Theoretical Investigation of Iterative Phase Retrieval Algorithm for Quasi-Optical Millimeter-Wave RF Beams

Sudheer Jawla, *Student Member, IEEE*, Jean-Philippe Hogge, and Stefano Alberti

Abstract—In this paper, we present a detailed analysis of the iterative phase retrieval approach (IPRA) for determining the phase profile of the output microwave beam of a gyrotron from known intensity patterns emphasizing the field propagation techniques which are used to propagate the RF field of the microwave beam between known intensity planes. The propagation method, based on first Rayleigh–Sommerfeld diffraction integral (RSDI), is solved using fast Fourier transform (FFT) technique and zero padding. It is observed that the use of FFT and, therefore, the discretization of the RSDI propagation kernel introduce aberrations in the propagated field due to the superposition of the original field with its replicated versions. This problem is solved by approximations leading to the Huygens–Fresnel propagation method which further imposes the restrictions on the distances of propagation depending on the size of the transverse plane used to discretize the intensity pattern. This constraint of the distance of propagation causes problem in the iterative phase retrieval approach (IPRA) when more than two intensity planes are used. A method based on interpolation is proposed to overcome this restriction. IPRA is then further discussed to optimize several parameters, such as plane separation, plane dimension, mesh size, and measurement accuracies, which become more of an issue during the measurements of infrared intensity thermograms of the output microwave beam.

Index Terms—Fourier transform, gyrotron, infrared intensity thermogram, millimeter wave, phase retrieval, zero padding.

I. INTRODUCTION

MAGNETICALLY confined fusion devices currently operating require high millimeter-wave power for electron cyclotron resonance heating in their plasma fusion experiments. Gyrotrons are the sources of this RF power. Microwaves generated from a gyrotron in high-order waveguide modes must be transformed into a low-order symmetrical mode for efficient transmission. The output microwave beam transformed in the form of fundamental Gaussian beam of proper waist size is then injected into a corrugated waveguide to maximize the coupling to an HE_{11} mode. However, small deviations from the optimized Gaussian beam profile can significantly increase the coupling to unwanted higher order modes. Therefore, it is required to have knowledge of the RF field properties in terms of determining the amplitude and phase profiles. Since the phase measurements are not available in such high-power applications, one has to rely on intensity measurements to calculate the phase profile [1]–[7]. The intensity measurements can be performed using an infrared thermography technique by irradiating a target material with the high-power microwave beam and measuring the spatial temperature elevation (which is proportional to the spatial RF power profile of the microwave beam) on the material using an infrared camera.

Although a number of methods [2], [7]–[9] have been studied for phase reconstruction, the error reduction approach seems to be most accurate in general and in particular for high-power gyrotrons. Other methods, such as the transport of intensity equation (TIE) [9], solve the real part of the wave equation, considering it as an elliptic partial differential equation. The main difficulty with this equation is its boundary behavior because the value of phase at the boundary is difficult to measure. Moreover, a proper phase solution must satisfy the other part of the wave equation. Furthermore, in TIE, an accurate estimate of the intensity derivative along the direction of propagation needs closely spaced intensity measurements (no specific criterion for the separation between the planes is defined, and at millimeter wavelength, there is no significant difference between very closely spaced planes). The moment irradiance approach [7] works well only if the beam pattern contains a minimal fraction of higher order modes.

The motivation of this paper is the development of an appropriate methodology to calculate the accurate phase profile of a microwave beam from measured intensity thermograms. For this purpose, IPRA is studied in more detail using theoretical beam profiles of the 170-GHz/2-MW coaxial cavity gyrotron [10]–[12], emphasizing more on an accurate field-propagation scheme and on some basic parameters, such as scan separation, plane dimension, number of planes used in IPRA, measurement accuracies (target material and its thickness), etc.

This paper is organized as follows. Section II describes, briefly, the IPRA. Section III discusses, in detail, the wave propagation schemes by calculating the full Rayleigh–Sommerfeld diffraction integral (RSDI) and analyzing the Huygens–Fresnel propagation picture. Section IV provides a thorough discussion on optimized parameters and results. The conclusions are presented in Section V.

The authors are with the Centre de Recherches en Physiques des Plasmas, Association Euratom–Confédération Suisse, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: sudheerkumar.jawla@epfl.ch; jean-philippe.hogge@epfl.ch; stefano.alberti@epfl.ch).

II. IPRA

The IPRA for phase reconstruction, as proposed by Katsenelenbaum and Semenov [1], is outlined briefly, as follows. The
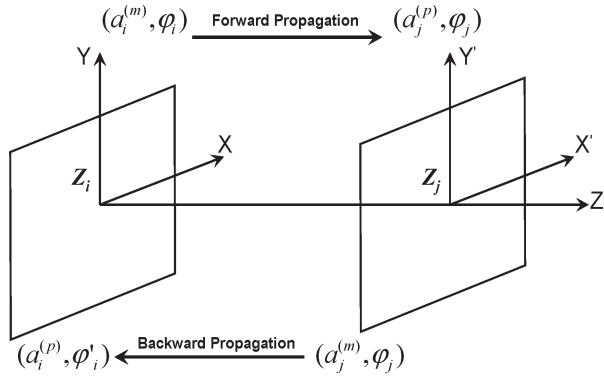
Fig. 1. Iterative method for phase reconstruction is shown. Field is propagated from z_i to z_j, and then, it is back propagated to the initial plane. Φ_i is the initial phase.

field of an electromagnetic wave at any ith plane, with position z_i, transverses to the direction of propagation of the wave and can be written as

$$f_i(x,y) = a_i(x,y) \exp[i\phi_i(x,y)] \qquad (1)$$

where $a_i(x,y)$ and $\varphi_i(x,y)$ are the wave amplitude and phase at the ith plane, respectively. The field $f_j(x',y')$ at a later distance z_j can be written as

$$f_j(x',y') = \int\!\!\int_\Omega f_i(x,y) P(x'-x, y'-y) dx\, dy \qquad (2)$$

where (x,y) represents the input plane coordinates and (x',y') represents the output plane coordinates at z_i and z_j, respectively. $P(x-x', y-y')$ is the propagator used to propagate the field from the plane at z_i to z_j. One can define P either as a Huygens–Fresnel propagator, which is a quadratic phase function, or a Fraunhofer propagator in the far field or a Green function as defined in the Rayleigh–Sommerfeld diffraction theory [15]. The integral is performed over the size Ω (domain in the real space) of the transverse plane. The measured field intensity, with an initial guess of phase at z_i, is propagated to the next plane at z_j, where the propagated intensity is replaced by the measured intensity and then back propagated to the previous plane location. This iteration method is shown in Fig. 1, where the superscripts m and p denote the measured and propagated field amplitudes in two planes z_i and z_j, respectively.

The criterion to terminate the iterative process is expressed as

$$E_{i,n} = \int\!\!\int_{\Omega_i} \left| a_i^{(m)} - a_{i,n}^{(p)} \right|^2 d\Omega_i < \varepsilon_i \qquad (3)$$

where $E_{i,n}$ means the error value after the nth iteration and ε_j is a set of given small values. One can show that $E_{i,n}$ is monotonically decreasing with n [1], [5]. The approach can be easily extended to any number of cross sections. The criterion to stop the iteration is discussed, in more detail, in Section IV.

III. FIELD-PROPAGATION METHODS

The success of the aforementioned iterative algorithm depends on the field-propagation scheme used. Therefore, it is

more appropriate to solve the full RSDI which is the most accurate and complete description of the propagation in scalar diffraction theory. We solve the first RSDI using a Fourier transform convolution method. After that, the effect of discretization on the propagated field is studied in detail when the RSDI propagation kernel is sampled during the Fourier transform approach. The discussion is further extended in detail, leading to the Huygens–Fresnel propagation approach in the paraxial regime where a spherical wave is replaced by a parabolic wave [14], [15]. This discussion is then applied to the IPRA in the next section.

