An inverse method for the design of TIR collimators to achieve a uniform color light beam

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computer Science

CASA-Report 12-09
April 2012

An inverse method for the design of TIR collimators
to achieve a uniform color light beam

by

C.R. Prins, J.H.M. ten Thije Boonkkamp, T.W. Tukker, W.L. IJzerman

An inverse method for the design of TIR collimators to achieve a uniform color light beam.

C.R. Prins[‡] [§] and J.H.M. ten Thije Boonkkamp[§]
Eindhoven University of Technology

T.W. Tukker and W.L. IJzerman
Philips Research, Philips Lighting

Abstract. Color over Angle (CoA) variation in the light output of white LEDs is a common and unsolved problem. In this article we introduce a new method to reduce CoA variation using a special collimator. The method is based on analytical inverse design methods. We present a numerical algorithm to solve the differential equations arising from this method and verify the results using Monte-Carlo raytracing.

1. Introduction

White LED technology is at the point of surpassing traditional light technologies such as compact fluorescent lamps in light output, light quality and efficiency. Since the 70's, the flux output per LED luminaire has increased by more than a factor 20 every decade, while the production cost per lumen has dropped by a factor 10 per decade [4]. LEDs are not sold separately, instead, they are built into a luminaire that consists of one or more LEDs, an optical system, electronics, heat sinks and a nice housing. The optical system is often a collimator which focuses the light in a specific direction. In this article we consider single LEDs in combination with a collimator.

Unfortunately, it is difficult to create an LED that emits light with a uniform white color. Such LEDs are created by coating a blue LED with a layer of yellow phosphor and possibly an additional layer of red phosphor. This phosphor coating converts part of the blue light into yellow or red light, resulting in white light. The fraction of the light that is converted depends on the distance that a light ray travels through the layers of phosphor, which is related to the emission angle of the light. Therefore the color of the emitted light is angle-dependent: the light emitted normal to the surface is more bluish white, while the light

---

[‡] Department of Applied Physics
[§] Department of Mathematics and Computer Science

emitted nearly parallel to the surface is more yellowish [10, p. 353-357]. This is called Color-over-Angle (CoA) variation.

A lot of research has been done to reduce this CoA variation by modifying the LED. It is often reduced by introducing bubbles in the phosphor layer [14] or by applying a dichroic coating [9], but these methods also reduce the efficiency and increase the production costs of the LED. If the LED is used in combination with a collimating optic, CoA variation can be reduced by using microstructures on top of the collimator. This is a widely adopted technique. However, microstructures introduce extra costs in the production process of the collimator and make the collimator look unattractive and broaden the light beam.

This article introduces an alternative method to reduce CoA variation, using freeform collimators. Reducing the CoA variation with freeform collimators has two advantages. First, the special optic does not introduce additional light loss, while modifications of the LED usually do. Second, expensive modifications to the LED are no longer necessary, resulting in cost reduction for the total optical system. Most color mixing methods are based on reducing color variation by mixing light from many different angles of the light source. Wang et al. [11] study the reduction of CoA variation using domes which mix light from two different angles. They note that it is theoretically possible to completely remove the CoA variation, but they do not show a proof. In this paper we will introduce a constructive method based on analytical inverse design methods [2, 6, 7] which shows that a complete removal of CoA variation is indeed possible.

Section 2 describes how to design a TIR (Total Internal Reflection) collimator with a specified output intensity distribution using inverse methods, and Section 3 explains how to include color mixing in this method. Subsequently, Section 4 discusses the numerical calculation and verification. Finally, concluding remarks are presented in Section 5.

2. Design of a TIR collimator using inverse methods

Before discussing color mixing, we will explain how to design a TIR collimator using inverse methods. A TIR collimator is a rotationally symmetric collimator, usually made of a transparent plastic like polycarbonate (PC) or polymethyl methacrylate (PMMA). A profile of a TIR collimator can be seen in Figure 1.

The design procedure consists of two steps: first we need to find relations between the angles t of rays leaving the source and the angles θ of rays leaving the collimator, the so-called transfer functions. Sub-
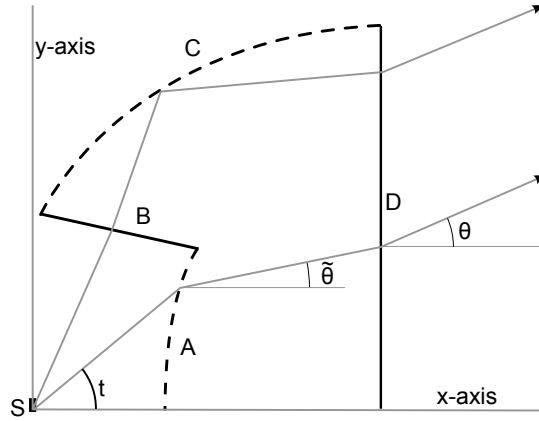
Figure 1. Profile of a TIR collimator. A full TIR collimator can be obtained by rotating the profile around the x-axis. The point light source S is located at the origin. The collimator consists of four distinct surfaces: A, B, C and D. The solid black lines indicate flat fixed surfaces, the dashed lines are the free surfaces that are to be calculated. All rays in the collimator are exactly once refracted or reflected by a free surface.

sequently we use these transfer functions to calculate the free surfaces A and C of the TIR collimator.

2.1. TRANSFER FUNCTIONS

A transfer function $\eta : [\theta_{\min}, \theta_{\max}] \rightarrow [t_{\min}, t_{\max}]$ describes a relation between the angle $t \in [t_{\min}, t_{\max}]$ of the light emitted from the source and the angle of the light $\theta \in [\theta_{\min}, \theta_{\max}]$ emitted from the TIR collimator. The direction of a ray of light is denoted by its counterclockwise angle with respect to the x-axis. The interval $[t_{\min}, t_{\max}]$ can be a subset of the emission angles of the source. Note that we could have chosen the transfer functions as mappings $[t_{\min}, t_{\max}]$ on $[\theta_{\min}, \theta_{\max}]$ as well. The latter choice is more common in literature, however, as we will see later, our choice is more appropriate for the purpose of color mixing. We choose the transfer functions to be strictly monotonic and thus invertible.

We find a transfer function using conservation of luminous flux. Let $I(t, u)$ [lm/sr] be the rotationally symmetric intensity distribution of the light source. Here $t \in [0, \pi/2]$ is the angle with respect to the symmetry axis (inclination), and $u \in [0, 2\pi)$ is the angle that rotates around the symmetry axis (azimuth). Because of the symmetry of the system, the intensity $I(t, u)$ is independent of u and denoted by $I(t)$. We introduce an effective intensity $\mathcal{I}(t)$, which is the flux per rad through the circular strip $[t, t + dt]$ on the unit sphere divided by 2π,

by integrating $I(t)$ over the angle u:

$$\mathcal{I}(t) := \frac{1}{2\pi} \int_0^{2\pi} I(t) \sin(t) \, \mathrm{d}u = I(t) \sin(t). \tag{1}$$

The effective intensity has units [lm/rad]. Similarly, there is an intensity distribution of the light exiting the collimator, denoted by $G(\theta, \phi)$, where $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$ are inclination and azimuth. We require this target distribution to be rotationally symmetric as well, and we denote it by $G(\theta)$. By integration over the angle ϕ, we find the effective target intensity distribution $\mathcal{G}(\theta) = G(\theta) \sin(\theta)$. The transfer function should be such that the intensity of the light exiting the collimator has the required intensity $\mathcal{G}(\theta)$. A more elaborate description of effective intensity distributions can be found in Maes [6].

