photonics

MDPI

Article

A Novel Sophia-SPGD (Stochastic Parallel Gradient Descent) Optimization Method for Wavefront Correction in WFS-Less AO (Wavefront Sensorless Adaptive Optics) Systems

Peng Chen [1], Wenjie Yang [1], Yongqi Ge [1], Zhiguang Zhang [2], Xianshuo Li [1], Zhengqing Qi [1] and Huizhen Yang [1,*]

[1] School of Network & Telecom Engineering, Jinling Institute of Technology, Nanjing 211169, China
[2] School of Electronic Engineering, Jiangsu Ocean University, Lianyungang 222005, China
* Correspondence: yanghz@jit.edu.cn

Abstract: In wavefront sensorless adaptive optics (WFS-less AO) systems, stochastic parallel gradient descent (SPGD) is the primary optimization method for correcting wavefront distortions. However, as the intensity of atmospheric turbulence interference increases, the fixed gain coefficient of the SPGD algorithm results in significant decreases in convergence speed and precision. Moreover, the algorithm is inclined to local optima, thus failing to satisfy the requirements for real-time wavefront distortion correction. To address these issues, this paper introduces a new optimization algorithm, Sophia optimized stochastic parallel gradient descent (Sophia-SPGD), which is based on second-order clipped stochastic optimization in deep learning. This algorithm computes the first-order and second-order moments of the performance metrics from its first and second gradients, respectively, and dynamically modulates the gain via a shearing mechanism to increase the convergence speed and diminish the probability of falling into local optima. Numerical simulations and experiments demonstrate that under strong turbulence conditions, the performance of Sophia-SPGD surpasses that of the traditional SPGD algorithm.

Keywords: wavefront sensorless adaptive optics; stochastic parallel gradient descent; second-order closed stochastic optimization

check for updates

Received: 18 February 2025
Revised: 26 March 2025
Accepted: 1 April 2025
Published: 2 April 2025

Citation: Chen, P.; Yang, W.; Ge, Y.; Zhang, Z.; Li, X.; Qi, Z.; Yang, H. A Novel Sophia-SPGD (Stochastic Parallel Gradient Descent) Optimization Method for Wavefront Correction in WFS-Less AO (Wavefront Sensorless Adaptive Optics) Systems. *Photonics* **2025**, *12*, 337. https://doi.org/10.3390/photonics12040337

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Adaptive optics (AO) technology, through deformable mirror (DM) or spatial light modulator (SLM) control methods, can correct optical systems in real time and reduce wavefront distortion. This technology is widely used in astronomical telescopes and free-space optical communication (FSO) to reduce the impact of atmospheric turbulence on light transmission [1–4]. AO systems are divided into conventional AO systems and wavefront sensorless (WFS-less) AO systems. WFS-less AO systems directly use the light intensity information of the image sensor to calculate performance indicators and, on the basis of the direct optimization of system performance indicators, obtain the control signals required by the wavefront corrector to achieve wavefront correction and control. This greatly reduces the complexity of the adaptive optical system and has a wider application space [5–7]. Currently, the performance of WFS-less AO systems mainly depends on the adopted wavefront correction control algorithm. This type of algorithm can be roughly divided into the following categories: Model-based control methods include geometric optics principles, nonlinear optimization, mode methods, etc. Intelligent-optimization-based methods that are used include the genetic algorithm (GA) [8,9], particle swarm optimization

(PSO) [10], differential evolution algorithm (DEA) [11], simulated annealing (SA) [12], Runge–Kutta optimizer (RUN) [13,14], and gradient-descent-based control methods, such as the sequential gradient descent algorithm, multielement high-frequency vibration, and stochastic parallel gradient descent algorithm (SPGD) [15–19]. Model control methods have faster convergence speeds, but they need to establish an accurate mathematical model on the basis of the physical characteristics of the system to design a control algorithm. However, intelligent optimization control methods have higher computational complexity [20–22]. In contrast, gradient descent control algorithms do not depend on the system model, are relatively simple in terms of computation, and have a lower dependence on optical systems.

SPGD is currently the most representative and widely applicable control algorithm because of its simple implementation and strong comprehensive correction ability. Conventional SPGD algorithms use fixed gain coefficients, resulting in a slow convergence speed or easy trapping in local optima, leading to a low correction performance. To optimize the problems existing in the above SPGD algorithm, with the continuous development of deep learning, optimizers such as momentum, AdaGrad, Adam, and Nadam have been successfully integrated into the SPGD algorithm, significantly accelerating the iteration process. These improved SPGD algorithms have been validated through both theoretical analysis and simulations, demonstrating their effectiveness in wavefront distortion correction [23–26]. The Sophia optimizer is a second-order optimizer that can achieve a faster convergence speed and better optimization effects in deep learning models [27]. The Sophia optimizer as a second-order stochastic optimization algorithm been designed to reduce the cost of model training and improve training efficiency. It uses a low-cost random diagonal approximation of the Hessian as a preconditioner and introduces a clipping mechanism to control the maximum value of the update scale. This approach combines the advantages of second-order optimization while maintaining computational feasibility. Compared with the traditional Adam optimizer, it performs better in terms of validation forward loss, total computational cost, and actual training time. It can halve the number of training steps under the same loss, effectively reducing the demand for total computational resources.

To address the issues of slow convergence speed and susceptibility to local optima in the SPGD algorithm, this paper proposes a novel adaptive gain stochastic parallel gradient descent algorithm (Sophia-SPGD). The algorithm incorporates the lightweight second-order optimizer Sophia, which is utilized in deep learning, into the SPGD. It approximates the gradient by leveraging changes in the image performance index. By incorporating an adaptive-learning-rate adjustment mechanism, first- and second-order momentum corrections, and a clipping mechanism to stabilize the control of the gradient descent direction and step size, Sophia-SPGD enables precise and rapid updates to the wavefront corrector voltage signals. Through the modification of gradients and learning rates, the algorithm enhances the robustness of the optimization process, significantly improving both the convergence speed and accuracy of wavefront correction. This is highly important for the real-time performance and lightweight nature of WFS-less AO.

2. Methods

As is illustrated in Figure 1, the WFS-less AO system uses a deformable mirror (DM) for wavefront correction, with a charge-coupled device (CCD) serving as the image sensor, alongside other components such as the wavefront control module. During the laser beam transmission process, the incoming laser beam passes through a nonuniform medium, causing distortions in the wavefront phase. The distorted laser beam is then reflected by the deformable mirror and passes through a converging lens to form an image on the CCD camera. The CCD camera captures this image and transmits it to the controller. The controller reads the far-field spot image and uses a blind optimization algorithm based

on the optimization metrics of the far-field spot image to generate control voltage signals. These signals are amplified by a high-voltage amplifier and drive the deformable mirror to complete a cycle of closed-loop correction.

