

# 激光长程传输中的大气湍流非共轭智能校正方法

王靖婷<sup>1)2)</sup> 刘霄汉<sup>1)</sup> 王帅<sup>1)2)</sup> 张建柱<sup>1)</sup> 谢晓钢<sup>1)</sup> 罗文<sup>1)</sup> 苏华<sup>1†)</sup>

1) (北京应用物理与计算数学研究所, 北京 100094)

2) (中国工程物理研究院研究生院, 北京 100088)

## 摘要

自适应光学系统广泛应用于天文观测、自由空间光通信等领域, 相位共轭技术是其核心校正原理之一, 通过变形镜加载探测像差的共轭相位实现相位畸变补偿。然而, 在激光水平传输场景中, 受大气湍流非共轭特性、湍流与衍射效应耦合影响, 探测像差伴随相位奇点, 导致传统共轭校正方法性能显著退化甚至失效。为此, 本文提出基于机器学习的智能校正方法, 旨在寻找优于理想相位共轭的最优补偿波面, 提升激光大气水平传输湍流校正效果。首先, 本文系统剖析随机并行梯度下降 (SPGD) 算法的校正机制, 针对其局部最优、收敛缓慢等缺陷, 提出融合自适应增益调整、动量机制与提前终止策略的改进方案。仿真结果表明, 改进型 SPGD 算法在强湍流环境下可有效提升激光能量集中度。进一步地, 探索了强化学习在相位优化中的应用, 初步结果验证其具备挖掘更优校正相位的潜力。本研究证实了水平传输中存在优于理想校正相位的波面, 为非共轭湍流校正提供了新算法思路与仿真支撑。

**关键词:** 自适应光学, 激光大气传输, 强化学习, 非共轭校正

**PACS:** 42.68.Bz, 42.30.-d

† 通信作者. E-mail: su\_hua@iapcm.ac.cn

第一作者. E-mail: 2352085250@qq.com

## 1 引言

光在大气中传输时, 受大气湍流影响, 其波前相位发生严重畸变, 导致光束扩展、闪烁和相干性下降<sup>[1]</sup>。自适应光学 (AO) 技术通过实时测量畸变波前并进行相位校正, 从而缓解光束质量退化程度<sup>[2]</sup>。传统 AO 的核心补偿原理是相位共轭技术, 即通过变形镜等器件在光路上施加共轭相位以补偿畸变波前, 在天文观测和斜程激光传输中已取得显著成功<sup>[3]</sup>。

然而, 在激光水平或近地面长距离传输<sup>[4]</sup>中, 与垂直或斜程路径<sup>[5-10]</sup>不同, 水平光束穿越的湍流区域范围广、折射率起伏强烈<sup>[1, 12]</sup>且分布在整个传输路径

上，导致湍流表现出非共轭特性<sup>[13-15]</sup>。这意味着接收端测量到的相位畸变并非光束所累积湍流效应的简单共轭。因此，传统 AO 校正方法<sup>[1]</sup>即使在理想共轭条件

下也无法改善光束质量，在菲涅尔数 ( $N_F = \frac{a^2}{\lambda L}$ ) 较小时，增加的共轭相位可能使光束质量进一步恶化。

在近地面自由空间光通信、激光雷达探测等实际工程应用<sup>[16]</sup>中，激光常需要在水平或近地面长距离大气路径中传输。在这些场景中，传统共轭校正方法性能严重退化，因此，探索适用于非共轭湍流环境的新型波前校正方法具有重要的工程价值和理论意义。

我们基于 EasyLaser<sup>[17]</sup>平台的数值仿真，系统探究了相位共轭原理在水平传输中的局限性<sup>[18]</sup>。结果表明，当  $N_F < 1.2$  时，理想相位校正与传统 AO 均无法提供有效改善；相位共轭原理基本失效。这一失效归因于分布式湍流对波前的强烈调制，使得共轭补偿假设不再成立。仅当  $N_F$  超过临界值 ( $N_F > 1.2$ ) 时，校正能力才逐渐恢复。

这些发现表明，在非共轭湍流条件下需要不依赖于相位共轭的优化校正策略。近年来，无模型优化算法与机器学习方法<sup>[19]</sup>在复杂波前校正任务<sup>[20]</sup>中展现出潜力。2009 年，中国科学院安徽光学精密机械研究所利用随机并行梯度下降 (SPGD) 算法直接优化变形镜控制电压，对激光水平传输中湍流效应引起的畸变校正有一定效果<sup>[21]</sup>；2021 年马圣杰等人基于改进型 SPGD 算法对大气湍流导致的涡旋光波前畸变进行校正<sup>[22]</sup>；2022 年，赵辉等人将深度学习中 AdaBelief 优化器的一阶动量和二阶动量集成到 SPGD 算法中，提高了算法的收敛速度<sup>[23]</sup>。总体来说，针对 SPGD 算法在湍流校正中的应用已有部分研究工作，但考虑的场景通常传输距离较短、湍流影响较弱，对于本文关心的水平长程传输问题研究尚不充分。

强化学习 (RL) 框架如 Soft Actor-Critic (SAC)，则提供了通过与湍流环境交互来学习最优校正策略的可能性。2020 年电子科技大学许振兴从传统 AO 控制方法与强化学习找到结合点并做探索性研究，建立了自学习智能控制模型<sup>[24]</sup>。目前强化学习在湍流校正中的应用已有了一定进展，但仍缺少针对非共轭校正问题的研究。

基于之前的工作，本文提出一种基于机器学习的方法，旨在寻找在水平长

程传输中优于理想共轭的湍流校正相位。我们首先分析了标准 SPGD 算法在此类场景中的局限性，并引入了自适应增益调整、动量机制与提前终止等改进策略，以提升收敛性并避免局部最优。进一步，我们探索了强化学习技术在相位优化中的应用，证明其在发现更优非共轭校正解方面的巨大潜力。本研究通过仿真验证，在水平传输场景中存在性能优于理想相位校正的波前调控机制；该发现不仅为突破相位共轭技术瓶颈的湍流校正研究提供了全新算法框架，更为相关领域的实验探索奠定了坚实基础。

## 2 基于 SPGD 的激光传输非共轭相位校正方法

### 2.1 方法原理

#### 2.1.1 标准随机并行梯度下降算法

在自适应光学波前整形中，随机并行梯度下降是一种直接优化波前校正器控制参数以最大化系统性能指标  $J$ （如斯特列尔比或环围能量）的无模型方法。其核心思想是通过对控制向量  $w$ （通常为变形镜的驱动电压或 Zernike 系数）施加随机扰动，并根据性能指标的变化来估计梯度方向<sup>[21, 25]</sup>。具体步骤如下图所示。

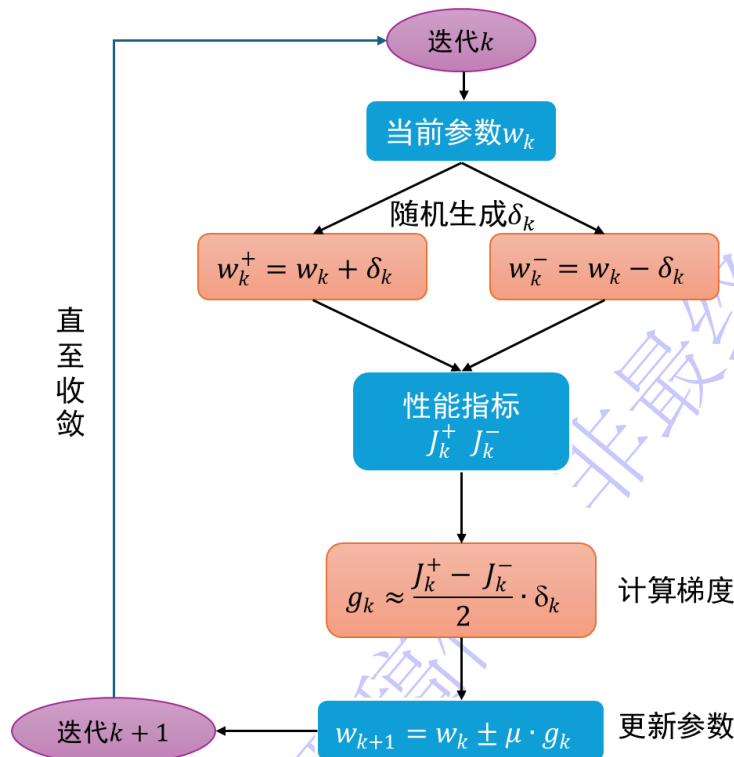


图1 标准SPGD算法流程图

Fig.1. Standard SPGD Algorithm Flowchart.

