Article

Design and Performance Analysis of NadamSPGD Algorithm for Sensor-Less Adaptive Optics in Coherent FSOC Systems

Li Xu ^{1,2}, Jianli Wang ¹, Leqiang Yang ¹ and Heng Zhang ^{1,2,*}

¹ Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China; xuli171@mails.ucas.edu.cn (L.X.); wangjianli@ciomp.ac.cn (J.W.); yangleqiang@ciomp.ac.cn (L.Y.)

² University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: zhangheng@ciomp.ac.cn

Abstract: Sensor-less adaptive optics (SLAO) based on stochastic parallel gradient descent (SPGD) is effective for the compensation of atmospheric turbulence in coherent free-space optical communication (CFSOC) systems. However, SPGD converges slowly and easily falls into local extremes. Therefore, we propose a novel NadamSPGD algorithm for efficient wavefront correction that combines Nesterov-accelerated adaptive moment estimation (Nadam) and SPGD. Specifically, Nesterov's accelerated gradient momentum (NAG) and adaptive gain coefficients are integrated to conventional SPGD to accelerate its convergence speed and avoid converging to extremum points. Theoretical analysis, numerical simulations and experimental results demonstrate that NadamSPGD can increase the convergence speed by ~50% and significantly improve the robustness of parameters, and thus more efficiently suppress the negative effects of atmospheric turbulence on mixing efficiency (ME) and bit error rate (BER). Our algorithm also presents better dynamic performance under strong turbulence and high Greenwood frequency conditions, and it is more suitable for real-time SLAO systems. This study proves that the NadamSPGD algorithm is suitable for SLAO in the CFSOC system and is a viable substitute for SPGD to improve the quality of optical communications.

Keywords: coherent free-space optical communication; sensor-less adaptive optics; stochastic parallel gradient descent; atmosphere turbulence

Citation: 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. <https://doi.org/10.3390/photonics9020077>

Received: 28 December 2021

Accepted: 26 January 2022

Published: 29 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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

The free space optical communication (FSOC) system has developed rapidly in modern communications due to the advantages of security, communication speed and license-free operation [1,2]. Recently, the coherent free space optical communication (CFSOC) has attracted more attention for its longer relay distance, higher sensitivity and better receiver selectivity compared with conventional FSOC [3–5]. However, the application of CFSOC is seriously hindered by atmospheric turbulence. The mixing efficiency (ME) and the bit error rate (BER) of the CFSOC system are severely degraded owing to the wavefront distortions caused at the receiver [6–8]. Adaptive optics (AO) is considered as one of the effective methods to compensate wavefront aberrations induced by atmospheric turbulence. Many applications of AO in CFSOC have been presented and made significant achievements [9–11]. In conventional AO systems, the Shack–Hartmann wavefront sensor (SH-WFS) is widely used and directly determines the system performance. However, due to the inherent shortcomings of its working principle, it is challenging to obtain satisfactory accuracy under strong scintillation or low optical power [12,13], which directly degrades system performances. Therefore, a sensor-less adaptive optics (SLAO) system based on multi-dimensional optimization algorithms is proposed. The SLAO system optimizes performance indicators of CFSOC directly depending on received images, and no longer needs wavefront reconstruction [14–18].

The multi-dimensional optimization algorithm has a significant impact on the performance of the SLAO system. Although various algorithms have been proposed to perform wavefront correction, the stochastic parallel gradient descent (SPGD) algorithm is most widely used in SLAO due to its simple model, easy implementation and few parameters [19,20]. However, SPGD has the shortcomings of slow convergence speed and easily falling into local extremes that limit its practical applications, especially in complex real-time systems [21]. In order to address the above problems, several attempts have been conducted in SPGD to speed up the convergence (decrease the iteration numbers) and/or avoid falling into local extremes. The decoupled SPGD (DSPGD) algorithm was proposed by Lachinova et al. to improve the convergence efficiency for compensation of atmospheric phase aberrations in the tiled fiber array system [22]. However, the available applications of DSPGD are limited by the requirement of prior knowledge of performance metrics. The modified SPGD based on the use of updating rules with finite memory and the frozen hypothesis was proposed by Gao et al. to correct the rapidly changing aero-optical aberrations in the AO system [23]. However, the modification increases the complexity of the algorithm, and the improvement effect is greatly affected by the perturbation. The multi-perturbation SPGD with the fast-decent mode and the modal basis updating mode was developed by Wu et al. to enhance the effectiveness of the algorithm [24]. However, this method needs to split the incoming beam into N sub-beams and use N wavefront correctors that increase the complexity of the optical system. The adaptive SPGD (ASPGD) integrating the momentum and adaptive gain coefficient estimation was proposed by Hu et al. to control a fast steering mirror (FSM) to achieve efficient fiber coupling [25]. The ASPGD avoids converging to the local extremum points and accelerates the convergence speed of SPGD to some extent. Yang et al. added pattern recognition in SPGD to check and prevent the algorithm from trapping into a local extreme in the incoherent beam combining system [26]. Song et al. modified SPGD with a momentum term derived from the Newtonian equation to improve the convergence speed and disturbance immunity of the coherent beam combining [27]. Ma et al. proposed an improved algorithm called adaptive gradient estimation SPGD for beam clean-up of a solid-state laser to obtain high output beam quality and increase the convergence speed and algorithm stability [28]. Summarizing, although the above studies obtained promising results, most of them were aimed at specific optical problems, lack applications in the CFSOC system, and cannot be applied to achieve efficient ME and decrease the BER directly. The feasibility of these algorithms in the CFSOC system requires further verification and implementation.