A. RSDI Calculation

To find the field at any point in the observation plane, one approaches the scalar diffraction theory and solves the full RSDI given by [14], [15]

$$f_j(x',y') = \frac{1}{2\pi} \int\!\!\int_\Omega f_i(x,y) \frac{\partial}{\partial z}\left(\frac{\exp(ikr)}{r}\right) dx\, dy \qquad (4)$$

where

$$r = \left[(x'-x)^2 + (y'-y)^2 + (\Delta z)^2 \right]^{\frac{1}{2}}$$

and k is the wavenumber given in terms of wavelength λ by $k = 2\pi/\lambda$ and $\Delta z = (z_j - z_i)$. The integration is over the region Ω in the input plane z_i, where $f_i(x,y)$ is nonzero. The area outside the region does not contribute to the integral, according to the boundary condition imposed on the problem. The diffraction integral in (4) can be understood in the form of a linear convolution as

$$f_j(x',y') = \int\!\!\int_\Omega f_i(x,y) g_{i\to j}(x'-x, y'-y) dx\, dy \qquad (5)$$

where

$$g_{i\to j}(x'-x, y'-y) = \frac{1}{2\pi} \frac{\partial}{\partial z}\left(\frac{\exp(ikr)}{r}\right) \qquad (6)$$

or, equivalently

$$g_{i\to j}(x'-x, y'-y) = \frac{2\Delta z}{r}\left(ik - \frac{1}{r}\right)\frac{\exp(ikr)}{4\pi r}. \qquad (7)$$

If the aforementioned equation is discretized on a grid of size $P \times Q$, then, for any point (x',y') in the observation plane z_j, one can write

$$f_j(x',y') = \sum_{p=1}^{P}\sum_{q=1}^{Q} f_i(x_p, y_q) g_{i\to j}(x'-x_p, y'-y_q)\Delta x \Delta y. \qquad (8)$$

This corresponds to a discrete linear convolution of f and g. Since the Fourier transform operator is linear, we can write

$$\begin{aligned} f_j(x',y') &= \mathcal{F}^{-1}\{\mathcal{F}\{f_j(x',y')\}\} \\ &= \mathcal{F}^{-1}\{\mathcal{F}\{f_i(x,y)\} \times \mathcal{F}\{g_{i\to j}(x'-x_p, y'-y_q)\}\}. \end{aligned} \qquad (9)$$

Therefore, one can use fast Fourier transform (FFT) to calculate this convolution. Although this convolution form is numerically easy for the calculation of the diffraction integral as compared with a direct integral approach, on the other hand, this method also provides some constraints in terms of distance of propagation, plane size, and the number of sampling points use to discretize a given scalar field. This is discussed in the following.

B. Discretization Analysis of RSDI and Huygens–Fresnel Propagation Picture

Based on a well-established plane wave decomposition technique, the RSDI propagation picture (solved in the previous section) can always be represented as a linear shift invariant system. Since the propagation of the wave field to any further plane involves a Fourier transform system, we therefore need to analytically study the discretization process of the field and its propagation in the form of a band-limited signal because one can easily verify that the Fourier transform of a discretized image is, within a scale factor, a periodic replication of the Fourier transform of the input image on a sampling grid. Following the usual procedure of sampling [15], [16], [18] for a band-limited function $f_j(x', y')$ and its recovery from its sampled version, we sample $f_j(x', y')$ by a 2-D array of Dirac delta functions situated on a rectangular grid with spacing $\Delta x'$ and $\Delta y'$ as

$$f_{js}(x', y') = \sum_{m=-M}^{M} \sum_{n=-N}^{N} f_j(m\Delta x', n\Delta y')\delta(x' - m\Delta x', y' - n\Delta y') \tag{10}$$

where the subscript "s" represents the sampled version of the field $f_j(x', y')$ and (M, N) represents the number of subdivisions in the x- and y-directions, respectively. Passing through a Fourier transform system, one can write

$$
\begin{aligned}
f_{js}&(x', y') \\
&= \mathcal{F}^{-1}\{\mathcal{F}\{f_{js}(x', y')\}\} \\
&= \mathcal{F}^{-1}\left\{(\nu_{x's}\nu_{y's}) \sum_{m=-M}^{M} \sum_{n=-N}^{N} F_j(\nu_{x'} - m\nu_{x's}, \nu_{y'} - n\nu_{y's})\right\}
\end{aligned}
\tag{11}
$$

where the sampling frequencies $\nu_{x's}$ and $\nu_{y's}$ are, respectively, the grid spacing in the x- and y-directions given by $\nu_{x's} = 1/\Delta x'$ and $\nu_{y's} = 1/\Delta y'$ and F_j is the Fourier transform of $f_j(x', y')$, which is given by

$$F_j(\nu_{x'}, \nu_{y'}) = F_i(\nu_x, \nu_y) \times G_{i\to j}(\nu_{x'-x}, \nu_{y'-y}). \tag{12}$$

Here, F_i is the Fourier transform of $f_i(x, y)$, and $G_{i\to j}(\nu_{x'-x}, \nu_{y'-y})$ is the Fourier transform of the RSDI propagation kernel given by the transfer function of the free space [15], using (6), as

$$\mathcal{F}\{g_{i\to j}(x'-x, y'-y)\} = \exp\left(2\pi i\Delta z\left(\frac{1}{\lambda^2} - \nu_x^2 - \nu_y^2\right)^{1/2}\right). \tag{13}$$

Therefore, using (12) and (13) and for the same grid spacing in both input and output planes, i.e., $\Delta x' = \Delta x$ and $\Delta y' = \Delta y$, one can write the sampled version of the propagated field in the output plane at z_j as

$$
\begin{aligned}
&f_{js}(x', y') \\
&= \mathcal{F}^{-1}\left\{
\begin{aligned}
&(\nu_{xs}\nu_{ys}) \sum_{m=-M}^{M} \sum_{n=-N}^{N} F_i(\nu_x - m\nu_{xs}, \nu_y - n\nu_{ys}) \\
&\times \exp\left(2\pi i\Delta z\left(\tfrac{1}{\lambda^2} - (\nu_x - m\nu_{xs})^2 - (\nu_y - n\nu_{ys})^2\right)^{\frac{1}{2}}\right)
\end{aligned}
\right\}.
\end{aligned}
\tag{14}
$$

As we know that, if the Nyquist sampling theorem is fulfilled, one can always recover the continuous image from the sampled images [16] using a low-pass filter with a frequency response given by $1/(\nu_{xs}\nu_{ys})$. Therefore, from this point onwards, we can remove the multiple factor of $(\nu_{xs}\nu_{ys})$ and write the field as continuous, instead of the sampled field. The back-propagated field can be written in a similar way by convolving the field in the output plane with the back-propagation kernel which is equal to the complex conjugate of the forward RSDI propagation kernel. The back-propagated field is therefore written as

$$
\begin{aligned}
f_i(x, y) &= \iint_\Omega f_j(x', y')g_{j\to i}(x - x', y - y')dx'dy' \\
&= \mathcal{F}^{-1}\{\mathcal{F}\{f_j(x', y')\} \times \mathcal{F}\{g_{j\to i}(x - x', y - y')\}\}
\end{aligned}
\tag{15}
$$

where $g_{j\to i} = g^*_{i\to j}$. Therefore, using (14) and (15), one can write the back-propagated field as

$$
\begin{aligned}
f_i(x, y) = \mathcal{F}^{-1}\Bigg\{ &\sum_{m=-M}^{M} \sum_{n=-N}^{N} F_i(\nu_x - m\nu_{xs}, \nu_y - n\nu_{ys}) \\
&\times \exp(2\pi i\Delta z\psi(\nu_x - m\nu_{xs}, \nu_y - n\nu_{ys})) \\
&\times \exp(-2\pi i\Delta z\psi(\nu_x, \nu_y)) \Bigg\}
\end{aligned}
\tag{16}
$$

where

$$
\begin{aligned}
\psi(\nu_x - m\nu_{xs}, \nu_y - n\nu_{ys}) &= \big(1/\lambda^2 - (\nu_x - m\nu_{xs})^2 \\
&\quad - (\nu_y - n\nu_{ys})^2\big)^{1/2}
\end{aligned}
\tag{17a}
$$
$$\psi(\nu_x, \nu_y) = \big(1/\lambda^2 - \nu_x^2 - \nu_y^2\big)^{1/2}. \tag{17b}$$

