האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

EVOLUTIONARILY STABLE STRATEGIES OF RANDOM GAMES AND THE FACETS OF RANDOM POLYTOPES

By

OHAD NAVON

Discussion Paper # 702 (September 2016)

מרכז פדרמן לחקר הרציונליות

THE FEDERMANN CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Edmond J. Safra Campus,
Jerusalem 91904, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Evolutionarily Stable Strategies of Random Games and the Facets of Random Polytopes

Ohad Navon*

Abstract

An evolutionarily stable strategy (ESS) is an equilibrium strategy that is immune to invasions by rare alternative (mutant) strategies. Unlike Nash equilibria, ESS do not always exist in finite games. In this paper we address the question of what happens when the size of the game increases: does an ESS exist for almost every" large game? We let the entries of an $n \times n$ game matrix be independently randomly chosen according to a symmetrical subexponential distribution F, and study the expected number of ESS with support of size d as $n \to \infty$. In a previous paper by Hart, Rinott and Weiss [6] it was shown that this limit is $\frac{1}{2}$ for $d = 2$.

This paper deals with the case of $d \geq 4$, and proves the conjecture in [6] (Section 6,c), that the expected number of ESS with support of size $d \geq 4$ is 0.

Furthermore, it discusses the classic problem of the number of facets of a convex hull of n random points in $\mathbb{R}^d$, and relates it to the above ESS problem. Given a collection of i.i.d. random points, our result implies that the expected number of facets of their convex hull converges to 2^d as $n \to \infty$.

*Institute of Mathematics, and Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. [249159]. Moreover, it was supported by the Einstein Institute of Mathematics, and Federmann Center for the Study of Rationality. This paper was written as the author's M.Sc. thesis under the supervision of Sergiu Hart

Acknowledgements

I wish to thank my supervisor Sergiu Hart for his extensive support throughout my studies. I could not have asked for better supervision.

My thanks to the Center for the Study of Rationality for financial and administrative support.

Special appreciation is due to Ilan Nehama and Yannai Gonczarowski for their helpful discussions and encouragement.

Last but not least, I wish to thank my family for accompanying me on this challenging journey.

Contents

Section 1

Introduction

The concept of evolutionarily stable strategy (ESS for short) was introduced by Maynard Smith and Price [8]. It refers to a strategy that, when played by the whole population, is immune to invasions by rare alternative ("mutant") strategies (see Section 4 for a precise definition).

Formally, an ESS corresponds to a symmetric Nash equilibrium that satisfies an additional stability requirement. Every (symmetric) finite game has a (symmetric) Nash equilibrium, but the same is not true for ESS. There are games with finitely many pure strategies that have no ESS. Moreover, the nonexistence of ESS is not an "isolated" phenomenon: it holds for open sets of games. For instance, the "rock-scissors-paper" game and all its small enough perturbations have no ESS.

This leads us to the question of what happens to the existence of ESS when the number of strategies is large. Does an ESS exist for "almost every" large game? Specifically, assuming that the payoffs in the game are identically independently distributed random variables, what is the probability that an ESS exists and what is the limit of this probability as the size of the game increases? For pure ESS, the answer to this question is simple: the probability that a pure ESS exists is $1-(1-\frac{1}{n})^n$, which converges to $1-\frac{1}{e} \simeq 63\%$ as $n \to \infty$, where n is the number of pure strategies.

But what about mixed ESS? In the present paper, we study mixed ESS with support of size d, called "d-point ESS". We find that, unlike for pure ESS, here the answer depends on the underlying distribution F from which the payoffs are drawn. We focus on subexponential distributions (e.g. pareto, lognormal, and more; see Chapter 2 for a precise definition).

To capture some intuition, consider the family of cumulative distribution functions $F_\alpha(x) = 1 - e^{-x^\alpha}$, for all $x \geq 0$, where $\alpha \in (0, 1)$. Our results imply that for every $d \geq 4$, the expected number of d-point ESS converges to 0 as $n \to \infty$. This result extends that of Hart, Rinott and Weiss [6], who showed that the expected number of two-point ESS converges to $1 - \frac{1}{\sqrt{e}}$.

An interesting consequence of our result concerns a classic problem originally studied by Rényi and Sulanke [7]. Given n i.i.d. random points in $\mathbb{R}^d$, what is the expected number of facets of their convex hull as $n \to \infty$? Assuming that the coordinates are random variables that are subexponential and symmetric, we show that the expected number of such facets converges to 2^d as $n \to \infty$.

The paper is organized as follows: Section 2 introduces subexponential distributions and their main properties. Section 3 studies random polytopes and proves the main result stated above. Section 4 discusses d-point ESS.

Section 2

Subexponential Distributions

Definition 2.1. *Let F be a probability distribution function and denote its tail by $G = 1-F$. F will be called a* ***subexponential distribution*** *if the following conditions are met:*

1. *Define $X_+ = max\{X, 0\}, Y_+ = max\{Y, 0\}$ for X, Y i.i.d.-F (identically independently distributed according to F); then*

$$\lim_{t\to\infty} \frac{Pr[X_+ + Y_+ > t]}{Pr[max\{X_+, Y_+\} > t]} = 1 \tag{2.1}$$

2. *For every $\epsilon > 0$ there exist $t_0 > 0$ and $c_0 > 1$ such that*

$$\forall t > t_0, c \in (1, c_0) : \frac{G(ct)}{G(t)^c} > 1 - \epsilon \tag{2.2}$$

The notion of subexponentiality was introduced by Chistyakov [2] in relation to the first condition mentioned above. Some examples include (see [5] for a more extensive list):

- Regularly varying tails: $G(x) = x^{-\alpha} l(x)$, where $\alpha \geq 0$ and l is a slowly varying function. That is, $\lim_{x\to\infty} \frac{l(cx)}{l(x)} = 1$ for every $c > 0$. In particular, these include:
  - Pareto: $G(x) = x^{-\alpha}$, for $x \in (1, \infty)$, where $\alpha > 0$.
  - Cauchy: $G(x) = \int_x^\infty \frac{1}{\pi(1+y^2)} \mathrm{d}y = \frac{1}{\pi} \arctan(x) + \frac{1}{2}$, for $x \in (0, \infty)$.
- Lognormal: $G(x) = \int_x^\infty \frac{1}{y\sqrt{2\pi}} e^{-\log^2 \frac{y}{2}} \mathrm{d}y$, for $x \in (0, \infty)$.
- Weibull: $G(x) = e^{-x^\alpha}$, for $x \in (0, \infty)$, where $\alpha \in (0, 1)$.
- ”Almost exponential”: $G(x) = e^{-x \log^{-\alpha} x}$, for $x \in (0, \infty)$, where $\alpha > 0$.

By the cumulative hazard function $g(t) := -\log G(t)$ in the second condition that for every $\epsilon > 0$ there exist $t_0 \equiv t_0(\epsilon)$ and $c_0 \equiv c_0(\epsilon) > 1$ such that $g(ct) \leq cg(t) + \epsilon$ for all $t > t_0$ and all $c \in (1, c_0)$. Therefore, a sufficient condition for the second part of definition 2.1 is that $\frac{g(t)}{t}$ is a non-increasing function for all large enough t. This is the case when g is concave (and so G is log-convex) or even star-concave (we will see in Lemma 2.2 below that (2.1) implies that $\frac{g(t)}{t} \to 0$ as $t \to \infty$). It is now easy to verify that all the distributions listed above satisfy both conditions.

Finally, subexponential distributions are closed under "tail equivalence." That is, if $1 - F(t) \backsim 1 - \hat{F}(t)$ as $t \to \infty$, then F is subexponential if and only if $\hat{F}$ is too. A partial proof is found in Theorem 3 in Teugels [9].

The following lemma collects a number of properties that will be used throughout the paper.

Lemma 2.2. *There exists $\gamma_t : \mathbb{R}_+ \to \mathbb{R}_+$ such that for every $a > 0$:*

1. $\lim_{t\to\infty} \gamma_t = 0$
2. $\lim_{t\to\infty} \gamma_t t = \infty$
3. $\lim_{t\to\infty} G(t)^{a\gamma_t} = 1$
4. $\lim_{t\to\infty} \frac{G\big((1+a\gamma_t)t\big)}{G(t)} = 1$

Proof. See Lemma 18 in [6] for (1)-(3), whereas (4) follows from (3) together with (2.2). □

Lemma 2.3. *Let $X_1, \ldots, X_d$ be i.i.d. with a common subexponential distribution F; then*

$$\lim_{t\to\infty} Pr[X_1 + \cdots + X_d > t] \lesssim dG(t) \tag{2.3}$$

That is, for every $\epsilon > 0$ there exists t_0 such that for every $t > t_0$ it holds that

$$Pr[X_1 + \cdots + X_d > t] < (1 + \epsilon)dG(t) \tag{2.4}$$

Proof. By Lemma 1 of Teugels [9] the above holds for $max\{X_i, 0\}$'s. The rest immediately follows as:

$$X_1 + \cdots + X_d \leq max\{X_1, 0\} + \cdots + max\{X_d, 0\} \tag{2.5}$$

□

Section 3

Facets of Random Polytopes

Let $P_1, \ldots, P_n$ be a collection of points in the d-dimensional real space $\mathbb{R}^d$. Moreover, let each P_i be a random point whose coordinates are i.i.d. random variables distributed according to a symmetrical continuous subexponential distribution F. Furthermore, we assume that P_i is independent of P_j for every $i \neq j$.

Consider the convex hull of $P_1, \ldots, P_n$. By Lemma A.1 it almost surely has a strictly positive d-dimensional volume. Denote by $\mathcal{F}_n$ the number of its facets. A facet is any $(d-1)$-dimensional hyperplane passing through some d points of $\{P_1, \ldots, P_n\}$ such that the convex hull lies entirely on one side of that hyperplane.

Theorem 3.1.

$$\forall d \in \mathbb{N} : \lim_{n\to\infty} \mathbb{E}[\mathcal{F}_n] = 2^d \tag{3.1}$$

3.1 Preliminaries

We estimate the probability that $P_1, \ldots, P_d$ to generate a hyperplane with a strictly positive outward normal that contains a facet of the convex hull. By Lemma A.2 (see Appendix) these points almost surely generate a $(d-1)$-dimensional hyperplane that does not pass through the origin. Denote that hyperplane by $\mathcal{H}$ and define:

$$\mathcal{H} = \left\{x \in \mathbb{R}^d : V \cdot x = 1\right\} \tag{3.2}$$

Notice that V is uniquely determined by $P_1, \ldots, P_d$. Moreover, it is a vector of random variables that are dependent on $\{P_i^j\}_{i,j=1}^d$. Lastly, it is a continuous random variable as it is linearly dependent on $\{P_i^j\}_{i,j=1}^d$.

Remark 3.2. *The term **outward** with respect to our hyperplane normal is relevant only in the case where $\mathcal{H}$ is a $(d-1)$-dimensional hyperplane that does not contain the origin. In* *such a case the direction that is referred to as "outward" denotes the half-space that does not contain the origin.*

We use the term *lies above* to refer to any point P such that

$$V \cdot P > 1 \tag{3.3}$$

and we do so only when $P_1, \ldots, P_d$ generate a $(d-1)$-dimensional hyperplane. Moreover, since F is symmetric, whenever $\mathcal{H}$ contains a facet the probability that it has a strictly positive outward normal is 2^{-d}.

Remark 3.3. *Henceforth we denote by $V \gg 0$ the event where $P_1, \ldots, P_d$ determine a unique hyperplane that does not pass through the origin and whose normal has strictly positive coordinates.*

Let $Q = (Q^1, \ldots, Q^d) \in \mathbb{R}^d$ be a random point where the Q^i are i.i.d.-F. We define the following random variable:

$$U := \begin{cases} Pr\Big[V \cdot Q > 1 \mid P_1, \ldots, P_d\Big] & : V \gg 0 \\ 1 & : \text{otherwise} \end{cases}$$

Let $\mathcal{F}_n^+$ be the number of facets of the convex hull with a positive outward normal.

