**Abstract**—We analyze the principles underlying *minimum variance distortionless response* (MVDR) beamforming in order to integrate it into a pixel-based algorithm. There is a challenge posed by the low echo signal-to-noise ratio (eSNR) when calculating beamformer contributions at pixels far away from the beam centreline. Together with the well-known scarcity of samples for covariance matrix estimation, this reduces the beamformer performance and degrades the image quality. To address this challenge, we implement the MVDR algorithm in two different ways. First, we develop the *conventional minimum variance pixel-based* (MVPB) beamformer that performs the MVDR after the pixel-based superposition step. This involves a combination of methods in the literature, extended over multiple transmits to increase the eSNR. Then we propose the *coherent MVPB* beamformer, where the MVDR is applied to data within individual transmits. Based on pressure field analysis, we develop new algorithms to improve the data alignment and matrix estimation, and hence overcome the low-eSNR issue. The methods are demonstrated on data acquired with an ultrasound open platform. The results show the coherent MVPB beamformer substantially outperforms the conventional MVPB in a series of experiments, including phantom and *in vivo* studies. Compared to the unified pixel-based beamformer, the newest delay-and-sum algorithm in [1], the coherent MVPB performs well on regions that conform to the diffuse scattering assumptions on which the minimum variance principles are based. It produces less good results for parts of the image that are dominated by specular reflections.

**Index Terms** - Ultrasound imaging, adaptive beamforming, minimum variance, pixel-based, spatial coherence, time delay, image quality.

I. INTRODUCTION

Beamforming is a vital component in the process of generating high quality ultrasound images. It determines the pattern of delays and weights applied to each channel to control the acoustic beam shape on transmit, and the similar delays and weights that create focal zones during receive. The goal of a beamformer is to create a uniformly narrow pulse-echo beam over the insonified region. The beamformer performance, therefore, is usually evaluated through figures of merit that measure the image quality, including echo signal-to-noise ratio (eSNR), contrast and spatial resolution.

In an earlier study [1], we developed pixel-based (PB) beamformers that improve image quality by removing the need for interpolation. The methods use dynamic delay-and-sum (DAS) algorithms to generate data at individual pixels of the image field. Performance is enhanced by increasing the amount of information retained from individual channel echo sequences and used to form the reconstructed image [2], [3]. The approach is comparable to the *synthetic transmit beam* (STB) [4] and *synthetic aperture* (SA) [5] techniques, where all received time-series waveforms are stored and reused to synthesize data at several scanlines or in different directions. The PB beamformers, however, do not attempt to cancel noise or off-target echoes that compromise the image resolution. Finding methods for further improvement is the subject of this work.

Array processing techniques can be used to improve image quality beyond that possible with DAS [6]. Among them, *minimum variance distortionless response* (MVDR) beamforming has been applied to ultrasound imaging by several research groups [7]–[11]. The MVDR approach improves lateral resolution by preserving the echo signal reflected from the main point-target while minimizing acquisition noise and interference from clutter. The beamforming weights can be calculated from the inverse of the data covariance matrix [13]. This allows the MVDR to be implemented directly on measurements at the array output. It is classified as *adaptive beamforming*, implying that the beamformer is optimized based on its incoming signals.

There is, however, a practical problem that must be solved in order to implement an MVDR beamformer. It involves estimating the data covariance matrix when there is only one sample of the data vector, or *snapshot*, available. Various approximations have been used to overcome this issue, including *spatial smoothing* and *diagonal loading* [10]. The first method reduces the receive aperture size so that multiple snapshots can be formed over the linear array, while the second regularizes the main diagonal of the estimated matrix to avoid ill-conditioning. Recently, spatial smoothing has been combined with *forward-backward averaging* to improve the matrix estimation [14], [15]. Estimation errors are, however, still unavoidable, potentially limiting performance.

In this study, we explore the possibility of combining the MVDR with PB beamformers to improve image qual-
ity. We name the combination minimum-variance pixel-based (MVPB) beamforming. Some similar approaches have been proposed in literature. In [16], Holfort et al. conducted a simulation study that integrated the MVDR with SA imaging. The beam was broad, generated with a single element or planewave transmission. In [17], Rabinovich et al. suggested the use of focused transmit beams for better lateral resolution. They combined the STB with the MVDR for cardiac imaging. The data, however, was accumulated only between two consecutive transmits. In this paper, we are using a highly focused beam and extend the algorithm among multiple transmits to find the most effective way of exploiting the information available in the echo data.

A key issue that arises when combining the MVDR with pixel-based superposition is which to apply first. If we apply the MVDR before data superposition there is a problem with the low eSNR. For an individual transmit-receive event, there is very little insonification far from the centreline of the beam. The data collected at such points, therefore, is dominated by noise and clutter. Low eSNR has been shown to have a negative effect on the MVDR performance in several studies [18], [19].

In the following section, we first develop a MVPB beamformer where the MVDR is performed after data superposition. It is a hybrid algorithm that results from a combination of existing methods from the literature [17]. We then propose an application of the MVDR to pre-summed data. To overcome the issue of low eSNR in each transmit, we exploit our pressure field analysis in [1] to form new algorithms for improving data alignment and matrix estimation. The MVPB beamformers are evaluated and compared in Section III through a series of experiments, including simulation, phantom, and in vivo studies. The results are explained and discussed in Section IV. The work is summarized with some concluding remarks in Section V.