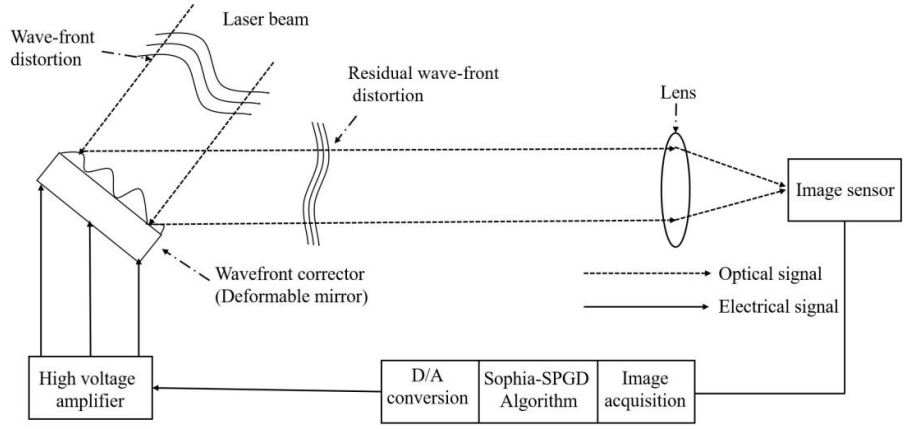


Figure 1. Schematic of the WFS-less AO system. The solid lines indicate electrical signals, the dashed lines indicate optical signals, and the jagged lines indicate wavefront disturbances.

The distortion compensation process of the deformable mirror is achieved by altering the shape of its deformable surface to counteract the target wavefront distortions. There are many methods for representing the surface shape and wavefront; the most direct method is to represent the surface shape as $S(x, y)$, which is the distance of a point at coordinates (x, y) from the horizontal plane. According to the working principle of the deformable mirror, the compensatory surface shape $\phi(x, y)$ produced by the mirror can be represented by a linear combination of the influence functions of each actuator, and the equations are as follows:

$$\phi(x, y) = \sum_{i=1}^{N} v_i S_i(x, y) \tag{1}$$

$$S_i(x, y) = \exp(\ln \omega(\frac{\sqrt{(x - x_i)^2 + (y - y_i)^2}}{d_0})^{\alpha}) \tag{2}$$

Here, v_i is the i-th control signal for the DM actuator, $S_i(x, y)$ is the influence function of the i-th DM actuator, (x, y) represents the coordinates on the wavefront plane, (x_i, y_i) represents the coordinates of the i-th actuator on the DM, d_0 represents the spacing between adjacent elements, ω represents the coupling value between elements, and α represents the Gaussian metric [13,14].

In the simulation experiment, a sum of 100 Kolmogorov phase screens under turbulence strengths of D/r_0 = 10, 15, and 20 are generated [28]. The phase screen comprises Zernike aberrations from the 3rd to the 104th order, excluding tilt terms. The statistical properties of the generated phase screen conform to the Kolmogorov spectrum, and the screens do not exhibit correlations between them. The turbulence intensity is denoted as D/r_0, where D is the telescope aperture and r_0 is the atmospheric coherence length. The optimization algorithm uses the mean radius (MR) and mean radius ratio (MRR) [29] as an optimization metric; the Strehl ratio (SR) [30] acts as an indicator metric. The MRR is the ratio between the average radius of the ideal far-field point spread function (PSF) and the average radius of the corrected PSF.

Allowing $\phi(x, y)$ to be the phase distribution of the wave front, in a circular aperture, the complex amplitude of the wave front $U(x, y)$ can be expressed as:

$$U(x, y) = \exp(i\phi(x, y)), \quad 0 \le \sqrt{(x^2 + y^2)} \le 1 \tag{3}$$

The relationship between the PSF and the complex amplitude of the wavefront can be described by means of Fourier transform. Performing a two-dimensional Fourier transform on the $U(x, y)$ yields its representation of $P(u, v)$ in the frequency domain, and the PSF can be obtained by calculating the square of the modulus of $P(u, v)$.

$$P(u, v) = \int \int_{-\infty}^{\infty} U(x, y) \exp[-i2\pi(ux + vy)]dxdy \tag{4}$$

$$PSF(u, v) = |P(u, v)|^2 \tag{5}$$

The formulas for the SR, mean radius ratio (MR), and MRR are given below:

$$SR = I(x, y)/I_0(x_0, y_0) \tag{6}$$

$$MR = \frac{\iint \sqrt{(x - x_0)^2 + (y - y_0)^2} I(x, y)dxdy}{\iint I(x, y)dxdy} \tag{7}$$

$$MRR = MR_{ideal}/MR \tag{8}$$

where $I(x, y)$ is the peak intensity of the image formed by the optical system with aberrations, $I_0(x_0, y_0)$ is the peak intensity of the ideal diffraction-limited image, and MRideal is the ideal mean radius. The SRs are frequently used in numerical simulations as the system's objective function because of their low computational demands and simple structure. SR values close to 1 indicate minimal aberration effects and a high image quality, whereas lower values suggest significant aberrations and a reduced image quality. The mean radius is an essential parameter for evaluating the quality of an optical system, as it measures the deviation of the actual wavefront from the ideal wavefront. A smaller mean radius signifies a closer approximation to the ideal, diffraction-limited wavefront, indicating a higher optical quality and enhanced imaging performance. Throughout the algorithm's iterative process, MR values are recorded. For analytical convenience, changes in the SR values are also recorded alongside.

In this paper, we introduce a Sophia-optimized SPGD algorithm, referred to as Sophia-SPGD, which integrates the Sophia optimizer from deep learning with the traditional SPGD algorithm. The variation in the performance metrics of the spot image is approximated as a gradient. The algorithm incorporates momentum, adaptivity, and gradient clipping techniques in the gradient descent process. By utilizing the gradient momentum method, the algorithm compares the variation in gradients from the current and previous iterations. If the variation in gradients is in the same direction, it generates a positive excitation, whereas opposing directions cause a negative excitation. This method effectively minimizes oscillations throughout the iteration process, enhancing the stability and efficiency of the optimization. The learning rate is adaptively adjusted to mitigate local convergence issues. Additionally, the algorithm incorporates gradient clipping to control the step size of the gradient descent, ensuring that the performance indices of spot images do not fluctuate dramatically in the later stages of the iterations. This approach streamlines the optimization process, enhancing the efficiency and stability of wavefront corrections in WFS-less adaptive optics systems. The Sophia-SPGD algorithm flowchart is shown in Figure 2.

