# Multi-Standard Quadratic Optimization Problems

Immanuel M. Bomze, Werner Schachinger

ISDS, University of Vienna, Austria

---

**Abstract**

A Standard Quadratic Optimization Problem (StQP) consists of maximizing a (possibly indefinite) quadratic form over the standard simplex. Likewise, in a multi-StQP we have to maximize a (possibly indefinite) quadratic form over the cartesian product of several standard simplices (of possibly different dimensions). Two converging monotone interior point methods are established. Further, we prove an exact cone programming reformulation for establishing rigid yet affordable bounds and finding improving directions.

# 1 Introduction

Standard quadratic optimization problems (maximizing a – possibly indefinite – quadratic form over a standard simplex)constitute a key problem class in quadratic optimization. They occur on several occasions as subproblems, and have numerous real-world applications, among them the maximum-clique problem and portfolio optimization [3]. A multi-standard quadratic optimization problem, likewise, consists of maximizing a quadratic form of the product of simplices. As such, this problem class has not yet been investigated in general set-ups, but if the aforementioned product is a power of one simplex, these problems arise under the name of *relaxation labelling processes* which are used in Computer Imaging and Pattern Recognition [9, 11, 13]. Also, if this one simplex is of dimension one, we arrive at another key class in quadratic programming, namely box-constrained QPs with applications, e.g., in the maximum-cut problem and other combinatorial optimization fields, be it as an exact reformulation or as a relaxation problem for rigid bounds.

To be precise, let $\Delta_i \subset \mathbb{R}^{n_i}$ be standard simplices in $\mathbb{R}^{n_i}$ with $n_i \geq 2$ as $i \in \{1, \ldots, m\}$, and put

$$\Lambda = \bigotimes_{i=1}^{m} \Delta_i \,,$$

a polytope in $\mathbb{R}^M$ with $M = \sum_{i=1}^m n_i \geq 2m$. Given an arbitrary symmetric $M \times M$ matrix $Q$, we consider the following problem as a *multi-Standard Quadratic Problem (multi-StQP) in maximization form*:

$$\max \left\{ z^\top Q z : z \in \Lambda \right\} \,. \tag{1}$$

We decompose $Q$ according to the structure of $\Lambda = \bigotimes_{i=1}^{m} \Delta_i$ into $n_i \times n_j$ submatrices $R_{ij}$ with $R_{ji} = R_{ij}^\top$ for all $i, j \in \{1, \ldots, m\}$, such that

$$z^\top Q z = \sum_{i,j=1}^{m} (x^i)^\top R_{ij} x^j \quad \text{for all } z = (x^1, \ldots, x^m) \in \Lambda \,, \text{ where } x^i \in \Delta_i \subset \mathbb{R}^{n_i} \,. \tag{2}$$

Any subset $S \subseteq \{1, \ldots, n_i\}$ generates a face $\Delta_i(S) = \{x \in \Delta_i : x_j = 0 \text{ if } j \notin S\}$ of $\Delta_i$. Likewise, the faces of $\Lambda$ are of the form

$$\Lambda(\bigotimes_{i=1}^{m} S_i) = \bigotimes_{i=1}^{m} \Delta_i(S_i) \tag{3}$$

for a collection of subsets $S_i \subseteq \{1, \ldots, n_i\}$, all $i \in \{1, \ldots, m\}$.

First we establish the first-order optimality conditions for problem (1). As the constraints are linear, we need no constraint qualifications to ensure that every local solution to (1) satisfy these KKT conditions. Now let $\bar{z} = (\bar{x}^i)$ be a local solution. For any $i$, there is a multiplier vector $u^i \in \mathbb{R}^{n_i}_+$ with $(u^i)^\top \bar{x}^i = 0$ and a scalar multiplier $\lambda_i \in \mathbb{R}$ such that

$$2 \sum_{j=1}^{m} R_{ij} \bar{x}^j + u^i + \lambda_i e^i = o\,, \quad i \in \{1, \dots, m\}\,, \tag{4}$$

where $e^i = [1, \dots, 1]^\top \in \mathbb{R}^{n_i}$, or, equivalently, for all $i \in \{1, \dots, m\}$

$$\left( \sum_{j=1}^{m} R_{ij} \bar{x}^j \right)_k \leq \sum_{j=1}^{m} (\bar{x}^i)^\top R_{ij} \bar{x}^j\,, \text{ all } k \in \{1, \dots, n_i\}\,, \text{ with equality if } \bar{x}^i_k > 0\,. \tag{5}$$

Equivalence of (4) with (5) follows by multiplying (4) by $(\bar{x}^i)^\top$, by observing $(e^i)^\top \bar{x}^i = 1$ and the complementary slackness condition $(u^i)^\top \bar{x}^i = 0$.

# 2 Replicator dynamics for multi-StQPs

For the case $m = 1$ we arrive at the classical StQP. An efficient interior-point local optimization procedure for StQPs is given by the so-called replicator dynamics. For multi-StQPs there are at least two options to generalize this procedure. Both have the general form

$$x^i[t+1] = \frac{1}{(x^i[t])^\top F_i[t]}\, x^i[t] * F_i[t]\,, \quad i \in \{1, \dots, m\}\,, \tag{6}$$

where $x * y = \text{Diag}\,(x) y$ denotes the dyadic (Hadamard) product between two vectors of the same length, and $F_i : \mathbb{R} \to \mathbb{R}^{n_i}_+$ are functions. System (6) is autonomous if $F_i[t] = G_i(z[t])$ for some functions $G_i : \Delta_i \to \mathbb{R}^{n_i}_+$.

