Highlights

An Even Tighter Bound for Shapley-Folkman-Starr Theorem

Haoyu Wu, Ao Tang

- Tighter error bound compared to previous results.
- Application in the course allocation problem.

An Even Tighter Bound for Shapley-Folkman-Starr Theorem

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

[a]Department of Mathematics, Hong Kong University of Science and Technology, 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 and apply the result to the course allocation problem.

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. The Shapley-Folkman lemma, as stated by Starr (1969), claims the existence of approximate equilibrium in an economy, assuming nonconvex preferences. For a set S, let conv$S$ define its convex hull.

Theorem 1.1. *(Shapley-Folkman lemma) Let S_i be subsets of $\mathbb{R}^m$, $1 \leq i \leq n$, and $S = \sum_{i=1}^n S_i$ 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.*

Zhou (1993) presented a novel proof of the Shapley-Folkman lemma, but the algebraic idea behind his proof was not explicitly expressed. This idea can be extended to derive a modification of the Shapley-Folkman lemma. To introduce this modification, it is necessary to establish the concept of the kth convex hull.

Definition. *Denote the kth convex hull of a set S, as $conv_k S$, the set of convex combinations 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 all convex combinations of at most k points of the set, as a subset of the convex hull, and gets larger as k increases. Figure 1 is an easy example to illustrate the concept of the 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 $(m+1)$th convex hull is just the whole convex hull by Carathéodory's theorem (Rockafellar (1970)) in $\mathbb{R}^m$. Based on this concept, we state the modified theorem,

Declarations of interest: none.

*Corresponding author

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

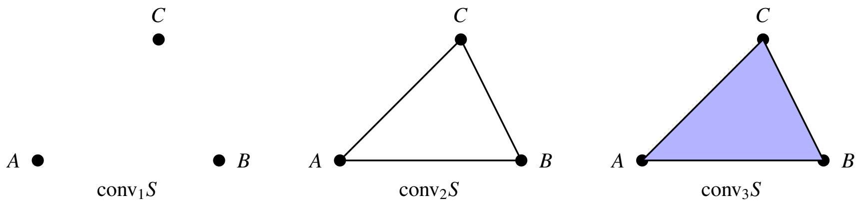

Figure 1: the kth convex hull of set $S = \{A, B, C\}$.

Theorem 1.2. *(Zhou (1993)[1]) Let S_i be subsets of $\mathbb{R}^m$, $1 \le i \le n$, and $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. For any $z \in convS$, there exists $k_i \in \mathbb{N}$ and $\sum_{i=1}^n k_i \le n + m$ such that $z = \sum_{i=1}^n z_i$ and $z_i \in conv_{k_i} S_i$.*

The theorem indicates a constraint on the k_i's. By examining the convexity via the concept of k-extreme, Bi and Tang (2020) generalized the theorem. A point in a convex set S is called k-extreme if it does not lie in the interior of any $(k+1)$-dimensional convex subset of S. The concept is formally defined as,

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

Generally, a point in a convex set is an extreme point if and only if it is 0-extreme and on the boundary if and only if $(m-1)$-extreme. Any point in the convex hull in $\mathbb{R}^m$ is m-extreme. In the concrete example $S = \{A, B, C\}$ in figure 1: the three vertexes (A, B, C) are 0-extreme, and the points on segments AB, AC, BC including endpoints are 1-extreme, and any point in the triangle is 2-extreme. Based on the concept of the k-extreme point and the previous concept of the kth convex hull, Bi and Tang (2020) proposed a refined Shapley-Folkman lemma:

Theorem 1.3. *Let S_i be subsets of $\mathbb{R}^m$, $1 \le i \le n$, and $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i. Assume z is a k-extreme point of $convS$, then there exist integers $1 \le k_i \le k+1$ with $\sum_{i=1}^n k_i \le n + k$ 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 any point in the convex hull is m-extreme, then Theorem 1.2 becomes a special case of the Theorem 1.3.

