Journal of Computational Physics **158,** 161–185 (2000)
doi:10.1006/jcph.1999.6407, available online at http://www.idealibrary.com on IDEAL®

Fast Evaluation of Two-Dimensional Transient Wave Fields

Mingyu Lu, Jianguo Wang, A. Arif Ergin, and Eric Michielssen

Department of Electrical and Computer Engineering, Center for Computational Electromagnetics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801
E-mail: mingyulu@students.uiuc.edu, jwang@extreme.ece.uiuc.edu, aergin@decwa.ece.uiuc.edu, michiels@decwa.ece.uiuc.edu

Received July 6, 1999

This paper presents a novel scheme to efficiently evaluate transient linear wave fields that are generated by two-dimensional (2D) source configurations. The scheme, termed the plane wave time domain algorithm (PWTD), realizes a diagonal translation operator for 2D transient wave fields through their representation in terms of Hilbert transformed plane wave expansions. Numerical results are presented that validate the algorithm and demonstrate its convergence properties. The proposed PWTD algorithm can be coupled to classical 2D time domain integral equation solvers in a two-level and multilevel setting. It is shown that analysis of a 2D surface scattering phenomenon, in which sources are represented in terms of N_s spatial and N_t temporal samples, based on two-level and multilevel PWTD augmented integral equation solvers, requires $O(N_s^{1.5} N_t \log N_t)$ and $O(N_s N_t \log N_s \log N_t)$ computational resources, respectively (as opposed to $O(N_s^2 N_t^2)$ for a classical solver). Therefore, these PWTD schemes render feasible the rapid integral equation based analysis of 2D transient scattering phenomena involving large surfaces.

Key Words: integral equations; time domain; plane wave time domain (PWTD); Hilbert transform; slant stack transform (SST); computational complexity.

1. INTRODUCTION

Accurate and efficient techniques for analyzing transient linear wave phenomena are of interest to disciplines ranging from acoustics to electromagnetics to geophysics [1, 2]. In the past, significant efforts have been expended on the development of integral equation based methods for analyzing transient two-dimensional (2D) surface scattering phenomena, the vast majority of which can be classified as marching on in time (MOT) schemes [3–6]. Unfortunately, all these methods suffer from a high computational complexity, which severely limits their application to practical, real-world problems.

0021-9991/00 $35.00

To accelerate the integral equation based analysis of three-dimensional (3D) transient scattering phenomena, our group recently developed the 3D plane wave time domain (PWTD) algorithm [7, 8], which can be considered the extension of the 3D frequency domain fast multipole method [9, 10] to the time domain. The 3D PWTD algorithm permits the *fast evaluation* of transient linear wave fields that are generated by *known* 3D source configurations. In addition, when coupled to 3D MOT solvers, the 3D PWTD algorithm permits the *fast reconstruction of unknown* 3D source configurations that generate the fields scattered from an object. In other words, 3D PWTD enhanced MOT solvers permit the rapid solution of transient 3D surface scattering problems.

In this paper, we propose a 2D PWTD algorithm, i.e., an extension of our previous work from 3D to 2D. The 2D time domain Green's function for the wave equation, in contrast to its 3D cousin, is not localized in time and has an infinitely long temporal tail, which complicates the construction of the 2D PWTD scheme. When used in conjunction with an MOT solver [3–5], the 2D PWTD method drastically accelerates the solution of transient 2D scattering problems.

Consider an impenetrable or penetrable but homogenous object of invariant cross section, extending along the z-axis, and residing in a homogeneous medium with wave speed c (Fig. 1). Assume that the maximum transverse linear dimension of this object is D, that the object is excited by a z-invariant transient incident field with temporal bandwidth ω_{max}, and that the surface sources induced by the incident field are represented using $N_s \propto D\omega_{max}/c$ and $N_t \propto T\omega_{max}$ spatial and temporal samples, where T is the total temporal duration of the analysis. The computational cost associated with the integral equation based analysis of this scattering problem using classical MOT methods scales as $O(N_s^2 N_t^2)$. The 2D PWTD algorithm developed in this paper adopts a plane wave expansion to arrive at a diagonal translation operator for transient 2D wave fields that permits the rapid evaluation of fields due to surface bound source densities. It will be shown that two-level and multilevel PWTD enhanced MOT algorithms permit the analysis of 2D scattering phenomena in $O(N_s^{1.5} N_t \log N_t)$ and $O(N_s N_t \log N_s \log N_t)$ operations. It is anticipated that these PWTD enhanced MOT schemes will render feasible the fast analysis of 2D transient scattering phenomena involving large and complex surfaces.

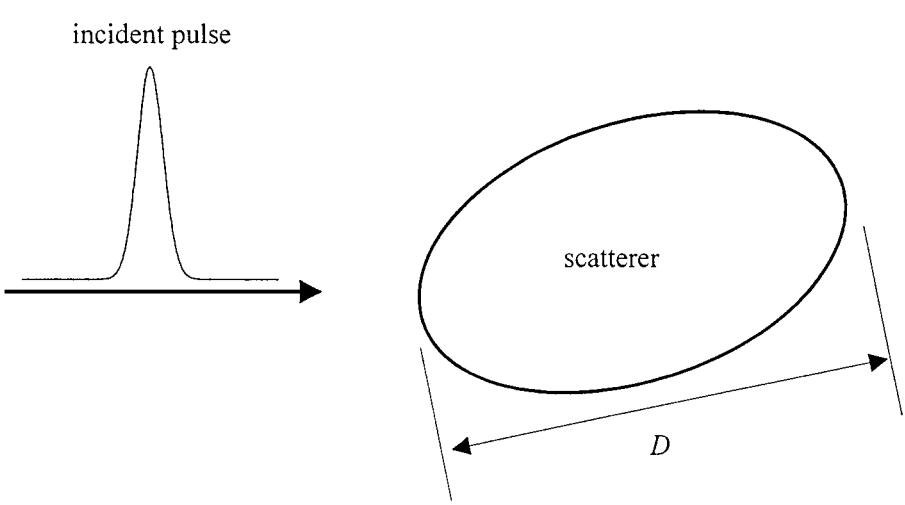

FIG. 1. Sketch map of two-dimensional scattering.

We wish to note that the proposed scheme is not the first attempt at constructing a fast method for evaluating 2D wave fields due to known sources. First, Crutchfield proposed a scheme that, in spirit, is similar to the one outlined in this paper [11]. In contrast to our scheme, however, Crutchfield's does not allow for a systematic elimination of ghost signals associated with time domain plane wave expansions. Second, Alpert *et al.* proposed a technique for rapidly evaluating nonreflecting boundary kernels for time domain wave propagation [12]. However, their scheme applies only to circular source distributions, whereas ours applies to arbitrarily shaped configurations.

This paper is organized as follows. The proposed PWTD scheme for reconstructing 2D wave fields due to known sources is presented in Section 2. This section also discusses the accuracy and convergence properties of this algorithm. The computational complexity of two-level and multilevel PWTD enhanced MOT solvers is analyzed in Section 3. Finally, Section 4 summarizes the conclusions of this study. A variety of definitions, including those for the Fourier and Hilbert transforms and a local interpolation function that is used extensively in this study, is provided in Appendix.

2. THE TWO-DIMENSIONAL PLANE WAVE TIME DOMAIN ALGORITHM

This section describes the 2D PWTD algorithm for evaluating transient fields due to two-dimensional source distributions. Subsection 2.1 introduces notation and analyzes the computational complexity of a traditional 2D MOT scheme. Subsection 2.2 describes a representation of the transient field produced by a two-dimensional source configuration in terms of a Hilbert transformed plane wave expansion and a condition that permits a ghost free recovery of the field from this expansion. The 2D PWTD algorithm, including a temporal sampling scheme for sources and fields, a closed form expression for a diagonal translation operator for 2D wave fields, and a method for efficiently evaluating the Hilbert transform within the context of the PWTD scheme, are discussed in Subsection 2.3. Finally, Subsection 2.4 describes a variety of numerical experiments that validate the algorithm.

2.1. *Preliminaries*

Consider a field $A(\rho, t)$ that satisfies the 2D wave equation,

$$\nabla^2 A(\rho, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} A(\rho, t) = -J(\rho, t), \tag{1}$$

where $J(\rho, t)$ is the source distribution and the operator $\nabla^2 = (\partial^2/\partial x^2 + \partial^2/\partial y^2)$.

The field at an observer ρ_o can be expressed as

$$A(\rho_o, t) = \int_C d\rho' J(\rho', t) * g(|\rho_o - \rho'|, t), \tag{2}$$

where C is the contour of the scatterer's cross section on which the sources reside, $*$ denotes temporal convolution, $g(\cdot,\cdot)$ is the 2D Green's function

$$g(\rho, t) = \frac{1}{2\pi}\frac{u(t - \rho/c)}{\sqrt{t^2 - (\rho/c)^2}}, \tag{3}$$

and $u(\cdot)$ is the Heaviside step function (Appendix).

Assume for the moment that $J(\rho, t)$ consists of a single line source located at ρ_s with temporal signature $f(t)$ of approximate duration T and bandwidth ω_{max}, i.e.,

$$J(\rho, t) = f(t)\delta(\rho - \rho_s), \tag{4}$$

where $f(t) \cong 0$ outside the interval $0 \le t \le T$. From Eq. (2), the field at ρ_o due to this source can be expressed as

$$A(\rho_o, t) = f(t) * g(\rho_{so}, t), \tag{5}$$

where $\rho_{so} = |\rho_{so}|$ and $\rho_{so} = \rho_o - \rho_s$.

To study the computational cost associated with the evaluation of Eq. (5), assume that the band-limited source signature $f(t)$ is represented in terms of $N_t \propto T\omega_{max}$ temporal samples. It is easily verified that, for arbitrary ρ_s and ρ_o, the cost of directly evaluating the right hand side of Eq. (5) for a total number of N_t temporal samples scales as $O(N_t^2)$. Next, consider the 2D transient scattering problem described in the Introduction to this paper. The computational cost associated with the integral equation based analysis of this problem scales as $O(N_t^2 N_s^2)$. This complexity estimate is arrived at as follows. MOT-based integral equation solvers construct consecutive temporal snapshots of the source distributions that generate the fields scattered by an object by requiring that the sum of the incident and scattered fields satisfies a given instantaneous boundary condition on the scatterer surface. By virtue of Eq. (5), the scattered fields depend on the past values of the surface sources. The evaluation of the instantaneous scattered field observed at a single location on the surface calls for the evaluation of the convolution in Eq. (5) for all the N_s sources and can be accomplished in $O(N_s N_t)$ operations. Therefore the cost of calculating the scattered fields over the scatterer's surface scales as $O(N_t N_s^2)$. Once these fields have been evaluated, instantaneous source distributions can be computed, which in turn permits the MOT solver to consider the next time step. The MOT scheme incurs the above cost for each of the N_t time steps, hence the $O(N_s^2 N_t^2)$ complexity estimate. This computational cost prohibits the analysis of large-scale scattering problems.

2.2. *The 2D Plane Wave Expansion*

To reduce the computational cost associated with the evaluation of the convolution in Eq. (5) for all source-observer pairs in a classical MOT algorithm, a plane wave expansion of the source field, similar to that introduced in [7, 8], is considered next.

