Ultrasound Matrix Imaging. I. The focused reflection matrix and the F-factor

William Lambert, Laura A. Cobus, Mathias Fink, Alexandre Aubry

Ultrasound Matrix Imaging. I. The focused reflection matrix and the F-factor.

William Lambert,[1,2] Laura A. Cobus,[1,3] Mathias Fink,[1] and Alexandre Aubry[1,*]

[1] *Institut Langevin, ESPCI Paris, CNRS, PSL University, 1 rue Jussieu, 75005 Paris, France*
[2] *SuperSonic Imagine, Les Jardins de la Duranne, 510 Rue René Descartes, 13857 Aix-en-Provence, France*
[3] *Dodd-Walls Centre for Photonic and Quantum Technologies and Department of Physics, University of Auckland, Private Bag 92019, Auckland 1010, New Zealand*

This is the first article in a series of two dealing with a matrix approach for aberration quantification and correction in ultrasound imaging. Advanced synthetic beamforming relies on a double focusing operation at transmission and reception on each point of the medium. Ultrasound matrix imaging (UMI) consists in decoupling the location of these transmitted and received focal spots. The response between those virtual transducers form the so-called focused reflection matrix that actually contains much more information than a raw ultrasound image. In this paper, a time-frequency analysis of this matrix is performed, which highlights the single and multiple scattering contributions as well as the impact of aberrations in the monochromatic and broadband regimes. Interestingly, this analysis enables the measurement of the incoherent input-output point spread function at any pixel of this image. A focusing criterion can then be built, and its evolution used to quantify the amount of aberration throughout the ultrasound image. In contrast to the standard coherence factor used in the literature, this new indicator is robust to multiple scattering and electronic noise, thereby providing a highly contrasted map of the focusing quality. As a proof-of-concept, UMI is applied here to the *in-vivo* study of a human calf, but it can be extended to any kind of ultrasound diagnosis or non-destructive evaluation.

I. INTRODUCTION

To investigate soft tissues in ultrasound imaging, a sequence of incident waves is used to insonify the medium. Inside the medium, the waves encounter short-scale fluctuations of acoustic impedance, generating back-scattered echoes that are used to build an ultrasound image. Conventionally, this estimation of the medium reflectivity is performed using the process of delay-and-sum (DAS) beamforming, which relies on a coherent summation over all the measured back-scattered echoes generated by any point of the medium. Each echo is selected by computing the time-of-flight associated with the forward and return travel paths of the ultrasonic wave between the probe and the image voxel. From a physical point of view, time delays in transmission are used to concentrate the ultrasound wave on a focal area whose size is ideally only limited by diffraction. Time delays at reception select echoes coming from this excited area. This process falls into the so-called confocal imaging techniques, meaning that, for each point of the image, a double focusing operation is performed.

The critical step of computing the time-of-flight for each insonification and each focal point is achieved in any clinical device by assuming the medium as homogeneous with a constant speed of sound. This assumption is necessary in order to achieve the rapidity required for real-time imaging; however, it may not be valid for some configurations in which long-scale fluctuations of the medium speed of sound impact wave propagation [1]. In soft tissues, such fluctuations are around 5%, as the speed of sound typically ranges from 1400 m/s (*e.g.* fat tissues) to 1650 m/s (*e.g.* skin, muscle tissues) [2]. In such situations, the incident focal spot spreads beyond the diffraction-limited area, the exciting pressure field at this focusing point is reduced, and undesired back-scattered echoes are generated by surrounding areas. In reception, the coherent summation is not optimal as it mixes echoes which originate from a distorted focal spot. These aberrating effects can strongly degrade the image resolution and contrast. For highly heterogeneous media, such aberrations may impact the diagnosis of the medical exam or limit the capability to image some organs. A classic example of this effect is liver imaging of difficult-to-image patients, in which the probe is placed between the ribs. Because the ultrasonic waves must travel through successive layers of skin, fat, and muscle tissue before reaching the liver, both the incident and reflected wave-fronts undergo strong aberrations, distort [3].

To assess the quality of focus and monitor the convergence of aberration correction techniques, Mallart and Fink [4] introduced a focusing parameter C, known as the coherence factor in the ultrasound community [5]. Based on the van Cittert-Zernike theorem, this indicator is linked to the spatial coherence of the back-scattered field measured by the probe for a focused transmit beam. This method assumes that images are dominated by ultrasonic speckle – a grainy, noise-like texture which is generated by a medium of random reflectivity. This assumption is in fact often valid in medical ultrasound, where soft tissues are composed of randomly distributed scatterers that are unresolved at ultrasonic frequencies. In that case, each transmitted focusing wave excites an ensemble of scatterers which are randomly distributed

* alexandre.aubry@espci.fr

within the focal spot. The back-scattered wavefront results from the random superposition of the echoes generated by each of those scatterers. If the focus is perfect, the time delays used to focus (in transmission and in reception) are exactly the same as the actual round-trip time-of-flight of the waves which arrive at, and are backscattered by, scatterers located within the resolution cell. The spatial coherence of the received signal is then maximal, and the associated coherence factor C tends towards 2/3 in the speckle regime. A decrease in the quality of focus is matched by a decrease in C; thus, this indicator is popular in the literature for the evaluation of ultrasonic focusing quality.

Very recently, ultrasound matrix imaging (UMI) has been proposed for a quantitative mapping of aberrations in ultrasound imaging [6]. Experimentally, the first step in this approach is to record the reflection matrix associated with the medium to be imaged. This matrix contains the response of the medium recorded by each transducer of the probe, for a set of illuminations. Depending on the problem one is facing, this matrix can be investigated in different bases. Here, for imaging purposes, the reflection matrix will be projected into a focused basis. In contrast with conventional ultrasound imaging that relies on confocal beamforming, the idea here is to apply independent focused beamforming procedures at the input and output of the reflection matrix. This process yields a focused reflection (FR) matrix that contains the responses between virtual transducers synthesized from the transmitted and received focal spots. The FR matrix holds much more information on the medium than does a conventional ultrasound image. Importantly, the resolution can be assessed locally for any pixel of the image. By comparing this resolution with the optimal diffraction-limited one, a new focusing criterion F is introduced to assess the local focusing quality, notably in the speckle regime. While this indicator shows some similarities with the coherence factor introduced by Mallart and Fink [4], the parameter F constitutes a much more sensitive probe of the focusing quality. In the quest for local correction of distributed aberrations [7–11], this parameter F can be of particular interest as it can play the role of a guide star for any pixel of the ultrasound image. More generally, the FR matrix is a key operator for imaging applications [11–15]. The accompanying article [16] will present in details a matrix method of aberration correction in which the FR matrix plays a pivotal role.

