

# A GENERAL REGULARIZED CONTINUOUS FORMULATION FOR THE MAXIMUM CLIQUE PROBLEM

JAMES T. HUNGERFORD<sup>1</sup> and FRANCESCO RINALDI<sup>2</sup>

ABSTRACT. In this paper, we develop a general regularization-based continuous optimization framework for the maximum clique problem. In particular, we consider a broad class of regularization terms that can be included in the classic Motzkin-Strauss formulation and we develop conditions that guarantee the equivalence between the continuous regularized problem and the original one in both a global and a local sense. We further analyze, from a computational point of view, two different regularizers that satisfy the general conditions.

## 1. INTRODUCTION.

Let  $G = (\mathcal{V}, \mathcal{E})$  be a simple undirected graph on vertex set  $\mathcal{V} = \{1, 2, \dots, n\}$  and edge set  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . Since  $G$  is simple and undirected  $(j, i) \in \mathcal{E}$  whenever  $(i, j) \in \mathcal{E}$ , and  $(i, i) \notin \mathcal{E}$  for any  $i \in \mathcal{V}$ . A *clique* in  $G$  is a subset  $C \subseteq \mathcal{V}$  such that  $(i, j) \in \mathcal{E}$  for every  $i, j \in C$  with  $i \neq j$ . In this paper, we consider the classical *Maximum Clique Problem* (MCP): find a clique  $C \subseteq \mathcal{V}$  such that  $|C|$  is maximum.

The Maximum Clique Problem has a wide range of applications (see [4, 20] and references therein) in areas such as social network analysis, telecommunication networks, biochemistry, and scheduling. The cardinality of a maximum clique in  $G$  is denoted  $\omega(G)$ . A clique  $C$  is said to be *maximal* if it is not contained in any strictly larger clique; that is, if there does not exist a clique  $D$  such that  $C \subset D$ .  $C$  is said to be *strictly maximal* if there do not exist vertices  $i \in C$  and  $j \notin C$  such that  $C \cup \{j\} \setminus \{i\}$  is a clique.

The MCP is NP-hard [12]. However, due in part to its wide applicability, a large variety of both heuristic and exact approaches have been investigated (see [4] for a thorough overview of formulations and algorithms going up to 1999; a more recent survey of algorithms is given in [20]). A significant number of the solution methods proposed (for example, [3, 5, 9, 14, 15, 18, 19]) are based on solving the following well-known continuous quadratic programming formulation of the MCP, due to Motzkin and Straus [15]:

$$(1) \quad \begin{aligned} & \max \quad \mathbf{x}^T \mathbf{A} \mathbf{x} \\ & \text{subject to} \quad \mathbf{x} \in \Delta, \end{aligned}$$

---

2010 *Mathematics Subject Classification.* Primary: 90C30; Secondary: 90C35, 90C20.

*Key words and phrases.* maximum clique, sparse optimization, support, concave minimization, Motzkin-Straus.

where  $\Delta$  is the  $n$ -dimensional simplex defined by

$$\Delta := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \text{ and } \mathbf{1}^\top \mathbf{x} = 1\}$$

and  $\mathbf{A} = (a_{ij})_{i,j \in \mathcal{V}}$  denotes the adjacency matrix for  $G$  defined by

$$a_{ij} = \begin{cases} 1, & (i, j) \in \mathcal{E} \\ 0, & (i, j) \notin \mathcal{E} \end{cases} \quad \forall i, j \in \mathcal{V}.$$

For any non-empty clique  $C$  we let  $\mathbf{x}(C) \in \Delta$  denote the corresponding characteristic vector (defined by  $x(C)_i = \frac{1}{|C|}$  whenever  $i \in C$  and  $x(C)_i = 0$  otherwise). The equivalence between MCP and (1) is given by the following theorem:

**Theorem 1.1** (Theorem 1 [15]). *The optimal objective value of (1) is*

$$1 - \frac{1}{\omega(G)}$$

*and  $\mathbf{x}(C)$  is a global maximizer of (1) for any maximum clique  $C$ .*

Solution approaches to MCP based on solving (1) include nonlinear programming methods [8] and methods based on discrete time replicator dynamics [3, 4, 5, 18]. Since (1) is NP-hard (by reduction to MCP), the computing time required to obtain a global maximizer can grow exponentially with the size of the graph; hence, finding a global maximizer may be impractical in many settings. On the other hand, iterative optimization methods will typically converge to a point satisfying the first-order optimality (Karush-Kuhn-Tucker) conditions. In general, verifying whether a first-order point of a quadratic program is even *locally* optimal is an NP-hard problem [16, 17]. However, it was shown in [9] that local optimality of a first-order point (in fact, any feasible point) in (1) can be ascertained in polynomial time.

In [19, Proposition 3], a characteristic vector for a clique was shown to satisfy the standard first-order optimality condition for (1) if and only if the associated clique is maximal. In [9, Theorem 2], the authors gave a characterization of the local optima of (1) and demonstrated a one-one correspondence between strict local maximizers and strictly maximal cliques. These results suggest the possibility of applying iterative optimization methods to (1) in order to *approximately* solve MCP (ie. to find large *maximal* cliques). However, one known [3, 18, 19] drawback of this approach in practice is the presence of “infeasible” or “spurious” local maximizers of (1) which are not characteristic vectors for cliques and from which a clique can not be recovered through any simple transformation. Such points are an undesirable property of the program, since they can cause continuous based heuristics to fail by terminating without producing a clique. In [3], the author addresses this issue by introducing the following regularized formulation (with  $\alpha = \frac{1}{2}$ ):

$$(2) \quad \begin{aligned} \max \quad & \mathbf{x}^\top \mathbf{A} \mathbf{x} + \alpha \|\mathbf{x}\|_2^2 \\ \text{subject to} \quad & \mathbf{x} \in \Delta. \end{aligned}$$

In contrast to (1), the local maximizers of (2) have been shown to be in one-one correspondence with the maximal cliques in  $G$  (see [3, Theorem 9]), and a replicator dynamics approach to solving (2) was shown to reduce the total number of algorithm failures by 30%, compared with a similar approach to solving (1). In [5], the authors enhanced the algorithm of [3], adding an annealing heuristic to obtain even stronger results. In addition, it was demonstrated that the correspondence between the local/global optima of (2) and MCP is maintained for any  $\alpha \in (0, 1)$ . A similar formulation and approach [2] has also been applied successfully to a weighted version of MCP.

In practice, the numerical performance of an iterative optimization method (in terms of speed and/or solution quality) may depend on the particular regularization term employed. In this paper we consider a broad class of regularization terms, and develop conditions under which the regularized program is equivalent to MCP in both a global and a local sense. We establish the equivalence in a step by step manner that reveals some of the underlying structural properties of (1). We provide two different examples of regularization terms satisfying the general conditions, and also give some preliminary computational results evaluating their effectiveness in terms of both speed and solution quality. Over the course of our analysis (see Section 2), we also correct an (apparently as yet unidentified) erroneous result in the literature linking maximal cliques to local maximizers in (1) by constructing an example of a maximal clique whose characteristic vector is not a local maximizer of (1).

The paper is organized as follows. In Section 2, we develop a general regularized formulation of MCP and provide conditions under which the global/local maximizers of the regularized program are in one-one correspondence with the maximum/maximal cliques in  $G$ . In Section 3, we report on some preliminary computational results comparing the performance of two new regularization terms with the one proposed by Bomze in [3]. We conclude in Section 4.

**Notation.**  $\mathbf{0}$  and  $\mathbf{1}$  denote column vectors whose entries are all 0 and all 1 respectively and  $\mathbf{I}$  denotes the identity matrix, where the dimensions should be clear from the context.  $\nabla f(\mathbf{x})$  denotes the gradient of  $f$ , a row vector, and  $\nabla^2 f(\mathbf{x})$  denotes the Hessian of  $f$ . For a set  $\mathcal{Z}$ ,  $|\mathcal{Z}|$  is the number of elements in  $\mathcal{Z}$ .  $\mathbf{e}_i \in \mathbb{R}^n$  denotes the  $i$ -th column of the  $n \times n$  identity matrix. If  $\{s_i\}_{i=1}^n \subset \mathbb{R}$  is a finite sequence of length  $n$ , then  $\text{Diag}(\{s_i\}_{i=1}^n)$  is the  $n \times n$  diagonal matrix whose  $(i, i)$ th entry is  $s_i$ . If  $\mathbf{x} \in \mathbb{R}^n$ , then  $\text{supp}(\mathbf{x})$  denotes the *support* of  $\mathbf{x}$ , defined by  $\text{supp}(\mathbf{x}) = \{i : x_i \neq 0\}$ . Given vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $[\mathbf{x}, \mathbf{y}] := \{t\mathbf{x} + (1-t)\mathbf{y} : t \in [0, 1]\}$ . For a given positive integer  $n$ , we denote the set  $\{1, 2, \dots, n\}$  by  $[n]$ .  $\mathcal{S}_n$  is the set of permutations of  $[n]$ . If  $\mathbf{B} \in \mathbb{R}^{n \times n}$  is a symmetric matrix, we write  $\mathbf{B} \succeq \mathbf{0}$  if  $\mathbf{B}$  is positive semidefinite,  $\mathbf{B} \succ \mathbf{0}$  if  $\mathbf{B}$  is positive definite, and  $\mathbf{B} \preceq \mathbf{0}$  (resp.  $\mathbf{B} \prec \mathbf{0}$ ) when  $-\mathbf{B} \succeq \mathbf{0}$  (resp.  $-\mathbf{B} \succ \mathbf{0}$ ). If  $\mathbf{X} \subseteq \mathbb{R}^n$  and  $f : \mathbf{X} \rightarrow \mathbb{R}$ , then a point  $\mathbf{x} \in \mathbf{X}$  is a *local* (resp. *strict local*) *maximizer* of the problem