To represent the source field in terms of a plane wave basis, the source signal $f(t)$ is broken up into N_v consecutive subsignals $f_v(t)$, $v = 0, 1, \ldots, N_v - 1$, as

$$f(t) = \sum_{v=0}^{N_v - 1} f_v(t). \tag{6}$$

As depicted in Fig. 2, it is assumed that $f_v(t) \cong 0$ for $t < t_{v1}$ and $t > t_{v2}$, i.e., that $f_v(t)$ starts at t_{v1} and ends at t_{v2}, and that all subsignals $f_v(t)$ are of equal duration $T_s = t_{v2} - t_{v1}$. A computational scheme for achieving decomposition (6) will be described in Subsection 2.3.1. Let $A_v(\rho_o, t)$ denote the field at ρ_o due to the line source at ρ_s with temporal signature $f_v(t)$, then

$$A_v(\rho_o, t) = f_v(t) * g(\rho_{so}, t), \tag{7}$$

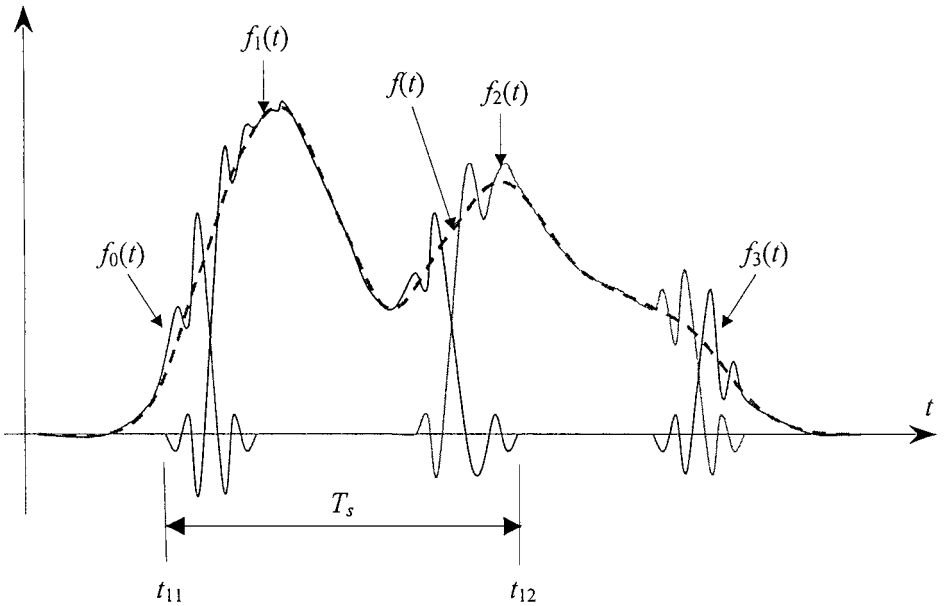

FIG. 2. The original signal and subsignals.

and therefore

$$A(\rho_o, t) = \sum_{v=0}^{N_v-1} A_v(\rho_o, t). \tag{8}$$

To construct a plane wave representation of the source field, consider the signal $I_{f_v}(\rho_{so}, t)$ defined as

$$I_{f_v}(\rho_{so}, t) = \frac{1}{4\pi} \int_0^{2\pi} f_v(t - \rho_{so} \cdot \hat{\mathbf{k}}(\phi)/c)\, d\phi, \tag{9}$$

where $\hat{\mathbf{k}}(\phi) = \hat{\mathbf{x}} \cos\phi + \hat{\mathbf{y}} \sin\phi$. It is easily shown that

$$I_{f_v}(\rho_{so}, t) = f_v(t) * I_\delta(\rho_{so}, t), \tag{10}$$

where $I_\delta(\rho_{so}, t)$ is

$$\begin{aligned} I_\delta(\rho_{so}, t) &= \frac{1}{4\pi} \int_0^{2\pi} \delta(t - \rho_{so} \cdot \hat{\mathbf{k}}(\phi)/c)\, d\phi \\ &= \frac{1}{2\pi} \frac{P(t, \rho_{so}/c)}{\sqrt{(\rho_{so}/c)^2 - t^2}}. \end{aligned} \tag{11}$$

Here, $P(\cdot,\cdot)$ is a unit pulse function (Appendix).

Now, let $\tilde{A}_v(\rho_o, t)$ denote the Hilbert transform (Appendix) of $I_{f_v}(\rho_{so}, t)$,

$$\tilde{A}_v(\rho_o, t) = \mathcal{H}\{I_{f_v}(\rho_{so}, t)\} = f_v(t) * \tilde{g}(\rho_{so}, t). \tag{12}$$

In the above equation, $\tilde{g}(\rho_{so}, t)$ denotes the Hilbert transform of $I_\delta(\rho_{so}, t)$ [2],

$$\tilde{g}(\rho_{so}, t) = \mathcal{H}\{I_\delta(\rho_{so}, t)\} = \frac{u(t - \rho_{so}/c) - u(-t - \rho_{so}/c)}{2\pi\sqrt{t^2 - (\rho_{so}/c)^2}}. \tag{13}$$

Comparison of Eqs. (3) and (13) establishes the following relationship between $g(\rho_{so}, t)$ and $\tilde{g}(\rho_{so}, t)$,

$$g(\rho_{so}, t) = \begin{cases} \tilde{g}(\rho_{so}, t), & t \geq 0 \\ 0, & t < 0. \end{cases} \tag{14}$$

From Eqs. (7), (12), and (13), it is seen that $\tilde{A}_v(\rho_o, t)$ is closely related to the actual field $A_v(\rho_o, t)$. Specifically, it follows from Eq. (13) that $\tilde{g}(\rho_{so}, t)$ is composed of two parts: $g(\rho_{so}, t)$ and an *image* with respect to the temporal origin. Therefore, $\tilde{A}_v(\rho_o, t)$ is also composed of two parts: the actual signal $A_v(\rho_o, t)$ and an image, which will be referred to as the *ghost signal*. This observation closely follows that in [2].

To exploit Eq. (12) in the construction of a fast scheme for evaluating $A_v(\rho_o, t)$, assume that ρ_s resides within a source circle centered around ρ_1, and that ρ_o resides within an observation circle centered around ρ_2 (Fig. 3). Assume that these two circles have the same radius R_s. Next, a scheme is described for representing the observer field in terms of Hilbert transformed plane wave expansions through a *diagonal* translation operator. The scheme is designed to facilitate its incorporation into MOT based integral equation solvers, which, as will be demonstrated below, guarantees that the above described ghost signal is automatically eliminated by time-gating.

If the vector connecting the source and the observation points is expressed as

$$\begin{aligned} \rho_{so} &= (\rho_o - \rho_2) + (\rho_2 - \rho_1) + (\rho_1 - \rho_s) \\ &= \rho_{2o} + \rho_{12} + \rho_{s1}, \end{aligned} \tag{15}$$

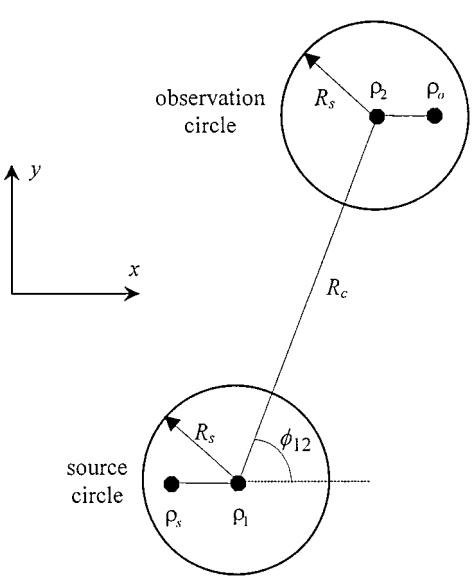

FIG. 3. Geometry under study.

then Eq. (9) can be cast in the form

$$I_{f_v}(\rho_{so}, t) = \frac{1}{4\pi}\int_0^{2\pi} \delta\left[t - \frac{\rho_{2o}\cdot\hat{\mathbf{k}}(\phi)}{c}\right] * \delta\left[t - \frac{\rho_{12}\cdot\hat{\mathbf{k}}(\phi)}{c}\right] * \delta\left[t - \frac{\rho_{s1}\cdot\hat{\mathbf{k}}(\phi)}{c}\right] * f_v(t)\, d\phi. \tag{16}$$

Using Eq. (12), and the fact that the Hilbert transform commutes with the temporal convolutions, the following expression for $\tilde{A}_l(\rho_o, t)$ results:

$$\tilde{A}_v(\rho_o, t) = \frac{1}{4\pi}\int_0^{2\pi} \delta\left[t - \frac{\rho_{2o}\cdot\hat{\mathbf{k}}(\phi)}{c}\right] * \mathcal{H}\left\{\delta\left[t - \frac{\rho_{12}\cdot\hat{\mathbf{k}}(\phi)}{c}\right] * \delta\left[t - \frac{\rho_{s1}\cdot\hat{\mathbf{k}}(\phi)}{c}\right] * f_v(t)\right\} d\phi. \tag{17}$$

This equation suggests the following four-stage scheme for evaluating $\tilde{A}_v(\rho_o, t)$:

(i) *Construction of outgoing rays.* For every direction $\hat{\mathbf{k}}(\phi)$, evaluate the rightmost convolution appearing in Eq. (17) between $f_v(t)$ and $\delta[t - \rho_{s1}\cdot\hat{\mathbf{k}}(\phi_n)/c]$. This operation imposes a direction dependent temporal shift on $f_v(t)$. The resulting signals are termed outgoing rays, and each ray describes a plane wave emanating from the source circle in the $\hat{\mathbf{k}}(\phi)$ direction.

(ii) *Construction of incoming rays.* For every direction $\hat{\mathbf{k}}(\phi)$, carry out the center convolution appearing in Eq. (17) between the outgoing rays and $\delta[t - \rho_{12}\cdot\hat{\mathbf{k}}(\phi)/c]$. This operation translates outgoing rays between the centers of the source and observation circles. The resulting signals are termed incoming rays, and each ray describes a plane wave impinging on the observer circle from direction $-\hat{\mathbf{k}}(\phi)$.

(iii) *Construction of Hilbert transformed incoming rays.* For every direction $\hat{\mathbf{k}}(\phi)$, evaluate the Hilbert transform of the incoming ray.

(iv) *Construction of the observer field.* For every direction $\hat{\mathbf{k}}(\phi)$, evaluate the leftmost convolution appearing in Eq. (17), i.e., convolve the Hilbert transform of the incoming rays with $\delta[t - \rho_{2o}\cdot\hat{\mathbf{k}}(\phi)/c]$ and add up the resulting signals for all $\hat{\mathbf{k}}(\phi)$, i.e., carry out the integration in Eq. (17). This operation projects the Hilbert transform of the incoming rays onto the observers by imposing appropriate temporal shifts.

