Dynamic string-averaging CQ-methods for the split feasibility problem with percentage violation constraints arising in radiation therapy treatment planning

Mark Brooke[1,*], Yair Censor[2] and Aviv Gibali[3]

[1]Department of Oncology, University of Oxford,
Oxford OX3 7DQ, UK
{mark.brooke@oncology.ox.ac.uk}
[2]Department of Mathematics, University of Haifa
Mt. Carmel, Haifa 3498838, Israel
{yair@math.haifa.ac.il}
[3]Department of Mathematics, ORT Braude College
Karmiel 2161002, Israel
{avivg@braude.ac.il}

January 14, 2020. Revised: December 18, 2020.

Abstract

In this paper we study a feasibility-seeking problem with percentage violation constraints. These are additional constraints, that are appended to an existing family of constraints, which single out certain subsets of the existing constraints and declare that up to a specified fraction of the number of constraints in each subset is allowed to be violated by up to a specified percentage of the existing bounds. Our motivation to investigate problems with percentage violation constraints comes from the field of radiation therapy treatment planning wherein the fully-discretized inverse planning problem is formulated as a split feasibility problem and the percentage violation constraints give rise to non-convex constraints. Following the CQ algorithm of Byrne (2002, *Inverse Problems*, Vol. **18**, pp. 441–53), we develop a string-averaging CQ method that uses only projections onto the individual sets which are half-spaces represented by linear inequalities. The question of extending our theoretical results to the non-convex sets case is still open. We describe how our results apply to radiation therapy treatment planning and provide a numerical example.

Keywords: String-averaging, CQ-algorithm, split feasibility, percentage violation constraints, radiation therapy treatment planning, dose-volume constraints, common fixed points, cutter operator.

[*]Corresponding author: mark.brooke@oncology.ox.ac.uk, cc: markdanielbrooke@gmail.com

1 Introduction

1.1 Motivation

In this work we are motivated by a linear split feasibility problem with percentage violation constraints arising in radiation therapy treatment planning. We first provide the background in general terms.

Inverse radiation therapy treatment planning (RTTP). This problem, in its fully-discretized modeling approach, leads to a linear feasibility problem. This is a system of linear interval inequalities

$$c \leq Ax \leq b, \tag{1}$$

wherein the "dose matrix" A is pre-calculated by techniques called in RTTP "forward calculation" or "forward planning" and the vector x is the unknown vector of "intensities" that, when used in setting up the treatment machine, will realize this specific "treatment plan". The vectors b and c contain upper and lower bounds on the total dose Ax permitted and required in volume elements (voxels) of sensitive organs/tissues and target areas, respectively, inside the irradiated body. The components of b and c are prescribed by the attending oncologist and given to the treatment planner.

Percentage violation constraints (PVCs). In general terms, these are additional constraints that are appended to an existing family of constraints. They single out certain subsets of the existing constraints and declare that up to a specified fraction of the number of constraints in each subset is allowed to be violated by up to a specified percentage of the existing bounds. Such PVCs are useful in the inverse problem of RTTP, mentioned above, where they are called "dose volume constraints" (DVCs). When the system of linear interval inequalities is inconsistent, that is, there is no solution vector that satisfies all inequalities, the DVCs allow the oncologist and the planner to relax the original constraints in a controlled manner to achieve consistency and find a solution.

Split feasibility. PVCs are, by their very nature, integer constraints which change the feasibility problem to which they are attached from being a continuous feasibility problem into becoming a mixed integer feasibility problem. An alternative to the latter is to translate the PVCs into constraints sets that are appended to the original system of linear interval inequalities but are formulated on the vectors Ax, rather than directly on x. This gives rise to a "split feasibility problem" which is split between two spaces: the space of "intensity vectors" x and the space of "dose vectors" $d := Ax$ in which A is the operator mapping one space onto the other.

Non-convexity. The constraints sets, that arise from the PVCs, in the space of "dose vectors" are non-convex sets but, due to their special form enable the calculation of orthogonal projections of points onto them. This opens the door for applying our proposed dynamic string-averaging CQ-method to the RTTP inverse problem with PVCs. Mathematical analysis for the case of non-convex sets remains an open question. Looking at it from the practical point of view one may consider also alternatives such as reformulating PVCs as ℓ_1-norm constraints. See, for example, Candès et al. (2008); Kim et al. (2013).

Group-structure of constraints. Each row in the system (1) represents a constraint on a single voxel. Lumping together constraints of voxels, according to the organ/tissue to which they belong, divides the matrix A and the whole system into "groups" of constraints, referred to below as "blocks of constraints" in a natural manner. These groups affect the formulation of the split feasibility problem at hand by demanding that the space of intensity vectors x be mapped separately by each group of rows of the matrix A into another space of dose vectors d.

1.2 Contribution

Motivated by the above we deal in this paper with the "multiple-operator split common fixed point problem" (MOSCFPP) defined next.

Problem 1 ***The multiple-operator split common fixed point problem (MOSCFPP)***. *Let $\mathcal{H}$ and $\mathcal{K}$ be two real Hilbert spaces, and let r and p be two natural numbers. Let $U_i : \mathcal{H} \to \mathcal{H}$, $1 \leq i \leq p$, and $T_j : \mathcal{K} \to \mathcal{K}$, $1 \leq j \leq r$, be given operators with nonempty fixed point sets $\operatorname{Fix}(U_i)$ and $\operatorname{Fix}(T_j)$, respectively. Further, let $A_j : \mathcal{H} \to \mathcal{K}$, for all $1 \leq j \leq r$, be given bounded linear operators. In addition let Φ be another closed and convex subset of $\mathcal{H}$. The MOSCFPP is:*

$$\text{Find an } x^* \in \Phi \text{ such that } x^* \in \cap_{i=1}^{p} \operatorname{Fix}(U_i) \text{ and,} \tag{2}$$

$$\text{for all } 1 \leq j \leq r,\ A_j x^* \in \operatorname{Fix}(T_j). \tag{3}$$

This problem formulation unifies several existing "split problems" formulations and, to the best of our knowledge, has not been formulated before. We analyze it and propose a "dynamic string-averaging CQ-method" to solve it, based on techniques used in some of those earlier formulations. We show in detail how this problem covers and applies to the linear split feasibility problem with DVCs in RTTP. Our convergence results about the dynamic string-averaging CQ-algorithm presented here rely on convexity assumptions. Therefore, there remains an open question whether our work can be expanded to cover the case of the non-convex constraints in the space of dose vectors d used in RTTP. Recent work in the field report on strides made in the field of projection methods when the underlying sets are non-convex, see, for example, Hesse et al. (2014); Bauschke et al. (2014); Attouch et al. (2013). This encourages us to expand the results presented here in the same way.

1.3 Structure of the paper

We begin by briefly reviewing relevant "split problem" formulations which have led to our proposed MOSCFPP and a "dynamic string-averaging CQ-method" to solve it. Starting from a general formulation of two concurrent inverse problems in different vector spaces connected by a bounded linear operator, we explore the inclusion of multiple convex constraint sets within each vector space. Defining operators that act on each of these sets allows us to formulate equivalent fixed point problems, which naturally leads to our MOSCFPP. We then provide some insight into how one may solve such a problem, using constrained minimization, or successive metric projections as part of a CQ-type method (Byrne, 2002). These projection methods form the basis of our "dynamic string-averaging CQ-method", which is introduced in Section 4. Important mathematical foundations for this method are provided in Section 3, which serve to describe the conditions under which the method converges to a solution in Section 5. Finally, we bring percentage violation constraints (PVCs) into our problem formulation (Section 6) and consolidate our work by providing examples of how the MOSCFPP and "dynamic string-averaging CQ-method" may be applied in RTTP (Section 7). A numerical example is provided on a synthetically created treatment plan, detailed in Section 8.

An important comment must be made here. The introduction of a new mathematical model for an application naturally calls for simulated numerical validation, particularly when a new algorithm is proposed. Here we present a rudimentary numerical example since more complex clinically-relevant treatment plans rely heavily on the medical physics context of the radiation therapy treatment planning problem. As such, they call for evaluation of the results in the context of the radiation therapy treatment planning problem itself and require a dedicated proper background and framework which are outside the scope of this paper. An extensive analysis of the methods presented in this paper, on a number of clinical treatment plans, will be published in an appropriate medical physics journal.

2 A brief review of "split problems" formulations and solution methods

The following brief review of "split problems" formulations and solution methods will help put our work in context. The review is non-exhaustive and focuses only on split problems that led to our new formulation that appears in Problem 1. Other split problems such as "the common solution of the split variational inequality problems and fixed point problems" (see, e.g., Lohawech et al., 2018) or "split Nash equilibrium problems for non-cooperative strategic games" (see, e.g., Li, 2019) and many others are not included here. The "split inverse problem" (SIP), which was introduced by Censor et al. (2012) (see also Byrne et al., 2012), is formulated as follows.

Problem 2 ***The split inverse problem (SIP)***. *Given are two vector spaces X and Y and a bounded linear operator $A : X \to Y$. In addition, two inverse problems are involved. The first one, denoted by IP_1, is formulated in the space X and the second one, denoted by IP_2, is formulated in the space Y. The SIP is:*

$$\text{Find an } x^* \in X \text{ that solves } IP_1 \text{ such that } y^* := Ax^* \in Y \text{ solves } IP_2. \tag{4}$$

The first published instance of a SIP is the "split convex feasibility problem" (SCFP) of Censor and Elfving (1994), which is formulated as follows.

Problem 3 ***The split convex feasibility problem (SCFP)***. *Let $\mathcal{H}$ and $\mathcal{K}$ be two real Hilbert spaces. Given are nonempty, closed and convex sets $C \subseteq \mathcal{H}$ and $Q \subseteq \mathcal{K}$ and a bounded linear operator $A : \mathcal{H} \to \mathcal{K}$. The SCFP is:*

$$\text{Find an } \; x^* \in C \text{ such that } Ax^* \in Q. \tag{5}$$

This problem was employed, among others, for solving an inverse problem in intensity-modulated radiation therapy (IMRT) treatment planning (see Censor et al., 2006; Davidi et al., 2015; Censor et al., 2005). More results regarding the SCFP theory and algorithms, can be found, for example, in Yang (2004); López et al. (2012); Gibali et al. (2018), and the references therein. The SCFP was extended in many directions to Hilbert and Banach spaces formulations. It was extended also to the following "multiple sets split convex feasibility problem" (MSSCFP).