标准 SPGD 算法结构简单，无需系统模型，但存在收敛速度慢、易陷入局

部最优解以及学习率 $\mu$ 选择困难等固有缺陷。在激光在大气中水平传输、非共轭湍流校正这类具有复杂、多峰优化景观的问题中，不利于探索优于理想相位校正的补偿波面。

### 2.1.2 针对非共轭湍流校正问题的SPGD算法

为解决标准 SPGD 算法的固有限制，本研究提出了一种集成多项改进策略的增强型 SPGD 算法，旨在提升其在水平传输非共轭湍流校正任务中的收敛性、鲁棒性与校正性能。

#### (1) 相位表征与扰动模型的改进

标准 SPGD 算法在处理水平传输强湍流时，常受限于低阶 Zernike 多项式对于实际复杂湍流的表征能力不足。为更精细地描述复杂的非共轭相位结构，本工作将 Zernike 模式基底从常规的表征前 15 阶的低阶相位扩展为更接近实际湍流相位的 65 阶<sup>[26]</sup>。因此，控制参数向量 $w$ 的维度也相应增至 65，即 $w \in R^{65}$ 。这使算法能够表征和校正更高空间频率的波前畸变，从而更全面地覆盖由强湍流引发的、超出传统相位共轭补偿范围的高阶相位分量，为寻找更优的校正波面提供了更大的解空间。

此外，考虑到在扩展的高维控制空间中，随机扰动的生成也需要符合物理真实性。若对所有阶次的 Zernike 系数施加相同幅度的扰动，会与真实大气湍流相位功率谱随空间频率增加而衰减的统计特性不符，因此可能导致优化过程在不符实际的方向上过度探索，效率低下。

为此，本算法在生成均匀随机扰动向量 $\delta_k$ 后，引入一个指数衰减因子 $\eta_i$ 对其进行调制：

$$\delta'_k = \delta_k \odot \eta \quad (1)$$

$$\eta_i = \exp(-\lambda \cdot i), i = 0, 1, \dots, N \quad (2)$$

其中， $\odot$ 表示逐元素乘法， $i$ 为 Zernike 模式的阶数， $\lambda > 0$ 为衰减系数。增加了指数衰减因子后，离焦、像散这类的低阶像差的扰动幅度相对较大，而更高频成分的高阶像差的扰动幅度被有效抑制。改进后的 SPGD 算法在迭代过程中的探索方向更符合 Kolmogorov 湍流理论，也就是实际湍流中相位扰动能量主要集中于低阶，高阶像差的影响较小。从而引导优化过程更快速地向有潜力的区域收敛，避免在高阶、小能量维度上的无效振荡，显著提升了优化起始阶段的稳定性

和整体的收敛速度。

## (2) 核心优化算法的改进

首先，将自适应矩估计（Adam）优化器融入 SPGD 的迭代更新框架<sup>[21]</sup>。Adam 优化器不再使用固定学习率，而是自适应调整每个参数的学习率，并利用梯度的一阶矩（均值）和二阶矩（未中心化方差）估计，显著提升了优化过程的收敛速度和稳定性。

在第 $k$ 次迭代获得梯度估计 $g_k$ 后，Adam 优化器的更新步骤如下：

$$m_k = \beta_1 m_{k-1} + (1 - \beta_1) g_k \quad (3)$$

$$v_k = \beta_2 v_{k-1} + (1 - \beta_2) g_k^2 \quad (4)$$

$$\widehat{m}_k = \frac{m_k}{1 - \beta_1^k} \quad (5)$$

$$\widehat{v}_k = \frac{v_k}{1 - \beta_2^k} \quad (6)$$

$$w_{k+1} = w_k - \alpha \cdot \frac{\widehat{m}_k}{\sqrt{\widehat{v}_k + \epsilon}} \quad (7)$$

其中， $\beta_1 \in [0,1)$ ， $\beta_2 \in [0,1)$ ，为衰减率，通常默认设为 0.9 和 0.999； $\alpha$ 为全局步长，也是 SPGD 算法中的初始学习率； $\epsilon$ 是为保证数值稳定性而添加的小常数，通常为 $10^{-8}$ ； $m^k$ 和 $v^k$ 是对初始迭代阶段偏差的校正项。加入了 Adam 优化器的 SPGD 算法在梯度幅值较大的参数方向上采取更谨慎地更新步长，而在梯度平缓的方向上加速更新，从而显著提升收敛效率。

由于水平传输湍流的非共轭特性，系统性能指标 $J(w)$ 相对于校正参数 $w$ 的响应函数具有显著的非凸性和多峰性。这意味着，基于瞬时随机梯度估计的标准 SPGD 更新，并不能保证搜索路径的单调改善，即标准 SPGD 算法中希望出现的 $J(w_{k+1}) \leq J(w_k)$ 在本仿真中并非必然成立。算法可能因单次不良更新而偏离当前最优区域，导致校正效果回退，进而影响收敛的稳定性与最终性能。

针对上述问题，改进后的算法增加了最优参数回滚机制。SPGD 算法在运行过程中始终保存历史最优参数向量 $w_{best}$ 、历史最优性能指标 $J_{best}$ 和它所对应的校正面型，每次迭代后，若当前参数 $w_{k+1}$ 对应的性能未能超越 $J_{best}$ ，则立即将参

数回滚至 $W_{best}$ ，校正面型将始终保持在迭代中最好的状态。这一机制确保了优化过程的单调非退化性，极大增强了算法在强扰动环境下的稳定性。

此外，在 SPGD 算法寻找比理想相位校正更好的湍流校正面型的时候，改进后的算法还增加了自适应学习率衰减与早停策略。若连续 $N_{stag}$ 次尝试后均不能提升 $J_{best}$ ，则判定优化陷入平台期。此时，全局学习率 $\alpha$ 将根据预设规则进行衰减。若学习率衰减至预设阈值后性能仍无改善，则优化过程自动终止。对算法改善后，有效平衡了寻优深度与计算效率。仿真中，最大停滞次数为 20，最小学习率阈值 0.0001，这两个数据根据预实验确定，可在收敛效率与搜索深度之间取得平衡。

### (3) 倾斜像差的实时传感与独立补偿

在优化高阶非共轭像差的同时，为了消除倾斜像差 (Tip-Tilt) 对优化目标的干扰，算法采用实时反馈进行解耦补偿。通过分析每一迭代步的远场光斑质心偏移量 $(p_x, p_y)$ ，结合预标定的 Zernike 倾斜响应系数，实时计算出所需的倾斜补偿相位 $\Phi_{tilt}$ 。倾斜相差的补偿相位独立于优化循环产生，并与优化得到的高阶校正相位线性叠加，从而确保算法核心专注于校正更复杂的非共轭高阶波前畸变，从而显著降低了 SPGD 算法的优化难度，并保障了系统对倾斜扰动的快速抑制能力。

综合上述改进，本算法通过控制相位表征的扰动模型、对 SPGD 核心算法增加更适用于湍流校正的约束条件以及倾斜像差独立于高阶进行解耦补偿的三种方法，构建了一个适用于非共轭湍流校正的高效波前优化 SPGD 算法框架，其整体流程如图所示。

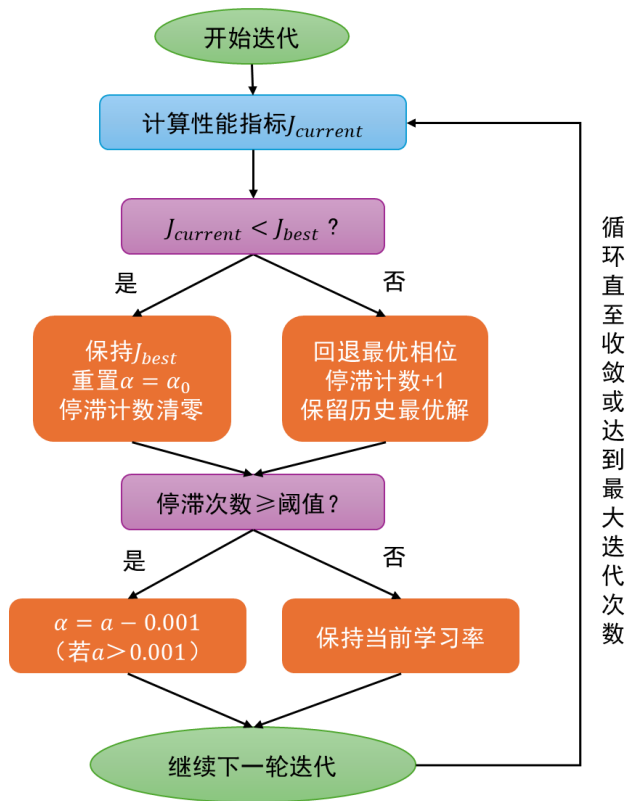


图2 面向非共轭湍流校正的改进型SPGD算法框架

Fig.2. An Improved SPGD Algorithm Framework for Non-Conjugate Turbulence Correction.

