Filomat 38:27 (2024), 9701–9710
https://doi.org/10.2298/FIL2427701P

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Hybrid optimization models based on S-iteration process

Milena J. Petrović[a], Vladimir Rakočević[b], Dejan Ilić[c]

[a]*University of Priština, Faculty of Sciences and Mathematics, Kosovska Mitrovica, Serbia*
[b]*Serbian Academy of Sciences and Arts, Kneza Mihaila 35, 11000 Belgrade, Serbia*
[c]*University of Niš, Faculty of Sciences and Mathematics, Višegradska 33, 18000 Niš, Serbia*

Abstract. The goal of this research is to set the theoretical frame for developing a new type of hybrid accelerated gradient optimization methods. Based on the S-iteration concept, an accelerated three-term hybrid method and its modification are developed. We prove that the presented iterative processes are well defined. Convergence properties of generated schemes are analysed. Observed characteristics regarding the strong convergence of the presented method and its modification are confirmed.

1. Introduction and preliminaries

Quasi-Newton methods started a positive move in upgrading the calculative sides of Newton's method. These type of methods inspired many authors to generate interesting method classes for solving unconstrained nonlinear optimization problems [3, 7, 8, 11, 17, 18, 31]. Instead of directly computing the Hessian and its inverse of the objective function, in the mentioned quasi-Newton type of methods adequate approximations of these elements are used. With that, the satisfying convergence rate is conserved. Regarding the research presented in this paper, the class of the accelerated gradient descent methods, shortly *AGD methods*, is specially important. The AGD methods use the acceleration parameter in each iteration, often denoted as $\gamma_k, \quad k \in \mathbb{N}$ is the number of iterations, to multiply with the iterative vector direction, i.e. the gradient g_k. Authors in [30] first segregated this class of optimization methods. This separation enabled clearer and easier investigation of this particular sort of iterations.

The general form of the AGD method is

$$x_{k+1} = x_k - \gamma_k^{-1} g_k t_k, \tag{1}$$

where x_{k+1} is the next iterative point, x_k the current one, g_k presents the gradient of the objective function and t_k is the iterative step length value. Parameter γ_k stays for the acceleration parameter of the posed optimization scheme. This crucial factor of the AGD method is basically obtained based on the features of

---

2020 *Mathematics Subject Classification*. Primary 47H10, 47H09,49M37; Secondary 90C30, 90C06, 65K99.
Keywords. fixed-point theorems, contraction-type mappings, Nonlinear programming, unconstrained optimization, line search, quasi-Newton methods, gradient descent methods, convergence rate, hybrid model.
Received: 16 November 2023; Revised: 21 March 2024; Accepted: 20 August 2024
Communicated by Marko Petković
The authors gratefully acknowledge support from the Project supported by Ministry of Science, Technological Development and Innovation of the Republic of Serbia (project no. 451-03-65/2024-03/200124 for Dejan Ilić and project no. 451-03-65/2024-03/200123 for Milena J. Petrović).
Email addresses: milena.petrovic@pr.ac.rs (Milena J. Petrović), vrakoc@ptt.rs (Vladimir Rakočević), ilicde@ptt.rs (Dejan Ilić)

the Quasi-Newton equation, but herein we omit this detailed procedure which can be found for example in [30]. In the AGD schemes, as we previously said, instead of function's Hessian $\nabla^2 f(\xi)$ its approximation is used. Very often, this approximation is given by some adequate scalar matrix and can be written as:

$$\nabla^2 f(\xi) \approx \gamma_{k+1} I, \tag{2}$$

where $\gamma_{k+1} = \gamma(x_k, x_{k+1})$ is the scalar which usually depends on the current and the previous iterative points and that has to be derived. Very often this important element of the AGD scheme is obtained using the first or second order Taylor expansion of the objective function, [4, 19–21, 23, 24, 30]. We list below only several accelerated parameter expressions including numbers of the relavant references:

$$\theta_k^{AGD} = -\frac{t_k g_k^T g_k}{t_k y_k^T g_k}, \qquad ([4])$$

$$\gamma_{k+1}^{ADD} = 2\frac{f(x_{k+1}) - f(x_k) - \alpha_k g_k^T\left(\alpha_k d_k - \gamma_k^{-1} g_k\right)}{\left(\alpha_k d_k - \gamma_k^{-1} g_k\right)^T\left(\alpha_k d_k - \gamma_k^{-1} g_k\right)}, \qquad ([24])$$

$$\gamma_{k+1}^{ADSS} = 2\frac{f(x_{k+1}) - f(x_k) + \left(\alpha_k {\gamma_k}^{-1} + \beta_k\right)\|g_k\|^2}{\left(\alpha_k {\gamma_k}^{-1} + \beta_k\right)^2\|g_k\|^2}, \qquad ([20])$$

$$\gamma_{k+1}^{HADSS} = 2\frac{f(x_{k+1}) - f(x_k) + \alpha\left(t_k \gamma_k^{-1} + p_k\right)\|g_k\|^2}{\alpha^2\left(t_k \gamma_k^{-1} + p_k\right)^2\|g_k\|^2}, \qquad ([21])$$

$$\gamma_{k+1}^{SM} = 2\gamma_k \frac{\gamma_k\left[f(x_{k+1}) - f(x_k)\right] + t_k\|g_k\|^2}{t_k^2\|g_k\|^2}, \qquad ([30])$$

$$\gamma_{k+1}^{HSM} = 2\gamma_k \frac{\gamma_k\left[f(x_{k+1}) - f(x_k)\right] + (\alpha_k + 1)t_k\|g_k\|^2}{(\alpha_k + 1)^2 t_k^2\|g_k\|^2}, \qquad ([23]).$$

