

Audibility of dispersion error in room acoustic finite-difference time-domain simulation as a function of simulation distance

Jukka Saarelna, Jonathan Botts, Brian Hamilton, and Lauri SaviojaMV

Citation: [The Journal of the Acoustical Society of America](#) **139**, 1822 (2016); doi: 10.1121/1.4945746

View online: <http://dx.doi.org/10.1121/1.4945746>

View Table of Contents: <http://asa.scitation.org/toc/jas/139/4>

Published by the [Acoustical Society of America](#)

Audibility of dispersion error in room acoustic finite-difference time-domain simulation as a function of simulation distance

Jukka Saarelna,^{a)} Jonathan Botts, Brian Hamilton, and Lauri Savioja

Department of Computer Science and Engineering, Aalto University School of Science, Espoo, Finland

(Received 24 July 2015; revised 21 December 2015; accepted 18 March 2016; published online 13 April 2016)

Finite-difference time-domain (FDTD) simulation has been a popular area of research in room acoustics due to its capability to simulate wave phenomena in a wide bandwidth directly in the time-domain. A downside of the method is that it introduces a direction and frequency dependent error to the simulated sound field due to the non-linear dispersion relation of the discrete system. In this study, the perceptual threshold of the dispersion error is measured in three-dimensional FDTD schemes as a function of simulation distance. Dispersion error is evaluated for three different explicit, non-staggered FDTD schemes using the numerical wavenumber in the direction of the worst-case error of each scheme. It is found that the thresholds for the different schemes do not vary significantly when the phase velocity error level is fixed. The thresholds are found to vary significantly between the different sound samples. The measured threshold for the audibility of dispersion error at the probability level of 82% correct discrimination for three-alternative forced choice is found to be 9.1 m of propagation in a free field, that leads to a maximum group delay error of 1.8 ms at 20 kHz with the chosen phase velocity error level of 2%.

© 2016 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

[<http://dx.doi.org/10.1121/1.4945746>]

[MV]

Pages: 1822–1832

I. INTRODUCTION

Simulation methods for the prediction of acoustic characteristics of rooms are common tools in the design process of critical listening rooms, performance spaces, and regular housing. Due to the limitations of geometric prediction methods, several wave-based methods have been proposed for room acoustic prediction^{1,2} although the computational load has been the limiting factor for a full audible bandwidth simulation on a large scale.

During recent years, finite difference methods have gained interest in room acoustic research. The finite-difference time-domain (FDTD) method has been studied for room acoustic prediction by several different authors. The FDTD method is simple to implement, and explicit FDTD schemes are easily parallelizable. The parallel nature of the method makes it possible to efficiently use distributed computing in the time stepping, and therefore allowing large domain sizes, which has made FDTD a viable option for wide bandwidth simulation. A general downside of the FDTD method is the numerical dispersion error. The dispersion error in FDTD schemes is such that the high frequency components travel with a different phase velocity than the low frequency components due to a non-linear phase response of the update operator. The dispersion error is also dependent on the direction of the propagating wave. This leads to difficulties in a possible correction of the error³ in the simulated response because different propagation paths have different dispersion characteristics. Therefore a certain amount of dispersion error is inevitable in the simulated responses.

Several different FDTD schemes have been proposed for room acoustic simulation. Most finite difference methods used for room acoustic simulation are explicit time-stepping schemes using a different approximation of the Laplacian. The trade-off is between more isotropic, accurate, and computationally intensive schemes and schemes that are easy to formulate, parallelize, and extend to general geometries. Schemes that are often found in comparison studies are *standard rectilinear* (SRL), *interpolated wideband* (IWB), and *close-cubic packed* (CCP).⁶ Analytic comparisons of different explicit schemes have been carried out by several authors,^{5,6} and implementation specific considerations have been studied.^{7,8} SRL is often preferred due to its simplicity, IWB because it has the highest simulation bandwidth, and the CCP scheme for its computational efficiency when implemented on the *face-centered cubic* (FCC) lattice.⁵

It is common for compact explicit schemes that the phase velocity is real and monotonically decreasing as a function of real wavenumber, but the behaviour of the phase velocity is slightly different as a function of temporal frequencies up to the Nyquist. The real and imaginary parts of the relative phase velocity of the SRL, CCP, and IWB schemes are illustrated in Fig. 1 as a function of normalized (normalized to the sampling frequency; this holds throughout this work, with the exception of Fig. 5) temporal frequency, for the worst-case directions of propagation respective to each scheme. It can be seen that the phase velocity profiles are generally complex, and the regions with non-zero imaginary parts are accompanied by increasing phase velocity. The point of transition from real to complex phase velocity is marked by a “cutoff frequency,” above which frequencies are attenuated exponentially.⁹ This cutoff frequency is

^{a)}Electronic mail: jukka.saarelna@aalto.fi

usually interpreted as the high frequency limit of the simulation bandwidth of the given scheme and therefore the complex part is neglected in phase velocity figures.⁶ It can be observed that the phase velocity of the SRL scheme has the lowest cutoff frequency and steepest decrease in the real part of the phase velocity. The IWB scheme has the cutoff frequency at the Nyquist frequency, but a relatively similar phase velocity profile with the CCP scheme up to the normalized frequency of 0.2.

The aim of the current study is to measure the maximum distance in the simulation domain that can be used between a planar source and a receiver position in a free field before the dispersion error is noticeable in comparison to an error-free reference signal. The measurement is made using an adaptive psychometric procedure. The motivation for the study is that to the present authors knowledge, no study has yet addressed the absolute level of perceivable dispersion error in room acoustic FDTD simulation although the method has been proposed for auralization purposes. For such applications, the knowledge of the perceptual threshold for the absolute error level is important. The sampling frequency is fixed to maintain a constant phase velocity error of 2% at a chosen bandwidth limit of 20 kHz. The error percentage is therefore at its maximum value at 20 kHz and decreasing monotonically towards the temporal frequency of 0 Hz (DC). The error percentage is chosen to be relatively low so that the sampling frequencies of the simulations would remain within practical limits. The dispersion error is introduced to the stimulus by evaluating impulse responses of different propagation distances in the simulation domain using the dispersion relations of each scheme. The usage of the dispersion relation to introduce the dispersion error to the stimulus signal is motivated by the computational cost of the FDTD simulation. The proposed dispersion filter can be evaluated for a wide range of source-receiver distances in

the time range of several hundred milliseconds, whereas a simulation of a single distance condition may take up to several hours to compute at the used sampling frequencies. As the current study is considering only the free field propagation in the direction of worst-case error, this approach is found valid and convenient. Additionally, due to the plane wave assumption, no magnitude deviation is present in the stimuli signal, and therefore it is possible to measure only the audibility of the all-pass characteristics of the error. Three different non-staggered FDTD schemes are evaluated, SRL, IWB and CCP, at their respective stability limits.

