A new theorem of alternatives leading to sufficient conditions for the superiorization guarantee question of Dynamic String-Averaging in the inconsistent case

Kay Barshad* and Yair Censor

Department of Mathematics, University of Haifa, Mt. Carmel, Haifa 3498838, Israel

✉ kaybarshad@gmail.com; ✉ yair@math.haifa.ac.il

November 6, 2025. Revised March 31, 2026. Revised July 13, 2026

Abstract

We study the Superiorization Methodology (SM) in the context of the General Dynamic String-Averaging (GDSA) method in the inconsistent case (that is, where the input operators don't have a common fixed point) which primarily aims at achieving convex feasibility while simultaneously reducing an objective function. In many scientific and real-world problems modeled as constrained minimization tasks, striving for the exact constrained optimum can be costly in terms of time, energy, and resources. Therefore, applying the SM can offer a practical and efficient alternative. In particular, we present a new "theorem of alternatives" for the superiorization method which leads to investigation of theoretical conditions under which the superiorized version of the GDSA algorithm converges to a "superior" feasible point, i.e., one with an objective function value that is smaller or equal to that produced by the unperturbed feasibility-seeking algorithm. While this question has only been partially addressed in the existing literature, we present new sufficient conditions that guarantee that the SM attains such a superior outcome.

($*$) Corresponding author: Kay Barshad, kaybarshad@gmail.com.

1 Introduction

The Superiorization Methodology (SM) embeds objective function reduction steps ("perturbations") into a feasibility-seeking algorithm ("the basic algorithm"), producing a "superiorized version of the basic algorithm". These perturbations aim to reduce the objective function values locally before the next feasibility-seeking iterative step. While the SM has been widely successful in practice across many large-scale applications, a full mathematical guarantee that the superiorized algorithm both converges to a feasible point and achieves a superior objective function value (that is, a smaller or equal value) compared to the unperturbed algorithm, in which all other settings are the same as in the perturbed one, remains unresolved.

Practical applications provide strong empirical evidence that the SM often achieves global objective function reduction. Many studies cited in [18] demonstrate this in real-world applications. Partial theoretical results exist, such as strong Fejér monotonicity (see, for example, [4, Theorem 5.4] and [23, Theorem 4.1]) and concentration of measure arguments ([22]), but counterexamples (see, for instance, [2], [6] and [39]) also show that the SM can fail to deliver the guarantee. The recent study described in [6] yields a necessary condition for the SM to fail. The central open question, therefore, is to identify precise conditions under which the SM is guaranteed to succeed.

The General Dynamic String-Averaging (GDSA) method has been recently introduced in [4], where its convergence properties were studied and it was shown there that the sequence of output operators of this method is Fejér monotone with respect to a certain feasible set. Although the results in [4] apply to the inconsistent case, where the input operators have no common fixed points, they are also valid in the consistent case, where the input operators do have a common fixed point.

In this paper we continue to focus on the inconsistent case, which (like in [4]) requires assumptions on the output operators. Namely, we present a new "theorem of alternatives" for the SM. It states that either the superiorized version of the GDSA algorithm outperforms the feasibility-seeking GDSA algorithm, or, starting from the same initialization point, the sequence of distances between their iterates is strictly decreasing. As a consequence of this theorem, we derive sufficient conditions guaranteeing that the SM algorithm, based on the GDSA scheme with negative (sub)gradient descent perturbations, provides a superior outcome compared to the corresponding unperturbed scheme under the same settings. These sufficient conditions provide insights into effective strategies for choosing the perturbation parameters, which we also demonstrate in our work.

It should be emphasized at the outset that the sufficient conditions presented in Corollaries 3.3 and 3.4 below are primarily theoretical guarantees. In general, they cannot be verified during the execution of the algorithm. Nevertheless, their structural meaning suggests heuristic practical indicators (discussed at the end of Section 3) that may be monitored or encouraged in implementations. These indicators are heuristic in nature and should not be interpreted as verifiable criteria equivalent to the sufficient conditions.

Our results are somewhat restricted in the consistent case, since they leave room for improvement by imposing admissibility assumptions on the input operators rather than on the output ones. We leave a detailed study of the consistent case for future work in a different framework.

From a computational perspective, SM is particularly valuable in large, sparse problems where traditional optimization methods become costly. For instance, the Linear Superiorization (LinSup) approach (see [19]) has been shown to outperform the Simplex algorithm on large-scale linear programs. The continuously updated bibliography in [18] further attests to SM's success across practical, large-scale applications.

The original motivation of the SM remains practical: the desire to reach a feasible point in the intersection of many constraints sets while achieving low-cost improvements of an objective function—not necessarily full optimization, but "satisficing"[1] reductions. Recent developments, such as randomized perturbations and step-size restarts, further enhance the SM's effectiveness without undermining its resilience to perturbations, such randomization and step-size restart strategies, are detailed in [2], [11], [22] and [27]. For more information concerning the SM see, for example,[6], [18], [20], [23], [30] and [39].

Modern studies showcase the practical impact of the SM in various situations where data of constrained optimization problems has to be considered. Guenter et al. in [24] compare the SM with regularization on large, sparse linear systems in image reconstruction. Fink et al. in [14] use bounded perturbation resilient mappings with SM-inspired perturbations for nonconvex multicast beamforming. Pakkaranang et al. in [11] integrate the SM into a modified proximal gradient algorithm for non-smooth composite optimization, which they apply to image recovery problems, achieving strong convergence results. Ma et al. in [1] extend derivative-free optimization algorithms to large-scale problems within an SM-related framework. Numerous other applications are summarized in [18], which include, inter alia, [30], [29], [32], [33] and [37], to mention but a few.

The rest of the paper is organized as follows. Section 2 provides background on the algo-

[1]Support for the reasoning of the SM may be borrowed from the American scientist and Nobel-laureate Herbert Simon (see [38]) who was in favor of 'satisficing' rather than 'maximizing'. Satisficing is a decision-making strategy that aims for a satisfactory or adequate result, rather than the optimal solution. This is because aiming for the optimal solution may necessitate needless expenditure of time, energy and resources. The term 'satisfice' was coined by Simon in 1956, see also: https://en.wikipedia.org/wiki/Satisficing.

rithms and operators under consideration, which is required for establishing our results. In Section 3 our new theorem of alternatives is formulated and proved which enables comparing the behavior of the superiorized algorithm with the unperturbed feasibility-seeking one. This is followed by our sufficient conditions for the guarantee question of the superiorization methodology for Dynamic String-Averaging in the inconsistent case, discussed above.

2 Preliminaries

Throughout this paper $\mathbb{N}$ denotes the set of natural numbers (starting from 0), and for any two integers m and n, with $m \leq n$, we denote by $\{m, m+1, \ldots, n\}$ the set of all integers between m and n. The real line is denoted by $\mathbb{R}$. For a set A, we denote by $|A|$ the cardinality of A. For a real Hilbert space $\mathcal{H}$, we use the following notations:

- $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathcal{H}$.
- $\| \cdot \|$ denotes the norm on $\mathcal{H}$ induced by $\langle \cdot, \cdot \rangle$.
- Id denotes the identity operator on $\mathcal{H}$.
- $\mathrm{Fix}T$ denotes the, possibly empty, set $\mathrm{Fix}T := \{x \in \mathcal{H} \mid T(x) = x\}$ of fixed points of an operator $T : \mathcal{H} \to \mathcal{H}$.
- For a nonempty and convex subset C of $\mathcal{H}$, we denote by P_C the (unique) metric projection onto C, the existence of which is guaranteed if C is, in addition, closed.
- For a convex function $\phi : \mathcal{H} \to \mathbb{R}$ and a point $x \in \mathcal{H}$, we denote by $\partial\phi(x)$ the subdifferential of ϕ at x, that is,

$$\partial\phi(x) := \{g \in \mathcal{H} \mid \langle g, y - x \rangle \leq f(y) - f(x) \text{ for all } y \in \mathcal{H}\}. \tag{2.1}$$

- $B(x, r)$ denotes the open ball centered at $x \in \mathcal{H}$ of radius $r > 0$.
- For a function $f : \mathcal{H} \to \mathbb{R}$ and a subset A of $\mathcal{H}$, we denote by $\underset{x \in A}{\mathrm{Argmin}} f(x)$ and $\underset{x \in A}{\mathrm{Argmax}} f(x)$, respectively, the set of minimizers of f and the set of maximizers of f on the set A.

We recall the following types of algorithmic operators. For more information on such operators, see, for example, [15].