As mentioned previously, the observer signal $\tilde{A}_v(\rho_o, t)$ constructed through the above sequence of operations will be corrupted by a ghost signal. To render the above scheme useful within an MOT based integral equation solver, assume that steps (ii) and (iii) are effected only at $t = t_{v2}$. In other words, assume that the signals $\tilde{A}_v(\rho_o, t)$ are explicitly time gated out for $t < t_{v2}$. Note that, for $t > t_{v2}$, the outgoing rays are completely formed and move out from the center of the source circle with speed c; hence, the above assumption simplifies the execution of steps (ii)–(iii). Next, note that the earliest time of arrival of the actual signal in the observer circle associated with a source with temporal signature $f_v(t)$ residing in the source circle is

$$\begin{aligned} t_a &= \frac{R_c - 2R_s}{c} + t_{v1} \\ &= \frac{R_c - 2R_s}{c} + t_{v2} - T_s. \end{aligned} \tag{18}$$

Since $\tilde{A}_v(\rho_o, t)$ will only be observed in the observer circle for $t \geq t_{v2}$, it is required that $t_a \geq t_{v2}$, as otherwise part of the true signal would be discarded, which would be an undesired byproduct of the above suggested scheme. Therefore, it follows from Eq. (18) that

$$T_s \leq \frac{R_c - 2R_s}{c}. \tag{19}$$

Equation (19) guarantees that none of the actual signal is dismissed due to time gating. However, what about the ghost signal? Will it be observed? From Fig. 4, which depicts $I_{f_v}(\rho_{so}, t)$ and $\tilde{A}_v(\rho_o, t)$, it is seen that the ghost signal vanishes after $t_{v2} - \rho_{so}/c < t_{v2}$. Hence, if translations only take place after $t = t_{v2}$, then condition (19) also guarantees that only actual fields $A_v(\rho_o, t)$ are observed, i.e., $A_v(\rho_o, t) = \tilde{A}_v(\rho_o, t)$ for $t > t_{v2}$. Therefore, the above described sequence of operations guarantees a ghost free recovery of the field due to an arbitrary configuration of sources distributed over the source circle throughout the observer circle.

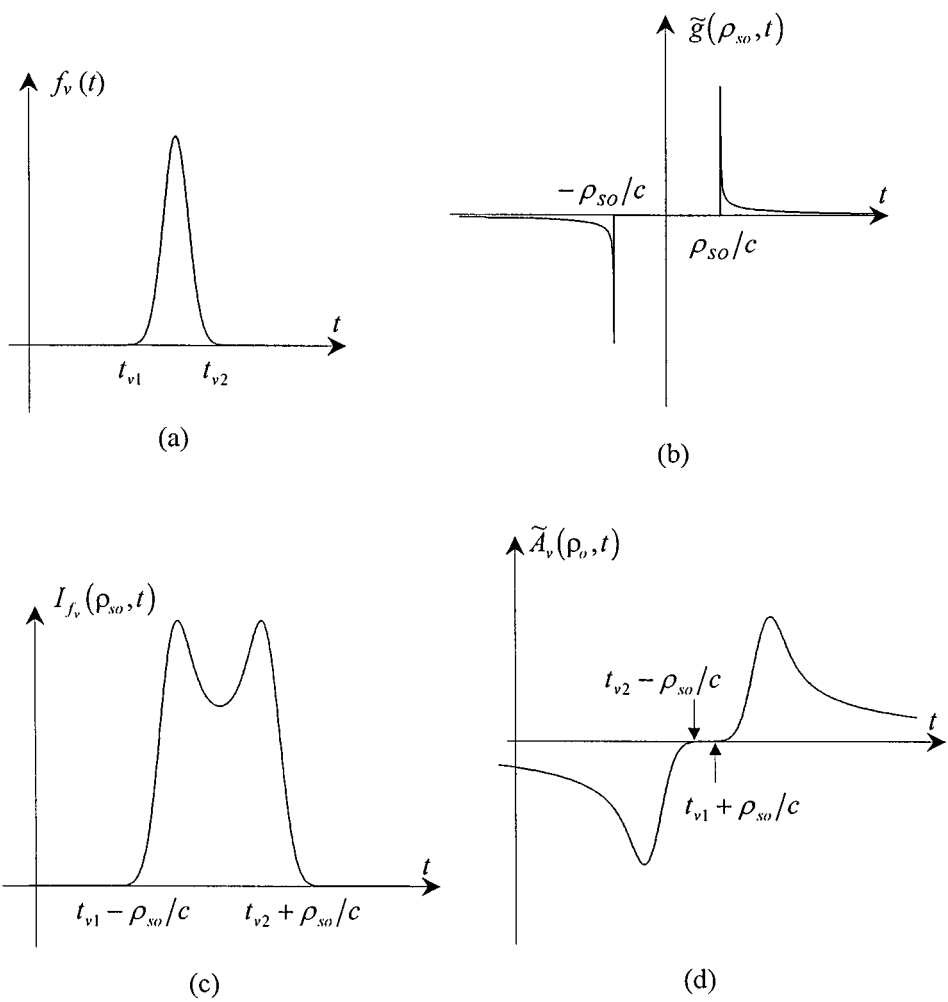

FIG. 4. Illustration for PWTD algorithm. (a) The source subsignal $f_v(t)$; (b) $\tilde{g}(\rho_{so}, t)$; (c) the plane wave expansion of $f_v(t)$; (d) result of the convolution between $f_v(t)$ and $\tilde{g}(\rho_{so}, t)$.

2.3. *Computational Aspects*

To efficiently evaluate $A_v(\rho_o, t)$ using the above described algorithm, computational schemes are needed to decompose the source signal into subsignals as described by Eq. (6), to carry out the convolutions and the angular integration appearing in Eq. (17), and to evaluate the Hilbert transform. Subsections 2.3.1, 2.3.2, and 2.3.3 address these issues in turn. The computational complexity of this scheme within an MOT solver is analyzed in Subsection 2.3.4.

2.3.1. Temporal sampling and representation. This section describes a technique for breaking up the source signal into a set of subsignals. This decomposition is called for because Eq. (19) imposes a maximum duration on the signal that can be translated ghost free from the source to the observer circle. Because $f(t)$ is band-limited, each subsignal $f_v(t)$ can be described by M_t samples of $f(t)$, as

$$f_v(t) = \sum_{k=vM_t}^{(v+1)M_t-1} f(k\Delta t)\psi_k(t), \tag{20}$$

where $\psi_k(t)$ is a time shifted local interpolant, and the choice of M_t will be discussed later. Many good choices of $\psi_k(t)$ exist; however, in this study, we choose

$$\psi_k(t) = P(t - k\Delta t, \omega_{max}, \chi_1, p_t), \tag{21}$$

where $P(t, \omega_{max}, \chi_1, p_t)$ is an approximate prolate spheroidal interpolant (Appendix). The interpolant $P(t, \omega_{max}, \chi_1, p_t)$ is band-limited to $\omega_f = \chi_1\omega_{max}$. The temporal oversampling ratio $\chi_1 > 1$ is related to the sampling rate as $\Delta t = \pi/(\chi_1\omega_{max})$. As stated in the Appendix, $P(t, \omega_{max}, \chi_1, p_t)$ is virtually time-limied and, for all practical purposes, only $2p_t + 1$ samples of $\psi_k(t)$ are required in the interpolation described by Eq. (20). The error of this local interpolation can be controlled and made arbitrarily small by increasing χ_1 or p_t. As a result, the above construction indeed guarantees that $f_v(t) \cong 0$ when $t < t_{v1}$ and $t > t_{v2}$, where

$$\begin{cases} t_{v1} = (vM_t - p_t)\Delta t \\ t_{v2} = [(v+1)M_t - 1 + p_t]\,\Delta t. \end{cases} \tag{22}$$

As each subsignal $f_v(t)$ is essentially time-limited with duration T_s, Eqs. (19) and (22) imply that M_t should satisfy

$$M_t \le \left\lfloor \frac{T_s}{\Delta t} - 2p_t + 1 \right\rfloor, \tag{23}$$

where the $\lfloor\cdot\rfloor$ denotes the nearest smaller interger.

The above scheme not only accomplishes the decomposition (6) (because of the local nature of the interpolants $\psi_k(t)$), but also permits the efficient evaluation of the convolutions and angular integrations appearing in Eq. (17) because the $\psi_k(t)$ has a finite bandwidth, as discussed next.

2.3.2. Diagonal translation operators. To efficiently evaluate the angular integral in Eq. (17), i.e., to replace the angular integration in Eq. (17) by a finite sum, consider the

signal $p(t, \phi)$ defined as

$$p(t, \phi) = \delta\left[t - \frac{\rho_{2o} \cdot \hat{\mathbf{k}}(\phi)}{c}\right] * \delta\left[t - \frac{\rho_{s1} \cdot \hat{\mathbf{k}}(\phi)}{c}\right] * f_v(t). \tag{24}$$

This signal can be interpreted as the transient far-field radiation pattern of a source located at $\rho_{s1} + \rho_{2o}$, which is always less than $2R_s$ in magnitude. Since the angular variation in the radiation pattern of such a source distribution is quasi band-limited [13, 14], $p(t, \phi)$ can be represented in terms of $2N_\phi + 1$ samples as [13]

$$p(t, \phi) = \sum_{n=-N_\phi}^{N_\phi} p(t, \phi_n) D(\phi - \phi_n), \tag{25}$$

where $D(\phi)$ represents the Dirichlet kernel

$$D(\phi) = \frac{\sin[(2N_\phi + 1)\phi/2]}{(2N_\phi + 1)\sin(\phi/2)} = \frac{1}{2N_\phi + 1} \sum_{m=-N_\phi}^{N_\phi} e^{jm\phi}, \tag{26}$$

and $\phi_n = n2\pi/(2N_\phi + 1)$.

In the above, the parameter N_ϕ can be chosen as

$$N_\phi = \left\lceil \chi_2' \frac{\omega_f}{c}(2R_s) \right\rceil, \tag{27}$$

where $\lceil \cdot \rceil$ denotes the nearest larger integer, and χ_2' is termed the *angular oversampling ratio*. The expansion in Eq. (25) converges very fast with increasing χ_2'. Note that N_ϕ is closely related to $\omega_f = \chi_1 \omega_{\max}$, the bandwidth of the subsignal $f_v(t)$, as well as χ_2'. Henceforth, the parameter $\chi_2 = \chi_1 \chi_2'$ will be used as the combined oversampling ratio in dealing with N_ϕ.

Combining Eqs. (16), (24), and (25) yields

$$\begin{aligned} I_{f_v}(\rho_{so}, t) &= \frac{1}{4\pi} \int_0^{2\pi} \delta\left[t - \frac{\rho_{12} \cdot \hat{\mathbf{k}}(\phi)}{c}\right] * \left[\sum_{n=-N_\phi}^{N_\phi} p(t, \phi_n) D(\phi - \phi_n)\right] d\phi \\ &= \sum_{n=-N_\phi}^{N_\phi} p(t, \phi_n) * \mathcal{T}_n(\rho_{12}, t) \\ &= \sum_{n=-N_\phi}^{N_\phi} \delta\left[t - \frac{\rho_{2o} \cdot \hat{\mathbf{k}}(\phi_n)}{c}\right] * \mathcal{T}_n(\rho_{12}, t) * \delta\left[t - \frac{\rho_{s1} \cdot \hat{\mathbf{k}}(\phi_n)}{c}\right] * f_v(t), \end{aligned} \tag{28}$$

where

$$\begin{aligned} \mathcal{T}_n(\rho_{12}, t) &= \frac{1}{4\pi} \int_0^{2\pi} \delta\left[t - \frac{\rho_{12} \cdot \hat{\mathbf{k}}(\phi)}{c}\right] D(\phi - \phi_n)\, d\phi \\ &= \frac{1}{4\pi(2N_\phi + 1)} \sum_{m=-N_\phi}^{N_\phi} \int_0^{2\pi} \delta\left[t - \frac{R_c \cos\phi}{c}\right] e^{jm(\phi - \phi_n + \phi_{12})}\, d\phi \\ &= \frac{1}{2\pi(2N_\phi + 1)} \frac{c}{R_c} \sum_{m=0}^{N_\phi} \varepsilon_m \cos[m(-\phi_n + \phi_{12})] \frac{T_m(ct/R_c)}{\sqrt{1 - (ct/R_c)^2}} P\left(t, \frac{R_c}{c}\right). \end{aligned} \tag{29}$$

The $\mathcal{T}_n(\rho_{12}, t)$ appearing in the above equation is termed the plane wave translation function between the centers of the source and the observation circles, $T_m(\cdot)$ is the Chebyshev polynomial of degree m, $R_c = |\rho_{12}|$, ϕ_{12} is the angle between ρ_{12} and $\hat{\mathbf{x}}$ (Fig. 3), and $\varepsilon_m = 1$ for $m = 0$; $\varepsilon_m = 2$ for $m \neq 0$ (Neumann number).