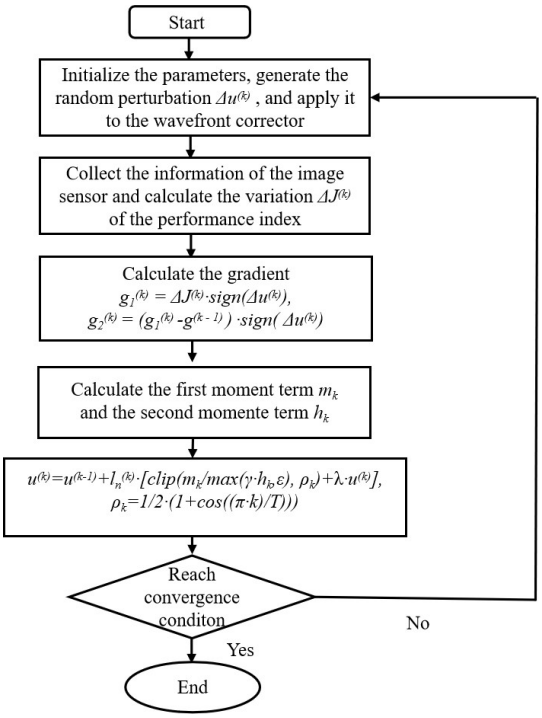


Figure 2. Flowchart of the Sophia-SPGD algorithm.

The core concept of the stochastic parallel gradient descent (SPGD) algorithm involves estimating the gradient of control parameters through changes in performance metrics ΔJ and randomized perturbation voltage vector Δu. The algorithm iteratively searches for control voltage vectors u along the direction of gradient descent until it finds a maximum value of the performance metrics. Here, the performance metric J is a function of the control voltage vectors, expressed as $J = J(u)$, where u represents the control voltages for the various units of the wavefront corrector.

The specific process is as follows: First, the system randomly generates a small perturbation voltage vector $\Delta u^{(k)} = \{\Delta u_1, \Delta u_2, \ldots, \Delta u_N\}^{(k)}$. Then, the positive perturbation voltages $u^{(k-1)} + \Delta u^{(k)}$ are applied to the DM to obtain the CCD images and calculate the positive performance metrics ${J_+}^{(k)}$, and the negative perturbation voltages $u^{(k-1)} - \triangle u^{(k)}$ are applied to the DM via a similar method to obtain the negative performance metrics ${J_-}^{(k)}$. The formula for calculating the variation in performance metrics is as follows:

$$\Delta J^{(k)} = {J_+}^{(k)} - {J_-}^{(k)} \tag{9}$$

$$u^{(k)} = u^{(k-1)} + lr \cdot \Delta J^{(k)} \cdot \Delta u^{(k)} \tag{10}$$

where lr represents the gain coefficient, which is generally a positive value; $u^{(k)} = \{u_1, u_2, \ldots, u_N\}^{(k)}$ is the voltage of the deformable mirror controller added during the k-th iteration; N denotes the number of corrective units in the wavefront corrector; and $\Delta u^{(k)}$ is the randomly disturbed voltage vector applied during the k-th iteration. The control voltage vector $u^{(k)} = \{u_1, u_2, \ldots, u_N\}^{(k)}$ is calculated in accordance with Formula (6) and applied to the deformable mirror. Images are collected to obtain the correction effect, and the k-th iteration is completed.

SPGD is pervasively recognized as an effective method for performance metric optimization; however, the use of a fixed gain coefficient lr limits its adaptability for wavefront correction in wavefront sensorless adaptive optics systems. This leads to suboptimal correction results, slow correction speeds, and an easy fall into the local optimum. To increase

the convergence speed and reduce the probability of falling into the local optimum, the momentum of inertia and adaptive learning rates are the most common considerations.

In the Sophia-SPGD algorithm, the first-order gradient and the second-order gradient are calculated according to Equations (11) and (12):

$$g_1^{(k)} = \Delta J^{(k)} \cdot sign(\Delta u^{(k)}) \tag{11}$$

$$g_2^{(k)} = (g_1^{(k)} - g_1^{(k-1)}) \cdot sign(\Delta u^{(k)}) \tag{12}$$

where $sign()$ is the sign function, which determines the sign of a number. The sign of $sign(\Delta u^{(k)})$ and its performance are determined by the direction of optimization of the performance metrics. If the performance metrics increase, $\Delta J^{(k)} sign(\Delta u^{(k)})$ takes a positive value; otherwise, it takes a negative value.

In the Sophia-SPGD algorithm, the introduction of a momentum term aims to accelerate learning and reduce oscillations. Here, m_k represents the first-order momentum term, and h_k represents the first-order momentum term. β_1 and β_2 are momentum hyperparameters, typically ranging between 0 and 1. The first-order momentum of the gradient m_k and the second-order momentum of the gradient according to h_k in Equations (13) and (14) are calculated as follows:

$$m_k = \beta_1 m_{k-1} + (1 - \beta_1) g_1^{(k)} \tag{13}$$

$$h_k = \beta_2 h_{k-1} + (1 - \beta_2) g_2^{(k)} \tag{14}$$

These formulas imply that the current momentum is a weighted average of the previous momentum and the current gradient. These two momentum terms play crucial roles in the parameter update process. By integrating past momentum with current gradient information, the algorithm progresses more steadily toward the optimal solution. This approach prevents overly sensitive parameter changes in situations of drastic gradient variations and enhances the convergence speed and robustness of the algorithm.

The primary role of adopting an adaptive learning rate is to balance the convergence speed and avoid becoming trapped in local optima during the optimization process. Adaptive-learning-rate algorithms automatically adjust the learning rate according to different parameters and formula conditions. In the initial stages of iteration, a higher learning rate is provided to accelerate convergence. As the iteration progresses and the performance indices gradually approach their extremum, the learning rate automatically decreases to prevent significant fluctuations near the optimal solution, thus ensuring stable convergence to the optimum. The relationship between the learning rate l_n and the number of iterations k is expressed as

$$l_n = lr_0 / (1 + \rho_0 \cdot k) \tag{15}$$

Here, lr_0 is a manually predefined initial learning rate. ρ_0 is the attenuation rate.

Introducing weight decay λ when updating control voltages can smooth the changes in parameter update, thereby accelerating the convergence speed. As discussed above, the Sophia-SPGD algorithm is used to update the control voltage vector computation formula as follows:

$$u^{(k+1)} = u^{(k)} + l_n \cdot [clip(\, m_k / max\{\gamma \cdot h_k, \epsilon\},\ 1/2 \cdot (1 + \cos(\pi \cdot k)/T)) + \lambda \cdot u^{(k)}] \tag{16}$$

where $u^{(k+1)}$ represents the voltage signal for the k iteration and where T is the total number of iterations. The clipping function $clip$ is used to limit $m_k / max\{\gamma \cdot h_k, \epsilon\}$ within the range of the adaptive function $(1/2 \cdot (1 + \cos(\pi \cdot k)/T))$, and $\varepsilon > 0$ is a very small constant set to 10^{-8} to avoid dividing by 0. In the above equation, the typical parameters are $lr_0 = 0.02 \sim 0.04$, $\rho_0 = 0.001$, $\gamma = 0.01$, $\beta_1 = 0.9$, $\beta_2 = 0.99$. Using an adaptive function

to limit the range of the clipping function can effectively increase the convergence accuracy in the later stages of iteration and prevent falling into a local optimum.