Definition 2.1. Let $T : \mathcal{H} \to \mathcal{H}$ be an operator and let $\lambda \in [0, 2]$. The operator $T_\lambda : \mathcal{H} \to \mathcal{H}$ defined by $T_\lambda := (1 - \lambda) \, Id + \lambda T$ is called a λ-*relaxation* of the operator T. The operator T_2 is called a *reflection* of the operator T.

The operator T_λ in this definition is an averaged mapping in the sense of Section 2 in [3].

Definition 2.2. An operator $T : \mathcal{H} \to \mathcal{H}$ is said to be *nonexpansive* (NE) if

$$(\forall x, y \in \mathcal{H}) \ \ \|T(x) - T(y)\| \leq \|x - y\| .$$

For $\lambda \in [0, 2]$, an operator $T : \mathcal{H} \to \mathcal{H}$ is said to be λ-*relaxed nonexpansive* if T is a λ-relaxation of a nonexpansive operator U, that is, $T = U_\lambda$.

Remark 2.3. Clearly, convex combinations and compositions of nonexpansive operators are also nonexpansive. Moreover, for each $\lambda \in [0, 1]$, the λ-relaxation of a nonexpansive operator is also nonexpansive.

Definition 2.4. We say that an operator $T : \mathcal{H} \to \mathcal{H}$ is *firmly nonexpansive* if

$$(\forall x, y \in \mathcal{H}) \, \langle T(x) - T(y), x - y \rangle \geq \|T(x) - T(y)\|^2 .$$

Theorem 2.5 ([15, Theorem 2.2.10])**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be an operator. The following conditions are equivalent:*

(i) *T is firmly nonexpansive.*

(ii) *T_λ is nonexpansive for each $\lambda \in [0, 2]$.*

(iii) *There exists a nonexpansive operator $N : \mathcal{H} \to \mathcal{H}$ such that $T = 2^{-1} (Id + N)$.*

Theorem 2.5 is essentially Proposition 11.2 on page 42 of the book by Goebel and Reich, [28].

Example 2.6. Given a nonempty, closed an convex subset C of $\mathcal{H}$, the metric projection P_C onto C is firmly nonexpansive (see, e.g., Theorem 2.2.21 in [15]). Moreover $\mathrm{Fix} P_C = C$.

Definition 2.7. For $\lambda \in [0, 2]$, an operator $T : \mathcal{H} \to \mathcal{H}$ is called λ-*relaxed firmly nonexpansive* if T is a λ-relaxation of a firmly nonexpansive operator U, that is, $T = U_\lambda = (1 - \lambda) \, Id + \lambda U$.

Definition 2.8. We say that an operator $T : \mathcal{H} \to \mathcal{H}$ is *quasi-nonexpansive* if

$$\|T(x) - z\| \leq \|x - z\|$$

for each $x \in \mathcal{H}$ and $z \in \mathrm{Fix} T$.

The notions of bounded regularity and approximate shrinking are needed for establishing the strong convergence of the methods which we discuss in this paper. The bounded regularity of a finite family of sets was studied in [8, Section 5] and [7]. This property was extended to an infinite family of sets in [4]. We recall it briefly below.

Definition 2.9. For a nonempty index set I, the family $\{C_i\}_{i\in I}$ of nonempty, closed and convex subsets of $\mathcal{H}$ with nonempty intersection C is *boundedly regular* if for any bounded sequence $\{x^k\}_{k=0}^{\infty}$ in $\mathcal{H}$, the following implication holds:

$$\lim_{k\to\infty} d\left(x^k, C_i\right) = 0 \text{ for each } i \in I \implies \lim_{k\to\infty} d\left(x^k, C\right) = 0.$$

The next proposition provides sufficient conditions for bounded regularity. We recall that a topological space X is locally compact if each $x \in X$ has a compact neighborhood with respect to the topology inherited from X.

Proposition 2.10 ([4, Proposition 3.2])**.** *Let $\{C_i\}_{i\in I}$ be a family of nonempty, closed and convex subsets of $\mathcal{H}$ with a nonempty intersection C. Then the following assertions hold:*

(i) *If there is $i_0 \in I$ for which the set C_{i_0} is a locally compact topological space (with respect to the norm topology inherited from $\mathcal{H}$), then the family $\{C_i\}_{i\in I}$ is boundedly regular.*

(ii) *If $\mathcal{H}$ is of finite dimension, then the family $\{C_i\}_{i\in I}$ is boundedly regular.*

We also recall the following notion of approximate shrinking which was extensively studied in [17]. This notion is also known in the literature as regularity (see, for instance, [10, Definition 7.1] or [16, Definition 3.1]). More information on boundedly regular operators and families of mappings, and their applications, can be found, for example, in Section 3 of [35].

Definition 2.11. A quasi-nonexpansive operator $T : \mathcal{H} \to \mathcal{H}$ *is approximately shrinking* if for each bounded sequence $\{x^k\}_{k=0}^{\infty}$ in $\mathcal{H}$, the following implication holds:

$$\lim_{k\to\infty} \left\|T\left(x^k\right) - x^k\right\| = 0 \implies \lim_{k\to\infty} d\left(x^k, \mathrm{Fix}T\right) = 0.$$

Example 2.12. Given a nonempty, closed an convex subset C of $\mathcal{H}$, the metric projection P_C onto C is approximately shrinking (see Example 3.5 in [17]).

Theorem 2.13 ([9, Theorem 16.7(ii)])**.** *Let $\phi : \mathcal{H} \to \mathbb{R}$ be convex and continuous function at the point $x \in \mathcal{H}$. Then the subgradient set $\partial\phi(x)$ is nonempty.*

Definition 2.14 ([4, Definition 2.24]). Let $\Gamma \subseteq \mathcal{H}$ be a given nonempty subset of $\mathcal{H}$ and $\{T_k\}_{k=0}^{\infty}$ be a sequence of operators, $T_k : \mathcal{H} \to \mathcal{H}$ for each $k \in \mathbb{N}$. The algorithm $x^{k+1} := T_k(x^k)$, for all $k \in \mathbb{N}$, 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 algorithm, converges in the norm of $\mathcal{H}$ to a point in Γ for all $x^0 \in \mathcal{H}$, then any sequence $\{y^k\}_{k=0}^{\infty}$ in $\mathcal{H}$ that is generated by the algorithm $y^{k+1} := T_k(y^k + \beta_k v^k)$, for all $k \in \mathbb{N}$, also converges in the norm of $\mathcal{H}$ to a point in Γ for all $y^0 \in \mathcal{H}$, provided that $\{\beta_k v^k\}_{k=0}^{\infty}$ are bounded perturbations, meaning that $\{\beta_k\}_{k=0}^{\infty}$ is a sequence of positive real numbers such that $\sum_{k=0}^{\infty} \beta_k < \infty$ and that the vector sequence $\{v^k\}_{k=0}^{\infty}$ is a bounded sequence in $\mathcal{H}$.

The notion of perturbation resilience relies on ideas presented in [12], [13] and [26], wherein it was shown that if iterates of a nonexpansive operator converge for any initial point, then its inexact iterates with summable errors also converge. For a more detailed history consult, e.g., [5].

Throughout the rest of the paper we refer, among other things, to the following setting defined below. We recall that for an arbitrary sequence of sets $\{A_k\}_{k=0}^{\infty}$, $\limsup_{k\to\infty} A_k := \cap_{n=0}^{\infty} \cup_{k=n}^{\infty} A_k$.

Let m be a positive integer. We consider a finite family $\{U_i\}_{i=1}^{m}$ of α_i-relaxed firmly nonexpansive operators, where $U_i : \mathcal{H} \to \mathcal{H}$ and $\alpha_i \in (0, 2]$ for each $i = 1, 2, \ldots, m$. Let $\{q_k\}_{k=0}^{\infty}$ be a bounded sequence of positive integers. By Theorem 2.5, the operator U_i is nonexpansive for each $i = 1, 2, \ldots, m$. Let $\mathcal{M} := \max_{k\in\mathbb{N}} q_k$ and define

$$\rho_{\{U_i\}_{i=1}^{m}} := \min\left\{\mathcal{M}^{-1} \min_{i\in\{1,2,\ldots,m\}} (2-\alpha_i)\,\alpha_i^{-1}, 1\right\} \leq 1,$$

where the inequality follows immediately from the definition, since $\rho_{\{U_i\}_{i=1}^{m}}$ is defined as the minimum of two numbers, one of which equals 1. Let $\{\Omega_k\}_{k=0}^{\infty}$ be a family of nonempty sets such that $\Omega_k \subset \{1, 2, \ldots, m\}^{\{1,2,\ldots,q_k\}}$. That is, Ω_k is a finite subset of the set of functions from $\{1, 2, \ldots, q_k\}$ to $\{1, 2, \ldots, m\}$ for each $k \in \mathbb{N}$. Since the sequence $\{q_k\}_{k=0}^{\infty}$ is bounded, the number of different elements in the family $\{\Omega_k\}_{k=0}^{\infty}$ is finite. For each $k \in \mathbb{N}$ and each $t \in \Omega_k$, set $V_k[t] := U_{t(q_k)} \cdots U_{t(2)} U_{t(1)}$ and let $\omega_k : \Omega_k \to (0, 1]$ be a function such that $\sum_{t\in\Omega_k} \omega_k(t) = 1$.