For future reference, note that the translation function $\mathcal{T}_n(\rho_{12}, t)$ does not depend on the source and observation locations. It is also noteworthy that a closed form expression for $\tilde{\mathcal{T}}_n(\rho_{12}, \omega)$, the Fourier transform (Appendix) of $\mathcal{T}_n(t)$, is available:

$$\tilde{\mathcal{T}}_n(\rho_{12}, \omega) = \frac{1}{2(2N_\phi + 1)} \sum_{m=0}^{N_\phi} \varepsilon_m (-j)^m J_m\left(\frac{\omega R_c}{c}\right) \cos[m(-\phi_n + \phi_{12})]. \qquad (30)$$

In this equation, $J_m(\cdot)$ is the mth order Bessel function.

Combining Eqs. (12) and (28), the following expression for $\tilde{A}_l(\rho_o, t)$ results:

$$\tilde{A}_v(\rho_o, t) = \sum_{n=-N_\phi}^{N_\phi} \delta\left[t - \frac{\rho_{2o} \cdot \hat{\mathbf{k}}(\phi_n)}{c}\right] * \mathcal{H}\left\{\mathcal{T}_n(\rho_{12}, t) * \delta\left[t - \frac{\rho_{s1} \cdot \hat{\mathbf{k}}(\phi_n)}{c}\right] * f_v(t)\right\}. \qquad (31)$$

Given the sampled representations of $f_l(t)$ in terms of local bandlimited interpolants, the rightmost convolution can be carried out directly in the time domain (care should be taken to avoid aliasing). Given in addition that the outgoing rays are band-limited and that both the outgoing rays and the translation function are essentially time-limited, the center convolution can be accomplished in the frequency domain through a fast Fourier transform (FFT). The leftmost convolution can again be carried out in the time domain (variations on this theme are possible).

2.3.3. Evaluation of the Hilbert transform. In this section, an efficient technique for evaluating the Hilbert transform within the context of Eq. (31) is described. Before doing so, the properties of the incoming rays—the signals within brackets in the discrete representation Eq. (31)—are examined more closely.

The outgoing rays that result from the rightmost convolution in Eq. (31) pass through the center of the source circle within the temporal interval $[t_{v1} - R_s/c, t_{v2} + R_s/c]$, because $\rho_{s1} \leq R_s$ (Fig. 5a). It can be seen from Eq. (29) that $\mathcal{T}_n(\rho_{12}, t)$ spans the temporal interval $[-R_c/c, R_c/c]$. Therefore, upon convolving $\mathcal{T}_n(\rho_{12}, t)$ with the outgoing rays, the temporal interval during which incoming rays pass through the center of the observation circle is $[t_{v1} - (R_s + R_c)/c, t_{v2} + (R_s + R_c)/c]$ (Fig. 5b). Although these incoming rays are time-limited, their Hilbert transform in general has an infinitely long temporal tail. The signal that results from Hilbert transforming the incoming ray, which consists of the actual field and the ghost signal, is also illustrated in Fig. 5b. The FFT cannot be directly invoked to compute the Hilbert transform because the spectral sampling rate has to be extremely high to produce accurate temporal waveforms. To overcome this difficulty, it is noted that the infinite temporal tail of the Hilbert transforms of the incoming rays is caused by the rapid variation of their spectral content around DC. Hence, a multi-resolution spectral representation of the incoming rays is considered.

Next, specific steps are described for efficiently evaluating the Hilbert transform within the context of the PWTD algorithm. To facilitate the description, the temporal origin is shifted to the center of the incoming rays. Then, the signal to be Hilbert transformed is

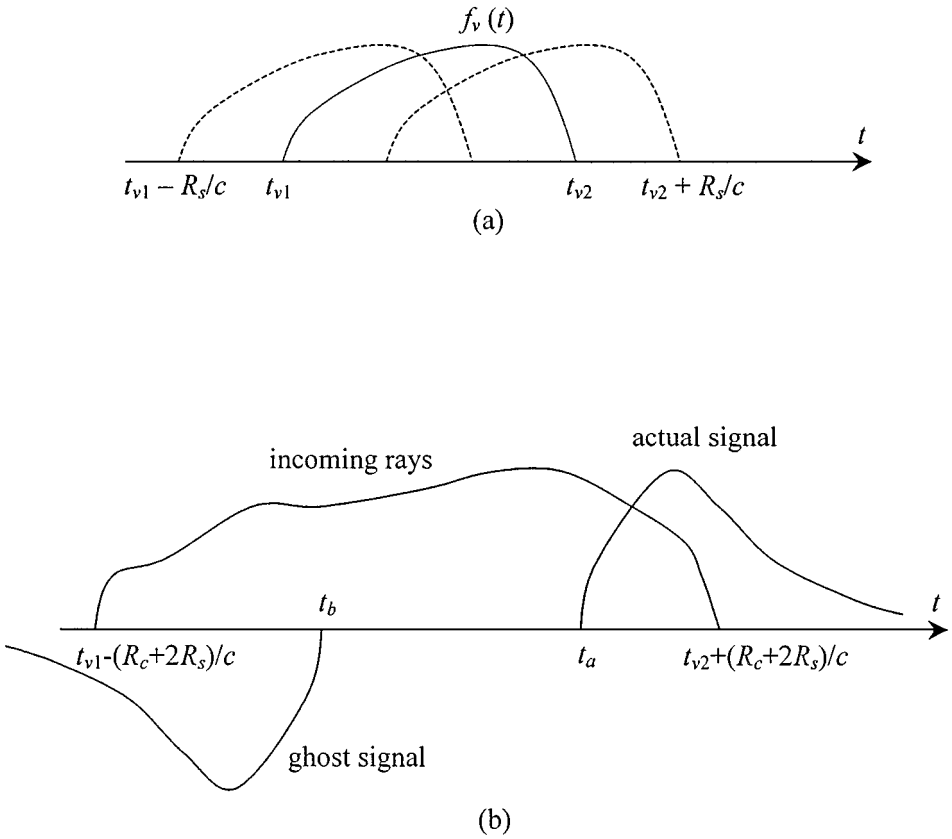

FIG. 5. Time sequence of plane wave expansion and Hilbert transform. (a) The relationship between subsignal and its outgoing rays; (b) The relationship of incoming rays and the results after Hilbert transform.

virtually time-limited within $[-t_f, t_f]$, and frequency-limited within $[-\omega_f, \omega_f]$, where $t_f = T_s/2 + (R_s + R_c)/c$.

First, the spectrum of the outgoing rays, sampled at a rate of $\Delta\omega = \pi/(\chi_3 t_f)$, is obtained using an FFT. This spectrum is then multiplied with that of the pertinent translation function given by Eq. (30). Here, $\chi_3 > 1$ is the *spectal oversampling rate.* The continuous spectrum can be recovered by convolving the sampled representation with the local spectral interpolant $P(\omega, t_f, \chi_3, p_s)$ (Appendix). This interpolant becomes vanishingly small and can be truncated when $|\omega| \geq p_s \Delta\omega$. The interpolation error can be controlled and made arbitrarily small by increasing p_s or χ_3. In what follows, however, it is assumed that $\chi_3 \leq 2$; hence, accuracy is attained by increasing p_s.

Second, the spectral samples are split into two sets. The first set consists of all samples with frequencies satisfying $|\omega| > p_s \Delta\omega$ and is termed the *high frequency set.* The second set consists of $2p_s + 1$ samples that reside in the frequency interval $[-p_s \Delta\omega, p_s \Delta\omega]$ and is termed the *low frequency set.* Similarly, the original spectrum is split into two spectra. The first spectrum results from convolving the high frequency set with $P(\omega, t_f, \chi_3, p_s)$ and is denoted $\mathrm{HF}^{(1)}$. The second spectrum results from convolving the low frequency set with $P(\omega, t_f, \chi_3, p_s)$ and is denoted $\mathrm{LF}^{(1)}$.

Third, the $\mathrm{LF}^{(1)}$ spectrum is sampled at a rate $\Delta\omega/2$. This operation results in $8p_s + 1$ samples. These samples are again split into high and low frequency sets. The former consists of the samples whose frequency satisfies $|\omega| > p_s \Delta\omega/2$, and the latter consists of $2p_s + 1$

samples that reside within the interval $[-p_s\Delta\omega/2,\ p_s\Delta\omega/2]$. The $\mathrm{LF}^{(1)}$ spectrum is then split into two spectra, $\mathrm{HF}^{(2)}$ and $\mathrm{LF}^{(2)}$, which result from convolving $P(\omega, 2t_f, \chi_3, p_s)$ with the above described sets of high and low frequency samples, respectively. This step is carried out recursively, a total of N_{rec} times. During the ith recursion, $i = 2, 3, \ldots, N_{rec}$, spectrum $\mathrm{LF}^{(i)}$ is sampled at a rate $\Delta\omega/2^i$, interpolated using $P(\omega, 2^i t_f, \chi_3, p_s)$, and split into high frequency spectrum $\mathrm{HF}^{(i+1)}$ and low frequency spectrum $\mathrm{LF}^{(i+1)}$. This recursive process above is illustrated in Fig. 6. The original spectrum is approximated by $N_{rec}+1$ spectra $\mathrm{HF}^{(1)}, \mathrm{HF}^{(2)}, \ldots, \mathrm{HF}^{(N_{rec})}$ and $\mathrm{LF}^{(N_{rec})}$.

