Highlights

Error Estimation via a Refined Shapley-Folkman Lemma

Haoyu Wu, Ao Tang

- Tighter error bound compared to previous results.
- Qualitative improvement in particular cases.

Error Estimation via a Refined Shapley-Folkman Lemma

Haoyu Wu[a,*], Ao Tang[b]

[a]*Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong,*
[b]*School of Electrical and Computer Engineering, Cornell University, Ithaca, 14853, New York, U.S.*

Abstract

Based on a refined Shapley-Folkman lemma, we derive a tighter error bound of the Shapley-Folkman-Starr theorem. Our new bound reaches a qualitatively tighter improvement in some particular cases.

Keywords: Shapley-Folkman lemma, Error-estimation

1. Introduction

The Shapley-Folkman lemma (Theorem 1.1) provides a characterization of the points within the convex hull of a Minkowski sum. Additionally, it establishes an estimation bound (Theorem 1.1.2) on the Hausdorff distance between the Minkowski sum and its convex hull. The original Shapley-Folkman lemma, as stated by Starr (1969), claims the existence of approximate equilibria in an economy, assuming nonconvex preferences.

Theorem 1.1. *(Shapley-Folkman lemma) Let S_i, $1 \le i \le n$, be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. For any $z \in convS$, there exists $z_i \in convS_i$ such that $z = \sum_{i=1}^n z_i$ and $z_i \in S_i$ except for at most* $\min\{m, n\}$ *values of i.*

The lemma could be applied on the estimation of error bounds. To state the conclusion, we define the measurement of the size of set $S \subset \mathbb{R}^m$, i.e. (outer) radius $\mathrm{rad}(S)$ and diameter $D(S)$,

$$\mathrm{rad}(S) = \inf_{y \in \mathbb{R}^m} \sup_{x \in S} |x - y|$$
$$D(S) = \sup_{x,y \in S} |x - y|$$

where $|\cdot|$ denote the l^2 norm in Euclidean space.

Corollary 1.1.1. *(Shapley-Folkman-Starr theorem)With the same assumption above and let d_H denote the Hausdorff distance,*

$$d_H^2(S, convS) \le \min\{m, n\}R^2 := \min\{m, n\}R^2$$

Declarations of interest: none.

*Corresponding author

Email addresses: `hwubs@connect.ust.hk` (Haoyu Wu), `atang@ece.cornell.edu` (Ao Tang)

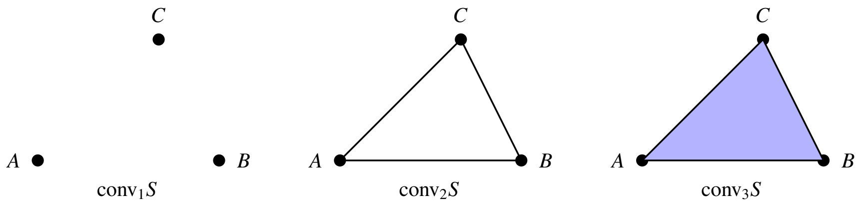

Figure 1: the kth-convex hull of set $S = \{A, B, C\}$

Where $R_i = \mathrm{rad}(S_i)$ is the outer radius of S_i. If we assume more that all S_i are compact, then the Hausdorff distance is achieveable. For any $z \in \mathrm{conv}S$, there is $x \in S$ such that

$$|x - z|^2 \leq \min\{m, n\} \max_{1 \leq i \leq n} R_i^2$$

The Shapley-Folkman lemma finds applications in various fields. Tardella (1990) utilized the Shapley-Folkman lemma to establish the Lyapunov convexity theorem in measure theory. In probability theory, the law of large numbers can be directly derived from the Shapley-Folkman lemma Artstein and Vitale (1975). Bi and Tang (2020) introduced a refined form of the Shapley-Folkman lemma (to be discussed in section 2), which distinguishes the convexity by the concept of k-extreme points and offers improvement compared to the original idea of applying the Shapley-Folkman lemma to estimate the duality gap of separable optimization problems Aubin and Ekeland (1976).