Problem 4 ***The multiple sets split convex feasibility problem (MSSCFP)***. *Let $\mathcal{H}$ and $\mathcal{K}$ be two real Hilbert spaces and r and p be two natural numbers. Given are sets C_i, $1 \leq i \leq p$ and Q_j, $1 \leq j \leq r$, that are closed and convex subsets of $\mathcal{H}$ and $\mathcal{K}$, respectively, and a bounded linear operator $A : \mathcal{H} \to \mathcal{K}$. The MSSCFP is:*

$$\text{Find an } x^* \in \cap_{i=1}^{p} C_i \text{ such that } Ax^* \in \cap_{j=1}^{r} Q_j. \tag{6}$$

Masad and Reich (2007) proposed the "constrained multiple set split convex feasibility problem" (CMSSCFP) which is phrased as follows (see also Latif et al. (2016)).

Problem 5 ***The constrained multiple set split convex feasibility problem (CMSSCFP)***. *Let $\mathcal{H}$ and $\mathcal{K}$ be two real Hilbert spaces and r and p be two natural numbers. Given are sets C_i, $1 \leq i \leq p$ and Q_j, $1 \leq j \leq r$, that are closed and convex subsets of $\mathcal{H}$ and $\mathcal{K}$, respectively, and for $1 \leq j \leq r$, given bounded linear operators $A_j : \mathcal{H} \to \mathcal{K}$. In addition let Φ be another closed and convex subset of $\mathcal{H}$. The CMSSCFP is:*

$$\text{Find an } x^* \in \Phi \text{ such that } x^* \in \cap_{i=1}^{p} C_i \text{ and } A_j x^* \in Q_j, \text{ for } 1 \leq j \leq r. \tag{7}$$

Another extension, due to Censor and Segal (2009), is the following "split common fixed points problem" (SCFPP).

Problem 6 ***The split common fixed points problem (SCFPP).*** *Let $\mathcal{H}$ and $\mathcal{K}$ be two real Hilbert spaces and r and p be two natural numbers. Given are operators $U_i : \mathcal{H} \to \mathcal{H}$, $1 \leq i \leq p$, and $T_j : \mathcal{K} \to \mathcal{K}$, $1 \leq j \leq r$, with nonempty fixed point sets $\mathrm{Fix}\,(U_i)$, $1 \leq i \leq p$ and $\mathrm{Fix}\,(T_j)$, $1 \leq j \leq r$, respectively, and a bounded linear operator $A : \mathcal{H} \to \mathcal{K}$. The SCFPP is:*

$$\text{Find an } x^* \in \cap_{i=1}^{p} \mathrm{Fix}(U_i) \text{ such that } Ax^* \in \cap_{j=1}^{r} \mathrm{Fix}(T_j). \tag{8}$$

Problems 3–6 are SIPs but, more importantly, they are special cases of our MOSCFPP of Problem 1.

Focusing in a telegraphic manner on algorithms for solving some of the above SIPs, we observe that the SCFP of Problem 3 can be reformulated as the constrained minimization problem:

$$\min_{x \in C} \frac{1}{2} \|P_Q(Ax) - Ax\|^2, \tag{9}$$

where P_Q is the orthogonal (metric) projection onto Q. (Note that the term "orthogonal projection" is used mainly for subspaces while the "metric" projection refers to any kind of sets (see, e.g., Cegielski, 2012, Section 2.2.4)). Since the objective function is convex and continuously differentiable with Lipschitz continuous gradients, one can apply the projected gradient method (see, e.g., Goldstein, 1964) and obtain Byrne's well-known *CQ-algorithm* (Byrne, 2002). The iterative step of the CQ-algorithm has the following structure:

$$x^{k+1} = P_C(x^k - \gamma A^\star (Id - P_Q) Ax^k), \tag{10}$$

where $A^\star$ stands for the adjoint ($A^\star = A^T$ transpose in Euclidean spaces) of A, γ is some positive number, Id is the identity operator, and P_C and P_Q are the orthogonal projections onto C and Q, respectively. For the MSSCFP of Problem 4, the minimization model considered in Censor et al. (2005), is

$$\min_{x \in \mathbb{R}^M} \left(\sum_{i=1}^{p} \mathrm{dist}^2(x, C) + \sum_{j=1}^{r} \mathrm{dist}^2(Ax, Q) \right), \tag{11}$$

leading, for example, to a gradient descent method which has an iterative simultaneous projections nature:

$$x^{k+1} = x^k - \gamma \sum_{i=1}^{p} \alpha_i \left(Id - P_{C_i}\right) x^k + \sum_{j=1}^{r} \beta_j A^\star \left(Id - P_{Q_j}\right) Ax^k, \tag{12}$$

where $\gamma \in \left(0, \frac{2}{L}\right)$ with

$$L := \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{r} \beta_j \|A\|_F^2 \tag{13}$$

where $\|A\|_F^2$ is the squared Frobenius norm of A.

Inspired by the above and the work presented in Penfold et al. (2017), we propose in the sequel a "dynamic string-averaging CQ-method" for solving the MOSCFPP of Problem 1.

3 Preliminaries

Through this paper $\mathcal{H}$ and $\mathcal{K}$ are two real Hilbert spaces and let $D \subset \mathcal{H}$. For every point $x \in \mathcal{H}$, there exists a unique nearest point in D, denoted by $P_D(x)$ such that

$$\|x - P_D(x)\| \leq \|x - y\|, \text{ for all } y \in D. \tag{14}$$

The operator $P_D : \mathcal{H} \to \mathcal{H}$ is called the *metric projection* onto D.

Definition 1 *Let $T : \mathcal{H} \to \mathcal{H}$ be an operator and $D \subset \mathcal{H}$.*

(i) The operator T is called ***Lipschitz continuous*** *on D with constant $L > 0$ if*

$$\|T(x) - T(y)\| \leq L\|x - y\|, \textit{ for all } x, y \in D. \tag{15}$$

(ii) The operator T is called ***nonexpansive*** *on D if it is 1-Lipschitz continuous.*

(iii) The ***Fixed Point set*** *of T is*

$$\mathrm{Fix}(T) := \{x \in \mathcal{H} \mid T(x) = x\}. \tag{16}$$

*(iv) The operator T is called c-**averaged** (c-av) (Baillon et al., 1978) if there exists a nonexpansive operator $N : D \to \mathcal{H}$ and a number $c \in (0, 1)$ such that*

$$T = (1 - c)Id + cN. \tag{17}$$

In this case we also say that T is c-av (Byrne, 2004). If two operators T_1 and T_2 are c_1-av and c_2-av, respectively, then their composition $S = T_1T_2$ is $(c_1 + c_2 - c_1c_2)$-av; (see Byrne, 2004, Lemma 2.2.)

*(v) The operator T is called ν-**inverse strongly monotone** (ν-ism) on D if there exists a number $\nu > 0$ such that*

$$\langle T(x) - T(y), x - y\rangle \geq \nu\|T(x) - T(y)\|^2, \textit{ for all } x, y \in D. \tag{18}$$

(vi) The operator T is called **firmly nonexpansive** *(FNE) on D if*

$$\langle T(x)-T(y), x-y\rangle \geq \|T(x)-T(y)\|^2, \text{ for all } x, y \in D, \tag{19}$$

A useful fact is that T is firmly nonexpansive if and only if its complement $Id-T$ is firmly nonexpansive. Moreover, T is firmly nonexpansive if and only if T is $(1/2)$-av (see Goebel and Reich (1984, Proposition 11.2) and Byrne (2004, Lemma 2.3)). In addition, T is averaged if and only if its complement $Id-T$ is ν-ism for some $\nu > 1/2$; (see, e.g., Byrne, 2004, Lemma 2.1).

(vii) The operator T is called **quasi-nonexpansive** *(QNE)*

$$\|T(x)-w\| \leq \|x-w\| \text{ for all } (x,w) \in \mathcal{H} \times \mathrm{Fix}(T) \tag{20}$$

(viii) The operator T is called is called a **cutter** *(also* **firmly quasi- nonexpansive***) ($T \in \mathfrak{T}$) if $\mathrm{Fix}(T) \neq \emptyset$ and*

$$\langle T(x)-x, T(x)-w\rangle \leq 0 \text{ for all } (x,w) \in \mathcal{H} \times \mathrm{Fix}(T). \tag{21}$$

(ix) Let $\lambda \in [0,2]$, the operator $T_\lambda := (1-\lambda)Id + \lambda T$ is called **λ-relaxation** *of the operator T. With respect to cutters above it is known that for $\lambda \in [0,1]$, the λ-relaxation of a cutter is also a cutter (see, e.g., Cegielski, 2012, Remark 2.1.32).*

(x) The operator T is called **ρ-strongly quasi-nonexpansive** *(ρ-SQNE), where $\rho \geq 0$, if $\mathrm{Fix}(T) \neq \emptyset$ and*

$$\|T(x)-w\| \leq \|x-w\| - \rho\|T(x)-x\|, \text{ for all } (x,w) \in \mathcal{H} \times \mathrm{Fix}(T). \tag{22}$$

A useful fact is that a family of SQNE operators with non-empty intersection of fixed point sets is closed under composition and convex combination (see, e.g., Cegielski, 2012, Corollary 2.1.47).

(xi) The operator T is called is called **demi-closed** *at $y \in \mathcal{H}$ if for any sequence $\{x^k\}_{k=0}^{\infty}$ in D such that $x^k \to \overline{x} \in D$ and $T(x^k) \to y$, we have $T(\overline{x}) = y$.*

Next we recall the well-known *Demi-closedness Principle* (Browder, 1965).

Lemma 1 *Let $\mathcal{H}$ be a Hilbert space, D a closed and convex subset of $\mathcal{H}$, and $N : D \to \mathcal{H}$ a nonexpansive operator. Then $Id - N$ (Id is the identity operator on $\mathcal{H}$) is* **demi-closed** *at $y \in \mathcal{H}$.*

Let $A : \mathcal{H} \to \mathcal{K}$ be a bounded linear operator with $\|A\| > 0$, and $C \subseteq \mathcal{H}$ and $Q \subseteq \mathcal{K}$ be nonempty, closed and convex sets. The operator $V : \mathcal{H} \to \mathcal{H}$ which is defined by

$$V := Id - \frac{1}{\|A\|^2} A^\star (Id - P_Q) A \tag{23}$$

is called a *Landweber operator* and $U : \mathcal{H} \to \mathcal{H}$ defined by

$$U := P_C V \tag{24}$$

is called a *projected Landweber operator* with V as in (23). See, e.g., Cegielski (2012, 2015, 2016).