## 2.2 数值仿真系统与参数设置

为检验第 2.1 节所述改进 SPGD 算法在水平传输非共轭湍流校正中的有效性，并探究优于理想相位共轭的补偿波面的存在性，本研究构建了一套闭环数值仿真系统。该系统以 EasyLaser 大气激光传输仿真功能为核心，通过采用多相位屏结合角谱传播的数值模型<sup>[17]</sup>，可精确再现光波在湍流路径上的衍射过程，是研究非共轭校正问题的可靠数值仿真基础。

本文中大气水平传输仿真采用多相位屏法模拟连续湍流路径<sup>[27]</sup>，沿传输方向上按照设定间距放置 15 层随机相位屏，每层相位屏基于 Kolmogorov 谱生成，其湍流强度由大气相干长度  $r_0$  控制，通过修改  $C_n^2$ <sup>[28]</sup> 变换相干长度；光场在相位屏之间的自由空间传播采用角谱传输算法，能够准确模拟光束在湍流中的衍射效应，接收端提取远场光斑的环围能量作为评价指标。该仿真方法已集成于 EasyLaser 平台并经多组实验数据验证<sup>[18, 29]</sup>。

### 2.2.1 仿真系统架构与流程

大气水平传输仿真系统如图所示，形成了“畸变生成、波前校正、传输评估、

优化反馈”的完整闭环，具体包含以下四个模块：

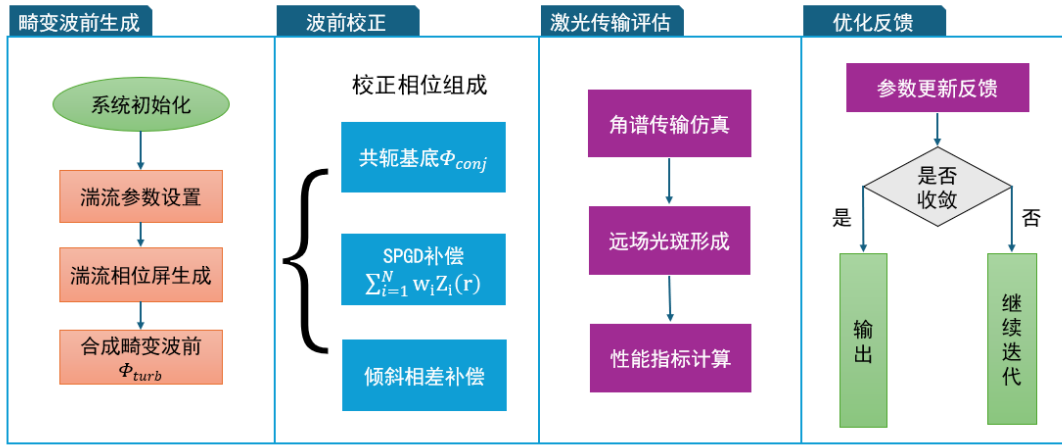


图3 激光长程传输湍流校正仿真系统架构

Fig.3. Architecture of a Laser Long-Range Transmission Turbulence Correction Simulation System.

(1) 大气湍流信道仿真模块

利用 EasyLaser 光束传输函数，模拟激光从发射端到接收端的传输。由于本研究的目的是探究优于理想相位共轭的补偿波面的存在性，因此将风速固定设置为 0，生成不随时间动态演变的静态湍流相位畸变  $\Phi_{turb}$ 。

(2) 自适应光学仿真模块

波前传感器：仿真中理想化地假设可从接收的信标光场中完美提取湍流引起的总相位畸变  $\Phi_{beacon}$ ，并对其解包裹与翻转<sup>[30]</sup>，得到理想相位共轭项。

$$\Phi_{conj} = -Unwrap(\Phi_{beacon}) \quad (8)$$

波前校正器：模拟一个连续面形变形镜。其施加的校正相位  $\Phi_{cor}$  由三部分线性叠加构成：

$$\Phi_{cor} = \underbrace{\Phi_{conj}}_{\text{共轭基底}} + \underbrace{\sum_{i=1}^N w_i Z_i(r)}_{\text{SPGD 优化补偿}} + \underbrace{(a_{1,c} Z_1 + a_{2,c} Z_2)}_{\text{倾斜相差补偿}} \quad (9)$$

其中，Zernike 系数向量  $\mathbf{w}$  为优化对象；倾斜补偿系数  $a_{1,c}$ ， $a_{2,c}$  根据每步迭代的远场光斑质心偏移量实时计算。

(3) 改进 SPGD 优化控制器模块

这部分为核心算法模块，按照第 2.1.2 节所述流程实现。在每一迭代步中，

控制器首先生成一个经指数衰减因子调制符合物理先验的随机扰动<sup>[31]</sup>，并驱动传输仿真模块完成正、负扰动下的两次完整模拟光传输计算，从而获取性能指标的差分。在此基础上控制器调用内部的 Adam 优化器更新内核，计算出对 Zernike 系数向量 $\mathbf{w}$ 的调整量。同时，控制器持续监测优化轨迹：若检测到性能提升停滞，则触发学习率衰减机制；若单次更新导致性能退化，则执行参数回滚，将系统状态恢复至历史最优。

#### (4) 性能诊断与评价模块

每迭代步结束后，调用函数对远场光场进行分析，获取以光斑质心为中心、在预设评价半径 $r_{bucket}$ （本文中设为理想衍射极限半径 $r_{diff}$ ）的环围区域内的平均功率密度 $P_c$ ，这一物理量直接反映了经过湍流扰动与波前校正后，光束能量在理想瞄准点附近的集中程度<sup>[32]</sup>，经过多种物理量作为优化目标的结果对比， $P_c$ 是最合适的，综合仿真结果，最终根据 $P_c$ 构建了损失函数 $J$ 。

$$J_k = 1 - \frac{P_c(k)}{P_c^{ideal}} \quad (10)$$

其中， $P_c^{ideal}$ 为无湍流、无像差条件下的理想衍射极限值。该损失函数 $J_k \in (0,1]$ 被实时反馈给 SPGD 优化控制器，作为估计性能梯度与更新校正参数的唯一依据。优化目标是 minimized  $J_k$ ，即等价于最大化能量集中度。

此外，该模块同步提取远场光斑的质心坐标 $(px, py)$ ，用于实时倾斜像差补偿。

为对校正性能进行客观、标准化的横向对比，定义了一个更适用于水平传输非共轭湍流校正的性能提升因子 $\alpha_c$ 作为核心评价指标：

$$\alpha_c = \frac{P_c^{corrected}}{P_c^{uncorrected}} \quad (11)$$

其中分子 $P_c^{corrected}$ 为采用某种校正方法后的稳态 $P_c$ 值，分母 $P_c^{uncorrected}$ 则为相同湍流条件下不作任何波前校正时的 $P_c$ 值。该因子定量表征了特定校正策略相对于“不校正”基线所带来的性能增益倍数，其值大于 1 即表明校正有效。

在整个研究框架中，损失函数 $J_k$ 与评价指标 $\alpha_c$ 扮演着截然不同但相辅相成的

角色：前者是优化器在迭代中追寻的直接、动态目标，后者则是用于最终结果分析、在不同条件与算法间进行公平比较的静态、标准化度量。

### 2.2.2 关键参数与仿真设计

所有仿真参数均基于典型的近地水平激光传输应用场景设定。系统仿真采用波长 $\lambda = 1064\text{nm}$ ，发射孔径直径 $D = 0.32\text{m}$ 。为系统研究非共轭效应的强弱对校正方法的影响，菲涅尔数 $N_F$ 被选为核心控制变量之一，通过调整传输距离 $L$ 实现 $N_F$ 在一定范围内的变化。

湍流强度由大气相干长度 $r_0$ 表征。本研究中 $r_0$ 的取值覆盖  $0.03\text{m}$  至  $0.16\text{m}$ ，相应的 $D/r_0$ 比值介于 2 至 10 之间，从而涵盖了从弱湍流到强湍流的典型大气条件 [33]。

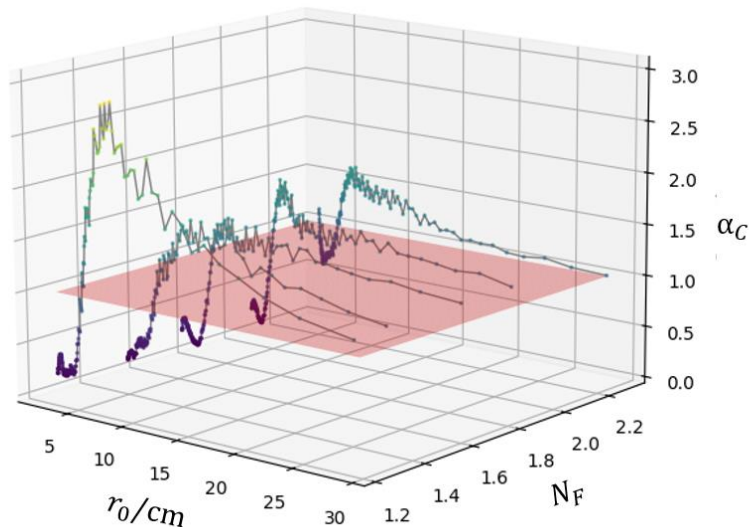


图4 不同水平场景下理想相位共轭校正的结果

Fig.4. Results of Ideal Phase-Conjugate Correction at Different Scene Levels.

