

The complexity of simple models – a study of worst and typical hard cases for the Standard Quadratic Optimization Problem

Immanuel M. Bomze* , Werner Schachinger* , Reinhard Ullrich*

* University of Vienna, Austria

Abstract

In a Standard Quadratic Optimization Problem (StQP), a possibly indefinite quadratic form (the simplest nonlinear function) is extremized over the standard simplex, the simplest polytope. Despite this simplicity, the nonconvex instances of this problem class allow for remarkably rich patterns of coexisting local solutions, which are closely related to practical difficulties in solving StQPs globally. In this study, we improve on existing lower bounds for the number of strict local solutions by a new technique to construct instances with a rich solution structure. Furthermore we provide extensive case studies where the system of supports (the so-called pattern) of solutions are analyzed in detail. Note that by naive simulation, in accordance to theory, most of the interesting patterns would not be encountered, since random instances have, with a high probability, quite sparse solutions (with singleton or doubleton supports), and likewise their expected numbers are considerably lower than in the worst case. Hence instances with a rich solution pattern are rather rare. On the other hand, by concentrating on (thin) subsets of promising instances, we are able to give an empirical answer on the size distribution of supports of strict local solutions to the StQP and their patterns, complementing average-case analysis of this NP-hard problem class.

Keywords Local Solutions; Quadratic Optimization; Evolutionary Stability; Global Optimization; Replicator Dynamics; Selection Stability

1 Introduction

1.1 Background and motivation

We consider the *Standard Quadratic Optimization Problem (StQP)* given by

$$\max_{\mathbf{x} \in \Delta^n} \mathbf{x}^\top \mathbf{A} \mathbf{x} \quad (1)$$

where \mathbf{A} is a symmetric $n \times n$ -matrix and Δ^n is the standard-simplex

$$\Delta^n = \left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1, x_i \geq 0 \text{ for all } i \in N \right\},$$

where $N = \{1, \dots, n\}$, also denoted as $[1 : n]$. Despite of its simplicity, this model is quite versatile. Applications are numerous, ranging from the famous Markowitz portfolio problem in Finance, Economics (evolutionary game theory) through Machine Learning (background-foreground clustering in image analysis) to life science applications, e.g. in Population Genetics (selection models) and Ecology (replicator dynamics). So it is not too surprising that the following questions are closely related (for detailed explanations and background see e.g. [3, 17, 24, 27]):

How many strict local solutions are there at most in a given StQP ?

How many evolutionarily stable states (*ESSs*) can coexist in a given (partnership) game ?

How many asymptotically stable fixed points can coexist for the replicator dynamics ?

How many stable equilibria can coexist in a one-locus, multi-allelic system ?

How many maximal cliques can coexist in an undirected graph ?

The last question sheds light on an important aspect of StQPs, namely the discrete combinatorial structure of this problem class of continuous optimization models. Indeed, for the subclass of StQPs based upon an adjacency matrix \mathbf{A} of an undirected graph, the answer to the last question is well known by the famous Moon/Moser bound [21]: asymptotically $\sqrt[3]{3}^n \approx (1.4422)^n$, a number exponential in the order n of the graph, and this bound is attained at the complete multipartite Turán graph $T(3k, k)$ with $n = 3k$.

However, the Moon/Moser bound is not valid for general symmetric matrices \mathbf{A} (with also non-binary entries). In this paper, we will push this exponential bound up, improving, and building, upon earlier investigations [9] where the basis ≈ 1.4592 was established, to a basis of ≈ 1.4933 , establishing a new world record, to the best of our knowledge. Note that a worst-case *upper* bound on the maximal number of strict local solutions is due to Sperner's theorem (see, e.g. [5, Section 6.3]), and asymptotically equals $\approx \frac{2^n}{1.25\sqrt{n}}$ (see (7) below), so that the basis necessarily is smaller than two. Relaxing further, one can establish without too much effort a sharp upper bound on the number of isolated critical points (satisfying the KKT condition) which is $2^n - 1$, the number of possible supports, an example being $\mathbf{A} = \mathbf{I}_n$, the $n \times n$ identity matrix: here, the barycenter of every face of Δ^n is a KKT point for (1).

To get the bound of $(15120)^{\frac{n}{24}} \approx (1.4933)^n$, we construct an instance with $n = 24$ having 15,120 coexisting strict local solutions, and an accordingly rich pattern structure. The methods employed may serve as good examples of the interaction/feedback loop between experimental and theoretical work in Applied Mathematics. Whether one views the experimental parts rather as empirical work, remains a matter of personal taste, but for sure our experiments add to the empirical evidence on the complexity of StQPs, and as we hope, to a significant

extent both from the worst-case and the typical-case point of view as (conditionally random) distributions of hard instances are investigated. In a sense, our experiments are similar to importance sampling technique in advanced simulation studies, in that the search (instance) space is swept thoroughly in interesting parts. See Subsection 1.3 below.

1.2 Looking over the fence; basic definitions and notation

The above mentioned close relations between the different fields Optimization, game theory and population/selection dynamics; as explained in [4], the optimization problem (1) is closely related to an *evolutionary game* with strategy set N in pairwise contests, with payoff matrix \mathbf{A} . If $\mathbf{A} = \mathbf{A}^\top$, this means that the row and the column player share the payoff equally, a partnership game. Likewise, the symmetric matrix $\mathbf{A} = [a_{ij}]$ could also collect the (incremental) fitness values for the allelic combination $\{i, j\} \in N \times N$ where N is the allele set for a single autosomal locus, an aspect we will not pursue further in this paper. Finally we may look at certain population dynamics called *replicator dynamics* defined as a system of coupled autonomous differential equations, a continuous-time dynamical system (a dot \dot{y} signifies derivative w.r.t. time t) and perform the usual qualitative equilibrium analysis:

$$\dot{x}_i(t) = x_i(t) ([\mathbf{A}\mathbf{x}(t)]_i - \mathbf{x}(t)^\top \mathbf{A}\mathbf{x}(t)) , \quad i \in N , \quad \mathbf{x}(0) = \mathbf{x}_0 \in \Delta^n ; \quad t \geq 0 . \quad (2)$$

We have the following equivalences for a point $\mathbf{x} \in \Delta^n$:

$$\left. \begin{aligned} & \mathbf{x} \quad \text{is a strict local maximizer of (1), i.e. strictly maximizes population overall fitness} \\ \iff & \mathbf{x} \quad \text{is an evolutionary stable strategy (ESS) for payoff matrix } \mathbf{A} \\ \iff & \mathbf{x} \quad \text{is an asymptotically stable fixed point for the dynamics (2).} \end{aligned} \right\} \quad (3)$$

For succinct proofs, we refer to [4, Theorem 10] where also equivalences for the weaker versions of solutions are stated and proved:

$$\left. \begin{aligned} & \mathbf{x} \quad \text{is a local maximizer of (1), i.e. maximizes population overall fitness} \\ \iff & \mathbf{x} \quad \text{is a neutrally stable strategy (NSS) for payoff matrix } \mathbf{A} \\ \iff & \mathbf{x} \quad \text{is a Lyapunov stable fixed point for the dynamics (2);} \end{aligned} \right\} \quad (4)$$

and

$$\left. \begin{aligned} & \mathbf{x} \quad \text{is a KKT point for (1), i.e. satisfies first-order necessary condition for local maximality} \\ \iff & \mathbf{x} \quad \text{is a Nash equilibrium strategy (NES) for payoff matrix } \mathbf{A} \\ \iff & \mathbf{x} \quad \text{is a saturated fixed point for the dynamics (2).} \end{aligned} \right\} \quad (5)$$

For us relevant are the first two equivalences in (3), (4) and (5), and in the following we will use both terms (strict local maximizer/ESS; local maximizer/NSS; KKT point/NES) interchangeably. For the readers' convenience, we will repeat the definitions of ESS, NSS, NES below (note that these also apply to non-symmetric square matrices $\mathbf{A} \neq \mathbf{A}^\top$, used for modeling non-partnership evolutionary games); to this end, we will introduce a bit of notation first: given any point $\mathbf{x} \in \Delta^n$, let

$$I(\mathbf{x}) = \{i \in N : x_i > 0\}$$

be the *support* of \mathbf{x} . The *extended support* is given by

$$J_A(\mathbf{x}) = \{i \in N : [A\mathbf{x}]_i = \mathbf{x}^\top A\mathbf{x}\}.$$

A vector $\mathbf{x} \in \Delta^n$ is called a (symmetric) *Nash Equilibrium State/Strategy (NES)*, if $\mathbf{x}^\top A\mathbf{x} \geq \mathbf{y}^\top A\mathbf{x}$ for all $\mathbf{y} \in \Delta^n$, i.e., if \mathbf{x} is a best reply to itself. It is easy to see that this implies $I(\mathbf{x}) \subseteq J_A(\mathbf{x})$, and that the set of alternative best replies \mathbf{y} to \mathbf{x} is given by $BR_A(\mathbf{x}) = \text{conv} \{\mathbf{e}_j : j \in J_A(\mathbf{x})\}$, where \mathbf{e}_j is the j -th column of I_n . In other words, all pure strategies actually used by \mathbf{x} , namely all $i \in I(\mathbf{x})$, must be alternative best replies.

A *Neutrally Stable State/Strategy (NSS)* is a *NES* which in addition satisfies $\mathbf{x}^\top A\mathbf{y} \geq \mathbf{y}^\top A\mathbf{y}$ for all $\mathbf{y} \in BR_A(\mathbf{x})$; in words, a state/strategy \mathbf{x} is *neutrally stable* if it is at least as good a reply to any \mathbf{y} than that \mathbf{y} to itself, for any alternative best reply \mathbf{y} to \mathbf{x} . Finally, an *NSS* is called *Evolutionarily Stable State/Strategy (ESS)*, if the last inequality is strict: $\mathbf{x}^\top A\mathbf{y} > \mathbf{y}^\top A\mathbf{y}$ for all $\mathbf{y} \in BR_A(\mathbf{x}) \setminus \{\mathbf{x}\}$: any alternative best reply \mathbf{y} to \mathbf{x} fares strictly worse against itself than the incumbent \mathbf{x} performs against \mathbf{y} .

We will use some more convenient notation: Given an $n \times n$ matrix A , we denote by

$$ESS(A) := \{\mathbf{x} \in \Delta^n : \mathbf{x} \text{ is an } ESS \text{ for } A\}$$

the set of all strict local maximizers of (1) if $A^\top = A$, cf. (3);

$$NSS(A) := \{\mathbf{x} \in \Delta^n : \mathbf{x} \text{ is an } NSS \text{ for } A\}$$

the set of all local maximizers of (1) if $A^\top = A$, cf. (4); and

$$NES(A) := \{\mathbf{x} \in \Delta^n : \mathbf{x} \text{ is an } NES \text{ for } A\}$$

the set of all KKT points of (1) if $A^\top = A$, cf. (5). Obviously, $ESS(A) \subseteq NSS(A) \subseteq NES(A)$. The last two sets are never empty, but they can be infinite, while the first one has to be always finite (but may be empty, e.g. for $A = O$). However, generically, the first two sets coincide [4, Corollary 14], in which case they are both nonempty and finite. More precisely, for every A and any $I \subseteq N$ there is at most one *ESS* \mathbf{x} with $I(\mathbf{x}) = I$. This follows immediately from the following non-inclusion result, see, e.g. [1]:

$$(\mathbf{x}, \mathbf{p}) \in NES(A) \times ESS(A) \quad \text{and} \quad I(\mathbf{x}) \subseteq J_A(\mathbf{p}) \quad \implies \quad \mathbf{x} = \mathbf{p}. \quad (6)$$

Hence we do not only know $|ESS(A)| \leq 2^n$ but rather, by Sperner's theorem about maximal antichains,

$$|ESS(A)| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor} \sim 2^n / \sqrt{\pi n}, \quad (7)$$

where we used Stirling's approximation formula. This is a consequence of (6), as the (*ESS*) *pattern* of A ,

$$\text{pattern}(A) := \{I(\mathbf{x}) : \mathbf{x} \in ESS(A)\}$$

forms an antichain w.r.t. set inclusion in the complete lattice of all subsets of N , which has n elements. Instances with a rich such pattern (and thus many coexisting strict local solutions) will be the main focus of our paper, extending and improving upon previous studies [9, 10, 11, 12, 25, 26]. Despite its discrete combinatorial nature, the pattern comprises all essential information on the solution set of the instance, at least generically; note that generically we have $ESS(A) = NSS(A)$. Indeed, given $I \in \text{pattern}(A)$, it is a mere matter of solving a linear equation system in $|I|$ variables to obtain the unique *ESS* \mathbf{x} such that $I(\mathbf{x}) = I$, see also Section 3.1.

1.3 Global solutions; worst-case complexity and average case results

Generically, there is only one global solution, but there are instances exhibiting exponentially many different (strict) global solutions. Consider for instance the complete multipartite Turán graph $T(3k, k)$ which has 3^k maximum cliques; if the matrix A is I_{3k} plus twice the adjacency matrix of $T(3k, k)$, then (1) has $3^k = (\sqrt[3]{3})^n$ strict global solutions, see [6].

Regarding the above equivalence results (3) on coexistence of many solutions we may rephrase them as follows: There is a constant $c \in [\sqrt[3]{3}, 2]$ such that

At least $\sim c^n$ strict global solutions can coexist for an StQP in n variables;

At least $\sim c^n$ distinct *ESSs* realize maximum welfare in a partnership game with n strategies;

At least $\sim c^n$ distinct selection equilibria with n alleles may yield maximum overall fitness in the population.

W.r.t. replicator dynamics in n state variables, there may coexist at least $\sim c^n$ asymptotically stable fixed points at which the Lyapunov function attains its global maximum.

It is yet unknown whether the maximal number of strict global solutions grows at the same rate or at a slower rate than the number of strict local solutions. This problem calls for future investigations which are beyond the scope of this paper.

Turning to classical worst-case complexity theory, while StQPs are NP-hard (e.g., by reduction to the maximum clique problem), they form a PTAS [4]. While for general QPs, determining one local solution is already NP-hard [22], this effort is polynomial-time for the StQP, at least in the maximum-clique case as the greedy algorithm is $\mathcal{O}(n)$; in practice, a very efficient local maximization algorithm for the general case was proposed in [23]. Obviously, determining *all* local solutions cannot be easier than solving the StQP. For a thorough discussion of these and related issues, we refer to [15].

So what is a typical hard instance ? Previous work on random instances of StQPs would discourage the naive random generation approach although this has been done in the past, for instance in [12] for adjacency matrices A of directed graphs of small order. In a series of papers [18, 19, 20], Kontogiannis and Spirakis looked at several continuous (e.g. Gaussian or uniform) i.i.d. distributions of entries of $n \times n$ instances \tilde{A} and proved, among other results, that $\mathbb{E} [|\text{pattern}(\tilde{A})|] = \mathbb{E} [|\text{ESS}(\tilde{A})|]$ does not grow faster than $\exp(0.138n)$ (which is way slower than e.g. $\sqrt[3]{3}^n$), and that

$$\mathbb{P} \left[\max \left\{ |I| : I \in \text{pattern}(\tilde{A}) \right\} \geq n^{2/3} \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

So instances with patterns of size $\sim c^n$ with $c \geq \sqrt[3]{3}$ exist, but they become increasingly rare as n gets large. Interesting complex (and therefore large) patterns would likely be missed or occur too rarely in a naive random sampling approach. On a related question, Chen and Peng obtained very recently in [14, 13] an even stronger probability bound, again using i.i.d. continuous distributions of a certain type, including exponential, uniform and Gaussian distribution, denoting by $x^*(A)$ a (generically unique) global solution of (1):

$$\mathbb{P} \left[|I(x^*(\tilde{A}))| \geq 3 \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Again, this shows that naive random sampling is not very promising; if we know that the support of the global solution is at most a doubleton (and this happens with a high probability), we simply have to search among all such $\mathcal{O}(n^2)$ doubletons and singletons; such instances occur with a high probability, and are no challenge

algorithmically. Note that due to (6), also the possible pattern structures are constrained as they cannot contain any supports containing a doubleton already detected as a member of the pattern. Therefore, the worst cases are hidden in regions of small probability, and we will explore that region more systematically than just random searching, exploiting structures guaranteeing complex patterns.

The paper is organized as follows: the central theoretical development is presented in Section 2, starting with robustness investigations in Subsection 2.1, combining small instances to larger ones in Subsection 2.2, introducing generating polynomials and perturbation methods in Subsections 2.3 and 2.4 as key ingredients for our arguments, and collecting the consequences for counting *ESSs* in Subsections 2.5 and 2.6. The last two Subsections 2.7 and 2.8 deal with intricate search strategies without which it would have been quite hard (if possible at all) to detect interesting instances in a systematic way. Thereafter, Section 3 applies the developed theory in an algorithmic approach to generate interesting patterns for several matrix classes, including a discussion of the achieved improvements of the status quo. A detailed study on cyclically symmetric instances of orders exceeding most of the previously investigated cases, including descriptive statistics and illustrations is presented in Section 4. These results shed some light on the “complexity of complexity” in this simplest class of hard optimization problems, and open a new perspective on the notion of an average case beyond naive random simulation.