In later years, Zhou (1993) presented a novel proof of the Shapley-Folkman lemma. However, the algebraic idea behind his proof was not explicitly expressed. Interestingly, this idea can be extended to derive a stronger version of the Shapley-Folkman lemma, a specialized case of Theorem 2 (to be discussed later), without characterizing the convexity. In order to introduce this strong version, it is necessary to establish the concept of the k-th convex hull.

Definition. *Denote the kth-convex hull of a set S, as $conv_k S$, the set of convex combination of at most k elements,*

$$conv_k S = \left\{ \sum_{i=1}^{k} a_i v_i : v_i \in S, 0 \leq a_i \leq 1, 1 \leq i \leq k, \sum_{i=1}^{k} a_i = 1 \right\}$$

The kth-convex hull of a set is just all convex combinations of at most k points of the set, as a subset of the convex hull and getting larger as k increases. Figure 1 is an easy example to illustrate the concept of kth-convex hull. In the case the three points set $S = \{A, B, C\}$, $\mathrm{conv}_1 S$ is just the set S, $\mathrm{conv}_2 S$ is the segments AB, AC, BC and $\mathrm{conv}_3 S$ is the whole triangle ABC. In general, $\mathrm{conv}_1 S = S$ and the $(m+1)$-convex hull is just the original convex hull by Carathodory's theorem Rockafellar (1970) in $\mathbb{R}^m$. Based on this concept, we state the theorem,

Theorem 1.2. *Let S_i, $1 \leq i \leq n$, be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. For any $z \in convS$, there exists $k_i \in \mathbb{N}$ and $\sum_{i=1}^n k_i \leq n+m$ such that $z = \sum_{i=1}^n z_i$ and $z_i \in conv_{k_i} S_i$.*

The strong version of the Shapley-Folkman lemma incorporates an additional constraint on the components of the Minkowski sum. Building upon this constraint, Budish and Reny (2020) proposed a new improvement in the error bound estimation utilizing the Shapley-Folkman lemma, achieving a quantitatively tighter bound.

Corollary 1.2.1. *Budish and Reny (2020) With the same assumption above and additional assumption that $n > m$ and all S_i are compact sets, for any $z \in convS$, there is $x \in S$ such that*

$$|x - z|^2 \leq \frac{mD^2}{4}$$

Where $D_i = \sup_{x,y \in S_i} |x - y|$ is the diameter of S_i and $D = \max_{1 \leq i \leq n} D_i$. Combining Jung's theorem (Lemma 1), it's an improved bound of the original Shapley-Folkman-Starr theorem with up to around 30% quantitative improvement.

In our study, we employ the refined Shapley-Folkman lemma (theorem 2) to discern the convexity of points in the convex hull of the Minkowski sum, while considering the new constraint outlined in theorem 2 to attain a tighter error bound. This new bound exhibits a degree improvement in specific cases. Moreover, we consider the case when the size(radius, diameter, etc) of components of the Minkowski sum is random variables instead of a given number, and we calculate the bound of expectations.

2. The refined Shapely-Folkman Lemma

To categorize the convexity, a point in a convex set S is called k-extreme if it lies in the interior of a k-dimensional convex subset of S, but it does not belong to any $(k+1)$-dimensional convex subset of S. The concept could be formally defined as,

Definition. *A point z in a convex set S is called a k-extreme point of S if there not exist $(k+1)$ linear independent vectors $d_1, d_2, \cdots, d_{k+1}$ such that $z \pm d_i \in S$ for $i = 1, 2, \cdots, k+1$.*

In general, a point in convex set is an extreme point if only if it's 0-extreme and on boundary if only if $(m-1)$-extreme. Any point in the convex hull in $\mathbb{R}^m$ is m extreme. In the concrete example in figure 1; A, B, C the three vertexes are 0-extreme and the points on segments AB, AC, BC without end points are 1-extreme, the remaining points are 2-extreme. Based on the concept of the k-extreme point and previous concept of kth convex hull, Bi and Tang (2020) proposed a refined Shapley-Folkman lemma.

