Computation of double freeform optical surfaces using a Monge–Ampère solver: Application to beam shaping

ELSEVIER

Computation of double freeform optical surfaces using a Monge–Ampère solver: Application to beam shaping

Check for updates

N.K. Yadav [a,*], J.H.M. ten Thije Boonkkamp [a], W.L. IJzerman [a,b]

[a] *CASA, Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513 5600 MB Eindhoven, The Netherlands*
[b] *Signify Research, High Tech Campus 7, 5656 AE Eindhoven, The Netherlands*

ARTICLE INFO

ABSTRACT

In this article, we present a formulation for the design of double freeform lens surfaces to control the intensity distribution of a laser beam with plane wavefronts. Double freeform surfaces are utilized to shape collimated beams. Two different layouts of the freeform lens optical system are introduced, i.e., a single lens with double freeform surfaces, and two separate lenses with two flat and two freeform surfaces. The freeform lens design problem can be formulated as a Monge–Ampère type differential equation with transport boundary condition, expressing conservation of energy combined with the law of refraction and the constraint imposed on the optical path length between source and target planes. Numerical solutions are computed using a generalized least-squares algorithm which is presented by Yadav et al. (2018). The algorithm is capable to compute two solutions of the Monge–Ampère boundary value problem, corresponding to either c-convex or c-concave freeform surfaces for both layouts. The freeform surfaces are validated for several numerical examples using a ray-tracer based on Quasi-Monte Carlo simulation.

1. Introduction

Laser beam shaping is the art of controlling the energy (intensity) distribution and/or the phase profile of the beam. The energy of a laser beam usually varies like a Gaussian function with the distance from the axis of symmetry. Many basic and industrial applications benefit from the use of laser beams with a specific shape, such as semiconductor lithography, laser printing, optical data processing, optical metrology, microvia drilling in the manufacturing of printed wiring boards, etc [1–4].

In illumination optics, freeform optical design problems are similar to laser beam shaping techniques. Freeform optics concerns the computation of optical surfaces that convert a given source distribution of light into a desired target distribution. Freeform optics is trending more and more in laser beam shaping due to its many degrees of freedom in design that can be used to obtain compact and dedicated beam shapes. In order to achieve an effective control on both the intensity distribution and phase profile of the laser beam, at least two freeform surfaces are needed [5,6].

In this article, we address the problem of collimated beam shaping by computing double freeform surfaces in order to achieve good control over the intensity distribution. We restrict ourselves to planar wavefronts. The problem of collimated beam shaping has two different layouts: one single lens with two freeform surfaces, and two separate lenses with two flat and two freeform surfaces. We will present an efficient numerical method which can be employed to deal with both layouts.

Applying the laws of geometrical optics, one can obtain the following mathematical relation for the two freeform surfaces with parallel ingoing and outgoing light rays, i.e.,

$$u_1(\mathbf{x}) + u_2(\mathbf{y}) = c(\mathbf{x}, \mathbf{y}), \tag{1}$$

where $u_1(\mathbf{x})$, $(\mathbf{x} \in \mathcal{S})$ and $u_2(\mathbf{y})$, $(\mathbf{y} \in \mathcal{T})$ define the location and shape of the freeform surfaces, with $\mathcal{S}$ and $\mathcal{T}$ the source and target domain, respectively. The function $c(\mathbf{x}, \mathbf{y})$ is referred to as the cost function in optimal mass transport, and is in our application a non-quadratic function in $|\mathbf{x} - \mathbf{y}|$, in short non-quadratic cost function. From this relation (1) we can determine the optical map $\mathbf{m} : \mathcal{S} \to \mathcal{T}$. Furthermore, combined with energy conservation, relation (1) gives rise to a non-standard Monge–Ampère type equation, subject to a transport boundary condition.

There are several numerical methods available which can be employed to compute freeform surfaces of optical systems governed by the standard Monge–Ampère equation corresponding to a quadratic cost function [7–10]. However, numerical methods for freeform design problems with a non-quadratic cost function are rare [5,11]. Recently, some articles have been published on the design of double freeform

---

surfaces with applications to beam shaping [12–16]. In [12,13], the authors presented Monge–Ampère solvers to design double freeform surfaces for the control of both the intensity distribution and phase profile of a laser beam. They discretized the standard Monge–Ampère equation using finite differences and solved the resulting nonlinear algebraic system using Newton iteration. As initial guess for Newton iteration, they use ray mapping method which is based on the quadratic cost function. Alternatively, ray mapping methods to compute double freeform surfaces are presented in [14–16]. These methods are two-stage procedures: first the ray mapping between source and target domains is calculated, and second, the freeform surfaces are computed iteratively from the mapping.

We have introduced an extended least-squares method for the Monge–Ampère equation with non-quadratic cost function describing a freeform lens in [5]. We employ this least-squares method for our laser beam shaping problem. The least-squares method is a two-stage procedure. First, we compute the optical map in a least squares sense from a constrained minimization problem for the energy balance, and second, we compute the shape of the freeform surfaces from the mapping and relation (1). The algorithm is very efficient and can handle very complicated target distributions, and is also capable to provide two solutions, one corresponding to c-convex and one to c-concave freeform surfaces. Our algorithm is very generic, can be applied to compute freeform surfaces of other optical systems governed by non-quadratic cost function. Target distributions can be arbitrary and complex such as image or far-field approximations [7,8]. Furthermore, the algorithm can be useful for other applications of optimal transport, such as the design of antennas or shape optimization [17,18].

We have organized our paper as follows. In Section 2 we give a mathematical formulation of the freeform beam shaping design problem. A short overview of the proposed numerical method is presented in Section 3. We apply the numerical method to two beam shaping test problems in Section 4, and concluding remarks are given in Section 5.

2. Problem formulation

The optical system for freefrom lens design has two different layouts: one single lens with two freeform surfaces, and two separate lenses with altogether two freefrom surfaces. The geometrical structure of both optical systems is shown schematically in Fig. 1. The source and target are situated in a medium of refractive index n_o (outer), and the refractive index between the freefrom surfaces is n_i (inner). If we choose $n_o = 1$ and $n_i > 1$ then Fig. 1 represents a single lens optical system with two freeform surfaces, and for the choice $n_i = 1$ and $n_o > 1$ it represents two separate lenses. Let $(x_1, x_2, z) \in \mathbb{R}^3$ denote the Cartesian coordinates with z the horizontal coordinate and $x = (x_1, x_2) \in \mathbb{R}^2$ the coordinates in the source plane $z = 0$, denoted by α_1, and let S be a bounded source domain in the plane α_1. The source S emits parallel light rays which propagate in the positive z-direction. The exitance of the source is given by $f(x)$ [W/m^2], $x \in S$. The target at a distance $\ell > 0$ from the plane α_1 is denoted by $\mathcal{T}$.

The ingoing light rays, starting in the medium of refractive index n_o, are intercepted by the first lens surface $\mathcal{L}_1$, defined as the graph of a function $z \equiv u_1(x)$, $x \in S$. The light rays are refracted off by the first freefrom surface $\mathcal{L}_1$, and refracted off again by the second freeform surface $\mathcal{L}_2$, and form the output beam, which is required to have parallel light rays in the positive z-direction. The second lens surface $\mathcal{L}_2$ is defined as the graph of a function $w \equiv \ell - z = u_2(y)$, $y \in \mathcal{T}$, where $y \equiv (y_1, y_2) \in \mathbb{R}^2$ are the Cartesian coordinates of the target plane α_2. After refraction at $\mathcal{L}_2$ the rays are again in the medium of refractive index n_o.

Our goal is to achieve the desired irradiance $g(y)$ [W/m^2] at the target plane $\alpha_2 : z = \ell$ corresponding to a given exitance $f(x)$, with a planar wavefront, i.e., after two refractions the refracted rays must form a parallel beam. We assume that both freefrom surfaces $\mathcal{L}_1$ and $\mathcal{L}_2$ are perfect lens surfaces and no energy is lost in the refraction.
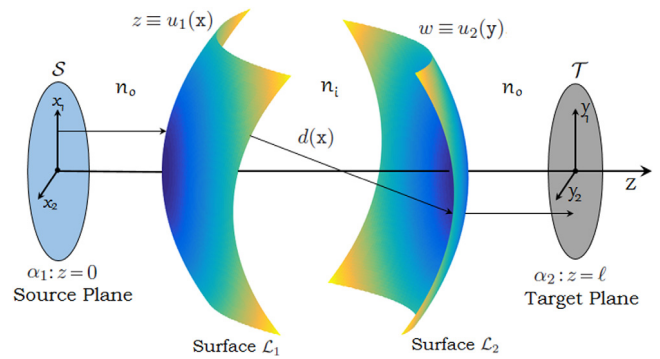

Fig. 1. Sketch of a freeform lens optical system.