II. METHODS

We consider an ultrasonic system where an $N_r$-element active aperture used for both transmit and receive during a pulse-echo sequence. Data at each image point P is generated from $N_t$ transmits that have incident beams passing through P.

A. Conventional MVPB Beamformer

In the first combination, the MVDR is applied to compounded data from multiple transmits that increase the eSNR. The receive aperture size is enlarged to $N = N_t + N_r - 1$. Let $\mathbf{x}(k)$ be a data vector collected from the $N$-element array at an instance time $k$. The vector can be modeled as

$$ \mathbf{x}(k) = \mathbf{s}_p(k) \mathbf{a}(\tau_p) + \mathbf{i}(k) + \mathbf{n}(k), $$

where $\mathbf{s}_p(k)$ is the signal waveform at P, $\mathbf{i}(k)$ and $\mathbf{n}(k)$ are interference from clutter and acquisition noise respectively, and $\mathbf{a}(\tau_p)$ is the complex time-delay vector.

Using the MVDR criterion for narrow-band signals, the beamforming weight vector $\mathbf{w}$ is calculated by [6]

$$ \mathbf{w} = \frac{\mathbf{R}^{-1}_{\tau_p} \mathbf{a}(\tau_p)}{\mathbf{a}^H(\tau_p) \mathbf{R}^{-1}_{\tau_p} \mathbf{a}(\tau_p)}, $$

(2)

where $\mathbf{R}_{\tau_p}$ is the $N \times N$ interference-plus-noise covariance matrix, and $(\cdot)^H$ stands for the Hermitian transpose. The beamformer output is obtained as

$$ y(k) = \mathbf{w}^H \mathbf{x}(k) = \sum_{m=1}^{N} w^*_m x_m(k). $$

(3)

Equations (2) and (3) reveal the strategy behind MVDR beamforming. Because echo signals have zero-mean, $\mathbf{R}_{\tau_p}$ is also the spatial correlation matrix of noise plus interference among the array elements, i.e., $\mathbf{R}_{\tau_p+n} = E[(\mathbf{i}(k) + \mathbf{n}(k))(\mathbf{i}(k) + \mathbf{n}(k))^H]$. The MVDR algorithm acts as a spatial filter that decorrelates noise and clutter in $\mathbf{x}(k)$ before delaying each of its components and adding them together. The signal $\mathbf{s}_p(k)$ is preserved by matching the decorrelation output to $\mathbf{a}(\tau_p)$. Thus $\mathbf{a}(\tau_p)$ defines the main signal and contains all the spatial characteristics of the filter [6].

Since $\mathbf{s}_p(k)$ remains the same through the MVDR filter, the decorrelation is still equivalent if it applies to all components of $\mathbf{x}(k)$. This can be proved by using the matrix inversion lemma, from which $\mathbf{w}$ is given by [13]

$$ \mathbf{w} = \frac{\mathbf{R}^{-1}_{\tau_p} \mathbf{a}(\tau_p)}{\mathbf{a}^H(\tau_p) \mathbf{R}^{-1}_{\mathbf{a}} \mathbf{a}(\tau_p)}, $$

(4)

where $\mathbf{R}_a \triangleq \sigma_a^2 \mathbf{a}(\tau_p) \mathbf{a}^H(\tau_p) + \mathbf{R}_{\tau_p+n}$ is the data covariance matrix, and $\sigma_a^2$ is the scattering strength at P. Equation (4) allows the MVDR to be implemented directly on $\mathbf{x}(k)$. However, beamformer performance depends on the accuracy with which it is possible to determine $\mathbf{a}(\tau_p)$ and estimate $\mathbf{R}_a$.

For each individual transmit sequence, we adopt the SA imaging model and assume that the sound waves propagating towards P are spherical [5], [20]. Based on this assumption, we calculate the time delay vector $\mathbf{a}(\tau_p)$ and refer to it as the conventional time delay.

The covariance matrix $\mathbf{R}_a$ is estimated using several approximations, including spatial smoothing and diagonal loading [10]. In spatial smoothing, the vector dimension is reduced to $L < N$ so that several snapshots of the data vector can be formed over the $N$-element array. This is based on an assumption that the backscattered field is stationary so that the spatial correlation depends only on array element relative offsets rather than their absolute positions. For the modified covariance matrix to be nonsingular, the number of snapshots should be greater than or equal to the matrix size [21], or $L \leq N/2$.

The matrix estimation can also be improved with forward-backward averaging [14], [15]. This technique exploits the structure of the data covariance matrix based on an assumption of a spatially stationary backscattered field. Under this assumption, $\mathbf{R}_a$ is persymmetric, i.e., $\mathbf{R}_a = \mathbf{J} \mathbf{R}_a^T \mathbf{J}$ where $\mathbf{J}$ is the exchange matrix (having ones on the counterdiagonal and zeros everywhere else). Specifically, we calculate $\mathbf{R}_a^{FW}$ and...
This using insights relating to the time delay calculation and backwards procedure, for which spatially different beam profile. Point P in region (IV) is treated the same as in region (II) and hence is not illustrated in the figure (from [1]).

\[
\hat{R}_x^{BW} \text{ by applying spatial smoothing to } x(k) \text{ in the forwards and backwards procedure, for which } \hat{R}_x^{BW} = \text{J} \left( \hat{R}_x^{FW} \right)^* \text{J}.
\]

By setting \( R_x = 0.5 \left( \hat{R}_x^{FW} + \hat{R}_x^{BW} \right) \), we have \( \hat{R}_x = \text{J} \hat{R}_x^* \text{J} \).