The luminous flux (in lm) that is emitted from the collimator between θ and $\theta + \mathrm{d}\theta$ must be equal to the luminous flux emitted from the source between $\eta(\theta)$ and $\eta(\theta + \mathrm{d}\theta)$. This leads to the following relation:

$$\mathcal{G}(\theta) \, \mathrm{d}\theta = \sigma \, \mathcal{I}(\eta(\theta)) \, \mathrm{d}\eta(\theta), \tag{2}$$

where $\sigma = -1$ for monotonically decreasing transfer functions and $\sigma = 1$ for monotonically increasing transfer functions. The transfer function can be calculated by transforming (2) into the differential equation:

$$\eta'(\theta) = \frac{\mathcal{G}(\theta)}{\sigma \, \mathcal{I}(\eta(\theta))}, \tag{3}$$

which can be integrated with initial value $\eta(\theta_{\min}) = t_{\min}$ (for increasing η) or $\eta(\theta_{\min}) = t_{\max}$ (for decreasing η).

We have to assure conservation of luminous flux for the entire optical system. Integration of (2) from $\theta_{\min}$ to $\theta_{\max}$ yields the following relation:

$$\int_{\theta_{\min}}^{\theta_{\max}} \mathcal{G}(\theta) \, \mathrm{d}\theta = \int_{t_{\min}}^{t_{\max}} \mathcal{I}(t) \, \mathrm{d}t. \tag{4}$$

The requirement (4) on the luminous flux in the system poses a restriction on the function $\mathcal{G}(\theta)$. If it is not satisfied, the system has no physical meaning.

2.2. Design of a TIR collimator for known transfer functions

When we have two transfer functions with ranges $[0, t_{\mathrm{av}}]$ and $[t_{\mathrm{av}}, \pi/2]$ for some $t_{\mathrm{av}} \in [0, \pi/2]$, we can design a TIR collimator. As shown in Figure 1, a TIR collimator consists of the surfaces A up to D. Light propagates through the collimator by two possible routes. In the

first route, light is refracted by surface B, reflected by total internal reflection by surface C and finally refracted by surface D. Light from the LED emitted at angles in the range $[t_{av}, \pi/2]$ will follow this route. In the second route, light is refracted by surface A and subsequently refracted once more by surface D. Light emitted in the range $[0, t_{av}]$ will follow this route.

Given a transfer function η, we can calculate the relation between the angle t and the angle $\tilde{\theta}$ of a ray before refraction by surface D. By Snell's law we have the relation $n \sin(\tilde{\theta}) = \sin(\theta)$, where n is the refractive index of the material of the collimator. This gives the relation

$$\tilde{\theta}(t) = \arcsin(\sin(\eta^{-1}(t))/n). \tag{5}$$
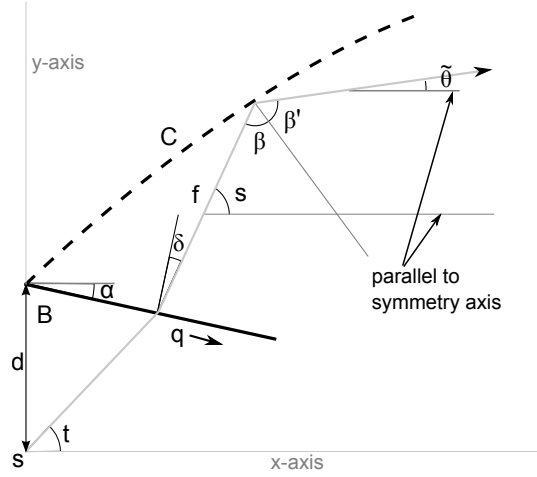
Figure 2. Calculation of the free surface C. This figure is a magnification of a part of Figure 1. The path of the light is denoted by the grey line.

First, consider the light through the route S-B-C-D. This light is refracted at surface B, reflected by C and finally refracted by D. Let d be the distance from the left end of surface B to the origin and α the clockwise angle of this surface with respect to the x-axis (see Figure 2). Let $(x_B(t), y_B(t))$ be the coordinates of the profile of surface B. We choose the parametrization such that a ray leaving the source at angle t will hit surface B at $(x_B(t), y_B(t))$ and leaves this surface at angle $s(t)$

with respect to the symmetry axis. We derive using basic geometry:

$$x_{\mathrm{B}}(t) = \frac{d}{\tan(t) + \tan(\alpha)}, \tag{6a}$$

$$y_{\mathrm{B}}(t) = \frac{d\tan(t)}{\tan(t) + \tan(\alpha)}, \tag{6b}$$

$$s(t) = \arccos\left(\frac{\cos(\alpha + t)}{n}\right) - \alpha, \tag{6c}$$

where (6c) was derived using Snell's law of refraction. Now we can calculate surface C using the 'generalized functional method' developed by Bortz and Shatz [2, 3]. They have derived the following differential equation, of which the variables are illustrated in Figure 2:

$$\frac{\mathrm{d}f}{\mathrm{d}q} = \frac{\mathrm{d}s}{\mathrm{d}q}\tan(\beta)\,f + \tan(\beta)\cos(\delta) - \sin(\delta). \tag{7}$$

Here f is the distance a light ray travels from surface B to surface C, q is the arc length along surface B, defined to be 0 at $t = \pi/2$, δ is the ray-emission angle measured counterclockwise with respect to the normal at the point of emission, and s is the angle of the ray leaving surface B with respect to the symmetry axis. The angle β depends on whether the optical surface is reflective or refractive. For a reflective surface we have [2]

$$\beta = \frac{1}{2}(\pi - \tilde{\theta} + s). \tag{8}$$

Usually we know q and s as functions of the emission angle t from the original light source. Multiplying (7) with $\mathrm{d}q/\mathrm{d}t$ gives:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\mathrm{d}s}{\mathrm{d}t}\tan(\beta)\,f + \frac{\mathrm{d}q}{\mathrm{d}t}\big(\tan(\beta)\cos(\delta) - \sin(\delta)\big). \tag{9}$$

From (6a)-(6c) we derive

$$q(t) = \frac{d}{\cos(\alpha)\tan(t) + \sin(\alpha)}, \tag{10a}$$

$$\delta(t) = s(t) - \frac{\pi}{2} - \alpha. \tag{10b}$$

Now we can calculate all the functions we need to solve equation (9). Substitute the relations (5), (6c), (8), (10a) and (10b) in equation (9) and calculate the function f. The position of the reflective surface is described by

$$x_{\mathrm{C}}(t) = x_{\mathrm{B}}(t) + f(t)\,\cos(s(t)), \tag{11a}$$

$$y_{\mathrm{C}}(t) = y_{\mathrm{B}}(t) + f(t)\,\sin(s(t)). \tag{11b}$$

Now consider the light through the route S-A-D. Similarly, we can calculate the position of surface A. Because the light incident in A comes directly from the point light source S, we can take $q(t) = 0$ and $s(t) = t$. This greatly simplifies (9). The refractive setting has the following expression for β [2]:

$$\tan(\beta) = \frac{\sin(\tilde{\theta} - t)}{1/n - \cos(\tilde{\theta} - t)}. \tag{12}$$

If $1/n - \cos(\tilde{\theta} - t) \leq 0$, then refraction from the angle t to the angle $\tilde{\theta}$ is physically not possible. Substituting $q(t)$, $s(t)$ and β in (9) yields the equation:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \left(\frac{\sin(\tilde{\theta}(t) - t)}{1/n - \cos(\tilde{\theta}(t) - t)} \right) f, \tag{13}$$

and we have the following description of surface A:

$$x_{\mathrm{A}}(t) = f(t) \cos(t), \tag{14a}$$
$$y_{\mathrm{A}}(t) = f(t) \sin(t). \tag{14b}$$

3. A color weighted TIR collimator

In Section 2 we have seen how to calculate the free surfaces of a TIR collimator for given source and target intensities. In this section we show how to incorporate color uniformity in this model. First we introduce some theory of the human perception of color, and derive a system of two coupled ordinary differential equations that is analogous to equation (3). Subsequently, we discuss the removable singularities arising in this system.