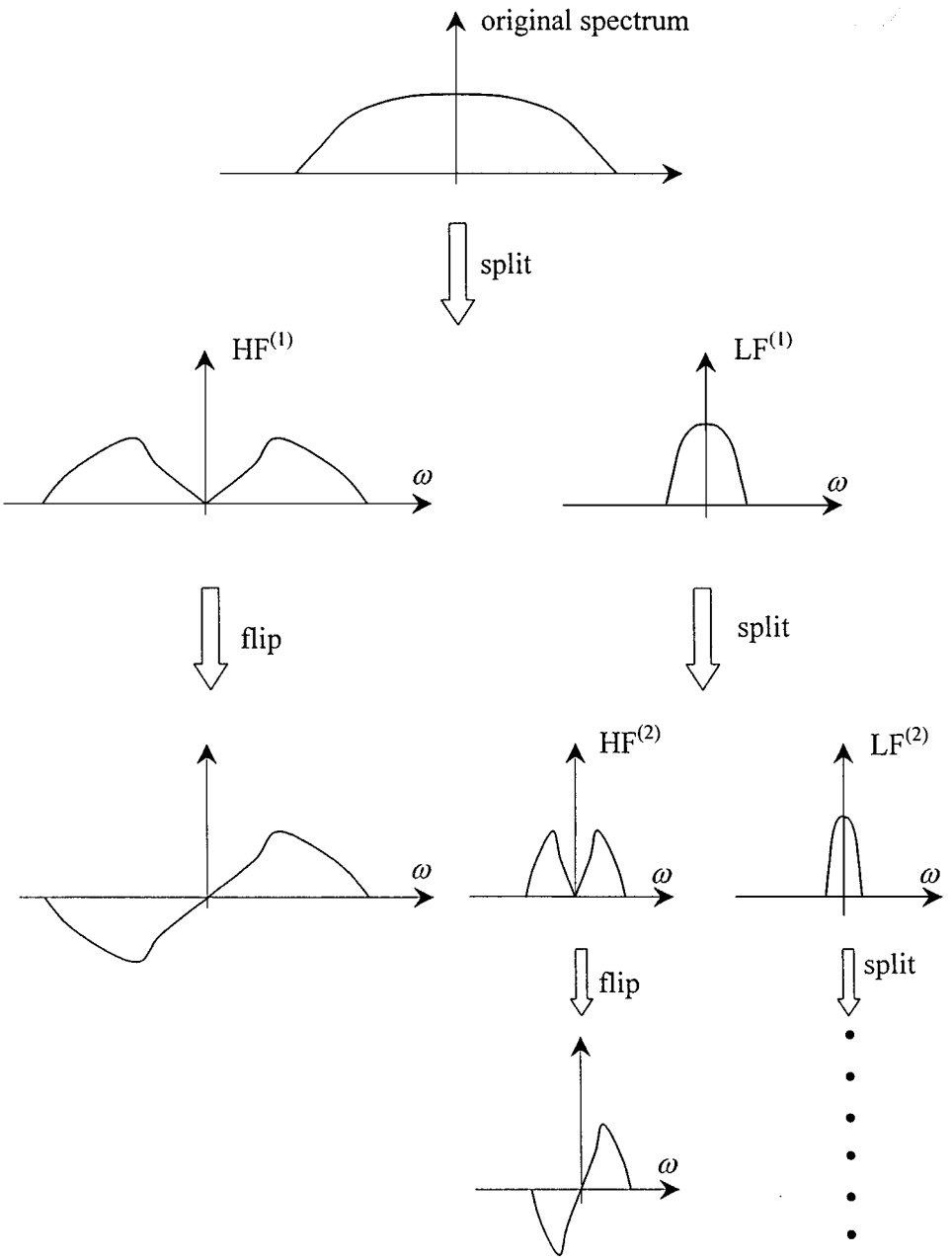

FIG. 6. Illustration of numerical evaluation of the Hilbert transform.

Finally, the Hilbert transform of the incoming rays can be approximated by Hilbert transforming the signals described by $\mathrm{HF}^{(i)}$, $i = 1, 2, \ldots, N_{rec}$. The temporal duration of the signal described by the spectrum $\mathrm{HF}^{(i)}$ is $[-2^{i-1}\chi_3 t_f, 2^{i-1}\chi_3 t_f]$. It is easily verified that the temporal duration of $\mathrm{sgn}(\omega)\mathrm{HF}^{(i)}$ is the same as that of $\mathrm{HF}^{(i)}$. Therefore, a conventional FFT scheme can be used to obtain the temporal signals corresponding to $\mathrm{sgn}(\omega)\mathrm{HF}^{(i)}$, $i = 1, 2, \ldots, N_{rec}$ (in practice, the use of the FFT is only required for $\mathrm{HF}^{(1)}$; the spectra $\mathrm{HF}^{(i)}$, $i = 2, 3 \ldots, N_{rec}$ are fully described by $8p_s + 1$ samples and their transform can be evaluated using a discrete Fourier transform). The sum of all these signals will approach the Hilbert transform of the outgoing rays with increasing N_{rec}. The energy of the error, which in essence is the energy of $\mathrm{LF}^{(N_{rec})}$, decreases exponentially fast with increasing N_{rec}.

2.3.4. Computational complexity. The computational cost associated with the evaluation of the observer field has four components, resulting from the four stage scheme described in Subsection 2.2.

(i) *Construction of outgoing rays.* At each temporal sampling point, the source is projected onto $O(N_\phi)$ outgoing rays. Each subsignal consists of $O(M_t)$ samples, and there are N_t/M_t subsignals. Therefore, the computational cost associated with the construction of the outgoing rays is

$$\begin{aligned} C_{ff}^1 &\propto (\#\text{ of rays}) \times (\#\text{ of subsignals}) \times (\text{cost of projecting a subsignal}) \\ &\propto N_\phi \frac{N_t}{M_t} M_t \\ &\propto N_\phi N_t. \end{aligned} \tag{32}$$

(ii) *Construction of incoming rays.* Translation occurs every M_t time steps. During translation, $O(N_\phi)$ outgoing rays are translated and $O(N_\phi)$ incoming rays constructed. The cost of one translation is $O(M_t \log M_t)$ provided that the convolution is effected using an FFT and that $T_s \propto R_c/c$. The latter assumption guarantees that the duration of the incoming rays is commensurate with that of the translation functions for the FFT to be an efficient means for carrying out the convolution. Therefore, the computational cost associated with the construction of the incoming rays is

$$\begin{aligned} C_{ff}^2 &\propto (\#\text{ of directions}) \times (\#\text{ of traslations per direction}) \times (\text{cost per translation}) \\ &\propto N_\phi \frac{N_t}{M_t} M_t \log M_t \\ &\leq N_\phi N_t \log N_t. \end{aligned} \tag{33}$$

(iii) *Construction of the Hilbert transforms of the incoming rays.* Each incoming ray is Hilbert transformed immediately following translation. The evaluation of the Hilbert transform requires an FFT to calculate the spectrum of the incoming ray, and additional filtering operations to decompose the spectrum of the incoming rays into a multiresolution basis. The cost of the FFT is $O(M_t \log M_t)$, because the lengths of incoming rays are proportional to M_t. The cost of decomposing the spectrum of the incoming rays into a multiresolution basis is of $O((N_{rec} + 1)(8p_s + 1))$ and is fixed. The inverse Fourier transform of the high frequency spectrum $\mathrm{HF}^{(1)}$ is carried out using an FFT, resulting in a cost

$O(M_t \log M_t)$. All the other high frequency spectra are stored, and their contribution to the Hilbert transformed incoming rays is computed using a discrete Fourier transform, resulting in a cost that is of $O((N_{rec}+1)(8p_s+1))$ per time step. In practice, their contributions are computed immediately upon translation for "all ray samples that reside over the observer circle," and new samples are added one by one, as the ray propagates through the observer sphere. The dominant cost in this sequence of operations is of $O(M_t \log M_t)$. In conclusion, the cost associated with the construction of the Hilbert transformed incoming rays is

$$\begin{aligned} C_{ff}^3 &\propto (\#\text{ of directions}) \times (\#\text{ of traslations per direction}) \times (\text{cost per Hilbert transform}) \\ &\propto N_\phi \frac{N_t}{M_t} M_t \log M_t \\ &\leq N_\phi N_t \log N_t. \end{aligned} \tag{34}$$

(iv) *Construction of the observer field.* Finally, the superposition of all Hilbert transformed incoming rays that impinge on a group are projected onto the individual observers and summed up. This operation is the counterpart of that described under (i); hence its cost scales as

$$C_{ff}^4 \propto N_\phi N_t. \tag{35}$$

The above cost estimates will be revisited in Section 3, where the computational complexity of PWTD augmented MOT solvers is studied.

2.4. *Numerical Validation and Discussion*

In this section, results from several numerical experiments are presented that validate the proposed algorithm. A single source and a single observation point are selected within the source and observation circles shown in Fig. 3, where $\phi_{12} = 90^\circ$, $R_c = 10$ m, and $R_s = 0.5$ m. The time signature of the source is Gaussian with $f_1^{(t)} = e^{-(t-t_0)^2/2\sigma^2}$, where $t_0 = 10^{-8}$, $\sigma = 10^{-9}$, and we safely choose $\omega_{max} = 7 \times 10^9$ rad/s. The following parameters determine the accuracy of the scheme: χ_1, χ_2, χ_3, p_t, p_s, and N_{rec}. In what follows, we fix $\chi_3 = 2$, and study the behavior of the error with respect to changes in the other parameters.

The field at the observation point computed by direct convolution (Eq. (2)), $A_D(t)$, is compared to that computed using the PWTD algorithm, $A_{PWTD}(t)$. The L2 error of the PWTD field is defined as

$$\text{L2 error} = \frac{\int_{-\infty}^{+\infty} [A_D(t) - A_{PWTD}(t)]^2 \, dt}{\int_{-\infty}^{+\infty} A_D^2(t) \, dt}. \tag{36}$$

In the first example, we pick $\chi_1 = 3.45$ and $p_t = 15$, which ensures that temporal sampling errors are negligible (errors of this type were studied withn the context of the 3D PWTD scheme, presented in [7]). The L2 error for different values of N_{rec} and p_s is shown in Fig. 7, where we assume $\chi_2 = 1$ and $N_\phi = 24$. As expected, for a fixed N_{rec}, the error reaches a lower limit as p_s is increased, and vice versa. Optimal combinations for N_{rec} and p_s can be determined from this figure.

Figure 8 shows the behavior of the instantaneous field error, $|A_D(t) - A_{PWTD}(t)|$ for $N_{rec} = 5$, 6, and 7, assuming that $p_s = 15$ and that all other parameters remain fixed at their

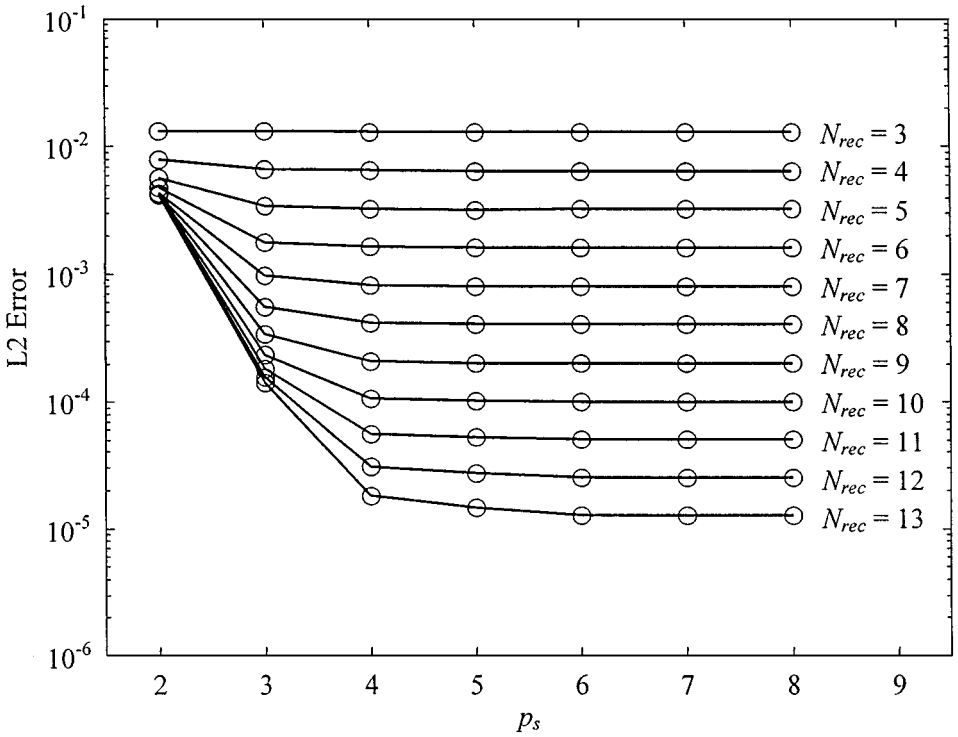

FIG. 7. The L2 error for different p_s and N_{rec}, when $\chi_1 = 3.45$, $p_t = 15$, and $\chi_2 = 1$.