参数选择的仿真依据源于对理想相位共轭校正基准性能的预先分析。如图 4 所示，当湍流较弱时，不同 $N_F$ 理想相位共轭的校正效果差异不大。考虑到本研究聚焦于相位共轭原理失效的挑战性场景，综合基准分析结果，后续的 SPGD 算法性能评估被有意设置在 $N_F$ 值较小、且 $r_0$ 值较小的参数区域。这些区域正是传统共轭校正性能受限、非共轭特性凸显的工况，从而为验证改进型 SPGD 算法在相位共轭原理失效区寻找更优校正波面的能力提供了明确的测试目标与对比基准。

### 2.3 结果与分析

为系统评估改进 SPGD 算法在不同传输条件下的校正性能，本研究设计了两组对比实验：传输距离  $L = 10 \text{ km}$  ( $N_F = 2.4$ ) 与  $L = 20 \text{ km}$  ( $N_F = 1.2$ )。前者对应衍射效应相对较弱的场景，后者则突出衍射效应显著、非共轭特性更为凸显的工况。通过对比两组实验结果，可以深入揭示传输距离（即菲涅尔数）对非共轭效应强弱及算法校正效果的影响规律。

### 2.3.1 传输距离 $L = 10 \text{ km}$ ( $N_F = 2.4$ ) 下的校正性能

图 5 展示了当传输距离固定为  $L = 10 \text{ km}$  ( $N_F = 2.4$ ) 时，性能提升因子  $\alpha_c$  随湍流相干长度  $r_0$  的变化规律。作为菲涅尔数较大的对比组，此场景下衍射效应相对较弱，为后续分析非共轭效应的增强提供了基准参考。

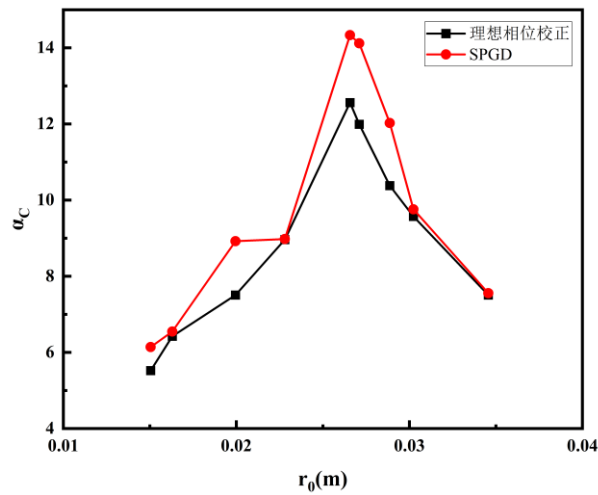


图5  $L = 10 \text{ km}$ 时改良SPGD算法和理想共轭相位的校正结果

Fig.5. Results of Improved SPGD Algorithm and Ideal Phase-Conjugate Correction at

$L = 10 \text{ km}$ .

实验结果表明，在弱至中等湍流区 ( $r_0$ 较大)，理想相位共轭本身已具备良好的校正效果，基本接近最优校正策略。在此区域，改进 SPGD 算法虽仍保持一定的性能优势，但提升幅度相对有限，二者差距并不显著。

随着湍流强度增强，理想相位共轭的校正效果有一定下降，而改进 SPGD 算法的  $\alpha_c$  依然持续优于理想共轭。这一结果初步验证了改进 SPGD 算法在传统方法失效区域的适用性。

### 2.3.2 传输距离 $L = 20 \text{ km}$ ( $N_F = 1.2$ ) 下的校正性能与算法优势深化

为进一步探究 $N_F$ 减小（即非共轭效应增强）对校正性能的影响，将传输距离增加至 $L = 20$  km ( $N_F = 1.2$ )，实验结果如图 6 所示。相较于 $L = 10$  km 的场景，整体校正性能曲线向湍流相干长度更大的方向发生了系统性偏移，且理想相位共轭的失效特征更为显著。

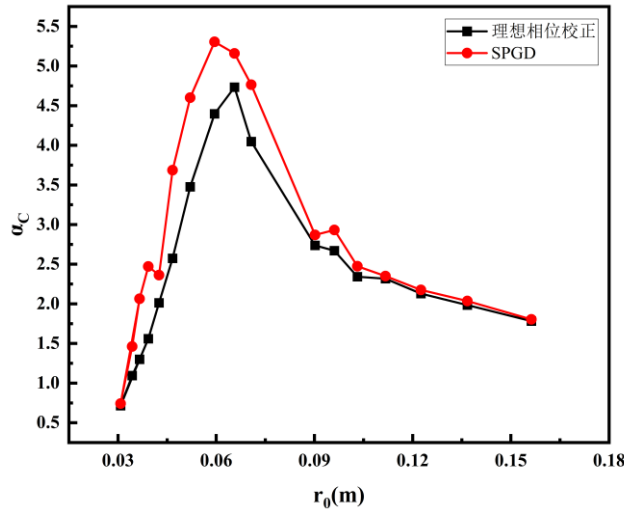


图6  $L = 20$  km时改良SPGD算法和理想共轭相位的校正结果

Fig.6. Results of Improved SPGD Algorithm and Ideal Phase-Conjugate Correction at  $L = 20$  km.

图 7(a)-(c)选取了  $L = 20$  km ( $N_F = 1.2$ ) 仿真中的三个关键工况，以直观展示不同湍流强度下的校正效果差异：

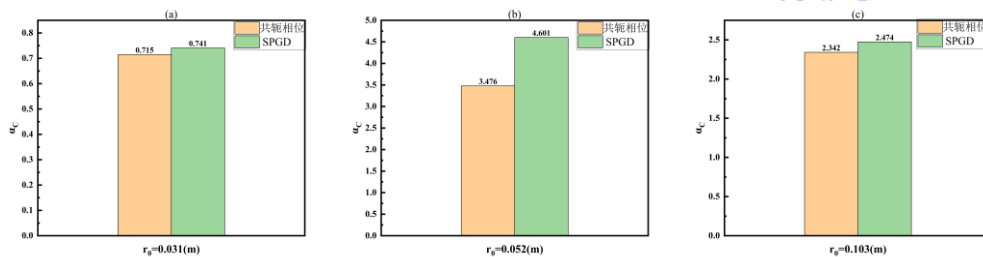


图7 不同湍流相干长度下改良SPGD算法和理想共轭相位的校正结果

(a)  $r_0 = 0.031$ m; (b)  $r_0 = 0.052$ m; (c)  $r_0 = 0.103$ m

Fig.7. Results of Improved SPGD and Ideal Phase-Conjugate Correction at Different

$r_0$ .

(a)  $r_0 = 0.031$ m; (b)  $r_0 = 0.052$ m; (c)  $r_0 = 0.103$ m

当 $r_0 = 0.031\text{m}$ 时，理想相位共轭的 $\alpha_c$ 已降至 0.71468，这意味着施加共轭相位后的能量集中度反而低于不校正的情况，表明在此强湍流条件下，理想相位共轭原理已完全失效，增加的共轭相位甚至引入了新的波前畸变。然而，改进 SPGD 算法以共轭相位为基底，仍能实现微弱提升，证实了其在相位共轭原理失效区仍具备一定的波前整形能力。

随着湍流强度减弱至 $r_0 = 0.052\text{m}$ 时，改进 SPGD 算法的优势逐渐放大。算法寻获的校正波面使得 $\alpha_c$ 达到 4.6，相较于理想相位校正的 $\alpha_c = 3.48$ ，性能提升超过 20%。这一显著增益直观地体现在远场光斑质量的改善上，SPGD 校正后的光斑能量集中度明显优于单纯共轭校正，表明算法成功探索到了传统方法未能触及的更优校正模式。

当湍流进一步减弱至 $r_0 = 0.103\text{m}$ 的中等湍流区，算法依然表现出一定的校正效果（ $\alpha_c = 2.4736$ ），而理想相位校正在此区域已恢复较好的校正性能，因此对比差异并不明显。

对比图 5 与图 6 可以揭示一个重要的规律：当菲涅尔数减小（传输距离增加，衍射效应增强）时，改进 SPGD 算法相较于理想相位共轭的绝对性能优势随之扩大。在 $L = 20\text{km}$ （ $N_F = 1.2$ ）的强非共轭区，SPGD 在 $r_0 = 0.052\text{m}$ 附近实现了超过 20%的最大相对提升；而在 $L = 10\text{km}$ （ $N_F = 2.4$ ）的场景下，最大提升幅度明显收窄。这表明，在相位共轭本身有效的弱衍射区，以梯度估计为基础的 SPGD 算法探索非共轭解空间的动机和能力可能受到限制，易收敛于与理想相位共轭相近的局部最优解。

### 2.3.3 讨论与小结

上述实验结果系统性地验证了改进 SPGD 算法在水平传输非共轭湍流校正中的核心价值：在传统相位共轭原理失效或性能受限的强湍流、小菲涅尔数区域，该算法能够有效地寻获更优的校正波面，显著提升光束质量。这一发现证实了在非共轭条件下，确实存在优于理想相位共轭的校正波面，且能够通过智能优化算法寻获。

然而，实验结果也揭示了 SPGD 算法的固有局限。其基于瞬时随机梯度估计的优化本质，使其在复杂、多峰的损失景观中易陷入局部最优，收敛速度受问

题条件影响较大。特别是在菲涅尔数较大（相位共轭本身有效）的场景下，算法的性能提升空间显著收窄。这为后续引入具备更强全局探索能力的优化方法提供了明确的实验依据与研究动机。

### 3 基于强化学习的激光传输非共轭相位优化校正方法

尽管前一部分提出的改进 SPGD 算法在非共轭湍流校正中展现了优于传统相位共轭的潜力，但其本质仍是一种基于即时随机梯度估计的局部优化方法。在强湍流引起的复杂、多峰优化景观中，SPGD 固有的对初始条件敏感、易陷入局部最优和探索策略缺乏系统性的缺陷，可能会阻碍全局更优的校正波面的发现。

为突破这一瓶颈，本部分引入一种基于最大熵强化学习（Maximum Entropy Reinforcement Learning）框架的先进优化模型。具体而言，采用 Soft Actor-Critic 算法，通过最大化累积奖励与策略熵的加权和，鼓励智能体在有效利用当前最优策略与主动探索未知状态之间达成平衡，特别适用于像湍流相位校正这样具有复杂、多峰优化特性的连续控制问题，为求解高维、连续的非共轭相位优化问题提供了新的方法论。

#### 3.1 演员-评论家算法

Soft Actor-Critic (SAC) 是一种基于最大熵框架的深度强化学习算法，由 Haarnoja 等人于 2018 年提出<sup>[34]</sup>。与传统的策略梯度方法不同，SAC 在优化目标中显式引入策略熵项，鼓励智能体在追求高累积奖励的同时保持动作的随机性，从而实现系统性的探索与利用平衡<sup>[35]</sup>。

