



Article Adaptive Gradient Estimation Stochastic Parallel Gradient Descent Algorithm for Laser Beam Cleanup

Shiqing Ma^{1,2,3}, Ping Yang^{1,2,*}, Boheng Lai^{1,2}, Chunxuan Su^{1,2,3}, Wang Zhao^{1,2}, Kangjian Yang^{1,2}, Ruiyan Jin^{1,2}, Tao Cheng^{1,2} and Bing Xu^{1,2}

- ¹ Key Laboratory on Adaptive Optics, Chinese Academy of Sciences, Chengdu 610209, China; mashiqing16@mails.ucas.ac.cn (S.M.); laiboheng@ioe.ac.cn (B.L.); suchunxuan16@mails.ucas.ac.cn (C.S.); zhaowang@ioe.ac.cn (W.Z.); yangkangjian@ioe.ac.cn (K.Y.); jinruiyan@ioe.ac.cn (R.J.); chengtao@ioe.ac.cn (T.C.); bingxu@ioe.ac.cn (B.X.)
- ² Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China
- ³ School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
 - * Correspondence: yangping@ioe.ac.cn

Abstract: For a high-power slab solid-state laser, obtaining high output power and high output beam quality are the most important indicators. Adaptive optics systems can significantly improve beam qualities by compensating for the phase distortions of the laser beams. In this paper, we developed an improved algorithm called Adaptive Gradient Estimation Stochastic Parallel Gradient Descent (AGESPGD) algorithm for beam cleanup of a solid-state laser. A second-order gradient of the search point was introduced to modify the gradient estimation, and it was introduced with the adaptive gain coefficient method into the classical Stochastic Parallel Gradient Descent (SPGD) algorithm. The improved algorithm accelerates the search for convergence and prevents it from falling into a local extremum. Simulation and experimental results show that this method reduces the number of iterations by 40%, and the algorithm stability is also improved compared with the original SPGD method.

Keywords: stochastic parallel gradient descent algorithm; beam cleanup; slab laser

1. Introduction

High-power solid-state lasers have been widely used in many fields such as machinery, medical, and defense [1–6]. With the development of lasers, applications such asspacebased laser space debris cleanup have put forward new requirements for laser systems. These scenarios require the laser system to have a high average output power and high output beam quality and require the system to be simple, compact, and miniaturized. However, the thermal effect of the laser introduces wavefront distortion, which leads to degradation of beam quality [7-9]. With the rapid development of gain materials, pump sources, laser extraction structures, and adaptive optics (AO) technology, the output power and beam quality of solid-state lasers have been significantly improved [10-16]. The adaptive optics technology measures the wavefront distortion through the wavefront sensor and uses the control system to control the deformable mirror to correct the wavefront distortion, which can effectively improve the laser beam quality [15,17]. Compared with classical AO, the wavefront sensor-less AO (WFSless AO) system does not need a wavefront sensor to measure the wavefront distortion; the control signal of the correction element is obtained by optimization algorithm, so it has a simpler structure and it is easier for the system to achieve lightweight requirements [18–21].

Commonly used optimization algorithms include the hill climbing algorithm [20], genetic algorithm [22], simulated annealing algorithm [23,24], stochastic parallel gradient descent (SPGD) algorithm [25–32], etc. Among these, the SPGD algorithm is more efficient



Citation: 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. https://doi.org/10.3390/ photonics8050165

Received: 13 April 2021 Accepted: 14 May 2021 Published: 19 May 2021

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



Copyright: © 2021 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/). and more applicable than other algorithms in AO systems and has been widely used for laser beam cleanup [27,33]. However, the convergence speed of the classical SPGD algorithm is slower, and it would have a probability of converging to a local extremum [34]. In recent years, a variety of improved algorithms aimed at increasing the speed of the algorithm were proposed and widely used to optimize the connection weights of deep neural networks, such as the AdaGrad algorithm [28], RMSProp algorithm [30], and Adam algorithm [31]. In 2020, Hu et al. proposed an adaptive stochastic parallel gradient descent method based on the Adam algorithm to achieve efficient fiber coupling [32]. Although these methods have achieved excellent results, most of them were aimed at specific problems in application and cannot be directly used to achieve beam cleanup. In addition, these algorithms introduce momentum and multiple hyperparameters, which increase the difficulty of calculation and parameter selection.