**Theorem 2.1** *If $(x^i[t])^\top F_i[t] > 0$ is ensured for all $i, t$, then dynamics (6) is well defined on the set $\Lambda$ and this set and all its faces $\Lambda(\bigotimes_{i=1}^m S_i)$ are forward-invariant.*

**Proof.** The first assertion follows from $e^\top (x * y) = e^\top \text{Diag}\,(x) y = x^\top y$, as this implies

$$(e^i)^\top x^i[t+1] = \frac{(e^i)^\top x^i[t] * F_i[t]}{(x^i[t])^\top F_i[t]} = 1\,.$$

Invariance of the faces follows from $(x * y)_i = 0$ whenever $x_i = 0$. □

### 2.1 Simultaneous replicator dynamics

Suppose that $Q$ has no negative entries and a positive diagonal. Note that by replacing $Q$ with $Q + \gamma E$ for large enough $\gamma > 0$, we can always achieve this. Here and in the sequel, $E = ee^\top$ with $e = (e^i) = [1, \dots, 1]^\top \in \mathbb{R}^M$. Note that convexity or concavity properties of the objective function $q$ over $\Lambda$ remain unaffected by adding or subtracting $\gamma E$.

Start with any $z(0) \in \operatorname{relint} \Lambda$ (this means $x^i_k(0) > 0$ for all $k$, all $i$), and define the following dynamics in discrete time $t$, specifying the right-hand side of (6):

$$F_i[t] = G_i(z[t]) \quad \text{with} \quad G_i(z) = \sum_{j=1}^{m} R_{ij} x^j \quad \text{for all } z = (x^1, \dots, x^m) \in \Lambda \,, \tag{7}$$

or, written in a more explicit form

$$x^i[t+1] = \frac{1}{(x^i[t])^\top \sum_{j=1}^m R_{ij} x^j[t]} \, x^i[t] * \left( \sum_{j=1}^m R_{ij} x^j[t] \right) , \quad i \in \{1, \dots, m\} \,. \tag{8}$$

**Theorem 2.2** *If $Q$ is non-negative with positive diagonal, then*

*(a) dynamics (8) is well defined on the set $\Lambda$ and this set and all its faces, as well as the relative interior of these faces, are forward-invariant;*

*(b) a point $\bar{z} = (\bar{x}^i) \in \Lambda$ is a fixed point under these dynamics if and only if for all $i \in \{1, \dots, m\}$, we have*

$$\left( \sum_{j=1}^m R_{ij} \bar{x}^j \right)_k = (\bar{x}^i)^\top \sum_{j=1}^m R_{ij} \bar{x}^j \,, \qquad \text{if } \bar{x}^i_k > 0 \,, \text{ all } k \in \{1, \dots, n_i\} \,. \tag{9}$$

**Proof.** Assertion (a) with forward-invariance of the faces follows from the positivity assumptions on $Q$, employing Theorem 2.1. Similarly we argue for the relative interior of the faces: indeed, from the assumptions we get $[G_i(z)]_k \geq [R_{ii} x^i]_k \geq r^{(ii)}_{kk} x^i_k$ if $z = (x^i) \in \Lambda$ and hence $x^i_k[t+1] \geq r^{(ii)}_{kk} \left(x^i_k[t]\right)^2 > 0$ if $x^i_k[t] > 0$. Here $r^{(ii)}_{kk} > 0$ is a diagonal element of $R_{ii}$ and hence of $Q$. Assertion (b) is a direct consequence of

$$\frac{x^i_k[t+1]}{x^i_k[t]} = \frac{\left( \sum_{j=1}^m R_{ij} x^j[t] \right)_k}{(x^i[t])^\top \sum_{j=1}^m R_{ij} x^j[t]} \,.$$

Indeed, the left-hand side above equals one for any $k$ with $\bar{x}_k^i > 0$, if and only if $z[t] = \bar{z}$ is fixed under (8). □

Observe that the stationarity conditions (9) are implied by the KKT conditions (5), but not vice versa. However, limit points of trajectories starting in relint $\Lambda$ are KKT points:

**Theorem 2.3** *Suppose $Q$ is non-negative with positive diagonal and assume that $\bar{z} = \lim_{t\to\infty} z[t]$ for some trajectory under (8) with $z_k[0] > 0$ for all $k$. Then $\bar{z} = (\bar{x}^i)$ satisfies (5).*

**Proof.** By continuity, it follows $\bar{z}[1] = \lim_{t\to\infty} z[t+1] = \lim_{t\to\infty} z[t] = \bar{z}[0] = \bar{z}$, so that $\bar{z}$ is fixed under (8). Hence it remains to show that

$$\left(\sum_{j=1}^{m} R_{ij}\bar{x}^j\right)_k \leq (\bar{x}^i)^\top \sum_{j=1}^{m} R_{ij}\bar{x}^j \text{ if } \bar{x}_k^i = 0\,. \tag{10}$$

