Optimum waist of localized basis functions in truncated series employed in some optical applications

Farshid Ghasemi* and Khashayar Mehrany

Department of Electrical Engineering, Sharif University of Technology, P.O. Box 11365-8639 Tehran, Iran

*Corresponding author: farshid.ghasemi@gatech.edu

Received 9 September 2009; revised 16 January 2010; accepted 23 January 2010; posted 1 February 2010 (Doc. ID 116797); published 3 March 2010

The waist parameter is a particularly important factor for functional expansion in terms of localized orthogonal basis functions. We present a systematic approach to evaluate an asymptotic trend for the optimum waist parameter in truncated orthogonal localized bases satisfying several general conditions. This asymptotic behavior is fully introduced and verified for Hermite–Gauss and Laguerre–Gauss bases. As a special case of importance, a good estimate for the optimum waist in projection of discontinuous profiles on localized basis functions is proposed. The importance and application of the proposed estimation is demonstrated via several optical applications. © 2010 Optical Society of America

OCIS codes: 000.4430, 050.1755, 350.4238, 060.2310, 130.2790, 130.5296.

1. Introduction

Exploitation of a truncated orthogonal series expansion as an approximation for the solution of a physical problem is a very well-known technique among physicists and engineers. The perpetual question propounded in this approach is which basis functions with what parameters should be used to achieve the highest accuracy by the least computational effort.

Once the set of basis functions is chosen, there is usually a set of free parameters left to be determined. In particular, the waist parameter for localized basis functions can play an important role in the accuracy of the expansion for nonsmooth functions, and its optimum determination has attracted much attention in different applications [1–7]. To determine the optimum waist for projecting a specific function on a specific set of localized basis functions, different suggestions have been proposed in the literature. For instance, Borghi *et al.* in [1] have proposed a rule of thumb for finding the optimum waist of two-dimensional (2D) Laguerre–Gauss (LG) basis functions in expanding the circ function. Liu *et al.* in [2] have studied the applicability of this rule in finding the optimum waist for one-dimensional (1D) Hermite–Gauss (HG) basis functions employed for expanding a step function. Some have also used structure-related parameters, e.g. normalized frequency of an optical waveguide, to determine the basis function parameters [8–11]. Some others have further contemplated on the problem for a class of signals and have minimized an upper bound for the squared error (e.g., see [3] and references therein).

Accurate determination of the most favorable parameters for basis functions, however, asks for numerical optimization techniques. The optimization process is, however, nonlinear and rather complex and, hence, calls for complicated optimization algorithms, e.g., iterative methods [5,6]. The difficulty gets even worse when the error function exhibits many local minima, increasing the probability that the optimization algorithm gets stuck in a certain minimum. To avoid such difficulties, some efforts have been made to analytically solve the nonlinear

0003-6935/10/081210-09$15.00/0

optimization problem, but the results are restricted to those circumstances where only a few basis functions are retained in the calculations [7]. Unfortunately, the optimized parameters may be considerably different when the number of retained basis functions is increased. Therefore, the optimized parameters may turn out to be nonoptimum when more basis functions are kept. In this context, a good estimate for the asymptotic behavior of the optimum parameter with respect to the number of basis functions is complementary to the already existing optimization formulations and obviates the need to follow a complicated optimization process.

To this end, the behavior of optimum waist parameter for a predetermined set of localized basis functions is analytically studied in this paper, where the following assumptions are made. The derivative of each basis function with respect to the free parameter is supposed to be in the same space spanned by the basis functions. The obtained results will, therefore, not be valid for discontinuous basis functions whose derivatives include the delta function. Furthermore, despite the generality of the proposed method, basis functions are assumed to form an orthogonal set and, thus, their derivative matrix (as defined in Appendix A) is a band matrix. This assumption is found very helpful in reducing the complexity of formulation. It is worth mentioning that the validity, as well as the applicability, of the proposed analytic approach in choosing the optimal free parameter is demonstrated for the HG and the LG bases. The rules of thumb as suggested by [1,2] are then extended and a systematic approach to evaluateing the asymptotic optimum parameter is presented. Finally, different practical applications are presented and the applicability and significance of the proposed method are exhibited.

2. Formulation

Assume $f(\mathbf{r})$ as a complex function in a complete square-integrable linear space $\mathbf{S} \subseteq L^2(\mathbb{R}^D)$, which is spanned by the set of basis functions $\psi_m(\mathbf{r}, w)$. D is the dimension of the Euclidean space. The sequence of coefficients $\{\alpha_m\}_{m=0}^{M-1}$ is to be determined in such a manner that $\sum_{m=0}^{M-1} \alpha_m \psi_m(\mathbf{r}, w)$ becomes the best approximation for $f(\mathbf{r})$ within S, minimizing the squared error defined by

$$e^2 = \int \left| f(\mathbf{r}) - \sum_{m=0}^{M-1} \alpha_m \psi_m(\mathbf{r}, w) \right|^2 d\mathbf{r}. \qquad (1)$$

Since the set of expansion coefficients $\{\alpha_m\}_{m=0}^{M-1}$ guarantees the best achievable approximation within S, the following set of linear algebraic equations must hold [12]:

$$\begin{aligned} \mathbf{G}^t{}_M \mathbf{A} &= \mathbf{B}_M, \\ \mathbf{G}_M &= [g_{i,k}]; \\ g_{i,k}(w) = \langle \psi_i, \psi_k \rangle &= \int \psi_i(\mathbf{r}, w) \psi_k^*(\mathbf{r}, w) \mathrm{d}\mathbf{r}, \\ \mathbf{B}_M = [b_i]; &\qquad b_i = \langle f, \psi_i \rangle, \\ A = [\alpha_i]; &\qquad i = 0, \ldots, M-1. \end{aligned} \qquad (2)$$