Proposition 3.4.

$$\forall d \geq 2 : \lim_{n \to \infty} \mathbb{E}[\mathcal{F}_n^+] = 1 \tag{3.4}$$

Notice that Proposition 3.4 implies Theorem 3.1. As V is continuous it is almost surely not parallel to any axis. Moreover, our distribution F is symmetric and there are exactly 2^d orthants in which V can be found.

We introduce a few additional notations. For every $t > 0$, we define the following sets:

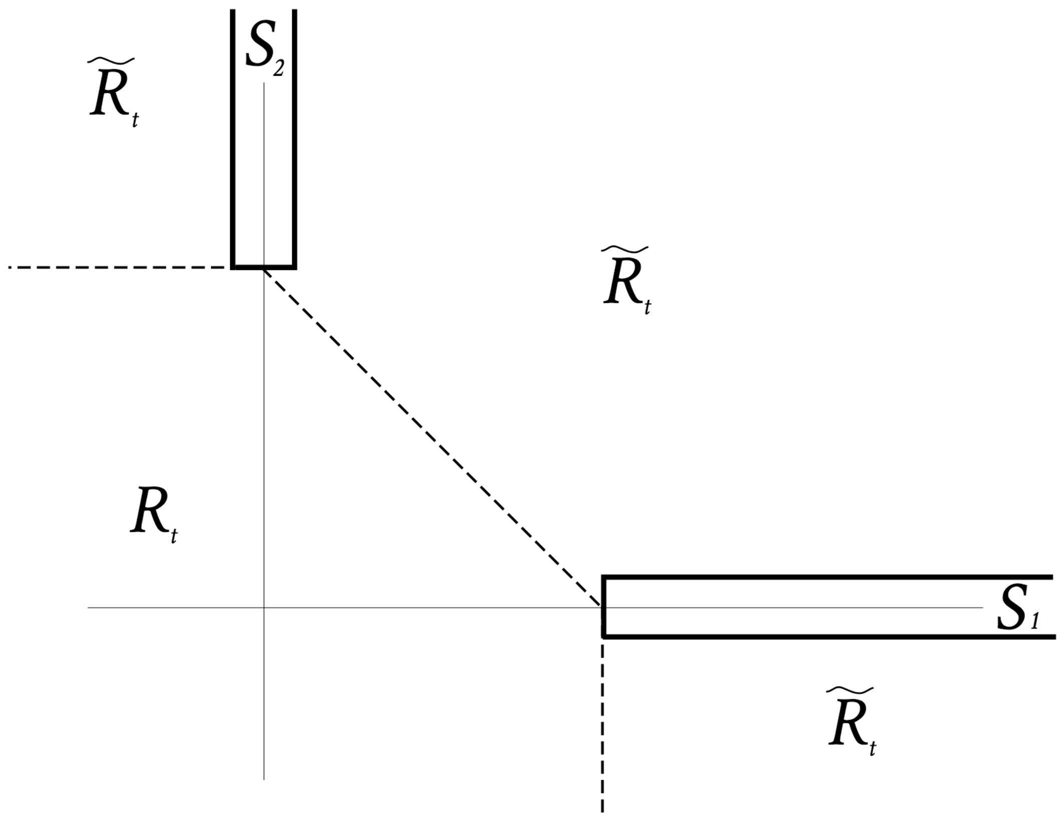
$$R_t = \Big\{x \in \mathbb{R}^d \mid \sum_{i=1}^{d} x^i < t; \forall i \leq d : x^i < t\Big\} \tag{3.5}$$

$$S_t^i = \Big\{x \in \mathbb{R}^d \mid x^i \geq t; \forall j \neq i : |x^j| \leq \gamma_t t\Big\} \tag{3.6}$$

$$S_t = \bigcup_{i=1}^{d} S_t^i \tag{3.7}$$

$$\tilde{R}_t = \mathbb{R}^d \setminus \Big(S_t \cup R_t\Big) \tag{3.8}$$

Note that R stands for *region*, whereas S stands for *sleeve*.



3.2 Main Theorem

In what follows, let $\epsilon > 0$.

Remark 3.5. *Notations:*

- *For* $0 < u < \frac{1}{1+\epsilon}$, *define* $t_{u,\epsilon}$ *such that* $G(t_{u,\epsilon}) = (1+\epsilon)u$. *Moreover, we say that* $t_{u,\epsilon}$ ***corresponds*** *to* u *(whenever it is clear from the context,* t *will stand for* $t_{u,\epsilon}$*).*
- *Recall that* $Q = (Q^1, \ldots, Q^d) \in \mathbb{R}^d$ *be a random point where the* Q^i *are i.i.d.-*F.
- *Denote by* $\mathcal{X}$ *the event that** $\mathcal{H}$ *has a positive outward normal and contains a facet.*
- $H(t) = 1 - F(-t) - G(t)$.

Lemma 3.6.

$$Pr\Big[\mathcal{X}\Big] = (n-d)\int_0^1 (1-u)^{n-d-1} F_U(u)\mathrm{d}u \tag{3.9}$$

*The hyperplane determined by $P_1, \ldots, P_d$.

Proof. Define $\mathcal{Y}$ as the event where:

- $\mathcal{H}$ does not pass through the origin.
- $V \gg 0$.
- For every $i > d$ it holds that $P_i \cdot V < 1$.

By the definition of U and the independence of the P_i's,

$$Pr\Big[\forall i > d : V \cdot P_i < 1 \mid P_1, \dots, P_d\Big] = (1 - U)^{n-d} \tag{3.10}$$

Therefore,

$$Pr\Big[\mathcal{Y}\Big] = \int_0^1 (1 - u)^{n-d} \mathrm{d}F_U(u) \tag{3.11}$$

Notice that for all $j > d$ the points P_j are independent of $P_1, \dots, P_d$, and hence of U too. Moreover, the integrand vanishes at $u = 1$, and so the atom there does not matter. Now integration by parts yields

$$\int_0^1 (1 - u)^{n-d} \mathrm{d}F_U(u) = \Big[(1 - u)^{n-d} F_U(u)\Big]_0^1 + (n - d) \int_0^1 (1 - u)^{n-d-1} F_U(u) \mathrm{d}u \tag{3.12}$$

Notice that the first term vanishes.

Notice that $\mathcal{X}$ is contained in the union of $\mathcal{Y}$ and the event where $\mathcal{H}$ passes through the origin. The latter event has a probability 0 by Lemma A.2 (see Appendix), which concludes the proof. □

Lemma 3.7. *As* $u \to 0$*,*

$$Pr\Big[Q \notin R_{t_{u,\epsilon}}\Big] = O(u) \tag{3.13}$$

Proof. For every $t > 0$,

$$Pr\Big[Q \notin R_t\Big] \leq Pr\Big[\sum_{i=1}^{d} Q^i > t\Big] + \sum_{i=1}^{d} Pr\Big[Q^i > t\Big] \tag{3.14}$$

By Lemma 2.3 there exists t_0 such that for every $t > t_0$ it holds that:

$$Pr\Big[\sum_{i=1}^{d} Q^i > t\Big] < (1 + \epsilon) dG(t) = (1 + \epsilon)^2 du = O(u) \tag{3.15}$$

Moreover,

$$\sum_{i=1}^{d} Pr\Big[Q^i > t\Big] = dG(t) = (1 + \epsilon) du = O(u) \tag{3.16}$$

Combining the two completes the proof. □

Lemma 3.8. *There exists* $u_0 \in (0,1)$ *such that for all* $0 < u < u_0$ *and every* $0 \ll v \in \mathbb{R}^d$ *it holds that*

if $Pr\Big[v \cdot Q > 1\Big] < u$ *then* $\forall i \le d : t_{u,\epsilon} < \frac{1}{v^i}$

Proof. Without loss of generality let $v^1 \ge v^2 \ge \cdots \ge v^d > 0$, let $t_{u,\epsilon} = t$.

If $t \ge \frac{1}{v^1}$ then,

$$\begin{aligned} Pr\Big[v \cdot Q > 1\Big] &\ge Pr\Big[v \cdot Q > v^1 t\Big] \\ &= Pr\Big[\frac{1}{v^1} \sum_{i=1}^{d} v^i Q^i > t\Big] \\ &\ge Pr\Big[\forall i > 1 : |Q^i| < \gamma_t t; Q^1 > (1 + (d-1)\gamma_t)t\Big] \\ &= \Big(1 - F(-\gamma_t t) - G(\gamma_t t)\Big)^{d-1} G(t + (d-1)\gamma_t t) \\ &= H(\gamma_t t)^{d-1} G(t + (d-1)\gamma_t t) \end{aligned}$$

By Lemma 2.2 we may pick t_0 such that for all $t > t_0$ it holds that

$$H(\gamma_t t)^{d-1} > (1+\epsilon)^{-\frac{1}{2}} \tag{3.17}$$

and so

$$G(t + (d-1)\gamma_t t) > (1+\epsilon)^{-\frac{1}{2}} G(t) \tag{3.18}$$

and so

$$Pr\Big[v \cdot Q > 1\Big] > \frac{G(t)}{1+\epsilon} = u \tag{3.19}$$

which completes the proof. □

Corollary 3.9. *There exists* $u_0 > 0$ *such that for all* $0 < u < u_0$ *it holds that if* $U < u$ *then* $P_i \notin R_{t_{u,\epsilon}}$ *for all* $i \le d$.

Proof. Pick u_0 according to Lemma 3.8. Let e^i be the i-th unit vector; then the point te^i lies below $\mathcal{H}$ for all $i \le d$, and so by Lemma A.7, does the entire set R_t. Therefore $\forall i \le d : P_i \notin R_t$, as these lie on $\mathcal{H}$. □

Lemma 3.10. *As* $u \to 0$,

$$Pr\Big[U < u\Big] = O(u^d) \tag{3.20}$$

Proof. By Corollary 3.9 there exists u_0 such that for all $u < u_0$ we have

$$Pr\Big[U < u\Big] \le Pr\Big[\forall i \le d : P_i \in \mathbb{R}^d \setminus R_t\Big] = \Big(Pr\Big[Q \notin R_t\Big]\Big)^d \tag{3.21}$$

and so by Lemma 3.7 we have

$$Pr\Big[U < u\Big] = O(u^d) \tag{3.22}$$

which completes the proof. □

Proposition 3.11. *Let $\alpha > 0$ and $u_0 > 0$ such that $F_U(u) < \alpha u^d$ for all $0 < u < u_0$; then,*

$$\lim_{n\to\infty} \mathbb{E}[\mathcal{F}_n^+] \le \alpha \tag{3.23}$$

Proof. Separating the integral in Lemma 3.6 into two parts yields

$$(n-d)\int_0^{u_0}(1-u)^{n-d-1}F_U(u)\mathrm{d}u + (n-d)\int_{u_0}^{1}(1-u)^{n-d-1}F_U(u)\mathrm{d}u \tag{3.24}$$

for the left summand,

$$\begin{aligned}(n-d)\textstyle\int_0^{u_0}(1-u)^{n-d-1}F_U(u)\mathrm{d}u &\le \alpha(1+\epsilon)(n-d)\textstyle\int_0^{u_0}(1-u)^{n-d-1}u^d\mathrm{d}u\\ &\le \alpha(1+\epsilon)(n-d)\textstyle\int_0^{1}(1-u)^{n-d-1}u^d\mathrm{d}u\\ &= \frac{\alpha(1+\epsilon)}{\binom{n}{d}}\end{aligned}$$

For a detailed computation of the last equality see Lemma A.3 in the Appendix. It is fairly easy to deduce that

$$(n-d)\int_0^{u_0}(1-u)^{n-d-1}F_U(u)\mathrm{d}u \le \alpha n^{-d} + \epsilon O(n^{-d}) \tag{3.25}$$

As for the right summand,

$$(n-d)\int_{u_0}^{1}(1-u)^{n-d-1}F_U(u)\mathrm{d}u \le (n-d)(1-u_0)^{n-d-1} \tag{3.26}$$

which is less than $\epsilon O(n^{-d})$.