In order to solve the above problems, this study first analyzes the ME and BER of the CFSOC system according to coherent communication theory, and establishes the relationship between system performance indicators and the fitness of the optimization algorithm in SLAO. Inspired by Nesterov-accelerated adaptive moment estimation (Nadam) in deep learning [29], a novel algorithm named NadamSPGD is proposed. The proposed NadamSPGD combines Nesterov's accelerated gradient (NAG) momentum and the adaptive gain coefficients with SPGD to improve the correction speed and robustness of SLAO without noticeably increasing the complexity of the algorithm. The wavefront correction effect of the proposed algorithm is analyzed through numerical simulations and laboratory experiments. The results demonstrate that NadamSPGD can significantly increase the convergence speed and robustness in SLAO, and the negative influence of atmospheric turbulence on ME and BER can be efficiently suppressed, which is of great significance to the CFSOC system.

The structure of this paper is as follows. In Section 2, system models and working principles of CFSOC and SLAO are described and analyzed. The basic principles of the SPGD algorithm and the Nadam optimizer are introduced, and a novel hybrid algorithm NadamSPGD making full use of their characteristics is proposed in Section 3. In Section 4, the phase aberration model based on the Zernike polynomial is established, and related simulations and comparisons are presented to demonstrate the improved performance of NadamSPGD in the CFSOC system. We also performed a 97-element closed-loop SLAO

experiment to verify the feasibility of the proposed algorithm. Finally, Section 5 draws the conclusions.

2. System Model and Theoretical Analysis

2.1. CFSOC System Model with SLAO

The architecture of the typical CFSOC system with SLAO is illustrated in Figure 1. At the transmitting terminal, the laser source is modulated to generate a carrier signal. At the receiving terminal, the received optical signal is mixed with local oscillation to generate an intermediate frequency signal. Then, the proper demodulator and digital signal processor are used to further process and complete the subsequent processing. During the transmission of the laser through the atmospheric link, the wavefront is distorted by atmospheric turbulence. The SLAO system is introduced in the receiving terminal to compensate the influence of atmospheric turbulence and improve the quality of the optical signal.

Figure 1. Typical CFSOC system with SLAO.

The schematic of the SLAO system composed of a beam steering unit (BSU) and a high-order aberration correction unit (HCU) is shown in Figure 2. In BSU, by quickly capturing and tracking the beam, large skew of the laser beam, such as tilts and jitters, is corrected through an FSM. In HCU, the laser carrier signal falls on the deformable mirror (DM) and then is divided into two beams by a beam splitter (BS). A high-speed camera (HSC) is used to capture speckle images, and the current performance index of the system is obtained by calculating the energy concentration rate of the images. Then, the selected optimization algorithm is executed in the high-order aberration correction computer (HCC) and generates a voltage control signal of the DM according to the performance index. Finally, the high-voltage amplifiers are used to amplify the signal to a suitable voltage range and control the DM to correct the distorted wavefront. After SLAO correction, the performance of the CFSOC system can be improved with higher ME and lower BER. In this study, we only consider the high-order aberration correction [9,16].

Figure 2. Schematic of SLAO system.

2.2. DM Model in SLAO

We have designed and manufactured a 97-element continuous surface DM (CSDM) as the wavefront corrector in this study. By changing the surface shape of the mirror in real time according to the control voltage, the CSDM can efficiently correct wavefront aberrations. Generally, the influence of the CSDM is estimated with a Gaussian function [30,31]:

$$S_j(x, y) = \exp \left\{ \ln \omega \left[\frac{1}{d} \sqrt{(x - x_j)^2 + (y - y_j)^2} \right]^\alpha \right\}, \quad (1)$$

where ω is the coupling coefficient determined by the size of the electrode driver and the CSDM, (x_j, y_j) is the center coordinate of the j th actuator, α is the Gaussian index and d is the normalized interval between adjacent actuators. The phase compensation $\varphi(x, y)$ produced by the CSDM with 97 actuators is expressed as:

$$\varphi(x, y) = \sum_{j=1}^{97} u_j S_j(x, y), \quad (2)$$

where u_j is the voltage of the j th actuator in the range of the maximum possible voltage. It can be seen that $\varphi(x, y)$ is linear with the voltages applied to the actuators.

2.3. Theoretical Basis of the CFSOC

In a CFSOC system, the ME and BER are effective indicators to evaluate the performance. Assuming that the local oscillator (LO) is a plane wave and the intensity of the received optical signal (OS) is uniform, based on the theory of coherent detection, the total optical power of the combined beam in CFSOC is given by [32]:

$$I = \int_S \left\{ A_O^2 + A_S^2 + 2A_O A_S \cos[2\pi(f_S - f_O) + \Delta\varphi] \right\} ds, \quad (3)$$

where $\Delta\varphi = \varphi_S - \varphi_O$ is the phase difference between OS and LO, f_S and f_O are their frequencies, and A_S and A_O represent the optical amplitudes of OS and LO, respectively. Generally, a signal symbol transmission time is less than 1 ns in the CFSOC system, while the Greenwood frequency (GF) is on the millisecond scale. According to Taylor's turbulence hypothesis, the phase aberrations $\Delta\varphi$ caused by atmospheric turbulence can be considered frozen during the detection time, and given by the expression:

$$\Delta\varphi = \varphi(r) + \varphi(t), \quad (4)$$

where $\varphi(r)$ represents the spatial part of phase aberrations caused by atmospheric turbulence, which is time-independent, and $\varphi(t)$ denotes the temporal part of phase in the optical

signal, which is space coordinate-independent. When in homodyne detection, $f_s = f_O$, then ME can be defined by:

$$ME = \left[\int_S A_O A_S \cos(\Delta\varphi) ds \right]^2 / \left(\int_S A_S^2 ds \int_S A_O^2 ds \right). \quad (5)$$