Suppose this is violated for some $k$ and $i$. By continuity, for some large enough $T$ and sufficiently small $\varepsilon > 0$ we then would have

$$\frac{x_k^i[t+1]}{x_k^i[t]} = \frac{\left(\sum_{j=1}^{m} R_{ij}x^j[t]\right)_k}{(x^i[t])^\top \sum_{j=1}^{m} R_{ij}x^j[t]} > 1+\varepsilon \quad \text{for all } t \geq T\,,$$

which would yield $x_k^i[t] \geq x_k^i[T](1+\varepsilon)^{t-T} \to \infty$ as $t \to \infty$, which is absurd. Hence the result. □

To enforce convergence of dynamics (8), we impose the following regularity condition:

$$\textit{There are finitely many fixed points under (8).} \tag{11}$$

**Theorem 2.4** *If $Q$ is non-negative with positive diagonal, then*

*(a) along all non-constant trajectories under (8) the objective is strictly increasing;*

*(b) under regularity condition (11), all trajectories starting in relint $\Lambda$ converge to a KKT point of (1).*

**Proof.** Let $P(z) = \frac{1}{2}z^\top Q z$. Then dynamics (8) reads, due to (2),

$$x_k^i[t+1] = \frac{x_k^i[t]\,(\partial P(z[t])/\partial x^i)_k}{(x^i[t])^\top\,(\partial P(z[t])/\partial x^i)}\,.$$

Since $P$ is a polynomial in $z$ with no negative coefficients, the Baum-Eagon Theorem [2] yields $P(z[t+1]) > P(z[t])$ unless $z[t+1] = z[t]$. Hence (a) is established. To prove (b), we follow [10], see Proposition 1 there: since $\Lambda$ is compact, (a) implies that the $\omega$-limit $A$ of $z[0]$, i.e., the set of all accumulation points of any (8)-trajectory $z[t]$, is compact (this is standard and follows by a straightforward diagonal subsequence argument), connected and consists entirely of fixed points. To establish the latter two properties, we first note that step size vanishes asymptotically:

$$\|z[t+1] - z[t]\| \to 0 \quad \text{as } t \nearrow \infty . \tag{12}$$

This is accomplished by the following observation: take $\bar{z} = \lim_{\nu\to\infty} z[t_\nu] \in A$ where $t_\nu$ is a subsequence of $\mathbb{N}$. Now we get for the immediate successor $\hat{z} = (\hat{x}^1, \ldots, \hat{x}^m)$ of $\bar{z} = (\bar{x}^1, \ldots, \bar{x}^m)$ under (8), i.e., for

$$\hat{x}^i = \frac{1}{(\bar{x}^i)^\top G_i(\bar{z})}\, \bar{x}^i * G_i(\bar{z}) , \quad i \in \{1, \ldots, m\} ,$$

that $z[t_\nu + 1] \to \hat{z}$, by continuity of all functions involved. Thus, on one hand, $P(\hat{z}) \geq P(\bar{z})$ from (a); on the other hand, since $t_\nu + 1 \leq t_{\nu+1}$ for all $\nu \in \mathbb{N}$, we have $P(z[t_\nu + 1]) \leq P(z[t_{\nu+1}])$, and passing to the limit as $\nu \nearrow \infty$, by continuity again, $P(\hat{z}) \leq P(\bar{z})$. Hence $P(\hat{z}) = P(\bar{z})$, and (a) yields $\hat{z} = \bar{z}$, which establishes (12) and also that $\bar{z}$ is a fixed point. Next suppose that $A$ decomposes into two disjoint compact sets $A_1$, $A_2$ with relatively open neighbourhoods $U_i$ which also are disjoint and have a positive distance to each other. For $\bar{z} \in A_1$ and $\tilde{z} \in A_2$, pick subsequences $t_\nu$ and $s_\nu$ such that $z[t_\nu] \to \bar{z}$ and $z[s_\nu] \to \tilde{z}$, as well as $t_\nu < s_\nu < t_{\nu+1}$, all $\nu$. Choose $N$ so large that for all $\nu \geq N$, $z[t_\nu] \in U_1$ and $z[s_\nu] \in U_2$ and also $z[r] \in U_1 \cup U_2$ if $t_\nu \leq r \leq s_\nu$. Now define

$$r_\nu = \max\{r \geq t_\nu : z[t] \in U_1 \text{ for all } t \in \{t_\nu, \ldots, r\}\} < s_\nu , \text{ all } \nu \geq N .$$

Then by construction $z[r_\nu] \in U_1$ but $z[r_\nu + 1] \in U_2$ for all $\nu \geq N$, contradicting (12). Hence $A$ is also connected. By (11), $A$ must however be finite. As any finite connected set is a singleton, any trajectory must converge. Finally we invoke Theorem 2.3, and the result follows. □

**Remark 2.1** *Recently Paul Tseng (personal communication) extended an ascent direction method from the StQP case treated in [3] to a much wider class of nonlinear optimization problems including all multi-StQPs, and showed that accumulation points of the so generated sequence are all KKT-points.*

*Proposition 4 in [3] then ensures convergence if different KKT points have different objective values. Since in any QP there are only finitely many objective values of KKT points by Lemma 4 in [4], this assumption implies that there are but finitely many KKT points.*