In conclusion,

$$Pr\Big[\mathcal{X}\Big] \le \alpha n^{-d} + \epsilon O(n^{-d}) \tag{3.27}$$

To complete the proof we define a collection of $\binom{n}{d}$ indicators. For each choice of $1 \le i_1 < \cdots < i_d \le d$, we define $\mathcal{X}_{\{i_k\}_{k=1}^d}$ as the event where:

- The convex-hull of $P_1, \ldots, P_n$ has a strictly positive volume.
- $\{P_{i_k}\}_{k=1}^d$ generate a hyperplane that does not pass through the origin and contains a facet of the convex hull that has a positive outward normal.

Now denote by $\mathbb{1}_{\{i_k\}_{k=1}^d}$ the indicator of $\mathcal{X}_{\{i_k\}_{k=1}^d}$.

As P_i and P_j are independent for all $i \ne j$, we have

$$Pr\Big[\mathcal{X}\Big] = Pr\Big[\mathbb{1}_{\{k\}_{k=1}^d} = 1\Big] \tag{3.28}$$

Recall that $\mathcal{F}_n^+$ is the number of facets of the convex hull with a positive outward normal. Then,

$$\mathcal{F}_n^+ = \sum_{1\le i_1<\cdots<i_d\le d} \mathbb{1}_{\{i_k\}_{k=1}^d} = \binom{n}{d}\mathbb{1}_{1,\ldots,d} \tag{3.29}$$

implying that

$$\mathcal{F}_n^+ = \binom{n}{d} Pr\Big[\mathcal{X}\Big] \tag{3.30}$$

Thus, by (3.27),

$$\mathcal{F}_n^+ \leq (\alpha + \epsilon)\binom{n}{d} O(n^{-d}) \leq \alpha + O(\epsilon) \tag{3.31}$$

which completes the proof. □

Similarly, we have

Proposition 3.12. *Let $\alpha > 0$ and $u_0 > 0$ such that $F_U(u) > \alpha u^d$ for all $0 < u < u_0$; then,*

$$\lim_{n\to\infty} \mathbb{E}[\mathcal{F}_n^+] \geq \alpha \tag{3.32}$$

The proof is essentially identical to that of Proposition 3.11.

The following proposition will yield a better understanding of the convex hull and particularly of its facets.

Definition 3.13. *If $V \gg 0$ then $\mathcal{H}$ is said to satisfy the* ***permutation property*** *if there exists a permutation π of $\{1, \ldots, d\}$ such that*

$$\forall i \leq d : P_i \in S_t^{\pi(i)} \tag{3.33}$$

Proposition 3.14. *Denote by $\mathcal{NPP}$ the event where $\mathcal{H}$* ***does not satisfy*** *the permutation property. Then, as $u \to 0$,*

$$Pr\Big[\mathcal{NPP}, U < u\Big] = \epsilon O(u^d) \tag{3.34}$$

We need the following lemmas to prove Proposition 3.14.

Lemma 3.15. *Define*

$$\hat{R}_{t_{u,\epsilon}} = \left\{x \in \mathbb{R}^d : \sum_{i=1}^{d} x_i > t\right\} \setminus S_t \tag{3.35}$$

Then, as $u \to 0$,

$$Pr\Big[Q \in \hat{R}_{t_{u,\epsilon}}\Big] = \epsilon O(u) \tag{3.36}$$

Proof. Notice that

$$Pr\Big[Q \in \{x \in \mathbb{R}^d : \sum_{i=1}^{d} x_i > t\}\Big] = Pr\Big[\sum_{i=1}^{d} Q^i > t\Big] \tag{3.37}$$

and that

$$Pr\Big[Q \in S_t\Big] = \sum_{i=1}^{d} Pr\Big[\forall j \neq i : |Q^j| < \gamma_t t; Q^i > (1 + (d-1)\gamma_t)t\Big] \tag{3.38}$$

Notice that for all $t > t_0$ large enough so that $\gamma_t < 1$, the sets S_t^i are disjoint. Thus:

$$Pr\Big[Q \in \hat{R}_t\Big] = Pr\Big[\sum_{i=1}^{d} Q^i > t\Big] - \sum_{i=1}^{d} Pr\Big[\forall j \neq i : |Q^j| < \gamma_t t; p^i > (1 + (d-1)\gamma_t)t\Big] \tag{3.39}$$

By Lemma 2.3, for every $t > t_1 > t_0$,

$$Pr\Big[\sum_{i=1}^{d} Q^i > t\Big] < (1+\epsilon)dG(t) = d(1+\epsilon)^2 u = du + \epsilon O(u) \tag{3.40}$$

Define α to be

$$\alpha = Pr\Big[\forall j \neq i : |Q^j| < \gamma_t t; Q^i > (1 + (d-1)\gamma_t)t\Big] \tag{3.41}$$

As α is independent of i we have

$$Pr\Big[Q \in \hat{R}_t\Big] \leq du + \epsilon O(u) - d\alpha \tag{3.42}$$

Notice that,

$$\alpha = H(\gamma_t t)^{d-1} G(t + (d-1)\gamma_t t) \tag{3.43}$$

By Lemma 2.2 we can choose $t_2 > t_1$ so that for all $t > t_2$ it holds that

$$G(t + (d-1)\gamma_t t) > \frac{G(t)}{\sqrt{1+\epsilon}} \tag{3.44}$$

and

$$H(\gamma_t t) > \frac{1}{\sqrt{1+\epsilon}} \tag{3.45}$$

Thus,

$$\alpha \geq \frac{G(t)}{1+\epsilon} = u \tag{3.46}$$

which implies that

$$Pr\Big[Q \in \hat{R}_t\Big] \leq du + \epsilon O(u) - du = \epsilon O(u) \tag{3.47}$$

□

Lemma 3.16. *As* $u \to 0$,

$$Pr\left[Q \in \tilde{R}_{t_{u,\epsilon}}\right] = \epsilon O(u) \tag{3.48}$$

Proof. Notice that

$$\tilde{R}_t \setminus \hat{R}_t \subset \bigcup_{i=1}^{d} \bigcup_{j \neq i} \left\{x \in \mathbb{R}^d \mid x^i > t, x^j < -\gamma_t t\right\} \tag{3.49}$$

Thus,

$$Pr\left[Q \in \tilde{R}_t \setminus \hat{R}_t\right] \leq \sum_{i=1}^{d} \sum_{j \neq i} Pr\left[Q^i > t, Q^j < -\gamma_t t\right] \tag{3.50}$$

which further translates to

$$Pr\left[Q \in \tilde{R}_t \setminus \hat{R}_t\right] \leq d(d-1)F(-\gamma_t t)G(t) \tag{3.51}$$

Now, for all $t > t_0$, it holds that $F(-\gamma_t t) \leq \epsilon$. Thus

$$Pr\left[Q \in \tilde{R}_t \setminus \hat{R}_t\right] \leq \epsilon d(d-1)G(t) = \epsilon d(d-1)(1+\epsilon)u = \epsilon O(u) \tag{3.52}$$

Combining the two we get

$$Pr\left[Q \in \tilde{R}_t\right] = Pr\left[Q \in \hat{R}_t\right] + Pr\left[Q \in \tilde{R}_t \setminus \hat{R}_t\right] = \epsilon O(u) \tag{3.53}$$

Therefore, by Lemma 3.15, the proof is completed. □

We get the following corollary from Lemma 3.16.

Corollary 3.17.

$$Pr\left[\exists i \leq d : P_i \notin S_{t_{u,\epsilon}}, U < u\right] = \epsilon O(u^d) \tag{3.54}$$

Proof. By Corollary 3.9 there exists u_0 such that if $U < u < u_0$, then $P_i \notin R_t$ for all $i \leq d$. Moreover, let t_0 correspond to u_0.

Hence, if $P_i \notin S_t$, then for all $t > t_0$,

$$P_i \in \tilde{R}_t \tag{3.55}$$

Furthermore, by Lemma 3.10,

$$Pr\left[\exists i \leq d : P_i \in \tilde{R}_t, U < u\right] \leq Pr\left[\exists i \leq d : P_i \in \tilde{R}_t; \forall j \neq i : P_j \notin R_t\right] \tag{3.56}$$

Notice that $\{P_i \in \tilde{R}_t\}$ is independent of $\{\forall j \neq i : P_j \notin R_t\}$. Moreover, recall that Q is a random point; then we have

$$Pr\Big[\exists i \le d : P_i \in \tilde{R}_t, U < u\Big] \le dPr\Big[Q \in \tilde{R}_t\Big]\Big(Pr\Big[Q \notin R_t\Big]\Big)^{d-1} \tag{3.57}$$

By Lemma 3.16,

$$Pr\Big[Q \in \tilde{R}_t\Big] = \epsilon O(u) \tag{3.58}$$

and by Lemma 3.7,

$$\Big(Pr\Big[Q \notin R_t\Big]\Big)^{d-1} = O(u^{d-1}) \tag{3.59}$$

Thus,

$$Pr\Big[Q \in \tilde{R}_t\Big]\Big(Pr\Big[Q \notin R_t\Big]\Big)^{d-1} = \epsilon O(u^d) \tag{3.60}$$

which completes the proof. □

Recall that if $P_1 \in S_t^k$ for some $k \le d$ then $P_1^k > t$.

Lemma 3.18. *Let $k_1, k_2 \le d$; then for every $a > 0$ as $u \to 0$,*

$$Pr\Big[\forall i : P_i \in S_{t_{u,\epsilon}}; P_1 \in S_{t_{u,\epsilon}}^{k_1}, P_2 \in S_{t_{u,\epsilon}}^{k_2}, 1 < \frac{p_1^{k_1}}{p_2^{k_2}} < 1 + a\gamma_{t_{u,\epsilon}}, U < u\Big] = \epsilon O(u^d) \tag{3.61}$$

Proof. Let $\forall$ be the event where:

- $\forall i : P_i \in S_t$
- $P_i \in S_t^{k_i}$ for $i = 1, 2$
- $1 < \frac{p_1^{k_1}}{p_2^{k_2}} < 1 + a\gamma_t$

Denote $\rho_t = 1 + a\gamma_t$. Thus,

$$Pr\Big[\forall, U < u \mid P_2\Big] \le Pr\Big[\forall i > 2 : P_i \notin R_t\Big]\Big(G(p_2^{k_2}) - G(\rho_t p_2^{k_2})\Big) \tag{3.62}$$

By Lemma 3.7,

$$Pr\Big[\forall, U < u \mid P_2\Big] \le O(u^{d-2})\Big(G(p_2^{k_2}) - G(\rho_t p_2^{k_2})\Big) \tag{3.63}$$

By definition 2.1 there exists t_1 such that for every $t > t_1$,

$$(1-\epsilon)G(p_2^{k_2})^{\rho_t} < G(\rho_t p_2^{k_2}) \tag{3.64}$$

Thus,

$$Pr\Big[\mathcal{Y}, U < u \mid P_2\Big] \le O(u^{d-2})\Big(G(p_2^{k_2}) - (1-\epsilon)G(p_2^{k_2})^{\rho_t}\Big) \tag{3.65}$$

Therefore,

$$Pr\Big[\mathcal{Y}, U < u\Big] \le O(u^{d-2}) \int_t^\infty (G(p_2^{k_2}) - (1-\epsilon)G(p_2^{k_2})^{\rho_t})\mathrm{d}F(p_2^{k_2}) \tag{3.66}$$

which further translates to

$$Pr\Big[\mathcal{Y}, U < u\Big] \le O(u^{d-2})\Big[-\frac{1}{2}G(p_2^{k_2})^2 - (1-\epsilon)\frac{-G(p_2^{k_2})^{1+\rho_t}}{1+\rho_t}\Big]_{p_2^{k_2}=t}^{\infty} \tag{3.67}$$

As $\lim_{p_2^{k_2}\to\infty} G(p_2^{k_2}) = 0$ we have

$$Pr\Big[\mathcal{Y}, U < u\Big] \le O(u^{d-2})\Big(0 + \frac{1}{2}G(t)^2 - (1-\epsilon)\frac{G(t)^{1+\rho_t}}{1+\rho_t}\Big) \tag{3.68}$$

By Lemma 2.2, we can choose $t_2 > t_1$ such that for all $t > t_2$, it holds that

$$\frac{G(t)^{1+\rho_t}}{1+\rho_t} > (1-\epsilon)\frac{G(t)^2}{2} \tag{3.69}$$

which implies that

$$Pr\Big[\mathcal{Y}, U < u\Big] \le O(u^{d-2})\Big(\frac{1}{2}G(t)^2 - (1-\epsilon)^2\frac{G(t)^2}{2}\Big) \tag{3.70}$$

or simply

$$Pr\Big[\mathcal{Y}, U < u\Big] \le O(u^{d-2})(2\epsilon - \epsilon^2)G(t^2) = \epsilon O(u^d) \tag{3.71}$$

which completes the proof. □

The following lemma gives us a better understanding of the sleeves S_t^i.