2 Constructing complex *ESS* patterns

2.1 Some robustness results

Denote by \mathcal{S}^n all symmetric matrices of order n , and by \mathcal{F}_n those $A \in \mathcal{S}^n$ such that (1) has only strict local maximizers (and therefore finitely many):

$$\mathcal{F}_n := \{A \in \mathcal{S}^n : NSS(A) = ESS(A)\} ;$$

by \mathcal{F}_n^d those matrices which remain in \mathcal{F}_n upon sufficiently small diagonal perturbations:

$$\mathcal{F}_n^d := \{A \in \mathcal{F}_n : \text{for some } \varepsilon > 0, A + \text{Diag}(\mathbf{d}) \in \mathcal{F}_n \text{ for all } \mathbf{d} \text{ with } \|\mathbf{d}\| < \varepsilon\} ;$$

and by \mathcal{F}_n^r those $A \in \mathcal{S}^n$ such that all *NESs* are quasistrict in the sense that \mathbf{e}_i is an alternative best answer to \mathbf{x} if and only if strategy i is used in \mathbf{x} with positive frequency/probability:

$$\mathcal{F}_n^r := \{A \in \mathcal{S}^n : I(\mathbf{x}) = J_A(\mathbf{x}) \text{ for all } \mathbf{x} \in NES(A)\} .$$

We have $\mathcal{F}_n^r \subseteq \mathcal{F}_n$ due to [4, Thm.13], and that \mathcal{F}_n^r is dense in \mathcal{S}^n [4, Cor.14]. Moreover, \mathcal{F}_n^r is open in \mathcal{S}^n ; indeed, considering its complement, suppose that $(B_\ell, \mathbf{q}_\ell) \rightarrow (A, \mathbf{q}) \in \mathcal{S}^n \times \Delta^n$ along a sequence, where $B_\ell \in \mathcal{S}^n \setminus \mathcal{F}_n^r$ and $\mathbf{q}_\ell \in NES(B_\ell)$ such that $I(\mathbf{q}_\ell) = I \neq J = J_{B_\ell}(\mathbf{q}_\ell)$ for all ℓ (and of course $I \subset J$). By continuity, we have for ℓ large enough, $I(\mathbf{q}) \subseteq I(\mathbf{q}_\ell) = I \subset J = J_{B_\ell}(\mathbf{q}_\ell) \subseteq J_A(\mathbf{q})$ and similarly

$$\mathbf{y}^\top A \mathbf{q} = \lim_{\ell \rightarrow \infty} \mathbf{y}^\top B_\ell \mathbf{q}_\ell \leq \lim_{\ell \rightarrow \infty} \mathbf{q}_\ell^\top B_\ell \mathbf{q}_\ell = \mathbf{q}^\top A \mathbf{q} \quad \text{for all } \mathbf{y} \in \Delta^n . \quad (8)$$

Hence $\mathbf{q} \in NES(A)$ with $I(\mathbf{q}) \neq J_A(\mathbf{q})$, so $A \in \mathcal{S}^n \setminus \mathcal{F}_n^r$, and \mathcal{F}_n^r is thus an open set in \mathcal{S}^n . Therefore we arrive at the inclusions

$$\mathcal{F}_n^r \subseteq \mathcal{F}_n^d \subseteq \mathcal{F}_n .$$

However, the set \mathcal{F}_n is not open for $n \geq 3$, as the following slight extension of Example 3 in [3] shows:

Example 1. Let $A := \begin{pmatrix} -E_n & \mathbf{o} \\ \mathbf{o}^\top & 0 \end{pmatrix}$ and $B_\epsilon := \begin{pmatrix} -E_n & \mathbf{o} \\ \mathbf{o}^\top & -\epsilon \end{pmatrix}$, where E_n is the all ones matrix of order n . Then $NSS(A) = ESS(A) = \{\begin{pmatrix} \mathbf{o} \\ 1 \end{pmatrix}\}$, and for $\epsilon > 0$ we have $NSS(B_\epsilon) = \left\{ \frac{1}{1+\epsilon} \begin{pmatrix} \epsilon \mathbf{y} \\ 1 \end{pmatrix} : \mathbf{y} \in \Delta^n \right\}$, an infinite set if $n \geq 2$.

Let $\mathcal{F} := \bigcup_{n \geq 1} \mathcal{F}_n$, $\mathcal{F}^d := \bigcup_{n \geq 1} \mathcal{F}_n^d$, $\mathcal{F}^r := \bigcup_{n \geq 1} \mathcal{F}_n^r$ and $\mathcal{S} := \bigcup_{n \geq 1} \mathcal{S}^n$.

Lemma 2. (a) If $A \in \mathcal{F}_n$, then for all $B \in \mathcal{F}_n$ sufficiently close to A we have

$$|\{K' \in \text{pattern}(B) : K \subseteq K'\}| \geq 1, \quad \text{for all } K \in \text{pattern}(A), \quad (9)$$

$$|\{K \in \text{pattern}(A) : K \subseteq K'\}| \leq 1, \quad \text{for all } K' \in \text{pattern}(B), \quad (10)$$

which implies

$$|\text{pattern}(B)| \geq |\text{pattern}(A)|.$$

(b) If $A \in \mathcal{F}_n^r$, then for all $B \in \mathcal{F}_n$ sufficiently close to A we have $B \in \mathcal{F}_n^r$ and

$$\text{pattern}(B) = \text{pattern}(A).$$

Proof. (a) For $A \in \mathcal{F}_n$ let $f(\mathbf{x}) = \mathbf{x}^\top A \mathbf{x}$. For $\mathbf{p} \in \Delta^n$ define the set $U_\delta(\mathbf{p})$ to be the connected component of $\{\mathbf{x} \in \Delta^n : f(\mathbf{x}) \geq f(\mathbf{p}) - \delta\}$ containing \mathbf{p} . Let sets $\tilde{U}_\delta(\mathbf{p})$ be defined analogously, but with $\tilde{f}(\mathbf{x}) := \mathbf{x}^\top B \mathbf{x}$ in place of f . For $\delta \geq 0$ those sets are nonempty and they shrink towards $\{\mathbf{p}\}$ as $\delta \searrow 0$. Indeed, the family of sets $(U_\delta(\mathbf{p}))_{\delta > 0}$ constitutes a neighborhood basis for \mathbf{p} . This is true in the more general setting of \mathbf{p} being a strict local maximizer on Δ^n of a continuous function $f : \Delta^n \rightarrow \mathbb{R}$: To begin with, there is $\epsilon > 0$ such that $f(\mathbf{x}) < f(\mathbf{p})$ for all $\mathbf{x} \in \Delta^n$ satisfying $0 < \|\mathbf{x} - \mathbf{p}\| \leq \epsilon$. Then, for all $\eta \leq \epsilon$ we are to find $\delta > 0$, such that $U_\delta(\mathbf{p}) \subseteq N_\eta(\mathbf{p}) := \{\mathbf{x} \in \Delta^n : \|\mathbf{x} - \mathbf{p}\| < \eta\}$. Fix such η , define the compact set $A_\eta := \{\mathbf{x} \in \Delta^n : \frac{\eta}{2} \leq \|\mathbf{x} - \mathbf{p}\| \leq \eta\}$ and observe $m_\eta := \max_{\mathbf{x} \in A_\eta} f(\mathbf{x}) < f(\mathbf{p})$. Now for any $\delta < f(\mathbf{p}) - m_\eta$, the set $\{\mathbf{x} \in \Delta^n : f(\mathbf{x}) \geq f(\mathbf{p}) - \delta\}$ is contained in the complement of A_η , which is the union of two disjoint connected components: $N_{\frac{\delta}{2}}$, and an unbounded one. Thus $U_\delta(\mathbf{p})$, the connected component of $\{\mathbf{x} \in \Delta^n : f(\mathbf{x}) \geq f(\mathbf{p}) - \delta\}$ containing \mathbf{p} , satisfies $U_\delta(\mathbf{p}) \subseteq N_{\frac{\delta}{2}} \subseteq N_\delta$. Continuing with the proof, choose $\delta > 0$ small enough such that $U_\delta(\mathbf{p}) \cap U_\delta(\mathbf{q}) = \{\}$ for all $\{\mathbf{p}, \mathbf{q}\} \subseteq ESS(A)$ with $\mathbf{p} \neq \mathbf{q}$, and that $I(\mathbf{x}) = I(\mathbf{p})$ for all $\mathbf{x} \in U_\delta(\mathbf{p})$ with $I(\mathbf{x}) \subseteq I(\mathbf{p})$. Since Δ^n is compact, for B close enough to A we have $\tilde{U}_{\frac{\delta}{2}}(\mathbf{p}) \subseteq U_\delta(\mathbf{p})$ for every $\mathbf{p} \in ESS(A)$, and

$$\max_{\mathbf{x} \in \tilde{U}_{\frac{\delta}{2}}(\mathbf{p})} \mathbf{x}^\top B \mathbf{x} \geq \mathbf{p}^\top B \mathbf{p} > \mathbf{p}^\top B \mathbf{p} - \frac{\delta}{2} = \mathbf{y}^\top B \mathbf{y},$$

for all $\mathbf{y} \in \partial(\Delta^n \setminus \tilde{U}_{\frac{\delta}{2}}(\mathbf{p}))$, so that any global maximizer $\tilde{\mathbf{p}}$ of the latter problem is a local maximizer of $\max_{\mathbf{x} \in \Delta^n} \mathbf{x}^\top B \mathbf{x}$. So by assumption of $B \in \mathcal{F}$, we get $\tilde{\mathbf{p}} \in ESS(B)$. Because of $\tilde{U}_{\frac{\delta}{2}}(\mathbf{p}) \cap F_{I(\mathbf{p})} \subseteq U_\delta(\mathbf{p}) \cap F_{I(\mathbf{p})} \subseteq \text{relint}(F_{I(\mathbf{p})})$, where $F_K := \{\mathbf{x} \in \Delta^n : x_j = 0 \text{ for } j \notin K\}$ denotes a face of the standard simplex, we have $I(\mathbf{p}) \subseteq I(\tilde{\mathbf{p}})$. With $K = I(\mathbf{p})$ and $K' = I(\tilde{\mathbf{p}})$ we have thus shown (9). Next, for any $\mathbf{r} \in ESS(B)$, we cannot have two $\{\mathbf{p}, \mathbf{q}\} \subseteq ESS(A)$ with $\mathbf{p} \neq \mathbf{q}$ such that $I(\mathbf{p}) \subseteq I(\mathbf{r})$ and $I(\mathbf{q}) \subseteq I(\mathbf{r})$. Otherwise, define $\mathbf{v} := \mathbf{p} - \mathbf{q} \neq \mathbf{o}$ and observe that due to $\{\mathbf{p}, \mathbf{q}\} \subseteq ESS(A)$ and $\mathbf{r} \in ESS(B)$ we have

$$\frac{d^2}{dh^2} f(\mathbf{q} + h\mathbf{v}) = 2\mathbf{v}^\top A \mathbf{v} > 0 \quad \text{and} \quad \frac{d^2}{dh^2} \tilde{f}(\mathbf{r} + h\mathbf{v}) = 2\mathbf{v}^\top B \mathbf{v} < 0,$$

which is absurd if we choose B so close to A such that for all of the finitely many pairs $\{\mathbf{p}, \mathbf{q}\} \subseteq ESS(A)$ with $\mathbf{v} = \mathbf{p} - \mathbf{q} \neq \mathbf{o}$, the strict inequality $\mathbf{v}^\top A \mathbf{v} > 0$ implies also $\mathbf{v}^\top B \mathbf{v} > 0$. This proves (10) and completes the proof of part (a).

(b) The first assertion follows because \mathcal{F}^r is open. Next we show $\text{pattern}(\mathbf{A}) \subseteq \text{pattern}(\mathbf{B})$ for $\mathbf{B} \in \mathcal{F}^r$ sufficiently close to \mathbf{A} . Indeed, the point $\tilde{\mathbf{p}}$ constructed in (a) above satisfies $(\mathbf{B}\tilde{\mathbf{p}})_i = \tilde{\mathbf{p}}^\top \mathbf{B}\tilde{\mathbf{p}}$ for $i \in I(\mathbf{p})$, because $I(\mathbf{p}) \subseteq I(\tilde{\mathbf{p}})$ for $\delta > 0$ small enough, and likewise $(\mathbf{B}\tilde{\mathbf{p}})_i - \tilde{\mathbf{p}}^\top \mathbf{B}\tilde{\mathbf{p}} < 0$ for $i \notin I(\mathbf{p})$ follows from a continuity argument, since $(\mathbf{A}\mathbf{p})_i - \mathbf{p}^\top \mathbf{A}\mathbf{p} < 0$ for $i \notin I(\mathbf{p}) = J_{\mathbf{A}}(\mathbf{p})$ by assumption. Therefore $I(\tilde{\mathbf{p}}) = I(\mathbf{p})$, and $\text{pattern}(\mathbf{A}) \subseteq \text{pattern}(\mathbf{B})$ follows. Finally, assume that there would be a sequence of matrices $\mathbf{B}_\ell \in \mathcal{F}^r$ converging to \mathbf{A} with $\mathbf{q}_\ell \in \text{ESS}(\mathbf{B}_\ell)$ such that $I(\mathbf{q}_\ell) = J_{\mathbf{B}_\ell}(\mathbf{q}_\ell) = K \notin \text{pattern}(\mathbf{A})$ for all ℓ . By compactness, we may and do assume without loss of generality that $\lim_{\ell \rightarrow \infty} \mathbf{q}_\ell = \mathbf{q} \in \Delta^n$ exists. It follows $I(\mathbf{q}) \subseteq I(\mathbf{q}_\ell) = J_{\mathbf{B}_\ell}(\mathbf{q}_\ell) \subseteq J_{\mathbf{A}}(\mathbf{q})$ for all ℓ large enough. Now $\mathbf{q} \in \text{NES}(\mathbf{A})$ by (8), but since $\mathbf{A} \in \mathcal{F}^r$ by assumption, we have $I(\mathbf{q}) = J_{\mathbf{A}}(\mathbf{q})$, and therefore $K = I(\mathbf{q}_\ell) = J_{\mathbf{B}_\ell}(\mathbf{q}_\ell) = I(\mathbf{q}) \in \text{pattern}(\mathbf{A})$ for ℓ large enough, which proves $\text{pattern}(\mathbf{B}) = \text{pattern}(\mathbf{A})$. \square

We now give examples showing that in Lemma 2(a) we can have strict inequalities in any of (9) and (10), resulting in $|\text{pattern}(\mathbf{B})| > |\text{pattern}(\mathbf{A})|$, as well as strict inclusion $K \subset K'$ in (9) and (10).

Example 3. (a) Let $\mathbf{A} := \begin{pmatrix} -1 & -2 & 0 \\ -2 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $\mathbf{B}_\epsilon := \begin{pmatrix} -1 & -2 & 0 \\ -2 & -1 & 0 \\ 0 & 0 & -\epsilon \end{pmatrix}$. Then $\text{NSS}(\mathbf{A}) = \text{ESS}(\mathbf{A}) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$, and for $\epsilon > 0$ we have $\text{ESS}(\mathbf{B}_\epsilon) = \left\{ \frac{1}{1+\epsilon} \begin{pmatrix} 6 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{1+\epsilon} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. For $K := \{3\} \in \text{pattern}(\mathbf{A})$ the l.h.s. of (9) is 2.

(b) Let $\mathbf{A} := \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\mathbf{B}_\epsilon := \mathbf{A} + \epsilon \mathbf{I}_3$. Then $\text{NSS}(\mathbf{A}) = \text{ESS}(\mathbf{A}) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$, and $\text{ESS}(\mathbf{B}_\epsilon) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} \right\}$ holds for $\epsilon > 0$. For $K' := \{1, 2\} \in \text{pattern}(\mathbf{B}_\epsilon)$ the l.h.s. of (10) is 0.

(c) Let $\mathbf{A} := \begin{pmatrix} -1 & 2 & 8 \\ 2 & -4 & -16 \\ 8 & -16 & -66 \end{pmatrix}$ and $\mathbf{B}_\epsilon := \mathbf{A} + \epsilon \mathbf{I}_3$. Then $\text{NSS}(\mathbf{A}) = \text{ESS}(\mathbf{A}) = \left\{ \begin{pmatrix} 2/3 \\ 1/3 \\ 0 \end{pmatrix} \right\}$, and for $0 < \epsilon \leq \frac{1}{10}$ we have $\text{ESS}(\mathbf{B}_\epsilon) = \left\{ \frac{1}{18-130\epsilon+3\epsilon^2} \begin{pmatrix} 12-80\epsilon+\epsilon^2 \\ 6-53\epsilon+\epsilon^2 \\ 3\epsilon+\epsilon^2 \end{pmatrix} \right\}$. Thus $K \subset K'$, with $K = \{1, 2\}$ and $K' = \{1, 2, 3\}$.

2.2 Pattern generation by combining matrices

We now are in a position to construct complex patterns based upon a collection of given symmetric matrices $\mathbf{A}, \mathbf{B}_1, \dots, \mathbf{B}_n$ of orders n, k_1, \dots, k_n . These will be used to construct a matrix of order $\bar{n} := \sum_{i=1}^n k_i$. Define

$$\mathbf{Q} = \begin{pmatrix} \boldsymbol{\eta}_{k_1}^\top & \mathbf{o}^\top & \dots & \mathbf{o}^\top \\ \mathbf{o}^\top & \boldsymbol{\eta}_{k_2}^\top & \dots & \mathbf{o}^\top \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{o}^\top & \mathbf{o}^\top & \dots & \boldsymbol{\eta}_{k_n}^\top \end{pmatrix} \quad (11)$$

where $\boldsymbol{\eta}_k$ is the all ones vector in \mathbb{R}^k , and \mathbf{o} stands for the zero vector of appropriate dimension. Then both $\mathbf{A}_{\mathbf{Q}} := \mathbf{Q}^\top \mathbf{A} \mathbf{Q}$ and $\mathbf{B} := \text{Diag}(\mathbf{B}_1, \dots, \mathbf{B}_n)$ are symmetric matrices of order \bar{n} , and, thinking of t as a large real parameter, we let

$$\mathbf{G}_t = t\mathbf{A}_{\mathbf{Q}} + \mathbf{B}.$$

Note that special cases of this construction have been considered before: The case $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\mathbf{B}_1, \mathbf{B}_2$ arbitrary has been used in the proof of [9, Thm. 1], and the case \mathbf{A} arbitrary of order n , $\mathbf{B}_1 = \mathbf{I}_2$, $\mathbf{B}_i = \mathbf{I}_1$ for $i \in [2:n]$, has been used in the proof of [10, Thm. 1].

