].

Theorem 3.1 (Shapley-Folkman) *Let $V_i \in \mathbb{R}^d$, $i = 1, \ldots, n$ be a family of subsets of $\mathbb{R}^d$. If*

$$x \in \mathbf{Co}\left(\sum_{i=1}^n V_i\right) = \sum_{i=1}^n \mathbf{Co}\left(V_i\right) \quad \textit{then} \quad x \in \sum_{[1,n]\setminus\mathcal{S}} V_i \ + \ \sum_{\mathcal{S}} \mathbf{Co}(V_i)$$

where $|\mathcal{S}| \leq d$.

This theorem has been used, for example, to prove existence of equilibria in markets with a large number of agents with non-convex preferences. Classical proofs usually rely on a dimension argument [Starr, 1969], but the one we recall in the supplementary materials is more constructive. It was also used to produce a priori bounds on the duality gap in [Aubin and Ekeland, 1976], see also [Ekeland and Temam, 1999, Bertsekas, 2014, Udell and Boyd, 2016] for a more recent discussion. The following result is similar in spirit to those [Aubin and Ekeland, 1976].

Proposition 3.2 *Suppose the functions f_i in* (P) *satisfy Assumption 2.1. There is a point $x^\star \in \mathbb{R}^d$ at which the primal optimal value of* (CoP) *is attained, such that*

$$\underbrace{\sum_{i=1}^n f_i^{**}(x_i^\star)}_{CoP} \ \leq \ \underbrace{\sum_{i=1}^n f_i(\hat{x}_i^\star)}_{P} \ \leq \ \underbrace{\sum_{i=1}^n f_i^{**}(x_i^\star)}_{CoP} \ + \ \underbrace{\sum_{i=1}^{m+1} \rho(f_{[i]})}_{\text{gap}} \tag{3}$$

where $\hat{x}^\star$ is an optimal point of (P) *and $\rho(f_{[1]}) \geq \rho(f_{[2]}) \geq \ldots \geq \rho(f_{[n]})$.*

The result above directly links the gap bound with the *number of nonzero coefficients in the conic combination* defining the solution $z^\star$ (see (9) in the supplementary material). The smaller this number, the tighter the gap bound. In fact, if we use the k^{th}-nonconvexity measure $\rho_k(f)$ in (1) instead of $\rho(f)$, the duality gap bound can be refined to

$$\text{gap} \leq \max_{\beta_i \in [1, m+2]} \left\{ \sum_{i=1}^n \rho_{\beta_i}(f_i) : \sum_{i=1}^n \beta_i = n + m + 1 \right\}.$$

Since $\rho_1(f) = 0$, this last bound can be significantly smaller, since the result in [Aubin and Ekeland, 1976] implicitly assumes that $\sum_{i=1}^n \beta_i = n + (m+2)(m+1)$, instead of $n + m + 2$ here.

More importantly, remark also that this bound is written in terms of *unstable quantities*, namely the number of linear constraints in $Ax \leq b$ and the number of nonzero coefficients in the exact conic representation of $z^\star \in \mathcal{G}_r^{**}$. In the sections that follow, we will seek to further tighten this bound by both simplifying the coupling constraints to reduce m using approximate extended formulations, and reducing the number of nonzero coefficients in the conic representation using approximate versions of Carathéodory's theorem.

4 Stable Bounds on the Duality Gap

The result of Aubin and Ekeland [1976] recalled above uses the Shapley-Folkman theorem to refine the conclusion of Proposition 2.3, and bounds the duality gap in problem (P) by

$$\text{gap} \leq \sum_{i=1}^{m+1} \rho(f_{[i]})$$

where m is the number of constraints $Ax \leq b$. As remarked by [Udell and Boyd, 2016], we can actually take m to be the number of *active* constraints at the optimum of problem (P), which can be substantially smaller than m but is hard to bound a priori. We can write a more stable version of the result of Aubin and Ekeland [1976] using approximate representations of the optimal solution in the Minkowski sum of epigraphs. We get the following result.

Theorem 4.1 *Suppose the functions f_i in* (P) *satisfy Assumption 2.1. There is a point $x^\star \in \mathbb{R}^d$ at which the primal optimal value of* (CoP) *is attained, and as in* (9) *we let*

$$z^\star = \sum_{i=1}^{n} \begin{pmatrix} f_i^{**}(x_i^\star) \\ A_i x_i^\star \end{pmatrix} + \begin{pmatrix} 0 \\ w-b \end{pmatrix}$$

with $w \in \mathbb{R}^m_+$ be the corresponding minimizer in (8)*. Suppose that we use an approximate conic representation of $z^\star$ using only $s \in [n, n+m+1]$ coefficients, writing*

$$\lambda(s) = \underset{\substack{\lambda_{ij}\geq 0 \\ z_{ij}\in\mathcal{F}_i}}{\operatorname{argmin}} \left\{ \left\| z^\star - \sum_{i=1}^{n}\sum_{j=1}^{m+2} \lambda_{ij} z_{ij} \right\| : \; \sum_{i=1}^{n} \mathbf{Card}(\lambda_i) \leq s, \; \mathbf{1}^T\lambda_i = 1, i = 1,\ldots,n \right\}$$

where $z_{ij} \in \mathcal{F}_i$ for $i = 1,\ldots,n$, $j = 1,\ldots,m+2$, and $u(s) = z^\star - \sum_{i=1}^{n}\sum_{j=1}^{m+2}\lambda_{ij}(s) z_{ij}$. We have the following bound on the solution of problem (pP)

$$\underbrace{\mathrm{h}_{CoP}(u_2(s))}_{\text{(pCoP)}} \;\leq\; \underbrace{\mathrm{h}_{P}(u_2(s))}_{\text{(pP)}} \;\leq\; \underbrace{\mathrm{h}_{CoP}(0)}_{\text{(CoP)}} \;+\; \underbrace{|u_1(s)| + \max_{\beta_i\in[1,m+2]} \left\{ \sum_{i=1}^{n} \rho_{\beta_i}(f_i) : \sum_{i=1}^{n}\beta_i = s \right\}}_{\text{gap}(s)}. \tag{4}$$

Furthermore, we can take m to be the number of active inequality constraints at $x^\star$.

The structure of this last bound differs from the previous ones because the perturbation u is acting on the epigraph formulation of (pP), so it induces an error on both the objective values (the first coefficient $u_1(s)$ in this epigraph representation) and on the contraints (the last m coefficients $u_2(s)$). This means that we now bound the gap on a perturbed version of problem (pP), with constraint perturbation size controlled by u_2. The tightness of the duality gap bound in (4) depends on two distinct quantities. The first, namely u, is a function of how much we can "compress" the convex approximation of $z^\star$ in (9). The second, controlled by the sum of the nonconvexity measures $\rho_{\beta_i}(f_i)$, measures the severity of the problem's lack of convexity. The sparsity parameter s controls the tradeoff between these two components to minimize the bound, and is bounded by n plus the number of active constraints. The results that follow will seek to make this tradeoff and all the quantities involved more explicit.

5 Coupling Constraints

The duality gap bounds in (3) or (4) heavily depend on the structure of the coupling constraints $Ax \leq b$ and exploiting this structure can lead to significant precision gain as detailed in what follows.

5.1 Active constraints & Helly theorems

As noticed by [Udell and Boyd, 2016], it suffices to consider only active constraints at the optimum when computing the duality gap bound in (3) or (4). This number can be significantly smaller than m. In particular, [Calafiore and Campi, 2005, Th. 2] or [Shapiro et al., 2009, Lem. 5.31] for example show $m \leq d$ using Helly's theorem. Bounds on the number of active constraints play a key role in solving chance constrained problems for example [Calafiore and Campi, 2005, Tempo et al., 2012, Zhang et al., 2015]. Let us write $A_I x \leq b_I$ the equations corresponding to active constraints at the optimum, where $b_I \in \mathbb{R}^{\tilde{m}}$. We will see in the next section that we can further reduce the number of inequalities defining active constraints by changing their representation.

5.2 Extended formulations

The duality gap bounds in (3) are written in terms of the number of linear constraints $Ax \leq b$ in problem (P). These constraints form a polytope $\mathcal{P}$ and the gap bound heavily depends on the *representation* of this polytope. Producing a more compact formulation of $\mathcal{P}$, i.e. one using less linear inequalities, would then make our duality gap bounds much more precise. One way to produce such compact representations is to use extended formulations. An *extended formulation* of the constraint polytope $\mathcal{P} = \{x \in \mathbb{R}^d : Ax \leq b\}$ writes it as the projection of another, potentially simpler, polytope with

$$\mathcal{P} = \{x \in \mathbb{R}^d : Bx + Cu \leq d, u \in \mathbb{R}^m\}$$

where $B \in \mathbb{R}^{q \times d}$, $C \in \mathbb{R}^{q \times m}$ and $d \in \mathbb{R}^q$. The *extension complexity* $xc(\mathcal{P})$ is the minimum number of inequalities of an extended formulation of the polytope $\mathcal{P}$. A fundamental result by [Yannakakis, 1991, Th. 3] connects extended formulations and nonnegative matrix factorization. Suppose the vertices of a polytope $\mathcal{P} = \{x \in \mathbb{R}^d : Ax \leq b\}$ are given by $\{v_1, \ldots, v_p\}$, we write S the *slack matrix* of $\mathcal{P}$, with

$$S_{ij} = b_i - (Av_j)_i, \quad \text{for } i = 1, \ldots, m, j = 1, \ldots, p.$$

By construction, S is a nonnegative matrix. [Yannakakis, 1991, Th. 3] shows that

$$\{x \in \mathbb{R}^d : Ax + Fy = b, y \geq 0\}$$

is an extended formulation of $\mathcal{P}$ if and only if S can be factored as $S = FV$ where $F \in \mathbb{R}_+^{m \times q}$ and $V \in \mathbb{R}_+^{q \times p}$ are both nonnegative. In particular, the smallest extended formulation of $\mathcal{P}$ corresponds to the lowest rank NMF of S, which means $xc(\mathcal{P}) = \textbf{Rank}_+(S)$, the nonnegative rank of S.