In our work, we developed an adaptive gradient estimation stochastic parallel gradient descent (AGESPGD) algorithm in the WFSless AO system for beam cleanup of solid-state lasers. In the iterative process of the algorithm, both the gradient and the second-order gradient are considered. The gradient value controls the direction of the iteration, and the second-order gradient is introduced to adaptively modify the gradient estimate factor to increase the convergence and prevent falling into local extremes. This article is organized as follows: Firstly, we analyze and derive the principle and iterative formula of the AGESPGD method. Second, this method is verified by a large number of simulations. Finally, the performance of the algorithm is tested in a real experimental system. Both simulation and experimental results prove that our proposed method has a significant improvement in convergence speed and robustness to search step-size.

2. Principle of the AGESPGD Algorithm

2.1. Model of Laser Beam Cleanup System

Wavefront sensor-less adaptive optics technology is a simple but effective method for beam cleanup in solid-state laser systems. The schematic of the WFSless AO system is shown in Figure 1. It consists of a deformable mirror (DM), a set of solid-state laser systems, a CCD camera, and a set of optical focusing lenses. The distorted beam is converged on the CCD camera through the focusing lens, and the camera can collect the degraded far-field spot.



Figure 1. Schematic of wavefront sensor-less adaptive optics (WFSless AO).

The basic principle of WFSless AO is using an optimization algorithm to search for the best voltage matrix $U = (u_1, u_2, u_3, \dots, u_n)$ (*n* is the number of DM actuator) to minimize the residual wavefront surface $E(x, y) = \Phi(x, y) - \Psi(x, y)$. $\Phi(x, y)$ represents the wavefront

aberration of the beam. Ψ (*x*,*y*) represents the surface shape of the DM driven by voltage matrix *U*. The expression of Ψ (*x*,*y*) is

$$\Psi(x,y) = \sum_{i=1}^{n} u_i f_i(x,y).$$
 (1)

where *i* is the number of actuators, and $f_i(x,y)$ is the surface function generated by applying a unit voltage to the *i*-th actuator. The beam quality of the degraded far-field image can be used as a merit function J(U) to evaluate the magnitude of the wavefront residual E(x, y). The merit function usually selects a certain evaluation index of beam quality, such as sharpness function, Strehl ratio (SR), power in the bucket (PIB), etc. This method does not require a complex wavefront sensor system to measure wavefront distortion. Therefore, the laser beam cleanup could be achieved by searching for the best voltage matrix of DM to optimize the merit function, as shown in Equation (2).

$$\max_{U} J(U) \tag{2}$$

2.2. AGESPGD

For the classic SPGD method, the convergence speed and direction are only determined by the gradient value of the search point. When the gradient of the merit function is variable, the algorithm performance is seriously degraded by this method. Specifically, the algorithm would repeatedly oscillate near the optimal point or even fall into the local optimal because the search speed accelerates as the gradient gradually increases. When the gradient gradually decreases, the algorithm convergence speed becomes extremely slow because the search speed also decreases. Therefore, we introduce a value of secondorder gradient of current point to predict the change of the gradient, which makes the convergence process more effective.

The classical SPGD algorithm originated from the gradient descent method in artificial neural networks [25,35,36]. The principle is to apply a combination of interference voltages $\Delta U = (\delta u_1, \delta u_2, \delta u_3, \dots, \delta u_n)$ to multiple actuators to obtain changes in the value of the merit function [33]:

$$\delta J = J^{+} - J = J(u_{1} + \delta u_{1}, u_{2} + \delta u_{2}, \cdots, u_{n} + \delta u_{n}) - J(u_{1}, u_{2}, \cdots, u_{n}),$$
(3)

where J^+ and J^- represent the value of the merit function after and without disturbance, respectively. Taylor expansion of Equation (3) can be obtained:

$$\delta J = \sum_{i=1}^{n} \frac{\partial J}{\partial u_i} \delta u_i + O([\delta u_i]).$$
⁽⁴⁾

By simply transforming and ignoring higher-order terms, we can obtain

$$\delta J \delta u_i = \frac{\partial J}{\partial u_i} (\delta u_i)^2 + \sum_{j \neq i}^n \frac{\partial J}{\partial u_j} \delta u_j \delta u_i + \cdots .$$
(5)