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) employed his refinement Theorem 1.3 for the estimation of the duality gap on the nonconvex optimization problem with separable objective functions, which offered improvement compared to the original idea of applying the Shapley-Folkman lemma from Aubin and Ekeland (1976).

Starr applied the Shapely-Folkman lemma to derive a corollary to measure the Hausdorff distance between a Minkowski sum of sets and its convex hull to claim the existence of approximate equilibrium in a nonconvex economy. To state the result, we define the measurement of the size of a set $S \subset \mathbb{R}^m$, i.e. (outer) radius $\text{rad}(S)$ and diameter $D(S)$:

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

[1]The theorem is summarized from but not directly stated in Zhou's proof of Shapley-Folkman lemma.

$$D(S) = \sup_{x,y\in S} |x-y|,$$

where $|\cdot|$ denotes the l^2 norm in Euclidean space throughout this paper. The Hausdorff distance with respect to usual Euclidean space d_H for $X, Y \subset \mathbb{R}^m$ is defined as

$$d_H(A,B) = \max\{\sup_{x\in X}\inf_{y\in Y} |x-y|, \sup_{y\in Y}\inf_{x\in X} |x-y|\}.$$

Corollary 1.4. *(Shapley-Folkman-Starr theorem) Let S_i be subsets of $\mathbb{R}^m$, $1 \le i \le n$, and $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i, then*

$$d_H^2(S, convS) \le \min\{m,n\}R^2 := \min\{m,n\} \max_{1\le i\le n} rad^2(S_i).$$

The Hausdorff distance is achievable if we assume more that all S_i are compact, i.e., for any $z \in \text{conv}S$, there is $x \in S$ such that

$$|x-z|^2 \le \min\{m,n\}R^2.$$

Theorem 1.2 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 Theorem 1.2.

Corollary 1.5. *(Budish and Reny (2020)) Let S_i be subsets of $\mathbb{R}^m$, $1 \le i \le n$, and $S = \sum_{i=1}^n S_i$ be the Minkowski sum of all S_i, and additional assumption that $n > m$,*

$$d_H^2(S, convS) \le \frac{m}{4} \max_{1\le i\le n} D^2(S_i),$$

if all S_i are compact sets, for any $z \in convS$, there is $x \in S$ such that

$$|x-z|^2 \le \frac{m}{4} \max_{1\le i\le n} D^2(S_i).$$

Combining Jung's theorem (lemma 1), it is an improved bound of the Shapley-Folkman-Starr Corollary 1.4 with up to around 30% quantitative improvement[2].

In our study, we employ Bi and Tang's refined Shapley-Folkman lemma (Theorem 1.3) to discern the convexity of points in a convex hull of the Minkowski sum while considering the new constraint outlined in Theorem 1.3 to attain a tighter error bound. We compare our result to the existing results and apply our result to the course allocation problem.

2. A Tighter Error Bound

Applying Theorem 1.3, we derive our main result as a new error bound.

Theorem 2.1. *Let S_i, $1 \le i \le n$, 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,*

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

[2] See section 3.1 in Budish and Reny (2020).

where $D = \max_{1 \le i \le n} D(S_i)$. 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 Shapley-Folkman-Starr Corollary 1.4 but we category the convexity by k to replace the dimension m in error upper bound. The latter bound is attained by jointly considering all $D(S_i)$ and constraints on k_i and will be a (qualitative) improvement compared to the first term in some particular cases. Specifically, the first term is solely concerned with the maximum value among all $D(S_i)$ (also referred to as the l_∞ norm), while the latter term emphasizes the summation of both $D(S_i)$ and $D^2(S_i)$ (corresponding to the l_1 and l_2 norms, respectively).