The current paper describes the mathematical construction and physical meaning of the FR matrix. In particular, it is shown how the FR matrix can be exploited to probe the local ultrasound focusing quality, independently of the medium reflectivity. The FR matrix was originally introduced in Ref. [6]; here, it is investigated in more detail, and its properties are demonstrated for *in-vivo* ultrasound imaging of the calf of a human subject. First, a time-frequency analysis of the FR matrix is performed, and its mathematical expression is derived rigorously for both the monochromatic and broadband regimes. The relative weight of single and multiple scattering contributions, as well as the impact of aberrations, is also discussed for each regime. While Ref. [6] only considered virtual transducers belonging to the same transverse plane, here we investigate the impulse responses between focusing points located at any depth [see Fig. 1(a)]. A local incoherent input-output point spread function (PSF) can be extracted from the FR matrix along both the axial and transverse directions. The ratio between the diffraction-limited resolution δx_0 and the transverse width w_x of the PSF yields the focusing parameter F. The frequency dependence of δx_0 is taken into account to obtain a quantitative estimation of the focusing quality with the parameter F. A F-map is then built and superimposed to the ultrasound image; its comparison with the state-of-the-art C-map shows a drastic gain in terms of contrast. This striking result is due to the fact that, unlike the coherence factor C, F is robust to both multiple scattering and electronic noise.

The paper is structured as follows: Section II presents the principle of the FR matrix illustrated with *in-vivo* ultrasonic data. In Sec. III, the FR matrix is defined in the monochromatic regime. A 2D local input-output PSF is extracted from the FR matrix. The manifestation of aberrations, out-of-focus echoes and multiple scattering in this monochromatic PSF is discussed. In Sec. IV, a coherent sum of the monochromatic FR matrices leads to a broadband FR matrix. This operation, which is equivalent to a time-gating process, suppresses contributions from out-of-focus reflectors and from multiple scattering events equally. In Sec. VI, a time-frequency analysis of the FR matrix is performed and shows, in particular, the impact of absorption and scattering on the FR matrix as a function of depth and frequency. In Sec. VII, a focusing parameter F is obtained by comparing the experimentally-measured PSD width with the ideal, diffraction-limited one. It is shown that the frequency-dependence of F should be taken into account for a quantitative mapping of the focusing quality. Finally a discussion section outlines the advantages of F compared to the coherence factor C widely used in the literature [4, 5]. Section IX presents conclusions and general perspectives.

II. EXPERIMENTAL PROCEDURE

UMI begins with an experimental recording of the reflection matrix, $\mathbf{R}$. In principle, this measurement can be achieved using any type of illumination (element-by-element [17], focused beams [18], *etc.*); here, a plane-wave acquisition sequence [19] has been arbitrarily chosen. The probe was placed in direct contact with the calf of a healthy volunteer, orthogonally to the muscular fibers. (This study is in conformation with the declaration of Helsinki). The acquisition was performed using a medical ultrafast ultrasound scanner (Aixplorer Mach-30, Supersonic Imagine, Aix-en-Provence, France) driv-

ing a $5-18$ MHz linear transducer array containing 192 transducers with a pitch $p = 0.2$ (SL18-5, Supersonic Imagine). The acquisition sequence consisted of transmission of 101 steering angles spanning from -25^o to 25^o, calculated for the hypothesis of a tissue speed of sound of $c_0 = 1580$ m/s [2]. The pulse repetition frequency was set at 1000 Hz. The emitted signal was a sinusoidal burst lasting for three half periods of the central frequency $f_c = 7.5$ MHz. For each excitation, the backscattered signal was recorded by the 192 transducers of the probe over a time length $\Delta t = 80$ μs at sampling frequency $f_s = 40$ MHz. Acquired in this way, the reflection matrix is denoted $\mathbf{R}_{u\theta}(t) \equiv R(u_{\text{out}}, \theta_{\text{in}}, t)$, where u_{out} defines the transverse position of the received transducer, θ_{in}, the incident angle of the emitted plane wave and t, the time-of-flight. In the following, the subscripts *in* and *out* will denote the transmitting and receiving steps of the reflection matrix recording process.

III. REFLECTION MATRIX IN THE FREQUENCY DOMAIN

The next step consists in projecting the reflection matrix into a focused basis. This step is greatly simplified by performing beamforming operations in the frequency domain, i.e. applying appropriate phase shifts to all frequency components of the received signals in order to realign them at each focal point. A matrix formalism is particularly suitable for this operation, because, in the frequency domain, the projection of data from the plane-wave or transducer bases to any focal planecan be achieved with a simple matrix product [6].

Consequently, a temporal Fourier transform should be first applied to the experimentally acquired reflection matrix $\mathbf{R}_{u\theta}(t)$:

$$\mathbf{R}_{u\theta}(f) = \int dt\ \mathbf{R}_{u\theta}(t)\ e^{j2\pi ft}. \tag{1}$$

with f the temporal frequency. Physically, the matrix $\mathbf{R}_{u\theta}(f)$ can be seen as the result of a three-step sequence:

- Insonification of the medium by a set of plane waves. This emission sequence can be described by a transmission matrix $\mathbf{P} = [P(\mathbf{r}, \theta)]$ defined between the plane wave and focused bases. Each column of this matrix corresponds to the pressure wave-field at each point $\mathbf{r} = \langle x, z \rangle$ of the medium when a plane wave is emitted form the probe with an angle of incidence θ.

- Scattering of each incident wave by the medium heterogeneities. This scattering can be modelled by a matrix $\mathbf{\Gamma}$ defined in the focused basis; in the single scattering regime, $\mathbf{\Gamma}$ is diagonal and its elements correspond to the medium reflectivity $\gamma(\mathbf{r}, f)$ at frequency f.

- Back-propagation of the reflected waves towards the transducers. This step can be described by the Green's matrix $\mathbf{G} = [G(u, \mathbf{r})]$ between the transducer and focused bases. Each line of this matrix corresponds to the wavefront that would be recorded by the array of transducers if a point source was introduced at a point $\mathbf{r} = (x, z)$ inside the sample.

Mathematically, the matrix $\mathbf{R}_{\mathbf{u}\theta}(f)$ can thus be expressed as the following matrix product:

$$\mathbf{R}_{u\theta}(f) = \mathbf{G}^{\top}(f) \times \mathbf{\Gamma}(f) \times \mathbf{P}(f). \tag{2}$$

The holy grail for imaging is to have access to the transmission and Green's matrices, $\mathbf{P}$ and $\mathbf{G}$. Their inversion or, more simply, their phase conjugation can enable a projection of the reflection matrix into a focused basis, thereby providing an estimation of the medium's reflectivity $\gamma(\mathbf{r})$ despite aberrations and multiple scattering. However, these matrices T and G are not directly accessible in most imaging configurations.