The implementation of the Sophia-PGD algorithm is comprehensively described in Algorithm 1.

Algorithm 1. Sophia-SPGD

Inputs: The learning rate lr, the initial learning rate lr_0, hyperparameters λ, γ, β_1, β_2, the constant ε, the initial momentum coefficient ρ_0.

Output: Calculated control voltage vectors $u_1, u_2, \ldots, u_N$

Set $u^{(k)} = \{u_1, u_2, \ldots, u_N\}^{(k)} = 0$, where N is the number of corrector channels. Set $m_0 = 0$, $v_0 = 0$, $h_0 = 0$.

for k = 1 to T **do**

Randomly generate perturbed voltages $\Delta u^{(k)} = \{\Delta u_1, \Delta u_2, \ldots, \Delta u_N\}^{(k)}$

Compute the perturbation of performance indicators ${J_+}^{(k)}$, ${J_-}^{(k)}$, and $\Delta J^{(k)} = {J_+}^{(k)} - {J_-}^{(k)}$

Compute $g_1^{(k)} = \Delta J^{(k)} \cdot sign(\Delta u^{(k)})$

$m_k = \beta_1 m_{k-1} + (1 - \beta_1) g_1^{(k)}$

$g_2^{(k)} = (g_1^{(k)} - g_1^{(k-1)}) \cdot sign(\Delta u^{(k)})$

$h_k = \beta_2 h_{k-1} + (1 - \beta_2) g_2^{(k)}$

$l_n = lr_0 / (1 + \rho_0 \cdot k)$

$u^{(k+1)} = u^{(k)} + l_n \cdot [clip(\, m_k / max\{\gamma \cdot h_k, \epsilon\},\ 1/2 \cdot (1 + \cos(\pi \cdot k)/T)) + \lambda \cdot u^{(k)}]$

End

To implement SPGD or any of its variants, a performance metric J must be specified. Here, the MR and SR are used to measure the correction performance of the algorithm and are defined as shown in Equations (3) and (4). Through numerical simulation, we can further observe and assess the performance of the Sophia-SPGD algorithm in practical scenarios to determine its effectiveness and advantages in enhancing correction performance.

3. Simulations and Analysis

To analyze the feasibility of the Sophia-SPGD algorithm, we first select the optimal parameters for each algorithm on the basis of many simulation experiments under various turbulence intensities. Then, we use both algorithms to correct the same set of randomly generated wavefront aberrations and compare the simulation results. The number of correction iterations is set at 1500. A 97-element DM is introduced to perform the simulations. Figure 3 shows the obtained averaged SR and MRR adaptation curves. Except for cases with turbulence intensity $D/r_0 = 5$, the correction capability of Sophia-SPGD surpasses that of SPGD, particularly in the presence of stronger turbulence.

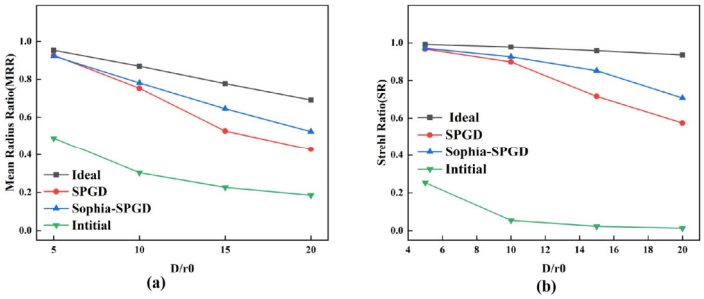


Figure 3. Convergence extreme values of (**a**) MRR and (**b**) SR after 1500 iterations of Sophia-SPGD and traditional SPGD under turbulence strengths D/r0 = 5, 10, 15, and 20.

During the simulation process, 100 frames of random wavefront aberrations under different turbulence intensities of D/r0 = 10, 15, and 20 were used as correction subjects to analyze the convergence speeds of the Sophia algorithm and the SPGD control algorithms. The average correction results of the 100 frames of random aberrations serve as the

experimental outcome, with each of the four algorithms iterating 1500 times. The average MRR and average SR adaptation curves are presented in Figure 4. Both the SPGD and Sophia-SPGD algorithms have sufficiently converged after 1500 iterations. Comparing the convergence curves of different algorithms under the same turbulence conditions, it is evident that at lower turbulence, Sophia does not demonstrate a significant advantage in terms of convergence speed over SPGD. However, a comparison of the convergence curves under varying turbulence conditions reveals that as the turbulence intensity increases, the convergence performance of the Sophia-SPGD control algorithm progressively exceeds that of the SPGD control algorithm. The Sophia-SPGD and SPGD algorithms can effectively optimize and compensate for aberrations to some extent. However, under high-turbulence conditions, the correction accuracy is degraded because of the limitations of the correction capacity of the deformable mirror.

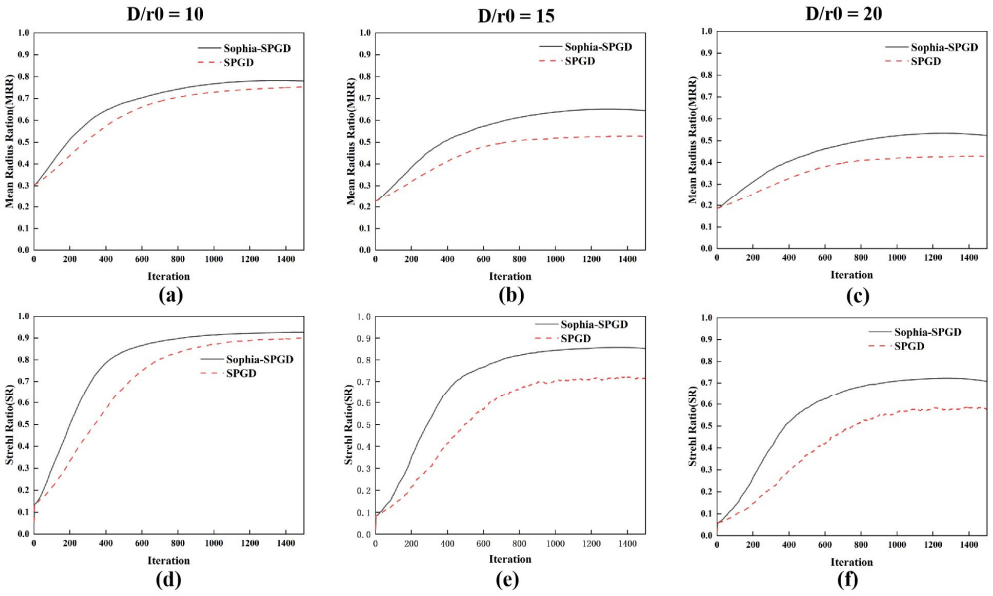


Figure 4. Adaptive process of the adaptive optical system with D/r0 = 10, 15, and 20 when the Sophia-SPGD and SPGD control algorithms are used. (**a**–**c**) MRR curves; (**d**–**f**) the corresponding SR curves.