Conditions \ Results	Starr's original result	Budish and Reny (2020)	Our results
In general	$\min\{n, m\}R^2$	N.A.	$\frac{1}{2}\min\left\{\frac{nmD^2}{n+m}, \sum_{i=1}^n D(S_i)^2 - \frac{\left(\sum_{i=1}^n D(S_i)\right)^2}{n+k}\right\}$
$n \ge m$	mR^2	$mD^2/4$	$\min\left\{mD^2/4, \frac{1}{2}\left[\sum_{i=1}^n D(S_i)^2 - \frac{\left(\sum_{i=1}^n D(S_i)\right)^2}{n+k}\right]\right\}$
$n \ge m$ & k-extreme	mR^2	$mD^2/4$	$\min\left\{kD^2/4, \frac{1}{2}\left[\sum_{i=1}^n D(S_i)^2 - \frac{\left(\sum_{i=1}^n D(S_i)\right)^2}{n+k}\right]\right\}$
m is $O(n)$	$O(nD^2)$	N.A.	$\min\left\{O(kD^2), O\left(\sum_{i=1}^n D(S_i)^2\right)\right\}$

Table 1: Comparison of different results

Corollary 2.2. *If $n \ge k$, the estimated bound $\frac{nkD^2}{2(n+k)}$ in Theorem 2.1 can be modified to $\frac{kD^2}{4}$.*

Since k is always less than m, the dimension of given Euclidean space, and combining $\frac{D}{2} \le R$, the Shapley-Folkman-Starr Corollary 1.4 is a direct corollary of our theorem. Furthermore, Corollary 2.2 asserts that the new bound is at least as effective as Corollary 1.5 in all cases, categorizing the convexity.

Corollary 2.3. *In general, we are unable to determine whether the first term is greater than or less than the latter term. However, the latter term is not larger than the first term in degree in the particular case that k is $O(n)$.*

It is difficult to compare the two terms without information on the sizes of components. Therefore, we naturally consider the sizes to be independent and identically distributed random variables.

Corollary 2.4. *Assume that $D(S_i)$'s are independent and identically distributed random variables, and we calculate expectation. The first term is $\frac{nk}{2(n+k)}\mathbb{E}(D^2)$ and the latter term is*

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

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

3. Example

We introduce an application related to the problem of allocating courses to students. We assume that there are m courses in total and n students, each of whom may enroll in no more than

T courses subject to credit constraints. Consequently, feasible bundles are represented as vectors in the form of $\{0,1\}^m$, where at most T entries are 1 and the remaining entries are 0.

Budish (2011) stated the existence of an allocation to students that is approximately competitive but may lack feasibility, with a bound on excess demands. Expanding upon the analysis, Budish and Reny (2020) provided an enhancement to this excess demand bound via his improved Shapley-Folkman-Starr theorem bound (Corollary 1.5). We contribute further by incorporating the concept of k-extremity into the course allocation model and employ our findings to approximate a refined bound. The primary objective of this section is to compare our novel bounds with the results delineated by Budish and Reny (2020) (Section 4).

In order to elucidate the concept of k-extremity within our course allocation model, we introduce the concept of dominant courses. These special courses, such as individual project-based courses, have only one quota each. Consequently, each student is restricted to enrolling in at most one dominant course, reflecting the realistic constraint that a college student is not allowed to work on more than one project simultaneously. It is also assumed that all students have a preference for bundles that include a dominant course over those that do not, and all students equivalently prefer all the dominant courses.

Under these premises, we posit that a competitive equilibrium exists within an $(m-l)$-dimensional convex subset of the total course allocation space (see Appendix B). As a result of these considerations, we offer a tighter bound on the excess demand with these extra assumptions. To illustrate, let's consider an example with $m = 100$ courses, of which $l = 30$ are dominant courses, and $n = 500$ students. Recognizing that not all students are in a position to take exactly T courses—owing to commitments such as part-time work or internships—we introduce a distribution for the number of courses students enroll in. Specifically, we assume a ratio of $1:2:4:2:1$ for the number of courses students will take, ranging from 1 to 5. This means that 50 students are enrolled in 1 course, 100 students are taking 2 courses, and so on.

