Journal of Global Optimization **13:** 369–387, 1998.

On Standard Quadratic Optimization Problems

IMMANUEL M. BOMZE*
I.S.O.C., University of Vienna, Wien, Austria

(Accepted in final form 30 June 1998)

Abstract. A standard quadratic optimization problem (QP) consists of finding (global) maximizers of a quadratic form over the standard simplex. Standard QPs arise quite naturally in copositivity-based procedures which enable an escape from local solutions. Furthermore, several important applications yield optimization problems which can be cast into a standard QP in a straightforward way. As an example, a new continuous reformulation of the maximum weight clique problem in undirected graphs is presented which considerably improves previous attacks both as numerical stability and interpretation of the results are concerned. Apparently also for the first time, an equivalence between standard QPs and QPs on the positive orthant is established. Also, a recently presented global optimization procedure (GENF - genetical engineering via negative fitness) is shortly reviewed.

Key words: Maximum clique, Optimality conditions, Portfolio selection, Quadratic programming

1. Introduction

A standard quadratic optimization problem (QP) consists of finding (global) maximizers of a quadratic form over the standard simplex, i.e. we consider (global) optimization problems of the form

$$\mathbf{x}'A\mathbf{x} \to \max! \quad \text{subject to } \mathbf{x} \in \Delta\,, \tag{1.1}$$

where A is an arbitrary symmetric $n \times n$ matrix; a$'$ denotes transposition; and Δ is the standard simplex in n-dimensional Euclidean space $\mathbb{R}^n$,

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n : x_i \geq 0 \text{ for all } i \in \nu,\ \mathbf{e}'\mathbf{x} = 1\}\,,$$

where $\mathcal{V} = \{1, \dots, n\}$ (of course, the region $\{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} \geq \mathbf{o}\,,\ \mathbf{e}'\mathbf{y} \leq 1\}$ can always be represented by $\Delta \subseteq \mathbb{R}^{n+1}$, introducing a slack variable). Here, $\mathbf{e} = \sum_{i\in\mathcal{V}} \mathbf{e}_i = [1, \dots, 1]'$ while $\mathbf{e}_i$ denotes the i-th standard basis vector in $\mathbb{R}^n$.

Note that the maximizers of (1.1) remain the same if A is replaced with $A+\gamma\mathbf{ee}'$ where γ is an arbitrary constant. So without loss of generality assume henceforth that all entries of A are non-negative. Furthermore, the question of finding maximizers of a general quadratic function $\mathbf{x}'Q\mathbf{x} + 2\mathbf{c}'\mathbf{x}$ over Δ can be homogenized in a similar way by considering the rank-two update $A = Q + \mathbf{ec}' + \mathbf{ce}'$ in (1.1) which has the same objective values.

* The author benefited from valuable suggestions by an anonymous referee.

Of course, quadratic optimization problems like (1.1) – even regarding the detection of local solutions – are NP-hard [32]. Nevertheless, there are several exact procedures which try to exploit favourable data constellations in a systematic way, and to avoid the worst-case behaviour whenever possible. One prototypical example for this type of algorithms is specified in Section 3 below.

The paper is organized as follows: Section 2 contains a concise review on the role of copositivity in characterizations of global optimality in general QPs. Theoretically in all cases, and for many important applications also practically, checking copositivity can be reformulated in terms of standard QPs, providing support for the justification of the terminology. Section 3 deals with an iterative procedure proposed in [14], which consists of two parts. At first, a local solution of (1.1) is generated by following the paths of feasible points provided by a dynamical system borrowed from evolutionary modelling; in the second step, the procedure escapes from an inefficient local maximizer in a way such that improvement in the objective is guaranteed. This step has also an interesting interpretation in terms of genetic evolution models.

In Section 4, an apparently new equivalence result is presented which shows under which conditions exact penalization is possible in order to get rid of the equality constraint defining the standard simplex: all local solutions (and the global ones, respectively) of the resulting QP on the positive orthant exactly correspond to local (and global) solutions of the standard QP. Section 5 is devoted to an important application, namely to find a clique of maximum weight in an undirected graph with weights on the nodes. Here, an exact one-to-one correspondence between the local/global solutions of a whole class of standard QPs, and maximal/maximum weight cliques is established for the first time, by introducing a new class of quadratic forms which can be viewed as a regularization of another class recently introduced by Ramana and his coworkers [29] who followed an idea of Lovász. In contrast to their approach, the proposed class avoids three shortcomings: the possibility of jamming along paths generated by a popular local optimization phase borrowed from population genetics; the occurrence of spurious solutions which cannot be traced back to the sought cliques; and the lack of one-to-one correspondence between local solutions and maximal cliques. Finally, Section 6 shortly indicates how (also indefinite) standard QPs can emerge in portfolio selection models along the lines of Markowitz [40].

2. From general to standard quadratic optimization problems

Consider the general quadratic maximization problem

$$f(\mathbf{x}) = \tfrac{1}{2}\mathbf{x}'Q\mathbf{x} + \mathbf{c}'\mathbf{x} \to \max! \quad \text{subject to} \quad \mathbf{x} \in M\,, \tag{2.2}$$

where $M = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ with A an $m \times n$ matrix and Q a symmetric $n \times n$ matrix. We now review a well-known characterization for global optimality of a Karush-Kuhn-Tucker point $\overline{\mathbf{x}}$ for (2.2) (see, e.g. [10, 15]). Using a novel approach via low-rank updates, the formulation presented here might be more appealing than previous ones.

First, add a trivial non-binding constraint, i.e. the most elementary strict inequality $0 < 1$, to obtain slacks $\bar{\mathbf{u}}$ as follows: denote by $\mathbf{a}_i'$ the i-th row of A and put $\mathbf{a}_0 = \mathbf{o}$. Similarly put $b_0 = 1$ and enrich $\overline{A} = [\mathbf{a_0}|A']' = [\mathbf{o}, \mathbf{a}_1, \ldots, \mathbf{a}_m]'$ as well as $\bar{\mathbf{b}} = [b_0|\mathbf{b}']' = [1, b_1, \ldots, b_m]'$. Finally, define $\bar{\mathbf{u}} = \bar{\mathbf{b}} - \overline{A}\bar{\mathbf{x}} \geq \mathbf{o}$.

Then perform, for any $i \in \{0, \ldots, m\}$, a rank-one update of $\overline{A}$ and a rank-two update of Q, using the current gradient $\mathbf{g} = \nabla f(\bar{\mathbf{x}}) = Q\bar{\mathbf{x}} + \mathbf{c}$ of the objective:

$$D_i = \bar{\mathbf{u}}\,\mathbf{a}_i' - \bar{u}_i \overline{A} \qquad \text{and} \qquad Q_i = -\mathbf{a}_i\,\mathbf{g}' - \mathbf{g}\,\mathbf{a}_i' - \bar{u}_i Q\,.$$

This gives a symmetric $n \times n$ matrix Q_i and a matrix D_i which is effectively $m \times n$ since its ith row is zero.

Denoting by $J(\bar{\mathbf{x}}) = \{0, \ldots, m\} \setminus I(\bar{\mathbf{x}})$ the set of all non-binding constraints, the following result is proved in [10]:

THEOREM 1 *A Karush-Kuhn-Tucker point $\bar{\mathbf{x}}$ of (2.2) is a global solution if and only if for all $i \in J(\bar{\mathbf{x}}) = \{i : \bar{u}_i > 0\}$,*

$$\mathbf{v}' Q_i \mathbf{v} \geq 0 \qquad \textit{if} \qquad D_i \mathbf{v} \geq \mathbf{o}\,,$$

i.e. iff Q_i are $D_i^{-1}(\mathbb{R}^m_+)$-copositive. If $\mathbf{v}' Q_i \mathbf{v} < 0$ for some $D_i \mathbf{v} \geq \mathbf{o}$, then

$$\tilde{\mathbf{x}} = \bar{\mathbf{x}} + \lambda \mathbf{v}$$

is an improving feasible point for $\lambda = \bar{u}_i / (\bar{\mathbf{a}}_i' \mathbf{v})$ (if $i = 0$, i.e. $\lambda = 1/0$, this means that (2.2) is unbounded).