$\max \{f(\mathbf{x}) : \mathbf{x} \in \mathbf{X}\}$  if there exists some  $\epsilon > 0$  such that  $f(\mathbf{x}) \geq f(\tilde{\mathbf{x}})$  (resp.  $f(\mathbf{x}) > f(\tilde{\mathbf{x}})$ ) for every  $\tilde{\mathbf{x}} \in \mathbf{X}$  with  $0 < \|\tilde{\mathbf{x}} - \mathbf{x}\|_2 < \epsilon$ .  $\mathbf{x}$  is an *isolated* local maximizer if there exists some  $\epsilon > 0$  such that  $\tilde{\mathbf{x}}$  is not a local maximizer for any  $\tilde{\mathbf{x}} \in \mathbf{X}$  with  $0 < \|\tilde{\mathbf{x}} - \mathbf{x}\|_2 < \epsilon$ .  $\text{conv}(\mathbf{X})$  denotes the convex hull of  $\mathbf{X}$ .  $\text{span}_+(\mathbf{X}) := \{\sum_{i=1}^k \alpha_i \mathbf{x}^i : k \in \mathbb{N}_+, \alpha_i \geq 0, \mathbf{x}^i \in \mathbf{X} \forall i \in [k]\}$ .  $\mathbb{N}_+$  denotes the set of positive natural numbers.

## 2. GENERAL REGULARIZED FORMULATION.

Consider the following problem:

$$(3) \quad \begin{aligned} \max \quad & f(\mathbf{x}) := \mathbf{x}^\top \mathbf{A}\mathbf{x} + \Phi(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \Delta, \end{aligned}$$

where  $\Phi : X \rightarrow \mathbb{R}$  is a twice continuously differentiable function defined on some open set  $X \supset \Delta$ . Throughout this section, we will also assume that  $\Phi$  satisfies the following conditions for every  $\mathbf{x} \in \Delta$ :

- (C1)  $\nabla^2 \Phi(\mathbf{x}) \succeq \mathbf{0}$
- (C2)  $\|\nabla^2 \Phi(\mathbf{x})\|_2 < 2$
- (C3)  $\Phi$  is constant on the set

$$\mathcal{P}(\mathbf{x}) := \{\tilde{\mathbf{x}} \in \Delta : \exists \sigma \in \mathcal{S}_n \text{ such that } \tilde{x}_i = x_{\sigma(i)} \forall i \in [n]\},$$

where  $\mathcal{S}_n$  is the set of permutations of  $[n]$ . Note that (C1) is equivalent to requiring that  $\Phi$  is convex at  $\mathbf{x}$  and (C2) is equivalent to  $\nabla^2 \Phi(\mathbf{x}) - 2\mathbf{I} \prec \mathbf{0}$  (a fact that will be used later). Also, since  $\Phi \equiv 0$  satisfies (C1) – (C3) trivially, the results of this section will hold in particular for the original unpenalized formulation (1) when no additional assumptions are made on  $\Phi$ .

We will establish the global equivalence between (3) and MCP through a series of intermediate results. For any clique  $C$  define the set

$$\Delta(C) := \{\mathbf{x} \in \Delta : \text{supp}(\mathbf{x}) \subseteq C\},$$

and let

$$(4) \quad \Delta^0 := \bigcup_{C \text{ clique}} \Delta(C) = \{\mathbf{x} \in \Delta : \text{supp}(\mathbf{x}) \text{ is a clique}\}.$$

**Lemma 2.1.** *Let  $\mathbf{x} \in \Delta^0$ . Then,*

1. *For any  $\tilde{\mathbf{x}} \in \mathcal{P}(\mathbf{x}) \cap \Delta^0$  we have*

$$f(\tilde{\mathbf{x}}) = f(\mathbf{x}).$$

2. *For any  $\mathbf{0} \neq \mathbf{d} \in \mathbb{R}^n$  such that  $\mathbf{x} + t\mathbf{d} \in \Delta^0$  for all sufficiently small  $t > 0$  we have*

$$\mathbf{d}^\top \nabla^2 f(\mathbf{x}) \mathbf{d} < 0.$$

*Proof.* First, we note that for any  $\mathbf{z} \in \mathbb{R}^n$  such that  $\text{supp}(\mathbf{z})$  is a clique, we have

$$\begin{aligned} \mathbf{z}^\top \mathbf{A} \mathbf{z} &= \sum_{i=1}^n \sum_{j=1}^n z_i a_{ij} z_j = \sum_{i \in \text{supp}(\mathbf{z})} \sum_{j \in \text{supp}(\mathbf{z}) \setminus \{i\}} z_i z_j \\ &= \sum_{i \in \text{supp}(\mathbf{z})} \left( \sum_{j \in \text{supp}(\mathbf{z})} z_i z_j - z_i^2 \right) \end{aligned} \tag{5}$$

(5)

$$= (\mathbf{1}^\top \mathbf{z})^2 - \mathbf{z}^\top \mathbf{z} . \tag{6}$$

Next let  $\mathbf{x} \in \Delta^0$ . We prove the two parts separately.

**Part 1.** For any  $\tilde{\mathbf{x}} \in \mathcal{P}(\mathbf{x}) \cap \Delta^0$  we have

$$\begin{aligned} f(\tilde{\mathbf{x}}) &= \tilde{\mathbf{x}}^\top \mathbf{A} \tilde{\mathbf{x}} + \Phi(\tilde{\mathbf{x}}) \\ &= (\mathbf{1}^\top \tilde{\mathbf{x}})^2 - \tilde{\mathbf{x}}^\top \tilde{\mathbf{x}} + \Phi(\tilde{\mathbf{x}}) \\ &= (\mathbf{1}^\top \mathbf{x})^2 - \mathbf{x}^\top \mathbf{x} + \Phi(\mathbf{x}) \\ &= \mathbf{x}^\top \mathbf{A} \mathbf{x} + \Phi(\mathbf{x}) \\ &= f(\mathbf{x}) , \end{aligned} \tag{7}$$

(7)

(8)

(9)

where (7) and (9) are due to (6), since  $\tilde{\mathbf{x}}, \mathbf{x} \in \Delta^0$ , and (8) is due to (C3) and the assumption that  $\tilde{\mathbf{x}} \in \mathcal{P}(\mathbf{x})$ .

**Part 2.** Let  $\mathbf{0} \neq \mathbf{d} \in \mathbb{R}^n$  be any vector such that  $\mathbf{x} + t\mathbf{d} \in \Delta^0$  for all sufficiently small  $t > 0$ . Then when  $t$  is sufficiently small  $\text{supp}(\mathbf{d}) \subseteq \text{supp}(\mathbf{x} + t\mathbf{d})$ , which implies that  $\text{supp}(\mathbf{d})$  is a clique; moreover, since  $\mathbf{1}^\top \mathbf{x} = \mathbf{1}^\top (\mathbf{x} + t\mathbf{d}) = 1$ , we have  $\mathbf{1}^\top \mathbf{d} = 0$ . So by (6) we have

$$\begin{aligned} \mathbf{d}^\top \nabla^2 f(\mathbf{x}) \mathbf{d} &= 2\mathbf{d}^\top \mathbf{A} \mathbf{d} + \mathbf{d}^\top \nabla^2 \Phi(\mathbf{x}) \mathbf{d} \\ &= 2[(\mathbf{1}^\top \mathbf{d})^2 - \mathbf{d}^\top \mathbf{d}] + \mathbf{d}^\top \nabla^2 \Phi(\mathbf{x}) \mathbf{d} \\ &= -2\mathbf{d}^\top \mathbf{d} + \mathbf{d}^\top \nabla^2 \Phi(\mathbf{x}) \mathbf{d} \\ &= -\mathbf{d}^\top [2\mathbf{I} - \nabla^2 \Phi(\mathbf{x})] \mathbf{d} \\ &< 0 , \end{aligned} \tag{10}$$