Using all of these approximations, we obtain an estimation of \( R_x \):

\[
\hat{R}_x = \frac{1}{2} \left( \hat{R}_x^{FW} + \hat{R}_x^{BW} + \gamma I \right),
\]

where \( \gamma \) is the diagonal loading parameter, and \( I \) is the \( N \times N \) identity matrix.

Combining the matrix estimation with the conventional time delay, we form a beamformer named the conventional MVPB.

**B. Coherent Time Delay Calculation**

In the next section, we develop a combination that performs the MVDR calculation before superposition. The receive aperture size is reduced to \( N_r \) and the MVDR outputs are subsequently superposed across \( N_t \) transmits. This strategy is only viable if we can improve the low eSNR of the data for points located far from the beam centreline. We seek to achieve this using insights relating to the time delay calculation and the data covariance matrix estimation.

Previously, we analyzed the transmit pressure field of a focused beam to find a better description for the transmit aperture. This provides a generalization of the virtual source element technique and is valid over the entire insonified region [22]. Based on those findings, we developed a unified PB beamformer and shown it improves the coherence among the selected signals. Thus, the data selection algorithm from this model can be benefit MVDR performance. We briefly summarize the main results of the analysis here; the details are provided in [1].

By taking the delays at each transmit element into account, we model the active aperture as an arc centered at the focus F with radius \( R_0 \). We also denote \( d \) as the focal depth, and \( R_1 \) and \( R_2 \) as distances from P to the edges of the arc. From F, we divide the imaging plane into four regions, denoted from (I) to (IV) clockwise, using the limited angle \( \alpha \) from the virtual source approach [22].

Figures (a)-(c) plot three different scenarios for P in regions from (I) to (III). A point in region (IV) will have the same treatment as in region (II), thus, it is not shown in the figures.

Let us denote \( \tau_{p,1}^{tr} \) and \( \tau_{p,2}^{tr} \) as the arrival times of the two pulses in the transmit wave-shape. For P in region (I), we have

\[
\tau_{p,1}^{tr} = \frac{d-a}{c} \text{ and } \tau_{p,2}^{tr} = \frac{R_{\text{max}}}{c} - \tau_0^{tr},
\]

where \( R_{\text{max}} = \max (R_1, R_2) \), and \( \tau_0^{tr} = (R_0 - d)/c \), which is the time lag between activation of the first (outermost) and the last (center) transmit elements. In this region, the first pulse (with \( \tau_{p,1}^{tr} \)) has much higher magnitude.

For P in region (III), we have

\[
\tau_{p,1}^{tr} = \frac{R_{\text{min}}}{c} - \tau_0^{tr} \text{ and } \tau_{p,2}^{tr} = \frac{d+a}{c},
\]

where \( R_{\text{min}} = \min (R_1, R_2) \). In this region, the wave-shape is dominated by the second pulse (with \( \tau_{p,2}^{tr} \)).

In region (II), we have

\[
\tau_{p,1}^{tr} = \frac{R_1}{c} - \tau_0^{tr} \text{ and } \tau_{p,2}^{tr} = \frac{R_2}{c} - \tau_0^{tr},
\]

where \( R_1 \) and \( R_2 \) are now the minimal and maximal distances from P to the aperture. In this region, the pulses have similar magnitudes to each other. The result is the same as for P in region (IV) except that now \( R_1 \) and \( R_2 \) are, respectively, the maximal and minimal distances from P to the arc. At the focus F, the maximal and minimal distances are the same and equal to \( R_1 \) and \( R_2 \), hence both pulses are merged into a highly focused pulse.

The receive time delay may be calculated straightforwardly based on the distance between P and the corresponding receiving element. When this is combined with the transmit travel times above, we form new two-way time delays that extract data associated with the pulses on the transmit wave-shape. We name these coherent time delays.
Unlike the conventional time delay formulae, the coherent delays depend on the relative positions of P and the active transmit aperture. This is, however, consistent with the idea of applying MVDR to data from individual transmits. There are also differences between the coherent delays and the time delays calculated in the unified PB beamformer (see Appendix). In the unified PB, we determine the time delay so that it refers to the highest pulse in regions (I) and (III).

Then we form a smooth transition between these two pulses when generating data outside these regions. However, in the present work, we use the fact that the coherent time delays are always associated with the high-energy reflected pulses over the entire insonified region. This enables us to extract two data vectors at each imaging point except at the focus. We show how to combine these data vectors with the MVDR beamforming process in the next section.

C. Coherent MVPB Beamformer

We develop a new MVPB beamformer that uses the coherent time delay calculation described above. Let \(a(\tau_{p,1})\) and \(a(\tau_{p,2})\) be the time delay vectors associated with the two pulses. Applying the MVDR using both time delay vectors, we generate the focusing signal at each transmit as

\[
y(k) = b_1 w_1^H x(k) + b_2 w_2^H x(k),
\]

where \(w_i\) is the MVDR weights calculated with the time delay vector \(a(\tau_{p,i})\) \((i = 1, 2)\). The coefficients \(b_1\) and \(b_2\) are selected to generate an artefact-free image. The selection is performed as follows.