The importance of accelerated factor of a certain AGD method is specially emphasised in [24]. In this paper, the authors generated the non-accelerated version of the AGD method they have presented. They did the comparative analysis and numerically confirmed significantly better performance profiles in favour of the accelerated version.

Another two, equally important, elements of the AGD schemes are 1. iterative search direction vector and 2. the value of the iterative step length. In the AGD schemes we chose the search vector to be the negative gradient direction i.e. $-g_k$. Step-size parameter is obtained mostly by applying some of the inexact line search procedures [5, 9, 13, 16, 26, 27, 29, 33–35] since these techniques are satisfyingly efficient and significantly less expensive, when the computational time is considered, then the exact line search procedure. In this paper we use Armijo's line search algorithm, so called *Backtracking*.

This paper is organized as such: in Section 2 the three-term $s-$ iterative process is used for determination of a new hybrid optimization model. In Section 3 the modification of the defined hybrid model is presented. Finely, the convergence properties of derived hybrid accelerated schemes are examined.

2. Hybrid correction of *SM* method based on *s*− iteration

The iterative processes of Picard, Mann and Ishikawa, which are, for real sequences $\{\alpha_k\}, \{\beta_k\} \in (0, 1)$, respectively listed below

$$\begin{aligned} &u_1 = u \in \mathbb{C}, \\ &u_{k+1} = Tu_k, \quad k \in \mathbb{N}, \qquad ([25]) \end{aligned}$$

$$\begin{aligned} &v_1 = v \in \mathbb{C}, \\ &v_{k+1} = (1 - \alpha_k)v_k + \alpha_k T v_k, \quad k \in \mathbb{N},, \qquad ([14]) \end{aligned}$$

$$\begin{aligned} &z_1 = z \in \mathbb{C}, \\ &z_{k+1} = (1 - \alpha_k)z_k + \alpha_k T y_k, \\ &y_k = (1 - \beta_k)z_k + \beta_k T z_k, \quad k \in \mathbb{N}, \qquad ([10]) \end{aligned}$$

motivated many authors in their researches [2, 6, 12]. The author in [12] presented a three-term iteration which overcomes those three. This process is defined as

$$\left\{\begin{array}{l} x_1 = x \in \mathbb{R}, \\ x_{k+1} = Ty_k, \\ y_k = (1-\alpha_k)x_k + \alpha_k T x_k, \quad k \in \mathbb{N} \end{array}\right\} \tag{3}$$

or, if written in aggregated form, hybrid Kahn's iterative scheme is

$$x_{k+1} = T[(1-\alpha_k)x_k + \alpha_k T x_k], \quad k \in \mathbb{N}. \tag{4}$$

In both of the last two relations (3) and (4), the operator $T : \mathbb{C} \to \mathbb{C}$ is a mapping defined on nonempty convex subset C of a normed space $\mathbb{E}$, while iterative outcomes x_k and y_k are the sequences derived by the relations (3). The sequence of positive numbers $\{\alpha_k\} \in (0, 1)$ is denoted as *correction parameter*. Numerical test results detected significantly better efficiency metrics than the metrics achieved by forerunners: Picard's, Mann's and Ishikawa's processes.

Iteration (3), i.e. (4), inspired some authors in developing new optimization hybrid models [21–23]. Taking concrete AGD method as operator T in (3) contributed with several efficient hybrid minimization methods. In this paper we present the development of one such model based on $S-$ iteration described in [1].

The S-iteration is defined on convex subset $\mathbb{C}$ of a linear space X. Taking a mapping $T : \mathbb{C} \to \mathbb{C}$, S-iteration is generated as a three term process in the following way

$$\begin{array}{l} x_1 = x \in \mathbb{C}, \\ x_{n+1} = (1-\alpha_n)Tx_n + \alpha_n T y_n, \\ y_n = (1-\beta_n)x_n + \beta_n T x_n \qquad n \in \mathbb{N}. \end{array} \tag{5}$$

In (5), $\{\alpha_n\}$ and $\{\beta_n\}$ are the sequences of real numbers satisfying

$$\begin{array}{l} \{\alpha_n\}, \{\beta_n\} \subset (0, 1) \\ \sum_{n=1}^{\infty} \alpha_n \beta_n (1-\beta_n) = \infty. \end{array} \tag{6}$$