(11)

where (11) follows from (C2).  $\square$

Now consider the following problem:

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \Delta(C) . \end{aligned} \tag{12}$$

**Proposition 2.1.** *The unique local (hence global) maximizer of (12) is  $\mathbf{x}(C)$ .*

*Proof.* Suppose by way of contradiction that there exist distinct local maximizers  $\mathbf{x}^1 \neq \mathbf{x}^2$  of (12). Then by Taylor's Theorem (see for example Proposition A.23 of [1])  $\exists \boldsymbol{\xi} \in [\mathbf{x}^1, \mathbf{x}^2] \subseteq \Delta$  such that

$$(13) \quad f(\mathbf{x}^2) = f(\mathbf{x}^1) + \nabla f(\mathbf{x}^1)\mathbf{d} + \frac{1}{2}\mathbf{d}^\top \nabla^2 f(\boldsymbol{\xi})\mathbf{d},$$

where  $\mathbf{d} = \mathbf{x}^2 - \mathbf{x}^1$ . Since  $\mathbf{x}^1, \mathbf{x}^2 \in \Delta$ , we have that  $\mathbf{x}^1 + t\mathbf{d} \in [\mathbf{x}^1, \mathbf{x}^2] \subseteq \Delta(C)$  for all sufficiently small  $t > 0$ . So by the standard first-order necessary local optimality condition (see for example, Section 1 of [10]), we have

$$(14) \quad \nabla f(\mathbf{x}^1)\mathbf{d} \leq 0.$$

Moreover, Part 2 of Lemma 2.1 implies

$$(15) \quad \mathbf{d}^\top \nabla^2 f(\boldsymbol{\xi})\mathbf{d} < 0.$$

Combining (13), (14), and (15), we obtain  $f(\mathbf{x}^2) < f(\mathbf{x}^1)$ . But then interchanging  $\mathbf{x}^2$  and  $\mathbf{x}^1$  the same argument can be used to show that  $f(\mathbf{x}^1) < f(\mathbf{x}^2)$ , a contradiction. Therefore, there is a unique local (hence global) maximizer of (12), say  $\mathbf{x}^*$ .

Next, we claim that  $\mathcal{P}(\mathbf{x}^*) \cap \Delta(C) = \{\mathbf{x}^*\}$ . Indeed, suppose by way of contradiction that  $\exists \tilde{\mathbf{x}} \in \mathcal{P}(\mathbf{x}^*) \cap \Delta(C)$  such that  $\tilde{\mathbf{x}} \neq \mathbf{x}^*$ . Since  $\text{supp}(\mathbf{x}^*) \subseteq C$  and  $C$  is a clique, Part 1 of Lemma 2.1 implies  $f(\tilde{\mathbf{x}}) = f(\mathbf{x}^*)$ . But then  $\tilde{\mathbf{x}}$  must be a global maximizer of (12), contradicting the uniqueness of  $\mathbf{x}^*$ . Hence, we must have  $\mathcal{P}(\mathbf{x}^*) \cap \Delta(C) = \{\mathbf{x}^*\}$ . Thus,  $x_i^* = x_j^*$  for any  $i, j \in C$ . Since  $\mathbf{x}^* \in \Delta(C)$ , this implies  $\mathbf{x}^* = \mathbf{x}(C)$ .  $\square$

**Remark 2.1.** *By Part 2 of Lemma 2.1, (12) is a strictly concave (and smooth) maximization problem. Thus, the uniqueness of the maximizer of (12) may be seen as following from standard results in the theory of convex optimization (for instance, Proposition B.4 in [1]).*

Next, consider the problem

$$(16) \quad \begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \Delta^0. \end{aligned}$$

**Proposition 2.2.** *A point  $\mathbf{x} \in \Delta^0$  is a local maximizer of (16) if and only if  $\mathbf{x} = \mathbf{x}(C)$  for some maximal clique  $C$ . Moreover, every local maximizer of (16) is strict.*

*Proof.* First, observe that for any local maximizer  $\mathbf{x}$  of (16), by (4) there exists some maximal clique  $C$  such that  $\mathbf{x} \in \Delta(C)$ ; and since  $\mathbf{x}$  is a local maximizer of (16), it is also a local maximizer of (12), which implies that  $\mathbf{x} = \mathbf{x}(C)$ , by Proposition 2.1. Thus, the proof will be complete when we show that every characteristic vector for a maximal clique is a strict local maximizer in (16). To this end, let  $C$  be a maximal clique, and suppose by way of contradiction that  $\mathbf{x}(C)$  is not a strict local maximizer of (16). Then, for every  $k \in \mathbb{N}_+$  there exists some  $\mathbf{x}^k \in \Delta^0$  with  $0 < \|\mathbf{x}^k - \mathbf{x}(C)\|_2 < 1/k$

such that  $f(\mathbf{x}^k) \geq f(\mathbf{x}(C))$ . Since there are only finitely many sets in the unions in (4), there must exist some clique  $C'$  and some subsequence  $(\mathbf{x}^{k_l})_{l=1}^\infty \subseteq (\mathbf{x}^k)_{k=1}^\infty$  such that  $\mathbf{x}^{k_l} \in \Delta(C')$  for each  $l \geq 1$ , with  $\mathbf{x}^{k_l} \rightarrow \mathbf{x}(C)$ . Hence,  $\mathbf{x}(C) \in \overline{\Delta(C')} = \Delta(C')$ , which implies  $C = \text{supp}(\mathbf{x}(C)) \subseteq C'$ . Since  $C$  is maximal, we must have that  $C = C'$ , and thus  $\mathbf{x}^{k_l} \in \Delta(C') = \Delta(C)$  for each  $l \geq 1$ . Thus,  $\mathbf{x}(C)$  is not a strict local maximizer of (12), contradicting Proposition 2.1. This completes the proof.  $\square$

**Proposition 2.3.** *If  $C^1$  and  $C^2$  are cliques, then*

$$|C^1| < |C^2| \Leftrightarrow f(\mathbf{x}(C^1)) < f(\mathbf{x}(C^2)).$$

*Proof.* Let  $C^1$  and  $C^2$  be cliques. First, suppose that  $|C^1| < |C^2|$ . Let  $C$  be any clique such that  $C \subset C^2$  and  $|C| = |C^1|$ . Then  $\mathbf{x}(C^1) \in \mathcal{P}(\mathbf{x}(C))$ . So, Part 1 of Lemma 2.1 implies  $f(\mathbf{x}(C^1)) = f(\mathbf{x}(C))$ . Moreover, by Proposition 2.1  $f(\mathbf{x}(C)) < f(\mathbf{x}(C^2))$ , since  $\mathbf{x}(C) \in \Delta(C^2)$ . Hence,  $f(\mathbf{x}(C^1)) < f(\mathbf{x}(C^2))$ . Conversely, suppose that  $f(\mathbf{x}(C^1)) < f(\mathbf{x}(C^2))$ . Then, by the proof of the forward direction we must have  $|C^1| \leq |C^2|$ . Moreover, if  $|C^1| = |C^2|$ , then  $\mathbf{x}(C^1) \in \mathcal{P}(\mathbf{x}(C^2))$  and Part 1 of Lemma 2.1 implies  $f(\mathbf{x}(C^1)) = f(\mathbf{x}(C^2))$ , a contradiction. Hence, we must have  $|C^1| < |C^2|$ .  $\square$

**Corollary 2.1.** *A point  $\mathbf{x} \in \Delta^0$  is a global maximizer of (16) if and only if  $\mathbf{x} = \mathbf{x}(C)$  for some maximum clique  $C$ .*

*Proof.* Let  $\mathbf{x} \in \Delta^0$ . Then  $\mathbf{x}$  is a global maximizer of (16) if and only if  $\mathbf{x}$  is a local maximizer and  $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$  for every local maximizer  $\bar{\mathbf{x}} \neq \mathbf{x}$ , which by Proposition 2.2 holds if and only if  $\mathbf{x} = \mathbf{x}(C)$  for some maximal clique  $C$  and  $f(\mathbf{x}(C)) \geq f(\mathbf{x}(\bar{C}))$  for every maximal clique  $\bar{C} \neq C$ . The corollary then follows from Proposition 2.3.  $\square$

**Proposition 2.4.** *For every clique  $C$ ,  $\mathbf{x}(C)$  is a global maximizer of (3) if and only if  $C$  is a maximum clique.*