Equations (16), (17a), and (17b) indicate that the back-propagated field $f_i(x, y)$ is a modified version of the original field modulated by the functions $\psi(\nu_x, \nu_y)$ and $\psi(\nu_x - m\nu_{xs}, \nu_y - n\nu_{ys})$. Equation (16) can further be simplified in a more compact form using the Taylor series expansion for the RHS of (17a) as

$$
\begin{aligned}
&\psi(\nu_x - m\nu_{xs}, \nu_y - n\nu_{ys}) \\
&= \psi(\nu_x, \nu_y)\left[1 - \frac{1}{2}b + \sum_{r=2}^{\infty} \frac{(-1)^r(2r)!}{(1 - 2r)(r!)^2 4^r}(-b)^r\right]
\end{aligned}
\tag{18}
$$

where

$$b = \frac{1}{\psi(\nu_x, \nu_y)} \left[(m\nu_{xs})^2 + (n\nu_{ys})^2 - 2(m\nu_x\nu_{xs} + n\nu_y\nu_{ys}) \right].$$
(19)

Using (18) and (19) in (16), we can write the discretized version of the back-propagated field when passed through a Fourier transform system. This can be written as in (20), which is shown at the bottom of the page, i.e., the back-propagated field is now a modulated version of the original input field. As one can see, (20) is difficult to solve accurately because of the higher order terms which generate aberration like effects; therefore, we will move to a fairly accurate solution of this equation by considering a small angle diffraction, which means that $1/\lambda \gg \nu_x, \nu_y$ for all contributing frequencies, and, therefore, by retaining terms only up to the parabolic terms, which leads us to the Huygens–Fresnel propagation approach. Nevertheless, an optimized solution of (20), with higher order terms in the expansion of the exponential factor, i.e., aberration terms, could be obtained by optimizing the parameters, e.g., size of the diffracting plane, number of sampling points, wavelength, etc. This will be the topic of a future study. In the following discussion, we will focus on the discretization analysis of the approximated solution of the earlier discussed RSDI, i.e., the Huygens–Fresnel propagation method.

To simplify (20), we retain terms up to the parabolic term; therefore, using (19), along with the Fourier shift theorem, one can write the back-propagated field in the parabolic approximation as

$$f_i(x, y) = \sum_{m=-M}^{M} \sum_{n=-N}^{N} f_i(x - \Delta z\lambda m\nu_{xs}, y - \Delta z\lambda n\nu_{ys})$$
$$\times \exp\left\{ -i\pi\Delta z\lambda \left(m^2\nu_{xs}^2 + n^2\nu_{ys}^2 \right) \right\}. \quad (21)$$

It is quite clear from (21) that the recovered back-propagated field in the Fresnel approximation is a modulated and replicated version of the original field. The process of discretization has introduced a periodicity in the back-propagated input field with a period of $(z\lambda\nu_{xs}, z\lambda\nu_{ys})$ in the x- and y-directions, respectively, and modulated it by a factor of

$$\exp\left\{ -i\pi\Delta z\lambda \left(m^2\nu_{xs}^2 + n^2\nu_{ys}^2 \right) \right\}. \quad (22)$$

To countercheck the accuracy of (21), one can easily derive a similar equation following the equivalent discretization process if one starts the analysis of the Huygens–Fresnel propagation method, where the convolution kernel is a quadratic phase function $h_{i \to j}(x, y)$ [15], [19] instead of a Green's function derivative (6) and is defined as

$$h_{i \to j}(x, y) = \frac{1}{i\lambda\Delta z} \exp\left\{ i2\pi\frac{\Delta z}{\lambda} \right\} \exp\left\{ i\pi\frac{x^2 + y^2}{\lambda\Delta z} \right\}. \quad (23)$$

Convolving this quadratic phase function with the input field and following the steps of (9)–(16), one reaches the equation similar to (21).

Coming back to (21), which is a modulated version of the original input field, for a square grid with Δx equal to Δy, if the grid spacing is chosen such that

$$(\nu_{xs})^2 = 2a/\lambda\Delta z \quad (24)$$

where a is an integer, then the modulating factor in (22) can be made equal to one for all (m, n) and the discrete version of the back-propagated field (21) is therefore dependent on the parameter a (an integer which represents the replication process) only for a scalar factor and a possible sign change. If $2a$ is made equal to the total number of sampling points M in both directions, i.e., $2a = M$, then (24) gives a condition for the distance of propagation as

$$\Delta z = \frac{M(\Delta x)^2}{\lambda} \Rightarrow M = \frac{L^2}{\lambda\Delta z} \quad (25)$$

where L is the transverse plane dimension. This equation implies that, if the number of sampling points M is chosen equal to the Fresnel number, then (21) will give the back-propagated field without replication and modulation provided that the transverse plane dimensions are the same at both positions. Therefore, ideally, one should propagate the field to a distance given by (25) in order to avoid the effects of aliasing. Fig. 2 shows a clearer picture of (25) in terms of distance of propagation, number of subdivisions, and the plane dimension. It is clear from Fig. 2 that, for larger plane dimensions (to have a negligible intensity at the boundary), we need a large number of sampling points to propagate the field to closer distances, and clearly, this will increase the computational time of the

$$f_i(x, y) = \mathcal{F}^{-1} \left\{ \sum_{m=-M}^{M} \sum_{n=-N}^{N} \underbrace{F_i(\nu_x - m\nu_{xs}, \nu_y - n\nu_{ys})}_{\text{translation}} \times \underbrace{\exp\left(-\pi i\Delta z \times \psi(\nu_x, \nu_y) \times b \right)}_{\text{parabolic term}} \right.$$

$$\left. \times \underbrace{\exp\left[2\pi i\Delta z \times \psi(\nu_x, \nu_y) \sum_{r=2}^{\infty} \frac{(-1)^r(2r)!}{(1-2r)(r!)^2 4^r}(-b)^r \right]}_{\text{aberration term}} \right\}. \quad (20)$$
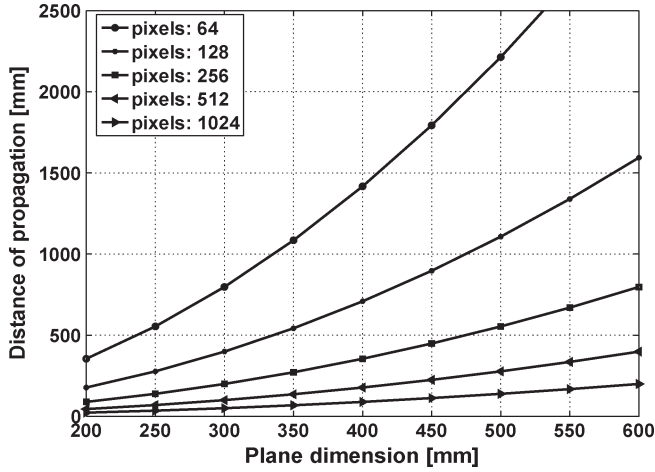
Fig. 2. Distance of propagation given by (25) is plotted w.r.t. the plane dimension for different numbers of sampling points.

IPRA algorithm when we propagate the field for more than one time during the minimization of the reconstruction error (3). Moreover, we want to be in the Rayleigh range for the Fresnel approximation to be more accurate in the paraxial region.

In our case of the 170-GHz/2-MW gyrotron output beam, the beam waist of the field calculated from the simulations [12] is 25.4 mm, which gives the Rayleigh range Z_R, equal to 1150 mm calculated by $Z_R = \pi w_0^2 / \lambda$, where w_0 is the radius of the beam at its waist. Therefore, we have to make a compromise between the chosen number of sampling points and the plane size for a given distance of propagation, keeping in mind that we could fulfill the constraint on the intensity (to be zero at the boundary). We should have a large number of sampling points if we choose a larger plane size (if one measures the field profile farther away from its beam waist) in complying with the propagation distance given by (25) and in order to remain in the Rayleigh range; this is computationally demanding. Therefore, it is always preferable to measure the beam field profile close to the beam waist location.

