Tractable Continuous Approximations for Constraint Selection via Cardinality Minimization

Miju Ahn*[1], Harsha Gangammanavar†[1], and David Troxell‡[2]

[1]Department of Operations Research and Engineering Management, Southern Methodist University, Dallas TX

[2]Department of Statistics, Stanford University, Stanford, CA

First submission: October 14, 2022.

Abstract

We study a cardinality minimization problem that simultaneously minimizes an objective function and the cardinality of unsatisfied soft constraints. This paper proposes two continuous approximation methods which reformulate the discrete cardinality as complementarity constraints and difference-of-convex functions. We show that, under suitable conditions, local and stationary solutions of the approximation problems recover local minimizers of the cardinality minimization problem. To demonstrate effectiveness, we apply the proposed methods to applications where violating as few preference conditions (or soft constraints) is desired. The performance of the new methods is compared to benchmark formulations used in practice. Our numerical study supports the use of methods based on our new approximations for cardinality minimization that produce comparable solutions while improving computational efficiency.

1 Introduction

In this paper, we study the threshold-based cardinality minimization problems (CMP) that take the following form:

$$\min \ \left\{ \ f_0(\mathbf{x}) + \lambda \|(\max\{f_i(\mathbf{x}) - \tau_i, 0\})_{i=1,\ldots,m}\|_0 \quad \text{subject to} \quad \mathbf{x} \in \mathcal{X} \ \right\}. \tag{1}$$

Here, $\lambda > 0$ is the penalty parameter and $\| \bullet \|_0$ is the ℓ_0-norm. The objective function $f_0(\cdot)$ and the constraints $f_i(\cdot)$ for all $i = 1, \ldots, m$, are convex functions. The set $\mathcal{X} \subset \mathbb{R}^n$ is a

*mijua@smu.edu

†harsha@smu.edu

‡dtroxell@stanford.edu

compact convex set that captures what we will refer to as the "hard constraints". These hard constraints can be attributed to the operational requirements that cannot be violated. On the other hand, we have a set of constraints $f_i(\mathbf{x}) \leq \tau_i$ for all $i = 1, \ldots, m$, that correspond to preference requirements that are allowed to be violated. The vector $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$ of these constraints captures the preference or threshold assigned by the decision maker. In this sense, the second set of constraints can be considered as "soft constraints". While the soft constraints are allowed to be violated, the desire is to minimize the number of violations given by $\|(\max\{f_i(\mathbf{x}) - \tau_i, 0\})_{i=1,\ldots,m}\|_0$.

A form of the problem in (1) that is of particular interest to us in this paper is

$$\min \ \left\{ \ f_0(\mathbf{x}) + \lambda \|(\max\{|x_i| - \tau_i, 0\})_{i=1,\ldots,n}\|_0 \quad \text{subject to} \quad \mathbf{x} \in \mathcal{X} \ \right\}. \tag{2}$$

We refer to the above as the threshold-based or *two-tailed cardinality minimization problem* and denote it as τ-CMP. We obtain the τ-CMP by setting $f_i(\mathbf{x}) = |x_i|$ for all $i = 1, \ldots, m$ in (1) with $m = n$. This problem essentially minimizes the count of unacceptably large elements of the vector $\mathbf{x}$.

It must be noted that the parameter τ_i can be absorbed into the function f_i and we could state the soft constraints in (1) as $f_i(\mathbf{x}) \leq 0$ for all i. While the analysis we present in this paper is applicable to this modification, our choice to retain τ_i as separate parameters is to emphasize that the decision-maker can exert a choice in selecting these parameter.

A special case of (2) studied in [14] which considers the cardinality of large valued elements of $\mathbf{x}$ exceeding the prescribed threshold $\boldsymbol{\tau}$. The authors refer to the special case as the cardinality of upper tail minimization problem. Another notable special case of (2) is when $\tau_i = 0$ for all $i = 1, \ldots, n$. The second term of (2) is often referred to as the ℓ_0-norm, denoted $\|\mathbf{x}\|_0$, which counts the number of nonzero components of $\mathbf{x}$. Despite its misleading name, ℓ_0-norm does not satisfy the properties of a norm. It is a discrete and nonconvex function, and directly minimizing a problem involving such a function is known to be computationally intractable. A predominant approach to solve such problems is by replacing the discontinuous function by the ℓ_1-norm given by $\|\mathbf{x}\|_1 = \sum_{j=1}^{n} |\mathbf{x}_j|$. This replacement results in a convex optimization problem that can be solved efficiently using off-the-shelf solvers. However, the solutions to the optimization problem with ℓ_1-norm result in suboptimal solutions in general. This observation motivated another stream of approaches that use continuous nonconvex surrogates for the ℓ_0-function. One such approximation related to our work is the capped-ℓ_1 function [11, 12]. This function approximates the ℓ_0-function by the ℓ_1-norm around the origin and a constant elsewhere. We apply a similar approximation for (2) in Section 2.2.1 where we further discuss the relationship. The nonconvex approximation methods showed superior performance when applied to various applications including image reconstruction [19], signal recovery [2, 4], and deep learning methods [3].

This paper is motivated by applications where generalizing above mentioned problems involving discrete cardinality can be beneficial. Such generalization allows the optimization problem

to selectively enforce soft constraints while minimizing the objective function of concern. Our methods introduce exact and approximate reformulations of the discrete problems (1) and (2). The overarching goal of our study is to provide computationally tractable formulations and understand how one can recover solutions to the discrete problem by solving the reformulations. The problems (1) and (2) with nonzero $\boldsymbol{\tau}$ arise in several application settings. The following application motivates our study.

1.1 Motivating application

Radiation therapy, specifically intensity-modulated radiation therapy (IMRT), has emerged as one of the principal treatment options for various types of cancer. IMRT is a minimally invasive treatment option where radiation of ionized beams is projected on a region of interest surrounding the tumor tissues. The region of interest includes other healthy tissues (organs at risk and normal body tissues) that are invariably exposed to radiation. Although the healthy tissues can repair themselves, limiting the exposure of organs at risk to within clinically acceptable thresholds is desirable. One achieves precise radiation delivery by shaping the dose pattern across the region of interest [8]. The desired dose pattern is generated using a multileaf collimator system by determining the angle and intensity of radiation for a set of beamlets.

We refer to the problem of designing a suitable dose pattern as the fluence map optimization problem. This problem can be formulated as a mathematical program. A clinician determines a prescription dose for the tumor and tolerance doses for organs at risk. We design the dose pattern for given values of prescription and tolerance doses. The angle and radiation amounts for the set of beamlets on the collimator constitute the decision vector of the program. The objective is to minimize the difference between radiation delivered to the tumor tissues and the prescription dose. This objective function $f_0(\cdot)$ is often modeled as a quadratic function. For an organ at risk i, the amount of dose delivered, captured by the function $f_i(\cdot)$, must be within the tolerable dose τ_i for $i = 1, \ldots, m$. With this, we see that the fluence map optimization problem takes the form of (1). We present the detailed mathematical model when we return to this problem in the numerical experiments.

1.2 Contribution

The main contributions of this work are threefold.

1. *Continuous approximations of the CMP.* We present two alternative continuous approximations of the CMP based on difference-of-convex (DC) programming. In the first approach, we rely upon mathematical programming with complementarity constraint (MPCC) reformulation of the CMP. This reformulation utilizes additional variables that indicate the occurrence and degree of violation, respectively. By relaxing the complementarity constraint using a penalty function that is DC-representable, we achieve a relaxed approximation of the CMP problem. We refer to this approach as the MPCC-DC. In contrast to MPCC-DC

approach, we directly develop a DC approximation of the CMP objective function in the second approach. Consequently, we refer to this approach as direct-DC. Both our approximations are amenable to applying the difference-of-convex algorithm, thereby providing a viable solution method.

2. *Analysis of Solutions.* We develop the relationships between solutions to models that result from the alternative continuous approximations. First, we show the equivalence of the MPCC and the CMP in terms of their optimal solutions. Our subsequent analysis shows that under suitable conditions, a locally optimal solution of the MPCC-DC formulation recovers a local solution of the CMP shown in (1). For direct-DC formulation that provides a lower bound on the CMP, we identify conditions under which we can solve the problem to global optimality. Finally, for the special case of CMP in (2), such a recovery of local optimum is achieved by a stationary solution of the direct-DC approach. We summarize these relationships between solutions of different formulations in Figure 1.

3. *Computational validation.* We perform extensive computational experiments to compare the efficacy of our approximations. For this purpose, we utilize a CMP that arises in IMRT planning, our motivating example. In addition, we also use a portfolio optimization problem that minimizes the number of scenarios where the desired target is not achieved. In small-scale instances of the portfolio optimization problem, our approximations lead to solutions that have comparable performance to globally optimal solutions while reducing the computational times. Motivated by these results, we perform extensive experiments with the IMRT planning instances for which alternative approaches fail to provide any solutions.

In addition to the numerical results presented in this paper, this work also provides analytical corroboration for our case study on a CMP problem arising in power systems planning and operations [18]. In this case study, we use the direct-DC approach to minimize an objective function that includes counting the number of transmission lines operated outside the acceptable thermal limits.

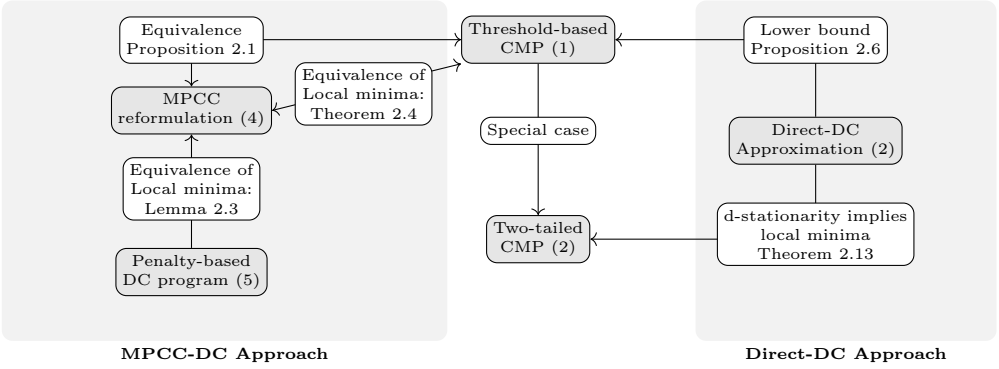


Figure 1: Schematic of the analysis in the paper showing different models and their relationships

1.3 Notation

We define $[m] = \{1, \ldots, m\}$. We use $\mathbf{x} = (x_i)_{i\in[n]}$ to denote a vector in $\mathbb{R}^n$ where x_i is the i-th component of $\mathbf{x}$. Let $\mathbf{1}_n \in \mathbb{R}^n$ denote a vector of all ones. The notation $\|\mathbf{x}\|_0$ denotes the cardinality of nonzero components of $\mathbf{x}$. We denote an open neighborhood of radius r and center $\mathbf{x}$ by $\mathcal{B}(\mathbf{x}, r)$.