In assessing the feasibility of course bundles for a student i, who intends to enroll in t_i courses, we define the diameter, D_i, as the greatest distance between any two feasible bundles, calculated to be $\sqrt{2t_i}$. Based on Corollary 1.4, Budish and Reny (2020) (Section 4) presented an upper bound of the total excess demand as $\sqrt{2Tm}/2$. For the scenario previously outlined, this bound translates to $\sqrt{1000}/2 = 15.8$, which yields a maximized average excess demand per course of 1.5 students[3]. Our results, Theorem 2.1 and Corollary 2.2, posit a more refined upper bound of total excess demands as,

$$\min\left\{\sqrt{2T(m-l)}/2,\ \sqrt{\sum_{i=1}^{n} D_i^2 - (\sum_{i=1}^{n} D_i)^2/(n+m-l)}/\sqrt{2}\right\}$$
$$= \min\{\sqrt{700}/2,\ \sqrt{(3000 - 1201^2/570)/2}\} = \min\{13.2, 15.3\} = 13.2,$$

in the assumed case. Consequently, the maximized average excess demand per course is reduced to 1.25 students[4]. When the number of students n is reduced to 150, the upper bounds presented by Budish and Reny (2020) and our first bound remain unchanged. However, our latter bound

[3]This scenario is realized when 50 courses have an excess demand of one student while the remaining 51 courses each have an excess demand of two students.

[4]This outcome occurs when 75 courses have an excess demand of one student, and the remaining 25 courses each have an excess demand of two students.

adjusts to $\sqrt{(900-360^2/220)/2}=12.5$, which further lowers the maximized average excess per course to 1.18 students[5].

In conclusion, our main theorem introduces a tighter bound in the context of course allocation problem when the concept of "dominant courses" is considered. Our first bound demonstrates an enhancement compared to the findings of Budish and Reny (2020), with the consideration of "dominant courses," improving both the maximum demand for any single course (from 15.8 to 13.2) and the maximized average excess per course (from 1.51 to 1.25), irrespective of the student population size. Furthermore, when the number of students and the distribution of the number of courses they are enrolled in are factored in, our latter bound potentially provides a superior refinement in both the maximum demand for any single course (from 13.2 to 12.5) and the maximized average excess per course (from 1.25 to 1.18), particularly when the number of students is less or when the distribution is positively skewed[6].

4. Conclusion

We distinguish the convexity in the convex hull of set $S=\sum_{i=1}^{n} S_i$ by the concept of k-extreme and apply the refined Shapley-Folkman lemma (Theorem 1.3), categorizing the n components into k_ith convex hull. Instead of the traditional way in which each $D(S_i)$ is directly estimated to be $\max_{1\le i\le n} D(S_i)$, we consider all $D(S_i)$ jointly and use the constraint on k_i to reach an even tighter bound. Our new bound presents degree improvement in certain cases, especially for the case when k is $O(n)$.

We raise a natural open question here. Under the same assumption that $S=\sum_{i=1}^{n} S_i \subset \mathbb{R}^m$, and for any point z in the convex hull conv$S$, we can see that there exists a vector $\mathbf{k}=(k_1,\cdots,k_n)^T \in (\mathbb{N}\cup\{0\})^n, \|\mathbf{k}\|_1 \le n+k$ such that $z\in\sum_{i=1}^{n}\text{conv}_{k_i}S$, and we denote $A(\mathbf{k})=\sum_{i=1}^{n}\text{conv}_{k_i}S\subset \text{conv}S$. 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 $A(\mathbf{k})$, S, and conv$S$. While the Shapely-Folkman-Starr theorem focuses on measurement about $d_H(S,\text{conv}S)$, it is fair to mention that Adiprasito (2020) (Theorem 1.1) stated a relevant result of measurement about $d_H(\text{conv}S,\text{conv}_kS)$: Let S be a subset of finite points in $\mathbb{R}^n$ and $k\in\mathbb{N}$, then,

$$d_H(\text{conv}S,\text{conv}_kS)\le\frac{D(S)}{\sqrt{2k}}.$$

Appendix A. Proofs in section 2

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

$$\sqrt{\frac{2(m+1)}{m}}rad(S)\le D(S)\le\frac{rad(S)}{2}.$$

[5] This is achieved when excess demand consists of one student for 82 courses and two students for the remaining 18 courses.