In the general case where $T : \mathcal{H} \to \mathcal{H}$ is quasi-nonexpansive and $A : \mathcal{H} \to \mathcal{K}$ is a bounded and linear operator with $\|A\| > 0$, a so-called *Landweber-type operator* (see, e.g., Cegielski, 2016) is defined by

$$V := Id - \frac{1}{\|A\|^2} A^\star (Id - T) A. \tag{25}$$

Note that (23) is a special case of (25), since P_Q is firmly nonexpansive, thus, quasi-nonexpansive.

4 The dynamic string-averaging CQ-method

In this section we present our "dynamic string-averaging CQ-method" for solving the MOSCFPP of Problem 1. It is actually an algorithmic scheme which encompasses many specific algorithms that are obtained from it by different choices of strings and weights. First, for all $j = 1, 2, \dots, r$, construct from the given data of Problem 1, the operators $V_j : \mathcal{H} \to \mathcal{H}$ defined by

$$V_j := Id - \gamma_j A_j^\star (Id - T_j) A_j, \tag{26}$$

where $\gamma_j \in \left(0, \frac{1}{L_j}\right)$, $L_j = \|A_j\|^2$. For quasi-nonexpansive T_j this definition coincides with that of "Landweber-type operators related to T_j" of Cegielski (2016, Definition 2) with a relaxation of γ_j.

For simplicity, and without loss of generality, we assume that $r = p$ in Problem 1. This is not restrictive since if $r < p$ we will define $T_j := Id$ for $r + 1 \leq j \leq p$, and if $p < r$ we will define $U_i := Id$ for $p + 1 \leq i \leq r$, which, in both cases, will not make any difference to the formulation of Problem 1.

Define $\Gamma := \{1, 2, \dots, p\}$ and for each $i \in \Gamma$ define the operator $R_i : \mathcal{H} \to \mathcal{H}$ by$R_i := U_i V_i$. An *index vector* is a vector $t = (t_1, t_2, \dots, t_q)$ such that $t_i \in \Gamma$ for all $i = 1, 2, \dots, q$. For a given index vector $t = (t_1, t_2, \dots, t_q)$ we denote its *length* by $\ell(t) := q$, and define the operator $Z[t]$ as the product of the individual operators R_i whose indices appear in the index vector t, namely,

$$Z[t] := R_{t_{\ell(t)}} R_{t_{\ell(t)-1}} \cdots R_{t_1}, \tag{27}$$

and call it a *string operator*. A finite set Θ of index vectors is called *fit* if for each $i \in \Gamma$, there exists a vector $t = (t_1, t_2, \ldots, t_q) \in \Theta$ such that $t_s = i$ for some $s \in \Gamma$.

Denote by $\mathcal{M}$ the collection of all pairs (Θ, w), where Θ is a fit finite set of index vectors and

$$w : \Theta \to (0, \infty) \text{ is such that } \sum_{t \in \Theta} w(t) = 1. \tag{28}$$

For any $(\Theta, w) \in \mathcal{M}$ define the convex combination of the end-points of all strings defined by members of Θ by

$$\Psi_{\Theta,w}(x) := \sum_{t \in \Theta} w(t) Z[t](x), \ x \in \mathcal{H}. \tag{29}$$

We fix a number $\Delta \in (0, 1/p)$ and an integer $\bar{q} \geq p$ and denote by $\mathcal{M}_* \equiv \mathcal{M}_*(\Delta, \bar{q})$ the set of all $(\Theta, w) \in \mathcal{M}$ such that the lengths of the strings are bounded and the weights are all bounded away from zero, namely,

$$\mathcal{M}_* := \{(\Theta, w) \in \mathcal{M} \mid \ell(t) \leq \bar{q} \text{ and } w(t) \geq \Delta \text{ for all } t \in \Theta\}. \tag{30}$$

The dynamic string-averaging CQ-method with variable strings and variable weights is described by the following iterative process.

Algorithm 1 ***The dynamic string-averaging CQ-method with variable strings and variable weights***

Initialization: *Select an arbitrary* $x^0 \in \mathcal{H}$,

Iterative step: *Given a current iteration vector* x^k *pick a pair* $(\Theta_k, w_k) \in \mathcal{M}_*$ *and calculate the next iteration vector by*

$$x^{k+1} = \Psi_{\Theta_k, w_k}(x^k). \tag{31}$$

The iterative step of (31) amounts to calculating, for all $t \in \Theta_k$, the strings' end-points

$$Z[t](x^k) = R_{i^t_{\ell(t)}} \cdots R_{i^t_2} R_{i^t_1}(x^k), \tag{32}$$

and then calculating

$$x^{k+1} = \sum_{t \in \Theta_k} w_k(t) Z[t](x^k). \tag{33}$$

This algorithmic scheme applies to x^k successively the operators $R_i := U_i V_i$ whose indices belong to the string t. This can be done in parallel for all strings and then the end-points of all strings are convexly combined, with weights that may vary from iteration to iteration, to form the next iterate x^{k+1}. This is indeed an algorithm provided that the operators $\{R_i\}_{i=1}^p$ all have algorithmic implementations. In this framework we get a *sequential algorithm* by allowing a single string created by the index vector $t = \Gamma$ and a *simultaneous algorithm* by the choice of p different strings of length one each containing one element of Γ. Intermediate structures are possible by judicious choices of strings and weights.

5 Convergence

Next we prove the equivalence between Problem 1 and a common fixed point problem which is not split, give a description of $\text{Fix}(V_j)$, and state a property of V_j.

Lemma 2 *Denote the solution set of Problem 1 by Ω and assume that it is nonempty. Then, for V_j as in (26),*

(i) $x^* \in \Omega$ if and only if x^* solves the common fixed point problem:

$$\textit{Find } x^* \in (\cap_{i=1}^p Fix(U_i)) \cap \left(\cap_{j=1}^r Fix(V_j)\right), \tag{34}$$

(ii) for all $j = 1, 2, \ldots, r$:

$$Fix\ (V_j) = \{x \in \mathcal{H} \mid A_j x \in Fix\ (T_j)\} = A_j^{-1}(Fix\ (T_j)), \tag{35}$$

where A_j^{-1} denotes here the inverse image (pre-image) of A_j. I.e., $A_j^{-1} : \mathcal{K} \to \mathcal{H}$ and for any $y \in \mathcal{K}$, $A_j^{-1}(y) := \{x \in \mathcal{H} \mid A_j x = y\}$;

(iii) if, in addition, all operators T_j are cutters then all V_j are cutters (i.e., are 1-SQNE),

(iv) if T_j is ρ-SQNE, $A_j \cap FixT_j \neq \emptyset$ (here we refer to A_j as the image set of A_j) and satisfies the demi-closedness principle then V_j also satisfies the demi-closedness principle.

Proof. (i) We need to show only that

$$x^* \in \cap_{j=1}^r \text{Fix}(V_j) \Leftrightarrow A_j x^* \in \text{Fix }(T_j) \text{ for all } j = 1, 2, \ldots, r. \tag{36}$$

Indeed, for any $j = 1, 2, \ldots, r$,

$$\begin{gathered} A_j x^* \in \text{Fix }(T_j) \Leftrightarrow A_j x^* - T_j A_j x^* = 0 \\ \Leftrightarrow A_j^\star (Id - T_j) A_j x^* = A_j^\star 0 \ \Leftrightarrow -\gamma_j A_j^\star (Id - T_j) A_j x^* = 0 \\ \Leftrightarrow x^* - \gamma_j A_j^\star (Id - T_j) A_j x^* = x^* \Leftrightarrow x^* \in \text{Fix}(V_j). \end{gathered} \tag{37}$$

(ii) Follows from (37).

(iii) To show that V_j is a cutter take $w \in \text{Fix}(V_j)$, $\gamma_j \in \left(0, \frac{1}{L_j}\right)$ and $\xi \in \mathcal{H}$.

$$\begin{aligned}
&\frac{1}{\gamma_j}\langle w - V_j(\xi), \xi - V_j(\xi)\rangle \\
&= \left\langle w - \xi - \gamma_j A_j^\star(T_j - Id)A_j\xi, A_j^\star(Id - T_j)A_j\xi\right\rangle \\
&= \left\langle w - \xi, A_j^\star(Id - T_j)A_j\xi\right\rangle + \gamma_j\|A_j^\star(Id - T_j)A_j\xi\|^2 \\
&= \langle A_j w - A_j\xi, (Id - T_j)A_j\xi\rangle + \gamma_j\|A_j^\star(Id - T_j)A_j\xi\|^2 \\
&= \langle A_j w - T_j(A_j\xi), (Id - T_j)A_j\xi\rangle + \gamma_j\|A_j^\star(Id - T_j)A_j\xi\|^2 \\
&\quad - \|(Id - T_j)A_j\xi\|^2.
\end{aligned} \tag{38}$$

Since T_j is a cutter and $A_j w \in \text{Fix }(T_j)$, we have

$$\langle A_j w - T_j(A_j\xi), (Id - T_j)A_j\xi\rangle \leq 0. \tag{39}$$

Also,

$$\gamma_j\|A_j^\star(Id - T_j)A_j\xi\|^2 \leq \gamma_j\|A_j\|^2\|(Id - T_j)A_j\xi\|^2 \leq \|(Id - T_j)A_j\xi\|^2, \tag{40}$$

for all $\gamma_j \in (0, 1/L_j)$. Using the above we get that

$$\langle w - V_j(\xi), \xi - V_j(\xi)\rangle \leq 0. \tag{41}$$

which proves that V_j is a cutter.

(iv) Proved in Cegielski (2016, Theorem 8(iv)). ■

The special case where in Problem 1 there is only one operator $A : \mathcal{H} \to \mathcal{K}$ and (3) is replaced by

$$\text{for all } 1 \leq j \leq r,\ Ax^* \in \text{Fix }(T_j) \tag{42}$$

which amounts to $Ax^* \in \cap_{j=1}^r \text{Fix }(T_j)$ was treated in the literature (see, e.g., Cegielski, 2015, 2016; Wang and Xu, 2011). The extensions to our more general case, necessitated by the application to RTTP at hand, follow the patterns in those earlier papers. In our convergence analysis we rely on the convergence result of Reich and Zalas (2016, Theorem 4.1) who, motivated by Censor and Tom (2003, Algorithm 3.3), invented and investigated the "modular string averaging (MSA) method" (Reich and Zalas, 2016, Procedure 1.1).