1.4 Organization

We organize the rest of the paper as follows. In section §2, we present the two alternative continuous approximations for the CMP in (1). We also develop the relationship between the solutions of the approximate reformulations and the true problem and analyze the special case of τ-CMP in this section. In section §3, we present the results from the numerical experiments conducted on the fluence map optimization problem. We summarize our conclusions in §4.

2 Tractable Continuous Approximations

This section presents two alternative approaches to tackle the problem in (1). Both our approaches result in continuous approximations of the original problem that involve the DC program of the following form

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) = g(\mathbf{x}) - h(\mathbf{x}), \tag{3}$$

where $\mathcal{X}$ is a convex closed set, and g and h are both convex. The first approach involves the DC approximation of an MPCC. On the other hand, in the second approach, we apply a DC approximation directly to the CMP objective function.

2.1 MPCC-DC approach

In the first approach, we reformulate the ℓ_0-function in (1) in terms of complementarity constraints. This maneuver results in an MPCC. The MPCC can in turn be expressed as a smooth continuous nonlinear program. In this section, we extend the reformulation approach developed in [7] to problems of the form in (1).

To obtain the so-called *full-complementarity* reformulation of (1), we introduce two auxiliary vectors $\boldsymbol{\eta} \in \mathbb{R}^m_+$ and $\boldsymbol{\xi} \in [0,1]^m$ that are complementary to one another, that is, they satisfy the Hadamard constraint $\boldsymbol{\eta} \circ \boldsymbol{\xi} = 0$. The variable η_i takes a positive value only if $f_i(\mathbf{x}) - \tau_i > 0$ and the complementarity requirement enforces the corresponding ξ_i to zero. Consequently, a penalty

of λ is incurred in the objective function. The resulting reformulation is given as

$$\begin{aligned} \min_{\mathbf{x}\in\mathcal{X}} \ & f_0(\mathbf{x}) + \lambda \mathbf{1}_m^\top(\mathbf{1}_m - \boldsymbol{\xi}) \\ \text{subject to } & \eta_i \geq f_i(\mathbf{x}) - \tau_i \quad \forall i \in [m] \\ & 0 \leq \xi_i \leq 1,\ \eta_i \geq 0 \quad \forall i \in [m], \\ & \boldsymbol{\eta} \circ \boldsymbol{\xi} = 0. \end{aligned} \tag{4}$$

We define a concatenated decision vector as $\mathbf{z} := (\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta})$ and the corresponding feasible region of the above problem without the complementarity constraint by $\mathcal{Z} := \{(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta}) \mid \mathbf{x} \in \mathcal{X}, \eta_i \geq f_i(\mathbf{x}) - \tau_i,\ 0 \leq \xi_i \leq 1,\ \eta_i \geq 0,\ \forall i = 1, \ldots, m\}$.

Based on the construction of the MPCC, it is not difficult to see that the component vector $\mathbf{x}^*$ of the optimal solution vector $(\mathbf{x}^*, \boldsymbol{\eta}^*, \boldsymbol{\xi}^*)$ of (4) is an optimal solution to (1). We formally establish this observation regarding the equivalence of the CMP problem in (1) and the MPCC in (4) in the following proposition.

Proposition 2.1. *If $\mathbf{x}^*$ is an optimal solution of (1) then there exists $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ such that $\mathbf{z}^* = (\mathbf{x}^*, \boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ is the optimal solution of (4). Conversely, if $\mathbf{z}^*$ is an optimal solution of (4) then $\mathbf{x}^*$ is an optimal solution of (1).*

Proof. To show the first statement, consider two possible cases for any given i: $f_i(\mathbf{x}^*) > \tau_i$ and $f_i(\mathbf{x}^*) \leq \tau_i$. For the former case, we must have η_i^* strictly positive which yields $\xi_i^* = 0$ to meet the complementarity constraint of (4). For the latter, we choose $\eta_i^* = 0$ and $\xi_i^* = 1$ to achieve optimality. With described procedure, the two problems achieve the same objective value. For the remaining, it suffices to show its contrapositive: if $\mathbf{x}^*$ is not an optimal solution of (1) then $\mathbf{z}^*$ is not optimal for (4) for any $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$. Since $\mathbf{x}^*$ not optimal, there exists $\tilde{\mathbf{x}} \in \mathcal{X}$ such that $f_0(\mathbf{x}^*) + \lambda\|(\max\{f_i(\mathbf{x}^*) - \tau_i, 0\})_{i=1,\ldots,m}\|_0 > f_0(\tilde{\mathbf{x}}) + \lambda\|(\max\{f_i(\tilde{\mathbf{x}}) - \tau_i, 0\})_{i=1,\ldots,m}\|_0$. By applying the above argument, we can show that there does not exist $\mathbf{z}^*$ that is optimal for (4). □

To tackle (4), we take the approach of [9] and [10] where the problems with complementarity constraints are reformulated as DC programs. This reformulation enables the application of the DCA. Specifically, we replace the complementarity constraint $\boldsymbol{\eta} \circ \boldsymbol{\xi} = 0$ in (4) by a piecewise penalty term in the objective function. This penalty term is given by

$$\rho(\boldsymbol{\eta}, \boldsymbol{\xi}) := \sum_{i=1}^{m} \min\{\eta_i, \xi_i\}.$$

Using the above penalty term, the full-complementarity problem in (4) can be written in the following form:

$$\min_{\mathbf{z}\in\mathcal{Z}} \ \left\{ f_0(\mathbf{x}) + \lambda \mathbf{1}_m^\top(\mathbf{1}_m - \boldsymbol{\xi}) + \gamma\rho(\boldsymbol{\eta}, \boldsymbol{\xi}) \right\} \tag{5}$$

where, $\gamma > \lambda > 0$ is another penalty parameter. By defining

$$\begin{aligned} g(\mathbf{z}) =& f_0(\mathbf{x}) + \lambda \mathbf{1}_n^\top (\mathbf{1}_n - \boldsymbol{\xi}) \\ h(\mathbf{z}) =& \gamma \sum_{i=1}^{m} \max\{-\eta_i, -\xi_i\}, \end{aligned}$$

we obtain a DC decomposition of the penalized objective function in (5). The following proposition captures the nature of solutions of (5).

Proposition 2.2. *For any $\mathbf{z} = (\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathcal{Z}$, we have the index sets $\mathcal{I}_0, \mathcal{I}_1$, and $\mathcal{I}_2$ are such that $\mathcal{I}_0 \cup \mathcal{I}_1 \cup \mathcal{I}_2 = [m]$ and $\mathcal{I}_0 \cap \mathcal{I}_1 \cap \mathcal{I}_2 = \emptyset$. These index sets satisfy*

i. $\eta_i = 0$ *and* $\xi_i = 1$ $\forall i \in \mathcal{I}_0$;

ii. $\eta_i > 1$ *and* $\xi_i = 0$ $\forall i \in \mathcal{I}_1$;

iii. $0 < \eta_i \leq 1$ *and* $\xi_i \in \{0, 1\}$ $\forall i \in \mathcal{I}_2$.

Proof. For any $\mathbf{x} \in \mathcal{X}$, the second and third terms in the objective of (5) are separable in i. Therefore, we can assess each soft constraint independently and decompose the set of soft constraints into mutually exclusive index sets $\mathcal{I}_0, \mathcal{I}_1$, and $\mathcal{I}_2$.

i. When $\eta_i = 0$, we have $\min\{\eta_i, \xi_i\} = 0$. Since we have a minimization objective with $\lambda > 0$, we must have $\xi_i = 1$.

ii. When $\eta_i > 1$, the $\min\{\eta_i, \xi_i\}$ is attained by ξ_i. Then the ith component in the objective reduces to $\lambda + (\gamma - \lambda)\xi_i$. Since $\gamma > \lambda$, we must have $\xi_i = 0$.

iii. Let $0 < \eta_i \leq 1$. When the $\min\{\eta_i, \xi_i\}$ is attained by η_i, then due to the minimization objective we must have $\xi_i = 1$. On the other hand, when the $\min\{\eta_i, \xi_i\}$ is attained by ξ_i, following the previous argument $\xi_i = 0$.

□

Lemma 2.3. *If $\bar{\mathbf{z}} = (\bar{\mathbf{x}}, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{\eta}})$ is a local minimizer of (5) that satisfies $\bar{\eta}_i = 0$ or $\bar{\eta}_i > 1$ $\forall i \in [m]$, then it is a local minimizer of the MPCC (4).*

Proof. Since $\bar{\mathbf{z}}$ is a local minimizer, there exists $\mathcal{B}(\bar{\mathbf{z}}, r)$. Let $\mathbf{z} = (\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathcal{B}(\bar{z}, r)$ be a solution feasible to the MPCC. Using the fact that $\boldsymbol{\eta} \circ \boldsymbol{\xi} = 0$ we have

$$\begin{aligned} f_0(\mathbf{x}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \boldsymbol{\xi}) =& \; f_0(\mathbf{x}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \boldsymbol{\xi}) + \gamma \sum_{i=1}^{m} \min\{\eta_i, \xi_i\} \\ \geq& \; f_0(\bar{\mathbf{x}}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \bar{\boldsymbol{\xi}}) + \gamma \sum_{i=1}^{m} \min\{\bar{\eta}_i, \bar{\xi}_i\} \\ =& \; f_0(\bar{\mathbf{x}}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \bar{\boldsymbol{\xi}}). \end{aligned}$$

The first inequality follows from the local minimizing property of $\bar{\mathbf{z}}$ with respect to the DC program (5). The last equality follows Proposition 2.2. Finally, noting that $\bar{\mathbf{z}} \in \mathcal{Z}$ and $\min\{\bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\xi}}\} = 0$ implies $\bar{\boldsymbol{\eta}} \circ \bar{\boldsymbol{\xi}} = 0$, we have that $\bar{\mathbf{z}}$ is feasible to the MPCC. With this, we have completed the proof. □

The following result establishes a means to identify a local minimum of the CMP.

