

Research Article

Xiaoshu Zhao, Haoze Lin, Huajin Chen*, Hongxia Zheng* and Jack Ng

Back-propagation-assisted inverse design of structured light fields for given profiles of optical force

<https://doi.org/10.1515/nanoph-2023-0101>

Received February 14, 2023; accepted April 19, 2023;

published online May 1, 2023

Keywords: back-propagation algorithm; inverse design; optical force; structured light field.

Abstract: Designing a monochromatic spatially-structured light field that recovers the pre-specified profile of optical force (OF) exerted on a particle is an inverse problem. It usually requires high dimensional optimization and involves lengthy calculations, thus remaining little studied despite decades of research on OF. We report here the first attempt to attack this inverse design problem. The modus operandi relies on the back-propagation algorithm, which is facilitated by the currently available machine learning framework, and, in particular, by an exact and efficient expression of OF that shows only polynomial and trigonometric functional dependence on the engineered parameters governing the structured light field. Two illustrative examples are presented in which the inversely designed structured light fields reproduce, respectively, a predefined spatial pattern of OF and a negative longitudinal OF in a transversely trapping area.

1 Introduction

Optical forces (OFs) exerted on an object illuminated by light have been extensively explored and exploited for intact manipulation of microparticles since the pioneering work of Ashkin in 1970s [1]. Despite decades of research [1–31] in many scientific fields ranging from optics to chemistry and biological science, the inverse design of monochromatic optical fields based on the pre-specified OF remains little explored, in particular, beyond the dipole limit. The earliest implementations of inverse design of some desired light patterns include finding holograms for optical trapping [32]. Later on, the technique was incorporated into commercial holographic trapping instruments as well as the design of tractor beams [33], but all limited to the Rayleigh (dipole) approximation. The inverse design of light fields capable of projecting a pre-specified profile of OF on particle of arbitrary size, however, remains challenging, due to the notoriously complex relationship between the OF and the parameters defining the light field, which impedes the optimization. This paper reports the first attempt to attack this problem.

Given a reasonably predefined spatial pattern of OF, in order to inversely design a structured optical field that can produce the desired pattern of OF on a micro-particle immersed in it, there are three issues to be solved.

The first issue is how to depict a general monochromatic optical field. Theoretically, a generic optical field can be described by its electric field of the form, see, e.g., [34] for a proof,

$$\mathbf{E} = \oint_{4\pi} \mathbf{e}_u e^{i\mathbf{k} \cdot \mathbf{r}} d\Omega_u, \quad (1)$$

where $\mathbf{u}$ is the real unit vector denoting the direction of the wave vector $\mathbf{k}$, and $\mathbf{e}_u$, depending on $\mathbf{u}$ but not on $\mathbf{r}$, is perpendicular to $\mathbf{u}$. The integration is over the unit sphere of the directions of wave vectors $\mathbf{k} = k\mathbf{u}$. The time dependence $e^{-i\omega t}$ is assumed. Numerically, Eq. (1) suggests

*Corresponding authors: **Huajin Chen**, School of Electronic Engineering, Guangxi University of Science and Technology, Liuzhou, Guangxi 545006, China; and Guangxi Colleges and Universities Key Laboratory of Microwave Communication and Micro-Nano Photoelectric Technology, Liuzhou, Guangxi 545006, China, E-mail: huajinchen13@fudan.edu.cn.

<https://orcid.org/0000-0002-8783-3691>; and **Hongxia Zheng**, School of Electronic Engineering, Guangxi University of Science and Technology, Liuzhou, Guangxi 545006, China; State Key Laboratory of Surface Physics and Department of Physics, Fudan University, Shanghai 200433, China; and Guangxi Colleges and Universities Key Laboratory of Microwave Communication and Micro-Nano Photoelectric Technology, Liuzhou, Guangxi 545006, China, E-mail: hxzheng18@fudan.edu.cn

Xiaoshu Zhao, State Key Laboratory of Surface Physics and Department of Physics, Fudan University, Shanghai 200433, China, E-mail: 20210190016@fudan.edu.cn

Haoze Lin, Department of Computer Science, Columbia University, New York, NY 10027, USA, E-mail: hl3459@columbia.edu

Jack Ng, Department of Physics, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China, E-mail: wuzh3@sustech.edu.cn

that a generic field can be cast into a summation of a myriad of plane waves

$$\mathbf{E} = \sum_{j=1}^{n_p} \mathbf{E}_j e^{i\mathbf{k}_j \cdot \mathbf{r}} \quad \text{with} \quad \mathbf{E}_j = E_0(p_j \hat{\boldsymbol{\theta}}_{k_j} + q_j \hat{\boldsymbol{\phi}}_{k_j}), \quad (2)$$

where n_p is the number of the constituent plane waves, $\mathbf{E}_j$ is the complex amplitude vector, $\mathbf{k}_j = k\hat{\mathbf{k}}_j$ is the real wave vector for the j th plane wave. Two complex numbers p_j and q_j characterize the amplitude, polarization, and phase at the position vector $\mathbf{r} = 0$. The unit vectors $\hat{\boldsymbol{\theta}}_{k_j}$ and $\hat{\boldsymbol{\phi}}_{k_j}$ denote, respectively, the directions of increasing polar angle and azimuthal angle in the spherical coordinate system for the j th wave vector $\mathbf{k}_j$ so that $\hat{\mathbf{k}}_j$, $\hat{\boldsymbol{\theta}}_{k_j}$, and $\hat{\boldsymbol{\phi}}_{k_j}$ form a right-handed triplet of unit vectors. With Eq. (2), one is, in principle, ready to recover arbitrary spatially-structured light [35, 36] by tuning the number n_p of the constituent plane waves, and the polar angle θ_j and azimuthal angle φ_j of $\mathbf{k}_j$, as well as the two complex numbers p_j and q_j for each constituent plane wave. Given all these parameters, with the use of spatial light modulators, most theoretically designed structured light fields can be realized experimentally, with the idea of angular spectrum representation [37].

The second issue is the calculation of the OF exerted on a particle illuminated by the structured light field given by Eq. (2). The inverse design usually involves extensive calculations [38–40], the “brute-force” calculation is obviously not feasible. The generalized Lorenz-Mie theory (GLMT) [41] works well for calculating OF on spherical particles, see, e.g., [42, 43]. However, for the structured light field in Eq. (2) that is composed of multiple interferential plane waves, the GLMT requires expanding each plane wave in a series of vector spherical wave functions [44, 45] to obtain the beam shape coefficients [41]. Although this can be worked out analytically, see, e.g., [41–44] it would require computing the associated Legendre functions and their derivatives. This impedes the analytical calculation of the gradient of loss (see, below) with respect to the engineered parameters with standard back-propagation [46] by direct use of the common-used machine learning frameworks [47, 48]. In this paper we adopt our recently-developed approach [34, 49–52], which calculates the OF exerted on a particle immersed in an arbitrary number of interferential plane waves without resorting to the vector spherical wave functions. The approach is further reformulated in such a way that the expression of the OF involves only the polynomials and trigonometric functions. So it fully facilitated the application of the back-propagation within the currently available machine learning frameworks, not having to resort to neural networks (NNs) for approximating the OF calculation. For comparison, we

have trained a NN to approximate the OF with high degrees of precision. The trained NN generates the results in time less by around 50 % compared with our reformulated approach. Nevertheless, it takes much time in training, and, to be worse, needs to be retrained when the size or scattering property of particle is changed. The details of the OF calculation formulation will be given in Section 2.

