A necessary condition for the guarantee of the superiorization method

Kay Barshad[1], Yair Censor[2], Walaa Moursi[1], Tyler Weames[1] and Henry Wolkowicz[1]

July 13, 2024

[1]Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
✉ kbarshad@uwaterloo.ca; walaa.moursi@uwaterloo.ca; tweames@uwaterloo.ca; hwolkowi@uwaterloo.ca

[2]Department of Mathematics, University of Haifa, Mt. Carmel, Haifa 3498838, Israel
✉ yair@math.haifa.ac.il

Abstract

We study a method that involves principally convex feasibility-seeking and makes secondary efforts of objective function values reduction. This is the well-known superiorization method (SM) where the iterates of an asymptotically convergent iterative feasibility-seeking algorithm are perturbed by objective function nonascent steps. We investigate the, yet only partially answered in the literature, question under what conditions one can guarantee that a sequence generated by an SM algorithm asymptotically converges to a feasible point whose objective function value is superior (meaning smaller or equal) to that of a feasible point reached by the unperturbed exactly same feasibility-seeking algorithm that the SM algorithm employs. We present a condition under which an SM algorithm that uses negative gradient descent steps in its perturbations fails to yield such a superior outcome. The significance of the discovery of this "negative condition" is that it necessitates that the inverse of this condition will have to be assumed to hold in any future guarantee result for the SM. The condition is important for practitioners who use the SM because it is avoidable in experimental work with the SM, thus increasing the success rate of the method in real-world applications.

Keywords. Feasibility-seeking; superiorization; bounded perturbations resilience; dynamic string-averaging; strict Fejér monotonicity; Guarantee question of superiorization;

1 Introduction

The superiorization method (SM) interlaces into a feasibility-seeking algorithm (called the "basic algorithm") objective function value reduction steps (called "perturbations") creating a, so called, "superiorized version of the basic algorithm". These steps cause the objective function to reach lower values locally, prior to performing the next feasibility-seeking iterations. A mathematical guarantee has not been found to date that the overall process of the superiorized version of the basic algorithm will not only retain its feasibility-seeking nature, but also accumulate and preserve globally the objective function reductions.

Numerous works that are cited in [9] show that this global function reduction of the SM occurs in practice in many real-world applications. In addition to a partial answer in [17] with the aid of **the concentration of measure principle** there is also the partial result of [20, Theorem 4.1] about **strict Fejér monotonicity** of sequences generated by an SM algorithm.

The guarantee problem of the SM is the question: under what conditions, if any, one can guarantee that a sequence generated by the superiorized version of a bounded perturbation resilient (see Definition 3.1 below) iterative feasibility-seeking algorithm converges to a feasible point that has objective function value smaller or equal to that of a point to which this algorithm would have converged if no perturbations were applied-everything else being equal.

In the preliminaries section we provide, for the reader's convenience, a compact brief review about the SM and bounded perturbations resilience of algorithms and present the dynamic string-averaging projection (DSAP) feasibility-seeking algorithmic scheme of [20] that is discussed here. Recent reviews on the topic appear in [25] and [10].

In spite of all the above, examples of cases where the SM fails have been constructed, see [2, 33]. In this context, "fails" means that the guarantee problem question stated above did not hold. So, the quest for recognizing the properties of such situations in order to make statements on the guarantee problem of the SM continues.

Proving mathematically a guarantee of global objective function value reduction of the SM, compared to running its feasibility-seeking algorithm without perturbations, will probably require some additional assumptions on the feasible set, on the objective function, on the parameters involved, or even on the set of permissible initialization points. We present in this note a "negative condition", namely, a condition under which an SM algorithm that is based on the dynamic string-averaging projection (DSAP) feasibility-seeking algorithmic scheme, which uses negative gradient descent steps in its perturbations, will fail to yield a superior outcome.

The significance of this negative condition is twofold. On one hand, it necessitates that a reverse statement that will nullify it will have to be assumed to hold in any potential future guarantee result for the SM. On the other hand, we show that when this condition does hold, then the SM algorithm can produce a superior result only if it converges to an optimal point. The latter statement is related to the alternative in Theorem 4.1(a) of [20], reproduced in Theorem 3.5 below.

Although this note is discussing theory, we care to make a comment about computational aspects. In numerical examples one should present large-size problems because for small problems many algorithms "work" and the performance differences start to show when the problems are very large and quite sparse. We care to mention this because this is why the SM was envisioned in the first place. For example, in [8] it is shown that the advantage of an SM algorithm (called there "LinSup" for Linear Superiorization) for the data of a linear program, over an LP algorithm (Malab's Simplex) increases as the problem sizes grow. Many papers referenced in the continuously updated bibliography page on the SM and perturbation resilience of algorithms [9] attest to the success of SM algorithms in a variety of real-world practical large-size problems.