*Taking a closer look to the proofs of Proposition 4 in [3], convergence of (8)-trajectories similarly follows also if assumption (11) is replaced by the condition that different fixed points have different objective values. However, an argument similar to Lemma 4 in [4] shows that the latter assumption implies (11). Generally speaking, convergence to a single point seems to be hard to obtain without imposing additional (convexity) assumptions also in other (general purpose) quadratic programming methods like affine-scaling type Newton-KKT interior-point algorithms [14], [1]. However, for single StQPs there is such a result: see [10].*

The results of Theorem 2.2 generalize the results in [11] on relaxation labelling processes in several respects: first, as mentioned above, relaxation labelling processes use a feasible set $\Lambda = \bigotimes_{i=1}^{m} \Delta_i$ with all simplices $\Delta_i$ being equal. Second, there is no result settling convergence to a single point in [11]. Results similar to (a) above for the continuous-time case of relaxation labelling processes can be found in [13].

## 2.2 Sequential replicator dynamics

We again suppose that $Q$ has no negative entries and a positive diagonal. Now, we allow for step-by-step updates of the blocks $x^i[t]$ as follows. Consider any order of the blocks, e.g., the given natural one $1 < 2 < \cdots < m$. Let us abbreviate for $z = (x^i) \in \Lambda$ and $y = (w^i) \in \Lambda$

$$\begin{aligned} H_1(z,y) &= \sum_{j=1}^{m} R_{1j}x^j + \left[\sum_{j=2}^{m}(x^1)^\top R_{1j}x^j\right] e^1 \text{ (which does not depend on } y\text{)}, \\ H_i(z,y) &= \sum_{j<i} \left(R_{ij}w^j + \left[(x^i)^\top R_{ij}w^j\right] e^i\right) + \sum_{j=i}^{m} R_{ij}x^j\,, \quad \text{for } i \in \{2,\dots,m\}\,. \end{aligned} \tag{13}$$

and define the following dynamics on $\Lambda$

$$x^i[t+1] = \frac{1}{x^i[t]^\top H_i(z[t], z[t+1])}\, x^i[t] * H_i(z[t], z[t+1])\,, \quad i \in \{1,\dots,m\}\,. \tag{14}$$

This is, in effect, dynamics (6) with $F_i[t] = H_i(z[t], z[t+1])$. Observe that $F_i[t]$ involves future states only via $x_j[t+1]$ with $j < i$ by construction.

Now, all properties of (8) also are shared by (14). For ease of notation, we here and in the sequel adopt the usual convention that $\sum_\emptyset = 0$.

**Theorem 2.5** *If $Q$ is non-negative with positive diagonal, then*

*(a) dynamics (14) is well defined on the set $\Lambda$ and this set and all its faces, as well as the relative interior of these faces, are forward-invariant;*

*(b) a point $\bar{z} = (\bar{x}^i) \in \Lambda$ is a fixed point under these dynamics if and only if (9) is satisfied. Hence (8) and (14) have the same sets of fixed points, and assumption (11) applies to (14) as well.*

**Proof.** Assertion (a) follows from the positivity assumptions on $Q$ as in Theorem 2.2. To show assertion (b), observe that $z = y$ implies

$$\begin{aligned} H_1(z,z) &= \textstyle\sum_{j=1}^m R_{1j}x^j + \left[\sum_{j>1} (x^i)^\top R_{ij} x^j\right] e^1 \,, \\ H_i(z,z) &= \textstyle\sum_{j=1}^m R_{ij}x^j + \left[\sum_{j<i} (x^i)^\top R_{ij} x^j\right] e^i \,, \end{aligned} \tag{15}$$

and therefore the denominators of (14) equal

$$\begin{aligned} (\bar{x}^1)^\top H_1(\bar{z},\bar{z}) &= \textstyle\sum_{j=1}^m (\bar{x}^1)^\top R_{1j}\bar{x}^j + \sum_{j>1} (\bar{x}^1)^\top R_{1j}\bar{x}^j \,, \\ (\bar{x}^i)^\top H_i(\bar{z},\bar{z}) &= \textstyle\sum_{j=1}^m (\bar{x}^i)^\top R_{ij}\bar{x}^j + \sum_{j<i} (\bar{x}^i)^\top R_{ij}\bar{x}^j \,. \end{aligned}$$

Now, argued similarly as in the proof of Theorem 2.2, the fixed point property amounts to

$$[H_i(\bar{z},\bar{z})]_k = (\bar{x}^i)^\top H_i(\bar{z},\bar{z}) \,, \qquad \text{if } \bar{x}^i_k > 0 \,, \text{ all } k \in \{1,\ldots,n_i\} \,. \tag{16}$$

Comparing (15) with (16), we see that the fixed point property exactly is (9). □

Contrasting to the previous dynamics, (14) is no longer autonomous, unless we stack $u[t] = (z[t], z[t+1])$ and consider its dynamics on $\Lambda \otimes \Lambda$. However, since the Baum-Eagon theorem cannot be applied directly, the next theorem must be proved in a different way. For that purpose, it is convenient to introduce the symmetric, time-dependent $n_i \times n_i$ matrices

$$\left.\begin{aligned} R_i[t] &= \sum_{j>i} \left[R_{ij}x^j[t](e^i)^\top + e^i(x^j[t])^\top R_{ji}\right] \quad \text{and} \\ S_i[t] &= \sum_{j<i} \left[R_{ij}x^j[t](e^i)^\top + e^i(x^j[t])^\top R_{ji}\right] \end{aligned}\right\} \tag{17}$$