According to Equation (5), ME is approximate to the Strehl ratio (SR) of the far-field images, defined by the ratio of far-field encircled energy to the diffraction limited encircled energy [33]. We only consider the spatial and temporal errors, then after A_O compensation, the mathematical expectation of the residual wavefront phase aberrations based on a CSDM can be expressed as [34–36]:

$$E(\sigma_\varphi^2) = E(\sigma_{fit}^2 + \sigma_{time}^2) = \left[\alpha_F \left(\frac{d}{r_0} \right)^{5/3} + \kappa \left(\frac{f_G}{f_{3dB}} \right)^{5/3} \right] (rad), \quad (6)$$

where σ_{fit}^2 is the spatial characteristic that represents the wavefront fitting error caused by the limited number of CSDM actuators, σ_{time}^2 is the temporal error due to the contradiction between fast-changing atmospheric turbulence and the finite closed-loop control bandwidth (CLCB) of the SLAO system, α_F is the fitting error coefficient, r_0 denotes the atmospheric coherent length, d is the equivalent interval of the actuators interval projected on the entrance pupil of the receiving antenna, κ is a constant which equals 1 for the plane wave, f_{3dB} denotes the CLCB and f_G represents the GF.

Thus, the relationship between ME and the average residual wavefront variance can be expressed as:

$$ME \propto SR = \exp \left\{ - \left[\alpha_f \left(\frac{d}{r_0} \right)^{5/3} + \kappa \left(\frac{f_G}{f_{3dB}} \right)^{5/3} \right] \right\}. \quad (7)$$

The BER of the synchronous binary phase shift keying (BPSK) coherent detection can be expressed as:

$$BER = \frac{1}{2} \operatorname{erfc} \left(\sqrt{ME \cdot 2N_p \delta} \right), \quad (8)$$

where the function erfc is the complementary error function, δ is the quantum efficiency of the receiver detector and N_p respects the number of photons received within a signal bit.

3. NadamSPGD Algorithm in SLAO

3.1. Fitness in SLAO

It is necessary to establish a connection between the fitness of optimization algorithms in SLAO and system evaluation indicators of the CFSOC system. First, we assume that the initial wavefront aberration of the laser carrier signal through the atmospheric channel is $\varphi_0(r, \theta)$, and solutions of the algorithm are 97-dimension vectors $u = \{u_1, u_2, \dots, u_{97}\}$; every component in the vectors respects the control voltage of each actuator in CSDM. In SLAO, the solutions are continuously updated and the compensation phase $\varphi(r, \theta)$ is generated according to Equation (2). Therefore, the residual phase aberration can be given by $\mathcal{C}\mathcal{E}(r, \theta) = \varphi_0(r, \theta) - \varphi(r, \theta)$. On the basis of the analysis in Section 2.3, this study takes the SR of $\mathcal{C}\mathcal{E}(r, \theta)$ as the fitness J in our algorithm to simplify the calculation [33,36,37]. In the following sections, the aim of the SLAO system is to optimize J to its ideal value. When the maximum J is obtained through the algorithm, we obtain the optimum voltage signals u and the best ME of the CFSOC system.

3.2. Conventional SPGD in SLAO

When using the conventional SPGD in SLAO, the compensation of the wavefront aberration can be described in the following steps. First, random tiny perturbation voltages $\Delta u^{(k)} = \{\Delta u_1, \Delta u_2, \dots, \Delta u_{97}\}$ that satisfy the Bernoulli distribution are applied to CSDM

control voltage vectors $u^{(k-1)} = \{u_1^{(k-1)}, u_2^{(k-1)}, \dots, u_{97}^{(k-1)}\}$ simultaneously to obtain the gradient estimation. The disturbances $\Delta u^{(k)}$ have fixed amplitude, i.e., $|\Delta u^{(k)}| = \Delta u$. Then, in brief, by using perturbed indicators values $J_{\pm}^{(k)} = J(u^{(k-1)} \pm \Delta u^{(k)})$, the variation of the performance metric can be calculated by $\Delta J^{(k)} = J_+^{(k)} - J_-^{(k)}$. Thereafter, the iterative formula for updating the CSDM control voltages can be expressed as:

$$u^{(k)} = u^{(k-1)} + \gamma \Delta u^{(k)} \Delta J^{(k)}, \quad (9)$$

where the superscript k denotes the number of iterations and γ is the positive gain coefficient. From Equation (9), as $u^{(k)}$ updates along the direction of the gradient descent, the performance metric J reaches an extremum after multiple iterations. The conventional SPGD only considers the current update vector, making it easy to form a local extremum and converge slowly if the gradient becomes flat or the curvature is large. Another limitation of SPGD is that all the optimization parameters use a single gain rate, and it is difficult to find a suitable gain rate value in real-world wavefront correction systems. For the above two problems in SPGD, a feasible solution is to add a momentum term to accumulate the past gradient encountered in the previous updates and set variable gain coefficients.

3.3. NadamSPGD

The Nesterov-accelerated adaptive moment estimation (Nadam) proposed by Dozat at the International Conference on Learning Representations (ICLR) in 2016 is an extension to adaptive moment estimation (Adam) that uses NAG momentum. Adam is an extension of gradient descent in neural networks that adds a first and second moment of the gradient, incorporates some inertia to updates and automatically adapts a learning rate for each parameter that is being optimized [38]. NAG is an extension to momentum terms; the update in NAG is performed using the gradient of the projected update to the parameter rather than the actual current variable value [39]. This has the effect of slowing down the search when the optimal value is located, rather than overshooting the minima as in traditional momentum. Nadam takes advantages of Adam and NAG, and can result in better performance of gradient-based optimization algorithms. Many studies have shown that in most cases, Nadam can improve the speed of convergence and the quality of learned models compared to a slew of related algorithms such as RMSProp, Momentum and Adam.