For each $k \in \mathbb{N}$, define $T_{(\Omega_k,\omega_k)} := \sum_{t\in\Omega_k} \omega_k(t)\, V_k[t]$. Define the set $I := \limsup_{k\to\infty} \{T_{(\Omega_k,\omega_k)}\}$, that is, $I = \limsup_{k\to\infty} A_k$, where the sequence of sets $\{A_k\}_{k=0}^{\infty}$ is defined by singletons, $A_k := \{T_{(\Omega_k,\omega_k)}\}$. Define a family of sets $\{C_T\}_{T\in I}$ by $C_T := \mathrm{Fix}T$ for each $T \in I$. Set $F :=$

$\cap_{k=0}^{\infty}\mathrm{Fix}T_{(\Omega_k,\omega_k)}$ and $C := \cap_{T\in I}C_T$ (in case I is empty $\cap_{T\in I}C_T = \mathcal{H}$ by definition). At this stage, no assumption is made on the non-emptiness of the sets I, F and C. Such assumptions are made in the sequel.

Remark 2.15. The operators $T_{(\Omega_k,\omega_k)}$, defined above, are "string-averaging operators" as first introduced in [21] and further studied in various forms and settings, see, for instance, Example 5.21 in [9], [31] and [34], to name but a few. In those and other papers, the index vector t is called "a string", the composite operator $V_k[t]$ is called "a string operator" and ω_k are called "weight functions".

We introduce the following definition of lim sup-admissibility of sequences of operators. The almost cyclic control [15, Definition 5.6.10] is a particular case of the admissible control [25, Definition 3.2] used in this definition, while this admissible control itself is a refinement of the repetitive control [15, Definition 5.6.11]. In particular, limsup-admissibility guarantees repetitive activation of all asymptotically relevant operators, without imposing a uniform window condition.

Definition 2.16 ([4, Definition 4.2]). We say that a sequence $\{T_k\}_{k=0}^{\infty}$of operators, $T_k : \mathcal{H} \to \mathcal{H}$ for each $k \in \mathbb{N}$, is lim sup-*admissible*, if $\{T_k\}_{k=0}^{\infty} \subset \limsup_{k\to\infty}\{T_k\}$ and for each $T \in \limsup_{k\to\infty}\{T_k\}$, there is an integer $M_T > 0$ such that $T \in \cup_{n=k}^{k+M_T-1}\left\{T_{(\Omega_n,w_n)}\right\}$ for all $k \in \mathbb{N}$.

Remark 2.17. Clearly, for each $k_0 \in \mathbb{N}$, $\limsup_{k\to\infty}\{T_k\} = \limsup_{k\to\infty}\{T_{k_0+k}\}$ and if a sequence $\{T_k\}_{k=0}^{\infty}$ of operators is lim sup-admissible, then the sequence $\{T_{k_0+k}\}_{k=0}^{\infty}$ is lim sup-admissible.

Remark 2.18. Observe that if, in the above setting, we require the sequence $\{\omega_k\}_{k=0}^{\infty}$ to attain a finite number of values, then the sequence $\left\{T_{(\Omega_k,\omega_k)}\right\}_{k=0}^{\infty}$ will also contain a finite number of operators. Thus, in order to ensure the existence of $k_0 \in \mathbb{N}$ such that the sequence $\left\{T_{\left(\Omega_{k_0+k},\omega_{k_0+k}\right)}\right\}_{k=0}^{\infty}$ is lim sup-admissible, we only need to require the existence of an integer $M_T > 0$, for each $T \in I$, such that $T \in \cup_{n=k}^{n+M_T-1}\left\{T_{(\Omega_n,w_n)}\right\}$ for all $k \in \mathbb{N}$.

Lemma 2.19 ([4, Lemma 4.7]). *Let $\{\lambda_k\}_{k=0}^{\infty}$ be a sequence of real numbers such that $\lambda_k \in \left[\varepsilon, 1+\rho_{\{U_i\}_{i=1}^m}-\varepsilon\right]$ for each $k \in \mathbb{N}$, where $\varepsilon > 0$. Then the operator $T_{(\Omega_k,\omega_k)\lambda_k}$ is a nonexpansive operator for each $k \in \mathbb{N}$.*

We consider the properties of the following algorithm which is the algorithmic framework for our unperturbed GDSA method.

Algorithm 2.20 ([4, Algorithm 4.5]). *Given $\varepsilon \in (0,1]$, $x^0 \in \mathcal{H}$ and a sequence $\left\{T_{(\Omega_k,\omega_k)}\right\}_{k=0}^{\infty}$ of operators, the algorithm is defined by the recurrence*

$$x^{k+1} := x^k + \lambda_k\left(T_{(\Omega_k,\omega_k)}\left(x^k\right) - x^k\right),$$

where $\lambda_k \in \left[\varepsilon, 1 + \rho_{\{U_i\}_{i=1}^m} - \varepsilon\right]$ *for each* $k \in \mathbb{N}$.

Let $\phi : \mathcal{H} \to \mathbb{R}$ be a convex and continuous real valued objective function. We investigate the following algorithm which provides a general framework for superiorization methods with negative (sub)gradient perturbation. It is the superiorized version of Algorithm 2.20.

Algorithm 2.21 ([4, Algorithm 5.1])**.** *Given* $y^0 \in \mathcal{H}$*, a sequence* $\{N_k\}_{k=0}^\infty$ *of positive integers, a sequence* $\{\lambda_k\}_{k=0}^\infty$ *of positive numbers and a family of positive real sequences* $\left\{\{\beta_{k,n}\}_{n=1}^{N_k}\right\}_{k=0}^\infty$ *such that* $\sum_{k=0}^\infty \sum_{n=1}^{N_k} \beta_{k,n} < \infty$*, the algorithm is defined by the recurrences*

$$y^{k+1} := y^k + \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n} + \lambda_k \left(T_{(\Omega_k,\omega_k)} \left(y^k + \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n} \right) - y^k - \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n} \right)$$

wherein

$$v^{k,n+1} := \begin{cases} -\left\|s^{k,n}\right\|^{-1} s^{k,n}, & \text{if } 0 \notin \partial\phi\left(y^k + \sum_{i=1}^n \beta_{k,i} v^{k,i}\right), \\ 0, & \text{if } 0 \in \partial\phi\left(y^k + \sum_{i=1}^n \beta_{k,i} v^{k,i}\right), \end{cases} \tag{2.2}$$

for each $k \in \mathbb{N}$ *and each* $n = 0, 1, \ldots, N_k - 1$*, where* $s^{k,n}$ *is a selection of the subgradient* $\partial\phi\left(y^k + \sum_{i=1}^n \beta_{k,i} v^{k,i}\right)$ *(which exists by Theorem 2.13) for each* $k \in \mathbb{N}$ *and each* $n = 0, 1, \ldots, N_k - 1$ *(recalling that, by definition,* $\sum_{i=1}^0 \beta_{k,i} v^{k,i} := 0$*).*

Theorem 2.22 ([4, Theorem 4.8])**.** *Assume that* $C \neq \emptyset$ *and that the sequence* $\left\{T_{(\Omega_k,\omega_k)}\right\}_{k=0}^\infty$ *is* $\limsup$*-admissible. Let* $\{\lambda_k\}_{k=0}^\infty$ *be a sequence of real numbers such that* $\lambda_k \in \left[\varepsilon, 1 + \rho_{\{U_i\}_{i=1}^m} - \varepsilon\right]$ *for each* $k \in \mathbb{N}$*, where* $\varepsilon > 0$*, and let* $x^0 = y^0 \in \mathcal{H}$*. Suppose that* $\left\{x^k\right\}_{k=0}^\infty$ *and* $\left\{y^k\right\}_{k=0}^\infty$ *are the sequences generated, respectively, by Algorithm 2.20 and Algorithm 2.21 with respect to the sequence* $\{\lambda_k\}_{k=0}^\infty$*. Then the following assertions hold:*

(i) *The sequence* $\left\{x^k\right\}_{k=0}^\infty$ *converges weakly to a point* $x \in C$*.*

(ii) *If each* $T \in I$ *is approximately shrinking and the family* $\{C_T\}_{T\in I}$ *is boundedly regular, then the convergence is (i) is strong.*

(iii) *The sequence*$\left\{x^k\right\}_{k=0}^\infty$ *is weakly (strongly, if the convergence in (i) is strong) bounded perturbations resilient with respect to* C*.*

Remark 2.23.