IV. MONOCHROMATIC FOCUSED REFLECTION MATRIX

The projection of the reflection matrix into a focused basis requires a preliminary step – the definition of transmission matrices which model wave propagation from the plane-wave basis, or transducer basis, to any focal point $\mathbf{r}$ in the medium. Here, we consider free-space transmission matrices $\mathbf{P_0}(f)$ and $\mathbf{G_0}(f)$, which correspond to plane waves and two-dimensional (2D) Green's functions [20] propagating in a fictitious homogeneous medium with constant speed-of-sound, c_0:

$$\begin{aligned} P_0(\theta, \mathbf{r}, f) &= \exp(i\mathbf{k} \cdot \mathbf{r}) \\ &= \exp[ik_0 (z\cos\theta + x\sin\theta)], \end{aligned} \tag{3a}$$

$$G_0(\mathbf{r}, u, f) = -\frac{i}{4}\mathcal{H}_0^{(1)}\left(k_0\sqrt{(x-u)^2 + z^2}\right). \tag{3b}$$

with $k_0 = 2\pi f/c_0$ the wave number. These transmission matrices are then used to beamform the reflection matrix in transmission and reception:

$$\mathbf{R_{rr}}(f) = \mathbf{G_0^*}(f) \times \mathbf{R}_{u\theta}(z) \times \mathbf{P_0^\dagger}(f), \tag{4}$$

where the symbols $*$, $\dagger$ and $\times$ stand for phase conjugate, transpose conjugate and matrix product, respectively. Each row of $\mathbf{P_0^\dagger} = [P_0^*(\theta_{\text{in}}, \mathbf{r}_{\text{in}})]^T$ defines the combination of plane waves that should be applied to focus at each input focusing point $\mathbf{r}_{\text{in}} = \{x_{\text{in}}, z_{\text{in}}\}$, at frequency f. Similarly, each column of $\mathbf{G_0^*} = [G_0^*(u_{\text{out}}, \mathbf{r}_{\text{out}})]$ contains the amplitude and phase that should be applied to the signal received by each transducer u_{out} in order to sum coherently the echoes coming from the output focusing point $\mathbf{r}_{\text{out}} = \{x_{\text{out}}, z_{\text{out}}\}$. Fig. 1(a) illustrates this matrix focusing process. Note that, for clarity, the input focus-

ing operation in Fig. 1(a) is represented by a cylindrical wave-front instead of a combination of plane waves. This is justified by the fact that a plane wave synthetic beamforming numerically mimics a focused excitation [19].

In contrast to standard synthetic ultrasound imaging, in which input and output focusing points coincide, the approach presented here decouples these points. In emission, the incident energy is concentrated at the focusing point $\mathbf{r}_{\text{in}}$; this point can thus be seen as a virtual source [21, 22]. Similarly, in reception, a virtual sensor is synthesized by selecting echoes coming from the focusing point $\mathbf{r}_{\text{out}}$. Therefore, each coefficient $R(\mathbf{r}_{\text{out}}, \mathbf{r}_{\text{in}}, f)$ of $\mathbf{R_{rr}}(f)$ contains the monochromatic responses of the medium between a set of virtual transducers. The position of each virtual transducer maps onto a pixel of the ultrasound image.

Fig. 2(a_2) shows the x-projection $\mathbf{R}_{xx}(z, f)$ of $\mathbf{R_{rr}}(f)$ at depth $z = 18$ mm and frequency $f = 6$ MHz. $\mathbf{R}_{xx}(z, f) = [R(x_{\text{out}}, x_{\text{in}}, f, z)]$ contains the responses between virtual transducers, $\mathbf{r}_{\text{in}} = (x_{\text{in}}, z)$ and $\mathbf{r}_{\text{out}} = (x_{\text{out}}, z)$, located at the same depth. Note that $\Delta x_{\max}$ has been fixed to avoid the spatial aliasing induced by the incompleteness of the plane wave illumination basis; $\Delta x_{\max}$ is inversely proportional to the angular step $\delta\theta$ of the plane wave illumination basis: $\Delta x_{\max} \sim \lambda_{\max}/(2\delta\theta)$, with λ the wavelength (See Supplementary Material, Section S1). Thus, the coefficients $R(x_{\text{out}}, x_{\text{in}}, f, z)$ associated with a transverse distance $|x_{\text{out}} - x_{\text{in}}|$ larger than a superior bound $\Delta x_{\max}$ are not displayed (See Supplementary Material, Section S1).

As displayed by Fig. 2(a_2), Fig. 2(a_2) shows that most of the signal in $\mathbf{R}_{xx}(f)$ is concentrated around its diagonal. This feature is characteristic of a predominant single scattering contribution [6]. Indeed, singly-scattered echoes can only originate from a virtual detector $\mathbf{r}_{\text{out}}$ which is located very near to virtual source $\mathbf{r}_{\text{in}}$. The diagonal elements of $\mathbf{R_{rr}}(f)$, which obey $\mathbf{r}_{\text{in}} = \mathbf{r}_{\text{out}}$, directly provide the image $\mathcal{I}(\mathbf{r}, f)$ which would be obtained via multi-focus (a.k.a. confocal) imaging at frequency f:

$$\mathcal{I}(\mathbf{r}, f) \equiv |R(\mathbf{r}, \mathbf{r}, f)|^2 . \tag{5}$$

An example of such an image is displayed in Fig. 2(a_1). Compared to a standard ultrasound image built from broadband signals (Fig. 2b_1), this monochromatic image displays poor axial resolution and is therefore difficult to exploit. However, in the following, the off-diagonal elements of $\mathbf{R_{rr}}$ are shown to provide valuable information on the physical properties of the medium, as well as on the wave focusing quality.

$\mathbf{R_{rr}}(f)$ can be expressed theoretically by injecting (2) into (4):

$$\mathbf{R_{rr}}(f) = \mathbf{H}_{\text{out}}^{\top}(f) \times \mathbf{\Gamma}(f) \times \mathbf{H_{in}}(f), \tag{6}$$

where

$$\mathbf{H}_{\text{in}}(f) = \mathbf{P}(f)\mathbf{P_0^{\dagger}}(f) \text{ and } \mathbf{H}_{\text{out}} = \mathbf{G}(f)\mathbf{G_0}^{\dagger}(f) \tag{7}$$