For the convenience of the readers we quote next in full details Procedure 1.1 and Theorem 4.1 of Reich and Zalas (2016). We adhere to the original notations of Reich and Zalas and later identify them with the notations of our work. Let $U_i : \mathcal{H} \to \mathcal{H}$ be a finite family of quasi-nonexpansive mappings where $i \in I := \{1, 2, \cdots, M\}$ and define $U_0 := Id$. The problem under investigation is the common fixed point problem of finding an $x \in C := \cap_{i \in I}\text{Fix}(U_i)$. The algorithmic scheme is

$$x^0 \in \mathcal{H},\ \ x^{k+1} = T_k x^k. \tag{43}$$

where the operator T_k depends on a chosen subset of the input operators U_i.

Reich and Zalas proposed Procedure 1.1 for constructing operators T_k (called "modules") is as follows. Fix $N \in \mathbb{N}$, for $n = 1, 2, \ldots, N$; let $\varepsilon \in (0, 1)$ be a fixed parameter; define modules $V_n := U_{-n}$ for all $n = -M, \ldots, 0$. For $n = 1, 2, \ldots, N$ define modules V_n by choosing one of the following cases:

(a) Relaxation: Fix a singleton $J_n = \{j_n\} \subseteq \{-M, \ldots, 0\}$ and a relaxation $\alpha_n \in [\varepsilon, 2 - \varepsilon]$, and set

$$V_n := Id + \alpha_n \left(V_{j_n} - Id\right). \tag{44}$$

(b) Convex combination: Fix a nonempty subset $J_n \subseteq \{-M, \ldots, n-1\}$ and weights $\omega_{j,n} \in [\varepsilon, 1 - \varepsilon]$ satisfying $\sum_{j \in J_n} \omega_{j,n} = 1$, and set

$$V_n := \sum_{j \in J_n} \omega_{j,n} V_j. \tag{45}$$

(c) Composition: Fix a "string" $J_n \subseteq \{-M, \ldots, n-1\}$ with length less than $M + n$ and set

$$V_n := \Pi_{j \in J_n} V_j. \tag{46}$$

Using the above Modular String Averaging (MSA) procedure of Reich and Zalas, by preforming N_k steps with parameter $\varepsilon_k > 0$, T_k is defined as the last module from the pool, that is, $T_k := V^k_{N_k}$. Such constructions of the operators T_k lead to various combination schemes such as: sequential, convex combination and composition. A string averaging (SA) scheme that is relevant to our method here is obtained by taking a convex combination of multiple compositions, as in Reich and Zalas (2016, Equation (1.12)).

Reich and Zalas Theorem 4.1 is quoted next.

Theorem 1 *Let $\left\{x^k\right\}_{k=0}^{\infty}$ be a sequence generated by the iterative method*

$$x^0 \in \mathcal{H}, \quad x^{k+1} = T_k(x^k) \tag{47}$$

and assume that:

(i) each operator U_i, $i \in I$ is a cutter;

(ii) $I \subseteq I_k \cup I_{k+1} \cup \cdots \cup I_{k+s-1}$, for each $k = 0, 1, 2, \ldots,$ and some $s \geq M - 1$;

(iii) the sequence $\{N_k\}_{k=0}^{\infty}$ is bounded.

If, for each $i \in I$, the operator U_i satisfies Opial's demi-closedness principle then the sequence $\{x^k\}_{k=0}^{\infty}$ converges weakly to some point in C.

If, for each $i \in I$, the operator U_i is approximately shrinking and the family $\mathcal{C} := \{\text{Fix } U_i \mid i \in I\}$ is boundedly regular then the sequence $\{x^k\}_{k=0}^{\infty}$ converges strongly to some point in C.

Our convergence theorem for the dynamic string-averaging CQ-method now follows.

Theorem 2 *Let $p \geq 1$ be an integer and suppose that Problem 1 with $r = p$ has a nonempty solution set Ω. Let $\{U_i\}_{i=1}^{p}$ and $\{T_i\}_{i=1}^{p}$ be cutters on Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, respectively. Further assume that $U_i - Id$ and $T_i - Id$ are demi-closed at zero for all i. Then any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 1 with $R_i := U_i V_i$ for all i, where V_i are defined as in (26), converges weakly to a point $x^* \in \Omega$.*

Proof.

First we identify the notations in our work with those in Reich and Zalas (2016).

(1) The operators $\{U_i\}_{i=1}^{M}$ of Reich and Zalas (2016, Theorem 4.1) are our $\{R_i\}_{i=1}^{p}$ where$R_i := U_i V_i$ as described in the beginning of Section 4 above.
(2) Our operators Ψ_{Θ_k, w_k} (31) are identified with the algorithmic operators T_k of Equation (1.12) in Reich and Zalas (2016).
(3) Our operators $\{U_i\}_{i=1}^{p}$ and $\{T_i\}_{i=1}^{p}$ are assumed to be cutters, then so are also $\{V_i\}_{i=1}^{p}$, by Lemma 2(iii). Hence, the composition operators $R_i := U_i V_i$ are ρ-SQNE for all i and, therefore, also our Ψ_{Θ_k, w_k} are ρ-SQNE for all k.
(4) We assume that our $U_i - Id$ and $T_i - Id$ are demi-closed at zero for all i, therefore, by Lemma 2(iv), $V_i - Id$ are also demi-closed at zero. So, our operators $R_i = U_i V_i$, as composition of demi-closed operators, are demi-closed, see for example (Cegielski, 2015, Theorem 4.2). Our operators $R_i = U_i V_i$ are identified with $\{U_i\}_{i=1}^{M}$ of Reich and Zalas (2016).

Next we show that our dynamic string-averaging CQ-method fits into the MSA (Reich and Zalas, 2016, Procedure 1.1) and that the assumptions of Reich and Zalas (2016, Theorem 4.1) hold.

Since we identify our Ψ_{Θ_k, w_k} from (31) with the right-hand side of Equation (1.12) of Reich and Zalas (2016) (being careful with regard to the duplicity of symbols that represent different things in that work and here), Algorithm 1 can be represented by the iterative process of Equation (1.2), or Equation (4.2), of Reich and Zalas (2016).

Next we show the validity of the assumptions needed by Reich and Zalas (2016, Theorem 4.1).

Assumption (i) of Reich and Zalas (2016, Theorem 4.1): The operators $\{U_i\}_{i=1}^{M}$ of Reich and Zalas (2016, Theorem 4.1) are our $R_i := U_i V_i$. Although our R_i are not necessarily cutters, the arguments in the proof of Reich and Zalas (2016, Theorem 4.1) are based on the strongly quasi-nonexpansiveness of the operators T_k there (our Ψ_{Θ_k,w_k}) and by Lemma 2(iii) above, our operators $\{V_i\}_{i=1}^p$ (defined in (26)) are cutters and this together with the assumption on our $\{U_i\}_{i=1}^p$ and $\{T_i\}_{i=1}^p$, yields that the composition operators $R_i := U_i V_i$ are ρ-SQNE for all i and, thus, so are also our Ψ_{Θ_k,w_k}.

Assumptions (ii)+(iii) of Reich and Zalas (2016, Theorem 4.1): Since the construction of the operators Ψ_{Θ_k,w_k} is based on $\mathcal{M}_*$ (30) which mandates a fit Θ, it guarantees that every index $i \in \Gamma$ appears in the construction of Ψ_{Θ_k,w_k} for all $k > 0$, thus, Assumption (ii) in Reich and Zalas (2016, Theorem 4.1) holds. Following the same reasoning, it is clear that the number of steps N_k, defined in the MSA (Reich and Zalas, 2016, Procedure 1.1), is bounded.

The weak convergence part of the proof of Reich and Zalas (2016, Theorem 4.1) requires that all (their) $\{U_i\}_{i=1}^M$ satisfy Opial's demi-closedness principle (i.e., that $U_i - Id$ are demi-closed at zero). In our case, we assume that $U_i - Id$ and $T_i - Id$ are demi-closed at zero for all i. By Lemma 2(iv) above $V_i - Id$ are also demi-closed at zero. So, we identify $\{U_i\}_{i=1}^M$ of Reich and Zalas (2016) with our U_is and V_is and construct first the operators $R_i = U_i V_i$, and then use them as the building bricks of the algorithmic operators Ψ_{Θ_k,w_k}.

Observe that in our proposed dynamic string-averaging scheme the weights are chosen, in every iteration k, so that $(\Theta_k, w_k) \in \mathcal{M}_*$ (see the iterative step of Algorithm 1). This requires, according to (30), that $w(t) \geq \Delta$ for all $t \in \Theta$, where $\Delta \in (0, 1/p)$ is a fixed positive number. Therefore, for any t it must hold that $\sum_{k=0}^{\infty} w_k(t) = \infty$, meaning that we "visit" every operator infinitely many times. This fully coincides with the assumption in (Reich and Zalas, 2016) that $w_k(i) \in [\varepsilon, 1-\varepsilon]$ for some $\varepsilon > 0$ which implies that $\sum_{k=0}^{\infty} w_k(i) = \infty$ for all i, in their notation.

Thus, the desired result is obtained. ■

Remark 1 *(i) If one assumes that the T_j operators are firmly nonexpansive, then similar arguments as in the proof of Moudafi (2011, Theorem 3.1) show that the V_j operators are also averaged and then Reich and Zalas (2016, Theorem 4.1) can be adjusted to hold for averaged operators.*

(ii) It is possible to propose inexact versions of Algorithm 1 following Reich and Zalas (2016, Theorem 4.5) and Combettes' "almost cyclic sequential algorithm (ACA)" (Combettes, 2001, Algorithm 6.1).

(iii) Our work can be extended to cover also underrelaxed operators, i.e., by defining $R_i := (U_i)_\lambda (V_i)_\delta$ for $\lambda, \delta \in [0,1]$. This is allowed due the fact that if an operator is firmly quasi-nonexpansive, then so is its relaxation.

(iv) Reich and Zalas (2016, Theorem 4.1) also includes a strong convergence part under some additional assumptions on their operators $\{U_i\}_{i=1}^M$. It is possible to adjust this theorem

for our case as well.

(v) We proposed here a general scheme that allows dynamic string averaging; the closest CQ variant appears in the work of Wang and Xu (2011, Theorem 3.1) where only sequential, cyclically controlled, iterations are allowed.