Here, the index i runs from 0 to $M-1$ and the superscripts t, T, and $*$ indicate the transpose, the conjugate transpose, and the complex conjugate, respectively. The minimum error e^2 can then be simplified as

$$e^2 = \langle f, f \rangle - \mathbf{B}_M^t \mathbf{G}_M^{-1} \mathbf{B}_M^*. \qquad (3)$$

Because w is a free parameter, it can be optimized to render the minimum error. The value of w can then be found by setting $\partial e^2 / \partial w = 0$. It will be assumed that the norm of each basis function $\langle \psi_m(\mathbf{r}, w) \rangle$ is independent of w. Using the chain rule and after some algebraic manipulations, $\partial e^2 / \partial w$ can be written as

$$\begin{aligned} \frac{\partial e^2}{\partial w} = &-\mathbf{B}_M^t (\mathbf{D}_M^T \mathbf{G}_M^{-1} + \mathbf{G}_M^{-1} \mathbf{D}_M) \mathbf{B}_M^* - \mathbf{B}_r^t \mathbf{D}_r^T \mathbf{G}_M^{-1} \mathbf{B}_M^* \\ &- \mathbf{B}_M^t \mathbf{G}_M^{-1} \mathbf{D}_r \mathbf{B}_r^*, \end{aligned} \qquad (4)$$

where $\mathbf{D}_M$, $\mathbf{D}_r$, and $\mathbf{B}_r$ are defined as in Appendix A. It is, however, assumed that the derivative of each basis function, $\partial \psi_m / \partial w$, resides in S. On the other hand, it can be shown that, when the basis is orthonormal, the following relation holds between $\mathbf{G}_M$ and $\mathbf{D}_M$ (see Appendix B):

$$\mathbf{G}_M^{-1} \mathbf{D}_M + \mathbf{D}_M^T \mathbf{G}_M^{-1} = 0. \qquad (5)$$

By substituting the preceding equation into Eq. (4),$\partial e^2 / \partial w$ can be further simplified as

$$\begin{aligned} \frac{\partial e^2}{\partial w} &= -2\mathrm{Re}\{\mathbf{B}_M^t \mathbf{G}_M^{-1} \mathbf{D}_r \mathbf{B}_r^*\} \\ &= -2\mathrm{Re}\left\{ \sum_{i=0}^{M-1} \sum_{k=M}^{\infty} \langle f, \psi_i \rangle \langle f, \psi_k \rangle^* (g_M^{-1})_i d_{i,k} \right\}. \end{aligned} \qquad (6)$$

Here, $(g_M^{-1})_i$ refers to (i,i) entry of $\mathbf{G}_M^{-1}$ (the inverse of a truncated Gram matrix). The zeros of this equation will be employed in Section 3 to discuss the displacement of the optimum waist parameter.

3. Asymptotic Behavior of Optimum Waist Parameter

The next problem is to determine the asymptotic trend of optimum w as the number of retained basis functions, M, tends to infinity. To study the asymptotic behavior of the optimum waist parameter, we should inspect the displacement of the zeros of $\partial e^2 / \partial w$, hereafter denoted by w_M^{opt}, as a function of M. It is, therefore, necessary to study the dependence of $\mathbf{D}_w$ and also the inner products $\langle f, \psi_i \rangle$ on w, whose characteristics govern the zeros of $\partial e^2 / \partial w$ in Eq. (6).

First, it can be shown that, if the waist parameter w just scales the basis functions linearly along the radial direction in $\mathbb{R}^D$ (as for HG basis) then the entries of $\mathbf{D}_w$, i.e., $d_{i,k}$ in Eq. (6), are proportional to $1/w$. Furthermore, if the basis functions depend merely on the radial distance, $|\bar{\mathbf{r}}|$ (as for LG basis), then all the entries become proportional to D/w. These two facts are later employed in this section to simplify the difference equation governing the zeros of $\partial e^2/\partial w$.

Second, we study the zeros of the other agent involved in Eq. (6), i.e., $F_M(w) \triangleq \langle f, \psi_M \rangle$, as the first step toward extraction of the asymptotic trend for the sought-after roots of Eq. (6). To this end, $F_M(w)$ is approximated around one of its zeros as

$$F_M(w) \approx a_M(w - w_M), \tag{7}$$

where w_M denotes an arbitrary zero of $F_M(w)$ and a_M is the first-order coefficient in Taylor expansion. It should be also noted that $F_M(w)$ might have several zeros and w_M should, therefore, have another index discriminating various zeros for a fixed M. Here, this index is considered to be fixed and is omitted for the sake of brevity.

Next, to obtain the relation among zeros for successive truncation orders, i.e., w_M, w_{M+1}, w_{M+2}, etc., we contemplate the derivative of this function with respect to w:

$$\frac{\partial}{\partial w} F_M(w) = \frac{\partial}{\partial w} \langle f, \psi_M \rangle = \sum_i d_{M,i} F_i(w). \tag{8}$$

Now, by using the fact that $d_{M,i}$ is proportional to D/w and substituting Eq. (7) into Eq. (8), the following two difference equations corresponding to the zeroth and first-order terms are obtained as follows:

$$\frac{1}{D} a_M = \sum_i \tilde{d}_{M,i} a_i, \tag{9a}$$

$$0 = \sum_i \tilde{d}_{M,i} a_i w_i, \tag{9b}$$

where $\tilde{d}_{M,i}$ is defined as the value of $d_{M,i}$ for $w = 1$.

Once the trend of w_M is extracted by solving the above-mentioned difference equations, the trend of w_M^{opt} follows the same line. This point is further discussed in Section 4.

4. Special Bases

In this section, the approach outlined in Section 3 is applied to the HG and the LG basis functions.