3.1. Transfer functions for color mixing

A lot of research has been done on perception of color by humans [8, 15]. It has been found that a beam of light can be fully described by its luminous flux (in lm) and two dimensionless chromaticity coordinates x and y with values between 0 and 1. Chromaticity coordinates can be used in color mixing calculations. Suppose we have two beams of light, numbered 1 and 2, with luminous fluxes L_1 and L_2 and color coordinates (x_1, y_1) and (x_2, y_2), respectively. In Malacara [8, p.57-58,103-105] we can find that the chromaticity coordinates $(x_{\mathrm{T}}, y_{\mathrm{T}})$ of
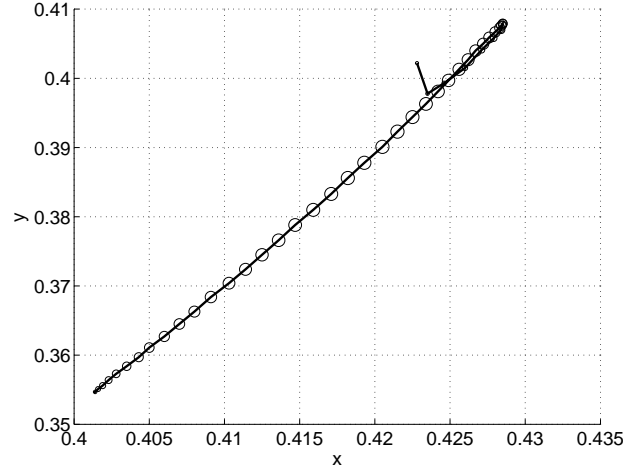
Figure 3. Scatter plot of the measured x and y chromaticity coordinates of the LED used in the numerical experiments in this article. The values are approximately on a straight line. The size of the circles corresponds to the effective intensity. The measured data in the lower left corner correspond to light emitted perpendicular to the surface of the LED, the upper right to the light emitted parallel to the surface of the LED.

the two colored light beams after mixing become:

$$x_{\mathrm{T}} = \frac{x_1 \, L_1/y_1 + x_2 \, L_2/y_2}{L_1/y_1 + L_2/y_2}, \tag{15a}$$

$$y_{\mathrm{T}} = \frac{L_1 + L_2}{L_1/y_1 + L_2/y_2}. \tag{15b}$$

Note that the resulting chromaticity coordinates $(x_{\mathrm{T}}, y_{\mathrm{T}})$ are on the straight line segment between (x_1, y_1) and (x_2, y_2). For infinitesimal angles $\mathrm{d}t_1$ and $\mathrm{d}t_2$, the luminous flux between t_1 and $t_1 + \mathrm{d}t_1$ and between t_2 and $t_2 + \mathrm{d}t_2$ is $\mathcal{I}(t_1) \, \mathrm{d}t_1$ and $\mathcal{I}(t_2) \, \mathrm{d}t_2$ respectively. The rule (15) translates to

$$x_{\mathrm{T}} = \frac{x(t_1) \, \mathcal{I}(t_1) \, \mathrm{d}t_1/y(t_1) + x(t_2) \, \mathcal{I}(t_2) \, \mathrm{d}t_2/y(t_2)}{\mathcal{I}(t_1) \, \mathrm{d}t_1/y(t_1) + \mathcal{I}(t_2) \, \mathrm{d}t_2/y(t_2)}, \tag{16a}$$

$$y_{\mathrm{T}} = \frac{\mathcal{I}(t_1) \, \mathrm{d}t_1 + \mathcal{I}(t_2) \, \mathrm{d}t_2}{\mathcal{I}(t_1) \, \mathrm{d}t_1/y(t_1) + \mathcal{I}(t_2) \, \mathrm{d}t_2/y(t_2)}. \tag{16b}$$

From measured data we have seen an *approximately linear relation* between $x(t)$ and $y(t)$, meaning that all chromaticity coordinates $(x(t), y(t))$ are on a straight line (see Figure 3). Therefore we assume a linear relation in this article. Also we have observed that $x(t)$ and $y(t)$ are, apart from a region at large angles that contains little luminous

flux (see Figure 3 and Figure 4b), increasing functions of t. For the sake of simplicity, we will assume that $x(t)$ and $y(t)$ are strictly increasing.

LED light sources are usually Lambertian or close to Lambertian, i.e., the intensity distribution $I(t, u)$ of the emitted light is proportional to $\cos(t)$, or has an intensity distribution close to this. Multiplying with a factor $\sin(t)$ for the rotationally symmetric setting, the LED will have in general an effective intensity distribution close to $\mathcal{I}(t) = \frac{1}{2}I_0 \sin(2t)$. The exact shape of $\mathcal{I}$ may vary, but we can always assume that $\mathcal{I}(0) = \mathcal{I}(\pi/2) = 0$, the right derivative $\mathcal{I}'_+(0) > 0$ and the left derivative $\mathcal{I}'_-(\pi/2) < 0$. Also we assume that $\mathcal{I}(t) > 0$ for all $0 < t < \pi/2$. Similarly, we have $\mathcal{G}(0) = 0$ and the right derivative $\mathcal{G}'_+(0) > 0$. Furthermore we will assume that $\mathcal{G}(\theta) > 0$ for all $0 < \theta \leq \theta_{\max}$.

To construct a collimator that mixes the light of the source in such a way that the color point at the target is constant and the required intensity at the target is $\mathcal{G}(\theta)$, we devide the interval $[\theta_{\min}, \theta_{\max}]$ in N different segments. Define an ordered list $0 = \tau_0 < \tau_1 < \ldots < \tau_N = \pi/2$. Each segment i is a subinterval $[\tau_{i-1}, \tau_i] \subset [0, \pi/2]\, (i = 1, 2, \ldots, N)$. The relation between the angles of emission from the light source and the angle of emission from the collimator is defined by the transfer functions $\eta_i : [0, \theta_{\max}] \to [\tau_{i-1}, \tau_i] \subset [0, \pi/2]$. For every angle θ at the far field target, there will be light from exactly one angle in each segment. The chromaticity of the light at this angle at the target will be a weighted average of the chromaticities of the light from the different segments. Every transfer function is a monotonic, invertible map by definition. For ease of notation, we will introduce the following convention: $\mathcal{I}_i(\theta) = \mathcal{I}(\eta_i(\theta))$ is the intensity of the light at the source (in segment i) that is directed to the angle θ. Similarly we write $x_i(\theta) = x(\eta_i(\theta))$ and $y_i(\theta) = y(\eta_i(\theta))$ for $i = 1, 2, \ldots, N$.