(vi) For the case of a two-set non-convex feasibility problem, Attouch et al. (2013, Theorem 5.3) propose a CQ variant but without a relaxation and if more than two non-convex sets are allowed, then a fully simultaneous method is obtained.

6 Percentage violation constraints (PVCs) arising in radiation therapy treatment planning

6.1 Transforming problems with a PVC

Given p closed convex subsets $\Omega_1, \Omega_2, \cdots, \Omega_p \subseteq \mathbb{R}^n$ of the n-dimensional Euclidean space $\mathbb{R}^n$, expressed as level sets

$$\Omega_j = \{x \in \mathbb{R}^n \mid f_j(x) \leq v_j\}, \text{ for all } j \in J := \{1, 2, \ldots, p\}, \tag{48}$$

where $f_j : \mathbb{R}^n \to \mathbb{R}$ are convex functions and v_j are some given real numbers, the convex feasibility problem (CFP) is to find a point $x^* \in \Omega := \cap_{j \in J} \Omega_j$. If $\Omega = \emptyset$ then the CFP is said to be inconsistent.

Problem 7 ***Convex feasibility problem (CFP) with a percentage-violation constraint (PVC) (CFP+PVC).*** *Consider p closed convex subsets $\Omega_1, \Omega_2, \cdots, \Omega_p \subseteq \mathbb{R}^n$ of the n-dimensional Euclidean space $\mathbb{R}^n$, expressed as level sets according to (48). Let $0 \leq \alpha \leq 1$ and $0 < \beta < 1$ be two given real numbers. The CFP+PVC is:*

Find an $x^* \in \mathbb{R}^n$ such that $x^* \in \cap_{j=1}^p \Omega_j$ and in up to a fraction α (i.e., $100\alpha\%$) of the total number of inequalities in (48) the bounds v_j may be potentially violated by up to a fraction β (i.e., $100\beta\%$) of their values.

A PVC is an integer constraint by its nature. It changes the CFP (48) to which it is attached from being a continuous feasibility problem into becoming a mixed integer feasibility problem. Denoting the inner product of two vectors in $\mathbb{R}^n$ by $\langle a, b \rangle := \sum_{i=1}^n a_i b_i$, the linear feasibility problem (LFP) with PVC (LFP+PVC) is the following special case of Problem 7.

Problem 8 ***Linear feasibility problem (LFP) with a percentage-violation constraint (PVC) (LFP+PVC).*** *This is the same as Problem 7 with f_j, for $j = 1, 2, \ldots, p$, in (48) being linear functions, meaning that the sets Ω_j are half-spaces:*

$$\Omega_j = \left\{x \in \mathbb{R}^n \mid \left\langle a^j, x \right\rangle \leq b_j \right\}, \textit{ for all } j \in J, \tag{49}$$

for a set of given vectors $a^j \in \mathbb{R}^n$ and b_j some given real numbers.

Our tool to "translate" the mixed integer LFP+PVC into a "continuous" one is the notion of sparsity norm, called elsewhere the zero-norm, of a vector $x \in \mathbb{R}^n$ which counts the number of nonzero entries of x, that is,

$$\|x\|_0 := |\{x_i \mid x_i \neq 0\}|, \tag{50}$$

where $|\cdot|$ denotes the cardinality, i.e., the number of elements, of a set. This notion has been recently used for various purposes in compressed sensing, machine learning and more. The rectifier (or "positive ramp operation") on a vector $x \in \mathbb{R}^n$ means that, for all $i = 1, 2, \ldots, n$,

$$(x_+)_i := \max(0, x_i) = \begin{cases} x_i, & \text{if } x_i > 0, \\ 0, & \text{if } x_i \leq 0. \end{cases} \tag{51}$$

Obviously, x_+ is always a component-wise nonnegative vector. Hence, $\|x_+\|_0$ counts the number of positive entries of x and is defined by

$$\|x_+\|_0 := |\{x_i \mid x_i > 0\}|. \tag{52}$$

We translate the LFP+PVC to the following.

Problem 9 ***Translated problem of LFP+PVC (for LFP with upper bounds).*** *For the data of Problem 8, let $A \in \mathbb{R}^{p\times n}$ be the matrix whose columns are formed by the vectors a^j and let $b \in \mathbb{R}^p$ be the column vector consisting of the values b_j, for all $j \in J$. The translated problem of LFP+PVC (for LFP with upper bounds) is:*

$$\text{Find an } x^* \in \mathbb{R}^n \text{ such that } \left\langle a^j, x^* \right\rangle \leq (1+\beta)b_j, \tag{53}$$

$$\text{for all } j \in J, \text{ and } \|(Ax^* - b)_+\|_0 \leq \alpha p. \tag{54}$$

The number of the violations in (53) is $\|(Ax^* - b)_+\|_0$ and $\|(Ax^* - b)_+\|_0 \leq \alpha p$ guarantees that the number of violations of up to β in the original row inequalities remains at bay as demanded. This is a split feasibility problem between the space $\mathbb{R}^n$ and the space $\mathbb{R}^p$ with the matrix A mapping the first to the latter. The constraints in $\mathbb{R}^n$ are linear (thus convex) but the constraint

$$x^* \in S := \{y \in \mathbb{R}^p \mid \|(y - b)_+\|_0 \leq \alpha p\} \tag{55}$$

is not convex. This makes Problem 9 similar in structure to, but not identical with, Problem 3.

Similarly, if the linear inequalities in Problem 9 are in an opposite direction, i.e., of the form $c_j \leq \left\langle a^j, x \right\rangle$, for all $j \in J$, then the translated problem of LFP+PVC will be as follows.

Problem 10 ***Translated problem of LFP+PVC (for LFP with lower bounds).*** *For the data of Problem 8, let $A \in \mathbb{R}^{p\times n}$ be the matrix whose columns are formed by the vectors a^j and let $c \in \mathbb{R}^p$ be the column vector consisting of the values c_j, for all $j \in J$. The translated problem of LFP+PVC (for LFP with lower bounds) is:*

$$\text{Find an } x^* \in \mathbb{R}^n \text{ such that } (1-\beta)c_j \leq \left\langle a^j, x^* \right\rangle, \tag{56}$$

$$\text{for all } j \in J, \text{ and } \|(c - Ax^*)_+\|_0 \leq \alpha p. \tag{57}$$

This is also a split feasibility problem between the space $\mathbb{R}^n$ and the space $\mathbb{R}^p$ with the matrix A mapping the first to the latter. The constraints in $\mathbb{R}^n$ are linear (thus convex) but the constraint

$$x^* \in T := \{y \in \mathbb{R}^p \mid \|(c-y)_+\|_0 \leq \alpha p\} \tag{58}$$

is again not convex.

6.2 Translated block LFP+PVC

Consider an $m \times n$ matrix A divided into blocks A_ℓ, for $\ell = 1, 2, \ldots, \Gamma$, with each block forming an $m_\ell \times n$ matrix and $\sum_{\ell=1}^{\Gamma} m_\ell = m$. Further, the blocks are assumed to give rise to block-wise LFPs of the two kinds; those with upper bounds, say for $\ell = 1, 2, \ldots, p$, and those with lower bounds, say for $\ell = p+1, p+2, \ldots, p+r$. PVCs are imposed on each block separately with parameters α_ℓ and β_ℓ, respectively, for all $\ell = 1, 2, \ldots, \Gamma$. The original block-LFP prior to imposing the PVCs is:

$$\begin{array}{ll} A_\ell x \leq b^\ell, & \text{for all } \ell = 1, 2, \ldots, p, \\ c^\ell \leq A_\ell x, & \text{for all } \ell = p+1, p+2, \ldots, p+r. \end{array} \tag{59}$$

Such constraints will be termed "hard dose constraints" (HDCs). After imposing the PVCs and translating the systems according to the principles of Problems 9 and 10 we obtain the translated problem of LFP+PVC for blocks.

Problem 11 ***Translated problem of LFP+PVC for blocks.*** *Find an $x^* \in \mathbb{R}^n$ such that*

$$\begin{array}{rl} A_\ell x^* \leq (1+\beta_\ell)b^\ell, & \textit{for all } \ell = 1, 2, \ldots, p, \\ (1-\beta_\ell)c^\ell \leq A_\ell x^*, & \textit{for all } \ell = p+1, p+2, \ldots, p+r, \\ \|(A_\ell x^* - b^\ell)_+\|_0 \leq \alpha_\ell m_\ell, & \textit{for all } \ell = 1, 2, \ldots, p, \\ \|(c^\ell - A_\ell x^*)_+\|_0 \leq \alpha_\ell m_\ell, & \textit{for all } \ell = p+1, p+2, \ldots, p+r. \end{array} \tag{60}$$

This is a split feasibility problem between the space $\mathbb{R}^n$ and the space $\mathbb{R}^m$ but with a structure similar to Problem 5 where, for $\ell = 1, 2, \ldots, \Gamma$, each A_ℓ maps $\mathbb{R}^n$ to $\mathbb{R}^{m_\ell}$. Again, it is not identical with Problem 5 because here the constraints in $\mathbb{R}^{m_\ell}$, for $\ell = 1, 2, \ldots, \Gamma$, are not convex. . Although Problem 11 defines an upper PVC on exactly p blocks and a lower

PVC on exactly r blocks, we can, without loss of generality, choose to define PVCs only on a subset of these blocks. For blocks without a PVC, the problem reverts to a standard LFP.

7 Application to radiation therapy treatment planning

The process of planning a radiotherapy treatment plan involves a physician providing dose prescriptions which geometrically constrain the distribution of dose deposited in the patient. Choosing the appropriate nonnegative weights of many individual beamlet dose kernels to achieve these prescriptions as best as possible is posed as a split inverse problem. We focus, for our purposes, on constraining the problem with upper and lower dose bounds, and dose volume constraints (DVCs), which we more generally refer to as PVCs in this work. DVCs allow dose levels in a specified proportion of a structure to fall short of, or exceed, their prescriptions by a specified amount. They largely serve to allow more flexibility in the solution space.