*Proof.* We will show that there exists a global maximizer of (3) which lies in  $\Delta^0$ . The proof will then follow from Corollary 2.1. To this end, let  $\mathbf{x}$  be any global maximizer of (3). If  $\mathbf{x} \in \Delta^0$ , then we are done. So, suppose instead that  $\mathbf{x} \notin \Delta^0$ . Then  $\text{supp}(\mathbf{x})$  is not a clique and there exist indices  $i \neq j \in \text{supp}(\mathbf{x})$  such that  $a_{ij} = 0$ . Next, for any  $t \in [-x_i, x_j]$ , let  $\mathbf{x}(t) := \mathbf{x} + t(\mathbf{e}_i - \mathbf{e}_j)$ , and observe that by Taylor's Theorem there exists

some  $\boldsymbol{\xi} \in [\mathbf{x}, \mathbf{x}(t)]$  such that

$$\begin{aligned}
 f(\mathbf{x}(t)) &= f(\mathbf{x}) + t\nabla f(\mathbf{x})(\mathbf{e}_i - \mathbf{e}_j) + \frac{t^2}{2}(\mathbf{e}_i - \mathbf{e}_j)^\top \nabla^2 f(\boldsymbol{\xi})(\mathbf{e}_i - \mathbf{e}_j) \\
 (17) \quad &= f(\mathbf{x}) + \frac{t^2}{2} \left[ (2a_{ii} + 2a_{jj} - 4a_{ij}) + (\mathbf{e}_i - \mathbf{e}_j)^\top \nabla^2 \Phi(\boldsymbol{\xi})(\mathbf{e}_i - \mathbf{e}_j) \right] \\
 (18) \quad &= f(\mathbf{x}) + \frac{t^2}{2} \left[ (\mathbf{e}_i - \mathbf{e}_j)^\top \nabla^2 \Phi(\boldsymbol{\xi})(\mathbf{e}_i - \mathbf{e}_j) \right] \\
 (19) \quad &\geq f(\mathbf{x}) .
 \end{aligned}$$

Here, (17) follows from the first-order optimality condition at  $\mathbf{x}$ , which implies that  $\nabla f(\mathbf{x})(\mathbf{e}_i - \mathbf{e}_j) = 0$  since  $\mathbf{x}(t)$  is feasible for all  $t \in [-x_i, x_j]$ . Equality (18) follows from the fact that  $a_{ij} = 0 = a_{ii} = a_{jj}$ . And (19) follows from (C1). Thus, setting  $t = x_j$ , we obtain another global maximizer  $\mathbf{x}(t) \in \Delta$  such that  $\text{supp}(\mathbf{x}(t)) = \text{supp}(\mathbf{x}) \setminus \{j\} \subset \text{supp}(\mathbf{x})$ . We may repeat this process, gradually reducing the size of  $\text{supp}(\mathbf{x})$  while maintaining global maximality, until  $\text{supp}(\mathbf{x})$  is a clique (possibly of size 1), at which point the proof is complete, since then  $\mathbf{x} \in \Delta^0$ .  $\square$

By Proposition 2.2, a one-one correspondence exists between the local maximizers of (16) and the maximal cliques in  $G$ . However, as is already well-known in the case when  $\Phi \equiv 0$  (see the discussions pertaining to “infeasible” or “spurious” local optima in [3, 18, 19]), if  $\Delta^0$  in (16) is relaxed to  $\Delta \supseteq \Delta^0$  (in fact,  $\Delta = \text{conv}(\Delta^0)$ , but we need not prove this here), there may exist local maximizers of (3) that are *not* characteristic vectors for cliques, which may cause iterative optimization methods for solving (3) to fail, terminating without producing a clique.

Conversely, when  $\Phi \equiv 0$  there may exist characteristic vectors for maximal cliques which are not local maximizers in (3). Indeed, in the graph  $G$  in Figure 1, the sets  $C = \{1, 2\}$  and  $\hat{C} = \{3, 4, 5\}$  are both maximal cliques; and since  $\mathbf{x}(C), \mathbf{x}(\hat{C}) \in \Delta$ , we have that for all sufficiently small  $t > 0$ ,  $\mathbf{x}(C) + t\mathbf{d} \in \Delta$ , where  $\mathbf{d} = \mathbf{x}(\hat{C}) - \mathbf{x}(C)$ . Moreover, it is easy to check that by computation one has that  $\nabla f(\mathbf{x})\mathbf{d} = 0$  and  $\mathbf{d}^\top \nabla^2 f(\mathbf{x})\mathbf{d} = \frac{1}{6} > 0$ , where  $\mathbf{x} = \mathbf{x}(C)$ . Hence, for all sufficiently small  $t > 0$

$$\begin{aligned}
 f(\mathbf{x} + t\mathbf{d}) &= f(\mathbf{x}) + t\nabla f(\mathbf{x})\mathbf{d} + \frac{t^2}{2}\mathbf{d}^\top \nabla^2 f(\mathbf{x})\mathbf{d} \\
 &= f(\mathbf{x}) + \frac{t^2}{2}\mathbf{d}^\top \nabla^2 f(\mathbf{x})\mathbf{d} \\
 &> f(\mathbf{x}) .
 \end{aligned}$$

Thus,  $\mathbf{x}(C)$  is not a local maximizer of (3), despite the fact that  $C$  is maximal. We note that the above is a counterexample to [9, Corollary 2].

Hence, there is not necessarily any relationship (in either direction) between the local optima of (1) and the maximal cliques in  $G$ . However, we will see in the next proposition that for any strictly convex  $\Phi$  satisfying (C1) – (C3), the local maximizers of (3) are in *one-one* correspondence with the

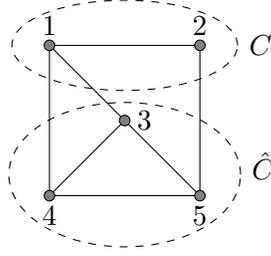


FIGURE 1. An example of a graph  $G$  and two maximal cliques  $C$  and  $\hat{C}$ . Here,  $\mathbf{x}(C)$  is not a local maximizer of (1) even though  $C$  is maximal.

characteristic vectors for maximal cliques. The proposition is based on three lemmas.

**Lemma 2.2.** *Let  $\mathbf{x} \in \Delta$  and let  $\mathcal{F}(\mathbf{x})$  denote the set of first-order feasible directions for (3) at  $\mathbf{x}$ , defined by*

$$\mathcal{F}(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{1}^\top \mathbf{d} = 0 \text{ and } d_i \geq 0 \text{ whenever } x_i = 0\}.$$

Then,

$$\mathcal{F}(\mathbf{x}) = \text{span}_+(\mathcal{F}(\mathbf{x}) \cap \mathcal{D}),$$

where  $\mathcal{D} = \bigcup_{\substack{i,j=1 \\ i \neq j}}^n \{\mathbf{e}_i - \mathbf{e}_j\}$ .

*Proof.* The lemma follows immediately from [10, Corollary 2.2] and the fact that  $\mathcal{D}$  is a reflective edge-description (defined in [10]) of  $\Delta$ .  $\square$

The next lemma is a restatement in the language of the present paper of the result [19, Proposition 3] which states that a characteristic vector for a maximal clique satisfies the first-order optimality conditions of (1).

**Lemma 2.3.** *If  $C$  is a maximal clique, then*

$$\mathbf{x}(C)^\top \mathbf{A} \mathbf{d} \leq 0 \quad \forall \mathbf{d} \in \mathcal{F}(\mathbf{x}(C)).$$

*Proof.* Let  $C$  be a maximal clique. We claim that for any  $\mathbf{d}^s \in \mathcal{F}(\mathbf{x}(C)) \cap \mathcal{D}$  we have  $\mathbf{x}(C)^\top \mathbf{A} \mathbf{d}^s \leq 0$ . Once this is shown, the proof will be complete, since by Lemma 2.2  $\forall \mathbf{d} \in \mathcal{F}(\mathbf{x}(C)) \exists k \in \mathbb{N}_+, \alpha_1, \alpha_2, \dots, \alpha_k \geq 0, \mathbf{d}^1, \mathbf{d}^2, \dots, \mathbf{d}^k \subseteq \mathcal{F}(\mathbf{x}(C)) \cap \mathcal{D}$  such that  $\mathbf{d} = \sum_{s=1}^k \alpha^s \mathbf{d}^s$ , and therefore

$$\mathbf{x}(C)^\top \mathbf{A} \mathbf{d} = \sum_{s \in [k]} \alpha^s \mathbf{x}(C)^\top \mathbf{A} \mathbf{d}^s \leq 0.$$

So, suppose that  $\mathbf{d}^s \in \mathcal{F}(\mathbf{x}(C)) \cap \mathcal{D}$ . Then  $\mathbf{d}^s = (\mathbf{e}_i - \mathbf{e}_j)$  for some  $i \neq j$ . Moreover, by definition of  $\mathcal{F}(\mathbf{x}(C))$  we have  $x_j(C) > 0$ ; that is,  $j \in C$ .

Hence,

$$\begin{aligned}
\mathbf{x}(C)^\top \mathbf{A} \mathbf{d}^s &= \sum_{t=1}^n \sum_{r=1}^n x_t(C) a_{tr} d_r^s \\
&= \frac{1}{|C|} \sum_{t \in C} (a_{ti} - a_{tj}) \\
(20) \qquad &= \frac{1}{|C|} \left[ \sum_{t \in C \setminus \{j\}} (a_{ti} - 1) + a_{ji} \right].
\end{aligned}$$

Here, (20) follows from the fact that  $C$  is a clique,  $j \in C$ , and  $a_{jj} = 0$ . We now consider two cases.