The light directed to an angle θ will come from N different angles $t_i = \eta_i(\theta)$ at the source. Similar to (2), we have conservation of luminous flux

$$\mathcal{G}(\theta)\, \mathrm{d}\theta = \sum_{i=1}^{N} \sigma_i \mathcal{I}(t_i)\, \mathrm{d}\eta_i(\theta), \tag{17}$$

and similar to (16), we have the target chromaticity denoted by $(x_{\mathrm{T}}, y_{\mathrm{T}})$:

$$x_{\mathrm{T}} = \frac{\sum_{i=1}^{N} \sigma_i \mathcal{I}_i(\theta)\, x_i(\theta)\, \eta_i'(\theta)\, \mathrm{d}\theta / y_i(\theta)}{\sum_{i=1}^{N} \sigma_i \mathcal{I}_i(\theta)\, \eta_i'(\theta)\, \mathrm{d}\theta / y_i(\theta)}, \tag{18a}$$

$$y_{\mathrm{T}} = \frac{\sum_{i=1}^{N} \sigma_i \mathcal{I}_i(\theta)\, \eta_i'(\theta)\, \mathrm{d}\theta}{\sum_{i=1}^{N} \sigma_i \mathcal{I}_i(\theta)\, \eta_i'(\theta)\, \mathrm{d}\theta / y_i(\theta)}. \tag{18b}$$

Here $\sigma_i = -1$ for monotonically decreasing η_i and $\sigma_i = 1$ for monotonically increasing η_i. Later we will show that $(x_{\mathrm{T}}, y_{\mathrm{T}})$ is a weighted

average chromaticity. The chromaticity coordinates $(x_\mathrm{T}, y_\mathrm{T})$ must be on the straight line relating $x(t)$ and $y(t)$. From the color mixing rule (15) we can see that the chromaticity resulting from mixing two beams must be on the straight line segment between the chromaticities of the original beams. Subsequently, either (18a) or (18b) is redundant. It is most convenient to drop (18a), because (18a) is slightly more complicated than (18b). From (17) and (18b) we derive the following system of differential equations:

$$\sum_{i=1}^{N} \sigma_i \, \eta_i'(\theta) \, \mathcal{I}_i(\theta) = \mathcal{G}(\theta), \tag{19a}$$

$$\sum_{i=1}^{N} \sigma_i \, \eta_i'(\theta) \, \mathcal{I}_i(\theta)/y_i(\theta) = \mathcal{G}(\theta)/y_\mathrm{T}. \tag{19b}$$

The equations (19) describe a system of two coupled ordinary differential equations. In the following, we choose $N = 2$, because otherwise the system is underdetermined.

From (18a) and (18b) we can derive expressions to calculate $(x_\mathrm{T}, y_\mathrm{T})$. First rewrite (18a) and (18b) as

$$\left(\sum_{i=1}^{2} \sigma_i \, \mathcal{I}_i(\theta) \, \mathrm{d}t_i/y_i(\theta) \right) x_\mathrm{T} = \sum_{i=1}^{2} \sigma_i \, \mathcal{I}_i(\theta) \, x_i(\theta) \, \mathrm{d}t_i/y_i(\theta), \tag{20a}$$

$$\left(\sum_{i=1}^{2} \sigma_i \, \mathcal{I}_i(\theta) \, \mathrm{d}t_i/y_i(\theta) \right) y_\mathrm{T} = \sum_{i=1}^{2} \sigma_i \, \mathcal{I}_i(\theta) \, \mathrm{d}t_i. \tag{20b}$$

The domains of the transfer functions are adjacent. Therefore, for any function $F(s)$, we have

$$\sum_{i=1}^{2} \sigma_i \int_{\eta_i(0)}^{\eta_i(\theta_{\max})} F(t) \, \mathrm{d}t = \int_{0}^{\pi/2} F(t) \, \mathrm{d}t, \tag{21}$$

and thus, integrating (20) yields

$$x_\mathrm{T} = \frac{\int_0^{\pi/2} \mathcal{I}(t) \, x(t)/y(t) \, \mathrm{d}t}{\int_0^{\pi/2} \mathcal{I}(t)/y(t) \, \mathrm{d}t}, \tag{22a}$$

$$y_\mathrm{T} = \frac{\int_0^{\pi/2} \mathcal{I}(t) \, \mathrm{d}t}{\int_0^{\pi/2} \mathcal{I}(t)/y(t) \, \mathrm{d}t}, \tag{22b}$$

which shows that $(x_\mathrm{T}, y_\mathrm{T})$ are weighted average chromaticity coordinates, indeed.

Before evaluating these relations, we will have to choose the values of σ_1 and σ_2. For now we will choose $\sigma_1 = 1$ and $\sigma_2 = -1$. This choice immediately implies the initial values of η_1 and η_2. The function η_1 should map $[0, \theta_{\max}]$ to an interval $[0, \tau_1]$ for some $\tau_1 \in (0, \pi/2)$ and should be monotonically increasing. This implies $\eta_1(0) = 0$. Similarly, the function η_2 should map $[0, \theta_{\max}]$ to the interval $[\tau_1, \pi/2]$ and η_2 should be monotonically decreasing, so $\eta_2(0) = \pi/2$. If the coefficient matrix of (19) is not singular, we can derive the following expression for η_1' and η_2':

$$\eta_1'(\theta) = \frac{\mathcal{G}(\theta)}{\mathcal{I}_1(\theta)} \frac{1/y_2(\theta) - 1/y_{\mathrm{T}}}{1/y_2(\theta) - 1/y_1(\theta)}, \tag{23a}$$

$$\eta_2'(\theta) = -\frac{\mathcal{G}(\theta)}{\mathcal{I}_2(\theta)} \frac{1/y_{\mathrm{T}} - 1/y_1(\theta)}{1/y_2(\theta) - 1/y_1(\theta)}. \tag{23b}$$

3.2. REMOVABLE SINGULARITIES

At $\theta = 0$ we have a division of 0 by 0 in (23) because $\mathcal{I}_1(0) = \mathcal{I}_2(0) = 0$ and $\mathcal{G}(0) = 0$.