Problem 11 describes the split feasibility problem as it applies in the context of radiation therapy treatment planning. Each block represents a defined geometrical structure in the patient, which is classified either as an *avoidance structure* or a *target volume*. An example of an avoidance structure is an organ at risk (OAR), in which one wishes to deposit minimal dose. An example of a target structure is the planning target volume (PTV), to which a sufficient dose is prescribed to destroy the tumoural tissue. If there are p avoidance structures, any number of blocks in $\{1, 2, \dots, p\}$ can have lower PVCs applied. Similarly, if there are r target volumes then any number of blocks in $\{p+1, p+2, \dots, p+r\}$ can have an upper PVC applied.

This problem can be formulated as the MOSCFPP described in Problem 1 as follows. For the data of Problem 11, define $\bar{\Gamma} \subseteq \{1, 2, \dots, p+r\}$ and for all $i = 1, 2, \dots, m_\ell$, let

$$C_\ell^i := \{x \in \mathbb{R}_+^n \mid \langle a_\ell^i, x \rangle \leq (1 + \beta_\ell) b_i^\ell\}, \tag{61}$$

for all $\ell \in \{1, 2, \dots, p\}$ where $\mathbb{R}_+^n$ is the nonnegative orthant, and

$$C_\ell^i := \{x \in \mathbb{R}_+^n \mid (1 - \beta_\ell) c_i^\ell \leq \langle a_\ell^i, x \rangle\}, \tag{62}$$

for all $\ell \in \{p+1, p+2, \dots, p+r\}$. Additionally, let

$$Q_\ell := \{A_\ell x = v \in \mathbb{R}^{m_\ell} \mid \|(v - b^\ell)_+\|_0 \leq \alpha_\ell m_\ell\}, \tag{63}$$

for all $\ell \in \{1, 2, \dots, p\} \cap \bar{\Gamma}$ and

$$Q_\ell := \{A_\ell x = v \in \mathbb{R}^{m_\ell} \mid \|(c^\ell - v)_+\|_0 \leq \alpha_\ell m_\ell\} \tag{64}$$

for all $\ell \in \{p+1, p+2, \dots, p+r\} \cap \bar{\Gamma}$. The above A_ℓ are blocks of the original matrix A and we denote by $A_\ell x = v$ the image of the vector x under A_ℓ.

Problem 12 ***Translated problem of MOSCFPP for RTTP.***

Let the operators $P_{C_\ell^i} : \mathbb{R}^n \to \mathbb{R}^n$ be orthogonal projections onto C_ℓ^i for all $\ell \in \{1, 2, \ldots, p + r\}$ and $i \in \{1, 2, \ldots, m_\ell\}$, and let $P_{Q_\ell} : \mathbb{R}^{m_\ell} \to \mathbb{R}^{m_\ell}$ be orthogonal projections onto Q_ℓ, for all $\ell \in \Gamma$. The translated MOSCFPP for RTTP is:

$$\begin{aligned} &\textit{Find an } x^* \in \mathbb{R}_+^n \textit{ such that } x^* \in \bigcap_{\ell=1}^{p+r} \bigcap_{i=1}^{m_\ell} \mathrm{Fix}(P_{C_\ell^i}) \textit{ and,} \\ &\textit{for all } \ell \in \Gamma,\ A_\ell x^* \in \mathrm{Fix}(P_{Q_\ell}). \end{aligned} \tag{65}$$

We seek a solution to Problem 12 using our dynamic string-averaging CQ-method, described in Algorithm 1. We define, for all $\ell \in \Gamma$,

$$V_\ell := Id - \gamma_\ell A_\ell^T (Id - P_{Q_\ell}) A_\ell, \tag{66}$$

where $\gamma_\ell \in \left(0, \frac{1}{L_\ell}\right)$, $L_\ell = \|A_\ell\|^2$ and A_ℓ^T is the transpose of A_ℓ.

Remark 2 *In practical use relaxation parameters play an important role:*

(i) Each projection operator $P_{C_\ell^i} : \mathbb{R}^n \to \mathbb{R}^n$ may be relaxed with a parameter $\lambda_\ell \in (0, 2)$ defined on the block $\ell \in \{1, 2, \ldots, p + r\}$.

(ii) The relaxation parameters λ_ℓ, as defined in (i), and γ_ℓ, as given in (66), are permitted to take any value within their bounds on any iterative step of Algorithm 1. That is, they may depend on (vary with) the iteration index k and, therefore, be labeled λ_ℓ^k and γ_ℓ^k.

(iii) The sets Q_ℓ are nonconvex and if for a given $\alpha_\ell m_\ell$ it is nonempty, then it is also closed and then projection onto Q_ℓ exists, is not necessarily unique, but can be calculated explicitly, see, e.g., (Penfold et al., 2017, Eq. (24)). For properties regarding similar sets see, e.g., (Beck, 2017, Subsection 6.8.3). A recent work of Hesse et al. (2014) includes an investigation of these questions, see Section III there. Answers about the sets Q_ℓ and projections onto them in the specific setting related to the radiation therapy treatment planning problem considered here are not yet available.

Tracking the percentage of elements in the current iteration of dose vectors $A_\ell x^k$ that are violating their constraints enables one to impose an adaptive version of Algorithm 1 using the comments in Remark 2. If, for example, one block has more PVC violations than LFP (dose limit constraints) violations then one could choose to alter the relaxation parameters at the next iteration, λ_ℓ^{k+1} and γ_ℓ^{k+1}, in order to place less emphasis on the projections onto C_ℓ^i.

8 Numerical implementation

8.1 Operator definitions

In Problem 12 we introduced the orthogonal projection operators $P_{C_\ell^i}$, which act in the space of the pencil beam intensity vector x, and P_{Q_ℓ}, which act in the space of the dose vector $A_\ell x$. Here we provide explicit formulae, as examples, for calculating these projections in practice. Given an arbitrary vector $z \in \mathbb{R}^n$ and some $\ell \in \{1, 2, \ldots, p+r\}$ and $i \in \{1, 2, \ldots, m_\ell\}$, if it is the case that z is not in C_ℓ^i then it must be projected onto the nearest hyperplane which defines the boundary of C_ℓ^i. Otherwise, no action is taken. If block ℓ represents an avoidance structure ($\ell \in \{1, 2, \ldots, p\}$) then the projection can be calculated by

$$P_{C_\ell^i}(z) = \begin{cases} z, & \langle a_\ell^i, z\rangle \leq (1+\beta_\ell)b_i^\ell, \\ z + \lambda_\ell \dfrac{(1+\beta_\ell)b_i^\ell - \langle a_\ell^i, z\rangle}{\langle a_\ell^i, a_\ell^i\rangle} a_\ell^i, & \langle a_\ell^i, z\rangle > (1+\beta_\ell)b_i^\ell, \end{cases} \tag{67}$$

where $\lambda_\ell \in (0, 2)$ is a user-selected relaxation parameter. Alternatively, if ℓ represents a target structure ($\ell \in \{p+1, p+2, \ldots, p+r\}$) then the projection can be similarly calculated using

$$P_{C_\ell^i}(z) = \begin{cases} z, & \langle a_\ell^i, z\rangle \geq (1-\beta_\ell)c_i^\ell, \\ z + \lambda_\ell \dfrac{(1-\beta_\ell)c_i^\ell - \langle a_\ell^i, z\rangle}{\langle a_\ell^i, a_\ell^i\rangle} a_\ell^i, & \langle a_\ell^i, z\rangle < (1-\beta_\ell)c_i^\ell. \end{cases} \tag{68}$$

Note that, since in the above $\lambda_\ell \in (0, 2)$ are used, then $P_{C_\ell^i}(z)$ are *relaxed* projections.

It is of interest to note that in clinical practice a structure may well have both an upper bound and a lower bound placed on the permitted dose. Such cases can be handled by simply defining two blocks for the same structure, one as an avoidance block, to which (67) applies, and one as a target block, to which (68) applies.

Projection of the dose vector onto Q_ℓ follows a slightly more elaborate procedure. We first define a helper set,

$$\overline{Q}_\ell := \{y \in \mathbb{R}^{m_\ell} \mid \|y_+\|_0 \leq \alpha_\ell m_\ell\}, \tag{69}$$

and describe the projection onto the set, $P_{\overline{Q}_\ell}$, by the following rules: for an arbitrary vector $y \in \mathbb{R}^{m_\ell}$, first count the number of positive entries, $\|y_+\|_0$. If $\|y_+\|_0 \leq \alpha_\ell m_\ell$ then the vector is in $\overline{Q}_\ell$ and no action is needed; $P_{\overline{Q}_\ell} = Id$, the identity operator. However, if $\|y_+\|_0 > \alpha_\ell m_\ell$ then $P_{\overline{Q}_\ell}$ replaces the $\lfloor(\|y_+\|_0 - \alpha_\ell m_\ell)\rfloor$ smallest positive components of y with zeros and leaves the others unchanged. We can now define P_{Q_ℓ} in terms of a projection onto the helper set. Given $v \in \mathbb{R}^{m_\ell}$,

$$P_{Q_\ell}(v) = \begin{cases} P_{\overline{Q}_\ell}(v - b^\ell) + b^\ell, & \ell \in \{1, 2, \ldots, p\} \cap \overline{\Gamma}, \\ -P_{\overline{Q}_\ell}(c^\ell - v) + c^\ell, & \ell \in \{p+1, p+2, \ldots, p+r\} \cap \overline{\Gamma}. \end{cases} \tag{70}$$

Since the sets $\overline{Q}_\ell$ are non-convex, the projection is not unique and so it might happen that only one of the possible values has to be chosen. The reader is referred to related results by Lu and Zhang (2012, Proposition 3.1), Hesse et al. (2014, Equation (20)) and Schaad (2010, Page 54).

8.2 Inverse planning algorithm

We provide here a practical example of how Algorithm 1 may be implemented for inverse planning in radiation therapy treatment planning. In this example we initialize each of the beamlet weights to unit intensity, $x^0 = (1, 1, \ldots, 1)^T$, before running through multiple cycles of an iterative scheme that is equivalent to a fully sequential Algorithm 1 with unit weights, $w_k = 1$ for all k, in (33). The pseudo-code of this procedure is detailed in Algorithm 2. The two "for" loop control cycles therein imply that the blocks, ℓ, may be chosen in any order, without replacement, and so may the voxels, i, within each block. Within each cycle, a nonnegativity constraint is enforced after all possible projections have been applied. This sets any unphysical negative entries in the beamlet intensity vector, x, to zero. In this example a preset number of cycles are performed before stopping and accepting the final solution. However, one may easily replace this by a tolerance-based stopping criterion.