The third issue concerns the optimization. To this end we design a loss function measuring the difference between the pre-specified OF and the calculated OF with the given engineered parameters chosen from n_p , θ_j , φ_j , p_j and q_j , subject to some physical constraints for the easiness of experimental realization. Then with the standard back-propagation, the gradient of the loss function with respect to the engineered parameters are obtained “analytically”. This is achieved by using the reformulated OF expression written in terms of only the polynomials and trigonometric functions while circumventing the much more complicated special functions in the framework of GLMT. Next, the ADAM optimizer [53, 54] is used to optimize the engineered parameters for decreasing the loss function. This is somewhat similar to the optimization process where the direct search algorithm is used to calculate holograms [55]. After some epochs of optimization, we arrive at a structured light field in the form of Eq. (2) which reproduces the desired spatial pattern of OF to a reasonable accuracy. Details of the optimization procedure will be presented in Section 3 by two illustrative examples.

2 Formulation

In this section, the Cartesian multipole expansion theory [34, 49–52] is first recapitulated for the calculation of the OF on a spherical particle. Then we reformulate the OF expression to a form that show only the polynomial and trigonometric functional dependence on the engineered parameters p_j , q_j , θ_j and φ_j . This is done for the purpose of the calculation of gradient of the loss function by direct use of the back-propagation within the current machine learning framework, avoiding the tedious mathematical calculations involving the associated Legendre functions and their derivatives in the framework of GLMT. For simplicity, we set $E_0 = k = \varepsilon_0 = \mu_0 = c = 1$ during derivation. The final expression for OF is thus in unit of $\varepsilon_0 \varepsilon_b E_0^2 / k^2$, where ε_b denotes the relative electric permittivity in the background medium.

According to the Cartesian multipole expansion theory [34, 49–51], the time-averaged OF $\langle F \rangle$ on a spherical particle illuminated by a generic structured light field given by Eq. (2) can be written as

$$\begin{aligned}
\langle F \rangle &= \sum_l \left[\langle F_{\text{int}}^{e(l)} \rangle + \langle F_{\text{int}}^{m(l)} \rangle + \langle F_{\text{rec}}^{e(l)} \rangle + \langle F_{\text{rec}}^{m(l)} \rangle + \langle F_{\text{rec}}^{x(l)} \rangle \right], \\
\langle F_{\text{int}}^{e(l)} \rangle &= -\frac{\pi(2l+1)}{l(l+1)} \text{Im} \sum_{i,j} a_l \left[Q_{l,i}^{(1)} Z_{ee,ij}^{(1)} - Q_{l,i}^{(2)} Z_{mm,ij}^{(1)} \right], \\
\langle F_{\text{int}}^{m(l)} \rangle &= -\frac{\pi(2l+1)}{l(l+1)} \text{Im} \sum_{i,j} b_l \left[Q_{l,i}^{(1)} Z_{mm,ij}^{(1)} - Q_{l,i}^{(2)} Z_{ee,ij}^{(1)} \right], \\
\langle F_{\text{rec}}^{e(l)} \rangle &= -\frac{\pi}{2(l+1)^2} \text{Im} \sum_{i,j} a_{l+1} a_l^* \left[R_{l,ij}^{(1)} Z_{ee,ij}^{(1)*} \right. \\
&\quad \left. - R_{l,ij}^{(2)} Z_{mm,ij}^{(1)*} - 4i R_{l,ij}^{(3)} S_{em,ij}^{(1)*} + R_{l,ij}^{(4)} Z_{mm,ij}^{(1)} \right. \\
&\quad \left. - R_{l,ij}^{(5)} Z_{mm,ij}^{(1)} + 4i R_{l,ij}^{(6)} S_{em,ij}^{(1)} \right], \\
\langle F_{\text{rec}}^{m(l)} \rangle &= -\frac{\pi}{2(l+1)^2} \text{Im} \sum_{i,j} b_{l+1} b_l^* \left[R_{l,ij}^{(1)} Z_{mm,ij}^{(1)*} \right. \\
&\quad \left. - R_{l,ij}^{(2)} Z_{ee,ij}^{(1)*} - 4i R_{l,ij}^{(3)} S_{em,ij}^{(1)*} + R_{l,ij}^{(4)} Z_{mm,ij}^{(1)} \right. \\
&\quad \left. - R_{l,ij}^{(5)} Z_{ee,ij}^{(1)} + 4i R_{l,ij}^{(6)} S_{em,ij}^{(1)*} \right], \\
\langle F_{\text{rec}}^{x(l)} \rangle &= -\frac{\pi(2l+1)}{2l^2(l+1)^2} \text{Im} \sum_{i,j} a_l b_l^* \left[R_{l,ij}^{(4)} \left(Z_{mm,ij}^{(1)} - Z_{ee,ij}^{(1)*} \right) \right. \\
&\quad \left. + R_{l,ij}^{(5)} \left(Z_{mm,ij}^{(1)*} - Z_{ee,ij}^{(1)} \right) \right. \\
&\quad \left. + 4i R_{l,ij}^{(6)} S_{em,ij}^{(1)*} + 4i R_{l,ij}^{(7)} S_{em,ij}^{(1)} \right].
\end{aligned} \tag{3}$$

where the superscript $*$ denotes the complex conjugate, a_l and b_l are the Mie coefficients [56], $Q_{l,ij}^{(n)}$ with $n = 1, 2$ and $R_{l,ij}^{(n)}$ with $n = 1, 2, \dots, 7$ are polynomials of $x_{ij} = \hat{\mathbf{k}}_i \cdot \hat{\mathbf{k}}_j$ given by Eq. (15) in Supplementary Information, and the vector quantities read

$$\begin{aligned}
\mathbf{S}_{em,ij}^{(1)} &= \mathbf{E}_i \times \mathbf{B}_j^* e^{i(\mathbf{k}_i - \mathbf{k}_j) \cdot \mathbf{r}}, \\
\mathbf{Z}_{ee,ij}^{(1)} &= \mathbf{G}_{ee,ij}^{(1)} - i \mathbf{S}_{em,ij}^{(1)}, \\
\mathbf{Z}_{mm,ij}^{(1)} &= \mathbf{G}_{mm,ij}^{(1)} - i \mathbf{S}_{em,ij}^{(1)*}, \\
\mathbf{G}_{ee,ij}^{(1)} &= -i(\mathbf{k}_j \cdot \mathbf{E}_i) \mathbf{E}_j^* e^{i(\mathbf{k}_i - \mathbf{k}_j) \cdot \mathbf{r}}, \\
\mathbf{G}_{mm,ij}^{(1)} &= -i(\mathbf{k}_i \cdot \mathbf{B}_j) \mathbf{B}_j^* e^{i(\mathbf{k}_i - \mathbf{k}_j) \cdot \mathbf{r}},
\end{aligned} \tag{4}$$

with $\mathbf{E}_i$ and $\mathbf{B}_i$ denoting, respectively, the complex amplitude vectors for the electric and magnetic fields of the i th plane wave.