Inspired by Nadam, in view of the above two deficiencies in SPGD, we incorporate the Nesterov momentum and the adaptive gain coefficient estimation into the conventional SPGD, and propose a novel NadamSPGD algorithm to accelerate the optimization speed in SLAO.

Firstly, the gradient in NadamSPGD for the current step is approximated as [25]:

$$g^{(k)} = \partial J^{(k)} / \partial u^{(k)} = \Delta J^{(k)} \cdot \Delta u^{(k)} / (\Delta u)^2. \quad (10)$$

Next, the first momentum term is introduced into SPGD to accelerate its convergence. The exponentially decaying moving average over the parameters is used to give higher weight to the more recent value. The momentum is updated using the hyper-parameter μ as:

$$m^{(k)} = \mu m^{(k-1)} + (1 - \mu) g^{(k)}. \quad (11)$$

Then, the initialization bias correction strategy is used to offset the instability that initializing $m^{(k)}$ to zero may create. As momentum is most effective with a warming schedule, we parameterize μ by k , as well, just like in the NAG algorithm for completeness. Mathematically, the relevant formula is as follows:

$$\hat{m}^{(k)} = \mu^{(k+1)} m^{(k)} / \left[1 - \prod_{i=1}^{k+1} \mu^{(i)}\right] + (1 - \mu^{(k)}) g^{(k)} / \left[1 - \prod_{i=1}^k \mu^{(i)}\right]. \quad (12)$$

The $\hat{m}^{(k)}$ is the final form of the Nesterov momentum term in NadamSPGD. It uses the change from the last iteration to calculate the projected position of the variable, and then uses the derivative of the projected position in the calculation of the new position for the variable. By calculating the gradient of the projected position, it is equivalent to adding a correction factor to the acceleration that has been accumulated. Logically speaking, this will produce a superior gradient update. Nesterov momentum is known to reduce the number of iterations required and improve the rate of convergence of the optimization algorithm.

Furthermore, in order to solve the single gain rate in SPGD, we adjust the adaptive gain rate for different parameters by involving the second momentum term using the hyper-parameter v :

$$n^{(k)} = vn^{(k-1)} + (1-v)\left(g^{(k)}\right)^2, \quad (13)$$

$$\hat{n}^{(k)} = n^{(k)} / (1 - v^k), \quad (14)$$

where hyper-parameter v controls the exponential decay rate of the moving average, and the biased second momentum term $n^{(k)}$ is then bias-corrected by Equation (14) to avoid being initialized to zero at the start of the search, resulting in the bias-corrected estimate $\hat{n}^{(k)}$. The $\hat{n}^{(k)}$ sums up the weighted square results of the past gradients that indicate the uncentered variance of the gradients. During the updating process, we divide $\hat{n}^{(k)}$ to search the suitable gain rate adaptively.

As discussed above, we update CSDM control voltage vectors using NadamSPGD as follows:

$$u^{(k)} = u^{(k-1)} + \eta \hat{m}^{(k)} / \left(\sqrt{\hat{n}^{(k)}} + \varepsilon \right), \quad (15)$$

where η is the learning rate and ε is a parameter to avoid division by zero error, usually set to 10^{-8} .

The implementation of the NadamSPGD algorithm for SLAO is comprehensively described in Algorithm 1.

Algorithm 1 Pseudo code of the NadamSPGD algorithm.

Pseudo Code of the NadamSPGD Algorithm

Input: The learning rate α , the hyper-parameters μ and v , the constant ε , the amplitude of random perturbation voltages Δu , and the maximal number of iterations N .

Output : Calculated the control voltage vectors of CSDM $u = \{u_1, u_2, \dots, u_{97}\}$.

```

1 : Initialize control voltage vectors  $u^{(0)}$ , the 1st momentum term  $m^{(0)}$ , the 2nd momentum term  $n^{(0)}$ 
2 : for  $k = 1, \dots, N$  do
3 :   Randomly generate the perturbed voltages obeying the Bernoulli distribution  $\Delta u^{(k)}$ 
4 :   Obtain the evaluation functions under perturbation voltage  $J_{\pm}^{(k)} = J(u^{(k-1)} \pm \Delta u^{(k)})$ 
5 :   Obtain the change of the evaluation function  $\Delta J^{(k)} = J_{+}^{(k)} - J_{-}^{(k)}$ 
6 :   Calculate the gradient  $g^{(k)}$  (see Equation (10))
7 :   Calculate the bias – corrected Nesterov momentum  $\hat{m}^{(k)}$  (see Equations (11) and (12))
8 :   Calculate the bias – corrected 2nd momentum term  $\hat{n}^{(k)}$  (see Equations (13) and (14))
9 :   Update the control voltage  $u^{(k)}$  (see Equation (15))
10 : end for
```

Theoretically, the NadamSPGD algorithm is better than SPGD in terms of the gradient estimation and adaptive gain coefficient. The Nesterov momentum factor can improve convergence and suppress oscillations. The adaptive gain coefficient is adjusted in real time during the iterations, which can improve the convergence speed of corrections. NadamSPGD can be considered as an extension of SPGD that takes full advantages of gradients without noticeably increasing implementation complexity. The following simulations focus on the correction speed and robustness of the two algorithms.