Algorithm 2 ***The dynamic string-averaging CQ-method: A pseudo-code example for RTTP***

Initialization: $x^0 = (1, 1, \ldots, 1)^T$, *cycle number* $k = 1$, *choose max cycles* N_{cycles};
while $k < N_{\text{cycles}}$ **do**
  for $\ell \in \{1, 2, \ldots, p, p+1, \ldots, p+r\}$ **do**
    if $\ell \in \overline{\Gamma}$ **then**
      Choose some $0 < \gamma_\ell < 2/\|A_\ell\|^2$;
      $x^k \leftarrow x^k - \gamma_\ell A_\ell^T (A_\ell x^k - P_{Q_\ell}(A_\ell x^k))$;
    end
    for $i \in \{1, 2, \ldots, m_\ell\}$ **do**
      $x^k \leftarrow P_{C_\ell^i}(x^k)$;
    end
    $x^{k+1} \leftarrow x_+^k$ *(enforce nonnegativity constraint)*;
    $k \leftarrow k + 1$;
  end
end

8.3 Numerical example

A two-dimensional pseudo-dose grid was created using *MATLAB*, version R2019a (The MathWorks, Inc., 2020). The grid is made of a matrix of dimensions 512×512 representing 262,144 pixels which altogether comprise an area of dosimetric interest. In a clinical treatment plan this would be the entire patient geometry and the pixels would be replaced by a large number of three-dimensional voxels. Without loss of generality, we assume two spatial dimensions for simplicity. In order to achieve a basic emulation of dose deposited by multiple beamlets, 1,156 Gaussian pseudo-dose kernels were uniformly distributed across the grid. Each kernel had a standard deviation of 20 pixels and an amplitude such that their sum produced a homogeneous intensity map, with a mean value of 50 units. Figure 1(a) shows a visualization of the intensity (pseudo-dose) matrix due to a single Gaussian kernel, with each dotted grid point representing the centre of one of the 1,156 kernels. Figure 1(b) shows the sum of all contributions. Note that each kernel contributes equally to the sum at this stage, prior to the inverse planning procedure. From this point on, for the proper RTTP context, we shall assume that pixel values directly correspond to "dose".

We have thus far introduced 1,156 different matrices of dimensions 512×512. In order to form a *dose-influence* matrix, A, for use in inverse planning, each matrix is collapsed to a single column vector with 262,144 entries, making sure to keep track of which indices corresponded to which spatial positions in the dose grid. The matrix A is formed by all column vectors and therefore has $262,144$ rows and 1156 columns.

A prescription composed of four hard dose constraints (HDCs), for minimum and maximum dose bounds, and three DVCs, shown in Table 1, was applied to three arbitrarily defined disjoint square regions. DVCs in Table 1 are written in the standard notation, $D_{V\%}$, which is the dose that is received by exactly $V\%$ of the structure. In the framework of this paper, an upper DVC on block ℓ is equivalent to writing $D_{100\alpha_\ell\%} \leq b^\ell$ and a lower DVC is equivalent to $D_{100\alpha_\ell\%} \geq c^\ell$. $D_{\max}$ and $D_{\min}$ represent the maximum and minimum dose constraints, respectively. The three defined square regions can be seen overlaying the

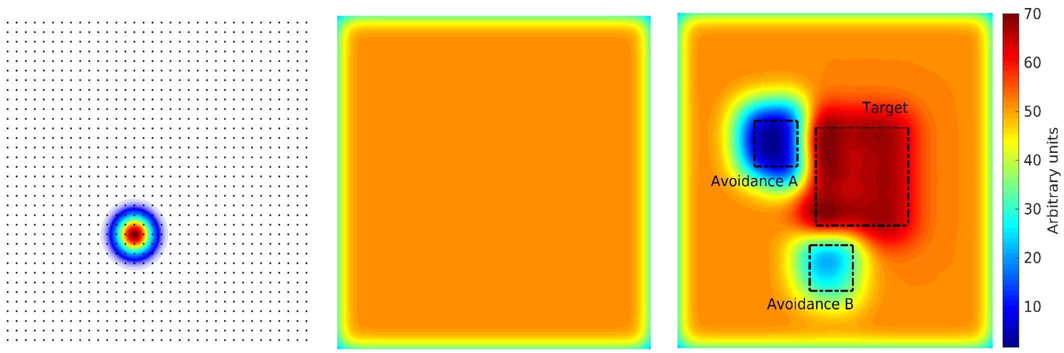
Table 1: Prescription chosen for the two-dimensional numerical example. Pseudo-dose units are arbitrary. $D_{V\%}$ represents the dose that is received by exactly V% of the structure. $D_{\max}$ and $D_{\min}$ represent the maximum and minimum dose constraints, respectively.

Structure	HDCs	DVCs
Avoidance A	$D_{\max} = 25$	$D_{10\%} \leq 20$
Avoidance B	$D_{\max} = 40$	$D_{25\%} \leq 30$
Target	$D_{\min} = 60$	$D_{90\%} \geq 65$
	$D_{\max} = 70$	

dose solution in Figure 1(c). These consist of two avoidance regions, “Avoidance A” and “Avoidance B”, and one target region, “Target”. The column indices of the matrix A corresponding to pixels inside the boundary of these regions can be used to form submatrices, A_1, A_2 and A_3, respectively.

We now have a framework in which Algorithm 2 can be applied. We have A_ℓ for $\ell \in \{1, 2, 3\}$ with $p = 2$ and $r = 1$, and we have $x^0 = (1, 1, \ldots, 1)^T$ with 1,156 entries. In this particular case, both a lower and upper bound on the dose have been prescribed for the “Target” structure. Therefore, we will actually use A_ℓ for $\ell \in \{1, 2, 3, 4\}$, where $A_4 = A_3$ and $\ell = 3$ corresponds to the minimum dose constraint while $\ell = 4$ corresponds to the maximum dose constraint.

Algorithm 2 was applied to the problem described above in order to reduce the dose in the avoidance structures and elevate it in the target structure, according to the prescription in Table 1. Forty cycles ($N_{\text{cycles}} = 40$) were used and the relaxation parameters, λ_ℓ and γ_ℓ, were set to their midrange values, 1 and $1/\|\|A_\ell\|^2$, respectively. Explicitly, $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$, $\gamma_1 = 1.546 \times 10^{-6}$, $\gamma_2 = 1.545 \times 10^{-6}$, and $\gamma_3 = \gamma_4 = 1.030 \times 10^{-6}$. Figure 1(c) shows a visualisation of the dose solution following the algorithmic procedure. It is common in the clinic to evaluate plans using their dose-volume histogram (DVH), which shows the percentage of each structure that has received a certain dose. Figure 2 shows a suitable DVH for this plan, with all prescriptions being approximately met. General convergence to the solution is indicated by a decrease in the total number of pixels violating the constraint imposed upon them, shown in the log-loss plot in Figure 3. Further, log-loss plots for all four types of constraints (minimum dose, maximum dose, lower DVC and upper DVC) are displayed in Figure 4. Again, these all show a general decrease in the number of



(a) (b) (c)

Figure 1: (a) A single Gaussian pseudo-dose kernel contribution shown at one grid point. (b) Homogeneous pseudo-dose of 50 units formed by superimposing all 1,156 Gaussian contributions. (c) Optimized pseudo-dose map showing the structures for which the prescription in Table 1 was applied.

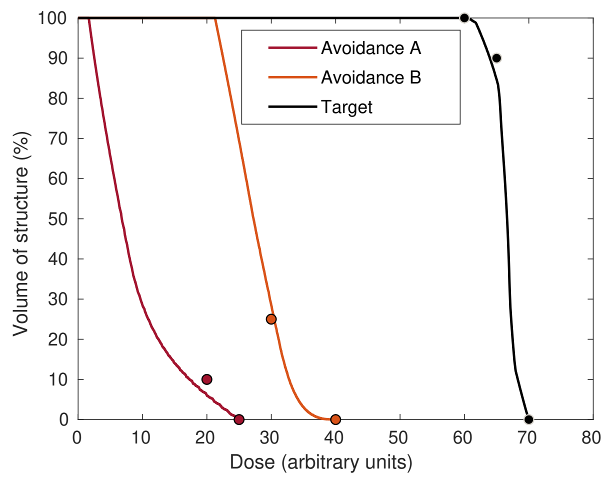


Figure 2: Cumulative dose-volume histogram (DVH) showing the percentage of each structure that has received a certain dose. HDC and DVC prescriptions are shown as filled circles.

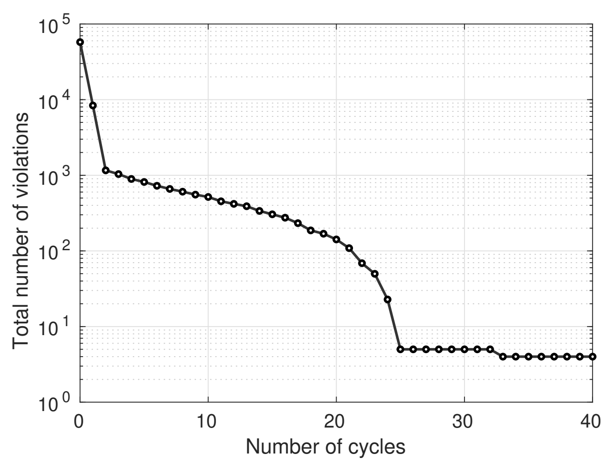


Figure 3: Number of total violations as a function of number of the algorithmic cycles. A decrease indicates improvement in meeting the prescription.

violations and, therefore, indicate that the solution gradually improves as the number of cycles increases.

As mentioned in Section 1.3, more extensive analysis in the context of radiation therapy treatment planning and, in particular, medical physics is necessary in order to justify the use of the proposed dynamic string-averaging CQ-method. This work is ongoing and will be published in an appropriate medical physics journal.