2.1. Ray map and optical surfaces

In this section, we first give an expression for the ray-trace map, and second, we derive the mathematical relation (1) for the location of the freeform surfaces using the laws of geometrical optics.

The mapping m can be derived by tracing a typical ray through the optical system. Let us consider a ray emitted from a position $x \in S$ on the source and propagating in the positive z-direction, let $\hat{s} = \hat{e}_z$ be the unit direction of the incident ray. The ray strikes the first freeform surface $\mathcal{L}_1$, refracts off in direction $\hat{t}$, strikes the second freeform surface $\mathcal{L}_2$, and refracts off, again in the direction $\hat{s}$. The unit surface normal of the first freeform surface $\mathcal{L}_1$, directed towards the light source, is given by

$$\hat{n}_1 = \frac{(\nabla u_1, -1)}{\sqrt{|\nabla u_1|^2 + 1}}. \tag{2}$$

Throughout this article, we use the convention that a hat denotes a unit vector. According to Snell's law [19,20], the direction $\hat{t} = \hat{t}(x)$ of the refracted ray can be expressed as

$$\hat{t} = \eta \hat{s} + F(|\nabla u_1|; \eta)\hat{n}_1, \tag{3}$$

where $\eta = n_o/n_i$ is the ratio of the corresponding refractive indexes and

$$F(z; \eta) = \frac{1}{\sqrt{z^2 + 1}} \left[\eta - \sqrt{1 + (1 - \eta^2)z^2} \right]. \tag{4}$$

If we write $\hat{t} = (t_1, t_2, t_3)^T$ then the first two components of the vector $\hat{t}$, can be written as a function of the third component:

$$\binom{t_1}{t_2} = (\eta - t_3)\nabla u_1. \tag{5}$$

The image on the target screen of the point $x \in S$ is the point $y \in \mathcal{T}$ under the ray trace map m, i.e., $y = m(x)$. This map can be obtained by the projection of $\hat{t}$ on the plane α_1, i.e.,

$$m(x) = x + \binom{t_1}{t_2} d(x), \tag{6}$$

where $d(x)$ is the distance between the freeform surfaces $\mathcal{L}_1$ and $\mathcal{L}_2$ along the ray refracted in the direction $\hat{t}(x)$. The total optical path length $L(x)$ corresponding to the ray emitted at a point $x \in S$, is given by

$$L(x) = n_o u_1(x) + n_i d(x) + n_o u_2(y). \tag{7}$$

The theorem of Malus and Dupin (the principle of equal optical path) states that the total optical path length between any two orthogonal wavefronts is the same for all rays [21, p.130]. As we deal with two parallel beams, the wavefronts coincide with the planes α_1 and α_2. Therefore, the total optical path length will be independent of the position vector x, i.e., $L(x) = L$. The horizontal distance ℓ between the source and the target plane is given by

$$\ell = u_1(x) + (\hat{s} \cdot \hat{t})d(x) + u_2(y). \tag{8}$$

Combining Eqs. (3), (7) and (8), we obtain the following expression

$$d(\mathrm{x}) = \frac{\beta}{n_\mathrm{i} - n_\mathrm{o} t_3}, \tag{9}$$

where $\beta = L - n_\mathrm{o}\ell$ is a constant. Note that $\beta < 0$ is possible if $n_\mathrm{o} > 1$, i.e., for the second layout (two lenses), whereas for the first layout (single lens) $\beta > 0$. Substituting (5) and (9) in (6), we have

$$\boldsymbol{m}(\mathrm{x}) = \mathrm{x} + \frac{\beta}{n_\mathrm{i}} \frac{\eta - t_3}{1 - \eta t_3} \nabla u_1. \tag{10}$$

Substituting t_3 in the above equation from the law of refraction (3), the mapping $\boldsymbol{m}$ is given by the relation

$$\boldsymbol{m}(\mathrm{x}) = \mathrm{x} - \frac{\beta}{n_\mathrm{i}} \frac{\nabla u_1(\mathrm{x})}{\sqrt{1 + (1 - \eta^2)|\nabla u_1|^2}}. \tag{11}$$

Next, we derive relation (1) for the location of the freeform surfaces. An alternative expression for the distance d reads

$$d^2 = \left(\ell - u_1(\mathrm{x}) - u_2(\mathrm{y})\right)^2 + |\mathrm{x} - \mathrm{y}|^2. \tag{12}$$

Thus, from Eqs. (7) and (12), we obtain

$$n_\mathrm{i}^2\left(\ell - u_1(\mathrm{x}) - u_2(\mathrm{y})\right)^2 + n_\mathrm{i}^2|\mathrm{x} - \mathrm{y}|^2 = \left(L - n_\mathrm{o} u_1(\mathrm{x}) - n_\mathrm{o} u_2(\mathrm{y})\right)^2,$$

which can be rewritten as

$$\left[u_1(\mathrm{x}) + u_2(\mathrm{y}) - \frac{n_\mathrm{o} L - n_\mathrm{i}^2 \ell}{n_\mathrm{o}^2 - n_\mathrm{i}^2}\right]^2 = \frac{n_\mathrm{i}^2 \beta^2}{(n_\mathrm{o}^2 - n_\mathrm{i}^2)^2} + \frac{n_\mathrm{i}^2}{n_\mathrm{o}^2 - n_\mathrm{i}^2}|\mathrm{x} - \mathrm{y}|^2,$$

and after elementary algebraic derivations, we obtain

$$u_1(\mathrm{x}) + u_2(\mathrm{y}) = c(\mathrm{x}, \mathrm{y}),$$

$$c(\mathrm{x}, \mathrm{y}) = \ell + \frac{n_\mathrm{o}\beta}{n_\mathrm{o}^2 - n_\mathrm{i}^2} \pm \frac{n_\mathrm{i}}{n_\mathrm{o}^2 - n_\mathrm{i}^2}\sqrt{\beta^2 + (n_\mathrm{o}^2 - n_\mathrm{i}^2)|\mathrm{x} - \mathrm{y}|^2}. \tag{13}$$

The function $c(\mathrm{x}, \mathrm{y})$ is known as the cost function in optimal mass transport theory [6]. If we choose the plus sign in $c(\mathrm{x}, \mathrm{y})$, the formulation represents the single lens case [5,22], otherwise if we choose the minus sign, it represents the two-lens case; for details see Appendix.

So far, we have obtained a mathematical formulation representing the freeform surfaces of the optical system and an expression for the ray-trace map $\boldsymbol{m}$. Next, we formulate a second order partial differential equation of Monge–Ampère type for the freeform surfaces.

2.2. Energy conservation

In the light transfer process global energy conservation must be satisfied, i.e.,

$$\iint_S f(\mathrm{x})\mathrm{dx} = \iint_{\mathcal{T}} g(\mathrm{y})\mathrm{dy}, \tag{14}$$

which is a constraint on f and g. Moreover, the mapping $\mathrm{y} = \boldsymbol{m}(\mathrm{x}) : S \rightarrow \mathcal{T}$ should satisfy energy conservation for each $\mathcal{A} \subset S$, i.e., local energy conservation, represented by

$$\iint_{\mathcal{A}} f(\mathrm{x})\mathrm{dx} = \iint_{\boldsymbol{m}(\mathcal{A})} g(\mathrm{y})\mathrm{dy}. \tag{15}$$

After a change of variables, we obtain the differential form of local energy conservation

$$f(\mathrm{x}) = g(\boldsymbol{m}(\mathrm{x}))|\det(\mathrm{D}\boldsymbol{m}(\mathrm{x}))|, \tag{16}$$

where $\mathrm{D}\boldsymbol{m}$ is the Jacobi matrix of the mapping $\boldsymbol{m}$, which measures the expansion/contraction of a tube of rays due to the two refractions. The accompanying boundary condition is derived from the condition that all light from the source domain S must be transferred into the target domain $\mathcal{T}$, and is given by

$$\boldsymbol{m}(\partial S) = \partial\mathcal{T}, \tag{17}$$

stating that the boundary of the source S is mapped to the boundary of the target $\mathcal{T}$. This is a consequence of the edge ray principle [23].