Since determining whether or not a matrix Q is $D^{-1}(\mathbb{R}^m_+)$-copositive amounts to the question whether the following homogeneous problem

$$\mathbf{v}'(-Q)\mathbf{v} \to \max! \qquad \text{subject to } D\mathbf{v} \geq \mathbf{o}\,,$$

is unbounded, it is only a small step from a copositivity check to a standard QP. Indeed, for 'nice' D, this problem is easily decomposable (see [25] and cf. [23]) into a few standard copositivity problems of the form

$$\mathbf{x}' A \mathbf{x} \to \max! \qquad \text{subject to } \mathbf{x} \geq \mathbf{o}\,,$$

where the constraint $\mathbf{e}'\mathbf{x} = \sum_i x_i = 1$ can be added without loss of generality, rendering a standard QP. In fact, in order to determine an improving feasible direction as in Theorem 1, it is not necessary to solve the latter problem to optimality, but rather sufficient to determine a feasible point $\mathbf{x} \in \Delta$ with $\mathbf{x}' A \mathbf{x} > 0$.

As an interesting aside, if the original problem (2.2) is itself already a standard QP, then all copositivity checks from Theorem 1 can be reduced into a single one (see, e.g. Theorem 7 of [13]): if $\overline{\mathbf{x}} \in \Delta$ is any feasible point, then $\overline{\mathbf{x}}$ is a global maximizer of $\mathbf{x}'Q\mathbf{x}$ over Δ if and only if the matrix

$$\overline{Q} = (\overline{\mathbf{x}}'Q\overline{\mathbf{x}})\mathbf{e}\mathbf{e}' - Q$$

is $\mathbb{R}^n_+$-copositive. Of course, if one wants to reduce this to a standard QP as above, one arrives at a trivial reformulation of global optimality, so that efficient reduction strategies are still necessary to obtain an improving feasible direction. In the context of checking copositivity, this can be accomplished by means of block pivoting as in [12] where emphasis is laid upon reducing the dimensionality of the problem at the cost of creating several subproblems of the same type. Transformed into the language of standard QPs [14], one arrives then at the G.E.N.F.* approach [18] described in the following section.

3. Global optimization of standard QPs: an evolutionary approach

First we concentrate on the evolutionary approach to local solutions of standard QPs (1.1). To this end, consider the following dynamical system operating on Δ:

$$\dot{x}_i(t) = x_i(t)[(A\mathbf{x}(t))_i - \mathbf{x}(t)'A\mathbf{x}(t)], \qquad i \in \mathcal{V}, \tag{3.3}$$

where a dot signifies derivative w.r.t. time t, and a discrete time version

$$x_i(t+1) = x_i(t)\frac{(A\mathbf{x}(t))_i}{\mathbf{x}(t)'A\mathbf{x}(t)}, \qquad i \in \mathcal{V}. \tag{3.4}$$

The *stationary points* under (3.3) and (3.4) coincide, and all local solutions of (1.1) are among these (see Theorem 2 below). Of course, there are quite many stationary points, e.g. all vertices $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of Δ. However, only those $\mathbf{x}$ are serious candidates for strict local solutions which are asymptotically stable, which means that every solution to (3.3) or (3.4) which starts close enough to $\mathbf{x}$, will converge to $\mathbf{x}$ as $t \to \infty$.

Both (3.3) and (3.4) arise in population genetics under the name *selection equations* where they are used to model time evolution of haploid genotypes, A being the (symmetric) fitness matrix, and $x_i(t)$ representing the relative frequency of allele i in the population. The Fundamental Theorem of Selection states that average fitness, i.e. the objective function $\mathbf{x}(t)'A\mathbf{x}(t)$ is (strictly) increasing over time along trajectories [24], and moreover every trajectory $\mathbf{x}(t)$ converges to a stationary point [38].

To formulate the results, we need some notions and notations. First, consider the *generalized Lagrangian*

$$L(\mathbf{x}; \lambda, \mu) = \frac{1}{2}\mathbf{x}'A\mathbf{x} + \lambda'\mathbf{x} + \mu(\mathbf{e}'\mathbf{x} - 1)$$

* Genetic Engineering via Negative Fitness – I owe this acronym to P. Marcotte (personal communication at the Trier WoGO '97).

of problem (1.1) where the multipliers λ_i and μ may have arbitrary sign. Call a critical point $\mathbf{x}$ of the generalized Lagrangian a *generalized* Karush-Kuhn-Tucker point if $L(\mathbf{x}; \lambda, \mu) = \frac{1}{2}\mathbf{x}'A\mathbf{x}$ irrespective of the sign of λ_i.

Next recall that a point $\mathbf{x} \in \Delta$ is said to be a *(symmetric) Nash (equilibrium) strategy* if and only if $\mathbf{y}'A\mathbf{x} \leq \mathbf{x}'A\mathbf{x}$ for all $\mathbf{y} \in \Delta$. Furthermore, a Nash strategy $\mathbf{x}$ is said to be a *neutrally stable strategy* (NSS) if and only if

$$\mathbf{y}'A\mathbf{x} = \mathbf{x}'A\mathbf{x} \quad \text{implies } \mathbf{x}'A\mathbf{y} \geq \mathbf{y}'A\mathbf{y}\,, \tag{3.5}$$

and an *evolutionarily stable strategy* (ESS) if and only if the inequality in (3.5) is strict for $\mathbf{y} \neq \mathbf{x}$.

For ease of reference, let us repeat now the characterization results from [16] which links three different fields: optimization theory, evolutionary game theory, and qualitative theory of dynamical systems.

THEOREM 2 *Let $A = A'$ be an arbitrary symmetric $n \times n$ matrix and $\mathbf{x} \in \Delta$. Consider the following properties:*

(a1) $\mathbf{x}$ is an evolutionarily stable strategy;
(a2) $\mathbf{x}$ is a strict local solution of (1.1);
(a3) $\mathbf{x}$ is an asymptotically stable stationary point of (3.3) and (3.4);
(b1) $\mathbf{x}$ is a neutrally stable strategy;
(b2) $\mathbf{x}$ is a local solution of (1.1);
(c1) $\mathbf{x}$ is a Nash strategy;
(c2) $\mathbf{x}$ is a Karush-Kuhn-Tucker point for (1.1);
(d1) $\mathbf{x}$ is a stationary point under (3.3) or (3.4);
(d2) $\mathbf{x}$ is a generalized Karush-Kuhn-Tucker point for (1.1).

Then the following implications and equivalences hold true:
(a1) $\Leftrightarrow$ (a2) $\Leftrightarrow$ (a3) $\Rightarrow$ (b1) $\Leftrightarrow$ (b2) $\Rightarrow$ (c1) $\Leftrightarrow$ (c2) $\Rightarrow$ (d1) $\Leftrightarrow$ (d2).

Note that a box-constrained QP (cf. the following section) in n variables can in general have up to 2^n strict local maximizers, whereas a standard QP has 'only'

$$\binom{n}{\lfloor n/2 \rfloor} \approx 2^n/\sqrt{\pi n/2}\,. \tag{3.6}$$

This is an important consequence of the characterization of strict solutions of (1.1) in terms of the ESS property: the system of supports $S = \{i \in \mathcal{V} : \overline{x}_i > 0\}$ of ESSs $\overline{\mathbf{x}}$ forms an antichain w.r.t. set inclusion $\subseteq$, yielding as an upper bound for the number of ESSs [11] the left-hand expression in (3.6), while the right-hand one is an approximation for large n using Stirling's formula. To formulate a sharpening of the antichain property, we need some more notation: for a subset $S \subseteq \mathcal{V}$, we shall denote the face of Δ corresponding to S by

$$\Delta_S = \{\mathbf{x} \in \Delta : x_i = 0 \text{ if } i \notin S\} \tag{3.7}$$

and its relative interior by

$$\Delta_S^\circ = \{\mathbf{x} \in \Delta_S : x_i > 0 \text{ if } i \in S\}\,. \tag{3.8}$$

The proofs of the following results can be found in [13, 18].

THEOREM 3 *If no principal minor of $A = A'$ vanishes, then with probability one any trajectory of (3.3) converges to a strict local solution $\overline{\mathbf{x}}$ of (1.1). Furthermore, if $\overline{\mathbf{x}} \in \Delta_S^\circ$ (i.e. if $S = \{i \in \mathcal{V} : \overline{x}_i > 0\}$), then*
(a) $\mathbf{y}'A\mathbf{y} < \overline{\mathbf{x}}'A\overline{\mathbf{x}}$ for all $\mathbf{y} \in \Delta_S$ with $\mathbf{y} \neq \overline{\mathbf{x}}$;
(b) Δ_S° is contained in the basin of attraction of $\overline{\mathbf{x}}$.