In regions (I) and (III) (see Figs. 1) where one pulse dominates the other in the transmit wave-shape, we select the data with the higher magnitude pulses only. Thus, \((b_1, b_2) = (1, 0)\) for \(P\) in region (I) and \((b_1, b_2) = (0, -1)\) for \(P\) in region (III). The minus sign is because of the opposite phases between the two pulses. In region (II) (and similarly in region (IV)), we choose \((b_1, b_2)\) for \(P\) using

\[
b_1 = \frac{|r_a - r_p|}{|r_b - r_a|} \quad \text{and} \quad b_2 = -\frac{|r_a - r_p|}{|r_a - r_b|},
\]

where \(r_a, r_b, \) and \(r_p\) are the position vectors of A, B, and P (see Fig. 1(c)). The origin of these position vectors is assumed to be at the center of the active aperture. Notice the minus sign for \(b_2\), again, to account for the phase opposition between the two pulses. This combination is shown to generate an artefact-free B-mode image.

In our pressure field analysis, the two pulses in the transmit wave-shape are both derived as scaled and delayed from the excitation pulse \([1]\). Each imaging point \(P\) (except \(F\)) can then be considered as illuminated by two pulses with arrival times \(\tau_{p,1}^r\) and \(\tau_{p,2}^r\). Data vectors extracted at these times, therefore, can be combined to estimate the covariance matrix. With two data vectors, the matrix is estimated by

\[
\hat{R}_x = \frac{1}{2} \left( \hat{R}_{x,1}^{FW} + \hat{R}_{x,1}^{BW} + \hat{R}_{x,2}^{FW} + \hat{R}_{x,2}^{BW} + \gamma I \right),
\]

where \(\hat{R}_{x,1}^{FW}\) and \(\hat{R}_{x,1}^{BW}\) are calculated from data associated with \(\tau_{p,i}\) \((i = 1, 2)\).

This new approach, using two data vectors, helps to increase the size of the estimation matrix. We name the method joint estimation to differentiate it from the calculation in (5). Together with the coherent time delay, we form a beamformer named coherent MVPB.

D. Multibeam MVPB beamformer

In [23], Vignon and Burcher developed a pixel-wise minimum variance beamformer that performed the MVDR before data superposition. The time delay was calculated by using the virtual source element approach, that is, the transmit time delay to each imaging point \(P\) was given by \(\tau_{p,1}^r\) in (6) for \(P\) shallower than \(F\) and given by \(\tau_{p,2}^r\) in (7) for \(P\) deeper than \(F\). The data covariance matrix was calculated by taking the average of the covariance matrices derived from individual transmit events. In this study, we implement this method and compare it to our proposed coherent MVPB beamformer.

First, we upgrade the time calculation to our coherent time delays which are valid outside the limited angle. Within each transmit event, two data vectors are extracted for pixel \(P\), associated with two transmit pulses. For the data vector associated with pulse \(\tau_{p,i}\) \((i = 1, 2)\), the MVDR weight vectors \(w_i\) are calculated with the covariance matrix given by

\[
R_{x,i} = \sum_{k=1}^{N_i} R_{x,i}^{(k)},
\]

where \(R_{x,i}^{(k)}\) is the covariance matrix of the data vector associated with pulse \(\tau_{p,i}\) in transmit event \(k\). Each \(R_{x,i}^{(k)}\) is estimated using Eq. (5). We name this method the multibeam MVPB beamformer.

To generate a smooth and artefact-free B-mode image, we also use the coefficients \((b_1, b_2)\) of the coherent MVPB beamformer to combine the outputs of the MVDR beamformers calculated with two data vectors. Thus, the major difference between the coherent MVPB and multibeam MVPB is in their covariance matrix calculations.

E. Evaluation Metrics

We evaluate each beamformer developed in our study based on the quality of the images it generates. This quality can be measured with three metrics: spatial resolution, eSNR, and contrast.

The spatial resolution can be measured through the response on the quality of the images it generates. This quality can be measured with three metrics: spatial resolution, eSNR, and contrast. The spatial resolution can be measured through the response of each beamformer to individual scatterers, or through the speckle size of the beamformed B-mode calculated from the spatial auto-covariance function of the RF data [24, 25]. Because minimum variance beamforming affects the lateral resolution, we are only interested in the lateral width of the correlation length \(S_{cx}\). In general, the smaller \(S_{cx}\) the better the beamformer performance. The eSNR is calculated using the average/difference over multiple beamformed RF frames of the same scan. It is measured from a kernel that contains no specular reflector to avoid interferences from side-lobes in the noise power.
We evaluate the contrast through a contrast ratio (CR) between a lesion and the background given by [15]

$$CR = \frac{I_{out} - I_{in}}{\sqrt{I_{out}^2 + I_{in}^2}},$$

where $I_{in}$ and $I_{out}$ are the mean intensities (in decibels) measured inside and outside the lesion, respectively. The term CR has a value of 1 for perfect contrast, and a value of 0 for no contrast between the lesion and background. The background kernel is selected to have the same area as the lesion. To minimise the effect of variations in the attenuation and diffraction of the ultrasound, we choose the kernel as a circular ring enclosing the lesion.