Next, we derive a Monge–Ampère type equation for the freeform surfaces using the energy conservation constraint (16) and relation (13) for the location of the freefrom surfaces. We assume that both freeform surfaces are either c-convex or c-concave functions [18, p.54]. The surfaces $z = u_1(\mathrm{x})$ and $\ell - z = u_2(\mathrm{y})$ are c-convex if

$$u_1(\mathrm{x}) = \max_{\mathrm{y} \in \mathcal{T}}\{c(\mathrm{x}, \mathrm{y}) - u_2(\mathrm{y})\}, \tag{18a}$$

$$u_2(\mathrm{y}) = \max_{\mathrm{x} \in S}\{c(\mathrm{x}, \mathrm{y}) - u_1(\mathrm{x})\}, \tag{18b}$$

alternatively, these are c-concave if

$$u_1(\mathrm{x}) = \min_{\mathrm{y} \in \mathcal{T}}\{c(\mathrm{x}, \mathrm{y}) - u_2(\mathrm{y})\}, \tag{19a}$$

$$u_2(\mathrm{y}) = \min_{\mathrm{x} \in S}\{c(\mathrm{x}, \mathrm{y}) - u_1(\mathrm{x})\}. \tag{19b}$$

For a continuously differentiable function c, the map $\mathrm{y} = \boldsymbol{m}(\mathrm{x})$ is implicitly given by

$$\nabla_\mathrm{x} c(\mathrm{x}, \boldsymbol{m}(\mathrm{x})) - \nabla u_1(\mathrm{x}), \tag{20}$$

under the condition that the Jacobi matrix $\boldsymbol{C} = \mathrm{D}_{\mathrm{xy}}c$, defined by

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 c}{\partial x_1 \partial y_1} & \dfrac{\partial^2 c}{\partial x_1 \partial y_2} \\ \dfrac{\partial^2 c}{\partial x_2 \partial y_1} & \dfrac{\partial^2 c}{\partial x_2 \partial y_2} \end{pmatrix}, \tag{21}$$

is invertible. For our optical systems the map $\boldsymbol{m}$ given in (11) satisfies relation (20) indeed, with the cost function c defined in (13). The matrix $\boldsymbol{C}$ is symmetric negative definite, i.e., $\det(\boldsymbol{C}) > 0$ and $\mathrm{tr}(\boldsymbol{C}) < 0$, which is a consequence of the fact that the function c depends on $|\mathrm{x} - \mathrm{y}|$ [5]. Since the cost function c is continuously differentiable, from relation (20), we deduce that

$$\boldsymbol{C}\mathrm{D}\boldsymbol{m}(\mathrm{x}) = \mathrm{D}^2 u_1(\mathrm{x}) - \mathrm{D}_{\mathrm{xx}}c(\mathrm{x}, \boldsymbol{m}(\mathrm{x})) \equiv \boldsymbol{P}(\mathrm{x}), \tag{22}$$

where $\mathrm{D}^2 u_1$ is the Hessian of u_1. The matrix $\boldsymbol{P}(\mathrm{x}) = \mathrm{D}^2 u_1(\mathrm{x}) - \mathrm{D}_{\mathrm{xx}}c(\mathrm{x}, \boldsymbol{m}(\mathrm{x}))$ is negative semi-definite for a c-concave pair (u_1, u_2) and positive semi-definite for a c-convex pair (u_1, u_2). In the following, we discuss the c-convex case only, thus we require the matrix $\boldsymbol{P}$ to be positive semi-definite. Substituting $\mathrm{D}\boldsymbol{m}$ from (22) in the energy balance (16) and by choosing the positive sign for the determinant, we obtain

$$\frac{\det(\boldsymbol{P}(\mathrm{x}))}{\det(\boldsymbol{C}(\mathrm{x}, \boldsymbol{m}(\mathrm{x})))} = \frac{f(\mathrm{x})}{g(\boldsymbol{m}(\mathrm{x}))}. \tag{23}$$

Recall that the 2×2 matrix $\boldsymbol{P}$ is positive semi-definite if and only if

$$\mathrm{tr}(\boldsymbol{P}) \geq 0 \quad \text{and} \quad \det(\boldsymbol{P}) \geq 0. \tag{24}$$

Because $\det(\boldsymbol{C}) > 0$ and the right hand side functions $f \geq 0$ and $g > 0$, it is obvious that $\det(\boldsymbol{P}) \geq 0$. So, the only requirement to verify is $\mathrm{tr}(\boldsymbol{P}) \geq 0$. A similar condition can be obtained for the c-concave case.

Next, we give a brief description of the least-squares algorithm [5] to solve the Monge–Ampère equation (23) subject to the boundary condition (17) and constraint (24).

3. Numerical method

First, we calculate the mapping $\boldsymbol{m}$ using the extended least-squares method as follows: we enforce the equality $\boldsymbol{C}\mathrm{D}\boldsymbol{m} = \boldsymbol{P}$ by minimizing the functional

$$J_\mathrm{I}(\boldsymbol{m}, \boldsymbol{P}) = \frac{1}{2}\iint_S \|\boldsymbol{C}\mathrm{D}\boldsymbol{m} - \boldsymbol{P}\|_\mathrm{F}^2 \mathrm{dx}. \tag{25}$$

The norm used in this functional is the Frobenius norm. Next, we address the boundary condition (17) by minimizing the functional

$$J_\mathrm{B}(\boldsymbol{m}, \boldsymbol{b}) = \frac{1}{2}\oint_{\partial S} |\boldsymbol{m} - \boldsymbol{b}|^2 \mathrm{d}s, \tag{26}$$

where $|.|$ denotes the ℓ_2-norm for vectors, and $\boldsymbol{b}$ is a vector function mapping ∂S to $\partial \mathcal{T}$. We combine the functionals J_{I} for the interior and J_{B} for the boundary domain by taking a weighted average

$$J(\boldsymbol{m}, \boldsymbol{P}, \boldsymbol{b}) = \alpha J_{\mathrm{I}}(\boldsymbol{m}, \boldsymbol{P}) + (1 - \alpha) J_{\mathrm{B}}(\boldsymbol{m}, \boldsymbol{b}). \tag{27}$$

The parameter α $(0 < \alpha < 1)$ controls the weight of the first functional compared to the second functional. The minimization of (27) gives us the mapping $\boldsymbol{m}$ which is implicitly related to the surface function u_1 as shown in relation (20). We minimize the functionals (25)–(27) iteratively starting with an adequate initial guess $\boldsymbol{m}^0$, which will be specified shortly. We perform the iteration

$$\boldsymbol{b}^{n+1} = \operatorname{argmin} J_{\mathrm{B}}(\boldsymbol{m}^n, \boldsymbol{b}), \tag{28a}$$

$$\boldsymbol{P}^{n+1} = \operatorname{argmin} J_{\mathrm{I}}(\boldsymbol{m}^n, \boldsymbol{P}), \tag{28b}$$

$$\boldsymbol{m}^{n+1} = \operatorname{argmin} J(\boldsymbol{m}, \boldsymbol{P}^{n+1}, \boldsymbol{b}^{n+1}). \tag{28c}$$

Thus, given the nth iterand $\boldsymbol{b}^n$, $\boldsymbol{P}^n$ and $\boldsymbol{m}^n$ we successively compute the new iterand $\boldsymbol{b}^{n+1}$, $\boldsymbol{P}^{n+1}$ and $\boldsymbol{m}^{n+1}$ as explained below, see Fig. 2.