While the nonnegative rank is again an unstable quantity, stable (approximate) versions of this result can be defined using *nested polytopes* [Pashkovich, 2012, Braun et al., 2012, Gillis and Glineur, 2012]. Given polytopes $\mathcal{P}, \mathcal{Q} \in \mathbb{R}^d$, an extended formulation of the pair $(\mathcal{P}, \mathcal{Q})$ is a polytope

$$\mathcal{K} = \max\{x \in \mathbb{R}^d : Ax + Fy = b, y \geq 0\}$$

such that $\mathcal{P} \subset \mathcal{K} \subset \mathcal{Q}$. Furthermore, suppose $\mathcal{P} = \textbf{Co}(\{v_1, \ldots, v_p\})$ and $\mathcal{Q} = \{x \in \mathbb{R}^d : Ax \leq b\}$, defining the slack matrix of the pair $(\mathcal{P}, \mathcal{Q})$ as $S_{ij} = b_i - (Av_j)_i$, for $i = 1, \ldots, m$, $j = 1, \ldots, p$, the result in [Braun et al., 2012, Th. 1] shows that the extension complexity of the pair satisfies $xc(\mathcal{P}, \mathcal{Q}) \leq \textbf{Rank}_+(S) + 1$. Overall, this means that we can replace m in Proposition 3.2 and Theorem 4.1 by the extension complexity of the polytope formed by the active constraints, which can be substantially smaller.

6 An Approximate Shapley-Folkman Theorem

We will now derive a version of the Shapley-Folkman result in Theorem 3.1 which only approximates x but where $\mathcal{S}$ is typically smaller.

6.1 Approximate Carathéodory Theorems

Recent activity around Carathéodory's theorem [Donahue et al., 1997, Vershynin, 2012, Dai et al., 2014] has focused on producing tight approximate versions of this result, where one aims at finding a convex combination using fewer elements, which is still a "good" approximation of the original element of the convex hull. The following theorem states an upper bound on the number of elements needed to achieve a given level of precision, using a randomization argument.

Theorem 6.1 (Approximate Carathéodory) *Let $V \subset \mathbb{R}^d$, $x \in \textbf{Co}(V)$ and $\varepsilon > 0$. We assume that V is bounded and we write D_p the quantity $D_p \triangleq \sup_{v \in V} \|v\|_p$. Then, there exists some $\hat{x} \in \textbf{Co}(V)$ and $m \leq 8pD_p^2/\varepsilon^2$ such that*

$$\|x - \hat{x}\| = \left\|x - \textstyle\sum_{i=1}^N \lambda_i v_i\right\|_p \leq \varepsilon,$$

for some $v_i \in V$, $\lambda_i > 0$ and $\mathbf{1}^\top \lambda = 1$.

This result is a direct consequence of Maurey's lemma [Pisier, 1981] and is based on a probabilistic approach which samples vectors v_i with replacement and uses concentration inequalities to control approximation error, but can also be seen as a direct application of Frank-Wolfe type algorithms to the projection problem minimize $\left\|x - \sum_{i=1}^N \lambda_i v_i\right\|^2$ in the variable $\lambda \in \mathbb{R}^n$. In the results that follow however, we will have $N = n + m + 1$, and we will seek approximations using s terms with $s \in [n, n+m+1]$ with n typically much bigger than m. Sampling with replacement does not provide precise enough bounds in this setting and we will use results from [Serfling, 1974] on sample sums *without replacement* to produce a more precise version of the approximate Carathéodory theorem that handles the case where a high fraction of the coefficients is sampled.

Theorem 6.2 *Let $x = \sum_{j=1}^N \lambda_j V_j$ for $V \in \mathbb{R}^{d \times N}$ and some $\lambda \in \mathbb{R}^N$ such that $\mathbf{1}^T \lambda = 1, \lambda \geq 0$. Let $\varepsilon > 0$ and write $R = \max\{R_v, R_\lambda\}$ where $R_v = \max_i \|\lambda_i V_i\|_\infty$ and $R_\lambda = \max_i |\lambda_i|$. Then, there exists some $\hat{x} = \sum_{j \in \mathcal{J}} \mu_j V_j$ with $\mu \in \mathbb{R}^m$ and $\mu \geq 0$, where $\mathcal{J} \subset [1, N]$ has size*

$$|\mathcal{J}| = 1 + N \, \frac{\log(2d)(\sqrt{N}\, R/\varepsilon)^2}{2 + \log(2d)(\sqrt{N}\, R/\varepsilon)^2}$$

and is such that $\|x - \hat{x}\|_\infty \leq \varepsilon$ and $|\sum_{j \in \mathcal{J}} \mu_j - 1| \leq \varepsilon$.

The result above uses Hoeffding-Serfling bounds to provide error bounds in ℓ_∞ norm. Recent results by [Bardenet et al., 2015] provide Bernstein-Serfling type inequalities where the radius R above can be replaced by a standard deviation. Since the vectors we consider here have a block structure coming from the epigraphs $\mathcal{F}_i$, we consider generic Banach spaces to properly fit the norm to this structure by extending this last result to arbitrary norms in $(2, D)$-smooth Banach spaces using a recent result by [Schneider, 2016] and show a more general version as Theorem 8.9 in the supplementary material. We also show a Bennett-Serfling like inequality in§8.7 which allow us to control the sampling ratio using a variance term. This means we the sampling ratio in Theorem 6.2 above can be replaced by

$$\alpha_m \geq \frac{2\ln(2/\delta_0)\big[2(D\sigma)^2 + \epsilon_0 R_v/(3N)\big]N}{\epsilon^2 + 2\ln(2/\delta_0)\big[2(D\sigma)^2\big]N},$$

where

$$\sigma \triangleq \frac{1}{\sum_{k=1}^m \frac{1}{(N-k)^2}} \Big|\Big| \Big(\sum_{k=1}^m \frac{1}{(N-k)^2} \mathbb{E}_{k-1} ||V_k - \mathbb{E}_{k-1}(V_k)||^2 \Big)^{1/2} \Big|\Big|_\infty ,$$

plays the role of the variance when sampling without replacement.

6.2 Approximate Shapley-Folkman Theorems

We now prove an approximate version of the Shapley-Folkman theorem, plugging approximate Carathéodory results inside the proof of Theorem 3.1.

Corollary 6.3 *Let $\varepsilon > 0$ and $V_i \in \mathbb{R}^d$, $i = 1, \ldots, n$ be a family of subsets of $\mathbb{R}^d$. Suppose*

$$x = \sum_{i=1}^n \sum_{j=1}^{d+1} \lambda_{ij} v_{ij} \;\in\; \sum_{i=1}^n \mathbf{Co}\,(V_i)$$

where $\lambda_{ij} \geq 0$ and $\sum_j \lambda_{ij} = 1$. We write $R_v = \max_{\{ij : \lambda_{ij} \neq 1\}} \|\lambda_{ij} v_{ij}\|$ and $R_\lambda = \max_{\{ij : \lambda_{ij} \neq 1\}} |\lambda_{ij}|$, for some norm $\|\cdot\|$ such that $(\mathbb{R}^d, \|\cdot\|)$ is $(2, D)$-smooth. There exists a point $\bar{x}$ and an index set $\mathcal{S} \subset [1, n]$ such that

$$\bar{x} \in \sum_{[1,n] \setminus \mathcal{S}} V_i \;+\; \sum_{i \in \mathcal{S}} \mathbf{Co}(V_i) \quad \textit{with} \quad \|x - \bar{x}\| \leq \sqrt{2d} \left(\frac{R_v}{R_\lambda} + M_V \right) \varepsilon$$

where $|\mathcal{S}| \leq m - d$ with

$$m = 1 + 2d \, \frac{c\,(DR_\lambda/\varepsilon)^2}{1 + c\,(DR_\lambda/\varepsilon)^2} \qquad \textit{and} \qquad M_V = \sup_{\substack{\|u\|_2 \leq 1 \\ v_i \in V_i}} \left\| \sum_i u_i v_i \right\|. \tag{5}$$

where $c > 0$ is an absolute constant.

We prove a slightly more general version of this result in Theorem 8.10 in the supplementary material. The result of Aubin and Ekeland [1976] recalled in Proposition 3.2 shows that the Shapley-Folkman theorem can be used in the bounds of Proposition 2.3 to ensure the set $\mathcal{S}$ is of size at most $m+1$, therefore providing an upper bound on the duality gap caused by the lack of convexity (see also [Ekeland and Temam, 1999, Bertsekas, 2014]). We now study what happens to these bounds when using the approximate Shapley-Folkman result in Corollary 6.3 instead of Theorem 3.1. Plugging these last results inside the duality gap bound in Theorem 4.1 yields the following result.

Corollary 6.4 *Suppose the functions f_i in* (P) *satisfy Assumption 2.1. There is a point $x^\star \in \mathbb{R}^d$ at which the primal optimal value of* (CoP) *is attained, and as in* (9) *we let*

$$z^\star = \sum_{i=1}^{n} \begin{pmatrix} f_i^{**}(x_i^\star) \\ A_i x_i^\star \end{pmatrix} + \begin{pmatrix} 0 \\ w-b \end{pmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{m+2} \lambda_{ij} z_{ij} + \begin{pmatrix} 0 \\ w-b \end{pmatrix}$$

with $w \in \mathbb{R}^m_+$ and $z_{ij} \in \mathcal{F}_i$, where $\lambda_{ij} \geq 0$, $\sum_j \lambda_{ij} = 1$. Call $R_v = \max_{\{ij:\lambda_{ij}\neq 1\}} \|\lambda_{ij} z_{ij}\|_2$ and $R_\lambda = \max_{\{ij:\lambda_{ij}\neq 1\}} |\lambda_{ij}|$. Let $\gamma > 0$, we have the following bound on the solution of problem (pP)

$$\underbrace{\mathrm{h}_{CoP}(u_2(s))}_{\text{(pCoP)}} \;\leq\; \underbrace{\mathrm{h}_{P}(u_2(s))}_{\text{(pP)}} \;\leq\; \underbrace{\mathrm{h}_{CoP}(0)}_{\text{(CoP)}} \;+\; \underbrace{|u_1(s)| + \max_{\beta_i \in [1,m+2]} \left\{ \sum_{i=1}^{n} \rho_{\beta_i}(f_i) : \sum_{i=1}^{n} \beta_i = s \right\}}_{\text{gap(s)}}.$$

where

$$\max\{|u_1(s)|, \|u_2(s)\|_2\} \leq \sqrt{2m}\,(R_v + R_\lambda M_V)\,\gamma \tag{6}$$

with

$$s = n + 1 + 2m\,\frac{c}{\gamma^2 + c} \quad \textit{and} \quad M_V = \sup_{\substack{\|u\|_2 \leq 1 \\ v_i \in \mathcal{F}_i}} \left\| \sum_i u_i v_i \right\|_2,$$

for some absolute constant $c > 0$.

