# Graph-based quadratic optimization: A fast evolutionary approach

Samuel Rota Bulò [a,*], Marcello Pelillo [a], Immanuel M. Bomze [b]

[a] DAIS – Università Ca' Foscari Venezia, Italy
[b] ISDS, University of Vienna, Austria

## ABSTRACT

Quadratic optimization lies at the very heart of many structural pattern recognition and computer vision problems, such as graph matching, object recognition, image segmentation, etc., and it is therefore of crucial importance to devise algorithmic solutions that are both efficient and effective. As it turns out, a large class of quadratic optimization problems can be formulated in terms of so-called "standard quadratic programs" (StQPs), which ask for finding the extrema of a quadratic polynomial over the standard simplex. Computationally, the standard approach for attacking this class of problems is to use *replicator dynamics*, a well-known family of algorithms from evolutionary game theory inspired by Darwinian selection processes. Despite their effectiveness in finding good solutions in a variety of applications, however, replicator dynamics suffer from being computationally expensive, as they require a number of operations per step which grows quadratically with the dimensionality of the problem being solved. In order to avoid this drawback, in this paper we propose a new population game dynamics (InImDyn) which is motivated by the analogy with infection and immunization processes within a population of "players." We prove that the evolution of our dynamics is governed by a quadratic Lyapunov function, representing the average population payoff, which strictly increases along non-constant trajectories and that local solutions of StQPs are asymptotically stable (i.e., attractive) points. Each step of InImDyn is shown to have a linear time/space complexity, thereby allowing us to use it as a more efficient alternative to standard approaches for solving StQPs and related optimization problems. Indeed, we demonstrate experimentally that InImDyn is orders of magnitude faster than, and as accurate as, replicator dynamics on various applications ranging from tree matching to image registration, matching and segmentation.

© 2011 Elsevier Inc. All rights reserved.

## 1. Introduction

Optimality is arguably one of the most pervasive and flexible meta-principles used in computer vision and pattern recognition, as it allows one to formulate real-world problems in a pure, abstract setting with solid theoretical as well as philosophical underpinnings, and permits access to the full arsenal of algorithms available in the optimization and operations research literature. On the other hand, although graphs have always been an important tool in computer vision because of their representational power and flexibility, there is now a renewed and growing interest toward explicitly formulating vision problems as graph optimization problems, and researchers are increasingly making use of sophisticated graph-theoretic concepts, results, and algorithms [18,3].

Among the variety of optimization problem families, quadratic optimization plays unquestionably a prominent role in computer vision as it naturally arises whenever abstract entities (e.g., pixels,

edgels, regions, etc.) exhibit mutual pairwise interactions. The maximum clique problem, for example, which finds applications in such problems as shape and object recognition [6,2,16,45,55], stereo correspondence [26], point pattern matching [36], and image sequence analysis [48], has been successfully addressed in terms of quadratic optimization via the Motzkin–Straus theorem, a result that has recently been generalized in various ways [41,57,52] and applied to pairwise clustering problems (see Section 2). Other important applications of quadratic programming can be found in [37,21,54].

As it turns out, a large class of quadratic optimization problems can be formulated in terms of *standard quadratic programs* (StQPs), which ask for finding the extrema of a quadratic polynomial over the standard simplex. Computationally, the standard approach to solving StQPs is to use *replicator dynamics*, a class of evolutionary game-theoretic algorithms inspired by Darwinian selection processes. Indeed, there exists an intimate connection between optimization and game theory, as it can be seen that the solutions of any StQP are in one-to-one correspondence to the equilibria of a particular class of two-player games, known as partnership, or doubly-symmetric games, whereby the players' payoffs are assumed to coincide [62,25]. Interestingly, replicator dynamics also

* Corresponding author.
  E-mail address: srotabul@dsi.unive.it (S. Rota Bulò).

arise independently in different branches of theoretical biology [25] and are closely related to the classical Lotka–Volterra equations from population ecology, while in population genetics they are known as selection equations [17]. Further, replicator dynamics turn out to be a special instance of a general class of dynamical systems introduced by Baum and Eagon [4] in the context of Markov chain theory and represent a special case of the well-known relaxation labeling processes for solving consistent labeling problems [49].

Although replicator dynamics have proven to be an effective technique in a variety of StQP applications [44,8,56,34,41], a typical problem associated with these algorithms is the scaling behavior with the dimensionality of the problem being solved. In particular, for a problem involving $N$ variables, the computational complexity of each replicator dynamics step is $\mathcal{O}(N^2)$, thereby hindering their use in large-scale applications, such as high-resolution image/video segmentation and matching. Previous attempts aimed at improving the computational time of the replicator dynamics can be found in works of Pelillo [42,43,47], where an exponential replicator model (a member of a larger class of "payoff-monotonic" game dynamics) has been employed in order to reduce the number of iterations needed for the algorithm to find a solution. However, despite requiring less iterations, the proposed solution still suffers from a per-step quadratic complexity.

