

Location-dependent Spatiotemporal Antialiasing in Photoacoustic Computed Tomography

Peng Hu, Lei Li, and Lihong V. Wang, *Fellow, IEEE*

Abstract—Photoacoustic computed tomography (PACT) images optical absorption contrast by detecting ultrasonic waves induced by optical energy deposition in materials such as biological tissues. An ultrasonic transducer array or its scanning equivalent is used to detect ultrasonic waves. The spatial distribution of the transducer elements must satisfy the spatial Nyquist criterion; otherwise, spatial aliasing occurs and causes artifacts in reconstructed images. The spatial Nyquist criterion poses different requirements on the transducer elements' distributions for different locations in the image domain, which has not been studied previously. In this research, we elaborate on the location dependency through spatiotemporal analysis and propose a location-dependent spatiotemporal antialiasing method. By applying this method to PACT in full-ring array geometry, we effectively mitigate aliasing artifacts with minimal effects on image resolution in both numerical simulations and in vivo experiments.

Index Terms—Photoacoustic computed tomography, spatial Nyquist criterion, location-dependent spatiotemporal antialiasing

I. INTRODUCTION

PHOTOACOUSTIC computed tomography (PACT) images biological tissues' optical absorption through detection of photon-induced ultrasonic waves [1]–[7]. PACT forms high-resolution images at greater depths than ballistic optical imaging [8] by using tissues' much lower scattering to ultrasonic waves than to photons. An ultrasonic transducer array or its scanning equivalent is often used to detect photon-induced ultrasonic waves. The detected signals are then used to recover tissues' optical absorption through image reconstruction [9]–[13]. In ultrasonic detection, the Nyquist sampling criterion must be satisfied in both spatial and temporal dimensions to avoid aliasing. The Nyquist criterion in the temporal dimension is typically satisfied because of the limited bandwidth of the ultrasonic transducer and the high temporal sampling frequency of the data acquisition system. However, sparse spatial sampling is commonly used to reduce system cost or scanning time, which may violate the Nyquist criterion in spatial dimensions (spatial Nyquist criterion) and cause artifacts in reconstructed images due to spatial aliasing [1], [14]–[16].

Various methods have been proposed to mitigate artifacts caused by spatial aliasing. In the image domain, total variation

(TV) regularization has been used in model-based iterative methods to mitigate noise and aliasing artifacts [12], [17]–[20]. In PACT, because the image domain is identical to the object domain, we will use them interchangeably henceforth. TV regularization shows high performance for piecewise smooth images [21], [22]. However, for PACT images with rich blood vessel structures, TV regularization tends to suppress vessels of small diameters. A regularization strategy specifically suited for vessel structures is needed. Deep learning has been proved effective in processing images with complex structures [23]–[28] and has shown an advantage in maintaining vessel structures [29]. However, a neural network is often system dependent and not universally applicable across different imaging systems or detection geometries. In the signal domain, temporal filtering and spatial interpolation have been used for antialiasing [15]. However, it is still a challenge to find a balance between mitigating aliasing artifacts and maintaining image resolution. Another method proposed by Cai *et al.* [30] mitigates aliasing artifacts by connecting the image domain and the signal domain. It identifies potential sources of aliasing signals in the image domain, maps the sources to the signal domain, suppresses all the signals in the mapped region, and uses the remaining signals for image reconstruction [30]. This method performs well if there exist only a few dominant sources of aliasing signals. As sources of aliasing signals increase, this method may cause substantial information loss.

To mitigate aliasing artifacts without compromising image resolution, we perform detailed spatiotemporal analysis for image subdomains here. We first reconstruct an image using the universal back-projection (UBP) method [9]. Applying a threshold to the reconstructed image, we identify the dominant sources of aliasing signals. Then we divide the whole image domain into multiple subdomains. We apply spatiotemporal analysis to source points, transducer locations, and each subdomain's reconstruction locations [15], revealing the spatial aliasing effects on the subdomain in detail. Next, we apply temporal filtering and spatial interpolation to signals so that the filtered signals satisfy the spatial Nyquist criterion specifically for this subdomain. We use the filtered signals to form an image in this subdomain. Repeating this process for all subdomains, we mitigate the aliasing artifacts for the whole image. We call this method location-dependent spatiotemporal antialiasing. Through numerical simulations, we demonstrate that the

Manuscript received October 12, 2022; revised November 19, 2022; accepted November 22, 2022. This project was sponsored by National Institutes of Health Grants R01 NS102213, U01 NS099717 (BRAIN Initiative), U01 EB029823 (BRAIN Initiative), and R35 CA220436 (Outstanding Investigator Award). Lihong V. Wang has financial interests in Microphotoacoustics, Inc., CalPACT, LLC, and Union Photoacoustic Technologies, Ltd., which did not support this work. (Corresponding author: Lihong V. Wang)

The authors are with the Caltech Optical Imaging Laboratory, Andrew and Peggy Cherng Department of Medical Engineering, Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: penghu@caltech.edu; leili@caltech.edu; LVW@caltech.edu).

proposed method effectively mitigates aliasing artifacts with minimal effects on the image resolution. We further validate this method through *in vivo* human breast imaging.

II. UBP METHOD AND LOCATION-DEPENDENT RECENTERING OF SIGNALS

In the forward model in PACT with a homogeneous medium, the signal detected by the element at the location \mathbf{r}_n and time t is expressed as [8], [15]

$$\hat{p}(\mathbf{r}_n, t) = \frac{1}{4\pi c^2} \sum_{m=1}^M v_m \frac{p_0(\mathbf{r}'_m)}{\|\mathbf{r}'_m - \mathbf{r}_n\|} h'_e \left(t - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} \right), \quad n = 1, 2, \dots, N, t \geq 0. \quad (1)$$

Here, we have M point sources distributed at $\mathbf{r}'_m, m = 1, 2, \dots, M$, and N point transducer elements distributed at $\mathbf{r}_n, n = 1, 2, \dots, N$; c is the speed of sound; v_m is the volume of the m -th source point; $p_0(\mathbf{r}'_m)$ is the initial pressure at \mathbf{r}'_m ; $h_e(t)$ is the ultrasonic transducer's electric impulse response (EIR), and $h'_e(t)$ denotes its time derivative. Based on the UBP method [9], the initial pressure can be reconstructed from the detected signals. In a spatially discrete form, the reconstruction can be expressed as

$$\hat{p}_0(\mathbf{r}'') \approx \sum_{n=1}^N w_n \hat{b} \left(\mathbf{r}_n, t = \frac{\|\mathbf{r}'' - \mathbf{r}_n\|}{c} \right), \quad \mathbf{r}'' \in D. \quad (2)$$

Here, $\hat{p}_0(\mathbf{r}'')$ is the reconstructed initial pressure at \mathbf{r}'' , D is the image domain, and $\hat{b}(\mathbf{r}_n, t) = 2\hat{p}(\mathbf{r}_n, t) - 2t \frac{\partial \hat{p}(\mathbf{r}_n, t)}{\partial t}$ is the back-projection term computed from the detected signals. The weights $w_n, n = 1, 2, \dots, N$ express the solid-angle term $\frac{d\Omega}{\Omega_0}$ in UBP [9]. Substituting (1) into (2), we obtain

$$\hat{p}_0(\mathbf{r}'') \approx \frac{1}{2\pi c^2} \sum_{n=1}^N w_n \sum_{m=1}^M v_m \frac{p_0(\mathbf{r}'_m)}{\|\mathbf{r}'_m - \mathbf{r}_n\|} \left(1 - \left(t + \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} \right) \frac{\partial}{\partial t} \right) h'_e \left(\frac{\|\mathbf{r}'' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} \right), \quad \mathbf{r}'' \in D. \quad (3)$$

As demonstrated in [15], spatial aliasing in PACT has two sources: spatial sampling and image reconstruction. Spatial aliasing in spatial sampling and image reconstruction can be explained by analyzing the step size of $\frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}$ (in (1)) and $\frac{\|\mathbf{r}'' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c}$ (in (3)), respectively, as n varies [15]. Here, we divide the whole image domain D into subdomains; then we analyze these two terms and develop antialiasing strategies for each subdomain.

For simplicity, we focus on 2D image reconstruction and consider only rectangular subdomains. For a subdomain D_{sub} of size $l_x \times l_y$ centered at $\mathbf{r}'_{\text{c,sub}}$, we shift the time t to t' according to $t' = t - \frac{\|\mathbf{r}'_{\text{c,sub}} - \mathbf{r}_n\|}{c}$; instead of analyzing $\hat{p}(\mathbf{r}_n, t)$, we analyze

$$\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t') = \hat{p} \left(\mathbf{r}_n, t' + \frac{\|\mathbf{r}'_{\text{c,sub}} - \mathbf{r}_n\|}{c} \right), \quad n = 1, 2, \dots, N, \quad (4)$$

which corresponds to a temporal recentering of signals based on the transducer elements' distances to the subdomain center $\mathbf{r}'_{\text{c,sub}}$. After the recentering, the signals originating from $\mathbf{r}'_{\text{c,sub}}$

arrive at all detectors at time 0, and the exact range of interest for t' is dynamically determined for each subdomain. This recentering is essential in antialiasing to minimize temporal filtering and thus image blurring. Substituting (1) into (4) yields

$$\begin{aligned} \hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t') &= \frac{1}{4\pi c^2} \sum_{m=1}^M v_m \frac{p_0(\mathbf{r}'_m)}{\|\mathbf{r}'_m - \mathbf{r}_n\|} \\ &\quad h'_e \left(t' - \left(\frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_{\text{c,sub}} - \mathbf{r}_n\|}{c} \right) \right), \\ &\quad n = 1, 2, \dots, N. \end{aligned} \quad (5)$$

III. SPATIOTEMPORAL ANTIALIASING FOR AN IMAGE SUBDOMAIN

Given the subdomain D_{sub} , we categorize our analysis into four cases with increasing complexities: without point sources outside D_{sub} , with a single point source outside D_{sub} , with multiple point sources outside D_{sub} , and with general sources outside D_{sub} . In all cases, we discuss image reconstruction only in D_{sub} .

A. Without point sources outside the subdomain

In the first case, without point sources outside the subdomain D_{sub} , we perform spatiotemporal analysis only for source points and reconstruction locations inside D_{sub} . Let \mathbf{r}' and \mathbf{r}'' be a source point and a reconstruction location, respectively, in D_{sub} . Let \mathbf{r} and \mathbf{r}_{adj} be two adjacent element locations, as shown in Fig. 1(a).

First, we analyze spatial aliasing in spatial sampling based on (5). We define

$$t'_r = \frac{\|\mathbf{r}' - \mathbf{r}\|}{c} - \frac{\|\mathbf{r}'_{\text{c,sub}} - \mathbf{r}\|}{c}, \quad (6)$$

and its step size when the element location \mathbf{r} changes to \mathbf{r}_{adj} :

$$\begin{aligned} \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c,sub}}, \mathbf{r}') &= |t'_{\mathbf{r}_{\text{adj}}} - t'_r| = \\ &\quad \left| \left(\frac{\|\mathbf{r}' - \mathbf{r}_{\text{adj}}\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}\|}{c} \right) - \right. \\ &\quad \left. \left(\frac{\|\mathbf{r}'_{\text{c,sub}} - \mathbf{r}_{\text{adj}}\|}{c} - \frac{\|\mathbf{r}'_{\text{c,sub}} - \mathbf{r}\|}{c} \right) \right|, \\ &\quad \mathbf{r}' \in D_{\text{sub}}. \end{aligned} \quad (7)$$

For any $\mathbf{r}' \in D_{\text{sub}}$ with $\|\mathbf{r}' - \mathbf{r}\| \neq \|\mathbf{r}' - \mathbf{r}_{\text{adj}}\|$, there exists a branch (blue-dotted curve in Fig. 1(a)) of a hyperbola crossing \mathbf{r}' and with \mathbf{r} and \mathbf{r}_{adj} as the foci. We denote either one of the intersection points (using the other one leads to the same result) between the branch and the boundary of D_{sub} as $\hat{\mathbf{r}'}$, as shown in Fig. 1(a). For convenience, we denote the boundary of D_{sub} as ∂D_{sub} in the following discussions. Based on one of the hyperbola's definitions, we have $\frac{\|\mathbf{r}' - \mathbf{r}_{\text{adj}}\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}\|}{c} = \frac{\|\mathbf{r}' - \mathbf{r}_{\text{adj}}\|}{c} - \frac{\|\hat{\mathbf{r}'} - \mathbf{r}\|}{c}$. For any $\mathbf{r}' \in D_{\text{sub}}$ with $\|\mathbf{r}' - \mathbf{r}\| = \|\mathbf{r}' - \mathbf{r}_{\text{adj}}\|$, \mathbf{r}' will be on the perpendicular bisector (black-dashed line in Fig. 1(a)) of the line segment with \mathbf{r} and \mathbf{r}_{adj} as endpoints. Here, we define

$$\tau_{D_{\text{sub}}}(\mathbf{r}, \mathbf{r}_{\text{adj}}) = \max_{\mathbf{r}' \in \partial D_{\text{sub}}} \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c,sub}}, \mathbf{r}'). \quad (8)$$