The dynamical systems (3.3) and (3.4) are frequently called *replicator dynamics*, and are well suited for implementation in practical applications, see [14, 16, 49]. This is reflected also in theory by the result that (3.3) is most efficiently approaching fixed points in the sense that it is a Shahshahani gradient system [53], cf. the proof of Proposition 6 below. The discrete time version (3.4) also corresponds to a particular instance of an algorithm widely popular in computer vision. These *relaxation labeling processes* are closely related to artificial neural network learning systems, and have found applications in a variety of practical tasks, e.g. to solve certain labeling problems arising in the 3-D interpretation of ambiguous line drawings [34, 48, 52]. Furthermore, dynamics (3.4) belongs to a class of dynamical systems investigated in [6, 7], which has proven to be useful in the speech recognition domain [36].

Although strictly increasing objective values are guaranteed as we follow trajectories under (3.3) or (3.4), we could get stuck in an inefficient local solution of (1.1). A global optimization procedure therefore must incorporate a decision maker at a higher level than ‘blind’ nature. One attempt is based on the reduction of problem dimension at the cost of generating a series of subproblems, which seems to be a promising approach in view of the NP-hardness in quadratic programming. This procedure is circumscribed by *genetic engineering via negative fitness* (G.E.N.F.) because of the following interpretation from [18].

From Theorem 3, a strict local solution $\overline{\mathbf{x}}$ of (1.1) must be a global one if all $\overline{x}_i > 0$. Consequently, at an inefficient local solution necessarily $\overline{x}_i = 0$ for some i. In the usual genetic interpretation, this means that some alleles die out during the selection process, and these are therefore unfit in the environment currently prevailing. The escape step now artificially re-introduces some alleles which would have gone extinct during the natural selection process. This is done via the negative fitness approach: remove all alleles which are not unfit, i.e. all $i \in S = \{i \in \mathcal{V} : \overline{x}_i > 0\}$. Then determine fitness *minimizers* in the reduced problem, i.e. consider problem (1.1) with A replaced by

$$\overline{A} = [\gamma_S - a_{ij}]_{i,j \in \mathcal{V}\setminus S}$$

where $\gamma_S = \max_{i,j\in\mathcal{V}\setminus S} a_{ij}$ is the maximum fitness of all extinct alleles. After having obtained a local solution $\mathbf{y}$ of this auxiliary problem, put

$$T = \{j \in \mathcal{V} \setminus S : y_j > 0\}$$

which can be viewed as the set of 'truly unfit' alleles. Now the following result can be shown [14]:

THEOREM 4 *Suppose $\overline{\mathbf{x}}$ is a local solution to the master problem (1.1) with surviving allele set $S = \{i \in \mathcal{V} : \overline{x}_i > 0\}$. Pick a disjoint set T of size $m \geq 1$ by 'negative genetic engineering' as above. For all $s \in S$ and $t \in T$, replace a_{si} with a_{ti}, remove all (other) unfit $j \in T$. Consider the reduced problem $\mathcal{P}_{t\to s}$, i.e. problem (1.1) in $n - m$ variables for the so obtained matrix $A_{t\to s}$.*
Then $\overline{\mathbf{x}}$ is a global solution to the master problem (1.1) if and only if for all $(s, t) \in S \times T$, the maximum of $\mathcal{P}_{t\to s}$ does not exceed the current best value $\overline{\mathbf{x}}' A\overline{\mathbf{x}}$. In the negative, i.e. if $\mathbf{u}' A_{t\to s}\mathbf{u} > \overline{\mathbf{x}}' A\overline{\mathbf{x}}$ for some $\mathbf{u} \in \mathbb{R}^{n-m}$ in the standard simplex, and if $j \in T$ is chosen such that

$$\sum_{p\notin T\cup\{s\}} a_{jp}u_p + \frac{1}{2}a_{jj}u_s \geq \sum_{p\notin T\cup\{s\}} a_{qp}u_p + \frac{1}{2}a_{qq}u_s \quad \textit{for all } q \in T\ ,$$

then a strictly improving feasible point $\tilde{\mathbf{x}}$ is obtained as follows:

$$\tilde{x}_q = \begin{cases} u_t & \textit{if } q = j\ , \\ 0 & \textit{if } q \in T \cup \{s\} \setminus \{j\}\ , \\ u_q & \textit{if } q \in \mathcal{V} \setminus T\ . \end{cases}$$

In view of the possible combinatorial explosion in effort with increasing number of variables, this dimension reducing strategy seems to be promising: if k is the size of S, the above result yields a series of km standard QPs in $n - m$ variables rather than in n. We are now ready to describe the algorithm which stops after finitely many repetitions, since it yields strict local solutions with strictly increasing objective values (cf. Theorem 3).

Algorithm

1. Start with $\mathbf{x}(0) = [1/n, \ldots, 1/n]$ or nearby, iterate (3.4) until convergence;
2. the limit $\overline{\mathbf{x}} = \lim_{t\to\infty} \mathbf{x}(t)$ is a strict local solution with probability one; call the escape procedure of Theorem 4 to improve the objective, if possible; denote the improving point $\tilde{\mathbf{x}}$;
3. repeat 1, starting with $\mathbf{x}(0) = \tilde{\mathbf{x}}$.

Of course, it remains to discuss what happens if the assumptions of Theorem 3 are not met. In this case, theory still guarantees convergence to a critical point of problem (1.1), but this point is not necessarily asymptotically stable, or neutrally stable, and therefore the result of step 2. above need not be a (strict) local solution to (1.1). The usual numerical cure (restart in a randomly chosen point close by) helps if there are only finitely many stationary points, but else jamming cannot be avoided with absolute certainty (see, e.g. Example 2.2 in [32], p. 58; and Figure 6, case 1 in [9] for a phase portrait under the replicator dynamics). Another case where too small perturbations could fail because there is a continuum of stationary points can be generated with the matrix

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}.$$

Here a trajectory $\mathbf{x}(t)$ in the relative interior of Δ exists which converges to $\mathbf{x} = [\frac{1}{2}, \frac{1}{2}, 0]'$. However, $\mathbf{x}'A\mathbf{x} = 2 < 2 + 6\varepsilon^2 = (\mathbf{x} + \varepsilon\mathbf{u})'A(\mathbf{x} + \varepsilon\mathbf{u})$ where $\mathbf{u} = [-2, 1, 1]'$, and $\mathbf{x} + \varepsilon\mathbf{u} \in \Delta$ if $0 < \varepsilon < \frac{1}{4}$, hence $\mathbf{x}$ is no local solution. With a high probability, perturbing $\mathbf{x}$ (e.g. to $\mathbf{x} + \varepsilon\mathbf{u}$ with arbitrarily small $\varepsilon > 0$) yields immediately the global solution $\mathbf{x}^* = [0, \frac{1}{2}, \frac{1}{2}]'$, but with (small) positive probability, namely if a starting point in the sector left and below the trajectory $\mathbf{x}(t)$ (see Figure 6, case 22 in [9]) is chosen, the process again gets jammed near $\mathbf{x}$.

Admitting only a finite number of stationary points is a robust feature among all possible replicator dynamics with arbitrary symmetric matrices A. But in applications with a special structure like the maximum clique problem one cannot resort to this argument. Indeed, a frequently used type of matrices lacks this property for certain graphs, and Example 2.2 in [32] is of this type. See the end of Section 5 for more details.

4. Standard QPs and quadratic optimization over the positive orthant

This section proposes an equivalence result which establishes exactness of penalization w.r.t. the constraint $\mathbf{e}'\mathbf{x} = 1$.