We initialize our minimization procedure by constructing an initial guess $\boldsymbol{m}^0$ which maps a bounding box of the source area S to a bounding box of the target area $\mathcal{T}$. Without loss of generality we assume the smallest bounding box of the source and target domain are rectangular and denote these by $[a_{\min}, a_{\max}] \times [b_{\min}, b_{\max}]$ and $[c_{\min}, c_{\max}] \times [d_{\min}, d_{\max}]$, respectively. Then the initial guess reads:

$$m_1^0 = \frac{x_1 - a_{\min}}{a_{\max} - a_{\min}} c_{\min} + \frac{a_{\max} - x_1}{a_{\max} - a_{\min}} c_{\max}, \tag{29a}$$

$$m_2^0 = \frac{x_2 - b_{\min}}{b_{\max} - b_{\min}} d_{\min} + \frac{b_{\max} - x_2}{b_{\max} - b_{\min}} d_{\max}. \tag{29b}$$

In this article, we give a brief description of the minimization of J_{I}, J, and the computation of u_1 and u_2; for a more mathematical description see [5].

Minimizing procedure for $\boldsymbol{P}$

We assume $\boldsymbol{m}$ fixed and minimize $J_{\mathrm{I}}(\boldsymbol{m}, \boldsymbol{P})$ over all matrices that satisfy (23) and (24). Since the integrand of $J_{\mathrm{I}}(\boldsymbol{m}, \boldsymbol{P})$ does not depend on derivatives of $\boldsymbol{P}$, the minimization procedure can be performed pointwise. Thus, we minimize $\|\boldsymbol{C}\boldsymbol{D} - \boldsymbol{P}\|_{\mathrm{F}}$ for each grid point $\mathrm{x} \in S$, where $\boldsymbol{D}$ is the central difference approximation of $\mathrm{D}\boldsymbol{m}$. This gives rise to the following constrained minimization problem

$$\text{minimize} \quad H_S(p_{11}, p_{22}, p_{12}) \equiv \frac{1}{2}\|\boldsymbol{Q}_S - \boldsymbol{P}\|_{\mathrm{F}}^2, \tag{30a}$$

$$\text{subject to} \quad \det(\boldsymbol{P}) = \frac{f}{g(\boldsymbol{m})} \det(\boldsymbol{C}), \tag{30b}$$

$$\operatorname{tr}(\boldsymbol{P}) \geq 0, \tag{30c}$$

where $\boldsymbol{Q} = \boldsymbol{C}\boldsymbol{D}$ and $\boldsymbol{Q}_S = (\boldsymbol{Q} + \boldsymbol{Q}^T)/2$. We minimize H_S instead of H because we are only interested in the minimizer, not in its value [5]. This problem can be solved using a Lagrangian multiplier λ, and the possible minimizer has to satisfy the following algebraic system

$$p_{11} + \lambda p_{22} = q_{11}, \tag{31a}$$

$$\lambda p_{11} + p_{22} = q_{22}, \tag{31b}$$

$$(1 - \lambda) p_{12} = \frac{1}{2}(q_{12} + q_{21}), \tag{31c}$$

$$p_{11} p_{22} - p_{12}^2 = \frac{f}{g} \det(\boldsymbol{C}). \tag{31d}$$

The system (31a)–(31c) is linear in p_{11}, p_{22} and p_{12}, and is regular if $\lambda \neq \pm 1$, for the detailed solution see [5].

Minimizing procedure for $\boldsymbol{m}$

Here, we assume $\boldsymbol{b}$ and $\boldsymbol{P}$ are fixed and minimize $J(\boldsymbol{m}, \boldsymbol{P}, \boldsymbol{b})$. The minimizer is obtained from

$$\delta J(\boldsymbol{m}, \boldsymbol{P}, \boldsymbol{b})[\boldsymbol{\eta}] = 0, \tag{32}$$
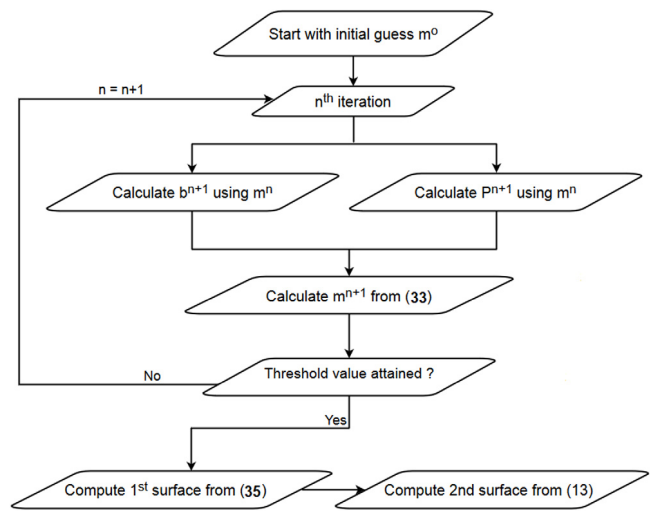

Fig. 2. Flow diagram of the proposed algorithm.

where δJ represents the first variation of J with respect to $\boldsymbol{m}$ in the direction $\boldsymbol{\eta}$. This gives rise to the following coupled elliptic system of PDEs with Robin boundary condition, i.e.,

$$\nabla \cdot (\boldsymbol{C}^T \boldsymbol{C} \mathrm{D}\boldsymbol{m}) = \nabla \cdot (\boldsymbol{C}^T \boldsymbol{P}), \quad \mathrm{x} \in S, \tag{33a}$$

$$(1 - \alpha)\boldsymbol{m} + \alpha(\boldsymbol{C}^T \boldsymbol{C} \nabla \boldsymbol{m}) \cdot \hat{\boldsymbol{n}} = (1 - \alpha)\boldsymbol{b} + \alpha \boldsymbol{C} \cdot \boldsymbol{P}\hat{\boldsymbol{n}}, \quad \mathrm{x} \in \partial S. \tag{33b}$$

We discretize the elliptic equations using the finite volume method, for the details see [5].

Calculation of freeform surfaces

We compute the first lens surface $z = u_1(\mathrm{x})$ from relation (20) in the least-squares sense, i.e.,

$$u_1(\mathrm{x}) = \operatorname{argmin} I(\phi), \quad I(\phi) = \frac{1}{2} \iint_S |\nabla \phi(\mathrm{x}) - \nabla_{\mathrm{x}} c(\mathrm{x}, \boldsymbol{m}(\mathrm{x}))|^2 \mathrm{dx}, \tag{34}$$

and this gives rise to the following Poisson problem with Neumann boundary condition:

$$\Delta u_1 = \nabla \cdot \nabla_{\mathrm{x}} c(\mathrm{x}, \boldsymbol{m}), \quad \mathrm{x} \in S, \tag{35a}$$

$$\nabla u_1 \cdot \hat{\boldsymbol{n}} = \nabla_{\mathrm{x}} c \cdot \hat{\boldsymbol{n}}, \quad \mathrm{x} \in \partial S. \tag{35b}$$

The Poisson equation is discretized using standard finite differences, and the discretized system is solved in Matlab using LU decomposition. The second freeform surface is calculated from relation (13), substituting the converged mapping $\boldsymbol{m}(\mathrm{x})$ and the first freeform surface $u_1(\mathrm{x})$. A summarizing flow diagram of the algorithm is presented in Fig. 2.

4. Test problems

As discussed earlier, the lens beam shaping problem has two different layouts: one single lens with two freeform surfaces, and two separate lenses also with two freeform surfaces. In this section, we provide numerical test results for both layouts. In addition, each layout has either c-convex or c-concave freeform surfaces. We apply our algorithm to achieve two different target distributions for the same Gaussian source distribution. The intensity distribution of a laser beam usually varies like a Gaussian function with the distance from the axis of symmetry, which makes it a challenging problem in optical design. As 99.7% of the total intensity is confined within 3σ of the mean, with σ the standard deviation [24, p.120], implying that about 55% of the source domain contains only 0.3% of the total flux to transport from the source to the target for our test problems.
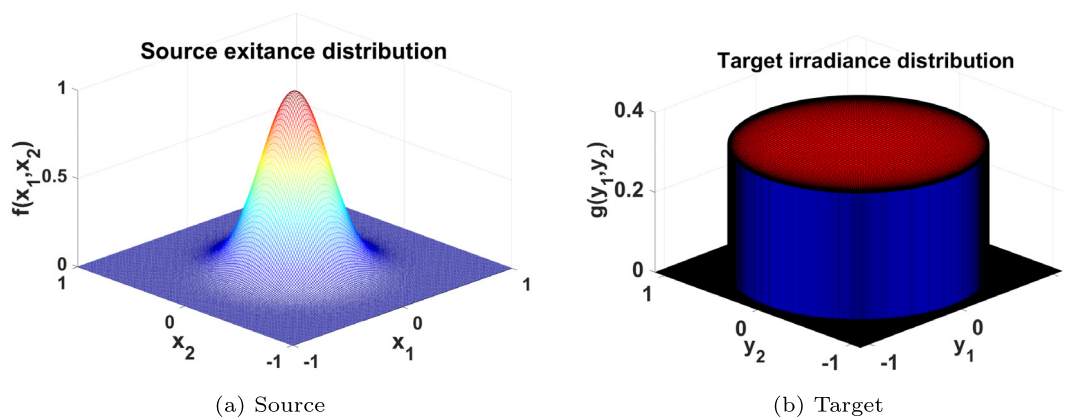
(a) Source

(b) Target

Fig. 3. Gaussian to circular top-hat: Source and target distributions.