In the case of the Huygens–Fresnel propagation scheme, if one wants to propagate a given field to arbitrary distances which are different from the distances given in (25), then one can follow two approaches, which can still act in accordance with the condition of propagation given by (25).

1) The field can be propagated to the distance given by (25), and then, it can be repeated again until the desired value of z is reached. However, this approach is time consuming.

2) For one-time propagation for a desired z value and fixed plane size L, one can interpolate the field to a grid size that corresponds to (25). However, this arbitrary z value can lead to a noninteger M value, and if we take the closest integer value for M, this difference of M, which prevents it from being an integer, is equivalent to the difference in the given z value. Therefore, we are violating (25), and this can introduce aliasing in the field amplitude and phase.

This is further discussed in detail, in view of IPRA, in the next section.

IV. DISCUSSION ON IPRA

Since IPRA is an error-minimization method, we can optimize several parameters to achieve a good compromise between the final results and the computational efforts. We discuss the effects of various parameters, which can either be altered during the computations or can be optimized during the free-space infrared measurements of the output microwave beam of the gyrotron. The main issues which affect the reconstruction algorithm and the accuracy of the result, which need to be incorporated during the measurements, measured data processing, and phase retrieval algorithm, are:

1) field-propagation method;
2) plane separation;
3) size of the planes and number of sampling points used;
4) position of the beam profile from its waist;
5) number of planes used in the algorithm;
6) initial phase guess during the computations;
7) dynamic range and the alignment of the consecutive beam profiles.

Few of these parameters can be given attention during the beam profile measurements; for example, the dynamic range (by the appropriate choice of the target material), plane separation and their size, etc., can be optimized during the reconstruction algorithm, such as the initial phase guess, number of planes, etc.

A. Field-Propagation Method and Its Effect on IPRA

1) Huygens–Fresnel Propagation Method: As we have explained in Section II, the IPRA requires the propagation and back propagation of the wave field between two or more number of planes until the algorithm converges to a global minimum. Therefore, we discussed the wave propagation methods in the previous section. Here, we further extend our explanation of the Huygens–Fresnel propagation method, emphasizing on IPRA. For this purpose, we take the theoretical beam profile of the 170-GHz/2-MW coaxial gyrotron at any arbitrary position (500 mm from the gyrotron window). We propagate this field to the different positions and back propagate it at the same position in order to see the effect of (25). Fig. 3 shows the electric field amplitude and phase at 500 mm from the gyrotron window for three different cases of distance of propagation and back propagation. Fig. 3(a) shows the theoretical field amplitude and phase at 500 mm from a window plane. Fig. 3(b)–(d) shows the back-propagated field at the same position corresponding to $z = t(M \Delta x^2)/\lambda$, where $t = (1.2, 0.9, 1)$. Here, Δx is the grid spacing corresponding to a square plane size of 450 mm $\times$ 450 mm, where λ is the wavelength. The contours of constant field amplitude are at every -3 dB from the peak value. As shown in the amplitudes and phase patterns in Fig. 3 for the values of z, which do not follow (25), i.e., when $t = 1.2$ and 0.9, the field profile is corrupted by the foldover frequencies which are introduced in the field by the process of modulation and replication, as we have discussed in the previous section. As shown in Fig. 3(d), where we follow (25), i.e., $t = 1$, we are able to remove the effects of modulation, and hence, we get exactly the same filed profile as that of the input field.
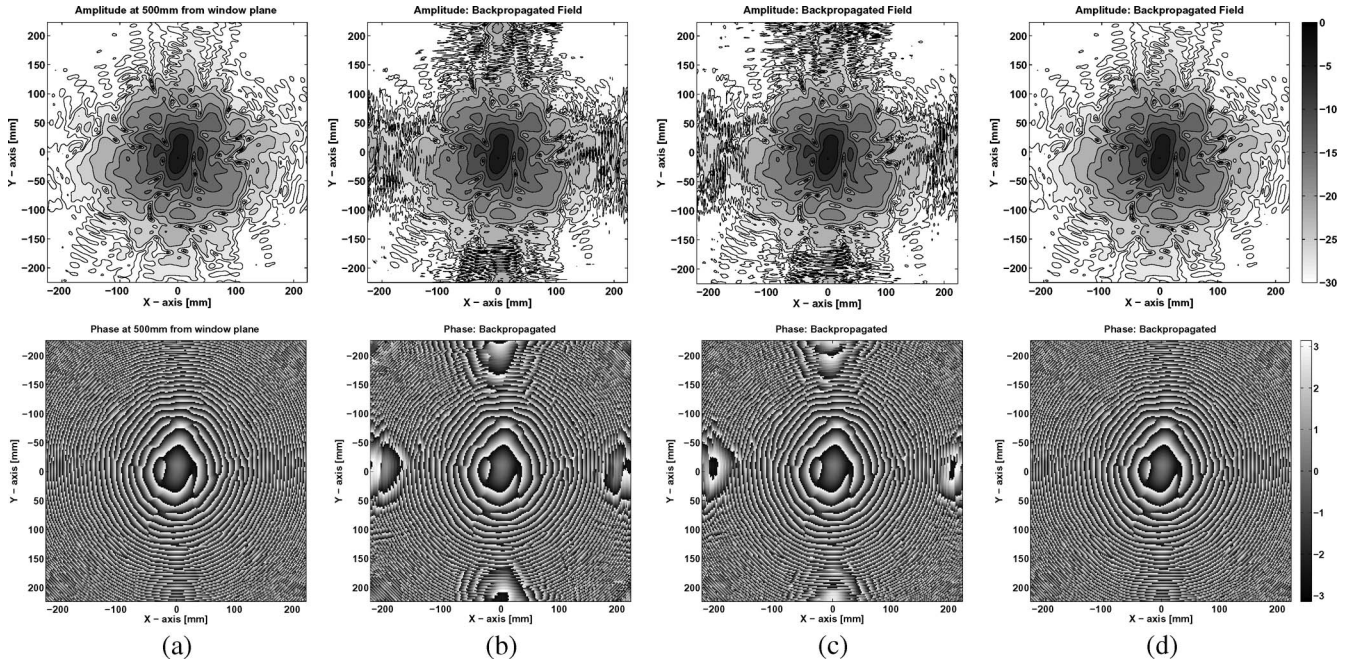
Fig. 3. (a) Amplitude and phase of the output Gaussian beam of the 170-GHz/2-MW gyrotron at 500 mm from the window. The plane dimension is 450 mm $\times$ 450 mm. (b) Beam is propagated/back propagated corresponding to $z = 1.2(M\Delta x^2)/\lambda = 537.9$ mm, where $M = 256$ is the number of subdivisions in both directions, Δx is the grid spacing, and λ is the wavelength. (c) Beam is propagated/back propagated corresponding to $z = 0.9(M\Delta x^2)/\lambda$. (d) Beam is propagated/back propagated corresponding to $z = (M\Delta x^2)/\lambda$. Isomagnitude contours of field amplitudes are plotted at every -3 dB increment from the peak value. The phase is wrapped in a 2π modulo and shown in radians.

For a better representation of the effect of small differences in the propagation distance z using (25), we take a fundamental Gaussian beam at an arbitrary position (e.g., corresponding to the parameters of the 170-GHz/2-MW gyrotron, i.e., beam waist equal to 25.4 mm at the window plane and frequency equal to 170 GHz) and propagate this field to a distance slightly different from the one given by (25). We can see how sensitive the phase of the propagated/back-propagated field is. Fig. 4 shows the phase profiles of the Gaussian beam field in the case of the Huygens–Fresnel propagation for three different values of z corresponding to $t = (1.001, 1.00001, 1)$. Here, the plane dimension is equal to 300 mm $\times$ 300 mm, and the number of sampling points is chosen equal to 512. Since the plane size is chosen such that the intensity can be considered as zero at the boundary, the amplitude modulation is not significant close to the boundary of the transverse plane; therefore, we only show the wrapped phase, and it is clear from Fig. 4 that the overlapping of the foldover frequencies start reducing as we approach the z value given by (25).