**Case 1:**  $i \in C$ . In this case, since  $a_{ii} = 0$ , the right hand side of (20) equals

$$(21) \qquad \frac{1}{|C|} \left[ \sum_{t \in C \setminus \{i, j\}} (a_{ti} - 1) + a_{ji} - 1 \right].$$

But since  $i \in C$  and  $C$  is a clique, we have that  $a_{ti} = 1$  for every  $t \in C \setminus \{i, j\}$ . Hence, (21) is equal to

$$\frac{1}{|C|} (a_{ji} - 1) \leq 0.$$

Thus,  $\mathbf{x}(C)^\top \mathbf{A} \mathbf{d}^s \leq 0$ .

**Case 2:**  $i \notin C$ . In this case, we have that (20) is equal to

$$(22) \qquad \frac{1}{|C|} \left[ \sum_{t \in C \setminus \{i, j\}} (a_{ti} - 1) + a_{ji} \right].$$

But since  $C$  is maximal and  $i \notin C$  there must exist some  $k \in C$  such that  $a_{ik} = 0$ . If  $k = j$ , then (22) is less than or equal to zero, since  $a_{ji} = 0$  and each of the terms in the first summation is less than or equal to zero. On the other hand, if  $k \neq j$ , then there exists a term in the first summation of (22) which is equal to  $-1$ , and hence, since  $a_{ji} \leq 1$ , (22) is less than or equal to zero. Thus,  $\mathbf{x}(C)^\top \mathbf{A} \mathbf{d}^s \leq 0$ . This completes the proof.  $\square$

**Lemma 2.4.** *Let  $\emptyset \neq S \subseteq \mathcal{V}$ . Then*

$$(23) \qquad \nabla \Phi(\mathbf{x}(S)) \mathbf{d} \leq 0 \quad \forall \mathbf{d} \in \mathcal{F}(\mathbf{x}(S)).$$

*Moreover, if  $\nabla^2 \Phi(\tilde{\mathbf{x}}) \succ \mathbf{0} \forall \tilde{\mathbf{x}} \in \Delta$ , then the inequality in (23) is strict whenever  $\text{supp}(\mathbf{d}) \not\subseteq S$ .*

*Proof.* First, observe that if (23) holds in the case when  $\nabla^2 \Phi(\tilde{\mathbf{x}}) \succ \mathbf{0} \forall \tilde{\mathbf{x}} \in \Delta$ , then it also holds for any  $\Phi$  (satisfying (C1) – (C3)). The argument is as follows: If  $\Phi$  satisfies (C1) – (C3), then for all sufficiently large  $k \in \mathbb{N}$ ,

the regularization function  $\Phi^{(k)} := \Phi + \frac{1}{k} \|\cdot\|_2^2$  also satisfies (C1) – (C3), and moreover for any  $\tilde{\mathbf{x}} \in \Delta$  we have that

$$\nabla^2 \Phi^{(k)}(\tilde{\mathbf{x}}) = \nabla^2 \Phi(\tilde{\mathbf{x}}) + \frac{2}{k} \mathbf{I} \succ \mathbf{0}.$$

Hence,

$$(24) \quad \nabla \Phi(\mathbf{x}(S)) \mathbf{d} + \frac{2}{k} \mathbf{x}(S)^\top \mathbf{d} = \nabla \Phi^{(k)}(\mathbf{x}(S)) \mathbf{d} \leq 0 \quad \forall \mathbf{d} \in \mathcal{F}(\mathbf{x}(S)).$$

But then, taking the limit of (24) as  $k \rightarrow \infty$  we obtain (23).

So suppose that  $\nabla^2 \Phi(\tilde{\mathbf{x}}) \succ \mathbf{0}$  for every  $\tilde{\mathbf{x}} \in \Delta$ . We first prove the following result for any  $\mathbf{x} \in \Delta$ :

$$(25) \quad \nabla \Phi(\mathbf{x})(\mathbf{e}_i - \mathbf{e}_j) < 0 \quad \forall i, j \in [n] \text{ such that } x_i < x_j.$$

To see this, let  $\mathbf{x} \in \Delta$ , suppose that  $i, j \in [n]$  are such that  $x_i < x_j$ , and let  $t := x_j - x_i > 0$  and  $\bar{\mathbf{x}} := \mathbf{x} + t(\mathbf{e}_i - \mathbf{e}_j)$ . Then  $\bar{\mathbf{x}} \in \mathcal{P}(\mathbf{x})$ , since  $\bar{x}_i = x_j$ ,  $\bar{x}_j = x_i$ , and  $\bar{x}_k = x_k \forall k \neq i, j$ . So, by (C3) we have  $\Phi(\bar{\mathbf{x}}) = \Phi(\mathbf{x})$ . But taking a Taylor expansion of  $\Phi$  about  $\mathbf{x}$ , we have that for some  $\boldsymbol{\xi} \in [\mathbf{x}, \bar{\mathbf{x}}] \subseteq \Delta$

$$\Phi(\mathbf{x}) = \Phi(\bar{\mathbf{x}}) = \Phi(\mathbf{x}) + t \nabla \Phi(\mathbf{x})(\mathbf{e}_i - \mathbf{e}_j) + \frac{t^2}{2} (\mathbf{e}_i - \mathbf{e}_j)^\top \nabla^2 \Phi(\boldsymbol{\xi})(\mathbf{e}_i - \mathbf{e}_j),$$

and hence

$$0 = t \nabla \Phi(\mathbf{x})(\mathbf{e}_i - \mathbf{e}_j) + \frac{t^2}{2} (\mathbf{e}_i - \mathbf{e}_j)^\top \nabla^2 \Phi(\boldsymbol{\xi})(\mathbf{e}_i - \mathbf{e}_j).$$

So since  $\nabla^2 \Phi(\boldsymbol{\xi}) \succ \mathbf{0}$  (by assumption) and  $t > 0$ , we have that

$$\nabla \Phi(\mathbf{x})(\mathbf{e}_i - \mathbf{e}_j) < 0.$$

Thus, (25) is proved. Moreover, note that by continuity of  $\nabla \Phi(\cdot)(\mathbf{e}_i - \mathbf{e}_j)$  we have in addition that

$$(26) \quad \nabla \Phi(\mathbf{x})(\mathbf{e}_i - \mathbf{e}_j) \leq 0 \quad \forall i, j \in [n] \text{ such that } x_i = x_j > 0.$$

In fact, by symmetry it is easy to see that the inequality in (26) must actually be an equality.

Now we prove the lemma. Let  $\emptyset \neq S \subseteq [n]$  and let  $\mathbf{d} \in \mathcal{F}(\mathbf{x}(S))$ . Then by Lemma 2.2 there exist  $k \in \mathbb{N}_+$ ,  $\mathbf{d}^1, \dots, \mathbf{d}^k \in \mathcal{F}(\mathbf{x}(S)) \cap \mathcal{D}$ , and  $\alpha^1, \alpha^2, \dots, \alpha^k \geq 0$  such that  $\mathbf{d} = \sum_{s=1}^k \alpha^s \mathbf{d}^s$ . Hence,

$$(27) \quad \nabla \Phi(\mathbf{x}(S)) \mathbf{d} = \sum_{s=1}^k \alpha^s \nabla \Phi(\mathbf{x}(S)) \mathbf{d}^s.$$

Next, note that by definition of  $\mathcal{D}$  for each  $s \in [k]$  there exist  $i, j \in [n]$  such that  $\mathbf{d}^s = (\mathbf{e}_i - \mathbf{e}_j)$ ; moreover, since  $(\mathbf{e}_i - \mathbf{e}_j) \in \mathcal{F}(\mathbf{x}(S))$ , we have that  $x_i \leq \frac{1}{|S|} = x_j$ . Hence, by (25), (26), and (27) we have that  $\nabla \Phi(\mathbf{x}(S)) \mathbf{d}^s \leq 0$  for each  $s \in [k]$ . So, by (27) we have  $\nabla \Phi(\mathbf{x}(S)) \mathbf{d} = \sum_{s=1}^k \alpha^s \nabla \Phi(\mathbf{x}(S)) \mathbf{d}^s \leq 0$ . Moreover, in the case where  $\text{supp}(\mathbf{d}) \not\subseteq S$ , there must exist some  $s \in [k]$  and some  $i, j \in [n]$  such that  $\mathbf{d}^s = (\mathbf{e}_i - \mathbf{e}_j)$  with  $i \notin S$  and  $\alpha^s > 0$ . By (25),

this implies  $\nabla\Phi(\mathbf{x}(S))\mathbf{d}^s < 0$ . Hence,  $\nabla\Phi(\mathbf{x}(S))\mathbf{d} < 0$ . This completes the proof.  $\square$