Once again, we can take m to be the number of active inequality constraints at $x^\star$. Note that in practice, not all solutions z^* are good starting points for the approximation result described above. Obtaining a good solution typically involves a "purification step" along the lines of [Udell and Boyd, 2016] for example.

7 Conclusion

The Shapley-Folkman theorem bounds the duality gap of separable optimization problems whose objective is a sum of a large number of loosely coupled terms. Our results show that the original gap bound in [Aubin and Ekeland, 1976] is highly conservative and can be relaxed in a number of ways, using e.g. sparse approximations of the solution in the epigraph, and more compact extended formulations of the coupling constraints. In particular, these results reformulate the duality gap bound in terms of *stable quantities*. While these stable bounds on the duality gap are still very conservative, they highlight the fact that *finite sum minimization* problems such as empirical risk minimization are often much *more robust to lack of convexity* than what naive bounds would predict.

References

Jean-Pierre Aubin and Ivar Ekeland. Estimates of the duality gap in nonconvex optimization. *Mathematics of Operations Research*, 1(3):225–245, 1976.

Matthieu Fradelizi, Mokshay Madiman, Arnaud Marsiglietti, and Artem Zvavitch. The convexification effect of minkowski summation. *Preprint*, 2017.

Ross M Starr. Quasi-equilibria in markets with non-convex preferences. *Econometrica: journal of the Econometric Society*, pages 25–38, 1969.

Robert J Aumann and Micha Perles. A variational problem arising in economics. *Journal of Mathematical Analysis and Applications*, 11:488–503, 1965.

Henri Berliocchi and Jean-Michel Lasry. Intégrandes normales et mesures paramétrées en calcul des variations. *Bulletin de la Société Mathématique de France*, 101:129–184, 1973.

Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. SIAM, 1999.

GS Lauer, NR Sandell, DP Bertsekas, and TA Posbergh. Solution of large-scale optimal unit commitment problems. *IEEE Transactions on Power Apparatus and Systems*, (1):79–86, 1982.

Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

Madeleine Udell and Stephen Boyd. Bounding duality gap for separable problems with linear constraints. *Computational Optimization and Applications*, 64(2):355–378, 2016.

G Pisier. Remarques sur un résultat non publié de B. Maurey. *Séminaire Analyse fonctionnelle (dit" Maurey-Schwartz")*, pages 1–12, 1981.

Robert J Serfling. Probability inequalities for the sum in sampling without replacement. *The Annals of Statistics*, pages 39–48, 1974.

Rémi Bardenet, Odalric-Ambrym Maillard, et al. Concentration inequalities for sampling without replacement. *Bernoulli*, 21(3):1361–1385, 2015.

Markus Schneider. Probability inequalities for kernel embeddings in sampling without replacement. In *Artificial Intelligence and Statistics*, pages 66–74, 2016.

Carlo Ciliberto, Alessandro Rudi, Lorenzo Rosasco, and Massimiliano Pontil. Consistent multitask learning with nonlinear output relations. *arXiv preprint arXiv:1705.08118*, 2017.

Nicolas Le Roux, Mark Schmidt, and Francis Bach. A stochastic gradient method with an exponential convergence rate for strongly-convex optimization with finite training sets. *arXiv preprint arXiv:1202.6258*, 2012.

Rie Johnson and Tong Zhang. Accelerating stochastic gradient descent using predictive variance reduction. In *Advances in neural information processing systems*, pages 315–323, 2013.

Aaron Defazio, Francis Bach, and Simon Lacoste-Julien. Saga: A fast incremental gradient method with support for non-strongly convex composite objectives. *arXiv preprint arXiv:1407.0202*, 2014.

Zeyuan Allen-Zhu and Yang Yuan. Improved svrg for non-strongly-convex or sum-of-non-convex objectives. In *International conference on machine learning*, pages 1080–1089, 2016.

Sashank J Reddi, Ahmed Hefny, Suvrit Sra, Barnabas Poczos, and Alex Smola. Stochastic variance reduction for nonconvex optimization. In *International conference on machine learning*, pages 314–323, 2016.

R. T. Rockafellar. *Convex Analysis*. Princeton University Press., Princeton., 1970.

Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. Convex analysis and minimization algorithms. 1993.

Yingjie Bi and Ao Tang. Refined shapely-folkman lemma and its application in duality gap estimation. *arXiv preprint arXiv:1610.05416*, 2016.

Giuseppe Calafiore and Marco C Campi. Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, 2005.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

Roberto Tempo, Giuseppe Calafiore, and Fabrizio Dabbene. *Randomized algorithms for analysis and control of uncertain systems: with applications*. Springer Science & Business Media, 2012.

Xiaojing Zhang, Sergio Grammatico, Georg Schildbach, Paul Goulart, and John Lygeros. On the sample size of random convex programs with structured dependence on the uncertainty. *Automatica*, 60:182–188, 2015.

Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, 1991.

Kanstantsin Pashkovich. *Extended formulations for combinatorial polytopes*. PhD thesis, Universitätsbibliothek, 2012.

Gabor Braun, Samuel Fiorini, Sebastian Pokutta, and David Steurer. Approximation limits of linear programs (beyond hierarchies). In *IEEE 53rd Annual Symposium on Foundations of Computer Science*, 2012.

Nicolas Gillis and François Glineur. On the geometric interpretation of the nonnegative rank. *Linear Algebra and its Applications*, 437(11):2685–2712, 2012.

Michael J Donahue, Christian Darken, Leonid Gurvits, and Eduardo Sontag. Rates of convex approximation in non-hilbert spaces. *Constructive Approximation*, 13(2):187–220, 1997.

Roman Vershynin. How close is the sample covariance matrix to the actual covariance matrix? *Journal of Theoretical Probability*, 25(3):655–686, 2012.

Dong Dai, Philippe Rigollet, Lucy Xia, Tong Zhang, et al. Aggregation of affine estimators. *Electronic Journal of Statistics*, 8(1):302–327, 2014.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Claude Lemaréchal and Arnaud Renaud. A geometric study of duality gaps, with applications. *Mathematical Programming*, 90(3):399–427, 2001.

Iosif Pinelis. Optimum bounds for the distributions of martingales in banach spaces. *The Annals of Probability*, pages 1679–1706, 1994.

Karthik Sridharan. A gentle introduction to concentration inequalities.

8 Supplementary Material

We now detail full proofs of the results discussed in the paper.

8.1 Duality & Convex Relaxations

We use the biconjugate to produce a convex relaxation of problem (P) written

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n f_i^{**}(x_i) \\ \text{subject to} & Ax \leq b \end{array} \tag{CoP}$$

in the variables $x_i \in \mathbb{R}^{d_i}$. Writing the epigraph of problem (P) as in [Boyd and Vandenberghe, 2004, §5.3] or [Lemaréchal and Renaud, 2001],

$$\mathcal{G} \triangleq \left\{ (x, r_0, r) \in \mathbb{R}^{d+1+m} : \sum_{i=1}^n f_i(x_i) \leq r_0,\ Ax - b \leq r \right\},$$

and its projection on the last $m+1$ coordinates,

$$\mathcal{G}_r \triangleq \left\{ (r_0, r) \in \mathbb{R}^{m+1} : (x, r_0, r) \in \mathcal{G} \right\}, \tag{7}$$

we can write the Lagrange dual function of (P) as

$$\Psi(\lambda) \triangleq \inf \left\{ r_0 + \lambda^\top r : (r_0, r) \in \mathcal{G}_r^{**} \right\}, \tag{8}$$

in the variable $\lambda \in \mathbb{R}^m$, where $\mathcal{G}^{**} = \overline{\mathbf{Co}(\mathcal{G})}$ is the closed convex hull of the epigraph $\mathcal{G}$ (the projection being linear here, we have $(\mathcal{G}_r)^{**} = (\mathcal{G}^{**})_r = \mathcal{G}_r^{**}$). We need constraint qualification conditions for strong duality to hold in (CoP) and we now recall the result in [Lemaréchal and Renaud, 2001, Th. 2.11] which shows that because the explicit constraints are affine here, the dual functions of (P) and (CoP) are equal. The (common) dual of (P) and (CoP) is then

$$\sup_{\lambda \geq 0} \Psi(\lambda) \tag{D}$$

in the variable $\lambda \in \mathbb{R}^m$. The following result shows that strong duality holds under mild technical assumptions.

Theorem 8.1 *[Lemaréchal and Renaud, 2001, Th. 2.11] The function $\Psi(\lambda)$ is also the dual function associated with* (CoP)*. Assuming that Ψ is not constant equal to $-\infty$ and that there is a feasible x in the relative interior of $\mathbf{dom}\left(\sum_{i=1}^n f_i^{**}\right)$ then Ψ attains its maximum and*

$$\max_\lambda \Psi(\lambda) = \inf \left\{ \sum_{i=1}^n f_i^{**}(x_i) : x \in \mathbb{R}^d,\ Ax \leq b \right\}$$

i.e. strong duality holds.

This last result shows that the convex problem (CoP) indeed shares the same dual as problem (P).

8.2 Perturbed Problems

In the next section, perturbed versions of problems (P) and (CoP) will emerge to quanitfy our approximation bounds. These are written respectively

$$\begin{array}{lll} \mathrm{h}_P(u) \triangleq & \text{min.} & \sum_{i=1}^n f_i(x_i) \\ & \text{s.t.} & Ax - b \leq u \\ & & x_i \in Y_i, \quad i = 1, \ldots, n, \end{array} \tag{pP}$$

in the variables $x_i \in \mathbb{R}^{d_i}$, with perturbation parameter $u \in \mathbb{R}^m$, and

$$\begin{array}{lll} \mathrm{h}_{CoP}(u) \triangleq & \text{min.} & \sum_{i=1}^n f_i^{**}(x_i) \\ & \text{s.t.} & Ax - b \leq u \end{array} \tag{pCoP}$$

in the variables $x_i \in \mathbb{R}^{d_i}$, with perturbation parameter $u \in \mathbb{R}^m$.