A. Hermite–Gauss Basis

HG functions are defined as

$$h_m(x, w) = \frac{1}{\sqrt{w\sqrt{\pi}}} \frac{1}{\sqrt{2^m m!}} H_m\left(\frac{x}{w}\right) \times \exp\left(\frac{-x^2}{2w^2}\right) \quad (x \in \mathbb{R}), \tag{10}$$

where H_m is the Hermite polynomial [13]. In this basis, Eqs. (9) result in fourth-order difference equations. Since even and odd orders become decoupled, without loss of generality, we can assume that M is even. By changing the variables as $M \to 2m$ and $a_{2m} \to a'_m$, Eq. (9a) now turns to

$$a'_m = \frac{-1}{2}\sqrt{2m(2m-1)}\, a'_{m-1} + \frac{1}{2}\sqrt{(2m+1)(2m+2)}\, a'_{m+1}, \tag{11}$$

and Eq. (9b) yields

$$\frac{w_{M+2}}{w_{M-2}} = \frac{\sqrt{M(M-1)}}{\sqrt{(M+1)(M+2)}} \frac{a_{M-2}}{a_{M+2}} = \frac{\sqrt{M(M-1)}}{\sqrt{(M+1)(M+2)}} \frac{a'_{m-1}}{a'_{m+1}}. \tag{12}$$

To acquire the asymptotic behavior of w_M, therefore, a'_{m-1}/a'_{m+1} is needed. To this end, the Birkhoff–Adams theorem [14] can be employed. Regarding the definitions in [14], Eq. (11) has two *normal solutions*. Since we are interested in the asymptotic behavior of a'_m, we will retain only the first two dominant terms of the solution with respect to $1/m$. In this way, the solution for a'_m relies on the initial conditions of Eq. (11), which, in turn, depend on the given function f. However, by plugging the solutions into Eq. (12), it would be observed that the initial conditions do not play any role in the first two dominant terms of a'_{m-1}/a'_{m+1}. The final result then reads as

$$\frac{a'_{m-1}}{a'_{m+1}} \sim 1 + O\left(\frac{1}{m^2}\right) \sim 1 + O\left(\frac{1}{M^2}\right), \tag{13}$$

where $\sim$ denotes the asymptotic behavior. Substituting Eq. (13) into Eq. (12) then leads to

$$\frac{w_{M+2}}{w_{M-2}} \sim \left(1 - \frac{1}{2M}\right) \sim \frac{\sqrt{M-2}}{\sqrt{M+2}}. \tag{14}$$

In the last step of deriving Eq. (14), we have cast the asymptotic behavior in the form of $f(M+2)/f(M-2)$, so that the contributions of w_{M+2} and w_{M-2} can be discriminated. Finally, for some proportionality constant μ, we have

$$w_M \approx \frac{\mu}{\sqrt{M}}. \tag{15}$$

Now, if the derivative matrix $\mathbf{D}_w$ is banded, which is the case for HG and LG bases, then $\partial e^2/\partial w$ can be expressed a

$$\frac{\partial e^2}{\partial w} = -2\mathrm{Re}\left\{\sum_{1\le\varepsilon_i\le W/2}\sum_{0\le\varepsilon_k\le W/2}\langle f,\psi_{M-\varepsilon_i}\rangle\langle f,\psi_{M+\varepsilon_k}\rangle^*(g_M^{-1})_{M-\varepsilon_i}d_{M-\varepsilon_i M+\varepsilon_k}\right\}, \tag{16}$$

where ε_i and ε_k do not exceed the half-bandwidth of $\mathbf{D}_w$, $W/2$. By replacing the linear approximation of $\langle f,\psi_i\rangle$ around its zero, Eq. (7), into Eq. (16), an approximate quadratic equation for the zeros of the error derivative is achieved. Whatever this equation is, its zeros will obey the same behavior as w_M. The reason is as follows: for each of the terms in the form of $a_{M+\varepsilon}(w-w_{M+\varepsilon})$, the zero position is scaled by $w_{M+1+\varepsilon}/w_{M+\varepsilon}$ (ε is either ε_i or ε_k). Since $\varepsilon \ll M$, $w_{M+1+\varepsilon}/w_{M+\varepsilon}$ is independent of ε, up to the second order of approximation regarding $1/M$. Similarly, $a_{M+1+\varepsilon}/a_{M+\varepsilon}$, $d_{M+1-\varepsilon_i,M+1+\varepsilon_k}/d_{M-\varepsilon_i,M+\varepsilon_k}$, and $(g_M^{-1})_{M-\varepsilon_i}/(g_{M+1}^{-1})_{M+1-\varepsilon_i}$ are identical for all $\varepsilon \ll M$ and follow the same vein. This means that, with increasing M by one unit, all the zeros and the coefficients are identically scaled and, as a result, the zeros of the mentioned approximate quadratic equation will be scaled in the same manner as w_M.

B. Laguerre–Gauss Basis

LG functions are defined as

$$\varphi_m(x) = \sqrt{2\frac{n!}{(n+k)!}}(2x)^{k/2}L_m^k(2x) \times \exp(-x) \quad (x\in\mathbb{R}^+),$$
$$\varphi_m^{(D)}(\mathbf{r},w) = c_D\frac{1}{w^{D/2}}\varphi_m\left(\left(\frac{|\mathbf{r}|}{w}\right)^D,1\right). \tag{17}$$

L_m^k is the associated Laguerre polynomial [15] and D is the space dimension. To have an orthonormal basis, $c_1 = 1$, $c_2 = 1/\sqrt{\pi}$, and $c_3 = \sqrt{3/(4\pi)}$. It must be noted that this definition of basis only encompasses the radial functionality in higher dimensions. In other words, it is a complete basis for rotationally symmetric functions in $\mathrm{L}^2(\mathbb{R}^D)$.