II. RELATED RESEARCH

Several authors have studied the perception of dispersion error in FDTD simulation with varying approaches. Southern *et al.*¹⁰ conducted a listening test where the dispersion errors in the side-diagonal and on-axis directions in a three-dimensional (3D) SRL scheme were compared. The authors used five different propagation distances and five different low-pass filter cutoff frequencies for each sample pair corresponding to the waveforms from side-diagonal and on-axis propagation directions. A hard point source¹¹ was used as the excitation. Sound samples used in the test were a trombone and a violin/cello phrases. It was concluded that the test subjects were able to discriminate between the two dispersion error levels when the low-pass filter had a normalized cutoff frequency of 0.12–0.15 or higher. The sampling frequency used in the simulation was 5000 Hz. No significant effect was found between the different propagation distances. Magnitude deviations between each sound sample pair varied from 0.7 to 8 dBs at the normalized cutoff frequencies of 0.06–0.18 used in the study, respectively. The corresponding group delay value differences between the two directions for plane wave propagation for the different distances used in the study were between 0.02 ms of the

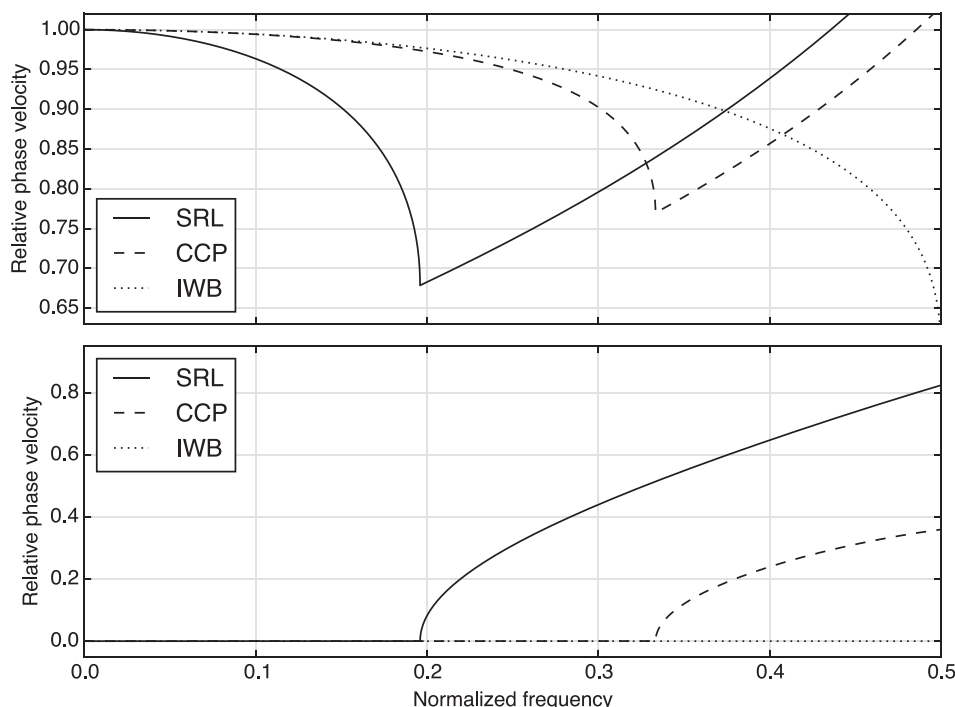


FIG. 1. The relative phase velocity of SRL, CCP, and IWB schemes as a function of normalized frequency. In the top plot is the real part of the phase velocity and in the bottom plot is the imaginary part. The value of the phase velocity is monotonically decreasing for real wavenumbers. Frequencies that have complex phase velocity are attenuated exponentially (Ref. 9).

lowest normalized cutoff frequency of 0.06 at the distance of 0.83 m and 16.08 ms of the normalized cutoff frequency of 0.18 and distance of 11.80 m. As the study compares two signals containing different degrees of error rather than a signal containing error to an error-free reference, it is inconclusive about the absolute error level that can be perceived. Although the authors speculate that the perceptual differences are due to the difference in the magnitude, the combined effect of group delay and magnitude error cannot be ruled out. In the current study the magnitude difference is not present due to a plane wave assumption.

López *et al.*¹² conducted listening tests where the subjects ability to discriminate between different mesh resolutions was measured in digital waveguide mesh simulation using simulated room responses. A room with a volume of 327 m³ and a constant frequency independent reflection coefficient of 0.8 was used in the simulation. Sampling frequencies of 20, 30, and 40 kHz were used. The sound samples used in the test were 15 s long male and female speech samples that were bandlimited to 5 kHz. It was concluded that the participants could not discriminate between the responses attained with the sampling frequencies 30 and 40 kHz. Twenty percent of the participants were available to discriminate between the simulations run with the sampling rates 20 and 30 kHz in the case of the male speech sample, and 30% in the case of female speech in an ABX comparison task. The study compares responses that include surface reflections that have different dispersion characteristics due to varying reflection paths and path lengths. The result is hard to interpret in the sense of absolute level of perceivable dispersion error, and therefore a more direct experiment design is proposed in this study.

The dispersion error in FDTD is characterized by the non-linear phase response of the update operator. The phase response of the operator manifests itself as strong frequency-dependent group and phase delay. Several existing studies of the perceptual threshold of group delay distortion in audio signals have been conducted^{13–17} which can give an indication of the perceptual threshold of the dispersion error found in FDTD simulation. In these studies, all-pass filters were used to generate specific group delay values at given frequency bands. The results of the studies agree that the perceptual limit of group delay is close to 2 ms in different frequency bands varying from 1 to 12 kHz. The group delay error of 2 ms corresponds to a 9.8 m propagation in the studied FDTD schemes at 20 kHz where the phase velocity error is fixed to 2%.

III. NUMERICAL DISPERSION

The partial differential equation of interest in room acoustics is the linear wave equation, which in 3D Cartesian coordinate system is given by

$$\frac{\partial^2 p}{\partial t^2} = c^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right), \quad (1)$$

where $p = p(x, y, z, t)$ is the acoustic pressure and c is the speed of sound, taken here to be 344 m/s. Equation (1) can be discretized by substituting the partial derivatives with finite differences. A family of compact explicit FDTD schemes for wave equation follows:^{6,18}

$$\delta_t^2 p_{k,l,m}^n = \lambda^2 [(\delta_x^2 + \delta_y^2 + \delta_z^2) + a(\delta_x^2 \delta_y^2 + \delta_x^2 \delta_z^2 + \delta_y^2 \delta_z^2) + b(\delta_x^2 \delta_y^2 \delta_z^2)] p_{k,l,m}^n, \quad (2)$$

where a and b are coefficient specific for each scheme, and $\lambda = c\Delta t/\Delta x$ denotes the Courant number, where Δx is the spatial step size and Δt the sampling interval. The coefficients a and b get values $a=0$, $b=0$ for the SRL scheme, $a=1/4$, $b=0$ for the CCP scheme, and $a=1/4$, $b=1/16$ for the IWB scheme. In this study, the Courant number is limited for stable free-field time stepping. The difference operators δ_t^2 , δ_x^2 , δ_y^2 , and δ_z^2 are defined as

$$\begin{aligned} \delta_t^2 p_{k,l,m}^n &= p_{k,l,m}^{n+1} - 2p_{k,l,m}^n + p_{k,l,m}^{n-1}, \\ \delta_x^2 p_{k,l,m}^n &= p_{k+1,l,m}^n - 2p_{k,l,m}^n + p_{k-1,l,m}^n, \\ \delta_y^2 p_{k,l,m}^n &= p_{k,l+1,m}^n - 2p_{k,l,m}^n + p_{k,l-1,m}^n, \\ \delta_z^2 p_{k,l,m}^n &= p_{k,l,m+1}^n - 2p_{k,l,m}^n + p_{k,l,m-1}^n, \end{aligned} \quad (3)$$

where $p_{k,l,m}^n = p(x, y, z, t)$, with $x = k\Delta x$, $y = l\Delta x$, $z = m\Delta x$, and $t = n\Delta t$. Assuming plane wave propagation, solutions on the grid are defined by the following dispersion relation that directly follows from Eq. (2):

$$\begin{aligned} s_t &= \lambda^2 [G(\hat{k}_x, \hat{k}_y, \hat{k}_z)], \\ G(\hat{k}_x, \hat{k}_y, \hat{k}_z) &= s_x + s_y + s_z - 4a(s_x s_y + s_x s_z + s_y s_z) \\ &\quad + 16bs_x s_y s_z, \end{aligned} \quad (4)$$

where ω is the angular frequency, and variables $s_t = \sin^2[\omega(\Delta t/2)]$, $s_x = \sin^2[\hat{k}_x(\Delta x/2)]$, $s_y = \sin^2[\hat{k}_y(\Delta x/2)]$, $s_z = \sin^2[\hat{k}_z(\Delta x/2)]$. Variables $\hat{k}_x = \hat{k} \cos \theta \cos \phi$, $\hat{k}_y = \hat{k} \sin \theta \cos \phi$, and $\hat{k}_z = \hat{k} \sin \phi$ are the wavenumber components where θ and ϕ represent the azimuth and elevation angles of the direction of propagation, respectively. By solving the variable ω from Eq. (4), the numerical dispersion relation for temporal frequency as a function of wavenumber components takes the form

$$\omega(\hat{k}_x, \hat{k}_y, \hat{k}_z) = \frac{2}{\Delta t} \arcsin \left(\lambda \sqrt{G(\hat{k}_x, \hat{k}_y, \hat{k}_z)} \right). \quad (5)$$