Lemma 3.19. *There exists u_0 such that for all $k \le d$ and any $Q_1, Q_2 \in \mathcal{H} \cap S_{t_{u,\epsilon}}^k$, whenever $u < u_0$ we have:*

If $U < u$ then $|1 - \frac{Q_1^k}{Q_2^k}| \le 3(d-1)\gamma_{t_{u,\epsilon}}$

Proof. Recall that $\mathcal{H}$ is represented by $V \cdot x = 1$ and the probability for a random point to lie above it is U. Moreover, let u_0 be so that for all $U < u < u_0$, by Lemma 3.8 for every $i \le d$, we have $\frac{1}{V^i} \ge t$ or $V^i \le \frac{1}{t}$, which implies that

$$\sum_{i \ne k}^{d} V^i \le \frac{d-1}{t} \tag{3.72}$$

Without loss of generality $Q_1^k > Q_2^k$. Furthermore, let e^k be the k-th unit vector. Since $Q_1, Q_2, \frac{1}{V^k}e^k \in \mathcal{H} \cap S_t^k$ we get

$$Q_1^k \leq \frac{1}{V^k}(1 + \sum_{i \neq k}^{d} \gamma_t t V^i) \tag{3.73}$$

$$Q_2^k \geq \frac{1}{V^k}(1 - \sum_{i \neq k}^{d} \gamma_t t V^i) \tag{3.74}$$

Thus,

$$\frac{Q_1^k}{Q_2^k} \leq \frac{1 + \sum_{i \neq k}^{d} \gamma_t t V^i}{1 - \sum_{i \neq k}^{d} \gamma_t t V^i} = \frac{1 + \gamma_t t \sum_{i \neq k}^{d} V^i}{1 - \gamma_t t \sum_{i \neq k}^{d} V^i}. \tag{3.75}$$

By (3.72),

$$\frac{Q_1^k}{Q_2^k} \leq \frac{1 + \gamma_t t \frac{d-1}{t}}{1 - \gamma_t t \frac{d-1}{t}} = \frac{1 + (d-1)\gamma_t}{1 - (d-1)\gamma_t} = 1 + \frac{2(d-1)\gamma_t}{1 - (d-1)\gamma_t} \tag{3.76}$$

Now pick $t_1 > t_0$ (where t_0 corresponds to u_0) such that for every $t > t_1$ we have

$$\gamma_t \leq \frac{1}{3(d-1)} \tag{3.77}$$

Thus,

$$1 - (d-1)\gamma_t \geq 1 - \frac{d-1}{3(d-1)} = \frac{2}{3} \tag{3.78}$$

or simply

$$\frac{2(d-1)\gamma_t}{1 - (d-1)\gamma_t} \leq 2(d-1)\gamma_t \frac{3}{2} = 3(d-1)\gamma_t \tag{3.79}$$

In conclusion, for every $t > t_1$,

$$\frac{Q_1^k}{Q_2^k} \leq 1 + 3(d-1)\gamma_t \tag{3.80}$$

□

Corollary 3.20. *If $\mathcal{H}$ has a strictly positive normal, then there exists t_0 such that for all $t > t_0$ and for every $k \leq d$ and any $Q_1, Q_2 \in \mathcal{H} \cap S_t^k$ we have*

If for every $i \leq d$ the i-th axis intersection† *of $\mathcal{H}$ is larger than t then $|1 - \frac{Q_1^k}{Q_2^k}| \leq 3(d-1)\gamma_t$*

†The i-th axis intersection of the hyperplane $\mathcal{H}$ is the unique point $\hat{Q} \in Span\{e^i\} \cap \mathcal{H}$. As $\mathcal{H}$ has a strictly positive normal, that point is indeed unique.

The proof is similar to that of Lemma 3.19. It is fairly straightforward to obtain (3.72) and the rest follows.

We can now find the probability that two distinct P_i's lie in the same sleeve.

Lemma 3.21. *For every $k \leq d$, as $u \to 0$,*

$$Pr\Big[\forall i \leq d : P_i \in S_t; P_1, P_2 \in S^k_{t_{u,\epsilon}}, U < u\Big] = \epsilon O(u^d) \tag{3.81}$$

Proof. Let $\mathcal{Y}$ be the event where $\forall i \leq d : P_i \in S_t$ and $P_1, P_2 \in S^k_t$. Recall that $\mathcal{H}$ is represented by $V \cdot x = 1$.

As $P_1, P_2 \in \mathcal{H}$, by Lemma 3.19,

$$|1 - \frac{P^k_1}{P^k_2}| \leq 1 + 3(d-1)\gamma_t \tag{3.82}$$

Hence by Lemma 3.18,

$$Pr\Big[\mathcal{Y}, U < u\Big] = \epsilon O(u^d) \tag{3.83}$$

completing our proof. □

Proposition 3.14 can now be proved.

Proof of Proposition 3.14. Recall that $\mathcal{NPP}$ is the event where $\mathcal{H}$ does not satisfy the permutation property. Notice that

$$Pr\Big[\mathcal{NPP}, U < u\Big] \leq Pr\Big[\mathcal{NPP}, \forall i \leq d : P_i \in S_t, U < u\Big] + Pr\Big[\mathcal{NPP}, \exists i : P_i \notin S_t, U < u\Big] \tag{3.84}$$

By Corollary 3.17,

$$Pr\Big[\mathcal{NPP}, \exists i : P_i \notin S_t, U < u\Big] \leq Pr\Big[\exists i : P_i \notin S_t, U < u\Big] = \epsilon O(u^d) \tag{3.85}$$

Now, if $\forall i \leq d : P_i \in S_t$ and $\mathcal{NPP}$ both hold, the fact that there does not exist a permutation π of $1, \ldots, d$ such that $P_m \in S^{\pi(m)}_t$ implies that there must exist $i \neq j$ and $k \leq d$ such that $P_i, P_j \in S^k_t$. Therefore,

$$\Big\{\mathcal{NPP}, \forall i \leq d : P_i \in S_t, U < u\Big\} \subset \bigcup_{i=1}^{d} \bigcup_{j \neq i} \bigcup_{k=1}^{d} \Big\{\forall m \leq d : P_m \in S_t; P_i, P_j \in S^k_t, U < u\Big\} \tag{3.86}$$

As the probability of the joined events is independent of i and j, we have

$$Pr\Big[\mathcal{NPP}, \forall i \leq d : P_i \in S_t, U < u\Big] \leq \sum_{i=1}^{d} \sum_{j \neq i} \sum_{k=1}^{d} Pr\Big[\forall m \leq d : P_m \in S_t; P_1, P_2 \in S^k_t, U < u\Big]$$

$$ \tag{3.87} $$

which implies that

$$Pr\Big[\mathcal{NPP}, \forall i \leq d : P_i \in S_t, U < u\Big] \leq d^2(d-1)Pr\Big[\forall m \leq d : P_m \in S_t; P_1, P_2 \in S_t^k, U < u\Big] \tag{3.88}$$

and so by Lemma 3.21,

$$Pr\Big[\mathcal{NPP}, \forall i \leq d : P_i \in S_t, U < u\Big] = \epsilon O(u^d) \tag{3.89}$$

which completes the proof. □

Remark 3.22. *Let π be a permutation of $\{1, \ldots, d\}$; then we define $\mathcal{L}_t^\pi$ as the following event:*

$$\mathcal{L}_t^\pi = \Big\{\forall i \leq d : P_i \in R_t^{\pi(i)}\Big\} \tag{3.90}$$

Lastly, we use the notation $\mathcal{L}_t$ whenever π is the identity permutation.

The following proposition will aid us in proving a lower bound on the expected number of facets for Proposition 3.4.

Proposition 3.23. *As $u \to 0$,*

$$Pr\Big[U < u\Big] \geq \Big(1 - O(\epsilon)\Big)u^d \tag{3.91}$$

Remark 3.24. *The notation $f(u) \geq \Big(1 - O(\epsilon)\Big)g(u)$ as $u \to 0$ stands for:*
There exist $u_0 > 0$ and $a > 0$ such that for every $u < u_0$, it holds that

$$f(u) > (1 - a\epsilon)g(u) \tag{3.92}$$

Lemma 3.25. *Let π be some permutation of $\{1, \ldots, d\}$; then, as $u \to 0$,*

$$Pr\Big[\mathcal{L}_{t_{u,\epsilon}}^\pi, U < u\Big] \geq \Big(1 - O(\epsilon)\Big)\frac{1}{d!}u^d \tag{3.93}$$

Firstly, we prove the lemma for the identity permutation. Secondly, the core idea of the proof is to choose a subset of $\Big\{U < u\Big\} \cap \mathcal{L}_t$ whose probability is easier to compute and then show that its probability is large enough. That subset will be

$$\mathcal{Y}_u = \Big\{\sum_{i=1}^{d} G\Big((1 - 3(d-1)\gamma_t)P_i^i\Big) \leq (1 - b\epsilon)u\Big\} \cap \mathcal{L}_t \tag{3.94}$$

where b is such that by Lemma 3.16 there exists u_0 so that for all $u < u_0$ we have

$$Pr\Big[Q \in \tilde{R}_t, U < u\Big] < \epsilon b u \tag{3.95}$$

We need the following lemmas to prove Lemma 3.25

Lemma 3.26. *There exists u_0 such that for all $u < u_0$ it holds that if $\mathcal{Y}_u$ holds, then $R_{t_{u,\epsilon}}$ lies entirely below $\mathcal{H}$.*

Proof. By Lemma A.7 it suffices to prove that for all $i \le d$ the i-th axis intersection of $\mathcal{H}$ is greater than t. Or, simply, for all $i \le d$ it holds that $\frac{1}{V^i} > t$. Without loss of generality $V^1 \ge V^2, \dots, V^d$. Now if $\mathcal{Y}_u$ holds then since G is strictly positive we have for every $k \le d$

$$G\Big((1 - 3(d-1)\gamma_t)P_k^k\Big) < \sum_{i=1}^{d} G\Big((1 - 3(d-1)\gamma_t)P_i^i\Big) \le (1 - b\epsilon)u < (1+\epsilon)u = G(t) \quad (3.96)$$

Thus,

$$P_k^k > (1 - 3(d-1)\gamma_t)P_k^k > t \quad (3.97)$$

Recall that $\sum_{i=1}^{d} V^i P_1^i = 1$ and $|P_1^j| \le \gamma_t t$ for all $j > 1$. Hence

$$V^1\Big(P_1^1 - (d-1)\gamma_t t\Big) \le V \cdot P_1 = 1 \quad (3.98)$$

Recall that $V \gg 0$ and thus dividing both sides by V^1 yields

$$\frac{1}{V^1} \ge P_1^1 - (d-1)\gamma_t t \quad (3.99)$$

and so by (3.97) we have

$$\frac{1}{V^1} \ge P_1^1(1 - (d-1)\gamma_t) \ge P_1^1(1 - 3(d-1)\gamma_t) > t \quad (3.100)$$

which means that R_t lies entirely below $\mathcal{H}$. □

Lemma 3.27. *Let Q be a random point in $\mathbb{R}^d$. There exist u_0 so that for all $u < u_0$, if $\mathcal{Y}_u$ holds and Q lies in $S_{t_{u,\epsilon}}$ above $\mathcal{H}$, then for all $i \le d$ we have $Q^i \ge \Big(1 - 3(d-1)\gamma_{t_{u,\epsilon}}\Big)P_i^i$.*

Proof. Consider the case where there exists some point $\hat{Q} \in S_t^i \cap \mathcal{H}$ such that $\hat{Q}^i = Q^i$; then, by Lemma 3.19,

$$|1 - \frac{\hat{Q}^i}{P_i^i}| \le 3(d-1)\gamma_t \quad (3.101)$$

which implies that

$$Q^i = \hat{Q}^i > \Big(1 - 3(d-1)\gamma_t\Big)P_i^i \quad (3.102)$$

Now consider the case where Q^i is less than $\hat{Q}^i$ for all $\hat{Q} \in S_t^i \cap \mathcal{H}$. By (3.74),

$$Q^i < \frac{1}{V^i}\Big(1 - \gamma_t t \sum_{j \ne i} V^j\Big) \quad (3.103)$$

Notice that

$$V \cdot Q = \sum_{j=1}^{d} V^j Q^j \leq V^i Q^i + \sum_{j \neq i} V^j \gamma_t t \tag{3.104}$$

Thus,

$$V \cdot Q < V^i \frac{1}{V^i}\Big(1 - \gamma_t t \sum_{j \neq i} V^j\Big) + \gamma_t t \sum_{j \neq i} V^j \tag{3.105}$$

or simply

$$V \cdot Q < 1 \tag{3.106}$$

which means that Q lies below $\mathcal{H}$.