The original motivation of the SM was, and still is, to handle situations in which reaching a feasible point in the nonempty intersection of finitely many sets, in a tractable manner, is the principal task. In such situations the SM provides a low cost perturbation that preserves the convergence of the perturbed feasibility-seeking algorithm to a feasible point while aiming to improve, aka superiorize, (reduce, not necessarily optimize) the value of an objective function. Methods to increase the effectiveness of the objective function reduction steps in SM algorithms, without invalidating the bounded perturbation resilience of the embedded feasibility-seeking algorithm, have recently been published. These include a randomization approach and an approach with restarts of the step-sizes.

The paper is structured as follows. In Section 2 we list a few (out of many) works that used the SM in practice and in Section 3 we provide, for the reader's convenience, a compact brief review about the SM and bounded perturbations resilience of algorithms, recent reviews appear in [25] and [10]. In that section we also summarize the dynamic string-averaging projection (DSAP) feasibility-seeking algorithmic scheme of [20], that is discussed in this paper. The negative condition and its consequences appear in Section 4.

2 Some applications that used the superiorization methodology

In [21] Guenter et al. consider the fully-discretized modeling of an **image reconstruction from projections** problem that leads to a system of linear equations which is huge and very sparse. Solving such systems, sometimes

3

under limitations on the computing resources, is and remains a challenge. The authors aim not only at solving the linear system resulting from the modeling alone, but consider the constrained optimization problem of minimizing an objective function subject to the modeling constraints. To do so, they recognize two fundamental approaches: (i) superiorization, and (ii) regularization. Within these two methodological approaches they evaluate 21 algorithms over a collection of 18 different phantoms (i.e., test problems), presenting their experimental results in very informative ways.

In [7] Fink et al. study the **nonconvex multi-group multicast beam-forming problem** with quality-of-service constraints and per-antenna power constraints. They formulate a convex relaxation of the problem as a semidefinite program in a real Hilbert space, which allows them to approximate a point in the feasible set by iteratively applying a bounded perturbation resilient fixed-point mapping. Inspired by the superiorization methodology, they use this mapping as a basic algorithm, and add in each iteration a small perturbation with the intent to reduce the objective value and the distance to nonconvex rank-constraint sets.

Pakkaranang et al. [3] construct a novel algorithm for solving **non-smooth composite optimization problems**. By using an inertial technique, they propose a modified proximal gradient algorithm with outer perturbations and obtain strong convergence results for finding a solution of a composite optimization problem. Based on bounded perturbation resilience, they present their algorithm with the superiorization method and apply it to image recovery problems. They provide numerical experiments that show the efficiency of the algorithm and compare it with previously known algorithms in signal recovery.

Especially interesting is the recent work of Ma et al. [1] who propose a novel decomposition framework for **derivative-free optimization (DFO) algorithms** which significantly extends the scope of current DFO solvers to larger-scale problems. They show that their proposed framework closely relates to the superiorization methodology.

Many more publications on practical applications of the SM can be found in [9], such as [28–32], to mention but a few.

3 Preliminaries: The superiorization methodology and the dynamic string-averaging projection (DSAP) feasibility-seeking algorithmic scheme

3.1 The superiorization methodology

The superiorization methodology (SM) was born when the terms and notions "superiorization" and "perturbation resilience", in the present context, first appeared in the 2009 paper [14] which followed its 2007 forerunner by Butnariu

et al. [4]. The ideas have some of their roots in the 2006 and 2008 papers of Butnariu et al. [5, 6] where it was shown that if iterates of a nonexpansive operator converge for any initial point, then its inexact iterates with summable errors also converge. Since its inception in 2007, the SM has evolved and gained ground. Recent publications on the SM are devoted to either weak or strong superiorization, without yet using these terms. They are [4, 11–14, 18, 19, 22, 23, 27] many of which contain a detailed description of the SM, its motivation, and an up-to-date review of SM-related previous work.

The Webpage[1] [9] is dedicated to superiorization and perturbation resilience of algorithms and contains a continuously updated bibliography on the subject. It is a source for the wealth of work done in this field to date, including two dedicated special issues of journals [16] and [24].

Our recent review [10] can serve as an introduction to the SM, also [26] and [25] are very helpful. Just to make the continued reading here more convenient for the reader we give below some of the fundamental notions of the SM.

Throughout the rest of the paper we consider a real Hilbert space X with the norm $\|\cdot\|$, a real-valued, convex and continuous function $\phi : X \to R$. For a point $z \in X$, we denote by $\partial\phi(z)$ the subgradient set of ϕ at z.

Definition 3.1. Bounded perturbation resilience (BPR). Let $\Gamma \subseteq X$ be a given nonempty subset of X. An algorithmic operator $\mathcal{A} : X \to X$ is said to be **bounded perturbations resilient with respect to** Γ if the following is true: If a sequence $\{x^k\}_{k=0}^\infty$, generated by the basic algorithm $x^{k+1} := \mathcal{A}(x^k)$, for all $k \geq 0$, converges to a point in Γ for all $x^0 \in X$, then any sequence $\{y^k\}_{k=0}^\infty$ of points in X that is generated by the algorithm $y^{k+1} = \mathcal{A}(y^k + \beta_k v^k)$, for all $k \geq 0$, also converges to a point in Γ for all $y^0 \in X$ provided that, for all $k \geq 0$, $\beta_k v^k$ are **bounded perturbations**, meaning that $\beta_k \geq 0$ for all $k \geq 0$ such that $\sum_{k=0}^\infty \beta_k < \infty$, and that the vector sequence $\{v^k\}_{k=0}^\infty$ is bounded. A basic algorithm of the form $x^{k+1} := \mathcal{A}(x^k)$, for all $k \geq 0$,is said to be **bounded perturbations resilient with respect to** Γ if its algorithmic operator $\mathcal{A}$ is bounded perturbations resilient with respect to Γ.