Theorem 2.4. *If $\bar{\mathbf{z}}$ is a local minimum of the MPCC in* (4)*, then $\bar{\mathbf{x}}$ is a local minimum of the CMP* (1)*.*

Proof. Since $\bar{\mathbf{z}}$ is a local minimum of (4), there exists a sufficiently small open neighborhood $\mathcal{B}(\bar{\mathbf{z}}, r)$ such that (i)

$$f_0(\mathbf{x}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \boldsymbol{\xi}) \geq f_0(\bar{\mathbf{x}}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \bar{\boldsymbol{\xi}})$$

and (ii) $\eta_i \neq 0$ whenever $\bar{\eta}_i \neq 0$, for all $\mathbf{z} \in \mathcal{B}(\bar{\mathbf{z}}, r)$. Define $\mathcal{S}(\boldsymbol{\eta}) = \{i \in [m] : \eta_i \neq 0\}$. Further, consider $\mathcal{A} = \{\mathbf{z} : f_0(\mathbf{x}) \geq f_0(\bar{\mathbf{x}}) - \epsilon,\ \forall\ 0 < \epsilon < \bar{\epsilon}\}$ for some $\bar{\epsilon} < \lambda$. The following two cases arise.

- Case 1: Consider $\mathbf{z} \in \mathcal{A} \cap \mathcal{B}(\bar{\mathbf{z}}, r) \cap \mathcal{Z}$ such that $\mathcal{S}(\boldsymbol{\eta}) = \mathcal{S}(\bar{\boldsymbol{\eta}})$. Feasibility of $\bar{\mathbf{z}}$ to MPCC (4) implies $\bar{\boldsymbol{\eta}} \circ \bar{\boldsymbol{\xi}} = 0$. Since $\mathcal{S}(\boldsymbol{\eta}) = \mathcal{S}(\bar{\boldsymbol{\eta}})$, we have $\boldsymbol{\eta} \circ \bar{\boldsymbol{\xi}} = 0$. Therefore, $(\mathbf{x}, \boldsymbol{\eta}, \bar{\boldsymbol{\xi}}) \in \mathcal{A} \cap \mathcal{B}(\bar{\mathbf{z}}, r) \cap \mathcal{Z}$ and feasible to MPCC. From the local minimum property of $\bar{\mathbf{z}}$ to MPCC, we have

  $$f_0(\mathbf{x}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \bar{\boldsymbol{\xi}}) \geq f_0(\bar{\mathbf{x}}) + \lambda \mathbf{1}_m^\top (\mathbf{1}_m - \bar{\boldsymbol{\xi}}).$$

  This implies

  $$f_0(\mathbf{x}) + \lambda \|\boldsymbol{\eta}\|_0 \geq f_0(\bar{\mathbf{x}}) + \lambda \|\bar{\boldsymbol{\eta}}\|_0 \qquad \forall (\mathbf{x}, \boldsymbol{\eta}, \bar{\boldsymbol{\xi}}) \in \mathcal{A} \cap \mathcal{B}(\bar{\mathbf{z}}, r) \cap \mathcal{Z}.$$

- Case 2: Consider $\mathbf{z} \in \mathcal{A} \cap \mathcal{B}(\bar{\mathbf{z}}, r) \cap \mathcal{Z}$ such that $\mathcal{S}(\boldsymbol{\eta}) \supset \mathcal{S}(\bar{\boldsymbol{\eta}})$. For such a $\mathbf{z}$, we have $\|\boldsymbol{\eta}\|_0 \geq \|\bar{\boldsymbol{\eta}}\|_0 + 1$ and consequently

  $$f_0(\mathbf{x}) + \lambda \|\boldsymbol{\eta}\|_0 \geq f_0(\bar{\mathbf{x}}) - \epsilon + \lambda(\|\bar{\boldsymbol{\eta}}\|_0 + 1) \geq f_0(\bar{\mathbf{x}}) + \lambda \|\bar{\boldsymbol{\eta}}\|_0 + \lambda - \epsilon \geq f_0(\bar{\mathbf{x}}) + \lambda \|\bar{\boldsymbol{\eta}}\|_0.$$

  The last inequality follows from $(\lambda - \epsilon) > 0$.

□

By combining the results in Lemma 2.3 and Theorem 2.4, we have the following result.

Theorem 2.5. *If $\bar{\mathbf{z}} = (\bar{\mathbf{x}}, \bar{\boldsymbol{\xi}}, \bar{\boldsymbol{\eta}})$ is a local minimizer of* (5) *that satisfies $\bar{\eta}_i = 0$ or $\bar{\eta}_i > 1\ \forall i \in [m]$, then it is a local minimizer of the CMP* (1)*.*

2.2 Direct-DC approach

Before we discuss the direct-DC approximation, we present the concepts of stationarity that are relevant for our purpose in this section. Consider the problem (3) where g and h are not necessarily differentiable. We first introduce the definition of a critical point.

Definition 2.1. A vector $\mathbf{x}^*$ is a critical point of (3) if

$$0 \in \partial g(\mathbf{x}^*) - \partial h(\mathbf{x}^*) + \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$$

where $\mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$ is the normal cone of $\mathcal{X}$ at the point $\mathbf{x}^*$ i.e., $\mathcal{N}_{\mathcal{X}}(\mathbf{x}^*) = \{\mathbf{v} \in \mathbb{R}^n \,|\, \mathbf{v}^T(\mathbf{x} - \mathbf{x}^*) \leq 0 \ \forall \mathbf{x} \in \mathcal{X}\}$. The sets $\partial g(\mathbf{x}^*)$ and $\partial h(\mathbf{x}^*)$ are the subdifferentials of g and h at the point $\mathbf{x}^*$, respectively.

Another kind of stationary solutions that plays a central role in our analysis is a directional stationary solution. We provide the formal definition below.

Definition 2.2. A vector $\mathbf{x}^*$ is a d(-irectional) stationary solution of (3) if the directional derivative of $f(\mathbf{x})$ at $\mathbf{x}^*$ is nonnegative for all feasible directions; i.e.,

$$f'(\mathbf{x}^*; \mathbf{x} - \mathbf{x}^*) \geq 0 \text{ for all } \mathbf{x} \in \mathcal{X}.$$

From the definitions, we can verify the relationship between the two types of stationary solutions. For the problem (3), d-stationarity always implies criticality. To see this, let $\mathbf{x}^*$ be a d-stationary solution and observe

$$\begin{aligned} 0 &\leq g'(\mathbf{x}^*; \mathbf{x} - \mathbf{x}^*) - h'(\mathbf{x}^*; \mathbf{x} - \mathbf{x}^*) \\ &= \max_{\mathbf{u} \in \partial g(\mathbf{x}^*)} \mathbf{u}^T(\mathbf{x} - \mathbf{x}^*) - \max_{\mathbf{w} \in \partial h(\mathbf{x}^*)} \mathbf{w}^T(\mathbf{x} - \mathbf{x}^*) \\ &\leq \max_{\mathbf{u} \in \partial g(\mathbf{x}^*)} \mathbf{u}^T(\mathbf{x} - \mathbf{x}^*) - \bar{\mathbf{w}}^T(\mathbf{x} - \mathbf{x}^*) \text{ for any } \bar{\mathbf{w}} \in \partial h(\mathbf{x}^*), \quad \forall \mathbf{x} \in \mathcal{X}. \end{aligned}$$

This indicates that there exists $\bar{\mathbf{u}} \in \partial g(\mathbf{x}^*)$ such that $\bar{\mathbf{w}} - \bar{\mathbf{u}} \in \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$. Hence $\mathbf{x}^*$ is a critical point.

The converse holds under a condition. A critical point $\mathbf{x}^*$ of (3) is a d-stationary solution if $h(\mathbf{x})$ is differentiable at $\mathbf{x}^*$. By definition, there exist $\bar{\mathbf{u}} \in \partial g(\mathbf{x}^*)$ and $\bar{\mathbf{v}} \in \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$ such that $\mathbf{0} = \bar{\mathbf{u}} - \nabla h(\mathbf{x}^*) + \bar{\mathbf{v}}$. This means for any $\mathbf{x} \in \mathcal{X}$, we have

$$\begin{aligned} 0 &\leq [\bar{\mathbf{u}} - \nabla h(\mathbf{x}^*)]^T (\mathbf{x} - \mathbf{x}^*) \\ &\leq \max_{\mathbf{u} \in \partial g(\mathbf{x}^*)} \mathbf{u}^T(\mathbf{x} - \mathbf{x}^*) - \nabla h(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \end{aligned}$$

which shows that $\mathbf{x}^*$ is a d-stationary solution.

Next, we propose an alternative approach to attain a continuous approximation of the CMP that directly uses a DC representation of the CMP objective function. For each condition

$\|\max\{f_i(\mathbf{x}) - \tau_i, 0\}\|_0$ in the problem (1), we approximate the discrete term by $g_i(\mathbf{x}) - h_i(\mathbf{x})$ where g_i and h_i are defined as

$$g_i(\mathbf{x}) = \max\left\{\frac{1}{\varepsilon}\Big(f_i(\mathbf{x}) - \tau_i\Big), 0\right\},$$
$$h_i(\mathbf{x}) = \max\left\{\frac{1}{\varepsilon}\Big(f_i(\mathbf{x}) - \tau_i\Big) - 1, 0\right\}.$$

Here, $\varepsilon > 0$ is an approximation parameter. The proposed DC program is then defined as

$$\min_{\mathbf{x}\in\mathcal{X}} \; f_0(\mathbf{x}) + \lambda \sum_{i=1}^{m} [\, g_i(\mathbf{x}) - h_i(\mathbf{x}) \,] \triangleq F(\mathbf{x}). \tag{8}$$

Observe that the approximation returns the same output as the discrete function when $f_i(\mathbf{x}) \leq \tau$ or $f_i(\mathbf{x}) \geq \tau + \varepsilon$. In the remaining case, the approximation returns a value between 0 and 1. This property immediately provides the following result.

Proposition 2.6. *For any $\varepsilon > 0$, the approximate DC program in* (8) *provides a lower bound to* (1).

Let f_0 and f_i for all $i \in [m]$ be differentiable. Furthermore, let the following assumptions hold.

(A1) There exists a scalar $\sigma \geq 0$ such that

$$f_0(\mathbf{x}) - f_0(\mathbf{y}) - \nabla f_0(\mathbf{y})^T(\mathbf{x} - \mathbf{y}) \geq \frac{\sigma}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

(A2) For each f_i, there exists $L_i \geq 0$ such that

$$0 \leq f_i(\mathbf{x}) - f_i(\mathbf{y}) - \nabla f_i(\mathbf{y})^T(\mathbf{x} - \mathbf{y}) \leq \frac{L_i}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

When f_0 is a strongly convex function, the penalty parameter λ can be carefully selected so that corresponding d-stationary solution achieves global optimality. We first show Lemma 2.7, which will be used to establish the global optimality of a d-stationary solution in Proposition 2.8.