III. RESULTS

In this study, our imaging results are generated using the ULA-OP ultrasound system developed at the MSD Laboratory, University of Florence, Italy [26]. A linear array probe (LA523, Esaote spa) has 192 elements each of dimensions $0.215 \times 6.0 \text{ mm}^2$ separated by a $0.03 \text{ mm}$ kerf. There are 64 elements active in each transmit and receive cycle. In the lateral direction, the transmit focus is always on the centreline of the active aperture with a focal length of 20 mm. In elevation, the aperture is weakly-focused using an acoustic lens also with a focal length of 20 mm. A three-cycle excitation voltage is applied to generate an ultrasound pulse with center frequency at 6.0 MHz and a 40% pulse-echo bandwidth. The echo data from all the receiving elements is stored in an array and can be individually accessed. The signals are sampled at 50 Msamples/s, resulting in an axial sampling interval of 0.0154 mm. The ULA-OP stores 2048 samples of each received waveform which covers data in a range of 31 mm. In the lateral direction the beam is stepped by an element pitch to acquire data with lateral spacing 0.245 mm. Envelope images are generated using a pixel-based approach with a pixel size of 0.0154 mm $\times$ 0.049 mm.

A. MVPB Implementations

We compare the MVPB beamformer with the unified PB beamformer, the newest DAS algorithm we developed in [1]. The PB beamformed images are generated with $N_t = 32$ transmits. In simulation and phantom study, the dynamic focused image is presented for reference.

The MVDR beamformer is applied with the reduced dimension $L$ less than or equal to the number of snapshots available [10], [15]. The diagonal loading parameter $\gamma$ varies from $1\%\Delta$ to $5\%\Delta$, where $\Delta = \text{Tr}(R)/L$ and $\text{Tr}(R)$ is the trace of the estimated covariance matrix. For the conventional MVPB algorithm where the receive aperture size $N_r = 95$, we set $L = 48$ and $\gamma = 5\%\Delta$. Because the coherent MVPB has the receive aperture size $N_r = 64$ but the covariance matrix is estimated from two data vectors, we set $L = 43$ and $\gamma = 1\%\Delta$. The multibeam MVPB is calculated with the covariance matrix rank less than or equal to $\min(N_t, N_r)$ [23]. Therefore, we set $L = 32$ and $\gamma = 1\%\Delta$ for this beamformer.

B. Simulated images

We first apply the beamformers to data generated by Field II [27], [28]. The data is simulated based on the parameters of the ULA-OP system. The numerical phantom has three pairs of backscattering points in the imaging field, distributed at depths of 20 mm, 22 mm, and 24 mm. Their lateral distances are 1 mm, 2 mm, and 3 mm, respectively. There is also a circular lesion, 5 mm in diameter, centered at $(z, x) = (22, 3)$ mm. The scatterer strength inside the lesion is 10 times less than those in the background.

The images shown in Figs. 2(a)-(b) are generated with dynamic focusing and unified PB beamforming. Those generated with the MVPB beamformers, including the conventional, coherent, and multibeam, are shown in Figs. 2(c)-(e), respectively. Compared to dynamic focusing, the conventional MVPB image shows a clearer speckle pattern. However, it has the
Fig. 3. The lateral response from each beamformer to scatterers at various depths (a) 20 mm, (b) 22 mm, and (c) 24 mm. The beamformers are: unified PB (Unified PB), conventional MVPB (Con. MVPB), coherent MVPB (Coh. MVPB), and multibeam MVPB (MB-MVPB). The legend in Fig. 4 applies to all the plots in this figure.

TABLE I

<table>
<thead>
<tr>
<th>Beamformer</th>
<th>Average FWHM at depth</th>
<th>CR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 mm</td>
<td>22 mm</td>
</tr>
<tr>
<td>Unified PB</td>
<td>0.341 mm</td>
<td>0.318 mm</td>
</tr>
<tr>
<td>Conventional MVPB</td>
<td>0.519 mm</td>
<td>0.576 mm</td>
</tr>
<tr>
<td>Coherent MVPB</td>
<td>0.161 mm</td>
<td>0.168 mm</td>
</tr>
<tr>
<td>Multibeam MVPB</td>
<td>0.256 mm</td>
<td>0.225 mm</td>
</tr>
</tbody>
</table>

Table showing the average full width at half maximum (FWHM) and contrast ratio (CR) measured on the simulated images.

Fig. 4. The lateral variation of beamformer response inside the lesion plotted at a depth of 22 mm.

poorest quality among the pixel-based images. We conjecture that the poor resolution is caused by the inconsistency of the conventional time delay calculation when used over a large number of transmits. Thus, we generate another image from this beamformer with only 8 transmits in Fig. 2(f) (Conventional MVPB*). The covariance matrix size is reduced to $L = 36$ and the diagonal loading is set to $\gamma = 1\%\Delta$. The image is now less smooth, but the point-targets show smaller lateral width and the lesion contrast is also higher.