Theorem 4. (a) Let $\bar{b}_i := \max_{\mathbf{x} \in \Delta^{k_i}} \mathbf{x}^\top \mathbf{B}_i \mathbf{x}$ for $i \in [1:n]$, so that $\bar{b}_i = \bar{\mathbf{p}}_i^\top \mathbf{B}_i \bar{\mathbf{p}}_i$ for some $\bar{\mathbf{p}}_i \in \Delta^{k_i}$. Then we have

$$ESS(G_t) = \bigcup_{K \subseteq [1:n]} \mathcal{E}_K, \text{ where}$$

$$\mathcal{E}_K := \left\{ \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix} : I(\boldsymbol{\alpha}) = K, \mathbf{p}_i \in ESS(\mathbf{B}_i) \text{ for } i \in K, \mathbf{p}_i = \bar{\mathbf{p}}_i \text{ for } i \notin K, \right. \\ \left. \boldsymbol{\alpha} \in ESS(t\mathbf{A} + \text{Diag}(\mathbf{b})), \text{ where } b_i := \begin{cases} \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i, & i \in K, \\ \bar{b}_i, & \text{else.} \end{cases} \right\}. \quad (12)$$

$$(b) \text{ With the same assumptions as in (a), we have } NSS(G_t) = \bigcup_{K \subseteq [1:n]} \mathcal{N}_K, \text{ where}$$

$$\mathcal{N}_K := \left\{ \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix} : I(\boldsymbol{\alpha}) = K, \mathbf{p}_i \in NSS(\mathbf{B}_i) \text{ for } i \in K, \mathbf{p}_i = \bar{\mathbf{p}}_i \text{ for } i \notin K, \right. \\ \left. \boldsymbol{\alpha} \in NSS(t\mathbf{A} + \text{Diag}(\mathbf{b})), \text{ where } b_i := \begin{cases} \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i, & i \in K, \\ \bar{b}_i, & \text{else.} \end{cases} \right\}. \quad (13)$$

$$(c) \text{ With the same assumptions as in (a), we have } NES(G_t) = \bigcup_{K \subseteq [1:n]} \mathcal{Q}_K, \text{ where}$$

$$\mathcal{Q}_K := \left\{ \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix} : I(\boldsymbol{\alpha}) = K, \mathbf{p}_i \in NES(\mathbf{B}_i) \text{ for } i \in K, \mathbf{p}_i = \bar{\mathbf{p}}_i \text{ for } i \notin K, \right. \\ \left. \boldsymbol{\alpha} \in NES(t\mathbf{A} + \text{Diag}(\mathbf{b})), \text{ where } b_i := \begin{cases} \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i, & i \in K, \\ \bar{b}_i, & \text{else.} \end{cases} \right\}. \quad (14)$$

Proof. We start with the observation that any $\mathbf{y} \in \Delta^{\bar{n}}$ can be written as $\mathbf{y} = \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix}$, where $\boldsymbol{\alpha} \in \Delta^n$ and $\mathbf{p}_i \in \Delta^{k_i}$ for $i \in [1:n]$. Further note that $\mathbf{Q}\mathbf{y} = \boldsymbol{\alpha}$ and $\mathbf{B}\mathbf{y} = [\alpha_i \mathbf{B}_i \mathbf{p}_i]_i$ holds. Next, for $\mathbf{z} = \begin{pmatrix} \beta_1 \mathbf{q}_1 \\ \vdots \\ \beta_n \mathbf{q}_n \end{pmatrix}$, we get

$$\begin{aligned} (\mathbf{y} - \mathbf{z})^\top \mathbf{G}_t \mathbf{y} &= t(\mathbf{Q}\mathbf{y} - \mathbf{Q}\mathbf{z})^\top \mathbf{A} \boldsymbol{\alpha} + \sum_{i \in [1:n]} [\alpha_i^2 \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i - \alpha_i \beta_i \mathbf{q}_i^\top \mathbf{B}_i \mathbf{p}_i] \\ &= t(\boldsymbol{\alpha} - \boldsymbol{\beta})^\top \mathbf{A} \boldsymbol{\alpha} + \sum_{i \in K} \alpha_i (\alpha_i - \beta_i) \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i + \sum_{i \in K} \alpha_i \beta_i (\mathbf{p}_i - \mathbf{q}_i)^\top \mathbf{B}_i \mathbf{p}_i \\ &= (\boldsymbol{\alpha} - \boldsymbol{\beta})^\top [t\mathbf{A} + \text{Diag}(\mathbf{b})] \boldsymbol{\alpha} + \sum_{i \in K} \alpha_i \beta_i (\mathbf{p}_i - \mathbf{q}_i)^\top \mathbf{B}_i \mathbf{p}_i, \end{aligned} \quad (15)$$

which is nonnegative for any $\mathbf{z} \in \Delta^{\bar{n}}$ if $\mathbf{y} \in \mathcal{Q}_K$, and in a similar way,

$$\mathbf{y}^\top \mathbf{G}_t \mathbf{y} - \mathbf{z}^\top \mathbf{G}_t \mathbf{z} = t[\boldsymbol{\alpha}^\top \mathbf{A} \boldsymbol{\alpha} - \boldsymbol{\beta}^\top \mathbf{A} \boldsymbol{\beta}] + \sum_{i \in [1:n]} \alpha_i^2 \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i - \sum_{i \in [1:n]} \beta_i^2 \mathbf{q}_i^\top \mathbf{B}_i \mathbf{q}_i. \quad (16)$$

Let $K = I(\mathbf{Q}\mathbf{y}) = I(\boldsymbol{\alpha})$. First suppose $\mathbf{y} \in \mathcal{E}_K$. If $\mathbf{z} \neq \mathbf{y}$ as above is close to \mathbf{y} , then \mathbf{q}_i is close to \mathbf{p}_i for all $i \in K$ while $\mathbf{q}_i \in \Delta^{k_i}$ is arbitrary for $i \notin K$, and $\boldsymbol{\beta}$ is close to $\boldsymbol{\alpha}$. Further,

$$\begin{aligned} \mathbf{y}^\top \mathbf{G}_t \mathbf{y} &= t\boldsymbol{\alpha}^\top \mathbf{A} \boldsymbol{\alpha} + \sum_{i \in [1:n]} \alpha_i^2 \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i = t\boldsymbol{\alpha}^\top \mathbf{A} \boldsymbol{\alpha} + \sum_{i \in K} \alpha_i^2 \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i + \sum_{i \notin K} \alpha_i^2 \bar{\mathbf{p}}_i^\top \mathbf{B}_i \bar{\mathbf{p}}_i \\ &\geq t\boldsymbol{\beta}^\top \mathbf{A} \boldsymbol{\beta} + \sum_{i \in K} \beta_i^2 \mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i + \sum_{i \notin K} \beta_i^2 \bar{\mathbf{p}}_i^\top \mathbf{B}_i \bar{\mathbf{p}}_i \geq t\boldsymbol{\beta}^\top \mathbf{A} \boldsymbol{\beta} + \sum_{i \in [1:n]} \beta_i^2 \mathbf{q}_i^\top \mathbf{B}_i \mathbf{q}_i = \mathbf{z}^\top \mathbf{G}_t \mathbf{z}, \end{aligned}$$

in case of (b), and where at least one of the inequalities is strict in case of (a), so that \mathbf{y} is a (strict) local maximizer of $\mathbf{z}^\top \mathbf{G}_t \mathbf{z}$, and one inclusion is shown. To show the reverse inclusion, let again $K = I(\mathbf{Q}\mathbf{y})$ and put all $\mathbf{q}_i = \mathbf{p}_i$, while $\boldsymbol{\beta}$ is close to $\boldsymbol{\alpha}$. Then obviously \mathbf{y} is close to \mathbf{z} , and thus (16) yields

$$\boldsymbol{\alpha}^\top (t\mathbf{A} + \text{Diag}(\mathbf{b})) \boldsymbol{\alpha} - \boldsymbol{\beta}^\top (t\mathbf{A} + \text{Diag}(\mathbf{b})) \boldsymbol{\beta} = \mathbf{y}^\top \mathbf{G}_t \mathbf{y} - \mathbf{z}^\top \mathbf{G}_t \mathbf{z} \geq 0$$

with strict inequality for $\boldsymbol{\beta} \neq \boldsymbol{\alpha}$ in case of (a), so that $\boldsymbol{\alpha} \in NSS(t\mathbf{A} + \text{Diag}(\mathbf{b}))$ in case b) or $\boldsymbol{\alpha} \in ESS(t\mathbf{A} + \text{Diag}(\mathbf{b}))$ in case (a) follows. Similarly, put $\boldsymbol{\beta} = \boldsymbol{\alpha}$ and all $\mathbf{q}_j = \mathbf{p}_j$ for all $j \in K \setminus \{i\}$ while $\mathbf{q}_i \neq \mathbf{p}_i$ is close to \mathbf{p}_i for some fixed $i \in K$. In this case, (16) implies

$$\alpha_i^2 (\mathbf{p}_i^\top \mathbf{B}_i \mathbf{p}_i - \mathbf{q}_i^\top \mathbf{B}_i \mathbf{q}_i) = \mathbf{y}^\top \mathbf{G}_t \mathbf{y} - \mathbf{z}^\top \mathbf{G}_t \mathbf{z} \geq 0$$

with strict inequality in case of (a), hence the result on \mathbf{p}_i follows. In a similar way, the remaining inclusions in (c) are shown by means of (15). This completes the proof. \square

2.3 The use of generating polynomials

For an $n \times n$ matrix $\mathbf{A} \in \mathcal{S}$ define the *pattern generating* polynomial $p_{\mathbf{A}}$ and the *support size generating* polynomial $q_{\mathbf{A}}$ by

$$p_{\mathbf{A}}(x_1, \dots, x_n) := \sum_{I \in \text{pattern}(\mathbf{A})} \prod_{i \in I} x_i, \quad q_{\mathbf{A}}(x) := \sum_{I \in \text{pattern}(\mathbf{A})} x^{|I|} = p_{\mathbf{A}}(x, \dots, x). \quad (17)$$

Then we have $|ESS(\mathbf{A})| = q_{\mathbf{A}}(1)$, and the mean support size of $ESS(\mathbf{A})$ is given by $\frac{q'_{\mathbf{A}}(1)}{q_{\mathbf{A}}(1)}$. Also, if $\mathbf{P} = (\mathbf{e}_{\pi(1)}, \dots, \mathbf{e}_{\pi(n)})$ is a permutation matrix, then for $\mathbf{A}' := \mathbf{P}^\top \mathbf{A} \mathbf{P}$ we have $\mathbf{p} \in ESS(\mathbf{A}) \Leftrightarrow \mathbf{P}^\top \mathbf{p} \in ESS(\mathbf{A}')$, and therefore $p_{\mathbf{A}'}(x_{\pi(1)}, \dots, x_{\pi(n)}) = p_{\mathbf{A}}(x_1, \dots, x_n)$ and $q_{\mathbf{A}'}(x) = q_{\mathbf{A}}(x)$.

Theorem 5. (a) If $\{\mathbf{A}, \mathbf{B}_1, \dots, \mathbf{B}_n\} \subset \mathcal{F}^r$, then for t big enough, we have $\mathbf{G}_t \in \mathcal{F}^r$.

Moreover, $p_{\mathbf{G}_t}(\mathbf{x}_1, \dots, \mathbf{x}_n) = p_{\mathbf{A}}(p_{\mathbf{B}_1}(\mathbf{x}_1), \dots, p_{\mathbf{B}_n}(\mathbf{x}_n))$, where $\mathbf{x}_i = (x_{i,1}, \dots, x_{i,k_i})$ for $i \in [1:n]$, and $q_{\mathbf{G}_t}(x) = p_{\mathbf{A}}(q_{\mathbf{B}_1}(x), \dots, q_{\mathbf{B}_n}(x))$. In particular we have $|ESS(\mathbf{G}_t)| = p_{\mathbf{A}}(q_{\mathbf{B}_1}(1), \dots, q_{\mathbf{B}_n}(1))$.

(b) If $\{\mathbf{A}, \mathbf{B}_1, \dots, \mathbf{B}_n\} \subset \mathcal{F}^d$, then for t big enough, we have $\mathbf{G}_t \in \mathcal{F}^d$ and $|ESS(\mathbf{G}_t)| \geq p_{\mathbf{A}}(q_{\mathbf{B}_1}(1), \dots, q_{\mathbf{B}_n}(1))$.

Proof. (a) There is a $T > 0$ such that for all $\mathbf{b} \in \mathcal{B} := \prod_{i \in [1:n]} \left[\min_{\mathbf{x} \in \Delta^{k_i}} \mathbf{x}^\top \mathbf{B}_i \mathbf{x}, \max_{\mathbf{x} \in \Delta^{k_i}} \mathbf{x}^\top \mathbf{B}_i \mathbf{x} \right]$, we have $t\mathbf{A} + \text{Diag}(\mathbf{b}) = t(\mathbf{A} + \frac{1}{t} \text{Diag}(\mathbf{b})) \in \mathcal{F}^r$ for $t \geq T$, by Lemma 2(b). Fix such a t and let $\mathbf{y} = \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix} \in NES(\mathbf{G}_t)$. Then $I(\mathbf{y}) \neq J_{\mathbf{A}}(\mathbf{y})$ would imply $I(\boldsymbol{\alpha}) \neq J_{\mathbf{A}}(\boldsymbol{\alpha})$ or $I(\mathbf{p}_i) \neq J_{\mathbf{A}}(\mathbf{p}_i)$ for some $i \in I(\boldsymbol{\alpha})$. But this is a contradiction since $\boldsymbol{\alpha}$ resp. \mathbf{p}_i are NES s of matrices in \mathcal{F}^r . Therefore $\mathbf{G}_t \in \mathcal{F}^r$. Again by Lemma 2(b), we have $\text{pattern}(t\mathbf{A} + \text{Diag}(\mathbf{b})) = \text{pattern}(\mathbf{A})$ for all $t \geq T$ and all $\mathbf{b} \in \mathcal{B}$. For such t we observe $ESS(\mathbf{G}_t) = \bigcup_{K \subseteq [1:n]} \mathcal{E}_K = \bigcup_{L \in \text{pattern}(\mathbf{A})} \mathcal{E}_L$. It is now convenient to identify $[1:n]$ with $\bigcup_{i \in [1:n]} (\{i\} \times [1:k_i])$. For $\boldsymbol{\alpha}$ and \mathbf{y} as above, we have $I(\mathbf{y}) = \{\{i\} \times I(\mathbf{p}_i) : i \in I(\boldsymbol{\alpha})\}$ and conclude

$$\begin{aligned} p_{\mathbf{G}_t}(\mathbf{x}_1, \dots, \mathbf{x}_n) &= \sum_{I \in \text{pattern}(\mathbf{G}_t)} \prod_{j \in I} x_j = \sum_{L \in \text{pattern}(\mathbf{A})} \sum_{I \in \{I(\mathbf{y}) : \mathbf{y} \in \mathcal{E}_L\}} \prod_{j \in I} x_j \\ &= \sum_{L \in \text{pattern}(\mathbf{A})} \prod_{k \in L} \left(\sum_{S \in \text{pattern}(\mathbf{B}_k)} \prod_{\ell \in S} x_{k,\ell} \right) = \sum_{L \in \text{pattern}(\mathbf{A})} \prod_{k \in L} p_{\mathbf{B}_k}(\mathbf{x}_k) \\ &= p_{\mathbf{A}}(p_{\mathbf{B}_1}(\mathbf{x}_1), \dots, p_{\mathbf{B}_n}(\mathbf{x}_n)). \end{aligned}$$

(b) Using $B_i \in \mathcal{F}^d$ for $i \in [1:n]$, we fix an $\varepsilon > 0$ such that for all $\mathbf{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$ with $\|\mathbf{d}\| \leq \varepsilon$ we have $B_i + \text{Diag}(d_i) \in \mathcal{F}$ where $\mathbf{d}_i \in \mathbb{R}^{k_i}$ for $i \in N$. Let now

$$\bar{\mathcal{B}} := \prod_{i \in [1:n]} \left[\min_{\mathbf{x} \in \Delta^{k_i}, \|\mathbf{d}\| \leq \varepsilon} \mathbf{x}^\top (B_i + \text{Diag}(\mathbf{d}_i)) \mathbf{x}, \max_{\mathbf{x} \in \Delta^{k_i}, \|\mathbf{d}\| \leq \varepsilon} \mathbf{x}^\top (B_i + \text{Diag}(\mathbf{d}_i)) \mathbf{x} \right].$$