Theorem 2.1. *Let S_i, $1 \leq i \leq n$, be subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n$ be the Minkowski sum of all S_i. Assume z is a k-extreme point of $convS$, then there exist integers $1 \leq k_i \leq k+1$ with $\sum_{i=1}^n k_i \leq n+k$ and points $z_i \in conv_{k_i} S_i$ such that $z = \sum_{i=1}^n z_i$.*

Let k be m and recall all points in the convex hull is m-extreme, Corollary 1.2.1 become a special case of Theorem 2.1. Based on the idea categorizing all z_i into k_i, we can get an improvement by considering all z_i jointly compared to the original Shapley-Folkman-Starr theorem 1.1.1.

3. Tighter Error Bound

Applying the refined Shapley-Folkman lemma, we derive our main result as a new estimated error bound.

Theorem 3.1. *Let S_i, $1 \leq i \leq n$, be compact subsets of $\mathbb{R}^m$, $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. Assume z is a k-extreme point of $convS$, then there exist $x \in S$ such that,*

$$|z - x|^2 \leq \frac{1}{2} \min \left\{ \frac{nkD^2}{n+k}, \sum_{i=1}^n D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k} \right\}$$

Where $D_i = \sup_{x,y\in S_i} |x-y|$ is the diameter of S_i and $D = \max_{1\le i\le n} D_i$. The first term has degree $O(\min\{n,k\}D^2) = O(\min\{n,k\}R^2)$ by Jung's inequality (Lemma 1), which is the same as the degree of the bound of the original Shapley-Folkman-Starr theorem 1.1.2 but we category the convexity by k instead of the dimension m. The latter bound is attained by jointly considering all D_i and constraints on k_i and will be an qualitative improvement compared to the first term in some particular cases. Specifically, the first term is solely concerned with the maximum value among all D_i (also referred to as the l_∞ norm), while the latter term emphasizes the summation of both D_i and D_i^2 (corresponding to the l_1 and l_2 norms, respectively).

Conditions \ Results	Corollary 1.1.1	Corollary 1.2.1	Theorem 3.1 and Corollary 3.1.1
In general	$\min\{n,m\}R^2$	N.A.	$\frac{1}{2}\min\left\{\frac{nmD^2}{n+m}, \sum_{i=1}^n D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k}\right\}$
$n \ge m$	mR^2	$mD^2/4$	$\min\left\{mD^2/4, \frac{1}{2}\left[\sum_{i=1}^n D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k}\right]\right\}$
$n \ge m$ & k-extreme	mR^2	$mD^2/4$	$\min\left\{kD^2/4, \frac{1}{2}\left[\sum_{i=1}^n D_i^2 - \frac{\left(\sum_{i=1}^n D_i\right)^2}{n+k}\right]\right\}$
m is $O(n)$	$O(nD^2)$	N.A.	$\min\left\{O(kD^2), O\left(\sum_{i=1}^n D_i^2\right)\right\}$

Table 1: Comparison of different results

Corollary 3.1.1. *If $n \ge k$, the estimated bound $\frac{nkD^2}{2(n+k)}$ in theorem 3.1 can be modified to $\frac{kD^2}{4}$.*

Since k is always smaller than m, the dimension of given Euclidean space, and combining $\frac{D}{2} \le R$, the original Shapley-Folkman-Starr theorem 1.1.1 is a direct corollary of our theorem. Furthermore, Corollary 3.1 asserts that the new bound is at least as effective as corollary 1.2.1 in all cases, categorizing the convexity.

Corollary 3.1.2. *In general, we are unable to determine whether the first term is greater than or less than the latter term. But the latter term is not larger than the first term in degree in the particular case $k \in O(n)$.*

Without information of sizes of components, it is difficult to compare the two terms. Therefore, we naturally consider that the sizes follow independent and identically distributed (i.i.d.).