(a) Single lens, c-convex case.

(b) Two lenses, c-convex case.

(c) Single lens, c-concave case.
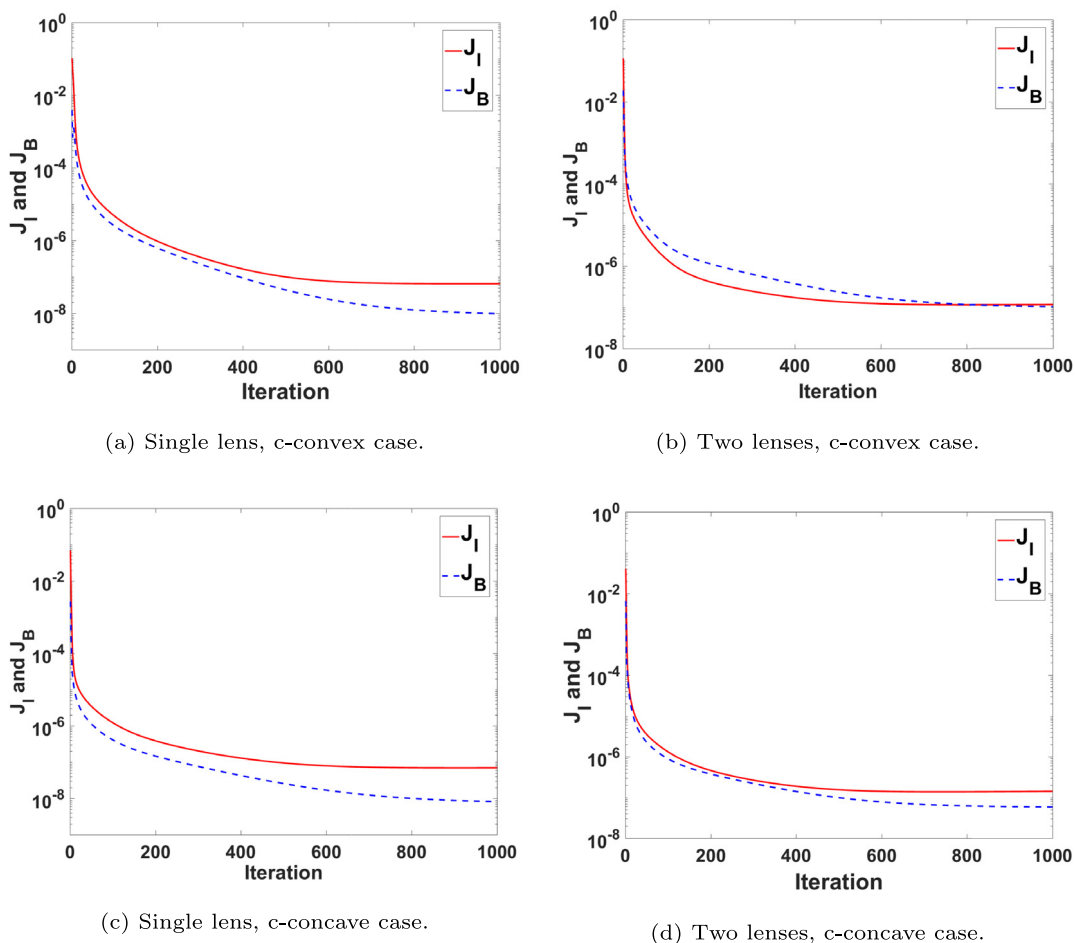

(d) Two lenses, c-concave case.

Fig. 4. Gaussian to circular top-hat: convergence history of the algorithm.

The first test problem concerns the transformation of a Gaussian input beam into a uniform circular top-hat target distribution, and in the second example, we show that our algorithm can also map the Gaussian distribution into a non-uniform distribution at the target. We will also present a comparison between both layouts. We compute all four possible optical systems for the first test problem, and for the second example, we execute our numerical algorithm to compute c-convex freeform surfaces for the single-lens system and c-concave freeform surfaces for the two-lens system.

4.1. From a Gaussian to a uniform circular top-hat distribution

In the first test problem, we compute freeform surfaces such that an incident Gaussian beam transform into a uniform top-hat circular beam as shown in Fig. 3. The source domain is given by $\mathcal{S} = [-1, 1] \times [-1, 1]$. The light source emits a parallel beam of light with Gaussian exitance, i.e.,

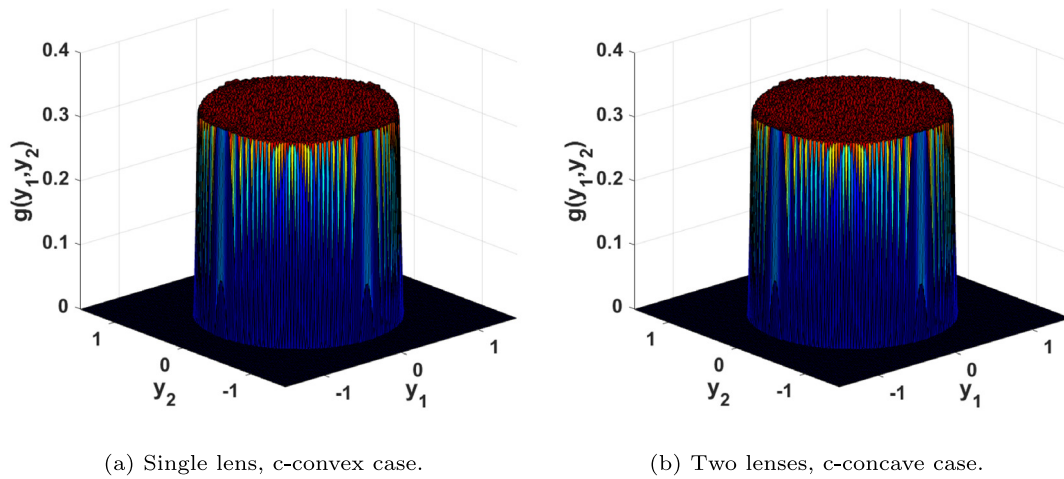$$f(\mathbf{x}) = A \exp\left(-\frac{1}{2}\left(\frac{|\mathbf{x}|}{\sigma}\right)^2\right). \tag{36}$$

(a) Single lens, c-convex case.

(b) Two lenses, c-concave case.

Fig. 5. Gaussian to circular top-hat: Achieved target irradiance distribution.

(a) Single lens, c-convex case.

(b) Two lenses, c-convex case.