Equations (3) and (4) represent an alternative approach for the forward calculation of the OF exerted on arbitrarily-sized spherical particle immersed in any given structured light described by Eq. (2). The formulation, which is based on the Cartesian multipole expansion, does not need to expand each plane wave in Eq. (2) in a series of the vector spherical

wave functions as in the spherical multipole-based formulation [41–43, 57–59]. It therefore leads to a greatly enhanced computational speed that is even comparable to a trained NN approximation. For the purpose of inverse design, we reformulate the OF expressions in a form that further facilitates the calculation of the gradient of loss function with respect to the engineered parameters based on the standard back-propagation [46]. The reformulated OF expression read, after lengthy algebra,

$$\langle F \rangle = \text{Re} \sum_{i,j} \langle F_{ij} \rangle e^{i(\mathbf{k}_i - \mathbf{k}_j) \cdot \mathbf{r}}, \tag{5}$$

where $\langle F_{ij} \rangle$ are $\mathbf{r}$ independent, and its Cartesian components are worked out to be quadratic form in terms of polarization parameters p and q ,

$$\begin{aligned}
F_{x,ij} &= N_{ij}^{(1)} p_i p_j^* + N_{ij}^{(2)} q_i q_j^* + N_{ij}^{(3)} p_i q_j^* + N_{ij}^{(4)} q_i p_j^*, \\
F_{y,ij} &= N_{ij}^{(5)} p_i p_j^* + N_{ij}^{(6)} q_i q_j^* + N_{ij}^{(7)} p_i q_j^* + N_{ij}^{(8)} q_i p_j^*, \\
F_{z,ij} &= N_{ij}^{(9)} p_i p_j^* + N_{ij}^{(10)} q_i q_j^* + N_{ij}^{(11)} p_i q_j^* + N_{ij}^{(12)} q_i p_j^*.
\end{aligned} \tag{6}$$

The coefficients $N_{ij}^{(n)}$, with $n = 1, 2, \dots, 12$, are products of polynomials of x_{ij} and the trigonometric functions of θ_i , φ_i , θ_j , and φ_j , befitting the back-propagation. Their explicit expressions are given in Supplementary Information by Eq. (16).

With Eqs. (5) and (6), the calculation of OF on a spatial grid of 18,040 points in the next section takes only a fraction of second on a mobile computer with NVIDIA GeForce GTX 1060 (6G), allowing for the expeditious optimization with the gradient descent method.

3 Demonstration results and discussion

In this section we present two demonstration examples for the inverse design of structured light fields from the pre-specified OF.

3.1 Trapping inside of a prespecified pattern

The first example concerns capturing particle transversely inside of a predefined pattern. The non-magnetic particle to be trapped transversely has relative permittivity $\epsilon_s = 2.53$ and radius $r_s = 0.5\lambda_0$ with $\lambda_0 = 532$ nm denoting the wavelength of the operating structured light in vacuum. The particle is immersed in water with a refractive index of 1.33. The OF exerted on the particle is decomposed into the longitudinal supporting OF $Y_z(x, y)$ in z direction (the particle is not trapped along z by the OF), which pushes the particle

against, say, the glass plate, and the transverse trapping OF $Y_t(x, y)$, which confines the particle inside of the predefined pattern of ‘FDU’. A rectangular area with x extending from 0 to 40.8λ and y from 0 to 17.4λ is set to be the region for optical manipulation. Here λ stands for the wavelength in the background medium (water). The area is divided into a grid of $205 \times 88 = 18,040$ points, with a resolution (grid size) of 0.2λ . Such an area is referred to as training area and the grid points inside referred to as training samples throughout the paper, since we are “training” a structured light field to reproduce a pre-designed pattern of OF.

Initially, the value of the longitudinal OF $Y_z(x, y)$ at $z = 0$ is set to 3 within the pattern ‘FDU’ and 0 otherwise, with its profile displayed in Figure 4(a) in Supplementary Information. To avoid the unphysical sharp change of OF at the edge of the pattern, we apply 8 discrete convolutions (see, e.g., [60]) to the $Y_z(x, y)$ profile and get a smoothed profile with practical physical feasibility. The convolution, padded with a 1×1 border of zeros and using 1×1 strides, has a kernel of 3×3 with all values set to $1/9$. The output size of each convolution stays the same as the input size, resulting in a smoothed profile of the longitudinal OF as exhibited in Figure 4(b). Each convolution here is indeed equivalent to averaging the value of longitudinal OF over $3 \times 3 = 9$ grid points. To confine the particle inside of the ‘FDU’ pattern, the values of the two Cartesian components of the transverse trapping force $Y_t(x, y)$ at $z = 0$ plane are set to be

$$\begin{aligned} Y_x(x, y) &= 6[Y_z(x + 0.2\lambda, y) - Y_z(x - 0.2\lambda, y)], \\ Y_y(x, y) &= 6[Y_z(x, y + 0.2\lambda) - Y_z(x, y - 0.2\lambda)], \end{aligned} \quad (7)$$

which has its profile as shown in Figure 4(c).

To realize the OF patterns as shown in Figure 4(b) and (c), we use $n_p = 100$ plane waves to construct the structured light of the form Eq. (2). To facilitate the implementation using the space light modulator, all the plane waves are supposed to have the electric field polarized parallel to the x - o - z plane.

This is achieved by setting p_i and q_i for the i th plane wave to

$$\begin{aligned} p_i &= \frac{\cos \varphi_i}{\sqrt{(\cos \varphi_i)^2 + (\cos \theta_i \sin \varphi_i)^2}} A e^{i\gamma_i}, \\ q_i &= -\frac{\cos \theta_i \sin \varphi_i}{\sqrt{(\cos \varphi_i)^2 + (\cos \theta_i \sin \varphi_i)^2}} A e^{i\gamma_i}, \end{aligned} \quad (8)$$

where θ_i and φ_i are, respectively, the polar angle and azimuthal angle of the i th wave vector, and γ_i is its phase at $z = 0$. All plane waves are set to have the same amplitude A . We optimize the structured light through the change of the engineered parameters θ_i , φ_i , γ_i , and A such that the

optimized light can reproduce the OF profiles shown in Figure 4(b) and (c). Initially, A is set to 0.05, and φ_i and γ_i are randomly distributed in the range of 0 to 2π , while θ_i randomly distributed between 0 and 0.08π . During the optimization, all θ_i are confined to be less than 0.16π . It is noted that given OF as displayed in Figure 4(b) and (c), we actually do not know what the value of A should be. So A is set to be an engineered parameter, though it does not change the pattern of OF.