In this paper we study a new population game dynamics, the *infection-immunization dynamics* (InImDyn), which avoids this drawback and leads to a remarkable computational gain over previous approaches. InImDyn is motivated by the analogy with infection and immunization processes within a population of "players." Intuitively, the evolutionary process can be interpreted as follows: as time passes by, an advertisement on the basis of the aggregate behavior of the population tells the agents that a certain strategy is *successful* or is *unsuccessful*. A strategy is successful if it is performing best in terms of payoff in the population, whereas it is unsuccessful if it is the worst performing strategy still alive in the population. Both variants will be taken into account: in contrast to the best-reply approach typically used in evolutionary game theory [25], which selects the strategy with highest average payoff, a successful strategy is chosen only if its *absolute* deviation from the average payoff is largest among all absolute deviations. Otherwise, the largest absolute deviation is provided by an unsuccessful strategy, and we move straight away from it by help of its *co-strategy* (to be defined below). In its most generic formulation, this phase encodes a particular selection function for infective strategies, which basically increases (decreases) the share of agents playing the successful (unsuccessful) strategy, as long as there is no barrier to the invasion. Hence, assuming that agents can gather information only about the announced strategy, they will be inclined to switch to the successful strategy, or abandon the one unsuccessful.

In the paper we prove that the evolution of our dynamics is governed by a quadratic Lyapunov function, representing the average population payoff, which strictly increases along any non-constant trajectory and that local solutions of StQPs are asymptotically stable (i.e., attractive) equilibrium points. We also show that each step of InImDyn has a linear time/space complexity, as opposed to the quadratic per-step complexity of replicator dynamics. We provide experimental evidence that the proposed algorithm is orders of magnitude faster than the standard algorithms on various graph-based computer vision applications, ranging from tree matching to image segmentation, matching and registration, while preserving the quality of the solutions found. Hence our approach can be considered an efficient and theoretically sound alternative to the replicator dynamics, that can be usefully employed in those graph-based computer vision and pattern recognition problems where computational complexity might be an issue, e.g., video

and high-resolution image segmentation, matching of large graphs, clustering of large datasets, etc.

The paper is organized as follows. In Section 2 we provide a short review of various graph-based problems that lead to an StQP formulation, while in Section 3 we summarize the basic concepts and results of evolutionary graph theory and replicator dynamics. Section 4 is devoted to the description of our new class of evolutionary dynamics and Section 5 describes a specific instance which exhibits a linear time/space complexity per step. In Section 6 we report on the experimental results, and we finally draw our conclusions in Section 7. A preliminary version of this work has been presented in [51].

## 2. Quadratic formulation for graph-theoretic problems

Many graph-theoretic problems can be formulated in terms of a *standard quadratic program* (StQP), which is defined as:

$$\text{maximize} \quad \mathbf{x}^\top Q \mathbf{x}$$
$$\text{subject to} \quad \mathbf{x} \in \Delta$$

where $Q \in \mathbb{R}^{n \times n}$ is a symmetric matrix, and $\Delta$ is the *standard simplex* of $\mathbb{R}^n$:

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

A large class of quadratic programming problems (QPs), instances of which arise frequently in computer vision and pattern recognition, can be rewritten in terms of StQPs. In fact, consider a general QP over a bounded polyhedron

$$\begin{aligned} \text{maximize} \quad & \tfrac{1}{2}\mathbf{x}^\top Q \mathbf{x} + \mathbf{c}^\top \mathbf{x} \\ \text{subject to} \quad & \mathbf{x} \in \Delta \end{aligned} \quad (1)$$

where $M = \text{conv}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq \mathbb{R}^n$ is the convex hull of the points $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

It is easy to see that the original QP in (1) can be written as the following StQP:

$$\begin{aligned} \text{maximize} \quad & \mathbf{y}^\top \widehat{Q} \mathbf{y} \\ \text{subject to} \quad & \mathbf{y} \in \Delta \end{aligned}$$

where $\widehat{Q} = \tfrac{1}{2}(V^\top Q V + \mathbf{e}^\top V^\top \mathbf{c} + \mathbf{c}^\top V \mathbf{e})$ and $V = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$.

Thus every QP over a polytope can be expressed as an StQP. Of course, this approach is practical only when the polytope is explicitly expressed in terms of its $k$ vertices (and when $k$ is not too large). This is the case of QPs over the $\ell^1$ ball, where $V = [I| - I]$, $I$ being the $n \times n$ identity matrix and $\Delta \subset \mathbb{R}^{2n}$ [11]. However, even for general QPs of the form:

$$\begin{aligned} \text{maximize} \quad & \tfrac{1}{2}\mathbf{x}^\top Q \mathbf{x} + \mathbf{c}^\top \mathbf{x} \\ \text{subject to} \quad & \mathbf{x} \in \mathbb{R}^n_+ \text{ and } A\mathbf{x} = \mathbf{b} \end{aligned}$$

we can use StQP as a relaxation without using all vertices (see [12] for details).

Now, we provide a short review of a few graph-theoretic problems that can be formulated in terms of a StQP, namely the maximum clique problem, graph/tree matching, and pairwise data clustering.

### 2.1. Maximum clique problem

Let $G = (V, E)$ be an undirected graph, where $V = \{1, \ldots, n\}$ is the set of vertices and $E \subseteq V \times V$ is the set of edges. The *order* of $G$ is the number of its vertices, and its *size* is the number of edges. Two vertices $i, j \in V$ are said to be *adjacent* if $(i, j) \in E$. The *adjacency matrix* of $G$ is the $n \times n$ symmetric matrix $A_G = (a_{ij})$ defined as $a_{ij} = 1$ if $(i, j) \in E$, and $a_{ij} = 0$ otherwise. A subset $C$ of vertices in $G$ is called