Specially, by taking $\alpha_n + \beta_n = 1$, we reduce conditions (6) to conditions (7)

$$\begin{array}{l} \{\alpha_n\} \subset (0, 1) \\ \sum_{n=1}^{\infty} \alpha_n^2 (1-\alpha_n) = \infty, \end{array} \tag{7}$$

and define a simpler S-iteration version involving only one sequence α_n of real numbers:

$$\begin{array}{l} x_1 = x \in \mathbb{R}, \\ x_{n+1} = (1-\alpha_n)Tx_n + \alpha_n T y_n, \\ y_n = \alpha_n x_n + (1-\alpha_n)Tx_n \qquad n \in \mathbb{N}. \end{array} \tag{8}$$

Similar to [23], we form the hybrid accelerated model by applying the AGD iteration from [30], so called *SM* method, as operator T, i.e.

$$Ty_k = y_k - \gamma_k^{-1} t_k g_k.$$

This way, (8) becomes

$$\begin{array}{ll} x_1 = x \in \mathbb{R}, & \\ x_{n+1} = (1-\alpha_n)Tx_n + \alpha_n T y_n = (1-\alpha_n)(x_n - \gamma_n^{-1} t_n g_n) + \alpha_n (y_n - \gamma_n^{-1} t_n g_n), & \\ y_n = \alpha_n x_n + (1-\alpha_n)(x_n - \gamma_n^{-1} t_n g_n), & n \in \mathbb{N}. \end{array} \tag{9}$$

Proposition 2.1. *The three-term iterative process* (9) *is equivalent to the accelerated gradient descent scheme*

$$x_{n+1} = x_n - (1 + \alpha_n - \alpha_n^2)\gamma_n^{-1} t_n g_n. \tag{10}$$

Remark 2.1. *We named the process* (10) ***SHSM*** *method.*

Before we prove the Proposition 2.1. we first derive the SHSM iterative value of the accelerated parameter. To achieve this goal we will apply a common approach used in [30]. Taking the second order Taylor expansion of the objective function on which the rule (10) is applied on, we get

$$f(x_{k+1}) \approx f(x_k) - \left(1 + \alpha_k - \alpha_k^2\right) t_k g_k^T \gamma_k^{-1} g_k + \frac{1}{2}\left(1 + \alpha_k - \alpha_k^2\right)^2 t_k^2 (\gamma_k^{-1} g_k)^T \nabla^2 f(\xi) \gamma_k^{-1} g_k, \tag{11}$$

Variable ξ is described as:

$$\xi \in [x_k, x_{k+1}], \quad \xi = x_k + \beta(x_{k+1} - x_k) = x_k - \beta\left(1 + \alpha_k - \alpha_k^2\right) t_k \gamma_k^{-1} g_k, \quad 0 \leq \beta \leq 1.$$

After applying the diagonal scalar matrix approximation of the function's Hessian $\nabla^2 f(\xi) \approx \gamma_{k+1} I$, relation (11) becomes

$$f(x_{k+1}) \approx f(x_k) - \left(1 + \alpha_k - \alpha_k^2\right) t_k \gamma_k^{-1} \|\mathbf{g}_k\|^2 + \frac{1}{2}\left(1 + \alpha_k - \alpha_k^2\right)^2 t_k^2 \gamma_k^{-2} \gamma_{k+1} \|\mathbf{g}_k\|^2. \tag{12}$$

From (12) we easily determine the accelerated parameter of the SHSM process

$$\gamma_{k+1}^{SHSM} = 2\gamma_k \frac{\gamma_k \left(f(x_{k+1}) - f(x_k)\right) + \left(1 + \alpha_k - \alpha_k^2\right) t_k \|g_k\|^2}{\left(1 + \alpha_k - \alpha_k^2\right)^2 t_k^2 \|g_k\|^2}. \tag{13}$$

The positiveness is the necessary condition that the accelerated parameter γ_{k+1}^{SHSM} has to satisfy. Otherwise, the Second-Order Necessary and Sufficient Conditions would not be fulfilled. In case $\gamma_{k+1}^{SHSM} < 0$ we simply put $\gamma_{k+1}^{SHSM} = 1$. This way the next iterative point is calculated by

$$x_{k+2} = x_{k+1} - \left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right) t_{k+1} g_{k+1}.$$

Since $\{\alpha_n\} \subset (0, 1)$ we have $1 + \alpha_{k+1} - \alpha_{k+1}^2 > 0$ and $0 < t_{k+1} < 1$, so the previous scheme presents a classical gradient descent iteration.

We now expose the proof of the Proposition 2.1.

Proof. [Proposition 2.1.] After substituting y_n into the expression that defines x_{n+1}, i.e. third equation of (9) into the second one, we get

$$\begin{aligned} x_{n+1} &= (1 - \alpha_n)(x_n - \gamma_n^{-1} t_n g_n) + \alpha_n[\alpha_n x_n + (1 - \alpha_n)(x_n - \gamma_n^{-1} t_n g_n) - \gamma_n^{-1} t_n g_n] \\ &= (1 - \alpha_n)(x_n - \gamma_n^{-1} t_n g_n) + \alpha_n^2 x_n + \alpha_n(1 - \alpha_n)(x_n - \gamma_n^{-1} t_n g_n) - \alpha_n \gamma_n^{-1} t_n g_n \\ &= x_n - \gamma_n^{-1} t_n g_n + \alpha_n^2 \gamma_n^{-1} t_n g_n - \alpha_n \gamma_n^{-1} t_n g_n \\ &= x_n - (1 + \alpha_n - \alpha_n^2)\gamma_n^{-1} t_n g_n. \end{aligned}$$