Then uniformly in $\mathbf{b} \in \bar{\mathcal{B}}$, $t\mathbf{A} + \text{Diag}(\mathbf{b}) = t(\mathbf{A} + \frac{1}{t}\text{Diag}(\mathbf{b})) \in \mathcal{F}$ holds for t big enough, since $\mathbf{A} \in \mathcal{F}^d$. We fix such t and assume $\bar{\mathbf{G}}_t := t\mathbf{A}_Q + \mathbf{B} + \text{Diag}(\mathbf{d}) \notin \mathcal{F}$ for some \mathbf{d} with $\|\mathbf{d}\| \leq \varepsilon$. Then $\mathbf{x}^\top \bar{\mathbf{G}}_t \mathbf{x}$ must have a non-strict local maximizer on $\Delta^{\bar{n}}$, in particular there have to be $\mathbf{y}, \mathbf{z} \subset \Delta^{\bar{n}}$, such that $\mathbf{y} \neq \mathbf{z}$ and all points $\mathbf{x}_\lambda := (1 - \lambda)\mathbf{y} + \lambda\mathbf{z}$, $0 \leq \lambda \leq 1$, are local maximizers of $\mathbf{x}^\top \bar{\mathbf{G}}_t \mathbf{x}$ on $\Delta^{\bar{n}}$. Write $\mathbf{y} = \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix}$ and $\mathbf{z} = \begin{pmatrix} \beta_1 \mathbf{q}_1 \\ \vdots \\ \beta_n \mathbf{q}_n \end{pmatrix}$ (where $\{\mathbf{p}_i, \mathbf{q}_i\} \subseteq NSS(B_i + \text{Diag}(\mathbf{d}_i)) = ESS(B_i + \text{Diag}(\mathbf{d}_i))$ for $i \in [1:n]$ and $\{\alpha, \beta\} \subset \Delta^n$), and consider $\mathbf{x}_{\lambda,i} := (1 - \lambda)\alpha_i \mathbf{p}_i + \lambda\beta_i \mathbf{q}_i$. We claim that for some $\bar{\mathbf{p}} \in ESS(B_i + \text{Diag}(\mathbf{d}_i))$ we have $\{\mathbf{x}_{\lambda,i} : 0 \leq \lambda \leq 1\} \subseteq LH(\bar{\mathbf{p}})$, the linear hull of that $\bar{\mathbf{p}}$. This is clear if $\min(\alpha_i, \beta_i) = 0$. For $0 \leq \lambda \leq 1$ and $\min(\alpha_i, \beta_i) > 0$ we have $((1 - \lambda)\alpha_i + \lambda\beta_i)^{-1} \mathbf{x}_{\lambda,i} \in ESS(B_i + \text{Diag}(\mathbf{d}_i))$, a finite set, whereas the left-hand side is continuous in λ and therefore constant. This implies $\mathbf{p}_i = \mathbf{q}_i$ and $\{\mathbf{x}_{\lambda,i} : 0 \leq \lambda \leq 1\} \subseteq LH(\mathbf{p}_i)$. So we may rewrite \mathbf{z} as $\mathbf{z} = \begin{pmatrix} \beta_1 \mathbf{p}_1 \\ \vdots \\ \beta_n \mathbf{p}_n \end{pmatrix}$. But then $(1 - \lambda)\alpha + \lambda\beta \in NSS(t\mathbf{A} + \text{Diag}(\mathbf{b}))$ for $0 \leq \lambda \leq 1$, where $\mathbf{b} \in \bar{\mathcal{B}}$ and therefore $t\mathbf{A} + \text{Diag}(\mathbf{b}) \in \mathcal{F}$. So $(1 - \lambda)\alpha + \lambda\beta \in ESS(t\mathbf{A} + \text{Diag}(\mathbf{b}))$, again a finite set, for $0 \leq \lambda \leq 1$. This is only possible if $\alpha = \beta$. So we have derived $\mathbf{y} = \mathbf{z}$, which is a contradiction. Therefore $\bar{\mathbf{G}}_t \in \mathcal{F}$ whenever $\|\mathbf{d}\| \leq \varepsilon$, and thus $\mathbf{G}_t = t\mathbf{A}_Q + \mathbf{B} \in \mathcal{F}^d$. The last assertion of (b) will be proved by demonstrating $|\mathcal{P}| \leq |ESS(\mathbf{G}_t)|$, where

$$\mathcal{P} := \left\{ \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix} : \alpha \in ESS(\mathbf{A}), \mathbf{p}_i \in ESS(\mathbf{B}_i) \text{ for } i \in I(\alpha), \mathbf{p}_i = \bar{\mathbf{p}}_i \text{ for } i \notin I(\alpha) \right\}. \quad (18)$$

The set \mathcal{P} need not coincide with $ESS(\mathbf{M})$ for some matrix \mathbf{M} , but it satisfies $|\mathcal{P}| = p_A(q_{B_1}(1), \dots, q_{B_n}(1))$, by (a). By Lemma 2(a), for t big enough and uniformly in $\mathbf{b} \in \bar{\mathcal{B}}$ the matrix $\mathbf{B} = \mathbf{B}_{t,\mathbf{b}} = \mathbf{A} + \frac{1}{t}\text{Diag}(\mathbf{b})$ will be close enough to \mathbf{A} to ensure that (9) and (10) hold for (\mathbf{A}, \mathbf{B}) . We fix such t . For $\mathbf{p} = \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix} \in \mathcal{P}$ with $\alpha \in ESS(\mathbf{A})$, $I(\alpha) = K$ and \mathbf{b} as in (12) we observe that, by (9), there is $\beta \in ESS(t\mathbf{A} + \text{Diag}(\mathbf{b}))$ such that $I(\alpha) \subseteq I(\beta)$, and $\mathbf{q} := \begin{pmatrix} \beta_1 \mathbf{q}_1 \\ \vdots \\ \beta_n \mathbf{q}_n \end{pmatrix} \in ESS(\mathbf{G}_t)$ by (12). Clearly $I(\mathbf{p}) \subseteq I(\mathbf{q})$, and thus

$$|\{K' \in \text{pattern}(\mathbf{G}_t) : K \subseteq K'\}| \geq 1, \quad \text{for all } K \in \{I(\mathbf{p}) : \mathbf{p} \in \mathcal{P}\}. \quad (19)$$

Now consider $\mathbf{q} = \begin{pmatrix} \beta_1 \mathbf{q}_1 \\ \vdots \\ \beta_n \mathbf{q}_n \end{pmatrix} \in ESS(\mathbf{G}_t)$. By Theorem 4(a) we have $\beta \in ESS(t\mathbf{A} + \text{Diag}(\mathbf{b}'))$ for some $\mathbf{b}' \in \bar{\mathcal{B}}$ defined as in the proof of (a), and $\mathbf{q}_i \in ESS(\mathbf{B}_i)$ for $i \in [1:n]$. But by (10), there is at most one $\alpha \in ESS(\mathbf{A})$ such that $I(\alpha) \subseteq I(\beta)$. If there exists such α , then $\mathbf{p} = \begin{pmatrix} \alpha_1 \mathbf{p}_1 \\ \vdots \\ \alpha_n \mathbf{p}_n \end{pmatrix} \in \mathcal{P}$, where $\mathbf{p}_i = \mathbf{q}_i$ for $i \in I(\alpha)$ and $\mathbf{p}_i = \bar{\mathbf{p}}_i$ for $i \notin I(\alpha)$, satisfies $I(\mathbf{p}) \subseteq I(\mathbf{q})$, and for any $\tilde{\mathbf{p}} \in \mathcal{P}$ with $\tilde{\mathbf{p}} \neq \mathbf{p}$ we have $I(\tilde{\mathbf{p}}) \not\subseteq I(\mathbf{q})$ (otherwise $I(\tilde{\mathbf{p}}) \subseteq I(\mathbf{p})$ implied $I(\tilde{\alpha}) \subseteq I(\beta)$ and $\tilde{\alpha} \in ESS(\mathbf{A})$). Therefore

$$|\{K \in \{I(\mathbf{p}) : \mathbf{p} \in \mathcal{P}\} : K \subseteq K'\}| \leq 1, \quad \text{for all } K' \in \text{pattern}(\mathbf{G}_t). \quad (20)$$

Relations (19) and (20) together culminate in $|\{I(\mathbf{p}) : \mathbf{p} \in \mathcal{P}\}| \leq |\text{pattern}(\mathbf{G}_t)|$, which is equivalent to $|\mathcal{P}| \leq |ESS(\mathbf{G}_t)|$, and thus the proof is complete. \square

An obvious modification of Theorem 5 concerns dropping the i th row and column, i.e. considering the case $k_i = 0$ in (11); in this case, the term p_{B_i} in the formula for p_A should be replaced with zero.

We can also replace some of the assumptions in Theorem 5 by weaker variants; indeed, the following implications hold true for t large enough:

- (a) $A \in \mathcal{F}^r$ and $B_i \in \mathcal{H} \implies G_t \in \mathcal{H}$, where $\mathcal{H} \in \{\mathcal{F}^r, \mathcal{F}^d, \mathcal{F}, \mathcal{S}\}$;
- (b) $A \in \mathcal{F}^d$ and $B_i \in \mathcal{H} \implies G_t \in \mathcal{H}$, where $\mathcal{H} \in \{\mathcal{F}^d, \mathcal{F}, \mathcal{S}\}$;
- (c) $A \in \mathcal{F}^r$ and $B_i \in \mathcal{S} \implies p_{G_t}(x_1, \dots, x_n) = p_A(p_{B_1}(x_1), \dots, p_{B_n}(x_n))$;
- (d) $A \in \mathcal{F}^d$ and $B_i \in \mathcal{S} \implies |ESS(G_t)| \geq p_A(q_{B_1}(1), \dots, q_{B_n}(1))$.

However, Example 1 shows that $\{A, B_i\} \subset \mathcal{F}$ does *neither* imply $G_t \in \mathcal{F}$ *nor* $|ESS(G_t)| \geq p_A(q_{B_1}(1), \dots, q_{B_n}(1))$ for all large t in general.

2.4 Emergence of complex patterns by explicit perturbation

The following corollary deals with a case where an explicit lower bound for t ensuring validity of the assertion of Theorem 5(a) can easily be constructed.

Corollary 6. *Let $A = E_n - I_n$ and $B_i \in \mathcal{F}_{k_i}$ for $i \in [1:n]$, with $p_i := |ESS(B_i)|$. Let $b := \max_{i \in [1:n]} \max_{x \in \Delta^{k_i}} x^\top B_i x$. Then for $t > b$ the matrix G_t satisfies $G_t \in \mathcal{F}$ and $|ESS(G_t)| = \prod_{i \in [1:n]} p_i$. If even $B_i \in \mathcal{F}_{k_i}^r$ holds for $i \in [1:n]$, then $G_t \in \mathcal{F}^r$.*

Proof. Consider $y = \begin{pmatrix} \alpha_1 p_1 \\ \vdots \\ \alpha_n p_n \end{pmatrix} \in \Delta^{\bar{n}}$, where $\alpha \in \Delta^n$, $p_i \in \Delta^{k_i}$ for $i \in [1:n]$, and $\bar{n} = \sum_{i=1}^n k_i$. Then we have

$$y^\top G_t y = t - \sum_{i=1}^n \sigma_i \alpha_i^2 = \alpha^\top (tE_n - \text{Diag}(\sigma)) \alpha \quad (21)$$

with $\sigma_i := t - p_i^\top B_i p_i > 0$, all $i \in [1:n]$. This shows that $y \in ESS(G_t)$ if and only if, for every $i \in [1:n]$, $p_i \in ESS(B_i)$, and for every choice of (p_1, \dots, p_n) , α is the unique ESS of $tE_n - \text{Diag}(\sigma)$, namely given by $\alpha_i = \frac{\mu}{\sigma_i}$ with $\mu = (\sum \sigma_i^{-1})^{-1} > 0$, hence α has full support by the choice of t which ensures $NSS(tE_n - \text{Diag}(\sigma)) = ESS(tE_n - \text{Diag}(\sigma)) = \{\alpha\}$ by (6). Using now $NSS(B_i) = ESS(B_i)$ for $i \in [1:n]$, we readily obtain $NSS(G_t) = ESS(G_t)$, i.e. $G_t \in \mathcal{F}$, from Theorem 4(a),(b). We further observe that $\text{pattern}(tE_n - \text{Diag}(\sigma)) = \text{pattern}(A) = \{[1:n]\}$, and thus $tE_n - \text{Diag}(\sigma) \in \mathcal{F}^r$, holds for $t > b$. Finally, along the lines of the proof of Theorem 5(a), we deduce $|ESS(G_t)| = \prod_{i \in [1:n]} p_i$, and $G_t \in \mathcal{F}^r$ in case of $B_i \in \mathcal{F}_{k_i}^r$ for $i \in [1:n]$. \square

Example 7. Let $A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $B_1 := \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \\ 3 & 3 & 0 \end{pmatrix}$ and $B_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, so that $ESS(B_1) = \left\{ \begin{pmatrix} 3/5 \\ 0 \\ 2/5 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/5 \\ 2/5 \end{pmatrix} \right\}$. There is only one more NES of B_1 , namely $\frac{1}{11}(3, 3, 5)^\top$ with full support, so that $B_1 \in \mathcal{F}_3^r$. Likewise, $ESS(B_2) = \{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \}$ with only one more NES , namely $\frac{1}{2}\eta_2$, and $B_2 \in \mathcal{F}_2^r$. On the other hand, $NES(A) = ESS(A) = \{ \frac{1}{2}\eta_2 \}$ with full support, so also $A \in \mathcal{F}_2^r$. Moreover $\max_{x \in \Delta^3} x^\top B_1 x = \frac{9}{5}$, $\max_{x \in \Delta^2} x^\top B_2 x = 1$ and $b := \max(\frac{9}{5}, 1) = \frac{9}{5}$. Then $t := 2 > b$ yields $G_2 = \begin{pmatrix} 1 & 0 & 3 & 2 & 2 \\ 0 & 1 & 3 & 2 & 2 \\ 3 & 3 & 0 & 2 & 2 \\ 2 & 2 & 2 & 1 & 0 \\ 2 & 2 & 2 & 0 & 1 \end{pmatrix}$, and from Corollary 6 we get $|ESS(G_2)| = 2 \cdot 2 = 4$. With $p_A(x_1, x_2) = x_1 x_2$, $p_{B_1}(x_1, x_2, x_3) = x_1 x_3 + x_2 x_3$, $q_{B_1}(x) = 2x^2$, $p_{B_2}(x_1, x_2) = x_1 + x_2$ and $q_{B_2}(x) = 2x$, Theorem 5(a) readily provides $p_{G_2}(x_1, x_2, x_3, x_4, x_5) = (x_1 x_3 + x_2 x_3)(x_4 + x_5)$ and $q_{G_2}(x) = 2x^2 \cdot 2x = 4x^3$, telling that G_2 has 4 ESS s, each of support size 3.

The next corollary shows that $t = 1$ is in some cases sufficiently large for the last assertion of Theorem 5(a) to hold, however $G_1 \in \mathcal{F}^r$ can not be guaranteed, even if we assume $A \in \mathcal{F}^r$ and $B_i \in \mathcal{F}^r$ for $i \in [1:n]$, see Example 9(b) below.

Corollary 8. *Let $A \in \mathcal{F}_n$ and $B_i \in \mathcal{F}_{k_i}$ for $i \in [1:n]$, with $p_i^\top B_i p_i = 0$ for every $i \in [1:n]$ and every $p_i \in ESS(B_i)$. Then for any $t > 0$ the matrix G_t satisfies $G_t \in \mathcal{F}$ and $|ESS(G_t)| = p_A(q_{B_1}(1), \dots, q_{B_n}(1))$. If in addition $p^\top A p = 0$ for every $p \in ESS(A)$, then this property is also shared by G_t : $q^\top G_t q = 0$ for every $q \in ESS(G_t)$.*

Proof. Let $t > 0$ and consider y as in the proof of Theorem 4, and as above. By (12), we have $y \in ESS(G_t)$ only if $p_i \in ESS(B_i)$ for all $i \in [1:n]$. Under the assumption of the corollary, we get, cf. (16), $y^\top G_t y = t \alpha^\top A \alpha$ with $\alpha \in ESS(A)$. Thus, in this case, $ESS(G_t) = \mathcal{P}$, with \mathcal{P} given in (18). Therefore $|ESS(G_t)| = |\mathcal{P}| = p_A(q_{B_1}(1), \dots, q_{B_n}(1))$. As we have $NSS(A) = ESS(A)$ and $NSS(B_i) = ESS(B_i)$ for $i \in [1:n]$, we readily obtain $NSS(G_t) = ESS(G_t)$, i.e. $G_t \in \mathcal{F}$, from Theorem 4(a),(b). Furthermore, the last assertion is a straightforward consequence of $y^\top G_t y = t \alpha^\top A \alpha$. \square

Example 9. (a) Let $A := \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$, $B_1 := \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$ and $B_2 := \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$. As in the previous example it can be checked that all three matrices belong to \mathcal{F}^r , that $ESS(A) = \{\frac{1}{2}\eta_2\}$, $ESS(B_1) = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ and $ESS(B_2) = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$. All the local maximizers of the three corresponding StQPs are global maximizers, and the global maximum equals 0 in all three cases. Therefore, by Corollary 8, all of the local maximizers (there are $p_A(3, 2) = 3 \cdot 2 = 6$ of them) of the StQP corresponding to $G_1 = \begin{pmatrix} -1 & -2 & -2 & 1 & 1 \\ -2 & -1 & -2 & 1 & 1 \\ -2 & -2 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -2 \\ 1 & 1 & 1 & -2 & -1 \end{pmatrix}$ are also global maximizers, and the global maximum equals 0.

(b) Let $A := \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ and $B_1 = B_2 := 2A$. Then all three matrices belong to \mathcal{F}^r , and by Corollary 8 we have $|ESS(G_1)| = p_A(q_{B_1}(1), \dots, q_{B_n}(1)) = 2 + 2 = 4$. Still $G_1 = \begin{pmatrix} 0 & -2 & -1 & -1 \\ -2 & 0 & -1 & -1 \\ -1 & -1 & 0 & -2 \\ -1 & -1 & -2 & 0 \end{pmatrix} \notin \mathcal{F}^r$, since $p = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix} \in NES(G_1)$ satisfies $I(p) = \{1, 2\} \neq \{1, 2, 3, 4\} = J_{G_1}(p)$. Clearly for t large enough, $G_t \in \mathcal{F}^r$ will hold.

2.5 Special pattern classes

Corollary 10. *Let $A \in \mathcal{F}_n^r$ and $B_i \in \mathcal{F}_{k_i}^r$ for $i \in [1:n]$ be given. Assume that each of these matrices has the property that its set of ESS consists of points that all have the same support size, i.e., the support size generating polynomials of those matrices are monomials. Then G_t will have that same property if t is large enough, and any of the following conditions is satisfied:*

(a) $q_A(x) = x^n$,

(b) There are $m, p_1, \dots, p_n \in \mathbb{N}$ such that $q_{B_i}(x) = p_i x^{m_i}$ holds for $i \in [1:n]$.