The disturbance voltage applied to each actuator is independently and randomly distributed, so we can get $\sum_{i \neq j}^{n} \frac{\partial J}{\partial u_i} \delta u_i \delta u_j = 0$ and an estimate of the gradient can be approximately written as

$$\overline{J'}_i \approx \delta J \delta u_i \approx \Delta J_i \delta u_i / (\delta u_i)^2.$$
(6)

According to mathematical proof, when the random characteristics of random disturbances are equal in amplitude $|\delta u_i| = du$, the random characteristics of each disturbance

degenerate to Bernoulli distribution and would not affect the convergence. According to the previous derivation, the iterative formula can be rewritten as

$$u_i^{t+1} = u_i^t - \gamma \overline{J'}_i = u_i^t - \gamma (J^+ - J) \delta u_i.$$
⁽⁷⁾

In practical applications, the bilateral disturbance strategy is usually used to improve the accuracy of gradient estimation [30], that is, to apply positive disturbance and negative disturbance to the actuator and measure the corresponding index function value. $\overline{J'}_i$ in Equation (7) can be written as

$$I'_{i} = (J^{+} - J^{-})\delta u_{i}.$$
(8)

where J^+ and J^- represent the value of the merit function after applying positive and negative disturbances, respectively. The relationship between J^+ , J^- , and J is described in Figure 2. Two virtual points (P_1 and P_2) are added between them, which represent the half-step length of the positive and negative disturbances, respectively, and their corresponding mathematical expressions for gradient estimation are

$$\bar{J}_{i1}' = (J - J^{-})\delta u_i/2, \bar{J}_{i2}' = (J^{+} - J)\delta u_i/2.$$
(9)



Figure 2. Schematic of the relationship between indexes.

From these two virtual points, we can calculate the second-order gradient at the point $\{u_i\}$ as

$$\overline{J''_i} \approx \left(\overline{J'_{i1}} - \overline{J'_{i2}}\right) \delta u_i
= \left[(J^+ - J) \delta u_i / 2 - (J - J^-) \delta u_i / 2 \right] / \delta u_i
= C (J^+ + J^- - 2J).$$
(10)

According to the previous analysis, the second-order gradient is introduced to control gradient estimation $(\overline{J'}_i)$. Therefore, by introducing a parameter $\eta = \frac{C}{J''}$ (*C* is constant) that is negatively related to the second-order gradient J'' into Equation (7), we get the iterative formula of the AGESPGD as follows:

$$u_i^{t+1} = u_i^t - \gamma \overline{J'}_i / \overline{J''}_i. \tag{11}$$

Algorithm 1 shows the process of the AGESPGD algorithm. The termination condition is generally selected to reach a certain value or reach a certain number of iterations.

Algorithm 1. The procedure of the Adaptive Gradient Estimation Stochastic Parallel Gradient Descent (AGESPGD) method

Require: the stochastic merit function $J(u_1, u_2, u_3, ..., u_n)$, the amplitude of disturbance δu , the gain rate γ , the certain value of J_0 , and the maximal number of iterations T.

Ensure: The optimal parameters γ and duInitialize the indexes J = 0, J' = 0, J'' = 0. Initialize the voltages $u_1, u_2, u_3, \dots, u_n$. **while** $t \leq T$ or $J \geq J_c$ **do** Randomly generate the values of $\Delta U_t = (\delta u^t_1, \delta u^t_2, \delta u^t_3, \dots, \delta u^t_n)$. Calculate the value of index $J^+ J^- J$. $\overline{J''_i} = (J^+ - J^-) \delta u_i$.. $\overline{J''_i} = J^+ + J^- - 2J$.. $u_i^{t+1} = u_i^t - \gamma \overline{J'_i} / \overline{J''_i}$. t = t + 1. **end while**

In order to verify our proposed AGESPGD method, we constructed a unary function f(x). This function has multiple local minima; its curve is shown in Figure 3, and its expression is given by Equation (12). From the simulation results, the classical SPGD only converge to the local minimum, while the AGESPGD method can converge to the optimal solution and only requires 88 iterations. From this simple simulation, we can see that our proposed AGESPGD algorithm converges faster and avoids the local extremum problem.

$$f(x) = (x-3)^2 + 3\sin(2\pi x) + 5.$$
(12)



Figure 3. The optimization trace of the Stochastic Parallel Gradient Descent (SPGD) and ASPGD for minimizing the function in Equation (12).