4. Simulation and Experiment

4.1. Simulation Analysis

Zernike polynomials are widely used to describe the distorted wavefront caused by atmospheric turbulence. The wavefront $\varphi_0(r, \theta)$ can be considered as the two-dimensional functions decomposed by Zernike polynomials in polar coordinates [40]:

$$\varphi_0(r, \theta) = a_0 + \sum_{i=1}^{\infty} a_i Z_i(r, \theta), \quad (16)$$

where a_i is the coefficient of the i th Zernike polynomial. The 0th term and $Z_1(r, \theta)$, $Z_2(r, \theta)$ represent the piston and tilt aberrations along the X and Y directions, respectively, and can be corrected by BSU directly. In our simulations, the wavelength λ is set to 635 nm and $D/r_0 = 10$, then the 3rd to 35th terms in Zernike polynomials are added as the distorted wavefront to imitate atmosphere turbulence. The randomly generated initial Zernike coefficients from $a_3 - a_{35}$ are given in Figure 3a. The corresponding distorted phase of the original wavefront and the original point spread function (PSF) are shown in Figure 3b,c, respectively. The original wavefront is seriously distorted, and the initial ME is 0.2277. In the next simulations, the voltages of the 97-element CSDM are algorithmically calculated to compensate this wavefront aberration to compare the performance of algorithms. To facilitate our observation, the simulation results treat the optimization objective as ME. Since the calculation time is related to the performance of hardware systems, the performance improvement of algorithms is generally evaluated by comparing the number of iterations. In our study, we follow this common practice [16–21].

Figure 3. Random set of simulated data: (a) initial Zernike coefficients of the wavefront aberration; (b) corresponding distorted phase of the original wavefront; (c) corresponding original PSF.

In Figure 4, the corresponding residual wavefront aberration and PSFs correction results under different iterations with NadamSPGD are presented. The ME increases from 0.2277 to 0.9099 after 400 iterations, that is, four times before compensation. Clearly, most of the distortions have been well compensated by NadamSPGD.

Figure 4. Wavefront correction results under different iterations: (a) residual wavefront aberration and the corresponding PSF after 100 iterations; (b) residual wavefront aberration and the corresponding PSF after 200 iterations; (c) residual wavefront aberration and the corresponding PSF after 300 iterations; (d) residual wavefront aberration and the corresponding PSF after 400 iterations.

Next, we further verify the performance improvement of the proposed algorithm by simulations. Considering the randomness of algorithms, we execute each simulation 100 times. The optimization curves of ME and system BER based on SPGD and NadamSPGD with the number of iterations under their optimal parameters are illustrated in Figure 5. The learning rate $\alpha = 0.1$, the hyper-parameters $\mu = 0.999$ and $v = 0.99$, $\mu^{(k)} = \mu(1 - 0.5 \times 0.96^{k/250})$ as suggested in Reference [29], the constant $\varepsilon = 10^{-8}$, the amplitude of random perturbation voltages $\Delta u = 0.5$, the quantum efficiency $\delta = 1$, $N_p = 12$ and the BER is calculated according to Equation (8).

Figure 5. Comparison of the optimization curves between SPGD and NadamSPGD: (a) ME curves; (b) system BER curves.

From Figure 5, both SPGD and our algorithm can effectively compensate the aberration, improve the ME, and decrease the system BER. We treated 0.8 as the index of ME to observe and compare the feature of the two algorithms intuitively [41]. The SPGD reaches the index after at least 166 iterations, and at most 242 iterations in the worst case, averaging at

193 iterations. NadamSPGD converges after at least 81 iterations, and at most 163 iterations, with an average of 112 iterations, which is 58.03% of SPGD. In addition, for the standard deviation, SPGD fluctuates greatly, as they merely depend on the random disturbance at each iteration and the current gradient. The NadamSPGD considers both the current gradient and historical gradients during the iteration process; thus, it reduces the impact of randomness. In summary, NadamSPGD not only converges faster than SPGD, but also has better robustness to the randomness of disturbances. Figure 5b shows that the system BER is substantially improved after the wavefront aberration is compensated with NadamSPGD. The BER dropped from approximately 10^{-3} to 10^{-10} after 149 iterations. Due to the inherent limitation of SPGD, the value of BER cannot be suppressed below 10^{-10} until 262 iterations.

The amplitude of random perturbation voltages has a great influence on the performance of gradient-based algorithms. To verify the robustness of the two algorithms to perturbations, we evaluate SPGD and NadamSPGD under the same settings as previous simulations, except changing Δu from 0.01 to 1. The results are shown in Figure 6, from which we can see that SPGD is extremely sensitive to Δu . When Δu is 0.01, the correction speed of SPGD is extremely slow (Figure 6a), and when Δu increases to 0.1, the ME curve based on SPGD only reaches 0.3308 after 400 iterations (Figure 6b). When Δu increases to 1, SPGD prematurely converges to the local optimum (Figure 6c). However, the change in Δu has almost no effect on the correction performance of NadamSPGD, and it still works well with $\Delta u \in [0.01, 1]$ (shown in Figure 6d–f). The results demonstrate that NadamSPGD has robustness in a large range of Δu values, which makes it easy to apply to practical applications of SLAO.

Figure 6. Comparison of SPGD and NadamSPGD under different Δu values: (a) SPGD curve when $\Delta u = 0.01$; (b) SPGD curve when $\Delta u = 0.1$; (c) SPGD curve when $\Delta u = 1$; (d) NadamSPGD curve when $\Delta u = 0.01$; (e) NadamSPGD curve when $\Delta u = 0.1$; (f) NadamSPGD curve when $\Delta u = 1$.

According to the Kolmogorov turbulence model, the ratio of the receiving antenna aperture D to the atmospheric coherence length r_0 (D/r_0) can characterize the intensity of the atmospheric turbulence. In order to further explore the correction performance of the

two algorithms under different turbulence intensities, we analyze the relationship between ME and the iteration number under different D/r_0 in Figure 7.