Proof. Let $p_i, m_i \in \mathbb{N}$ be such that $q_{B_i}(x) = p_i x^{m_i}$ holds for $i \in [1:n]$ and assume that t is large enough to ensure validity of the assertion of Theorem 5 (a). In case (a) we have $p_A(x_1, \dots, x_n) = \prod_{i=1}^n x_i$ and therefore $q_{G_t}(x) = \prod_{i=1}^n q_{B_i}(x) = \left(\prod_{i=1}^n p_i\right) x^{\sum_{i=1}^n m_i}$, and in case (b) we have $q_A(x) = r x^\ell$ for some $r, \ell \in \mathbb{N}$, which means that p_A is homogeneous of degree ℓ , and therefore $q_{G_t}(x) = p_A(p_1 x^m, p_2 x^m, \dots, p_n x^m) = p_A(p_1, p_2, \dots, p_n) x^{m\ell}$. As $q_{G_t}(x)$ turns out to be a monomial in both cases, the proof is complete. \square

Example 11. Let $A = (a_{i,j}) \in \mathcal{S}^9$ be defined by $a_{i,i} = 0, a_{i,i \oplus 1} = 35, a_{i,i \oplus 2} = 12, a_{i,i \oplus 3} = a_{i,i \oplus 4} = 20$, for $i \in [1:9]$, where \oplus is addition modulo 9, but with 0 represented by 9, so that e.g. $2 \oplus 8 \oplus 8 = 9$. The matrix A is an instance of a cyclically symmetric matrix, which are considered in detail in the next section. It can be checked that $A \in \mathcal{F}_9^r$, and that

$$\text{pattern}(A) = \{\{1, 2, 3, 4, 5\} \oplus i : i \in [1:9]\} \cup \{\{1, 2, 3, 4, 7\} \oplus i : i \in [1:9]\} \cup \{\{1, 2, 3, 6, 7\} \oplus i : i \in [1:9]\},$$

where we denote $K \oplus i := \{k \oplus i : k \in K\}$. So $|\text{ESS}(A)| = 27$, with all ESSs of the same support size 5, i.e. $q_A(x) = 27x^5$. We want to construct a matrix $G_t \in \mathcal{F}_{12}^r$ having many ESSs with common support size 5, using for B_1, \dots, B_9 the matrices $l_1, l_1, l_1, l_1, l_1, l_1, l_2, l_2, l_2$ in some order. Thus $q_{B_i}(x) \in \{x, 2x\}$, and so condition (b) of Corollary 10 is satisfied with $m = 1$. As it turns out, for $t = 3$ and the order specified above we have $G_t \in \mathcal{F}^r$ and $104 = p_A(1, 1, 1, 1, 1, 1, 2, 2, 2) = \max\{p_A(x) : x \in \{1, 2\}^9, \mathbf{\eta}_9^\top x = 12\}$. Accordingly $q_{G_t}(x) = 104x^5$.

Knowing $q_A(x)$ and $q_{B_i}(1), q'_{B_i}(1)$ for $i \in [1:n]$ (but not $p_A(x_1, \dots, x_n)$), does in general not allow to determine $q_{G_t}(1)$ and $q'_{G_t}(1)$ for t large enough. There are however some cases where we can obtain partial results, which we now present.

Corollary 12. Let $A \in \mathcal{F}_n^r$ and $B_i \in \mathcal{F}_{k_i}^r$ for $i \in [1:n]$ be given, and let t large enough to ensure validity of the assertion of Theorem 5(a) for all matrices $G_{t,\pi} := tA_Q + \text{Diag}(B_{\pi(1)}, \dots, B_{\pi(n)})$, with $\pi \in \mathcal{S}_n$, where \mathcal{S}_n is the set of permutations of $[1:n]$, and G_t clearly is one of those $G_{t,\pi}$.

(a) Assume $q_A(x) = x^n$. Then $q_{G_t}(1) = \prod_{i=1}^n q_{B_i}(1)$ and $\frac{q'_{G_t}(1)}{q_{G_t}(1)} = \sum_{i=1}^n \frac{q'_{B_i}(1)}{q_{B_i}(1)}$.

(b) Assume that $q_A(x) = rx^\ell$ for some $r, \ell \in \mathbb{N}$, and that there is $m > 0$ such that $\frac{q'_{B_i}(1)}{q_{B_i}(1)} = m$ for all $i \in [1:n]$. Then we have $\frac{q'_{G_t}(1)}{q_{G_t}(1)} = m\ell$.

(c) Assume that there are $a, b \in \mathbb{N}$ such that $q_{B_i}(1) = a$ and $q'_{B_i}(1) = b$ for $i \in [1:n]$, and let $m := \frac{b}{a}$. Then $q_{G_t}(1) = q_A(a)$ and $\frac{q'_{G_t}(1)}{q_{G_t}(1)} = m \frac{aq'_A(a)}{q_A(a)} \geq m \frac{q'_A(1)}{q_A(1)}$.

(d) In general, there is $\pi \in \mathcal{S}_n$ such that $q_{G_{t,\pi}}(1) \geq q_A\left(\sqrt[n]{\prod_{i=1}^n q_{B_i}(1)}\right)$.

(e) Assume that $q_A(x) = rx^\ell$ for some $r, \ell \in \mathbb{N}$, and $q_{B_1}(1) \geq \dots \geq q_{B_n}(1)$. Then there is $\pi \in \mathcal{S}_n$ such that $q_{G_{t,\pi}}(1) \geq \lceil \frac{r\ell}{n} \rceil q_{B_1}(1) \left(\prod_{i=2}^n q_{B_i}(1)\right)^{\frac{\ell-1}{n-1}} + \left(r - \lceil \frac{r\ell}{n} \rceil\right) \left(\prod_{i=2}^n q_{B_i}(1)\right)^{\frac{\ell}{n-1}}$.

Proof. For the proof of (a), note that $p_A(x) = \prod_{i=1}^n x_i$, and therefore $q_{G_t}(1) = \prod_{i=1}^n q_{B_i}(1)$. Moreover, $x_i \frac{\partial p_A}{\partial x_i}(x) = p_A(x)$ for $i \in [1:n]$, and thus

$$q'_{G_t}(1) = \sum_{i=1}^n \frac{\partial p_A}{\partial x_i}(q_{B_1}(1), \dots, q_{B_n}(1)) q'_{B_i}(1) = p_A(q_{B_1}(1), \dots, q_{B_n}(1)) \sum_{i=1}^n \frac{q'_{B_i}(1)}{q_{B_i}(1)} = q_{G_t}(1) \sum_{i=1}^n \frac{q'_{B_i}(1)}{q_{B_i}(1)}.$$

Regarding (b), recall that p_A is homogeneous of degree ℓ , and we use that, together with $q'_{B_i}(1) = m q_{B_i}(1)$ for $i \in [1:n]$, in

$$\begin{aligned} q'_{G_t}(1) &= \sum_{i=1}^n \frac{\partial p_A}{\partial x_i}(q_{B_1}(1), \dots, q_{B_n}(1)) q'_{B_i}(1) = m \sum_{i=1}^n \frac{\partial p_A}{\partial x_i}(q_{B_1}(1), \dots, q_{B_n}(1)) q_{B_i}(1) \\ &= m \ell p_A(q_{B_1}(1), \dots, q_{B_n}(1)) = m \ell q_{G_t}(1). \end{aligned}$$

For (c) we use $\sum_{i=1}^n \frac{\partial p_{\mathbf{A}}}{\partial x_i}(a, \dots, a) = q'_{\mathbf{A}}(a)$, and the fact, that $\frac{aq'_{\mathbf{A}}(a)}{q_{\mathbf{A}}(a)}$ is increasing for $a \geq 0$. We turn to (d), choose $\pi \in \mathcal{S}_n$ randomly according to the uniform distribution, and consider the random variable $q_{\mathbf{G}_{t,\pi}}(1)$. Fix k , let Y be uniformly distributed on $\{y \in \{0, 1\}^n : \boldsymbol{\eta}_n^\top y = k\}$, and note that, by Jensen's inequality,

$$\begin{aligned} \mathbb{E} \prod_{i=1}^k q_{\mathbf{B}_{\pi(i)}}(1) &= \mathbb{E} \prod_{i=1}^n (q_{\mathbf{B}_i}(1))^{Y_i} = \mathbb{E} \exp \left(\sum_{i=1}^n Y_i \log(q_{\mathbf{B}_i}(1)) \right) \\ &\geq \exp \left(\mathbb{E} \sum_{i=1}^n Y_i \log(q_{\mathbf{B}_i}(1)) \right) = \exp \left(\sum_{i=1}^n \frac{k}{n} \log(q_{\mathbf{B}_i}(1)) \right) = \left(\prod_{i=1}^n q_{\mathbf{B}_i}(1) \right)^{\frac{k}{n}}. \end{aligned}$$

Using now linearity of expectation in (17), we arrive at $E := \mathbb{E} q_{\mathbf{G}_{t,\pi}}(1) \geq q_{\mathbf{A}} \left(\sqrt[n]{\prod_{i=1}^n q_{\mathbf{B}_i}(1)} \right)$. Since clearly there has to be $\pi \in \mathcal{S}_n$ satisfying $q_{\mathbf{G}_{t,\pi}}(1) \geq E$, this finishes the proof of d).

For the proof of (e), note that from $r\ell = \sum_{i=1}^n |\{s \in \text{pattern}(\mathbf{A}) : i \in s\}|$ we may deduce that there is $i^* \in [1:n]$ that is contained in at least $\lceil \frac{r\ell}{n} \rceil$ elements of $\text{pattern}(\mathbf{A})$. Choose $\pi \in \{\tau \in \mathcal{S}_n : \tau(i^*) = 1\}$ randomly according to the uniform distribution. Then, as before,

$$\mathbb{E} q_{\mathbf{G}_{t,\pi}}(1) \geq \sum_{i^* \in s \in \text{pattern}(\mathbf{A})} q_{\mathbf{B}_1}(1) \left(\prod_{i=2}^n q_{\mathbf{B}_i}(1) \right)^{\frac{\ell-1}{n-1}} + \sum_{i^* \notin s \in \text{pattern}(\mathbf{A})} \left(\prod_{i=2}^n q_{\mathbf{B}_i}(1) \right)^{\frac{\ell}{n-1}},$$

and putting everything together, (e) is proved. \square

Example 13. Assume we are told that there is $\mathbf{A} \in \mathcal{F}_{10}^r$ satisfying $q_{\mathbf{A}}(x) = 10x^8$, but $\text{pattern}(\mathbf{A})$ is not revealed. We want to use that information to infer the existence of a matrix $G_t \in \mathcal{F}_{21}^r$ having many ESSs with common support size 8. For $(\mathbf{B}_i)_{i \in [1:10]}$ we are going to use \mathbf{l}_3 and 9 copies of \mathbf{l}_2 in some order. From Corollary 12(e) we deduce that for t large enough and some order of the matrices \mathbf{B}_i we can have $|\text{ESS}(\mathbf{G}_t)| \geq 8 \cdot 3 \cdot 2^7 + 2 \cdot 2^8 = 3584$. Now there is a cyclically symmetric matrix $\mathbf{A} \in \mathcal{F}_{10}^r$, defined by $a_{i,i} = 0, a_{i,i \oplus 1} = a_{i,i \oplus 2} = a_{i,i \oplus 3} = 13, a_{i,i \oplus 4} = a_{i,i \oplus 5} = 8$, for $i \in [1:10]$. We have $\text{pattern}(\mathbf{A}) = \{[1:8] \oplus i : i \in [1:10]\}$ and moreover, for t large enough, $p_{\mathbf{G}_t}(\mathbf{x}) = 3584$ for any $\mathbf{x} \in \{2, 3\}^{10}$, such that $\boldsymbol{\eta}_{10}^\top \mathbf{x} = 21$, which shows that the inequality in Corollary 12(e) can be sharp in some cases.

2.6 Consequences for counting ESSs

While in Section 3 we will employ the results of this section to find matrices of orders 10 to 24 with very large numbers of ESSs, we use them now to obtain exponential growth (in n) of the maximal number of ESSs of matrices in certain subclasses of \mathcal{F} .

We first recall some results from [9]. Define U_n to be the *largest number of ESSs* that any $n \times n$ -matrix can have:

$$U_n := \max \{|\text{ESS}(\mathbf{A})| : \mathbf{A} \in \mathcal{S}^n\}$$

where \mathcal{S}^n denotes the set of all symmetric $n \times n$ matrices. We are interested in lower bounds for U_n , and we will exploit the following theorem from [9].

Theorem 14. The sequence $\{U_n\}$ is such that $\lim_{n \rightarrow \infty} U_n^{\frac{1}{n}}$ exists. Furthermore, denoting this limit by γ , $U_n^{\frac{1}{n}} \leq \gamma$ for all n .

Proof. See [9, Thm. 2]. \square

The proof technique in [9] builds upon the inequality $U_{n+m} \geq U_n U_m$ and Fekete's Subadditivity Lemma (in its superadditive variant). It can also be used to establish the following results.

Theorem 15. *Let*

$$U_n^{(d)} := \max \{ |ESS(A)| : A \in \mathcal{F}_n^d \},$$

$$U_n^{(r)} := \max \{ |ESS(A)| : A \in \mathcal{F}_n^r \},$$

$$U_n^{(g)} := \max \{ |ESS(A)| : A \in \mathcal{F}_n^r, \text{ every } \mathbf{p} \in ESS(A) \text{ is a global solution of (1)} \},$$

$$U_n^{(s)} := \max \{ |ESS(A)| : A \in \mathcal{F}_n^r, |\{K\} : K \in \text{pattern}(A)\}| = 1 \}.$$

- (a) For any $\bullet \in \{d, r, g, s\}$ there is $\gamma^{(\bullet)} \leq \gamma$ such that $\lim_{n \rightarrow \infty} (U_n^{(\bullet)})^{\frac{1}{n}} = \gamma^{(\bullet)}$, and $(U_n^{(\bullet)})^{\frac{1}{n}} \leq \gamma^{(\bullet)}$ holds for all n .
(b) If $(U_n^{(\bullet)})^{\frac{1}{n}} \geq c$ for $n \in \{n_1, \dots, n_k\}$, and for some N every integer $\geq N$ is in the sumset $n_1 \mathbb{N}_0 + \dots + n_k \mathbb{N}_0$, then $(U_n^{(\bullet)})^{\frac{1}{n}} \geq c$ for all $n \geq N$.

Proof. (a) To establish $U_{n+m}^{(\bullet)} \geq U_n^{(\bullet)} U_m^{(\bullet)}$ we choose B_1 of order n and B_2 of order m such that $|ESS(B_1)| = U_n^{(\bullet)}$ and $|ESS(B_2)| = U_m^{(\bullet)}$. Furthermore we let $A = E_2 - 2I_2$ (with $A \in \mathcal{F}^r$ and $p_A(x) = x_1 x_2$), construct G_t of order $n+m$, and observe that for t large enough we will have $U_{n+m}^{(\bullet)} \geq |ESS(G_t)| \geq U_n^{(\bullet)} U_m^{(\bullet)}$, by invoking Theorem 5 b) in case of $\bullet = d$, Theorem 5 a) in case of $\bullet = r$, Corollary 8 in case of $\bullet = g$ (here we may w.l.o.g. assume, by adding multiples of the all ones matrix to the matrices B_i , that $\mathbf{p}^\top B_i \mathbf{p} = 0$ for all $\mathbf{p} \in ESS(B_i)$), also note that $\mathbf{q} = (\frac{1}{2})$ is the only ESS of A and satisfies $\mathbf{q}^\top A \mathbf{q} = 0$), and Corollary 10 in case of $\bullet = s$. Now use Fekete's Subadditivity Lemma. Since the feasible sets in the maximization problems defining the constants $\gamma^{(\bullet)}$ are all subsets of \mathcal{S}^n , it is clear that we have $\gamma^{(\bullet)} \leq \gamma$.

(b) Let $n \geq N$. Then, for some $\ell_1, \dots, \ell_k \in \mathbb{N}_0$ we have $n = \sum_{i=1}^k \ell_i n_i$. Denote $m := \sum_{i=1}^k \ell_i$, let $A := E_k - I_k$ and let B_1, \dots, B_m consist of ℓ_i copies of C_i for $i \in [1:k]$, where C_i is of order n_i and satisfies $|ESS(C_i)| \geq c^{n_i}$. Then, for t large enough, $U_n^{(\bullet)} \geq |ESS(G_t)| \geq \prod_{i=1}^k |ESS(C_i)|^{\ell_i} \geq c^n$, by Theorem 5. \square

We remark that besides the obvious $\max(\gamma^{(g)}, \gamma^{(s)}) \leq \gamma^{(r)} \leq \gamma^{(d)} \leq \gamma$ we do not know of further inequalities relating these constants, and in particular not, if any of the stated inequalities is strict.

2.7 Cyclically Symmetric Matrices

We will employ symmetry transformations of the coordinates of vectors given by cyclic permutation (see also [7], where this notation has been introduced), denoting by $a \oplus b$, $a \ominus b$ and $a \odot b$ the result of addition, subtraction and multiplication modulo n . To keep in line with standard notation, we consider the remainders $[1:n]$ instead of $[0:n-1]$, e.g. $1 \oplus (n-1) = n$, see also Example 11. To be more precise, let P_i be the square $n \times n$ permutation matrix which effects $P_i x = [x_{i \oplus j}]_{j \in [1:n]}$ for all $x \in \mathbb{R}^n$ (for example, if $n = 3$ then $P_2 x = [x_3, x_1, x_2]^\top$). Obviously $P_i = (P_1)^i$ for all integers i (recall P^{-3} is the inverse matrix of PPP), $P_i^\top = P_{n-i} = P_i^{-1}$ and $P_n = I_n$. A *circulant matrix* $S = C(a)$ based on a vector $a \in \mathbb{R}^n$ (as its last column rather than the first) is given by

$$S = [P_{n-1}a, P_{n-2}a, \dots, P_1a, a].$$

If $S = C(a)$ is symmetric it is called *cyclically symmetric*, and that holds whenever

$$a_i = a_{n-i} \quad \text{for all } i \in [1:n-1]. \quad (22)$$

It is easy to see that any circulant matrix $S = C(a)$ satisfies $P_i^\top S P_i = S$ for all $i \in [1:n]$, and this is the key to their use in finding matrices with many *ESSs*.