Good focussing is only achieved when the adaptive beamformer is applied to data at individual transmits using our new algorithms. The coherent MVPB image shows improvements in the resolution of the point-targets over the unified PB at all depths. The lesion also shows a higher contrast. Meanwhile, the multibeam MVPB image shows a higher CR but less improvement in spatial resolution. To enable detailed comparison, we plot the lateral beam profiles of each PB method in Fig. 3. This shows the responses to the six point-targets at the three different depths. At each depth, we measure the average beamwidth at $-6$ dB, and summarize the results in Table I. The table also contains the lesion CRs. Both coherent and multibeam MVPB beamformers show improvements over the unified PB in spatial resolution and contrast. In the table, the figures-of-merit calculated for the conventional MVPB image are the lowest.

We also compare the lesion shape reconstructed with different beamformers by plotting, in Fig. 4, the cross sections going through the lesion at a depth of 22 mm. The lesion edges are located at 0.5 mm and 5.5 mm, as indicated by the arrows in the plot. The plot shows that the coherent MVPB has the sharpest edges while the conventional MVPB profile has the most blurring.

C. Phantom study

We demonstrate the beamformers experimentally on data acquired with the ULA-OP system [26], used to scan a tissue-mimicking phantom. The phantom has speeds of sound ranging from 1538 m/s to 1551 m/s and a background attenuation coefficient slope of $0.55$ dB cm$^{-1}$ MHz$^{-1}$. The scanned lesion is a hypoechoic circular target, 5 mm in diameter, positioned at a depth of 23 mm. The attenuation coefficient slope is $0.2$ dB cm$^{-1}$ MHz$^{-1}$ inside the lesion. Imaging results are shown in Figs. 5(a)–(e), generated respectively with dynamic focusing, unified PB, conventional, coherent, and multibeam MVPB. All beamformers are applied to the same data.

Among the pixel-based algorithms, the conventional MVPB has the lowest spatial resolution and lesion contrast. Its speckle size is more than twice that measured in the coherent MVPB image. Compared to the simulation, the improvements offered by the coherent and multibeam MVPB over the unified PB are reduced. They have equivalent CRs while the speckle pattern of the multibeam MVPB is on a par with that of the unified PB. Enhancements in the spatial resolution are observed in the coherent MVPB image which also has the smallest speckle size. The metrics, $S_{cx}$ and CR, are measured and shown in Table II along with the eSNR. The eSNR values are calculated over twelve beamformed RF frames from the same scanning sequence. They are equivalent for all the PB beamformers.
Fig. 5. Images of an idealised lesion generated by different beamformers: (a) Dynamic focusing, (b) Unified PB, (c) conventional MVPB, (d) coherent MVPB, (e) multibeam MVPB, and (f) conventional MVPB with only 8 transmits (conventional MVPB*). All images are log-compressed and displayed with a dynamic range of 40 dB.

<table>
<thead>
<tr>
<th>Beamformers</th>
<th>(S_{xx})</th>
<th>CR</th>
<th>eSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unified PB</td>
<td>0.44±0.02 mm</td>
<td>0.52</td>
<td>35.4±2.1 dB</td>
</tr>
<tr>
<td>Conventional MVPB</td>
<td>0.71±0.08 mm</td>
<td>0.39</td>
<td>32.3±1.1 dB</td>
</tr>
<tr>
<td>Coherent MVPB</td>
<td>0.33±0.03 mm</td>
<td>0.52</td>
<td>33.3±1.3 dB</td>
</tr>
<tr>
<td>Multibeam MVPB</td>
<td>0.43±0.01 mm</td>
<td>0.53</td>
<td>35.5±2.1 dB</td>
</tr>
</tbody>
</table>

*Azimuthal resolution \(\lambda f = 0.33\) mm.

Similar to the simulation, we generate the image with the conventional MVPB* and show the result in Fig. 5(f). Compared to the conventional MVPB image, the conventional MVPB* (with \(N_t = 8\)) has a slightly smaller speckle sizes although this is hard to detect visually. The metric \(S_{xx}\) measured on this image is 0.65±0.06 mm. However, there are sharp transitions between pulse-echo sequences that create some block-artefacts on the generated image. These can be reduced by using linear interpolation to combine data between adjacent transmits [17]. This strategy, however, is not a focus of our study.

### D. In vivo study

Minimum variance beamforming is known to be sensitive to phase aberration. The human body has a variety of sound-speed variations and complicated structures. In this limited in vivo study, we demonstrate our MVPB algorithms on data acquired from a carotid scan. In clinical practice ultrasound is used to detect plaque developed inside the common carotid arteries. Plaque can limit the blood supply to the brain and increase the risk of stroke. We use the ULA-OP system to scan the right side in the neck of a healthy 53 year-old volunteer. This data was obtained with appropriate ethical clearance and informed consent.

The images in the study are shown in Figs. 6(a)–(d), respectively for the unified PB, conventional MVPB, coherent MVPB, and multibeam MVPB. They are transverse views of two blood vessels in the neck, including the internal jugular vein (IJV) and the common carotid artery (CCA). The CCA carries the oxygen-rich blood cells up to the brain and the IJV drains the deoxygenated cells down to the heart. The thyroid gland (TG) is shown just around the CCA. All images are generated with the same data. We compare the details resolution of each image by showing magnified views of the same region on the TG, enclosed by the white rectangles. The magnified views are shown on the right column of the figure.