Theorem 1. *Assume $\mathcal{G}_+'(0) > 0$, $\mathcal{I}_+'(0) > 0$ and $\mathcal{I}_-'(\pi/2) < 0$. Also assume $\eta_1'(0) \neq 0$ and $\eta_2'(0) \neq 0$. At $\theta = 0$ we have*

$$\eta_1'(0) = \sqrt{\frac{\mathcal{G}_+'(0)}{\mathcal{I}_+'(0)} \frac{1/y(\pi/2) - 1/y_{\mathrm{T}}}{1/y(\pi/2) - 1/y(0)}}, \tag{24a}$$

$$\eta_2'(0) = -\sqrt{-\frac{\mathcal{G}_+'(0)}{\mathcal{I}_-'(\pi/2)} \frac{1/y_{\mathrm{T}} - 1/y(0)}{1/y(\pi/2) - 1/y(0)}}. \tag{24b}$$

Proof. Because of the properties of $\mathcal{I}(t)$, $\eta_1(\theta)$ and $\eta_2(\theta)$ mentioned earlier, we have that $\mathcal{I}_1(\theta) \to 0$ and $\mathcal{I}_2(\theta) \to 0$ when $\theta \downarrow 0$. Also we have $\mathcal{G}(\theta) \to 0$. We can calculate $\lim_{\theta \downarrow 0} \mathcal{G}(\theta)/\mathcal{I}_1(\theta)$ and $\lim_{\theta \downarrow 0} \mathcal{G}(\theta)/\mathcal{I}_2(\theta)$ using l'Hôpital's rule:

$$\lim_{\theta \downarrow 0} \frac{\mathcal{G}(\theta)}{\mathcal{I}(\eta_1(\theta))} = \lim_{\theta \downarrow 0} \frac{\frac{\mathrm{d}}{\mathrm{d}\theta} \mathcal{G}(\theta)}{\frac{\mathrm{d}}{\mathrm{d}\theta} \mathcal{I}(\eta_1(\theta))} = \frac{\mathcal{G}_+'(0)}{\mathcal{I}_+'(0)\, \eta_1'(0)}, \tag{25a}$$

$$\lim_{\theta \downarrow 0} \frac{\mathcal{G}(\theta)}{\mathcal{I}(\eta_2(\theta))} = \lim_{\theta \downarrow 0} \frac{\frac{\mathrm{d}}{\mathrm{d}\theta} \mathcal{G}(\theta)}{\frac{\mathrm{d}}{\mathrm{d}\theta} \mathcal{I}(\eta_2(\theta))} = \frac{\mathcal{G}_+'(0)}{\mathcal{I}_-'(\pi/2)\, \eta_2'(0)}. \tag{25b}$$

Note that $\mathcal{G}_+'(0)$, $\mathcal{I}_+'(0)$ and $\mathcal{I}_-'(\pi/2)$ are the left and right derivatives as defined earlier. Evaluating (23) for $\theta \downarrow 0$ with substitution of (25) yields (24). Also note that the expressions under the square root are always strictly positive because $1/y(t)$ is a monotonic function, $\mathcal{G}_+'(0) > 0$ and $\sigma_i \mathcal{I}'(\eta_i(0)) > 0$. $\square$

12

Lemma 1. $\eta_1(\theta) \leq \eta_2(\theta)$ *for all* $\theta \in [0, \theta_{\max}]$.

Proof. Suppose $\eta_1(\theta) > \eta_2(\theta)$ for some $0 \leq \theta \leq \theta_{\max}$. Integration of (19a) yields

$$\sum_{i=1}^{2} \sigma_i \int_{\eta_i(0)}^{\eta_i(\theta)} \mathcal{I}(t)\,\mathrm{d}t = \int_0^{\theta} \mathcal{G}(\psi)\,\mathrm{d}\psi. \tag{26}$$

Using (26) we have:

$$\int_0^{\theta} \mathcal{G}(\psi)\,\mathrm{d}\psi = \int_0^{\eta_1(\theta)} \mathcal{I}(t)\,\mathrm{d}t + \int_{\eta_2(\theta)}^{\pi/2} \mathcal{I}(t)\,\mathrm{d}t >$$
$$\int_0^{\pi/2} \mathcal{I}(t)\,\mathrm{d}t = \int_0^{\theta} \mathcal{G}(\psi)\,d\psi, \tag{27}$$

where the last equality arises from conservation of luminous flux. This shows that such values for $\eta_1(\theta)$ and $\eta_2(\theta)$ cannot be a solution to the integral equations, and we can conclude $\eta_1(\theta) \leq \eta_2(\theta)$. $\qquad\square$

Now we define t_{av} such that $y(t_{\mathrm{av}}) = y_{\mathrm{T}}$. We will see that $\eta_1(\theta_{\max}) = \eta_2(\theta_{\max}) = t_{\mathrm{av}}$.

Lemma 2. $\eta_1(\theta_{\max}) = \eta_2(\theta_{\max}) = t_{\mathrm{av}}$.

Proof. From Lemma 1 we deduce that $\eta_1(\theta_{\max}) \leq \eta_2(\theta_{\max})$. From (26) we can see that conservation of luminous flux is not satisfied for $\eta_1(\theta_{\max}) < \eta_2(\theta_{\max})$, and we conclude that $\eta_1(\theta_{\max}) = \eta_2(\theta_{\max})$. Integrate (19a) and (19b) from θ to $\theta_{\max}$ to find

$$\sum_{i=1}^{2} \sigma_i \int_{\eta_i(\theta)}^{\eta_i(\theta_{\max})} \mathcal{I}(t) \frac{y_{\mathrm{T}}}{y(t)}\,\mathrm{d}t = \sum_{i=1}^{2} \sigma_i \int_{\eta_i(\theta)}^{\eta_i(\theta_{\max})} \mathcal{I}(t)\,\mathrm{d}t. \tag{28}$$

Using $\eta_1(\theta_{\max}) = \eta_2(\theta_{\max})$ this simplifies to

$$\int_{\eta_1(\theta)}^{\eta_2(\theta)} \mathcal{I}(t) \frac{y_{\mathrm{T}}}{y(t)}\,\mathrm{d}t = \int_{\eta_1(\theta)}^{\eta_2(\theta)} \mathcal{I}(t)\,\mathrm{d}t. \tag{29}$$

Suppose that for some θ, $\eta_1(\theta) > t_{\mathrm{av}}$. Because $y(t)$ is monotonically increasing, we have $y_{\mathrm{T}}/y(t) < 1$ for $t \geq \eta_1(\theta)$, and (29) is not satisfied. Similarly, (29) is not satisfied if $\eta_2(\theta) < t_{\mathrm{av}}$. We can conclude that $\eta_1(\theta) \leq t_{\mathrm{av}}$ and $\eta_2(\theta) \geq t_{\mathrm{av}}$ Combining these two results for $\theta = \theta_{\max}$, we may conclude

$$\eta_1(\theta_{\max}) = t_{\mathrm{av}} = \eta_2(\theta_{\max}). \tag{30}$$

$\qquad\square$

Figure 5b shows the consequence of the last two lemmas: the graph of η_1 is always below η_2 and both graphs end at t_{av}. This gives rise to a second singularity at $\theta \uparrow \theta_{\max}$: both $y_1(\theta)$ and $y_2(\theta)$ approach y_{T} and again we have division of 0 by 0 in (23).