The loss function L is then set to be the mean square deviation with an $L2$ regularization of A , reading

$$L = \frac{1}{N} \sum_{(x,y)} \left| \mathbf{a} \cdot [F(x, y) - Y(x, y)] \right|^2 + \eta A^2, \quad (9)$$

where N denotes the number of training samples. $Y(x, y)$ and $F(x, y)$ are the designed value (see, Figure 4(b) and (c)) and the calculated value of the OF, respectively. The vector $\mathbf{a}$ is used to adjust the proportion of the tolerance in transverse and longitudinal OFs. η is a hyperparameter for regularization with respect to amplitude. The summation runs over all $N = 18,040$ grid points (training samples) in the training area. In the optimization (training), we set $\mathbf{a} = (2.5, 2.5, 1)$ and $\eta = 200$.

The ADAM optimizer [53, 54] is used to decrease the loss function L , with the gradient of L with respect to the engineered parameters θ_i , φ_i , γ_i , and A obtained analytically with the standard back propagation [46], available in the machine learning framework like Tensorflow [47]. The optimization is completed within a few minutes with learning rate 0.001. The profiles of the longitudinal and transverse OFs exerted on the particle, the light intensity in the x - o - y plane, the loss function L versus epoch, and the distribution of the wave vectors for the $n_p = 100$ plane waves are shown in Figure 1. In Figure 5 in Supplementary Information we also show the gradient (conservative) and scattering (non-conservative) OFs based on the decomposition formulation proposed in [49–51, 59]. The scattering force dominates the longitudinal supporting OF, while the transverse trapping force stems mainly from the gradient force, with its optical potential shown in Figure 5(e).

The transverse trapping of particle inside of the pre-specified pattern is not limited to $z = 0$ plane. Particle can be so captured within the range of $-3\lambda \lesssim z \lesssim 3\lambda$, as shown in Figure 6 in Supplementary Information, where the results at $z = \pm 3\lambda$ are displayed. In addition, the optimized structured light obtained for particle with $r_s = 0.5\lambda_0$ also work quite well for a particle of radius ranging from $0.1\lambda_0$ to $1.0\lambda_0$. A case which yields a structured light field that can confine particle of radius $r_s = 1.2\lambda_0$ inside of the ‘FDU’ pattern is also shown in Figure 7 in Supplementary Information.

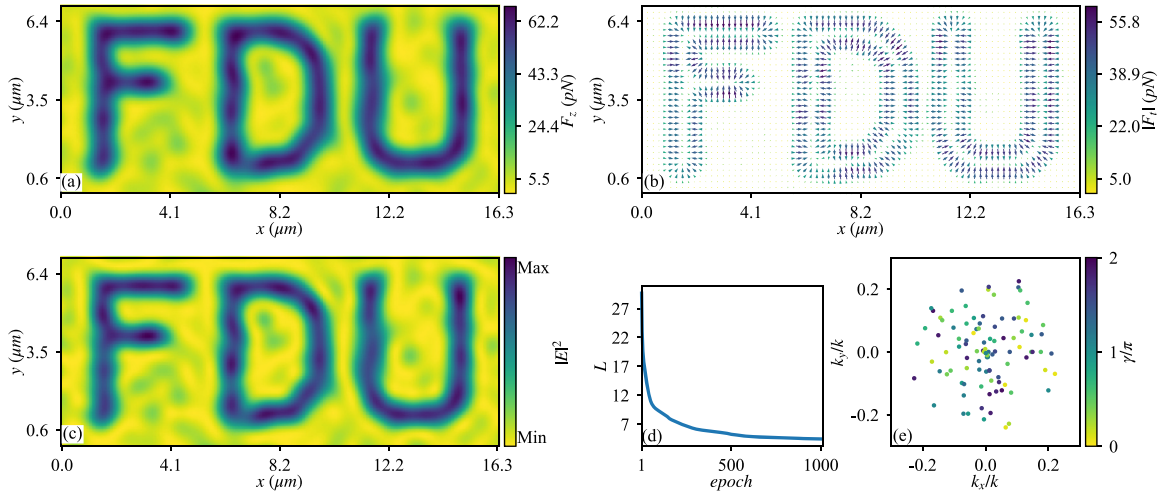


Figure 1: The calculated longitudinal monotonic OF F_z (a) and transverse trapping OF F_t (b) exerted on a particle with $r_s = 0.5\lambda_0$ and $\varepsilon_s = 2.53$, placed at $z = 0$ in water and illuminated by the structured light optimized for producing the ‘FDU’ pattern of OF. Each plane wave has $A E_0 = 8.68 \times 10^5$ V/m. In panel (b) the arrows denote the directions of F_t , while their lengths and colors denote the magnitudes of F_t . (c) The profile of $|E|^2$ of the optimized structured light in the x - y plane. (d) The change of the loss function versus epoch of optimization. (e) The wave vectors and phases of the constituent plane waves, where k_x and k_y of each point denote the transverse Cartesian components of the wave vector, with the color denoting the phase γ , see, Eq. (8).

3.2 Optical pulling force

As the second demonstration example, in this subsection, we apply the inverse design algorithm to find some scenarios of optical pulling.

Pulling force against the wave propagation was first theoretically discussed in acoustic wave by Marston [61]. The earliest experimental demonstration of an optical pulling force was reported in [62], where optical pulling is achieved by concurring gradient and scattering forces in solenoid beams. The optical pulling force (negative OF) originated exclusively from the scattering force was first identified in the case of particle illuminated by Bessel beams [12, 13, 22], which can be modeled by a superposition of plane waves with their wave vectors $\mathbf{k}_i$ lying on a conical surface. That means, the z components of all wave vectors are the same, giving in principle rise to a non-diffractive optical field where the OF stays unchanged along z axis. As a demonstration example, for simplicity, we assume that all wave vectors of the constituent plane waves making up the structured light have an equal polar angle.

We choose to recover a simple case as in [63], where a nonmagnetic particle with a permittivity $\varepsilon_s = 7$ is located in vacuum and illuminated by a zero-order transverse magnetic (TM) Bessel beam with cone angle $\alpha = 70^\circ$. The particle size is $kr_s = 2.7$. We optimize a structured light field consisting of 6 plane waves to achieve the negative longitudinal OF in a transversely trapped area on the same particle. In practice, such a pulling force is expected to be realizable on a

particle fabricated by materials such as silicon carbide (SiC) or titanium dioxide (TiO₂), which have refractive indices 2.65 and 2.95, respectively, at wavelength around 530 nm [64].