Lemma 16. *Let a problem $\max_{x \in \Delta^n} x^\top S x$ be given and, let M be an arbitrary permutation matrix with $M^\top S M = S$. If x^* is a solution to the problem, then Mx^* is also a solution of the problem. The two vectors need not differ from each other, though, if additional symmetry prevails.*

Proof. $(Mx^*)^\top S (Mx^*) = (x^*)^\top (M^\top S M) x^* = (x^*)^\top S x^*$. □

Due to the structure of the matrices $C(a)$ any found *ESS* leads to potentially another $n - 1$ *ESSs*, where the involved permutation matrices are P_1, \dots, P_{n-1} , if we are cautious enough to break above-mentioned symmetry. Then the number of *ESSs* found in a game with a cyclically symmetric matrix must be a sum of multiples of the prime factors of n (or 0 or 1). When n is prime, this leads to the nice property that the game has 0,1 or a multiple of n *ESSs*. We add here a result linking our construction of matrices G_t to cyclic symmetry.

Proposition 17. *Let $A \in \mathcal{F}_n^r$ and $\bar{B} \in \mathcal{F}_k^r$, with A, \bar{B} both cyclically symmetric, and let $B_i := \bar{B}$ for $i \in [1:n]$. Then G_t is congruent to a cyclically symmetric matrix via a permutation matrix.*

Proof. Let $N := nk$ and define the $N \times N$ matrix $M := \sum_{i=0}^{n-1} \sum_{j=0}^{k-1} E_{1+ik+j, 1+i+jn}$, where $E_{\ell, m} := e_\ell e_m^\top$ has a single entry 1 in row ℓ and column m and 0s elsewhere. Note that M is a permutation matrix, since $\{1+ik+j : i \in [0:n-1], j \in [0:k-1]\} = \{1+i+jn : i \in [0:n-1], j \in [0:k-1]\} = [1:N]$. Let $H_t := M^\top G_t M$. Clearly, H_t is symmetric. H_t is indeed cyclically symmetric, since we can prove $(H_t)_{\ell, m} = (H_t)_{\ell \oplus_N 1, m \oplus_N 1}$ for any $\ell, m \in [1:N]$, where \oplus_m denotes addition modulo m with remainders in $[1:m]$: Let $\ell := 1+i+jn$ and $m := 1+\iota+jn$, with $i, \iota \in [0:n-1]$ and $j, j \in [0:k-1]$. Then

$$(H_t)_{\ell, m} = e_{1+i+jn} M^\top G_t M e_{1+\iota+jn} = e_{1+ik+j} (tQ^\top A Q + B) e_{1+\iota+k+j} = t e_{1+\iota} A e_{1+\iota} + \delta_{i, \iota} e_{1+j} \bar{B} e_{1+j}, \quad (23)$$

where $\delta_{\cdot, \cdot}$ denotes the Kronecker delta. We now distinguish three cases:

- a) $i < n-1, \iota < n-1$ implies $(H_t)_{\ell+1, m+1} = t e_{2+\iota} A e_{2+\iota} + \delta_{i, \iota} e_{1+j} \bar{B} e_{1+j} = (H_t)_{\ell, m}$, because of cyclic symmetry of A .
- b) $i = n-1, \iota < n-1$: In this case $\ell+1 = 1+0+(j+1)n$, $\delta_{i, \iota} = \delta_{0, \iota+1} = 0$ and $(H_t)_{\ell \oplus_N 1, m+1} = t e_{2 \oplus_N \iota} A e_{2 \oplus_N \iota} = (H_t)_{\ell, m}$, again by cyclic symmetry of A . The case $i < n-1, \iota = n-1$ is treated analogously.
- c) $i = \iota = n-1$ implies $(H_t)_{\ell \oplus_N 1, m \oplus_N 1} = t e_{2 \oplus_N i} A e_{2 \oplus_N i} + e_{2 \oplus_N j} \bar{B} e_{2 \oplus_N j} = (H_t)_{\ell, m}$, because of cyclic symmetry of A and \bar{B} . □

Equation (23) with $m = N$ and therefore $\iota = n-1, j = k-1$ can be used to find the last column of the cyclically symmetric matrix H_t , see the following example.

Example 18. *Let $A := C([3, 7, 7, 3, 1]^\top)$ and $\bar{B} := I_3 - E_3 = C([-1, -1, 0]^\top)$. Then G_1 and the cyclically symmetric matrix $H_1 = C([3, 7, 7, 3, 1+(-1), 3, 7, 7, 3, 1+(-1), 3, 7, 7, 3, 1+0]^\top) = C([3, 7, 7, 3, 0, 3, 7, 7, 3, 0, 3, 7, 7, 3, 1]^\top)$ have equal numbers of *ESS*. Verifying $A, \bar{B} \in \mathcal{F}^r$ as well as $q_A(x) = 5x^3$, $q_{\bar{B}}(x) = 3x$, and invoking Corollary 8, we deduce $|ESS(H_1)| = q_{G_1}(1) = 5 \cdot 3^3 = 135$.*

In the vein of Theorem 15 we define $U_n^{(c)} := \max \{|ESS(A)| : A \in \mathcal{F}_n^r, A \text{ is cyclically symmetric}\}$, but unfortunately do not know, if $U_{n+m}^{(c)} \geq U_n^{(c)} U_m^{(c)}$ holds in general (we doubt it), or if $\lim_{n \rightarrow \infty} (U_n^{(c)})^{\frac{1}{n}}$ exists. Still we can define $\gamma^{(c)} := \limsup_{n \rightarrow \infty} (U_n^{(c)})^{\frac{1}{n}}$, which satisfies $\gamma^{(c)} \leq \gamma$. Also, if $(U_k^{(c)})^{\frac{1}{k}} \geq \bar{\gamma}$ for some k , then $(U_n^{(c)})^{\frac{1}{n}} \geq \bar{\gamma}$ for all $n \in k\mathbb{N}$.

2.8 Restricting the search space

We started experimenting to search cyclically symmetric matrices $C(\mathbf{a})$ for *ESSs* on different n , where we set $a_n = 0$ in every case. This reduces the degrees of freedom and does not do any harm, since multiples of E_n can be added to or subtracted from the game matrix without changing the game. So in total we have $\lfloor \frac{n}{2} \rfloor$ variables for every n , and for these variables we allowed integers.

This approach leads to good results for smaller n (say ≤ 12), but for larger n this procedure became prohibitively slow. For these instances, our idea is not only to exploit cyclic permutations inherent to the matrices $C(\mathbf{a})$, but to enforce $M^\top C(\mathbf{a}) M = C(\mathbf{a})$ for one or more additional permutation matrices M . This leads to more restrictions on the degrees of freedom for constructing the vector \mathbf{a} .

Empirically for $n \in [13 : 24]$ it turned out that using the following construction is a good choice. Define $n \times n$ matrices $E_{i,j} = \mathbf{e}_i \mathbf{e}_j^\top$, let k and n be mutually prime and define

$$M_k = \sum_{i=1}^n E_{k \odot i, i}.$$

Example 19. For $n = 5$ and $k = 3$ we get $M_3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

If we require $M_k^\top C(\mathbf{a}) M_k = C(\mathbf{a})$, then further restrictions on the a_i result. Therefore, if too many restrictions are imposed, then \mathbf{a} becomes trivial in that a_i is constant across all $i \in [1 : n - 1]$, and if we do not use enough restrictions by this construction, all a_i may have different values.

Example 20. Consider $n = 15$ and M_2 . By cyclical symmetry it is required that $a_i = a_{15-i}$ for $i \in [1 : 14]$. Since $C(\mathbf{a}) = M_2^\top C(\mathbf{a}) M_2$, we can calculate the further restrictions $a_i = a_{2i}$ for $i \in [1 : 14]$. These result in $\mathbf{a} = (a_1, a_1, a_3, a_1, a_5, a_3, a_1, a_1, a_3, a_5, a_1, a_3, a_1, a_1, 0)^\top$ by keeping the smallest indices.

A careful choice of these further restrictions was successfully applied to various cyclically symmetric matrices for different n , which lead to good results, see Subsection 3.2 for details.

3 Coexistence of many *ESSs* for several matrix classes

3.1 The challenge to find all *ESSs*

An algorithm to detect all *ESSs* of a given game (i.e., matrix) was already sketched in [2]. We briefly summarize its implementation here, focusing on the necessary modifications.

Since for every support set $I \in 2^{[1:n]} \setminus \emptyset$ there exists at most one *ESS* \mathbf{x} with $I(\mathbf{x}) = I$, the idea of the algorithm is to search for *ESSs* on (potentially) all the $2^n - 1$ supports of the game. This is done exploiting equation (6), so a full support search only happens when the game does not exhibit any *ESSs* at all.

On every searched support the following two steps are carried out:

1. FINDEQ - Find a serious candidate for a strict local maximizer.
2. CHECKSTAB - Check if the candidate is really a strict local maximizer.

The algorithm we implemented differs from the original approach in two points. The original algorithm suggests to search the whole power set of $[1 : n]$ by selecting sets \mathcal{J}_{\min} and \mathcal{J}_{\max} , being minimal (maximal) with respect to

the set inclusion, and this is done iteratively (see p.317 in [2]). From a computational point of view - especially when one expects to find many *ESS*s - it is faster to search only “from the smallest supports upwards” (i.e. neglecting the \mathcal{J}_{\max} -sets), since it is cheaper to run FINDEQ on small supports. The second difference is the way the sets are chosen - not by them being minimal with respect to the set inclusion - but by them having the minimal number of elements.

We coded this algorithm as multi-threaded for rational matrices using exact arithmetics, thus avoiding any roundoff errors. Our system setup is an Intel i7-4930K CPU with 6 kernels, 16GB RAM and a SSD hard disk, where the output of the algorithm is stored in a database.

Note that this implementation has been utilized in the literature before, see [7] and [8], although in a slightly different context.

To give an estimate, this setup can search and record up to 5000 matrices of order $n = 9$ per second, but this depends largely on the complexity of the rational numbers and on the pattern of the game.

FINDEQ - Finding a serious candidate for an *ESS* The question is how to find a suitable candidate for an *ESS*. Admitting *all NES* as candidates is too much (since there could be infinitely many), and also from a computational point of view it is wise to restrict the set of candidates of a game-matrix \mathbf{A} , denoted $K(\mathbf{A})$, such that $ESS(\mathbf{A}) \subseteq K(\mathbf{A}) \subseteq NES(\mathbf{A})$.

In the original paper [2] the characterization of extremal equilibria with the help of polyhedra is used. To this end the vertices of a polyhedron have to be found, which can be accomplished by linear programming techniques. This original method is far too slow for practical applications, so we adapted it by introducing Proposition 21, admitting every *NES* as candidate for an *ESS*, if it is the only *NES* on the currently considered support.

Define the restriction of a matrix \mathbf{A} on an index set I as \mathbf{A}_I , i.e. $\mathbf{A}_I = (a_{ij})$ for $i, j \in I$, and analogously the restriction of a vector \mathbf{x} on I is defined as \mathbf{x}_I .

Proposition 21. *Let a $n \times n$ game-matrix \mathbf{A} and a support $I \subseteq [1 : n]$ be given.*

(a) *Then $\mathbf{p} \in NES(\mathbf{A})$ with $I(\mathbf{p}) = I$ if and only if $(\mathbf{p}_I, \mathbf{p}^\top \mathbf{A} \mathbf{p})$ is a solution to the linear system*

$$\begin{aligned} \mathbf{A}_I \mathbf{x}_I - v \boldsymbol{\eta}_{|I|} &= 0 \\ \boldsymbol{\eta}_{|I|}^\top \mathbf{x}_I &= 1 \end{aligned} \tag{24}$$

in variables $(\mathbf{x}_I, v) \in \mathbb{R}^{|I|} \times \mathbb{R}$, that also satisfies the inequalities

$$\begin{aligned} x_i &> 0 \quad \forall i \in I \quad \text{as well as} \\ (\mathbf{A} \mathbf{x})_j &\leq v \quad \forall j \in [1 : n] \setminus I. \end{aligned} \tag{25}$$

(b) *If $\mathbf{p} \in ESS(\mathbf{A})$ with $I(\mathbf{p}) = I$, then the solution to equation (24) is unique.*

Proof. (a) Recall that $\mathbf{x} \in \Delta^n$ with $I(\mathbf{x}) = I$ is a *NES* if and only if $(\mathbf{A} \mathbf{x})_i = \mathbf{x}^\top \mathbf{A} \mathbf{x} \quad \forall i \in I$ and $(\mathbf{A} \mathbf{x})_i \leq \mathbf{x}^\top \mathbf{A} \mathbf{x} \quad \forall i \in [1 : n] \setminus I$, then (a) is just a reformulation of this statement.

(b) Applying the non inclusion result of (6) it is clear that existence of a unique solution of the system of equations (24) and inequalities (25) is a necessary condition for there being an *ESS* \mathbf{p} with $I(\mathbf{p}) = I$. It remains to show that it is impossible to happen that the system of equations (24) has infinitely many solutions which then get reduced to exactly one by inequalities (25).

Assume that the solution set \mathcal{L}_I of equations (24) is infinite, which implies $|I| > 1$. Assuming that there is an *ESS* \mathbf{p} with $I(\mathbf{p}) = I$ there are only two possibilities: either $\mathcal{L}_I \cap \Delta^n = \{\mathbf{p}\}$ or $|\mathcal{L}_I \cap \Delta^n| = \infty$. In the first case every neighborhood of \mathbf{p} would contain further points of $\mathcal{L}_I \cap \Delta^n$ contradicting $|\mathcal{L}_I \cap \Delta^n| = 1$. In the second case there will be a line segment L such that $\mathbf{p} \in L \subseteq \mathcal{L}_I \cap \Delta^n$, and there will be $v \in \mathbb{R}$ such that $\mathbf{q}^\top \mathbf{A} \mathbf{q} = v$ for all $\mathbf{q} \in L$, which means that \mathbf{p} is not a strict local maximizer, contradicting $\mathbf{p} \in \text{ESS}(\mathbf{A})$. \square

Proposition 21 gives us an efficient method to find the candidates in $K(\mathbf{A})$: for the considered support set solve the linear system in (24), if there exists a unique solution check the inequalities, if they hold then the solution is in $K(\mathbf{A})$. If there is no unique solution to the system (24) then there exists no (reasonable) candidate for an *ESS* with support I .

Note that whenever for all nonempty $I \subseteq [1 : n]$ the determinant of system (24) is nonzero, then $\mathbf{A} \in \mathcal{F}^d$.

CHECKSTAB - Checking which candidates are really *ESS*s Verification of the *ESS* property of the candidates is a quite cumbersome task, since it potentially involves checks for copositivity [2, Section 3]. It is interesting to note that empirically only a very small percentage of *ESS*s are detected by copositivity checks, a check for positive definiteness is sufficient in almost all cases, see Section 4.2. This is in line with genericity (i.e. the considered matrix being $\in \mathcal{F}_n$), but since we are working on discrete data, it is not straightforward.

One remark from a computational point of view: for generic matrices, the method of [2, Thm.3.3] is a recursive version of Sylvester’s minorant criterion. This approach gets very slow for large n , so it is advantageous to use a Cholesky decomposition instead for checking positive definiteness in these cases.

3.2 Methods and results

A first lower bound for γ , as defined in Theorem 14, has been given in [16] (due to the above mentioned connection with the Moon/Moser result [21]) with $\gamma \geq 3^{\frac{1}{3}}$, and it was conjectured that this holds with equality. However, it was shown in [25] that this is not the case: a 7×7 -cyclically symmetric matrix (see Section 2.7) was found which showed 14 *ESS*s instead of 12 predicted by the conjecture.

The latest improvement of the lower bound (to our knowledge) is $\gamma \geq 30^{\frac{1}{9}}$ and was stated in [9], there a 9×9 principal submatrix of an 11×11 -cyclically symmetric matrix with 30 *ESS*s was presented.

We used the methods described in Section 2.8 on cyclically symmetric matrices, and the results for $n \in [5 : 24]$ can be found in Table 1. The first three columns show the maximal number of *ESS*s we found together with the smallest vector \mathbf{x} which generates the matrix $\mathbf{C}(\mathbf{a})$ for every n , in the sense that $a_i = x_i$ for $i \in [1 : \lfloor \frac{n}{2} \rfloor]$. The remaining a_i are generated by $a_i = a_{n-i}$ for all $i \in [1 : n-1]$ and $a_n = 0$. The term *smallest* here is defined as the one minimizing $\|\cdot\|_\infty$ from all found vectors. If there was more than one smallest vector \mathbf{x} , we chose one of them arbitrarily.

The fourth column gives the lower bound for γ implied by the stated matrix.

The column $|I|$ states the support size of all found *ESS*s. An entry k for some matrix \mathbf{M} means that we have $\{|I(p)| : p \in \text{ESS}(\mathbf{M})\} = \{k\}$. Interestingly for all matrices listed in Table 1 only one support size occurs, also for n which are not prime. One would assume that the support sizes for the best results are around $\lfloor \frac{n}{2} \rfloor$, but enough instances deviate from that, especially when n gets larger. We can only conjecture that either we

n	\times	$U_n \geq$	$\gamma \geq$	$ I $	n	\times	$U_n \geq$	$\gamma \geq$	$ I $
5	1,3	5	1.37973	3	15	9,9,5,9,3,5,9	360	1.48054	7
6	1,-1,1	9	1.44225	2	16	7,5,4,4,4,5,7,-1	512	1.47683	6
7	1,5,8	14	1.45792	3	17	$\alpha, \beta, \gamma, \delta, \gamma, \beta, \alpha, \epsilon$ ^a	493	1.44013	9
8	5,5,13,11	20	1.45422	4	18	8,7,5,5,5,5,7,8,-1	1152	1.47938	7
9	1,5,10,13	30	1.45923	3	19	13,31,31,27,31,27,13,13,27	1444	1.46654	7
10	13,19,25,7,5	50	1.47876	4	20	13,13,13,8,8,8,13,13,13,-1	2560	1.48051	8
11	1,4,6,6,5	66	1.46357	5	21	15,15,7,15,15,7,7,15,7,15	4410	1.49123	9
12	293, 64, 179, 196, 64, 262	105	1.47378	4	22	2,2,4,4,5,5,4,4,2,2,-1	5632	1.48076	9
13	7,18,18,10,7,10	143	1.46486	5	23	$a, b, c, d, e, e, d, c, b, a, f$ ^b	2507	1.40537	13
14	4,3,2,2,3,4,-1	224	1.47189	5	24	15,15,7,15,15,7,15,7,15,15,7	15120	1.4933	10

^awhere $(\alpha, \beta, \gamma, \delta, \epsilon) = (1302, 831, 1222, 860, 518)$.