(i) In particular, if U_i is a $2\left(\mathcal{M}+1\right)^{-1}$-relaxed firmly nonexpansive, then $\rho_{\{U_i\}_{i=1}^m}=1$ and $\lambda_k\in[\varepsilon,2-\varepsilon]$ for each $k\in\mathbb{N}$. If U_i is a nonexpansive for each $i=1,2,\ldots,m$, then $\rho_{\{U_i\}_{i=1}^m}=0$ and $\lambda_k\in[\varepsilon,1-\varepsilon]$ for each $k\in\mathbb{N}$.

(ii) If the space $\mathcal{H}$ is of a finite dimension, then the convergence in *(i)* is strong.

(iii) Theorem 2.22*(ii)* presents a somewhat analogous result to Theorem 4.1 in [36], where the consistent case was considered, that is input operators were assumed to have a common fixed point. Theorem 2.22*(ii)* applies here to the inconsistent case, where the assumption of a common fixed point of the input operators is replaced by the non-emptiness of the intersection of fixed point sets of the output operators.

The following corollary is a direct consequence of Theorem 2.22.

Corollary 2.24. *Let $y^0\in\mathcal{H}$. Assume that the sequence $\left\{T_{(\Omega_k,\omega_k)}\right\}_{k=0}^{\infty}$ is* lim sup*-admissible and $C\neq\emptyset$. Let $\{\lambda_k\}_{k=0}^{\infty}$ be a sequence of real numbers such that $\lambda_k\in\left[\varepsilon,1+\rho_{\{U_i\}_{i=1}^m}-\varepsilon\right]$ for each $k\in\mathbb{N}$, where $\varepsilon>0$. Then the sequence $\left\{y^k\right\}_{k=0}^{\infty}$ generated by Algorithm 2.21 converges weakly to a point $y\in C$. If, in addition, each $T\in I$ is approximately shrinking and the family $\{C_T\}_{T\in I}$ is boundedly regular, then $\left\{y^k\right\}_{k=0}^{\infty}$ converges strongly to a point $y\in C$.*

In practice, the conditions of bounded regularity and limsup-admissibility are often satisfied in common settings. For instance, bounded regularity holds automatically in finite-dimensional Hilbert spaces. The limsup-admissibility is typically satisfied for sequences of operators generated by periodic or cyclic control schemes, or more generally when the selection of operators avoids pathological accumulation of "inactive" operators over iterations. These observations indicate that the assumptions used in our convergence analysis, while abstract, are applicable in a wide range of practical settings.

3 Sufficient conditions for the guarantee question

The guarantee question in our case is to ascertain under which conditions Algorithm 2.21 performs "better" than Algorithm 2.20, in the sense that the value of the objection function ϕ at the limit point of Algorithm 2.21 is smaller than or equal to the value of ϕ at the limit point of Algorithm 2.20. Below we investigate conditions under which we can establish a mathematical guarantee for this question. We begin by proving the following auxiliary result, which is a modification of Lemma 5.3 in [4].

Lemma 3.1. *For an arbitrary nonempty subset C of $\mathcal{H}$, a convex and continuous function $\phi : \mathcal{H} \to \mathbb{R}$ and $x, y \in \mathcal{H}$ such that $\phi(y) > \phi(x)$, there exist real numbers $r_1 > 0$, $r_2 > 0$ and a neighborhood V of x so that for each $\overline{y} \in B(y, r_1)$ and $v \in \partial\phi(\overline{y})$, the following assertions are satisfied:*

(i) $0 \notin \partial\phi(\overline{y})$ *and for each* $\overline{z} \in B(x, r_2)$,

$$\left\langle \|v\|^{-1} v, \overline{z} - \overline{y} \right\rangle < 0. \tag{3.1}$$

(ii) *For each* $z \in V$,

$$\left\langle \|v\|^{-1} v, z - \overline{y} \right\rangle < -2^{-1} r_2. \tag{3.2}$$

(iii) *Let p be a non-negative integer. Assume that $\{\alpha_n\}_{n=1}^{p}$ is a finite sequence of positive real numbers such that $\sum_{n=1}^{p} a_n < 2^{-1} r_1$ and $\{v^n\}_{n=1}^{p} \subset \mathcal{H} \setminus \{0\}$ is a sequence such that $v^n \in \partial\phi\left(\overline{y} - \sum_{i=1}^{n-1} \alpha_i \left\|v^i\right\|^{-1} v^i\right)$ for each $n = 1, 2, \ldots, p$. If, in addition, $\overline{y} \in B\left(y, 2^{-1} r_1\right)$, then*

$$\left\| \overline{y} - \sum_{n=1}^{p} \alpha_n \|v^n\|^{-1} v^n - z \right\|^2 \le \|\overline{y} - z\|^2 - \sum_{n=1}^{p} (r_2 - \alpha_n)\, \alpha_n \tag{3.3}$$

for each $z \in V$ (by definition $\sum_{n=1}^{0} \alpha_n \|v^n\|^{-1} v^n := \sum_{n=1}^{0} (r_2 - \alpha_n)\, \alpha_n := 0$).

Proof. Since $\phi(y) - \phi(x) > 0$. By the continuity of ϕ, there exist $r_1 > 0$ and $r_2 > 0$ such that

$$\phi(\overline{y}) - \phi(\overline{z}) > 0, \tag{3.4}$$

for each $\overline{y} \in B(y, r_1)$ and $\overline{z} \in B(x, r_2)$. Let $\overline{y} \in B(y, r_1)$ and $v \in \partial\phi(\overline{y})$. Set $V := B\left(x, 2^{-1} r_2\right)$.

(i) In view of (3.4) and (2.1), we have for each $\overline{z} \in B(x, r_2)$,

$$\langle v, \overline{z} - \overline{y} \rangle < 0.$$

It follows that $v \neq 0$ and (3.1) holds. Since v is an arbitrary element of $\partial\phi(\overline{y})$, it follows that $0 \notin \partial\phi(\overline{y})$.

(ii) Let $z \in V$. Define $\overline{z} := z + 2^{-1} r_2 \|v\|^{-1} v$. Then by the triangle inequality, $\overline{z} \in B(x, r_2)$ and by (3.1),

$$\left\langle \|v\|^{-1} v, z + 2^{-1} r_2 \|v\|^{-1} v - \overline{y} \right\rangle = \left\langle \|v\|^{-1} v, \overline{z} - \overline{y} \right\rangle < 0$$

and (3.2) follows.

(iii) Assume that $\overline{y} \in B\left(y, 2^{-1}r_1\right)$. Since $\sum_{n=1}^{p-1} a_n < 2^{-1}r_1$, we obtain

$$\left(\overline{y} - \sum_{n=1}^{p-1} \alpha_n \left\|v^n\right\|^{-1} v^n\right) \in B\left(y, r_1\right). \tag{3.5}$$

It is true that

$$v^p \in \partial\phi\left(\overline{y} - \sum_{n=1}^{p-1} \alpha_n \left\|v^n\right\|^{-1} v^n\right). \tag{3.6}$$

Let $z \in V$. The proof of (3.3) is by induction on p. Clearly, (3.3) is true for $p = 0$. Assume that $p > 0$. Then by the induction hypothesis, (3.5), (3.6) and *(ii)* above,

$$\begin{aligned}
\left\|\overline{y} - \sum_{n=1}^{p} \alpha_n \left\|v^n\right\|^{-1} v^n - z\right\|^2 &= \left\|\overline{y} - \sum_{n=1}^{p-1} \alpha_n \left\|v^n\right\|^{-1} v^n - \alpha_p \left\|v^p\right\|^{-1} v^p - z\right\|^2 \\
&= \left\|\overline{y} - \sum_{n=1}^{p-1} \alpha_n \left\|v^n\right\|^{-1} v^n - z\right\|^2 \\
&+ 2\alpha_p \left\langle \left\|v^p\right\|^{-1} v^p, z - \left(\overline{y} - \sum_{n=1}^{p-1} \alpha_n \left\|v^n\right\|^{-1} v^n\right)\right\rangle + \alpha_p^2 \\
&\le \left\|\overline{y} - z\right\|^2 - \sum_{n=1}^{p-1} \left(r_2 - \alpha_n\right) \alpha_n - \alpha_p r_2 + \alpha_p^2 \\
&= \left\|\overline{y} - z\right\|^2 - \sum_{n=1}^{p} \left(r_2 - \alpha_n\right) \alpha_n.
\end{aligned}$$

Lemma 3.1 is now proved. ■

The following key "theorem of alternatives" provides insights into the behavior of the superiorized algorithm (Algorithm 2.20 above) in comparison with the feasibility-seeking one (Algorithm 2.21 above). This result is a new theorem and differs fundamentally from the theorem of alternatives in [23, Theorem 4.1] which establishes Fejér-monotonicity of the sequence generated by the superiorized algorithm in the consistent case, under the assumption that the algorithm does not outperform the unperturbed one and that only metric projections are employed.