As the OF keeps invariable with z , we only need to specify the OF in the x - y plane. So we first choose a circle with a radius of $0.44\lambda_0$ centered at the origin as the training area, and select training samples evenly in the training area with the resolution of $0.08\lambda_0$. The number of training samples reduces thus to $N = 97$ and the value of the OF $Y(x, y)$ is prespecified according to

$$Y(x, y) = -2.14\hat{\rho} - 2\hat{z}, \quad (10)$$

where $\hat{\rho}$ and $\hat{z}$ are unit base vectors in the cylindrical coordinate system. Obviously, a particle subject to this OF can be stably trapped transversely near the origin, while being pulled towards light source by the negative longitudinal OF.

Next, we choose the polar angle of all 6 wave vectors to be $\theta_i = 70^\circ$ [63], set the azimuthal angle of the i th plane wave to $\varphi_i = (i - 1)\pi/3$ to retain cylindrical symmetry. The polarization parameters are rewritten as

$$\begin{aligned} p_i &= A_i \cos \beta_i e^{i\gamma_i^{(1)}}, \\ q_i &= A_i \sin \beta_i e^{i\gamma_i^{(2)}}, \end{aligned} \quad (11)$$

where A_i , β_i , $\gamma_i^{(1)}$, and $\gamma_i^{(2)}$ are engineered parameters, denoting the real amplitude, the polarization and the phase at $z = 0$ of each plane wave, respectively.

Similar to Eq. (9), the loss function L is defined by

$$L = \frac{1}{N} \sum_{(x,y)} |\mathbf{a} \cdot (\mathbf{F}(x, y) - \mathbf{Y}(x, y))|^2 + \eta \sum_i A_i^2, \quad (12)$$

where $\eta = 4$, the summation is over all $N = 97$ training samples, and the hyperparameter $\mathbf{a}$ is chosen to

$$\mathbf{a} = \begin{cases} (1, 1, 10), & k\sqrt{x^2 + y^2} \leq 0.53, \\ (1, 1, 0), & \text{otherwise.} \end{cases} \quad (13)$$

Again, the ADAM optimizer is used to decrease the loss function L . The gradient of L with respect to the engineered parameters are obtained again by the back-propagation. The final results turn out to be $q_i = 0$ with all p_i being equal, corresponding to the TM field [63], in which the magnetic field is polarized in a plane normal to z . The results indicate that all the 6 constituent plane waves have the same amplitude, phase, and polarization. The profiles of the longitudinal and transverse OFs, light intensity, and change of loss function versus training epoch are shown in Figure 2.

Due to the limited number of constituent plane waves, the structured light, and thus the transverse spatial distribution of OF show obvious periodicity, resulting in a periodic array of trapped area with negative OF. Figure 2(a)–(c) actually exhibit a period in the x - y plane. Besides, a comparison between Figure 2(b) and (c) suggests that the trapping originates from the gradient force, as confirmed by the calculated gradient and scattering forces (not shown here). The particle is trapped at the intensity maxima.

Next, let us look at a larger particle with radius $r_s = 0.5\lambda_0$ and $\varepsilon_s = 2.53$ in vacuum.

To generate a transverse trapped area with longitudinal pulling OF, we still use Eq. (10) for OF and Eq. (12) for loss function, with $\eta = 5$ and

$$\mathbf{a} = \begin{cases} (1, 1, 20), & k\sqrt{x^2 + y^2} \leq 1.3, \\ (1, 1, 0), & \text{otherwise.} \end{cases} \quad (14)$$

The radius of the training area is increased to $0.84\lambda_0$. The resolution (grid size) keeps to be $0.08\lambda_0$ and the number of training samples (grid points that are included in the calculation of loss) becomes $N = 349$. The optimization becomes more difficult as the particle gets larger. Therefore, we increase the number of plane waves that make up the structured light to $n_p = 10$ and all polar angle to $\theta = 0.44\pi$ (79.2°). The azimuthal angle of the i th plane wave is set to $\varphi_i = (i - 1)\pi/5$, still for the purpose of cylindrical symmetry. Rewriting the polarization parameters as shown in Eq. (11), we have the engineered parameters $\beta_i, \gamma_i^{(1)}, \gamma_i^{(2)}$, and A_i , with A_i confined to be greater than 0.05.

With the ADAM optimizer and back-propagation, we optimize the structured light, with learning rate equal to 0.001, to produce the desired OF. The optimized OFs, light intensity and change of loss function are shown in Figure 3. The engineered parameters of each plane wave optimized for the desired OF are shown in Table 1 in Supplementary Information.

Different from the previous case shown in Figures 2 and 3 implies that the particle is transversely trapped at light intensity minimum, a somewhat counter-intuitive phenomenon due to the effect of particle size. This is of no surprise. Because for particle of finite size, the gradient force is proportional to the gradient of light intensity that is a weighted average over the volume occupied by the particle, instead of the intensity exactly at particle center. In Figure 3, the particle centered at the origin actually occupies a volume with the largest averaged intensity and is thus trapped therein.

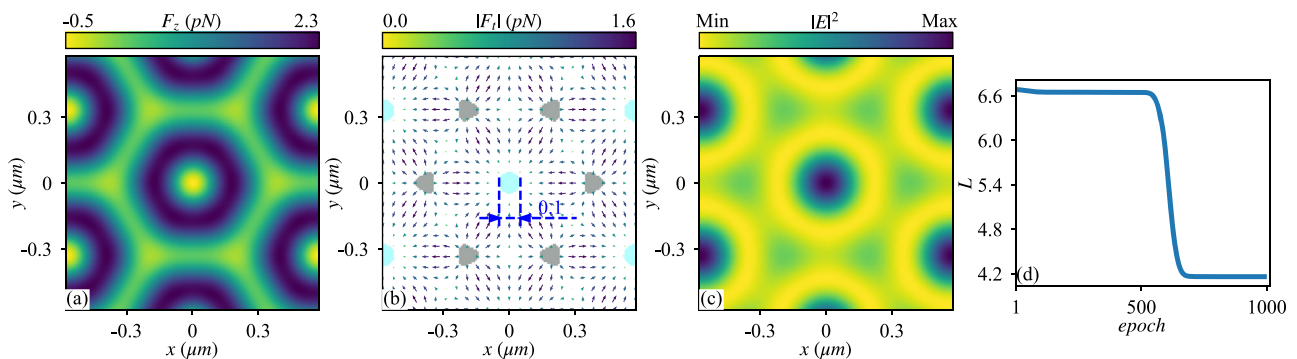


Figure 2: The same as Figure 1 except that the particle is placed in vacuum and has a radius of $r_s = 0.43\lambda_0$ and $\varepsilon_s = 7$ while the structured light is optimized to produce a transverse trapped area in the central part with a negative longitudinal OF. In panel (b), the shaded parts denote an area of negative longitudinal OF, with the colored ones exhibiting stable transverse trapping. The optimized structured light consists of 6 plane waves, all of which have the magnetic field perpendicular to z . The particle is trapped at the intensity maximum basically by the gradient OF.