We choose $\hat{\mathbf{r}}'$ as one of the intersection points between the perpendicular bisector and ∂D_{sub} . Thus, for any $\mathbf{r}' \in D_{\text{sub}}$, replacing \mathbf{r}' with $\hat{\mathbf{r}}'$ in $\tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}')$ does not change its value, yielding

$$\tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}') = \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c},\text{sub}}, \hat{\mathbf{r}}') \leq \tau_{D_{\text{sub}}}(\mathbf{r}, \mathbf{r}_{\text{adj}}), \quad \mathbf{r}' \in D_{\text{sub}}. \quad (9)$$

Through (9), we simplify the estimation of the upper limit of $\tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}')$ from searching \mathbf{r}' in D_{sub} to searching \mathbf{r}' on ∂D_{sub} , which reduces the computation cost by one dimension. Due to spatiotemporal coupling, as shown in the term $t - \left(\frac{\|\mathbf{r}'_m - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_{\text{c},\text{sub}} - \mathbf{r}_n\|}{c}\right)$ in (5), the upper cutoff frequency $f_{\text{c},\text{SS}}$ for spatial sampling must meet the Nyquist criterion:

$$f_{\text{c},\text{SS}} < \frac{1}{2\tau_{D_{\text{sub}}}(\mathbf{r}, \mathbf{r}_{\text{adj}})}. \quad (10)$$

Denoting

$$f_{c,D_{\text{sub}}}(\mathbf{r}_n) = \min_{\mathbf{r}_{n'} \text{ is adjacent to } \mathbf{r}_n} \frac{1}{2\tau_{D_{\text{sub}}}(\mathbf{r}_n, \mathbf{r}_{n'})}, \quad n, n' = 1, 2, \dots, N, \quad (11)$$

we can remove aliasing in the spatial sampling by processing signals of the element located at \mathbf{r}_n using a lowpass filter with an upper cutoff frequency $f_{c,D_{\text{sub}}}(\mathbf{r}_n)$, $n = 1, 2, \dots, N$. The filter is implemented as a third-order lowpass Butterworth filter combined with a sinc filter with the same upper cutoff frequency.

Next, we analyze spatial aliasing in the image reconstruction based on (3). For this analysis, we first estimate the upper limit of the step size of $\frac{\|\mathbf{r}'' - \mathbf{r}\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}\|}{c}$ between two adjacent element locations \mathbf{r} and \mathbf{r}_{adj} :

$$\tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}', \mathbf{r}'') = \left| \left(\frac{\|\mathbf{r}'' - \mathbf{r}_{\text{adj}}\|}{c} - \frac{\|\mathbf{r}'' - \mathbf{r}\|}{c} \right) - \left(\frac{\|\mathbf{r}' - \mathbf{r}_{\text{adj}}\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}\|}{c} \right) \right|, \quad \mathbf{r}', \mathbf{r}'' \in D_{\text{sub}}. \quad (12)$$

Based on (9), we use the triangular inequality to obtain

$$\begin{aligned} \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}', \mathbf{r}'') &\leq \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}') + \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}'') \\ &\leq 2\tau_{D_{\text{sub}}}(\mathbf{r}, \mathbf{r}_{\text{adj}}), \mathbf{r}', \mathbf{r}'' \in D_{\text{sub}}. \end{aligned} \quad (13)$$

From (9) and (13) as well as (11), we conclude that removing aliasing in the image reconstruction can be accomplished by additional lowpass filtering with an upper cutoff frequency $\frac{f_{c,D_{\text{sub}}}(\mathbf{r}_n)}{2}$, $n = 1, 2, \dots, N$. We observe that this cutoff frequency is half the value required for removing aliasing in the spatial sampling. This observation agrees with our previous finding [15], where the whole image domain was globally analyzed. This filtering would further compromise the spatial resolution.

Fortunately, the additional filtering is avoided by spatial interpolation if aliasing in the spatial sampling is removed first. We denote the denser element locations after the spatial interpolation as $\mathbf{r}_{\beta,n}$, $n = 1, 2, \dots, \beta N$ with β being an integer. Note that the denser element locations coincide with the physical element locations at $\mathbf{r}_n = \mathbf{r}_{\beta,\beta(n-1)+1}$, $n = 1, 2, \dots, N$. For each t' , the recentered signals ($\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$) from all the transducer elements form a vector of length N . We apply fast Fourier transform (FFT) to the vector and pad zeros following

the highest frequency components to form a new vector of length βN . Then we apply inverse FFT to the new vector to finish the spatial interpolation. Updating (11) with the denser locations, we obtain

$$f_{c,D_{\text{sub}},\beta}(\mathbf{r}_{\beta,n}) = \min_{\mathbf{r}_{\beta,n'} \text{ is adjacent to } \mathbf{r}_{\beta,n}} \frac{1}{2\tau_{D_{\text{sub}}}(\mathbf{r}_{\beta,n}, \mathbf{r}_{\beta,n'})}, \quad n, n' = 1, 2, \dots, \beta N. \quad (14)$$

From (7) and (8), we see that the larger the value of β , the closer the adjacent element locations, the smaller the value of $\tau_{D_{\text{sub}}}(\mathbf{r}_{\beta,n}, \mathbf{r}_{\beta,n'})$, and the larger the value of $f_{c,D_{\text{sub}},\beta}(\mathbf{r}_{\beta,n})$. According to (13) and (14), we can filter the signals of the n -th channel with upper cutoff frequency $\frac{f_{c,D_{\text{sub}},\beta}(\mathbf{r}_{\beta,n})}{2}$ to remove aliasing in the image reconstruction. To avoid compromising spatial resolution further, we choose β such that

$$\frac{f_{c,D_{\text{sub}},\beta}(\mathbf{r}_{\beta,\beta(n-1)+1})}{2} \geq f_{c,D_{\text{sub}}}(\mathbf{r}_n), n = 1, 2, \dots, N. \quad (15)$$

Thus, after the first temporal filtering with the upper cutoff frequency of $f_{c,D_{\text{sub}}}(\mathbf{r}_n)$ to remove aliasing in the spatial sampling and the spatial interpolation with factor β , we no longer need to perform additional temporal filtering for image reconstruction. Because a general subdomain is off-centered in the image domain, the spatial interpolation is applied to the recentered signals $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$ instead of the original signals $\hat{p}(\mathbf{r}_n, t)$. In our previous study for the whole image domain [15], we have $\beta = 2$. In this research, the minimal β satisfying (15) is obtained through numerical computations for each subdomain.

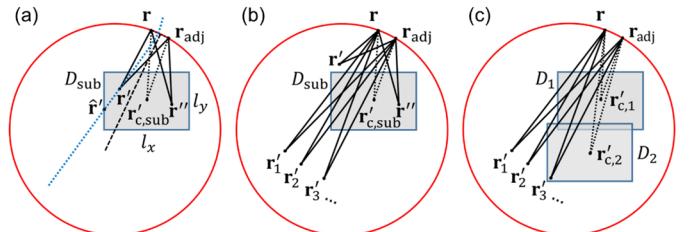


Fig. 1. Location-dependent spatiotemporal analysis. (a) No point sources outside the subdomain. A full-ring transducer array (red circle), an image subdomain (rectangle with gray interior and blue boundary, denoted as D_{sub} , centered at $\mathbf{r}'_{\text{c},\text{sub}}$, and of size $l_x \times l_y$), two adjacent element locations \mathbf{r} and \mathbf{r}_{adj} , a source point \mathbf{r}' inside D_{sub} , and a reconstruction location \mathbf{r}'' inside D_{sub} . There exists a hyperbola with \mathbf{r} and \mathbf{r}_{adj} as the foci, and with one branch (blue dotted curve) crossing \mathbf{r}' . One of the branch's intersection points with the boundary of D_{sub} is denoted as $\hat{\mathbf{r}}'$. This graph is used in the spatiotemporal analysis for D_{sub} without point sources outside. (b) One or multiple point sources outside the subdomain. A full-ring transducer array, an image subdomain D_{sub} centered at $\mathbf{r}'_{\text{c},\text{sub}}$, two adjacent element locations \mathbf{r} and \mathbf{r}_{adj} , and a reconstruction location \mathbf{r}'' inside D_{sub} . We have a single source point (\mathbf{r}') or multiple source points ($\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \dots$) outside D_{sub} . This graph is used in the spatiotemporal analysis for D_{sub} with a single or multiple point sources outside. (c) Multiple point sources outside multiple subdomains. A full-ring transducer array, two image subdomains D_1 and D_2 (centered at $\mathbf{r}'_{\text{c},1}$ and $\mathbf{r}'_{\text{c},2}$, respectively), and a group of source points: $\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \dots$

In summary, to implement spatiotemporal antialiasing in D_{sub} , we first apply location-dependent temporal filtering (LDTF) to the recentered signals $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$ of each element with an upper cutoff frequency $f_{c,D_{\text{sub}}}(\mathbf{r}_n)$, and obtain $\hat{p}_{D_{\text{sub}},\text{LDTF}}(\mathbf{r}_n, t')$, $n = 1, 2, \dots, N$. Then we apply spatial interpolation with a factor β to the filtered signals and obtain $\tilde{p}_{D_{\text{sub}}}(\mathbf{r}_{\beta,n}, t')$, $n = 1, 2, \dots, \beta N$. After reversing the temporal

recentering of the signals based on the relation

$$\tilde{p}(\mathbf{r}_{\beta,n}, t) = \tilde{p}_{D_{\text{sub}}} \left(\mathbf{r}_{\beta,n}, t - \frac{\|\mathbf{r}'_{\text{c},\text{sub}} - \mathbf{r}_{\beta,n}\|}{c} \right),$$

$$n = 1, 2, \dots, \beta N, t \geq 0, \quad (16)$$

we reconstruct the image in D_{sub} using (2) for the denser element locations.

B. With a single point source outside the subdomain

We increase the complexity of our theory by adding a single point source at \mathbf{r}' outside the subdomain D_{sub} , as shown in Fig. 1(b). For the single source point \mathbf{r}' , (5) reduces to

$$\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t') = \frac{vp_0(\mathbf{r}')}{4\pi c^2 \|\mathbf{r}' - \mathbf{r}_n\|}$$

$$h'_e \left(t' - \left(\frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_{\text{c},\text{sub}} - \mathbf{r}_n\|}{c} \right) \right),$$

$$n = 1, 2, \dots, N, \quad (17)$$

where $p_0(\mathbf{r}')$ is the initial pressure at $\mathbf{r}' \in D \setminus D_{\text{sub}}$ and v is the volume of the source point. Similarly, (3) reduces to

$$\hat{p}_0(\mathbf{r}'') \approx \frac{vp_0(\mathbf{r}')}{2\pi c^2} \sum_{n=1}^N \frac{w_n}{\|\mathbf{r}' - \mathbf{r}_n\|} \left(1 - \left(t + \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} \right) \frac{\partial}{\partial t} \right)$$

$$h'_e \left(\frac{\|\mathbf{r}'' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} \right), \mathbf{r}'' \in D_{\text{sub}}. \quad (18)$$

Here, we confine the image reconstruction in D_{sub} by letting $\mathbf{r}'' \in D_{\text{sub}}$.

We first analyze spatial aliasing in spatial sampling for signals from the source point \mathbf{r}' . Based on (17), the spatial aliasing is determined by the step size of $\frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_{\text{c},\text{sub}} - \mathbf{r}_n\|}{c}$ as n varies. Using (7), we express the step size as $\tau(\mathbf{r}_n, \mathbf{r}_{n'}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}')$ for adjacent element locations \mathbf{r}_n and $\mathbf{r}_{n'}$, and we define the upper cutoff frequency as

$$f_{c,SS,D_{\text{sub}},OS}(\mathbf{r}_n, \mathbf{r}') = \min_{\mathbf{r}_{n'} \text{ is adjacent to } \mathbf{r}_n} \frac{1}{2\tau(\mathbf{r}_n, \mathbf{r}_{n'}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}')},$$

$$n, n' = 1, 2, \dots, N. \quad (19)$$

To remove aliasing in spatial sampling (SS in the subscript) for signals from the source point \mathbf{r}' outside the subdomain (OS in the subscript), we apply lowpass filtering to $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$ with the above upper cutoff frequency. We assume that the value of $h'_e(t)$ is nonzero only for t in an interval $[0, T_e]$ ($[0, 1.8 \mu\text{s}]$ for this research), which is often small for the transducers used in PACT. To minimize unwanted smoothing of signals, we filter $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$ only for $t' - \left(\frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}'_{\text{c},\text{sub}} - \mathbf{r}_n\|}{c} \right) \in [0, T_e]$.