^bwhere $(a, b, c, d, e, f) = (27478, 22664, 10976, 25676, 18552, 17939)$.

Table 1: Results of using the algorithm on cyclically symmetric matrices

did not find the matrices with the largest number of *ESS*s possible, or that additional properties of the game matrices restrict the support-patterns further.

Note that we did not spend an equal amount of computation time on each n , nor was the number of searched matrices the same for every n . In that sense the results among different n are not comparable.

3.3 Discussion and consequences

The best lower bound for γ we obtained (for cyclically symmetric matrices) by finding 24×24 -matrices with 15120 *ESS*s. Then $\gamma \geq 15120^{\frac{1}{24}} \approx 1.4933 > 30^{\frac{1}{9}} \approx 1.4592$, the best value known in the literature before. Note that also the cases $n \in \{10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22\}$ would beat this value.

For the case $n = 7$ we are able to provide smaller integers than stated in [25].

As it can be seen, for 9×9 -matrices it is not necessary to take the principal submatrix of a cyclically symmetric 11×11 -matrix as done in [9], the use of a 9×9 -cyclically symmetric matrix is sufficient, which also exhibits much smaller values than the one in [9].

While for $n \leq 12$ we put no further restrictions on the entries of cyclically symmetric matrices of order n (for $n = 12$ we investigated the neighborhood of the matrix $C([69, 14, 42, 46, 14, 62, 14, 46, 42, 14, 69, 0]^T)$ with 93 *ESS*s and luckily found the one listed in the table), we took different routes for $n \geq 13$. Our best matrix for $n = 15$ satisfies $C(a) = M_2^T C(a) M_2$, and was found by restricting the search to a three parameter family, as explained in Example 20. Similarly, for $n \in \{13, 19, 21, 24\}$ our best matrices obey corresponding symmetries expressed by permutation matrices M_5, M_7, M_2 and M_5 , respectively, with additional further constraints for $n \in \{21, 24\}$ to restrict the search to two parameter families. For $n \in \{14, 16, 18, 20, 22\}$ we searched for matrices A of order $k \in \{7, 8, 9, 10, 11\}$ satisfying $q_A(x) = kx^{k-2}$. Such matrices are known to exist for any $k \geq 3$, see [9, Thm. 5], but we needed such matrices with integer entries. Then Corollary 8 is applied with $B_i = I_2 - E_2$ for $i \in [1:k]$, and $t = 1$, yielding a matrix G_1 of order $2k$ satisfying $|ESS(G_1)| = q_{G_1}(1) = k2^{k-2}$,

and then Proposition 17 yields a cyclically symmetric matrix of that same number of ESS s, which we list in the table. Neither method worked sufficiently well for the primes $n \in \{17, 23\}$. So we tried to mimic what gave good results for $n \in \{14, 16, 18, 20, 22\}$ by enforcing symmetry also within the first and second half of the vector \mathbf{a} (i.e. within the vector \mathbf{x}), disregarding the two middle elements of \mathbf{a} . What we were able to obtain is still not overwhelming, but we have included the results for the sake of completeness. We do not know of any better alternatives for these orders.

Summarizing this, cyclically symmetric matrices are a good starting point to find matrices with many local maximizers. For many n our matrices give lower bounds for γ that are better than the ones in the existing literature.

Note that $U_5 \geq 5$ is not the best lower bound which exists, symmetric 5×5 -matrices are known which yield $U_5 \geq 6$, e.g. the matrix from Example 9(a).

matrix	n	\mathbf{x}	$ ESS(C_i) $	$ I $	n	matrix	$U_n \geq$	$\gamma \geq$	$ I $
C_1	10	13,19,25,7,5	50	4	11	$G(C_1, 1_9 2_1)$	70	1.47142	4
C_2	9	35,12,20,20	27	5	13	$G(C_2, 1_5 2_4)$	157	1.47542	5
C_3	10	93,35,53,53,53	35	6	14	$G(C_3, 1_6 2_4)$	233	1.47604	6
					16	$G(C_3, 1_4 2_6)$	536	1.48106	6
					17	$G(C_3, 1_3 2_7)$	784	1.47997	6
C_4	15	9,9,5,9,3,5,9	360	7	18	$G(C_4, 1_{12} 2_3)$	1164	1.48024	7
					19	$G(C_4, 1_{11} 2_4)$	1694	1.47891	7
C_5	21	15,15,7,15,15,7,7,15,7,15	4410	9	22	$G(C_5, 1_{20} 2_1)$	6300	1.48832	9
					23	$G(C_5, 1_{19} 2_2)$	9060	1.4861	9

Table 2: Base matrices C_i and symmetric matrices improving upon Table 1

Also for some other n we can do better than in Table 1 by dropping the requirement of cyclic symmetry, as will be shown now. We take an $n \times n$ matrix A from a small set of promising cyclically symmetric “base” matrices found by computer search (such as those listed in the left half of Table 2, promising meaning “a lot of ESS s”, or “quite a lot of ESS s of large support size”), with $B_i \in \{I_1 - E_1, I_2 - E_2, I_3 - E_3\}$ for $i \in [1:n]$, or A from the set $\{E_n - nI_n : n \geq 2\}$, with B_i “record holders” of smaller order, and apply Theorem 5, in case we are looking for general symmetric matrices, respectively Corollary 8, in case we are looking for matrices, such that all ESS s are global maximizers of (1), respectively Proposition 17, in case we are looking for cyclically symmetric matrices. The improvements we found for symmetric matrices of orders $n \in [6 : 24]$ are listed in the right half of Table 2. They are all in the spirit of Examples 9(a) and 11, where by $G(A, 1_\ell 2_m)$ we denote G_t in case that $t = 1$, and that $B_i = I_1 - E_1$ for $i \in [1:\ell]$ and $B_i = I_2 - E_2$ for $i \in [\ell+1:\ell+m]$. For all the listed orders we got best results by grouping the matrices $B_i = I_2 - E_2$ at the end of the diagonal, but we do not believe that this is a general principle. To illustrate our approach a bit more, we consider the order $n = 21$ in more detail. Had we not found the matrix with 4410 ESS s, our best result in Table 1 would have been a matrix congruent to G_t with $A = E_3 - 3I_3$ and $B_1 = B_2 = B_3$ equal to the matrix of order 7 in Table 1 yielding $14^3 = 2744$ ESS s. Good results for Table 2 would have been G_t with $A = E_2 - 2I_2$ and B_1, B_2 equal to the matrices of order 10 and 11 in Table 1 yielding $50 \cdot 66 = 3300$ ESS s, and the matrix from Example 13 with 3584 ESS s.

All the matrices from Tables 1 and 2 are contained in \mathcal{F} . We checked a stronger condition than that, namely if the system of linear equations in Proposition 21 has at most one solution. This was done for all $2^n - 1$ support sets per matrix. That way we verified that all matrices C_i from Table 2 and all matrices from Table 1, with the possible exception of the matrices of orders $n \in \{14, 16\}$, are contained in \mathcal{F}^d . Containment of the remaining matrices ($n \in \{14, 16\}$ from Table 1 and right half of Table 2) in \mathcal{F} was inferred via Corollary 8 from $\{C([4, 3, 2, 2, 3, 4, 0]), C([7, 5, 4, 4, 4, 5, 7, 0]), C_1, \dots, C_5\} \subset \mathcal{F}$. As regards our record holder of order 24: we checked that it belongs to \mathcal{F}^r .

Among the symmetric matrices with the most *ESSs* that we found for orders 5 to 24 there are some with only global maximizers of the corresponding StQP. These are the matrices of orders $n \in \{5, 6, 15, 18, 19, 20, 24\}$ given in Example 9(a) and listed in Tables 1 and 2.

We now provide some conclusions that we can draw from Theorem 15 and Tables 1 and 2.

Corollary 22. (a) $\min(\gamma, \gamma^{(r)}, \gamma^{(d)}, \gamma^{(g)}, \gamma^{(s)}, \gamma^{(c)}) \geq \sqrt[24]{15120} \approx 1.4933$,

(b) $U_n \geq \left(\sqrt[12]{105} \cdot \sqrt[13]{157} \right)^{\frac{n}{2}}$ for $n \geq 13$, where $\left(\sqrt[12]{105} \cdot \sqrt[13]{157} \right)^{\frac{1}{2}} \approx 1.4746$,

(c) $U_n \geq \left(\sqrt[22]{6300} \right)^n$ for $n \geq 420$, where $\sqrt[22]{6300} \approx 1.4883$.

4 Detailed study on cyclically symmetric matrices

4.1 General experimental setup

The theory detailed before enables us “more than just educated guesses” on the solution complexity of StQPs, which were impossible to obtain by naive random sampling or brute force enumeration: indeed, by these methods, it is highly unlikely to encounter even one of the interesting instances. To gain insight how a large and representative number of matrices with a potentially complex solution structure behave, we therefore developed the resulting refined mathematical experiments which provide an illustration and a statistical evaluation for cyclically symmetric matrices for three different n , namely $n \in \{6, 7, 19\}$.

All the matrices we generated and tested are of the form $C(\mathbf{a})$ where

$$\begin{aligned} \mathbf{a} &= (a, b, c, b, a, 0)^\top && \text{for } n = 6, \\ \mathbf{a} &= (a, b, c, c, b, a, 0)^\top && \text{for } n = 7, \text{ and} \\ \mathbf{a} &= (a, b, b, c, b, c, a, a, c, c, a, a, c, b, c, b, b, a, 0)^\top && \text{for } n = 19. \end{aligned} \tag{26}$$

We chose these three n primarily because they all have three degrees of freedom, $n \in \{6, 7\}$ naturally, and for $n = 19$ by using the restrictions described in Section 2.8; note that $\mathbf{M}_7^\top C(\mathbf{a}) \mathbf{M}_7 = C(\mathbf{a})$. We wanted to include the larger order $n = 19$ since such evaluation has (to our knowledge) never been published before.

Since multiplying a matrix with a positive factor does not change the number of *ESSs* and their structure, just the direction and not the length of \mathbf{a} is important. Therefore we used points on the unit sphere as input for \mathbf{a} , employing the following procedure: take a Lambert-projection of the unit sphere, which is an equal-area projection (to prevent graphical biases in the following illustrations) with $\theta \in [-\pi, \pi]$ on the horizontal and $\sin(\phi) \in [-1, 1]$ on the vertical axis, where (θ, ϕ) are polar-coordinates of the form $x = \cos(\phi) \cos(\theta)$, $y = \cos(\phi) \sin(\theta)$, $z = \sin(\phi)$.

We performed another coordinate transformation, building an orthonormal system such that the north pole of the sphere points into direction $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, i.e.

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & 1 & \sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \\ 0 & -2 & \sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The reason for this transformation lies in a remarkable symmetry around direction $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ for $n \in \{7, 19\}$. Just note that in both cases the transformation $M_2^T C(a) M_2$, while not leaving $C(a)$ invariant, has the same effect on $C(a)$ as cyclically permuting $a \rightarrow b \rightarrow c \rightarrow a$.

To approximate a uniform distribution on the sphere we generated regularly spaced data points on the θ - $\sin(\phi)$ -plane, giving 1201×401 data points. Note that points on the left and right boundary are calculated twice due to periodicity, and also we need only half of the points on the top and bottom boundary. After correcting for that, we end up with exactly 480000 data points. Moreover, in the case $n \in \{7, 19\}$, we only need to consider one third of those points, due to the above mentioned symmetry. The sets of *ESSs* and patterns of the others can be inferred from the ones of that one third. Still we actually did not make use of the latter fact, and took all 480000 points into consideration. “Rounding” these points to points with rational coordinates is the explanation for not every entry in Table 3 corresponding to $n \in \{7, 19\}$ being divisible by 3.

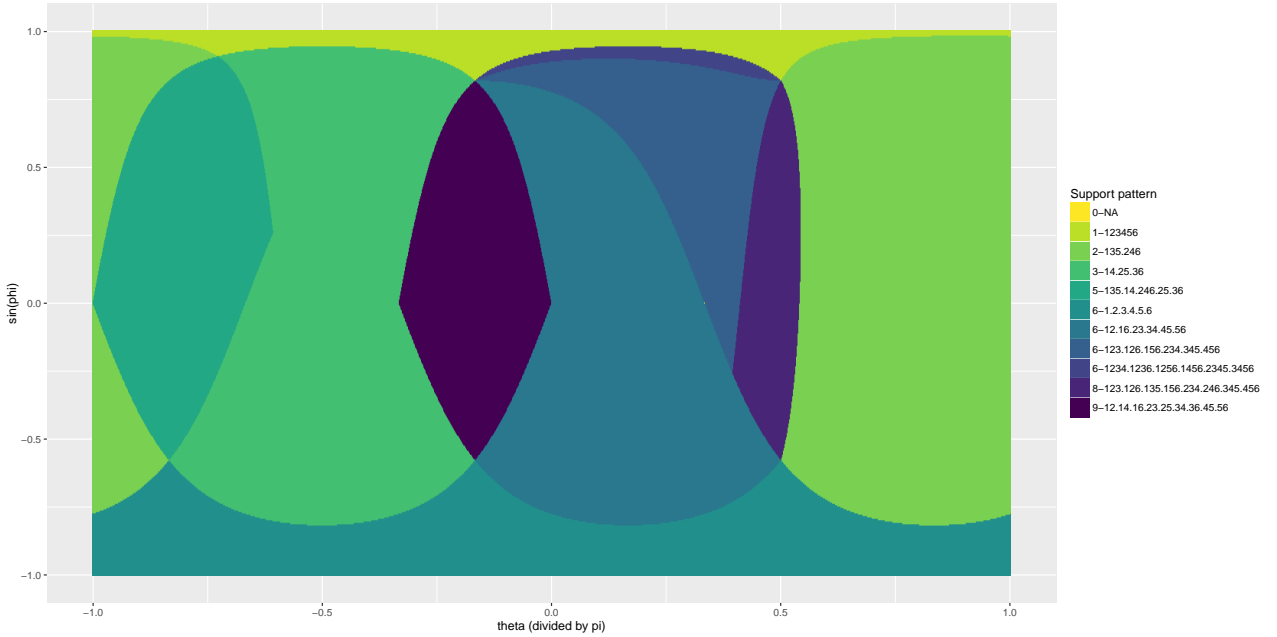


Figure 1: Whole search space for $n = 6$

Figure 1 shows *ESSs* and support-patterns for $n = 6$ plotted on the θ - $\sin(\phi)$ -map. The legend must be read like that: 2-135.246 means 2 *ESSs* with pattern $\{\{1, 3, 5\}, \{2, 4, 6\}\}$. The regions with many *ESSs* (meaning 8 and 9) take a quite large part in the middle of the map, that distinguishes them from the following plots, where only small regions for matrices with many *ESSs* can be found. There are some symmetries noticeable, but we will not pursue them further.

Figure 2 shows the support structure for $n = 7$, while Figure 3 shows the number of *ESSs* for $n = 19$. Here the number of different occurring patterns would be too large to illustrate them in a meaningful way. Both figures illustrate the whole search space on the left side, while the right side “zooms” into the plot to magnify an interesting region.

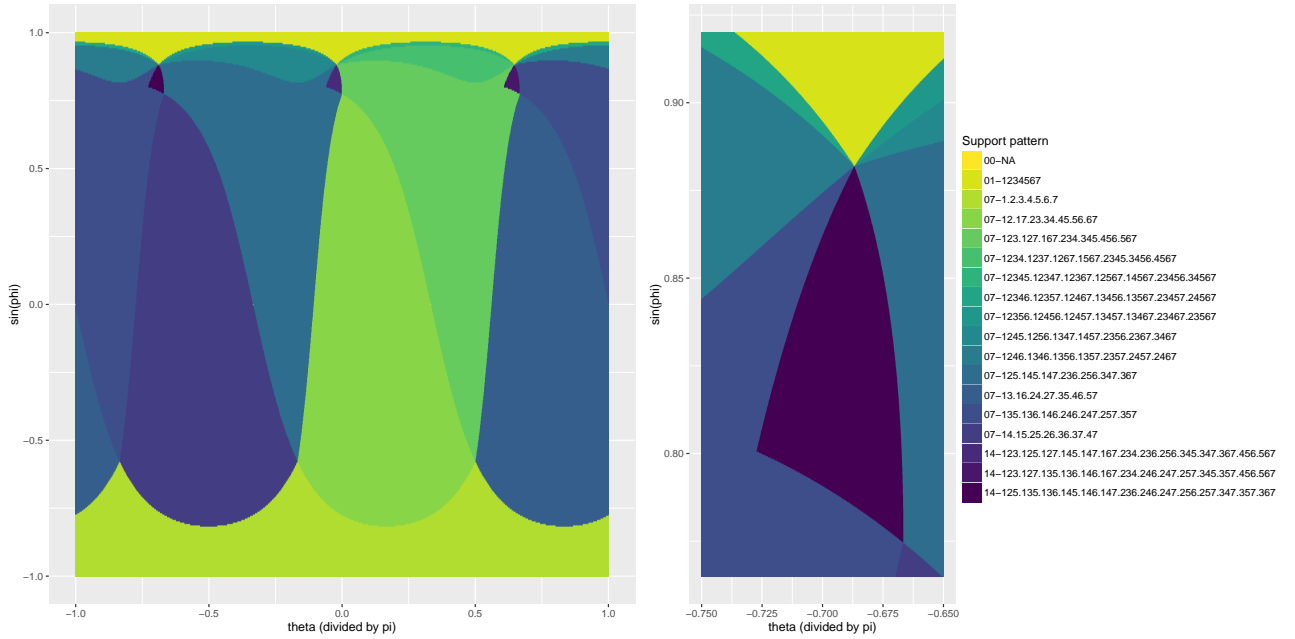


Figure 2: Whole search space and zoom-in for $n = 7$

It is interesting to see that all the regions for the different patterns are clearly separated (for all n), and they are connected for all patterns, except for zero *ESS*s. The regions with one *ESS* lie in the positive orthant (w.r.t. the standard basis) for all $n \in \{6, 7, 19\}$. The regions with zero *ESS*s have measure zero on the unit sphere, they form arcs and isolated points on the θ - $\sin(\phi)$ -plane. But these points are rarely hit due to the construction of the data, see also Section 4.2. Isolated points of zero *ESS*s can be seen in the plots upon a very close look.