Theorem 2. *The singularity of (23) at $\theta = \theta_{\max}$ can be removed with*

$$\eta_1'(\theta_{\max}) = \frac{\mathcal{G}(\theta_{\max})}{2\,\mathcal{I}(t_{\mathrm{av}})} > 0, \tag{31a}$$

$$\eta_2'(\theta_{\max}) = -\frac{\mathcal{G}(\theta_{\max})}{2\,\mathcal{I}(t_{\mathrm{av}})} < 0. \tag{31b}$$

Proof. Using l'Hôpital's rule and the chain rule for differentiation, we derive from (23):

$$\eta_1'(\theta_{\max}) = \lim_{\theta \uparrow \theta_{\max}} \eta_1'(\theta) \quad = \frac{\mathcal{G}(\theta_{\max})}{\mathcal{I}(t_{\mathrm{av}})} \frac{\eta_2'(\theta_{\max})}{\eta_2'(\theta_{\max}) - \eta_1'(\theta_{\max})}, \tag{32a}$$

$$\eta_2'(\theta_{\max}) = \lim_{\theta \uparrow \theta_{\max}} \eta_2'(\theta) \quad = \frac{\mathcal{G}(\theta_{\max})}{\mathcal{I}(t_{\mathrm{av}})} \frac{\eta_1'(\theta_{\max})}{\eta_2'(\theta_{\max}) - \eta_1'(\theta_{\max})}. \tag{32b}$$

Solving these for $\eta_1(\theta_{\max})$ and $\eta_2(\theta_{\max})$ yields (31). Here we ignored the possibility $\eta_1'(\theta_{\max}) = \eta_2'(\theta_{\max}) = 0$, but substitution in (19) shows that the derivatives of η_1 and η_2 must be non-zero indeed, if $\mathcal{G}(\theta_{\max}) > 0$. $\square$

4. Numerical results

The method described in Section 2 has been tested for an LED with a particularly high CoA variation. The chromaticity and intensity of this LED has been measured, and this data has been used to calculate the free surfaces of a TIR collimator. Because degrees are more common in optics, the calculations and experiments in this section are expressed in degrees instead of radians. We evaluated the performance of the TIR collimator using the raytracing software LightTools [1].

4.1. MODELLING OF THE LED

The LED has been measured using a goniophotometer. A goniophotometer is a device which measures intensity, chromaticity, and many other characteristics of light at different solid angles. Our LED was measured at 46 different angles t (with respect to the surface normal), between 0 and 90 degrees, and at 4 different angles u. The values were

averaged over the angles u because we assume rotational symmetry of the system.

The measured intensity of the LED has been fitted in Matlab using the polynomials $t^i - 90^i$ with $i = 2, \ldots, 7$. This set of polynomials has been chosen because they are zero at $t = 90°$. Furthermore, $I(t, u)$ is smooth, therefore the set of polynomials must have zero derivative with respect to t at $t = 0°$. A linear least squares fit[5] yields coefficients C_i. The following effective intensity function has been used in the Matlab calculation:

$$\mathcal{I}(t) = \sin(t) \left(\sum_{i=2}^{7} C_i \left(t^i - 90^i \right) \right). \tag{33}$$

The x and y chromaticity values have been fitted using the polynomials t^i for $i = 0, 2, 3, 4, 5, 6, 7$. These polynomials have been chosen because they have zero derivative at $t = 0$. The linear least squares fit yields coefficients D_i^x and D_i^y. The chromaticity functions used in the Matlab calculations are

$$x(t) = D_0^x + \sum_{i=2}^{7} D_i^x t^i, \tag{34a}$$

$$y(t) = D_0^y + \sum_{i=2}^{7} D_i^y t^i. \tag{34b}$$

The polynomials in (33), (34a) and (34b) have been used in solving the differential equations (23). Plots of the polynomials can be seen in Figure 4a and 4b. The monotonicity assumed in Subsection 3.1 is not fulfilled at angles larger than 70°, but because of the small luminous flux within this range, this does not affect the numerical solution method and the solvability for the ODE.

In the raytracing program LightTools, a 3D model was made to simulate the LED. The model was constructed using 46 emitting surfaces. Every surface emits light within a range of angles. The surfaces $k = 2, 3, \ldots, 45$ emit light within a range of angles $(2(k-1)-1, 2(k-1)+1)$, with the intensity and chromaticity coordinates corresponding to the measurement at angle $t = 2(k - 1)$. Surface 1 and 46 emit light corresponding the measured data at angle 0 and 90, but emit only in the ranges $(0, 1)$ and $(89, 90)$. A far-field receiver is added to the model, which measures the angles of the emitted light rays and calculates the intensity and chromaticity pattern from a Monte-Carlo raytracing simulation.

A comparison of the measured data, the least squares fit and the raytracing results of the LightTools model of the LED without collimator can be seen in Figure 4a and 4b. The measured data show an
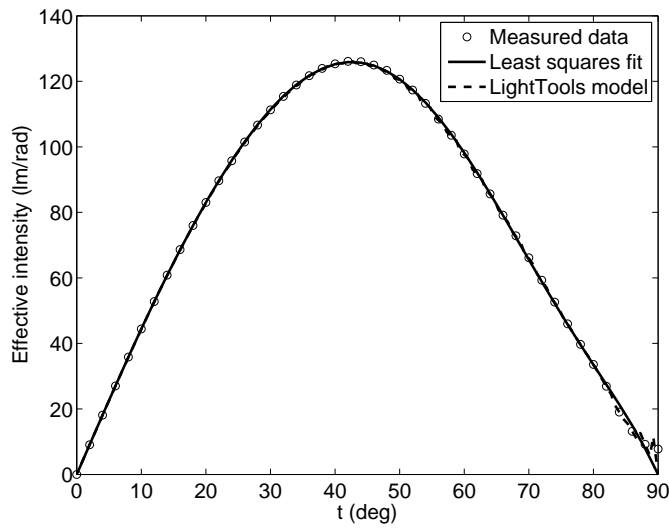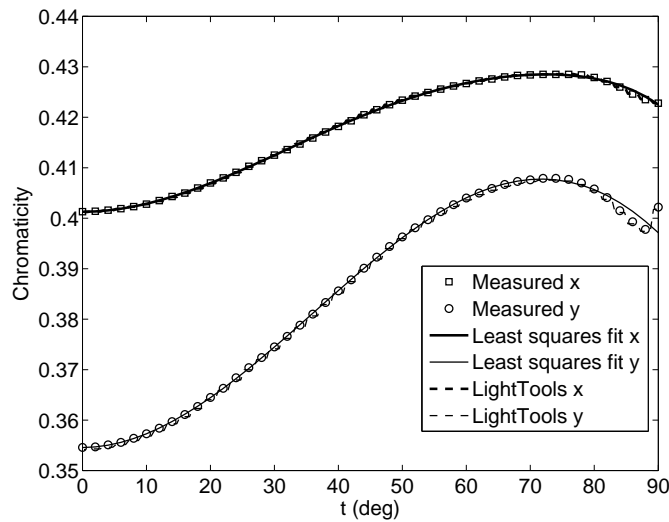
(a) Effective intensity $\mathcal{I}$ as function of t.

(b) Chromaticity coordinates x and y as function of t.

Figure 4. Comparison of measured data, least squares fit and LightTools model of the LED.

intensity pattern close to Lambertian, but display some irregularity at large angles, both in intensity and in chromaticity values. These irregularities are due to imperfections in the measurements. A scatter plot of the measured x and y chromaticity coordinates has been shown earlier in Figure 3. The plot shows the near-linear relationship between x and y, indeed.