Then we analyze spatial aliasing in the image reconstruction for signals from the source point \mathbf{r}' based on (18). We estimate the upper limit of the step size of $\frac{\|\mathbf{r}'' - \mathbf{r}_n\|}{c} - \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}$ as n varies using

$$\tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}', \mathbf{r}'') \leq \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}', \mathbf{r}'_{\text{c},\text{sub}}) + \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}'_{\text{c},\text{sub}}, \mathbf{r}'')$$

$$\leq \tau(\mathbf{r}, \mathbf{r}_{\text{adj}}, \mathbf{r}', \mathbf{r}'_{\text{c},\text{sub}}) + \tau_{D_{\text{sub}}}(\mathbf{r}, \mathbf{r}_{\text{adj}}),$$

$$\mathbf{r}' \in D \setminus D_{\text{sub}}, \mathbf{r}'' \in D_{\text{sub}}. \quad (20)$$

To remove aliasing in the image reconstruction (IR in the following subscript) for signals from the source point \mathbf{r}' , we can

apply a second lowpass filtering to $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$ with an upper cutoff frequency

$$f_{c,IR,D_{\text{sub}},OS}(\mathbf{r}_n, \mathbf{r}') = \frac{1}{2 \left(\tau(\mathbf{r}_n, \mathbf{r}_{n'}, \mathbf{r}', \mathbf{r}'_{\text{c},\text{sub}}) + \tau_{D_{\text{sub}}}(\mathbf{r}_n, \mathbf{r}_{n'}) \right)},$$

$$n, n' = 1, 2, \dots, N. \quad (21)$$

To avoid the second lowpass filtering, after the first lowpass filtering, we apply spatial interpolation to the filtered signals and obtain the interpolated signals at virtual locations $\mathbf{r}_{\beta,n}, n = 1, 2, \dots, \beta N$. Updating (21) with these virtual locations, we obtain

$$f_{c,IR,D_{\text{sub}},OS,\beta}(\mathbf{r}_{\beta,n}, \mathbf{r}') = \frac{1}{2 \left(\tau(\mathbf{r}_{\beta,n}, \mathbf{r}_{\beta,n'}, \mathbf{r}', \mathbf{r}'_{\text{c},\text{sub}}) + \tau_{D_{\text{sub}}}(\mathbf{r}_{\beta,n}, \mathbf{r}_{\beta,n'}) \right)},$$

$$n, n' = 1, 2, \dots, \beta N. \quad (22)$$

Here, we still have the relation $\mathbf{r}_n = \mathbf{r}_{\beta,\beta(n-1)+1}, n = 1, 2, \dots, N$. We choose a minimal β such that

$$f_{c,IR,D_{\text{sub}},OS,\beta}(\mathbf{r}_{\beta,\beta(n-1)+1}, \mathbf{r}') \geq f_{c,SS,D_{\text{sub}},OS}(\mathbf{r}_n, \mathbf{r}'),$$

$$n = 1, 2, \dots, N. \quad (23)$$

The factor β is obtained through numerical computations.

In practice, we have signals from both the subdomain D_{sub} and the source point \mathbf{r}' . First, we rewrite (19) (for the source point \mathbf{r}'), which applies to only the time domain subsets, as

$$f_{c,SS,D_{\text{sub}},OS}(\mathbf{r}_n, t', \mathbf{r}') = \begin{cases} f_{c,SS,D_{\text{sub}},OS}(\mathbf{r}_n, \mathbf{r}'), & t' - t'_0 \in T_e \\ f_{c,IS}, & \text{else} \end{cases},$$

$$n = 1, 2, \dots, N, \quad (24)$$

which applies to the whole time domain. Here, $t'_0 = \frac{\|\mathbf{r}' - \mathbf{r}_n\|}{c}$ denotes the recentered first arrival time from \mathbf{r}' to \mathbf{r}_n , and $f_{c,IS}$ denotes the upper cutoff frequency of the imaging system (IS in the subscript). Next, we combine the upper cutoff frequencies in (11) (for the subdomain D_{sub}) and (24) to yield the following upper cutoff frequency of the recentered signal $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$:

$$f_{c,D_{\text{sub}},r'}(\mathbf{r}_n, t') = \min\{f_{c,D_{\text{sub}}}(\mathbf{r}_n), f_{c,SS,D_{\text{sub}},OS}(\mathbf{r}_n, t', \mathbf{r}')\}. \quad (25)$$

For the n -th element at time t' , by applying lowpass filtering with the above upper cutoff frequency, we remove aliasing in spatial sampling for signals from both the subdomain D_{sub} and the source point \mathbf{r}' . This LDTF process is computationally intensive if implemented directly. In this research, we provide an efficient implementation through precomputation and interpolation, as shown in Appendix A. Based on (31) in Appendix A, we express the filtered signals as $\hat{p}_{D_{\text{sub}},\text{LDTF},r'}(\mathbf{r}_n, t')$. Further, we apply to the filtered signals a spatial interpolation with the minimal factor β satisfying (15) and (23), and reverse the recentering of the interpolated signals. We finish the reconstruction in the subdomain D_{sub} based on (2) for the virtual locations and denote the reconstructed image in D_{sub} as $\hat{p}_{0,D_{\text{sub}},r'}(\mathbf{r}''), \mathbf{r}'' \in D_{\text{sub}}$.

C. With multiple point sources outside the subdomain

We further extend our theory to include multiple point sources outside the subdomain D_{sub} . We denote the set of source points

outside D_{sub} as $G = \{\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}'_3, \dots\}$, as shown in Fig. 1(b), and update the upper cutoff frequency as follows:

$$f_{c,D_{\text{sub}},G}(\mathbf{r}_n, t') = \min \left\{ f_{c,D_{\text{sub}}}(\mathbf{r}_n), \min_{\mathbf{r}' \in G} f_{c,SS,D_{\text{sub}},OS}(\mathbf{r}_n, t', \mathbf{r}') \right\}, \quad n = 1, 2, \dots, N. \quad (26)$$

To remove spatial aliasing in the spatial sampling for signals from the subdomain D_{sub} and the source points in G , we apply lowpass filtering with the above upper cutoff frequency to the recentered signal $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$ of the n -th element at time t' , and obtain $\hat{p}_{D_{\text{sub}},LDTF,G}(\mathbf{r}_n, t')$ (31) in Appendix A). Then we find the minimal β such that (15) and (23) are satisfied for all $n = 1, 2, \dots, N$ and $\mathbf{r}' \in G$. Spatially interpolating the filtered signals with the factor β , reversing the recentering of the interpolated signals, and using (2) for the denser locations, we reconstruct the image for the subdomain D_{sub} , denoted as $\hat{p}_{0,D_{\text{sub}},G}(\mathbf{r}''), \mathbf{r}'' \in D_{\text{sub}}$.

D. With general sources outside the subdomain

In the fourth case, we consider general sources outside the subdomain D_{sub} . A direct method for LDTF with general sources is selecting all voxels outside D_{sub} as source points (grouped as G) and using (26) to obtain the upper cutoff frequencies for lowpass filtering. However, this direct method causes severe blurring in the reconstructed images due to unwanted filtering. To minimize unwanted filtering during spatiotemporal antialiasing, we select multiple sets of sparsely distributed source points in the image domain (Appendix B), denoted as G_1, G_2, \dots, G_J . We repeat the process in the previous case for each G_j and obtain an image of the subdomain D_{sub} , denoted as $\hat{p}_{0,D_{\text{sub}},G_j}(\mathbf{r}''), \mathbf{r}'' \in D_{\text{sub}}, j = 1, 2, \dots, J$. The final image for the subdomain D_{sub} is obtained through averaging

$$\hat{p}_{0,D_{\text{sub}}}(\mathbf{r}'') = \frac{1}{J} \sum_{j=1}^J \hat{p}_{0,D_{\text{sub}},G_j}(\mathbf{r}''), \mathbf{r}'' \in D_{\text{sub}}. \quad (27)$$

For $\mathbf{r}'' \in D \setminus D_{\text{sub}}$, we define $\hat{p}_{0,D_{\text{sub}}}(\mathbf{r}'')$ as zero.

IV. DIVISION OF IMAGE DOMAIN AND MOSAICKING OF SUBDOMAIN IMAGES

We divide the whole image domain D into multiple subdomains D_1, D_2, \dots, D_I , as shown in Appendix B. In Fig. 1(c), we depict two subdomains with a group of outside point sources. To form the whole image, these subdomains must satisfy

$$D = D_1 \cup D_2 \cup \dots \cup D_I. \quad (28)$$

To mitigate artifacts caused by pixel-value mismatch on subdomain boundaries, we overlap adjacent subdomains by a length of ξ_{DD} (Fig. 8(b) in Appendix B). Then, for each subdomain D_i , we repeat the process described in the fourth case above to obtain $\hat{p}_{0,D_i}(\mathbf{r}''), i = 1, 2, \dots, I$. Finally, we mosaic these subdomain images to form the whole image:

$$\hat{p}_0(\mathbf{r}'') \approx \sum_{i=1}^I \hat{w}_{l_x(D_i), l_y(D_i), \xi_{\text{DD}}}(\mathbf{r}'' - \mathbf{r}'_{c,i}) \hat{p}_{0,D_i}(\mathbf{r}''), \mathbf{r}'' \in D. \quad (29)$$

Here, $l_x(D_i)$ and $l_y(D_i)$ denote the sizes of the rectangle D_i in x -axis and y -axis directions, respectively, and $\mathbf{r}'_{c,i}$ is the center

of D_i . The normalized weight function \hat{w} is defined by (34) in Appendix B.

In summary, we have the general workflow of the location-dependent spatiotemporal antialiasing for PACT, shown in Fig. 2.

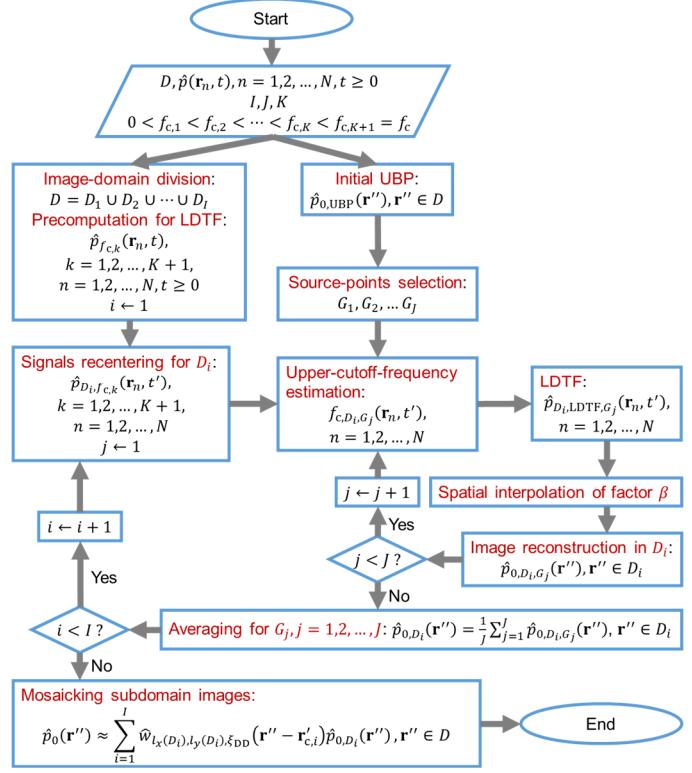


Fig. 2. Workflow of the location-dependent spatiotemporal antialiasing for PACT.

V. NUMERICAL SIMULATIONS WITH POINT SOURCES

A. Spatial aliasing in the image domain and signal domain

Before applying spatiotemporal antialiasing, we first visualize the spatial aliasing in both the image domain and signal domain through simulations with point sources. In fact, we were inspired to propose LDTF by observing the connection between spatial aliasing in these two domains. We use the MATLAB k-wave toolbox [31] for the 2D forward simulation with a ring transducer array of radius $R = 110$ mm. We let the frequency range of the transducer be from 0.1 MHz to 4.5 MHz (2.3-MHz central frequency, 191% one-way bandwidth, the upper cutoff frequency $f_{c,IS} = 4.5$ MHz) and the number of transducer elements be $N = 512$. We set the speed of sound as $c = 1.5 \text{ mm} \cdot \mu\text{s}^{-1}$. The shorter cutoff wavelength of this transducer is $\lambda_c = \frac{c}{f_{c,IS}} \approx 0.33$ mm.

In this simulation, non-zero initial pressure exists only at three point sources A, B, and C located at \mathbf{r}_A , \mathbf{r}_B , and \mathbf{r}_C , respectively. We reconstruct an image of the initial pressure from the simulated signals using the UBP method, shown in Fig. 3(a) with the three points labeled. Both this image and the ring array are centered at the origin. We denote the one-way Nyquist zone $S_1 = \{\mathbf{r}' \mid \|\mathbf{r}'\| < \frac{N\lambda_c}{4\pi}\}$ for the ring array [15] as a

blue-solid circle in Fig. 3(a), where the radius of S_1 is $\frac{N\lambda_c}{4\pi} \approx 13.6$ mm. For source points in S_1 , there is no aliasing in the spatial sampling. Closeup images of both subdomains (D_1 and D_2) in the red-dashed box and the yellow-dashed box (Fig. 3(b) and (c)) show aliasing artifacts.