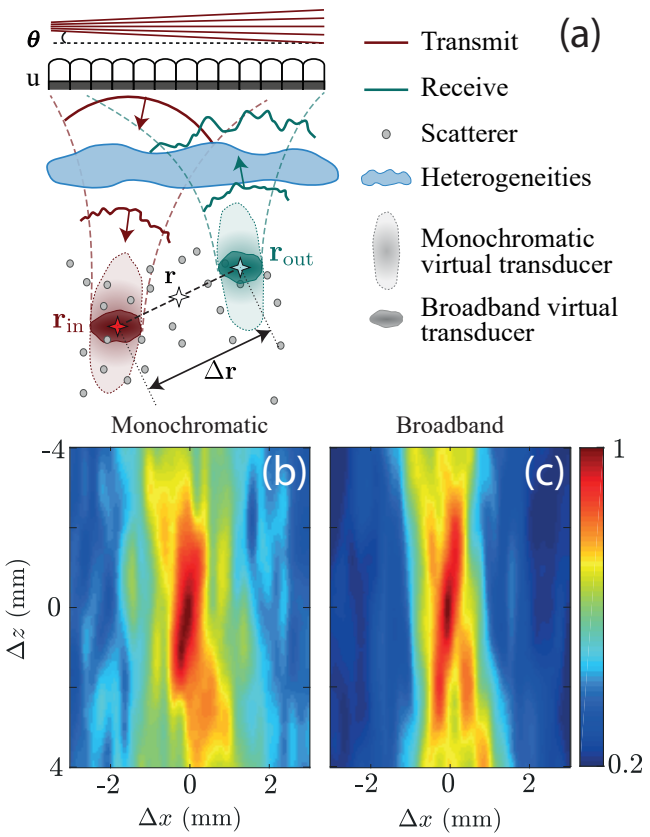

FIG. 1. Principle of the FR matrix. (a) UMI consists in splitting the location of the transmitted ($\mathbf{r}_{\text{in}}$) and received focusing ($\mathbf{r}_{\text{out}}$) points both in the axial and transverse directions, thereby synthesizing virtual transducers that can act as source and detector, respectively at any point in the medium. In a monochromatic regime, the synthesized virtual transducer displays an elongated shape in the $z-$direction because of diffraction. In the broadband domain, the axial resolution is inversely proportional to the signal bandwidth, giving a much thinner virtual transducer in the $z-$direction. (b)-(c) Monochromatic ($f = 7.5$ MHz) and broadband common mid-point intensity profile (10) averaged over a set of common mid-points $\mathbf{r}$ contained in the white rectangle displayed in Fig. 2(a_1) and (a_2). Both profiles have been normalized by their maximum.

are the input and output focusing matrices, respectively. Each column of $\mathbf{H}_{\text{in}}(f) = [H_{\text{in}}(\mathbf{r}, \mathbf{r}_{\text{in}}, f)]$ corresponds to one focused illumination, and contains the monochromatic PSF at emission – the spatial amplitude distribution of the input focal spot resulting from that particular illumination. Similarly, each column of $\mathbf{H}_{\text{out}}(f) = [H_{\text{out}}(\mathbf{r}, \mathbf{r}_{\text{out}}, f)]$ contains the spatial amplitude distribution of an output focal spot. As will be shown in the following, the spatial extension of these focal spots is a direct manifestation of the mismatch between the actual velocity distribution in the medium (corresponding to the propagation model described by transmission matrices $\mathbf{T}$ and $\mathbf{P}$), and the spatially-invariant velocity model associated with the free space transmission matrices $\mathbf{T}_0$ and