Corollary 3.1.3. *Consider the case that D_i's are i.i.d. and we estimate the bound of expectation. The first term is $\frac{nk}{2(n+k)}\mathbb{E}(D^2)$. The latter term is*

$$\frac{nk}{n+k}\mathbb{E}(D_i^2) + \frac{n(n-1)Var(D_i)}{n+k}$$

In case $k \in O(n)$ or $n\mathrm{Var}(D) \in O(k\mathbb{E}(D_i^2))$, the latter term has degree improvement $\frac{\mathbb{E}(D_i^2)}{\mathbb{E}(D^2)} = \frac{\mathbb{E}(D_i^2)}{\mathbb{E}(\max_{1\le i\le n} D_i^2)}$.

4. Example

Considering the structure of the two terms, a disparity emerges between the expected maximum of these i.i.d. and their average or average squared values as n increases. An enhancement in the latter term arises when all D_i follow a certain distribution that allows for positive infinity but with finite mean and squared mean. It's naturally to think of half-normal distribution and exponential distribution, and we also consider the outlier case, among which the latter term will have a qualitative improvement compared to the first term (with the same degree of the existing bound).

4.1. Outlier Case

In case that $D_1, D_2, \cdots, D_{n-1}$ are bounded above by a but $D_n = b$ such that $b >> a$. The first term will be $\frac{nk}{2(n+k)}b^2 \in O(\frac{nkb^2}{n+k})$. The latter term is, by our method,

$$
\begin{aligned}
\sum_{i=1}^{n} \frac{k_i - 1}{2k_i} D_i^2, k_i \in \mathbb{N}, \sum_{i=1}^{n} k_i \leq n + k &\leq \sum_{i=1}^{n-1} \frac{k_i - 1}{2k_i} a^2 + \frac{k_n - 1}{2k_n} b^2, k_i \in \mathbb{N}, \sum_{i=1}^{n} k_i \leq n + k \\
&\leq \frac{1}{2}\left[b^2 + (n-1)a^2 - \frac{((n-1)a + b)^2}{n+k} \right] \\
&\leq \frac{1}{2}\left[\frac{n+k-1}{n+k} b^2 + \frac{(n-1)(k+1)a^2 - 2(n-1)ab}{n+k} \right] \\
&\in O\left(b^2 + \frac{nk}{n+k} a^2 \right)
\end{aligned}
$$

The second last inequality is from the Cauchy-Schwarz inequality with a degree improvement at least $\min\{b^2/a^2, nk/(n+k)\}$.

4.2. Half-normal Distribution

Let X follow a normal distribution $N(0, \sigma^2)$ and $Y = |X|$ follows a half-normal distribution,

$$
f_Y(y, \sigma) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \exp\left(-\frac{y^2}{2\sigma^2} \right)
$$

We can calculate that,

$$
\mathbb{E}(Y) = \sigma\sqrt{\frac{2}{\pi}}, \mathbb{E}(Y^2) = \sigma^2,\ \mathbb{E}(\max_{i \in [n]} Y_i) = \sigma\sqrt{2\log n} + O(1)
$$

Consider that all D_i's are i.i.d. of the random variable Y, $D = \max_{1 \leq i \leq n} D_i$. The estimated bound of expectations,

$$
\mathbb{E}\left(\frac{nk}{n+k} D^2 \right) \geq \frac{nk}{n+k} \mathbb{E}^2(D) \in O(\frac{nk\sigma^2}{n+k} \log n)
$$

$$
\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n+k}\right) &= \sum_{i=1}^{n} \mathbb{E}(D_i^2) - \frac{\mathbb{E}(\sum_{i=1}^{n} D_i^2) + \mathbb{E}(\sum_{1\le i,j\le n, i\neq j} D_i D_j)}{n+k} \\
&= \sigma^2 \left(n - \frac{n + \frac{2}{\pi}n^2 - \frac{2}{\pi}n}{n+k}\right) \\
&= \sigma^2 \frac{\left(1-\frac{2}{\pi}\right)(n^2-n) + nk}{n+k} \\
&\le \sigma^2 \frac{n^2 - n + nk}{n+k} \in O\left(n\sigma^2\right)
\end{aligned}
$$

With more assumption that k is $O(n)$, the first term is least $O(n \log n\, \sigma^2)$ and the latter term has qualitative improvement $\frac{1}{\log n}$ compared to the first term.