8.3 The Shapley-Folkman Theorem

We now recall Carathéodory's result, and its conic formulation, which underpin all the other results in this section.

Theorem 8.2 (Carathéodory) *Let $V \subset \mathbb{R}^n$, then $x \in \mathbf{Co}(V)$ if and only if*

$$x = \sum_{i=1}^{n+1} \lambda_i v_i$$

for some $v_i \in V$, $\lambda_i \geq 0$ and $\mathbf{1}^\top \lambda = 1$.

Similarly, if we write $\mathbf{Po}(V)$ the conic hull of V, with $\mathbf{Po}(V) = \{\sum_i \lambda_i v_i : v_i \in V, \lambda_i \geq 0$, we have the following result (see e.g. [Rockafellar, 1970, Cor. 17.1.2]).

Theorem 8.3 (Conic Carathéodory) *Let $V \subset \mathbb{R}^n$, then $x \in \mathbf{Po}(V)$ if and only if*

$$x = \sum_{i=1}^{n} \lambda_i v_i$$

for some $v_i \in V$, $\lambda_i \geq 0$.

Theorem 8.4 (Shapley-Folkman) *Let $V_i \in \mathbb{R}^d$, $i = 1, \ldots, n$ be a family of subsets of $\mathbb{R}^d$. If*

$$x \in \mathbf{Co}\left(\sum_{i=1}^{n} V_i\right) = \sum_{i=1}^{n} \mathbf{Co}\left(V_i\right)$$

then

$$x \in \sum_{[1,n]\setminus\mathcal{S}} V_i \; + \; \sum_{\mathcal{S}} \mathbf{Co}(V_i)$$

where $|\mathcal{S}| \leq d$.

Proof. Suppose $x \in \sum_{i=1}^{n} \mathbf{Co}\left(V_i\right)$, then by Carathéodory's theorem we can write $x = \sum_{i=1}^{n} \sum_{j=1}^{d+1} \lambda_{ij} v_{ij}$ where $v_{ij} \in V_i$ and $\lambda_{ij} \geq 0$ with $\sum_{j=1}^{d+1} \lambda_{ij} = 1$. These constraints can be summarized as

$$z = \sum_{i=1}^{n} \sum_{j=1}^{d+1} \lambda_{ij} z_{ij}$$

where $z \in \mathbb{R}^{d+n}$ and

$$z = \begin{pmatrix} x \\ \mathbf{1}_n \end{pmatrix}, \quad z_{ij} = \begin{pmatrix} v_{ij} \\ e_i \end{pmatrix}, \quad \text{for } i = 1, \ldots, n \text{ and } j = 1, \ldots, d+1,$$

with $e_i \in \mathbb{R}^n$ is the Euclidean basis. Since z is a conic combination of the z_{ij}, there exist coefficients $\mu_{ij} \geq 0$ such that $z = \sum_{i=1}^{n} \sum_{j=1}^{d+1} \mu_{ij} z_{ij}$ and at most $d+n$ coefficients μ_{ij} are nonzero. Then, $\sum_{j=1}^{d+1} \mu_{ij} = 1$ means that a single $\mu_{ij} = 1$ for $i \in [1, n] \setminus \mathcal{S}$ where $|\mathcal{S}| \leq d$ (since $n + d$ nonzero coefficients are spread among n sets, with at least one nonzero coefficient per set), and $\sum_{j=1}^{d+1} \mu_i v_{ij} \in V_i$ for $i \in [1, n] \setminus \mathcal{S}$. ■

8.4 Duality Gap Bounds

Proposition 8.5 *Suppose the functions f_i in* (P) *satisfy Assumption 2.1. There is a point $x^\star \in \mathbb{R}^d$ at which the primal optimal value of* (CoP) *is attained, such that*

$$\underbrace{\sum_{i=1}^{n} f_i^{**}(x_i^\star)}_{CoP} \;\leq\; \underbrace{\sum_{i=1}^{n} f_i(\hat{x}_i^\star)}_{P} \;\leq\; \underbrace{\sum_{i=1}^{n} f_i^{**}(x_i^\star)}_{CoP} \;+\; \underbrace{\sum_{i\in\mathcal{S}} \rho(f_i)}_{\text{gap}}$$

with $\hat{x}^\star$ is an optimal point of (P)*, and*

$$\mathcal{S} \triangleq \{i : \ (f_i^{**}(x_i^\star), A_i x_i^\star) \notin \mathbf{Ext}(\mathcal{F}_i)\}$$

where $\mathcal{F}_i \subset \mathbb{R}^{m+1}$ is defined as

$$\mathcal{F}_i = \left\{(f_i^{**}(x_i), A_i x_i) : \ x_i \in \mathbb{R}^{d_i}\right\}$$

writing $A_i \in \mathbb{R}^{m \times d_i}$ the i^{th} block of A.

Proof. Using [Lemaréchal and Renaud, 2001, Cor. A.6], we know

$$\mathcal{G}_r^{**} = \left\{(r_0, r) \in \mathbb{R}^{m+1} : \ \sum_{i=1}^n f_i^{**}(x_i) \leq r_0, \ Ax - b \leq r\right\}.$$

Since $\mathcal{G}_r^{**}$ is closed by construction and the sets $\mathcal{F}_i$ are closed by Assumption 2.1, there is a point $x^\star \in \mathcal{G}_r^{**}$ which attains the primal optimal value in (CoP). We write the corresponding minimizer of (8) in $\mathcal{G}_r^{**}$ as

$$z^\star = \sum_{i=1}^n \begin{pmatrix} f_i^{**}(x_i^\star) \\ A_i x_i^\star \end{pmatrix} + \begin{pmatrix} 0 \\ w - b \end{pmatrix} \tag{9}$$

with $w \in \mathbb{R}_+^m$, which we summarize as

$$z^\star = \sum_{i=1}^n z^{(i)} + \begin{pmatrix} 0 \\ w - b \end{pmatrix},$$

where $z^{(i)} \in \mathcal{F}_i$. Since $f_i^{**}(x) = f_i(x)$ when $x \in \mathbf{Ext}(\mathcal{F}_i)$ because $\mathbf{epi}(f^{**}) = \overline{\mathbf{Co}(\mathbf{epi}(f))}$ when Assumption 2.1 holds, we have

$$\begin{aligned}
\overbrace{\sum_{i=1}^n f_i^{**}(x_i^\star)}^{CoP} \quad &= \quad \sum_{i \in [1,n] \setminus \mathcal{S}} f_i^{**}(x_i^\star) + \sum_{i \in \mathcal{S}} \sum_{j \in [1,m+2]} \lambda_{ij} f_i^{**}(x_{ij}^\star) \\
&= \quad \sum_{i \in [1,n] \setminus \mathcal{S}} f_i(x_i^\star) + \sum_{i \in \mathcal{S}} \sum_{j \in [1,m+2]} \lambda_{ij} f_i^{**}(x_{ij}^\star) \\
&\geq \quad \sum_{i \in [1,n] \setminus \mathcal{S}} f_i(x_i^\star) + \sum_{i \in \mathcal{S}} f_i^{**}(\tilde{x}_i) \\
&\geq \quad \sum_{i \in [1,n] \setminus \mathcal{S}} f_i(x_i^\star) + \sum_{i \in \mathcal{S}} f_i(\tilde{x}_i) - \sum_{i \in \mathcal{S}} \rho(f_i) \\
&\geq \quad \underbrace{\sum_{i=1}^n f_i(\hat{x}_i^\star)}_{P} \ - \ \sum_{i \in \mathcal{S}} \rho(f_i)
\end{aligned}$$

calling $\tilde{x}_i = \sum_{j \in [1,m+2]} \lambda_{ij} x_i^\star$, where $\lambda_{ij} \geq 0$ and $\sum_j \lambda_{ij} = 1$. The last inequality holds because the points $x_i^\star, \tilde{x}_i$ are feasible for (P), i.e.

$$\sum_{i \in [1,n] \setminus \mathcal{S}} A_i x_i^\star + \sum_{i \in \mathcal{S}} \sum_{j \in [1,m+2]} \lambda_{ij} A_i x_{ij}^\star \leq b,$$

means that

$$\sum_{i \in [1,n] \setminus \mathcal{S}} A_i x_i^\star + \sum_{i \in \mathcal{S}} A_i \tilde{x}_i \leq b,$$

which yields the desired result. ■

Proposition 8.6 *Suppose the functions f_i in* (P) *satisfy Assumption 2.1. There is a point $x^\star \in \mathbb{R}^d$ at which the primal optimal value of* (CoP) *is attained, such that*

$$\underbrace{\sum_{i=1}^n f_i^{**}(x_i^\star)}_{CoP} \;\leq\; \underbrace{\sum_{i=1}^n f_i(\hat{x}_i^\star)}_{P} \;\leq\; \underbrace{\sum_{i=1}^n f_i^{**}(x_i^\star)}_{CoP} \;+\; \underbrace{\sum_{i=1}^{m+1} \rho(f_{[i]})}_{\text{gap}}$$

where $\hat{x}^\star$ is an optimal point of (P) *and $\rho(f_{[1]}) \geq \rho(f_{[2]}) \geq \ldots \geq \rho(f_{[n]})$.*

Proof. Notice that the closed convex hull $\mathcal{G}_r^{**}$ of the epigraph of problem (P) can be written as a Minkowski sum, with

$$\mathcal{G}_r^{**} = \sum_{i=1}^n \mathcal{F}_i + (0, -b) + \mathbb{R}_+^{m+1}, \quad \text{where} \quad \mathcal{F}_i = \left\{(f_i^{**}(x_i), A_i x_i) : \; x_i \in \mathbb{R}^{d_i}\right\} \subset \mathbb{R}^{m+1}$$

The Krein-Milman theorem shows

$$\mathcal{G}_r^{**} = \sum_{i=1}^n \mathbf{Co}\left(\mathbf{Ext}(\mathcal{F}_i)\right) + (0, -b) + \mathbb{R}_+^{m+1}.$$