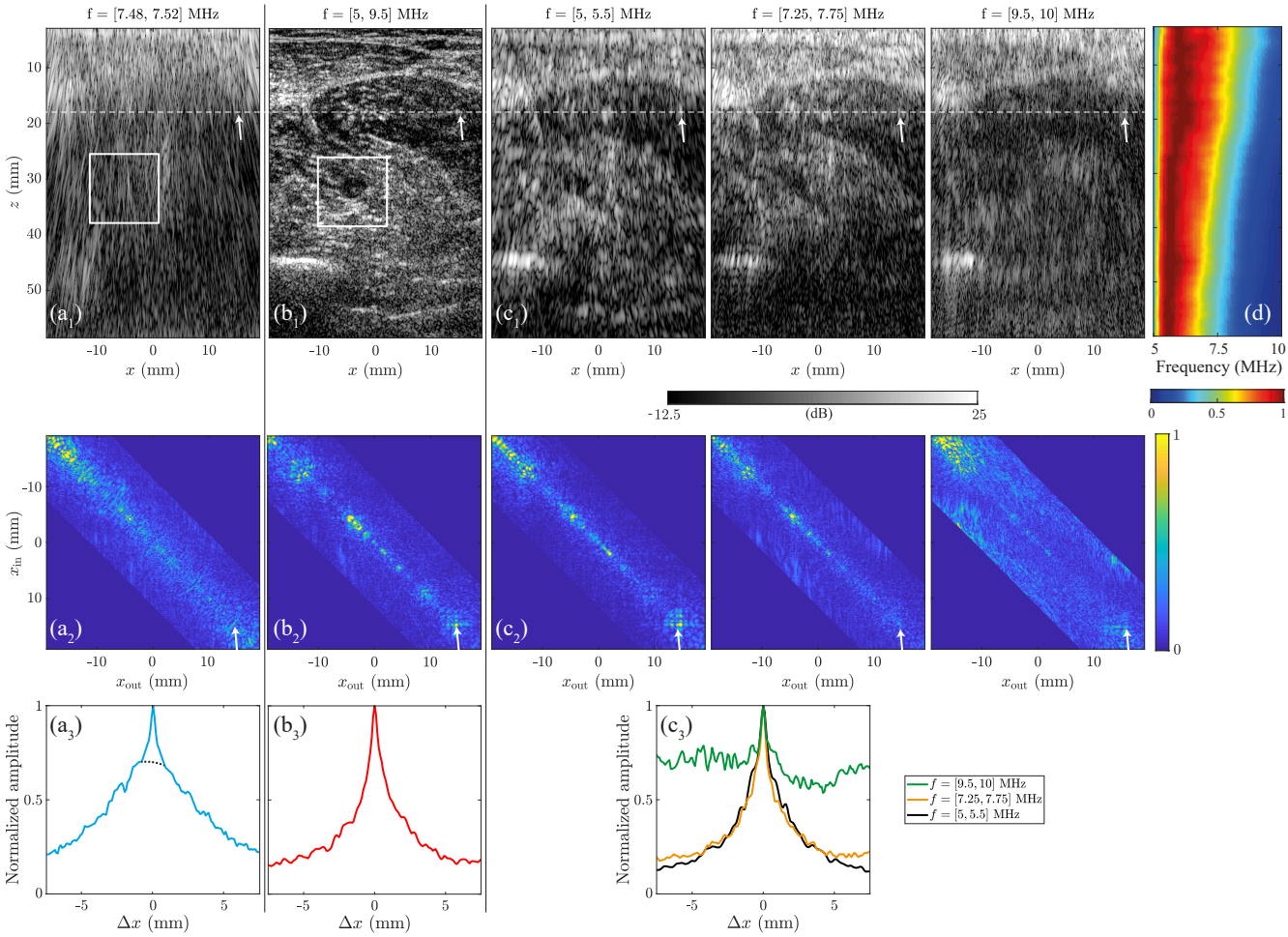

FIG. 2. UMI in a monochromatic, broadband and multi-frequency regime. (a) Monochromatic regime (f =7.5 MHz): (a_1) Multi-focus image (dB-scale), (a_2) reflection matrix $\mathbf{R}_{xx}(z)$ at z = 18 mm [see white dashed line in (a)], and (a_3) common mid-point intensity profile $I(z, \Delta x)$ ((10)) averaged at the same depth. (b) Broadband regime (5-10 MHz): (b_1) Multi-focus image (dB-scale), (b_2) Broadband reflection matrix $\mathbf{R}_{xx}(z)$ at z = 18 mm, and (b_3) common mid-point intensity profile $I(\mathbf{r}, \Delta x)$ averaged at the same depth. (c) Multi frequency regime (5-5.5 MHz, 7.25-7.75 MHz, 9.5-10 MHz, from left to right): (c_1) Multi-focus images (dB-scale), (c_2) multi-frequency reflection matrices $\mathbf{R}_{xx}(z)$ at z = 18 mm, and (c_3) common mid-point intensity profiles $I(\mathbf{r}, \Delta x)$ averaged at the same depth. (d) Spectrogram of the confocal signals versus depth z extracted from the diagonal elements of $\mathbf{R}_{xx}(z)$ (linear scale). The confocal spectrum is normalized by its maximum at each depth. The white rectangle in (a_1) and (a_2) accounts for the position of CMPs considered for the computation of the CMP intensity profiles displayed in Fig. 1(b) and (c), respectively.

$\mathbf{P}_0$. For lighter notation in the rest of this section, the frequency dependence f of each physical quantity is made implicit.

Equation (6) can be rewritten in terms of matrix coefficients as follows:

$$R(\mathbf{r}_{\text{out}}, \mathbf{r}_{\text{in}}) = \int d\mathbf{r}\; H_{\text{out}}(\mathbf{r}, \mathbf{r}_{\text{out}})\; \gamma(\mathbf{r})\; H_{\text{in}}(\mathbf{r}, \mathbf{r}_{\text{in}}). \quad (8)$$

This equation confirms that the diagonal coefficients of $\mathbf{R}_{\mathbf{rr}}$ form the confocal image at frequency f. Each pixel of this image results from a convolution between the sample reflectivity $\gamma(\mathbf{r})$ and an imaging PSF, $H(\mathbf{r}, \mathbf{r}')$, which is itself a product of the input and output PSFs: $H(\mathbf{r}, \mathbf{r}') = H_{\text{in}}(\mathbf{r}, \mathbf{r}') \times H_{\text{out}}(\mathbf{r}, \mathbf{r}')$. As $H_{\text{in}}(\mathbf{r}, \mathbf{r}_{\text{in}})$ and $H_{\text{out}}(\mathbf{r}, \mathbf{r}_{\text{out}})$ define the characteristic size of each virtual source at $\mathbf{r}_{\text{in}}$ and sensor at $\mathbf{r}_{\text{out}}$, their support thus dictates the resolution of the confocal image. In the absence of aberration, the transverse and axial dimensions of these focal spots, $\delta x_0(\mathbf{r})$ and $\delta z_0(\mathbf{r})$, are only limited by diffraction [23]:

$$\delta x_0(\mathbf{r}) = \frac{\lambda}{2\sin[\beta(\mathbf{r})]}\,,\; \delta z_0(\mathbf{r}) = \frac{2\lambda}{\sin^2[\beta(\mathbf{r})]}, \quad (9)$$

with $\beta(\mathbf{r})$ the maximum angle by which each focal point is illuminated or seen by the array of transducers. In ultrasound imaging, the radiation pattern of the transducers usually limits the numerical aperture to $\leq 25°$.

The virtual transducers thus typically display a characteristic elongated shape in the $z-$direction ($\delta x_0 << \delta z_0$), which accounts for the poor axial resolution exhibited by the monochromatic image in Fig. 2(a_1).

V. COMMON MID-POINT INTENSITY

The off-diagonal points in $\mathbf{R_{rr}}$ can be exploited for a quantification of the focusing quality at any pixel of the ultrasound image. To that aim, the relevant observable is the intensity profile along each anti-diagonal of $\mathbf{R_{rr}}$ [6]:

$$I(\mathbf{r}_\mathrm{m}, \Delta\mathbf{r}) = \left|R(\mathbf{r}_\mathrm{m} + \Delta\mathbf{r}/2, \mathbf{r}_\mathrm{m} - \Delta\mathbf{r}/2)\right|^2 . \qquad (10)$$

All pairs of points on a given anti-diagonal have the same midpoint $\mathbf{r}_\mathrm{m} = (\mathbf{r}_\mathrm{out} + \mathbf{r}_\mathrm{in})/2$, but varying spacing $\Delta\mathbf{r} = (\mathbf{r}_\mathrm{out} - \mathbf{r}_\mathrm{in})$. In the following, $I(\mathbf{r}_\mathrm{m}, \Delta\mathbf{r})$ is thus referred to as the common-midpoint (CMP) intensity profile. To express this quantity theoretically, we first make an isoplanetic approximation in the vicinity of each CMP $\mathbf{r}_\mathrm{m}$. This means that waves which focus in this region are assumed to have travelled through approximately the same areas of the medium, thereby undergoing identical phase distortions [11, 16]. The input and/or output PSFs can then be considered to be spatially invariant within this local region. Mathematically, this means that, in the vicinity of each common mid-point $\mathbf{r}_\mathrm{m}$, the spatial distribution of the input or output PSFs, $H_\mathrm{in/out}(\mathbf{r}, \mathbf{r}_\mathrm{in/out})$, only depends on the relative distance between the point $\mathbf{r}$ and the focusing point $\mathbf{r}_\mathrm{in/out}$, such that:

$$H_\mathrm{in/out}(\mathbf{r}, \mathbf{r}_\mathrm{in/out}) = H_\mathrm{in/out}(\mathbf{r} - \mathbf{r}_\mathrm{in/out}, \mathbf{r}_\mathrm{m}). \qquad (11)$$

We next make the assumption that, as is often the case in ultrasound imaging, scattering is due to a random distribution of unresolved scatterers. Such a speckle scattering regime can be modelled by a random reflectivity:

$$\langle \gamma(\mathbf{r_1})\gamma^*(\mathbf{r_2})\rangle = \langle|\gamma|^2\rangle\delta(\mathbf{r_2} - \mathbf{r_1}), \qquad (12)$$

where $\langle ...\rangle$ denotes an ensemble average and δ is the Dirac distribution. By injecting (8), (11) and (12) into (10), the following expression can be found for the CMP intensity:

$$I(\mathbf{r}_\mathrm{m}, \Delta\mathbf{r}) = \int d\mathbf{r}\ \left|H_\mathrm{out}\left(\mathbf{r} - \Delta\mathbf{r}/2, \mathbf{r}_\mathrm{m}\right)\right|^2 \left|H_\mathrm{in}\left(\mathbf{r} + \Delta\mathbf{r}/2, \mathbf{r}_\mathrm{m}\right)\right|^2 \left|\gamma(\mathbf{r} + \mathbf{r}_\mathrm{m})\right|^2. \qquad (13)$$