4.3. Exponential Distribution

In case k is $O(n)$ and D_i's are i.i.d. following distribution exp(1), i.e. the probability density function $f(x) = e^{-x}$ for $x \ge 0$ and equal 0 else.

$$
\mathbb{E}(D_i) = \mathrm{Var}(D_i) = 1, \mathbb{E}(D_i^2) = \mathbb{E}^2(D_i) + \mathrm{Var}(D_i) = 2
$$

$$
\mathbb{E}(D) = \sum_{i=1}^{n} \frac{1}{i}, \mathrm{Var}(D) = \sum_{i=1}^{n} \frac{1}{i^2} \Longrightarrow \mathbb{E}(D^2) = \left(\sum_{i=1}^{n} \frac{1}{i}\right)^2 + \sum_{i=1}^{n} \frac{1}{i^2} \in O(\log^2 n)
$$

.

By Corollary 3.1.3, the latter bound has an qualitative improvement $\frac{1}{\log^2 n}$ compared to the first bound.

5. Conclusion

We distinguish the convexity in the set $S = \sum_{i=1}^{n} S_i$ by the concept of k-extreme and apply the refined Shapley-Folkman lemma from Bi and Tang (2020), categorizing the n components into k_ith-convex hull. Instead of the traditional way in which each D_i is directly estimated to be $\max_{1\le i\le n} D_i$, we consider all D_i jointly and use the constraint on k_i to reach another bound (that is, the latter term). The new bound will have a qualitative improvement in some particular cases, especially the case $k \in O(n)$.

5.1. Prospect

Consider the same assumption that $S = \sum_{i=1}^{n} S_i \subset \mathbb{R}^m$. For any point z in the convex hull conv$S$, we can see that there is a vector $\mathbf{k} = (k_1, \cdots, k_n)^T \in (\mathbb{N} \cup \{0\})^n$, $\|\mathbf{k}\|_1 \le n + k$ such that $z \in \sum_{i=1}^{n} \mathrm{conv}_{k_i} S$, and we denote $A(\mathbf{k}) = \sum_{i=1}^{n} \mathrm{conv}_{k_i} S \subset \mathrm{conv} S$. This implies that any point within the convex hull will belong to at least one of the sets $A(\mathbf{k})$. Naturally, we are interested in exploring the relationships (e.g., Hausdorff distance, size of overlapping, etc.) among all these sets $A(\mathbf{k})$ and the original S and conv$S$. It's fair to mention that Adiprasito (2020) shows that a related result of measurement about $d_H(\mathrm{conv}S, \mathrm{conv}_k S)$.

Appendix A. Proof

Lemma 1. *For any set $S \subset \mathbb{R}^m$,*

$$\sqrt{\frac{2(m+1)}{m}} R \leq D \leq \frac{R}{2}$$

Proof. The first inequality is Jung's Theorem Jung (1901) and the last inequality is trivial in Euclidean space.