As shown in Fig. 4, the phase is extremely sensitive to z, and both propagated and back-propagated fields are affected. Therefore, during the measurements of the beam profile of the output microwave beam of the gyrotron, either one makes the measurements at specific z positions and use them in IPRA for phase retrieval with the Huygens–Fresnel propagation scheme for a fixed distance condition (25) or makes measurements at random positions along the beam propagation direction and then uses Huygens–Fresnel propagation scheme in IPRA with the propagation done in steps or interpolate the field to a different grid size M which corresponds to (25) for one-time propagation to a fixed distance. However, as we have explained earlier, it is not obvious that the value of M is an integer

[for interpolation to a different grid which complies with (25)] and this difference of M from being an integer is equivalent to a small difference in z, which will bring in the aliasing effects in the propagated/back-propagated field.

Alternatively, in the case of field propagation/back propagation without following (25), the propagated field can be multiplied by an appropriate window function in order to remove the effect of sidelobes, which occurs due to replication; however, in this case, we lose some information of the field close to the boundary. Moreover, it is difficult to determine the exact spatial position of the start over of foldover frequencies due to modulation and replication. In this case, the effect of the window function needs to be studied in advance (depending on the type of the window function) because the wave field is a complex function and the engraving function is a real function. The Huygens–Fresnel propagation method can also be solved using a zero-padding method by padding equal zeros in both the x- and y-directions, as explained in [19]. This procedure is used to remove the effect of foldover frequencies (as shown in Figs. 3 and 4), which gives the overlapping behavior at the boundary (the amount of overlapping depends on the sizes of the convolving kernels). The method of windowing is apparently equal to the zero padding of the propagated/back-propagated field.

Concluding the aforementioned Huygens–Fresnel propagation scheme for IPRA, the following can be stated.

1) If we use two planes in IPRA, we can easily use the Huygens–Fresnel propagation method, following the condition of propagation.

2) For more than two planes in IPRA or to propagate to any arbitrary distance, we can propagate in steps by assuming some intermediate plane. If the distance z is an integer multiple of values given by (25), then we can propagate in
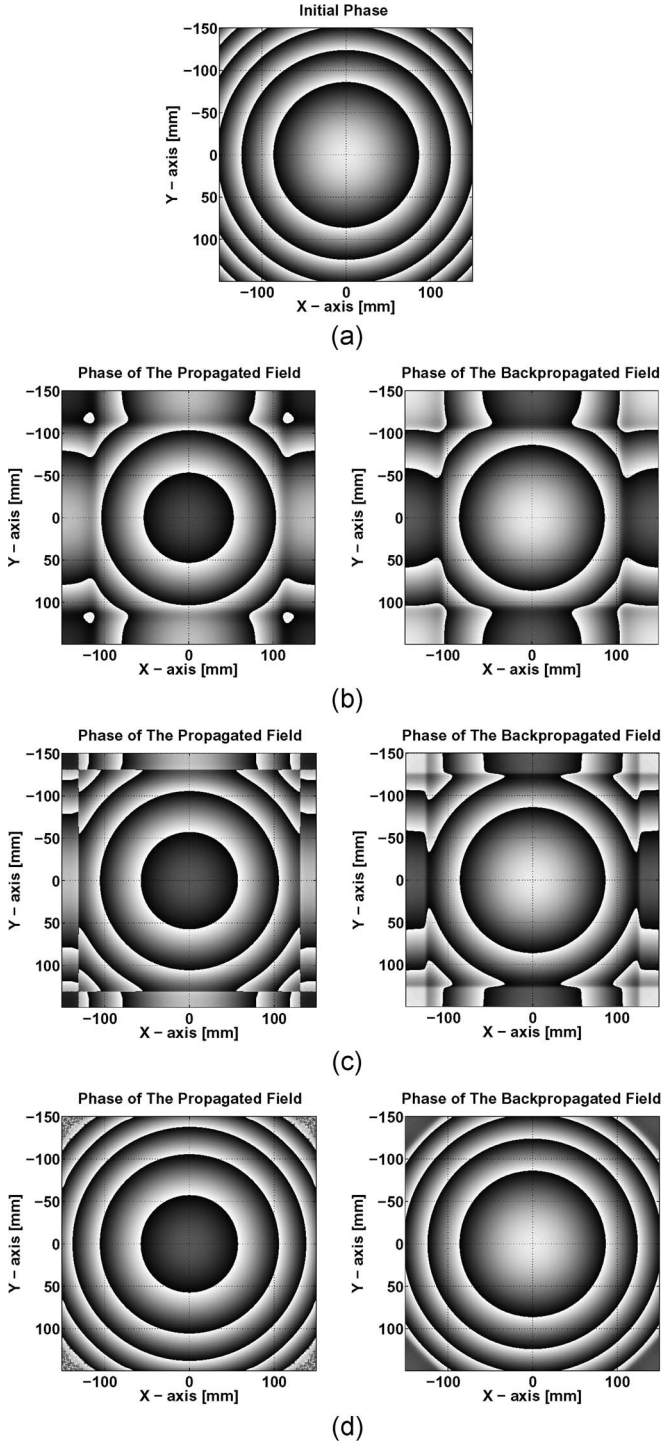
Fig. 4. Phase of a fundamental Gaussian beam at some arbitrary position. The parameters are $w_0 = 25.4$ mm, plane size = 300 mm × 300 mm, and frequency = 170 GHz. (a) Theoretical phase at the initial position. (b), (c), and (d) Propagated and back-propagated fields corresponding to $z = 1.001(M\Delta x^2)/\lambda$, $z = 1.00001(M\Delta x^2)/\lambda$, and $z = (M\Delta x^2)/\lambda$, respectively.

steps by using some virtual intermediate plane. However, if z is not an integer multiple, then we can change the grid size M to match the condition (25). In this case, if M is not an exact integer, then this deviation of M from being an integer is equivalent to a small change in the z position, and we have shown in Fig. 4 that the phase is

extremely sensitive to a tiny change in z. Propagation in steps or interpolation at each iteration will also increase the computational complexity of the algorithm.

Hence, if the beam profile measurements are done at those values of z, which satisfy (25), then the Huygens–Fresnel propagation scheme can be used in IPRA for an exact propagation without computational complexity. For more than two planes, the propagation without any intermediate step or interpolation is valid (following condition of propagation) only when the propagation/back propagation is done among the consecutive planes. For example, when we use three planes in the IPRA, we should follow the propagation procedure, as $1 \rightarrow 2 \rightarrow 3$, and back propagation, as $3 \rightarrow 2 \rightarrow 1$. If the back propagation is followed in a way $3 \rightarrow 1$, then one has to interpolate the field to a new M (which can be a noninteger) at each iteration during the IPRA in order to follow (25), and this can be more time consuming.

2) RSDI Method: Coming back to (9), we calculate the diffraction integral by using a zero-padding method [17]. Since we know that, when we are given two discrete Fourier transforms (DFTs) of finite length, we cannot just multiply them together because DFTs are periodic and they have nonzero values for $x \geq \Omega$; thus, the multiplication of these two DFTs will be nonzero for $x \geq \Omega$. Therefore, (5) gives the circular convolution of f and g instead of a linear convolution [16], [17]. One can easily overcome this difficulty by using the zero-padding technique. If the support of f is $P \times P$ and of g is $Q \times Q$, then the minimum periodicity for both type of convolutions (circular and linear) to lead the same result is $(P + Q - 1) \times (P + Q - 1)$; therefore, in order to avoid the wraparound effect due to circular convolution, we add zeros to the field $f_i(x, y)$ to $(P + Q - 1) \times (P + Q - 1)$ as

$$\begin{bmatrix} [f_i(x,y)]_{P \times P} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{(P+Q-1) \times (P+Q-1)} \quad (26)$$

and extend the propagation kernel $g_{i \rightarrow j}(x' - x, y' - y)$ to a grid size of $(P + Q - 1) \times (P + Q - 1)$ as

$$[g_{i \rightarrow j}(x_{p+q-1}, y_{p+q-1})]_{(P+Q-1) \times (P+Q-1)},$$
$$p = 1, 2, \ldots, P \text{ and } q = 1, 2, \ldots, Q. \quad (27)$$