To observe the wavefront correction effects of the SPGD and Sophia-SPGD algorithms more objectively, a typical phase screen generated under a turbulence strength of D/r0 = 10 is selected. The uncorrected wavefront, its corresponding 3rd–18th Zernike coefficients and initial uncorrected PSF, are provided in Figure 5a, b, and c, respectively. The simulation results obtained via the SPGD algorithm are shown in Figure 5d–f, and the simulation results obtained via the Sophia-SPGD algorithm are shown in Figure 5h–i. After 500 iterations, both of the above algorithms eliminate the wavefront aberration effectively, and the corresponding far-field point spread function (PSF) takes shape well. Judging from the distribution of the Zernike coefficients corresponding to the wavefront residuals after correction by the two algorithms, the correction performance of the Sophia-SPGD algorithm is better than that of the SPGD algorithm.

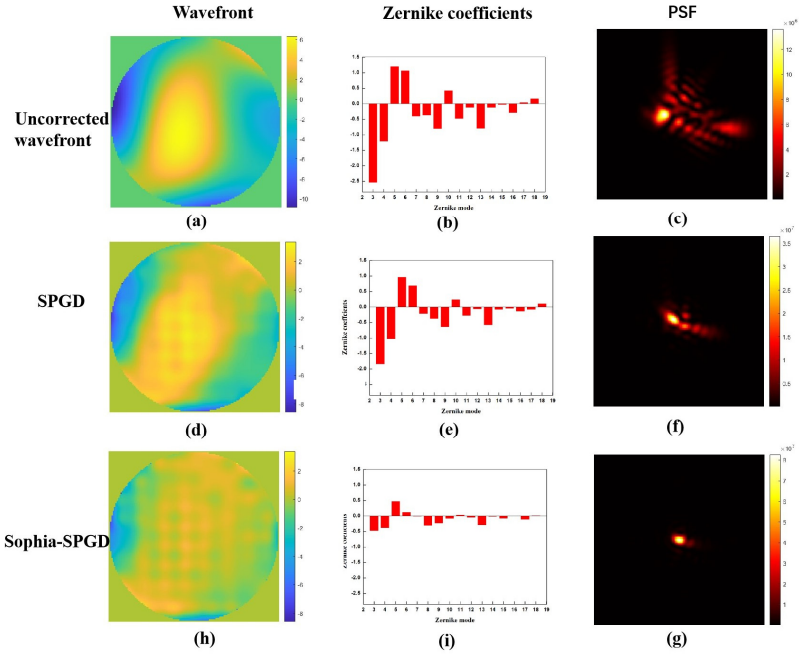


Figure 5. Wavefront aberration correction and the corresponding far-field point spread function (PSF) of a typical phase screen generated under turbulence strength $D/r_0 = 10$. The number of iterations is set to 500. (**a**) Uncorrected wavefront. (**d**) Wavefront corrected with SPGD. (**h**) Wavefront corrected with Sophia-SPGD. (**b**,**e**,**i**) The corresponding Zernike coefficients of wavefront. (**c**,**f**,**g**) The corresponding PSFs of wavefront, respectively.

In order to investigate whether the algorithms fluctuate in the later iterations of the wavefront correction, we give the SR iteration curves of the SPGD and Sophia optimization algorithms, respectively (Figure 6). The atmospheric turbulence strength is D/r = 15 for all of these realizations. It is shown that, during the iteration, the SR curve of conventional SPGD exhibits fluctuations. In contrast, the SR curve of the Sophia-SPGD algorithm is relatively smooth during the iteration and shows no significant fluctuations in the later stage of the iteration.

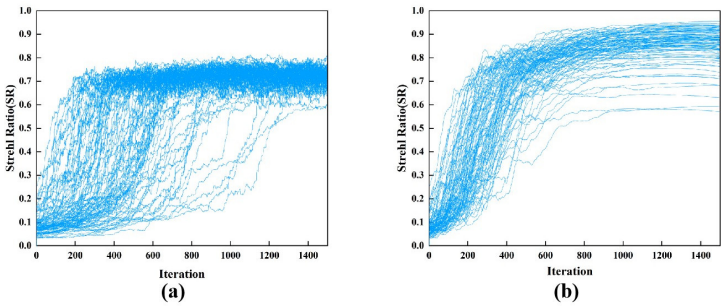


Figure 6. The SR curves of 100 realizations of (**a**) SPGD and (**b**) Sophia-SPGD under turbulence strength D/r0 = 15.

In addition, to better verify the stability of the algorithm, we use the Euclidean norm to measure the parameter update of the control voltage vector $\Delta u^{(k)}$ during the iteration. Figure 7 illustrates the $\left\|\Delta u^{(k)}\right\|$ in each step size of each algorithm throughout the iteration process. During the iteration of SPGD, the changes in adjacent step sizes are very drastic. In contrast, during the iteration of Sophia-SPGD, the changes in adjacent step sizes are gentle and tend to be stable in the later stage of the iteration.

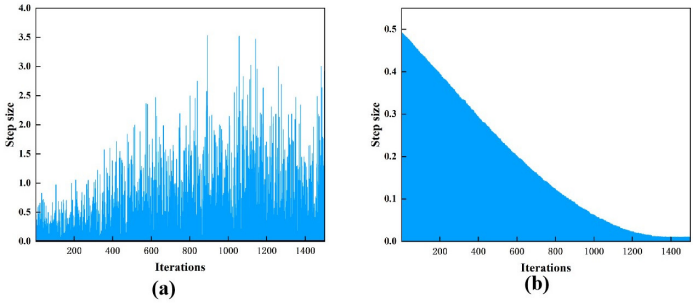


Figure 7. Step size $\Delta u^{(k)}$ of the k-th iteration during the whole iterative procedure of (**a**) SPGD and (**b**) Sophia-SPGD algorithms.

Convergence speed is a critical metric in the control of adaptive optics systems. Under three different conditions of strong turbulence, the initial average MR values were 0.304, 0.227, and 0.186, respectively; after correction by the SPGD control algorithm, the average MRR values at the convergence extremum were 0.753, 0.527, and 0.428, respectively. Calculating 80% of the MRR correction range under different turbulence conditions yielded values of 0.663, 0.467, and 0.380. These values serve as benchmarks to analyze and compare the number of iterations required by the two control algorithms to achieve the same correction effect, as shown in Table 1. The simulation data demonstrate that the Sophia algorithm is more efficient than the SPGD algorithm. At a turbulence strength of D/r0 = 10, the correction speed of the Sophia-SPGD algorithm is 35.6% faster than that of the SPGD algorithm; at D/r0 = 15, it is 75% faster; and at D/r0 = 20, it is 79.5% faster.