Then define the symmetric, time-dependent $n_i \times n_i$ matrices

$$\left.\begin{aligned} Q_1[t] &= R_{11} + R_1[t]\,, \\ Q_i[t] &= S_i[t+1] + R_{ii} + R_i[t]\,, \quad \text{if } i \in \{2,\dots,m-1\}\ , \\ Q_m[t] &= S_m[t+1] + R_{mm}\,. \end{aligned}\right\} \tag{18}$$

The separation of the $i=1$ block in the definition of (13) may seem strange at first sight, but the vector $\left[\sum_{j=2}^{m}(x^1)^\top R_{1j}x^j\right] e^1$ with all coordinates equal added there has the sole purpose to render $Q_1[t]$ a symmetric matrix.

Now it is straightforward to see that for any $w^i \in \Delta_i$, we have

$$\left.\begin{aligned} Q_i[t]w^i \quad = \quad & \sum_{j<i}\left(R_{ij}x^j[t+1] + (w^i)^\top R_{ij}x^j[t+1]\, e^i\right) + \\ & R_{ii}w^i + \sum_{j>i}\left(R_{ij}x^j[t] + (w^i)^\top R_{ij}x^j[t]\, e^i\right) \end{aligned}\right\} \tag{19}$$

as well as

$$(w^i)^\top Q_i[t]w^i = (w^i)^\top R_{ii}w^i + 2\sum_{j<i}(w^i)^\top R_{ij}x^j[t+1] + 2\sum_{j>i}(w^i)^\top R_{ij}x^j[t]\,. \tag{20}$$

**Theorem 2.6** *If $Q$ is non-negative with positive diagonal, then*

*(a) along all non-constant trajectories under (14), the objective is strictly increasing;*

*(b) under regularity condition (11), all trajectories starting in relint $\Lambda$ converge to a KKT point of (1).*

**Proof.** From (19) it is immediate that $H_i(z[t], z[t+1]) = Q_i[t]x^i[t]$ so that the dynamics (14) can be rewritten in a seemingly decoupled way as

$$x^i[t+1] = \frac{1}{(x^i[t])^\top Q_i[t]x^i[t]}\, x^i[t] * \left(Q_i[t]x^i[t]\right)\ , \quad i \in \{1,\dots,m\}\ . \tag{21}$$

Note that the actual coupling acts through time-dependence of the matrices $Q_i[t]$ which by definition have no negative entries with positive diagonal. However, the monotonicity result taken from Theorem 2.4 in the case $m=1$, i.e., for single StQPs, and one time-step $t \mapsto t+1$, for which $Q_i[t]$ can be considered as a constant symmetric matrix, implies

$$\delta_i = (x^i[t+1])^\top Q_i[t]x^i[t+1] - (x^i[t])^\top Q_i[t]x^i[t] > 0 \quad \text{unless } x^i[t+1] = x^i[t]\,, \tag{22}$$

for all $i \in \{1, \ldots, m\}$, and hence, by summing over these indices,

$$\sum_{i=1}^{m} \delta_i > 0 \quad \text{unless } z[t+1] = z[t]\,. \tag{23}$$

Now we calculate $\delta_i$ replacing $w^i$ with $x^i[t+1]$ and $x^i[t]$ in (20), respectively:

$$\begin{aligned} & \delta_i \\ = \quad & (x^i[t+1])^\top R_{ii} x^i[t+1] + 2\sum_{j<i}(x^i[t+1])^\top R_{ij} x^j[t+1] \\ + \quad & 2\sum_{j>i}(x^i[t+1])^\top R_{ij} x^j[t] \\ - \quad & (x^i[t])^\top R_{ii} x^i[t] - 2\sum_{j<i}(x^i[t])^\top R_{ij} x^j[t+1] \\ - \quad & 2\sum_{j>i}(x^i[t])^\top R_{ij} x^j[t] \end{aligned} \tag{24}$$

Hence, unless $z[t+1] = z[t]$, we arrive at

$$\begin{aligned} 0 \quad < \quad & \sum_{i=1}^{m} \delta_i \\ = \quad & \sum_{i,j}(x^i[t+1])^\top R_{ij} x^j[t+1] + 2\sum_i \sum_{j>i}(x^i[t+1])^\top R_{ij} x^j[t] \\ - \quad & \sum_{i,j}(x^i[t])^\top R_{ij} x^j[t] - 2\sum_i \sum_{j<i}(x^i[t])^\top R_{ij} x^j[t+1] \\ = \quad & z[t+1]^\top Q z[t+1] - z[t]^\top Q z[t] \\ + \quad & 2\sum_{i,j:j>i}(x^i[t+1])^\top R_{ij} x^j[t] - 2\sum_{i,j:i>j}(x^j[t+1])^\top R_{ij}^\top x^i[t]\,. \end{aligned} \tag{25}$$

Now, observing $R_{ij}^\top = R_{ji}$ and interchanging the roles of $i$ and $j$, we see that the two sums in the last row above are identical. Therefore

$$z[t+1]^\top Q z[t+1] - z[t]^\top Q z[t] > 0 \quad \text{unless } z[t+1] = z[t]\,,$$

which establishes the monotonicity result (a). Assertion (b) is shown exactly as in Theorem 2.4. □

We conclude this section by providing a sufficient condition for the regularity condition (11).