In the proof of the next proposition, we will use the following well-known second-order sufficient optimality condition (see [10]): A point  $\mathbf{x} \in \Delta$  is a local maximizer of (3) if

$$(28) \quad \nabla f(\mathbf{x})\mathbf{d} \leq 0 \quad \forall \mathbf{d} \in \mathcal{F}(\mathbf{x}),$$

and

$$(29) \quad \mathbf{d}^\top \nabla^2 f(\mathbf{x})\mathbf{d} < 0 \quad \forall \mathbf{0} \neq \mathbf{d} \in \mathcal{C}(\mathbf{x}),$$

where  $\mathcal{C}(\mathbf{x})$  is the critical cone at  $\mathbf{x}$  defined by

$$\mathcal{C}(\mathbf{x}) := \{\mathbf{d} \in \mathcal{F}(\mathbf{x}) : \nabla f(\mathbf{x})\mathbf{d} = 0\}.$$

**Proposition 2.5.** *Suppose that  $\nabla^2\Phi(\tilde{\mathbf{x}}) \succ \mathbf{0}$  for every  $\tilde{\mathbf{x}} \in \Delta$ . Then a point  $\mathbf{x} \in \Delta$  is a local maximizer of (3) if and only if  $\mathbf{x} = \mathbf{x}(C)$  for some maximal clique  $C$ . Moreover, every local maximizer of (3) is strict.*

*Proof.* Suppose that  $\nabla^2\Phi(\tilde{\mathbf{x}}) \succ \mathbf{0}$  for every  $\tilde{\mathbf{x}} \in \Delta$ . First, we claim that every local maximizer of (3) lies in  $\Delta^0$ . To this end, let  $\mathbf{x}$  be any local maximizer of (3). If  $\text{supp}(\mathbf{x})$  is a clique, then clearly  $\mathbf{x} \in \Delta^0$ . So, suppose by way of contradiction that  $\text{supp}(\mathbf{x})$  is not a clique. By applying an argument similar to the one given in the proof of Proposition 2.4, there exist indices  $i \neq j \in \text{supp}(\mathbf{x})$  such that for all  $t \in [-x_i, x_j]$  we have  $\mathbf{x}(t) = \mathbf{x} + t(\mathbf{e}_i - \mathbf{e}_j) \in \Delta$  and  $f(\mathbf{x}(t)) > f(\mathbf{x})$ , where the strict inequality here follows from (18) and the fact that  $\nabla^2\Phi(\xi) \succ \mathbf{0}$  for any  $\xi \in [\mathbf{x}, \mathbf{x}(t)] \subseteq \Delta$ . But this contradicts the fact that  $\mathbf{x}$  is a local maximizer of (3). Hence,  $\text{supp}(\mathbf{x})$  is a clique and  $\mathbf{x} \in \Delta^0$ .

So, every local maximizer of (3) lies in  $\Delta^0$  and is therefore a local maximizer of (16). Thus, by Proposition 2.2 every local maximizer of (3) is equal to  $\mathbf{x}(C)$  for some maximal clique  $C$ . Since there are only finitely many maximal cliques in  $G$ , there are only finitely many local maximizers of (3), which implies that every local maximizer of (3) is isolated, and is therefore a strict local maximizer (see for instance [9]).

To complete the proof, we must show that for any maximal clique  $C$ ,  $\mathbf{x}(C)$  is a local maximizer of (3). To see this, let  $C$  be a maximal clique and let  $\mathbf{x} := \mathbf{x}(C)$ . First, we show that the first-order condition (28) holds. Let  $\mathbf{d}^1 \in \mathcal{F}(\mathbf{x})$ . By Lemma 2.3, we have  $\mathbf{x}^\top \mathbf{A}\mathbf{d}^1 \leq 0$ , and by Lemma 2.4,  $\nabla\Phi(\mathbf{x})\mathbf{d}^1 \leq 0$ . Hence,

$$(30) \quad \nabla f(\mathbf{x})\mathbf{d} = 2\mathbf{x}^\top \mathbf{A}\mathbf{d}^1 + \nabla\Phi(\mathbf{x})\mathbf{d}^1 \leq 0.$$

So, we will be done when we show that the second-order condition (29) holds. To see this, let  $\mathbf{0} \neq \mathbf{d}^2 \in \mathcal{C}(\mathbf{x})$  be arbitrary. Then,

$$(31) \quad 0 = \nabla f(\mathbf{x})\mathbf{d}^2 = 2\mathbf{x}^\top \mathbf{A}\mathbf{d}^2 + \nabla\Phi(\mathbf{x})\mathbf{d}^2 \leq \nabla\Phi(\mathbf{x})\mathbf{d}^2,$$

where the last inequality follows from Lemma 2.3, since  $\mathbf{d}^2 \in \mathcal{F}(\mathbf{x})$ . Thus, by the second statement in Lemma 2.4 we must have  $\text{supp}(\mathbf{d}^2) \subseteq C$ , implying that  $\mathbf{x} + t\mathbf{d}^2 \in \Delta(C) \subseteq \Delta^0$  for all sufficiently small  $t > 0$ . And so by Part 2 of Lemma 2.1 we have

$$\mathbf{d}^{2\top} \nabla^2 f(\mathbf{x}) \mathbf{d}^2 < 0 .$$

This completes the proof.  $\square$

### 3. PRELIMINARY NUMERICAL RESULTS.

In this section, we conduct some preliminary numerical experiments on three different regularization functions satisfying the conditions outlined in Section 2 in order to give an indication of the potential impact of different regularization terms on the performance of a local optimization algorithm applied to (3). If in practice a maximum clique (rather than merely a maximal clique) is sought, the local optimization algorithm we employ in our experiments would need to be incorporated into a global optimization framework, such as branch and bound, in order to ensure convergence to a global maximizer.

**3.1. Regularization functions.** We considered the following three regularization terms, with the indicated choices of parameters:

$$(32) \quad \Phi_B(\mathbf{x}) := \frac{1}{2} \|\mathbf{x}\|_2^2,$$

$$(33) \quad \Phi_1(\mathbf{x}) := \alpha_1 \|\mathbf{x} + \epsilon \mathbf{1}\|_p^p, \quad \epsilon > 0, \quad p > 2, \quad 0 < \alpha_1 < \frac{2}{p(p-1)(1+\epsilon)^{p-2}},$$

$$(34) \quad \Phi_2(\mathbf{x}) := \alpha_2 \sum_{i=1}^n (e^{-\beta x_i} - 1), \quad \beta > 0, \quad 0 < \alpha_2 < \frac{2}{\beta^2}.$$

Here,  $\Phi_B$  is the 2-norm regularization function introduced by Bomze et al. [3], and  $\Phi_1$  is a generalization of  $\Phi_B$  to p-norms where  $p > 2$ .  $\Phi_2$  is a well-known (for instance, see [6]) approximation of the following non-smooth function:

$$(35) \quad \tilde{\Phi}(\mathbf{x}) = -\alpha_2 \|\mathbf{x}\|_0 ,$$

where  $\|\mathbf{x}\|_0 = \text{supp}(\mathbf{x})$ . The motivation behind the choice of  $\Phi_2$  is as follows. By definition of  $\tilde{\Phi}$ , maximizing  $\mathbf{x}^\top \mathbf{A} \mathbf{x} + \tilde{\Phi}$  over  $\Delta$  is closely related to the problem of finding a solution to (1) which has the smallest support (ie. the maximum sparsity). Following the argument laid out in the proof of Proposition 2.4, from any global maximizer of (1) which is *not* a characteristic vector for a maximum clique, there exists a path leading to another global maximizer which *is* a characteristic vector for a maximum clique and whose support is strictly smaller than that of the starting point. Hence, the global maximizers of (1) which have the smallest support are necessarily the characteristic vectors for maximum cliques. Thus,  $\Phi_2$  is a somewhat natural choice in our present context.