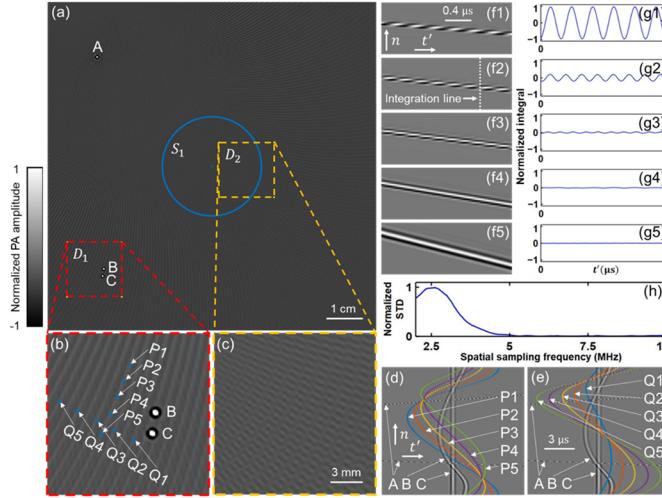


Fig. 3. Spatial aliasing in the image domain and signal domain. (a) Reconstructed image of three point sources (A, B, and C) with the one-way Nyquist zone S_1 (blue-solid circle). (b) and (c) Closeup subsets of (a) in the (b) red-dashed box and (c) the yellow-dashed box. Reconstruction locations P1–P5 along an artifact streak and Q1–Q5 from different artifact streaks are picked in (b) to identify the sources of artifacts. (d) Signals recentered based on (5) with $\mathbf{r}'_{c,1} = \mathbf{r}_C$. The signals used in the reconstructions at locations P1–P5 are marked by dotted curves with different colors. (e) The same recentered signals as in (d) but marked by colored-dotted curves showing signals used for the reconstructions at locations Q1–Q5. (f1)–(g1), (f2)–(g2), (f3)–(g3), (f4)–(g4), and (f5)–(g5) Signals acquired with spatial sampling frequencies of (f1) 3.33 MHz, (f2) 4.00 MHz, (f3) 5.00 MHz, (f4) 6.67 MHz, and (f5) 10.00 MHz, and their integration values along the respective dashed vertical lines ((g1)–(g5)). (h) Normalized STD of the integration value versus the spatial sampling frequency.

To identify the sources of aliasing artifacts in the signal domain, we recenter the detected signals based on (5) by letting $\mathbf{r}'_{c,1} = \mathbf{r}_C$. The recentered signals are shown in both Fig. 3(d) and (e), with the horizontal direction representing the time (t') and the vertical direction as the element index (n). For better visualization, the recentered signals are truncated in the temporal dimension while still containing all signals from the subdomain D_1 . Because the spatial Nyquist criterion is satisfied after recentering, signals from both B and C are smooth in both spatial and temporal dimensions. By contrast, signals from A appear dashed in Fig. 3(d) and (e) because the spatial Nyquist criterion is violated.

We show that these dashed patterns are sources of aliasing artifacts in Fig. 3(b) by visualizing the connection between the image domain and signal domain. In fact, from (2), we know that reconstruction at each point in the image domain using the UBP method is a weighted integration of a subset in the signal domain. In subdomain D_1 , we pick reconstruction locations P1–P5 along an artifact streak and locations Q1–Q5 on different artifact streaks, as shown in Fig. 3(b). Subsets for integrations in the signal domain for the two groups of reconstruction locations are shown as colored-dotted curves in Fig. 3(d) and (e), respectively. In Fig. 3(d), the five colored-dotted curves all

intersect with the signals from point A at approximately the same point. Spatial aliasing at this point in the signal domain is transformed by the UBP method to an artifact streak in the image domain. In other words, sharing the same point in the signal domain corresponds to sharing the same artifact streak in the image domain. In Fig. 3(e), the five colored-dotted curves intersect with the signals from point A at different points, which correspond to different artifact streaks in the image domain. In summary, aliasing artifacts in a subdomain are caused by dashed portions of signals that are recentered to this subdomain. For the subdomain D_2 , the spatially dashed portions of recentered signals may come from all three points, which are far from D_2 . Thus, the artifacts in Fig. 3(c) have more complex patterns than those in Fig. 3(b).

We further intuitively explain spatial aliasing. The colored-dotted curves in Fig. 3(d) and (e) intersect with not only the signals from point A but also the signals from points B and C. Signals from points B and C do not contribute to the aliasing artifacts in Fig. 3(b), which can be explained by the Nyquist criterion based on (5), as well as by the following intuitive geometric description. For each combination of a source point and a reconstruction location, the signals from the source point and the integration subset intersect. Both the signals and the integration subset have temporal step sizes when the element index n varies. It is the difference between the two temporal step sizes at the intersection points that determines the amplitude of the aliasing artifact at the reconstruction location. Signals with differential temporal step sizes of 0.30 μ s, 0.25 μ s, 0.20 μ s, 0.15 μ s, and 0.10 μ s (corresponding to spatial sampling frequencies 3.33 MHz, 4.00 MHz, 5.00 MHz, 6.67 MHz, and 10.00 MHz, respectively) are shown in Fig. 3(f1)–(f5), respectively. In the simplest case that the temporal step size of the integration subset is approximately zero within a sufficiently small subdomain, we let the temporal step size of the signals vary. Thus, the integration subsets in Fig. 3(f1)–(f5) are vertical lines as indicated by the white-dotted line in Fig. 3(f2). Normalized integration values (integrals) at different times in Fig. 3(f1)–(f5) are shown in Fig. 3(g1)–(g5), respectively. We see that the oscillation amplitude of the integral decreases as the spatial sampling frequency increases. Comparing integrations in Fig. 3(f1) and Fig. 3(f5), we see a key difference: at different times in Fig. 3(f1), the values along an integration line are dominated by positive values or negative values, resulting in fluctuating integrals; however, in Fig. 3(f5), there is no such dominance, resulting in a close-to-zero integral at any time. Thus, intuitively, the amplitudes of aliasing artifacts (caused by signals from a given source point) in the image domain are determined by the dominance of values of certain signs in the integration subsets, which are further determined by the relative temporal step sizes of the signals and integration subsets at their intersection points. Using standard deviation (STD) to quantify the amplitude of integrals in Fig. 3(g1)–(g5), we obtain the amplitude's dependency on the spatial sampling frequency, as shown in Fig. 3(h). Based on the Nyquist criterion, the amplitude is negligible for a spatial sampling frequency greater than $2f_{c,IS} = 9.0$ MHz.

B. Image reconstruction with spatiotemporal antialiasing

To mitigate aliasing artifacts, our previous work [15] used spatial interpolation and radius-dependent temporal filtering (RDTF) for a centered subdomain without strong outside sources, and here we propose LDTF for off-centered subdomains. We consider image reconstructions for only the two subdomains, D_1 and D_2 , shown in Fig. 3(a).

We first visualize the final signals used for the reconstruction of the subdomain D_1 in each method. For UBP without antialiasing, the signals ($\hat{p}_{D_1}(\mathbf{r}_n, t')$) recentered based on (5) with $\mathbf{r}'_{c,1} = \mathbf{r}_C$ are shown in Fig. 4(a1) (the same as in Fig. 3(d) and (e)). For UBP with spatial interpolation, the original signals ($\hat{p}(\mathbf{r}_n, t)$) are interpolated in spatial dimension then recentered for better comparison, as shown in Fig. 4(a2). For UBP with RDTF and spatial interpolation, the original signals ($\hat{p}(\mathbf{r}_n, t)$) are processed first by RDTF then by spatial interpolation. The results are further recentered for better comparison, as shown in Fig. 4(a3). For UBP with LDTF and spatial interpolation, the recentered signals ($\hat{p}_{D_1}(\mathbf{r}_n, t')$) are processed first by LDTF with only one set of source points $G = \{\mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C\}$ shown in Fig. 3(a) then by spatial interpolation, resulting in Fig. 4(a4).

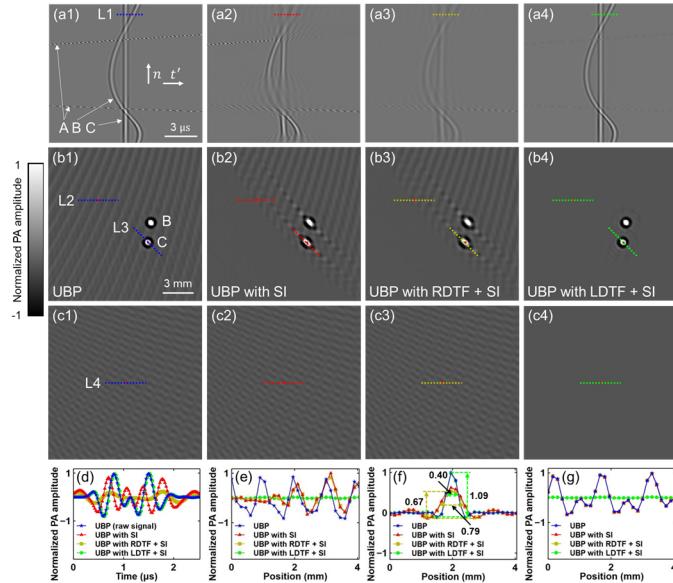


Fig. 4. Spatiotemporal antialiasing for the image reconstruction of point sources. (a1)–(a4) Final spatiotemporal signals used for reconstructions of the subdomain D_1 in UBP, UBP with spatial interpolation (SI), UBP with RDTF and SI, and UBP with LDTF and SI. (b1)–(b4) Images of the subdomain D_1 reconstructed using (b1) UBP, (b2) UBP with SI, (b3) UBP with RDTF and SI, and (b4) UBP with LDTF and SI. (c1)–(c4) Images of the subdomain D_2 reconstructed using the four methods to show artifacts only. (d)–(g) Comparisons of the values along dashed lines (d) L1, (e) L2, (f) L3, and (g) L4, respectively, for the four methods. The full width at half maximum (FWHM) of the main lobe in (f) is 0.79 mm for UBP with RDTF and SI, and 0.40 mm for UBP with LDTF and SI. The amplitudes are 0.67 and 1.09, respectively.

The image of the subdomain D_1 reconstructed using UBP, UBP with spatial interpolation, UBP with RDTF and spatial interpolation, and UBP with LDTF and spatial interpolation are shown in Fig. 4(b1)–(b4), respectively. Aliasing artifacts in the subdomain D_1 are caused by signals from the source point A. These artifacts appear in the whole subdomain in Fig. 4(b1) and partially in Fig. 4(b2) and (b3). They are substantially mitigated

in the whole subdomain in Fig. 4(b4). In addition, the images of point sources B and C in Fig. 4(b4) maintain the isotropy shown in (b1) while those in Fig. 4(b2) and (b3) are anisotropically blurred. These differences in aliasing artifacts and image blurring are determined by the final signals used for reconstruction in the different methods. Comparing Fig. 4(a1) and (a2), we see that the signals from A (sources of aliasing artifacts in D_1) are partially mitigated by spatial interpolation and the signals from B and C are partially disrupted by spatial interpolation. These partial mitigation and disruption are caused by the fact that the spatial Nyquist criterion is only partially satisfied in the original signals ($\hat{p}(\mathbf{r}_n, t)$). The partial mitigation of the signals from A causes partial mitigation of the aliasing artifacts in D_1 , whereas the partial disruption of the signals from B and C causes anisotropic blurring in D_1 , as shown in Fig. 4(b2). RDTF further blurs signals from all the three points, as shown in Fig. 4(a3); however, large numbers of aliasing artifacts remain, as shown in Fig. 4(b3). In contrast, LDTF with spatial interpolation substantially suppresses signals from A while maintaining signals from B and C. Thus, as shown in Fig. 4(b4), LDTF substantially mitigates aliasing artifacts and maintains the spatial resolution.

Repeating the four reconstruction methods for the subdomain D_2 yields the images shown in Fig. 4(c1)–(c4). Aliasing artifacts in D_2 are caused by signals from all the three point sources. Neither RDTF nor spatial interpolation effectively mitigates signals from these points, resulting in abundant aliasing artifacts in the reconstructed images, as shown in Fig. 4(c1)–(c3). By contrast, LDTF with spatial interpolation substantially suppresses signals from all the three points, resulting in markedly reduced aliasing artifacts, as shown in Fig. 4(c4).

For quantitative comparisons, we first pick a line L1 labeled in Fig. 4(a1) with the values along the line for the four methods shown in Fig. 4(d). In comparison to the raw signal, signals from B and C are distorted by spatial interpolation, blurred by RDTF, but maintained by LDTF with spatial interpolation. To compare the effects in the image domain, we draw two lines L2 and L3 in the subdomain D_1 , as shown in Fig. 4(b1). The values along lines L2 and L3, respectively, for the four methods, are compared in Fig. 4(e) and (f). LDTF with spatial interpolation is superior in antialiasing as shown in Fig. 4(e) while maintaining the spatial resolution as shown in Fig. 4(f). However, spatial interpolation anisotropically blurs the image through signal disruption; RDTF further blurs the image through temporal lowpass filtering. These blurring effects in the image domain are shown in Fig. 4(f). To compare aliasing artifacts in D_2 , we show the values along a line L4 labeled in Fig. 4(c1) for the four methods in Fig. 4(g). We see that LDTF with spatial interpolation is still superior for antialiasing in D_2 . Note that RDTF and spatial interpolation's ineffectiveness for antialiasing in D_1 and D_2 is not contradictory with the results in [15], where aliasing artifacts are mainly caused by signals from sources close to or inside the one-way Nyquist zone S_1 .