Figure 7. Comparison of SPGD and NadamSPGD under different turbulence intensities: (a) correction curves using SPGD and NadamSPGD when $D/r_0 = 5$; (b) correction curves using SPGD and NadamSPGD when $D/r_0 = 15$; (c) number of iterations required for ME reaching to 0.8 of the two algorithms and the difference between them under different D/r_0 .

From Figure 7, it is obvious that NadamSPGD can effectively correct turbulence with different intensities. As the intensity of turbulence increases, the correction speed gradually slows down. When $D/r_0 = 5$, NadamSPGD achieves an ME of 0.8 after 38 iterations, while SPGD needs 55 iterations to achieve the equal correction effect. When D/r_0 increases to 15, NadamSPGD converges after 291 iterations, while SPGD requires 594 iterations. However, the gap between the number of iterations required for two algorithms becomes larger with the increasing turbulence intensity. The results show that the advantages of NadamSPGD become more significant with the increase in turbulence intensity.

Based on the theoretical analysis in Section 2.3, we know that the GF of the atmosphere and the CLCB of SLAO severely affect the performance of the CFSOC system. When $d = 0.06$, $\alpha_F = 0.28$, $r_0 = 0.15$ and $\kappa = 1$, the relationships between the ME, BER and CLCB under different GFs were obtained and are shown in Figure 8.

Figure 8. Relationships between ME, BER and CLCB under different GFs: (a) relationship between ME and CLCB under different GFs; (b) relationship between BER and CLCB under different GFs.

As Figure 8 illustrates, under different GFs, the higher the CLCB of the SLAO is, the higher the ME and lower the BER that will be achieved. Thus, a higher CLCB is necessary to guarantee the communication quality with the increase in GF. Since the resonant frequency of CSDM is very high (>5 kHz), its response time is extremely short, and its impact on CLCB is negligible compared to the computational delay introduced by algorithms [42]. As

a result, the computational delay of the algorithm used for correction is the dominant factor in determining the CLCB. Generally, a more efficient algorithm means fewer iterations and higher CLCB. To further show the ability of NadamSPGD in the time domain, we assume that the number of iterations is inversely proportional to CLCB and the previous simulation results of NadamSPGD have achieved 100 Hz CLCB, considering the processing capacity of the FPGA and GPU-based high-performance processing platform. Thereafter, the ME and BER at different GFs can be calculated according to Equations (7) and (8), and the results are shown in Figure 9.

Figure 9. Comparison of ME and BER between SPGD and NadamSPGD under different GFs. (a) ME of the two algorithms under different GFs. (b) BER of the two algorithms under different GFs.

Figure 9 illustrates that as the GF increases, the dynamic ability of SLAO based on the two algorithms degrades. However, the attenuation of SPGD is more serious than NadamSPGD, meaning that the proposed algorithm has better dynamic performance and is more suitable for real-time SLAO systems.

4.2. Experiment

In this study, we also compare the two algorithms based on our SLAO experimental platform to analyze and demonstrate the performance improvement of NadamSPGD in the actual systems. Figure 10 presents the setup and photograph of the SLAO experimental platform.

Figure 10. Setup and photograph of the SLAO experiment platform. (a) Optical layout of the platform. The optical signal is colored in orange and the electronic signal is colored in blue for easy distinction. (b) Photograph of the platform. The focal lengths of L1–L4 lenses are 10 mm, 10 mm, 5 mm and 15 mm, respectively.

As shown in Figure 10, the SLAO experimental platform is constructed based on an auto-collimating optical system. An optical fiber-coupled laser source with a wavelength of 635 nm is used to emit the laser beam. The laser beam is collimated by lens L1, then

reflected by the BS and further expanded by L3 and L4 to match the size of the 97-element CSDM. A field stop is added between lenses L3 and L4 to suppress the stray light. After correcting by the CSDM, the beam passes through the L4, L3 and BS again, and is narrowed and split. Finally, the beam reaches the lens L2, and the speckle images are captured by the HSC. The computer is used to process the captured images, execute the algorithms and control the CSDM through the driving circuits to compensate the aberrations.

In our experiments, the performances of SPGD and NadamSPGD are evaluated under the same initial conditions for a fair comparison. A set of CSDM initial control voltages are generated according to Zernike coefficients to generate the initial wavefront aberration, and then the control voltages are algorithmically calculated to gradually flatten the CSDM itself. The flattening process of CSDM can simulate the correction process of atmospheric turbulence. The energy concentration rate of speckle images is calculated to reflect the performance of SLAO. The optimal setting for SPGD is $\Delta u = 0.5$ and $\gamma = 2$, and the optimal setting for NadamSPGD is $\Delta u = 0.5$, $\alpha = 0.1$, $\mu = 0.999$, $v = 0.99$ and $\varepsilon = 10^{-8}$. The experimental results are shown in Figure 11, from which we find that SPGD corrects slowly. After 300 iterations, the initial ME reaches 0.5331 from 0.1142, which significantly affects the system performance. However, NadamSPGD dynamically adjusts the gain factor according to the gradient estimation to achieve a rapid convergence. The ME using NadamSPGD reaches 0.8828 after 300 iterations, which is 1.66 times faster than SPGD.

Figure 11. Comparison of correction results using SPGD and NadamSPGD on the SLAO experiment platform: (a) initial far-field image before correction; (b) SPGD correction result after 300 iterations; (c) SPGD residual wavefront ME as a function of iterations; (d) initial far-field image before correction; (e) NadamSPGD correction result after 300 iterations; (f) NadamSPGD residual wavefront ME as a function of iterations.