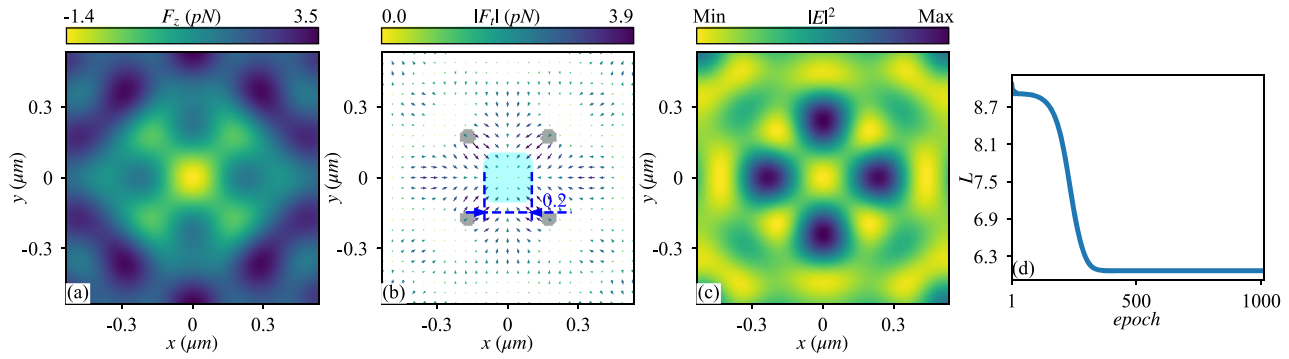


Figure 3: The same as Figure 2 except that the particle has a radius of $r_s = 0.5\lambda_0$ and $\epsilon_s = 2.53$ while the optimized structured light is composed of 10 plane waves with a diversity of polarization as presented in Table 1 in Supplementary Information and the first plane wave has $A_1 E_0 = 8.68 \times 10^5$ V/m. It produces a trapped area with negative longitudinal OF denoted by the central shaded area in panel (b). Although the dielectric particle is transversely trapped by the gradient OF, it is stabilized at the intensity minimum, contrast to the conventional case shown in Figure 2.

The optimized p_i and q_i in Table 1 imply $p(\varphi_i + \pi) \approx p(\varphi_i)$ and $q(\varphi_i + \pi) \approx q(\varphi_i)$, suggestive of the structured light propagating with a vortex factor $e^{2i\varphi}$. We have thus tried to generate a trapped area with pulling OF on such a particle using the Bessel beams of zero- and second-order, respectively. Both beams consist of the superimposed TM and transverse electric (TE) modes, with the relative amplitude and phase of the two modes optimized for our purpose. The optimal results are shown in Figure 8 in Supplementary Information. While the zero-order beam fails to produce a transversely trapped area for the pulling OF, the second-order beam, which carries orbital angular momentum, exhibits a vortex in the OF field, likely ruining the transverse confinement [7] of the particle to the area of pulling OF.

4 Conclusions

In this paper, we have demonstrated a back-propagation-based inverse design approach for devising a structured light field to achieve the predefined pattern of OF exerted on a particle of arbitrary size. The approach is built on reformulating the OF expression in a closed form in terms of only the elementary functions, avoiding the special functions within the conventional framework of GLMT. It thus permits the direct use of the back-propagation to analytically compute the gradients of the loss function with respect to the engineered parameters that govern the structured light field, without resorting to NNs. Besides facilitating the analytical calculation of the gradients, the NN is used also to expedite many physics simulations by its excellence in function approximation. Our algorithm is exact in the OF calculation, with the computing speed comparable to the trained NN. Then, with appropriately designed loss

function, the ADAM optimizer is applied to optimize the engineered parameters, leading eventually to a structured light field that reproduce the predefined profile of the OF with reasonable accuracy. Some demonstration examples are presented to illustrate the approach. They are completed on a mobile computer with NVIDIA GeForce GTX 1060 (6G), indicating that it is not hardware demanding as an approach to attack the inverse design problem that usually needs large computation resources. In addition, the algorithm supports the adjustment of the training samples as well as the change of the number of waves that form the structured light, showing great flexibility. Also, our approach can be directly applied to reproduce the three-dimensional pre-designed patterns of OF for both transverse and longitudinal optical manipulation [65–67]. Work along this line is in progress.

Finally, our inverse design algorithm can benefit further from the ideas in the theory of NNs to improve its functionality. Here are two examples. (i) Regularization plays an important role in our optimization towards the appropriate structured light. Without regularization, the amplitude of the light field tends to increase, resulting in, sometimes, a much smaller OF in the training area than in other areas, or, in some other cases, leading to very large amplitude difference between different constituent plane waves. This is obvious not conducive to optical manipulation or experimental implementation of the designed structured light by spatial light modulators. This is reminiscent of the ‘overfitting’ [68] in NNs. (ii) In this paper, we have given the approximate wave vector directions before training. When doing the inverse design for new OF requirements, we can start with a larger number of constituent plane waves and pre-train without regularization. The plane waves with higher amplitude in the training usually suggest their more

significant effect on producing the desired OF. We may imitate ‘pruning’ [69, 70] to remove those plane waves with low amplitude, then introduce the regularization, and restart the optimization with the approximate range of wave vectors adopted from the survived ones.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This work was supported by National Natural Science Foundation of China (12074084, 12174076, 12204117, and 12074169), Natural Science Foundation of Guangxi Province (2022JG110006 and 2021GXNSFDA196001), Open Project of State Key Laboratory of Surface Physics in Fudan University (KF2022_15), Stable Support Plan Program of Shenzhen Natural Science Fund (20200925152152003).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