VI. NUMERICAL SIMULATIONS OF BLOOD VESSEL PHANTOMS

We further validate the proposed method using two numerical blood vessel phantoms. The first numerical phantom consists of simple blood vessel structures, as shown in Fig. 5(a1). Two subsets of the simple numerical phantom enclosed in the red-dashed box and the yellow-dashed box, respectively, are shown in Fig. 5(b1) and (c1).

In the point source simulations presented in the above section, we select only two subdomains and all the three source points for spatiotemporal antialiasing, as shown in Fig. 4(a4), (b4), and (c4). For a general numerical phantom with blood vessel structures, multiple parameters are used to control the image reconstruction. As shown in Appendix B, the selection of source points is controlled by subdomain size l_{SP} , candidacy ratio α , and the number of sets J ; while the image-domain division is controlled by subdomain size l_{DD} and overlapping size ξ_{DD} . Additionally, we reduce T_e in (24) to an optimal value T_o to minimize unwanted blurring of images. For more efficient tuning of the ranges, we scale T_o for each subdomain to a balanced location-dependent parameter $T_b(D_i)$, $i = 1, 2, \dots, I$, as proposed in Appendix C. In summary, six parameters $(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_o)$ need to be tuned for best reconstruction. In Appendix D, we propose a parameter-tuning strategy based on alternating-direction optimization.

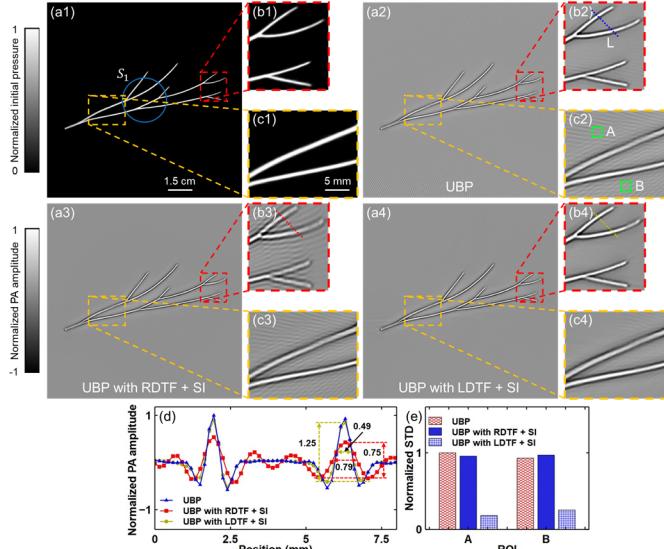


Fig. 5. Applying spatiotemporal antialiasing to the image reconstruction of a numerical phantom with simple blood vessel structures. (a1) A numerical phantom consisting of simple blood vessel structures. The one-way Nyquist zone S_1 with a radius of 13.6 mm is marked by a blue-solid circle. (b1) and (c1) Closeup subsets of the simple numerical phantom enclosed in the red-dashed box and the yellow-dashed box, respectively. (a2)–(c2), (a3)–(c3), and (a4)–(c4) Images of the simple numerical phantom reconstructed using UBP, UBP with RDTF and spatial interpolation (SI), and UBP with LDTF and SI, respectively, and their closeup subsets. A line L and two small regions A and B are picked for comparisons of the three methods. (d) Comparisons of the values along the line L for the three methods. The FWHM of one main lobe is 0.79 mm for UBP with RDTF and SI, and 0.49 mm for UBP with LDTF and SI. The amplitudes are 0.75 and 1.25, respectively. (e) Comparisons of the STDs of pixel values in regions A and B for the three methods.

Applying the parameter-tuning strategy to the first numerical phantom with simple structures, we choose

$(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_o)$ to be (0.6 mm, 0.02, 36, 18 mm, 1.8 mm, 1.2 μ s). Reconstructions of the simple numerical phantom using UBP, UBP with RDTF and spatial interpolation, and UBP with LDTF and spatial interpolation are shown in Fig. 5(a2)–(a4), respectively. Reconstructions of the subdomains in Fig. 5(b1) and (c1) using the three methods are shown in Fig. 5(b2)–(c2), (b3)–(c3), and (b4)–(c4), respectively. Comparing Fig. 5(c2)–(c4), we see that LDTF with spatial interpolation is more effective than RDTF with spatial interpolation in mitigating aliasing artifacts, which agrees with the comparison for point sources shown in Fig. 4(b1)–(b4). Moreover, comparing Fig. 5(b2) and (b3), we observe that RDTF with spatial interpolation compromises image resolution and introduces additional artifacts during antialiasing. These observations also agree with the results in Fig. 4(b1)–(b4) and are mainly caused by the fact that the spatial Nyquist criterion is partially unsatisfied after RDTF, as explained in Fig. 4(a1)–(a4). In summary, LDTF with spatial interpolation has better performance than RDTF with spatial interpolation in both mitigating aliasing artifacts and maintaining image resolution.

For quantitative comparisons, we draw a line L (marked in Fig. 5(b2)) in the red-boxed subdomain. The values along the line L for the three methods are shown in Fig. 5(d). We observe that RDTF with spatial interpolation blurs the image and introduces artifacts, whereas LDTF with spatial interpolation maintains the image resolution and maintains the low-amplitude background. It needs to be noted that aliasing artifacts only appear for certain combinations of source point and reconstruction location. The low-amplitude background around L in Fig. 5(b2) means that aliasing artifacts in this region are caused only by low-amplitude point sources. For quantitative comparisons of aliasing artifacts, we pick two small regions A and B (marked in Fig. 5(c2)) in the yellow-boxed subdomain. The amplitudes of aliasing artifacts in A and B are quantified by STDs of the pixel values in them. We compare these STDs in Fig. 5(e), which further validates that LDTF with spatial interpolation outperforms RDTF with spatial interpolation in mitigating aliasing artifacts.

The second numerical phantom is shown in Fig. 6(a1), which consists of blood vessels with complex structures. Two subsets of the complex numerical phantom are shown in Fig. 6(b1) and (c1), respectively. Applying the parameter-tuning strategy to the complex numerical phantom, we choose $(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_o)$ to be (3.6 mm, 0.08, 36, 18 mm, 1.8 mm, 1.2 μ s). Reconstructions of the complex numerical phantom using UBP, UBP with RDTF and spatial interpolation, and UBP with LDTF and spatial interpolation are shown in Fig. 6(a2)–(a4), respectively. The subdomain images reconstructed by the three methods are shown in Fig. 6(b2)–(c2), (b3)–(c3), and (b4)–(c4), respectively. Comparing these subdomain images, we see that LDTF with spatial interpolation more effectively mitigates aliasing artifacts than RDTF with spatial interpolation. Moreover, using numerical simulations of the complex numerical phantom, we demonstrate the advantage of using the location-dependent parameter $T_b(D_i)$ over using a constant parameter T_o across all subdomains for temporal filtering

(Appendix E).

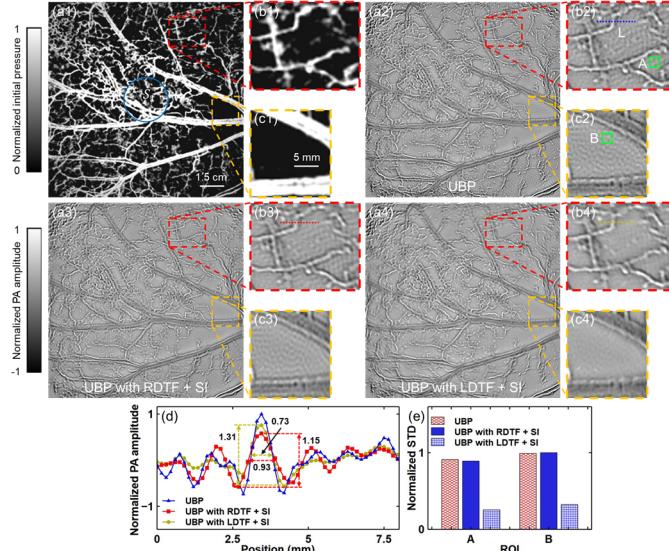


Fig. 6. Applying spatiotemporal antialiasing to the image reconstruction of a numerical phantom with complex blood vessel structures. (a1)–(c1) A numerical phantom consisting of complex blood vessel structures, and its closeup subsets. (a2)–(c2), (a3)–(c3), and (a4)–(c4) Images of the complex numerical phantom reconstructed using UBP, UBP with RDTF and spatial interpolation (SI), and UBP with LDTF and SI, respectively, and their closeup subsets. (d) Comparisons of the values along the line L marked in (b2) for the three methods. The FWHM of the dominant lobe is 0.93 mm for UBP with RDTF and SI, and 0.73 mm for UBP with LDTF and SI. The amplitudes are 1.15 and 1.31, respectively. (e) Comparisons of the STDs of pixel values in regions A and B marked in (b2) and (c2), respectively, for the three methods.

For quantitative comparisons, we draw a line L (marked in Fig. 6(b2)) in the red-boxed subdomain and pick two small regions A and B (marked in Fig. 6(b2) and (c2), respectively) in different subdomains for the three methods. The values along the line L are shown in Fig. 6(d) while the STDs of the pixel values in A and B are compared in Fig. 6(e). In Fig. 6(d), we still see that RDTF with spatial interpolation blurs the image and introduces new artifacts; whereas LDTF with spatial interpolation blurs the image to a smaller degree and mitigates the aliasing artifacts. The blurring effect of LDTF with spatial interpolation is more obvious in the complex numerical phantom than in the simple numerical phantom. In fact, compared with the simple numerical phantom, more source points are selected for LDTF in the complex numerical phantom, which results in the filtering of more signals and more blurring of the image. Thus, for images with complex structures, we find a balance between mitigating aliasing artifacts and maintaining image resolution by tuning the parameters. Because of this balance, aliasing artifacts still appear in Fig. 6(b4). In Fig. 6(e), we still observe that LDTF with spatial interpolation is better than RDTF with spatial interpolation in mitigating aliasing artifacts.

Due to the more intricate temporal filtering for antialiasing, the proposed method has a significantly higher computation cost. On a computer with Windows 10 Home and Intel® Core™ i7-6700 CPU @ 3.40 GHz, the reconstructions of the simple numerical phantom through UBP, UBP with RDTF and spatial interpolation, and UBP with LDTF and spatial interpolation (single-thread implementations) take 17.6 s, 98.8 s, and 682.7 s

(average values for 10 repetitions), respectively. For the complex numerical phantom, the computation times of the first two methods do not change but the third one takes 1345.5 s due to the differences in reconstruction parameters. All methods can be accelerated through GPU. For example, using an NVIDIA GeForce GTX 1050 Ti graphics card, we reduce the computation time of the UBP method from 17.6 s to 0.54 s. Although not demonstrated in this study, an efficient GPU acceleration of the proposed method is preferred for faster parameter tuning and image reconstruction in future studies.

VII. IN VIVO EXPERIMENTS

Finally, we apply the proposed LDTF with spatial interpolation to human breast imaging *in vivo*. The imaging system has been reported by Lin *et al.* [2], in which a 512-element full-ring ultrasonic transducer array (Imasonic, Inc., 110-mm radius, 2.25-MHz central frequency, 95% one-way bandwidth) was used. In our previous study [15], the cutoff frequency has been estimated to be $f_{c,IS} \approx 3.80$ MHz. Here we use the speed of sound $c = 1.49 \text{ mm} \cdot \mu\text{s}^{-1}$. A cross-sectional image of a breast *in vivo* is reconstructed using UBP and shown in Fig. 7(a1). The result of UBP with RDTF and spatial interpolation is shown in Fig. 7(a2).

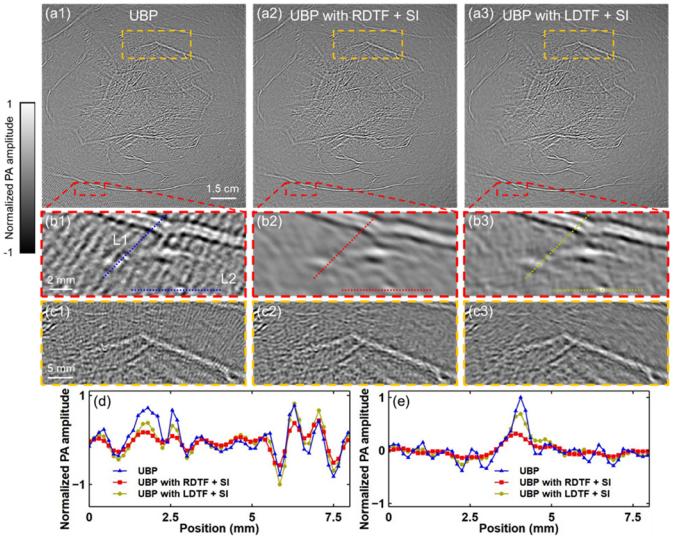


Fig. 7. Applying spatiotemporal antialiasing to human breast imaging *in vivo*. (a1)–(a3) Images of a human breast cross section *in vivo* reconstructed using (a1) UBP, (a2) UBP with RDTF and spatial interpolation (SI), and (a3) UBP with LDTF and SI. Two subdomains in the red-dashed box and the yellow-dashed box, respectively, are picked for comparisons of the three methods. (b1)–(c1), (b2)–(c2), and (b3)–(c3) Closeup images of the two subdomains for the three methods, respectively. Lines L1 and L2 are picked for comparisons. (d) and (e) Values along lines L1 and L2, respectively, for the three methods.