The relative phase velocity can be expressed as the ratio of the angular frequency and numerical wavenumber

$$\hat{v}(\hat{k}_x, \hat{k}_y, \hat{k}_z) = \frac{\omega(\hat{k}_x, \hat{k}_y, \hat{k}_z)}{c \sqrt{\hat{k}_x^2 + \hat{k}_y^2 + \hat{k}_z^2}}. \quad (6)$$

The relative phase velocity is plotted as a function of wavenumber components in Fig. 2. It can be observed that the maximum phase velocity error occurs in the axial direction in the case of the SRL scheme, and in the diagonal direction for CCP and IWB schemes. In the direction on minimum

error, all the schemes have linear dispersion relation at the stability limit, which means that the phase velocity is constant for all frequencies.

A. Phase velocity as a function of frequency

The directions in which most errors occur are the axial and diagonal directions as can be observed from Fig. 2. The analytic form of the phase velocity as a function of temporal frequency for these two cases can be derived from Eq. (4).

In an on-axis direction, only one of the wavenumber components is non-zero. By picking the axial direction as \hat{k}_x , it follows that $\hat{k}_y = \hat{k}_z = 0$, and Eq. (4) takes the form

$$\sin^2\left(\omega \frac{\Delta t}{2}\right) = \lambda^2 \sin^2\left(\hat{k}_x \frac{\Delta x}{2}\right), \quad (7)$$

from which the wavenumber component can be solved as a function of frequency

$$\hat{k}_x = \frac{2}{\Delta x} \arcsin\left[\frac{1}{\lambda} \sin\left(\omega \frac{\Delta t}{2}\right)\right]. \quad (8)$$

This form applies for all the schemes studied in this work although there is no phase velocity error in CCP and IWB schemes in the axial direction at the stability limit $\lambda = 1$, which results in a linear dispersion relation.

To solve the phase velocity in the diagonal direction, a similar approach that was used by van Walstijn and Kowalczyk¹⁹ is used here. In the diagonal direction the wavenumber components are equal and therefore $\hat{k}_x = \hat{k}_y = \hat{k}_z = \hat{k}_d/3$.

The frequency domain expression of the spatial difference operator in the diagonal direction can be expressed as

$$s_d = \sin^2\left(\frac{1}{\sqrt{3}} \hat{k}_d \frac{\Delta x}{2}\right) = \sin^2\left(\sqrt{3} \hat{k}_d \frac{\Delta x}{6}\right). \quad (9)$$

By substituting the functions of the different wavenumber components in Eq. (4) with Eq. (9), a general form of the explicit update equation in frequency domain for a plane wave propagating in the diagonal direction is achieved

$$s_t = \lambda^2 (3s_d - 4a \times 3s_d^2 + 16b \times s_d^3), \quad (10)$$

and that for the different schemes read

$$\text{SRL: } \lambda^2 3s_d - s_t = 0, \quad (11)$$

$$\text{CCP: } -\lambda^2 4a \times 3s_d^2 + \lambda^2 3s_d - s_t = 0, \quad (12)$$

$$\text{IWB: } \lambda^2 16b \times s_d^3 - \lambda^2 4a \times 3s_d^2 + \lambda^2 3s_d - s_t = 0. \quad (13)$$

Variable s_d is then solved. For clarity, the roots are named \hat{s}_d . For the SRL scheme, the linear equation has one root

$$\hat{s}_d = \frac{1}{3\lambda^2} s_t. \quad (14)$$

For CCP, a quadratic equation has to be solved. By substituting $a = \frac{1}{4}$ and $\lambda = 1$ as specified by Kowalczyk and van Walstijn,⁶ two roots are achieved,

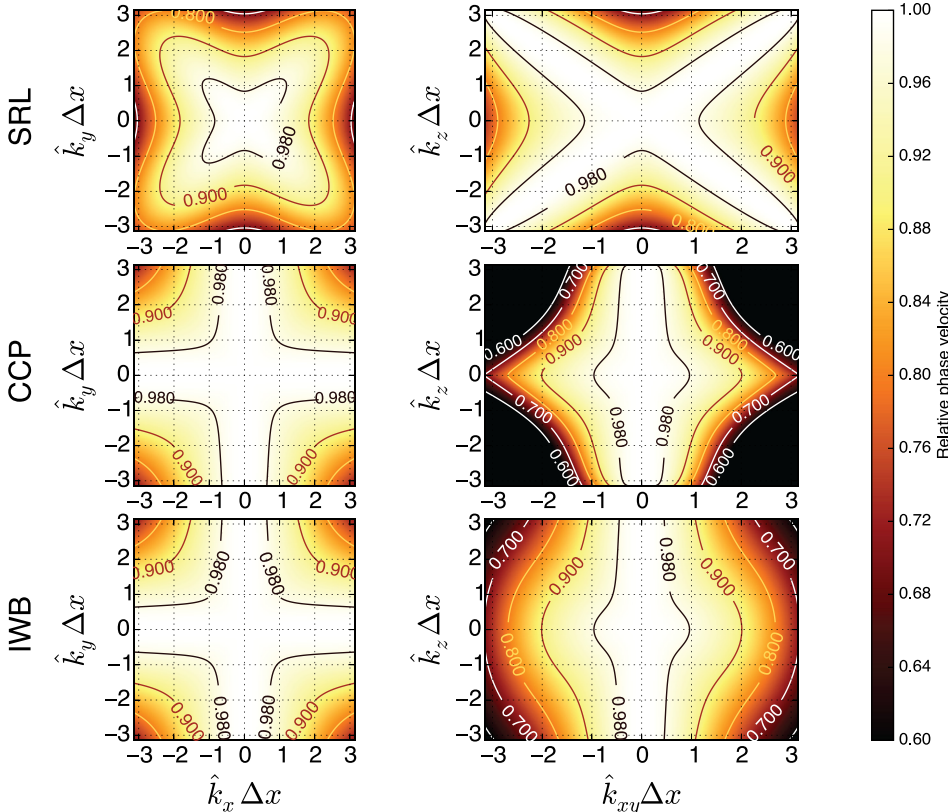


FIG. 2. (Color online) Relative phase velocity as a function of wavenumber components. On the left the phase velocity is plotted as a function of \hat{k}_x and \hat{k}_y with $\hat{k}_z = 0$. On the right as a function of $\hat{k}_{xy} = \hat{k}_x = \hat{k}_y$ and \hat{k}_z . The wave numbers are scaled according to the stability limit of the Courant number of each scheme.

$$\hat{s}_d = \frac{1}{6} \left(3 \pm \sqrt{9 - 12s_t} \right). \quad (15)$$

For IWB, that is a cubic equation. Roots for Eq. (13) with the substitutions $b = \frac{1}{16}$, $a = \frac{1}{4}$, and $\lambda = 1$ are

$$\hat{s}_d = \begin{cases} \sqrt[3]{s_t - 1} + 1, \\ 1 - \frac{1}{2} (1 \pm j\sqrt{3}) \sqrt[3]{s_t - 1}. \end{cases} \quad (16)$$