Lemma 2.7. *Let (A2) hold. For each h_i, we have*

$$h_i'(\mathbf{x}; \mathbf{x} - \mathbf{y}) - h_i'(\mathbf{y}; \mathbf{x} - \mathbf{y}) \leq \frac{L_i}{\varepsilon}\|\mathbf{x} - \mathbf{y}\|_2^2 \quad \forall \mathbf{x}, \mathbf{y}, \in \mathcal{X}. \tag{9}$$

Proof. Since each f_i has a Lipschitz continuous gradient,

$$\begin{aligned}
\frac{1}{\varepsilon}\nabla f_i(\mathbf{x})^T(\mathbf{x} - \mathbf{y}) &\leq \frac{1}{\varepsilon}\nabla f_i(\mathbf{y})^T(\mathbf{x} - \mathbf{y}) + \left| \frac{1}{\varepsilon}\nabla f_i(\mathbf{x})^T(\mathbf{x} - \mathbf{y}) - \frac{1}{\varepsilon}\nabla f_i(\mathbf{y})^T(\mathbf{x} - \mathbf{y}) \right| \\
&\leq \max\left\{ \frac{1}{\varepsilon}\nabla f_i(\mathbf{y})^T(\mathbf{x} - \mathbf{y}), 0 \right\} + \frac{1}{\varepsilon}\left| \Big(\nabla f_i(\mathbf{x}) - \nabla f_i(\mathbf{y})\Big)^T(\mathbf{x} - \mathbf{y}) \right| \\
&\leq \max\left\{ \frac{1}{\varepsilon}\nabla f_i(\mathbf{y})^T(\mathbf{x} - \mathbf{y}), 0 \right\} + \frac{L_i}{\varepsilon}\|\mathbf{x} - \mathbf{y}\|_2^2.
\end{aligned}$$

The right side of the inequality is nonnegative hence

$$\max\left\{\frac{1}{\varepsilon}\nabla f_i(\mathbf{x})^T(\mathbf{x}-\mathbf{y}),0\right\}-\max\left\{\frac{1}{\varepsilon}\nabla f_i(\mathbf{y})^T(\mathbf{x}-\mathbf{y}),0\right\}$$
$$= h_i'(\mathbf{x};\mathbf{x}-\mathbf{y})-h_i'(\mathbf{y};\mathbf{x}-\mathbf{y}) \le \frac{L_i}{\varepsilon}\|\mathbf{x}-\mathbf{y}\|_2^2.$$

□

With above lemma, we show the global optimality of d-stationary solutions under assumptions (A1) and (A2).

Proposition 2.8. *Let (A1), with $\sigma > 0$, and (A2) hold. If λ satisfies $\sigma\varepsilon \ge 2\lambda\sum_{i=1}^{m} L_i$, then any d-stationary solution $\mathbf{x}^*$ of (8) is also a global minimum.*

Proof. Recall that we denote the objective function of (8) as $F(\mathbf{x})$. Since $F'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*) \ge 0$ for all $\mathbf{x}\in\mathcal{X}$, we have

$$\begin{aligned}F(\mathbf{x})-F(\mathbf{x}^*) &\ge F(\mathbf{x})-F(\mathbf{x}^*)-F'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*)\\ &= f_0(\mathbf{x})-f_0(\mathbf{x}^*)-f_0'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*)+\\ &\quad \lambda\sum_{i=1}^{m}[g_i(\mathbf{x})-g_i(\mathbf{x}^*)-g_i'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*)]-\\ &\quad \lambda\sum_{i=1}^{m}[h_i(\mathbf{x})-h_i(\mathbf{x}^*)-h_i'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*)].\end{aligned}$$

Using assumption (A1) for f_0 and convexity of g_i, we have

$$\begin{aligned}F(\mathbf{x})-F(\mathbf{x}^*) &\ge \frac{\sigma}{2}\|\mathbf{x}-\mathbf{x}^*\|_2^2-\lambda\sum_{i=1}^{m}\{h_i(\mathbf{x})-h_i(\mathbf{x}^*)-h_i'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*)\} \qquad (10)\\ &\ge \frac{\sigma}{2}\|\mathbf{x}-\mathbf{x}^*\|_2^2-\lambda\sum_{i=1}^{m}\{h_i'(\mathbf{x};\mathbf{x}-\mathbf{x}^*)-h_i'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*)\}.\end{aligned}$$

The last inequality is due to convexity of h_i. By applying Lemma 2.7, we deduce

$$F(\mathbf{x})-F(\mathbf{x}^*) \ge \left(\frac{\sigma}{2}-\frac{\lambda}{\varepsilon}\sum_{i=1}^{m}L_i\right)\|\mathbf{x}-\mathbf{x}^*\|_2^2 \quad \forall \mathbf{x}\in\mathcal{X}.$$

□

A similar result is shown in [1][Proposition 3.1]. Therein, the concave component of the DC program is defined as $\max_{i\in\mathcal{I}} h_i(\mathbf{x})$ where each h_i is a differentiable convex function. While the problem in [1] is a generalization of (8), the result in Proposition shows that a d-stationary point $\mathbf{x}^*$ achieves global optimality over a restricted set. Although this result is well-known in the DC literature, we include it for completeness and to showcase the particular form it takes in the context of CMP.

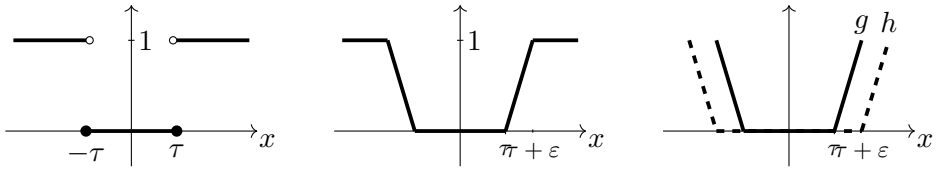


Figure 2: Let $x \in \mathbb{R}$. (Left) graph of $\|\max(|x| - \tau, 0)\|_0$; (Center) graph of the approximation function $g(x) - h(x)$ where $g(x) = \max\{\frac{1}{\varepsilon}(|x| - \tau), 0\}$ and $h(x) = \max\{\frac{1}{\varepsilon}(|x| - \tau) - 1, 0\}$; (Right) the DC components of the approximation: $g(x)$ is shown with solid line and $h(x)$ is shown with dashed line.

For a general DC program shown in (3), a d-stationary solution yields a local minimum if h is a piecewise affine function. This is due to the fact that, if h is piecewise affine, there exists a neighborhood $\mathcal{B}(\mathbf{x}^*, r)$ such that $h(\mathbf{x}) - h(\mathbf{x}^*) = h'(\mathbf{x}^*; \mathbf{x} - \mathbf{x}^*)$ for all $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, r)$ [6, Section 4.2]. Applying to (10) in the proof above, we obtain the following result.

Corollary 2.9. *Let f_i be a piecewise affine convex function for all $i \in [m]$. Under (A1), a d-stationary solution $\mathbf{x}^*$ of (8) is a local minimum.*

2.2.1 The special case of two-tailed cardinality minimization problem

The problem (2) is defined for a special case where $f_i(\mathbf{x}) = |x_i|$ for $i \in [n] = [m]$. When $\mathbf{x} \geq 0$, the term $\|(\max\{|x_i| - \tau_i, 0\})_{i=1,\ldots,m}\|_0$ reduces to the cardinality upper tail introduced in [14]. Several interesting applications motivate the study of this particular cardinality condition. One example is the security-constrained economic dispatch problem in power systems we studied in [18], and another is the IMRT planning problem which we formally introduce in Section 3.

Consider vectors $\boldsymbol{\ell}, \mathbf{u} \in \mathbb{R}^n$ such that $\ell_i \leq -\tau_i$ and $\tau_i \leq u_i$ for all $i \in [n]$. Define a box constraint $C \triangleq \{\mathbf{x} \,|\, \boldsymbol{\ell} \leq \mathbf{x} \leq \boldsymbol{u}\}$. For the current section, we consider a special case of (2) where (i) $\mathcal{X}$ is a box constraint $\mathcal{X} = \mathcal{C}$, and (ii) the objective function f_0 is differentiable. Applying the DC approximation presented in (8) to the problem (2) yields the DC program

$$\min_{\mathbf{x} \in \mathcal{X}} \; f_0(\mathbf{x}) + \lambda \sum_{i=1}^{n} \left[\max\left\{ \frac{1}{\varepsilon} \left(|x_i| - \tau_i \right), 0 \right\} - \max\left\{ \frac{1}{\varepsilon} \left(|x_i| - \tau_i \right) - 1, 0 \right\} \right]. \tag{11}$$

Figure 2 illustrates one-dimensional two-tailed cardinality function $\|\max\{|x| - \tau, 0\}\|_0$ for a scalar x and the DC approximation function. When $\tau = 0$, the former is referred to as the ℓ_0-function, and the latter reduces to the capped-ℓ_1 penalty [13]. Motivated by a recent work which studies the capped-ℓ_1 function for group structured sparsity problems [12], we establish the recovery of local solutions of (2) through the stationary solutions of (11). Let us assume that

(A3) There exists $\kappa \geq 0$ such that $\sup_{\mathbf{x} \in \mathcal{X}} \|\nabla f_0(\mathbf{x})\|_\infty \leq \kappa$.

Proposition 2.10. *Under Assumption (A3), let $\varepsilon\kappa < \lambda$ hold. If $\mathbf{x}^*$ is a critical point of* (11) *then either $|x_i^*| \le \tau_i$ or $|x_i^*| \ge \tau + \varepsilon$ for all $i \in [n]$.*

Proof. Suppose there exists j such that $\tau_j < |x_j^*| < \tau_j + \varepsilon$. By definition of the critical point, we have

$$0 = [\nabla f_0(\mathbf{x}^*)]_j + \lambda \left(\frac{1}{\varepsilon} \text{sign}(x_j^*) \right) + v_j \text{ for some } \mathbf{v} \in \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$$

where $[\nabla f_0(\mathbf{x}^*)]_j$ is the j-th component of $\nabla f_0(\mathbf{x}^*)$. Since $\sum_{i=1}^{n} v_i(x_i - x_i^*) \le 0$ for all $x_i \in [\ell_i, u_i]$ we must have $v_i = 0$ for all i. The condition $\varepsilon\kappa < \lambda$ yields contradiction. □

The above result immediately indicates that, under certain conditions, a critical point of (11) obtains the same objective value for the approximation (11) and the CMP (2). Due to the relationship between a critical point and a d-stationary solution, Proposition 2.10 also applies to any d-stationary point of (11). Before proceeding, let us define a problem:

$$\begin{aligned} &\min_{\mathbf{x}} \ f_0(\mathbf{x}) \\ &\text{subject to } \widehat{\mathcal{X}}(\mathbf{x}^*) \triangleq \{\mathbf{x} \in \mathcal{X} \,|\, |x_i| \le \tau_i \text{ whenever } |x_i^*| \le \tau_i \ \forall i \in [n] \}. \end{aligned} \tag{12}$$

The above is a convex program where the constraint set is a subset of $\mathcal{X}$. The next two results connect the solutions of problems (11) and (2) through the optimal solution of (12).