Lastly, we remain with the case where Q^i is greater than $\hat{Q}^i$ for all $\hat{Q} \in S_t^i \cap \mathcal{H}$. This implies that if Q lies above $\mathcal{H}$ and it is inside S_t^i, then we have

$$Q^i \geq P_i^i \geq \Big(1 - 3(d-1)\gamma_t\Big) P_i^i \tag{3.107}$$

□

Lemma 3.28. *As* $u \to 0$*,*

$$\mathcal{Y}_u \subset \Big\{U < u\Big\} \tag{3.108}$$

Proof. If $\mathcal{Y}_u$ holds by Lemma 3.26 for all $u < u_0$, we have

$$U = Pr\Big[V \cdot Q > 1 \mid P_1, \ldots, P_d\Big] \leq Pr\Big[Q \in \tilde{R}_t\Big] + \sum_{i=1}^{d} Pr\Big[Q \in S_t^i, \mathcal{Y}_u\Big] \tag{3.109}$$

and so, by Lemma 3.27,

$$U \leq Pr\Big[Q \in \tilde{R}_t\Big] + \sum_{i=1}^{d} Pr\Big[Q^i \geq (1 - 3(d-1)\gamma_t) P_i^i\Big] \tag{3.110}$$

Furthermore,

$$\sum_{i=1}^{d} Pr\Big[Q^i \geq (1 - 3(d-1)\gamma_t) P_i^i\Big] = \sum_{i=1}^{d} G\Big((1 - 3(d-1)\gamma_t) P_i^i\Big) \tag{3.111}$$

Since $\mathcal{Y}_u$ holds

$$\sum_{i=1}^{d} G\Big((1 - 3(d-1)\gamma_t) P_i^i\Big) < (1 - b\epsilon) u \tag{3.112}$$

whereas, by Lemma 3.16,

$$Pr\Big[Q \in \tilde{R}_t\Big] = \epsilon O(u) \tag{3.113}$$

Put differently, for every $u < u_1 < u_0$ we have

$$Pr\Big[Q \in \tilde{R}_t\Big] < \epsilon b u \tag{3.114}$$

and so

$$U < \Big(b\epsilon + (1 - b\epsilon)\Big)u = u \tag{3.115}$$

In conclusion, that for every $u < u_1$

$$\mathcal{Y}_u \subset \Big\{U < u\Big\} \tag{3.116}$$

□

Lemma 3.29.

$$Pr\Big[\mathcal{Y}_u\Big] \geq \Big(1 - O(\epsilon)\Big)\frac{1}{d!}u^d \tag{3.117}$$

Proof. Define the event

$$\mathcal{Z}_u = \Big\{\sum_{i=1}^{d} G\Big((1 - 3(d-1)\gamma_t)P_i^i\Big) \leq (1 - b\epsilon)u\Big\} \cap \neg\mathcal{L}_t \tag{3.118}$$

Hence

$$Pr\Big[\mathcal{Y}_u\Big] = Pr\Big[\sum_{i=1}^{d} G\Big((1 - 3(d-1)\gamma_t)P_i^i\Big) \leq (1 - b\epsilon)u\Big] - Pr\Big[\mathcal{Z}_u\Big] \tag{3.119}$$

For the left summand,

$$Pr\Big[\sum_{i=1}^{d} G\Big((1 - 3(d-1)\gamma_t)P_i^i\Big) \leq (1 - b\epsilon)u\Big] \leq Pr\Big[\sum_{i=1}^{d} G(P_i^i) \leq (1 - b\epsilon)u\Big] \tag{3.120}$$

and so, by Lemma A.5,

$$Pr\Big[\sum_{i=1}^{d} G(P_i^i) \leq (1 - b\epsilon)u\Big] = \frac{1}{d!}\Big((1 - b\epsilon)u\Big)^d = \frac{1}{d!}u^d(1 - b\epsilon)^d \tag{3.121}$$

Notice that there exists $a > 0$ and ϵ_0 such that for every $\epsilon < \epsilon_0$ it holds that

$$(1 - b\epsilon)^d > 1 - a\epsilon \tag{3.122}$$

Thus,

$$Pr\Big[\sum_{i=1}^{d} G(P_i^i) \leq (1-b\epsilon)u\Big] = \Big(1 - O(\epsilon)\Big)\frac{1}{d!}u^d \tag{3.123}$$

and so

$$Pr\Big[\mathcal{Y}_u\Big] - Pr\Big[\mathcal{Z}_u\Big] = \Big(1 - O(\epsilon)\Big)\frac{1}{d!}u^d \tag{3.124}$$

As for the right summand in (3.119), since $P_i^i > (1 - 3(d-1)\gamma_t)P_i^i > t$ for all $i \leq d$, if $\mathcal{Z}_u$ holds, $\neg\mathcal{L}_t$ merely translates to $\exists i \leq d : P_i \in \tilde{R}_t$. Notice that for all $k \leq d$ we have

$$\Big\{\sum_{i=1}^{d} G\Big((1-3(d-1)\gamma_t)P_i^i\Big) \leq (1-b\epsilon)u\Big\} \subset \Big\{\sum_{i\neq k} G\Big((1-3(d-1)\gamma_t)P_i^i\Big) \leq (1-b\epsilon)u\Big\} \tag{3.125}$$

Furthermore, the right-hand side is independent of the event $\{P_k \in \tilde{R}_t\}$. Thus,

$$Pr\Big[\mathcal{Z}_u\Big] \leq \sum_{k=1}^{d} Pr\Big[P_k \in \tilde{R}_t\Big] Pr\Big[\sum_{i\neq k} G\Big((1-3(d-1)\gamma_t)P_i^i\Big) \leq (1-b\epsilon)u\Big] \tag{3.126}$$

and so, by Lemma 3.16,

$$Pr\Big[\mathcal{Z}_u\Big] \leq d\epsilon O(u) Pr\Big[\sum_{i\neq k} G\Big((1-3(d-1)\gamma_t)P_i^i\Big) \leq (1-b\epsilon)u\Big] \tag{3.127}$$

Now, by Lemma A.5,

$$Pr\Big[\mathcal{Z}_u\Big] \leq d\epsilon O(u)\frac{1}{(d-1)!}(1-b\epsilon)^{d-1}u^{d-1} \tag{3.128}$$

Notice that

$$(1-b\epsilon)^{d-1} = 1 - O(\epsilon) \tag{3.129}$$

and so

$$Pr\Big[\mathcal{Z}_u\Big] \leq \epsilon O(u^d) \tag{3.130}$$

and so, by (3.124), the proof is completed. □

Proof of Lemma 3.25. Recall that by the definition of $\mathcal{Y}_u$

$$\mathcal{Y} \subset \mathcal{L}_t \tag{3.131}$$

Thus, by Lemma 3.28,

$$\mathcal{Y} \subset \mathcal{L}_t \cap \Big\{U < u\Big\} \tag{3.132}$$

and so, by Lemma 3.29, the proof is completed. □

Proof of Proposition 3.23. To begin with,

$$Pr\Big[U < u\Big] = Pr\Big[U < u, \exists\pi : \mathcal{L}_t^\pi\Big] + Pr\Big[U < u, \nexists\pi : \mathcal{L}_t^\pi\Big] \tag{3.133}$$

By Lemma 3.25,

$$Pr\Big[U < u, \exists\pi : \mathcal{L}_t^\pi\Big] = \Big(1 - O(\epsilon)\Big)d!\frac{1}{d!}u^d \tag{3.134}$$

For all t large enough such that $\gamma_t < 1$, the events $\mathcal{L}_t^\pi$ are mutually disjoint for all permutations. Thus,

$$Pr\Big[U < u\Big] \geq \Big(1 - O(\epsilon)\Big)u^d \tag{3.135}$$

□

The following proposition, taken together with the above, will enable us to prove Theorem 3.1, it remains to prove the following proposition.

Proposition 3.30. *Let π be a permutation of $\{1, \ldots, d\}$; then, as $u \to 0$,*

$$Pr\Big[\mathcal{L}_{t_{u,\epsilon}}^\pi, U < u\Big] \leq \Big(1 + O(\epsilon)\Big)\frac{1}{d!}u^d \tag{3.136}$$

Remark 3.31. *The notation $f(u) \leq \Big(1 + O(\epsilon)\Big)g(u)$ as $u \to 0$ stands for:*
There exist $u_0 > 0$ and $a > 0$ such that for every $u < u_0$ it holds that

$$f(u) < \Big(1 + a\epsilon\Big)g(u) \tag{3.137}$$

Before delving into the proof of Proposition 3.30, we use it to prove Theorem 3.1.

Proof of Theorem 3.1. Taking summation over the group of permutations of $\{1, \ldots, d\}$,

$$Pr\Big[U < u\Big] = Pr\Big[U < u, \exists\pi : \mathcal{L}_t^\pi\Big] + Pr\Big[U < u, \nexists\pi : \mathcal{L}_t^\pi\Big] \tag{3.138}$$

Proposition 3.30 and Proposition 3.14 together yield

$$Pr\Big[U < u\Big] \leq \Big(1 + O(\epsilon)\Big)d!\frac{1}{d!}u^d = \Big(1 + O(\epsilon)\Big)u^d \tag{3.139}$$

whereas, by Proposition 3.23,

$$Pr\Big[U < u\Big] = \Big(1 - O(\epsilon)\Big)u^d \tag{3.140}$$

Plugging $\alpha = 1$ into Proposition 3.11 and its corollary completes the proof. □

Lemma 3.32. *Let* $\theta \in (0,1), 0 \ll v \in \mathbb{R}^d$*; then*

$$Pr\Big[v \cdot Q \geq 1\Big] \geq \sum_{i=1}^{d} G\Big((1+(d-1)\theta)\frac{1}{v^i}\Big) \prod_{j \neq i} H(\theta \frac{1}{v^j}) \tag{3.141}$$

Proof. We have

$$\bigcup_{i=1}^{d} \Big\{Q^i \geq (1+(d-1)\theta)\frac{1}{v^i}, \forall j \neq i : |Q^j| \leq \theta \frac{1}{v^j}\Big\} \subset \Big\{v \cdot Q \geq 1\Big\} \tag{3.142}$$

The above union is disjoint and we have

$$Pr\Big[v \cdot Q \geq 1\Big] \geq \sum_{i=1}^{d} Pr\Big[Q^i \geq (1+(d-1)\theta)\frac{1}{v^i}, \forall j \neq i : |Q^j| \leq \theta \frac{1}{v^j}\Big] \tag{3.143}$$

Clearly,

$$Pr\Big[v \cdot Q \geq 1\Big] \geq \sum_{i=1}^{d} G\Big((1+(d-1)\theta)\frac{1}{v^i}\Big) \prod_{j \neq i} H(\theta \frac{1}{v^j}) \tag{3.144}$$

which completes the proof. □

We are now ready to prove Proposition 3.30.