To smooth the intensity fluctuations due to the random reflectivity, a spatial average over a few resolution cells is required while keeping a satisfactory spatial resolution. To do so, a spatially averaged intensity profile $I_\mathrm{av}(\mathbf{r}_\mathrm{m}, \Delta\mathbf{r})$ is computed in the vicinity of each point $\mathbf{r}_\mathrm{m}$, such that

$$I_\mathrm{av}(\mathbf{r}_\mathrm{m}, \Delta\mathbf{r}) = \left\langle W_L(\mathbf{r} - \mathbf{r}_\mathrm{m}) I(\mathbf{r}, \Delta\mathbf{r})\right\rangle_\mathbf{r} \qquad (14)$$

where the symbol $\langle ...\rangle_\mathbf{r}$ denotes the spatial average and $W_L(\mathbf{r})$ is a spatial window function, such that

$$W_L(\mathbf{r}) = \begin{cases} 1 & \text{for } |\mathbf{r}| < L/2 \\ 0 & \text{otherwise.} \end{cases} \qquad (15)$$

This spatial averaging process leads to replace $|\gamma(\mathbf{r})|^2$ in the last equation by its ensemble average $\langle|\gamma|^2\rangle$. $I_\mathrm{av}(\mathbf{r}_\mathrm{m}, \Delta\mathbf{r})$ then directly provides the convolution between the incoherent input and output PSFs, $|H_\mathrm{in}|^2$ and $|H_\mathrm{out}|^2$:

$$I_\mathrm{av}(\mathbf{r}_\mathrm{m}, \Delta\mathbf{r}) \propto \left[|H_\mathrm{in}|^2 \overset{\Delta\mathbf{r}}{\circledast} |H_\mathrm{out}|^2\right](\Delta\mathbf{r}, \mathbf{r}_\mathrm{m}). \qquad (16)$$

where the symbol $\overset{\Delta\mathbf{r}}{\circledast}$ denotes a spatial convolution over $\Delta\mathbf{r}$ – the relative position of $\mathbf{r}_\mathrm{in}$ with respect to $\mathbf{r}_\mathrm{out}$. Note that this formula only holds in the speckle regime; for a specular reflector, the CMP intensity profile is equivalent to the intensity of the coherent input-output PSF, $|H_\mathrm{in} \overset{\Delta\mathbf{r}}{\circledast} H_\mathrm{out}|^2$ [6]. In either case, this quantity is gives two interesting pieces of information: firstly, the intensity enhancement $I_\mathrm{av}(\mathbf{r}_\mathrm{m}, \mathbf{0})$ of the CMP intensity profile at $\Delta\mathbf{r} = \mathbf{0}$ is a direct measure of the average overlap between virtual source and detector pairs in the isoplanetic region. Secondly, the spatial extension of the CMP is linked to the lateral and axial dimensions of the input and output PSFs. In particular, the transverse full width at half maximum (FWHM) of the CMP, $w_x(\mathbf{r}_\mathrm{m})$, is a direct indicator of the focusing quality at each point $\mathbf{r}_\mathrm{m}$ of the medium. In the absence of aberrations,

$$H_\mathrm{in/out}(\Delta x, \Delta z = 0, \mathbf{r}_\mathrm{m}) = \mathrm{sinc}(2\pi\Delta x \sin\left[\beta(\mathbf{r}_\mathrm{m})\right]/\lambda), \qquad (17)$$

where $\mathrm{sinc}(x) = \sin(x)/x$, and $w_x(\mathbf{r}_\mathrm{m})$ is roughly equal to the diffraction-limited resolution $\delta x_0(\mathbf{r}_\mathrm{m})$.

Fig. 1(b) displays an example of a two dimensional CMP intensity profile. The corresponding cross-section of this profile at $\Delta z = 0$ is displayed in Fig. 2(a_3). This cross-section has been averaged over a set of CMPs contained in the white rectangle of Fig. 2(a_1). Similarly to the input/output PSFs H_in and H_out (9), the incoherent input-output PSF $|H_\mathrm{in}|^2 \overset{\Delta\mathbf{r}}{\circledast} |H_\mathrm{out}|^2$ displays a cigar-like shape. However, while its axial FWHM $w_z(\mathbf{r}_\mathrm{m})$ is close to the diffraction limit ($w_z(\mathbf{r}_\mathrm{m}) \sim \delta z_0 \sim 2.2$ mm, with $\beta = 25°$), its transverse FWHM $w_x(\mathbf{r}_\mathrm{m})$ is far from being ideal ($w_x(\mathbf{r}_\mathrm{m}) \sim 1$ mm $>> \delta x_0 \sim 0.25$ mm). This poor resolution can be explained by several potential effects: First, long-scale variations of the speed-of-sound can give rise to aberrations that distort the input and output PSFs [11]. Second, if there are spurious echos from out-of-focus scatterers (located above or below the focal plane), then the expansion of the resulting input/beams will be greater than those originating exactly at the focal plane [Fig. 1(a)]. This will result in an enlargement of the resulting input/output PSFs. Finally, short-scale heterogeneities may induce multiple scattering events that give rise to an incoherent back-

ground in the CMP intensity profile [6]. The prevalence of multiple scattering events can be estimated from the amount of signal at off-diagonal elements – an incoherent background which is a combination of multiple scattering contributions and electronic noise. The contributions of these two effects can be separated by examining the spatial reciprocity of measured ultrasonic echos, *i.e* the symmetry of $\mathbf{R_{rr}}$ [6]. For the FR matrix displayed in Fig. 2(a_2), the degree of symmetry of the off-diagonal coefficients of $\mathbf{R_{rr}}$ is close to 50%. Therefore, the incoherent background of -14 dB in Fig. 2(a_3) contains multiple scattering and electronic noise in equal proportion. However, these contributions still constitute a problem for imaging; in the monochromatic regime under examination here, is it extremely difficult to discriminate between the effects of aberration, multiple scattering, and singly-scattered echoes taking place out-of-focus. In the next section, we show that the contributions from multiple scattering and from out-of-focus echos can be greatly reduced via a time-gating operation. This filter improves the axial resolution of the ultrasound imaging.

VI. BROADBAND REFLECTION MATRIX

Under the matrix formalism, time gating can be performed by building a broadband FR matrix $\overline{\mathbf{R}}_{\mathbf{rr}}$. In the following, we show that besides improving the axial resolution and contrast of the ultrasound image, $\overline{\mathbf{R}}_{\mathbf{rr}}$ allows a clear distinction between the contributions from single and multiple scattering.