Knowing $\{\alpha_n\} \subset (0, 1)$, we easily conclude that

$$1 + \alpha_n - \alpha_n^2 > 1. \tag{14}$$

Estimation (14) in conjunction with the fact that $\gamma_k^{-1} > 0$ is the k-th iterative acceleration parameter proves that the iteration (10) is an accelerated gradient descent scheme. □

2.1. Optimal initial step size parameter of the SHSM process

We chose to derive the iterative step length value t_n in SHSM method (10) based on the inexact Backtracking procedure:

Algorithm 2.1. *The backtracking with starrting value* $t = 1$

Require: *Objective function* $f(x)$, *the direction* d_k *of the search at the point* x_k *and numbers* $0 < \sigma < 0.5$ *and* $\beta \in (0,1)$.

1: $t = 1$.

2: *While* $f(x_k + td_k) > f(x_k) + \sigma t g_k^T d_k$, *take* $t := t\beta$.

3: *Return* $t_k = t$.

However, we improve the Backtracking initial value $t = 1$ in a similar way as described in [19]. This initial correction of the step length parameter is described in the following lemma.

Proposition 2.2. *Optimal initial step length value of the Backtracking line search algorithm applied in the iteration* (10) *is*

$$t = \frac{1}{1 + \alpha_k - \alpha_k^2}. \tag{15}$$

Proof. Suppose $\gamma_{k+1} = \gamma_{k+2}$ and under this assumption consider the merit function

$$\Phi_{k+1}(t) = f(x_{k+1}) - \left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right) t\gamma_{k+1}^{-1}\|\mathbf{g}_{k+1}\|^2 + \frac{1}{2}\left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right)^2 t^2\gamma_{k+1}^{-2}\gamma_{k+1}\|\mathbf{g}_{k+1}\|^2.$$

Obviously, this quadratic function is convex when $\gamma_{k+1} > 0$ and the value of its gradient $\nabla(\Phi_{k+1}(t)) = \left\{\Phi_{k+1}(t)'_t\right\}$ is

$$\begin{aligned}(\Phi_{k+1})'_t &= -\left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right)\gamma_{k+1}^{-1}\|\mathbf{g}_{k+1}\|^2 + \left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right)^2 t\gamma_{k+1}^{-2}\gamma_{k+1}\|\mathbf{g}_{k+1}\|^2\\ &= \left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right)\gamma_{k+1}^{-1}\|\mathbf{g}_{k+1}\|^2\left(\left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right)t - 1\right).\end{aligned}$$

Supposing $\gamma_{k+1} > 0$, the function $\Phi_{k+1}(t)$ decreases, i.e. $\left\{\Phi_{k+1}(t)'_t\right\} < 0$, when

$$\left(1 + \alpha_{k+1} - \alpha_{k+1}^2\right)t - 1 < 0,$$

which leads to the final conclusion

$$\{\Phi_{k+1}(t)'_t\} < 0 \Leftrightarrow t < \frac{1}{1 + \alpha_{k+1} - \alpha_{k+1}^2}$$

and

$$\nabla(\Phi_{k+1}(t)) = \{0\} \Leftrightarrow t = \frac{1}{1 + \alpha_{k+1} - \alpha_{k+1}^2}.$$

□

According to the Proposition 2.2. we modify the first step in Algorithm (2.1) with $t = \frac{1}{1+\alpha_{k+1}-\alpha_{k+1}^2}$.

Algorithm 2.2. *The backtracking with starrting value* $t = \frac{1}{1+\alpha_{k+1}-\alpha_{k+1}^2}$

Require: *Objective function* $f(x)$, *the direction* d_k *of the search at the point* x_k *and numbers* $0 < \sigma < 0.5$ *and* $\beta \in (0,1)$.

1: $t = \frac{1}{1+\alpha_{k+1}-\alpha_{k+1}^2}$.

2: *While* $f(x_k + td_k) > f(x_k) + \sigma t g_k^T d_k$, *take* $t := t\beta$.

3: *Return* $t_k = t$.

2.2. The algorithm of the SHSM process defined by (10)

After analysis we have provided previously, we are now able to present the algorithm of the SHSM optimization process.

Algorithm 2.3. *The SHSM algorithm defined by* (10) *and* (13).

Require: *Function* $f(x)$, $\{\alpha_n\} \subset (0,1)$ *defined by* (7), *initial point* $x_0 \in dom(f)$.

1: *Set* $k = 0$ *and calculate* $f(x_0)$, $g_0 = \nabla f(x_0)$, *set* $\gamma_0 = 1$.
2: *Check the test criteria; if stopping criteria are fulfilled then stop the algorithm; otherwise, go to the next step.*
3: *Applying Algorithm 2.2: Compute the value of step size* $t_k \in (0,1]$ *taking* $d_k = -\gamma_k^{-1} g_k$.
4: *Determine* $x_{k+1} = x_n - (1 + \alpha_n - \alpha_n^2)\gamma_n^{-1} t_n g_n$, $f(x_{k+1})$ *and* $g_{k+1} = \nabla f(x_{k+1})$.
5: *Compute* γ_{k+1}^{shsm}, *approximation of the Hessian of function* f *at the point* x_{k+1} *using* (13).
6: *If* $\gamma_{k+1}^{shsm} < 0$ *take* $\gamma_{k+1}^{shsm} = 1$.
7: $k := k + 1$, *go to the step 2.*
8: *Return* x_{k+1} *and* $f(x_{k+1})$.