3. Simulation

We performed numerical simulations to verify the application of our proposed AGE-SPGD method in a laser beam cleanup system. The scheme of the beam cleanup system is shown in Figure 1. The wavefront aberration $\Phi(x, y)$ of the input beam is generated by the Zernike polynomial [37]. $\Phi(x, y)$ can be described as a linear combination of various orders of the Zernike polynomial:

$$\Phi(x,y) = \sum_{i=1}^{j} a_i Z_i(x,y).$$
(13)

 $Z_i(x, y)$ is the *i* order Zernike polynomial expression, and a_i is the corresponding coefficient of polynomials. The aberration represented by Z_0 is the overall phase shift, which will not cause degradation of the beam quality. The aberration Z_1 and Z_2 respectively represent the tilt in two directions, which can be corrected by beam pointing control. Therefore, in the simulation, wavefront aberration was simulated using 3~5 order Zernike polynomial. All coefficients were randomly generated. The relationship between the far-field intensity distribution of the beam and the wavefront aberration can be expressed as [25]

$$I(x',y') = \left| \iint A_0(x,y) \exp\{j[\Phi(x,y) - \Psi(x,y)]\} \bullet \exp^{\frac{jk[(x'-x)^2 + (y'-y)^2]}{2f}} dxdy \right|^2.$$
(14)

where (x',y') is the rectangular coordinate in the input plane of the camera, (x,y) is the pupil plane coordinate, $k=2\pi/\lambda$ is wave number, λ is the wavelength of the input wavefront, fis the focal length of the optical focusing lens, and j is the imaginary unit. $A_0(x,y)$ is the amplitude of the input wavefront, which is generally set to be unity. In the simulation, power in the bucket (PIB) was used as the optimization merit function [38,39], and its expression is represented by Equation (15), where S_0 is the area of the bucket.

$$J = \frac{\iint_{S_0} I(x', y') dx dy}{\iint_{S} I(x', y') dx dy}.$$
 (15)

The parameters of the simulation system were set as follows: beam aperture D = 80 mm, wavelength $\lambda = 1064$ nm, optical lens f = 5 m, camera pixel size $pix = 7.4 \mu$ m. A total of 150 sets of random aberrations were generated; one part was used to obtain the best parameters of the algorithm, and the other part was used to compare the performance of the two methods. The iteration curve is shown in Figure 4, and the parameters of the two methods were set as (SPGD: $\delta_u = 0.01 \ y = 10$ and AGESPGD: $\delta_u = 0.1 \ \gamma = 0.05$). The classical SPGD requires an average of 215 iterations to complete convergence. Our developed AGESPGD method only needed an average of 124 iterations to complete the convergence in 100 simulation experiments, and its average convergence was about 40% faster than the classical SPGD.



Figure 4. Iterative curves of AGESPGD and SPGD methods.

Subsequently, we also analyzed the peak–valley (PV) value and root-mean-square (RMS) value of the residual wavefront after correction, and the beam quality (timesdiffraction-limit factor, β) of the far-field spot. Figure 5 shows the comparison result of a certain random aberration. For the classical SPGD method, the average result of 100 random aberrations was that the PV and RMS were improved from 2.3 and 0.127 to 0.27 and 0.0025, respectively (unit is wavelength $\lambda = 1064$ nm), and the beam quality improved from 3.84 times the diffraction limit to 1.37 times the diffraction limit. For the AESPGD method, the PV and RMS were improved from 2.3 and 0.122 to 0.27 and 0.0028, respectively (unit is wavelength $\lambda = 1064$ nm), and the beam quality was improved from 3.84 times the diffraction limit to 1.35 times the diffraction limit.



Figure 5. Simulation result of AGESPGD and SPGD methods.

To further analyze the performance of the two methods, we used different values of δu to evaluate the two methods. The iteration curve is shown in Figure 6. The values of δ_u were set to 0.025, 0.01, and 0.005. The AGESGD method can converge stably for different values of δ_u . However, the classical SPGD method would show the phenomenon of slow convergence or oscillating back and forth. Overall, the AGESPGD method exhibited better performance in terms of convergence and robustness to δu than the classic SPGD in these simulations.