As a motivating introduction, we consider a box-constrained QP (BCQP), i.e. problem (2.2) with the special feasible set $M = \{\mathbf{x} : a_i \leq x_i \leq b_i\,,\ \text{all}\ i \in \mathcal{V}\}$. A recent survey of BCQP is [26]. For ease of exposition, let us now switch to minimization problems. Every BCQP can be written in the following form

$$\frac{1}{2}\mathbf{y}'Q\mathbf{y} + \mathbf{c}'\mathbf{y} \to \min! \qquad \text{subject to } -\mathbf{e} \leq \mathbf{y} \leq \mathbf{e}\,. \tag{4.9}$$

Here $M = [-1, 1]^n$, and following the nice homogenization idea of Ye [55], one can get rid of the linear term, at the cost of introducing an additional variable: indeed, consider the quadratic form in $(t, \mathbf{y}) \in [-1, 1]^{n+1}$

$$\overline{f}(t, \mathbf{y}) = \frac{1}{2}\mathbf{y}'Q\mathbf{y} + t\mathbf{c}'\mathbf{y}$$

Then one minimizer $(t^*, \mathbf{y}^*)$ of $\overline{f}$ over $[-1, 1]^{n+1}$ always satisfies $t^* = -1$ (if $\mathbf{c}'\mathbf{y}^* \geq 0$) or $t^* = 1$ (if $\mathbf{c}'\mathbf{y}^* \leq 0$), so that either $\mathbf{y}^*$ or $-\mathbf{y}^*$ is a solution to (4.9).

Hence we can study without loss of generality the homogeneous BCQP

$$\frac{1}{2}\mathbf{y}'Q\mathbf{y} \to \min ! \qquad \text{subject to } -\mathbf{e} \le \mathbf{y} \le \mathbf{e}\,. \tag{4.10}$$

Now shift the feasible box M by putting $\mathbf{p} = \frac{1}{2}(\mathbf{e} + \mathbf{y})$ which yields a BCQP on $M = [0, 1]^n$. Now, if Q has only non-negative entries, we arrive after dropping unnecessary variables and rescaling at the BCQP

$$\frac{1}{2}\mathbf{p}'C\mathbf{p} - \mathbf{e}'\mathbf{p} \to \min ! \qquad \text{subject to } \mathbf{o} \le \mathbf{p} \le \mathbf{b}\,. \tag{4.11}$$

These transformations lead us to investigate the orthant QP

$$h(\mathbf{p}) = \frac{1}{2}\mathbf{p}'C\mathbf{p} - \mathbf{e}'\mathbf{p} \to \min ! \qquad \text{subject to } \mathbf{p} \ge \mathbf{o}\,. \tag{4.12}$$

Note that $\mathbf{p} = \mathbf{o}$ can never be a local solution of (4.11) and (4.12). Further, let us assume that the objective is strictly convex along rays emanating from the origin, which amounts to imposing strict $\mathbb{R}^n_+$-copositivity of C:

$$\mathbf{p}'C\mathbf{p} > 0 \qquad \text{for all } \mathbf{p} \ge \mathbf{o} \text{ with } \mathbf{p} \ne \mathbf{o}\,.$$

This property implies that also problem (4.12) is bounded from below.

THEOREM 5 *Assume that C is strictly $\mathbb{R}^n_+$-copositive. Then local and global solutions of (4.12) and (1.1) with $A = \gamma\mathbf{e}\mathbf{e}' - C$ are related as follows (again choose γ so that A has no negative entries):*

(a) If $\overline{\mathbf{x}} \in \Delta$ is a local solution of (1.1), i.e., a local minimizer of $g(\mathbf{x}) = \mathbf{x}'C\mathbf{x}$ on Δ, then $\overline{\mathbf{p}} = (1/g(\overline{\mathbf{x}}))\overline{\mathbf{x}}$ is a local solution of of (4.12).

(b) If $\overline{\mathbf{p}} \ge \mathbf{o}$ is a local solution of (4.12), then $\overline{\mathbf{x}} = (1/\mathbf{e}'\overline{\mathbf{p}})\overline{\mathbf{p}}$ is a local solution of of (1.1).

(c) The objective values in cases (a) and (b) are related by

$$\frac{1}{g(\overline{\mathbf{x}})} = -2h(\overline{\mathbf{p}})\,.$$

Hence a global solution of (4.12) corresponds to one of (1.1), and vice versa.

Proof. (a) For arbitrary $\mathbf{p} \ge \mathbf{o}$ with $\mathbf{p} \ne \mathbf{o}$ let $\Phi(\mathbf{p}) = (1/\mathbf{e}'\mathbf{p})\mathbf{p}$. Then Φ maps the domain indicated continuously on Δ, and hence $V = \Phi^{-1}(U)$ is a neighbourhood of $\overline{\mathbf{p}} = (1/g(\overline{\mathbf{x}}))\overline{\mathbf{x}}$ if U is a neighbourhood of $\overline{\mathbf{x}}$ in Δ satisfying $g(\mathbf{x}) \ge g(\overline{\mathbf{x}})$ for all $\mathbf{x} \in U$. We now claim that $h(\mathbf{p}) \ge h(\overline{\mathbf{p}})$ for all $\mathbf{p} \in V$. Indeed, using $\mathbf{p}'C\mathbf{p} > 0$ and some algebra, it is easy to derive from $[\mathbf{p}'C\mathbf{p} - \mathbf{e}'\mathbf{p}]^2 \ge 0$ the inequality

$$h(\mathbf{p}) \ge -\frac{(\mathbf{e}'\mathbf{p})^2}{2\mathbf{p}'C\mathbf{p}} \qquad \text{for all } \mathbf{p} \ge \mathbf{o} \text{ with } \mathbf{p} \ne \mathbf{o}\,. \tag{4.13}$$

Now $g(\overline{\mathbf{x}}) \le g(\Phi(\mathbf{p})) = (1/(\mathbf{e}'\mathbf{p})^2)g(\mathbf{p}) = (-\frac{1}{2})[-((\mathbf{e}'\mathbf{p})^2/2\mathbf{p}'C\mathbf{p}])^{-1}$ so that by (4.13)

$$-\frac{1}{2g(\overline{\mathbf{x}})} \le h(\mathbf{p})$$

for all $\mathbf{p} \in V$. What remains to show is the assertion in (c). But this is immediate from $\mathbf{e}'\overline{\mathbf{x}} = 1$.

(b) If $\overline{\mathbf{p}}$ is a local solution of (4.12), then necessarily also the Karush-Kuhn-Tucker conditions are satisfied due to linearity of the constraints. Hence, there is a vector $\mathbf{r} \ge \mathbf{o}$ such that $\mathbf{r}'\overline{\mathbf{p}} = 0$ and $\nabla h(\overline{\mathbf{p}}) = C\overline{\mathbf{p}} - \mathbf{e} = \mathbf{r}$, which in turn entails $g(\overline{\mathbf{p}}) = \overline{\mathbf{p}}'C\overline{\mathbf{p}} = \mathbf{e}'\overline{\mathbf{p}}$ and $h(\overline{\mathbf{p}}) = -\dfrac{1}{2}\mathbf{e}'\overline{\mathbf{p}}$. On the other hand, by definition of $\overline{\mathbf{x}}$ we now get

$$g(\overline{\mathbf{x}}) = \frac{1}{(\mathbf{e}'\overline{\mathbf{p}})^2} g(\overline{\mathbf{p}}) = \frac{1}{\mathbf{e}'\overline{\mathbf{p}}}\,,$$

and thus again the relation in (c) is established. Next define $\Psi(\mathbf{x}) = \frac{1}{g(\mathbf{x})}\mathbf{x}$ which maps Δ continuously into the nonnegative orthant. Hence, $\Psi(\overline{\mathbf{x}}) = (\mathbf{e}'\overline{\mathbf{p}})\overline{\mathbf{x}} = \overline{\mathbf{p}}$ implies that $U = \Psi^{-1}(V)$ is a neighbourhood of $\overline{\mathbf{x}}$ in Δ provided that V is one of $\overline{\mathbf{p}}$ in the positive orthant such that $h(\mathbf{p}) \ge h(\overline{\mathbf{p}})$ for all $\mathbf{p} \in V$. Consequently,

$$-\frac{1}{2g(\overline{\mathbf{x}})} = h(\overline{\mathbf{p}}) \le h(\Psi(\mathbf{x})) = \frac{1}{2[g(\mathbf{x})]^2} g(\mathbf{x}) - \frac{1}{g(\mathbf{x})}\mathbf{e}'\mathbf{x} = -\frac{1}{2g(\mathbf{x})}$$

for all $\mathbf{x} \in U$, which shows (b). The remaining assertion in (c) is immediate. □

5. The search for a maximum weight clique

As an important application for standard QPs let us address the search for a maximum weight clique. For convenient reference, let us start with the unweighted case. Consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n nodes. A *clique* S is a subset of the *node set* $\mathcal{V}$ which corresponds to a complete subgraph of $\mathcal{G}$ (i.e., any pair of nodes in S is an edge in $\mathcal{E}$, the *edge set*). A clique S is said to be *maximal* if there is no larger clique containing S. A (maximal) clique is said to be a *maximum* clique if it contains most elements among all cliques. For a concise pre-1995 survey on the maximum clique problem see [47].