The zoom-in in Figure 2 magnifies the top left region with 14 *ESS*s, in the sense that we want to analyze the structure of this region in more detail. We generated 1015×501 data points in the interval $\theta \in [-0.75\pi, -0.65\pi]$ and $\sin(\phi) \in [0.765, 0.92]$, which again simulates a uniform distribution. The separation of the different support patterns is clearly visible, with an outstanding point at approximately $(-0.68\pi, 0.88)$ where many of the separated regions meet. It would be an interesting further task to analyze the structure of such extraordinary points in detail, since we encountered them also for $n = 19$ as well as when experimenting with two and four degrees of freedom.

The zoom-in in Figure 3 again magnifies the top left region with many *ESS*s, but with a far higher magnification than in Figure 2, using $\theta \in [-0.9010867\pi, -0.90108655\pi]$ and $\sin(\phi) \in [0.95118965, 0.95118975]$. Only approximately 50000 data points were generated because the runtime increased heavily due to the high magnification. Our best result, 1444 *ESS*s, has not been found when scanning the whole search space, it is only visible in the zoom-in on the right side. Again it can be seen that many separated regions are meeting at one point.

4.2 A quantitative evaluation

Table 3 shows the frequency of each pattern size, summarized from the 480000 data points generated for each n as described above.

For $n = 6$ it is clear that the number of *ESS*s must be the sum of multiples of 2 and 3, since they are the prime

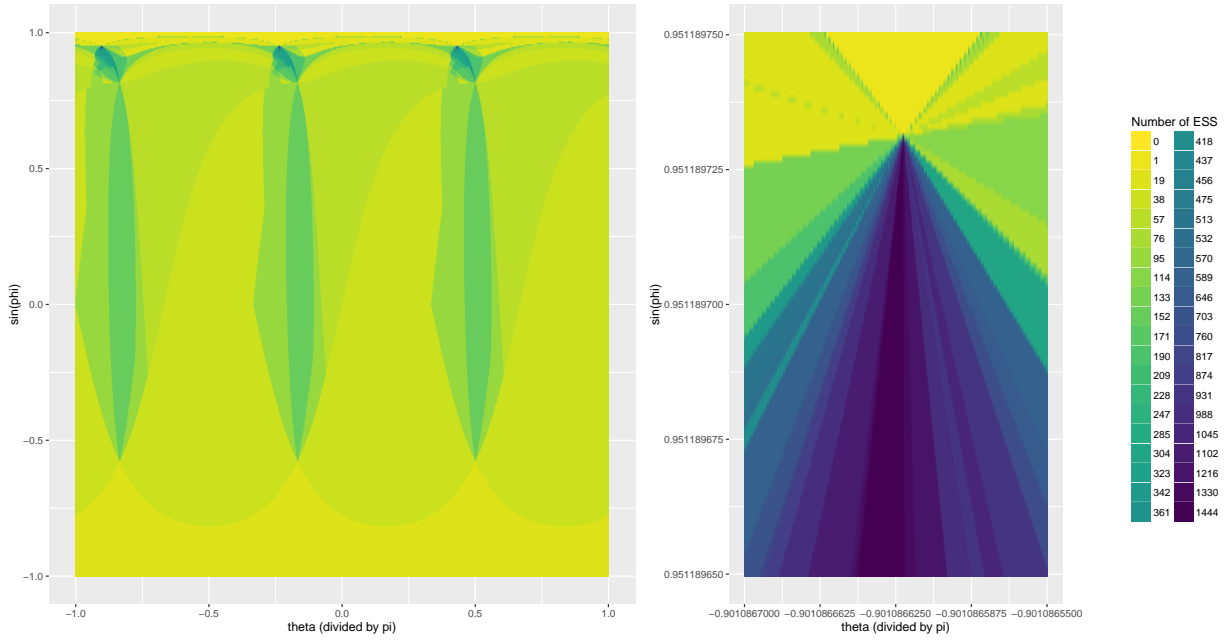


Figure 3: Whole search space and zoom-in for $n = 19$

n=6		n=7		n=19							
$ ESS $	count	$ ESS $	count	$ ESS $	count	$ ESS $	count	$ ESS $	count	$ ESS $	count
0	1	0	3	0	1	133	2073	304	453	646	3
1	16078	1	12804	1	4968	152	38181	323	564	760	9
2	115617	7	465897	19	65106	171	609	342	3	874	3
3	89006	14	1296	38	220787	190	1458	361	60	1102	3
5	43985			57	89343	209	234	418	3		
6	165105			76	1062	228	306	437	6		
8	16167			95	54510	247	24	475	9		
9	34041			114	207	285	12	532	3		

Table 3: Distribution of pattern sizes

factors. The distribution peaks at pattern size 6, while also many matrices were found with 2 and 3 ESS s. The frequency of the largest pattern size ($|ESS| = 9$), is also high compared to the $n \in \{7, 19\}$ ($|ESS| = 14$ and $|ESS| = 1102$).

For prime numbers $n = 7$ and $n = 19$ the number of ESS s is a multiple of n (or 0 or 1), leading to a quite limited spectrum of pattern sizes. For $n = 7$ the distribution concentrates most of its weight on pattern size 7, all other entries are almost neglectable. For $n = 19$ the spectrum has more variety. The distribution peaks at pattern size 38, and further spikes at pattern sizes 19, 57, 95 and 152 are noticeable, but every entry above 152 has almost no mass in the distribution.

For pattern size 1 the frequency is quite the same for $n = \{6, 7\}$ while it drops more than 50% for $n = 19$.

Figure 4 shows the distribution of the support size grouped by the number of ESS s. Interestingly, for the largest pattern sizes the support size is less than the expected $\left(\left\lfloor \frac{n}{2} \right\rfloor\right)$. This leads to the conjecture that either we did not find the maximal number of ESS s possible or that there are structural restrictions such that $\left(\left\lfloor \frac{n}{2} \right\rfloor\right)$ is not optimal.

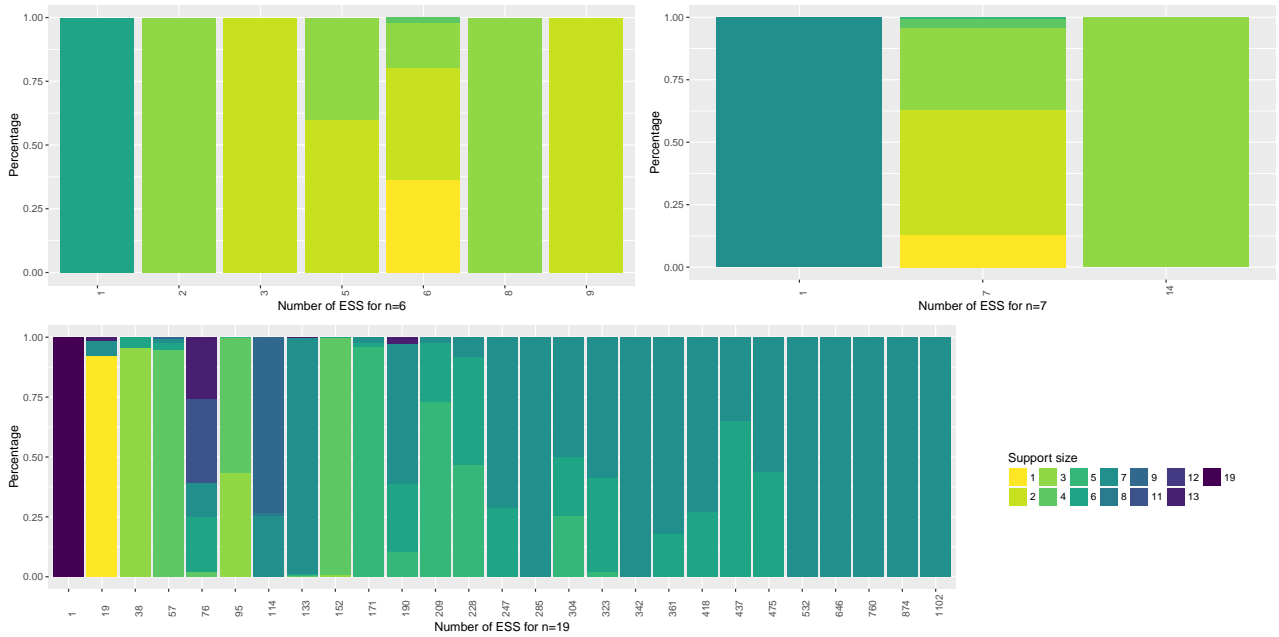


Figure 4: Distribution of the support size

For $n = 19$ support sizes 7 and 6 dominate for large pattern sizes, showing that for more than half of the different bins only these two support sizes are present.

Figure 5 sheds light on the outcome of CHECKSTAB from Section 3.1 grouped by the support size.

The most interesting feature in our opinion is the lack of *ESSs* verified by copositivity checks, all generated matrices are contained in \mathcal{F}_n . We never found any matrices $\notin \mathcal{F}_n$ in our experiments for $n \in \{6, 7\}$, but for $n = 19$ they exist, e.g. set $\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ then the resulting matrix has 19 *ESSs* all verified by copositivity.

Another interesting topic is the ratio of “True” and “False” regarding the different n and different support sizes. Almost all candidates of matrices of order 7 are verified as *ESSs*, while the opposite is true for $n = 19$. There support size 7 is an exception, more than half of the candidates are *ESSs* - an interesting fact considering the prevalence of support size 7 in matrices with many *ESSs*.

It is known that the number of *ESSs* with support size $n - 1$ is ≤ 2 , see [26], and this is the reason that for $n = 6$ support size 5 does not occur, since 5 and 6 are coprime. The same holds true for $n = 7$ with support size 6.

5 Conclusion

In this study, we improved upon existing lower bounds for the number of strict local solutions of the simplest problem class of NP-hard instances, the Standard Quadratic Optimization Problem, which is a central class with manifold direct applications.

We achieved this by new techniques to construct instances with a rich solution structure. Furthermore we provide extensive case studies where the systems of supports (the so-called patterns) of solutions are analyzed in detail. While by naive simulation most of the interesting patterns would not be encountered, since random instances have, with a high probability, quite sparse solutions, and likewise their expected numbers are considerably lower than in the worst case, we detected rather rare instances with a rich solution pattern.

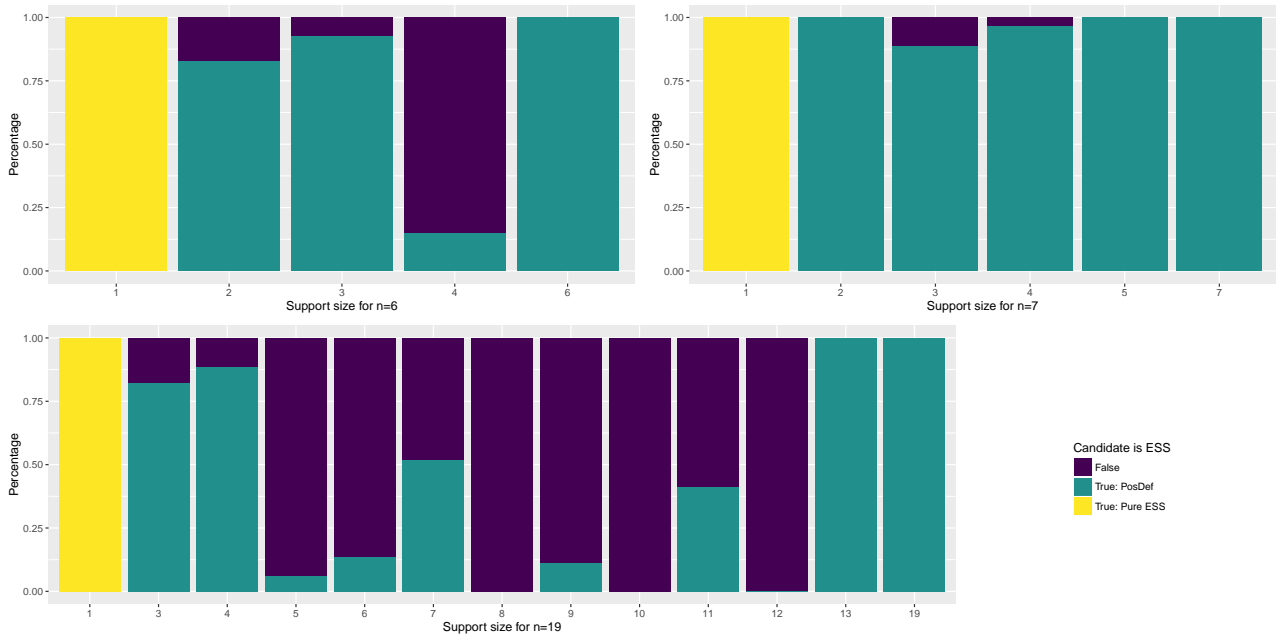


Figure 5: Distribution of the results of CHECKSTAB

Concentrating on these subsets of promising instances, we are able to give an empirical answer on the size distribution of supports of strict local solutions to the StQP and their patterns, complementing classical average-case analysis and enabling a new perspective on this notion.

References

- [1] David Timothy Bishop and Chris Cannings. Models of animal conflict. *Advances in Applied Probability*, 8(4):616–621, 1976.
- [2] Immanuel M. Bomze. Detecting all evolutionarily stable strategies. *J. Optim. Theory Appl.*, 75(2):313–329, 1992.
- [3] Immanuel M. Bomze. Regularity versus degeneracy in dynamics, games, and optimization: a unified approach to different aspects. *SIAM Rev.*, 44(3):394–414, 2002.
- [4] Immanuel M. Bomze and Etienne de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *J. Global Optim.*, 24(2):163–185, 2002.
- [5] Immanuel M. Bomze and Benedikt M. Pötscher. *Game theoretical foundations of evolutionary stability*, volume 324 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, 1989.
- [6] Immanuel M. Bomze. On standard quadratic optimization problems, *J. Global Optim.*, 13(4):369–387, 1998.
- [7] Immanuel M. Bomze, Werner Schachinger, and Reinhard Ullrich. From seven to eleven: completely positive matrices with high cp-rank. *Linear Algebra Appl.*, 459:208–221, 2014.
- [8] Immanuel M. Bomze, Werner Schachinger, and Reinhard Ullrich. New lower bounds and asymptotics for the cp-rank. *SIAM J. Matrix Anal. Appl.*, 36(1):20–37, 2015.
- [9] Mark Broom, Chris Cannings, and Glenn T. Vickers. On the number of local maxima of a constrained quadratic form. *Proc. R. Soc. Lond. A*, 443:573–584, 1993.
- [10] Mark Broom, Chris Cannings, and Glenn T. Vickers. Sequential methods for generating patterns of ESS’s. *J. Math. Biol.*, 32:597–615, 1994.

- [11] Mark Broom, Chris Cannings, and Glenn T. Vickers. ESS patterns: Adding pairs to an ESS. *Math. Biosci.*, 36(1):21–33, 1996.
- [12] Chris Cannings and Glenn Taylor Vickers. Patterns of ESS's II. *Journal of Theoretical Biology*, 132(4):409–420, 1988.
- [13] Xin Chen and Jiming Peng. New analysis on sparse solutions to random Standard Quadratic Optimization Problems and extensions. *Math. Oper. Res.*, 40(3):725–738, 2015.
- [14] Xin Chen, Jiming Peng, and Shuzhong Zhang. Sparse solutions to random Standard Quadratic Optimization Problems. *Math. Programming*, 141(1):273–293, 2013.
- [15] Etienne de Klerk. The complexity of optimizing over a simplex, hypercube or sphere: a short survey. *Central Eur. J. Oper. Res.*, 16(2):111–125, 2008.
- [16] Josef Hofbauer and Karl Sigmund. *The theory of evolution and dynamical systems: mathematical aspects of selection*. Cambridge University Press, Cambridge UK, 1988.
- [17] Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge University Press, Cambridge UK, 1998.
- [18] Spyros Kontogiannis and Paul Spirakis. Counting stable strategies in random evolutionary games. In X. Deng and D. Du, editors, *Proceedings of the 16th Annual International Symposium on Algorithms and Computation (ISAAC 2005)*, volume 3827 of *LNCS*, pages 839–848. Springer, Heidelberg, 2006.
- [19] Spyros C. Kontogiannis and Paul G. Spirakis. On the support size of stable strategies in random games. *Theoretical Computer Science*, 410(8-10):933–942, 2009.
- [20] Spyros C. Kontogiannis and Paul G. Spirakis. Well supported approximate equilibria in bimatrix games. *Algorithmica*, 57(2):653–667, 2010.
- [21] John W. Moon and Leo Moser. On cliques in graphs. *Israel J. Math.*, 3(2):23–28, 1965.
- [22] Panos M. Pardalos and Stephen A. Vavasis. Quadratic programming with one negative eigenvalue is NP-hard. *J. Global Optim.*, 1(1):15–22, 1991.
- [23] Samuel Rota-Bulò and Immanuel M. Bomze. Infection and immunization: a new class of evolutionary game dynamics. *Games Econ. Behav.*, 71(special issue in honor of J.F.Nash, jr.):193–211, 2011.
- [24] William H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge, Mass., 2010.
- [25] Glenn Taylor Vickers and Chris Cannings. On the number of stable equilibria in a one-locus, multi-allelic system. *Journal of Theoretical Biology*, 131(3):273–277, 1988.
- [26] Glenn Taylor Vickers and Chris Cannings. Patterns of ESS's I. *Journal of Theoretical Biology*, 132(4):387–408, 1988.
- [27] Jörgen W. Weibull. *Evolutionary game theory*. MIT Press, Cambridge, Mass., 1995.