Now by equating the solutions (14), (15), and (16) with the form (9),

$$\begin{aligned} \sin^2 \left(\sqrt{3} \hat{k}_d \frac{\Delta x}{6} \right) &= \hat{s}_d \\ \Leftrightarrow \hat{k}_d &= \frac{6}{\sqrt{3} \Delta x} \arcsin \left(\sqrt{\hat{s}_d} \right). \end{aligned} \quad (17)$$

The numerical, relative phase velocity can now be written as

$$\hat{v}_d(\omega) = \frac{\omega}{\hat{k}_d(\omega)c} = \frac{\sqrt{3} \omega \frac{\Delta t}{2}}{6 \arcsin(\sqrt{\hat{s}_d})}. \quad (18)$$

In the case of multiple roots in Eqs. (15) and (16), the one is chosen that results in numerical wavenumber of 0 at DC when evaluated with Eq. (17). For the CCP scheme, we consider the wavenumber components \hat{k}_x , \hat{k}_y , and \hat{k}_z to be within the wavenumber cell of the FCC lattice (a truncated octahedron)⁴ as is appropriate for this scheme.⁵

The impulse response between a source plane and a receiver in the simulation domain can be solved using the plane wave solution and the numerical wavenumbers (8) and (17),

$$\hat{F} = e^{-j\omega t} e^{-j\hat{k}(\omega)d}. \quad (19)$$

Evaluating the plane wave with frequencies $[0, \pi]$ using a time value of $t=0$, a frequency-domain representation of a plane wave that has propagated a distance d in the simulation domain in the direction of the solved wavenumber is achieved. Using inverse Fourier transform, a time-domain impulse response corresponding to the propagation is achieved.⁹ The impulse response can be used to introduce the

dispersion error to an arbitrary source signal via convolution. This truncated impulse response is referred to hereafter as a *dispersion filter*.

The group delay of the dispersion filter can be evaluated directly from the waveform (19) using the definition of group delay

$$\tau_{\text{group}} = -\frac{d\phi(\omega)}{d\omega}, \quad (20)$$

where $\phi(\omega) = \arg[\hat{F}(\omega)]$. In this study, the group delay error is evaluated by subtracting the group delay value of the first frequency bin after the DC component from the group delay value of the frequency bin of interest.

B. Filter validation

The proposed dispersion filter is validated against corresponding FDTD simulations to show that the approach can be used to quantify the dispersion error in the simulation domain. The evaluation of the dispersion filter is done by assigning a plane of sources into the simulation domain with a delta function as the source function, recording the response at a predetermined location, and comparing the recorded response to a response generated by the dispersion filter representing a propagation of the distance between the source plane and the recording location. The edge dimension dim of the cubic simulation domain was 548 nodes, where node refers to a single element of the domain.

The source locations are illustrated in Fig. 3. For the SRL scheme, the plane wave is introduced to the simulation domain by assigning a set of hard sources at positions where the Cartesian coordinate values satisfy the condition $x = dim/2$. In the case of the CCP schemes, the source plane is introduced as a set of hard sources at positions where the Cartesian coordinate values satisfy the condition $x + y = z$. The distance d_r between the source plane and the receiver position was set to 17 nodes.

In the case of the IWB scheme, the stencil points are a combination of three different stencils using the axial, side-diagonal, and diagonal nodes. Therefore in order to introduce a source plane into the domain, the stencil has to be “filled” using three different distances normal to the initial source plane.

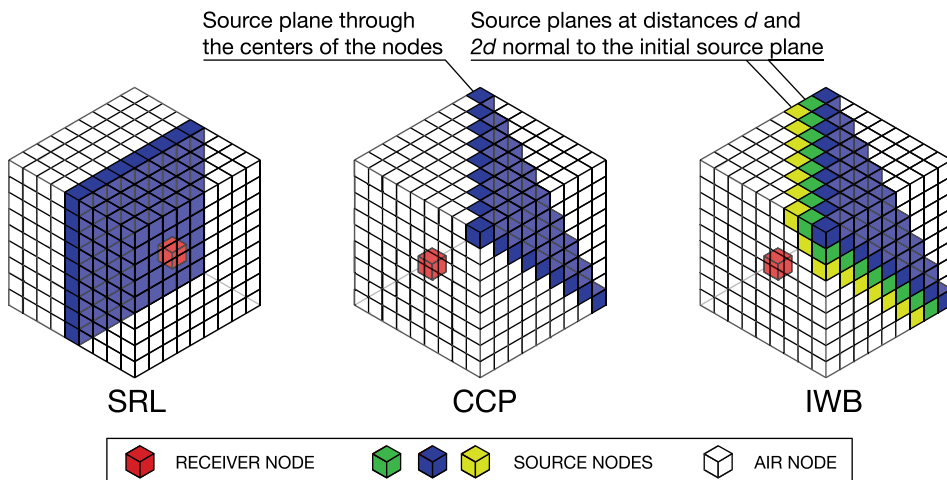


FIG. 3. (Color online) Source and receiver positions in the simulation for filter validations. For IWB two additional source planes are applied at distances $d = 1/\sqrt{3}$ and $2 \times d$ from the initial source plane location using the dispersion filter.

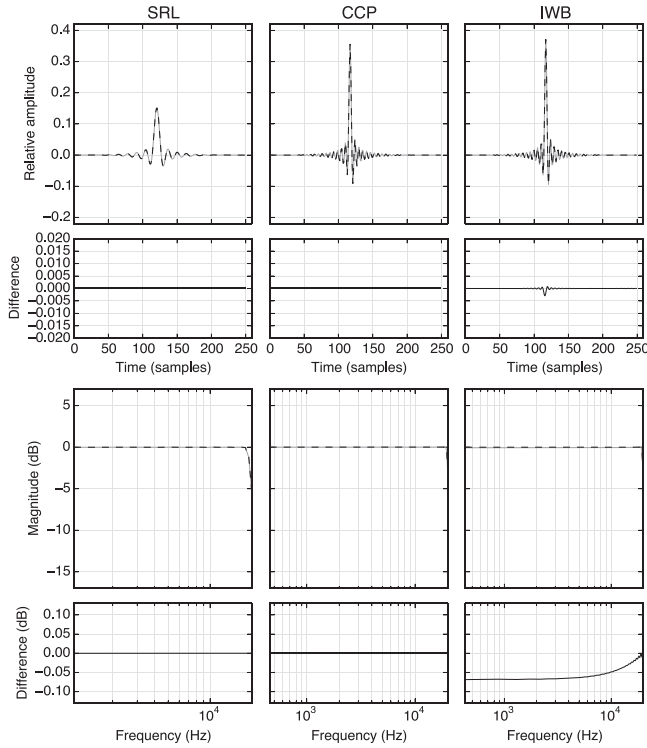


FIG. 4. Comparison of waveforms and magnitude spectrums generated with the dispersion filter (solid gray lines) and with simulation (dashed black lines).

The initial source plane is going through the center points of the node positions satisfying $x + y = z$. For the node positions satisfying $x + y = z - 1$ the distance to the initial source plane is $d = 1/\sqrt{3}$, and for node positions satisfying $x + y = z - 2$ the distance to the initial source plane is $d_2 = 2/\sqrt{3} = 2 \times d$. The source functions for the different planes of sources are solved using the derived dispersion filter. The distance d_r between the source plane and the receiver position was set to $\sqrt{3} \times 10^2$ for the CCP and IWB schemes.

Figure 4 shows the time-domain and magnitude responses generated using the dispersion filter in comparison to the simulated response. Both responses have been filtered with a low-pass filter (Hamming window, filter length of 200 taps) with a cutoff frequency of 20 kHz assuming the same sampling frequencies for the schemes that are used in the experiment (see Sec. IV A). It can be seen that the waveforms are almost identical. In the case of IWB, there is some deviation between the dispersion filter response and the simulated waveform. The deviation is likely due to the non-

exact source functions. The deviation is still small as the magnitude response deviates less than 0.1 dB. From the results of the evaluation, it is concluded that the dispersion filter represents the simulation with accuracy that is adequate for the perceptual evaluation of the dispersion error.