Table 1. Comparison of the average value of the MRR and the number of times that 80% of the local optimum is reached between the two algorithms under different turbulence strengths.

MRR	$D/r_0 = 10$		$D/r_0 = 15$		$D/r_0 = 20$	
	Iterations	**Scaled**	**Iterations**	**Scaled**	**Iterations**	**Scaled**
SPGD	609	1.356	553	1.75	594	1.795
Sophia-SPGD	449	1	316	1	331	1

Under three different conditions of strong turbulence, the initial average SR values were 0.136, 0.084, and 0.058; after correction by the SPGD control algorithm, the average MR values at the convergence extremum were 0.899, 0.714, and 0.575, respectively. Calculating 80% of the SR correction range under different turbulence conditions yielded values of 0.746, 0.573, and 0.472. These values serve as benchmarks to analyze and compare the number of iterations required by the two control algorithms to achieve the same correction effect, as shown in Table 2. At a turbulence strength of D/r0 = 10, Sophia-SPGD's correction speed is 65.9% faster than that of SPGD; at D/r0 = 15, it is 80.4% faster; and at D/r0 = 20, it is 93.2% faster.

Table 2. Comparison of the average value of the SR and the number of times that 80% of the local optimum is reached between the two algorithms under different turbulence strengths.

SR	$D/r_0 = 10$		$D/r_0 = 15$		$D/r_0 = 20$	
	Iterations	**Scaled**	**Iterations**	**Scaled**	**Iterations**	**Scaled**
SPGD	594	1.659	606	1.804	686	1.932
Sophia-SPGD	358	1	336	1	355	1

In conclusion, the convergence results from both the MR and SR indicate that as the turbulence intensity increases, the convergence effectiveness becomes constrained.

Notably, under conditions of strong turbulence, Sophia-SPGD consistently achieves faster convergence than SPGD.

4. Experiments and Results

To evaluate the correction capability of the wavefront sensorless adaptive optics system based on the Sophia algorithm, the aberrations within the experimental system were selected as the reference aberrations. These aberrations were used to control a 97-element piezoelectric deformable mirror, and experiments were conducted to verify the correction performance for both point targets and extended targets. The deformable mirror used was a high-speed deformable mirror with 97 piezoelectric elements manufactured by ALPAO, France, with an effective aperture of 22.5 mm and an actuator spacing of 2.5 mm. The CCD camera utilized was the Prime BSI Express Scientific camera, which was developed and produced by Teledyne Photometrics. The experimental results were evaluated via the MR as a performance metric.

In the point target correction experiment, the experiment setup was carried out referring to the optical path shown in Figure 1. We employed an adjustable-power semiconductor laser (wavelength of 635 nm) as the light source. During the experiment, the laser emitted a 22.5 mm diameter beam through a collimating system, matching the effective aperture of the deformable mirror. The beam is reflected by an anamorphic mirror, passes through a beam reduction mirror (12.5×), and finally focuses on the target surface of the CCD camera through a focusing lens (f = 150 mm). An image acquisition card captured the speckle information detected by the CCD and transferred it to a MATLAB 2022b program on the computer for analysis. The program calculates the average radius of the speckles within a selected area as a performance metric for the algorithm. The required control voltages for the deformable mirror were converted by a D/A card, amplified by a high-voltage amplifier, and applied to 97 elements of the mirror, causing it to deform and optimize the speckle radius. The algorithm iterates until it reaches a predetermined number of iterations or converges to an optimal value. The algorithm used in this experiment was the same as that used in the previously described simulation experiment.

Figure 8a displays the far-field intensity distributions before correction under system static aberration. The diffraction limit of the experimental system used was calculated to be 19.8 pixels. Visually, the far-field speckle patterns reveal that the SPGD algorithm improved the MR from 89 pixels to 36 pixels, whereas the Sophia-SPGD algorithm increased it from 89 pixels to 27 pixels. Additionally, the number of iterations required to reach a far-field speckle size of three times the diffraction limit was 152 and 133 for the SPGD and Sophia-SPGD algorithms, respectively. Three-dimensional far-field intensity distributions are shown in Figure 8b,c, where it can be clearly seen that the largest value of central energy was obtained with the Sophia-SPGD-algorithm's correction. Figure 9 shows the experimentally obtained MR curves during correction via the SPGD and Sophia-SPGD algorithms.

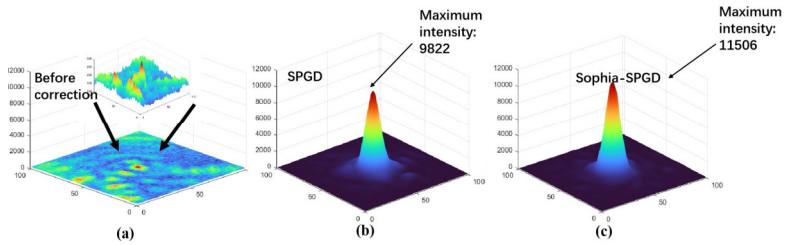


Figure 8. Comparison of 3D far-field intensity distributions: (**a**) before correction under static aberration; (**b**,**c**) after correction with the SPGD and Sophia-SPGD algorithms.

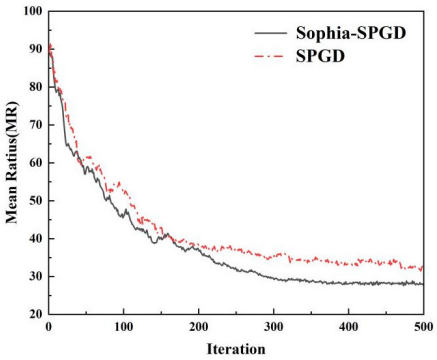


Figure 9. MR curves for the SPGD and Sophia-SPGD algorithms.

The extended-target-correction experimental system layout is shown in Figure 10. We used an LED with a center wavelength of 570 nm as the light source. The extended target was a resolution test board, model GH-YP832, Guangzhou, China which consisted of 29 sets of four-direction stripes with line widths from 2.5 μm to 500 μm. The diffraction limit of the constructed optical system was 25 μm. The anamorphic mirror was a 97-unit MEMS DM made by ALPAO, Grenoble, France, for aberration correction. The camera was a 16-bit CDD camera. Lens L1 has a focal length of 400 mm, and L2 has a focal length of 100 mm, both of which form a 4f system. This process is similar to point target tests: an image acquisition card collects information about the extended target from the CCD camera and transfers it to MATLAB 2022b programs on a computer for analysis. The program evaluates the algorithm's performance via an image frequency evaluation function within a selected area. The control voltages for the deformable mirror, which are calculated via the algorithm, are converted and amplified before being applied to the 97 elements of the mirror, which induces deformation to optimize the performance metrics. The algorithm iterates until it reaches a predetermined number of iterations or converges to the optimal value. The algorithm used in this experiment was the same as that employed in previous simulation experiments.