Now, since $\mathcal{F}_i \subset \mathbb{R}^{m+1}$, the Shapley Folkman Theorem 3.1 shows that the point $z^\star \in \mathcal{G}_r^{**}$ in (9) satisfies

$$z^\star \in \sum_{[1,n]\setminus\mathcal{S}} \mathbf{Ext}(\mathcal{F}_i) \;+\; \sum_{\mathcal{S}} \mathbf{Co}\left(\mathbf{Ext}(\mathcal{F}_i)\right)$$

for some set $\mathcal{S} \subset [1, n]$ with $|\mathcal{S}| \leq m+1$. This means that we can take $|\mathcal{S}| \leq m+1$ in Proposition 2.3 and yields the desired result. ■

Theorem 8.7 *Suppose the functions f_i in* (P) *satisfy Assumption 2.1. There is a point $x^\star \in \mathbb{R}^d$ at which the primal optimal value of* (CoP) *is attained, and as in* (9) *we let*

$$z^\star = \sum_{i=1}^n \begin{pmatrix} f_i^{**}(x_i^\star) \\ A_i x_i^\star \end{pmatrix} + \begin{pmatrix} 0 \\ w - b \end{pmatrix}$$

with $w \in \mathbb{R}_+^m$ be the corresponding minimizer in (8)*. Suppose that we use an approximate conic representation of $z^\star$ using only $s \in [n, n+m+1]$ coefficients, writing*

$$\lambda(s) = \underset{\substack{\lambda_{ij}\geq 0 \\ z_{ij}\in\mathcal{F}_i}}{\operatorname{argmin}} \left\{ \left\| z^\star - \sum_{i=1}^n \sum_{j=1}^{m+2} \lambda_{ij} z_{ij} \right\| : \; \sum_{i=1}^n \mathbf{Card}(\lambda_i) \leq s, \; \mathbf{1}^T \lambda_i = 1, i = 1, \ldots, n \right\}$$

where $z_{ij} \in \mathcal{F}_i$ for $i = 1, \ldots, n$, $j = 1, \ldots, m+2$, and $u(s) = z^\star - \sum_{i=1}^n \sum_{j=1}^{m+2} \lambda_{ij}(s) z_{ij}$. We have the following bound on the solution of problem (pP)

$$\underbrace{\mathrm{h}_{CoP}(u_2(s))}_{\text{(pCoP)}} \;\leq\; \underbrace{\mathrm{h}_P(u_2(s))}_{\text{(pP)}} \;\leq\; \underbrace{\mathrm{h}_{CoP}(0)}_{\text{(CoP)}} \;+\; \underbrace{|u_1(s)| + \max_{\beta_i\in[1,m+2]} \left\{ \sum_{i=1}^n \rho_{\beta_i}(f_i) : \sum_{i=1}^n \beta_i = s \right\}}_{\text{gap(s)}}.$$

Furthermore, we can take m to be the number of active inequality constraints at $x^\star$.

Proof. Let $\bar{z} = \sum_{i=1}^n \sum_{j=1}^{m+2} \lambda_{ij}(s) z_{ij}$. By construction, this point satisfies

$$\sum_{i=1}^n z_1^{(i)} = \sum_{i=1}^n \bar{z}_1^{(i)} + u_1(s) = \sum_{i=1}^n f_i^{**}(x_i) + u_1(s), \quad \text{and} \quad \sum_{i=1}^n \bar{z}_{[2,m+1]}^{(i)} - b \leq u_2(s),$$

where $z^{(i)}_{[2,m+1]} = A_i x_i^\star$. Since $f_i^{**}(x) = f_i(x)$ when $x \in \mathbf{Ext}(\mathcal{F}_i)$ because $\mathbf{epi}(f^{**}) = \mathbf{Co}(\mathbf{epi}(f))$ when Assumption 2.1 holds, we have

$$\begin{aligned}
\overbrace{\sum_{i=1}^{n} f_i^{**}(x_i^\star)}^{CoP} &= \sum_{i\in[1,n]\setminus\mathcal{S}} f_i^{**}(x_i) + \sum_{i\in\mathcal{S}}\sum_{j\in[1,m+2]} \lambda_{ij} f_i^{**}(x_{ij}) + u_1(s)\\
&= \sum_{i\in[1,n]\setminus\mathcal{S}} f_i(x_i) + \sum_{i\in\mathcal{S}}\sum_{j\in[1,m+2]} \lambda_{ij} f_i^{**}(x_{ij}) + u_1(s)\\
&\geq \sum_{i\in[1,n]\setminus\mathcal{S}} f_i(x_i) + \sum_{i\in\mathcal{S}} f_i^{**}\left(\tilde{x}_i\right) + u_1(s)\\
&\geq \sum_{i\in[1,n]\setminus\mathcal{S}} f_i(x_i) + \sum_{i\in\mathcal{S}} f_i\left(\tilde{x}_i\right) - \sum_{i\in\mathcal{S}}\rho(f_i) + u_1(s)\\
&\geq \underbrace{\sum_{i=1}^{n} f_i(x_i)}_{pP} \; - \; \sum_{i\in\mathcal{S}}\rho(f_i) + u_1(s)
\end{aligned}$$

calling $\tilde{x}_i = \sum_{j\in[1,m+2]} \lambda_{ij} x_i$, where $\lambda_{ij} \geq 0$ and $\sum_j \lambda_{ij} = 1$. The last inequality holds because the points $\tilde{x}_i$ are feasible for (pP) with perturbation $u_2(s)$, i.e.

$$\sum_{i\in[1,n]\setminus\mathcal{S}} A_i x_i + \sum_{i\in\mathcal{S}}\sum_{j\in[1,m+2]} \lambda_{ij} A_i x_{ij} \leq b + u_2(s),$$

means that

$$\sum_{i\in[1,n]\setminus\mathcal{S}} A_i x_i + \sum_{i\in\mathcal{S}} A_i \tilde{x}_i \leq b + u_2(s),$$

which yields the desired result. ■

8.5 Approximate Carathéodory

Theorem 8.8 *Let $x = \sum_{j=1}^{N} \lambda_j V_j$ for $V \in \mathbb{R}^{d\times N}$ and some $\lambda \in \mathbb{R}^N$ such that $\mathbf{1}^T\lambda = 1, \lambda \geq 0$. Let $\varepsilon > 0$ and write $R = \max\{R_v, R_\lambda\}$ where $R_v = \max_i \|\lambda_i V_i\|_\infty$ and $R_\lambda = \max_i |\lambda_i|$. Then, there exists some $\hat{x} = \sum_{j\in\mathcal{J}} \mu_j V_j$ with $\mu \in \mathbb{R}^m$ and $\mu \geq 0$, where $\mathcal{J} \subset [1, N]$ has size*

$$|\mathcal{J}| = 1 + N\,\frac{\log(2d)(\sqrt{N}\,R/\varepsilon)^2}{2 + \log(2d)(\sqrt{N}\,R/\varepsilon)^2}$$

and is such that $\|x - \hat{x}\|_\infty \leq \varepsilon$ and $|\sum_{j\in\mathcal{J}} \mu_j - 1| \leq \varepsilon$.

Proof. Denote by $x_i \triangleq \sum_{j=1}^{N} \lambda_i V_{ij}$. Let

$$S_m^{(i)} = \sum_{j\in\mathcal{J}} \lambda_j V_{ij}$$

where $\mathcal{J}$ is a random subset of $[1, N]$ of size m. A Serfling-like concentration inequality will give

$$\mathbf{Prob}\left(\left|\frac{1}{m}S_m^{(i)} - \frac{1}{N}x_i\right| \geq \varepsilon\right) \leq f(\varepsilon)\,.$$

Hence for any $\varepsilon > 0$

$$\mathbf{Prob}\left(\left|\frac{N}{m}S_m^{(i)} - x_i\right| \geq \varepsilon\right) \leq f(\varepsilon/N)\,.$$

In particular [Serfling, 1974, Cor 1.1] shows

$$\mathbf{Prob}\left(\left|\frac{N}{m}S_m^{(i)} - x_i\right| \geq \varepsilon\right) \leq \exp\left(\frac{-\alpha_m\varepsilon^2}{2N(1-\alpha_m)R_v^2}\right)$$

where $\alpha_m = (m-1)/N$ is the sampling ratio. A union bound then means that setting

$$\frac{\alpha_m}{1-\alpha_m} \geq \frac{\log(2d)(\sqrt{N}\,R_v)^2}{2\varepsilon^2}$$

or again

$$\alpha_m \geq \frac{\log(2d)(\sqrt{N}\,R_v)^2/2\varepsilon^2}{1+\log(2d)(\sqrt{N}\,R_v)^2/2\varepsilon^2}$$

ensures $\|x-\hat{x}\|_\infty \leq \varepsilon$ with probability at least $1/2$. A similar reasoning, picking this time

$$S_m^{(i)} = \sum_{j\in\mathcal{S}} \lambda_j,$$

ensures $\mu = \frac{N}{m}\lambda$ satisfies $|\sum_{j\in\mathcal{S}}\mu_j - 1| \leq \varepsilon$ with probability at least $1-1/2d$ since $R = \max\{R_v, R_\lambda\}$, which yields the desired result. ■

Theorem 8.9 (Approximate Carathéodory with High Sampling Ratio) *Let $x = \sum_{j=1}^N \lambda_j V_j$ for $V \in \mathbb{R}^{d\times N}$ and some $\lambda \in \mathbb{R}^N$ such that $\mathbf{1}^T\lambda = 1, \lambda \geq 0$. Let $\varepsilon > 0$ and write $R = \max\{R_v, R_\lambda\}$ where $R_v = \max_i \|\lambda_i V_i\|$ and $R_\lambda = \max_i |\lambda_i|$, for some norm $\|\cdot\|$ such that $(\mathbb{R}^d, \|\cdot\|)$ is $(2,D)$-smooth. Then, there exists some $\hat{x} = \sum_{j\in\mathcal{J}} \mu_j V_j$ with $\mu \in \mathbb{R}^m$ and $\mu \geq 0$, where $\mathcal{J} \subset [1,N]$ has size*

$$|\mathcal{J}| = 1 + N\,\frac{c(\sqrt{N}\,D\,R/\varepsilon)^2}{1+c(\sqrt{N}\,D\,R/\varepsilon)^2}$$

for some absolute constant $c > 0$, and is such that $\|x-\hat{x}\| \leq \varepsilon$ and $|\sum_{j\in\mathcal{J}}\mu_j - 1| \leq \varepsilon$.