Motzkin and Straus showed in [41] that $(1 - f^*)^{-1}$ is the size of a maximum clique if f^* denotes the optimal objective value of the standard QP (1.1) with $A = A_{\mathcal{G}}$, the adjacency matrix of the graph $\mathcal{G}$, i.e. of

$$f(\mathbf{x}) = \mathbf{x}'A_{\mathcal{G}}\mathbf{x} \to \max! \qquad \text{subject to } \mathbf{x} \in \Delta\,. \tag{5.14}$$

This approach has served as the basis of many clique-finding algorithms and also has been used to determine theoretical bounds on the maximum clique size [16, 19, 28, 46, 49]. However, since the (local) solutions of (5.14) lack strictness, it is difficult to identify a maximum clique from the optimal objective value f^* of (5.14). In a different approach followed in [13, 14, 16, 18] with promising empirical results, the regularized version $\widehat{A_{\mathcal{G}}} = A_{\mathcal{G}} + \frac{1}{2} I_n$ is used instead of $A_{\mathcal{G}}$ in (5.14), where I_n denotes the $n \times n$ identity matrix: $\widehat{A_{\mathcal{G}}} = [a_{ij}]_{i,j \in \mathcal{V}}$ with

$$a_{ij} = \begin{cases} \frac{1}{2} & \text{if } i = j; \\ 1 & \text{if } (i, j) \in \mathcal{E}; \\ 0 & \text{otherwise.} \end{cases} \tag{5.15}$$

For problem (1.1) with $A = \widehat{A_{\mathcal{G}}}$ from (5.15), $\mathbf{x}$ is a local solution to (1.1) if and only if $\mathbf{x} = 1/k \sum_{i \in S} \mathbf{e}_i$, where S is a maximal clique of size k. Then the objective is $\mathbf{x}'\widehat{A_{\mathcal{G}}}\mathbf{x} = 1 - (1/2k)$. Hence, S is a maximum clique of $\mathcal{G}$ if and only if $\mathbf{x}$ is the global solution to (1.1).

Now let us proceed to the weighted case. Given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a weight vector $\mathbf{w} = [w_1, \ldots, w_n]'$ of positive weights w_i associated to the nodes $i \in \mathcal{V}$, every clique S in $\mathcal{G}$ has a weight $W(S) = \sum_{i \in S} w_i$. The maximum weight clique problem (MWCP) consists of finding a clique in the graph which has largest total weight (note that the maximum weight clique does not necessarily have largest cardinality). It is clear that the classical unweighted version turns out to be a special case when the weights assigned to the nodes are all equal, and for this reason the MWCP has at least the same computational complexity as its unweighted counterpart (but see [2, 4] for classes of graphs for which MWCP is solvable in polynomial time).

The MWCP has important applications in fields such as computer vision, pattern recognition and robotics, where graphs are employed as a convenient means of representing high-level pictorial information [5]. Graphs arising in these applications often contain numerical attributes on their nodes and/or edges. In these cases, the matching problem becomes one of finding a maximum weight clique in the so-called association graph (see, e.g. [31, 54]). Among the many approaches to attack the MWCP are [1, 3, 37, 42, 43, 45], to mention just a few (see also [17, 35] for a more detailed account).

In a recent paper, Ramana and his coworkers [29] exploited an idea of Lovász in considering the following class of symmetric $n \times n$ matrices: let

$$\mathcal{M}(\mathcal{G}) = \left\{ (b_{ij})_{i,j \in \mathcal{V}} : b_{ij} \geq \frac{b_{ii} + b_{jj}}{2} \text{ if } (i, j) \notin \mathcal{E}\,, b_{ij} = 0, \text{ otherwise} \right\},$$

and put

$$\mathcal{M}(\mathbf{w}, \mathcal{G}) = \left\{ B = (b_{ij})_{i,j \in \mathcal{V}} \in \mathcal{M}(\mathcal{G}) : b_{ii} = \frac{1}{w_i} \text{ and } b_{ij} = b_{ji} \text{ for all } i, j \right\}.$$

They use the standard QP (1.1) with $A = \gamma \mathbf{e}\mathbf{e}' - B$ for any $B \in \mathcal{M}(\mathbf{w}, \mathcal{G})$ (again choose γ so that A has no negative entries) in that they show that the optimal objective value of the program

$$\mathbf{x}'B\mathbf{x} \to \min! \qquad \text{subject to } \mathbf{x} \in \Delta \tag{5.16}$$

equals $1/W(S^*)$ for any $B \in \mathcal{M}(\mathbf{w}, \mathcal{G})$, where S^* is a clique with maximum weight. Still the difficulty remains to identify S^* from a solution $\mathbf{x}^*$ of (5.16).

Moreover, the problem in this formulation (5.16), with a matrix B drawn from the Motzkin–Straus class $\mathcal{M}(\mathbf{w}, \mathcal{G})$, has some *structural* drawbacks. One of them is that maximal cliques S having less than maximum weight are not necessarily reflected by local solutions to (5.16) for every $B \in \mathcal{M}(\mathbf{w}, \mathcal{G})$. As an elementary example consider the graph consisting of just two disconnected nodes with different weights, where the objective in (5.16) can be a non-constant affine function although both singletons are maximal cliques. Hence, a major drawback is the lack of one-to-one correspondence between local solutions to (5.16) and maximal cliques. This is also reflected by the occurrence of jamming during local optimization procedures based on the replicator dynamics mentioned at the end of Section 3.

Therefore, a regularized variant of the formulation is proposed here, for which it is possible to sharpen the above results. Another advantage of the regularization already indicated is that spurious solutions (i.e. those from which one cannot extract the cliques directly), which can occur with the Motzkin–Straus class and which are discussed in detail in [17], are automatically prohibited with the proposed alternative approach.

Now, instead of the Motzkin–Straus class $\mathcal{M}(, \mathcal{G})w$ here a different class $\mathcal{C}(\mathbf{w}, \mathcal{G})$ of matrices is considered to be used as input data for problem (5.16): let

$$\mathcal{C}(\mathcal{G}) = \left\{(c_{ij})_{i,j\in\mathcal{V}} : c_{ij} \geq c_{ii} + c_{jj} \text{ if } (i, j) \notin \mathcal{E}\,,\, c_{ij} = 0, \text{ otherwise } \right\},$$

and consider

$$\mathcal{C}(\mathbf{w}, \mathcal{G}) = \left\{C = (c_{ij})_{i,j\in\mathcal{V}} \in \mathcal{C}(\mathcal{G}) : c_{ii} = \frac{1}{2w_i} \text{ and } c_{ij} = c_{ji} \text{ for all } i, j\right\}.$$

For the unweighted case where $\mathbf{w} = \mathbf{e}$, this regularization goes back to a different characterization of maximal cliques by Comtet [21], and has been thoroughly studied in [13]. Numerical experiences with this approach on a larger scale are reported in [16, 18]. The class $\mathcal{C}(\mathbf{w}, \mathcal{G})$ represents a natural extension to the weighted case.