5. Conclusions

In this paper, a novel NadamSPGD algorithm combining Nadam and SPGD is proposed to compensate wavefront aberrations more effectively in CFSOC. The theoretical analysis and numerical simulations demonstrate that the negative influence of varying degrees of atmospheric turbulence on ME and BER of the CFSOC system can be suppressed by NadamSPGD. Specifically, by integrating the NAG momentum and adaptive gain coefficients into the conventional SPGD, the proposed algorithm can not only accelerate the correction speed by approximately 50%, but also improve the robustness of parameters over a large range ($\Delta u \in [0.01, 1]$) without noticeably increasing the complexity of the

algorithm. Simultaneously, the stronger the turbulence intensity, the more obvious are the advantages of NadamSPGD. In addition, NadamSPGD exhibits improved dynamic capabilities as the Greenwood frequency increases, and therefore, it is more suitable for real-time SLAO systems. Finally, the effectiveness of the proposed algorithm is evaluated on our SLAO experiment platform, and the experimental results indicate that NadamSPGD converges much faster. In conclusion, NadamSPGD is more effective for SLAO to improve the communication quality of CFSOC systems, and is a good substitute for SPGD.

From the findings of this paper, researchers can design SLAO systems with excellent performance in CFSOC based on NadamSPGD. The proposed algorithm may shed light on the application of SLAO-based wavefront correction technology, such as astronomical observation, fiber laser coherent synthesis and biological microscopic imaging. In the future, we will build a high-performance processing platform based on FPGA and GPU, and apply the NadamSPGD to dynamic aberration correction experiments.

Author Contributions: Conceptualization, L.X. and H.Z.; methodology, J.W.; software, L.X.; validation, L.Y. and H.Z.; data curation, H.Z.; writing—original draft preparation, L.X.; writing—review and editing, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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 conflict of interest.

References

1. Jaiswal, A.; Bhatnagar, M.R. Free-Space optical communication: A Diversity-Multiplexing tradeoff perspective. *IEEE Trans. Inform. Theory* **2019**, *65*, 1113–1125. [\[CrossRef\]](#)
2. Qin, D.; Wang, Y.; Zhou, T. Performance analysis of hybrid radio frequency and free space optical communication networks with cooperative spectrum sharing. *Photonics* **2021**, *8*, 108. [\[CrossRef\]](#)
3. Chen, M.; Liu, C.; Rui, D.; Xian, H. Performance verification of adaptive optics for satellite-to-ground coherent optical communications at large zenith angle. *Opt. Express* **2018**, *26*, 4230–4242. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Tian, R.; Wu, Z.; Ma, S.; Gu, Y.; Li, X. Design and performance analysis of probabilistically shaped QAM signals for coherent FSO systems with Gamma-Gamma turbulence channels. *Appl. Sci.* **2021**, *11*, 9805. [\[CrossRef\]](#)
5. Li, M.; Cvijetic, M. Coherent free space optics communications over the maritime atmosphere with use of adaptive optics for beam wavefront correction. *Appl. Opt.* **2015**, *54*, 1453–1462. [\[CrossRef\]](#)
6. Belmonte, A. Influence of atmospheric phase compensation on optical heterodyne power measurements. *Opt. Express* **2008**, *16*, 6756–6767. [\[CrossRef\]](#)
7. Ma, J.; Li, K.; Tan, L.; Yu, S.; Cao, Y. Performance analysis of satellite-to-ground downlink coherent optical communications with spatial diversity over gamma-gamma atmospheric turbulence. *Appl. Opt.* **2015**, *54*, 7575–7585. [\[CrossRef\]](#)
8. Zuo, L.; Dang, A.; Ren, Y.; Guo, H. Performance of phase compensated coherent free space optical communications through non-kolmogorov turbulence. *Opt. Commun.* **2011**, *284*, 1491–1495. [\[CrossRef\]](#)
9. Cao, J.; Zhao, X.; Liu, W.; Gu, H. Performance analysis of a coherent free space optical communication system based on experiment. *Opt. Express* **2017**, *25*, 15299–15312. [\[CrossRef\]](#)
10. Huang, J.; Mei, H.; Deng, K.; Kang, L.; Zhu, W.; Yao, Z. Signal to noise ratio of free space homodyne coherent optical communication after adaptive optics compensation. *Opt. Commun.* **2015**, *356*, 574–577. [\[CrossRef\]](#)
11. Takenaka, H.; Toyoshima, M.; Takayama, Y. Experimental verification of fiber-coupling efficiency for satellite-to-ground atmospheric laser downlinks. *Opt. Express* **2012**, *20*, 15301–15308. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Primmerman, C.; Price, T.; Humphreys, R.; Zollars, B.; Barclay, H.; Herrmann, J. Atmospheric-compensation experiments in strong-scintillation conditions. *Appl. Opt.* **1995**, *34*, 2081–2088. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Weyrauch, T.; Vorontsov, M. Atmospheric compensation with a speckle beacon in strong scintillation conditions: Directed energy and laser communication applications. *Appl. Opt.* **2005**, *44*, 6388–6401. [\[CrossRef\]](#)
14. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Zhang, S.; Wang, R.; Wang, Y.; Mao, H.; Xu, G.; Cao, Z.; Xuan, L. Extending the detection and correction abilities of an adaptive optics system for free-space optical communication. *Opt. Commun.* **2021**, *482*, 126571. [\[CrossRef\]](#)