During the calculations of RSDI, we used the expression of g given in (7). The propagated or back-propagated field in the observation plane can be recovered from the convolution product by retaining the values in the lower right half of the convolved matrix. Fig. 5 shows the propagated and back-propagated fields calculated by using the aforementioned zero-padding technique. The propagated field can be compared directly to the Huygens–Fresnel propagation method shown in Fig. 3. The fields in Fig. 5 are propagated to the same position as in Fig. 3. Fig. 5(b)–(d) can be compared directly with Fig. 3(b)–(d). The dispersion like effects in Fig. 5(b)–(d) can be understood from (20) which contains all higher order terms in the exponential factor. If we compare Fig. 5(d) with Fig. 3(d) of the Huygens–Fresnel propagation, where we could remove all the effects of modulation term by following the condition of propagation (which is also partially valid in case
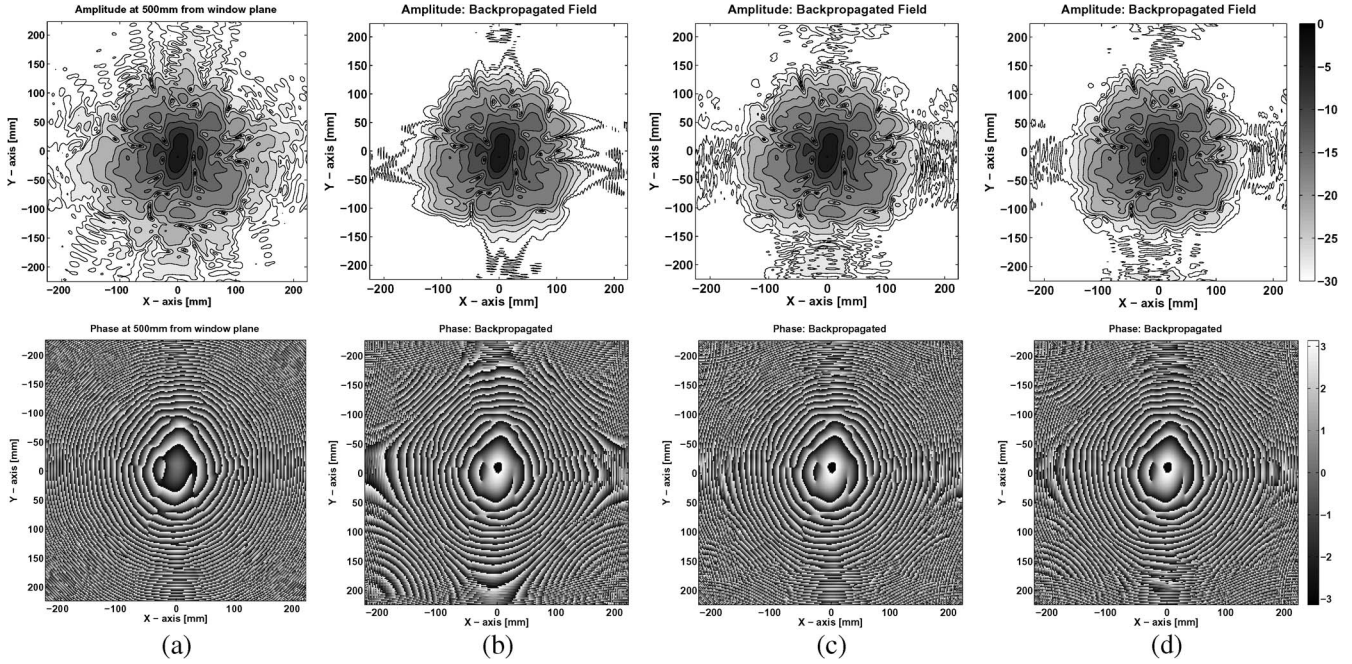
Fig. 5. Propagated and back-propagated fields of the 170-GHz/2-MW gyrotron by solving the full RSDI. (a) Amplitude and phase at 500 mm from the window. (b), (c), and (d) Beam is propagated/back propagated corresponding to $z = 537.9$, 403.4, and 448.2 mm, respectively. (b)–(d) can be directly compared with Fig. 3(b)–(d), respectively. Isomagnitude contours of field amplitudes are plotted at every -3 dB increment from the peak value. The phase is wrapped in a 2π modulo and shown in radians.

of RSDI), some aberration effects can still be seen in Fig. 5(d). Therefore, for an exact propagation scheme, one needs to solve (20) appropriately. Using the RSDI method in IPRA also produces some distortions in the reconstructed phase. Hence, the best possible way for the field-propagation method in IPRA is either using the Huygens–Fresnel propagation method for fixed distances or solving the RSDI (20) by optimizing z, L, and M; other than that will always introduce some distortions in the retrieved phase.

In order to check the accuracy of the back-propagated fields and, hence, the reconstructed fields, we calculate the matching coefficient for initial and back-propagated fields in the form of a cross correlation coefficient given by

$$CCF = \sum_{i,j} u_1(i,j) u_2(i,j). \tag{28}$$

Here, $u_1(i,j)$ is assigned the modulus of the normalized back-propagated field, and $u_2(i,j)$ is the modulus of the normalized theoretical field. The cross correlation function (CCF) is normalized so that $0 \leq CCF \leq 1$. Fig. 6 shows the CCF on a linear scale for different values of the propagation distance in terms of a parameter t defined as $z_{i \to j} = t(M\Delta x^2)/\lambda$. The CCF is shown for two different values of the plane size and for the two methods of propagations, namely, RSDI and Huygens–Fresnel propagation. The field is propagated/back propagated for different values of the parameter "t." The correlation is maximum (equal to one in the case of the Huygens–Fresnel propagation) for t equal to one. For larger values of t which corresponds to the larger distance of propagation and back propagation, the field expands, and it is no longer negligible at the boundary; therefore, the CCF is not accurate for these values because we lose the information at the
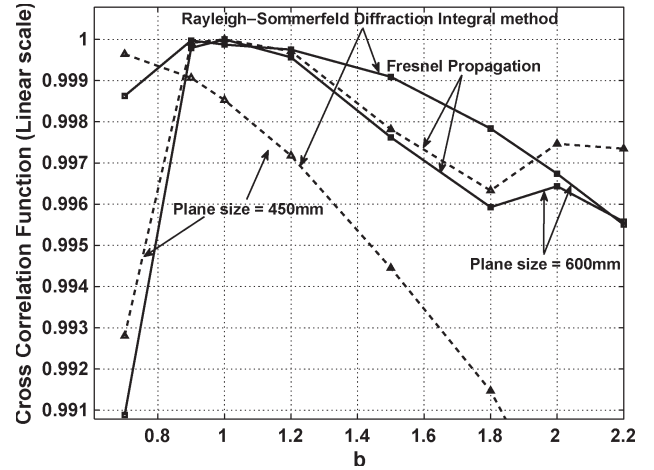

Fig. 6. CCF given by (28) as a function of the parameter "t" for RSDI and Huygens–Fresnel propagation methods.

boundary. Moreover, we do not know the behavior of the higher order terms in case of exact RSDI propagation; therefore, for smaller plane dimensions, the cross correlation coefficient for RSDI propagation deviates from the other values.