By contrast to the Motzkin–Straus class $\mathcal{M}(\mathbf{w}, \mathcal{G})$, we can prove for the Comtet class $\mathcal{C}(\mathbf{w}, \mathcal{G})$ again the one-to-one correspondence of strict local solutions of (5.16) and cliques S, if $B \in \mathcal{M}(\mathbf{w}, \mathcal{G})$ is replaced with a matrix $C \in \mathcal{C}(\mathbf{w}, \mathcal{G})$. This is done as in the unweighted case, by just adapting the arguments in [29] to this more appropriate setting. Furthermore, it will be shown below that in contrast to the Motzkin–Straus class, the occurrence of spurious solutions can be ruled out: *every* local solution to (5.16) with $C \in \mathcal{C}(\mathbf{w}, \mathcal{G})$ is strict and necessarily coincides with the *weighted barycenter* $\mathbf{x}^S \in \Delta_S^\circ$ of Δ_S – recall definitions (3.7) and (3.8) – which is a vector with coordinates

$$x_i^S = \begin{cases} w_i/W(S) & \text{if } i \in S\,, \\ 0 & \text{otherwise.} \end{cases}$$

For the proof of this result we need the following inequality which in its essence is used already by [29], but for $\mathcal{M}(\mathbf{w}, \mathcal{G})$ rather than for $\mathcal{C}(\mathbf{w}, \mathcal{G})$. Here, a simple argument is provided for both cases:

PROPOSITION 6 *If S is a clique and $\mathbf{y} \in \Delta_S$ then*

$$2\,\mathbf{y}'C\mathbf{y} = \mathbf{y}'B\mathbf{y} \geq 1/W(S)$$

for all $B \in \mathcal{M}(\mathbf{w}, \mathcal{G})$ and all $C \in \mathcal{C}(\mathbf{w}, \mathcal{G})$. Equality is obtained if and only if $\mathbf{y} = \mathbf{x}^S$.

Proof. Denote by $\langle ., . \rangle_{\mathbf{w},S}$ the Shahshahani inner product (cf. [30]):

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{w},S} = \sum_{i \in S} \frac{x_i\, y_i}{w_i}\,, \qquad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n\,,$$

and by $\|\mathbf{x}\|_{\mathbf{w},S} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{w},S}}$ the corresponding pseudo-norm. Then $\langle \mathbf{x}^S, \mathbf{y} \rangle_{\mathbf{w},S} = 1/W(S)$ for any $\mathbf{y} \in \Delta_S$ and hence also $\|\mathbf{x}^S\|^2_{\mathbf{w},S} = 1/W(S)$. By the inequality of Cauchy, Schwarz and Bunyakovsky,

$$\sum_{i \in S} \frac{y_i^2}{w_i} = \|\mathbf{y}\|^2_{\mathbf{w},S} \;\geq\; \langle \mathbf{x}^S, \mathbf{y} \rangle^2_{\mathbf{w},S}\, \|\mathbf{x}^S\|^{-2}_{\mathbf{w},S} = 1/W(S)$$

with equality if and only if $\mathbf{y} = \mathbf{x}^S$ (recall that $\mathbf{y} \in \Delta_S$ always). On the other hand, if S is a clique and $\mathbf{y} \in \Delta_S$, then

$$2\,\mathbf{y}'C\mathbf{y} = \mathbf{y}'B\mathbf{y} = \sum_{i \in S} \frac{y_i^2}{w_i}$$

for any $B \in \mathcal{M}(\mathbf{w}, \mathcal{G})$ and any $C \in \mathcal{C}(\mathbf{w}, \mathcal{G})$, which shows the assertion. □

THEOREM 7 *Let $\mathcal{G}$ be an arbitrary graph with positive weight vector $\mathbf{w} \in \mathbb{R}^n$, and consider a matrix $C \in \mathcal{C}(\mathbf{w}, \mathcal{G})$ in place of B for problem (5.16). Then the following assertions hold:*

- *(a) A vector $\mathbf{x} \in \Delta$ is a local solution to problem (5.16) if and only if $\mathbf{x} = \mathbf{x}^S$, where S is a maximal clique.*
- *(b) A vector $\mathbf{x} \in \Delta$ is a global solution to problem (5.16) if and only if $\mathbf{x} = \mathbf{x}^S$, where S is a maximum weight clique.*

Moreover, all local (and hence global) solutions to (5.16) are strict.

Proof. Suppose that $\mathbf{x} = \mathbf{x}^S$ where S is a maximal clique. Then for all $r \in \mathcal{V} \setminus S$ we have $(i, r) \notin \mathcal{E}$ for at least one $i \in S$ and hence $(C\mathbf{x})_r \geq (1 + w_i/w_r)/[2W(S)] > 1/[2W(S)]$, entailing for $A = \mathbf{ee}' - C$

$$(A\mathbf{x})_r = 1 - (C\mathbf{x})_r < 1 - 1/[2W(S)] = 1 - (C\mathbf{x})_j = (A\mathbf{x})_j = \mathbf{x}'A\mathbf{x}$$

for all $j \in S$ (the last equality follows by $\mathbf{x} \in \Delta_S$). Now suppose that $\mathbf{y} \in \Delta$ satisfies $\mathbf{y}'A\mathbf{x} = \mathbf{x}'A\mathbf{x}$. Then the strict inequalities above entail $y_r = 0$ for all $r \notin S$, i.e. $\mathbf{y} \in \Delta_S$. Hence, by Proposition 6 and the above arguments,

$$\begin{aligned} \mathbf{x}'A\mathbf{y} &= \mathbf{y}'A\mathbf{x} = \textstyle\sum_{j\in S} y_j(\mathbf{x}'A\mathbf{x}) = \mathbf{x}'A\mathbf{x} = 1 - \frac{1}{2W(S)} \\ &\geq 1 - \mathbf{y}'C\mathbf{y} = \mathbf{y}'(\mathbf{ee}' - C)\mathbf{y} = \mathbf{y}'A\mathbf{y} \end{aligned} \tag{5.17}$$

with equality only if $\mathbf{y} = \mathbf{x}^S = \mathbf{x}$, so that $\mathbf{x}$ satisfies the ESS property for A. Theorem 2 now yields strict local optimality of $\mathbf{x}$. Next suppose that $\mathbf{x}$ is a local solution to (5.16). Let $S = \{i \in \mathcal{V} : x_i > 0\}$ be the support of $\mathbf{x}$ and assume that S is no clique (hence $\mathbf{x}$ is no vertex), which means that there are $i, j \in S$ such that $i \neq j$ and $(i, j) \notin \mathcal{E}$, which by definition of $\mathcal{C}(\mathbf{w}, \mathcal{G})$ entails the strict inequality

$$c_{ij} + c_{ji} \geq 2[c_{ii} + c_{jj}] > c_{ii} + c_{jj}\,.$$

Now designate by $\mathbf{v}$ the vector with coordinates

$$v_k = \begin{cases} 1 & \text{if } k = i \\ -1 & \text{if } k = j \\ 0 & \text{otherwise} \end{cases}$$

put $\mathbf{x}(\alpha) = \mathbf{x} + \alpha\mathbf{v}$ for $\alpha \in \mathbb{R}$, and denote by s the straight line $s = \{\mathbf{x}(\alpha) : \alpha \in \mathbb{R}\}$. Then the objective function $\mathbf{x}'C\mathbf{x}$ is strictly concave along s, since

$$[\mathbf{x}(\alpha)]'C[\mathbf{x}(\alpha)] = \mathbf{x}'C\mathbf{x} + 2\alpha\,\mathbf{x}'C\mathbf{v} + \alpha^2[c_{ii} + c_{jj} - c_{ij} - c_{ji}]\,.$$

On the other hand, the straight line s passes the relative interior Δ_S° in $\mathbf{x} = \mathbf{x}(\alpha = 0)$ and therefore its intersection with Δ is a segment containing $\mathbf{x}$ in its relative interior, which is absurd in view of local optimality of $\mathbf{x}$ for (5.16) and strict concavity. Hence S has to be a clique. Now join $\mathbf{x}$ with $\mathbf{x}^S$ by a straight line and repeat the preceding arguments. Then Proposition 6 yields $\mathbf{x} = \mathbf{x}^S$. Finally, S has to be a maximal clique. Indeed, suppose that T is a larger maximal clique containing S. By positivity of weights, we have $W(T) > W(S)$. Proposition 6 shows

$$[\mathbf{x}^S]'C[\mathbf{x}^S] = 1/[2W(S)] > 1/[2W(T)] = [\mathbf{x}^T]'C[\mathbf{x}^T]\,.$$

Now join $\mathbf{x}^S$ and $\mathbf{x}^T$ with a straight line, which intersects Δ_T also in points of the form $(1-\alpha)\mathbf{x}^T + \alpha\mathbf{x}^S$ with $\alpha < 0$ close to zero. Quadraticity of the objective along this line; local optimality of $\mathbf{x} = \mathbf{x}^S$; and optimality of $\mathbf{x}^T$ on the entire face Δ_T

yields a contradiction. This establishes property (a), and the remaining assertions follow as simple consequences. $\square$