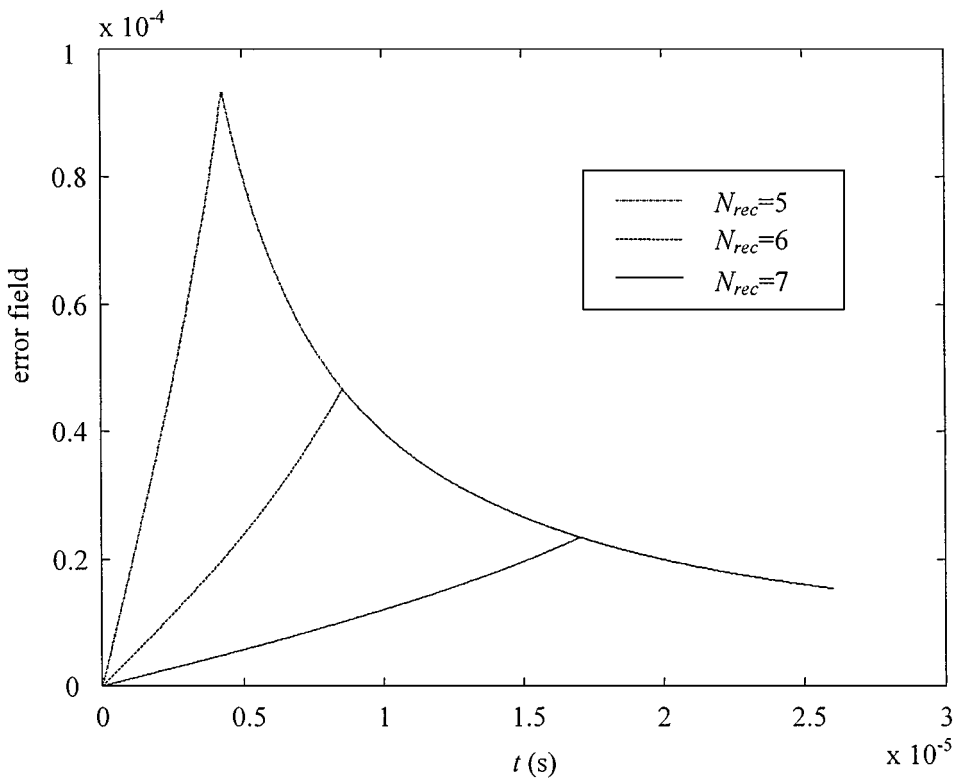

FIG. 8. Instantaneous error for $N_{rec} = 5$, 6, and 7, when $p_s = 15$, $\chi_1 = 3.45$, $p_t = 15$, and $\chi_2 = 1$.

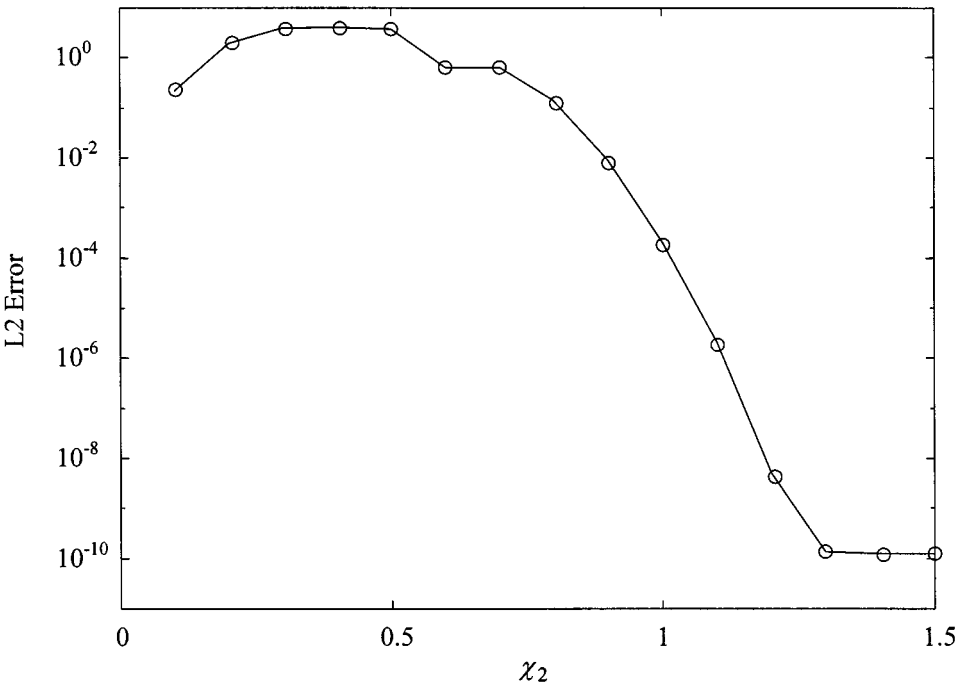

FIG. 9. L2 error for different χ_2, when $N_{rec} = 6$, $p_s = 7$, $\chi_1 = 3.45$, $p_t = 15$.

above values. For these parameter choices, the error is mainly due to the choice of N_{rec}, and Fig. 8 shows that the instantaneous error decreases when N_{rec} increases. The PWTD field vanishes after $T_t(N_{rec})$, and when $t > T_t(N_{rec})$, the error shown is nothing but the tail of the actual field. The instantaneous error reaches its maximum at $T_t(N_{rec})$. It can also be shown that when N_{rec} increases by one, $T_t(N_{rec})$ increases by about a factor of two, and the energy in the error field drops by a factor of two.

Next, the effect of the choice of χ_2 is studied.The time signature of the source is a modulated Gaussian pulse, $f_2(t) = \cos(\omega_0 t)e^{-(t-t_0)^2/2\sigma^2}$, where $\omega_0 = 10^9$ rad/s. The maximum frequency $\omega_{\max}$ is chosen to be 1.25×10^9 rad/s. When $N_{rec} = 6$, $p_s = 7$, and all the other parameters remain the same as before, the L2 error for different χ_2 is plotted in Fig. 9. It is observed that the error decreases rapidly with increasing χ_2, until the error becomes dictated by the other parameters.

Finally, we study the behavior of the error field throughout the observer circle. In this experiment, six sources with identical temporal signatures $f_1^{(t)}$ are distributed evenly over the perimeter of the source circle. In Fig. 10, the L2 error distribution is plotted throughout a 3 m by 3 m box, whose center coincides with that of the observation circle. The general logarithm of the L2 error is shown in different colors. Four different groups of parameters are chosen as in Table I. The minimum L2 errors in the observation box are also shown in the table. This experiment demonstrates that through the proper choice of parameters, the error introduced by the PWTD scheme can be made arbitrarily small.

3. COMPUTATIONAL COMPLEXITY OF PWTD ENHANCED MOT SCHEMES

This section describes two-level and multilevel PWTD enhanced MOT solvers. To this end, consider the 2D scatterer shown in Fig. 1. As discussed above, assuming that the

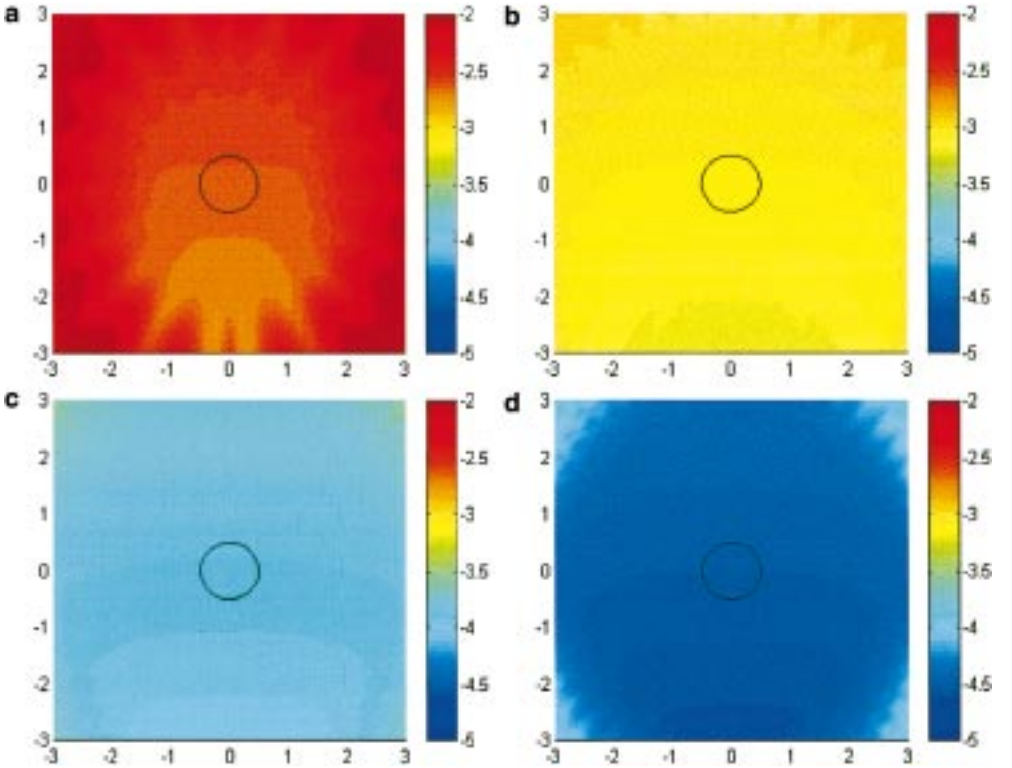

FIG. 10. The L2 error distribution within an observation box. The parameters are chosen as in Table I.

current on the scatterer is described in terms of N_s spatial unknowns for a total of N_t time steps, the computational cost associated with the analysis of this scattering problem using classical MOT schemes scales as $O(N_t^2 N_s^2)$. In the two-level PWTD enhanced MOT scheme, the scatterer is divided into a large number of small subscatterers and the PWTD scheme is invoked to evaluate interactions between the vast majority of subscatterer pairs. To arrive at a multilevel PWTD enhanced MOT scheme, this algorithm is cast in a divide and conquer framework. The proposed schemes are quite similar to frequency domain fast multipole algorithms [9, 10]. Subsections 4.1 and 4.2 describe two-level and multilevel PWTD enhanced MOT schemes, with computational complexities of $O(N_s^{1.5} N_t \log N_t)$ and $O(N_s N_t \log N_s \log N_t)$, respectively.