Figure 6. Simulation comparison of the AGESPGD and SPGD under different δ_u .

4. Experiment

In this section, the AGESPGD method was additionally studied in a real beam cleanup system. The experimental configuration is shown in Figure 7. The number of actuators of the deformable mirror (DM) was 59, and the laser used was an all-solid-state single-longitudinal-mode infrared laser with a wavelength of 1064 nm and an aperture of 80 mm. The beam passed through DM1, TM, and DM2 in sequence. DM1 was used to generate analog aberrations, and TM was used to control light stability to prevent speckle jitter. DM2 was used for aberration correction. Then, the beam entered a far-field imaging system that included a set of focusing lenses with an equivalent focal length of 5 m and a CCD camera (camera parameters: BAUMER, TXG03, CCD, 656×494 pixels, 7.4 µm pixel pitch).



Figure 7. Experimental system for laser beam cleanup.

We set the same initial point for the two methods and set the best parameter combination, respectively. DM1 generated a series of aberrations that were calculated from the wavefront aberrations collected by a real slab solid-state laser. Figure 8 shows 10 sets of iteration curves with different aberrations. It can be seen from the experimental results that the classical SPGD method required about 200 iterations to achieve a stable value. The AGESPGD algorithm only needed about 120 times, and the convergence was increased by nearly 40%. Figure 9 shows the far-field spots corrected by the two respective methods. After correction, the beam quality was improved from 3.71 times diffraction limit to 1.63 imes by the classical SPGD method and from 3.71 times to 1.61 times by the AGESPGD method. Experiments showed that the convergence of the AGESPGD method was greatly improved compared with the classical SPGD method.



Figure 8. Experimental results of laser beam cleanup.



Figure 9. Far-field intensity distribution of laser beam cleanup experiment.

Figure 10 shows the iterative curves of the two methods under different search steps δ_u . When δ_u was within the interval of [0.001, 0.05], the AGESPGD method could converge stably, while the classical SPGD method could not. Experimental results showed that the AGESPGD method not only had a much faster convergence rate than classical SPGD, but also had a better robustness to δ_u .



Figure 10. Iterative curves of the two methods under different search steps δ_u .

5. Conclusions

In this article, an adaptive gradient estimation stochastic parallel gradient descent (AGESPGD) method is developed for solid-state laser beam cleanup. The gradient estimate is modified by the second-order gradient of the search point. This modification is combined with adaptive gain estimation to improve the classical SPGD. The improved algorithm is able to accelerate the convergence speed and prevent it from falling to local extremum. The simulation and experiment results show that our developed method converges much faster than the classical SPGD method, and the robustness of the algorithm in searching step-size is also improved. Compared with the classical SPGD, the iteration number of the AGESPGD is reduced by 40%, and this method can converge stably when the search step is in the range of [0.001, 0.05]. At the same time, this method can improve the beam quality from 3.71 to 1.61 times the diffraction limits.