For the LG basis, the steps to be taken are very similar to those for the HG basis. In this case, Eqs. (9) result in two second-order equations:

$$a_M = \frac{-D}{2}\sqrt{M(M+k)}a_{M-1} + \frac{D}{2}\sqrt{(M+1)(M+k+1)}a_{M+1}, \tag{18a}$$

$$0 = \frac{-D}{2}\sqrt{M(M+k)}a_{M-1}w_{M-1} + \frac{D}{2}\sqrt{(M+1)(M+k+1)}a_{M+1}w_{M+1}. \tag{18b}$$

By employing the Birkhoff–Adams theorem and retaining the first two dominant terms in $1/M$, we will observe the following trend:

$$w_M \approx \frac{\mu}{M^{1/D}}, \tag{19}$$

which, by similar discussions, is expected for the optimum waist parameter, as well. As proposed by Eq. (19), the k defined in Eq. (17) does not appear in asymptotic trend of the optimum waist.

5. Practical Applications

In this section, general aspects of the asymptotic optimum waist are numerically investigated in several optical applications. These numerical observations are instructive for making judicious estimations of the optimum waist in practical applications.

A. Application in Permittivity Profile Expansion

As the first example, we have chosen a three-dimensional (3D) spherical defect to be expanded in the 3D LG basis with $k = 0$. Employing the expansion of the structure is a common practice in some versions of the Galerkin method, such as the localized function method (LFM) [16]. To express the defect structure, we considered a function with a unit value inside the unit sphere and vanishing elsewhere.

In Fig. 1, the approximation in Eq. (19) with $\mu = 1$, i.e., $w = 1/\sqrt[3]{M}$, is compared to the optimum waist obtained by direct search. As it can be seen, choosing the distance of discontinuity from origin ($r = 1$) as μ in Eq. (19), is a good approximation. Similarly, it was observed that, for expanding functions with a single discontinuity at r_0 (step, circ, and spherical defect), $\mu = r_0$ gives a good approximation for 1D HG and LG bases, as defined in Eqs. (10) and (17).

It is worth highlighting that the error increases sharply when the waist becomes smaller than the optimum value (Fig. 2). This fact has been repeatedly observed for LG and 1D HG bases with different truncation orders, and for different functions with

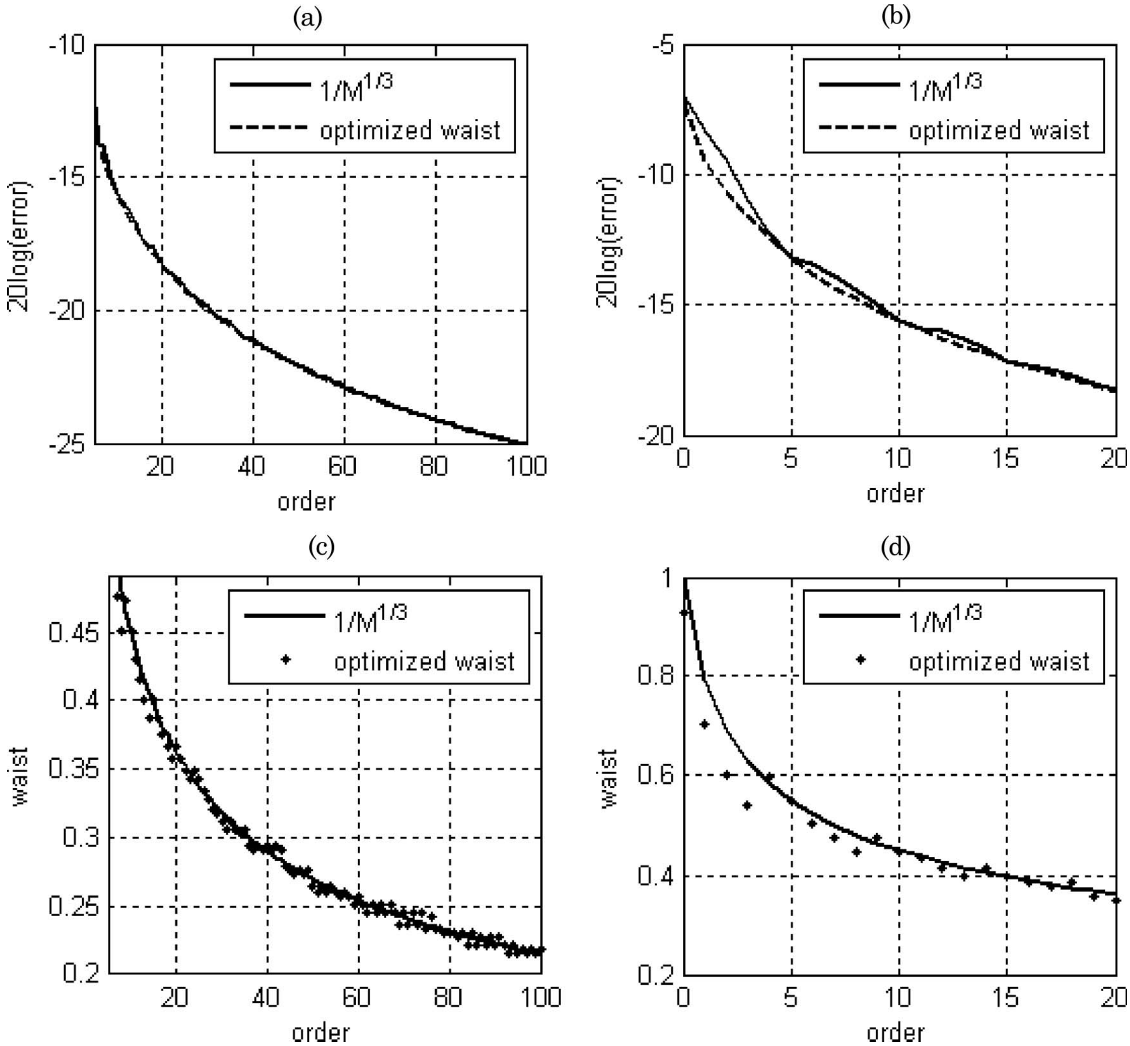

Fig. 1. Expansion of a sphere defect using the 3D LG basis. (a) Solid curve, error by $w = 1/\sqrt[3]{M}$; dashed curve, error by optimum waist. (b) Enlarged view of (a). (c) Solid curve, $w = 1/\sqrt[3]{M}$; dots, optimum waist. (d) Enlarged view of (c).