3. Modified s-hybrid correction of the SM process

Before we evaluate the modified *SHSM* iterative form based on *SM* operator, we first prove that the $n-$th degree of operator $T \equiv SM$ has the following expression

$$T^n x_n = x_n - n\gamma_n^{-1} t_n g_n. \tag{16}$$

Proposition 3.1. *$n-$th degree of the SM$-$ operator is expressed by* (16).

Proof. Using the principle of mathematical induction, we know that for

$$k = 1: \quad T^1 x_n = x_n - 1 \cdot \gamma_n^{-1} t_n g_n.$$

Assuming that the statement holds for $k = n$, i.e.

$$T^n x_n = x_n - n \cdot \gamma_n^{-1} t_n g_n,$$

we are proving that it is valid for $k = n + 1$, as well:

$$\begin{aligned} T^{(n+1)} x_n &= T(T^n x_n) = T(x_n - n \cdot \gamma_n^{-1} t_n g_n) \\ &= x_n - n\gamma_n^{-1} t_n g_n - \gamma_n^{-1} t_n g_n \\ &= x_n - (n+1)\gamma_n^{-1} t_n g_n, \end{aligned}$$

which proves (16). □

Now we define the modified *SHSM* iteration process $\{x_n\}$ as following:
for $x = x_1 \in \mathbb{R}^n$

$$x_{n+1} = S(x_n, \alpha_n, T^n), \quad n \in \mathbb{N}, \quad \{\alpha_n\} \in (0,1). \tag{17}$$

In three-term notation, this modified hybrid process can be displayed as:

$$\begin{aligned} &x_1 = x \in \mathbb{C}, \\ &x_{n+1} = (1 - \alpha_n) T^n x_n + \alpha_n T^n y_n, \\ &y_n = \alpha_n x_n + (1 - \alpha_n) T^n x_n \qquad n \in \mathbb{N}. \end{aligned} \tag{18}$$

Similar to Proposition 2.1. we defined the aggregated form of the modified SHSM method.

Proposition 3.2. *Three-term iterative process* (18) *is equivalent to accelerated gradient descent scheme*

$$x_{n+1} = x_n - (1 + \alpha_n - \alpha_n^2) n\gamma_n^{-1} t_n g_n. \tag{19}$$

Proof. As exposed in the proof of Proposition 2.1. we substitute y_n into the expression that defines x_{n+1}, i.e. third equation of (18) into the second one, and there we have

$$\begin{aligned} x_{n+1} &= (1-\alpha_n)(x_n - n\gamma_n^{-1}t_ng_n) + \alpha_n(y_n - n\gamma_n^{-1}t_ng_n) \\ &= (1-\alpha_n)(x_n - n\gamma_n^{-1}t_ng_n) + \alpha_n[\alpha_nx_n + (1-\alpha_n)(x_n - n\gamma_n^{-1}t_ng_n) - n\gamma_n^{-1}t_ng_n] \\ &= (1-\alpha_n)(x_n - n\gamma_n^{-1}t_ng_n) + \alpha_n[\alpha_nx_n + x_n - n\gamma_n^{-1}t_ng_n - \alpha_nx_n + n\alpha_n\gamma_n^{-1}t_ng_n - n\gamma_n^{-1}t_ng_n] \\ &= x_n - n\gamma_n^{-1}t_ng_n - \alpha_nx_n + n\alpha_n\gamma_n^{-1}t_ng_n + \alpha_nx_n - n\alpha_n\gamma_n^{-1}t_ng_n + n\alpha_n^2\gamma_n^{-1}t_ng_n - n\alpha_n\gamma_n^{-1}t_ng_n \\ &= x_n - (1+\alpha_n-\alpha_n^2)n\gamma_n^{-1}t_ng_n. \end{aligned}$$

The sequence $\{\alpha_n\} \subset (0,1)$, which implies that $1+\alpha_n-\alpha_n^2 > 1$. With that, the variable $\gamma_k^{-1} > 0$ stays for the iterative acceleration parameter of the iteration (19). All stated confirms that the (19) is an accelerated gradient descent process. □

Remark 3.1. *We named the process* (19) ***modified SHSM*** *method.*

Proposition 3.3. 1. *Accelerated parameter of the modified SHSM iteration is given by the following expression*

$$\gamma_{k+1}^{modSHSM} = 2\gamma_k \frac{\gamma_k\left(f(x_{k+1}) - f(x_k)\right) + \left(1+\alpha_k-\alpha_k^2\right)nt_k\|g_k\|^2}{\left(1+\alpha_k-\alpha_k^2\right)^2 n^2t_k^2\|g_k\|^2}; \tag{20}$$

2. *Optimal initial Backtracking step length value of the modified SHSM process is given by*

$$t = \frac{1}{n\left(1+\alpha_k-\alpha_k^2\right)}. \tag{21}$$

Proof. Statement 1. can be proved analogously as in Proposition 2.2. The procedure that confirms statement 2. is similar to the relevant analysis of deriving the accelerated parameter exposed in Section 2. □

For the end of this section we display the algorithm of the modified SHSM method.