In contrast, Theorem 3.2 below compares, under the above assumption, the distances between the iterates of the sequences produced by the superiorized algorithm with those of the unperturbed algorithm in the inconsistent case, for more general operators $\left\{T_{(\Omega_k,\omega_k)}\right\}_{k=0}^{\infty}$ defined in the previous section. This comparison enables us to derive sufficient conditions for addressing the guarantee question of the superiorization method in the present setting.

Theorem 3.2. *Suppose that $\phi : \mathcal{H} \to \mathbb{R}$ is a convex and continuous function and $\{\lambda_k\}_{k=0}^{\infty} \subset \left[\varepsilon, 1 + \rho_{\{U_i\}_{i=1}^m} - \varepsilon\right]$, where $\varepsilon > 0$, is a sequence of positive numbers. Let $x^0 \in \mathcal{H}$ and $\left\{x^k\right\}_{k=0}^{\infty}$ be a sequence generated by Algorithm 2.20 with respect to the sequence $\{\lambda_k\}_{k=0}^{\infty}$, converging strongly to a point $x \in C$. Assume that $y^0 = x^0$ and that the sequence $\left\{y^k\right\}_{k=0}^{\infty}$, generated by the corresponding Algorithm 2.21, converges strongly to a point $y \in C$. Then exactly one of the following two alternatives holds:*

(i) *$\phi(y) \leq \phi(x)$.*

or

(ii) *$\phi(y) > \phi(x)$ and there exists $k_0 \in \mathbb{N}$ such that $\left\|y^{k+1} - x^{k+1}\right\| < \left\|y^k - x^k\right\|$ for all natural $k \geq k_0$.*

Proof. Assume that $\left\{y^k\right\}_{k=0}^{\infty}$ converges strongly to a point y such that $\phi(y) > \phi(x)$. By Lemma 3.1, there exist real numbers $r_1 > 0$, $r_2 > 0$ and a neighborhood V of x such that each $\overline{y} \in B(y, r_1)$ and $v \in \partial\phi(\overline{y})$ satisfy its assertions. By using the strong convergence of $\left\{y^k\right\}_{k=0}^{\infty}$ to y, the strong convergence of $\left\{x^k\right\}_{k=0}^{\infty}$ to x and the convergence of the series $\sum_{k=0}^{\infty}\sum_{n=1}^{N_k}\beta_{k,n}$, choose $k_0 \in \mathbb{N}$ such that

$$y^k \in B\left(y, 2^{-1}r_1\right) \text{ and } x^k \in V \tag{3.7}$$

and

$$\sum_{n=1}^{N_k}\beta_{k,n} < \min\left\{2^{-1}r_1, r_2\right\} \tag{3.8}$$

for each integer $k \geq k_0$. This yields, for each $k \geq k_0$,

$$y^k + \sum_{i=1}^{n-1}\beta_{k,i}v^{k,i} \in B(y, r_1)$$

for each $n = 1, 2, \ldots, N_k$, and consequently, by Lemma 3.1*(i)*,

$$0 \notin \partial\phi\left(y^k + \sum_{i=1}^{n-1}\beta_{k,i}v^{k,i}\right) \tag{3.9}$$

for each $n = 1, 2, \ldots, N_k$. Let $k \geq k_0$ be an integer. By (2.2) and (3.9),

$$v^{k,n} = -\left\|s^{k,n-1}\right\|^{-1}s^{k,n-1}, \tag{3.10}$$

where

$$s^{k,n-1} \in \partial\phi\left(y^k + \sum_{i=1}^{n-1} \beta_{k,i} v^{k,i}\right) \tag{3.11}$$

for each $n = 1, 2, \ldots, N_k$. Set $p := N_k$, $z := x^k$ and $\overline{y} := y^k$. For each $n = 1, 2, \ldots, p$, define $\alpha_n := \beta_{k,n} > 0$ and $v^n := s^{k,n-1}$. Then, by (3.7), (3.8), (3.9), (3.10) and (3.11), we have that $\overline{y} \in B\left(y, 2^{-1}r_1\right)$, $z \in V$, $\sum_{n=1}^{p} \alpha_n < 2^{-1}r_1$, $\{v^n\}_{n=1}^{p} \subset \mathcal{H}\backslash\{0\}$ and $v^n \in \partial\phi\left(\overline{y} - \sum_{i=1}^{n-1} \alpha_i \left\|v^i\right\|^{-1} v^i\right)$ for each $n = 1, 2, \ldots, p$. Since the operator $T_{(\Omega_k,\omega_k)\lambda_k}$, which is the λ_k-relaxation of $T_{(\Omega_k,\omega_k)}$, is nonexpansive by Lemma 2.19 and $x^{k+1} = T_{(\Omega_k,\omega_k)\lambda_k}(z)$, we obtain from Lemma 3.1*(iii)* that

$$\begin{aligned}
\left\|y^{k+1} - x^{k+1}\right\|^2 &= \left\|T_{(\Omega_k,\omega_k)\lambda_k}\left(y^k + \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n}\right) - T_{(\Omega_k,\omega_k)\lambda_k}(z)\right\|^2 \le \left\|y^k + \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n} - z\right\|^2 \\
&= \left\|\overline{y} - \sum_{n=1}^{p} \alpha_n \left\|v_n\right\|^{-1} v_n - z\right\|^2 \le \left\|\overline{y} - z\right\|^2 - \sum_{n=1}^{p} \left(r_2 - \alpha_n\right)\alpha_n \\
&= \left\|y^k - x^k\right\|^2 - \sum_{n=1}^{N_k} \left(r_2 - \beta_{k,n}\right)\beta_{k,n}.
\end{aligned}$$

Since $\sum_{n=1}^{N_k} \beta_{k,n} < r_2$ by (3.8), the result follows and the proof of the theorem is complete. ■

Based on negation of the alternative (ii) of this new theorem of alternatives, we derive sufficient conditions in the next corollary.

Corollary 3.3. *Let $\{\lambda_k\}_{k=0}^{\infty} \subset \left[\varepsilon, 1 + \rho_{\{U_i\}_{i=1}^m} - \varepsilon\right]$, where $\varepsilon > 0$, be a sequence of positive numbers. Let $x^0 \in \mathcal{H}$ and $\left\{x^k\right\}_{k=0}^{\infty}$ be a sequence generated by Algorithm 2.20 with respect to the sequence $\{\lambda_k\}_{k=0}^{\infty}$ converging strongly to a point $x \in C$. Assume that $y^0 = x^0$ and that the sequence $\left\{y^k\right\}_{k=0}^{\infty}$, generated by the corresponding Algorithm 2.21 converges strongly to a point $y \in C$. Suppose that there exists a strictly increasing sequence $\{\gamma_k\}_{k=0}^{\infty}$ of natural numbers such that*

$$\left\|T_{\left(\Omega_{\gamma_k},\omega_{\gamma_k}\right)\lambda_{\gamma_k}}\left(y^{\gamma_k} + \sum_{n=1}^{N_{\gamma_k}} \beta_{\gamma_k,n} v^{\gamma_k,n}\right) - T_{\left(\Omega_{\gamma_k},\omega_{\gamma_k}\right)\lambda_{\gamma_k}}\left(x^{\gamma_k}\right)\right\| \ge \left\|y^{\gamma_k} - x^{\gamma_k}\right\| \tag{3.12}$$

for each $k \in \mathbb{N}$. Then $\phi(y) \le \phi(x)$.

Proof. Clearly,

$$y^{\gamma_k+1} = T_{(\Omega_{\gamma_k},\omega_{\gamma_k})\lambda_{\gamma_k}}\left(y^{\gamma_k} + \sum_{n=1}^{N_{\gamma_k}} \beta_{\gamma_k,n} v^{\gamma_k,n}\right).$$

The result now follows from Theorem 3.2. ■

The following corollary provides a stronger sufficient condition for the superiorization guarantee question.