16. Li, Z.; Cao, J.; Zhao, X.; Liu, W. Swarm intelligence for atmospheric compensation in free space optical communication—Modified shuffled frog leaping algorithm. *Opt. Laser Technol.* **2015**, *66*, 89–97. [[CrossRef](#)]
17. Gu, H.; Liu, M.; Liu, H.; Yang, X.; Liu, W. An algorithm combining convolutional neural networks with SPGD for SLAO in FSOC. *Opt. Commun.* **2020**, *475*, 126243. [[CrossRef](#)]
18. He, X.; Zhao, X.; Cui, S.; Gu, H. A rapid hybrid wave front correction algorithm for sensor-less adaptive optics in free space optical communication. *Opt. Commun.* **2018**, *429*, 127–137. [[CrossRef](#)]
19. Cao, J.; Zhao, X.; Li, Z.; Liu, W.; Song, Y. Stochastic parallel gradient descent laser beam control algorithm for atmospheric compensation in free space optical communication. *Optik* **2014**, *125*, 6142–6147. [[CrossRef](#)]
20. Huang, Z.; Tang, X.; Zhang, D.; Wang, X.; Hu, Q.; Li, J.; Liu, C. Coherent beam combination of ten fiber arrays via stochastic parallel gradient descent algorithm. *J. Opt. Technol.* **2015**, *82*, 16–20. [[CrossRef](#)]
21. 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. [[CrossRef](#)] [[PubMed](#)]
22. Lachinova, S.L.; Vorontsov, M.A. Performance analysis of an adaptive phase-locked tiled fiber array in atmospheric turbulence conditions. *Target-in-the-Loop Atmos. Track. Imaging Compens. II* **2005**, 5895, 589500.
23. Gao, Q.; Jiang, Z.; Yi, S.; Xie, W.; Liao, T. Correcting the aero-optical aberration of the supersonic mixing layer with adaptive optics: Concept validation. *Appl. Opt.* **2012**, *51*, 3922–3929. [[CrossRef](#)]
24. Wu, K.; Sun, Y.; Huai, Y.; Jia, S.; Chen, X.; Jin, Y. Multi-perturbation stochastic parallel gradient descent method for wavefront correction. *Opt. Express* **2015**, *23*, 2933–2944. [[CrossRef](#)] [[PubMed](#)]
25. Hu, Q.; Zhen, L.; Mao, Y.; Zhu, S.; Zhou, X.; Zhou, G. Adaptive stochastic parallel gradient descent approach for efficient fiber coupling. *Opt. Express* **2020**, *28*, 13141–13154. [[CrossRef](#)] [[PubMed](#)]
26. Yang, G.; Liu, L.; Jiang, Z.; Wang, T.; Guo, J. Improved SPGD algorithm to avoid local extremum for incoherent beam combining. *Opt. Commun.* **2017**, *382*, 547–555. [[CrossRef](#)]
27. Song, J.; Li, Y.; Che, D.; Guo, J.; Wang, T. Coherent beam combining based on the SPGD algorithm with a momentum term. *Optik* **2020**, *202*, 163650. [[CrossRef](#)]
28. Ma, S.; Yang, P.; Lai, B.; Su, C.; Zhao, W.; Yang, K.; Jin, R.; Cheng, T.; Xu, B. Adaptive Gradient Estimation Stochastic Parallel Gradient Descent Algorithm for Laser Beam Cleanup. *Photonics* **2021**, *8*, 165. [[CrossRef](#)]
29. Dozat, T. Incorporating Nesterov momentum into Adam. In Proceedings of the International Conference on Learning Representations, San Juan, Puerto Rico, 2–4 May 2016.
30. Alda, J.; Boreman, G. Zernike-based matrix model of deformable mirrors: Optimization of aperture size. *Appl. Opt.* **1993**, *32*, 2431–2438. [[CrossRef](#)]
31. Huang, L.; Rao, C.; Jiang, W. Modified Gaussian influence function of deformable mirror actuators. *Opt. Express* **2008**, *16*, 108–114. [[CrossRef](#)]
32. Liu, C.; Chen, S.; Li, X.; Xian, H. Performance evaluation of adaptive optics for atmospheric coherent laser communications. *Opt. Express* **2014**, *22*, 15554–15563. [[CrossRef](#)] [[PubMed](#)]
33. Liu, C.; Chen, M.; Chen, S.; Xian, H. Adaptive optics for the free-space coherent optical communications. *Opt. Commun.* **2016**, *361*, 21–24. [[CrossRef](#)]
34. Greenwood, D. Bandwidth specification for adaptive optics systems. *J. Opt. Soc. Am.* **1977**, *67*, 390–393. [[CrossRef](#)]
35. Tyson, R.; Wizinowich, P. *Principles of Adaptive Optics*; CRC: Boca Raton, FL, USA, 1991.
36. Huang, J.; Liu, C.; Deng, K.; Yao, Z.; Xian, H.; Li, X. Probability of the residual wavefront variance of an adaptive optics system and its application. *Opt. Express* **2016**, *24*, 2818–2928. [[CrossRef](#)] [[PubMed](#)]
37. Mahajan, V. Strehl ratio for primary aberrations in terms of their aberration variance. *J. Opt. Soc. Am.* **1983**, *73*, 860–861. [[CrossRef](#)]
38. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
39. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the importance of initialization and momentum in deep learning. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013.
40. Noll, R.J. Zernike polynomials and atmospheric-turbulence. *J. Opt. Soc. Am.* **1976**, *66*, 207–211. [[CrossRef](#)]
41. Cui, S.; Zhao, X.; He, X.; Gu, H. A Quick Hybrid Atmospheric-interference Compensation Method in a WFS-less Free-space Optical Communication System. *Curr. Opt. Photonics* **2018**, *2*, 612–622.
42. Yang, L.; Yao, K.; Wang, J.; Cao, J.; Lin, X.; Liu, X.; Liu, W.; Gu, H. Performance analysis of 349-element adaptive optics unit for a coherent free space optical communication system. *Sci. Rep.* **2019**, *9*, 1–11. [[CrossRef](#)] [[PubMed](#)]