Among these advanced beamforming methods, the conventional MVPB shows poorest performance. Besides the poor resolution, the image also has artefacts in the near-field region. These are similar to the artefacts shown in the DAS beamformed image generated with the same conventional time delay in our earlier study (see Fig. 4(b) in [1]). They are linked to the effects of the quadratic phase factor. Details of this factor can be found in [29].

Similar to the phantom study, there are small improvements of the multibeam MVPB over the unified PB. They have almost the same speckle patterns, as shown in Figs. 6(e) and (h). The improvement offered by the MVDR filter can only be observed in the coherent MVPB image. In its magnified
Fig. 6. Ultrasound images of the carotid artery (transverse view) generated with different beamformers, including (a) unified PB beamforming, (b) conventional MVPB, (c) coherent MVPB, and (d) multibeam MVPB. Images (e)–(h) are the magnified views of the regions enclosed by the white rectangles on images (a)–(d), respectively. All images are log-compressed and displayed with a dynamic range of 45 dB.

view, Fig. 6(g), the coherent MVPB shows much finer speckles compared to those from the unified PB. The two regions of IJV and CCA on the image in Fig. 6(c) are slightly wider, indicating the effects from the MVDR decorrelation. Some other features, such as the boundaries of the CCA (indicated by the white arrows on the dynamic focused image), however appear less smooth and less hyperechoic in the B-scan from the coherent MVPB. This reduction in contrast for features of this sort has also been observed in other studies with the standard MVDR beamformer (see Figs. 9 and 12, in [15]).

IV. DISCUSSION

The image results show that the coherent and multibeam MVPB beamformers offer significant improvements over the conventional MVPB. They clearly suggest that the MVDR should be applied to data before superposition. The improvements can be explained as follows. The MVDR enhances image quality by decorrelating the spatial coherence among the main signals and off-targets echoes. That coherence measures the correlation among the phases of the sound waves received on transducer elements, and it depends on the characteristics of the source [32]. Unlike other modalities, such as astronomical imaging where the sources are self-radiating [33], scattering sources in ultrasound imaging cause reflections of the incident beam [34]. Therefore, the coherence among backscattered signals depends on the transmit beam profile. In data acquisition, when the beam is shifted laterally, its profile is not the same with respect to a specific scatterer. Thus, the MVDR should be performed within each transmit for better decorrelation.

The improvements of the multibeam MVPB over the conventional MVPB mainly come from the coherent time delay calculation. The image quality achieved with the multibeam MVPB, however, is similar to that from the unified PB in both phantom and in vivo studies. This indicates the poor perfor-
formance of the MVDR when calculating the covariance matrix through a spatial compounding among multiple transmits. For each transmit event, the covariance matrix captures information relating to the relative positions of interfering scatterers with respect to the main target point \( P \). This interference is different from one transmit to another. By taking the average of these matrices, especially over a large number of transmits, we suppress statistical information and make the covariance matrix closer to the identity. This process reduces the spatial resolution of the MVDR approach.

The resolution is significantly improved with the coherent MVPB, when the joint estimation technique is used to extend the covariance matrix size. This also offers a high contrast ratio while keeping the eSNR at the same level. The joint estimation is equivalent to calculating the covariance matrix by using one data vector, then regularizing the result with signals obtained from the other pulse. In the worse case, when the signals from the second pulse are dominated totally by noise and clutter, the regularization can be considered as adding diagonal loading from the Gaussian noise and interference. The new calculation of the covariance matrix is most beneficial to the generation of MVDR beamformed data in regions (II) and (IV) of each transmit, when both pulses are small and have similar magnitudes (see Fig. 2(a), [1]). As a result, the coherent MVPB offers greater benefits near the focal depth where regions (I) and (III) become narrower, and the highly focused signal leads to data being in regions (II) and (IV).

Note that the two coherent time delays refer to the same data vector at the focus \( F \) where the two pulses on the wave-shape are merged together. At this point, the joint estimation does not follow the rule that the number of snapshots available should be greater than the matrix size. This may cause an unstable MVDR filter at the focus. This has not been an issue in our experiments, but if necessary the diagonal loading could be increased to further protect against ill-conditioning.

The MVDR beamformer, when combined with PB beamforming, does not increase the diagnostic information in the data. It enhances the image quality by removing undesired information, such as noise and clutter generated by sidelobes. The filter, however, also may suppress useful signals generated by the mainlobe when scanning targets that are larger than the beamwidth or regions that give rise to specular scattering or coherent scattering [30], [31]. The efficacy of MVPB beamforming depends on specific features of the scattering objects.

In a Field II simulation, all the data is generated from incoherent scattering. The results show the increased performance of the coherent MVPB over the unified PB beamformer. In the phantom study, there is less improvement in the CR, probably caused by a reduction in the specular scattering from the lesion boundary. The feature-dependent performance of the coherent MVPB beamformer can be seen clearer in the in vivo study, where the imaging environment contains more complicated structures. First, we have the thyroid gland (TG) which contains scatterers much smaller than the wavelength. It can be considered as a random medium and the acoustic fields backscattered are dominated by incoherent scattering. Inside this region, the coherent MVPB image has higher quality. However, the non-random structures, such as the boundary of the CCA, are shown less well. The discontinuities in the boundary demonstrate the degradation of features that exhibit coherent scattering. Because coherent scattering can make an important contribution to image quality, its attenuation in MVDR-based algorithms should be considered when the algorithm is applied in ultrasound imaging.