Corollary 3.4. *Let $\phi : \mathcal{H} \to \mathbb{R}$ be a convex and continuous function and $\{\lambda_k\}_{k=0}^{\infty} \subset \left[\varepsilon, 1 + \rho_{\{U_i\}_{i=1}^{m}} - \varepsilon\right]$, where $\varepsilon > 0$, be a sequence of positive numbers. Assume that $x^0 \in \mathcal{H}$ and $\left\{x^k\right\}_{k=0}^{\infty}$ is a sequence generated by Algorithm 2.20 with respect to the sequence $\{\lambda_k\}_{k=0}^{\infty}$ converging strongly to a point $x \in C$. Assume that that the sequence $\left\{y^k\right\}_{k=0}^{\infty}$, generated by the corresponding Algorithm 2.21 converges strongly to a point $y \in C$. Suppose that there exists a strictly increasing sequence $\{\gamma_k\}_{k=0}^{\infty}$ of natural numbers so that for each $k \in \mathbb{N}$, the parameters $\left\{\{\beta_{k,n}\}_{n=1}^{N_k}\right\}_{k=0}^{\infty}$ are such that*

$$\begin{aligned}\left\langle y^{\gamma_k} - T_{(\Omega_{\gamma_k},\omega_{\gamma_k})\lambda_{\gamma_k}}\left(x^{\gamma_k}\right), y^{\gamma_k} - T_{(\Omega_{\gamma_k},\omega_{\gamma_k})\lambda_{\gamma_k}}\left(y^{\gamma_k} + \sum_{n=1}^{N_{\gamma_k}} \beta_{\gamma_k,n} v^{\gamma_k,n}\right)\right\rangle \\ \leq \left\langle y^{\gamma_k} - x^{\gamma_k}, x^{\gamma_k} - T_{(\Omega_{\gamma_k},\omega_{\gamma_k})\lambda_{\gamma_k}}\left(x^{\gamma_k}\right)\right\rangle.\end{aligned} \tag{3.13}$$

Then $\phi(y) \leq \phi(x)$.

Proof. By (3.13), we see that

$$\left\langle y^{\gamma_k} - x^{\gamma_k+1}, y^{\gamma_k} - y^{\gamma_k+1}\right\rangle \leq \left\langle y^{\gamma_k} - x^{\gamma_k}, x^{\gamma_k} - x^{\gamma_k+1}\right\rangle \tag{3.14}$$

Now we use the properties of the inner product and (3.14) to deduce

$$\begin{aligned}\left\|y^{\gamma_k+1} - x^{\gamma_k+1}\right\|^2 &= \left\|y^{\gamma_k+1} - y^{\gamma_k}\right\|^2 + \left\|y^{\gamma_k} - x^{\gamma_k+1}\right\|^2 - 2\left\langle y^{\gamma_k+1} - y^{\gamma_k}, x^{\gamma_k+1} - y^{\gamma_k}\right\rangle \\ &= \left\|y^{\gamma_k+1} - y^{\gamma_k}\right\|^2 + \left\|y^{\gamma_k} - x^{\gamma_k}\right\|^2 + \left\|x^{\gamma_k} - x^{\gamma_k+1}\right\|^2 \\ &+ 2\left\langle y^{\gamma_k} - x^{\gamma_k}, x^{\gamma_k} - x^{\gamma_k+1}\right\rangle - 2\left\langle y^{\gamma_k} - x^{\gamma_k+1}, y^{\gamma_k} - y^{\gamma_k+1}\right\rangle \\ &\geq \left\|y^{\gamma_k} - x^{\gamma_k}\right\|^2.\end{aligned}$$

The result now follows from Corollary 3.3. ■

Although the inequalities in Corollaries 3.3 and 3.4 cannot, in general, be verified during the execution of the algorithm, their structural meaning suggests practical indicators that can be monitored or encouraged in implementations.

For example, one may require that each perturbation step increases the distance between the iterations of the superiorized algorithm and the feasibility-seeking one, or that each perturbation increases the angular deviation between the update direction of the superiorized algorithm and the update direction that the feasibility-seeking algorithm would produce when starting from the same iterate. In addition, monitoring feasibility measures along the superiorized iterates compared to the unperturbed algorithm can serve as an informal check that the perturbations guide the iterates toward a limit point with an objective function value that is not larger than that of the limit point of the iterates of the unperturbed algorithm.

These considerations are reflected in the parameter selection strategies described in the sequel, which are designed to promote the qualitative behavior prescribed by the theoretical conditions, even though the conditions themselves are not explicitly verified.

The following examples illustrate concrete settings in which the abstract assumptions in Corollaries 3.3 and 3.4 can be checked explicitly.

Example 3.5 (Simultaneous method)**.** Let $\varepsilon > 0$ be a real number. For all $k \in \mathbb{N}$, set $q_k := 1$, $\Omega_k := \{1,2,\ldots,m\}^{\{1,2,\ldots,q_k\}}$ and $\omega_k := \overline{\omega}$ for a fixed $\overline{\omega} : \{1,2,\ldots,m\}^{\{1,2,\ldots,q_k\}} \to (0,1]$. Clearly, for each $i \in \{1,2,\ldots,m\}$, there is a unique string $t^i \in \{1,2,\ldots,m\}^{\{1,2,\ldots,q_k\}}$ such that $t^i(1) = i$. Hence we can define the mapping $\omega : \{1,2,\ldots,m\} \to (0,1]$ by $\omega_i := \overline{\omega}\left(t^i\right)$ for each $i \in \{1,2,\ldots,m\}$. In this case Algorithm 2.20 with the above provisions provides a fully-simultaneous method, that is,

$$T_{(\Omega_k,\omega_k)} = \sum_{t\in\Omega_k} \overline{\omega}(t)\, U_{t(1)} = \sum_{i=1}^{m} \overline{\omega}\left(t^i\right) U_{t^i(1)} = \sum_{i=1}^{m} \omega_i U_i \tag{3.15}$$

for each $k \in \mathbb{N}$ and $C = F = \operatorname{Fix} \sum_{i=1}^{m} \omega_i U_i$. We see from (3.15) that the sequence $\left\{T_{(\Omega_k,\omega_k)}\right\}_{k=0}^{\infty}$ is lim sup-admissible. Hence, under the assumption that $\operatorname{Fix} \sum_{i=1}^{m} \omega_i U_i \neq \emptyset$ we obtain, by Theorem 2.22, the weak convergence of this fully-simultaneous method, with parameters $\{\lambda_k\}_{k=0}^{\infty} \subset \left[\varepsilon, 1 + \rho_{\{U_i\}_{i=1}^m} - \varepsilon\right]$, to a point in C. In particular, when $U_i = P_{C_i}$, where C_i is a nonempty, closed and convex subset of $\mathcal{H}$ for each $i = 1,2,\ldots,m$, we obtain the well-known simultaneous projection method, see, for example, [15, Subsection 5.4]. In this case $\rho_{\{U_i\}_{i=1}^m} = 1$ (by Example 2.6 and Remark 2.23(a)), $I = \left\{\sum_{i=1}^{m} \omega_i P_{C_i}\right\}$ and

$$C = F = \operatorname{Fix} \sum_{i=1}^{m} \omega_i P_{C_i} = \operatorname*{Argmin}_{x\in\mathcal{H}} f(x), \tag{3.16}$$

where $f:\mathcal{H}\to\mathbb{R}$ is a, so called, proximity function defined by $f:=2^{-1}\sum_{i=1}^{m}\omega_i\left\|P_{C_i}-Id\right\|^2$. For the proof of the last equality in (3.16), see, for instance, Theorem 4.4.6 in [15]. By Theorem 2.22*(i)*, the simultaneous projection method with parameters $\{\lambda_k\}_{k=0}^{\infty}\subset[\varepsilon,2-\varepsilon]$ converges weakly to a point in $\operatorname*{Argmin}_{x\in\mathcal{H}} f(x)$. If, in addition, the operator $\sum_{i=1}^{m}\omega_i P_{C_i}$ is approximately shrinking, then (since the family $\{C_T\}_{T\in I}=\{C\}$ is boundedly regular), this convergence is strong by Theorem 2.22*(ii)*.