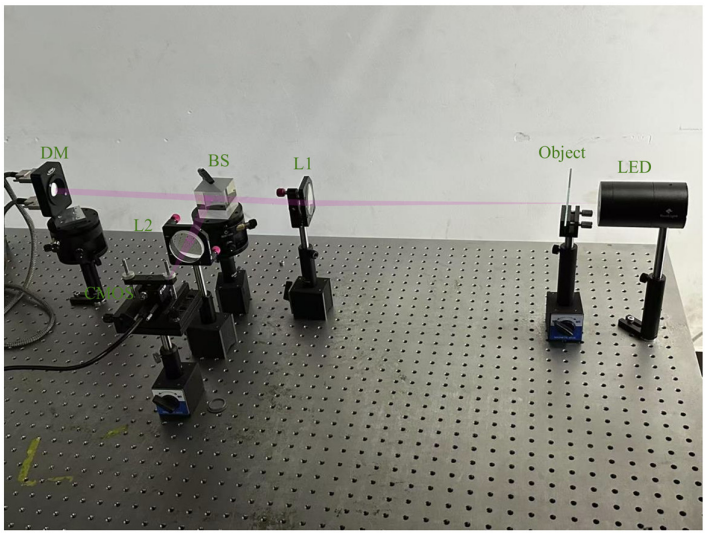


Figure 10. Experimental system layout.

Figure 11a shows that the original stripe image is blurred in detail. After correction with the SPGD algorithm (Figure 11b), although there was some improvement in image detail, blurriness remained. Figure 11c shows that the newly proposed Sophia-SPGD algorithm successfully removed the blurred details, resulting in a clearer overall structure that more closely approximated the ideal diffraction-limited image. Image quality was

assessed via the custom-modified frequency assessment function F. The F value of the image before correction was 4.1275; after correction, the F value of the SPGD algorithm was 4.6399; and the F value of the Sophia-SPGD algorithm was 5.8653.

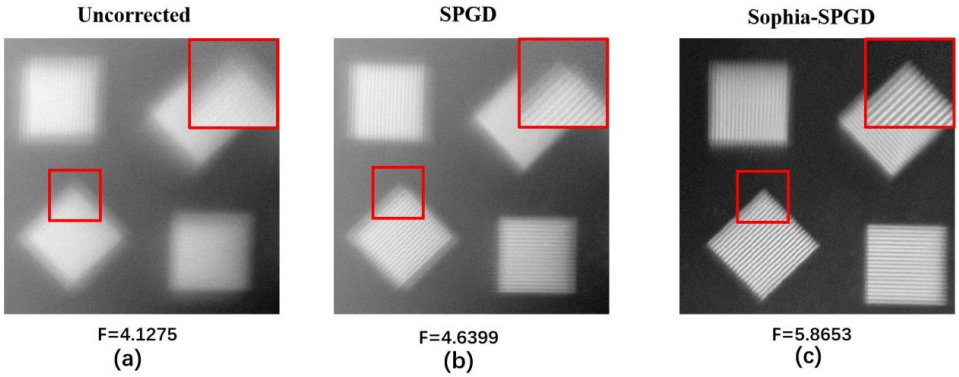


Figure 11. Reconstruction results of object imaging degraded by aberrations in the system. (**a**) Initial uncorrected image, (**b**) correction image via SPGD, (**c**) correction image via Sophia-SPGD.

F value curves after correction via the SPGD and Sophia-SPGD algorithms are shown in Figure 12. Compared with SPGD, Sophia-SPGD requires fewer iterations to achieve the same image quality evaluation metric. The images corrected by the SPGD algorithm were unsatisfactory, whereas the images corrected by Sophia-SPGD had satisfactory visual results.

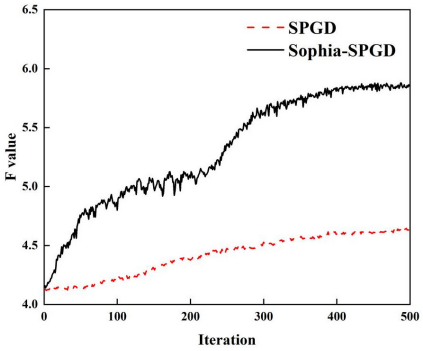


Figure 12. The F value curves for the SPGD and Sophia-SPGD algorithms.

5. Conclusions

This paper introduces an improved SPGD algorithm, termed Sophia-SPGD, aimed at addressing the poor adaptability of the gradient update within the algorithm, thereby enhancing the convergence speed and accuracy of the SPGD algorithm and reducing the likelihood of falling into local optima. The Sophia-SPGD algorithm approximates the minor changes in system performance indices as gradients, calculates the first and second moments of these gradients, and introduces mechanisms for weight decay and adaptive shearing limit updates. This allows for adaptive adjustment of the gain coefficient, improving the system's convergence accuracy, accelerating the convergence speed, and reducing the probability of settling into local minima. Through numerical simulations and experiments, our algorithm was compared with the traditional SPGD algorithm. The results demonstrate superior performance under various atmospheric turbulence conditions compared with the SPGD algorithm. The Sophia-SPGD algorithm can reduce the number of iterations and the probability of falling into local optima. Moreover, the parameter adaptability of the Sophia-SPGD algorithm is stronger than that of the SPGD algorithm. Overall, the Sophia-SPGD algorithm proposed in this paper features a rapid convergence speed and is less prone to local optima, making it more promising for practical applications.