The Comtet class $\mathcal{C}(\mathbf{w}, \mathcal{G})$ is isomorphic to the positive orthant in $\binom{n}{2} - e$ dimensions where e is the cardinality of $\mathcal{E}$. This class is a polyhedral pointed cone with its apex given by the matrix $C(\mathbf{w})$ with entries

$$c_{ij}(\mathbf{w}) = \begin{cases} \frac{1}{2w_i} & \text{if } i = j\,, \\ \frac{1}{2w_i} + \frac{1}{2w_j} & \text{if } i \neq j \text{ and } (i, j) \notin \mathcal{E}\,, \\ 0 & \text{otherwise.} \end{cases} \tag{5.18}$$

Observe that in the unweighted case, $C(\mathbf{e}) = \mathbf{e}\mathbf{e}' - \widehat{A_{\mathcal{G}}} = \widehat{A_{\overline{\mathcal{G}}}}$, the Comtet-regularized adjacency matrix of the complement graph $\overline{\mathcal{G}}$. This reflects the elementary property that an independent set of $\mathcal{G}$ is a clique of $\overline{\mathcal{G}}$. So, while the local maximizers of $\mathbf{x}'\widehat{A_{\mathcal{G}}}\mathbf{x}$ over Δ are exactly the barycenters $\mathbf{x}^S$ of maximal cliques S of $\mathcal{G}$, the local minimizers of $\mathbf{x}'\widehat{A_{\mathcal{G}}}\mathbf{x}$ over Δ are exactly the barycenters $\mathbf{x}^T$ of maximal independent sets T of $\mathcal{G}$. Note that within the Motzkin–Straus class $\mathcal{M}(\mathbf{e}, \mathcal{G})$, there is no matrix with this straightforward interpretation.

It can be shown [18] that all principal minors of $C(\mathbf{e})$ are not vanishing, so that jamming under the replicator dynamics is prevented due to Theorem 3. Hence, for $A = \widehat{A_{\mathcal{G}}}$ as specified in (5.15), the algorithm sketched in Section 3 is almost surely finite. Example 2.2 in [32] shows that one cannot hope for the same nice result in the program (5.14). Of course, one could try to transfer this to the weighted case, and this approach could be backed by the observation that for small $n \leq 3$, the matrices $C(\mathbf{w})$ are all non-singular for all choices of positive weight vectors $\mathbf{w}$. Also, if $\overline{\mathcal{G}}$ is complete, one can show that $C(\mathbf{w})$ is non-singular, regardless of the size of n. Admittedly, all these cases are of restricted interest in real-world applications.

However, for general graphs $\mathcal{G}$ and (irrational) weights this property is not always satisfied, as the tridiagonal 4×4-matrix $C(\mathbf{w})$ with $\mathbf{w} = [1, 1, 2/(7 + \sqrt{45}), 2/(7 + \sqrt{45})]'$ shows.* On the other hand, an (intelligent) grid search over integer values of $1/2w_i$ up to 1000 for tridiagonal 4×4-matrices did not yield any singular matrix of that kind, which might suggest that any matrix $C(\mathbf{w})$ with rational w_i is non-singular. To the best of my knowledge, I am not aware of any clue to this question. But even if $C(\mathbf{w})$ were singular, this does not alter the desired one-to-one correspondence between local solutions of (5.16) and suitable maximal cliques as in Theorem 7.

6. Portfolio selection as a standard QP

The familiar mean/variance portfolio selection problem (see, e.g. [39, 40]) can be formalized as follows: suppose there are n securities to invest in, at an amount

* I owe this example to A. Neumaier.

expressed in relative shares $x_i \geq 0$ of an investor's budget. Thus, the budget constraint reads $\mathbf{e}'\mathbf{x} = 1$, and the set of all feasible portfolio (investment plans) is given by Δ. Now, given the expected return m_i of security i during the forthcoming period, and an $n \times n$ covariance matrix V across all securities, the investor faces the multiobjective problem to maximize expected return $\mathbf{m}'\mathbf{x}$ and simultaneously minimize the risk $\mathbf{x}'V\mathbf{x}$ associated by her decision $\mathbf{x}$.

One of the most popular approaches to such type of problems in general applications is that the user prespecifies a parameter β which in her eyes balances the benefit of high return and low risk, i.e. consider the parametric QP

$$f_\beta(\mathbf{x}) = \mathbf{m}'\mathbf{x} - \beta\mathbf{x}'V\mathbf{x} \to \max! \qquad \text{subject to } \mathbf{x} \in \Delta\,. \tag{6.19}$$

Note that for fixed β, this is again a standard QP (cf. Section 1). Moreover in theory the matrix V is, as an exact covariance matrix, positive semidefinite (although it could be singular in many applications, see [40]), so that (6.19) is a convex problem. Hence in the algorithm in Section 3 the escape step would be superfluous, so there were no repetitions (note that the transformation proposed in Section 1 could destroy positive semidefiniteness on the whole of $\mathbb{R}^n$, but the objective function would still be concave over Δ). On the other hand, securities usually are highly correlated, and in time-series analysis one frequently encounters the situation that some of the most reliable estimators $\tilde{V}$ of the unknown covariance matrix V lack semidefiniteness properties [51, pp. 134ff]. More recently, also the econometric community has been aware of this phenomenon [44], which of course is a nuisance but nevertheless can be handled by the proposed procedure. Anyhow, the question remains how to choose β. In finance applications, the notion of market portfolio is used to determine a reasonable value for this parameter. This emerges more or less from an exogenous artefact, namely by introducing a completely risk-free asset which is used to scale return versus risk. For details see, e.g., Chapter 3 of [33].

Following a recent result of Best and Ding [8] who consider the problem

$$\max_{\beta>0}\max_{\mathbf{x}\in\Delta}\frac{1}{\beta}f_\beta(\mathbf{x})\,, \tag{6.20}$$

also a purely endogenous derivation of market portfolio seems to be possible: they show how optimal solutions $(\beta^*, \mathbf{x}^*)$ for (6.20) emerge from a single standard QP (1.1) with, e.g. $A = 2\mathbf{m}\mathbf{m}' - V$. Hopefully some empirical results with this approach on the Austrian stock market can be reported soon.

Of course, there are many approaches to solve the parametric QP (6.19), e.g. in [8, 40]. Given we know the solutions as specified in [8]

$$\mathbf{x}^*(\beta) = \mathbf{g}_i + \beta\mathbf{h}_i\,, \qquad \text{if } \beta_{i-1} < \beta < \beta_i\,,\ i \in \{1, \dots, t\}$$

with $0 = \beta_0 < \beta_1 < \dots < \beta_t = \infty$ and known vectors $\mathbf{g}_i$ and $\mathbf{h}_i$, it is natural to ask which of the investor's utility structure would yield the same outcome [20, 27]. In the present context, this question (referring to a nonparametric function class giving the same solution) can be rephrased as a typical application of hypersensitivity analysis in connection with composite quadratic programming, which will be addressed in a forthcoming paper [22].