Proof. We use [Schneider, 2016, Th. 1] instead of [Serfling, 1974, Cor 1.1] in the proof of Theorem 6.2. This means imposing

$$\alpha_m \geq \frac{c(\sqrt{N}\,R\,D/\varepsilon)^2}{1+c(\sqrt{N}\,R\,D/\varepsilon)^2}$$

Finally, $R = \max\{R_v, R_\lambda\} \geq R_\lambda$ ensures that the Hoeffding like bound in [Serfling, 1974] also holds, with $|\sum_{j\in\mathcal{S}}\mu_j - 1| \leq \varepsilon$, and yields the desired result. ■

Theorem 8.10 (Approximate Shapley-Folkman) *Let $\varepsilon, \beta, \gamma > 0$ and $V_i \in \mathbb{R}^d$, $i = 1, \ldots, n$ be a family of subsets of $\mathbb{R}^d$. Suppose*

$$x = \sum_{i=1}^n \sum_{j=1}^{d+1} \lambda_{ij} v_{ij} \;\in\; \sum_{i=1}^n \mathbf{Co}\,(V_i)$$

where $\lambda_{ij} \geq 0$ and $\sum_j \lambda_{ij} = 1$. We write $R = \max\{\beta R_v, \gamma R_\lambda\}$ where $R_v = \max_{\{ij:\lambda_{ij}\neq 1\}} \|\lambda_{ij}v_{ij}\|$ and $R_\lambda = \max_{\{ij:\lambda_{ij}\neq 1\}} |\lambda_{ij}|$, for some norm $\|\cdot\|$ such that $(\mathbb{R}^d, \|\cdot\|)$ is $(2,D)$-smooth. Then there exists a point $\hat{x} \in \mathbb{R}^d$, coefficients $\mu_i \geq 0$ and index sets $\mathcal{S}, \mathcal{T} \subset [1,n]$ with $\mathcal{S}\cap\mathcal{T} = \emptyset$ such that $q \triangleq |\mathcal{S}| + |\mathcal{T}| \leq d$, and

$$\hat{x} \in \; \textstyle\sum_{[1,n]\setminus(\mathcal{S}\cup\mathcal{T})} V_i \;+\; \sum_{i\in\mathcal{T}} \mu_i V_i \;+\; \sum_{i\in\mathcal{S}} \mu_i \mathbf{Co}(V_i)$$

with

$$\|x-\hat{x}\| \leq \frac{q}{\beta}\varepsilon, \qquad \left|\sum_{i\in\mathcal{S}\cup\mathcal{T}} \mu_i - q\right| \leq q\varepsilon \qquad \textit{and} \qquad \left(\sum_{i\in\mathcal{S}\cup\mathcal{T}} (\mu_i-1)^2\right)^{1/2} \leq \frac{q}{\gamma}\varepsilon.$$

where $|\mathcal{S}| \leq (m - |\mathcal{T}|)/2$ with

$$m = 1 + (d+q)\,\frac{c(\sqrt{d+q}\,R/q\varepsilon)^2}{1+c(\sqrt{d+q}\,R/q\varepsilon)^2}.$$

hence, in particular, $|\mathcal{S}| \leq m - q$.

Proof. If $x \in \sum_{i=1}^n \mathbf{Co}(V_i)$, as in the proof of Theorem 3.1 above, we can write

$$z = \sum_{i=1}^{n} \sum_{j=1}^{d+1} \lambda_{ij} z_{ij}$$

where $z \in \mathbb{R}^{d+n}$ and

$$z = \begin{pmatrix} \beta x \\ \gamma \mathbf{1}_n \end{pmatrix}, \quad z_{ij} = \begin{pmatrix} \beta v_{ij} \\ \gamma e_i \end{pmatrix}, \quad \text{for } i = 1, \ldots, n \text{ and } j = 1, \ldots, d+1,$$

with $e_i \in \mathbb{R}^n$ is the Euclidean basis, $\gamma, \beta > 0$ and by the classical Carathéodory bound, at most $d + n$ coefficients λ_{ij} are nonzero (note the extra scaling factors $\gamma, \beta > 0$ here compared to Theorem 3.1). Let us call $\mathcal{I} \subset [1, n]$ the set of indices such that $i \in \mathcal{I}$ iff at least two coefficients in $\{\lambda_{ij} : j \in [1, d+1]\}$ are nonzero. As in Theorem 3.1, we must have $|\mathcal{I}| \leq d$. We write

$$\frac{y}{|\mathcal{I}|} = \sum_{i \in \mathcal{I}} \sum_{j=1}^{d+1} \frac{\lambda_{ij}}{|\mathcal{I}|} z_{ij}$$

where $\sum_{i \in \mathcal{I}} \sum_{j=1}^{d+1} \lambda_{ij}/|\mathcal{I}| = 1$ and at most $d + |\mathcal{I}|$ coefficients λ_{ij} are nonzero. We will apply the result of Theorem 8.9 twice here with radius R/q where $q = |\mathcal{I}|$. Once on the upper block of the vectors z_{ij} using the norm $\|\cdot\|$ and then on the lower blocks of these vectors (corresponding to the constraints on λ_{ij}), using the ℓ_2 norm to exploit the fact that these lower blocks have comparatively low ℓ_2 radius.

Theorem 8.9 applied to the upper block of y/q and of the vectors z_{ij} shows that with probability higher than $1/2$ there exists some $\hat{x}/|\mathcal{I}| = \sum_{i \in \mathcal{I}} \sum_{j=1}^{d+1} \mu_{ij} v_{ij}$ with $|\sum_{i \in \mathcal{I}} \sum_{j=1}^{d+1} \mu_{ij} - 1| \leq \varepsilon$, $\mu \geq 0$, where at most m coefficients μ_{ij} are nonzero and

$$\left\| x - \sum_{i \in [1,n] \setminus \mathcal{I}} v_i - \hat{x} \right\| \leq |\mathcal{I}| \varepsilon / \beta.$$

for some $v_i \in V_i$. Then, Theorem 8.9 applied to the lower block of the vectors z_{ij} shows that with probability higher than $1/2$ the weights μ_{ij} sampled above satisfy

$$\left(\sum_{i \in \mathcal{I}} \left(\sum_{j=1}^{d+1} |\mathcal{I}| \mu_{ij} - 1 \right)^2 \right)^{1/2} \leq \tfrac{|\mathcal{I}|}{\gamma} \varepsilon.$$

with the ℓ_2 norm being $D = 1$ smooth. Setting $\mathcal{I} = \mathcal{S} \cup \mathcal{T}$, and since m nonzero coefficients are spread among q sets, we have $|\mathcal{S}| \leq m - q$. Setting $\mu_i = \sum_j |\mathcal{I}| \mu_{ij}$ then yields the desired result. ■

Corollary 8.11 *Let $\varepsilon > 0$ and $V_i \in \mathbb{R}^d$, $i = 1, \ldots, n$ be a family of subsets of $\mathbb{R}^d$. Suppose*

$$x = \sum_{i=1}^{n} \sum_{j=1}^{d+1} \lambda_{ij} v_{ij} \; \in \; \sum_{i=1}^{n} \mathbf{Co}(V_i)$$

where $\lambda_{ij} \geq 0$ and $\sum_j \lambda_{ij} = 1$. We write $R_v = \max_{\{ij : \lambda_{ij} \neq 1\}} \|\lambda_{ij} v_{ij}\|$ and $R_\lambda = \max_{\{ij : \lambda_{ij} \neq 1\}} |\lambda_{ij}|$, for some norm $\|\cdot\|$ such that $(\mathbb{R}^d, \|\cdot\|)$ is $(2, D)$-smooth. There exists a point $\bar{x}$ and an index set $\mathcal{S} \subset [1, n]$ such that

$$\bar{x} \in \sum_{[1,n] \setminus \mathcal{S}} V_i + \sum_{i \in \mathcal{S}} \mathbf{Co}(V_i) \quad \textit{with} \quad \|x - \bar{x}\| \leq \sqrt{2d} \left(\frac{R_v}{R_\lambda} + M_V \right) \varepsilon$$

where $|\mathcal{S}| \leq m - d$ with

$$m = 1 + 2d \frac{c\,(DR_\lambda/\varepsilon)^2}{1 + c\,(DR_\lambda/\varepsilon)^2} \qquad \textit{and} \qquad M_V = \sup_{\substack{\|u\|_2 \leq 1 \\ v_i \in V_i}} \left\| \sum_i u_i v_i \right\|.$$

where $c > 0$ is an absolute constant.

Proof. Theorem 8.10 means there exists $\hat{x} \in \mathbb{R}^d$, coefficients $\mu_i \geq 0$ and index sets $\mathcal{S}, \mathcal{T} \subset [1, n]$ such that

$$\begin{aligned}\hat{x} \;\in\; & \sum_{[1,n]\setminus(\mathcal{S}\cup\mathcal{T})} V_i \;+\; \sum_{i\in\mathcal{T}} \mu_i V_i \;+\; \sum_{i\in\mathcal{S}} \mu_i \mathbf{Co}(V_i) \\ \subset\; & \sum_{[1,n]\setminus\mathcal{S}} V_i \;+\; \sum_{i\in\mathcal{S}} \mathbf{Co}(V_i) \;+\; \sum_{i\in\mathcal{T}} (\mu_i - 1)\, V_i \;+\; \sum_{i\in\mathcal{S}} (\mu_i - 1)\, \mathbf{Co}(V_i)\end{aligned}$$

with

$$\left(\textstyle\sum_{i\in\mathcal{I}} (\mu_i - 1)^2\right)^{1/2} \leq \tfrac{q}{\gamma}\varepsilon. \qquad \text{and} \qquad \|x - \hat{x}\| \leq \tfrac{q}{\beta}\varepsilon$$

where $q \triangleq |\mathcal{S}| + |\mathcal{T}| \leq d$. Saturating the max term in R in Theorem 8.9 means setting $\beta R_v = \gamma R_\lambda$. Setting $\gamma = q/\sqrt{d+q}$ then yields $\|x - \hat{x}\| \leq \sqrt{d+q}\frac{R_v}{R_\lambda}\varepsilon$ and

$$\left(\textstyle\sum_{i\in\mathcal{I}} (\mu_i - 1)^2\right)^{1/2} \leq \sqrt{d+q}\,\varepsilon.$$

and the fact that

$$v \in \sum_{i\in\mathcal{T}} (\mu_i - 1)\, V_i \;+\; \sum_{i\in\mathcal{S}} (\mu_i - 1)\, \mathbf{Co}(V_i)$$

means

$$\|v\| \leq M_V \left(\textstyle\sum_{i\in\mathcal{I}} (\mu_i - 1)^2\right)^{1/2}$$

and yields the desired result. ■

8.6 Separable Constrained Problems

Here, we briefly show how to extend our previous to problems with separable *nonlinear* constraints. We now focus on a more general formulation of optimization problem (P), written