Author Contributions: Conceptualization, S.M. and P.Y.; Data curation, S.M.; Formal analysis, S.M.; Investigation, S.M., B.L., and C.S.; Methodology, S.M.; Resources, P.Y.; Software, S.M.; Validation, S.M., P.Y., and B.L.; Visualization, S.M.; Writing—original draft, S.M.; Writing—review and editing, S.M., P.Y., B.L., C.S., W.Z., K.Y., R.J., T.C., and B.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- 1. Xu, J.; Zhao, S.; Hou, R.; Zhan, S.; Shi, L.; Wu, J.; Fang, S.; Li, Y. Laser jamming technique research based on combined fiber laser. *Appl. Opt.* **2009**, *41*, 486–494.
- 2. Liu, S.; Farahmand, P.; Kovacevic, R. Optical monitoring of high power direct diode laser cladding. *Opt. Laser Technol.* 2014, 64, 363–376. [CrossRef]
- 3. Webster, P.J.L.; Wright, L.G.; Ji, Y.; Galbraith, C.M.; Kinross, A.W.; Vlack, C.V.; Fraser, J.M. Automatic laser welding and milling with in situ inline coherent imaging. *Opt. Lett.* **2014**, *39*, 6217–6220. [CrossRef]
- Bindra, K.S.; Upadhyaya, B.N. Applications of High Power Solid State Lasers in Nuclear Power Programme. Proc. Natl. Acad. Sci. India Sect. A 2018, 88, 375–386. [CrossRef]
- 5. Kawahito, Y.; Wang, H.; Katayama, S.; Sumimori, D. Ultra high power (100 kW) fiber laser welding of steel. *Opt. Lett.* **2018**, 43, 4667–4670. [CrossRef]
- Pietrzak, A.; Zorn, M.; Meusel, J.; Huelsewede, R.; Sebastian, J. Development of highly efficient laser bars emitting at around 1060 nm for medical applications. In *High-Power Diode Laser Technology XVI*; International Society for Optics and Photonics: San Francisco, CA, USA, 2018.
- Zelenogorsky, V.; Solovyov, A.; Kozhevatov, I.; Kamenetsky, E.; Rudenchik, E.; Palashov, O.; Silin, D.; Khazanov, E.J.A.O. High-precision methods and devices for in situ measurements of thermally induced aberrations in optical elements. *Appl. Opt.* 2006, 45, 4092. [CrossRef] [PubMed]
- 8. Antipov, O.L.; Anashkina, E.; Fedorova, K.A.J.Q.E. Electronic and thermal lensing in diode end-pumped Yb: YAG laser rods and discs. *Quantum Electron.* 2009, *39*, 1131. [CrossRef]
- 9. Slezak, O.; Lucianetti, A.; Divoky, M.; Sawicka, M.; Mocek, T. Optimization of Wavefront Distortions and Thermal-Stress Induced Birefringence in a Cryogenically-Cooled Multislab Laser Amplifier. *IEEE J. Quantum Electron.* **2013**, *49*, 960–966. [CrossRef]
- 10. Martin, W.S.; Chernoch, J.P. Multiple Internal Reflection Face-Pumped Laser. U.S. Patent 3,633,126, 4 January 1972.
- 11. Guofu, W. New Development of Laser Crystal Materials. J. Synth. Cryst. 1998, 27, 390-395.
- Redmond, S.; McNaught, S.; Zamel, J.; Iwaki, L.; Bammert, S.; Simpson, R.; Weiss, S.; Szot, J.; Flegal, B.; Lee, T. 15 kW neardiffraction-limited single-frequency Nd: YAG laser. In *Conference on Lasers and Electro-Optics*; Optical Society of America: Washington, DC, USA, 2007.
- 13. Klimek, D.E.; Mandl, A. Nd: YAG Ceramic ThinZag (Registered Trademark) High-Power Laser Development; Textron Defense Systems: Wilmington, MA, USA, 2011.
- 14. Peilin, L.; Liu, Q.; Fu, X.; Gong, M. Large-aperture end-pumped Nd: YAG thin-disk laser directly cooled by liquid. *Chin. Opt. Lett.* **2013**, *11*, 041408. [CrossRef]
- 15. Lizhi, D.; Bing, L.B.X.; Ping, Y. Recent progress of adaptive beam cleanup of solid-state slab lasers in In-stitute of Optics and Electronics, Chinese Academy of Sciences. *Opto-Electron. Eng.* **2018**, *45*, 170539.
- 16. Yading, G.; Qinjun, P.; Yong, B.; Zhongzheng, C.; Yang, L.; Lin, Z.; Chongfeng, S.; Lei, Y.; Baoshan, W.; Jian, X.; et al. 24.6 kW near diffraction limit quasi-continuous-wave Nd: YAG slab laser based on a stable–unstable hybrid cavity. *Opt. Lett.* **2020**, 45, 1136–1139.
- 17. Tyson, R.K. Principles of Adaptive Optics; CRC Press: Boca Raton, FL, USA, 2015.
- 18. Muller, R.A.; Buffington, A. Real-time correction of atmospherically degraded telescope images through image sharpening. *J. Opt. Soc. Am.* **1974**, *64*, 1200–1210. [CrossRef]
- Murray, L.P.; Dainty, C.; McGaughey, D.R. Wavefront sensor-less adaptive optics-Image correction through sharpness maximization. In *Frontiers in Optics*; Optical Society of America: Washington, DC, USA, 2006.
- Ping, Y.; Yuan, L.; Mingwu, A.; Shijie, H.; Bing, X. A wavefront sensor-less adaptive optical system for a solid-state laser. *Opt Lasers Eng.* 2008, 46, 517–521.
- 21. Liu, Y.; Ma, J.; Li, B.; Chu, J. Hill-climbing algorithm based on Zernike modes for wavefront sensorless adaptive optics. *Opt. Eng.* **2013**, *52*, 016601. [CrossRef]
- 22. Poland, S.P.; Wright, A.J.; Girkin, J.M. Evaluation of fitness parameters used in an iterative approach to aberration correction in optical sectioning microscopy. *Appl. Opt.* **2008**, *47*, 731–736. [CrossRef]
- Planchon, T.A.; Amir, W.; Field, J.; Durfee, C.; Squier, J.; Rousseau, P.; Albert, O.; Mourou, G. Adaptive correction of a tightly focused, high-intensity laser beam by use of a third-harmonic signal generated at an interface. *Opt. Lett.* 2006, *31*, 2214–2216. [CrossRef] [PubMed]