4.2. Computation of the transfer functions

An example calculation has been done using the data as seen in Figure 4a and 4b. The target intensity was chosen to be a Gaussian profile[13] with Full Width at Half Maximum (FWHM)[12] at 20°. This yields an effective target intensity

$$\mathcal{G}(\theta) = C \sin(\theta) \exp\left(-4\ln(2)\left(\frac{\theta}{\theta_{\mathrm{FWHM}}}\right)^2\right), \qquad (35)$$

with $0 \leq \theta \leq 1.25\,\theta_{\mathrm{FWHM}}$, $\theta_{\mathrm{FWHM}} = 20°$ and C chosen such that

$$\int_0^{1.25\theta_{\mathrm{FWHM}}} \mathcal{G}(\theta)\,\mathrm{d}\theta = \int_0^{90} \mathcal{I}(t)\,\mathrm{d}t. \qquad (36)$$

We can solve the differential equations (23) using for example ode45 in Matlab. For small values of $\mathcal{I}_1(\theta)$ and $\mathcal{I}_2(\theta)$ at small θ, the relations (24) should be used. However, for $\theta \to \theta_{\max}$, the numerical solution (Figure 5a) shows a singlarity that cannot be solved using (31). At this point, the graphs of $\eta_1(\theta)$ and $\eta_2(\theta)$ approach the line $t = t_{\mathrm{av}}$. Because of slight calculation errors in the calculation of y_{T}, t_{av} and in solving the differential equation, the η-functions will not converge exactly to t_{av}. When one of the η-functions crosses the line $t = t_{\mathrm{av}}$, the sign of the derivative of the other η-function changes. From that point on, the error rapidly increases.

A solution to this problem is implementing a custom ODE-solver, which recalculates y_{T} every step in θ. This does not significantly alter the value of y_{T}, but the slight correction stabilizes the solver significantly. For this experiment, a custom Runge-Kutta method was implemented. The method calculates η_1 and η_2 at discrete levels $\theta_0, \theta_1, \ldots, \theta_{N_{\mathrm{s}}}$ with fixed step size. At every θ_i, the value of y_{T} is recalculated over the angles that have not yet been integrated, according to:

$$y_{\mathrm{T}i} = \frac{\int_{\eta_1(\theta_i)}^{\eta_2(\theta_i)} \mathcal{I}(t)\,\mathrm{d}t}{\int_{\eta_1(\theta_i)}^{\eta_2(\theta_i)} \mathcal{I}(t)/y(t)\,\mathrm{d}t}. \qquad (37)$$

Then, in every step of the Runge-Kutta algorithm, the value $y_{\mathrm{T}i}$ is used instead of y_{T} in (23). This method stabilizes the solver. The running time of this algorithm is a few seconds for $N_{\mathrm{s}} = 500$. The solution for the example problem found with this method can be seen in Figure 5b.
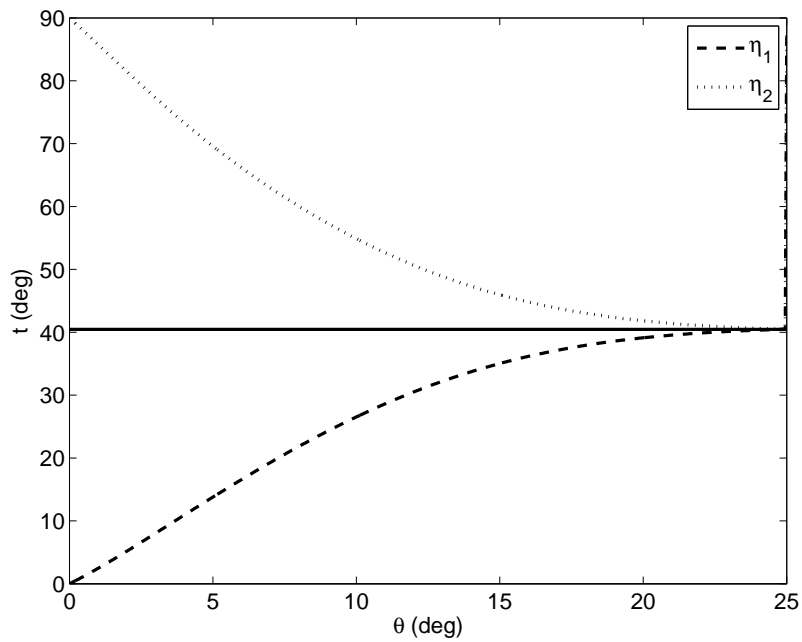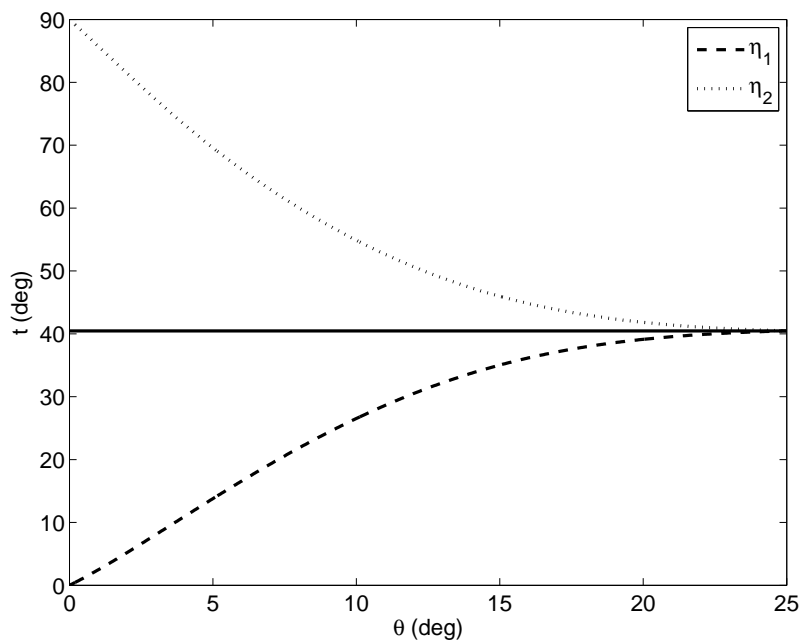
(a) Solution by ode45 from matlab.

(b) Solution by a y_T-recalculating ODE solver with $N_\mathrm{s} = 500$.

Figure 5. Solutions of (23) using fitted data and target intensity $\mathcal{G}(\theta)$ with Full Width Half Max at 20 degrees with two different ODE solvers. Note the instability at $\theta = \theta_{\max}$ in the ode45 solution.

4.3. EVALUATION OF THE TIR COLLIMATOR

Subsequently, a TIR collimator was calculated using the acquired transfer functions, using the method described in Subsection 2.2. This collimator was converted into a LightTools model. Every free surface was discretized by 500 points. An impression of the LightTools model of this collimator can be seen in Figure 6. The opening in the bottom of the collimator has a radius of 5 mm, and the collimator has a height of 9.47 mm.
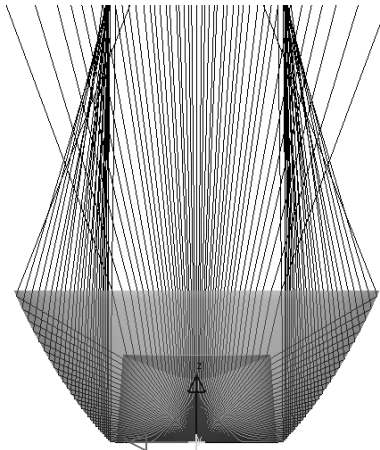

Figure 6. An impression of the LightTools model of the designed TIR collimator.