IV. EXPERIMENTAL SETUP

A. Stimuli

For these experiments, the reference signal in all cases is an unprocessed sound sample. The stimulus is a filtered version of the reference, where the filter corresponds to the dispersion error at a given distance in the simulation domain in the direction that introduces the most error in the scheme. The error for each scheme is calculated by evaluating at which normalized frequency Eq. (18) corresponds to 2% phase velocity error. This normalized frequency value is then used to scale the sampling frequency so that the normalized frequency value corresponds to 20 kHz. The normalized frequency where the 2% error occurs is different for each scheme used. For SRL, this normalized frequency, denoted hereafter as f_{\max} , is 0.076, for CCP it is 0.178, and for IWB it is 0.186, which results in temporal sampling rates of 264 030, 113 860, and 107 780 Hz, respectively. The phase velocity plots for the different schemes are presented in Fig. 5 in a similar manner as in Fig. 1 with the difference that the frequency axis is normalized so that unity refers to the f_{\max} of each scheme. The plot show that the phase and group velocity profiles of the different schemes are very similar to each other when the maximum phase velocity error is fixed to 2%.

The error waveforms generated with the dispersion filter were resampled to a sampling rate of 48 kHz and then introduced to the stimuli via convolution. Both the reference and the stimulus signal were low-pass filtered with a finite impulse response filter (Chebyshev window with a stop-band attenuation of 80 dB, filter length of 20 taps) with a cutoff frequency of 20 kHz. The loudness level of the reference and the stimulus were normalized according to the root mean square value of the signals.²⁰

The source material consisted of a synthetic click-like sound, referred to hereafter as “Click,” and a short phrase of male speech referred to hereafter as “Speech.” The rationale behind the selection was to achieve an absolute threshold using a sample of which the error is easy to discriminate, and additionally a threshold representing a scenario where the source material does not have a flat magnitude response, and has characteristics familiar to the subjects.

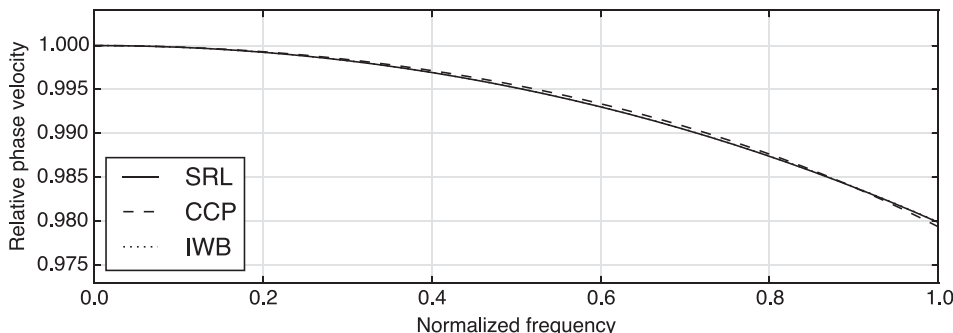


FIG. 5. Real parts of the phase velocities of the three schemes used in this study as a function of normalized frequency. The frequency axis is scaled so that unity represents the normalized frequency where the phase velocity error is 2% for each scheme.

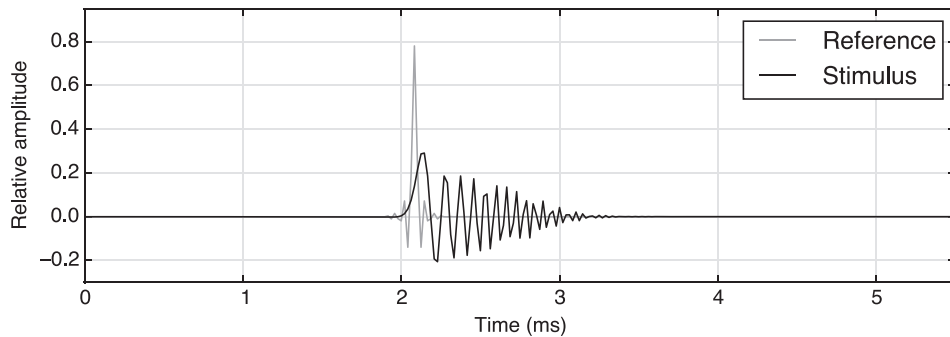


FIG. 6. The reference sample *Click* and a waveform generated by the dispersion filter of the SRL scheme for distance of 5 m. The propagation delay has been removed from the dispersion filtered waveform for comparison. It can be seen that the transient of the *Click* is smeared in time-domain as an effect of the continuous frequency dependent group delay.

The click-like sound was a $21\ \mu\text{s}$ long impulse. The reference and the stimuli signal used in the test for the SRL scheme for a propagation distance of 5 m is illustrated in Fig. 6. As can be seen, the transient of the *Click* is smeared due to dispersion. The reference and the stimulus in Fig. 6 are low-pass filtered with the previously mentioned filter.

The spectrogram representation of the sound sample *Speech* is illustrated in Fig. 7. The sample is a male voice delivering a phrase: “*My voice is recorded in an anechoic chamber.*” The figure shows that most of the frequency content is concentrated below 5 kHz, and frequencies over 15 kHz are non-existent. Three positions in the spectrogram show a higher level of frequencies at the bandwidth of 4–12 kHz: from 0.5 to 0.6 s, from 0.75 to 0.8 s, and from 2.1 to 2.2 s. These time windows correspond to the end of word “*voice,*” the end of word “*is,*” and to the beginning of the word “*chamber,*” respectively.

B. Subjects

Eight subjects participated in the test. No hearing abnormalities were reported by the participants. All of the participants were working in the field of acoustics and had previously attended listening tests of some sort.

C. Procedure

An adaptive staircase test using the QUEST (Ref. 21) method was used with the *3-alternative forced choice* (3AFC) procedure. The procedure uses assumptions that the psychometric function which is investigated has the same shape under all conditions, the value of the threshold does

not change during the experiment, and that individual trials are statistically independent.²¹ It can be assumed that these conditions are met in the experiment design.

In this study, the absolute threshold is measured and therefore the estimate for the slope of the psychometric function is not of interest. The probability threshold for 82% correct discrimination was measured. A higher probability level is selected due to reduced variance in the threshold estimate that leads to a more efficient procedure.²² As the shape of the psychometric function is not known, an estimate which corresponds to a relative shallow slope was used in order to spread the stimulus levels more widely around the “true” threshold value²³ (Weibull psychometric function, $\beta = 0.32$). The 3AFC procedure was chosen to reduce the variance of the threshold estimates in comparison to the 2-alternative forced choice procedure.²² The 4-alternative forced choice procedure was not considered due to the possible strain of memory and increased experiment time.²⁴

The 3AFC procedure was implemented as follows. Three test samples, consisting of two reference samples and one stimuli sample, were presented in random order with 1000 ms pause between each sample. The subject was then asked to specify which of the presented samples contained an audible dispersion error. The sound samples were presented to the subject only once.