To use LDTF with spatial interpolation in UBP for image reconstruction, we tune the parameters $(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_0)$. This parameter tuning is simplified by comparing the parameters used for the numerical blood vessel phantoms. In those simulations, we use $(l_{SP}, \alpha) = (0.6 \text{ mm}, 0.02)$ for the simple numerical phantom, $(l_{SP}, \alpha) = (3.6 \text{ mm}, 0.08)$ for the complex numerical phantom, and $(J, l_{DD}, \xi_{DD}, T_0) = (36, 18 \text{ mm}, 1.8 \text{ mm}, 1.2 \mu\text{s})$ for both phantoms. Considering that the structural complexities of the simple numerical phantom and the complex numerical phantom occupy a complexity range

large enough for our study, we still use $(J, l_{DD}, \xi_{DD}, T_o) = (36, 18 \text{ mm}, 1.8 \text{ mm}, 1.2 \mu\text{s})$ for the breast imaging. Following the same parameter-tuning protocol, we select $(l_{SP}, \alpha) = (1.8 \text{ mm}, 0.04)$ for the breast imaging. In summary, we use $(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_o) = (1.8 \text{ mm}, 0.04, 36, 18 \text{ mm}, 1.8 \text{ mm}, 1.2 \mu\text{s})$ for the reconstruction of the breast image. The result of UBP with LDTF and spatial interpolation is shown in Fig. 7(a3).

For better comparisons, we select two subdomains in a red-dashed box and a yellow-dashed box, respectively, as shown in Fig. 7(a1)–(a3). Closeup images of the subdomains are shown in Fig. 7(b1)–(c1), (b2)–(c2), and (b3)–(c3) for the three methods, respectively. Comparing Fig. 7(b1)–(b3), we see that both RDTF with spatial interpolation and LDTF with spatial interpolation mitigate aliasing artifacts. However, RDTF with spatial interpolation compromises image resolution. From Fig. 7(c1)–(c3), we see that LDTF with spatial interpolation is more effective than RDTF with spatial interpolation in mitigating the aliasing artifacts. For quantitative comparisons, we pick two lines L1 and L2 in the red-boxed subdomain (Fig. 7(b1)). Pixel values along these two lines, respectively, for the three methods are shown in Fig. 7(d) and (e), which further validate that LDTF with spatial interpolation is more effective than RDTF with spatial interpolation in both mitigating aliasing artifacts and maintaining image resolution.

VIII. CONCLUSIONS AND DISCUSSION

In this research, we proposed an antialiasing method for PACT based on LDTF with spatial interpolation, which exhibits better performance in mitigating aliasing artifacts while maintaining image resolution. We applied this method to UBP and validated it through numerical simulations and *in vivo* experiments. To apply this method, we first divide the image domain into subdomains and select multiple groups of source points with maximum amplitudes from an initial image reconstructed using UBP. Then for each subdomain and each group of source points, we temporally filter the signals from the source points that overlap with signals from the subdomain. We recentered signals for this subdomain, apply spatial interpolation to the recentered signals, and use them to reconstruct the image in the subdomain. In this process, doing temporal filtering only for signals from source points with high amplitudes is essential for mitigating the dominant aliasing artifacts while minimizing unwanted blurring of the image. Location-dependent recentering of signals before spatial interpolation is essential for protecting signals from the subdomain of interest during spatial interpolation and maintaining image resolution. The proposed method outperforms our previous method based on RDTF with spatial interpolation in mitigating aliasing artifacts and maintaining image resolution.

To get the best performance of LDTF with spatial interpolation, we analyzed the sensitivities of all the parameters. We found that parameters $(J, l_{DD}, \xi_{DD}, T_o)$ are relatively insensitive for this study whereas parameters (l_{SP}, α) are sensitive. Thus, in all numerical simulations and *in vivo* experiments, we have the same choice of the four parameters $(J, l_{DD}, \xi_{DD}, T_o)$ and only vary the other two (l_{SP}, α) . For a

certain application, (l_{SP}, α) can be varied for different preferences for mitigating aliasing artifacts and maintaining image resolution. Also, we proposed just one strategy for determining the location-dependent parameter $T_b(D_i)$ given the parameter T_o and a group of source points, which can be further improved in future studies. Moreover, in applications with subdomain illuminations or patterned illuminations, we can select the source points based on prior knowledge of the illumination patterns to achieve more efficient LDTF.

The spatiotemporal analysis in this research relies on the homogeneous-medium assumption, which is satisfied in the numerical simulations. The effectiveness of LDTF with spatial interpolation in human breast imaging *in vivo* further validates that our theory applies to approximately homogenous media. For strongly inhomogeneous media, such as in transcranial PACT, more studies need to be done to demonstrate and improve the method's performance.

LDTF with spatial interpolation is applicable to other image reconstruction methods and other detection geometries. In fact, we can use LDTF with spatial interpolation as a filter for preprocessing and use another method for reconstruction. For example, we can use LDTF with spatial interpolation before a model-based iterative method with TV regularization to reduce the requirement of the regularization parameter and minimize unwanted blurring. We also can use LDTF with spatial interpolation before a deep neural network to reduce the network's burden in antialiasing, which potentially makes the network more robust. Importantly, the proposed method does not rely on a specific transducer array geometry, thus is directly applicable to other geometries, such as a linear array. Moreover, the proposed method is not limited to 1D arrays for 2D imaging. Through dimension decomposition, the method is applicable to 2D arrays, such as arrays of spherical, cylindrical, and planar geometries, allowing for spatiotemporal antialiasing in 3D imaging. An efficient GPU acceleration of the proposed method is preferred for faster parameter tuning and image reconstruction in future studies.

APPENDIX A AN EFFICIENT IMPLEMENTATION OF LDTF

If implemented directly, processing signal $\hat{p}_{D_{\text{sub}}}(\mathbf{r}_n, t')$ of each element at time t' using a lowpass filter with an upper cutoff frequency of f is computationally intensive. For fast reconstruction, we give an efficient implementation of the LDTF through precomputation and interpolation. Before the reconstruction of any subdomain, we process the original signals $\hat{p}(\mathbf{r}_n, t)$ of each element using lowpass filters with upper cutoff frequencies of $f_{c,k}, k = 1, 2, \dots, K + 1$ satisfying $0 < f_{c,1} < f_{c,2} < \dots < f_{c,K} < f_{c,K+1} = f_c$. Here, a lowpass filter with an upper cutoff frequency means a third-order lowpass Butterworth filter followed by a sinc filter with the same upper cutoff frequency. We denote the filtered signals as $\hat{p}_{f_{c,k}}(\mathbf{r}_n, t), k = 1, 2, \dots, K + 1$. For reconstruction of a subdomain D_{sub} , we recenter the filtered signals based on (4) and obtain

$$\hat{p}_{D_{\text{sub}}, f_{c,k}}(\mathbf{r}_n, t') = \hat{p}_{f_{c,k}}\left(\mathbf{r}_n, t' + \frac{\|\mathbf{r}'_{c,\text{sub}} - \mathbf{r}_n\|}{c}\right),$$

$$n = 1, 2, \dots, N, k = 1, 2, \dots, K + 1. \quad (30)$$

For a general cutoff frequency $f > 0$, we obtain the filtered signals through the following linear interpolation:

$$\hat{p}_{D_{\text{sub}}, f}(\mathbf{r}_n, t') = \begin{cases} \frac{f}{f_{c,1}} \hat{p}_{D_{\text{sub}}, f_{c,1}}(\mathbf{r}_n, t'), & 0 < f \leq f_{c,1} \\ \frac{f_{c,k+1} - f}{f_{c,k+1} - f_{c,k}} \hat{p}_{D_{\text{sub}}, f_{c,k}}(\mathbf{r}_n, t') + \frac{f - f_{c,k}}{f_{c,k+1} - f_{c,k}} \hat{p}_{D_{\text{sub}}, f_{c,k+1}}(\mathbf{r}_n, t'), & f_{c,k} < f \leq f_{c,k+1}, \\ \hat{p}_{D_{\text{sub}}, f_c}(\mathbf{r}_n, t'), & f > f_{c,k+1} = f_c \\ n = 1, 2, \dots, N, f > 0. & \end{cases} \quad (31)$$

In practice, we let the upper cutoff frequencies $f_{c,k}, k = 1, 2, \dots, K + 1$ be dense enough so that further increasing their density has minor effects on the reconstructed images. As an application, we substitute $f_{c,D_{\text{sub}}, \mathbf{r}'}(\mathbf{r}_n, t')$ ($f_{c,D_{\text{sub}}, G}(\mathbf{r}_n, t')$) for f in (31) to obtain $\hat{p}_{D_{\text{sub}}, \text{LDTF}, \mathbf{r}'}(\mathbf{r}_n, t')$ ($\hat{p}_{D_{\text{sub}}, \text{LDTF}, G}(\mathbf{r}_n, t')$).

APPENDIX B

SELECTION OF SOURCE POINTS AND DIVISION OF IMAGE DOMAIN

For LDTF with spatial interpolation, we select multiple groups of source points in the image domain and divide the image domain into subdomains for reconstructions. To select source points, we first reconstruct an image using the UBP method and select the αM pixels with the largest absolute values as source-point candidates, shown as white pixels in Fig. 8(a). Here, M is the number of all pixels, and α is the candidacy ratio. Then we divide the image domain into subdomains of size $l_{\text{SP}} \times l_{\text{SP}}$ for further selection. In each subdomain, if there exist source-point candidates, we randomly select one; otherwise, we do not select. The selected source points in all subdomains form a group of source points for LDTF. Repeating this random selection J times, we obtain J groups of source points: G_1, G_2, \dots, G_J . The first two groups are shown in Fig. 8(a) as blue dots and red triangles, respectively.

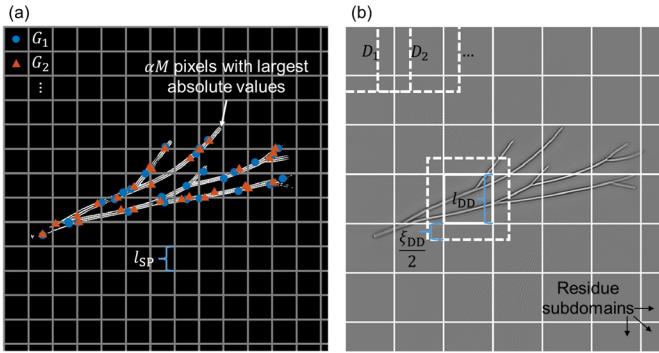


Fig. 8. Selection of source points and division of image domain. (a) Selection of source points from a UBP reconstructed image. The white pixels indicate the αM pixels with the largest absolute pixel values in the image. Subdomains have a size of $l_{\text{SP}} \times l_{\text{SP}}$ and are visualized by a grid. Two groups of selected source points are shown as blue dots (G_1) and red triangles (G_2), respectively. (b)

Domain division in a UBP reconstructed image. Initially, subdomains have a size of $l_{\text{DD}} \times l_{\text{DD}}$ (residual subdomains are smaller), shown as a grid. Then each subdomain is extended by $\frac{\xi_{\text{DD}}}{2}$ out of its boundaries inside the image domain.

For image-domain division, we start from one corner of the image domain and choose a square subdomain for every distance l_{DD} in each dimension, as shown in Fig. 8(b). The residual subdomains are rectangles whose sizes are determined by the residual lengths. To mitigate artifacts caused by pixel value mismatch on subdomain boundaries, we extend each subdomain by $\frac{\xi_{\text{DD}}}{2}$ outside of its boundaries inside the image domain. The extended subdomains are denoted as D_1, D_2, \dots, D_I . To mosaic the subdomain images, we define a 2D weight function

$$w_{l_x, l_y, \xi}(\mathbf{r}) = w_{l_x, l_y, \xi}(x, y) = w_{l_x, \xi}(x)w_{l_y, \xi}(y), \quad (32)$$

where the 1D weight function is defined as

$$w_{l, \xi}(x) = \begin{cases} 1, & |x| \leq \frac{l}{2} \\ 1 - \frac{2|x| - l}{\xi}, & \frac{l}{2} < |x| \leq \frac{l + \xi}{2} \\ 0, & |x| > \frac{l + \xi}{2} \end{cases} \quad (33)$$

We normalize the weight functions for these subdomains as

$$\begin{aligned} \hat{w}_{l_x(D_i), l_y(D_i), \xi_{\text{DD}}}(\mathbf{r}'' - \mathbf{r}'_{c,i}) = \\ \frac{w_{l_x(D_i), l_y(D_i), \xi_{\text{SD}}}(\mathbf{r}'' - \mathbf{r}'_{c,i})}{\sum_{i'=1}^I w_{l_x(D_{i'}), l_y(D_{i'}), \xi_{\text{DD}}}(\mathbf{r}'' - \mathbf{r}'_{c,i'})}, \\ i = 1, 2, \dots, I, \end{aligned} \quad (34)$$

where $l_x(D_i)$ and $l_y(D_i)$ denote the sizes of the rectangle D_i in x -axis and y -axis directions, respectively, and $\mathbf{r}'_{c,i}$ is the center of D_i . Then, we mosaic images in subdomains D_1, D_2, \dots, D_I through (29) to form the whole image in D .