Algorithm 3.2 is a superiorized version of the basic algorithm governed by $\mathcal{A}$.

3.2 Strict Fejér monotonicity of sequences generated by an SM algorithm with the dynamic string-averaging projection (DSAP) feasibility-seeking algorithmic scheme

The Dynamic String-Averaging Projection (DSAP) method of [19] constitutes a family of algorithmic operators that can play the role of $\mathcal{A}$ in Algorithm 3.2. For each $i = 1, 2, \ldots, m$, denote by $P_i := P_{C_i}$ the metric projection onto the set

---

[1]http://math.haifa.ac.il/yair/bib-superiorization-censor.html#top, last updated on June 15, 2024 with 186 items.

Algorithm 3.2. *Superiorized version of the basic algorithm governed by $\mathcal{A}$.*
*(0) **Initialization**: Let N be a natural number and let $y^0 \in X$ be an arbitrary user-chosen vector.*
*(1) **Iterative step**: Given a current iteration vector y^k pick an $N_k \in \{1, 2, \ldots, N\}$ and start an inner loop of calculations as follows:*
*(1.1) **Inner loop initialization**: Define $y^{k,0} = y^k$.*
*(1.2) **Inner loop step**: Given $y^{k,n}$, as long as $n < N_k$, do as follows:*
(1.2.1) Pick a $0 < \beta_{k,n} \leq 1$ in a way that guarantees that

$$\sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n} < \infty.$$

(1.2.2) Pick an $s^{k,n} \in \partial\phi(y^{k,n})$ and define direction vectors $v^{k,n}$ as follows:

$$v^{k,n} := \begin{cases} -\dfrac{s^{k,n}}{\|s^{k,n}\|}, & \text{if } 0 \notin \partial\phi(y^{k,n}), \\ 0, & \text{if } 0 \in \partial\phi(y^{k,n}). \end{cases}$$

(1.2.3) Calculate the perturbed iterate

$$y^{k,n+1} := y^{k,n} + \beta_{k,n} v^{k,n} \tag{3.1}$$

and if $n + 1 < N_k$ set $n \leftarrow n + 1$ and go to (1.2), otherwise go to (1.3).
(1.3) Exit the inner loop with the vector y^{k,N_k}
(1.4) Calculate

$$y^{k+1} := \mathcal{A}(y^{k,N_k})$$

set $k \leftarrow k + 1$ and go back to (1).

C_i. An **index vector** is a vector $t = (t_1, t_2, \ldots, t_q)$ such that $t_s \in \{1, 2, \ldots, m\}$ for all $s = 1, 2, \ldots, q$, whose length is $\ell(t) = q$. The composition of the individual projections onto the sets whose indices appear in the index vector t is $P[t] := P_{t_q} \cdots P_{t_2} P_{t_1}$, called a **string operator**.

A finite set Ω of index vectors is called **fit** if for each $i \in \{1, 2, \ldots, m\}$, there exists a vector $t = (t_1, t_2, \ldots, t_q) \in \Omega$ such that $t_s = i$ for some $s \in \{1, 2, \ldots, q\}$. Denote by $\mathcal{M}$ the collection of all pairs (Ω, w), where Ω is a finite fit set of index vectors and $w : \Omega \to (0, \infty)$ is such that $\sum_{t \in \Omega} w(t) = 1$.

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

$$P_{\Omega, w}(x) := \sum_{t \in \Omega} w(t) P[t](x), \ x \in X.$$

Let $\Delta \in (0, 1/m)$, and fix an arbitrary integer $\bar{q} \geq m$. Denote by $\mathcal{M}_* \equiv \mathcal{M}_*(\Delta, \bar{q})$ the set of all $(\Omega, w) \in \mathcal{M}$ such that the lengths of the strings are bounded and the weights are all bounded away from zero, i.e.,

$$\mathcal{M}_* := \{(\Omega, w) \in \mathcal{M} \mid \ell(t) \leq \bar{q} \text{ and } w(t) \geq \Delta, \ \forall \, t \in \Omega\}.$$

The convergence properties and bounded perturbation resilience of the DSAP method were analyzed in [19].