B. Effect of Other Parameters and Error Analysis of the Reconstructed Field

If one uses the Huygens–Fresnel propagation method as the field propagation in IPRA, then the relationship between the plane separations, size of the plane, and number of sampling points is clear from (25) and has been discussed in the previous section. One additional condition can be put together with the condition of propagation, i.e., the field intensity ratio between
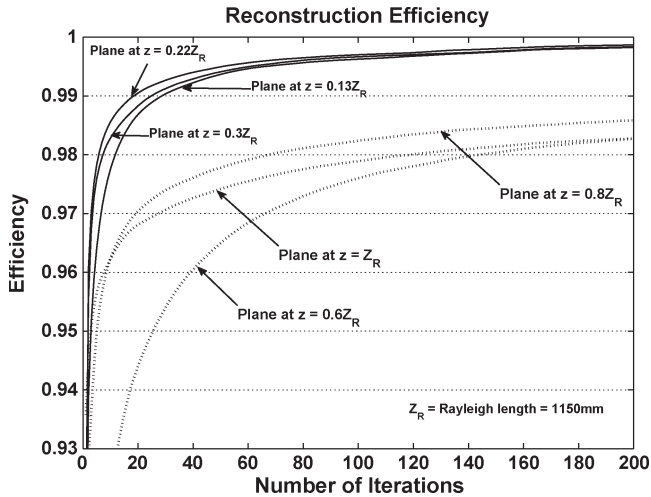
Fig. 7. Reconstruction efficiency in terms of CCF for two different cases of IPRA. In the first case, IPRA is evaluated using planes at 0.13 Z_R, 0.215 Z_R, and 0.3 Z_R mm. The plane size, in this case, is 300 mm × 300 mm, and the number of subdivisions in both directions is 512. In the second case, IPRA is evaluated using planes at 0.6 Z_R, 0.8 Z_R, and Z_R mm. The plane size, in this case, is 450 mm × 450 mm, and the number of subdivisions in both directions is 512. $Z_R = 1150$ mm is the Rayleigh range.

the boundary value and the maximum value should be in excess of ~ -50 dB. Evidently, the output beam field pattern would be broader if one goes far away from the beam waist location; therefore, one needs a larger plane size to fulfill the condition of zero intensity at the boundary. As shown in Fig. 2, the larger plane size requires a large number of sampling points if one wants closely spaced field patterns (this is obvious from the point of view of an accurate axial intensity variation and the efficiency of the reconstruction algorithm). On the other hand, if we are closer to the beam waist location, the field pattern would be smaller, and one needs a smaller plane size to fulfill the condition of zero intensity at the boundary and, hence, the smaller scan separation which can provide better axial variation of intensity. This is shown in Fig. 7, where the reconstruction efficiency is plotted in terms of the CCF for two different cases of plane size, scan separation, and position of the beam patterns w.r.t. the beam waist location. The distances of propagation are shown in terms of the Rayleigh range given by 1150 mm. The reconstruction algorithm is evaluated using the Huygens–Fresnel propagation method and following (25). The field was propagated between the three planes such that $1 \rightarrow 2 \rightarrow 3$ for the forward propagation and $3 \rightarrow 1$ for the backward propagation. During back propagation, the field was interpolated to a different grid to satisfy (25). As one can clearly see that in the case when the beam patterns are closely spaced and positioned closer to the beam waist location, the efficiency is better, compared to the other case, when the separation between the planes is larger and they are located farther from the beam waist location. Moreover, we used a bigger plane size of 450 mm × 450 mm in the latter case compared with the plane size of 300 mm × 300 mm of the first case. We can further justify this statement by analyzing the reconstruction error during IPRA.

We define a function called an error function calculated by summing the squared difference of the calculated modulus of
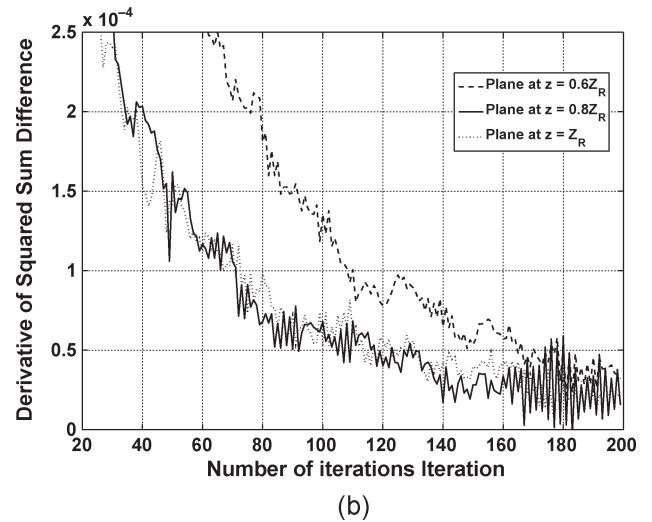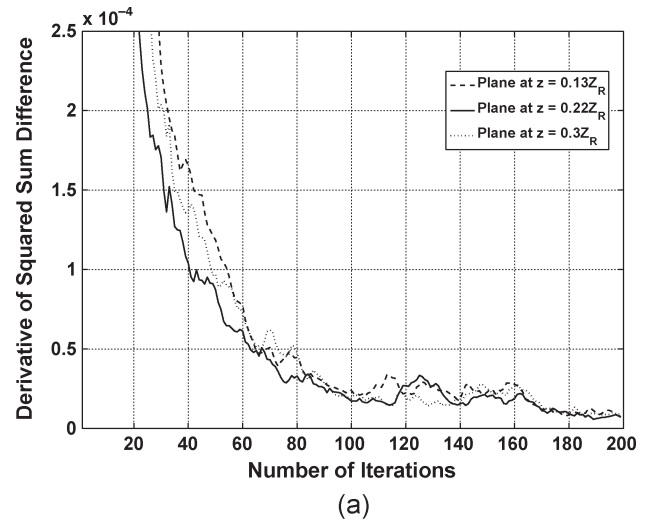

(a)

(b)

Fig. 8. Derivative of error function plotted w.r.t. the number of iterations. (a) Square plane size = 300 mm; number of sampling points = 512; and distances from the beam waist location are 0.13 Z_R, 0.215 Z_R, and 0.3 Z_R mm. (b) Square plane size = 450 mm; number of sampling points = 512; and distances from the beam waist location are 0.6 Z_R, 0.8 Z_R, and Z_R mm. $Z_R = 1150$ mm is the Rayleigh range.

the theoretical field $E_{\text{theory}}(i,j)$ and the reconstructed field $E_{\text{rec}}(i,j)$ at every point of the transverse plane as

$$\text{error} = \frac{\sum_{i,j} (|E_{\text{theory}}(i,j)| - |E_{\text{rec}}(i,j)|)^2}{\sum_{i,j} |E_{\text{theory}}(i,j)|^2}. \quad (29)$$

The derivative of this error function is shown in Fig. 8 for the aforementioned two cases shown in Fig. 7. One can easily see that the convergence in the first case is faster and smoother compared with the other case. Moreover, the algorithm converges to the higher value in the first case. The iteration algorithm was terminated manually after 200 iterations, irrespective of the value of the convergence at that point. The value of 200 was chosen arbitrarily from the inspection of the reconstructed amplitude as a best compromise between the computational effort and the quality of the result. Ideally, the error function derivative should reach a saturation level monotonically and
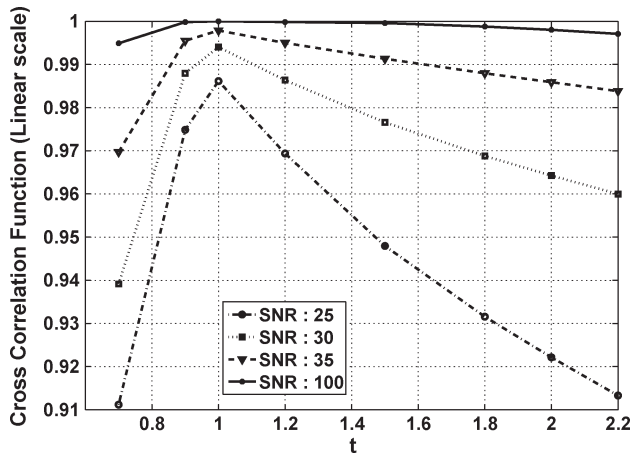
Fig. 9. Effect of noise on the CCF for different values of the SNR. Field is propagated/back propagated using the RSDI for a plane size of 600 mm.

should be terminated at its global minima. From a comparison between the two cases, one can conclude that the closely spaced planes situated close to the beam waist position (for a fixed number of sampling points and, therefore, smaller grid step size) provide more accurate results compared with the situation when the planes are located farther from each other and from the beam waist location.