Lemma 2.11. *Let Assumption (A3) hold. If $\mathbf{x}^*$ is a d-stationary solution of* (11) *computed with $\lambda > \varepsilon\kappa$, then $\mathbf{x}^*$ is the global minimizer of* (12).

Proof. Denote $g_i(x_i) = \max\{\frac{1}{\varepsilon}(|x_i| - \tau_i), 0\}$ and $h_i(x_i) = \max\{\frac{1}{\varepsilon}(|x_i| - \tau_i) - 1, 0\}$. By Proposition 2.10, we either have $|x_i^*| \le \tau_i$ or $|x_i^*| \ge \tau_i + \varepsilon$ for any i. When $|x_i^*| < \tau_i$ or $|x_i^*| > \tau + \varepsilon$, the derivatives of g_i and h_i are equal. The only points of interest are when $|x_i^*| = \tau_i$ and $|x_i^*| = \tau_i + \varepsilon$ where one of g_i and h_i is nondifferentiable. Therefore, by d-stationarity, $\mathbf{x}^*$ satisfies

$$\begin{aligned} 0 \le \nabla f_0(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) &+ \lambda \sum_{i:\, |x_i^*| = \tau_i} g_i'(x_i^*; x_i - x_i^*) \\ &+ \lambda \sum_{i:\, |x_i^*| = \tau_i + \varepsilon} \left[\frac{1}{\varepsilon} \text{sign}(x_i^*)(x_i - x_i^*) - h_i'(x_i^*; x_i - x_i^*) \right] \ \forall \mathbf{x} \in \mathcal{X} \end{aligned}$$

Using the fact $h_i'(x_i^*; x_i - x_i^*) = \max_{a \in \partial h_i(x_i^*)} a^T(x_i - x_i^*)$ and $\partial h_i(x_i^*) = \left[0, \frac{1}{\varepsilon}\right]$ at $x_i^* = \tau_i + \varepsilon$ and $\partial h_i(x_i^*) = \left[-\frac{1}{\varepsilon}, 0\right]$ at $x_i^* = -(\tau_i + \varepsilon)$, we verify that the last term in the above is nonpositive. Hence $\mathbf{x}^*$ satisfies,

$$\begin{aligned} 0 &\le \nabla f_0(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) + \lambda \sum_{i:\, |x_i^*| = \tau_i} g_i'(x_i^*; x_i - x_i^*) \\ &\le \nabla f_0(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) + \lambda \sum_{i:\, |x_i^*| = \tau_i} g_i(x_i) - g_i(x_i^*) \ \forall \mathbf{x} \in \mathcal{X}, \end{aligned}$$

where the last inequality follows from convexity of g_i. Noting that for any i such that $|x_i^*| = \tau_i$, we have $g(x_i) = g(x_i^*)\ \forall\ \mathbf{x} \in \widehat{\mathcal{X}}(\mathbf{x}^*)$ and combining with convexity of f_0, we deduce that $\mathbf{x}^*$ is the global minimizer of f_0 on the set $\widehat{\mathcal{X}}(\mathbf{x}^*)$. □

The next result interprets the local minimizer of (2) in terms of the problem (12). This result is motivated by [12][Proposition 2.6].

Lemma 2.12. $\mathbf{x}^*$ *is a local minimizer of* (2) *if and only if* $\mathbf{x}^*$ *is the global minimizer of* (12).

Proof. Consider a sufficiently small neighborhood of $\mathbf{x}^*$, $\mathcal{B}(\mathbf{x}^*, \bar{r})$, where

$$\left\|(\max\{|x_i^*| - \tau_i, 0\})_{i\in[n]}\right\|_0 = \left\|(\max\{|x_i| - \tau_i, 0\})_{i\in[n]}\right\|_0 \tag{13}$$

for all $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, \bar{r}) \cap \widehat{\mathcal{X}}(\mathbf{x}^*)$. On the other hand, by local minimality of $\mathbf{x}^*$, there exists $r \in (0, \bar{r}]$ such that

$$f_0(\mathbf{x}^*) + \lambda \left\|(\max\{|x_i^*| - \tau_i, 0\})_{i\in[n]}\right\|_0 \leq f_0(\mathbf{x}) + \lambda \left\|(\max\{|x_i^*| - \tau_i, 0\})_{i\in[n]}\right\|_0 \tag{14}$$

for all $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, r) \cap \mathcal{X}$. Since $\mathbf{x}^* \in \widehat{\mathcal{X}}(\mathbf{x}^*) \subseteq \mathcal{X}$, (14) applies to any $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, r) \cap \widehat{\mathcal{X}}(\mathbf{x}^*)$ which must be a nonempty set. Combining (13) and (14), we show that $\mathbf{x}^*$ is a local minimizer of (12), which is also a global minimizer due to the convexity of the problem.

To show the other direction, consider $\mathbf{x}^*$ which is a global minimizer of (12). For any $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, \bar{r}) \cap \widehat{\mathcal{X}}(\mathbf{x}^*)$, we have (14), achieving local minimality on the restricted set $\widehat{\mathcal{X}}(\mathbf{x}^*)$. If $\mathbf{x}^*$ satisfies $|x_i^*| > \tau_i$ for all i, then $\widehat{\mathcal{X}}(\mathbf{x}^*) = \mathcal{X}$ and we complete the proof. Otherwise, consider the case of $\mathcal{X} \setminus \widehat{\mathcal{X}}(\mathbf{x}^*)$; this corresponds to the case where there is at least one i such that $|x_i^*| \leq \tau_i$. There exists $r' \in (0, \bar{r})$ and $\mathcal{B}(\mathbf{x}^*, r')$ such that the following conditions hold: (i) for any $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, r')$, we have $f_0(\mathbf{x}^*) \leq f_0(\mathbf{x}) + \lambda$ due to continuity of f_0 ; and (ii)

$$\left\|(\max\{|x_i^*| - \tau_i, 0\})_{i\in[n]}\right\|_0 + 1 \leq \left\|(\max\{|x_i| - \tau_i, 0\})_{i\in[n]}\right\|_0; \tag{15}$$

for all $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, r') \cap (\mathcal{X} \setminus \widehat{\mathcal{X}}(\mathbf{x}^*))$. Hence for any $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*, r') \cap (\mathcal{X} \setminus \widehat{\mathcal{X}}(\mathbf{x}^*))$, we have

$$\begin{aligned} f_0(\mathbf{x}^*) + \lambda \left\|(\max\{|x_i^*| - \tau_i, 0\})_{i\in[n]}\right\|_0 &\leq f_0(\mathbf{x}) + \lambda \left(\left\|(\max\{|x_i^*| - \tau_i, 0\})_{i\in[n]}\right\|_0 + 1\right) \\ &\leq f_0(\mathbf{x}) + \lambda \left\|(\max\{|x_i| - \tau_i, 0\})_{i\in[n]}\right\|_0. \end{aligned}$$

This shows that $\mathbf{x}^*$ is the local minimizer of (2). □

Based on Lemma 2.11 and Lemma 2.12, we show that a d-stationary solution of the DC approximation problem (11) achieves local minimality of (2).

Theorem 2.13. *Let Assumption (A3) hold. If* $\mathbf{x}^*$ *is a d-stationary solution of* (11) *computed with* $\lambda > \varepsilon\kappa$*, then* $\mathbf{x}^*$ *is a local minimizer of* (2).

3 Computational Study

In this section, we report the results of the numerical experiments illustrating the benefits of continuous approximations. For this purpose, we use the IMRT planning problem presented in §1.1 as our motivating example. To showcase the performance of the continuous approximations, we considered two alternative formulations of the CMP. These include a mixed-integer programming (MIP) model where we use a binary variable to encode when we violate a constraint. This approach is similar to the one used to model sample average approximation of the chance constraint problems [15]. The second formulation is the extension of the MPCC in [7] that we introduced in the context of our MPCC-DC approach. Unfortunately, we could not solve the MIP and MPCC model instances of the IMRT planning problem due to their large-scale nature. As an alternative, we use a portfolio optimization problem to facilitate a comparison of different modeling approaches.

3.1 Problem description

We first provide the mathematical programming formulation of the two problems and then show the numerical results.

3.1.1 Portfolio optimization problem

We consider a problem of selecting a portfolio of n assets at a minimum cost. The decision vector $\mathbf{x}$ whose element, say x_j represents the percentage of total wealth we invest in asset $j \in [n]$. The cost of purchasing these assets is c_j, and our desired total investment return is r. We model the random return for assets as a random vector $\tilde{\mathbf{a}} \in \mathbb{R}^n$, i.e., we expect a return of $\tilde{a}_j$ if we invest the entire budget in asset $j \in [n]$. We consider m independent realizations (scenarios) of random asset-specific returns that we denote by $\{\mathbf{a}^i\}_{i=1}^m$. Therefore, in addition to minimizing the cost, we want to reduce the number of scenarios under which the desired return is not met. We model this problem as follows.

$$\min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} + \lambda \sum_{i\in[m]} \|r - (\mathbf{a}^i)^\top \mathbf{x}\|_0 \;\text{ subject to }\; \mathbf{1}^\top \mathbf{x} = 1, \mathbf{x} \geq 0. \tag{16}$$

Using the earlier notation, we have $f_0(\mathbf{x}) = \mathbf{c}^\top \mathbf{x}$, $f_i(\mathbf{x}) = -(\mathbf{a}^i)^\top \mathbf{x}$ and $\tau_i = r$ for all $i \in [m]$, and $\mathcal{X} = \{\mathbf{x} \mid \mathbf{1}^\top \mathbf{x} = 1, \mathbf{x} \geq 0\}$.

The portfolio selection problem, formulated as a chance-constrained problem, has previously been studied in [15]. There, the authors solve a sample average approximation of the chance-constrained problem. We use their data set that is available in SIPLIB. The data set comprises ten independently generated problem instances with different cost vectors $\mathbf{c}$ and the sample of random realizations of asset-specific returns. All instances model have $n = 20$ assets and have a maximum sample size of $m = 200$. In our experiments, we generate MIP, MPCC, and DC

instances with varying sample sizes ($m = 50, 100, 150$, and 200). The detailed formulations are presented in Appendix A.1.