Algorithm 3.3. *The DSAP method with variable strings and variable weights*
Initialization: *Select an arbitrary $x^0 \in X$,*
Iterative step: *Given a current iteration vector x^k pick a pair $(\Omega_k, w_k) \in \mathcal{M}_*$ and calculate the next iteration vector x^{k+1} by*

$$x^{k+1} := P_{\Omega_k, w_k}(x^k).$$

Theorem 3.4 ([19, Theorem 12]). *Assume that $\{C_i\}_{i=1}^m$ is a family of nonempty, closed and convex subsets of X with a nonempty intersection C. Let $\{\beta_k\}_{k=0}^\infty$ be a sequence of non-negative numbers such that $\sum_{k=0}^\infty \beta_k < \infty$, let $\{v^k\}_{k=0}^\infty \subset X$ be a norm-bounded sequence, let $\{(\Omega_k, w_k)\}_{k=0}^\infty \in \mathcal{M}_*$ for all $k = 0, 1, \ldots$. Then any sequence $\{y^k\}_{k=0}^\infty$ generated by the iterative formula*

$$y^{k+1} := P_{\Omega_k, w_k}\left(y^k + \beta_k v^k\right)$$

converges in the norm of X and its limit belongs to C. That is, Algorithm 3.3 converges to a point in C and its algorithmic operator is bounded perturbation resilient with respect to C.

Using the algorithmic operator of the DSAP feasibility-seeking Algorithm 3.3 as the algorithmic operator $\mathcal{A}$ in Algorithm 3.2, we recover the following main Theorem 4.1 of [20].

Theorem 3.5 ([20, Theorem 4.1]). *Assume that $\{C_i\}_{i=1}^m$ be a family of nonempty, closed and convex subsets of X with a nonempty intersection C. Set $C_{\min} := \{x \in C \mid \phi(x) \leq \phi(y) \text{ for all } y \in C$. Let $C_* \subseteq C_{\min}$ be a nonempty subset of $C_{\min}$, let $r_0 \in (0, 1]$ and $\bar{L} \geq 1$ be such that, for all $x \in C_*$ and all y such that $||x - y|| \leq r_0$,*

$$|\phi(x) - \phi(y)| \leq \bar{L}||x - y||,$$

and suppose that $\{(\Omega_k, w_k)\}_{k=0}^\infty \subset \mathcal{M}_*$. Then any sequence $\{y^k\}_{k=0}^\infty$, generated by Algorithm 3.2, the superiorized version of the DSAP algorithm, converges in the norm of X to a $y^* \in C$ and exactly one of the following two alternatives holds:

(a) $y^* \in C_{\min}$;

(b) $y^* \notin C_{\min}$ and there exist a natural number k_0 and a $c_0 \in (0, 1)$ such that for each $x \in C_*$ and for each integer $k \geq k_0$,

$$\|y^{k+1} - x\|^2 \leq \|y^k - x\|^2 - c_0 \sum_{n=1}^{N_k - 1} \beta_{k,n}.$$

This shows that $\{y^k\}_{k=k_0}^\infty$ is strictly Fejér-monotone with respect to C_*, i.e., that for every $x \in C_*$, the inequality $\|y^{k+1} - x\|^2 < \|y^k - x\|^2$ holds for all $k \geq k_0$, because $c_0 \sum_{n=1}^{N_k-1} \beta_{k,n} > 0$. The strict Fejér-monotonicity however does not guarantee convergence to a constrained minimum point, but only says that the so-created feasibility-seeking sequence $\{y^k\}_{k=0}^\infty$ has the additional property of getting strictly closer, without necessarily converging, to the points of a subset of the solution set of the constrained minimization problem.

Published experimental results repeatedly confirm that global reduction of the value of the objective function ϕ is indeed achieved, without losing the convergence toward feasibility, see [4, 11–14, 18, 19, 22, 23, 27]. In some of these cases the SM returns a lower value of the objective function ϕ than an exact minimization method with which it is compared, e.g., [15].

4 The "negative condition" on the superiorization method

We consider the dynamic string-averaging projection (DSAP) feasibility-seeking algorithmic scheme of [20] and its superiorized version, as presented above in Subsection 3.2. Speaking specifically about Algorithm 4.1 of [20], we know that according to Theorem 3.5, under the assumptions of this theorem, exactly one of two things must happen, i.e., alternative (a) or alternative (b). This is a non-constructive theorem because it tells nothing about when each of the alternatives can occur.

If alternative (a) is the case, then it is correct to say that the sequence generated by the superiorized version of the bounded perturbation resilient

feasibility-seeking Algorithm 3.3 (i.e., Algorithm 4.1 of [20]), converges to a feasible point that has objective function value smaller or equal to that of a point to which this algorithm would converge if no perturbations were applied. This would be true in this case because $y^* \in C_{\min}$ and the feasibility-seeking algorithm cannot do better.

The question remains for the case of alternative (b). Namely, can we give conditions under which, if alternative (b) holds, we will have that $\phi(y^*) \leqslant \phi(x^*)$, where x^* is the limit of the same feasibility-seeking algorithm that is used in this SM algorithm when it is run without perturbations, but everything else being equal, such as initialization point, relaxation parameters, order of constraints within the sweeps, etc.