In the frequency domain, the FR matrix is built by dephasing each RF signal in order to make scattering paths going through both input and output focal spots constructively interfere (4). A coherent sum over the overall bandwidth Δf can then be performed to build a broadband FR matrix:

$$\overline{\mathbf{R}}_{\mathbf{rr}}(\Delta f) = \frac{1}{\Delta f}\int_{f_-}^{f_+} df\ \mathbf{R_{rr}}(f) \tag{18}$$

with $f_\pm = f_c \pm \Delta f/2$ and f_c the central frequency of the RF signal bandwidth. In our experiment, $f_c = 7.5$ MHz and $\Delta f = 5$ MHz. One row of the broadband FR matrix corresponds to the situation in which the transmitted waves are focused at $\mathbf{r}_{\rm in}$, creating a virtual source while the virtual detector probes the spatial spreading of this virtual source at the expected ballistic time, *i.e* at time $t = 0$ in the focused basis [Fig. 1(a)]. Thus, the sum of monochromatic FR matrices over the entire bandwidth can be interpreted as a time-gating operation in which echos originating from a certain range of times-of-flight are extracted.

With the time-gating applied (18), the axial resolution of the virtual transducers should be drastically improved [see Fig. 1(a)]. To prove this assertion, we now derive an expression for the broadband FR matrix within the formalism of this work. For sake of simplicity and analytical tractability, paraxial and isoplanatic (11) approximations are made. The monochromatic PSFs can be decomposed as follows:

$$H_{\rm in/out}(\mathbf{r},\mathbf{r}',f) = \overline{H}_{\rm in/out}(\mathbf{r}-\mathbf{r}',\mathbf{r}_{\rm m},f)\ e^{j2\pi f(z-z')/c} \tag{19}$$

where $\overline{H}_{\rm in/out}$ represents the envelope of the PSF. Injecting (8) and (19) into (18) leads to the following expression for the coefficients of $\mathbf{R_{rr}}(\Delta f)$ (See Supplementary Material, Section S2.):

$$\begin{aligned}\overline{R}(\mathbf{r}_{\rm out},\mathbf{r}_{\rm in},\Delta f) = \int d\mathbf{r}\ e^{j2\pi f_c(2z-z_{\rm in}-z_{\rm out})/c} \\ \mathrm{sinc}\left[\frac{\pi\Delta f}{c}(2z-z_{\rm in}-z_{\rm out})\right] \\ \overline{H}_{\rm out}(\mathbf{r}-\mathbf{r}_{\rm out},\mathbf{r}_{\rm m})\ \gamma(\mathbf{r})\ \overline{H}_{\rm in}(\mathbf{r}-\mathbf{r}_{\rm in},\mathbf{r}_{\rm m}).\end{aligned} \tag{20}$$

where we have assumed, in first approximation, that $\overline{H}_{\rm in/out}$ is constant over the frequency bandwidth. The occurrence of the sinc factor in the integrand of the last equation shows that, in the broadband regime, the axial resolution $\overline{\delta z}_0$ is dictated by the frequency bandwidth, such that

$$\overline{\delta z}_0 \sim \frac{c}{2\Delta f}. \tag{21}$$

Fig. 2(b_1) shows the ultrasound image $\overline{\mathcal{I}}(\mathbf{r},\Delta f)$ built from the diagonals of $\overline{\mathbf{R}}_{\mathbf{rr}}(\Delta f)$ (5). The gain in resolution, compared to the original monochromatic image in [Fig. 2(a_1)], is clearly visible. While the monochromatic image displays an elongated speckle that is not very sensitive to the structural details of the calf, the coherent sum of (18) drastically improves the axial resolution and contrast of the image, revealing the micro-architecture of the calf tissues.

Fig. 2(b_2) shows the cross-section $\overline{\mathbf{R}}_{xx}(\Delta f, z)$ of the broadband FR matrix $\overline{\mathbf{R}}_{\mathbf{rr}}(\Delta f)$ at depth $z = 18$ mm [dotted white line of Fig. 2(b_1)]. Compared to its monochromatic counterpart [Fig. 2(a_2)], the single scattering contribution along the diagonal of $\overline{\mathbf{R}}_{\mathbf{rr}}$ is enhanced with respect to the off-diagonal coefficients. This enhancement is due to the time-gating procedure; contributions from scatterers which sit above and below the focal plane have been eliminated, so the remaining singly-scattered echoes are located near the diagonal of $\overline{\mathbf{R}}_{\mathbf{rr}}$. Note that the multiply-scattered echoes whose time-of-flight differs from the ballistic time are also removed in the same proportion. The single-to-multiple scattering ratio is thus preserved through the time gating operation.

The improvement in transverse resolution can also be seen in the broadband CMP intensity profile displayed in Fig. 1(c), and in its cross-section shown in Fig. 2(b_3). Compared to the monochromatic regime [Figs. 1(b) and 2(a2)], the CMP cross-section now displays a confocal, steep peak on top of a flat multiple scattering background. Surprisingly, although we are in a broadband regime, the 2D focal spot in Fig. 1(c) still exhibits a

cigar-like shape. To understand the reason for this, we now express the ensemble average of the CMP intensity profile (19) in the broadband regime under the paraxial approximation (see Supplementary Section S2):

$$\overline{I}_{\rm av}(\mathbf{r}_{\rm m},\Delta\mathbf{r},\Delta f)=\frac{A}{\Delta f}\int_{f_-}^{f_+}df\left[|\overline{H}_{\rm in}|^2\overset{\Delta x}{\circledast}|\overline{H}_{\rm out}|^2\right](\Delta\mathbf{r},\mathbf{r}_{\rm m},f), \tag{22}$$

where A is a constant. Equation 22 shows that, in the broadband regime, the CMP intensity profile corresponds to the sum over the frequency bandwidth of the incoherent input-output PSF $|\overline{H}_{\rm in}|^2\overset{\Delta x}{\circledast}|\overline{H}_{\rm out}|^2$. This explains why the axial resolution has not been improved by the time gating process, as it was for the broadband FR matrix (Eq. 20). Nevertheless, the CMP intensity evolution along Δx still offers a way to estimate the transverse resolution of the imaging PSF in the broadband regime. Indeed, Fig. 2(b3) shows that the transverse resolution remains far from being optimized: $\overline{w}_x(\Delta f)\sim 0.5$ mm $>> \delta x_0(f_c)\sim 0.25$ mm.