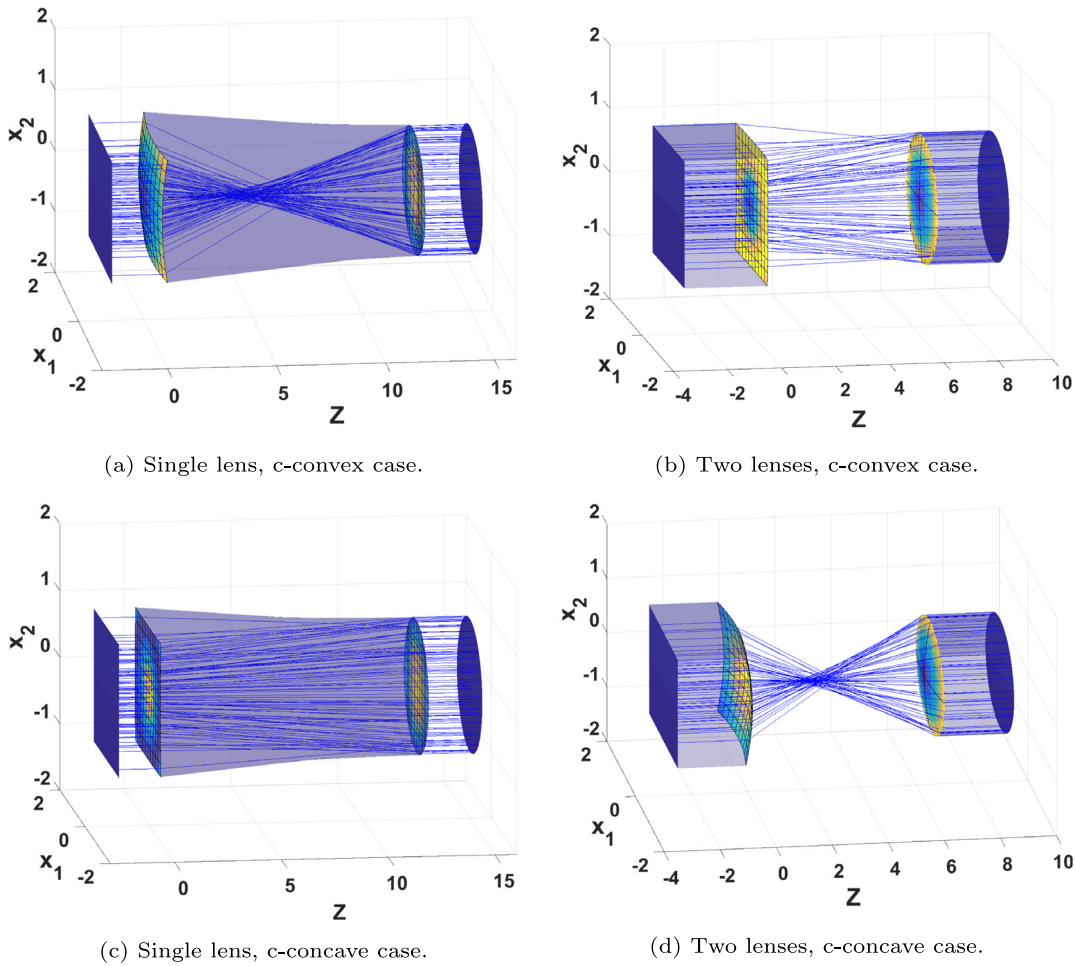(c) Single lens, c-concave case.

(d) Two lenses, c-concave case.

Fig. 6. Gaussian to circular top-hat: the freeform lens systems with 100 random rays.

In our numerical calculations, we use the parameter values $A = 3.1415 \times 10^{-1}$ and $\sigma^2 = 0.05$. The irradiance $g(y)$ of the target $\mathcal{T}$ is given by

$$g(y_1, y_2) = \begin{cases} \dfrac{1}{\pi} & \text{if } y_1^2 + y_2^2 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{37}$$

Note that we have chosen the parameter A in such a way that the source and target distributions satisfy the global energy balance (14). For the first layout: one single lens, we have the refractive indexes $n_i = 1.5, n_o = 1$ and select the parameter $\beta = 2\pi$, and for the second layout: two separate lenses, the refractive indexes are $n_i = 1, n_o = 1.5$ and we opt the parameter $\beta = -\pi$. For the both layouts, we fix the weight parameter $\alpha = 0.65$.
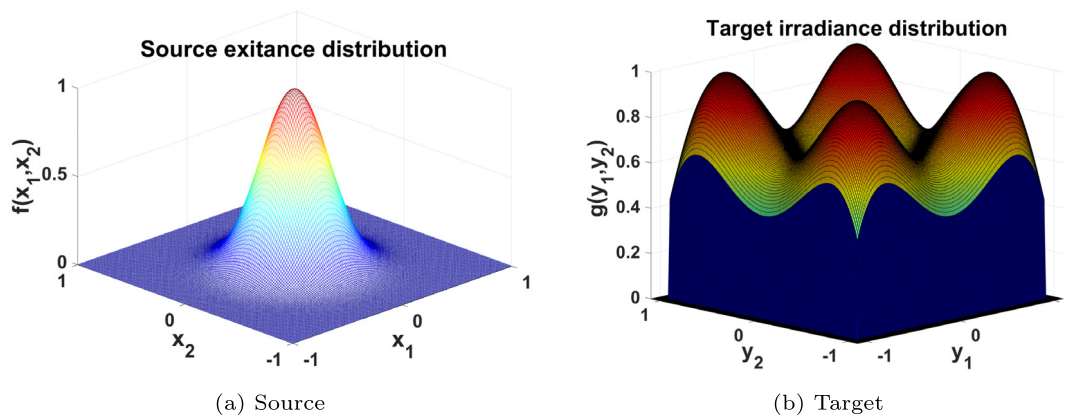
(a) Source

(b) Target

Fig. 7. Gaussian to non-uniform: Source and target distributions.

We discretize the source domain S uniformly in 200×200 grid points. To compare both layouts, we provide the convergence history of the algorithm for all four possible optical systems for both layouts, i.e., c-convex and c-concave freeform surfaces, in Fig. 4. The residuals J_I and J_B stall after around 1000 iterations at a value of approximately 10^{-7}, and the algorithm shown similar convergence for the both layouts, although for the second layout the algorithm converges a little slower than for the single lens problem. The algorithm is very memory and time efficient. All our numerical calculations were performed on a laptop with 4 GB of RAM and core i5 processor. The calculation time for 1000 iterations is about 380 s.

The freeform surfaces are validated using a ray tracing algorithm [8]. We run the algorithm for 10 million uniformly distributed quasi-random points on the source to verify the irradiance pattern at the target. The achieved target irradiance for 10 million rays is shown in Fig. 5 for both layouts, viz. c-convex surfaces of the single lens system and c-concave surfaces of the two-lens system. The distributions for the other two systems look almost the same. The figures show that the irradiance distribution is quite uniform. Indeed, $e = 2.3 \times 10^{-6}$, where the error e is defined as

$$e = \frac{1}{N} \sum_{i=1}^{N} (td_i - ta_i)^2, \qquad (38)$$

where N represents the number of grid cells, td the desired target irradiance and ta the achieved ray trace irradiance. A better result can be obtained by decreasing the grid size and increasing the number of rays traced through the system.

The freeform lens optical systems are shown in Fig. 6 with transparent side walls and 100 random rays including the source and target plane. We have computed all four possible optical systems for both layouts, viz. c-convex and c-concave freeform surfaces on their respective domains. For the single lens, c-convex case the freeform surfaces appear convex with respect to their domains but for the c-concave case the surfaces look very flat. For the two-lens system it is just the other way. The mapping for the c-convex case of the single lens system and for the c-concave case of the two-lens system turn around the irradiance pattern of the source on the target, i.e., rays cross in between two freeform surfaces and the envelop of the rays form a caustic surface.

4.2. From a Gaussian to a non-uniform distribution

In the second test problem, we show that our algorithm can also provide a non-uniform target distribution for a given source distribution. Here, we challenge our algorithm to convert the Gaussian exitance $f(x)$, as given by expression (36) into a non-uniform square shaped irradiance at the target $\mathcal{T}$ as shown in Fig. 7. The irradiance at the target is given
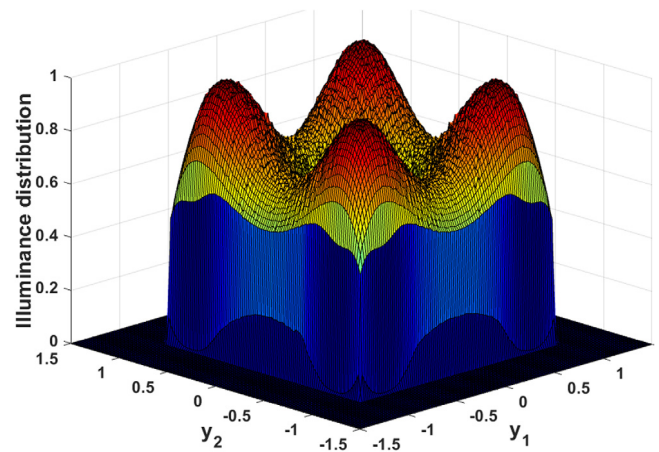

Fig. 8. Gaussian-to-non-uniform: Achieved target irradiance distribution.

by the function

$$g(y_1, y_2) = Bh(y_1)h(y_2),$$

$$h(y) = 1 - 5\left(\frac{5y}{7} - \frac{1}{2}\right)^2 \left(\frac{5y}{7} + \frac{1}{2}\right)^2. \qquad (39)$$

Here, we scale the target distribution by the parameter $B = 2.9061$ in order to satisfy the energy balance (14). The extended least-squares algorithm exhibits similar convergence as for the first test problem. The irradiance pattern at the target (peaks and lows of the polynomial) is validated by the ray tracing algorithm, and can be seen in Fig. 8, for 10 million uniformly distributed quasi-random points on the source domain.