- [1] A. Ashkin, “Acceleration and trapping of particles by radiation pressure,” *Phys. Rev. Lett.*, vol. 24, pp. 156–159, 1970.
- [2] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu, “Observation of a single-beam gradient force optical trap for dielectric particles,” *Opt. Lett.*, vol. 11, no. 5, pp. 288–290, 1986.
- [3] A. Ashkin, “Optical trapping and manipulation of neutral particles using lasers,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, pp. 4853–4860, 1997.
- [4] D. G. Grier, “A revolution in optical manipulation,” *Nature*, vol. 424, no. 6950, pp. 810–816, 2003.
- [5] A. Ashkin, *Optical Trapping and Manipulation of Neutral Particles Using Lasers*, Singapore, World Scientific, 2006.
- [6] J. R. Moffitt, Y. R. Chmela, S. B. Smith, and C. Bustamante, “Recent advances in optical tweezers,” *Annu. Rev. Biochem.*, vol. 77, p. 205, 2008.
- [7] J. Ng, Z. F. Lin, and C. T. Chan, “Theory of optical trapping by an optical vortex beam,” *Phys. Rev. Lett.*, vol. 104, p. 103601, 2010.
- [8] F. Fazal and S. Block, “Optical tweezers study life under tension,” *Nat. Photonics*, vol. 5, pp. 318–321, 2011.
- [9] K. Dholakia and T. Čížmár, “Shaping the future of manipulation,” *Nat. Photonics*, vol. 5, pp. 335–342, 2011.
- [10] M. Padgett and R. Bowman, “Tweezers with a twist,” *Nat. Photonics*, vol. 5, no. 6, pp. 343–348, 2011.
- [11] M. L. Juan, M. Righini, and R. Quidant, “Plasmon nano-optical tweezers,” *Nat. Photonics*, vol. 5, no. 6, pp. 349–356, 2011.
- [12] J. Chen, J. Ng, Z. F. Lin, and C. T. Chan, “Optical pulling force,” *Nat. Photonics*, vol. 5, no. 9, pp. 531–534, 2011.
- [13] A. Novitsky, C. W. Qiu, and H. F. Wang, “Single gradientless light beam drags particles as tractor beams,” *Phys. Rev. Lett.*, vol. 107, no. 20, p. 203601, 2011.
- [14] S. Sukhov and A. Dogariu, “Negative nonconservative forces: optical ‘tractor beam’ for arbitrary objects,” *Phys. Rev. Lett.*, vol. 107, p. 203602, 2011.
- [15] R. W. Bowman and M. J. Padgett, “Optical trapping and binding,” *Rep. Prog. Phys.*, vol. 76, p. 026401, 2013.
- [16] O. Brzobohatý, V. Karásek, M. Šiler, L. Chvátal, T. Čížmár, and P. Zemánek, “Experimental demonstration of optical transport, sorting and self-arrangement using a ‘tractor beam,’” *Nat. Photonics*, vol. 7, no. 2, pp. 123–127, 2013.
- [17] A. Dogariu, S. Sukhov, and J. Sáenz, “Optically induced ‘negative forces,’” *Nat. Photonics*, vol. 7, pp. 24–27, 2013.
- [18] K. Y. Bliokh, A. Y. Bekshaev, and F. Nori, “Extraordinary momentum and spin in evanescent waves,” *Nat. Commun.*, vol. 5, no. 3, p. 3300, 2014.
- [19] Y. B. Aleksandr, Y. K. Bliokh, and F. Nori, “Transverse spin and momentum in two-wave interference,” *Phys. Rev. X*, vol. 5, p. 011039, 2015.
- [20] D. Gao, W. Ding, M. Nieto-Vesperinas, et al., “Optical manipulation from the microscale to the nanoscale: fundamentals, advances and prospects,” *Light: Sci. Appl.*, vol. 6, p. e17039, 2017.
- [21] S. Sukhov and A. Dogariu, “Non-conservative optical forces,” *Rep. Prog. Phys.*, vol. 80, p. 112001, 2017.
- [22] X. Li, J. Chen, Z. F. Lin, and J. Ng, “Optical pulling at macroscopic distances,” *Sci. Adv.*, vol. 5, no. 3, p. eaau7814, 2019.
- [23] Y. K. Jiang, H. Z. Lin, X. Li, J. Chen, J. J. Du, and J. Ng, “Hidden symmetry and invariance in optical forces,” *ACS Photonics*, vol. 6, pp. 2749–2756, 2019.
- [24] H. J. Chen, H. X. Zheng, W. L. Lu, S. Y. Liu, J. Ng, and Z. F. Lin, “Lateral optical force due to the breaking of electric-magnetic symmetry,” *Phys. Rev. Lett.*, vol. 125, no. 7, p. 073901, 2020.
- [25] Y. Yang, Y. Ren, M. Chen, Y. Arita, and C. Rosales-Guzmán, “Optical trapping with structured light: a review,” *Adv. Photonics*, vol. 3, p. 034001, 2021.
- [26] X. Wu, R. Ehehalt, G. Razinskas, T. Feichtner, J. Qin, and B. Hecht, “Light-driven microdrones,” *Nat. Nanotechnol.*, vol. 17, no. 5, pp. 477–484, 2022.
- [27] F. Nan, X. Li, S. Zhang, J. Ng, and Z. Yan, “Creating stable trapping force and switchable optical torque with tunable phase of light,” *Sci. Adv.*, vol. 8, no. 11, p. eadd6664, 2022. <https://www.science.org/doi/10.1126/sciadv.add6664>.
- [28] I. A. Favre-Bulle, A. B. Stilgoe, E. K. Scott, and H. Rubinsztein-Dunlop, “Optical trapping in vivo: theory, practice, and applications,” *Nanophotonics*, vol. 8, no. 6, pp. 1023–1040, 2019.
- [29] Y. Zhou, X. H. Xu, Y. N. Zhang et al., “Observation of high-order imaginary Poynting momentum optomechanics in structured light,” *Proc. Natl. Acad. Sci.*, vol. 119, no. 44, 2022, Art. no. e2209721119.
- [30] M. M. Li, S. H. Yan, Y. N. Zhang, and B. L. Yao, “Generation of controllable chiral optical fields by vector beams,” *Nanoscale*, vol. 12, no. 28, pp. 15453–15459, 2020.
- [31] X. H. Xu and M. Nieto-Vesperinas, “Azimuthal imaginary Poynting momentum density,” *Phys. Rev. Lett.*, vol. 123, no. 23, p. 233902, 2019.
- [32] J. E. Curtis, B. A. Koss, and D. G. Grier, “Dynamic holographic optical tweezers,” *Opt. Commun.*, vol. 207, nos. 1–6, pp. 169–175, 2002.
- [33] A. Yevick, D. B. Ruffner, and D. G. Grier, “Tractor beams in the Rayleigh limit,” *Phys. Rev. A*, vol. 93, no. 4, p. 043807, 2016.
- [34] Y. K. Jiang, H. J. Chen, J. Chen, J. Ng, and Z. F. Lin, “Universal relationships between optical force/torque and orbital versus spin momentum/angular momentum of light,” 2015, Preprint. <https://doi.org/10.48550/arXiv.1511.08546>.