TABLE I
The Parameters and Minimum L2 Error in Fig. 10

	χ_1	χ_2	p_t	p_s	N_{rec}	Minimum L2 error
(a)	2.0	0.4	8	5	6	0.0017
(b)	3.0	0.6	12	6	7	6.07×10^{-4}
(c)	3.45	0.8	15	7	10	9.92×10^{-5}
(d)	3.45	1.0	15	8	12	2.51×10^{-5}

3.1. *Two-Level PWTD Enhanced MOT Scheme*

To efficiently evaluate the field due to transient sources that reside on a 2D object using the two-level PWTD algorithm, the object is enclosed into a fictitious square box, which is further subdivided into a large number of equally sized square boxes of circumscribing radius R_s. The set of sources that reside within a nonempty box is termed a group. If N_g denotes the total number of nonempty boxes, then the average number of sources per group is $M_s = N_s/N_g$. In addition, if the scatterer is discretized using a fixed number of basis functions per shortest wavelength in the incident pulse, then it can be shown that $M_s \propto R_s\omega_{max}/c$ and $N_\phi \propto M_s$ (see Eq. (27)). Let $R_{c,\alpha\alpha'}$ denote the distance between the centers of the boxes associated with groups α and α', $\alpha, \alpha' = 1, 2, \ldots, N_g$. A pair of groups (α, α') is termed a far-field pair if $R_{c,\alpha\alpha'} > \beta R_s$ and a near-field pair if $R_{c,\alpha\alpha'} \leq \beta R_s$ where β is a constant that is typically chosen in the range $3 \leq \beta \leq 6$. Let *FF* and *NF* denote the sets of all far- and near-field pairs, respectively. The numbers of far-field pairs scale as $O(N_g^2)$, and the number of near-field pairs as $O(N_g)$.

If two groups constitute a far-field pair, the 2D PWTD scheme is invoked to evaluate the field due to sources associated with one group at observers in the other group, and vice versa. Let $M_{t,min}$ denote the number of samples that describe the longest possible subsignal that can be used within the PWTD scheme when considering the far-field pair with the smallest $R_{c,\alpha\alpha'}$. It follows from Eqs. (19) and (23) that $M_{t,min}$ is given by

$$M_{t,min} = \min_{(\alpha,\alpha')\in FF}\{\lfloor (R_{c,\alpha\alpha'} - 2R_s)/(c\Delta t) - 2p_t + 1\rfloor\}. \tag{37}$$

If, for a given choice of p_t, Eq. (37) yields a negative $M_{t,min}$, then β needs to be increased to enlarge the near-field region. If Eq. (37) yields a positive $M_{t,min}$, then, at lest for a sufficiently large β, $M_{t,min}$ will be of $O(R_s\Delta t/c)$. The number of samples that feature in the PWTD based evaluation of fields exchanged between other far-field pairs is computed as

$$M_{t,\alpha\alpha'} = n_{\alpha\alpha'} M_{t,min}, \tag{38}$$

where the integer $n_{\alpha\alpha'}$ is given by

$$n_{\alpha\alpha'} = \left\lfloor \frac{(R_{c,\alpha\alpha'} - 2R_s)/(c\Delta t) - 2p_t + 1}{M_{t,min}} \right\rfloor. \tag{39}$$

Equation (38) guarantees that outgoing rays that are comprised of $O(M_{t,\alpha\alpha'})$ samples featuring in the PWTD based computation of fields exchanged between groups α and α' can be constructed by splicing $n_{\alpha\alpha'}$ shorter rays, each described by $O(M_{t,min})$ samples. This construction results in considerable computational savings because outgoing rays need not be constructed from scratch when considering different far-field pairs. With this proviso in mind, Eq. (38) also maximizes the length of the outgoing rays for each far-field pair. This implies that the convolution of the outgoing rays and translation functions can be effected using an FFT in $O(M_{t,\alpha\alpha'} \log M_{t,\alpha\alpha'})$ operations. Indeed, the translation functions span the temporal interval $[-R_{c,\alpha\alpha'}/c,\ R_{c,\alpha\alpha'}/c]$ (see Eq. (29)), which, in view of the above construction, can be shown to be proportional to $M_{t,\alpha\alpha'}\Delta t$. Therefore, both the outgoing rays and the translation fuctions are described by $O(M_{t,\alpha\alpha'})$ samples, and hence they can be convolved efficiently using an FFT.

The computational cost associated with the evaluation of all far-field interactions has four components, resulting from the four-stage scheme described above.

(i) *Construction of outgoing rays.* At each temporal sampling point, each and every source in every group is projected onto outgoing rays. Therefore, the computational cost associated with the construction of the outgoing rays is

$$\begin{aligned} C_{ff}^{TL,1} &\propto (\#\text{ of far-field groups}) \times (\#\text{ of source points per group}) \times C_{ff}^{1} \\ &\propto N_g M_s N_\phi N_t \\ &\propto \frac{N_s^2}{N_g} N_t. \end{aligned} \tag{40}$$

(ii) *Construction of incoming rays.* Incoming rays are constructed for all far-field pairs. Therefore, the computational cost associated with the construction of the incoming rays is

$$\begin{aligned} C_{ff}^{TL,2} &\propto \sum_{\alpha,\alpha'} C_{ff}^{2} \\ &\leq N_g^2 N_\phi N_t \log N_t \\ &\propto N_g N_s N_t \log N_t. \end{aligned} \tag{41}$$

(iii) *Construction of the Hilbert transforms of the incoming rays.* Hilbert transformed incoming rays are constructed immediately following a translation, i.e., following the construction of the incoming rays. Therefore, the cost associated with the construction of the Hilbert transformed incoming rays is

$$\begin{aligned} C_{ff}^{TL,3} &\propto \sum_{\alpha,\alpha'} C_{ff}^{3} \\ &\leq N_g^2 N_\phi N_t \log N_t \\ &\propto N_g N_s N_t \log N_t. \end{aligned} \tag{42}$$

(iv) *Construction of the observer field.* Finally, the superposition of all Hilbert transformed incoming rays that impinge on a group are projected onto the individual observers and summed up. This operation is the counterpart of that described under (i); hence its cost scales as

$$C_{ff}^{TL,4} \propto \frac{N_s^2}{N_g} N_t. \tag{43}$$

If two groups constitute a near-field pair, the PWTD scheme cannot be invoked directly. There reside approximately M_s sources in each group, and the M_s^2 interactions between sources and observers residing in these two groups need to be evaluated one by one. Although the computational cost for classically evaluating one such interaction scales as $O(N_t^2)$ (see Subsection 2.1), this cost can be reduced to $O(N_t \log N_t)$. This cost reduction is achieved by invoking the PWTD scheme for each near-field source-observer pair, in the limit $R_s \to 0$.

It follows that the computational cost for evaluating the near-field scales as

$$\begin{aligned} C_{nf}^{TL} &\propto \sum_{(\alpha,\alpha')\in NF} (\#\text{ of interactions for one near field pair}) \times C_{nf} \\ &\propto N_g M_s^2 N_t \log N_t \\ &\propto \frac{N_s^2}{N_g} N_t \log N_t. \end{aligned} \tag{44}$$

In conclusion, the total computational cost C^{TL} for a two-level scheme is

$$C^{TL} = C_{ff}^{TL,1} + C_{ff}^{TL,2} + C_{ff}^{TL,3} + C_{ff}^{TL,4} + C_{nf}^{TL}. \tag{45}$$

It can be verified that the choice $N_g \propto \sqrt{N_s}$ minimizes C^{TL}, and that with this choice of group size C^{TL} is of $O(N_s^{1.5} N_t \log N_t)$.

3.2. *Multilevel PWTD Enhanced MOT Scheme*

Cost savings beyond those achieved by the two-level PWTD enhanced MOT scheme result from casting the PWTD scheme into a divide and conquer framework, whereby groups are aggregated into larger entitites before translation.

As in the two-level scheme, the scatterer is enclosed in a square box. This box is recursively subdivided into four equally sized square boxes, a total of N_l times. A box that is subdivided into smaller boxes is termed the *parent* of the *child* boxes that result from the operation. The smallest box is termed a *level one* box, and the sources residing within a level one box are termed a *level one group*; higher level boxes and groups are defined similarly. For levels $l = 1, \ldots, N_l$, let $N_g(l)$ denote the number of groups (nonempty boxes), $M_s(l)$ the average number of sources in each group, $R_s(l)$ the circumscribing radius of a level l box, and $N_\phi(l)$ the number of directions required in the construction of the translation functions. It follows from the discussion of the two-level scheme that $N_\phi(l) \propto R_s(l)\omega_{max}/c \propto M_s(l)$. It is obvious that $R_s(l+1) = 2R_s(l)$, that $M_s(l+1) = 2M_s(l)$, and that $N_g(l+1) = N_g(l)/2$. In addition, N_l is always chosen such that a level one group contains $O(1)$ sources, which implies that $N_l \approx \log N_s$, that $N_g(1) \propto N_s$, that $M_s(1) \propto 1$, that $N_g(N_l) \propto 1$, and that $M_s(N_l) \propto N_s$; it then also follows that $N_g(l) M_s(l) \propto N_s$.

Next, consider all group pairs (α, α'), where both α and α' reside at the same level. A pair (α, α') is termed a *level l far-field pair* provided that $R_{c,\alpha\alpha'} > \beta R_s(l)$ and that their respective parent groups do not from a level $l+1$ far-field pair. A pair of groups that resides at level one and whose respective parent groups do not form a level two far-field pair is termed a *near-field pair*. The number of far-field pairs at level l is proportional to $N_g(l)$, and the number of near-field pairs is proportional to $N_g(1) \propto N_s$. In what follows, let $FF(l)$ and NF denote the sets of all level l far- and near-field pairs, respectively. The multilevel PWTD scheme will effect interactions between all far-field pairs.

As in the two-level algorithm, the PWTD scheme is invoked to evaluate all interactions involving far-field pairs. In the multilevel scheme, the number of samples that describe the longest possible subsignal that can be used within the PWTD scheme when considering the nearest by level one far-field pairs, $M_{t,min}(1)$, follows from Eq. (37) with FF replaced by $FF(1)$. Defining $M_{t,min}(l+1) = 2M_{t,\,min}(l)$, the number of samples that feature in the PWTD based evaluation of fields exchanged between level l far-field pair (α, α'), $M_{t,\alpha\alpha'}$ is computed using Eqs. (38) and (39) with $M_{t,min}$ replaced by $M_{t,min}(l)$. Like in the two-level

algorithm, this construction again guarantees that outgoing rays at level l can be constructed effectively by splicing fundamental rays described by $O(M_{t,min}(l))$ samples. In addition, fundamental rays at level $l+1$ can be obtained by splicing two fundamental outgoing rays at level l.