$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^n f_i(x_i) \\ \text{subject to} & \sum_{i=1}^n g_i(x_i) \leq b, \\ & x_i \in Y_i, \quad i = 1, \ldots, n,\end{array} \tag{cP}$$

where the g_i's take values in $\mathbb{R}^m$. We assume that the functions g_i are lower semicontinuous. Since the constraints are not necessarily affine anymore, we cannot use the convex envelope to derive the dual problem. The dual now takes the generic form

$$\sup_{\lambda \geq 0} \Psi(\lambda), \tag{cD}$$

where Ψ is the dual function associated to problem (cP). Note that deriving this dual explicitly may be hard. As for problem (P), we will also use the perturbed version of problem (cP), defined as

$$\mathrm{h}_{cP}(u) \triangleq \begin{array}{ll}\text{min.} & \sum_{i=1}^n f_i(x_i) \\ \text{s.t.} & \sum_{i=1}^n g_i(x_i) - b \leq u \\ & x_i \in Y_i, \quad i = 1, \ldots, n,\end{array} \tag{p-cP}$$

in the variables $x_i \in \mathbb{R}^{d_i}$, with perturbation parameter $u \in \mathbb{R}^m$. We let $\mathrm{h}_{cD} \triangleq \mathrm{h}_{cP}^{**}$ and in particular, solving for $\mathrm{h}_{cD}(0)$ is equivalent to solving problem (cD). Using these new definitions, we can formulate a more general bound for the duality gap (see [Ekeland and Temam, 1999, Appendix I, Thm. 3] for more details).

Proposition 8.12 *Suppose the functions f_i and g_i in* (cP) *are such that all $(f_i + \mathbf{1}^\top g_i)$ satisfy Assumption 2.1. Then, one has*

$$\mathrm{h}_{cD}((m+1)\bar{\rho}_g) \leq \mathrm{h}_{cP}((m+1)\bar{\rho}_g) \leq \mathrm{h}_{cD}(0) + (m+1)\bar{\rho}_f,$$

where $\bar{\rho}_f = \sup_{i\in[1,n]} \rho(f_i)$ and $\bar{\rho}_g = \sup_{i\in[1,n]} \rho(g_i)$.

Proof. Similar to Proposition 3.2, using the graph of h_{cP} instead of the $\mathcal{F}_i$'s. ■

We then get a direct extension of Corollary 6.4, as follows.

Corollary 8.13 *Suppose the functions f_i and g_i in* (cP) *are such that all $(f_i + \mathbf{1}^\top g_i)$ satisfy Assumption 2.1. There exist points $x_{ij}^\star \in \mathbb{R}^{d_i}$ and $w \in \mathbb{R}^m$ such that*

$$z^\star = \sum_{i=1}^{n} \sum_{j=1}^{m+2} \lambda_{ij}(f_i(x_{ij}^\star), g_i(x_{ij}^\star)) + (0, -b + w),$$

attains the minimum in (cD)*, where $\lambda_{ij} \geq 0$ and $\sum_j \lambda_{ij} = 1$. Call $R_v = \max_{\{ij:\lambda_{ij} \neq 1\}} \|\lambda_{ij} z_{ij}\|_2$ and $R_\lambda = \max_{\{ij:\lambda_{ij} \neq 1\}} |\lambda_{ij}|$. Let $\gamma > 0$, we have the following bound on the solution of problem* (cP)

$$\begin{aligned}\underbrace{\mathrm{h}_{cD}(u_2(s) + (m+1)\bar{\rho}_g \mathbf{1})}_{\text{(cD)}} &\leq \underbrace{\mathrm{h}_P(u_2(s) + (m+1)\bar{\rho}_g \mathbf{1})}_{\text{(p-cP)}} \\ &\leq \underbrace{\mathrm{h}_{cD}(0)}_{\text{(cD)}} + \underbrace{|u_1(s)| + \max_{\beta_i \in [1, m+2]} \left\{ \sum_{i=1}^{n} \rho_{\beta_i}(f_i) : \sum_{i=1}^{n} \beta_i = s \right\}}_{\text{gap(s)}}.\end{aligned}$$

where $\bar{\rho}_g = \sup_{i \in [1,n]} \rho(g_i)$ and

$$\max\{|u_1(s)|, \|u_2(s)\|_2\} \leq \sqrt{2m}\ (R_v + R_\lambda M_V)\, \gamma$$

with

$$s = n + 1 + 2m \frac{c}{\gamma^2 + c} \quad \textit{and} \quad M_V = \sup_{\substack{\|u\|_2 \leq 1 \\ v_i \in \mathcal{F}_i}} \left\| \sum_i u_i v_i \right\|_2,$$

for some absolute constant $c > 0$.

For simplicity, we have used coarse bounds on $\rho(g_i)$ but these can be relaxed to stable quantities using techniques matching those used on the objective in the previous sections.

8.7 Sterfling-Bennett Inequalities in (2,D) smooth Banach Spaces

We prove a Sterfling-Bennett inequality in Theorem 8.17 below. This concentration inequality allows to rewrite the bound involving the quantity R in Theorem 6.2 with a term taking into account the variance of V, hence leading to an approximate Caratheodory version for high sampling ratio and low variance.

Consider $V = \{\boldsymbol{v}_1; \ldots; \boldsymbol{v}_N\}$, a set of N vectors in a $(2, D)-$Banach space with norm $|| \cdot ||$ and $V_1, \ldots, V_n$, the random variables resulting from a sampling without replacement. $R_v \triangleq \sup_i ||\boldsymbol{v}_i||$ is the *range* of V. We introduce a specific notion of variance related to that sampling scheme as follows

$$\sigma \triangleq \frac{1}{\sum_{k=1}^{m} \frac{1}{(N-k)^2}} \Big|\Big| \Big(\sum_{k=1}^{m} \frac{1}{(N-k)^2} \mathbb{E}_{k-1} ||V_k - \mathbb{E}_{k-1}(V_k)||^2 \Big)^{1/2} \Big|\Big|_\infty, \tag{10}$$

where we write $|| \cdot ||_\infty$ for essential supremum to simplify notations. We identify it as a variance because it is a convex combination of the terms $\mathbb{E}_{k-1}||V_k - \mathbb{E}_{k-1}(V_k)||^2$. For $k = 1$, it is exactly the variance of V, while when $k = N - 1$ it is not much different from the diameter of the set V. This is the natural notion algebraically arising from the sampling without replacement. Nevertheless, one can notice that when the index k increases the weights also do, thus putting more emphasis on diameter-like measures rather than on variance-like measures.

Our goal is to upper bound, with a function depending on both σ^2 and R_v, the following probability

$$\mathbb{P}\Big(\Big|\Big| \frac{1}{m} \sum_{i=1}^{m} V_i - \mu \Big|\Big| \leq \epsilon \Big). \tag{11}$$

It is called *Sterfling* because the quality of the bound will depend on the sampling ratio.

Schneider [2016] shows an Hoeffding-Sterfling bound (i.e. not depending on σ^2) on $(2, D)-$Banach spaces, while [Bardenet et al., 2015] provided a Bernstein-Sterfling bound for real-valued random variable. Here we expand the result of [Schneider, 2016] to the case of Bennet-Sterfling inequality in $(2, D)-$Banach spaces. We exploit the forward martingale [Serfling, 1974, Bardenet et al., 2015, Schneider, 2016] associated to the sampling without replacement and plug it into a sligthly modified result from [Pinelis, 1994].

For completeness of the result, we recall the definition of $(2, D)-$ Banach spaces [Schneider, 2016, Definition 3] and we refer to [Schneider, 2016, section 3] for examples of such Banach spaces.

Definition 8.14 *A Banach space $(\mathcal{B}, ||\cdot||)$ is $(2, D)-$smooth if it a Banach space and there exists $D > 0$ such that*

$$||\boldsymbol{x}+\boldsymbol{y}||^2 + ||\boldsymbol{x}-\boldsymbol{y}||^2 \leq 2||\boldsymbol{x}||^2 + 2r||\boldsymbol{y}||^2 \ , \tag{12}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{B}$.

Using Banach spaces allows to endow our space with non-Euclidean norms which can lead to important gains in measuring the variance.

8.7.1 Forward Martingale when Sampling without Replacement

Consider $(M_k)_{k\in\mathbb{N}}$ the following random process

$$M_k = \begin{cases} \frac{1}{N-k}\sum_{i=1}^{k}(V_i - \mu) & 1 \leq k \leq m \\ M_n & \text{for } k > m \ . \end{cases} \tag{13}$$

It is a standard result that $(M_k)_{k\in\mathbb{N}}$ defines a forward martingale [Serfling, 1974, Bardenet et al., 2015, Schneider, 2016] w.r.t. the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$ defined as:

$$\mathcal{F}_k = \begin{cases} \sigma(V_1, \dots, V_k) & 1 \leq k \leq m \\ \sigma(V_1, \dots, V_n) & \text{for } k > m \ . \end{cases} \tag{14}$$

Importantly we also have the two following relations [Schneider, 2016, (3) and (5)]

$$M_k - M_{k-1} = \frac{V_k - \mathbb{E}_{k-1}(V_k)}{N-k} \tag{15}$$

$$||M_k - M_{k-1}|| \leq \frac{R_v}{N-k} \ . \tag{16}$$

8.7.2 Bennet for Martingales in Smooth Banach Spaces

We recall a sligthly modified version of [Pinelis, 1994, Theorem 3.4.]. This theorem is the analogous on martingales evolving on Banach spaces of Bennet concentration inequality for sums of real independent random variables.