强化学习描述了一个智能体 (Agent) 与环境 (Environment) 通过交互进行学习的范式，其核心要素包括：

策略 (Policy): 记作  $\Pi(a|s)$ ，定义了智能体在状态  $s$  下选择动作  $a$  的概率分布。策略是智能体行为的基本准则，可以看作从状态到动作的映射函数。

状态 (State): 记作  $s \in \mathcal{S}$ ，描述了环境在当前时刻的配置和信息。在相位校正任务中，状态包含远场光斑图和探测到的湍流畸变相位。

动作 (Action): 记作  $a \in \mathcal{A}$ ，是智能体对环境施加的干预。在相位校正任务中，动作对应波前校正器需要施加的校正相位参数。

奖励 (Reward): 记作  $r(s, a) \in \mathcal{R}$ ，是智能体在状态  $s$  下执行动作  $a$  后从环境获得的即时反馈信号。奖励函数的设计直接决定了智能体的优化目标。

累积回报 (Return): 记作  $R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k}$ , 其中  $\gamma \in [0, 1]$  为折扣因子。强化学习的核心目标是找到最优策略  $\Pi$ , 使得期望累积回报最大化。

SAC 采用 Actor-Critic 架构, 将策略学习与价值评估解耦为两个相互协作的模块:

Actor (策略网络): 记作  $\pi_{\phi}(a_t | s_t)$ , 其中  $\phi$  为网络参数。Actor 根据当前状态  $s_t$  输出动作  $a_t$  的概率分布, 在连续控制问题中通常输出高斯分布的均值和标准差。在相位校正任务中, Actor 负责生成校正相位对应的控制参数。

Critic (价值网络): 记作  $Q_{\theta}(s_t, a_t)$ , 其中  $\theta$  为网络参数。Critic 评估在状态  $s_t$  下执行动作  $a_t$  的期望累积回报, 为 Actor 的策略更新提供指导。如果一个动作被 Critic 评估为高价值, Actor 将增加选择该动作的概率; 反之则降低其概率。SAC 通常使用两个 Critic 网络 (双 Q 网络) 来缓解 Q 值过估计问题。在每次更新时, 取两个网络输出的较小值作为目标 Q 值, 能有效提升算法的稳定性和最终性能。

SAC 的核心创新在于其最大熵目标。与传统强化学习仅最大化累积奖励不同, SAC 的优化目标  $J(\pi)$  同时鼓励策略  $\pi$  保持高熵值:

$$J(\pi) = E_{(s_t, a_t) \sim \rho(\pi)} \left[ \sum_t \gamma^t (r(s_t, a_t) + \alpha H(\pi(\cdot | s_t))) \right] \quad (12)$$

其中,  $H(\pi(\cdot | s_t))$  表示策略在状态  $s_t$  下的熵;  $\alpha > 0$  为熵权重系数。该目标函数驱使智能体在寻求高奖励的同时, 尽可能随机地探索动作空间, 从而避免了类似 SPGD 算法过早收敛陷入局部最优, 为在复杂相位空间中系统性地发现潜在更优解提供了理论保障。

熵权重系数  $\alpha$  的取值对算法性能至关重要。SAC 将其转化为一个带约束的优化问题, 要求策略熵的均值保持在一个目标值  $H_{target}$  附近。该机制能动态平衡探索与利用: 在训练初期或校正效果陷入平台时增大  $\alpha$  值以加强探索; 当策略趋于稳定在高效区域时减小  $\alpha$  值以专注利用。

### 3.2 基于SAC算法的非共轭相位优化校正

#### 3.2.1 整体框架

本研究构建的基于 SAC 算法的非共轭相位优化校正框架如图 8 所示。该框架将湍流校正问题形式化为一个马尔可夫决策过程，其中智能体（控制器）与环境（大气湍流信道与光学系统）持续交互。

首先，环境将当前时刻包含畸变波前相位和远场光斑信息的状态反馈给 SAC 智能体。智能体内部的策略网络根据该状态，输出一组用于驱动变形镜的 Zernike 系数的动作。该校正相位被施加到光学系统中，激光经校正后在大气湍流中传输。环境在接收到动作后，计算新的远场光斑，并产生一个奖励信号，同时状态转移至下一时刻。智能体的价值网络则负责评估在当前状态下采取特定动作的长期期望回报。所有这些交互经验都会被存入经验回放缓冲区。SAC 算法通过从缓冲区中随机采样历史数据进行学习，并周期性地软更新目标网络，以稳定训练过程。通过这个闭环的“交互、学习、优化”过程，SAC 智能体能够在不依赖系统模型的情况下，通过最大化累积奖励与策略熵，学习到一个可以输出优于传统相位共轭的校正策略。

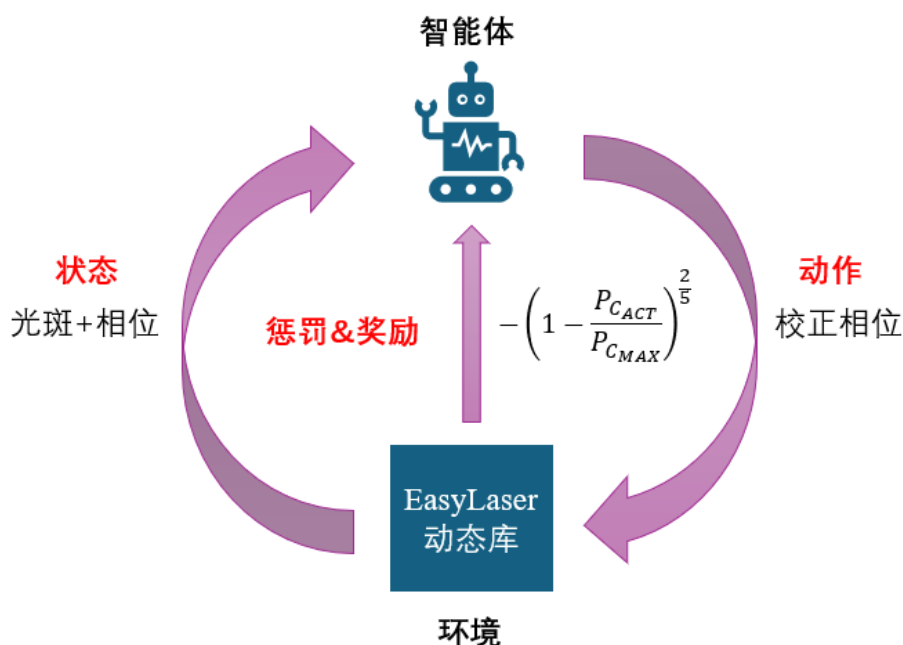


图8 SAC算法在相位校正中的应用机制

Fig.8. Application Mechanism of the SAC Algorithm in Phase Correction.

#### 3.2.2 策略网络和价值网络结构

为实现从高维光学图像到高维连续动作的端到端学习，本工作设计了专门

的策略网络和价值网络，其结构分别如图 9(a)和图 9(b)所示。

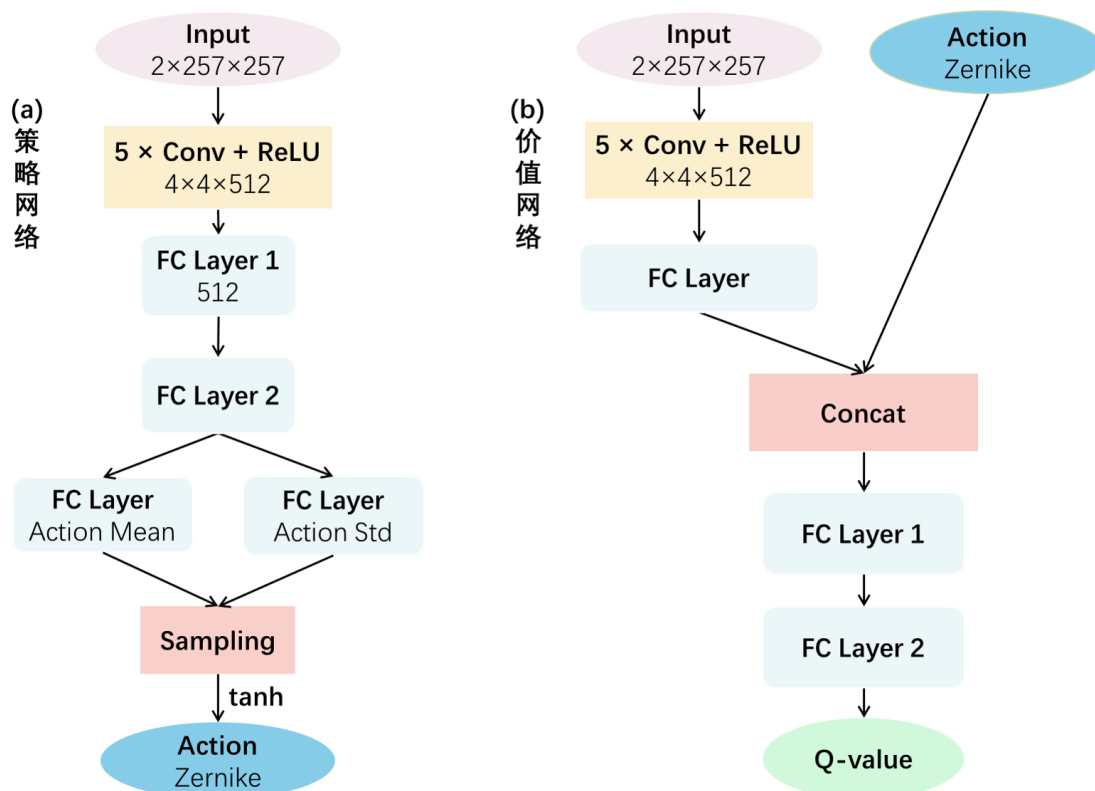


图9 策略网络与价值网络结构。(a) Actor网络；(b) Critic网络

Figure.9. Policy and value network architectures. (a) Actor network; (b) Critic network