We have presented the single lens system with c-convex freeform surfaces in Fig. 9(a) and the two-lens system with c-concave freeform surfaces in Fig. 9(b) with transparent side walls and 100 random rays including the target distribution at the target plane. In these figures the freeform surfaces look like a perfect convex or concave surface, however, the target distribution is non-uniform and we computed c-convex or c-concave surfaces only. To have a closer look at the lens surfaces, we have computed gradient surfaces for the single lens system, and plotted the modulus of the gradient in Fig. 10. Since, the source distribution is given by a Gaussian function, the first surface is convex as there is only one critical point as is evidenced by gradient plot Fig. 10(a). However, the target distribution is given by a polynomial with nine extrema, which can be seen from the gradient plot in Fig. 10(b), showing that the second freeform surface is a c-convex surface.
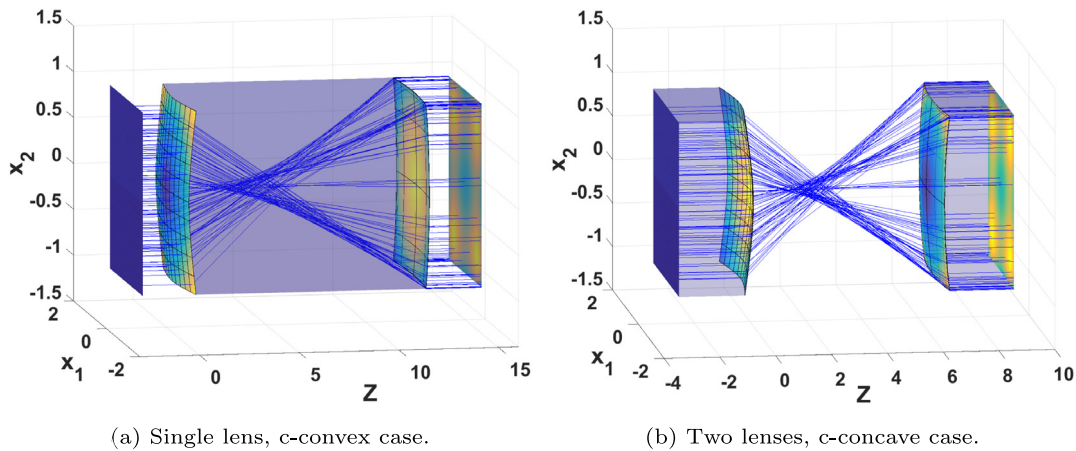
(a) Single lens, c-convex case.

(b) Two lenses, c-concave case.

Fig. 9. Gaussian to non-uniform distribution: the freeform lens systems.

(a) Gradient of the first surface.
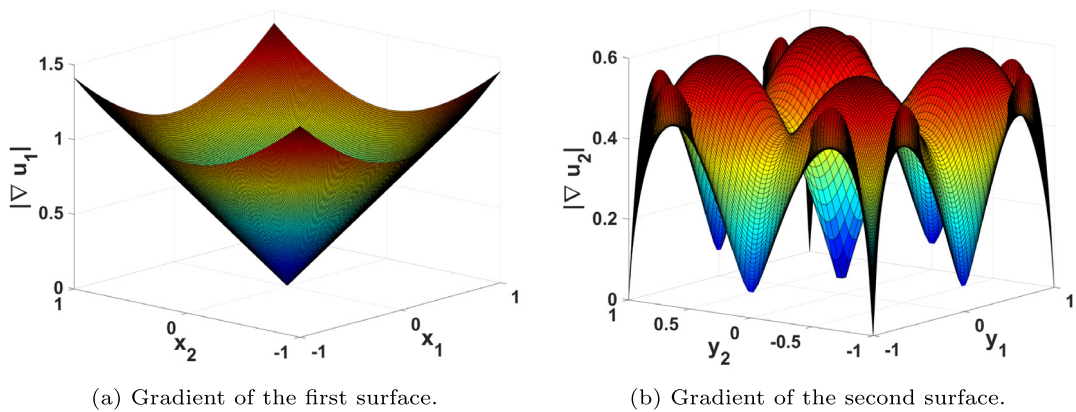
(b) Gradient of the second surface.

Fig. 10. Gaussian to non-uniform distribution, single lens system: modulus of gradients of the freeform surfaces.

5. Conclusions and discussion

We have presented a generic method for designing double freeform lens surfaces: "a single lens with two freefrom surfaces" and "two separate lenses with two freeform surfaces", in order to have a good control over the irradiance distribution of parallel light rays with planar wavefronts. Using properties of geometrical optics, i.e., the law of refraction, the principle of equal optical path length for the ingoing and outgoing wavefronts, and energy conservation, the double freeform surfaces design problems are formulated as a PDE of Monge–Ampère type with non-quadratic cost function coupled with the transport boundary condition. An extended least-squares method [5] is employed to compute the numerical solution. We have computed both c-convex and c-concave freeform surfaces for both layouts.

In various industrial problems there is a need to focus a laser beam to a well-defined shape and size with a uniform flat top irradiance distribution. From numerical simulations, it is evident that our algorithm is capable to convert a Gaussian distribution into a flat top-hat distribution at the target. Furthermore, beyond the laser applications, our algorithm can also achieve a non-uniform target distribution for a given source distribution. The design examples show clearly that the least-squares method presented in this article can be used in beam shaping.

The numerical algorithm can also be applied to the design of optical systems in illumination optics, e.g., for LED lighting. Moreover, the algorithm can be generalized to double freeform surfaces for wavefronts different from the planar case, such as, e.g., spherical wavefronts or cylindrical wavefronts, and for saddle point surfaces.

Appendix. The cost function

Here, we elaborate the properties of the cost function $c(\mathrm{x}, \mathrm{y})$ defined in relation (13), i.e.,

$$c(\mathrm{x}, \mathrm{y}) = \ell + \frac{n_\mathrm{o}\beta}{n_\mathrm{o}^2 - n_\mathrm{i}^2} \pm \frac{n_\mathrm{i}}{n_\mathrm{o}^2 - n_\mathrm{i}^2} \sqrt{\beta^2 + (n_\mathrm{o}^2 - n_\mathrm{i}^2)|\mathrm{x} - \mathrm{y}|^2}. \qquad (A.1)$$

The sign in front of the square root is unknown yet. To determine this we proceed as follows. Using Eq. (7) (with $L(\mathrm{x}) = L$) and (12), we can show that the term under the square root in (A.1) can be expressed as

$$\beta^2 + (n_\mathrm{o}^2 - n_\mathrm{i}^2)|\mathrm{x} - \mathrm{y}|^2 = \frac{1}{n_\mathrm{o}^2}\left(n_\mathrm{i}\beta + d(n_\mathrm{o}^2 - n_\mathrm{i}^2)\right)^2 \geq 0.$$

Substituting this expression in Eq. (A.1), we obtain

$$c(\mathrm{x}, \mathrm{y}) = \ell + \frac{n_\mathrm{o}\beta}{n_\mathrm{o}^2 - n_\mathrm{i}^2} \pm \frac{n_\mathrm{i}}{n_\mathrm{o}^2 - n_\mathrm{i}^2} \frac{1}{n_\mathrm{o}}\left|n_\mathrm{i}\beta + d(n_\mathrm{o}^2 - n_\mathrm{i}^2)\right|. \qquad (A.2)$$