3.1.2 Intensity-modulated radiation therapy for cancer treatment

At the beginning of radiation therapy, a computed tomography (CT) scan provides information regarding the current state of the cancer patient. The CT scan reveals the volume of the tumor and its position in the region of interest relative to other organs at risk (OARs). The information in a CT scan is represented in terms of three-dimensional volume elements called voxels (akin to pixels in two-dimensional images). We will use $\mathcal{V} := \mathcal{T} \cup \mathcal{H}$ to denote the set of all voxels which can be decomposed into voxels corresponding to the tumor denoted by $\mathcal{T}$ and healthy tissues denoted by $\mathcal{H}$. We assume the region of interest comprises m different organ types, some of which are designated OARs. Therefore, we can obtain a further decomposition of healthy tissues as $\mathcal{H} = \cup_{i=1}^{m} \mathcal{H}_i$, where $\mathcal{H}_i$ capture voxels for organ $i = 1, \ldots, m$. A treatment prescription includes a target amount of radiation for the tumor and upper limits on the amount of radiation considered acceptable for all OARs. We denote the former by μ and the latter by τ_i for $i = 1, 2, \ldots, m$ (measured in unit of radiation known as Grays (Gy)).

A multileaf collimator is used to precisely conform radiation delivery to the tumor structure. Radiation is delivered by aligning the collimator beam, comprising a set of beamlets, at different gantry angles. We denote by $\mathcal{B}$ the set of beamlets. We can control the gantry angle and the intensity of beamlets to obtain the desired dose pattern. Therefore, these constitute our primary decision variables. The amount of radiation delivered to a voxel $v \in \mathcal{V}$ is a nonlinear function of gantry angle and radiation amount. For a given gantry angle, say θ, we compute a dose deposition matrix $\mathbf{D}(\theta)$. The dose deposition matrix has a dimension of $|\mathcal{V}| \times |\mathcal{B}|$ with individual elements given by d_{vb}. We denote the intensity of a beamlet by $\mathbf{y} := (y_b)_{b \in \mathcal{B}}$. Several alternative optimization models that differ in the objective and constraints imposed are used in practice for radiation therapy planning (see [5]). For our purpose here, we consider a model that explicitly aims to achieve a precise dosage for tumor voxels while keeping the number of healthy voxels receiving more than the prescribed dosage to a minimum. We state such an optimization problem as follows:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \quad & \sum_{v \in \mathcal{T}} (x_v - \mu)^2 + \lambda \sum_{i=1}^{m} \left\| (\max\{|x_v| - \tau_i, 0\})_{v \in \mathcal{H}_i} \right\|_0 \qquad (17) \\ \text{subject to } & x_v = \sum_{b \in \mathcal{B}} d_{vb} y_b \qquad \forall v \in \mathcal{V}, \\ & \underline{y} \leq y_b \leq \bar{y} \qquad \forall b \in \mathcal{B}. \end{aligned}$$

The first set of constraints captures the amount of radiation delivered to a voxel $v \in \mathcal{V}$. The bounds on beamlet intensity are intended to capture the physical limitations of the collimator. Through the first term in the objective, the program aims to achieve precision toward tumor voxels. Mathematically, we state this term as minimizing the deviation in the amount of radiation delivered to the tumor from the prescribed dose μ. In the second term of the objective function,

the individual summands represent the count of voxels of an OAR i that exceeds the prescription τ_i.

For our numerical experiments, we use the head-and-neck cancer dataset provided in [16]. The dataset includes five instances of head-and-neck cancer, each with a CT scan and a dose deposition matrix. In addition to the tumor, the CT scan of the region of interest includes four OARs (spinal cord, brainstem, left and right parotids) and unspecified normal body tissue. The prescription tumor dose is 70 Gy, and the radiation threshold for the OARs are 45, 50, and 28 Gy, respectively. Each of the five datasets has varying numbers of healthy and tumor voxels, as detailed in Table 1. We experiment with the same set of hyperparameters for all the models.

Dataset	Healthy Voxels (%)	Tumor Voxels (%)	Beamlets
1	67386 (71.0)	27576 (29.0)	3910
2	67270 (67.8)	31930 (32.2)	3888
3	76160 (67.7)	36320 (32.3)	4128
4	53176 (70.4)	22372 (29.6)	3003
5	64713 (69.3)	28638 (30.7)	3256

Table 1: Attributes for each of the five head-and-neck cancer datasets located in [16]

3.2 Experiment setup

The experiments were conducted on a computer using a 3.2 GHz 8-Core Intel processor with 32GB of RAM, running Mac OS Big Sur version 11.6. To solve the DC approximations of CMP problem, we use the DC algorithm (DCA) [11, 17]. For completion, here we describe this algorithm using a general convex constrained DC program, $\min_{\mathbf{x}\in\mathcal{X}} g(\mathbf{x}) - h(\mathbf{x})$:

1. Initialize $\mathbf{x}^0 \in \mathcal{X}$. Set $k = 0$.
2. Iteratively update

$$\mathbf{x}^{k+1} \in \arg\min_{\mathbf{x}\in\mathcal{X}} g(\mathbf{x}) - h(\mathbf{x}^k) - (\mathbf{v}^k)^T(\mathbf{x} - \mathbf{x}^k)$$

   where $\mathbf{v}^k \in \partial h(\mathbf{x}^k)$.
3. Stop if prescribed termination criteria are satisfied.

We implemented the DC algorithm on MATLAB and use a randomly generated initial solution vector $\mathbf{x}^0$.

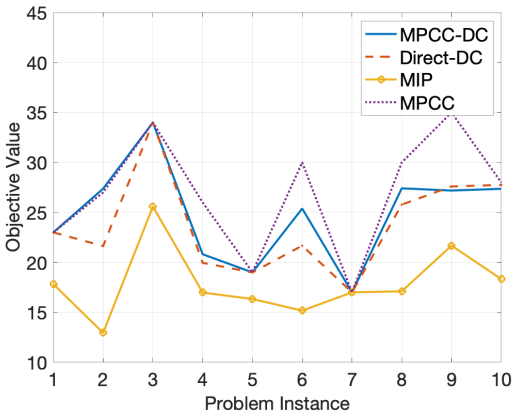
3.3 Numerical results

We present the numerical results for the two problems described earlier in this section.

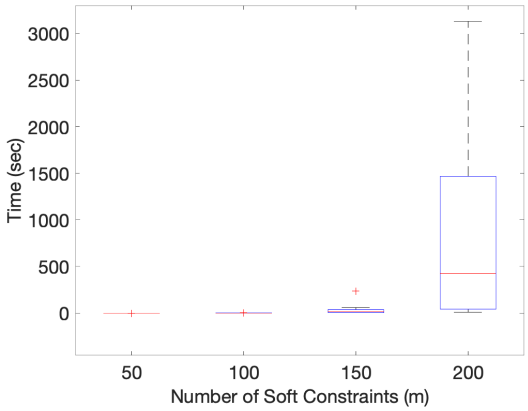
3.3.1 Comparison of alternative approaches

We first provide a comparison of the alternative formulations of the CMP problem. For this comparison, we solve all ten instances of the portfolio optimization problem for a different number of soft constraints. We report the objective function value obtained by each approach and the total computational time. The results are presented in Figure 3.

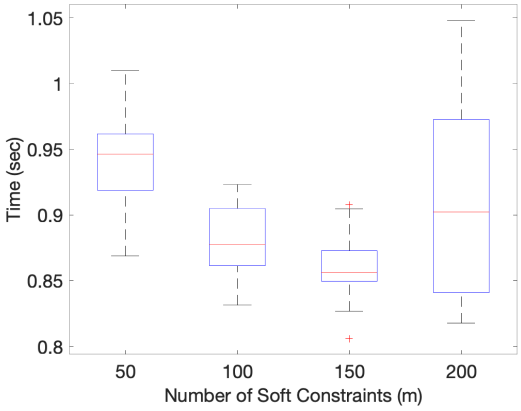
To compare the solution quality, we experimented using problem instances with $m = 50$ soft constraints. We could not solve the problem to optimality for all instances and all approaches (particularly MPCC) for larger m, making a comparison across all methods impossible. For each problem instance, we see in Figure 3a that MIP formulation provides a globally optimal solution. Therefore, the optimal value is the lowest among the four approaches. The two DC (MPCC-DC and direct-DC) approaches perform at least as well as the MPCC approach. Furthermore, MPCC and DC approaches could identify a globally optimal solution only for Instance 7, otherwise identifying just local solutions for other instances. Finally, the performance of the two DC approaches is comparable with the direct-DC approach marginally outperforming



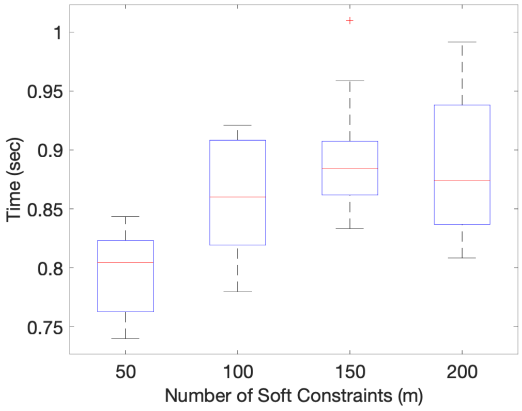
(a) Objective Value for $m = 50$



(b) MIP



(c) MPCC-DC



(d) Direct-DC

Figure 3: Objective function value and computational time comparison of alternative methods on portfolio optimization problem

the MPCC-DC.

Figures 3b, 3c, and 3d show the computational time for the MIP, MPCC-DC, and direct-DC approaches, respectively. In the MIP formulation, the number of binary variables and functional constraints increases with the number of soft constraints. This increase in problem size results in a significant increase in the computational time, as seen in the figure. While the computational time for DC approaches for $m = 50$ and $m = 100$ are higher (on average) than MIP, this time does not increase as much when we raise the number of soft constraints to $m = 150$ or $m = 200$. For instance, in the worst case, the computational time for MIP was 3130 seconds, and the same for DC approaches is around one second. The two DC approaches are comparable even in terms of computational time (see Figures 3c and 3d).

3.3.2 Performance of Direct-DC approach for IMRT planning

In our previous experiment, we see that the direct-DC approach marginally outperforms the MPCC-DC approach in terms of solution quality while providing comparable performance in terms of computational time. Therefore, we employ the direct-DC approach for our experiments with the large-scale problem of IMRT planning. We investigate the performance of the direct-DC approach relative to the quadratic programming-based benchmark formulations prevalent in radiation therapy planning. The detailed formulation of these benchmarks and the direct-DC approximation are provided in Appendix A.2.