Theorem 8.15 (Pinelis) *Suppose $(M_k)_{k\in\mathbb{N}}$ is a martingale of a $(2, D)-$smooth separable Banach space and that there exists $(a, b) \in \mathbb{R}_+^*$ such that*

$$\big|\big|\sup_k ||M_k - M_{k-1}||\big|\big|_\infty \quad \leq \quad a \tag{17}$$

$$\big|\big|\big(\sum_{j=1}^{\infty}\mathbb{E}_{j-1}||M_j - M_{j-1}||^2\big)^{1/2}\big|\big|_\infty \quad \leq \quad b/D \ , \tag{18}$$

then for all $\eta \geq 0$,

$$\mathbb{P}(\sup_k ||M_k|| \geq \eta) \leq 2\exp\big(-\frac{\eta^2}{2(b^2+\eta a/3)}\big) \ . \tag{19}$$

Proof. In the proof of [Pinelis, 1994, theorem 3.4.], we have

$$\mathbb{P}(\sup_k ||M_k|| \geq \eta) \leq 2\exp\big(-\lambda\eta + \frac{\exp(\lambda a) - 1 - \lambda a}{a^2}b^2\big) \ . \tag{20}$$

Besides, from [Sridharan, equation (16)] we have

$$\inf_{\lambda>0} \big[-\lambda\epsilon + (e^{-\lambda} - \lambda - 1)c^2 \big] \le -\frac{\epsilon^2}{2(c^2+\epsilon/3)} \ .$$

We can rewrite (20) as

$$\begin{aligned} \mathbb{P}(\sup_k ||M_k|| \ge \eta) &\le 2\exp\big(-\lambda a \frac{\eta}{a} + (\exp(\lambda a) - 1 - \lambda a)\frac{b^2}{a^2} \big) \\ &\le 2\exp\big(-\frac{\eta^2}{2(b^2 + \eta a/3)} \big) \ . \end{aligned}$$

[Pinelis, 1994] uses the exact minimization on λ which leads to a better but non standard form for the Bennet concentration inequality. ■

8.7.3 Bennet-Sterfling in Smooth Banach Spaces

The following lemma allows to identify the constants (a, b) appearing in theorem 8.15.

Lemma 8.16

$$\big|\big| \sup_k ||M_k - M_{k-1}|| \big|\big|_\infty \le \frac{R_v}{N-m} \tag{21}$$

$$\big|\big| \big(\sum_{j=1}^{\infty} \mathbb{E}_{j-1} ||M_j - M_{j-1}||^2 \big)^{1/2} \big|\big|_\infty \le \sigma \frac{\sqrt{m}}{\sqrt{(N-m-1)N}} \ , \tag{22}$$

with σ as in (10).

Proof. (21) directly follows from (16). Because of (15), we have

$$\sum_{k=1}^{\infty} \mathbb{E}_{k-1}(||M_k - M_{k-1}||^2) = \sum_{k=1}^{m} \frac{1}{(N-k)^2} \mathbb{E}_{k-1}(||V_k - \mathbb{E}_{k-1}(V_k)||^2) \ .$$

Because of (10), we have,

$$\sum_{k=1}^{\infty} \mathbb{E}_{k-1}(||M_k - M_{k-1}||^2) = \sigma^2 \sum_{k=1}^{m} \frac{1}{(N-k)^2} \ .$$

Because of Lemma 2.1. in [Serfling, 1974], we have

$$\begin{aligned} \sum_{k=1}^{m} \frac{1}{(N-k)^2} &= \sum_{k=N-m-1+1}^{N-1} \frac{1}{k^2} \\ &\le \frac{m}{N(N-m-1)} \ . \end{aligned}$$

It leads to

$$\sum_{k=1}^{\infty} \mathbb{E}_{k-1}(||M_k - M_{k-1}||^2) \le \sigma^2 \frac{m}{N(N-m-1)} \ .$$

■

Theorem 8.17 *Consider V a discrete set of N vectors in a $(2, D)-$Banach space and $(V_i)_{i=1,\dots,m}$ the random variables obtained by sampling without replacements m elements of V. For any $\epsilon > 0$,*

$$\mathbb{P}\Big(\Big|\Big| \frac{1}{m}\sum_{i=1}^{m} V_i - \mu \Big|\Big| \ge \epsilon \Big) \le 2\exp\Big(-\frac{m\epsilon^2}{2\big(2D^2 \frac{N-m}{N}\sigma^2 + \epsilon R_v/3\big)} \Big) \ , \tag{23}$$

with μ the mean of V, $R_v \triangleq \sup_{\boldsymbol{v}\in V} ||\boldsymbol{v}||$, and

$$\sigma^2 \triangleq \frac{1}{\sum_{k=1}^{m} \frac{1}{(N-k)^2}} \Big|\Big| \Big(\sum_{k=1}^{m} \frac{1}{(N-k)^2} \mathbb{E}_{k-1} \big|\big| V_k - \mathbb{E}_{k-1} V_k \big|\big|^2 \Big)^{1/2} \Big|\Big|_\infty \ . \tag{24}$$

Proof. Using Theorem 8.15 with the forward martingale (13), we have for any $\eta > 0$,

$$\begin{aligned} \mathbb{P}\Big(\frac{1}{N-m}\Big|\Big|\sum_{i=1}^{m}(V_i-\mu)\Big|\Big| \geq \eta\Big) &\leq \mathbb{P}(\sup_i ||M_i|| \geq D) \\ \mathbb{P}\Big(\Big|\Big|\frac{1}{m}\sum_{i=1}^{m}V_i-\mu\Big|\Big| \geq \frac{N-m}{m}\eta\Big) &\leq 2\exp\big(-\frac{\eta^2}{2(b^2+\eta a/3)}\big)\ . \end{aligned} \tag{25}$$

Because of lemma 8.16, $a = \frac{R_v}{N-m}$ and $b = D\sigma\frac{\sqrt{n}}{\sqrt{N(N-m-1)}}$ is a good choice and leads to

$$\begin{aligned} \mathbb{P}\Big(\Big|\Big|\frac{1}{m}\sum_{i=1}^{m}V_i-\mu\Big|\Big| \geq \frac{N-m}{m}\eta\Big) &\leq 2\exp\big(-\frac{m}{(N-m)^2}\frac{m\epsilon}{2\big(D^2\frac{m}{N(N-m-1)}\sigma^2+\frac{m}{(N-m)^2}\epsilon R_v/3\big)}\big) \\ &\leq 2\exp\big(-\frac{m\epsilon}{2\big(2D^2\frac{N-m}{N}\sigma^2+\epsilon R_v/3\big)}\big)\ , \end{aligned}$$

for any $\eta > 0$ with $\epsilon = \frac{N-m}{m}\eta$. ■

8.7.4 Approximate Caratheodory with High Sampling Ratio and Low Variance

The primary tool for proving Approximate Caratheodory is to find a lower bound on the sampling ratio sufficient for the tail of the distribution at given level ϵ_0 not to exceed a given probability δ_0. With the Bennet-Sterfling inequality, we express a lower bound in the following lemma.

Lemma 8.18 *In the setting of Theorem 8.17, for any $\delta_0 \in]0,1[$ and $\epsilon_0 > 0$, if the sampling ratio α_m satisfies*

$$\alpha_m \geq \frac{2\ln(2/\delta_0)\big[2(D\sigma)^2+\epsilon_0 R_v/3\big]/N}{\epsilon_0^2+2\ln(2/\delta_0)\big[2(D\sigma)^2\big]/N}\ , \tag{26}$$

we have

$$\mathbb{P}\Big(\Big|\Big|\frac{1}{m}\sum_{i=1}^{m}V_i-\mu\Big|\Big| \geq \epsilon_0\Big) \leq \delta_0\ . \tag{27}$$

Proof. Given $\delta_0 \in]0,1[$ and $\epsilon_0 > 0$, we are looking for a sampling ratio $\alpha_m = \frac{m}{N}$ such that

$$\mathbb{P}\Big(\Big|\Big|\frac{1}{m}\sum_{i=1}^{m}V_i-\mu\Big|\Big| \geq \epsilon_0\Big) \leq \delta_0\ . \tag{28}$$

With Bennet-Sterfling concentration inequality, it is sufficient to find α_m such that

$$\begin{aligned} 2\exp\big(-\frac{m\epsilon^2}{2\big(2D^2\frac{N-m}{N}\sigma^2+\epsilon R_v/3\big)}\big) &\leq \delta_0 \\ -\frac{N\alpha_m\epsilon^2}{2(D\sigma)^2(1-\alpha_m)+\epsilon R_v/3} &\leq 2\ln(\delta_0/2) \\ \alpha_m\epsilon^2 &\geq -\frac{2}{N}\ln(\delta_0/2)\big[2(D\sigma)^2(1-\alpha_m)+\epsilon R_v/3\big] \\ \alpha_m\big[\epsilon^2-\frac{2}{N}2(D\sigma)^2\ln(\delta_0/2)\big] &\geq -\frac{2}{N}\ln(\delta_0/2)\big[2(D\sigma)^2+\epsilon R_v/3\big] \\ \alpha_m &\geq -\frac{\frac{2}{N}\ln(\delta_0/2)\big[2(D\sigma)^2+\epsilon R_v/3\big]}{\epsilon^2-\frac{2}{N}\ln(\delta_0/2)2(D\sigma)^2}\ . \end{aligned}$$

For (27) to be true, it is sufficient that α_m satisfies the following,

$$\alpha_m \geq \frac{2\ln(2/\delta_0)\big[2(D\sigma)^2+\epsilon_0 R_v/3\big]/N}{\epsilon_0^2+2\ln(2/\delta_0)\big[2(D\sigma)^2\big]/N}\ . \tag{29}$$

which is the desired result. ■

Using the normalization of Theorem 8.8, we get

$$\alpha_m \quad \geq \quad \frac{2\ln(2/\delta_0)\big[2(D\sigma)^2 + \epsilon_0 R_v/(3N)\big]N}{\epsilon_0^2 + 2\ln(2/\delta_0)\big[2(D\sigma)^2\big]N}\,. \tag{30}$$

and the leading term is controlled by the variance.