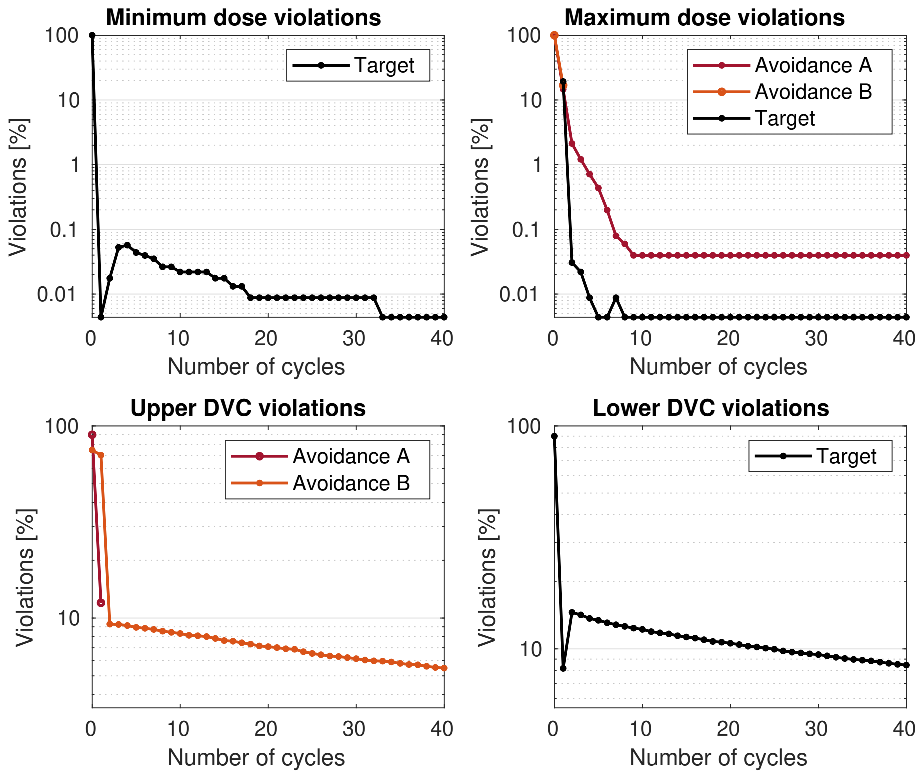


Figure 4: Percentage of violations as a function of the number of algorithmic cycles, shown separately for HDCs (minimum and maximum doses) and DVCs. An upper DVC is that which is applied to an avoidance structure while a lower DVC is that which is applied to a target structure.

9 Conclusions

We introduced a new split feasibility problem called "the multiple-operator split common fixed point problem" (MOSCFPP). This problem generalizes some well-known split feasibility problems such as the split convex feasibility problem, the split common fixed point problem and more. Following the recent work of Penfold et al. (2017), and motivated from the field of radiation therapy treatment planning, the MOSCFPP involves additional so-called Percentage Violation Constraints (PVCs) that give rise to non-convex constraints sets. A new string-averaging CQ method for solving the problem is introduced, which provides the user great flexibility in the weighting and order in which the projections onto the individual sets are executed.

List of abbreviations

RTTP Radiation therapy treatment planning

PVC Percentage violation constraint

DVC Dose volume constraint

MOSCFPP Multiple-operator split common fixed point problem

SIP Split inverse problem

SCFP Split convex feasibility problem

IMRT Intensity modulated radiation therapy

MSSCFP Multiple sets split convex feasibility problem

CMSSCFP Constrained multiple set split convex feasibility problem

SCFPP Split common fixed points problem

FNE Firmly nonexpansive

SQNE Strongly quasi-nonexpansive

MSA Modular string averaging

ACA Almost cyclic sequential algorithm

CFP Convex feasibility problem

LFP Linear feasibility problem

OAR Organ at risk

PTV Planning target volume

HDC Hard dose constraint

DVH Dose-volume histogram

Acknowledgments

We thank Scott Penfold, Reinhard Schulte and Frank Van den Heuvel for their help and encouragement of our work on this project. We are grateful to the reviewers for their constructive and helpful comments on the previous version of this paper. This work was supported by Cancer Research UK, grant number C2195/A25197, through a CRUK Oxford

Centre DPhil Prize Studentship and by the ISF-NSFC joint research program grant No. 2874/19. All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

References

Attouch, H., Bolte, J., Svaiter, B.F., 2013. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. *Mathematical Programming* 137, 91-129.

Baillon, J.B., Bruck, R.E., Reich, S., 1978. On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. *Houston Journal of Mathematics* 4, 1–9.

Bauschke, H.H. , Phan, H.M., Wang, X., 2014. The method of alternating relaxed projections for two nonconvex sets. *Vietnam Journal of Mathematics* 42, 421–450.

Beck, A., 2017. *First-Order Methods in Optimization*, *MOS-SIAM Series on Optimization*. SIAM, Philadelphia, PA, USA, 2017.

Browder, F.E., 1965. Fixed-point theorems for noncompact mappings in Hilbert space. *Proceedings of the National Academy of Sciences of the United States of America* 53, 1272–1276.

Byrne, C., 2002. Iterative oblique projection onto convex sets and the split feasibility problem. *Inverse Problems* 18 441–453.

Byrne, C., 2004. A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Problems* 20, 103–120.

Byrne, C., Censor, Y., Gibali, A., Reich, S., 2012. The split common null point problem. *Journal of Nonlinear and Convex Analysis* 13, 759–775.

Candès, E.J., Wakin, M.B., Boyd, S.P., 2008. Enhancing sparsity by reweighted l_1 minimization. *Journal of Fourier Analysis and Applications* 14, 877–905.

Cegielski, A., 2012. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, *Lecture Notes in Mathematics*. Vol. 2057. Springer, Heidelberg, Germany.

Cegielski, A., 2015. General method for solving the split common fixed point problem. *Journal of Optimization Theory and Applications* 165, 385–404.

Cegielski, A., 2016. Landweber-type operator and its properties. In *A Panorama of Mathematics: Pure and Applied*, *Contempory Mathematics* Vol. 658. American Mathematical Society, Providence, RI, pp. 139–148.

Censor, Y., Bortfeld, T., Martin, B., Trofimov, A., 2006. A unified approach for inversion problems in intensity-modulated radiation therapy. *Physics in Medicine and Biology* 51, 2353–2365.

Censor, Y., Elfving, T., 1994. A multiprojection algorithm using Bregman projections in a product space. *Numerical Algorithms* 8, 221–239.

Censor, Y., Elfving, T., Kopf, N., Bortfeld, T., 2005. The multiple-sets split feasibility problem and its applications for inverse problems. *Inverse Problems* 21, 2071–2084.

Censor, Y., Gibali, A., Reich, S., 2012. Algorithms for the split variational inequality problem. *Numerical Algorithms* 59, 301–323.

Censor, Y., Segal, A., 2009. On the string averaging method for sparse common fixed-point problems. *International Transactions in Operational Research* 16, 481–494.

Censor, Y., Tom, E., 2003. Convergence of string-averaging projection schemes for inconsistent convex feasibility problems. *Optimization Methods and Software* 18, 543–554.

Combettes, P.L., 2001. Quasi-Fejérian analysis of some optimization algorithms. In D. Butnariu, Y. Censor and S. Reich (Editors), *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, *Elsevier Science Publishers*, Amsterdam, The Netherlands, 2001, pp. 115–152.

Davidi, R., Censor, Y., Schulte, R. W., Geneser, S., Xing, L., 2015. Feasibility-seeking and superiorization algorithms applied to inverse treatment planning in radiation therapy. *Contemporary Mathematics* 636, 83–92.

Gibali, A., Liu, L.W., Tang, Y.C., 2018. Note on the modified relaxation CQ algorithm for the split feasibility problem. *Optimization Letters* 12, 817–830.

Goebel, K., Reich, S., 1984. *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*, *Monographs and Textbooks in Pure and Applied Mathematics.* Vol. 83. Marcel Dekker, Inc., New York.

Goldstein, A.A., 1964. Convex programming in Hilbert space. *Bulletin of the American Mathematical Society* 70, 709–710.

Hesse, R., Luke, D. R., Neumann, P., 2014. Alternating projections and Douglas-Rachford for sparse affine feasibility. *IEEE Transactions on Signal Processing* 62, 4868–4881.

Kim, H., Becker, S., Lee, R., Lee, S., Shin, S., Candes, E., Xing, L., Li, R., 2013. Improving IMRT delivery efficiency with reweighted L1-minimization for inverse planning. *Medical Physics* 40, 071719. doi: 10.1118/1.4811100.

Latif, A., Vahidi, J., Eslamian, M., 2016. Strong convergence for generalized multiple-set split feasibility problem. *Filomat* 30, 459–467.

Li, J., 2019. Split equilibrium problems for related games and applications to economic theory. *Optimization* 68, 1203–1222.

Lohawech, P., Kaewcharoen, A., Farajzadeh, A., 2018. Algorithms for the common solution of the split variational inequality problems and fixed point problems with applications. *Journal of Inequalities and Applications* (2018) 2018:358 https://doi.org/10.1186/s13660-018-1942-1.

López, G., Martín-Márquez, V., Wang, F., Xu, H.K., 2012. Solving the split feasibility problem without prior knowledge of matrix norms. *Inverse Problems* 28, 085004.

Lu, Z., Zhang, Y., 2012. Sparse approximation via penalty decomposition methods. *SIAM Journal on Optimization* 23, 2448–2478.

Maass, K., Kim, M., Aravkin, A., 2019. A nonconvex optimization approach to IMRT planning with dose-volume constraints. Available on: https://arxiv.org/abs/1907.10712. (Accessed: 17 December 2020.)

Masad, E., Reich, S., 2007. A note on the multiple-set split convex feasibility problem in Hilbert space. *Journal of Nonlinear and Convex Analysis* 8, 367–371.

Moudafi, A., 2010. The split common fixed-point problem for demicontractive mappings. *Inverse Problems* 26, 055007.

Moudafi, A., 2011. Split monotone variational inclusions. *Journal of Optimization Theory and Applications* 150, 275–283.

Penfold, S., Zalas, R., Casiraghi, M., Brooke, M., Censor, Y., Schulte, R., 2017. Sparsity constrained split feasibility for dose-volume constraints in inverse planning of intensity-modulated photon or proton therapy. *Physics in Medicine and Biology* 62, 3599–3618.

Reich, S., Zalas, R., 2016. A modular string averaging procedure for solving the common fixed point problem for quasi-nonexpansive mappings in Hilbert space. *Numerical Algorithms* 72, 297–323.

Schaad, J., 2010. Modeling the 8-queens problem and Sudoku using an algorithm based on projections onto nonconvex sets. *Masters thesis, The University of British Columbia.*

Wang, F., Xu, H.K., 2011. Cyclic algorithms for split feasibility problems in Hilbert spaces. *Nonlinear Analysis* 74, 4105–4111.

Yang, Q., 2004. The relaxed CQ algorithm solving the split feasibility problem. *Inverse Problems* 20, 1261–1266.