The desire to distinguish between the alternatives (a) or (b) of Theorem 3.5 leads us to the next lemma which gives a condition under which a limiting feasible point cannot belong to the solution set of the constrained minimization problem $\min \{\phi(x) \mid x \in C\}$. This is the "negative condition" eluded to above, under which an SM algorithm that is based on the dynamic string-averaging projection (DSAP) feasibility-seeking algorithmic scheme, which uses negative subgradient descent steps in its perturbations, will fail to yield a superior outcome. The notation $d(\hat{c}, D)$ below stands for the distance between the point $\hat{c}$ and the set D.

Lemma 4.1. *Let $\{C_i\}_{i=1}^m$ be a family of nonempty, closed and convex subsets of X with a nonempty intersection C, let $D \subset C$ be a nonempty subset of C and let $r \geq 1$ be a real number. Assume that $\{y^k\}_{k=0}^\infty$ is any sequence, generated by Algorithm 3.2 with positive step-sizes $\left\{ \{\beta_{k,n}\}_{n=0}^{N_k-1} \right\}_{k=0}^\infty$, direction vectors $\left\{ \{v^{k,n}\}_{n=0}^{N_k-1} \right\}_{k=0}^\infty$ and $y^0 \in X$ as an arbitrary initialization point.*
(i) If $\hat{c} \in C$ is a point for which

$$\left\| y^0 - \hat{c} \right\| \leq (r-1) \sum_{n=0}^{N_0-1} \beta_{0,n}, \tag{4.1}$$

then the sequence $\{y^k\}_{k=0}^\infty$ satisfies for all $k \geq 1$,

$$\left\| y^k - \hat{c} \right\| \leq r \sum_{\ell=0}^{k-1} \sum_{n=0}^{N_\ell-1} \beta_{\ell,n}. \tag{4.2}$$

(ii) If, additionally,

$$d(\hat{c}, D) > r \sum_{\ell=0}^{\infty} \sum_{n=0}^{N_\ell-1} \beta_{\ell,n} \tag{4.3}$$

holds, then the limit point y^* of the sequence $\{y^k\}_{k=0}^\infty$, which exists according to Theorem 3.5, obeys $y^* \notin D$.

9

Proof. *(i)* The proof is by induction on k. For $k = 1$, we use (3.1), (4.1), the nonexpansivity of the operator P_{Ω_0, w_0} and the boundedness of the sequence $\{v^{0,n},\}_{n=0}^{\infty}$ to obtain (since $\hat{c} \in C$)

$$
\begin{aligned}
\left\| y^k - \hat{c} \right\| &= \left\| y^1 - \hat{c} \right\| = \left\| P_{\Omega_0, w_0} \left(y^{0, N_0} \right) - P_{\Omega_0, w_0} \left(\hat{c} \right) \right\| \\
&\leq \left\| y^0 + \sum_{n=0}^{N_0 - 1} \beta_{0,n} v^{0,n} - \hat{c} \right\| \leq \left\| y^0 - \hat{c} \right\| + \left\| \sum_{n=0}^{N_0 - 1} \beta_{0,n} v^{0,n} \right\| \\
&\leq \left\| y^0 - \hat{c} \right\| + \sum_{n=0}^{N_0 - 1} \beta_{0,n} \leq r \sum_{n=0}^{N_0 - 1} \beta_{0,n} = r \sum_{\ell=0}^{k-1} \sum_{n=0}^{N_\ell - 1} \beta_{\ell,n}.
\end{aligned}
$$

Suppose that $k > 1$ and make the inductive assumption that

$$
\left\| y^{k-1} - \hat{c} \right\| \leq r \sum_{\ell=0}^{k-2} \sum_{n=0}^{N_\ell - 1} \beta_{\ell,n}. \tag{4.4}
$$

The operator P_{Ω_k, w_k} is nonexpansive for each k, and since $\hat{c} \in C$, recalling (3.1) and using the inductive assumption (4.4) above (since $r \geq 1$), we have

$$
\begin{aligned}
\left\| y^k - \hat{c} \right\| &= \left\| P_{\Omega_{k-1}, w_{k-1}} \left(y^{k-1, N_{k-1}} \right) - P_{\Omega_{k-1}, w_{k-1}} \left(\hat{c} \right) \right\| \\
&= \left\| P_{\Omega_{k-1}, w_{k-1}} \left(y^{k-1} + \sum_{n=0}^{N_{k-1} - 1} \beta_{k-1,n} v^{k-1,n} \right) - P_{\Omega_{k-1}, w_{k-1}} \left(\hat{c} \right) \right\| \\
&\leq \left\| y^{k-1} + \sum_{n=0}^{N_{k-1} - 1} \beta_{k-1,n} v^{k-1,n} - \hat{c} \right\| \leq \left\| y^{k-1} - \hat{c} \right\| + \sum_{n=0}^{N_{k-1} - 1} \beta_{k-1,n} \\
&\leq r \sum_{\ell=0}^{k-2} \sum_{n=0}^{N_\ell - 1} \beta_{\ell,n} + r \sum_{n=0}^{N_{k-1} - 1} \beta_{k-1,n} = r \sum_{\ell=0}^{k-1} \sum_{n=0}^{N_\ell - 1} \beta_{\ell,n}.
\end{aligned}
$$