Lemma 2. *Consider the set $S = \{s_0, s_1, ..., s_{k-1}\} \subset \mathbb{R}^m$, the cardinality of S, denote as $|S|$, $|S| = k$, then $rad^2(S) \leq \frac{k-1}{2k} D^2(S)$, where $D(S) = \sup_{x,y \in S} |x - y|$, is the diameter of S.*

Proof. For any convex combination z of S,

$$\begin{aligned}
z &= \lambda_0 s_0 + \lambda_1 s_1 + \cdots \lambda_{k-1} s_{k-1} \\
&= (1 - (\lambda_1 + \lambda_2 + \cdots + \lambda_{k-1})\, s_0 + \lambda_1 s_1 + \cdots \lambda_{k-1} s_{k-1} \\
&= s_0 + \lambda_1(s_1 - s_0) + \lambda_2(s_2 - s_0) \cdots \lambda_{k-1}(s_{k-1} - s_0) \\
&\in \{s_0\} \oplus \operatorname{span}\{s_1 - s_0, s_2 - s_0, \cdots s_{k-1} - s_0\}
\end{aligned}$$

Which implies conv$S$ is contained in a vector space that is isomorphic to $\mathbb{R}^{k-1}$. By Lemma 1, $\operatorname{rad}^2(S) \leq \frac{k-1}{2k} D^2(S)$.

Lemma 3. *(Cauchy-Schwarz Inequality) Given $a_1, a_2, \cdots, a_n \in \mathbb{R}^+$, $b_1, b_2, \cdots, b_n \in \mathbb{R}^+$,*

$$\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2 \geq \left(\sum_{i=1}^{n} a_i b_i \right)^2$$

Proof. It's a well-known result and the proof is omitted.

Proof of Theorem 1.2. Recall that conv$S$ = conv $\sum_{i=1}^n S_i = \sum_{i=1}^n$ convS_i. For any $z \in$ conv$S$, there exists $z_i \in$ conv$S_i, m_i \in \mathbb{N}, x_{ij} \in S_i, 1 \leq i \leq n$ such that

$$\begin{aligned}
&z = z_1 + z_2 + \cdots z_n \in \mathbb{R}^m \\
&z_i = \sum_{j=1}^{m_i} a_{ij} x_{ij}, \quad 1 \leq i \leq n, \sum_{j=1}^{m_i} a_{ij} = 1, a_{ij} \geq 0, 1 \leq i \leq n
\end{aligned}$$

Let $e_j, 1 \leq j \leq n$, denote the standard basis of $\mathbb{R}^n$,

$$\begin{aligned}
y &:= z \oplus (e_1 + e_2 + \cdots + e_n)^T \in \mathbb{R}^{m+n} \\
y &= \sum_{i=1}^{n} z_i \oplus e_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m_i} a_{ij} x_{ij} \right) \oplus e_i \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m_i} a_{ij} \left(x_{ij} \oplus e_i \right)
\end{aligned}$$

Notice that $y \in \mathbb{R}^{m+n}$, by Caratheodory theorem Rockafellar (1970) for conic combination, there exist non-negative b_{ij} with at most $m + n$ of b_{ij}'s strictly greater than 0,

$$y=\sum_{i=1}^{n}\sum_{j=1}^{m_i} b_{ij}(x_{ij}\oplus e_i)$$

By the construction of last n entries of y, we have $\sum_{j=1}^{m_i} b_{ij} = 1$ for $1 \le i \le n$. Let $k_i = \|(b_{i1}, b_{i2}, \cdots, b_{im_i})^T\|_0$ be the value of the count measure in each i, and reorder these positive term,

$$z=\sum_{i=1}^{n} x_i, \quad x_i=\sum_{j=1}^{k_i} b_{ij}x_{ij}\in \mathrm{conv}_{k_i}S_i, \quad \sum_{i=1}^{n} k_i \le m+n$$

Proof of Theorem 3.1. For any k-extreme point z, by Theorem 2.1, there exists $\{z_i\}_{i=1}^n$ such that

$$z=\sum_{i=1}^{n} z_i, z_i\in \mathrm{conv}_{k_i}S_i, k_i\in\mathbb{N}, \sum_{i=1}^{n} k_i \le n+k$$