Denote by $\mathcal{F} := \{z \in \Lambda : z \text{ satisfies (9) }\}$ the set of fixed points under (8) or (14). Define moreover the $m \times M$ matrix

$$H = \begin{bmatrix} (e^1)^\top & o^\top & \cdots & o^\top \\ o^\top & (e^2)^\top & \cdots & o^\top \\ \vdots & & \ddots & \vdots \\ o^\top & o^\top & \cdots & (e^m)^\top \end{bmatrix}. \tag{26}$$

Let $f = \Lambda(\bigotimes_{i=1}^m S_i)$ denote some face of $\Lambda$ as defined in (3) and let $Q_f$ be the symmetric matrix resulting from $Q$ by dropping all rows and columns which do not belong to $\bigotimes_{i=1}^m S_i$. Similarly, $H_f$ results from dropping the respective columns, so that $Q_f$ and $H_f$ are matrices of dimension $n \times n$ and $m \times n$, respectively, if $\bigotimes_{i=1}^m S_i$ has $n$ elements, where clearly $m \leq n \leq M$.

**Theorem 2.7** *We have*

*(a)* $|\mathcal{F}| < \infty \quad \Longleftrightarrow \quad |\mathcal{F} \cap \operatorname{relint} f| \leq 1$ *for all faces* $f \subseteq \Lambda$.

*(b) If* $\det Q \neq 0$ *and* $\det(HQ^{-1}H^\top) \neq 0$ *then* $|\mathcal{F} \cap \operatorname{relint} \Lambda| \leq 1$.

**Proof.** To prove the ($\Longleftarrow$)-part of (a) we first observe that $\Lambda$ is the finite (disjoint) union of its extremal points and the relative interiors of its one- and higher-dimensional faces. Each of those sets contains at most one fixed point, so the total number of fixed points has to be finite.
Now we turn to the proof of the ($\Longrightarrow$)-part of (a). We assume that for some face $f$ there are $y, z \in \mathcal{F} \cap \operatorname{relint} f$ with $z \neq y$. Without loss of generality we may and do assume that $f = \Lambda$. Now note that $z = (x^1, \ldots, x^m) \in \operatorname{relint} \Lambda$ satisfies (9) if and only if $\sum_j R_{ij} x^j = v_i e^i$ for some $v_i \in \mathbb{R}$, all $i \in \{1, \ldots, m\}$, or, equivalently,

$$Qz = \begin{bmatrix} v_1 e^1 \\ \vdots \\ v_m e^m \end{bmatrix} = H^\top v \quad \text{for some } v \in \mathbb{R}^m . \tag{27}$$

In fact, $v = H(z * Qz)$ by $z \in \Lambda$ and the definition of $H$, but this is of no concern here. Of course, also $Qy = H^\top u$ holds for some $u \in \mathbb{R}^m$, so that for the whole segment given by $w_\alpha = \alpha y + (1 - \alpha) z \in \operatorname{relint} \Lambda$ we get $Q w_\alpha = H^\top [\alpha u + (1 - \alpha) v]$ and hence any convex combination of the fixed points $y, z$ is also a fixed point, i.e., there are infinitely many fixed points.
Assertion (b) is proved by assuming $\det Q \neq 0$ and $\det(HQ^{-1}H^\top) \neq 0$, and that there are $y, z \in \mathcal{F} \cap \operatorname{relint} \Lambda$. From (27) we know that there are $u, v \in \mathbb{R}^m$ such that $Qy = H^\top u$, $Qz = H^\top v$. Therefore $Q(z - y) = H^\top a$, where $a = v - u$. Since $\det Q \neq 0$ we have

$$z - y = Q^{-1} H^\top a, \tag{28}$$

and multiplying that equation by $H$ from the left, we obtain $o = Hz - Hy = HQ^{-1}H^\top a$. Since $\det(HQ^{-1}H^\top) \neq 0$ we deduce that $a = o \in \mathbb{R}^m$, and then (28) implies $z = y$, so there is at most one fixed point in relint $\Lambda$. □

**Corollary 2.1** *If* $\det Q_f \neq 0$ *and* $\det(H_f Q_f^{-1} H_f^\top) \neq 0$ *hold for all faces* $f \subseteq \Lambda$, *then* $|\mathcal{F}| < \infty$.

In particular, the regularity condition (11) is satisfied for all matrices $Q$ from an open dense subset of the set of all symmetric $M \times M$-matrices.

# 3 Conic reformulation of multi-StQPs

Now suppose that a variant of replicator dynamics above or some other local optimization procedure gives a local solution. Then, to assess its quality if the problem is non-convex, we would like to obtain a rigid upper bound which can be calculated by some procedure which is not too costly. Note that by a variant of the method in [6], one can prove that multi-StQPs belong to the class of polynomial-time approximation schemes (PTAS). Indeed, for fixed $r$, the number of rational grid points with largest denominator $r$ contained in the feasible set $\Lambda$ is bounded by a polynomial of degree $r+1$ in $M = \sum_{i=1}^m n_i$.