We can also qualitatively discuss the effects of the dynamic range of the field intensity and initial phase guess on the IPRA. However, the appropriate discussion of these parameters can only be explained with the measured beam profiles [13], where one can truly investigate them in a more realistic way in the presence of the measurement noise. To give an insight into the effect of noise on the reconstruction algorithm and propagation distance using theoretical data, we added Gaussian white noise to the field for different values of signal-to-noise ratio (SNR) and calculated the CCF corresponding the case of RSDI for a plane size of 600 mm, as shown in Fig. 6. This is shown in Fig. 9, where the CCF is plotted with respect to the distance of propagation in terms of the parameter "b" for different values of SNR. It is clearly shown in Fig. 9 that the effect of noise on the CCF is not symmetric for different distances and that, regardless of the value of the SNR, the maximum of CCF is always found at the position ($t = 1$) which follows from (25).

More dynamics in the measured field intensity can provide more detailed and accurate reconstruction results. This parameter clearly depends on the target material used for measuring the beam profile during free-space infrared measurements. The choice of the material depends on the reflection and absorption of the microwave field at the frequency used and the temperature linearity w.r.t. the irradiating power.

The phase used to initially propagate the field for the first time would mainly affect the convergence of the algorithm. The resemblance of the initial phase with the true phase would help in faster and better convergence. The initial approximating phase can be taken as a flat, random, or quadratic phase. A better approximation of the initial phase can be made by using an irradiance of moment method. One can also calculate the initial phase by optimizing the coefficients of a 2-D quadratic phase function, which can minimize the propagation error

defined by (29), although this error minimization is virtually the same as IPRA; however, combining the two similar approaches can provide more confidence in the reconstructed result.

V. Conclusion

This paper discussed various parameters which affect the retrieval of phase from amplitude measurements using IPRA, mainly the field-propagation method. We solved the full RSDI using FFT, and the effect of the discretization is studied in detail. The Huygens–Fresnel propagation method is also discussed, and it is observed that using the Huygens–Fresnel propagation method for a fixed distance of propagation provides more accurate results and that the modulation effects in the propagated field can be avoided. A method for arbitrary distance of propagation using the Huygens–Fresnel propagation method is suggested; moreover, its implications on the propagated and back-propagated fields are discussed. Various parameters which can affect the reconstruction results were studied, and it is found that, for a correct phase reconstruction, it is preferred that the beam profile measurements be spaced based on (25) in the case of HFP. In the case of RSDI, it is difficult to find the exact condition on the distance of propagation from (20); however, during the analysis, it was observed that the minimum distance of propagation was approximately equal to the distance given by (25). To get more accurate results, measurements should be carried out within one or two Rayleigh lengths and, if possible, close to the beam waist location. One has to make a good compromise among all the parameters during IPRA — the mesh density, larger plane size (so that the field amplitude becomes almost zero (~ -50 dB) at the boundary), smaller scan separation, number of planes used, field close to the beam waist, larger dynamic range, initial phase guess, and properly aligned field profile.

References

[1] B. Z. Katsenelenbaum and V. V. Semenov, "Synthesis of phase correctors shaping a specified field," *Radio Eng. Electron. Phys.*, vol. 12, pp. 223–230, 1967.
[2] A. V. Chirkov, G. G. Denisov, and N. L. Aleksandrov, "3D wavebeam field reconstruction from intensity measurements in a few cross sections," *Opt. Commun.*, vol. 115, no. 5/6, pp. 449–452, 1995.
[3] A. V. Chirkov and G. G. Denisov, "Methods of wavebeam phase front reconstruction using intensity measurements," *Int. J. Infrared Millim. Waves*, vol. 21, no. 1, pp. 83–90, Jan. 2000.
[4] S. P. Antipov, A. A. Bogdashov, A. V. Chirkov, and G. G. Denisov, "Comparison of wavebeam phase front retrieval methods based on iterative algorithm and irradiance moments," *Int. J. Infrared Millim. Waves*, vol. 24, no. 10, pp. 1677–1685, Oct. 2003.
[5] D. R. Denison, T. S. Chu, M. A. Shapiro, and R. J. Temkin, "Gyrotron internal mode converter reflector shaping from measured field intensity," *IEEE Trans. Plasma Sci.*, vol. 27, no. 2, pp. 512–519, Feb. 1999.
[6] H. Idei, T. Shimozuma, M. A. Shapiro, T. Notake, S. Kubo, and R. J. Temkin, "Experimental verification of phase retrieval of quasi-optical millimeter-wave beams," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 11, pp. 3899–3905, Nov. 2006.

[7] J. P. Anderson, M. A. Shapiro, R. J. Temkin, and D. R. Denison, "Phase retrieval of gyrotron beams based on irradiance moments," *IEEE Trans. Microw. Theory Tech.*, vol. 50, no. 6, pp. 1526–1535, Jun. 2002.

[8] C. Lee, J. Lowengrub, J. Rubinstein, and X. Zheng, "Phase reconstruction by the weighted least action principle," *J. Opt. A, Pure Appl. Opt.*, vol. 8, no. 3, pp. 279–289, Mar. 2006.

[9] T. E. Gureyev and K. A. Nugent, "Phase retrieval with the transport of intensity equation. II. Orthogonal series solution for nonuniform illumination," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 13, no. 8, pp. 1670–1682, Aug. 1996.

[10] J. P. Hogge, S. Alberti, A. Arnold, D. Bariou, P. Benin, T. Bonicelli, A. Bruschi, R. Chavan, S. Cirant, O. Dumbrajs, D. Fasel, F. Gandini, E. Giguet, T. Goodman, R. Heidinger, M. Henderson, S. Illy, J. Jin, C. Lievin, R. Magne, P. Marmillod, P. L. Mondino, A. Perez, B. Piosczyk, L. Porte, T. Rzesnicki, M. Santinelli, M. Thumm, M. Q. Tran, and I. Yovchev, "Development of a 2-MW, CW coaxial gyrotron at 170 GHz and test facility for ITER," *J. Phys.: Conf. Ser.*, vol. 25, pp. 33–44, 2005.

[11] B. Piosczyk, G. Dammertz, O. Dumbrajs, O. Drumm, S. Illy, J. Jin, and M. Thumm, "A 2-MW, 170-GHz coaxial cavity gyrotron," *IEEE Trans. Microw. Theory Tech.*, vol. 32, no. 3, pp. 413–417, Jun. 2004.

[12] *Technical Report—The European 170 GHz, 2MW Coaxial Cavity Gyrotron*. Private Communication.

[13] S. K. Jawla, J.-P. Hogge, S. Alberti, B. Piosczyk, T. Rzesnicki, and C. Liévin, "Beam profile measurements of the output microwave beam of 2MW/170 GHz coaxial-cavity gyrotron for ITER," in *Proc. 35th IEEE ICOPS*, Karlsruhe, Germany, 2008, p. 249.

[14] M. Born and E. Wolf, *Principles of Optics*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[15] J. W. Goodman, *Introduction to Fourier Optics*. New York: McGraw-Hill, 1968.

[16] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1990.

[17] [Online]. Available: http://cnx.org/content/m10987/latest/

[18] L. Onural, "Exact analysis of the effects of sampling of the scalar diffraction field," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 24, no. 2, pp. 359–367, Feb. 2007.

[19] M. Sypek, C. Prokopowicz, and M. Gorecki, "Image multiplying and high-frequency oscillations effects in the Fresnel region light propagation simulation," *Opt. Eng.*, vol. 42, no. 11, pp. 3159–3164, Nov. 2003.

Sudheer Jawla (S'08) was born in Uttar Pradesh, India, in 1982. He received the B.Sc. degree from Ch. Charan Singh University, Meerut, India, in 2002 and the M.Sc. degree in physics from Indian Institute of Technology, New Delhi, India, in 2004. Since 2005, he has been working toward the Ph.D. degree in high-power gyrotrons in the Gyrotron Research Group, Centre de Recherches en Physiques des Plasmas, Association Euratom—Confédération Suisse, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

His research interests include quasi-optical millimeter-wave components for high-power gyrotrons, lasers and microwave driven particle acceleration.

Jean-Philippe Hogge, photograph and biography not available at the time of publication.

Stefano Alberti, photograph and biography not available at the time of publication.