Fundamental outgoing rays are constructed as follows. At level 1, the outgoing rays are constructed by directly projecting sources onto a set of directions, as in the two-level algorithm. At higher levels, outgoing rays are constructed by (i) splicing two consecutive outgoing rays, (ii) interpolating these rays to a set of directions that is roughly twice as dense ($N_\phi(l+1) = 2N_\phi(l)$), and (iii) translating these rays from the box center to that of its parent box (this means that every ray is time delayed or advanced based on its direction w.r.t. the vector connecting the box centers involved). A similar process, but in reverse order, is carried out when going down the multilevel tree: incoming rays at lower levels are constructed from those at higher levels via ray resection (the opposite of ray splicing) and anterpolation (or filtering) to reduce the number of angular components.

As always, the cost for computing all interactions between far-field elements comprises four components:

(i) *Construction of outgoing rays.* At level 1, outgoing rays are constructed as in the two-level scheme. At level $l > 2$, the outgoing rays are constructed as described above. The cost of splicing, interpolating, and projecting at all levels is

$$\begin{aligned} C_{ff}^{ML,1} &\propto \sum_{(\alpha,\alpha')\in FF(1)} C_{ff}^1 + \sum_{l=2}^{N_l} \sum_{(\alpha,\alpha')\in FF(l)} \{(\#\text{ of directions at level } l) \\ &\quad \times (\#\text{ of projections at level } l) \times (\text{length of each projection at level } l)\} \\ &\propto N_g(1)N_\phi(1)M_{t,min}(1) + \sum_{l=2}^{N_l} \left\{ N_g(l)N_\phi(l)\frac{N_t}{M_{t,min}(l)} M_{t,min}(l) \right\} \\ &\propto \sum_{l=1}^{N_l} \{N_g(l)M_s(l)N_t\} \\ &\propto N_s N_t \log N_s. \end{aligned} \tag{46}$$

(ii) *Construction of incoming rays.* At level l, translations for a far-field group pair (α, α') occur every $M_{t,\alpha,\alpha'}(l)$ time steps with a complexity of $O(N_\phi(l)N_t \log N_t)$, as discussed in the two-level case, hence

$$\begin{aligned} C_{ff}^{ML,2} &\propto \sum_{l=1}^{N_l} \left\{ \sum_{(\alpha,\alpha')\in FF(l)} C_{ff}^2 \right\} \\ &\le \sum_{l=1}^{N_l} \left\{ \sum_{(\alpha,\alpha')\in FF(l)} N_\phi(l)N_t \log N_t \right\} \\ &\propto \sum_{l=1}^{N_l} \{N_g(l)M_s(l)N_t \log N_t\} \\ &\propto \sum_{l=1}^{N_l} \{N_s N_t \log N_t\} \\ &\propto N_s N_t \log N_s \log N_t. \end{aligned} \tag{47}$$

(iii) *Construction of Hilbert transformed incoming rays.* As in the two-level algorithm, a Hilbert transform is carried out after each translation. The computational cost of this operation scales as

$$\begin{aligned} C_{ff}^{ML,3} &\propto \sum_{l=1}^{N_l} \left\{ \sum_{(\alpha,\alpha')\in FF(l)} C_{ff}^{3} \right\} \\ &\leq \sum_{l=1}^{N_l} \left\{ \sum_{(\alpha,\alpha')\in FF(l)} N_\phi(l) N_t \log N_t \right\} \\ &\propto \sum_{l=1}^{N_l} \{N_g(l) M_s(l) N_t \log N_t\} \\ &\propto \sum_{l=1}^{N_l} \{N_s N_t \log N_t\} \\ &\propto N_s N_t \log N_s \log N_t. \end{aligned} \tag{48}$$

(iv) *Construction of observer field.* The cost of this operation scales as that of $C_{ff}^{ML,1}$:

$$C_{ff}^{ML,4} \propto N_s N_t \log N_s. \tag{49}$$

For a near-field pair (α, α'), the interactions between the spatial unknowns within them should be evaluated in a point to point manner. Similar to the near-field part in Subsection 4.1, the cost for computing one near-field interaction can be reduced to $O(N_t)$ by using a Hilbert transformed plane wave expansion. The computational cost for near-field evaluations therefore scales as

$$\begin{aligned} C_{nf}^{ML} &\propto \sum_{(\alpha,\alpha')\in NF} (\#\text{ of interactions}) \times (\text{cost of computing one interaction}) \\ &\propto N_g(1) M_s^2(1) N_t \\ &\propto N_s N_t. \end{aligned} \tag{50}$$

It is seen that $C_{ff}^{ML,2}$ and $C_{ff}^{ML,3}$ scale worse than all other costs involved in the multilevel PWTD. It is concluded that the total cost associated with a multilevel PWTD enhanced MOT solver, viz.

$$C^{ML} = C_{ff}^{ML,1} + C_{ff}^{ML,2} + C_{ff}^{ML,3} + C_{ff}^{ML,4} + C_{nf}^{ML}, \tag{51}$$

scales as $O(N_s N_t \log N_s \log N_t)$.

4. CONCLUSIONS

This paper outlined the 2D PWTD algorithm which permits the rapid evaluation of linear transient wave fields generated by 2D source configuration. A four-stage aggregation-translation-disaggration scheme was developed that relies on an expansion of the wave field in terms of a Hilbert transformed plane wave spectrum. Numerical experiments show that the error of this algorithm can be controlled and be made arbitrarily small by the proper

choice of parameters. The proposed PWTD scheme can be coupled to classical time domain integral equation based solvers. For a two-dimensional surface scattering problem modeled in terms of N_s spatial unknowns for a total of N_t time steps, two-level and multilevel integral equation based solvers, developed on the basis of the method proposed in this paper, exhibit the computational complexities of $O(N_s^{1.5} N_t \log N_t)$ and $O(N_s N_t \log N_s \log N_t)$, respectively. Therefore, these PWTD schemes make feasible the rapid integral equation based analysis of scattering from large surfaces. These schemes can also be used in the construction of fast global absorbing boundary condition for truncating grids used in the finite difference modeling of wave phenomena. We are currently studying these and various other applications of these schemes.

APPENDIX: VARIOUS DEFINITIONS

In this paper, $u(x)$ and $P(x, y)$ denote Heaviside step and pulse functions defined as

$$u(x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0, \end{cases} \tag{A.1}$$

$$P(x, y) = u(x + y) - u(x - y). \tag{A.2}$$

In addition, $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ define the Fourier transform and its inverse, respectively,

$$\begin{cases} \tilde{f}(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t}\, dt \\ f(t) = \mathcal{F}^{-1}\{\tilde{f}(\omega)\} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{f}(\omega) e^{j\omega t}\, d\omega. \end{cases} \tag{A.3}$$

The Hilbert transform is denoted by $\mathcal{H}(\cdot)$,

$$\mathcal{H}\{f(t)\} = \frac{1}{\pi} \fint_{-\infty}^{+\infty} \frac{f(\alpha)}{t - \alpha}\, d\alpha, \tag{A.4}$$

where the symbol $\fint$ stands for principle value integral. The Hilbert transform can be also be expressed as

$$\mathcal{H}\{f(t)\} = \mathcal{F}^{-1}\{-j \operatorname{sgn}(\omega) \mathcal{F}[f(t)]\}, \tag{A.5}$$

where $\mathrm{Im}(\cdot)$ denotes the "imaginary part of," and $\operatorname{sgn}(x)$ is the sign function

$$\operatorname{sgn}(x) = \begin{cases} 1, & x > 0 \\ -1, & x < 0. \end{cases} \tag{A.6}$$

Finally, the following interpolation function is a variant of the one originally proposed by Knab [15],

$$P(t, \omega_{max}, \chi, p) = \frac{\omega_0}{\omega_f} \operatorname{sinc}(\omega_0 t) \frac{\sin\left\{\Omega p \Delta t \sqrt{(t/p\Delta t)^2 - 1}\right\}}{\sinh(\Omega p \Delta t) \sqrt{(t/p\Delta t)^2 - 1}}, \tag{A.7}$$

where

$$\begin{aligned} \omega_f &= \chi \omega_{max} \\ \Delta t &= \frac{\pi}{\omega_f} \\ \omega_0 &= \omega_{max} \frac{(\chi + 1)}{2} \\ \Omega &= \omega_{max} \frac{(\chi - 1)}{2}. \end{aligned} \tag{A.8}$$

The function $P(t, \omega_{max}, \chi, p)$ can be used to locally interpolate functions of bandwidth ω_{max} in terms of $2p+1$ samples. $P(t, \omega_{max}, \chi, p)$ is virtually time-limited and vanishes for $|t| > p\Delta t$. $P(t, \omega_{max}, \chi, p)$ is also band-limited to ω_f. The truncation error due to local interpolation decreases exponentially fast with increasing p.

ACKNOWLEDGMENTS

This research was supported by the Air Force Office for Scientific Research via the MURI program under Contract F49620-96-1-0025, the Defense Advanced Research Project Agency under Contract F30602-972-0328, and the Gebze Institute of Technology, Turkey.

REFERENCES

1. L. B. Felsen (Ed.), *Transient Electromagnetic Fields* (Springer-Verlag, Berlin, 1976).
2. M. Tygel and P. Hubral, *Transient Waves in Layered Media* (Elsevier, Amsterdam, 1987).
3. N. J. Damaskos, R. T. Brown, J. R. Jameson, and P. L. E. Uslenghi, Transient scattering by resistive cylinders, *IEEE Trans. Antennas Propag.* **33**, 21 (1985).
4. D. A. Vechinski and S. M. Rao, Transient scattering from two-dimensional dielectric cylinders of arbitrary shape, *IEEE Trans. Antennas Propag.* **40**, 1054 (1992).
5. D. A. Vechinski and S. M. Rao, Transient scattering from dielectric cylinders: E-field, H-field, and combined field solutions, *Radio Sci.* **27**, 611 (1992).
6. A. G. Tijhuis, Toward a stable marching-on-in-time method for two-dimensional transient electromagnetic scattering problems, *Radio Sci.* **19**, 1311 (1984).
7. A. A. Ergin, B. Shanker, and E. Michielssen, Fast evaluation of three-dimensional transient wave fields using diagonal translation operators, *J. Comput. Phys.* **146**, 157 (1998).
8. A. A. Ergin, B. Shanker, and E. Michielssen, The plane wave time domain algorithm for the fast analysis of transient wave phenomena, *IEEE Antennas Propag. Mag.*, in press.
9. R. Coifman, V. Rokhlin, and S. Wandzura, The fast mltipole method for the wave equation: A pedestrian prescription, *IEEE Antennas Propag. Mag.* **35**, 7 (1993).
10. J. M. Song and W. C. Chew, Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering, *Microwave Opt. Technol. Lett.* **10**, 14 (1995).
11. W. Y. Crutchfield, *Boundary Decomposition: A Fast Method for Simulating Wave Propagation*, Exxon Production Research Company, 1989 (unpublished).
12. B. Alpert, L. Greengard, and T. Hagstrom, Rapid evaluation of nonreflecting boundary kernels for time-domain wave propagation, *SIAM J. Numer. Anal.*, in press.
13. O. M. Bucci and G. Franceschetti, On the spatial bandwidth of scattered fields, *IEEE Trans. Antennas Propag.* **35**, 1445 (1987).
14. O. M. Bucci and G. Franceschetti, On the degrees of freedom of scattered fields, *IEEE Trans. Antennas Propag.* **37**, 918 (1989).
15. J. J. Knab, Inerpolation of band-limited functions using the approximate prolate series, *IEEE Trans. Inform. Theory* **25**, 717 (1979).