APPENDIX C

LOCATION-DEPENDENT PARAMETER $T_b(D_i)$ FOR TEMPORAL FILTERING

The parameter T_o determines the range for temporal filtering: as T_o increases, wider temporal ranges of signals are filtered. In practice, we tune T_o to find a balance between mitigating aliasing artifacts and maintaining image resolution. In numerical simulations, we observe that for the same T_o , different amounts of signals are filtered for the reconstructions of different subdomains. This observation means that an optimal choice of T_o for one subdomain may not be optimal for another subdomain, which makes the tuning of T_o location-dependent and computationally intensive. To make the tuning of T_o more efficient, we propose a strategy to scale T_o for different subdomains automatically.

Before an adjustment of T_o , we first quantify the amount of temporal filtering for each subdomain through numerical simulation. In an image domain with a size of $120 \times 120 \text{ mm}^2$, we have 14×14 evenly distributed source points, shown as white dots in Fig. 9(a). We choose $l_{\text{DD}} = 12 \text{ mm}$ and $\xi_{\text{DD}} = 1.8 \text{ mm}$ for image-domain division, which results in 100 subdomains. Two of the subdomains, D_1 and D_2 , are marked in Fig. 9(a) as a red-dashed box and a yellow-dashed box, respectively. We conduct a forward simulation to visualize the

signals from the source points, which occupy a subset of the signal domain marked as white pixels in Fig. 9(b). Here, we use a relatively small parameter $T_o = 0.3 \mu\text{s}$ in the simulation to avoid saturation of the occupancy. In the reconstruction of a subdomain using the UBP method, signals from only a subset are used. For subdomains D_1 and D_2 , the used signal subsets are marked by two red curves and two yellow curves, respectively, in Fig. 9(b). The signals from the centers of the two subdomains are marked by two blue curves, respectively. Based on (4), we recenter signals in the two subsets with respect to the subdomain centers, as shown in Fig. 9(c1) and (c2), respectively. For each subdomain, the corresponding signal subset is partially occupied by signals from the source points: for D_1 and D_2 , the occupancy ratios are 0.405 and 0.606, respectively. We have shown that temporal filtering is determined not only by the occupancy ratio but also by the relative temporal step sizes. Here, as an approximation, we ignore the effects of the relative temporal step sizes and use only the occupancy ratio to quantify the amount of temporal filtering.

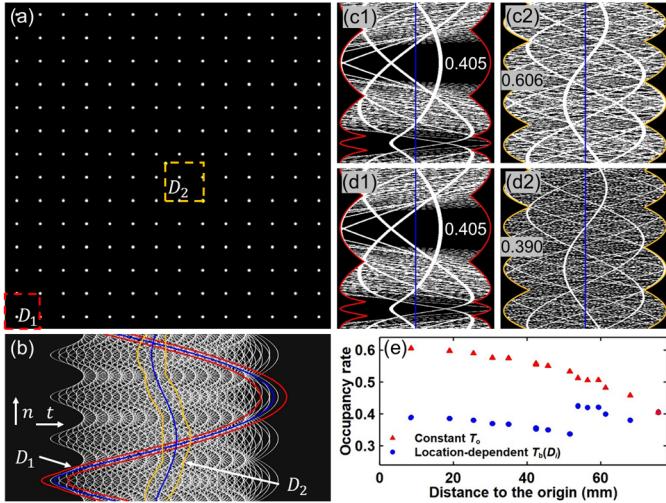


Fig. 9. Location-dependent parameter $T_b(D_i)$ for temporal filtering. (a) Evenly distributed source points, shown as white dots, and two subdomains D_1 and D_2 , marked by a red-dashed box and a yellow-dashed box, respectively. (b) The subset of the signal domain (white pixels) that is occupied by signals (with $T_o = 0.3 \mu\text{s}$) from the source points in (a). Signals from subdomains D_1 and D_2 occupy a region between two red curves and a region between two yellow curves, respectively. Signals from the centers of D_1 and D_2 are indicated by two blue curves. (c1) and (c2) Subsets of (b) recentered for D_1 and D_2 , respectively, showing a difference between the occupancy rates of signals from the source points: 0.405 for D_1 and 0.606 for D_2 . (d1) and (d2) The occupancy rates for D_1 and D_2 (0.405 and 0.390, respectively) after using the location-dependent parameter $T_b(D_i)$.

Next, we use the occupancy ratio to adjust T_o for each subdomain to achieve a low variation of occupancy ratios across subdomains. We denote the occupancy ratio of the subdomain D_i as o_i , $i = 1, 2, \dots, I$. Instead of using the same T_o for all subdomains, we use

$$T_b(D_i) = \left(\frac{\min_{1 \leq i \leq I} o_i}{o_i} \right)^\gamma T_o \quad (35)$$

for temporal filtering in the reconstruction of the subdomain D_i , $i = 1, 2, \dots, I$. Here, we use γ to account for the occupancy ratio's nonlinear dependency on T_o due to the overlapping of signals from different source points. After multiple tests, we choose $\gamma = 1.8$ to achieve a low variation of occupancy ratios

across subdomains. Applying the location-dependent parameter $T_b(D_i)$ to all the 100 subdomains, we obtain a new set of occupancy ratios. For subdomain D_1 , the occupancy ratio does not change. For subdomain D_2 , the occupancy ratio reduces from 0.606 to 0.390. The new signal occupancies for D_1 and D_2 are visualized in Fig. 9(d1) and (d2), respectively. The occupancy ratios for the constant T_o and the location-dependent $T_b(D_i)$ are compared in Fig. 9(e). From this comparison, we see that the variation of the occupancy ratios across subdomains is reduced by using the location-dependent parameter $T_b(D_i)$, which is controlled only by T_o for a given group of source points. In practice, we only tune T_o in (35), then the amount of temporal filtering in each subdomain is automatically adjusted by using $T_b(D_i)$. Thus, we simplify tuning T_o for each subdomain to tuning a single T_o for all subdomains.

APPENDIX D

PARAMETER SENSITIVITY ANALYSIS BASED ON ALTERNATING-DIRECTION OPTIMIZATION

The purpose of parameter tuning is to balance mitigating aliasing artifacts with maintaining image resolution. For this purpose, we quantify the amplitude of aliasing artifacts and image resolution for different choices of parameters. In a numerical phantom, we identify regions with zero initial pressure in the ground-truth image as background. For each reconstructed image, we use the STD of the background pixel values to quantify the amplitude of aliasing artifacts. As for image resolution, we use values along lines of interest in the reconstructed image for comparison. Location-dependent spatiotemporal antialiasing is affected by parameters used in the source points selection (l_{SP} , α , J), the image-domain division (l_{DD} , ξ_{DD}), and the temporal filtering T_o . Here, by tuning these parameters for image reconstruction of the simple numerical phantom, we analyze the proposed method's sensitivity to these parameters. This analysis is then used to guide the parameter tuning for the complex numerical phantom and the *in vivo* experiments.

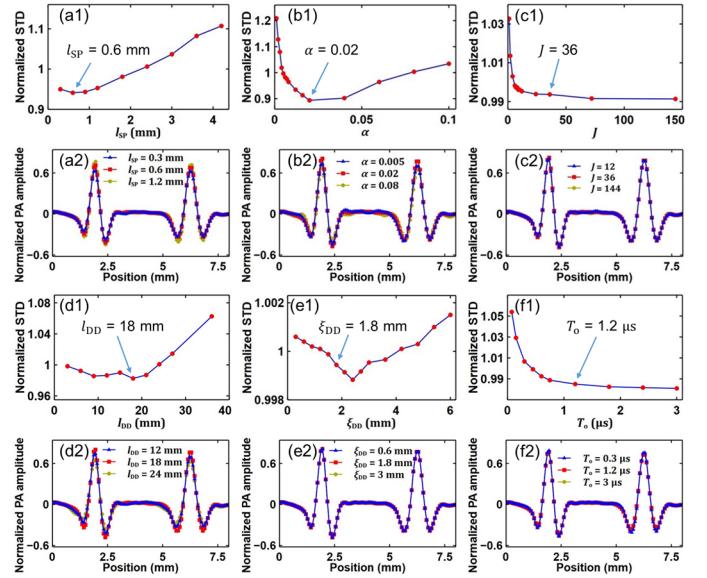


Fig. 10. Parameter sensitivity analysis. (a1)-(a2), (b1)-(b2), (c1)-(c2), (d1)-(d2), (e1)-(e2), and (f1)-(f2) Normalized STDs of the background pixel values and

values along the line L1 in Fig. 5(c2) for different choices of a parameter (l_{SP} , α , J , l_{DD} , ξ_{DD} , or T_o , respectively), while others are kept as constants.

Exploring all combinations of the six parameters is computationally intensive. We use alternating-direction optimization for more efficient tuning, meaning we tune one parameter each time without changing others. We start the tuning process by letting $(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_o) = (1 \text{ mm}, 0.01, 10, 10 \text{ mm}, 1 \text{ mm}, 1 \mu\text{s})$, which is an order-of-magnitude empirical estimation. After multiple iterations, we obtain the best choice $(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_o) = (0.6 \text{ mm}, 0.02, 36, 18 \text{ mm}, 1.8 \text{ mm}, 1.2 \mu\text{s})$ for the numerical phantom with simple blood vessel structures.

To justify the choice, we first tune l_{SP} while keeping $(\alpha, J, l_{DD}, \xi_{DD}, T_o) = (0.02, 36, 18 \text{ mm}, 1.8 \text{ mm}, 1.2 \mu\text{s})$. For aliasing artifacts, we obtain the STDs of the background pixel values for different values of l_{SP} and normalize them by dividing the average STD, as shown in Fig. 10(a1). For image resolution, we compare the values along the line L in Fig. 5(b2) for l_{SP} of 0.3 mm, 0.6 mm, and 1.2 mm, respectively, as shown in Fig. 10(a2). As we see in Fig. 10(a1), for $l_{SP} > 0.6 \text{ mm}$, the STD increases as l_{SP} increases. In fact, a greater value of l_{SP} means fewer source points for spatiotemporal antialiasing, which results in a greater amplitude of aliasing artifacts and less blurring of the image, as shown in Fig. 10(a1) and (a2), respectively. As a balance, we choose $l_{SP} = 0.6 \text{ mm}$. Then we tune α while letting $(l_{SP}, J, l_{DD}, \xi_{DD}, T_o) = (0.6 \text{ mm}, 36, 18 \text{ mm}, 1.8 \text{ mm}, 1.2 \mu\text{s})$. For mitigating aliasing artifacts, $\alpha = 0.02$ is the best choice, as shown in Fig. 10(b1). For $\alpha < 0.02$, a smaller α means fewer source points for spatiotemporal antialiasing, which results in a higher amplitude of aliasing artifacts. For $\alpha > 0.02$, a greater α means more source point candidates including low-amplitude ones. However, in the random-selection step, having more low-amplitude source points means that the dominant source points are less likely to be selected. As a result, the spatiotemporal antialiasing is less effective, leading to a higher amplitude of aliasing artifacts. For maintaining image resolution, $\alpha = 0.02$ still turns out to be the best choice, as shown in Fig. 10(b2). Thus, we choose $\alpha = 0.02$. The tuning of J is simpler. As shown in Fig. 10(c1) and (c2), the increase of J reduces the STD but has minor effects on image resolution. Considering that the computational time has linear dependency on J , we choose $J = 36$ for a balance between antialiasing performance and computational efficiency. Further, we tune l_{DD} while keeping $(l_{SP}, \alpha, J, \xi_{DD}, T_o) = (0.6 \text{ mm}, 0.02, 36, 1.8 \text{ mm}, 1.2 \mu\text{s})$. As we see in Fig. 10(d1), the STD is relatively small for $l_{DD} < 20 \text{ mm}$. For $l_{DD} > 20 \text{ mm}$, the STD increases as l_{DD} increases, which is explained by the fact that the advantage of the location dependency in our proposed method is mitigated as the subdomain size increases. In Fig. 10(d2), we see that $l_{DD} = 18 \text{ mm}$ is the best choice for maintaining image resolution. Combining the observations in Fig. 10(d1) and (d2), we choose $l_{DD} = 18 \text{ mm}$. Next, we tune ξ_{DD} while letting $(l_{SP}, \alpha, J, l_{DD}, T_o) = (0.6 \text{ mm}, 0.02, 36, 18 \text{ mm}, 1.2 \mu\text{s})$. As we see in Fig. 10(e1) and (e2), changing the value of ξ_{DD} has minor effects on both aliasing artifacts and image resolution. We let ξ_{DD} be one-tenth of l_{DD} : 1.8 mm. Finally, the tuning of T_o for the simple numerical phantom is simple. As shown in Fig. 10(f1) and (f2), the increase of T_o reduces the STD but has minor effects on image resolution. We choose T_o to be 1.2 μs . Further increasing

T_o has minor benefits in mitigating aliasing artifacts and may blur other regions of the image.