finite support. The reason is that, at the optimum waist, the width of the function and the width of the highest-order basis function are comparatively the same. Hence, by increasing the waist from its optimum value, some parts of the higher-order basis functions will gradually fall outside the support of the function. Thus, the weight of these basis functions must be small to obtain the optimum

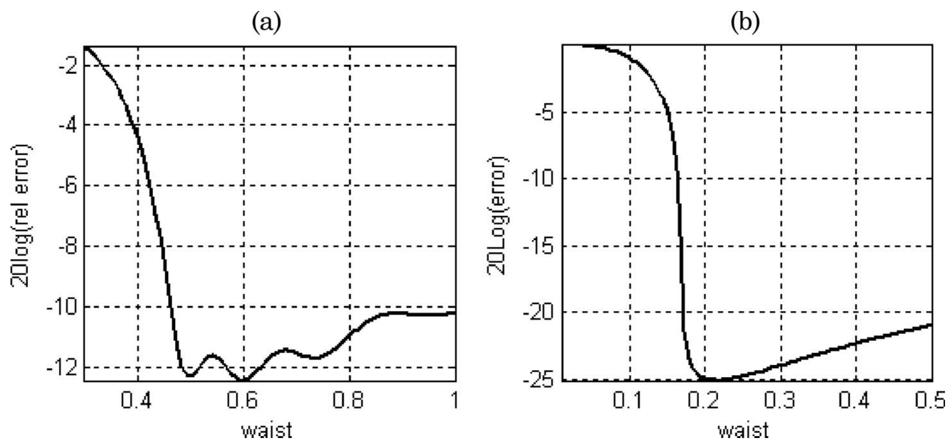

Fig. 2. Error versus waist: (a) order = 4, (b) order = 100.

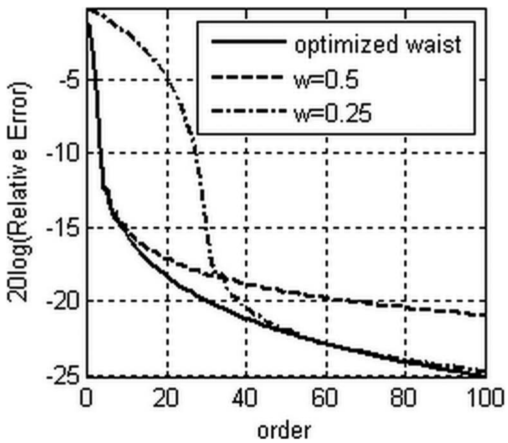

Fig. 3. Error versus order: solid curve, optimum waist; dashed curve, $w = 0.5$, dashed-dotted curve, $w = 0.25$.

expansion. In other words, the effective number of basis functions in the expansion will be lowered. On the other hand, when $w < w_{\mathrm{opt}}$, the effective widths of all basis functions become smaller than the support of the function. In this way, the end part of the support cannot be approximated by any basis function, leading to a sharp increase in the error. This fact is found to be useful in determining the optimum waist when expanding discontinuous functions.

In Fig. 3, the error of the optimum waist is compared to that of two fixed nonoptimum waists. When the fixed waist is smaller than w_{opt} in a specific order, as for $w = 0.25$ and orders of 10 to 40, the error sharply detaches from the minimum error. On the other hand, when the fixed waist is larger than the optimum one, as for $w = 0.5$ and orders of 10 to 100, the error slightly degrades. This is in accordance with what is already demonstrated and described in Fig. 2.

B. Application in Bragg Fiber Structure

In this section, the expansion of a Bragg fiber structure [17] in the 2D LG basis with $k = 0$ is studied. The considered structure [Fig. 4(a)] contains six discontinuities and its functionality can be expressed as the sum of six simple step functions. Although the total expansion is not simply a superposition of single expansions for single steps, we know that the error for reconstruction of a single step drastically blows up when the waist becomes smaller than its optimum value. Like the previous example, for a single discontinuity at $r = 3$, the suggestion $\tilde{w} = 3/\sqrt{M}$ gives a satisfactory approximation of the optimum waist. If we take the waist smaller than $\tilde{w}$, the error in reconstruction of the farthest discontinuity sharply increases and dominates other errors. On the other hand, if the waist is more than $\tilde{w}$, the reconstruction error increases for all the discontinuities. Hence, it appears that this $\tilde{w}$ must work well as an estimation for the optimum waist. The error of this estimation is compared to the error floor in Fig. 4.