The listening test consisted of a training session and two experiment sessions. In the training session, the participant was familiarized with the test routine and taught to listen specifically for the dispersion error. Two short trial sequences using the 3AFC procedure were presented to the subject starting from an easily noticeable error level and progressing

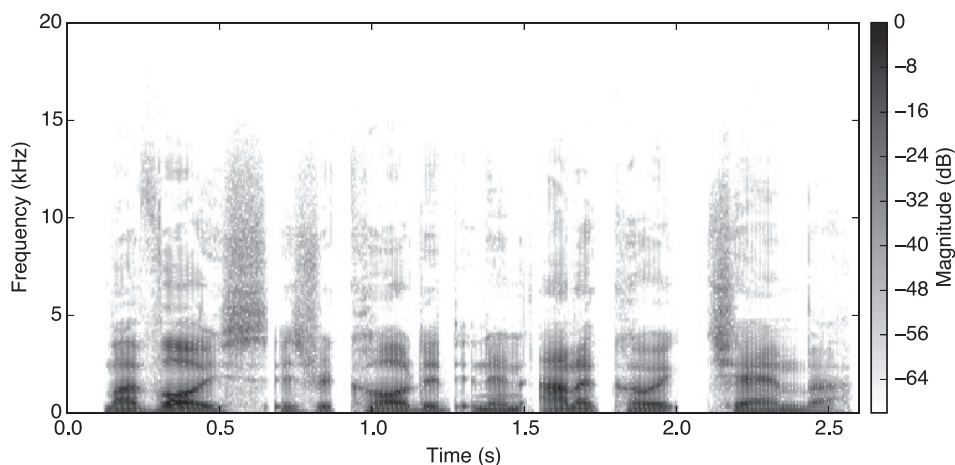


FIG. 7. Spectrogram of the sample *Speech*. The sample does not have virtually any frequency components above 15 kHz.

toward a low error level. The participant was given feedback for correct discrimination.

Two experiment sessions for each sound sample consisted of three interleaved staircase routines containing the trials for the different schemes. Subsequent discrimination tasks were randomly picked from the staircase routines of each scheme. After every six trials, a trial with an easily noticeable dispersion error was given in order to remind the participant what to listen to. The order of the sessions with different samples was randomized between subjects in order to balance possible learning and fatigue effects. The number of trials in each staircase routine was limited to 30. Feedback for correct discrimination was not given in the experiment sessions.

All of the experiments were completed using Sennheiser HD 650 headphones (Wedemark, Germany) connected to a Motu UltraLite-mk3 Hybrid audio interface (Cambridge, MA) on a laptop computer. The participant was seated in an isolated room. The headphone model has a diffuse field response; the frequency response at the eardrum is close to that which would be achieved with a loudspeaker with a flat frequency response in a diffuse field. The levels of the reference samples were calibrated using a B&K (Nærum, Denmark) 4153 artificial ear connected to a B&K 2250 sound level meter. The levels of both reference samples were set to $LC_{peak} = 90$ dB. A peak measure was used because the *Click* sample contains most of its energy in a short transient. The sample *Speech* had the levels $LAF_{max} = 75$ dB and $LAS_{max} = 66$ dB with the given peak level. LC_{peak} is the C-weighted peak *sound pressure level* (SPL) (no integration time constant), LAF_{max} the maximum value of A-weighted SPL with 125 ms integration, and LAS_{max} the maximum value of A-weighted SPL with 1000 ms integration during the playback of the sample. This SPL was found suitable not to strain the hearing of the test subjects excessively. The signals were played back diotically. An open-source library for psychophysics experiments was used to implement the test.²⁵

Additionally, each participant was asked to fill out a questionnaire after each experiment. In the questionnaire it was asked which characters in the sound sample led to discrimination, and at which part of the stimulus signal these characters occurred.

V. EXPERIMENT

A. Results

The measured thresholds for the different sound samples and different schemes are illustrated in Fig. 8, and the descriptive statistics for the different conditions are presented in Table I. The threshold value is the mean value of the probability distribution function of the threshold that is measured by the QUEST procedure. As can be observed from the figure, the threshold estimates for different sound samples differ noticeably from each other. The threshold values for different schemes in each sound sample group do not have such a clear difference.

The results of the following statistical tests are presented in Table II. A Shapiro-Wilk test for normality²⁶ was

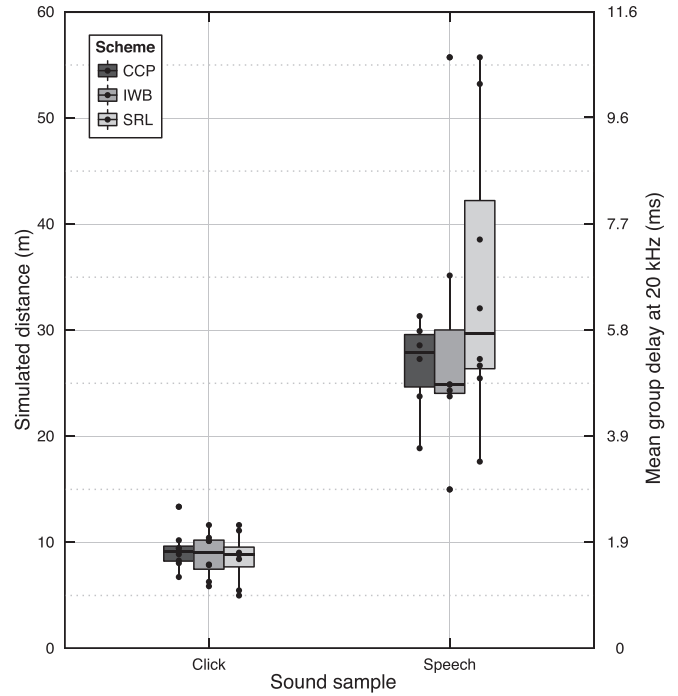


FIG. 8. The threshold observations for the different sound samples. Results for different schemes are indicated with a different color. Individual observations are indicated as points.

TABLE I. The descriptive statistics of the measured thresholds of the audibility of dispersion error. The unit is meters of propagation in a free field.

Sample	Scheme	μ (m)	σ (m)	Min (m)	Max (m)
Click	SRL	8.71	2.35	5.18	11.98
Click	CCP	9.56	2.00	6.75	13.71
Click	IWB	8.90	2.12	6.00	11.87
Speech	SRL	35.53	14.06	17.84	56.73
Speech	CCP	37.51	19.99	19.38	71.08
Speech	IWB	34.09	17.81	15.34	66.58

TABLE II. Results of statistical tests performed on the observation. *Within* indicates that the observations are compared within-subjects in the category.

Sample	Scheme	Test	Results
Click	SRL	Shapiro-Wilk	$W = 0.952$, $p = 0.460$
Click	CCP	Shapiro-Wilk	$W = 0.703$, $p = 0.281$
Click	IWB	Shapiro-Wilk	$W = 0.944$, $p = 0.580$
Speech	SRL	Shapiro-Wilk	$W = 0.927$, $p = 0.249$
Speech	CCP	Shapiro-Wilk	$W = 0.827$, $p = 0.008$
Speech	IWB	Shapiro-Wilk	$W = 0.876$, $p = 0.057$
Within	SRL	Wilcoxon	$V = 0$, $p = 0.008$
Within	CCP	Wilcoxon	$V = 0$, $p = 0.008$
Within	IWB	Wilcoxon	$V = 0$, $p = 0.008$
Click	Within	Levene	$F(2, 21) = 0.305$, $p = 0.74$
Speech	Within	Levene	$F(2, 21) = 0.038$, $p = 0.96$
Click	Within	ANOVA	$F(16, 2) = 0.364$, $p = 0.437$
Speech	Within	ANOVA	$F(16, 2) = 0.948$, $p = 0.476$
Speech	Within	Friedman	$\chi^2(2) = 5.56$, $p = 0.072$

performed for each group of observations (scheme and sound sample). The test indicated that the null hypothesis of normality cannot be rejected except for the threshold observations for the sound sample *Speech* measured for the CCP scheme.

As the variances of the sample groups differ, and hypothesis for normality was rejected for one of the groups, a paired student T-test cannot be performed between the sample groups. Instead, a Wilcoxon signed-rank test was performed²⁷ between observations in the different sound sample groups independently for each scheme group. The null hypothesis of the test is that the mean ranks of the groups are the same. The test indicates that the null hypothesis should be rejected, and that the result is highly significant. The observations of the two experiments are therefore analysed separately using *analysis of variance* (ANOVA).