We evaluate the performance in terms of the dosage received by tumorous and healthy voxels. Two primary questions guide this investigation.

- How many healthy voxels receive an amount of radiation exceeding their prescribed threshold, and what are the magnitudes of these violations?
- How many tumorous voxels do not receive as much radiation as prescribed, and how closely do these dosage patterns align with the prescribed amount?

Since the primary goal of IMRT is to keep healthy voxels free of exposure while administering sufficient radiation to tumorous voxels, we gain a complete view of each model's efficacy by investigating these two questions.

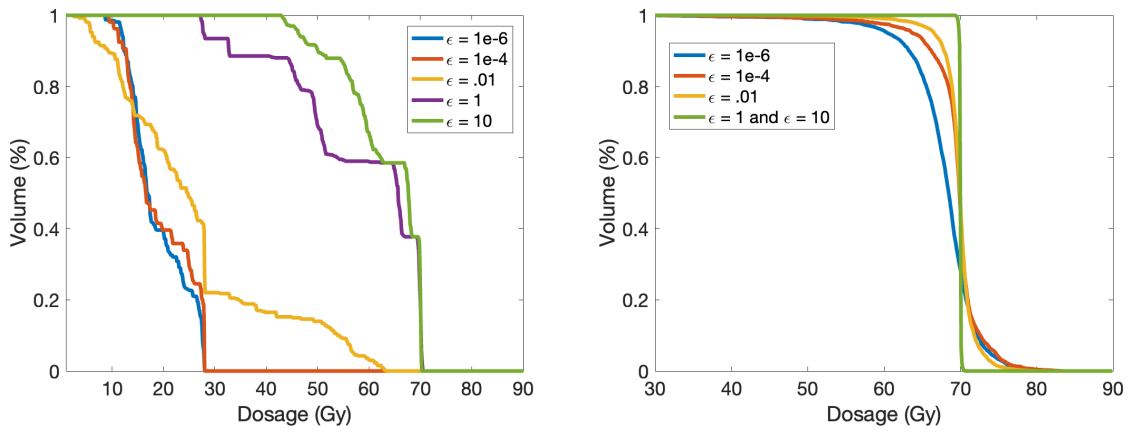
The only input parameter for direct-DC approach is the value of ε, so we begin by studying the effect of ε on the radiation pattern. We chose $\varepsilon = \{10^{-6}, 10^{-4}, 0.01, 1, 10\}$ for this experiment. For each value, we use the decision vector obtained when the DCA is terminated to compute the dosage received by the OAR and voxels. We present these results in the dose-volume histogram in Figure 4. A dose-volume histogram depicts the percentage of voxels that receive at least as much radiation as shown in the horizontal axis. In Figure 4a, we clearly see that with smaller ε values we ensure no healthy voxels in the left parotid organ receive radiation beyond its prescribed threshold of 28 Gy. Conversely, however, we see in the rightward plot that the solutions with the same ε values deviate more from the prescribed threshold for tumorous voxels.

This trade-off between healthy and tumor voxel radiation amounts persists for larger ε values as well. When $\varepsilon \geq 1$, the program ensures tumorous voxels receive radiation dosages very close to the prescribed threshold of 70, yet many healthy voxels receive excess radiation in the process. A similar observation was made in other OARs and datasets. For the remainder of experiments with the direct-DC approach, we use $\varepsilon = 1$ as it demonstrated a reasonable compromise between radiation precision for tumor and overdosing the healthy tissues.

We note that the MIP and MPCC methods failed to converge to any solutions within the user-defined time limit of one hour. Therefore, we use a quadratic-penalty (QP) and a one-sided quadratic-penalty (1-QP) models to serve as viable benchmarks for comparison. These models are presented in Appendix A. Based on hyperparameter selection procedures, we identified $\alpha = 1 - 10^{-6}$ for QP, $\bar{\lambda} = 0.01$ for the 1-QP, and $\varepsilon = 1$ for DC model instances.

First, we investigate the number of healthy voxel violations and the degree of these violations for all three methods. We summarize these results in Figure 5. While the percent total of healthy voxels that receive excess radiation is less than 3.5% for all datasets for the DC method, it reaches as high as 16.8% for the QP. Furthermore, the DC method has equal or lesser percentage of healthy voxels receiving excess radiation than both benchmark methods in all 5 datasets. In addition to the number of healthy violations, we can also view the severity of each violation.

For each method, the line plot in Figure 5 shows the average amount of violations as a percentage of their respective prescribed threshold (computed across all the OAR). The figure shows that the average magnitude of violations are greater for the DC when compared to 1-QP for all five datasets. This behavior is expected as the penalty term for the DC method penalizes the number while the 1-QP penalize the degree of violations. The QP model, on the other hand, equally penalizes deviation on both sides of the prescribed dosage. Therefore, we do not notice any specific pattern in terms of overdosing the healthy voxels. Our results show that in many cases where healthy voxels receive excess radiation, the radiation amount is $0\% - 5\%$ above the prescribed threshold. For some healthy voxels, we observe radiation levels as high as 150%



(a) OAR: Left parotid gland ($\tau_i = 28$ Gy) (b) Tumor ($\mu = 70$ Gy)

Figure 4: Dosage volume histograms for varying degrees of ε in the DC method for Dataset 1.

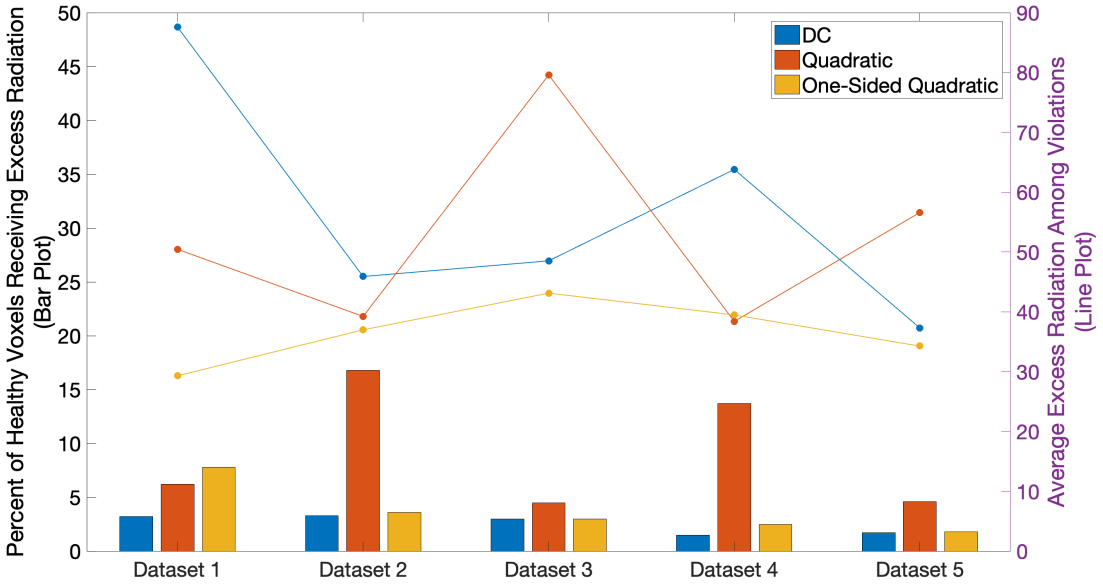


Figure 5: Percent of total healthy voxels receiving excess radiation (bar plot) and corresponding average violation magnitude as a percentage above the prescribed threshold (line plot).

greater than the prescribed threshold. However, excessive overdosing is far less ubiquitous in the DC and 1-QP methods than the the QP method. Therefore, we see that not only does the DC method consistently lead to fewer healthy voxels over-radiation, but it drastically reduces occurrences of severe burning as well. IMRT not only aims to protect healthy tissue, but it also must apply the correct amount of radiation to tumor cells. We also investigate the performance of our models on radiation to tumor voxels to answer our second question. Since all three models use the quadratic penalty to enforce precision for tumor tissue, we notice that about 50% of the tumor voxels are underdosed. Furthermore, the amount by which these tumor voxels are underdosed is also comparable across the methods.

Overall, a clear trade-off exists between adherence to dose prescriptions for healthy tumorous voxels as hyperparameter values change for the DC method. The DC method results in fewer healthy voxels that receive excess radiation beyond the prescribed threshold. While the DC violations are typically more severe on average, the number of severely overdosed voxels are fewer when compared to the QP method. Furthermore, all three methods perform similarly in regards to radiation precision to tumor voxels. Noting too that the MIP formulation failed to provide a solution to any dataset within the set timeframe, our experiments suggest that the DC method is an effective approach for IMRT planning problems involving cardinality minimization.

4 Conclusions

In this paper, we study an optimization problem which emphasizes on minimizing the cardinality of unsatisfied constraints. The problem is motivated by an application which aims to

deliver as many soft constraints in addition to minimizing the objective function subject to hard constraints. To provide computationally viable solution methods, we introduce continuous reformulations that approximate the discrete cardinality. Our analysis show that local solutions of the cardinality minimization problem are obtainable by computing local and stationary solutions of the proposed reformulations. We study effectiveness of the new methods by comparing their computational performance to alternative formulations. The numerical results indicate that the proposed methods produce comparable solutions while reducing the computation time. Our study on the IMRT dataset demonstrates that our methods are capable of computing solutions with minimal number of unsatisfied soft conditions as desired.

A Benchmark and Approximate Formulations

In this section, we present all the formulations of the two problems used in our numerical experiments.

A.1 Portfolio optimization problem

We use the complementarity reformulation-based approach for the portfolio optimization problem in §3.1.1. To benchmark the developed continuous approximation, we use the MPCC problem formulation which is stated as follows

$$\begin{aligned} \min_{\mathbf{x},\boldsymbol{\eta},\boldsymbol{\xi}} \ & \mathbf{c}^\top \mathbf{x} + \lambda \mathbf{1}^\top (\mathbf{1} - \boldsymbol{\xi}) \\ \text{subject to } & \eta_i \geq r - (\mathbf{a}^i)^\top \mathbf{x} \qquad \forall i \in [m] \\ & \mathbf{1}^\top \mathbf{x} = 1, \boldsymbol{\eta} \circ \boldsymbol{\xi} = 0 \\ & \mathbf{x} \geq 0,\ 0 \leq \xi_i \leq 1,\ \eta_i \geq 0 \qquad \forall i \in [m]. \end{aligned} \tag{18}$$

Since f_i are affine functions of $\mathbf{x}$ for all i, the above is a linear program with complementarity constraints for which efficient off-the-shelf solvers are available.