C. Application in the Beam Propagation Method

Optimal free parameter selection in orthonormal bases can be accomplished by minimizing the upper bound of the quadratic truncation error for a class of

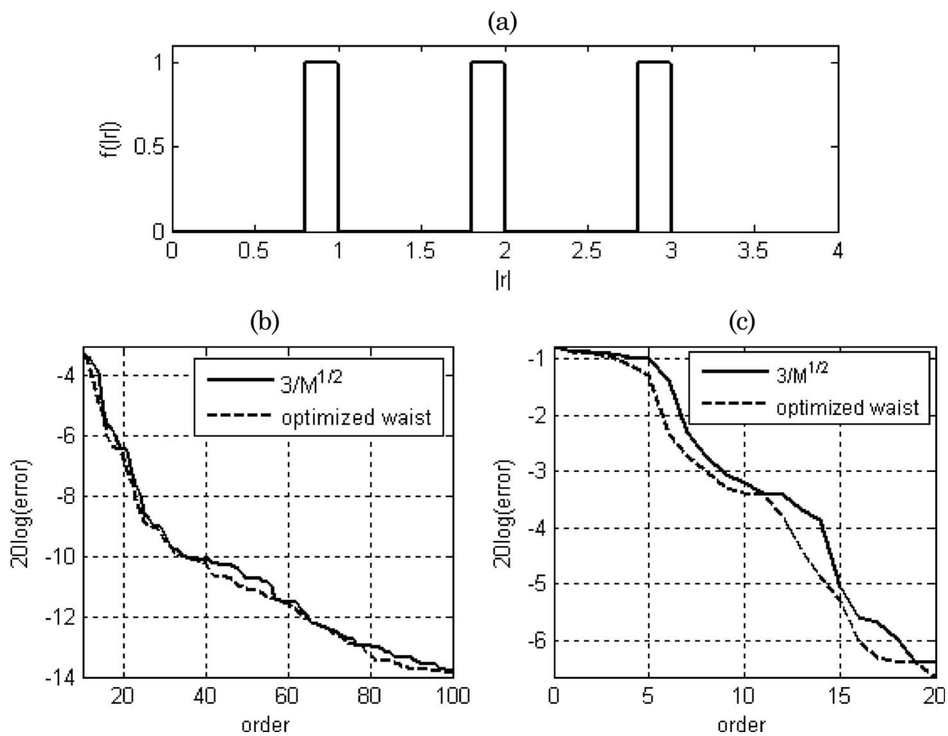

Fig. 4. Expansion of Bragg fiber structure in the 2D LG basis. (a) Bragg structure. (b) Dashed curve, error by optimum waist; solid curve, error by $w = 3/\sqrt{M}$. (c) Enlarged view of (b) for small orders.

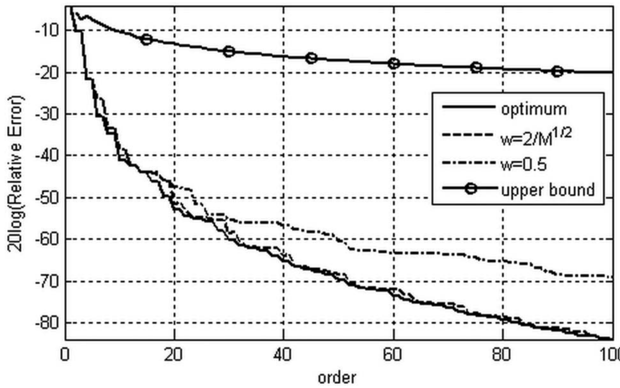

Fig. 5. Expansion of raised-cosine slab in 1D-HG basis. Error is plotted versus order. Solid curve, optimum waist; dashed curve, $w = 2/\sqrt{M}$; dashed-dotted curve, $w = 0.5$; circles, minimized upper bound [3].

functions. Den Brinker *et al.* in [3] used this approach to find the optimal waist parameter for several bases, including HG and LG. This suggestion, although interesting for signal processing applications, has several major shortcomings. First, the relation proposed in [3] calls for computing the norm of the derivative of the function. This norm does not exist for a discontinuous function. Even for slightly smoothened jumps, this norm is large and results in an obviously nonoptimum suggestion for the waist parameter. Second, the upper bound minimized in this approach is not tight enough for many instances of practical interest. Although [18] has used this optimal waist successfully, it should be noted that the function this reference expands is the real mode of a slab waveguide. For such bell-shaped functions as guided modes in waveguides, the HG expansion essentially converges fast and is less vulnerable to the choice of optimum waist.

As an example, we deal with the analysis of the reflection from a slab waveguide using BPM [19]. This paper has embedded 1D HG expansion in the beam propagation method (BPM) formulation. Since the

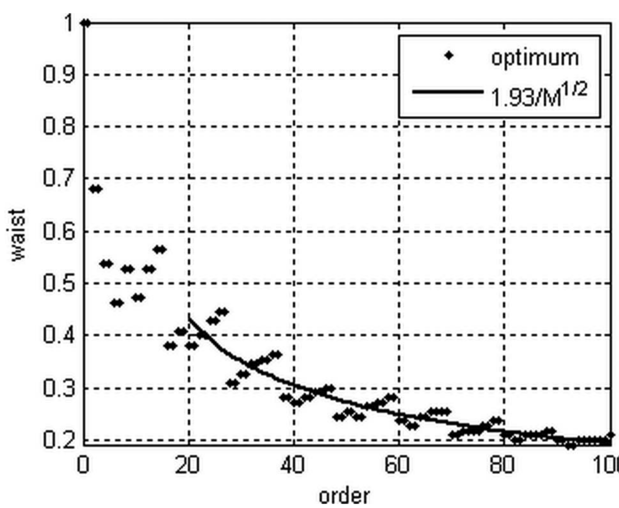

Fig. 6. Expansion of raised-cosine slab in the 1D HG basis: dots, optimum waist; solid curve, $1.93/\sqrt{M}$.

incident field is assumed to be a HG beam, it is suitably reconstructed by few HG functions. But the structure has the shape of a smoothened step and its expansion entails many more basis functions. Here we examine the optimum waist for expanding the structure.