To perform ANOVA, the cases should be independent, the variances should be equal, and the residuals are normally distributed. The cases compared here are assumed to be independent for different test subjects. Levene's test for relative variation²⁸ shows that the null hypothesis of equal variances cannot be rejected between the different schemes in each sound sample group.

A within-subjects single factor ANOVA was performed on the threshold observations of the experiments made with the different sound samples with scheme as a factor (SRL, CCP, and IWB). No significant evidence was found that the thresholds for the different schemes differ. As the normality test failed for the CCP scheme with the sound sample "*Speech*," additionally a non-parametric within-subjects Friedman rank sum test was performed to compare the variances between the schemes. The result of the Friedman test indicates that the null hypothesis cannot be rejected. The null hypothesis in this case is that apart from the effect of the participant, the threshold values for different schemes originate from the same distribution.

The results of the questionnaire indicated that the participants had made similar notions on the *Click* sample whereas different participants had concentrated on different parts of the sample *Speech*. For the sound sample *Click* the participants reported frequency chirps, and sweeps for high error levels, and change in timbre, color, and attack for low error levels. For the sound sample *Speech*, the participants

reported high frequency artefacts: modulation, ringing, metallic, and springy character and change in timbre. The participants reported of listening to syllables, some of the participants reported on listening to the word "*Voice*," and some of the participants the words "*Anechoic*," and "*Chamber*."

B. Discussion

The main result that the observations give evidence to is that the three different schemes do not have significantly different threshold values when the phase velocity error is fixed to 2% at 20 kHz. The result is important since the schemes have different computational requirements and theoretical efficiencies. The presented result indicates that the more complex CCP and IWB schemes have similar perceptual limits at a much lower sampling frequency than the SRL scheme. Additionally, this result indicates that the audibility of dispersion error in the studied schemes may be predicted using the phase velocity contours as they are remarkably similar in Fig. 5.

The second main result is the measured one; the mean perceptual threshold for the distance between the source and the receiver in the given conditions using the studied FDTD schemes is 9.1 m of propagation (1.8 ms group delay error at 20 kHz) for the sample *Click* and 35.7 m of propagation (6.9 ms group delay error at 20 kHz) for the sound sample *Speech*. The threshold value represents propagation in a free field in the worst-case direction of each scheme, and therefore its impact on a complete room response cannot be exhaustively concluded. The result gives an indication to the general limit of dispersion error in a sense that if the whole room response is simulated with less than the measured threshold, dispersion error in a free field is likely to go unnoticed.

The threshold values measured with the sound sample *Speech* had a greater variance than the thresholds measured with the sound sample *Click*. The threshold values for the sound sample *Speech* were also consistently higher. An obvious reason for the higher threshold values is the band limited spectrum of the sample. At 15 kHz the sound sample had hardly any energy as can be observed from Fig. 7, whereas at 12 kHz, there are several time moments where the level is high. The group delay error values for 15 and 12

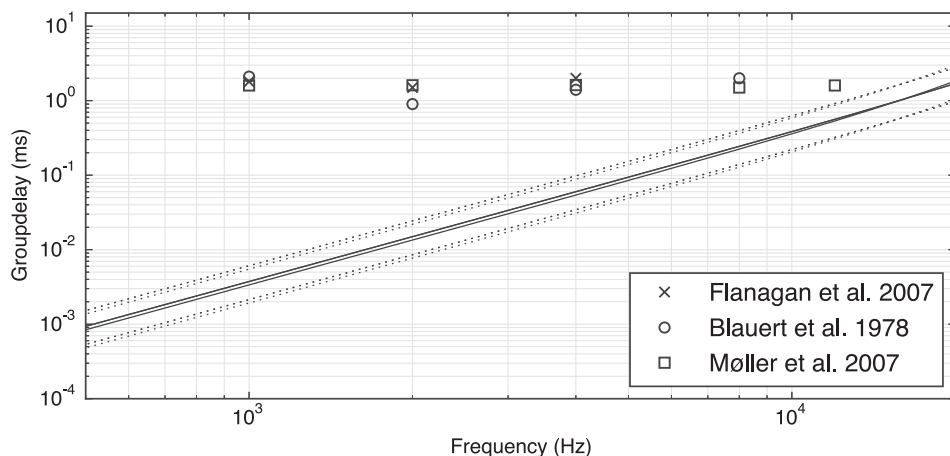


FIG. 9. The group delay error curves of the different schemes at different simulation distances: 5.18 m (lower dashed black lines), 9.06 m (solid black lines), and 14.76 m (higher dashed black lines). Peak group delay thresholds from different studies are plotted with different markers adapted from Ref. 14. Note that the probability levels for correct discrimination are not equal between the presented studies and should be interpreted only for guidance.

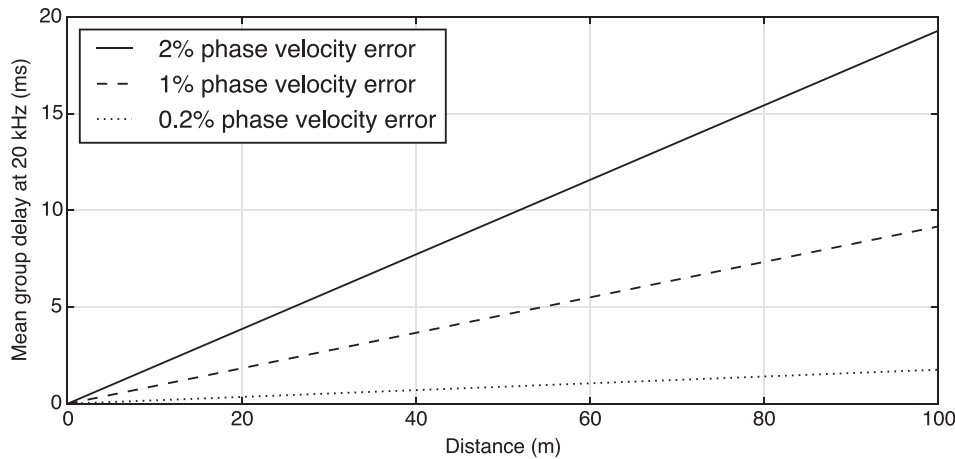


FIG. 10. The mean group delay errors of SRL, CCP, and IWB schemes evaluated at different distances at 20 kHz. Phase velocity error is fixed to the noted percentage at 20 kHz.

kHz with the given phase velocity error of 2% and mean distance of propagation (mean over the mean threshold of the different schemes from Table I, 35.7 m) are 3.5 and 2.2 ms, respectively.

The greater variance of the threshold values measured with the sound sample *Speech* may have been affected by the different approaches of what the participants have used for the discrimination. From the results of the questionnaire it was evident that the participants reported similar artefacts but the time moment in the sample which the participants concentrated on in order to discriminate the error varied. This may have been one of the reasons for a greater variance, although the magnitude of this factor cannot be concluded.

The group delay error curves for the observed maximum, minimum, and mean propagation distances for the sound sample *Click* from Table I are illustrated in Fig. 9. In comparison, the measured thresholds for group delays at different frequency bands from previous group delay studies are illustrated with different markers. Although the dispersion error of a FDTD simulation has generally different characteristics in comparison to a narrow bandwidth group delay of an all-pass filter, the results indicate that the mean group delay error value (1.8 ms at 20 kHz) measured in this work is in line with the measured threshold values in existing studies. It should be noted that the probabilities of correct discrimination in the different studies vary. Blauert and Laws measured the relative frequency of 50% correct whether the participants heard a difference between the stimulus and the reference.¹⁴ Flanagan *et al.* reported 75% threshold for correct discrimination.¹⁶ Møller *et al.* measured the threshold of the group delay error using the method of adjustment.¹⁷ In the current study the probability threshold of 82% of correct discrimination for 3AFC was measured, and therefore the connection of the results of previous studies should be taken only as a directive evidence.