Denote $T_i \subset S_i$ such that $z_i \in \mathrm{conv}T_i$, $|T_i| = k_i$, i.e. T_i consists of the k_i elements such that z_i is in their convex combination. Let $T = \sum_{i=1}^n T_i$ and we have $z \in \mathrm{conv}T$. Denote D_i as the diameter of S_i, that is, $D_i = \sup_{x,y\in S_i} |x-y|$. We apply the original Shapley-Folkman-Starr theorem to T and $\mathrm{conv}T$,

$$\begin{aligned}\inf_{w\in T}|z-w|^2 \le \sum_{i=1}^{n}\mathrm{rad}^2(T_i) &\le \sum_{i=1}^{n}\frac{k_i-1}{2k_i}\mathrm{D}(T_i), k_i\in\mathbb{N}, \sum_{i=1}^{n}k_i\le n+k\\ &\le \sum_{i=1}^{n}\frac{k_i-1}{2k_i}D_i^2, k_i\in\mathbb{N}, \sum_{i=1}^{n}k_i\le n+k\\ &\le \sum_{i=1}^{n}\frac{k_i-1}{2k_i}D_i^2, k_i>0, \sum_{i=1}^{n}k_i\le n+k\end{aligned}$$

The third last inequality is from the lemma 2. $T_i \subset S_i$ implies $T \subset S$,

$$\inf_{y\in S}|z-y|^2 \le \inf_{w\in T}|z-w|^2 \le \sum_{i=1}^{n}\frac{k_i-1}{2k_i}D_i^2, k_i>0, \sum_{i=1}^{n}k_i\le n+k$$

Then we will prove the two terms of our result separately.
Claim $\sum_{i=1}^n \frac{k_i-1}{k_i}D_i^2 \le \frac{nk}{n+k}D^2$:

$$\begin{aligned}\sum_{i=1}^{n}\frac{k_i-1}{k_i}D_i^2 \le \sum_{i=1}^{n}\frac{k_i-1}{k_i}D^2 &= nD^2 - D^2\sum_{i=1}^{n}\frac{1}{k_i}\\ &\le D^2\left(n-\frac{(\sum_{i=1}^n 1)^2}{\sum_{i=1}^n k_i}\right)\le \frac{nk}{n+k}D^2\end{aligned}$$

The second last inequality is from the Cauchy-Schwarz inequality.
Claim $\sum_{i=1}^n \frac{k_i-1}{k_i}D_i^2 \le \sum_{i=1}^n D_i^2 - \frac{(\sum_{i=1}^n D_i)^2}{n+k}$:

$$
\begin{aligned}
\sum_{i=1}^{n} \frac{k_i - 1}{k_i} D_i^2 &= \sum_{i=1}^{n} D_i^2 - \sum_{i=1}^{k} \frac{D_i^2}{k_i} \\
&\leq \sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{\sum_{i=1}^{n} k_i} \\
&\leq \sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n+k}
\end{aligned}
$$

The last inequality is from the Cauchy-Schwarz inequality. Since all S_i are compact, the existene of x is guaranteed.

Proof of Corollary 3.1.1. It's enough to show that the following optimal program with optimal value $\frac{k}{2}$, given $n > k$,

$$
\begin{aligned}
&\max \sum_{i=1}^{n} \frac{k_i - 1}{k_i} \\
&s.t.\, k_i \in \mathbb{N}, \sum_{i=1}^{n} k_i \leq n + k
\end{aligned}
$$

Noticed that $f(x) = 1 - 1/x$ is concave and increasing on $[1, \infty)$ and there are at least k of k_i strictly greater than 1. By the concavity and constraints, the optimal solution is in the case that j of $k_i's$ are equal a and $n - j$ of $k_i's$ equal $a + 1$ for some positive integer a and the sum of k_i is tightly equal $n + k$.

$$
ja + (n - j)(a + 1) = k \iff an - j = k
$$

Notice that $n > k$ and $j < n$, which only allows $a = 1$, hence $j = n - k$. This is the necessary condition to reach maximum value, and $a = 1, j = n - k$ is the only candidate, which claims that the optimal value is $(n - k)\frac{1-1}{1} + k\frac{2-1}{2} = \frac{k}{2}$.