In summary, for the simple numerical phantom, the tuning of every parameter is a robust process without abrupt changes. The performance of the location-dependent spatiotemporal antialiasing is sensitive to l_{SP} and α but insensitive to ξ_{DD} . The sensitivity is low for $J \geq 36$, $l_{DD} \leq 20 \text{ mm}$, and $T_o \geq 1.2 \mu\text{s}$. These observations serve as guidance for parameter tuning of the complex numerical phantom and *in vivo* experiments. In fact, we use the same $(J, l_{DD}, \xi_{DD}, T_o)$ for these reconstructions and only tune (l_{SP}, α) . For another imaging system, we can update $(J, l_{DD}, \xi_{DD}, T_o)$ accordingly through numerical simulations and only tune (l_{SP}, α) for different datasets.

APPENDIX E

THE ADVANTAGE OF USING THE LOCATION-DEPENDENT PARAMETER $T_b(D_i)$ OVER USING A CONSTANT T_o FOR TEMPORAL FILTERING

We choose to use the location-dependent parameter $T_b(D_i)$ for temporal filtering in this research. Here, based on numerical simulations of the complex numerical phantom, we demonstrate the choice's advantage over using the same parameter T_o for all subdomains. We use parameter values $(l_{SP}, \alpha, J, l_{DD}, \xi_{DD}, T_o) = (3.6 \text{ mm}, 0.08, 36, 18 \text{ mm}, 1.8 \text{ mm}, 1.2 \mu\text{s})$ in these simulations.

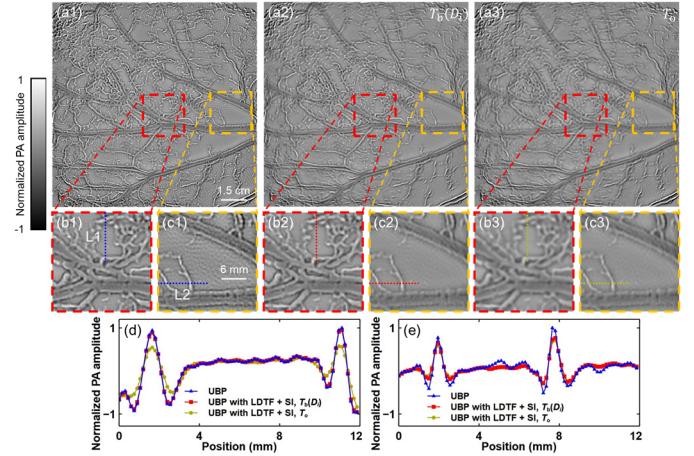


Fig. 11. Comparison of using the location-dependent range $T_b(D_i)$ and using a constant T_o for temporal filtering. (a1) Reconstructed image of the complex numerical phantom through UBP. (a2) and (a3) Reconstructed images through UBP with LDTF and spatial interpolation (SI), using the location-dependent parameter $T_b(D_i)$ and a constant parameter T_o , respectively. Two subdomains in a red-dashed box and a yellow-dashed box, respectively, are picked for comparisons. (b1)-(c1), (b2)-(c2), and (b3)-(c3) Closeup images of the two subdomains for the three methods, respectively. (d) and (e) Comparisons of values along lines L1 and L2, respectively, for the three methods.

The reconstructed image of the complex numerical phantom using the UBP method without spatiotemporal antialiasing is shown in Fig. 11(a1). Using the location-dependent parameter $T_b(D_i)$ for temporal filtering in the proposed spatiotemporal antialiasing, we obtain the image shown in Fig. 11(a2). Using the same parameter T_o across all subdomains for temporal filtering, we obtain the image shown in Fig. 11(a3). We pick two subdomains, marked by a red-dashed box and a yellow-dashed box, respectively, in the three images for comparisons. Closeup images of the subdomains are shown in Fig. 11(b1)-

(c1), (b2)-(c2), and (b3)-(c3), respectively. Comparing Fig. 11(b1)–(b3), we see that the image resolution is maintained in the central region by using the location-dependent parameter $T_b(D_i)$; whereas the image is blurred by using a constant T_o across subdomains. Comparing Fig. 11(c1)–(c3), we notice similar performances of using the location-dependent parameter $T_b(D_i)$ and using a constant T_o in mitigating aliasing artifacts in the peripheral region. Quantitatively, we pick two lines, L1 and L2, in the two subdomains (shown in Fig. 11(b1) and (c1), respectively) and compare the values on them for the three methods in Fig. 11(d) and (e), respectively. As we see in Fig. 11(d), the image resolution in the central region is not affected by using the location-dependent parameter $T_b(D_i)$ but compromised by using a constant T_o . For the peripheral region, we see similar performances of the two choices in Fig. 11(e).

In summary, to achieve similar performances in mitigating aliasing artifacts (most abundant in the peripheral region), using the location-dependent parameter $T_b(D_i)$ is better than using a constant T_o in maintaining image resolution (most obvious in the central region).

ACKNOWLEDGMENT

We thank Li Lin for his helpful discussion in this study.

REFERENCES

- [1] L. Li, L. Zhu, C. Ma, L. Lin, J. Yao, L. Wang, K. Maslov, R. Zhang, W. Chen, J. Shi, and others, “Single-impulse panoramic photoacoustic computed tomography of small-animal whole-body dynamics at high spatiotemporal resolution,” *Nat. Biomed. Eng.*, vol. 1, no. 5, p. 0071, 2017.
- [2] L. Lin, P. Hu, J. Shi, C. M. Appleton, K. Maslov, L. Li, R. Zhang, and L. V. Wang, “Single-breath-hold photoacoustic computed tomography of the breast,” *Nat. Commun.*, vol. 9, no. 1, p. 2352, 2018.
- [3] L. Lin, P. Hu, X. Tong, S. Na, R. Cao, X. Yuan, D. C. Garrett, J. Shi, K. Maslov, and L. V. Wang, “High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical translation,” *Nat. Commun.*, vol. 12, no. 1, pp. 1–10, 2021.
- [4] D. Razansky, A. Buehler, and V. Ntziachristos, “Volumetric real-time multispectral optoacoustic tomography of biomarkers,” *Nat. Protoc.*, vol. 6, no. 8, pp. 1121–1129, 2011.
- [5] R. A. Kruger, C. M. Kuzniak, R. B. Lam, D. R. Reinecke, S. P. Del Rio, and D. Steed, “Dedicated 3D photoacoustic breast imaging,” *Med. Phys.*, vol. 40, no. 11, p. 113301, 2013.
- [6] V. Ntziachristos and D. Razansky, “Molecular imaging by means of multispectral optoacoustic tomography (MSOT),” *Chem. Rev.*, vol. 110, no. 5, pp. 2783–2794, 2010.
- [7] J. G. Laufer, E. Z. Zhang, B. E. Treeby, B. T. Cox, P. C. Beard, P. Johnson, and B. Pedley, “In vivo preclinical photoacoustic imaging of tumor vasculature development and therapy,” *J. Biomed. Opt.*, vol. 17, no. 5, p. 056016, 2012.
- [8] L. V. Wang and H. Wu, *Biomedical optics: principles and imaging*. John Wiley & Sons, 2012.
- [9] M. Xu and L. V. Wang, “Universal back-projection algorithm for photoacoustic computed tomography,” *Phys. Rev. E*, vol. 71, no. 1, p. 016706, 2005.
- [10] B. E. Treeby and B. T. Cox, “k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields,” *J. Biomed. Opt.*, vol. 15, no. 2, p. 021314, 2010.
- [11] A. Rosenthal, D. Razansky, and V. Ntziachristos, “Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography,” *IEEE Trans. Med. Imaging*, vol. 29, no. 6, pp. 1275–1285, 2010.
- [12] K. Wang, R. Su, A. A. Oraevsky, and M. A. Anastasio, “Investigation of iterative image reconstruction in three-dimensional optoacoustic tomography,” *Phys. Med. Biol.*, vol. 57, no. 17, pp. 5399–5423, Sep. 2012.
- [13] K. Mitsuhashi, J. Poudel, T. P. Matthews, A. Garcia-Uribe, L. V. Wang, and M. A. Anastasio, “A forward-adjoint operator pair based on the elastic wave equation for use in transcranial photoacoustic computed tomography,” *SIAM J. Imaging Sci.*, vol. 10, no. 4, pp. 2022–2048, 2017.
- [14] Y. Xu, M. Xu, and L. V. Wang, “Exact frequency-domain reconstruction for thermoacoustic tomography. II. Cylindrical geometry,” *IEEE Trans. Med. Imaging*, vol. 21, no. 7, pp. 829–833, 2002.
- [15] P. Hu, L. Li, L. Lin, and L. V. Wang, “Spatiotemporal Antialiasing in Photoacoustic Computed Tomography,” *IEEE Trans. Med. Imaging*, vol. 39, no. 11, pp. 3535–3547, 2020.
- [16] N. Davoudi, X. L. Deán-Ben, and D. Razansky, “Deep learning optoacoustic tomography with sparse data,” *Nat. Mach. Intell.*, vol. 1, no. 10, pp. 453–460, Oct. 2019.
- [17] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, “Full-Wave Iterative Image Reconstruction in Photoacoustic Tomography With Acoustically Inhomogeneous Media,” *IEEE Trans. Med. Imaging*, vol. 32, no. 6, pp. 1097–1110, Jun. 2013.
- [18] K. Mitsuhashi, J. Poudel, T. P. Matthews, A. Garcia-Uribe, L. V. Wang, and M. A. Anastasio, “A Forward-Adjoint Operator Pair Based on the Elastic Wave Equation for Use in Transcranial Photoacoustic Computed Tomography,” *SIAM J. Imaging Sci.*, vol. 10, no. 4, pp. 2022–2048, Jan. 2017.
- [19] S. Arridge, P. Beard, M. Betcke, B. Cox, N. Huynh, F. Lucka, O. Ogunlade, and E. Zhang, “Accelerated high-resolution photoacoustic tomography via compressed sensing,” *ArXiv Prepr. ArXiv160500133*, 2016.
- [20] M. Pérez-Liva, J. L. Herranz, J. M. Udiá, E. Miller, B. T. Cox, and B. E. Treeby, “Time domain reconstruction of sound speed and attenuation in ultrasound computed tomography using full wave inversion,” *J. Acoust. Soc. Am.*, vol. 141, no. 3, pp. 1595–1604, Mar. 2017.
- [21] A. Chambolle, “An algorithm for total variation minimization and applications,” *J. Math. Imaging Vis.*, vol. 20, no. 1, pp. 89–97, 2004.
- [22] A. Beck and M. Teboulle, “Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems,” *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, 2009.
- [23] N. Davoudi, X. L. Deán-Ben, and D. Razansky, “Deep learning optoacoustic tomography with sparse data,” *Nat. Mach. Intell.*, vol. 1, no. 10, pp. 453–460, Oct. 2019.
- [24] N. Davoudi, N. Davoudi, B. Lafci, B. Lafci, A. Özbeck, A. Özbeck, X. L. Deán-Ben, X. L. Deán-Ben, D. Razansky, and D. Razansky, “Deep learning of image- and time-domain data enhances the visibility of structures in optoacoustic tomography,” *Opt. Lett.*, vol. 46, no. 13, pp. 3029–3032, Jul. 2021.
- [25] V. A. Kelkar and M. Anastasio, “Prior Image-Constrained Reconstruction using Style-Based Generative Models,” in *Proceedings of the 38th International Conference on Machine Learning*, 2021, pp. 5367–5377.
- [26] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, “Image reconstruction by domain-transform manifold learning,” *Nature*, vol. 555, no. 7697, pp. 487–492, Mar. 2018.
- [27] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, “Deep learning for undersampled MRI reconstruction,” *Phys. Med. Biol.*, vol. 63, no. 13, p. 135007, Jun. 2018.
- [28] T. Würfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, and A. K. Maier, “Deep Learning Computed Tomography: Learning Projection-Domain Weights From Image Domain in Limited Angle Problems,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1454–1463, Jun. 2018.
- [29] N. Davoudi, X. L. Deán-Ben, and D. Razansky, “Deep learning optoacoustic tomography with sparse data,” *Nat. Mach. Intell.*, vol. 1, no. 10, pp. 453–460, 2019.
- [30] C. Cai, X. Wang, K. Si, J. Qian, J. Luo, and C. Ma, “Streak artifact suppression in photoacoustic computed tomography using adaptive back projection,” *Biomed. Opt. Express*, vol. 10, no. 9, pp. 4803–4814, 2019.
- [31] B. E. Treeby and B. T. Cox, “k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields,” *J. Biomed. Opt.*, vol. 15, no. 2, p. 021314, 2010.