Algorithm 3.1. *The backtracking with starrting value* $t = \frac{1}{n\left(1+\alpha_k-\alpha_k^2\right)}$

Require: *Objective function $f(x)$, the direction d_k of the search at the point x_k and numbers $0<\sigma<0.5$ and $\beta \in (0,1)$.*
1: $t = \frac{1}{n\left(1+\alpha_k-\alpha_k^2\right)}$.
2: *While $f(x_k+td_k) > f(x_k) + \sigma t g_k^T d_k$, take $t := t\beta$.*
3: *Return $t_k = t$.*

Algorithm 3.2. *The modified SHSM algorithm defined by* (19) *and* (20).

Require: *Function $f(x)$, $\{\alpha_n\} \subset (0,1)$ defined by* (7)*, initial point $x_0 \in dom(f)$.*
1: *Set $k=0$ and calculate $f(x_0)$, $g_0 = \nabla f(x_0)$, set $\gamma_0 = 1$.*
2: *Check the test criteria; if stopping criteria are fulfilled then stop the algorithm; otherwise, go to the next step.*
3: *Applying Algorithm 3.1: Compute the value of step size $t_k \in (0,1]$ taking $d_k = -\gamma_k^{-1}g_k$.*
4: *Determine $x_{k+1} = x_n - (1+\alpha_n-\alpha_n^2)n\gamma_n^{-1}t_ng_n$, $f(x_{k+1})$ and $g_{k+1} = \nabla f(x_{k+1})$.*
5: *Compute γ_{k+1}, approximation of the Hessian of function f at the point x_{k+1} using* (20).
6: *If $\gamma_{k+1}^{shsm} < 0$ take $\gamma_{k+1}^{shsm} = 1$.*
7: *$k := k+1$, go to the step 2.*
8: *Return x_{k+1} and $f(x_{k+1})$.*

4. Convergence properties of the SHSM processes

In [28] authors introduced the class of nearly Lipschitzian mappings as a generalization of Lipschitzian mappings and later studied this class in [1]. Further we list some relevant definitions from [1] needed for forthcoming statements regarding the *SHSM* process.

Definition 4.1. *If the C is a nonempty subset of a Banach space and $\{a_n\}$ is a sequence in $[0, \infty)$ such that $a_n \to 0$, than a mapping $T : C \to C$ is said to be nearly Lipschitzian, with respact to $\{a_n\}$, if for each $n \in \mathbb{N}$, there exists a constant $k_n \geq 0$ such that for all $x, y \in C$*

$$\|T^n x - T^n y\| \leq k_n(\|x - y\| + a_n). \tag{22}$$

If we denote by $\eta(T^n)$ the infimum of constants k_n for which (22) holds, then for a nearly Lipschitzian mapping T with the sequence $\{(a_n,)\eta(T^n)\}$ we said to be

1. *nearly nonexpansive if $\eta(T^n) = 1$ for $n \in \mathbb{N}$;*
2. *nearly asymptotically nonexpansive if $\eta(T^n) \geq 1$ for $n \in \mathbb{N}$ and $\lim_{n\to\infty} \eta(T^n) = 1$;*
3. *nearly uniformly $k-$Lipschitzian if $\eta(T^n) \leq k$ for $n \in \mathbb{N}$;*
4. *nearly uniform $k-$contraction if $\eta(T^n) \leq k < 1$ for $n \in \mathbb{N}$.*

Remark 4.1. *Number $\eta(T^n)$ is called* nearly Lipschitzian constat*.*

Definition 4.2. *Mapping $T : C \to C$, where C is a nonempty subset of a Banach space X, is said to be asymptotic $k-$contraction mapping with sequence $\{a_n\}$ if*

$$\|T^n x - T^n y\| \leq (k + a_n)\|x - y\| + a_n.$$

In [1] authors concluded that if a mapping is a contraction then it is an asymptotic $k-$contraction. With that, asymptotic $k-$contraction implies uniformly $k-$contraction.

Next theorem confirms that iterative process (19) converges strongly. The motivation for this result arose from Theorem (3.7) in [1]. and from [15].