We first establish a conic reformulation using the convex, non-polyhedral cone of **completely positive** $M \times M$ matrices

$$\mathcal{C} = \left\{ X : X = FF^\top \text{ for some } M \times k \text{ matrix } F \text{ with no negative entries } \right\} .$$

It is well known that with respect to the Frobenius duality $X \bullet Y = \text{trace } (XY)$ of symmetric $M \times M$ matrices, this cone is the dual cone of the cone of **copositive** $M \times M$ matrices

$$\mathcal{C}^* = \left\{ Y = Y^\top : x^\top Y x \geq 0 \text{ for all } x \in \mathbb{R}^M_+ \right\} .$$

Using the matrix $H$ defined in (26), any $z = (x^i) \in \Lambda$ satisfies by definition $Hz = [(e^i)^\top x^i] = \bar{e} = [1, \ldots, 1]^\top \in \mathbb{R}^m$. Hence $z$ gives rise to a rank-one matrix $Z = zz^\top$ which belongs to $\mathcal{C}$ and satisfies $HZH^\top = \bar{e}(\bar{e})^\top$. This observation enables us to generalize Lemma 3 of [5] in the following way, establishing an exact conic reformulation of the multi-StQP:

**Theorem 3.1** *The (possibly non-convex) multi-StQP (1) is equivalent to a linear problem over the completely positive cone:*

$$\max \left\{ z^\top Q z : z \in \Lambda \right\} = \max \left\{ Q \bullet Z : HZH^\top = \bar{e}(\bar{e})^\top \, , \; Z \in \mathcal{C} \right\} , \qquad (29)$$

*and, by strong duality, to the following linear problem over the copositive cone:*

$$\inf \left\{ \sum_{i,j} X_{ij} = (\bar{e})^\top X \bar{e} : H^\top X H - Q \in \mathcal{C}^* \, , \; X = X^\top \textit{ an } m \times m \textit{ matrix} \right\} . \qquad (30)$$

**Proof.** First we show that the set

$$\mathcal{Z} = \left\{ Z \in \mathcal{C} : HZH^\top = \bar{e}(\bar{e})^\top \right\}$$

is the convex hull of all rank-one matrices $zz^\top$ with $z \in \Lambda$ (the preceding observation exactly means $zz^\top \in \mathcal{Z}$ if $z \in \Lambda$). Indeed, consider any $Z = FF^\top \in \mathcal{Z}$, where $F$ is an $n \times k$ matrix without negative entries. Then $(HF)(HF)^\top = HZH^\top = \bar{e}(\bar{e})^\top$ is of rank one, therefore also the $m \times k$ matrix $HF$ is of rank one and without negative entries, i.e., of the form $uv^\top$ for some vectors $u \in \mathbb{R}^m_+ \setminus \{o\}$ and $v \in \mathbb{R}^k_+ \setminus \{o\}$. But $HZH^\top = \bar{e}(\bar{e})^\top$ implies $uu^\top = (v^\top v)^{-1}\bar{e}(\bar{e})^\top$ so that all entries of $u$ must be equal (and positive). So without loss of generality we may and do assume that $u = \bar{e}$ and $v^\top v = 1$. But $HF = \bar{e}v^\top$ by definition (26) means $v_s = (HF)_{is} = \sum_{j\in I_i} F_{js}$ for all $i$, where the sets $I_i = \left\{ \sum\limits_{r<i} n_r + 1, \ldots, \sum\limits_{r\leq i} n_r \right\}$ form a partition of the index set $\{1, \ldots, M\}$ as $i$ ranges over $\{1, \ldots, m\}$. Next put $F = (f^1, \ldots, f^k)$ with $f^s \in \mathbb{R}^M_+$ and remember that $Z$ factorizes into $Z = FF^\top = \sum_{s=1}^k f^s(f^s)^\top$, so we may and do assume that all $f^s \neq o$, whence

$$0 < \sum_{j=1}^M f^s_j = \sum_{i=1}^m \sum_{j\in I_i} F_{js} = v_s m$$

which implies $v_s > 0$ for all $s$. Now put $z^s = \frac{1}{v_s} f^s \in \mathbb{R}^M_+$. From $v_s = (HF)_{is} = \sum_{j\in I_i} F_{js}$ for all $i$ it is immediate that $z^s \in \Lambda$ for all $s$, so we end up with a representation of $Z$ as a convex combination of rank-one matrices of the desired type:

$$Z = \sum_{s=1}^k f^s(f^s)^\top = \sum_{s=1}^k (v_s)^2 z^s(z^s)^\top ,$$

and $\sum_s (v_s)^2 = v^\top v = 1$. Hence $\mathcal{Z} = \operatorname{conv}\left\{zz^\top : z \in \Lambda\right\}$ (which is compact, since $\Lambda$ is compact, hence both maxima in (29) are attained), and from standard convexity arguments and the identity $Q \bullet (zz^\top) = z^\top Qz$ we arrive at (29):

$$\max\left\{z^\top Qz : z \in \Lambda\right\} = \max\left\{Q \bullet Z : Z \in \mathcal{Z}\right\} .$$

To obtain (30), we invoke standard duality theory (see for instance [7]), observing that (29) is feasible and (30) is strictly feasible: indeed, take $X = (\gamma + 1)\bar{e}(\bar{e})^\top$ where $\gamma = \max_{i,j} Q_{ij}$, then $[H^\top XH]_{ij} = \gamma + 1$ for all $i, j$ which implies $\min_{ij}[H^\top XH - Q]_{ij} = 1 > 0$ so that $H^\top XH - Q$ in the interior of $\mathcal{C}^*$. □

This primal-dual pair of conic programs has particular properties which are discussed in detail in [12]: if $m \geq 2$, then

(a) Slater's condition for (29) is always violated;

(b) if $Q$ has a representation $Q = H^\top \bar{Q} H$ for some symmetric $m \times m$ matrix $\bar{Q}$, then the set of minimizers of (30) is nonempty;

(c) if the set of minimizers of (30) is nonempty, it is always unbounded;

(d) however, the minimum of (30) need not be attained in general.