Author Contributions: Conceptualization, P.C. and H.Y.; methodology, P.C.; software, W.Y. and Z.Z.; validation, Y.G., X.L. and Z.Q.; resources, H.Y.; writing—original draft preparation, P.C.; writing—review and editing, H.Y.; funding acquisition, H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Laboratory on Adaptive Optics, China (grant number FNLAO-24-MS-O01); the National Natural Science Foundation of China (grant numbers 12473081 and U2141255); and the Jinling Institute of Science and Technology High-level Talent Introduction Program (grant number jit-b-201813).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tyson, R.K.; Frazier, B.W. *Principles of Adaptive Optics*; CRC Press: Boca Raton, FL, USA, 2022.
2. Madec, P.Y. Overview of deformable mirror technologies for adaptive optics and astronomy. In *Adaptive Optics Systems III*; SPIE: Bellingham, WA, USA, 2012; Volume 8447, pp. 22–39.
3. Martínez, N.; Ramos, L.F.R.; Sodnik, Z. Simulating the performance of adaptive optics techniques on FSO communications through the atmosphere. In *Laser Communication and Propagation Through the Atmosphere and Oceans VI*; SPIE: Bellingham, WA, USA, 2017; Volume 10408, pp. 49–57.
4. Li, M.; Gao, W.; Cvijetic, M. Slant-path coherent free space optical communications over the maritime and terrestrial atmospheres with the use of adaptive optics for beam wavefront correction. *Appl. Opt.* **2017**, *56*, 284–297. [CrossRef] [PubMed]
5. Primot, J. Theoretical description of Shack–Hartmann wave-front sensor. *Opt. Commun.* **2003**, *222*, 81–92. [CrossRef]
6. Antonello, J.; van Werkhoven, T.; Verhaegen, M.; Truong, H.H.; Keller, C.U.; Gerritsen, H.C. Optimization-based wavefront sensorless adaptive optics for multiphoton microscopy. *J. Opt. Soc. Am. A* **2014**, *31*, 1337–1347. [CrossRef] [PubMed]
7. Yang, H.; Soloviev, O.; Verhaegen, M. Model-based wavefront sensorless adaptive optics system for large aberrations and extended objects. *Opt. Express* **2015**, *23*, 24587–24601. [CrossRef]
8. Zhang, Y.; Zhang, H.; Yuan, G. A Phase Recovery Technique Using the Genetic Algorithm for Aberration Correction in a Coherent Imaging System. *Sensors* **2023**, *23*, 7679. [CrossRef]
9. Yang, P.; Hu, S.J.; Chen, S.Q.; Yang, W.; Xu, B.; Jiang, W.H. Research on the phase aberration correction with a deformable mirror controlled by a genetic algorithm. In *Journal of Physics: Conference Series, International Symposium on Instrumentation Science and Technology, 8–12 August 2006, Harbin, China*; IOP Publishing: Bristol, UK, 2006; Volume 48, No. 1.
10. Yue, W.; Jin, G.; Yang, X. Adaptive Particle Swarm Optimization for Automatic Design of Common Aperture Optical System. *Photonics* **2022**, *9*, 807. [CrossRef]
11. Hu, X.; Yang, D.; Dou, J.; Yang, Z.; Liu, Z. Differential evolution algorithm-based aberration correction method for the interference of vortex beams. *Appl. Phys. B Laser Opt.* **2022**, *128*, 133. [CrossRef]
12. Li, Z.; Cao, J.; Zhao, X.; Liu, W. Atmospheric compensation in free space optical communication with simulated annealing algorithm. *Opt. Commun.* **2015**, *338*, 11–21. [CrossRef]
13. Yang, H.; Zang, X.; Zhang, Z.; Liu, J. Wavefront Correction System Based on RUN Optimization Algorithm. *Acta Photonica Sin.* **2023**, *52*, 1111004.
14. Yang, H.; Zang, X.; Chen, P.; Hu, X.; Miao, Y.; Yan, Z.; Zhang, Z. An Efficient Method for Wavefront Aberration Correction Based on the RUN Optimizer. *Photonics* **2024**, *11*, 29. [CrossRef]
15. Vorontsov, M.A.; Sivokon, V.P. Stochastic parallel-gradient-descent technique for high-resolution wave-front phase-distortion correction. *J. Opt. Soc. Am. A* **1998**, *15*, 2745–2758.
16. Zhao, H.; Lv, D.; An, J.; Kuang, K.; Yu, M.; Zhang, T. Meta-heuristic SPGD algorithm in spatial light wavefront distortion correction. *Infrared Laser Eng.* **2022**, *51*, 20210759.
17. Li, J.; Wen, L.; Liu, H.; Wei, G.; Cheng, X.; Li, Q.; Ran, B. A Novel SPGD Algorithm for Wavefront Sensorless Adaptive Optics System. *IEEE Photonics J.* **2023**, *15*, 7801109.
18. Zhao, H.; An, J.; Yu, M.; Lv, D.; Kuang, K.; Zhang, T. Nesterov-accelerated adaptive momentum estimation-based wavefront distortion correction algorithm. *Appl. Opt.* **2021**, *60*, 7177–7185. [PubMed]

19. Zhang, H.; Xu, L.; Guo, Y.; Cao, J.; Liu, W.; Yang, L. Application of AdamSPGD algorithm to sensor-less adaptive optics in coherent free-space optical communication system. *Opt. Express* **2022**, *30*, 7477–7490.
20. Song, H.; Fraanje, R.; Schitter, G.; Kroese, H.; Vdovin, G.; Verhaegen, M. Model-based aberration correction in a closed-loop wavefront-sensor-less adaptive optics system. *Opt. Express* **2010**, *18*, 24070–24084.
21. Dong, B.; Li, Y.; Han, X.-L.; Hu, B. Dynamic Aberration Correction for Conformal Window of High-Speed Aircraft Using Optimized Model-Based Wavefront Sensorless Adaptive Optics. *Sensors* **2016**, *16*, 1414. [CrossRef]
22. Verstraete, H.R.G.W.; Wahls, S.; Kalkman, J.; Verhaegen, M. Model-based sensor-less wavefront aberration correction in optical coherence tomography. *Opt. Lett.* **2015**, *40*, 5722–5725. [CrossRef]
23. Yang, G.; Liu, L.; Jiang, Z.; Guo, J.; Wang, T. Incoherent beam combining based on the momentum SPGD algorithm. *Opt. Laser Technol.* **2018**, *101*, 372–378.
24. Song, J.; Li, Y.; Che, D.; Wang, T. Numerical and experimental study on coherent beam combining using an improved stochastic parallel gradient descent algorithm. *Laser Phys.* **2020**, *30*, 085102.
25. Fang, Z.; Xu, X.; Li, X.; Yang, H.; Gong, C. SPGD algorithm optimization based on Adam optimizer. In *AOPC 2020: Optical Sensing and Imaging Technology*; Luo, X., Jiang, Y., Lu, J., Liu, D., Eds.; International Society for Optics and Photonics, SPIE: Bellingham, WA, USA, 2020; Volume 11567, p. 115672S.
26. Xu, L.; Wang, J.; Yang, L.; Zhang, H. Design and Performance Analysis of NadamSPGD Algorithm for Sensor-Less Adaptive Optics in Coherent FSOC Systems. *Photonics* **2022**, *9*, 77. [CrossRef]
27. Liu, H.; Li, Z.; Hall, D.; Liang, P.; Ma, T. Sophia: A scalable stochastic second-order optimizer for language model pretraining. *arXiv* **2023**, arXiv:2305.14342.
28. Assémat, F.; Wilson, R.W.; Gendron, E. Method for simulating infinitely long and non stationary phase screens with optimized memory storage. *Opt. Express* **2006**, *14*, 988–999. [PubMed]
29. Jiang, P.; Liang, Y.; Xu, J.; Mao, H. A new performance metric on sensorless adaptive optics imaging system. *Optik* **2016**, *127*, 222–226.
30. Bobroff, N.; Rosenbluth, A.E. Evaluation of highly corrected optics by measurement of the Strehl ratio. *Appl. Opt.* **1992**, *31*, 1523–1536.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.