策略网络如图 9(a)所示，其输入为环境提供的状态，具体为二维观测数据，包括光学系统的畸变相位图和远场光斑图像，维度为  $2 \times 257 \times 257$ 。该输入首先经过一个卷积神经网络编码器 ( $N_{CNN}$ )，该编码器由五个卷积层和一个全连接层组成，用于提取高维空间特征。第一个卷积层使用  $8 \times 8$  的卷积核，步长为 4，将输入的空间维度降至  $64 \times 64$ ，并输出 32 个特征通道。随后四个卷积层均使用  $4 \times 4$  的卷积核，步长为 2，特征通道数依次增加至 64、128、256 和 512。每个卷积层后都应用了 ReLU 激活函数以引入非线性特征，并缓解梯度消失问题。编码器的输出为一个  $512 \times 4 \times 4$  的特征图，经展平后通过一个全连接层映射为 512 维的特征向量。该特征向量随后分支为两个独立的全连接层，分别输出动作均值和动作标准差，维度均为对应 Zernike 多项式第 4 至第 104 项的动作空间维度 101。最终，通过正态分布重参数化采样并使用 tanh 函数将动作值限制在  $[-0.6, 0.6]$  范围内，得到最终的 Zernike 系数动作。同时，网络还计算该动作的对数概率，用于软策略更新中的熵正则化项。

价值网络采用双 Q 网络结构以缓解 Q 值过估计问题，如图 9(b)所示。每个

Q 网络的输入均为状态和动作的组合。其中，状态的处理方式与策略网络相同，首先经过一个结构一致的卷积编码器提取特征。随后，将编码器输出的状态特征与维度为 101 的动作向量进行拼接。拼接后的向量依次通过两个全连接隐藏层，最终输出一个标量 Q 值。在 SAC 算法更新过程中，目标 Q 值通过计算两个 Q 网络输出的最小值得到，从而保证了价值评估的稳定性和准确性。

### 3.2.3 训练参数配置

训练过程中，使用经验回放缓冲区存储智能体与环境交互产生的五元组数据：状态、动作、奖励、下一状态、终止标记。缓冲区容量设为 10000 条样本，以保证数据的多样性。训练更新仅在缓冲区中样本数超过最小阈值 1000 时开始，每次从中随机采样 64 条数据用于梯度更新，以打破数据间的相关性并提高样本利用率。

本训练共进行 10000 个回合，每个回合包含一次完整的单步环境交互。核心超参数设置如下：策略网络、价值网络和熵系数的学习率均设为  $3 \times 10^{-4}$ ；目标网络的软更新参数  $\tau = 0.005$ ；目标熵根据经验设置为 -101，以自动调整策略的随机性。该设置保证了算法在训练初期具备足够的探索能力，并在收敛阶段逐步降低策略的不确定性。该超参数组合结合 SAC 算法特点，使得高维连续动作控制问题能够有效收敛。

图 10 展示了在典型强湍流条件 ( $C_n^2 = 0.035$ ) 下，训练过程中每 100 个回合的平均奖励变化曲线。如图所示，在约 0-2000 回合的训练初期，由于策略的随机性较高，平均奖励值较低且波动较大。随着训练进行，智能体通过经验回放不断优化策略，平均奖励呈现出稳步上升的趋势，并在约 6000 回合后逐渐进入收敛状态，奖励值稳定在高位。这表明智能体已经学习到一个稳定的、高性能的校正策略。训练完成后，我们保存了在验证集上表现最佳的策略模型用于后续的仿真验证。

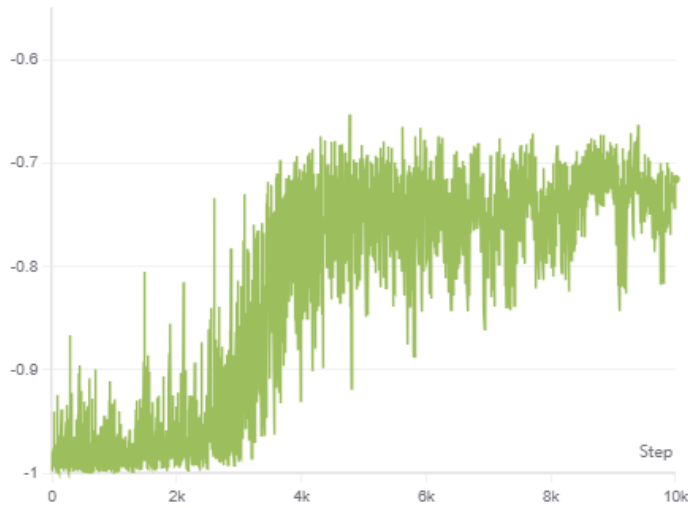


图10 强湍流条件下的训练奖励曲线

Figure.10. Training reward curve under strong turbulence conditions.

### 3.3 强化学习算法校正效果的仿真验证

为验证强化学习算法在非共轭湍流校正中的实际效能，我们利用与第 2.3 节相同的动态库和相同的仿真参数下完成了一系列对比实验。图 11 和图 12 系统地展示了在水平传输 ( $L = 20\text{km}$ ,  $N_F = 1.2$ ) 场景下，强化学习算法与 SPGD 算法的校正效果。

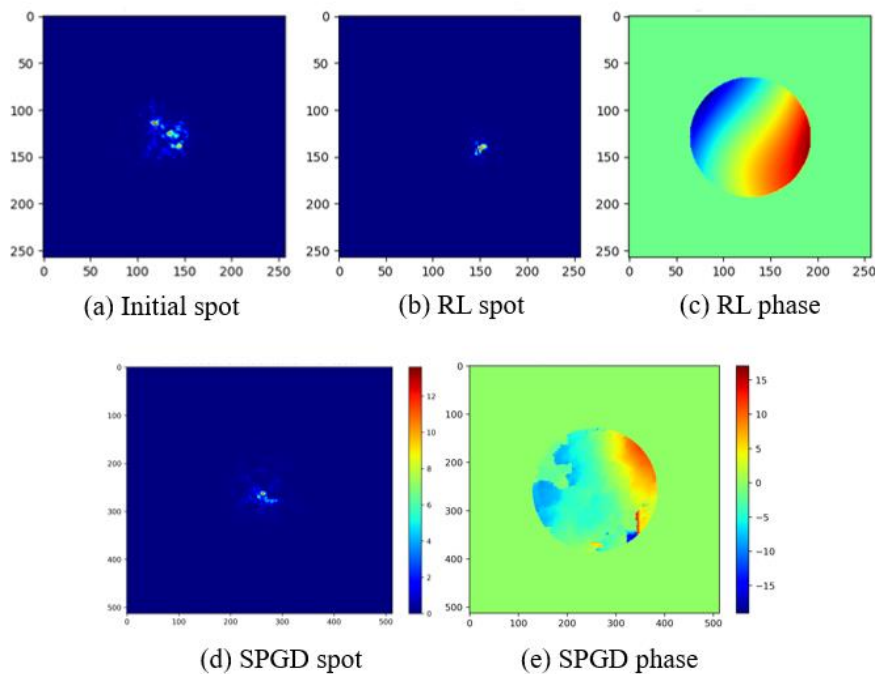


图11 水平传输场景下强化学习算法与SPGD算法的校正结果

(a) 初始光斑；(b) 强化学习校正后的光斑；(c) 强化学习的校正相位；(d) SPGD 校正后的光斑；(e) SPGD的校正相位

Fig.11. Correction results of reinforcement learning algorithm and SPGD algorithm in horizontal transmission scenarios.

(a) Initial spot; (b) Spot after reinforcement learning correction; (c) Reinforcement learning correction phase; (d) Spot after SPGD correction; (e) SPGD correction phase

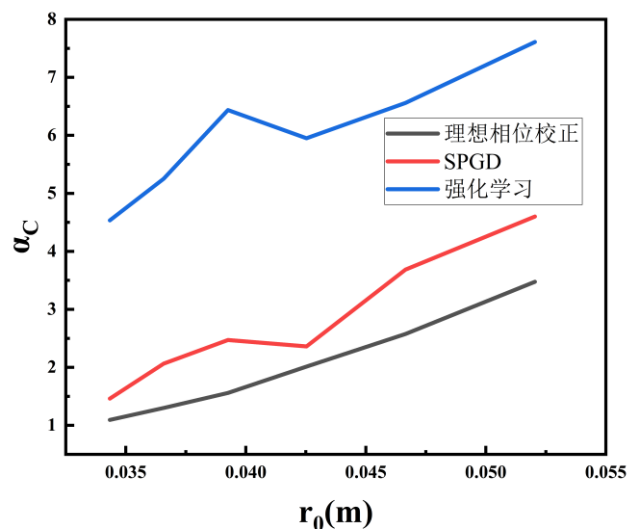


图12  $L = 20\text{km}$ ,  $N_F = 1.2$ 时不同情况下的湍流影响

Fig.12.  $L = 20\text{km}$ ,  $N_F = 1.2$  Turbulence Effects Under Different Conditions.