Notice that the CMP problem in (16) admits an equivalent MIP. This is stated as

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z}} \ & \mathbf{c}^\top \mathbf{x} + \lambda \sum_{i \in [m]} z_i \\ \text{subject to } & (\mathbf{a}^i)^\top \mathbf{x} \geq (1 - z_i) r \qquad \forall i \in [m] \\ & \mathbf{1}^\top \mathbf{x} = 1, \mathbf{x} \geq 0 \\ & z_i \in \{0, 1\} \quad i \in [m]. \end{aligned} \tag{19}$$

A.2 Intensity-modulated radiation therapy planning problem

For the IMRT problem in (17), we introduce two models to serve as a benchmark for our approach. The first model is prevalent in radiation therapy planning literature that focuses only on radiation precision. In this model, we penalize deviation from the prescribed dosage for healthy organs and tumors. This model formulation is stated as:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \ & \alpha \sum_{v\in\mathcal{T}} (x_v - \mu)^2 + (1-\alpha) \sum_{v\in\mathcal{H}} (x_v - \tau)^2 \\ \text{subject to } & x_v = \sum_{b\in\mathcal{B}} d_{vb} y_b \qquad \forall v \in \mathcal{V} \\ & \underline{y} \le y_b \le \bar{y} \qquad \forall b \in \mathcal{B}. \end{aligned} \tag{20}$$

The hyperparameter $\alpha \in [0,1]$ used in the objective function allows us to control the emphasis between the tumor and healthy organs. We refer (20) as the "quadratic penalty" (QP) model.

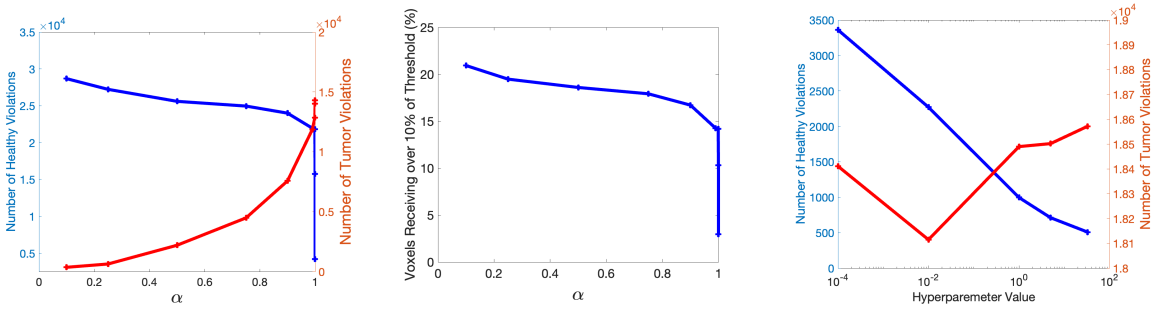
The CMP model for IMRT planning in (17) minimizes the number of healthy voxels receiving excess dosage. Alternatively, we can minimize the extent of over-dosage (beyond the prescribed dosage) for healthy voxels. The following model formulation, the "One-sided Quadratic Penalty" (1-QP), utilizes such an objective. We state it as:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y},\mathbf{z}} & \sum_{v\in\mathcal{T}} (x_v - \mu)^2 + \bar{\lambda} \sum_{v\in\mathcal{H}} z_v^2 \\ \text{subject to: } & x_v = \sum_{b\in\mathcal{B}} d_{vb} y_b \qquad \forall v \in \mathcal{V} \\ & z_v \ge x_v - \tau \qquad \forall v \in \mathcal{H} \\ & z_v \ge 0 \qquad \forall v \in \mathcal{H} \\ & \underline{y} \le y_b \le \bar{y} \qquad \forall b \in \mathcal{B}. \end{aligned} \tag{21}$$

Like the CMP model (17), and unlike the QP model (20), the above program does not penalize any deviation from prescriptions for healthy voxels if the amount of radiation received by such voxels is less than the corresponding prescribed threshold. Note that the parameter $\bar{\lambda}$ plays a similar role as λ in the CMP.

For completion, we next present the approximate model obtained by employing the direct-DC approach presented in §2.2.

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \ & \sum_{v\in\mathcal{T}} (x_v - \mu)^2 + \\ & \qquad \lambda \sum_{v\in\mathcal{H}} \left[\max\left\{ \frac{1}{\varepsilon}\Big(x_v - \tau\Big), 0 \right\} - \max\left\{ \frac{1}{\varepsilon}\Big(x_v - \tau\Big) - 1, 0 \right\} \right] \\ \text{subject to } & x_v = \sum_{b\in\mathcal{B}} d_{vb} y_b \qquad \forall v \in \mathcal{V} \\ & \underline{y} \le y_b \le \overline{y} \qquad \forall b \in \mathcal{B}. \end{aligned} \tag{22}$$



(a) Number of violations, QP

(b) Percentage of all healthy voxels receiving radiation amounts over 10% greater than the prescribed threshold, QP.

(c) Number of violations, 1-QP

Figure 6: Number of violations for healthy voxels and for tumorous voxels as hyperparameter value changes. QP method, Dataset 1

As with the portfolio optimization problem, the IMRT problem (17) also admits a mixed-binary program as an equivalent formulation. Unfortunately, we could not obtain solutions from this formulation for any dataset using the default integer programming solvers on Gurobi. Therefore, we did not consider the mixed-binary programs in our numerical experiments.

A.3 Benchmark methods and hyperparameters

In this section, we discuss hyperparameter selection for the benchmark methods detailed in section A.2. First, the QP method experiences a clear trade-off between the adherence to healthy voxels' prescriptions and tumorous voxels' prescriptions as hyperparameter values change (see Figure 6a). We see more healthy voxel violations with hyperparameter values near 1 even though these hyperparameter values indicate more emphasis on the objective term relating to healthy voxels. This counterintuitive nature is attributed to the relative increase in average violation magnitude; for hyperparameter values near 1, each violation magnitude is relatively small. Conversely, for hyperparemeter values near 0, the quadratic nature of the penalty decays and instead "sacrifices" some voxels so that large quantities of other voxels may meet requirements. We also note that the QP method tends to result in large quantities of healthy voxels receiving radiation levels over 10% greater than the prescribed threshold (see Figure 6b). Any intensity-modulated radiation therapy treatment with this quality is unacceptable. Therefore, in §3.3 we use $\alpha = 1 - (1e - 6)$ as this value typically results in approximately $3\% - 5\%$ of total healthy voxels receiving over 10% of the prescribed threshold. While this $3\% - 5\%$ value is chosen arbitrarily, it nevertheless provides a more practical solution to use in a real intensity-modulated radiation therapy setting.

Next, the 1-QP method also experiences a trade-off as hyperparameter values change. However, this method results in less extreme burning of healthy voxels as compared to the QP method. Also, the trade-off is less counterinuitive than in the QP method. Here, as less weight is placed on the terms relating to healthy voxels, fewer tumorous voxels receive insufficient ra-

diation and more healthy voxels receive excess radiation. Figure 6c details how the adherence to these two objectives change as hyperparameter values change. We note that $\bar{\lambda} = .01$ results in similar radiation dosage patterns for tumorous voxels to the DC method results. Therefore, we choose this hyperparameter value for comparisons enumerated in §3.3.

References

[1] M. Ahn, J.S. Pang, and J. Xin. Difference-of-convex learning: directional stationarity, optimality, and sparsity. *SIAM Journal on Optimization*, 27(3):1637–1665, January 2017.

[2] A. Al-Shabili, Y. Feng, and I. Selesnick. Sharpening sparse regularizers via smoothing. *IEEE Open Journal of Signal Processing*, 2:396–409, 2021.

[3] K. Bui, F. Park, S. Zhang, Y. Qi, and J. Xin. Structured sparsity of convolutional neural networks via nonconvex sparse group regularization. *Frontiers in Applied Mathematics and Statistics*, 6, February 2021.

[4] E. Candés, M. Wakin, and S. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, 2008.

[5] M. Ehrgott, C. Guler, H. Hamacher, and L. Shao. Mathematical optimization in intensity modulated radiation therapy. *Annals of Operations Research*, 175(1):309–365, 2010.

[6] F. Facchinei and J.S. Pang, editors. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer New York, 2004.

[7] M. Feng, J.E. Mitchell, J.S. Pang, A. Waechter, and X. Shen. Complementarity formulations of ℓ_0-norm optimization problems. *Pacific Journal of Optimization*, 14(2):273–305, 2018.

[8] D. Jaffray. Image-guided radiotherapy: from current concept to future perspectives. *Nature Reviews Clinical Oncology*, 9(12):688–699, 2012.

[9] F. Jara-Moroni, J.S. Pang, and A. Wächter. A study of the difference-of-convex approach for solving linear programs with complementarity constraints. *Mathematical Programming*, 169(1):221–254, 2018.

[10] H.A. Le Thi and T. Pham Dinh. On solving linear complementarity problems by DC programming and DCA. *Computational Optimization and Applications*, 50(3):507–524, 2011.

[11] H.A. Le Thi, T. Pham Dinh, H.M. Le, and X.T. Vo. DC approximation approaches for sparse optimization. *European Journal of Operational Research*, 244(1):26–46, 2015.

[12] W. Li, W. Bian, and K. C. Toh. DC algorithms for a class of sparse group ℓ_0 regularized optimization problems. *arXiv preprint arXiv:2109.05251*, 2021.

[13] M. Nikolova. Local strong homogeneity of a regularized estimator. *SIAM Journal on Applied Mathematics*, 61(2):633–658, 2000.

[14] M. Norton, A. Mafusalov, and S. Uryasev. Cardinality of upper average and its application to network optimization. *SIAM Journal on Optimization*, 28(2):1726–1750, January 2018.

[15] B. Pagnoncelli, S. Ahmed, and A. Shapiro. Sample average approximation method for chance constrained programming: theory and applications. *Journal of Optimization Theory and Applications*, 142(2):399–416, 2009.

[16] F. Saberian, A. Ghate, and M. Kim. Spatiotemporally optimal fractionation in radiotherapy. *INFORMS Journal on Computing*, 29(3):422–437, 2017.

[17] B. Sriperumbudur and G. Lanckriet. A proof of convergence of the concave-convex procedure using Zangwill's theory. *Neural Computation*, 24(6):1391–1407, 06 2012.

[18] D. Troxell, M. Ahn, and H. Gangammanavar. A cardinality minimization approach for security-constrained economic dispatch. *IEEE Transactions on Power Systems*, 37(5):3642–3652, 2021.

[19] C. Wang, M. Tao, J. Nagy, and Y. Lou. Limited-angle CT reconstruction via the L_1/L_2 minimization. *SIAM Journal on Imaging Sciences*, 14(2):749–777, 2021.