Theorem 4.1. *Assume that $\{x_n\}$ is a sequence generated by (19) and x^* is a unique minimizer of the modified SHMS process. Then the next statements are valid:*

(a) *$\|x_{n+1} - x^*\| \leq k\|x_n - x^*\| + k(k + 1)a_n$, for all $n \in \mathbb{N}$, where $\{a_n\}$ is a sequence such that $\sum_{n=1}^{\infty} a_n < \infty$*

(b) *sequence $\{x_n\}$ converges strongly to x^*.*

Proof. According to results from [1], since the modified SHSM process is a contraction it is a nearly $k-$contraction as well. So the following estimations hold for $p \in F(T) = \{x \mid Tx = x\} \neq \emptyset$

$$\begin{aligned}
\|x_{n+1} - p\| &\leq (1 - \alpha_n)\|T^n x_n - p\| + \alpha_n\|T^n y_n - p\| \\
&\leq (1 - \alpha_n)k(\|x_n - p\| + a_n) + \alpha_n k(\|y_n - p\| + a_n) \\
&= k[(1 - \alpha_n)\|x_n - p\| + \alpha_n\|y_n - p\| + a_n] \\
&\leq k[(1 - \alpha_n)\|x_n - p\| + \alpha_n\|[\alpha_n\|x_n - p\| + k(1 - \alpha_n)(\|x_n - p\| + a_n)] + a_n] \\
&= k[(1 - \alpha_n)\|x_n - p\| + \alpha_n^2\|x_n - p\| + k(1 - \alpha_n)\alpha_n\|x_n - p\| + (1 - \alpha_n)\alpha_n k a_n + \alpha_n a_n] \\
&= k[1 - \alpha_n + \alpha_n^2 + k\alpha_n - k\alpha_n^2]\|x_n - p\| + k[\alpha_n a_n k - \alpha_n^2 a_n k + \alpha_n a_n] \\
&= k[(1 - \alpha_n)(1 + \alpha_n k) + \alpha_n^2]\|x_n - p\| + k[\alpha_n a_n k - \alpha_n^2 a_n k + \alpha_n a_n] \\
&\leq k[(1 - \alpha_n)(1 + \alpha_n) + \alpha_n^2]\|x_n - p\| + k[\alpha_n a_n k + \alpha_n a_n] \\
&\leq k\|x_n - p\| + k(k + 1)\alpha_n a_n \quad \to \quad k\|x_n - p\|
\end{aligned} \tag{4.2}$$

when $n \to \infty$.

According to Lemma 1 in [32]

$$\lim_{n\to\infty} \|x_n - p\| \equiv A > 0.$$

Therefore, from (23) we have that

$$A \leq kA$$

when $n \to \infty$, which is a contradiction. So, the sequence $\{x_n\}$ converges strongly to $p \equiv x^*$. □

By its construction, the *SHSM* process is gradient direction descending. Further more, Backtracking line search procedure and its exit condition provides that in each iteration $f(x_{k+1}) \leq f(x_k)$, so there is $k \subset (0,1)$ such that $f(x_{k+1}) = kf(x_k)$,

Proposition 4.1. *For real number $k \subset (0,1)$, fixed point p and the process* (10) *the next estimation is valid*

$$\|x_{n+1} - p\| \leq k[1 - \alpha_n(1-\alpha_n)(1-k)]\|x_n - p\|. \tag{23}$$

Proof. Starting with the general form (8) of *SHSM* process, we have the following inequalities

$$\begin{aligned}
\|x_{n+1} - p\| &= \|(1-\alpha_n)k(x_n - p) + \alpha_n k(y_n - p)\| \\
&\leq (1-\alpha_n)k\|x_n - p\| + \alpha_n k\|y_n - p\| \\
&= k[(1-\alpha_n)k\|x_n - p\| + \alpha_n\|y_n - p\|] \\
&= k[(1-\alpha_n)\|x_n - p\| + \alpha_n\|\alpha_n(x_n - p) + k(1-\alpha_n)(x_n - p)\|] \\
&\leq k[(1-\alpha_n)\|x_n - p\| + \alpha_n^2\|x_n - p\| + k(1-\alpha_n)\alpha_n\|x_n - p\|] \\
&= k[1 - \alpha_n + \alpha_n^2 + k\alpha_n - k\alpha_n^2]\|x_n - p\| \\
&= k[1 - \alpha_n(1-\alpha_n)(1-k)]\|x_n - p\|,
\end{aligned}$$

which proves (23). □

5. Conclusion

New iterative optimization rules arrived from the $S-$iteration three-term process. Applying the adequately computed accelerated parameter and the optimal step length value, calculated by the Backtracking algorithm, we defined hybrid AGD schemes and proved their strong convergence features. The presented models can be applied on different optimization methods for further examinations, comparisons and improvements. Additionally, this research can be further studied regarding its application on the sets of uniformly convex functions and strictly quadratic functions. Numerical examinations of the proposed methods would be possible, in case they are convergent on these sets. Then, they would be comparable to the class of Khan's hybrid methods whose convergence properties are already confirmed on these sets of functions.

References

[1] R. P. Agarwal, Donal O'Regan, and D. R. Sahu, *Iterative construction of fixed points of nearly asymptotically nonexpansive mappings*, Journal of Nonlinear and Convex Analysis,**8(1)** (2007)
[2] M.Aliyari, M. Gabeleh, E. Karapinar *Mann and Ishikawa iterative processes for cyclic relatively nonexpensive mappings in uniformly convex Banach spaces*, Journal of Nonlinear and Convex Analysis, **22(4)** (2021) 699-713
[3] N. Andrei, *Nonlinear Conjugate Gradient Methods for Unconstrained Optimization*, Part of the book series: Springer Optimization and Its Applications (SOIA, volume **158**) 2020
[4] N. Andrei, *An acceleration of gradient descent algorithm with backtracking for unconstrained optimization*, Numer. Algor. **42** (2006) 63–73
[5] L. Armijo, *Minimization of functions having Lipschitz first partial derivatives*, Pac. J. Math, **6** (1966) 1–3
[6] Ü . Aksoy, I.M. Fulga, E. Karapinar *On fixed points of iterative contractive mappings defined via implicit relation*, Optimization, **70(5-6)** (2021) 1009-1027