The function to be expanded is a raised cosine. Comparing to [19], we have considered a wider $\cos^2$ section, so that the suggestion of [3] can produce a reasonable estimation. The function is defined as

$$f(x) = \begin{cases} 1 & |x| < 1 \\ \cos^2\left(\frac{\pi}{2}\left(|x| - 1\right)\right) & 1 < |x| < 2 \end{cases}. \tag{20}$$

There are two discontinuities at $|x| = 1$ and $|x| = 2$ in the second derivative of this function. The second derivative of the expansion includes the HG function of orders 0 to $M + 1$ with only M degrees of freedom in their coefficients. However, if we were willing to expand the second derivative efficiently with $M + 2$ basis functions, the optimum waist was $\tilde{w} = 2/\sqrt{M+2} \approx 2/\sqrt{M}$. The optimum waist as suggested by [3] is $\tilde{w} = 0.5$. In Fig. 5, the relative error is plotted against the highest retained order. For orders up to 15, these two approximations generate almost the same error. But in higher accuracies (error < −50 dB), the difference is conspicuous. For order = 100, they differ by almost 13 dB. Also included in this figure is the upper bound minimized by [3]. Using the notations of [3], the plotted curve represents

$$\sqrt{\frac{F(w)}{M}} \cong \sqrt{\frac{\sum_{n=0}^{M-1} n a_n^2(w)}{M\|f\|^2}}, \tag{21}$$

with $w = 0.5$. As can be seen, this upper bound is not tight enough. The optimum waist is demonstrated in Fig. 6. The relation $\alpha/\sqrt{M}$ is fitted (by least squares) to the optimum waist for orders from 20 to 100, which has led to $\alpha_{\rm opt} = 1.93$.

D. Application in Studying Cutoff of Planar Waveguides

As the last example, we make use of the results reported in [20] to demonstrate the benefits of the introduced optimum waist. This paper computes the cutoff frequency of 1D waveguides by expanding the solution of the TE wave equation in a 1D HG basis. Under such a condition, the Galerkin method and the variational formulation yield the same system of equations [21]. Hence the free parameters must be optimized in such a way that the cutoff frequency is minimized.

Table 1. Optimum Waist in Fig. 3 of [20] (Middle Row) and Estimated Optimum Waist (Lower Row)

M	25	50	75	100
$W = \frac{1}{\alpha}$ at minimum	0.143	0.100	0.080	0.067
$\frac{1}{\sqrt{2M}}$	0.141	0.100	0.081	0.071

We do not know the explicit form of the field and thus we cannot apply the proposed method and be sure that the optimal choice of the free parameters is made. It is, however, possible to make use of the proposed method and find reasonable results. Fortunately, we know that it is infinitely differentiable everywhere, except at the boundaries. At the boundaries, which are positioned at $x = 0$ and $x = 1$, the second derivative is discontinuous. This knowledge helps us to propose the optimum waist. Sharma and Meunier, in [20], have expanded the field versus merely even orders of HG. So by retaining M basis functions, even orders ranging from 0 to $2(M-1)$ are employed. Hence, it can be expected that $\tilde{w} = 1/\sqrt{2M}$ renders a good estimation of the optimum waist. In Fig. 3 of [20], the cutoff frequency is plotted against the parameter α, which is equivalent to $1/w$ in our nomenclature. Table 1 contains approximate values of $w = 1/\alpha$ for which the cutoff frequency has assumed its minimum. This data is extracted from Fig. 3 in [20]. By taking advantage of the *a priori* estimation $\tilde{w} = 1/\sqrt{2M}$, the need to numerically search for the minimum point is obviated. Putting away the nonlinear minimization procedure thanks to this *a priori* approximation, only a simple linear eigenvalue problem remains to be solved.

6. Summary and Conclusion

We have introduced a systematic approach to extract the asymptotic behavior of the optimum parameter for the basis functions complying with several general constraints. The result has been worked out for the waist of HG and LG bases. For functions including discontinuity in their profiles or their derivatives, an explicit estimation of the optimum waist has been proposed. This estimation extends the knowledge and observations on individual discontinuities to multiple discontinuity functions.

Although our discussion has been focused on expansion of known functions, partial information from the essence of the function to be expanded can be adequate to conjecture the optimum waist. Especially when the dominant error of the whole problem emanates from expansion error (e.g., as for variational formulations), the knowledge about discontinuity positions in the solution or its higher-order derivatives can be of assistance.

As demonstrated by miscellaneous examples, the estimation is beneficiary in many practical cases. However, it fails for spectrally sparse or band-limited functions. Another difficulty is expected to arise if index-hopping occurs. We have not observed a major change in the index of the optimum zero in waist optimization for the HG (or LG) basis. However, this is not the case if some other parameters, such as the basis chirp factor (as defined in [22]) are to be optimized. The chirp factor can be considered as a basis parameter, like waist, and can be treated in the same way. Our primitive experiments on the chirped HG basis have shown remarkable alterations in optimum index.

Appendix A

We assumed that the set of basis functions $\{\psi_m\}_{m=0}^{\infty}$ is complete in $\mathbf{S}$. Hence if we assume that the derivative of each basis function, $\partial\psi_m/\partial w$, lies in S, the following differentiation operator can be defined in a matrix form:

$$\frac{\partial}{\partial w}\boldsymbol{\Psi} = \mathbf{D}_w\boldsymbol{\Psi}, \qquad \boldsymbol{\Psi} = [\psi_i], \tag{A1}$$

where $\boldsymbol{\Psi}$ is a column vector containing all basis functions. $\mathbf{D}_w$ is an infinite matrix and is partitioned as follows:

$$\mathbf{D}_w = \begin{bmatrix} \mathbf{D}_M & \mathbf{D}_r \\ \mathbf{D}_s & \mathbf{D}_p \end{bmatrix}. \tag{A2}$$

Here, $\mathbf{D}_M$ is an M-by-M block of the infinite matrix $\mathbf{D}_w$. On the other hand, $\mathbf{B}$ is an infinite column vector whose entries represent $\langle f, \psi_M\rangle$ and can be similarly partitioned as

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_M \\ \mathbf{B}_r \end{bmatrix}, \tag{A3}$$

Here, $\mathbf{B}_M$ is the M-by-1 subvector of the infinite column vector $\mathbf{B}$.