References

1. Alizadeh, F. (1991), A sublinear-time randomized parallel algorithm for the maximum clique problem in perfect graphs, *ACM-SIAM Symposium on Discrete Algorithms* 2: 188–194.
2. Balas, E., Chvátal, V. and Nesetril, J. (1987), On the maximum weight clique problem, *Math. Oper. Res.* 12: 522–535.
3. Balas, E. and Xue, J. (1991), Minimum weighted coloring of triangulated graphs, with application to maximum weight vertex packing and clique finding in arbitrary graphs, *SIAM J. Comput.* 20: 209–221.
4. Balas, E. and Yu, C.S. (1989), On graphs with polynomially solvable maximum-weight clique problem, *Networks* 19: 247–253.
5. Ballard, D.H. and Brown, C.M. (1982), *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ.
6. Baum, L.E. and Eagon, J.A. (1967), An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology, *Bull. Amer. Math. Soc.* 73: 360–363.
7. Baum, L.E. and Sell, G.R. (1968), Growth transformations for functions on manifolds, *Pacif. J. Math.* 27: 211–227.
8. Best, M.J. and Ding, B. (1997), Global and local quadratic minimization, *J. Global Optimiz.* 10: 77–90.
9. Bomze, I.M. (1983), Lotka-Volterra equation and replicator dynamics: a two-dimensional classification, *Biol. Cybern.* 48: 201–211.
10. Bomze, I.M. (1992), Copositivity conditions for global optimality in indefinite quadratic programming problems, *Czechoslovak J. Oper. Res.* 1: 7–19.
11. Bomze, I.M. (1992), Detecting all evolutionarily stable strategies, *J. Optimization Theory Appl.* 75: 313–329.
12. Bomze, I.M. (1996), Block pivoting and shortcut strategies for detecting copositivity, *Linear Algebra and its Appl.* 248: 161–184.
13. Bomze, I.M. (1997), Evolution towards the maximum clique, *J. Global Optimiz.* 10: 143–164.
14. Bomze, I.M. (1997), Global escape strategies for maximizing quadratic forms over a simplex, *J. Global Optimiz.* 11: 325–338.
15. Bomze I.M. and Danninger, G. (1994), A finite algorithm for solving general quadratic problems, *J. Global Optimiz.* 4: 1–16.
16. Bomze, I.M., Pelillo, M. and Giacomini, R. (1997), Evolutionary approach to the maximum clique problem: empirical evidence on a larger scale, in: I.M. Bomze, T. Csendes, R. Horst and P.M. Pardalos (eds.), *Developments in Global Optimization*, pp. 95–108. Kluwer Academic Publishers, Dordrecht/Boston/London.
17. Bomze, I.M., Pelillo, M. and Stix, V. (1998), Approximating the maximum weight clique: an evolutionary game theory approach, in preparation.
18. Bomze, I.M. and Stix, V. (1998), Genetical engineering via negative fitness: evolutionary dynamics for global optimization, to appear in *Annals of O.R.*
19. Budinich, M. (1997), Properties of the adjacency matrix and the maximum clique problem, presented at *MAX-CLIQUE'97 – A Workshop on the Maximum Clique Problem and Its Applications*, Trieste, Italy.
20. Chamberlain, G. (1983), A characterization of the distributions that imply mean-variance utility functions, *J. Econ. Theo.* 29: 185–201.
21. Comtet, L. *Advanced Combinatorics*, Reidel, Dordrecht.

22. Churilov, L., Bomze, I.M. and Sniedovich, M. (1997), Hypersensitivity analysis for portfolio selection, in preparation.
23. Cohen, J. and Hickey, T. (1979), Two algorithms for determining volumes of convex polyhedra, *J. ACM* 26: 401–414.
24. Crow, J.F. and Kimura, M. (1970), *An introduction to population genetics theory,* Harper & Row, New York.
25. Danninger, G. (1990), A recursive algorithm to detect (strict) copositivity of a symmetric matrix, in: U. Rieder, P. Gessner, A. Peyerimhoff and F.J. Radermacher (eds.), *Methods of Operations Research*, Vol. 62, pp. 45–52. Hain, Meisenheim.
26. De Angelis, P.L., Pardalos, P.M. and Toraldo, G. (1997), Quadratic programming with box constraints, in: I.M. Bomze, T. Csendes, R. Horst and P.M. Pardalos (eds.), *Developments in Global Optimization,* pp. 73–93, Kluwer Academic Publishers, Dordrecht/Boston/London.
27. Epstein, L. (1985), Decreasing risk aversion and mean-variance analysis, *Econometrica* 53: 945–962.
28. Gibbons, L.E., Hearn, D.W. and Pardalos, P.M. (1996), A continuous based heuristic for the maximum clique problem, in: D. Johnson and M. Trick (eds.), *Cliques, Coloring, and Satisfiability: Second DIMACS Implementation Challenge*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 26, American Mathematical Society.
29. Gibbons, L.E., Hearn, D.W., Pardalos, P.M. and Ramana, M.V. (1997), Continuous characterizations of the maximum clique problem, *Math. Oper. Res.* 22: 754–768.
30. Hofbauer, J. and Sigmund, K. *The Theory of Evolution and Dynamical Systems*, Cambridge University Press, Cambridge, UK.
31. Horaud, R. and Skordas, T. (1989), Stereo correspondence through feature grouping and maximal cliques, *IEEE Trans. Pattern Anal. Machine Intell.* 11: 1168–1180.
32. Horst, R., Pardalos, P.M. and Thoai, N.V., (1995), *Introduction to Global Optimization,* Kluwer Academic Publishers, Dordrecht/Boston/London.
33. Huang, C.-F. and Litzenberger, R.H. (1988), *Foundations for Financial Economics*, 6th print, North-Holland, Amsterdam.
34. Hummel, R.A. and Zucker, S.W. (1983), On the foundations of relaxation labeling processes, *IEEE Trans. Pattern Anal. Machine Intell.* 5: 267–287.
35. Knuth, D.E. (1994), The sandwich theorem, *Electronic J. Combinatorics* A1.
36. Levinson, S.E., Rabiner, L.R. and Sondhi, M.M. (1983), An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition, *Bell Syst. Tech. J.* 62: 1035–1074.
37. Loukakis, E. and Tsouros, C. (1983), An algorithm for the maximally internally stable set in a weighted graph, *Int. J. Computer Math.* 13: 117–129.
38. Lyubich, Yu., Maistrowskii, G.D. and Ol'khovskii, Yu.G. (1980), Selection-induced convergence to equilibrium in a single-locus autosomal population, *Problems of Information Transmission* 16: 66–75.
39. Markowitz, H.M. (1952), Portfolio selection, *J. of Finance* 7: 77–91.
40. Markowitz, H.M. (1995), The general mean-variance portfolio selection problem, in: S.D. Howison, F.P. Kelly and P. Wilmott (eds.), *Mathematical Models in Finance,* pp. 93–99. Chapman & Hall, London.
41. Motzkin, T.S. and Straus, E.G. (1965), Maxima for graphs and a new proof of a theorem of Turán, *Canad. J. Math.* 17: 533–540.
42. Nemhauser, G.L. and Sigismondi, G. (1992), A strong cutting plane/branch-and-bound algorithm for node packing, *J. Opl. Res. Soc.* 43, 443–457.
43. Nemhauser, G.L. and Trotter, L.E. (1975), Vertex packings: Structural properties and algorithms, *Math. Programming* 8: 232–248.
44. Newey, W.K. and West, K.D. (1987), A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55: 703–708.

45. Pardalos, P.M. and Desai, N. (1991), An algorithm for finding a maximum weighted independent set in an arbitrary graph, *Int. J. Computer Math.* 38: 163–175.
46. Pardalos, P.M. and Phillips, A.T. (1990), A global optimization approach for solving the maximum clique problem, *Int. J. Computer Math.* 33: 209–216.
47. Pardalos, P.M. and Xue, J. (1994), The maximum clique problem, *J. Global Optimiz.* 4: 301–328.
48. Pelillo, M. (1994), On the dynamics of relaxation labeling processes, *Proc. IEEE Int. Conf. Neural Networks*, Orlando, FL, pp. 1006–1011.
49. Pelillo, M. (1995), Relaxation labeling networks for the maximum clique problem, *J. Artif. Neural Networks* 2: 313–327.
50. Pelillo, M. and Jagota, A. (1995), Feasible and infeasible maxima in a quadratic program for maximum clique, *J. Artif. Neural Networks* 2: 411–420.
51. Pötscher, B.M. and Prucha, I.R. (1997), *Dynamic Nonlinear Econometric Models – Asymptotic Theory*, Springer, Berlin.
52. Rosenfeld, A., Hummel, R.A. and Zucker, S.W.(1976), Scene labeling by relaxation operations, *IEEE Trans. Syst. Man Cybern.* 6: 420–433.
53. Sigmund, K. (1987), Game dynamics, mixed strategies, and gradient systems, *Theor. Pop. Biol.* 32: 114–126.
54. Venkateswar, V. and Chellappa, R. (1995), Hierarchical stereo and motion correspondence using feature groupings, *Int. J. Computer Vision* 15: 245–269.
55. Ye, Y. (1997), Approximating quadratic programming with quadratic constraints, *Working Paper, Department of Management Science, University of Iowa, Iowa City, IA 52242.*