[7] J.E. Deniss, J.S. Kowalik, R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*; Prentice-Hall, Englewood Cliffs, New Jerssey, 1983
[8] R. Fletcher, *Practical methods of optimization* Prentice-Hall,Wiley, New York, 2000
[9] A.A. Goldstein, *On steepest descent*, SIAM J. Control **3** (1965) 147–151
[10] S. Ishikawa, *Fixed points by a new iteration method*, Proc. Am. Math. Soc. **44** (1974) 147-150
[11] S.L.S Jacoby, S. L. S., J.S Kowalik, J.T. Pizzo, *Iterative Methods for Nonlinear Optimization Problems*, Prentice-Hall, Inc, Englewood, New Jersey, 1977.
[12] S.H. Khan, *A Picard-Mann hybrid iterative process*, Fixed Point Theory and Applications **69** (2013)
[13] C. Lemaréchal, *A view of line search*, In: A. Auslander, W. Oetti, J. Stoer (eds) Optimization and Optimal Control, Springer, Berlin, (1981) 59-78
[14] W.R. Mann, *Mean value methods in iterations*, Proc. Am. Math. Soc. **4** (1953) 506-510
[15] G. Marino, B. Scardamaglia, E. Karapinar, *Strong convergence theorem for strict pseudo-contractions in Hilbert spaces*, Journal of Inequalities and Applications, **2016(134)** (2016)
[16] J.J. Moré and D.J. Thuente, *On line search algorithm with guaranteed sufficient decrease*, Mathematics and Computer Science Division Preprint MCS-P153-0590, Argone National Laboratory, Argone, 1990
[17] J. Nocadal, S.J. Wright, *Numerical Optimization*, Springer-Verlag New York, Inc, 1999
[18] J.M. Ortega, W.C. Rheinboldt, *Iterative Solution Of Nonlinear Equation in Several Variables*, Academic Press, 1970
[19] S. Panić, M.J. Petrovic, M.M. Carević *Initial improvement of the hybrid accelerated gradient descent process*,Bulletin of the Australian Mathematical Society, Volume **98(2)** (2018) 331-338
[20] M.J. Petrović, *An accelerated Double Step Size method in unconstrained optimization*, Applied Math. Comput. **250** (2015) 309–319
[21] M.J. Petrović, *Hybridization Rule Applied on Accelerated Double Step Size Optimization Scheme*, Filomat, **33(3)** (2019) 655–665
[22] M. J. Petrović, D. Valjarević, D. Ilić, A. Valjarević *An Improved Modification of Accelerated Double Direction and Double Step-Size Optimization Schemes*,Mathematics, **10(2)** (2022)
[23] M.J. Petrović, V. Rakočević, N. Kontrec et al., *Hybridization of accelerated gradient descent method*,Numer. Algor. **79(3)** (2018) 769–786
[24] M.J. Petrović, P.S. Stanimirović, *Accelerated Double Direction method for solving unconstrained optimization problems*, Mathematical Problems in Engineering, Article ID 965104, 8 pages (2014)
[25] E. Picard, *Memoire sur la theorie des equations aux derivees partielles et la methode des approximations successives*, J. Math. Pures Appl. **6**(1890) 145-210
[26] F.A. Potra, Y. Shi, *Efficient line search algorithm for unconstrained optimization*, J. Optim. Theory Appl. **85** (1995) 677-704
[27] M.J.D. Powell, *Some global convergence properties of a variable-metric algorithm for minimization without exact line search*, AIAM-AMS Proc., Philadelphia **9** (1976) 53-72
[28] D. R. Sahu, *Fixed points of demicontinuous nearly Lipschitzian mappings in Banach spaces*, Comment. Math. Univ. Carolin. **46** (2005) 653-666.
[29] Z.
J. Shi, *Convergence of line search methods for unconstrained optimization*, Appl. Math. Comput. **157** (2004) 393-405
[30] P.S. Stanimirović, M.B. Miladinović, *Accelerated gradient descent methods with line search*, Numer. Algor. **54** (2010) 503-520
[31] W. Sun, Y.X. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer, New York, 2006
[32] K.K. Tan and H.K. Xu, *Approximating fixed points of nonexpansive mappings by the Ishikawa iteration process*, J. Math. Anal. Appl. **178** (1993) 301-308.
[33] M.N. Vrahatis, G.S. Androulakis, J.N. Lambrinos and G.D. Magoulas, *A class of gradient unconstrained minimization algorithms with adaptive step-size*, J. Comp. and Appl. Math. **114** (2000) 367-386
[34] P. Wolfe, *Convergence conditions for descent methods*, SIAM Rev. **11** (1968) 226-235
[35] Y. Yuan, *A new stepsize for the steepest descent method*, Research report, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, 2004