Appendix B

The partial derivative of the Gram matrix, $\partial G/\partial w$, can be expressed as

$$\mathbf{G} = \langle \boldsymbol{\Psi}, \boldsymbol{\Psi}^t\rangle, \tag{B1}$$

$$\Rightarrow \frac{\partial \mathbf{G}}{\partial w} = \mathbf{D}_w\mathbf{G} + \mathbf{G}\mathbf{D}_w^T. \tag{B2}$$

The basis is orthonormal so that the norm of basis functions $\psi_m(\bar{\mathbf{r}}, w)$ is independent of w. Hence, the partial derivative of the Gram matrix, which is independent of w, is null. Consequently, multiplying Eq. (B2) from both sides by $\mathbf{G}^{-1}$ results in

$$\mathbf{G}^{-1}\mathbf{D}_w + \mathbf{D}_w^T\mathbf{G}^{-1} = 0. \tag{B3}$$

It is worth noting that, for an orthonormal basis, the Gram matrix is unitary, and the above-mentioned equation reveals the antisymmetric essence of a partial derivative matrix, i.e., the fact that $d_{mn} = -d_{nm}^*$. Furthermore, $\mathbf{G}^{-1}$ is unitary and we can truncate the matrices:

$$\mathbf{G}_M^{-1}\mathbf{D}_M + \mathbf{D}_M^T\mathbf{G}_M^{-1} = 0. \tag{B4}$$

References

1. R. Borghi, F. Gori, and M. Santarsiero, “Optimization of Laguerre–Gauss truncated series,” Opt. Commun. **125**, 197–203 (1996).
2. Y. Liu and B. Lu, “Truncated Hermite–Gauss series expansion and its application,” Optik (Jena) **117**, 437–442 (2006).

3. A. C. den Brinker and H. J. W. Belt, “Optimal free parameters in orthonormal approximations,” IEEE Trans. Signal Process. **46**, 2081–2087 (1998).
4. P. Lazaridis, G. Debarge, and P. Gallion, “Discrete orthogonal Gauss–Hermite transform for optical pulse propagation analysis,” J. Opt. Soc. Am. B **20**, 1508–1513 (2003).
5. F. Chiadini, G. Panariello, and A. Scaglione, “Variational analysis of matched-clad optical fibers,” J. Lightwave Technol. **21**, 96–105 (2003).
6. T. Rasmussen, J. H. Povlsen, A. Bjarklev, O. Lumholt, B. Pedersen, and K. Rottwitt, “Detailed comparison of two approximate methods for the solution of the scalar wave equation for a rectangular optical waveguide,” J. Lightwave Technol. **11**, 429–433 (1993).
7. I. A. Erteza and J. W. Goodman, “A scalar variational analysis of rectangular dielectric waveguides using Hermite–Gaussian modal approximations,” J. Lightwave Technol. **13**, 493–506 (1995).
8. R. L. Gallawa, I. C. Goyal, and A. K. Ghatak, “Fiber spot size: a simple method of calculation,” J. Lightwave Technol. **11**, 192–197 (1993);
9. Z. H. Wang and J. P. Meunier, “Comments on ‘Fiber spot size: a simple method of calculation,’” J. Lightwave Technol. **13**, 1593 (1995).
10. R. L. Gallawa, I. C. Goyal, and A. K. Ghatak, “Optical waveguide modes: an approximate solution using Galerkin’s method with Hermite–Gauss basis functions,” IEEE J. Quantum Electron. **27**, 518–522 (1991).
11. J. P. Meunier, Z. H. Wang, and S. I. Hosain, “Evaluation of splice loss between two non identical single mode graded fibers,” IEEE Photonics Technol. Lett. **6**, 998–1000 (1994).
12. D. G. Dudley, *Mathematical Foundations for Electromagnetic Theory* (Oxford U. Press, 1994).
13. G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists* (Academic, 2005).
14. S. Elaydi, *An Introduction to Difference Equations* (Springer, 2005).
15. M. Martinelli and P. Martelli, “Laguerre mathematics in optical communications,” Opt. Photonics News **19**, 30–35 (2008).
16. T. M. Monro, D. J. Richardson, N. G. R. Broderick, and P. J. Bennett, “Holey optical fibers: an efficient modal model,” J. Lightwave Technol. **17**, 1093–1102 (1999).
17. S. Guo and S. Albin, “Comparative analysis of Bragg fibers,” Opt. Express **12**, 198–207 (2004).
18. V. García-Muñoz and M. A. Muriel, “Hermite–Gauss series expansions applied to arrayed waveguide gratings,” IEEE Photonics Technol. Lett. **17**, 2331–2333 (2005).
19. R. McDuff, “Matrix method for beam propagation using Gaussian Hermite polynomials,” Appl. Opt. **29**, 802–808 (1990).
20. A. Sharma and J. P. Meunier, “Cutoff frequencies in planar optical waveguides with arbitrary index profiles: an efficient numerical method,” Opt. Quantum Electron. **34**, 377–392 (2002).
21. A. Sharma and J-P. Meunier, “On the scalar modal analysis of optical waveguides using approximate methods,” Opt. Commun. **281**, 592–599 (2008).
22. P. Lazaridis, G. Debarge, and P. Gallion, “Exact solutions for linear propagation of chirped pulses using a chirped Gauss–Hermite orthogonal basis,” Opt. Lett. **22**, 685–687 (1997).