图 11 直观对比了在强湍流条件下不同场景下的远场光斑形态。图 11(a)中未校正的光斑因湍流畸变而严重扩散碎裂。在图 11(d)中应用 SPGD 算法后，光斑质量有所改善，但仍存在明显的能量弥散。与之形成鲜明对比的是，经强化学习算法优化后的校正光斑图 11(b)呈现出显著更优的能量集中度与更接近衍射极限的形状。这种视觉上的显著差异，源于两种方法所施加的校正相位的本质不同。

如图 11(c)与图 11(e)所示，SPGD 算法的补偿相位连续性较差，而强化学习算法则获得了一个不会受相位奇点影响、结构上更为平滑的校正相位  $\phi_{SAC}$ ，补全了 SPGD 算法未能探索的部分。

此外，SPGD 全局性较差，只能在提供的共轭相位基底附近寻找，易陷入局部最优解，而强化学习全局搜索能力强，不需要提供相位基底，寻找到的补偿相位对非共轭湍流畸变效果大大提升。

为评估强化学习在不同湍流强度下的普适性优势，图 12 绘制了性能提升因子  $\alpha_C$  随  $r_0$  变化的曲线。结果显示，在测试范围内，强化学习算法的  $\alpha_C$  值始终大于 1，并且校正效果明显好于 SPGD 算法。尤其在相位共轭趋于失效的强湍流区，强化

学习的优势更为明显。例如，在 $r_0 = 0.031\text{m}$ 时，强化学习将 $\alpha_c$ 从 SPGD 算法结果的 1.04 提升至 4.53，说明即使在相位共轭原理完全失效的强湍流区，强化学习仍能够实现稳定且较大的性能增益。这一系统性的量化结果，强有力地证实了强化学习不仅能在特定点例上、更能在宽泛条件下，稳定地寻获优于理想相位校正和 SPGD 算法的校正解。

综合视觉与定量分析，本仿真清晰地表明，基于最大熵强化学习的 SAC 算法，通过其强大的全局探索与序列决策能力，能够有效突破传统梯度类优化方法的局部最优限制，为水平传输非共轭湍流校正问题提供了一种性能更优的解决方案。

本部分的初步研究表明，以 SAC 为代表的先进强化学习算法，为解决非共轭湍流校正这一复杂优化问题提供了富有前景的新途径。后续工作将致力于更深入的系统性实验、参数优化，以及算法实时性的提升。

### 3.4 强化学习算法在非共轭相位优化中的理论优势

相比于第 2 章所述的改进 SPGD 算法，强化学习方法在解决非共轭相位优化问题上展现出多方面的理论优势<sup>[36]</sup>，这些优势从根本上有望突破传统梯度类方法的固有局限。

#### (1) 全局探索能力

SPGD 算法的本质是单点梯度下降：每次迭代仅基于当前参数点的局部梯度估计进行一次更新，其探索行为局限于当前参数附近的小邻域内，且由随机扰动幅度控制。在非共轭湍流引起的复杂、多峰损失景观中，这种局部探索策略极易陷入与初始值邻近的局部最优解，难以跨越势垒触及更优的全局区域。

相比之下，强化学习的最大熵框架引导的是对整个策略分布的优化，而非单点参数。目标函数中的熵正则项明确鼓励策略保持高随机性，使智能体在训练过程中能够系统性地探索远离当前策略均值的动作。最大熵目标将会在策略优化与分布匹配之间寻求平衡，赋予算法更强的跳出局部最优的能力，从而更有可能发现 SPGD 无法触及的全局更优校正波面。

#### (2) 离线策略学习

SPGD 作为一种在线无模型优化方法，其每次梯度估计仅利用当前正负扰动下的两次性能测量，历史交互数据被完全丢弃。

强化学习则通过经验回放池存储历史交互数据，算法能够从大量过往经验

中反复学习，更重要的是，离线策略允许智能体从任意历史策略生成的数据中学习价值函数，这使得算法能够在不与环境交互的情况下进行多次内部更新，大幅加速了价值估计的收敛过程。对于非共轭湍流校正这类高维连续控制问题，样本效率的提升意味着在有限交互次数内能够获得更优的校正策略。

### (3) 长期累积奖励与延迟奖励处理

SPGD 的梯度估计基于瞬时性能指标差分，其更新完全依赖于当前单步扰动的即时效果。这种短视的优化视角在处理具有延迟奖励特性的问题时存在根本性缺陷：某些校正策略可能需要多步连续调整才能显现出显著的性能提升而 SPGD 无法识别这类序列价值，因为每一步的即时回报可能为负，导致算法放弃具有长期潜力的优化方向。

强化学习则从长期累积奖励的视角评估动作价值。通过价值函数对未来累积回报的估计，智能体能够识别那些即时回报为负但长期收益为正的关键动作。在波前校正问题中，这意味着算法可以执行一系列“暂时性”性能下降的校正，以换取最终显著超越局部最优的全局解。算法能够在短期牺牲与长期收益之间做出理性权衡。这种序列决策能力，正是突破非共轭湍流校正瓶颈的关键所在。

从物理机制来看，激光水平长程传输中的湍流与衍射非线性作用，使得信标光到达探测端时相位连续性被破坏甚至出现相位奇点。显然，在发射端不连续相位的简单线性共轭，并不能使得光束在与非线性的“对抗”中获得优势。相反，具备优势的补偿相位应该在物理上具有更连续平滑的结构，而 RL 方法从传输中与湍流连续分布“对抗”的视角，表明了此类具有非线性演化竞争优势补偿相位的存在性。

## 4 结 论

本文针对水平传输中湍流非共轭性导致的传统相位共轭校正失效问题，系统探究了机器学习驱动的波前优化方法。理论与仿真证实，非共轭场景中存在性能优于理想相位共轭的校正波面，突破了“共轭相位即最优”的传统认知，确立了该领域研究的可行性与目标。针对随机并行梯度下降算法，提出融合高阶 Zernike 表征、Kolmogorov 谱扰动衰减、Adam 优化器及参数回滚与早停策略的改进方案，在  $N_F = 1.2$  的非共轭湍流条件下，将能量集中度提升因子  $\alpha_c$  较理想相位校正提升 20% 以上，同时明确了其在小  $N_F$  场景下受局部最优解约束的局限性。为突破梯度类方法瓶颈，引入强化学习算法，在  $r_0 = 0.031\text{m}$  的强湍流区域（SPGD 性能饱和

区)实现 $\alpha_c$ 达 4.53, 显著超越改进 SPGD 的 1.04, 验证了强化学习在高维、连续、非凸相位优化中的巨大潜力。综上, 本研究通过“验证可行性、提升有效性、探索优越性”的完整路径, 为水平传输非共轭湍流校正提供了切实算法方案与实验依据, 揭示了机器学习(尤其是强化学习)在自适应光学复杂波前控制中的广阔应用前景。未来实验验证可以在室内搭建等效湍流模拟装置, 验证改进 SPGD 和 RL 算法在可控环境下的校正效果, 然后开展水平大气外场实验, 采集实测湍流数据并进行算法性能评估, 重点验证 RL 算法在真实湍流中的适应性与实时性。

## 参考文献

- [1] Rao R Z 2006 *J. Atmos. Environ. Opt.* **1** 2 (in Chinese) [饶瑞中 2006 大气与环境光学学报 **1** 2]
- [2] Jiang N, Li X Y, Niu C H, Liu X 2022 *Laser Technol.* **46** (in Chinese) [姜楠, 李晓英, 牛春晖, 刘鑫 2022 激光技术 **46**]
- [3] Su Y 2005 *Wuli* **34** 583 (in Chinese) [苏毅 2005 物理 **34** 583]
- [4] Wang M J, Wei Y F, Ke X Z 2019 *Acta Phys. Sin.* **68** 140 (in Chinese) [王明军, 魏亚飞, 柯熙政 2019 物理学报 **68** 140]
- [5] Primmerman C, Fouche D 1976 *Applied Optics* **15** 990
- [6] Hardy J W 1998 *Adaptive optics for astronomical telescopes* (Oxford: Oxford University Press)
- [7] Ellerbroek B, Britton M C, Dekany R G, Gavel D T, Herriot G, Macintosh B A, Stoesz J A 2005 *Astronomical Adaptive Optics Systems and Applications II* San Diego, USA, August 3–4, 2005 p20
- [8] Rao C-H, Jiang W-H, Fang C, Ling N, Zhou W-C, Ding M-D, Zhang X-J, Chen D-H, Li M, Gao X-F 2003 *Chin. J. Astron. Astrophys.* **3** 576
- [9] Jiang W, Li M, Tang G, Ling N, Li M, Zheng D 1995 *Opt. Eng.* **34** 15
- [10] Jiang W H 2025 *Opto-Electron. Eng.* **45** 170489 (in Chinese) [姜文汉 2025 光电工程 **45** 170489]
- [11] Wu L, Ying J J, Geng B 2008 *Laser Infrared* **38** 974 (in Chinese) [武琳, 应家驹, 耿彪 2008 激光与红外 **38** 974]
- [12] Wang K D 2021 *Ph. D. Dissertation* (Xi'an: University of Chinese Academy of