Here, we need to check the sign of the expression $n_\mathrm{i}\beta + d(n_\mathrm{o}^2 - n_\mathrm{i}^2)$. Substituting d from Eq. (9), this expression becomes

$$n_\mathrm{i}\beta + d(n_\mathrm{o}^2 - n_\mathrm{i}^2) = n_\mathrm{o}\beta\frac{n_\mathrm{o} - n_\mathrm{i}t_3}{n_\mathrm{i} - n_\mathrm{o}t_3}. \qquad (A.3)$$

Substituting t_3 in the above relation from the law of refraction (3), we obtain

$$n_\mathrm{i}\beta + d(n_\mathrm{o}^2 - n_\mathrm{i}^2) = -\frac{n_\mathrm{i}n_\mathrm{o}\beta}{\sqrt{n_\mathrm{i}^2 + (n_\mathrm{i}^2 - n_\mathrm{o}^2)|\nabla u_1|^2}}. \qquad (A.4)$$

Let us first consider a single lens with two freeform surfaces, thus $n_\mathrm{o} = 1$ and $n_\mathrm{i} > 1$. The constant $\beta = L - \ell > 0$ is called "reduced optical path

length" [5]. Expression (A.4) shows that we have to choose the minus sign for the absolute value in expression (A.2), and consequently the plus sign in expression (A.1) (more details can be found in [5]), which results in

$$c(\mathrm{x},\mathrm{y}) = \ell + \frac{\beta}{1-n_{\mathrm{i}}^2} + \frac{n_{\mathrm{i}}}{1-n_{\mathrm{i}}^2}\sqrt{\beta^2 + (1-n_{\mathrm{i}}^2)|\mathrm{x}-\mathrm{y}|^2}. \tag{A.5}$$

Next, we check the sign of $n_{\mathrm{i}}\beta + d(n_{\mathrm{o}}^2 - n_{\mathrm{i}}^2)$ in expression (A.2) for the second layout, i.e., for the optical system with two different lenses. For this case, $n_{\mathrm{o}} > 1$ and $n_{\mathrm{i}} = 1$, and the constant $\beta = L - n_{\mathrm{o}}\ell < 0$. Thus, expression (A.4) shows that we have to choose the plus sign for the absolute value in expression (A.2), which results in

$$c(\mathrm{x},\mathrm{y}) = \ell + \frac{n_{\mathrm{o}}\beta}{n_{\mathrm{o}}^2 - 1} \pm \frac{\beta + (n_{\mathrm{o}}^2 - 1)d}{n_{\mathrm{o}}(n_{\mathrm{o}}^2 - 1)}. \tag{A.6}$$

Now, we substitute d from relation (7) in the above expression, which gives

$$c(\mathrm{x},\mathrm{y}) = \ell + \frac{n_{\mathrm{o}}\beta}{n_{\mathrm{o}}^2 - 1} \pm \left(\ell + \frac{n_{\mathrm{o}}\beta}{n_{\mathrm{o}}^2 - 1} - (u_1(\mathrm{x}) + u_2(\mathrm{y}))\right). \tag{A.7}$$

In the above equation, the right hand side equals the left hand side for the minus sign, therefore we have to choose minus sign in (A.1). Thus the mathematical expression for the two-lens case reads

$$c(\mathrm{x},\mathrm{y}) = \ell + \frac{n_{\mathrm{o}}\beta}{n_{\mathrm{o}}^2 - 1} - \frac{1}{n_{\mathrm{o}}^2 - 1}\sqrt{\beta^2 + (n_{\mathrm{o}}^2 - 1)|\mathrm{x}-\mathrm{y}|^2}. \tag{A.8}$$

References

[1] J.R. Leger, Laser Beam Shaping, in Micro-Optics: Elements, Systems and Applications, Taylor & Francis, London, 1997, pp. 223–257.

[2] F.M. Dickey, S.C. Holswade, D.L. Shealy, Laser Beam Shaping Applications, CRC Press, Taylor & Francis Group, Boca Raton, 2006.

[3] B. Adrien, B. Axel, W. Rolf, S. Jochen, L. Peter, High resolution irradiance tailoring using multiple freeform surfaces, Opt. Express 21 (9) (2013) 10563–10571.

[4] C. Dunsky, High-speed microvia formation with UV solid-state lasers, Proc. IEEE 90 (10) (2002) 1670–1680.

[5] N.K. Yadav, J.H.M. ten Thije Boonkkamp, W.L. I.J.zerman, A Monge-Ampère Problem with Non-Quadratic Cost Function: Application to Freeform Lens Design, J. Sci. Comput., in press.

[6] V. Oliker, Designing freeform lenses for intensity and phase control of coherent light with help from geometry and mass transport, Arch. Ration. Mech. Anal. 201 (3) (2011) 1013–1045.

[7] C.R. Prins, R. Beltman, J.H.M. ten Thije Boonkkamp, W.L. IJzerman, T.W. Tukker, A least-squares method for optimal transport using the Monge–Ampère equation, SIAM J. Sci. Comput. 37 (6) (2015) B937–B961.

[8] N.K. Yadav, J.H.M. ten Thije Boonkkamp, W.L. I.J.zerman, A least-squares method for the design of two-reflector optical system, submitted for publication.

[9] K. Brix, Y. Hafizogullari, A. Platen, Designing illumination lenses and mirrors by the numerical solution of Monge–Ampère equations, J. Opt. Soc. Amer. A 32 (10) (2015) 803–837.

[10] T. Glimm, V. Oliker, Optical design of two-reflector systems, the Monge–Kantorovich mass transfer problem and Fermat's principle, Indiana Univ. Math. J. 53 (5) (2004) 11070–11078.

[11] J. Rubinstein, G. Wolansky, Intensity control with a free-form lens, J. Opt. Soc. Amer. A 24 (2) (2007) 463–469.

[12] S. Chang, R. Wu, Z. Zheng, Design beam shapers with double freeform surfaces to form a desired wavefront with prescribed illumination pattern by solving a Monge–Ampère type equation, J. Opt. 18 (2016) 125602.

[13] Y. Zhang, R. Wu, Z. Zheng, H. Li, X. Liu, Double freeform surfaces design for laser beam shaping with Monge–Ampère equation method, Opt. Commun. 331 (2014) 297–305.

[14] Z. Feng, L. Huang, G. Jin, M. Gong, Beam shaping system design using double freeform optical surfaces, Opt. Express 21 (2013) 14728–14735.

[15] Z. Feng, L. Huang, G. Jin, M. Gong, Designing double freeform optical surfaces for controlling both irradiance and wavefront, Opt. Express 21 (2013) 28693–28701.

[16] C. Bösel, N.G. Worku, H. Gross, Ray-mapping approach in double freeform surface design for beam shaping beyond the paraxial approximation, Appl. Opt. 56 (13) (2017) 3679–3688.

[17] N.K. Yadav, Monge–Ampère Problems with Non-Quadratic Cost Function: Application to Freeform Optics (Ph.D. thesis), Eindhoven University of Technology, ISBN: 978-90-386-4574-2, 2018.

[18] C. Villani, Optimal Transport, Old and New, Springer, Berlin, 2009.

[19] C.R. Prins, Inverse Methods for Illumination Optics (Ph.D. thesis), Eindhoven University of Technology, ISBN: 978-90-386-3662-7, 2014.

[20] A.S. Glassner, An Introduction to Ray Tracing, Academic Press Ltd, London, 1991.

[21] M. Born, E. Wolf, Principles of Optics, fifth ed., Pergamon Press, Headington Hill Hall, 1975.

[22] N.K. Yadav, J.H.M. ten Thije Boonkkamp, W.L. I.J.zerman, A least-squares method for a Monge-Ampère equation with non-quadratic cost function applied to optical design, in: A.F. Radu, et al. (Eds.), Numerical Mathematics and Advanced Applications-ENUMATH 2017, in: Lecture Notes in Computational Science and Engineering, no. 126, Springer, Switzerland, 2018.

[23] H. Ries, A. Rabl, Edge-ray principle of nonimaging optics, J. Opt. Soc. Amer. A 11 (10) (1994) 2627–2632.

[24] D.C. Montgomery, G.C. Runger, Applied statistics and probability for engineers, sixth ed., John Wiley & Sons, Hoboken, 2014.