This completes the induction and shows that *(i)* indeed holds.
(ii) Assume to the contrary that $y^* \in D$. Taking the limit, as $k \to \infty$, on both sides of (4.2) yields

$$
\left\| y^* - \hat{c} \right\| \leq r \sum_{\ell=0}^{\infty} \sum_{n=0}^{N_\ell - 1} \beta_{\ell,n}
$$

which contradicts (4.3), thus proving that y^* cannot be in D. $\qquad\square$

Remark 4.2. Lemma 4.1 shows that if (4.2) and (4.3) hold, then alternative (a) of Theorem 3.5 cannot hold. The case $r = 1$ is of theoretical interest only because then the initialization point y^0 is feasible, contrary to the prevailing situation in applications wherein the feasibility-seeking is initialized outside the feasible set C.

The next corollary provides an insight into a necessary choice of an initialization point of Algorithm 3.2 in order to establish its convergence to a point in D.

Corollary 4.3. *Under the assumptions of Lemma 4.1, assume that a sequence $\{y^k\}_{k=0}^{\infty}$ generated by Algorithm 3.2 with step-sizes $\left\{\{\beta_{k,n}\}_{n=0}^{N_k-1}\right\}_{k=0}^{\infty}$ converges to a point $y^* \in D$ and there exists a point $\hat{c} \in C$ such that*

$$d\left(\hat{c}, D\right) > \sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n}. \tag{4.5}$$

Then the following inequality holds

$$\left\|y^0 - \hat{c}\right\| \geq \left(d\left(\hat{c}, D\right) \left(\sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n}\right)^{-1} - 1\right) \sum_{n=0}^{N_0-1} \beta_{0,n}. \tag{4.6}$$

Proof. Clearly,

$$d\left(\hat{c}, D\right) > r \sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n}$$

for any real $1 \leq r < d\left(\hat{c}, D\right) \left(\sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n}\right)^{-1}$. Since the limit point of any sequence generated by Algorithm 3.2 belongs to D, we must have, by Lemma 4.1, that

$$\left\|y^0 - \hat{c}\right\| > (r-1) \sum_{n=0}^{N_0-1} \beta_{0,n}$$

for each such r. (4.6) now follows by (4.5). □

Remark 4.4. Note that Lemma 4.1 and Corollary 4.3 are, in particular, true for $D = C_{\min} := \{x \in C \mid \phi(x) \leq \phi(y) \text{ for all } y \in C\}$.

Under the assumption (4.5), Corollary 4.3 provides a non-trivial necessary condition, namely, (4.6), for the convergence of any sequence generated by Algorithm 3.2 to a point in D. However, the aforesaid condition is not sufficient to this end, even if the assumption (4.5) holds. This observation is demonstrated in following example.

Example 4.5. Let $X := \mathbb{R}$, let $\phi : X \to X$ be defined by $\phi(x) := x^2$ for each $x \in X$ and define $C := [0, 10]$. Clearly, C is closed and convex subset of $\mathbb{R}$. Note that ϕ is 20-Lipschitz continuous on C, that is, $\|\phi(x) - \phi(x')\| \leq 20 \|x - x'\|$ for each $x, x' \in C$. Set $D := C_{\min} = \{0\}$, $\hat{c} := 8 \in C$ and define the sequence of step-sizes $\{\beta_{k,0}\}_{k=0}^{\infty}$ of length 1 (that is, $N_k = 1$ for each $k = 0, 1 \ldots$) by $\beta_{k,0} := 2^{-k}$ for each $k = 0, 1 \ldots$. We then have

$$d\left(\hat{c}, D\right) = 8 > 2 = \sum_{k=0}^{\infty} \beta_{k,0} = \sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n}.$$

Now choose any $y^0 \geq 13$. Then we have

$$\sum_{n=0}^{N_0-1} \beta_{0,n} = \beta_{0,0} = 1 \text{ and } \left(d\left(\hat{c}, D\right) \left(\sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n}\right)^{-1} - 1\right) = 3,$$

11

therefore,

$$\left\| y^0 - \hat{c} \right\| \geq 3 = \left(d\left(\hat{c}, D\right) \left(\sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n} \right)^{-1} - 1 \right) \sum_{n=0}^{N_0-1} \beta_{0,n}.$$

However, the sequence $\left\{ y^k \right\}_{k=0}^{\infty}$, generated by Algorithm 3.2, satisfies for each positive integer k,

$$y^k = 10 - \sum_{\ell=1}^{k-1} 2^{-\ell} \underset{k\to\infty}{\to} 9 \notin D.$$

Similarly, if we choose in this example $C := [0, 3 \cdot 2^{-1}]$, $\hat{c} := 1 \in C$ and keep all other settings being the same, then Condition (4.5) fails and Condition (4.6) trivially holds. But the sequence $\left\{ y^k \right\}_{k=0}^{\infty}$, generated by Algorithm 3.2, satisfies for each positive integer k,

$$y^k = 3 \cdot 2^{-1} - \sum_{\ell=1}^{k-1} 2^{-\ell} \underset{k\to\infty}{\to} 2^{-1} \notin D.$$

Taking D in Corollary 4.3 to be a level-set of the function ϕ within C, the next corollary follows.