- Liu, Y.; Ma, J.; Li, B.; Hu, Y.; Chu, J. Comparison of optimization algorithms for adaptive optics system without a wavefront sensor. In Proceedings of the 6th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Large Mirrors and Telescopes, Xiamen, China, 26–29 April October 2012; International Society for Optics and Photonics: Bellingham, WA, USA, 2012.
- 25. 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. [CrossRef]
- 26. Piatrou, P.; Roggemann, M. Beaconless stochastic parallel gradient descent laser beam control: Numerical experiments. *Appl. Opt.* **2007**, *46*, 6831–6842. [CrossRef]
- 27. Ping, Y.; Yu, N.; Xiang, L.; Bing, X.; Xinyang, L.; Lizhi, D.; Hu, Y.; Wenjing, L.; Wenhan, J.; Lei, L. Enhancement of the beam quality of non-uniform output slab laser amplifier with a 39-actuator rectangular piezoelectric deformable mirror. *Opt. Express.* **2010**, *18*, 7121–7130.
- 28. Duchi, J.; Hazan, E.; Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **2011**, *12*, 2121–2159.
- 29. Ma, H.; Zhou, Q.; Xu, X.; Du, S.; Liu, Z. Full-field unsymmetrical beam shaping for decreasing and homogenizing the thermal deformation of optical element in a beam control system. *Opt. Express* **2011**, *19*, A1037–A1050. [CrossRef]
- 30. Tieleman, T.; Hinton, G. Lecture 6.5-rmsprop, coursera: Neural networks for machine learning. Coursera 2012, 4, 26–31.
- 31. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization; Computer Science: Kowloon Tong, Hong Kong, 2014.
- Qintao, H.; Liangli, Z.; Yao, M.; Shiei, Z.; Xi, Z.; Guozhong, Z. Adaptive stochastic parallel gradient descent approach for efficient fiber coupling. Opt. Express 2020, 28, 13141–13154.
- 33. Huizhi, Y.; Xinyang, L.; Wenhan, J. Comparison of several stochastic parallel optimization control algorithms for adaptive optics system. *High Power Laser Part. Beams* **2008**, *20*, 11–16.
- 34. 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. [CrossRef]
- 35. Jabri, M.; Flower, B. Weight perturbation: An optimal architecture and learning technique for analog VLSI feedforward and recurrent multilayer networks. *IEEE Trans. Neural Netw.* **1992**, *3*, 154–157. [CrossRef] [PubMed]
- Alspector, J.; Meir, R.; Yuhas, B.; Jayakumar, A.; Lippe, D. A parallel gradient descent method for learning in analog VLSI neural networks. In *Advances in Neural Information Processing Systems*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1993; pp. 836–844.
- 37. Noll, R.J. Zernike polynomials and atmospheric turbulence. JOsA 1976, 66, 207–211. [CrossRef]
- 38. Bo, C.; Huizhen, Y.; Jinbao, Z. Performances index and convergence speed of parallel gradient descent algorithm in adaptive optics of point source. *Acta Opt. Sin.* **2009**, *29*, 1143–1148.
- 39. Xiang, W.; Yang, P.; Wang, S.; Xu, B.; Liu, H. Underwater image enhancement based on red channel weighted compensation and gamma correction model. *Opto-Electron. Adv.* **2018**, *1*, 180024. [CrossRef]