The model of the LED with the TIR collimator was evaluated by tracing 10^6 rays in non-dispersive mode. To ensure point-source behavior, the LED model was reduced in size to 0.01 mm by 0.01 mm. The results of the raytracing can be seen in Figure 7. The effective intensity shows the expected profile of a sine times a Gaussian between 0 and 21 degrees. The peak between 21 and 24 degrees is, judging from the chromaticity value of this light, due to a too high intensity coming from the reflective surface, possibly due to imperfections of the surface interpolation in LightTools. The chromaticity values are on two straight lines. The values between 0 and 1 degree are slightly too high. Part of this light originates from large angles from the LED, which is where the measurements show some irregularity, and especially the y-chromaticity values are a bit higher in the measurements than in the least squares fit. The chromaticity of the emitted light is entirely contained within one MacAdams ellipse, so the color difference is not noticeable for the human eye.
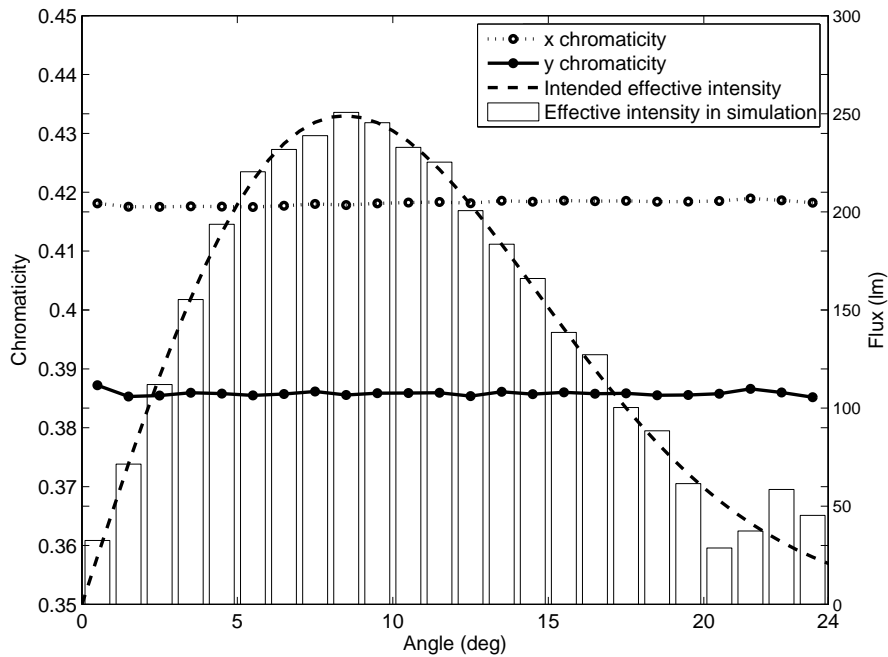
Figure 7. Results of the simulation with the two-segment collimator using Light-Tools. Here the source has a width of 0.01 mm. The bar plot shows the effective intensity. The open and closed dotted lines show the x and y chromaticity values.

5. Conclusion

A new approach has been introduced for reducing CoA variation in LED lighting systems based on inverse design methods. A system of two coupled ordinary differential equations was derived. Several characteristics of the solution were discussed in Section 3, and numerical issues of integrating the ODE system were resolved in Section 4. The method has been implemented for an LED spotlight with a TIR collimator. The collimator has been tested for a point source by Monte-Carlo raytracing in the raytracing software LightTools. The result of the experiment shows a reduction of the CoA variation to within the limits of visual perception (Figure 7).

In this article it has been shown that angular color mixing in the far field is possible for a point light source using inverse methods. Also, it has been shown that it is possible to remove all color variation by mixing light from only two different emission angles from the light source.

In future research, we want to include the finite size of the light source in the method. LEDs have a diameter of typically a millimeter,

which is too large compared to the optics size to be considered a point source. Furthermore, we are interested in solutions of (19) for $N > 2$. This would allow the calculation of transfer functions for a broader range of collimators, and allows for more design freedom of TIR collimators. Finally, an equivalent method for LEDs that have a curved x-y-chromaticity characteristic would be of interest.

References

1. 'ORA LightTools product website'. `http://www.opticalres.com/lt/ltprodds_f.html`. Accessed February 27, 2012.

2. Bortz, J. and N. Shatz: 2006, 'Generalized functional method of nonimaging optical design'. In: R. Winston & P. Benitez (ed.): Nonimaging Optics and Efficient Illumination Systems III, Vol. 6338 of Proc. SPIE.

3. Bortz, J. and N. Shatz: 2010, 'Mathematical relationships between the generalized functional, edge-ray and sms design methods'. In: R. Winston & J. M. Gordon (ed.): Nonimaging Optics: Efficient Design for Illumination and Solar Concentration, Vol. 7785 of Nonimaging Optics: Efficient Design for Illumination and Solar Concentration.

4. Haitz, R. and Y. T. Jeffrey: 2011, 'Solid-state lighting: 'The case' 10 years after and future prospects'. Physica Status Solidi A **208**, 17–29.

5. Heath, M.: 2005, Scientific Computing. An Introductory Survey. McGraw Hill, 2nd edition.

6. Maes, M.: 1997, 'Mathematical methods for reflector design'. Ph.D. thesis, University of Amsterdam.

7. Maes, M. and A. Janssen: 1991, 'A Note on cylindrical reflector design'. Optik **88**, 177–181.

8. Malacara, D.: 2002, Color Vision and Colorimetry. Theory and Applications. SPIE Press.

9. Mueller, G. O., 'Luminescent ceramic for a light emitting device'. United States patent US7361938 B2.

10. Schubert, F.: 2006, Light-Emitting Diodes. Cambridge University Press, 2nd edition.

11. Wang, K., D. Wu, F. Chen, Z. Liu, X. Luo, and S. Liu: 2010, 'Angular color uniformity enhancement of white light-emitting diodes integrated with freeform lenses'. Optics Letters **35**, 1860–1862.

12. Weisstein, E., 'Full Width at Half Maximum'. `http://mathworld.wolfram.com/FullWidthatHalfMaximum.html`.

13. Weisstein, E., 'Gaussian Function'. `http://mathworld.wolfram.com/GaussianFunction.html`.

14. Wu, H., N. Narendran, Y. Gu, and A. Bierman: 2007, 'Improving the Performance of Mixed-Color White LED Systems by Using Scattered Photon Extraction Technique'. In: Proc. SPIE, Vol. 6669 of Proc. SPIE.

15. Wyszecki, G. and W. Stiles: 2000, Color Science. Concepts and Methods, Quantitative Data and Formulae. John Wiley and Sons, INC, 2nd edition.

PREVIOUS PUBLICATIONS IN THIS SERIES:

Number	Author(s)	Title	Month
12-05	J.H.M. ten Thije Boonkkamp J. van Dijk L. Liu K.S.C. Peerenboom	Extension of the complete flux scheme to systems of conservation laws	March '12
12-06	E.N.M. Cirillo A. Muntean	Dynamics of pedestrians in regions with no visibility – a lattice model without exclusion	March '12
12-07	J.M.L. Maubach W.H.A. Schilders	Micro- and macro-block factorizations for regularized saddle point systems	Apr. '12
12-08	K. Kumar I.S. Pop F.A. Radu	Convergence analysis for a conformal discretization of a model for precipitation and dissolution in porous media	Apr. '12
12-09	C.R. Prins J.H.M. ten Thije Boonkkamp T.W. Tukker W.L. IJzerman	An inverse method for the design of TIR collimators to achieve a uniform color light beam	Apr. '12