Set $\mathcal{H}:=\mathbb{R}^2$ with the usual inner product and $m:=2$. Let $C_1:=\left\{u\in\mathbb{R}^2\mid u_1=1\right\}$ and $C_2:=\left\{u\in\mathbb{R}^2\mid u_1=-1\right\}$ be two parallel hyperplanes. Define $U_i:=P_{C_i}$ for each $i=1,2$, and $\lambda_k:=1$ for each $k\in\mathbb{N}$. Assume that $x^0:=y^0:=\begin{pmatrix}0\\0\end{pmatrix}$ and let $\phi:\mathcal{H}\to\mathbb{R}$ be a convex function defined by $\phi(u):=-u_2$ for each $u\in\mathcal{H}$. Clearly, $C_1\cap C_2=\emptyset$ and the subgradient set of ϕ at each point is $\left\{\begin{pmatrix}0\\-1\end{pmatrix}\right\}$. Let $\left\{\{\beta_{k,n}\}_{n=1}^{N_k}\right\}_{k=0}^{\infty}$ be an arbitrary sequence of positive numbers such that $\sum_{k=0}^{\infty}\sum_{n=1}^{N_k}\beta_{k,n}<\infty$. Set $\overline{\omega}:\{1,2\}^{\{1\}}\to(0,1]$ by $\overline{\omega}(t):=2^{-1}$ for each string $t\in\{1,2\}^{\{1\}}$ and we obtain the function $\omega:\{1,2\}\to(0,1]$, induced by $\overline{\omega}$, which attains a value of 2^{-1} for each $i\in\{1,2\}$. Then

$$T_{(\Omega_k,\omega_k)\lambda_k}=T_{(\Omega_k,\omega_k)}=\sum_{i=1}^{m}\omega_i P_{C_i}=2^{-1}P_{C_1}+2^{-1}P_{C_2}$$

and

$$C=F=\operatorname{Fix}\sum_{i=1}^{2}2^{-1}P_{C_i}=\operatorname*{Argmin}_{x\in\mathcal{H}} f(x).$$

In these settings the sequence $\left\{x^k\right\}_{k=0}^{\infty}$ generated by Algorithm 2.20 satisfies $x^k=\begin{pmatrix}0\\0\end{pmatrix}$ for each $k\in\mathbb{N}$ and the sequence $\left\{y^k\right\}_{k=0}^{\infty}$ generated by Algorithm 2.21 satisfies $y^k=\begin{pmatrix}0\\ \sum_{i=0}^{k-1}\sum_{n=1}^{N_{k-1}}\beta_{i,n}\end{pmatrix}$ for each $k\in\mathbb{N}$ (recalling that by definition, $\sum_{i=0}^{k-1}\sum_{n=1}^{N_{k-1}}\beta_{i,n}=0$ for $k=0$). As a result

$$\left\| \sum_{i=1}^{2} 2^{-1} P_{C_i}\left(y^k + \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n}\right) - \sum_{i=1}^{2} 2^{-1} P_{C_i}\left(x^k\right) \right\| = \left\|y^{k+1} - x^{k+1}\right\| = \sum_{i=0}^{k} \sum_{n=1}^{N_k} \beta_{i,n}$$
$$> \sum_{i=0}^{k-1} \sum_{n=1}^{N_{k-1}} \beta_{i,n} = \left\|y^k - x^k\right\|$$

and

$$\begin{aligned}
&\left\langle y^k - T_{(\Omega_k,\omega_k)\lambda_k}\left(x^k\right), y^k - T_{(\Omega_k,\omega_k)\lambda_k}\left(y^k + \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n}\right)\right\rangle \\
&= \left\langle y^k - x^{k+1}, y^k - y^{k+1}\right\rangle \\
&= \left\langle \begin{pmatrix} 0 \\ \sum_{i=0}^{k-1} \sum_{n=1}^{N_{k-1}} \beta_{i,n} \end{pmatrix}, \begin{pmatrix} 0 \\ -\sum_{n=1}^{N_k} \beta_{k,n} \end{pmatrix} \right\rangle \\
&\quad - \sum_{i=0}^{k-1} \sum_{n=1}^{N_{k-1}} \beta_{i,n} \cdot \sum_{n=1}^{N_k} \beta_{k,n} < 0 \\
&= \left\langle y^k - x^k, x^k - T_{(\Omega_k,\omega_k)\lambda_k}\left(x^k\right)\right\rangle.
\end{aligned}$$

for each $k \in \mathbb{N}$. By setting $\gamma_k := k$ for each $k \in \mathbb{N}$, we see that either of the inequalities which appear in Corollary 3.3 and Corollary 3.4 are satisfied. Therefore, the superiorization algorithm performs at least as well as the unperturbed one. Indeed, denote $L := \sum_{k=0}^{\infty} \sum_{n=1}^{N_k} \beta_{k,n} > 0$. Then $\lim_{k\to\infty} x^k = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\lim_{k\to\infty} y^k = \begin{pmatrix} 0 \\ L \end{pmatrix}$, while

$$\phi\begin{pmatrix} 0 \\ L \end{pmatrix} = -L < 0 = \phi\begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

More generally, in the above settings assume that ϕ is a convex and continuously differentiable function such that $\frac{\partial \phi}{\partial u_2}\left(y^0\right) \neq 0$. By using the continuous differentiability of ϕ, choose the perturbation steps $\{\beta_{k,n}\}_{n=1}^{N_k}$ small enough so that $\sum_{n=1}^{N_k} \beta_{k,n} < k^{-2}$ and

$$\frac{\partial \phi}{\partial u_2}\left(y^k\right) \cdot \frac{\partial \phi}{\partial u_2}\left(\begin{pmatrix} 0 \\ y_2^k + \sum_{i=1}^{n} \beta_{k,i} v_2^{k,i} \end{pmatrix}\right) > 0 \tag{3.17}$$

for each $n = 1, 2, \ldots, N_k$, where $v^{k,n}$ is defined recursively by (2.2), that is,

$$v^{k,i} := -\nabla \phi\left(y^k + \sum_{j=1}^{i-1} \beta_{k,i} v^{k,j}\right) \left\| \nabla \phi\left(y^k + \sum_{j=1}^{i-1} \beta_{k,i} v^{k,j}\right)\right\|^{-1}$$

for each $i = 1, 2, \ldots, n$.

Since for each $k \in \mathbb{N}$, $y_1^k = 0$, we see, in the light of (3.17), that for all $k \in \mathbb{N}$, either $y_2^k \geq 0$ (if $\frac{\partial \phi}{\partial u_2}\left(y^0\right) < 0$) or $y_2^k \leq 0$ (if $\frac{\partial \phi}{\partial u_2}\left(y^0\right) > 0$) and

$$\left\|y^{k+1} - x^{k+1}\right\| = \left\|y^{k+1}\right\| = \left|y_2^k + \sum_{i=1}^{N_k} \beta_{k,i} v_2^{k,i}\right| > \left|y_2^k\right| = \left\|y^k\right\| = \left\|y^k - x^k\right\|.$$

and

$$\left\langle y^k - T_{(\Omega_k,\omega_k)\lambda_k}\left(x^k\right), y^k - T_{(\Omega_k,\omega_k)\lambda_k}\left(y^k + \sum_{n=1}^{N_k} \beta_{k,n} v^{k,n}\right)\right\rangle$$

$$= \left\langle y^k - x^{k+1}, y^k - y^{k+1}\right\rangle = \left\langle \begin{pmatrix} 0 \\ y_2^k \end{pmatrix}, \begin{pmatrix} 0 \\ -\sum_{i=1}^{N_k} \beta_{k,i} v_2^{k,i} \end{pmatrix}\right\rangle \leq 0$$

$$= \left\langle y^k - x^k, x^k - T_{(\Omega_k,\omega_k)\lambda_k}\left(x^k\right)\right\rangle.$$

Again by setting $\gamma_k := k$ for each $k \in \mathbb{N}$, we see that either of the inequalities which appear in Corollary 3.3 and Corollary 3.4 are satisfied. Therefore, the superiorization algorithm performs no worse than the unperturbed one. Observe that the above discussion extends to any finite-dimensional Hilbert space $\mathcal{H}$.

We next discuss how practitioners should test the step-sizes of the perturbations in the SM to follow the sufficient conditions presented here in quest of the superiorization guarantee problem. While Conditions (3.12) and (3.13) provide a rigorous guarantee that the superiorized algorithm performs at least as well as the unperturbed algorithm, they are formulated in terms that are not always directly computable in practice. Nevertheless, these conditions admit a clear structural interpretation: they require that each perturbation step remains compatible with the underlying convergence mechanism of the feasibility-seeking algorithm and not compromise convergence with respect to the objective function value at the limit point.

Motivated by this interpretation, we propose practical strategies for selecting the perturbations that aim to emulate the behavior prescribed by (3.12) and (3.13). Although these strategies do not verify the assumptions explicitly, they are designed so that, on an iteration-by-iteration basis, the superiorized algorithm follows the corresponding theoretical model as closely as possible.