[6] While the results presented by Budish and Reny (2020) and our first bound consider only the maximum of the diameters, our latter bound takes into account the diameters of all feasible bundles for each student.

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

Lemma 2. *Consider the set* $S = \{s_0, s_1, ..., s_{k-1}\} \subset \mathbb{R}^m$*, and denote the cardinality of* S *as* $|S|$*, then* $rad^2(S) \le \frac{k-1}{2k} D^2(S)$.

Proof. For any convex combination z of S, there exists $0 \le \lambda_i \le 1$, $\sum_{i=0}^{k-1} \lambda_i = 1$, such that

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

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

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

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

Proof. It is a well-known result, and the proof has been omitted.

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

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

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

$$y := z \oplus (e_1 + e_2 + \cdots + e_n)^T \in \mathbb{R}^{m+n},$$

$$y = \sum_{i=1}^n z_i \oplus e_i = \sum_{i=1}^n \left(\sum_{j=1}^{m_i} a_{ij} x_{ij} \right) \oplus e_i = \sum_{i=1}^n \sum_{j=1}^{m_i} a_{ij} \left(x_{ij} \oplus e_i \right).$$

Notice that $y \in \mathbb{R}^{m+n}$, by théodory 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 such that,

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

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 2.1. For any k-extreme point z, by Theorem 1.3, there exists $\{z_i\}_{i=1}^n$ such that,

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

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 combinations. Let $T = \sum_{i=1}^n T_i$ and we have $z \in \mathrm{conv} T$. We apply the Shapley-Folkman-Starr Corollary 1.4 to T and $\mathrm{conv} T$, for any $k_i \in \mathbb{N}, \sum_{i=1}^n k_i \leq n + k, 1 \leq i \leq n$,

$$\inf_{w \in T} |z - w|^2 \leq \sum_{i=1}^n \mathrm{rad}^2(T_i) \leq \sum_{i=1}^n \frac{k_i - 1}{2k_i} D^2(T_i),$$

the second inequality is from the lemma 2. $T_i \subset S_i$ implies $T \subset S$, hence,

$$\inf_{y \in S} |z - y|^2 \leq \inf_{w \in T} |z - w|^2 \leq \frac{k_i - 1}{2k_i} D^2(T_i) \leq \sum_{i=1}^n \frac{k_i - 1}{2k_i} D^2(S_i).$$

Then, we will prove the two terms of our result separately. Let $D := \max_{1 \leq i \leq n} D(S_i)$, for any $k_i \in \mathbb{N}, \sum_{i=1}^n k_i \leq n + k, 1 \leq i \leq n$, we firstly claim $\sum_{i=1}^n \frac{k_i - 1}{k_i} D^2(S_i) \leq \frac{nk}{n+k} D^2$,

$$\sum_{i=1}^n \frac{k_i - 1}{k_i} D^2(S_i) \leq \sum_{i=1}^n \frac{k_i - 1}{k_i} D^2 = nD^2 - D^2 \sum_{i=1}^n \frac{1}{k_i} \leq D^2 \left(n - \frac{(\sum_{i=1}^n 1)^2}{\sum_{i=1}^n k_i} \right) \leq \frac{nk}{n + k} D^2.$$

The second last inequality is from the Cauchy-Schwarz inequality.