- Sciences) (in Chinese) [王凯迪 2021 博士学位论文 (西安: 中国科学院大学)]
- [13] Wang Y J, Wu Y, Wang C, Hou Z H, Yang G, Gong Z B, Wang C H, Xian H, Li X Y, Wu X B 1998 *Chin. J. Quantum Electron.* **15** 164 (in Chinese) [王英俭, 吴毅, 汪超, 候再红, 杨刚, 龚知本, 王春红, 鲜浩, 李新阳, 吴旭斌 1998 量子电子学报 **15** 164]
- [14] Jiang W H, Yan P Y, Li M Q, Dai Z C 1990 *Acta Opt. Sin.* **10** 558 (in Chinese) [姜文汉, 严佩英, 李明全, 戴子昌 1990 光学学报 **10** 558]
- [15] Hardy J W, Lefebvre J E, Koliopoulos C 1977 *J. Opt. Soc. Am.* **67** 360
- [16] Liu H T, Mao M J 2019 *Acta Phys. Sin.* **68** 074205 (in Chinese) [刘厚通, 毛敏娟 2019 物理学报 **68** 074205]
- [17] Xie X G, Zhang J Z, Yue Y F, An J Z, Zhang F Z 2013 *High Power Laser Part. Beams* **25** 2536 (in Chinese) [谢晓钢, 张建柱, 岳玉芳, 安建祝, 张飞舟 2013 强激光与粒子束 **25** 2536]
- [18] Wang J, Zhang J, Wang S, Liu X, Zhang F, Su H 2026 *Fifth International Computational Imaging Conference (CITA 2025)* Suzhou, China, September 19–21, 2025 p293 (SPIE)
- [19] Cheng T, Guo S C, Wang N, Zhao M M, Wang S, Yang P 2023 *Chin. J. Lasers* **50** 1101008 (in Chinese) [程涛, 郭思成, 王宁, 赵孟孟, 王帅, 杨平 2023 中国激光 **50** 1101008]
- [20] Liu Y T, Xu M, Fu X H, Fu G W 2024 *Acta Phys. Sin.* **73** 203 (in Chinese) [刘宇韬, 徐苗, 付兴虎, 付广伟 2024 物理学报 **73** 203]
- [21] Fan L, Qiao C H, Feng X X, Ma H M, Fan C Y 2009 *J. Atmos. Environ. Opt.* **4** 183 (in Chinese) [范玲, 乔春红, 冯晓星, 马慧敏, 范承玉 2009 大气与环境光学学报 **4** 183]
- [22] Ma S J, Hao S Q, Zhao Q S 2021 *Acta Opt. Sin.* **41** 0601001 (in Chinese) [马圣杰, 郝士琦, 赵青松 2021 光学学报 **41** 0601001]
- [23] Zhao H, Kuang K D, Lü D K, Yu M J, An J, Zhang T Q 2022 *Infrared Laser Eng.* **51** 20210697 (in Chinese) [赵辉, 邝凯达, 吕典楷, 余孟洁, 安静, 张天骐 2022 红外与激光工程 **51** 20210697]
- [24] Xu Z X 2020 *Ph. D. Dissertation* (Chengdu: University of Electronic Science and Technology of China) (in Chinese) [许振兴 2020 博士学位论文 (成都: 电子科技大学)]

大学]

[25]Hu L, Hu S, Gong W, Si K 2020 *Opt. Lett.* **45** 3741

[26]Xu Z X, Yang P, Cheng T, Xu B, Li H P 2020 *Chin. J. Lasers* **47** 179 (in Chinese)

[许振兴, 杨平, 程涛, 许冰, 李和平 2020 中国激光 **47** 179]

[27]Wu X Q, Yang Q K, Huang H H, Qing C, Hu X D, Wang Y J 2023 *Acta Phys. Sin.*

**72** 049201 (in Chinese) [吴晓庆, 杨期科, 黄宏华, 青春, 胡晓丹, 王英俭 2023 物理学报 **72** 049201]

[28]Chen X Y, Dong Y, Guan J Y, Li S T, Yu Y J, Lü Y F 2014 *Acta Phys. Sin.* **63** 219

(in Chinese) [陈薪羽, 董渊, 管佳音, 李述涛, 于永吉, 吕彦飞 2014 物理学报 **63** 219]

[29]Zhang J Z, Zhang F Z, Su H, Hu P, Xie X G, Luo W 2021 *Acta Phys. Sin.* **70** 122

(in Chinese) [张建柱, 张飞舟, 苏华, 胡鹏, 谢晓钢, 罗文 2021 物理学报 **70** 122]

[30]Zhang F 2003 *Chin. J. Quantum Electron.* **20** 656

[31]Churnside J H, Lataitis R J 1990 *Appl. Opt.* **29** 926

[32]Wang Y J 1996 *Ph. D. Dissertation* (Hefei: Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences) (in Chinese) [王英俭 1996 博士学位论文 (合肥: 中国科学院安徽光学精密机械研究所)]

[33]Ge X L, Huang Y B, Fan C Y 2006 *J. Atmos. Environ. Opt.* **1** 27 (in Chinese) [葛

筱璐, 黄印博, 范承玉 2006 大气与环境光学学报 **1** 27]

[34]Haarnoja T, Tang H, Abbeel P, Levine S 2017 *Proceedings of the 34th International Conference on Machine Learning (ICML 2017)* Sydney, Australia, August 6–11, 2017 p1352 (PMLR)

[35]Haarnoja T, Zhou A, Hartikainen K, Tucker G, Ha S, Tan J, Kumar V, Zhu H, Gupta A, Abbeel P 2018 arXiv:1812.05905 [cs.LG]

[36]Schulman J, Chen X, Abbeel P 2017 arXiv:1704.06440 [cs.LG]

## Non-Conjugate Intelligent Correction Method for Atmospheric Turbulence in Long-Range Laser Propagation

WANG Jingting<sup>1)2)</sup> LIU Xiaohan<sup>1)</sup> WANG Shuai<sup>1)2)</sup> ZHANG Jianzhu<sup>1)</sup> XIE Xiaogang<sup>1)</sup>

1) (Institute of Applied Physics and Computational Mathematics, Beijing 100094, China)

2) (Graduate School of China Academy of Engineering Physics, Beijing 100088, China)

#### Abstract

Adaptive optics (AO) systems are indispensable for mitigating atmospheric turbulence-induced wavefront distortions in applications such as astronomical observation and free-space optical communication. The phase conjugation principle, which applies the conjugate of the measured aberration via a deformable mirror, serves as the foundational correction mechanism. However, in horizontal long-range laser propagation, the distributed and intense nature of atmospheric turbulence, coupled with diffraction effects, gives rise to strong non-conjugate characteristics. Under these conditions, the measured wavefront often contains phase singularities, causing conventional phase conjugation to degrade severely or even exacerbate beam distortion when the Fresnel number is small, thus fundamentally limiting its effectiveness. To transcend this bottleneck, this paper proposes a machine learning-driven intelligent correction framework designed to identify an optimal compensation wavefront that outperforms the ideal phase-conjugate solution. Firstly, we systematically deconstruct the standard stochastic parallel gradient descent (SPGD) algorithm, identifying its inherent limitations, comprising susceptibility to local optima, slow convergence, and sensitivity to initial conditions, in the context of non-convex, multi-peaked optimization landscapes induced by strong turbulence. To address these, we introduce a suite of enhancements: high-order Zernike polynomial representation (up to 65 orders) to expand the solution space, a Kolmogorov-spectrum-based exponential decay perturbation model to align exploration with physical priors, integration of the Adam optimizer for adaptive moment estimation and learning rate adjustment, an optimal parameter rollback mechanism for ensuring monotonic performance non-degradation, and an early stopping strategy with adaptive learning rate decay for computational efficiency. Our comprehensive simulations, based on a multi-phase screen wave-optics propagation model on the EasyLaser platform, systematically evaluate the improved SPGD algorithm across varying turbulence strengths and Fresnel numbers. The results demonstrate that the improved SPGD

achieves a performance enhancement factor of over 20% compared to ideal phase conjugation at a Fresnel number of 1.2 in moderate-to-strong turbulence, conclusively verifying the existence of superior non-conjugate correction wavefronts. However, its performance improvement margin narrows significantly in weaker diffraction regimes, exposing the inherent local-search limitations of the gradient-based SPGD framework. To achieve more robust global optimization, we further pioneer the application of a deep reinforcement learning (RL) approach based on the Soft Actor-Critic (SAC) algorithm. By formulating the phase correction task as a Markov decision process and integrating convolutional neural network-based policy and value networks, the RL agent learns a direct mapping from high-dimensional wavefront and far-field intensity images to optimal Zernike correction coefficients. The maximum entropy framework of SAC enables systematic global exploration, effectively escaping local optima that trap gradient-based methods. Remarkably, under strong turbulence conditions where both ideal phase conjugation and the improved SPGD algorithm fail (e.g., demonstrating a performance factor of only 1.04 by SPGD), the RL-driven controller achieves a performance enhancement factor of 4.53, restoring a near-diffraction-limited far-field spot. This study establishes a transformative paradigm for turbulence correction by fundamentally moving beyond the phase conjugation principle, confirming the practical feasibility and profound potential of reinforcement learning for complex, high-dimensional, continuous wavefront control in adaptive optics.

**Keywords:** Adaptive optics, Laser atmospheric propagation, Reinforcement learning,

Non-conjugate correction

---

† Corresponding author. E-mail: su\_hua@iapcm.ac.cn  
The first author. E-mail: 2352085250@qq.com