Corollary 4.6. *Under the assumptions of Lemma 4.1, pick any $x^* \in C$ together with $y^0 \in X$ and define $D := \{ c \in C \mid \phi(c) \leq \phi(x^*) \}$. Choose positive step-sizes $\left\{ \{ \beta_{k,n} \}_{n=0}^{N_k-1} \right\}_{k=0}^{\infty}$ and $\hat{c} \in C$ such that*

$$d\left(\hat{c}, D\right) > \sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n} \tag{4.7}$$

and

$$\left\| y^0 - \hat{c} \right\| < \left(d\left(\hat{c}, D\right) \left(\sum_{k=0}^{\infty} \sum_{n=0}^{N_k-1} \beta_{k,n} \right)^{-1} - 1 \right) \sum_{n=0}^{N_0-1} \beta_{0,n}. \tag{4.8}$$

If y^* is the limit point of a sequence $\left\{ y^k \right\}_{k=0}^{\infty}$ generated by Algorithm 3.2, then $y^* \notin D$, that is, $\phi(y^*) > \phi(x^*)$.

5   Concluding comments

Corollary 4.6 shows the following. If y^* is the limit point of a sequence generated by the superiorization algorithm and if x^* is the limit point of the same algorithm without perturbations, everything else being equal, then under conditions (4.7) and (4.8), the superiorization algorithm fails to reach a feasible point with objective function value smaller or equal than that of the point reached by the feasibility-seeking algorithm used by that superiorization

algorithm. As mentioned above, a reverse of the negative condition will have to be included in any future mathematical guarantee claim for the SM. In future practical implementations of users of the SM it would be advisable to choose the initialization point far enough from the feasible set, so as to avoid the negative condition from occurring.

Data Availability. Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Compliance with Ethical Standards. The authors have no potential conflicts of interest to declare that are relevant to the content of this article.

Competing Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] S. Amaran, S.J. Bury, K. Ma, S. Rajagopalan and N.V. Sahinidis, Decomposition in derivative-free optimization, *Journal of Global Optimization*, **81** (2021), 269–292. 2

[2] F.J. Aragón-Artacho, Y. Censor, A. Gibali and D. Torregrosa-Belén, The superiorization method with restarted perturbations for split minimization problems with an application to radiotherapy treatment planning, *Applied Mathematics and Computation*, **440** (2023), Article 127627. Corrigendum, *Applied Mathematics and Computation*, **476**, Article 128788 (2024). 1

[3] V. Berinde, P. Kumam, N. Pakkaranang and Y.I. Suleiman, Superiorization methodology and perturbation resilience of inertial proximal gradient algorithm with application to signal recovery, *The Journal of Supercomputing*, **76** (2020), 9456–9477. 2

[4] D. Butnariu, R. Davidi, G.T. Herman and I.G. Kazantsev, Stable convergence behavior under summable perturbations of a class of projection methods for convex feasibility and optimization problems, *IEEE Journal of Selected Topics in Signal Processing,* **1** (2007), 540–547. 3.1, 3.2

[5] D. Butnariu, S. Reich and A.J. Zaslavski, Convergence to fixed points of inexact orbits of Bregman-monotone and of nonexpansive operators in Banach spaces, in: H.F. Nathansky, B.G. de Buen, K. Goebel, W.A. Kirk,

and B. Sims (Editors), *Fixed Point Theory and its Applications*, (Conference Proceedings, Guanajuato, Mexico, 2005), Yokahama Publishers, Yokahama, Japan, pp. 11–32, 2006. 3.1

[6] D. Butnariu, S. Reich and A.J. Zaslavski, Stable convergence theorems for infinite products and powers of nonexpansive mappings, *Numerical Functional Analysis and Optimization,* **29** (2008), 304–323. 3.1

[7] R.L.G. Cavalcante, J. Fink, and S. Stanczak, Multi-group multicast beamforming by superiorized projections onto convex sets, *IEEE Transactions on Signal Processing*, **69** (2021), 5708–5722. 2

[8] Y. Censor, Can linear superiorization be useful for linear optimization problems? *Inverse Problems*, **33** (2017), 044006 (22pp). 1

[9] Y. Censor, Superiorization and perturbation resilience of algorithms: A bibliography compiled and continuously updated. http://math.haifa.ac.il/yair/bib-superiorization-censor.html#top. 1, 2, 3.1

[10] Y. Censor, Superiorization: The asymmetric roles of feasibility-seeking and objective function reduction, *Applied Set-Valued Analysis and Optimization*, **5** (2023), 325–346. 1, 3.1

[11] Y. Censor, R. Davidi, E. Garduño and G.T. Herman, Superiorization: An optimization heuristic for medical physics, *Medical Physics,* **39** (2012), 5532–5546. 3.1, 3.2