Next, we show that each of the regularization functions above satisfies the conditions of Section 2. For any  $\delta \in (0, \epsilon)$ , where  $\epsilon > 0$  is the value used in the definition of  $\Phi_1$ , let  $X := \text{conv}(\cup_{i=1}^n \mathcal{B}_\delta(\mathbf{e}_i)) \subseteq \mathbb{R}^n$ , where  $\mathcal{B}_\delta(\mathbf{e}_i) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{e}_i\|_2 < \delta\}$ . Then,  $X$  is open,  $X \supset \Delta$ , and  $\Phi_B$ ,  $\Phi_1$ , and  $\Phi_2$  are each well-defined on  $X$ . Moreover, it is easy to check that  $\Phi_B$ ,  $\Phi_1$ , and  $\Phi_2$  are each twice continuously differentiable over  $X$  and that for any  $\mathbf{x} \in \Delta$  we have

$$(36) \quad \begin{aligned} \nabla^2 \Phi_B(\mathbf{x}) &= \mathbf{I} \succ \mathbf{0}, \\ \nabla^2 \Phi_1(\mathbf{x}) &= \alpha_1 p(p-1) \text{Diag}(\{(x_i + \epsilon)^{p-2}\}_{i=1}^n) \succ \mathbf{0}, \text{ and} \\ \nabla^2 \Phi_2(\mathbf{x}) &= \alpha_2 \beta^2 \text{Diag}(\{e^{-\beta x_i}\}_{i=1}^n) \succ \mathbf{0}, \end{aligned}$$

where the positive definiteness of the Hessians in (36) follows from the choice of parameters. Hence,  $\Phi_B$ ,  $\Phi_1$ , and  $\Phi_2$  each satisfy (C1) strictly. Next, observe that for any  $\mathbf{x} \in \Delta$  we have

$$\begin{aligned} \|\nabla^2 \Phi_B(\mathbf{x})\|_2 &= 1 < 2, \\ \|\nabla^2 \Phi_1(\mathbf{x})\|_2 &= \alpha_1 p(p-1) \max \{(x_i + \epsilon)^{p-2}\}_{i=1}^n \\ &< \frac{2}{(1 + \epsilon)^{p-2}} \max \{(x_i + \epsilon)^{p-2}\}_{i=1}^n \\ &\leq \frac{2}{(1 + \epsilon)^{p-2}} (1 + \epsilon)^{p-2} = 2, \text{ and} \\ \|\nabla^2 \Phi_2(\mathbf{x})\|_2 &= \alpha_2 \beta^2 \max \{e^{-\beta x_i}\}_{i=1}^n \\ &< 2 \max \{e^{-\beta x_i}\}_{i=1}^n \\ &\leq 2. \end{aligned}$$

Thus, (C2) is satisfied for each of  $\Phi_B$ ,  $\Phi_1$ , and  $\Phi_2$ . That (C3) holds follows easily from the fact that  $\Phi_B$ ,  $\Phi_1$ , and  $\Phi_2$  are each separable and the coefficients associated with the terms  $x_i$  are independent of  $i$ .

**3.2. The testing set.** In the experiments, we considered different families of widely used maximum-clique instances belonging to the DIMACS benchmark [11]:

- **C family:** Random graphs Cx.y, where  $x$  is the number of nodes and  $y$  the edge probability;
- **DSJC family:** Random graphs DSCJx.y. Here again,  $x$  is the number of nodes and  $y$  the edge probability;
- **brock family:** Random graphs with cliques hidden among nodes that have a relatively low degree;
- **gen family:** Artificially generated graphs with large, known embedded clique;
- **hamming family:** hamming-a-b are graphs on  $a$ -bit words with an edge if and only if the two words are at least hamming distance  $b$  apart;

TABLE 1. DIMACS instances used in the tests.

instance	best known	nodes	edges	graph degrees		best degrees	
				median	iqr	median	iqr
C125.9	34	125	6963	112	5	114.5	4.75
C250.9	44	250	27984	224	6	227	5
C500.9	57	500	112332	449	9	455	9
C1000.9	68	1000	450079	900	13	907	11.25
C2000.5	16	2000	999836	999	30	1006	11.5
C2000.9	80	2000	1799532	1800	18	1803	15.25
DSJC500_5	13	500	125248	250	16	259	14
DSJC1000_5	15	1000	499652	500	20	503	23
brock200_2	12	200	9876	99	10	101	11
brock200_4	17	200	13089	131	8	134	6
brock400_2	29	400	59786	299	10	299	9
brock400_4	33	400	59765	299	11	299	9
brock800_2	24	800	208166	521	18	516.5	20.25
brock800_4	26	800	207643	519	18.25	512	20.25
gen200_p0.9_44	44	200	17910	180	8	179.5	4.25
gen200_p0.9_55	55	200	17910	179	7.25	179	5.5
gen400_p0.9_55	55	400	71820	360	13.25	359	6
gen400_p0.9_65	65	400	71820	361	14	359	9
gen400_p0.9_75	75	400	71820	359	13	359	8
hamming8-4	16	256	20864	163	0	163	0
hamming10-4	40	1024	434176	848	0	848	0
keller4	11	171	9435	110	8	112	17
keller5	27	776	225990	578	38	578	33
keller6	59	3361	4619898	2724	50	2724	50
p_hat300-1	8	300	10933	73	39	103	20
p_hat300-2	25	300	21928	146.5	73	213	18
p_hat300-3	36	300	33390	224	38	251	15.25
p_hat700-1	11	700	60999	174.5	87	250	22.5
p_hat700-2	44	700	121728	353	177.5	508	31.5
p_hat700-3	62	700	183010	526	89	602	14
p_hat1500-1	12	1500	284923	383	197	509	82
p_hat1500-2	65	1500	568960	763	387	1100	37
p_hat1500-3	94	1500	847244	1132.5	192	1297.5	25.75

- **keller family**: Instances based on Keller’s conjecture [13] on tilings using hypercubes;
- **p\_hat family**: Random graphs generated with the p-hat generator, which is a generalization of the classical uniform random graph generator. Graphs generated with p-hat have wider node degree spread and larger cliques than uniform graphs.

In Table 1, we report the names of the instances used (**instance**), the best known solutions (**best known**), the number of nodes and edges in the instances (**nodes** and **edges**), and the median and interquartile range related to the graph degrees (**median** and **iqr** in column **graph degrees**) as well as the median and iqr of the degrees of the nodes lying in the best known solution (**median** and **iqr** in column **best degrees**).

**3.3. Experiments.** To conduct our tests, we developed a multistart framework in MATLAB that uses a hybrid algorithm as local optimizer. It combines the `fmincon` solver with the Frank-Wolfe method [7]. For each instance, we ran 100 trials each with a different randomly generated point in  $\Delta$  as a starting guess. The same starting guesses were used for all formulations. Since the iterates of the Frank-Wolfe method we used are only guaranteed

to converge to a point satisfying the first-order conditions, we omitted the trials in which the final iterate was not a true local optimizer from the statistical computations in the tables below. (Another potential way of dealing with this issue, which we leave to a future work, would be to take a step in an ascent direction whenever the final iterate is not a local maximizer, and then rerun the algorithm using the new point as a starting guess.) Once a local maximizer  $\mathbf{x}^*$  is obtained, the associated clique is constructed by taking  $C^* := \text{supp}(\mathbf{x}^*)$ . In our experiments, we used the parameter values  $p = 3$ ,  $\epsilon = 10^{-9}$ , and  $\beta = 5$ . Furthermore,  $\alpha_1$  and  $\alpha_2$  were suitably chosen in order to satisfy condition (33) and (34) respectively. All the tests were performed on an Intel Core i7-3610QM 2.3 GHz, 8GB RAM.

In Table 2, we report the largest clique size obtained (**max**), mean (**mean**), standard deviation (**std**) and the average CPU time (**CPU time**) over the 100 runs for each instance and each of the three formulations. When the largest clique obtained is the same as the largest known clique size, the result is reported in bold. Observe that with the exception of the **keller** instances, the average clique size obtained from using either of  $\Phi_1$  or  $\Phi_2$  was strictly larger than the average obtained from using  $\Phi_B$ . Overall,  $\Phi_1$  performed slightly better than  $\Phi_2$  yielding a strictly larger clique in 17 out of the 33 instances. Next, taking a look at the results related to the **C** and **p\_hat** families, we notice that as the number of nodes increases (and also the number of edges increases) finding a solution close to the best known gets harder and harder for all three formulations. This is likely due to the simplicity of the global optimization approach that we use to solve the problem. However, for the smaller instances in these groups  $\Phi_1$  and  $\Phi_2$  performed quite well. In particular, the formulation using  $\Phi_1$  found the largest known clique size in 6 of the instances. The **DSJC**, **brock**, and **gen** families all confirm the good behavior of the proposed formulations. Indeed, in all cases the solutions found were closer (sometimes significantly) to the best known clique than the ones found using  $\Phi_B$ .