There are some limitations to our study. The MVPB beamformer is based on an assumption of narrow-band signals, whereas ultrasound echo data should really be considered as a wide-band signal. The received signals among transducer elements, therefore, are correlated not only in the spatial but also in the temporal domain. Thus ideally, joint spatial and temporal filtering should be performed. The signals for beamforming are collected under assumptions of no attenuation and homogenous sound speed. Performance of the MVPB beamformers could be enhanced if compensation were introduced to correct for wave-front distortions caused by variations in attenuation and sound-speed.

V. SUMMARY AND CONCLUSIONS

One conclusion of this work is that the MVDR beamformer, when applied to data accumulated from multiple transmits, should be performed before superposition. Following this approach, we form the coherent MVPB beamformer with new algorithms for time delay calculation and covariance matrix estimation. These algorithms are based on the pressure field analysis in [1]. They help the beamformer become tolerant of the low eSNR at positions far from the beam centreline. In simulations and experiments, the coherent MVPB shows a major improvement in spatial resolution over the unified PB, the best DAS beamformer in [1].

We show experimentally how the MVDR beamformer can attenuate specular scattering in the backscattered field. The coherent MVPB, therefore, becomes less effective when specular reflection is important to the sonographic features. In general, the coherent MVPB works best when applied to diffuse random scatterers or for speckle tracking when incoherent scattering dominates the field. The method may have benefits in the context of breast cancer screening. In these applications, the tissue structures of greatest interest for lesion discrimination are the microvasculature, cell nuclei, and extracellular protein fibers composed of collagen and elastin [35]. Many of the structures are poorly organized and at a scale smaller than the wavelength. Such soft tissues are often considered as diffuse, weakly-scattering random media. In other applications that are dominated by object boundaries or coherent scattering, the correlations among the scatterers become important. In such cases, the unified PB beamformer may deliver the sonographic features with higher quality.

Both coherent MVPB and unified PB beamformer were developed based on the conventional architecture of ultrasound imaging systems. This has the advantage that any complexity in implementation relates principally to the data processing. With the development of low-cost high-performance computing, the methods can therefore be integrated into existing commercial instruments for use in clinical applications.
This appendix summarizes the unified PB beamformer developed in our earlier study [1]. It is still a DAS algorithm but relies on the two-pulse characterization presented in Section II-B. The focusing signal at point \( P \) (see Figs. 1) is given by

\[
g(x_P) = \sum_{i} \sum_{j} v_{i,j} r_{i,j}(t_P) w(t_P, t), \tag{A-1}
\]

where \( N_r \) is the number of transmits, \( N_r \) is the number of receiving elements on the active aperture, \( v_{i,j}(t) \) is the waveform or echo RF trace received on element \( j \) with transmit \( i \), and the \( v_{i,j} \)'s are apodization coefficients, \( \tau_{i,j}(t) \) is the round trip time delay used to extract signals from the received waveform \( r_{i,j}(t) \).

In each pulse-echo sequence, we use the time delay to select the signal associated with the higher magnitude pulse on the transmit wave-shape. As the receive time delay can be calculated straightforwardly, the selection depends on the transmit time delay \( \tau_{i,j} \). For \( P \) in region (I), we set \( \tau_{i,j} = \tau_{i,j}^{t_1} \), given in (6). For \( P \) in region (III), we set \( \tau_{i,j} = \tau_{i,j}^{t_2} \), given in (7). To generate an artefact-free image, it is very important to have a smooth transition from the first pulse to the second pulse. Thus, for \( P \) in region (II) (and in region (IV)), we calculate \( \tau_{i,j}^{tr} \) as an heuristic interpolation between \( \tau_{i,j}^{t_1} \) and \( \tau_{i,j}^{t_2} \) as

\[
\tau_{i,j}^{tr} = \frac{|r_a - r_{b}|}{|r_a - r_{b}|} \tau_{i,j}^{t_1} + \frac{|r_a - r_{b}|}{|r_a - r_{b}|} \tau_{i,j}^{t_2}, \tag{A-2}
\]

where \( r_a, r_b, r_p \) are vector positions of points \( A, B, \) and \( P \) (see Fig. 1(c)). The time delays \( \tau_{i,j}^{t_1} \) and \( \tau_{i,j}^{t_2} \) are given by Eq. (8).

This strategy incorporates the use of apodization coefficients \( v_{i,j} \) to suppress the backscattered noise in regions (II) and (IV) where there is not much insonification [1]. With a particular transmit beam described in Section III, we set \( v_{i,j} \) equal to one when \( P \) is in regions (I) and (III). As \( P \) moves laterally into regions (II) and (IV), \( v_{i,j} \) remains one within an element pitch, and linearly reduced to zero at a distance of three pitches. The unified PB beamformer was shown to enhance the resolution and eSNR over the entire imaging region [1].

ACKNOWLEDGMENT

We thank Dr. Laurence Berman of Addenbrooke’s Hospital, Cambridge, UK for performing the in vivo scan and advising on its interpretation. We are also grateful to the anonymous reviewers for their insightful comments and helpful suggestions.

REFERENCES