The relationship of the mean group delay error of the schemes studied in this work and the simulation distance is illustrated in Fig. 10 with different phase velocity error percentages. A linear relationship between phase velocity error and the slope of the curve can be observed. As can be seen from the figures, the group delay error increases linearly as a function of distance in all cases. The implication of this is that with any value of phase velocity error, there is a distance

at which the group delay error will become audible. Therefore the widely used metric of “low-enough” phase velocity error does not apply by itself. From the results of this study, we propose that a more meaningful approach in quantifying the dispersion error is to first specify the propagation distance of interest, and then specify what group delay error is acceptable at that distance. These two values will then determine the phase velocity error, and subsequently the sampling frequency that is needed to attain the phase velocity error level at the bandwidth of interest.

There are several factors that have been previously noted by different authors that can reduce the perceptibility of the error in more complex schemes. Air absorption is a significant source of attenuation at high frequencies, and therefore the frequency components containing a dispersion error are likely to be attenuated to an extent that they are unperceivable. The inclusion of reflections from room geometry can also reduce the effect of the dispersion error, as different propagation paths that contain less error are combined to the response. Last, limiting the bandwidth of the FDTD simulation result using low-pass filtering will remove delayed high frequency components from a chosen cutoff frequency and may make the remaining error unperceivable. With such an approach, the high frequency response must be simulated with other methods.

VI. CONCLUSIONS

The audibility of dispersion error in FDTD simulation was measured for three commonly employed FDTD schemes as a function of simulation distance in a free field. The dispersion error was evaluated using the numerical wavenumbers of each scheme. It was found that the thresholds for the different schemes are not significantly different when the phase velocity error is fixed to 2% at 20 kHz. The thresholds for different sound samples were found to be significantly different. The perceptual threshold for the dispersion error in a free field at the probability level of 82% for correct discrimination for 3AFC was found to be 9.1 m of propagation that corresponds to a group delay error value of 1.8 ms at 20 kHz with a fixed phase velocity error of 2% at the same frequency limit. The implication of the result to simulated room responses cannot be concluded exhaustively due to the exclusion of air absorption and surface reflections from the

study. The results give a general indication of audibility of dispersion error and may serve as reference in terms of perception of group delay error in FDTD.

Several areas of further research remain. The masking effect of air absorption and surface reflections are likely to be significant factors in the audibility of the dispersion error in full room responses, and should be investigated. Additionally, the band-limitation of FDTD results in low frequencies can lead to plausible results when combined with other simulation methods.

ACKNOWLEDGMENTS

Sebastian Prepelita and Pierre Chebeau are thanked for commenting on the manuscript. The research leading to these results received funding from the Academy of Finland, Project No. 265824.

- ¹M. Aretz, "Specification of realistic boundary conditions for the FE simulation of low frequency sound fields in recording studios," *Acta Acust. Acust.* **95**, 874–882 (2009).
- ²N. Raghuvanshi, R. Narain, and M. Lin, "Efficient and accurate sound propagation using adaptive rectangular decomposition," *IEEE Trans. Vis. Comput. Graphics* **15**, 789–801 (2009).
- ³L. Savioja and V. Välimäki, "Reducing the dispersion error in the digital waveguide mesh using interpolation and frequency-warping techniques," *IEEE Trans. Audio, Speech, Lang. Process.* **8**, 184–194 (2000).
- ⁴D. Peterson and D. Middleton, "Sampling and reconstruction of wave-number-limited functions in N-dimensional Euclidean spaces," *Inf. Control* **5**, 279–323 (1962).
- ⁵B. Hamilton and S. Bilbao, "On finite difference schemes for the 3-D wave equation using non-Cartesian grids," in *Proceedings of Sound and Music Computing Conference*, Stockholm, Sweden (2013), pp. 592–599.
- ⁶K. Kowalczyk and M. van Walstijn, "Room acoustics simulation using 3-D compact explicit FDTD schemes," *IEEE Trans. Audio, Speech, Lang. Process.* **19**, 34–46 (2011).
- ⁷L. Savioja, "Real-time 3D finite-difference time-domain simulation of low- and mid-frequency room acoustics," in *13th International Conference on Digital Audio Effects*, Graz, Austria (2010), pp. 77–84.
- ⁸B. Hamilton and C. J. Webb, "Room acoustics modelling using GPU-accelerated finite difference and finite volume methods on a face-centered cubic grid," in *16th International Conference on Digital Audio Effects*, Maynooth, Ireland (2013), pp. 336–343.
- ⁹J. B. Schneider and C. L. Wagner, "FDTD dispersion revisited: Faster-than-light propagation," *IEEE Microwave Guided Wave Lett.* **9**, 54–56 (1999).
- ¹⁰A. Southern, D. Murphy, T. Lokki, and L. Savioja, "The perceptual effects of dispersion error on room acoustic model auralization," in *Proceeding Forum Acusticum*, Aalborg, Denmark (2011), pp. 1553–1558.
- ¹¹A. Taflov, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 2nd ed. (Artech House, Boston, 2000), Chap. 5, pp. 175–233.
- ¹²M. Cobos, J. Escolano, J. J. López, and B. Pueo, "Subjective effects of dispersion in the simulation of room acoustics using digital waveguide mesh," in *124th Convention of Audio Engineering Society*, Amsterdam, Netherlands (2008), Paper No. 7471.
- ¹³J. H. Patterson and D. M. Green, "Discrimination of transient signals having identical energy spectra," *J. Acoust. Soc. Am.* **48**, 894–905 (1970).
- ¹⁴J. Blauert and P. Laws, "Group delay distortions in electroacoustical systems," *J. Acoust. Soc. Am.* **63**, 1478–1483 (1978).
- ¹⁵J. Deer, P. Bloom, and D. Preis, "Perception of phase distortion in all-pass filters," *J. Audio Eng. Soc.* **33**, 782–786 (1985).
- ¹⁶S. Flanagan, B. C. Moore, and M. A. Stone, "Discrimination of group delay in click-like signals presented via headphones and loudspeakers," *J. Audio Eng. Soc.* **53**, 593–611 (2005).
- ¹⁷H. Möller, P. Minnaar, S. K. Olesen, F. Christensen, and J. Plogsties, "On the audibility of all-pass phase in electroacoustical transfer functions," *J. Audio Eng. Soc.* **55**, 113–134 (2007).
- ¹⁸A. Gourlay and A. Mitchell, "A classification of split difference methods for hyperbolic equations in several space dimensions," *SIAM J. Numer. Anal.* **6**, 62–71 (1969).
- ¹⁹M. van Walstijn and K. Kowalczyk, "On the numerical solution of the 2D wave equation with compact FDTD schemes," in *11th International Conference on Digital Audio Effects*, Espoo, Finland (2008), pp. 205–212.
- ²⁰G. A. Soulodre, "Evaluation of objective loudness meters," in *116th Convention of Audio Engineering Society*, Berlin, Germany (2004), Paper No. 6161.
- ²¹A. B. Watson and D. G. Pelli, "Quest: A Bayesian adaptive psychometric method," *Percept. Psychophys.* **33**, 113–120 (1983).
- ²²S. P. McKee, S. A. Klein, and D. Y. Teller, "Statistical properties of forced-choice psychometric functions: Implications of probit analysis," *Percept. Psychophys.* **37**, 286–298 (1985).
- ²³B. Treutwein, "Adaptive psychophysical procedures," *Vis. Res.* **35**, 2503–2522 (1995).
- ²⁴M. R. Leek, "Adaptive procedures in psychophysical research," *Percept. Psychophys.* **63**, 1279–1292 (2001).
- ²⁵J. W. Peirce, "PsychoPy—Psychophysics software in Python," *J. Neurosci. Methods* **162**, 8–13 (2007).
- ²⁶S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika* **52**, 591–611 (1965).
- ²⁷F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.* **1**, 80–83 (1945).
- ²⁸B. B. Schultz, "Levene's test for relative variation," *Systematic Biol.* **34**, 449–456 (1985).