[12] also provides conditions which ensure dual attainability. One sufficient condition is that (1) has a unique solution.

Note that the copositive representation differs from that proposed recently by Burer for a more general setup in [8]: there, the matrix cones contain $(M+1) \times (M+1)$ matrices. To be more precise, this copositive representation would read, for our purposes, $\max \left\{ z^\top Q z : z \in \Lambda \right\} =$

$$\max \left\{ Q \bullet Z : \operatorname{diag}\,(HZH^\top) = \bar{e}\,,\ Hz = \bar{e}\,,\ \begin{bmatrix} 1 & z^\top \\ z & Z \end{bmatrix} \in \mathcal{C}' \right\}\,, \tag{31}$$

with $\mathcal{C}'$ denoting the cone of completely positive $(M+1) \times (M+1)$ matrices. As is well known, in any conic program the number of linear constraints is a decisive parameter for performance of the usual interior-point methods. Now there are only $2m+1$ linear constraints in (31) as opposed to $\binom{m+1}{2}$ in (29). On the other hand, the difference between these numbers, $\frac{m^2-3m}{2} - 1$, can be dominated by $M \geq 2m$ (and will be so for sure if $m \in \{2, \ldots, 8\}$). Note that $M$ is the lower bound for the number of additional linear constraints used in any SDP-based approximation hierarchy of $\mathcal{C}'$ as opposed to the same approximation of $\mathcal{C}$. For instance, to get a matrix $\begin{bmatrix} 1 & z^\top \\ z & Z \end{bmatrix}$ with no negative entries, one has to require $z \in \mathbb{R}^M_+$ in addition to the non-negativity of $Z$. Thus, it is *a priori* not evident which copositive representation is more favorable in practical implementations. This comparison will be an interesting topic for future investigations.

However, for both representation variants, we have at hand all approximation results for $\mathcal{C}$ and $\mathcal{C}^*$, for instance see [6]. Particularly, the dual formulation (30) can yield cheap bounds if we find an $X$ feasible to (30) with $\sum_{i,j} X_{ij}$ close to the infimum. Similarly, a variant of the hybrid strategy proposed in [5] could produce improving feasible points or directions, i.e., escaping from inefficient local solutions or KKT points.

# References


[1] Absil, P.-A.; Tits, A.: Newton-KKT interior-point methods for indefinite quadratic programming. Computational Optimization and Applications **36**, 5–41 (2007).

[2] Baum, L.; Eagon, J.A.: An inequality with applications to statistical prediction for function of Markov processes and to a model of ecology. Bulletin of the American Mathematical Society **73**, 360–363 (1967).

[3] Bomze, I.M.: On standard quadratic optimization problems. Journal of Global Optimization **13**, 369–387 (1998).

[4] Bomze, I.M.; Danninger, G.: A finite algorithm for solving general quadratic problems. Journal of Global Optimization **4**, 1–16 (1994).

[5] Bomze, I.M.; Dür, M.; de Klerk, E.; Quist, A.; Roos, C.; Terlaky, T.: On copositive programming and standard quadratic optimization problems. Journal of Global Optimization **18**, 301–320 (2000).

[6] Bomze, I.M.; de Klerk, E.: Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. Journal of Global Optimization **24**, 163–185 (2002).

[7] Boyd, S.; Vandenberghe, L.: *Convex Optimization.* Cambridge Univ. Press, Cambridge, UK (2004).

[8] Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. Preprint, University of Iowa, available at http://www.optimization-online.org/DB_HTML/2006/10/1501.html, last accessed Dec.06 (2006).

[9] Hummel,R.A.; Zucker, S.W.: On the foundations of relaxation labeling processes. IEEE Trans. Pattern Anal. Machine Intell. **5** 267-287 (1983).

[10] Losert, V.; Akin, E.: Dynamics of games and genes: Discrete versus continuous time. Journal of Mathematical Biology **17**, 241–251 (1983).

[11] Pelillo, M.: The dynamics of nonlinear relaxation labeling processes. Journal of Mathematical Imaging and Vision **7** 309-323 (1997).

[12] Schachinger, W.; Bomze, I.M.: A conic duality Frank–Wolfe type theorem via exact penalization in quadratic optimization. Technical Report **2006-10**, ISDS, University of Vienna (2006).

[13] Torsello, A.; Pelillo, M.: Continuous-time relaxation labeling processes. Pattern Recognition **33**, 1897–1908 (2000).

[14] Tseng, P.: Convergence Properties of Dikin's Affine Scaling Algorithm for Nonconvex Quadratic Minimization. Journal of Global Optimization **30**, 285–300 (2004).