Proof of Proposition 3.30. Without loss of generality let π be the identity permutation of $\{1, \ldots, d\}$. Recall that $\mathcal{H}$ is represented by $V \cdot x = 1$. Pick u_0 such that for all $u < u_0$ by Lemma 3.19 we get, for all $i \leq d$,

$$|1 - \frac{P_i^i}{(\frac{1}{V^i})}| \leq 3(d-1)\gamma_t \tag{3.145}$$

Hence,

$$\frac{1}{V^i} < \Big(1 + 3(d-1)\gamma_t\Big)P_i^i \tag{3.146}$$

Notice that by Lemma 3.32,

$$U = Pr\Big[Q \cdot V \geq 1\Big] \geq \sum_{i=1}^{d} G\Big((1+(d-1)\gamma_t)\frac{1}{V^i}\Big) \prod_{j \neq i} H(\gamma_t \frac{1}{V^j}) \tag{3.147}$$

Moreover, pick $u_1 < u_0$ so that, by Lemma 3.8, for all $u < u_1$ and $i \leq d$,

$$\frac{1}{V^i} > t \tag{3.148}$$

Thus,

$$U \geq \sum_{i=1}^{d} G\Big((1+(d-1)\gamma_t)\frac{1}{V^i}\Big)\prod_{j\neq i} H(\gamma_t t) \tag{3.149}$$

By (3.146),

$$U \geq \sum_{i=1}^{d} G\Big((1+3(d-1)\gamma_t)(1+(d-1)\gamma_t)P_i^i\Big)\prod_{j\neq i} H(\gamma_t t) \tag{3.150}$$

which further translates to

$$U \geq \sum_{i=1}^{d} G\Big((1+4(d-1)\gamma_t+3(d-1)^2\gamma_t^2)P_i^i\Big)\prod_{j\neq i} H(\gamma_t t) \tag{3.151}$$

Denote

$$\alpha = \sum_{i=1}^{d} G\Big((1+4(d-1)\gamma_t+3(d-1)^2\gamma_t^2)P_i^i\Big) \tag{3.152}$$

and also

$$\rho_t = (4d-3)\gamma_t \tag{3.153}$$

By Lemma 2.2, pick $t_2 > t_1$ (where t_1 corresponds to u_1), so that for all $t > t_2$

$$3(d-1)^2\gamma_t^2 < \gamma_t \tag{3.154}$$

Hence, for all $t > t_2$,

$$\alpha > \sum_{i=1}^{d} G\Big((1+\rho_t)P_i^i\Big) \tag{3.155}$$

Furthermore, again by Lemma 2.2, pick $t_3 > t_2$ so that for all $t > t_3$,

$$\alpha > \frac{1}{\sqrt{1+\epsilon}}\sum_{i=1}^{d} G(P_i^i)^{1+\rho_t} \tag{3.156}$$

Thus,

$$U > \frac{1}{\sqrt{1+\epsilon}}\sum_{i=1}^{d} G(P_i^i)^{1+\rho_t}\prod_{j\neq i} H(\gamma_t t) \tag{3.157}$$

By Hölder's inequality corollary (Lemma A.4), toggether with $q = 1+\rho_t$, $p = 1$, we have

$$\Big(\sum_{i=1}^{d} G(P_i^i)\Big)\Big(\sum_{i=1}^{d} G(P_i^i)^{1+\rho_t}\Big)^{\frac{-1}{1+\rho_t}} \leq d^{1-\frac{1}{1+\rho_t}} \tag{3.158}$$

and so

$$d^{-\rho_t}\Big(\sum_{i=1}^{d} G(P_i^i)\Big)^{1+\rho_t} \le \sum_{i=1}^{d} G(P_i^i)^{1+\rho_t} \tag{3.159}$$

Hence,

$$U > \frac{d^{-\rho_t}}{\sqrt{1+\epsilon}}\Big(\sum_{i=1}^{d} G(P_i^i)\Big)^{1+\rho_t} \prod_{j\neq i} H(\gamma_t t) \tag{3.160}$$

Pick $t_4 > t_3$ so that for every $t > t_4$ it holds that

$$d^{-\rho_t} \prod_{j\neq i} H(\gamma_t t) > \frac{1}{\sqrt{1+\epsilon}} \tag{3.161}$$

Therefore,

$$U > \frac{1}{1+\epsilon}\Big(\sum_{i=1}^{d} G(P_i^i)\Big)^{1+\rho_t} \tag{3.162}$$

Recall that $G(t) = (1+\epsilon)u$ and $U < u$, and so

$$G(t) > \Big(\sum_{i=1}^{d} G(P_i^i)\Big)^{1+\rho_t} \tag{3.163}$$

or simply

$$G(t)^{\frac{1}{1+\rho_t}} > \sum_{i=1}^{d} G(P_i^i) \tag{3.164}$$

By Lemma A.5, with $s = G(t)^{\frac{1}{1+\rho_t}}$,

$$Pr\Big[\mathcal{L}_t, U < u\Big] \le Pr\Big[\sum_{i=1}^{d} G(P_i^i) \le G(t)^{\frac{1}{1+\rho_t}}\Big] \le \frac{1}{d!} G(t)^{\frac{d}{1+\rho_t}} \tag{3.165}$$

By Lemma 2.2, pick $t_5 > t_4$ so that for all $t > t_5$ we have

$$G(t)^{\frac{d}{1+\rho_t}} < (1+\epsilon) G(t)^d \tag{3.166}$$

Therefore,

$$Pr\Big[\mathcal{L}_t, U < u\Big] < \frac{1+\epsilon}{d!} G(t)^d = \frac{1}{(1+\epsilon)^{d-1} d!} u^d \tag{3.167}$$

Notice that there exist $b > 0$ and ϵ_0 such that for all $\epsilon < \epsilon_0$ we have

$$\frac{1}{(1+\epsilon)^{d-1}} = \frac{1}{1 + \sum_{k=1}^{d-1} \binom{d-1}{k} \epsilon^k} < 1 + b\epsilon \tag{3.168}$$

which completes our proof. □

Section 4

Evolutionarily Stable Strategies

4.1 Preliminaries

Recall that a game $R \in \mathbb{R}^{n\times n}$ is called *symmetric* if, for every choice of two pure strategies i, j, the payoffs for the row and column players are $R(i, j)$ and $R(j, i)$, respectively.

Definition 4.1. *Let $R \in \mathbb{R}^{n\times n}$ be a symmetric game. The strategy $p \in \triangle(n) \subset \mathbb{R}^n$ is called an **evolutionarily stable strategy** if the following conditions hold:*

$$\forall q \in \triangle(n) \qquad :R(p, p) \geq R(q, p) \tag{4.1}$$

$$\forall p \neq q \in \triangle(n) \qquad :R(q, p) = R(p, p) \Rightarrow R(q, q) < R(p, q) \tag{4.2}$$

This definition is equivalent to the requirement that for every $q \neq p$, there exist an ”invasion barrier” $b(q) > 0$ such that

$$\forall \epsilon \in \big(0, b(q)\big) : R(p, (1-\epsilon)p + \epsilon q) > R(q, (1-\epsilon)p + \epsilon q) \tag{4.3}$$

The interpretation of this inequality is that any small enough proportion ϵ of q-mutants cannot successfully invade a p-population, since every q-mutant average payoff is strictly less than that of the existing population.

Throughout this section, we regard a symmetric game $R \in \mathbb{R}^{n\times n}$ whose entries are i.i.d. according to a symmetrical continuous subexponential distribution function F, whose tail is denoted by G. Moreover, in order to induce a geometrical viewpoint, we denote $R(i, j) = X_i^j$.

Lemma 4.2. *The following conditions are equivalent for $p \in \triangle(n)$ with $supp(p) = \big\{1, \ldots, d\big\}$:*

1. *p is a Nash equilibrium.*

2. *For every $1 \leq i, j \leq d$ and $k > d$ we have $R(i, p) = R(j, p) \geq R(k, p)$.*

3. *The set of points* $\{(X_j^1, \ldots, X_j^d)\}_{j=1}^d \subset \mathbb{R}^d$ *form a hyperplane* $\mathcal{H}$ *that lies above, as defined in (3.3), the points* $\{(X_j^1, \ldots, X_j^d)\}_{j>d}$.

The proof is outlined in A.6 in the Appendix.

Remark 4.3. *In the case that the conditions above hold, since* R *is a game whose entries are generated by a continuous distribution with probability* 1*, the inequality above is strict and for all* $j \notin supp(p)$*, and the point* $(X_j^{i_1}, \ldots, X_j^{i_d})$ *lies strictly below** $\mathcal{H}$*. Moreover,* p *is the normal of* $\mathcal{H}$*.*

The following characterization of ESS will play a key role in what follows.

Lemma 4.4. *Let* $p \in \triangle(n)$ *with* $supp(p) = \{1, \ldots, d\}$*; then* p *is an ESS if and only if*

$$\forall 0 \neq r \in \mathbb{R}^n, supp(r) \subset supp(p), \sum_{i=1}^{d} r_i = 0 : rRr^T < 0 \tag{4.4}$$

Proof. Assume p is an ESS, let r satisfy the conditions in the lemma. It suffices to prove that $(\epsilon r)R(\epsilon r)^T < 0$ for some $\epsilon > 0$. Therefore, we may assume that $p + r \in \triangle(n)$. Let $q := p + r$.

Clearly, $supp(q) \subset supp(p)$, and since p is a Nash equilibrium it holds that $R(q, p) = R(p, p)$. Since p is an ESS is follows that $R(q, q) < R(p, q)$, or, simply ,

$$R(q, q) - R(p, q) = (q - p)Rq^T < 0 \tag{4.5}$$

Recall that $q = p + r$; hence

$$rR(p + r)^T < 0 \tag{4.6}$$

Moreover, denote by e^i the i-th unit vector. Since p is a Nash equilibrium, we know that

$$Rp^T \in \text{Span}\{e^i\}_{i=1}^d \tag{4.7}$$

Since $\sum_{i=1}^d r_i = 0$ we get $rRp^T = 0$, and hence $rRr^T < 0$.

Suppose on the contrary that p is a Nash equilibrium satisfying the conditions mentioned in the lemma. Let $q \in \triangle(n)$ such that $supp(q) \subset supp(p)$. Choose

$$r = p - q \tag{4.8}$$

to get

$$(p - q)R(p - q)^T < 0 \tag{4.9}$$

*The point $Q \in \mathbb{R}^d$ *lies strictly below* $\mathcal{H}$, if $p \cdot Q < 1$.

Notice that

$$(p-q)R(p-q)^T = R(p,p) - R(q,p) - R(p,q) + R(q,q) \tag{4.10}$$

However, $R(q,p) = R(p,p)$, and hence $R(q,q) < R(p,q)$. □

Remark 4.5. *The previous lemma is essentially given in van Damme [3] (page 221).*

4.2 Expected Number of ESS

It was proven by Hart, Rinott and Weiss in [6] that the expected number of two-point ESS converges to $\frac{1}{2}$ as n approaches infinity. We consider the d-dimensional case for $d \geq 4$, and attempt to prove that in such a case the result is 0.

Denote by $\mathcal{N}_d^n$ the expected number of ESS with support of size d in an n-dimensional symmetric game R.