- [35] H. Rubinsztein-Dunlop, F. Andrew, M. V. Berry, et al., “Roadmap on structured light,” *J. Opt.*, vol. 19, no. 1, p. 013001, 2017.
- [36] A. Forbes, M. de Oliveira, and M. R. Dennis, “Structured light,” *Nat. Photonics*, vol. 15, no. 4, pp. 253–262, 2021.
- [37] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, New York, Cambridge University Press, 1995.
- [38] J. Peurifoy, Y. C. Shen, J. Li, et al., “Marin,” “Nanophotonic particle simulation and inverse design using artificial neural networks,” *Sci. Adv.*, vol. 4, no. 6, p. eaar4206, 2018.
- [39] S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, “Deep learning enabled inverse design in nanophotonics,” *Nanophotonics*, vol. 9, no. 5, pp. 1041–1057, 2020.
- [40] K. Yao, R. Unni, and Y. B. Zheng, “Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale,” *Nanophotonics*, vol. 8, no. 3, pp. 339–366, 2019.
- [41] G. Gouesbet and G. Gréhan, *Generalized Lorenz-Mie Theories*, 2nd ed. Berlin, Springer, 2017.
- [42] J. Ng, Z. F. Lin, C. T. Chan, and P. Sheng, “Photonic clusters formed by dielectric microspheres: numerical simulations,” *Phys. Rev. B*, vol. 72, no. 11, p. 085130, 2005. <https://link.aps.org/doi/10.1103/PhysRevB.72.085130>.
- [43] X. Li, Y. N. Liu, Z. F. Lin, J. Ng, and C. T. Chan, “Non-Hermitian physics for optical manipulation uncovers inherent instability of large clusters,” *Nat. Commun.*, vol. 12, no. 11, p. 6597, 2021.
- [44] Y. L. Xu, “Electromagnetic scattering by an aggregate of spheres,” *Appl. Opt.*, vol. 34, pp. 4573–4588, 1995.
- [45] G. Gouesbet and J. A. Lock, “On the description of electromagnetic arbitrary shaped beams: the relationship between beam shape coefficients and plane wave spectra,” *J. Quant. Spectrosc. Radiat. Transfer*, vol. 162, pp. 18–30, 2015.
- [46] D. Rumelhart, G. Hinton, and R. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, pp. 533–536, 1986.
- [47] M. Abadi, A. Agarwal, P. Barham, et al., “TensorFlow: large-scale machine learning on heterogeneous distributed systems,” 2016. Preprint.
- [48] P. Adam, G. Sam, F. Massa, et al., “PyTorch: an imperative style, high-performance deep learning library,” *Adv. Neural Inf. Process. Syst.*, vol. 32, pp. 8024–8035, 2019.
- [49] Y. K. Jiang, J. Chen, J. Ng, and Z. F. Lin, “Decomposition of optical force into conservative and nonconservative components,” 2016. Preprint.
- [50] X. N. Yu, Y. K. Jiang, H. J. Chen, S. Y. Liu, and Z. F. Lin, “Approach to fully decomposing an optical force into conservative and nonconservative components,” *Phys. Rev. A*, vol. 100, no. 13, p. 033821, 2019. <https://link.aps.org/doi/10.1103/PhysRevA.100.033821>.
- [51] H. X. Zheng, X. Li, Y. K. Jiang, J. Ng, Z. F. Lin, and H. J. Chen, “General formulations for computing the optical gradient and scattering forces on a spherical chiral particle immersed in generic monochromatic optical fields,” *Phys. Rev. A*, vol. 101, no. 5, p. 053830, 2020.
- [52] H. X. Zheng, X. Li, H. J. Chen, and Z. F. Lin, “Selective transport of chiral particles by optical pulling forces,” *Opt. Express*, vol. 29, no. 26, pp. 42684–42695, 2021.
- [53] D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” 2014. Preprint. <https://doi.org/10.48550/arXiv.1412.6980>.
- [54] S. Ruder, “An overview of gradient descent optimization algorithms,” 2016. Preprint. <https://doi.org/10.48550/arXiv.1609.04747>.
- [55] M. Polin, K. Ladavac, S. Lee, Y. Roichman, and D. G. Grier, “Optimized holographic optical traps,” *Opt. Express*, vol. 13, no. 15, pp. 5831–5845, 2005.
- [56] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, New York, John Wiley, 1998.
- [57] J. P. Barton, D. R. Alexander, and S. A. Schaub, “Theoretical determination of net radiation force and torque for a spherical particle illuminated by a focused laser beam,” *J. Appl. Phys.*, vol. 66, pp. 4594–4602, 1989.
- [58] Ø. Farsund and B. U. Felderhof, “Force, torque, and absorbed energy for a body of arbitrary shape and constitution in an electromagnetic radiation field,” *Physica A*, vol. 227, pp. 108–130, 1990.
- [59] H. X. Zheng, X. N. Yu, W. L. Lu, J. Ng, and Z. F. Lin, “GCforce: decomposition of optical force into gradient and scattering parts,” *Comput. Phys. Commun.*, vol. 237, pp. 188–198, 2019.
- [60] D. Vincent and V. Francesco, “A guide to convolution arithmetic for deep learning,” 2016. Preprint.
- [61] P. L. Marston, “Axial radiation force of a Bessel beam on a sphere and direction reversal of the force,” *J. Acoust. Soc. Am.*, vol. 120, no. 6, pp. 3518–3524, 2006.
- [62] S. Lee, Y. Roichman, and D. G. Grier, “Optical solenoid beams,” *Opt. Express*, vol. 18, no. 7, pp. 6988–6993, 2010.
- [63] N. Wang, J. Chen, S. Y. Liu, and Z. F. Lin, “Dynamical and phase-diagram study on stable optical pulling force in Bessel beams,” *Phys. Rev. A*, vol. 87, no. 6, p. 063812, 2013.
- [64] E. D. Palik, *Handbook of Optical Constants of Solids*, New York, Academic Press, 1998.
- [65] Y. S. Liang, S. H. Yan, Z. J. Wang, et al., “Simultaneous optical trapping and imaging in the axial plane: a review of current progress,” *Rep. Prog. Phys.*, vol. 83, no. 3, p. 032401, 2020.
- [66] T. Y. Li, X. H. Xu, B. Y. Fu, et al., “Integrating the optical tweezers and spanner onto an individual single-layer metasurface,” *Photonics Res.*, vol. 9, no. 6, pp. 1062–1068, 2021.
- [67] Y. H. Hu, J. J. Kingsley-Smith, M. Nikkhou, et al., “Structured transverse orbital angular momentum probed by a levitated optomechanical sensor,” 2022. Preprint. <https://doi.org/10.48550/arXiv.2209.09759>.
- [68] D. M. Hawkins, “The problem of overfitting,” *J. Chem. Inf. Comput. Sci.*, vol. 44, no. 1, pp. 1–12, 2022.
- [69] H. Wang, C. Qin, Y. Bai, Y. L. Zhang, and Y. Fu, “Recent advances on neural network pruning at initialization,” 2021. Preprint. <https://doi.org/10.48550/arXiv.2103.06460>.
- [70] R. Reed, “Pruning algorithms-a survey,” *IEEE Trans. Neural Netw.*, vol. 4, no. 5, pp. 740–747, 1993.

Supplementary Material: This article contains supplementary material (<https://doi.org/10.1515/nanoph-2023-0101>).