#### 4. CONCLUSIONS.

We described a general regularized continuous formulation for the MCP and developed conditions which guarantee an equivalence between the original problem and the continuous reformulation in both a global and a local sense. We have also proved the results in a step by step manner which we hope reveals some of the underlying structural properties of the formulation. We further proposed two specific regularizers that satisfy the general conditions given in the paper, and compared the two related continuous formulations with the one proposed in [3] on different families of widely used maximum-clique instances belonging to the DIMACS benchmark. The numerical results, albeit still preliminary, seem to confirm the effectiveness of

TABLE 2. Results obtained on the DIMACS instances.

instances	$\Phi_B$				$\Phi_1$				$\Phi_2$			
	max	mean	std	CPU time	max	mean	std	CPU time	max	mean	std	CPU time
C125.9	<b>34</b>	32.83	0.92	0.13	<b>34</b>	33.17	0.38	0.69	<b>34</b>	33.22	0.42	0.25
C250.9	40	37.08	1.30	0.43	<b>44</b>	40.79	1.09	1.31	<b>44</b>	40.77	1.13	0.77
C500.9	49	45.51	1.46	1.80	54	51.03	1.57	5.14	54	50.92	1.68	3.36
C1000.9	60	54.40	2.26	10.01	63	58.90	1.91	23.57	63	59.16	1.94	15.76
C2000.5	13	11.46	0.70	56.09	15	12.90	1.08	101.40	15	12.95	0.94	66.69
C2000.9	69	61.75	1.86	56.67	67	63.35	1.61	106.22	68	63.89	1.96	76.71
DSJC500.5	10	9.43	0.61	2.20	11	10.45	0.55	4.59	12	10.19	0.44	2.36
DSJC1000.5	13	10.68	0.97	10.29	14	11.86	0.79	25.53	14	11.53	1.20	13.69
brock200_2	9	8.00	0.58	0.25	10	9.30	0.67	0.64	10	9.04	0.21	0.37
brock200_4	14	12.60	0.87	0.30	15	13.44	0.63	1.35	15	13.40	0.64	0.43
brock400_2	23	19.84	1.00	1.16	24	22.14	1.26	3.49	24	21.80	1.23	1.60
brock400_4	22	19.77	1.33	1.10	24	21.47	1.15	2.62	24	21.46	1.05	1.51
brock800_2	18	15.60	1.12	6.14	20	17.32	0.96	15.58	20	17.28	0.94	7.81
brock800_4	17	15.22	0.94	6.53	20	17.12	1.05	17.79	19	17.08	1.00	8.83
gen200_p0.9_44	36	33.37	1.25	0.31	40	37.51	1.05	1.45	40	37.43	1.12	0.49
gen200_p0.9_55	40	36.93	1.02	0.31	41	38.95	1.14	1.06	41	38.98	1.20	0.41
gen400_p0.9_55	48	44.19	1.57	0.96	51	49.01	0.87	3.22	51	49.03	0.86	1.64
gen400_p0.9_65	48	43.97	2.28	1.13	51	48.05	1.75	4.14	51	48.25	1.90	1.79
gen400_p0.9_75	46	42.93	1.41	1.05	49	47.76	1.04	5.69	50	47.65	1.38	2.91
hamming8-4	<b>16</b>	13.61	2.27	0.43	<b>16</b>	15.73	1.00	0.66	<b>16</b>	15.73	1.01	0.51
hamming10-4	34	30.54	1.23	10.69	<b>40</b>	33.45	1.37	35.23	<b>40</b>	33.47	1.42	18.99
keller4	8	7.17	0.38	0.19	7	7.00	0.00	0.73	9	7.02	0.20	0.18
keller5	16	15.02	0.15	4.40	15	15.00	0.00	7.54	15	15.00	0.00	4.39
keller6	34	32.86	1.46	388.05	34	33.83	0.61	290.79	34	32.80	1.48	222.96
p_hat300-1	7	7.00	0.00	0.54	<b>8</b>	8.00	0.00	0.88	<b>8</b>	8.00	0.00	0.52
p_hat300-2	24	24.00	0.00	0.55	<b>25</b>	24.01	0.10	1.89	24	24.00	0.00	0.44
p_hat300-3	33	31.15	0.76	0.62	<b>36</b>	33.39	0.70	1.86	<b>36</b>	33.20	0.64	0.90
p_hat700-1	9	7.24	0.62	3.97	9	8.15	0.43	13.31	9	7.94	0.55	6.15
p_hat700-2	43	41.53	0.75	4.49	<b>44</b>	43.61	0.70	10.07	<b>44</b>	43.50	0.72	4.74
p_hat700-3	60	58.71	0.87	4.77	61	58.99	0.69	10.55	61	59.21	0.80	7.73
p_hat1500-1	11	9.00	0.86	23.26	11	9.36	1.12	41.27	11	9.83	0.65	25.40
p_hat1500-2	62	59.09	1.10	29.06	62	60.79	0.87	44.76	64	62.30	1.15	30.42
p_hat1500-3	87	81.90	1.80	31.76	88	86.99	0.39	225.12	92	88.21	1.98	38.19

the proposed regularizers; that is, when a local optimization method is applied to the new regularized formulations, cliques of high quality can often be obtained in reasonable computational times.

#### ACKNOWLEDGMENTS.

This work has been partly developed while the first author was an ER of the “MINO: Mixed-Integer Nonlinear Optimization” program funded by the European Union. The authors therefore gratefully acknowledge the financial support of the European Union’s Horizon 2020 Marie Curie Network for Initial Training (ITN) programme under grant agreement No. 316647.

#### REFERENCES

1. D. P. Bertsekas, *Nonlinear programming*, Athena Scientific, Belmont, MA, 1999.
2. I. R. Bomze, M. Pelillo, and V. Stix, *Approximating the maximum weight clique using replicator dynamics*, IEEE Transactions on Neural Networks **11** (2000), no. 6, 1228–1241.
3. Immanuel M Bomze, *Evolution towards the maximum clique*, Journal of Global Optimization **10** (1997), no. 2, 143–164.
4. Immanuel M. Bomze, Marco Budinich, Panos M. Pardalos, and Marcello Pelillo, *The maximum clique problem*, pp. 1–74, Springer US, Boston, MA, 1999.

5. Immanuel M. Bomze, Marco Budinich, Marcello Pelillo, and Claudio Rossi, *Annealed replication: a new heuristic for the maximum clique problem*, Discrete Applied Mathematics **121** (2002), no. 1–3, 27 – 49.
6. P.S. Bradley, O.L. Mangasarian, and J.B. Rosen, *Parsimonious least norm approximation*, Computational Optimization and Applications **11** (1998), 5 – 21.
7. Marguerite Frank and Philip Wolfe, *An algorithm for quadratic programming*, Naval research logistics quarterly **3** (1956), no. 1-2, 95–110.
8. L. E. Gibbons, D. W. Hearn, and P. M. Pardalos, *A continuous based heuristic for the maximum clique problem*, Second DIMACS Implementation Challenge (D. S. Johnson and M. Trick, eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 26, American Mathematical Society, USA, 1993, pp. 103–124.
9. L. E. Gibbons, D. W. Hearn, P. M. Pardalos, and M. V. Ramana, *Continuous characterizations of the maximum clique problem*, Math. Oper. Res. **22** (1997), 754–768.
10. W. W. Hager and J. T. Hungerford, *Optimality conditions for maximizing a function over a polyhedron*, Math. Program. **145** (2014), 179–198.
11. David S. Johnson and Michael A. Trick, *Cliques, coloring, and satisfiability: second dimacs implementation challenge, october 11-13, 1993*, vol. 26, American Mathematical Soc., 1996.
12. R.M. Karp, *Reducibility among combinatorial problems*, Complexity of Computer Computations: Proc. of a Symp. on the Complexity of Computer Computations (New York, NY), Plenum Press, 1972, pp. 85–103.
13. Ott-Heinrich Keller, *Über die lückenlose erfüllung des raumes mit würfeln.*, Journal für die reine und angewandte Mathematik **163** (1930), 231–248.
14. Antonina Kuznetsova and Alexander Strekalovsky, *On solving the maximum clique problem*, Journal of Global Optimization **21** (2001), no. 3, 265–288.
15. T. S. Motzkin and E. G. Straus, *Maxima for graphs and a new proof of a theorem of Turán*, Canad. J. Math. **17** (1965), 533–540.
16. K. G. Murty and S. N. Kabadi, *Some NP-complete problems in quadratic and linear programming*, Math. Program. **39** (1987), 117–129.
17. P. M. Pardalos and G. Schnitger, *Checking local optimality in constrained quadratic programming can be NP-hard*, Oper. Res. Lett. **7** (1988), 33–35.
18. Marcello Pelillo, *Relaxation labeling networks for the maximum clique problem*, J. Artif. Neural Netw. **2** (1996), no. 4, 313–328.
19. Marcello Pelillo and Arun Jagota, *Feasible and infeasible maxima in a quadratic program for maximum clique*, J. Artif. Neural Netw. **2** (1996), no. 4, 411–420.
20. Qinghua Wu and Jin-Kao Hao, *A review on algorithms for maximum clique problems*, European Journal of Operational Research **242** (2015), no. 3, 693 – 709.

RACETRAC STORE SUPPORT CENTER, 200 GALLERIA PKWY SE, ATLANTA, GA 30339  
*E-mail address:* [jamesthungerford@gmail.com](mailto:jamesthungerford@gmail.com)

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PADOVA, VIA TRIESTE 63, 35121  
 PADOVA - ITALY  
*E-mail address:* [rinaldi@math.unipd.it](mailto:rinaldi@math.unipd.it)