[12] Y. Censor, R. Davidi, S. Geneser, R.W. Schulte and L. Xing, Feasibility-seeking and superiorization algorithms applied to inverse treatment planning in radiation therapy, *Contemporary Mathematics*, **636** (2015), 83–92. 3.1, 3.2

[13] Y. Censor, R. Davidi and G.T. Herman, Perturbation resilience and superiorization of iterative algorithms, *Inverse Problems,* **26** (2010), 065008. 3.1, 3.2

[14] Y. Censor, R. Davidi, and G.T. Herman, Perturbation-resilient block-iterative projection methods with application to image reconstruction from projections, *International Transactions in Operational Research,* **16** (2009), 505–524. 3.1, 3.2

[15] Y. Censor, R. Davidi, G.T. Herman, R.W. Schulte and L. Tetruashvili, Projected subgradient minimization versus superiorization, *Journal of Optimization Theory and Applications,* **160** (2014), 730–747. 3.2

[16] Y. Censor, G.T. Herman and M. Jiang, Guest Editors, *Superiorization: Theory and Applications*, Special Issue of the journal *Inverse Problems*, Volume **33**, Number 4, April 2017. 3.1

[17] Y. Censor and E. Levy, An analysis of the superiorization method via the principle of concentration of measure, *Applied Mathematics and Optimization*, **83** (2021), 2273–2301. 1

[18] Y. Censor, S.N. Penfold, A.B. Rosenfeld and R.W. Schulte, Total variation superiorization schemes in proton computed tomography image reconstruction, *Medical Physics,* **37** (2010), 5887–5895. 3.1, 3.2

[19] Y. Censor and A.J. Zaslavski, Convergence and perturbation resilience of dynamic string-averaging projection methods, *Computational Optimization and Applications,* **54** (2013), 65–76. 3.1, 3.2, 3.4, 3.2

[20] Y. Censor and A.J. Zaslavski, Strict Fejér monotonicity by superiorization of feasibility-seeking projection methods, *Journal of Optimization Theory and Applications*, **165** (2015), 172–187. 1, 3.2, 3.5, 4

[21] S. Collins, M. Guenter, A. Ogilvy, W. Hare and A. Jirasek, Superiorization versus regularization: A comparison of algorithms for solving image reconstruction problems with applications in computed tomography, *Medical Physics*, **49** (2022), 1065–1082. 2

[22] R. Davidi, G.T. Herman and T. Nikazad, Accelerated perturbation-resilient block-iterative projection methods with application to image reconstruction, *Inverse Problems,* **28** (2012), 035005 (19pp). 3.1, 3.2

[23] E. Garduño and G.T. Herman, Superiorization of the ML-EM algorithm, *IEEE Transactions on Nuclear Science,* **61** (2014), 162–172. 3.1, 3.2

[24] A. Gibali, G.T. Herman and C. Schnörr, Guest Editors, *Superiorization versus Constrained Optimization: Analysis and Applications*, Special Issue of *Journal of Applied and Numerical Optimization*, Volume **2**, Number 1, April 2020. 3.1

[25] G.T. Herman, Problem structures in the theory and practice of superiorization, *Journal of Applied and Numerical Optimization*, **2** (2020) 71–76. 1, 3.1

[26] G.T. Herman, Superiorization for image analysis, in: *Combinatorial Image Analysis*, Lecture Notes in Computer Science Vol. **8466**, Springer, 2014, pp. 1–7. 3.1

[27] G.T. Herman and R. Davidi, Image reconstruction from a small number of projections, *Inverse Problems,* **24** (2008), 045011. 3.1, 3.2

[28] Y. Jia, N. McMichael, P. Mokarzel, B. Thompson, D. Si and T. Humphries, Superiorization-inspired unrolled SART algorithm with U-Net generated perturbations for sparse-view and limited-angle CT reconstruction, *Physics in Medicine & Biology*, **67** (2022), 245004. 2

[29] N. Janakiev, *Superiorized Algorithms for Medical Image Processing: Comparison of Shearlet-based Secondary Optimization Criteria*, Marshall Plan Scholarship Report, Salzburg University of Applied Sciences, Salzburg, Austria, and The Graduate Center of the City University of New York, NY, USA, September 2016. 2

[30] O. Langthaler, *Incorporation of the Superiorization Methodology into Biomedical Imaging Software*, Marshall Plan Scholarship Report, Salzburg University of Applied Sciences, Salzburg, Austria, and The Graduate Center of the City University of New York, NY, USA, September 2014. 2

[31] S. Luo and T. Zhou, Superiorization of EM algorithm and its application in single-photon emission computed tomography (SPECT), *Inverse Problems and Imaging*, **8** (2014), pp. 223–246. 2

[32] S. da Silva Oliveira, *Parallelized Superiorization Method for History Matching Problems Using Seismic Priors and Smoothness in Parts*, Ph.D. Dissertation, Department of Informatics and Applied Mathematics, The Federal University of Rio Grande do Norte, Brazil, 2023. 2

[33] D. Torregrosa-Belén, *Splitting Algorithms for Structured Optimization*, Ph.D. Dissertation, Department of Mathematics, University of Murcia, Spain, 2024. 1