Proof of Corollary 3.1.2.

$$
\begin{aligned}
&\frac{nk}{n+k} D^2 \leq \sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n+k} \\
\iff &nD^2 - \sum_{i=1}^{n} D_i^2 \leq \frac{n^2 D^2 - (\sum_{i=1}^{n} D_i)^2}{n+k} \\
\iff &\sum_{i=1}^{n} (D - D_i)(D + D_i) \leq \frac{\sum_{i=1}^{n} (D + D_i) \sum_{i=1}^{n} (D - D_i)}{n+k}
\end{aligned}
$$

It's only true that for large k. The left-hand side of the last inequality is greater than the right-hand side when $k = 0$, as Chebyshev's inequality.

In case $k \in O(n)$, the first term is in $O(nD^2)$, and the latter term is less than $\sum_{i=1}^{n} D_i^2$ is not greater than nD^2 in degree.

Proof of Corollary 3.1.3. It's enough that,

$$\mathbb{E}\left(\sum_{i=1}^{n} D_i^2 - \frac{\left(\sum_{i=1}^{n} D_i\right)^2}{n+k}\right) = n\mathbb{E}(D_i^2) - \frac{\mathbb{E}\left[(\sum_{i=1}^{n} D_i)^2\right]}{n+k}$$
$$= n\mathbb{E}(D_i^2) - \frac{n\mathbb{E}(D_i^2) + n(n-1)\mathbb{E}^2(D_i)}{n+k}$$
$$= \frac{nk}{n+k}\mathbb{E}(D_i^2) + \frac{n^2\mathbb{E}(D_i^2) - n\mathbb{E}(D_i^2) - n^2\mathbb{E}^2(D_i) + n\mathbb{E}^2(D_i)}{n+k}$$
$$= \frac{nk}{n+k}\mathbb{E}(D_i^2) + \frac{n(n-1)\mathrm{Var}(D_i)}{n+k}$$

References

Adiprasito, K., B.I.M.N.e.a., 2020. Theorems of carathéodory, helly, and tverberg without dimension. Discrete Comput Geo 64, 233–258. URL: https://doi.org/10.1007/s00454-020-00172-5.

Artstein, Z., Vitale, R.A., 1975. A Strong Law of Large Numbers for Random Compact Sets. The Annals of Probability 3, 879 – 882. URL: https://doi.org/10.1214/aop/1176996275, doi:10.1214/aop/1176996275.

Aubin, J.P., Ekeland, I., 1976. Estimates of the duality gap in nonconvex optimization. Mathematics of Operations Research 1, 225–245. URL: http://www.jstor.org/stable/3689565.

Bi, Y., Tang, A., 2020. Duality gap estimation via a refined shapley–folkman lemma. SIAM Journal on Optimization 30, 1094–1118. URL: https://doi.org/10.1137/18M1174805, doi:10.1137/18M1174805.

Budish, E., Reny, P.J., 2020. An improved bound for the shapley–folkman theorem. Journal of Mathematical Economics 89, 48–52. URL: https://www.sciencedirect.com/science/article/pii/S030440682030046X, doi:https://doi.org/10.1016/j.jmateco.2020.04.003.

Jung, H., 1901. Ueber die kleinste kugel, die eine räumliche figur einschliesst. Journal für die reine und angewandte Mathematik 123, 241–257. URL: http://eudml.org/doc/149122.

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

Starr, R.M., 1969. Quasi-equilibria in markets with non-convex preferences. Econometrica 37, 25–38. URL: http://www.jstor.org/stable/1909201.

Tardella, F., 1990. A new proof of the lyapunov convexity theorem. SIAM Journal on Control and Optimization 28, 478–481. URL: https://doi.org/10.1137/0328026, doi:10.1137/0328026, arXiv:https://doi.org/10.1137/0328026.

Zhou, L., 1993. A Simple Proof of the Shapley-Folkman Theorem. Economic Theory 3, 371–372. URL: https://ideas.repec.org/a/spr/joecth/v3y1993i2p371-72.html.