Theorem 4.6.

$$\forall d \geq 4 : \lim_{n\to\infty} \mathcal{N}_d^n = 0 \tag{4.11}$$

Remark 4.7. *We use the same notations as in the previous section, where $p_i^j = X_i^j$ and $P_i = (p_i^1, \ldots, p_i^d)$. Henceforth, we assume $supp(p) = \{1, \ldots, d\}$, and denote by $\mathcal{H}$ the hyperplane generated by the points $\{(X_j^1, \ldots, X_j^d)\}_{j \leq d}$. Again by Lemma A.1, $\mathcal{H}$ is almost surely uniquely determined. Now, if p is a Nash equilibrium, then it is simply the normal of $\mathcal{H}$; that is, $x \cdot p = c$ is the algebraic expression representing $\mathcal{H}$. Recall that U is roughly the probability for a random point to lie above $\mathcal{H}$.*

Remark 4.8. *Denote by $\mathcal{L}_t^\pi$ the event where $\mathcal{H}$ satisfies the permutation property, with the corresponding permutation π. Moreover, denote by $\mathcal{ESS}$ the event that $\{1, \ldots, d\}$ generate an ESS.*

Remark 4.9. *Lastly, we say that t corresponds to w whenever $G(t) = (1+\epsilon)w$.*

To prove Theorem 4.6, we define a new random variable, similar to U of the previous section. Recall that $\mathcal{ESS}$ is the event where the points $P_1, \ldots, P_d$ generate an ESS. Define

$$W := \begin{cases} Pr\Big[P_{d+1} \cdot V > 1 \mid P_1, \ldots, P_d\Big] & : V \gg 0, \mathcal{ESS} \\ 1 & : \text{otherwise} \end{cases}$$

Recall that $V \gg 0$ is the event where $P_1, \ldots, P_d$ determine a unique hyperplane that does not pass through the origin whose normal is strictly positive, as outlined in Remark 3.3.

Moreover, notice that for every $w < 1$, if $W < w$ then the points $P_1, \ldots, P_d$ generate an ESS.

Clearly, for all $P_1, \ldots, P_d$ we have $U \leq W$, which implies that for all $u > 0$,

$$\left\{W < u\right\} \subset \left\{U < u\right\} \tag{4.12}$$

and similarly

$$F_W(u) \leq F_U(u) \tag{4.13}$$

Remark 4.10. *Whenever $\mathcal{ESS}$ holds we denote by $p \in \triangle(n)$ the normalized version of V; that is, $\sum_{i=1}^{d} p^i = 1$ and for all $i \leq n$ we have $p^i \geq 0$.*

The following lemma is similar to Lemma 3.6.

Lemma 4.11.

$$Pr\left[\mathcal{ESS}\right] = (n-d)\int_0^1 (1-w)^{n-d-1} F_W(w)\mathrm{d}w \tag{4.14}$$

Proof. Denote by $\mathcal{Y}$ the event where:

- For all $i > d$ it holds that $V \cdot P_i < 1$.
- $V \gg 0$.
- $\mathcal{ESS}$ holds.

Notice that if any of the last two conditions does not hold we have $W = 1$, which implies that

$$Pr\left[\mathcal{Y} \mid P_1, \ldots, P_d\right] = (1-W)^{n-d} \tag{4.15}$$

Recall that if $\mathcal{ESS}$ holds then p is also a Nash equilibrium. Thus, by Lemma 4.2, $\mathcal{H}$ contains a facet of the convex hull of $P_1, \ldots, P_n$, and since the probability that $\mathcal{H}$ contains the origin is zero we have

$$Pr\left[\mathcal{ESS}\right] = Pr\left[\mathcal{Y}\right] = \int_{w=0}^{1} (1-w)^{n-d}\mathrm{d}F_W(w) \tag{4.16}$$

The rest of the proof is identical to that of Lemma 3.6. □

We need to reformulate Proposition 3.11 for our current needs.

Lemma 4.12. *Let $\alpha > 0$ and $w_0 > 0$ such that $F_W(w) < \alpha w^d$ for all $0 < w < u_0$; then*

$$\lim_{n\to\infty} \mathbb{E}[\mathcal{N}_d^n] \leq \alpha \tag{4.17}$$

Proof. Using 4.11 now allows us to use similar arguments to those in the proof of Proposition 3.11. □

Proposition 4.13. *Let π be a permutation of $\{1, \dots, d\}$; then, as $w \to 0$,*

$$\forall d \geq 4 : Pr\Big[\mathcal{L}_t^\pi, W < w\Big] = \epsilon O(w^d) \tag{4.18}$$

Proof. Let $\epsilon > 0$, and define $\mathcal{Y}_t$

$$\mathcal{Y}_t = \mathcal{L}_t^\pi \cap \Big\{W < w\Big\} \tag{4.19}$$

where w corresponds to t. Furthermore, choose w_0 so that by Lemma 3.8, for all $W < w < w_0$ and $i \leq d$, the i-th axis intersection† of $\mathcal{H}$ is larger than t. Finally, chosse $t_1 > t_0$ (where t_0 corresponds to w_0) so that for every $t > t_1$ it holds that $\gamma_t < \frac{1}{4d^4}$. We deal with four cases of the permutation π separately.

1. **There exists i such that $\pi(i) = i$.**
   Whenever $\mathcal{Y}_t$ holds and $\gamma_t < 1$:
   $$\begin{aligned} R(i,i) &= X_i^i \\ &= p^i X_i^i + (1-p^i)X_i^i \\ &> p^i X_i^i + (1-p^i)\gamma_t t \\ &= p^i X_i^i + \sum_{j\neq i} p^j \gamma_t t \\ &\geq p^i X_i^i + \sum_{j\neq i} p^j X_j^i \\ &\geq R(p,i) \end{aligned}$$
   which implies that $\neg\mathcal{ESS}$, and $W = 1$.

2. **π has two disjoint non trivial cycles**
   Let $l, k > 1$ be such that $(i_1 \ \dots \ i_k), (j_1 \ \dots \ j_l)$ are two disjoint cycles of π. Choose

   $$r = \sum_{s=1}^{k} e^{i_s} - k e^{j_1} \tag{4.20}$$

   Finally, define $i_{k+1} = i_1$ and recall that $\mathcal{L}_t^\pi$ implies that

   $$\forall i, j \neq \pi(i) : |X_i^j| \leq \gamma_t t \tag{4.21}$$

† The i-th axis intersection of the hyperplane $\mathcal{H}$ is the unique point $\hat{Q} \in Span\Big\{e^i\Big\} \cap \mathcal{H}$. As $\mathcal{H}$ has a strictly positive normal, that point is indeed unique.

Notice that

$$rRr^T = \sum_{s=1}^{d} \sum_{q=1}^{d} r^s r^q X_s^q \tag{4.22}$$

For all $x \le k$, we have $X_{i_s}^{i_{s+1}} > t$, and recall that since $\gamma_t < \frac{1}{4d^4} < \frac{1}{d^2k^2}$ we have

$$rRr^T > \sum_{s=1}^{k} X_{i_s}^{i_{s+1}} - d^2k^2\gamma_t t > kt - t > 0 \tag{4.23}$$

which implies that $\neg\mathcal{ESS}$ and $W = 1$.

3. **$d = 4$, and π has a unique cycle of length 4.**
   Without loss of generality $\pi = (2\ 3\ 4\ 1)$. Notice that we are free to label the coordinates as we wish, as long as the labels do not repeat themselves. Moreover, if $\mathcal{L}_t^\pi$ holds, then each P_i is found in some S_t^j. By our labeling freedom and without loss of generality, we have

$$X_1^2 > X_4^1 > X_3^4 > X_2^3 \tag{4.24}$$

Define

$$r = (e^1 + e^2 - e^3 - e^4) \tag{4.25}$$

A quick computation shows that

$$rRr^T < 0 \Rightarrow X_1^2 + X_3^4 < X_2^3 + X_4^1 + d^2\gamma_t t \tag{4.26}$$

For every $W < w < 1$, we know that $P_1, \ldots, P_4$ generate an ESS and so (4.26) holds. Divide (4.26) by X_2^3 to get

$$\frac{X_1^2 + X_3^4}{X_2^3} < 1 + \frac{X_4^1 + d^2\gamma_t t}{X_2^3} \tag{4.27}$$

Subtracting $\frac{X_1^2}{X_2^3}$ yields

$$\frac{X_3^4}{X_2^3} < 1 + \frac{X_4^1 - X_1^2 + d^2\gamma_t t}{X_2^3} \tag{4.28}$$

Since $X_1^2 > X_4^1$ and $X_3^4 > X_2^3$, we have

$$1 < \frac{X_3^4}{X_2^3} < 1 + \frac{d^2\gamma_t t}{X_2^3} \tag{4.29}$$

Lastly, as $X_2^3 > t$ we have

$$1 < \frac{X_3^4}{X_2^3} < 1 + d^2\gamma_t \tag{4.30}$$

Denote

$$\alpha = Pr\Big[U < w, \forall i \leq d : P_i \in S_t; 1 < \frac{X_3^4}{X_2^3} < 1 + d^2\gamma_t\Big] \tag{4.31}$$

Note that

$$Pr\Big[W < w, \mathcal{L}_t^\pi, 1 < \frac{X_3^4}{X_2^3} < 1 + d^2\gamma_t\Big] \leq \alpha \tag{4.32}$$

As in the notations of the previous section,

$$\mathcal{L}_t^\pi \subset \Big\{\forall i \leq d : P_i \in S_t\Big\} \tag{4.33}$$

By Lemma 3.18,

$$\alpha \leq \epsilon O(w^d) \tag{4.34}$$

Thus,

$$Pr\Big[W < w, \mathcal{L}_t^\pi, \frac{X_3^4}{X_2^3} < 1 + d^2\gamma_t\Big] \leq \epsilon O(w^d) \tag{4.35}$$

4. **$d \geq 5$ and π has a cycle of length d.**
   Without loss of generality $\pi = (2 \ \dots \ d \ 1)$. Pick $r = e^1 + e^2 - 2e^4$. A simple calculation shows that

   $$rRr^T \geq (X_1^2 - 4d^2\gamma_t t) > 0 \tag{4.36}$$

   Recall that $\gamma_t < \frac{1}{4d^4}$, which implies that $\neg\mathcal{ESS}$ holds.

□

We are ready for the main theorem.

Proof of Theorem 4.6. Notice that,

$$Pr\Big[W < w\Big] = Pr\Big[W < w, \exists\pi : \mathcal{L}_t^\pi\Big] + Pr\Big[W < w, \nexists\pi : \mathcal{L}_t^\pi\Big] \tag{4.37}$$

By Proposition 3.14,

$$Pr\Big[U < u, \nexists\pi : \mathcal{L}_t^\pi\Big] = \epsilon O(u^d) \tag{4.38}$$

Since $U \leq W$,

$$Pr\Big[W < u, \nexists\pi : \mathcal{L}_t^\pi\Big] \leq Pr\Big[U < u, \nexists\pi : \mathcal{L}_t^\pi\Big] \tag{4.39}$$

which implies that

$$Pr\Big[W < w, \nexists\pi : \mathcal{L}_t^\pi\Big] = \epsilon O(w^d) \tag{4.40}$$

By Proposition 4.13, for every permutation π we have

$$Pr\Big[W < w, \mathcal{L}_t^\pi\Big] = \epsilon O(w^d) \tag{4.41}$$

Thus,

$$Pr\Big[W < w, \exists\pi : \mathcal{L}_t^\pi\Big] = \epsilon d! O(w^d) = \epsilon O(w^d) \tag{4.42}$$

In conclusion,

$$Pr\Big[W < w\Big] = \epsilon O(w^d) \tag{4.43}$$

By Lemma 4.12 together with $\alpha = \epsilon$, the proof is completed. □

Section 5

Discussion

In this section we discuss some of the related literature, and end with a number of comments, conjectures, and open problems.

- **Facets and equilibria.** Bárány, Vempala and Vetta [1] used the connection between Nash equilibria and facets of random polytopes to find Nash equilibria in random games. In using this connection to find ESS, we emphasize that the number of facets of a random polytope and the number of d-point ESS of a random game have different distributions. Accordingly, this paper only considers the expected number of d-point ESS (of those facets) as the number of strategies (points that generate the convex hull) approaches infinity.