Then we claim $\sum_{i=1}^n \frac{k_i - 1}{k_i} D^2(S_i) \leq \sum_{i=1}^n D^2(S_i) - \frac{\left(\sum_{i=1}^n D^2(S_i)\right)^2}{n+k}$,

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

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

Proof of Corollary 2.2. it is enough to show that the following optimal program with optimal value $k/2$, given $n > k$,

$$\max \sum_{i=1}^n \frac{k_i - 1}{k_i}$$

$$s.t. \, k_i \in \mathbb{N}, \sum_{i=1}^n k_i \leq n + k.$$

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 maximized value, and $a = 1, j = n-k$ is the only candidate, which claims that the optimal value is $(n-k)(1-1)/1 + k(2-1)/2 = k/2$.

Proof of Corollary 2.3.

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

It is only valid 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 $\sum_{i=1}^{n} D(S_i)^2$ is not greater than nD^2 in degree.

Proof of Corollary 2.4. It is enough to show that,

$$
\begin{aligned}
&\mathbb{E}\left(\sum_{i=1}^{n} D^2(S_i) - \frac{\left(\sum_{i=1}^{n} D(S_i)\right)^2}{n+k}\right) = n\mathbb{E}(D^2(S_i)) - \frac{\mathbb{E}\left[(\sum_{i=1}^{n} D(S_i)^2\right]}{n+k} \\
&= n\mathbb{E}(D^2(S_i)) - \frac{n\mathbb{E}(D^2(S_i)) + n(n-1)\mathbb{E}^2(D(S_i))}{n+k} \\
&= \frac{nk}{n+k}\mathbb{E}(D^2(S_i)) + \frac{n^2\mathbb{E}(D^2(S_i)) - n\mathbb{E}(D^2(S_i)) - n^2\mathbb{E}^2(D(S_i)) + n\mathbb{E}^2(D(S_i))}{n+k} \\
&= \frac{nk}{n+k}\mathbb{E}(D^2(S_i)) + \frac{n(n-1)\mathrm{Var}(D(S_i))}{n+k}.
\end{aligned}
$$

Appendix B. Proof in Section 3

The proof here principally follows in Budish (2011) (Appendix A, Proof of Theorem 1) and modifies his proof with the constraints of l dominant goods.

Firstly, for any given budget vector $\mathbf{b}'$, without loss of generality, let $b'_1 \geq b'_2 \geq \cdots \geq b'_n$ and the taxes $(\tau'_{ix})_{i \in S, x \in 2^C}$ satisfy additional property that $i < j$ implies $\min_x b'_i + \tau'_{ix} > \max_x b'_j + \tau'_{jx}$. The assumption can be achieved due to the finite numbers of agents and schedules.

Secondly, without loss of generality, we assume that the first l entries of the space of allocation correspond to the l dominant goods. We restrict the value of the first l entries of the price space to the same constant value $(\min_x\{b'_l + \tau'_{lx}\} + \max_x\{b'_{l+1} + \tau'_{(l+1)x}\})/2$, which ensures only the first l agents affordable to the dominant goods, and the rest agents unaffordable. The first l entries of the constrained price space have no excess demand (first l entries of $\mathbf{z}$ always equal 0), and we can find a fixed point $\mathbf{p}^*$ in the same logic as in Budish (2011), appendix A, Proof of Theorem 1, step 2 and 3; and keep the same way in step 4 to 6.

Finally, we want to argue that there exists a $(m-l)$-extreme market-clearing excess demand vector in the convex hull of $A(6)$ in step 7, and we can employ our main theorem to find a tighter bound in step 8, comparing to $\sqrt{\sigma M}/2$. It can be done by setting the δ in step 7 small enough (less than $(\min_x\{b'_l+\tau'_{lx}\}-\max_x\{b'_{l+1}+\tau'_{(l+1)x}\})/4$) such that any point in $B_\delta(\mathbf{p}^*)$ will not change the demands of the first l dominant goods, and the non-dominant goods will have a market-clearing excess demand vector in a $(m-l)$-dimension convex hull. (The changing demands $\mathbf{z}(\mathbf{P}^\phi)$ only affect the non-dominant goods.)

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., 2011. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. Journal of Political Economy 119, 1061–1103. URL: http://www.jstor.org/stable/10.1086/664613.

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.