The role of the theoretical conditions (3.12) and (3.13) is, therefore, not merely to certify correctness in special cases, but rather to provide insight into how perturbations should be chosen so as to remain aligned with the convergence behavior of the unperturbed algorithm while promoting improved objective function values. The practical strategies proposed here are steered by this insight.

Assume that $\{\delta_k\}_{k=0}^{\infty}$ and $\{\tau_k\}_{k=0}^{\infty}$ are sequences of positive numbers such that $\sum_{k=1}^{\infty} \delta_k < \infty$ and $\tau_k < \delta_k$ for each $k \in \mathbb{N}$. Choose $x^0 := y^0$ as arbitrary starting points. For each $k \in \mathbb{N}$ such that $k \bmod 2 = 0$, let β_k be a selection of the set

$$\operatorname*{Argmin}_{\beta \in [\tau_k, \delta_k]^{N_k} \mid \sum_{n=1}^{N_k} \beta_n \le \delta_k} \left\| T_{(\Omega_k, \omega_k)\lambda_k} \left(y^k + \sum_{n=1}^{N_k} \beta^n v^{k,n} \right) - T_{(\Omega_k, \omega_k)\lambda_k} \left(x^k \right) \right\|,$$

and for each $k \in \mathbb{N}$ such that $k \bmod 2 = 1$, let β_k be a selection of the set

$$\operatorname*{Argmax}_{\beta \in [\tau_k, \delta_k]^{N_k} \mid \sum_{n=1}^{N_k} \beta_n \le \delta_k} \left\| T_{(\Omega_k, \omega_k)\lambda_k} \left(y^k + \sum_{n=1}^{N_k} \beta^n v^{k,n} \right) - T_{(\Omega_k, \omega_k)\lambda_k} \left(x^k \right) \right\|.$$

Let $\left\{x^k\right\}_{k=0}^{\infty}$ be a sequence generated by Algorithm 2.20 and let $\left\{y^k\right\}_{k=0}^{\infty}$ be a sequence generated by Algorithm 2.21, where $\beta_{k,n} := \beta_k^n$ for each $k \in \mathbb{N}$ and each $n = 1, \ldots, N_k$. In this way, choosing minimal possible values of $\left\|y^k - x^k\right\|$ in odd iterations and maximal possible values of $\left\|y^k - x^k\right\|$ in even ones, we follow the theoretical model described in Corollary 3.3.

Alternatively, in the above settings for each $k \in \mathbb{N}$, we can choose β_k to be a selection of the set

$$\operatorname*{Argmin}_{\beta \in [\tau_k, \delta_k]^{N_k} \mid \sum_{n=1}^{N_k} \beta_n \le \delta_k} \left\langle y^k - T_{(\Omega_k, \omega_k)\lambda_k} \left(x^k \right), y^k - T_{(\Omega_k, \omega_k)\lambda_k} \left(y^k + \sum_{n=1}^{N_k} \beta^n v^{k,n} \right) \right\rangle$$

and generate sequences $\left\{x^k\right\}_{k=0}^{\infty}$ and $\left\{y^k\right\}_{k=0}^{\infty}$ by, respectively, Algorithm 2.20 and Algorithm

2.21, where in Algorithm 2.21 $\beta_{k,n} := \beta_k^n$ for each $k \in \mathbb{N}$ and each $n = 0, 1, \ldots, N_k$. In this way we minimize the left side of (3.13) and thus follow the theoretical model described in Corollary 3.4.

Acknowledgments. We thank Walaa Moursi and Henry Wolkowicz from the Department of Combinatorics and Optimization at the University of Waterloo for some fruitful inspiring discussions in the early stages of this investigation. We gratefully acknowledge the enlightening and constructive comments of the two anonymous referees which helped us improve the paper.

Author Contribution declaration. Both authors Kay Barshad and Yair Censor worked on the research, on developing the ideas and both wrote the manuscript and reviewed it.

Funding. The work of Kay Barshad and Yair Censor is supported by the Cooperation Program in Cancer Research of the German Cancer Research Center (DKFZ) and Israel's Ministry of Innovation, Science and Technology (MOST). The work of Yair Censor was supported also by the U.S. National Institutes of Health Grant Number R01CA266467.

Data Availability. No data was used for the research described in the article.

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

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

[3] Baillon, J.B., Bruck, R.E., Reich, S.: On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. Houston Journal of Mathematics **4**, 1–9 (1977)

[4] Barshad, K., Censor, Y.: General perturbation resilient dynamic string-averaging for inconsistent problems with superiorization. Journal of Optimization Theory and Applications **207, 9** (2025)

[5] Barshad, K., Censor, Y.: Strong convergence, perturbation resilience and superiorization of generalized modular string-averaging with infinitely many input operators. Numerical Algorithms (2026). URL https://arxiv.org/abs/2603.14270. Accepted for publication

[6] Barshad, K., Censor, Y., Moursi, W., Weames, T., Wolkowicz, H.: A necessary condition for the guarantee of the superiorization method. Optimization Letters **19**, 1699–1714 (2025)

[7] Bauschke, H.H., Borwein, J.M.: On the convergence of von Neumann's alternating projection algorithm for two sets. Set-Valued Analysis **1**, 185–212 (1993)

[8] Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Review **38**, 367–426 (1996)

[9] Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, New York, USA (2017)

[10] Bauschke, H.H., Noll, D., Phan, H.M.: Linear and strong convergence of algorithms involving averaged nonexpansive operators. Journal of Mathematical Analysis and Applications **421**, 1–20 (2015)

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

[12] Butnariu, D., Reich, S., Zaslavski, A.J.: Convergence to fixed points of inexact orbits of Bregman-monotone and of nonexpansive operators in Banach spaces. In: Nathansky, H.F., de Buen, B.G., Goebel, K., Kirk, W.A., Sims B. (eds.) Fixed Point Theory and its Applications, Conference Proceedings, Guanajuato, Mexico, 2005, pp. 11–32. Yokahama Publishers, Yokahama, Japan (2006)

[13] Butnariu, D., Reich, S., Zaslavski, A.J.: Asymptotic behavior of inexact orbits for a class of operators in complete metric spaces. Journal of Applied Analysis **13**, 1–11 (2007)

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

[15] Cegielski, A.: Iterative Methods for Fixed Point Problems in Hilbert Spaces. Springer, Berlin, Germany (2012)

[16] Cegielski, A., Reich, S., Zalas, R.: Regular sequences of quasi-nonexpansive operators and their applications. SIAM Journal on Optimization **28**, 1508–1532 (2018)

[17] Cegielski, A., Zalas, R.: Properties of a class of approximately shrinking operators and their applications. Fixed Point Theory **15**, 399–426 (2014)

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

[19] Censor, Y.: Can linear superiorization be useful for linear optimization problems? Inverse Problems **33**, 044,006 (2017)

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

[21] Censor, Y., Elfving, T., Herman, G.T.: Averaging strings of sequential iterations for convex feasibility problems. In: D. Butnariu, Y. Censor, S. Reich (eds.) Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, pp. 101–114. Elsevier Science Publishers, Amsterdam, The Netherlands (2001)

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

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

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

[25] Combettes, P.L.: Hilbertian convex feasibility problem: convergence of projection methods. Appl. Math. Optim. **35**, 311–330 (1997)

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

[27] Ertürk, M., Salkim, A.: Superiorization and bounded perturbation resilience of a gradient projection algorithm solving the convex minimization problem. Optimization Letters **17**, 1957–1978 (2023)

[28] Goebel, K., Reich, S.: Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings. Marcel Dekker, New York and Basel (1984)

[29] Janakiev, N.: 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 (2016)

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

[31] Kong, Y., Pajoohesh, H., Herman, G.T.: String-averaging algorithms for convex feasibility with infinitely many sets. Inverse Problems **35**, 034001 (2019)

[32] Langthaler, O.: 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 (2014)

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

[34] Nikazad, T., Abbasi, M., Mirzapour, M.: Convergence of string-averaging method for a class of operators. Optimization Methods and Software **31**, 1189–1208 (2016)

[35] Reich, S., Salinas, Z.: Metric convergence of infinite products of operators in Hadamard spaces. Journal of Nonlinear and Convex Analysis **18**, 331–345 (2017)

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

[37] da Silva Oliveira, S.: 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)

[38] Simon, H.A.: Rational choice and the structure of the environment. Psychological Review **63**, 129–138 (1956)

[39] Torregrosa-Belén, D.: Splitting Algorithms for Structured Optimization: Theory and Applications. Ph.D. dissertation, Department of Mathematics, University of Murcia, Spain (2024). Dissertation available at: Digitum: Repositorio Institucional de la Universidad de Murcia. https://portalinvestigacion.um.es/documentos/66674326f44d1159d56c317c?lang=en