- **The case of $d = 3$.** The result implies that the expected number of three-point ESS cannot converge to a number larger than $\frac{1}{3}$. This is due to Proposition 3.14, which says that "almost every" ESS should has the permutation property. Moreover, Lemma 4.4 shows that only derangements can generate an ESS for subexponential distributions (see the proof of Theorem 4.6). A preliminary informal analysis suggests the following conjecture:
$$\lim_{n\to\infty} \mathcal{E}_3^n \in [0, \frac{1}{3}] \tag{5.1}$$
where there exist distributions so that the above interval is tight. Furthermore, the distribution for the number of three-point ESS is $\leq Poisson(\frac{1}{3})$. See Devroye [4] for the oscillation of the expected number of extreme points of a random set.

- **The number of any ESS.** Extrapolating from the above conjecture and the results of Hart, Rinott and Weiss (HRW for short) [6], we conclude the distribution for the number of any ESS is bounded by:
$$Poisson(1 + \frac{1}{2} + \frac{1}{3}) = Poisson(\frac{11}{6}) \tag{5.2}$$

Thus,

$$Pr\Big[\text{there exists an ESS}\Big] \leq 1 - e^{-\frac{11}{6}} \cong 0.84 \tag{5.3}$$

- **Other families of distributions.** A larger family of distributions was investigated by HRW [6]. They were able to show that the expected number of facets of a random polygon in the plane converges to infinity as the number of points increases. Moreover, the authors were able to show that their result implies that the expected number of two-point ESS for such distributions behaves in the same manner. We suspect that there exist distributions for which the expected number of facets and d-point ESS both converge to infinity. Approaching the problem from a different angle, we may ask what happens when there are distinct distributions for each column, or for each coordinate (of the random points that generate the convex hull).

Appendix

Lemma A.1. *Let $k \leq d$, let F be a continuous distribution, and let $v_1, \ldots, v_k$ be vectors in $\in \mathbb{R}^d$. Moreover, assume the coordinates of the above vectors are i.i.d.-F. Thus, the above vectors are almost surely independent.*

Proof. Let $A \in \mathbb{R}^{d \times d}$ where $\forall i, j \leq n : a_{ij}$ are i.i.d.-F. Then it suffices to prove that $Pr\Big[|A| = 0\Big] = 0$. We do so by induction on d. The base case follows from the continuity of F. Now, assume the lemma golds for $d - 1$. Let $S(d)$ be the group of d-permutations. Recall that

$$|A| = \sum_{\sigma \in S(d)} \text{sign}(\sigma) \prod_{i=1}^{d} a_{i,\sigma(i)} \tag{A.4}$$

Let A_{11} be the first principal minor of A, and denote by v_i^j the j-th coordinate of v_i. Rewriting the above yields

$$|A| = 0 \iff v_1^1 |A_{11}| = - \sum_{\sigma \in S(d):\sigma(1) \neq 1} \prod_{i=1}^{d} a_{i,\sigma(i)} \tag{A.5}$$

By the induction hypothesis, $|A_{11}| \neq 0$ almost surely. Hence, for every value of the right-hand side of (A.5), there exists only a single value of v_1^1 such that $|A| = 0$. In conclusion, $|A| \neq 0$ almost surely. □

Lemma A.2. *Let $P_1, \ldots, P_d$ be points in $\in \mathbb{R}^d$. Moreover, assume that the coordinates of these points are i.i.d.-F for a continuous distribution F. Thus, the above vectors almost surely generate a unique hyperplane that does not pass through the origin.*

Proof. Define $V_i := P_1 - P_i$ for $i \geq 2$. By Lemma A.1, these $d-1$ vectors are almost surely independent. Whenever that is the case there exists a hyperplane $\mathcal{H}$ containing P_i for all $i \leq d$ and parallel to V_i for all $i > 1$. Moreover, when such an $\mathcal{H}$ exists there exists $0 \neq \hat{V}$ perpendicular to it.

If the origin lies on $\mathcal{H}$ then it is a linear subspace containing each P_i as a vector. By Lemma A.1, the vectors $P_1, \ldots, P_d$ are almost surely independent, which implies that $\hat{V} \perp P_i$ for all $i \leq d$ or, simply, $\hat{V} = 0$, in contradiction to the above. □

Lemma A.3. *Let $d < n \in \mathbb{N}$; then,*

$$(n-d)\int_0^1 (1-u)^{n-d-1}u^d \mathrm{d}u = \frac{1}{\binom{n}{d}} \tag{A.6}$$

Proof. Let

$$\xi = (n-d)\int_0^1 (1-u)^{n-d-1}u^d \mathrm{d}u \tag{A.7}$$

Using integration by parts repeatedly yields

$$\begin{aligned}
\xi &= (n-d)\Big[u^d \tfrac{-(1-u)^{n-d}}{n-d}\Big]_0^1 - (n-d)d\tfrac{-1}{n-d}\int_0^1 (1-u)^{n-d}u^{d-1}\mathrm{d}u \\
&= 0 + d\int_0^1 (1-u)^{n-d}u^{d-1}\mathrm{d}u \\
&= d\Big[u^{d-1}\tfrac{-(1-u)^{n-d+1}}{n-d+1}\Big]_0^1 - d(d-1)\tfrac{-1}{n-d+1}\int_0^1 (1-u)^{n-d+1}u^{d-2}\mathrm{d}u \\
&= \tfrac{d(d-1)}{n-d+1}\Big[u^{d-2}\tfrac{-(1-u)^{n-d+2}}{n-d+2}\Big]_0^1 - \tfrac{-d(d-1)(d-2)}{(n-d+1)(n-d+2)}\int_0^1 (1-u)^{n-d+2}u^{d-3}\mathrm{d}u \\
&= \prod_{k=0}^{d-1}\tfrac{(d-k)}{(n-k)}\int_0^1 (1-u)^{n-d+(d-1)}\mathrm{d}u \\
&= \prod_{k=0}^{d-1}\tfrac{(d-k)}{(n-k)}\Big[\tfrac{-1}{n}(1-u)^n\Big]_0^1 \\
&= \tfrac{1}{n}\prod_{k=0}^{d-1}\tfrac{(d-k)}{(n-d+k)} \\
&= \tfrac{d!}{\prod_{k=0}^{d-1}(n-k)}\tfrac{(n-d)!}{(n-d)!} \\
&= \tfrac{d!(n-d)!}{n!} \\
&= \tfrac{1}{\binom{n}{d}}
\end{aligned}$$

□

Lemma A.4. *For all $q > p > 0, x \in \mathbb{R}^d$ it holds that*

$$\frac{\|\bar{x}\|_p}{\|\bar{x}\|_q} \leq \frac{d^{\frac{1}{p}}}{d^{\frac{1}{q}}} \tag{A.8}$$

Proof. Let $r = \frac{q}{p} > 1$, and so, by Hölder's inequality,

$$\sum_{i=1}^{d}|x|^p \leq \Big(\sum_{i=1}^{d}(|x|^p)^r\Big)^{\frac{1}{r}}\Big(\sum_{i=1}^{d}1^{\frac{r}{r-1}}\Big)^{1-\frac{1}{r}} = \Big(\sum_{i=1}^{d}|x|^q\Big)^{\frac{p}{q}}d^{1-\frac{p}{q}} \tag{A.9}$$

Taking everything to the power of $\frac{1}{p}$ completes the proof. □

Lemma A.5. *Let $X^1, \ldots, X^m$ be i.i.d. random variables distributed according to a continuous distribution F for some $m \in \mathbb{N}$; then for every $s \in (0,1)$, the following holds:*

$$Pr\Big[\sum_{i=1}^{m}G(X^i) \leq s\Big] = \frac{1}{m!}s^m \tag{A.10}$$

Proof. The proof is by induction on m. The base case is trivial. Now, assume the lemma holds for m and all $s \in (0, 1)$. A well known lemma from probability theory claims that

$$F(X^i) \sim U(0, 1) \tag{A.11}$$

Define $Y_i = G(X^i)$ and hence

$$Y^i \sim U(0, 1) \tag{A.12}$$

Thus,

$$Pr\Big[\sum_{i=1}^{m} G(X^i) \leq s\Big] = Pr\Big[\sum_{i=1}^{m} Y^i \leq s\Big] \tag{A.13}$$

Notice that the set

$$\Big\{y \in \mathbb{R}^m \mid \sum_{i=1}^{m} y^i \leq s, \forall i \leq d : y^i \geq 0\Big\} \tag{A.14}$$

is an m-dimensional scaled simplex whose volume is

$$\frac{1}{m!} s^m \tag{A.15}$$

□

Lemma A.6. *The following conditions are equivalent for* $p \in \triangle(n)$ *with* $supp(p) = \{1, \ldots, d\}$*;*

1. *p is a Nash equilibrium.*
2. *For every $1 \leq i, j \leq d$ and $k > d$ we have $R(i, p) = R(j, p) \geq R(k, p)$.*
3. *The set of points $\{(X_j^1, \ldots, X_j^d)\}_{j=1}^d \subset \mathbb{R}^d$ form a hyperplane $\mathcal{H}$ that lies above, as defined in (3.3), the points $\{(X_j^1, \ldots, X_j^d)\}_{j>d}$.*

Proof. $1 \iff 2$

It is immediate from the definition of Nash equilibrium.

$2 \iff 3$

Denote by $\hat{p}$ the projection of p to the subspace of $\mathbb{R}^n$ of dimension d containing solely the coordinates in $supp(p)$. Notice that

$$\mathcal{H} = \Big\{x \in \mathbb{R}^d \mid x \cdot \hat{p} = R(1, p)\Big\} \tag{A.16}$$

The rest now follows. □

Lemma A.7. *Let* $v \gg 0$ *and* $c > 0$ *such that* $\left\{x \in \mathbb{R}^d \mid x \cdot v = c\right\} = \mathcal{G} \subset \mathbb{R}^d$ *be a hyperplane with a strictly positive normal. If for all* $i \leq d$ *the point* te^i *lies below* $\mathcal{G}$*, then* R_t *(as defined in (3.5)) lies entirely below* $\mathcal{G}$.

Proof. Suppose that $q_1, q_2 \in \mathbb{R}^d$ both lie below $\mathcal{G}$; then every convex combination of the two lies below $\mathcal{G}$ as well. Thus every convex combination of $te^1, \ldots, te^d$ lies below $\mathcal{G}$. Moreover, any point in $\mathbb{R}^d$ can be written as $z - u$, where $0 \ll u$ and z is a convex combination of $te^1, \ldots, te^d$. Now, since $v \gg 0$ we have

$$(w - u) \cdot v < c \tag{A.17}$$

We conclude that R_t lies entirely below $\mathcal{G}$. □

References

[1] I. Bárány, S. Vempala, and A. Vetta, *Nash equilibria in random games*, In Proc. 2005 46th Annual IEEE Symposium on Foundations of Computer Science, Focs'05 (2005), p. Pittsburgh 123–131.

[2] V. Chisyakov, *A theorem on sums of independent positive random variables and its applications to branching random processes*, Theory of Probability and Its Applications, 9 (1964), pp. 640 – 648.

[3] E. V. Damme, *Stability and perfection of nash equilibria*, Springer, New York, (1991).

[4] L. Devroye, *On the oscillation of the expected number of extreme points of a random set*, Statist. Probab. Lett, 11 (1991), pp. 281–286.

[5] C. Goldie and C. Klüppelberg, *Subexponential distributions, in a practical guide to heavy tail*, Birkhäuser Boston, Basel (1998), pp. 435 – 459.

[6] S. Hart, Y. Rinott, and B. Weiss, *Evolutionarily stable strategies of random games, and the vertices of random polygons*, Annals of Applied Probability, 18 (2008), pp. 259–264, 275–287.

[7] A. Rényi and R. Sulanke, *Über die konvexe hülle von n zufällig gewählten punkten*, Zeitschrift Wahrscheinlichkeitstheorie und verwandte Gebiete, 2 (1963), pp. 75 – 84.

[8] J. M. Smith and G. R. Price, *The logic of animal conflict*, Nature, 246 (1973), pp. 15 – 18.

[9] J. L. Teugels, *The class of subexponential distributions*, The Annals of Probability, 3 (1975), pp. 1000 – 1011.