In the next section, we aim to develop a better way to evaluate transverse resolution in the broadband regime. We will define a quantitative parameter to measure the focusing quality at any pixel of the ultrasound image. This observable compares the experimental $\overline{\delta x}_0(\Delta f)$ to the ideal diffraction-limited case. To make our measurement quantitative, the theoretical prediction of the diffraction-limited $\overline{\delta x}_0(\Delta f)$ should be as accurate as possible; in the following, we work towards this accuracy by developing a theoretical time-frequency analysis of the FR matrix.

VII. TIME-FREQUENCY ANALYSIS OF THE FOCUSED REFLECTION MATRIX

A time-frequency analysis of the FR matrix is required to investigate the evolution of absorption and scattering as a function of frequency. To do so, the coherent sum of the monochromatic FR matrices [Eq. 18] can be performed over a smaller bandwidth δf centered on a given frequency f:

$$\overline{\mathbf{R}}_{\mathbf{rr}}(f,\delta f)=\frac{1}{\delta f}\int_{f-\delta f}^{f+\delta f}df'\;\mathbf{R}_{\mathbf{rr}}(f') \tag{23}$$

We have shown that the axial dimension $\overline{\delta z}_0$ of the virtual transducers is inversely proportional to the frequency bandwidth δf (Eq. 21). Thus, a compromise must to be made between the spectral and axial resolutions. Here, the following choice has been made: $\delta f = 0.5$ MHz and $\overline{\delta z}_0 = 3$ mm.

Fig. 2(c) shows the ultrasound images [Fig. 2(c_1)], FR matrices [Fig. 2(c2)] and CMP profiles [Fig. 2(c3)] for three different frequency bandwidths: 5-5.5 MHz, 7.25-7.75 MHz, 9.5-10 MHz. The axial resolution in each ultrasound image is of course deteriorated compared to the broadband image [Fig. 2(b_1)]; nevertheless, the time-frequency analysis of the FR matrices yields the evolution of the SNR versus depth and frequency. At $z = 18$ mm, for instance, the FR matrix at $f = 9.75$ MHz exhibits a tiny confocal intensity enhancement on top of a predominant noise background (SNR$\sim$3dB). Conversely, the FR matrices at $f = 5.25$ and 7.5 MHz exhibit a CMP intensity profile which more closely resembles its broadband counterpart. This weak SNR at 9.75 MHz can be partially explained by the finite bandwidth of the transducers ($5-10$ MHz). Interestingly, absorption losses undergone by ultrasonic waves in soft tissues also have a strong impact on the ultrasound image. Fig. 2(d) illustrates the effect of absorption by displaying the spectrum of the confocal signal, $\left\langle\overline{\mathcal{I}}(\mathbf{r},f,\delta f)\right\rangle_x$, as a function of depth. This spectrum shifts towards low frequencies as a function of depth. This frequency shift is characteristic of absorption losses in soft tissue, for which the attenuation coefficient is assumed to exhibit a linear dependence on frequency [2].

The time-frequency analysis of the FR matrix also allows the observation of frequency fluctuations in $\overline{\mathcal{I}}(\mathbf{r},f,\delta f)$ which are not related to absorption. For instance, while a bright scatterer [white arrow in Fig. 2(b_1)] is clearly visible in the broadband or low frequency regimes around $x = 12.5$ mm and $z = 18$ mm, its presence is not revealed by the FR matrices at higher frequencies. A time-frequency analysis can thus be of interest for characterization purposes, as the frequency response of bright scatterers is directly related to their size and composition.

Whether due to absorption or scattering, the variation of the temporal frequency spectrum of back-scattered echoes has a strong impact on the local resolution of the ultrasound images. In the next section, we show how to incorporate this frequency dependence in the theoretical expression of the spatial resolution, in order to establish a more precise quantitative focusing parameter using the CMP intensity profile.

VIII. THE LOCAL FOCUSING CRITERION

In this section, a local focusing criterion is established for the broadband ultrasound image. For the sake of lighter notation, the dependence of each physical quantity with respect to Δf is omitted. Aberrations caused by medium heterogeneities degrade the resolution of the ultrasound image and induce a spreading of singly-scattered echoes over the off-diagonal coefficients of $\overline{\mathbf{R}}_{\mathbf{rr}}$. In the speckle regime, it is difficult to determine by eye whether the image is aberrated, and if so, which areas are the most impacted. Interestingly, the transverse width $\overline{w}_x(\mathbf{r}_{\rm m})$ of the CMP intensity profile can yield an unambiguous answer to this question. In the speckle regime, this profile yields the convolution between the incoherent input-output PSF averaged over the frequency bandwidth (Eq. 22). While the incoherent

input-output PSF is not exactly equal to the confocal imaging PSF, (Eq. 20), it nevertheless fully captures the impact of transverse aberrations. It thus constitutes a relevant observable for assessing focusing quality.

A precise measurement of the transverse width of the CMP intensity profile, $\overline{w}_x(\mathbf{r}_\mathrm{m})$, requires an estimate of the background noise level, $I_b(\mathbf{r}_\mathrm{m})$. In practice, $I_b(\mathbf{r}_\mathrm{m})$ is computed by averaging the CMP profile away from the confocal peak, such that $I_b = \langle \overline{I}_\mathrm{av}(\mathbf{r}_\mathrm{m}, \Delta x)\rangle_{|\Delta x|>w_c}$, where w_c is the overall support of the confocal peak. In the case considered here, $w_c \sim 17\lambda_c \sim 3.5$ mm (see Fig. 2b3), where $\lambda_c \approx 0.21$ mm is the central wavelength defined at $f_c = 7.5$ MHz. $\overline{w}_x(\mathbf{r}_\mathrm{m})$ is then deduced as the FWHM of the CMP intensity profile $\overline{I}_\mathrm{av}(\mathbf{r}_\mathrm{m}, \Delta x)$ after substraction of the background noise $I_b(\mathbf{r}_\mathrm{m})$.

The width $\overline{w}_x(\mathbf{r}_\mathrm{m})$ of the imaging PSF is dictated by two distinct phenomena: diffraction and aberration. In the ideal case (*i.e* no aberrations), the image resolution $\overline{\delta x}_0(\mathbf{r}_\mathrm{m})$ is only impacted by diffraction and, more precisely, dictated by the angular aperture $\beta(\mathbf{r}_\mathrm{m})$ (Eq. 9). The angular aperture tends to decrease decrease at larger depths and nearer to the edge of the image. In the presence of aberrations, both diffraction and wave-front distortions alter the imaging PSF. In order to provide an indicator that is only sensitive to aberrations, the estimator $\overline{w}_x(\mathbf{r}_\mathrm{m})$ of the image resolution should be compared to its ideal value $\overline{\delta x}_0(\mathbf{r}_\mathrm{m})$ (Eq. 9) at each point of the ultrasound image. To do so, a local focusing criterion $F(\mathbf{r}_\mathrm{m})$ is defined:

$$F(\mathbf{r}_\mathrm{m}) = \overline{\delta x}_0(\mathbf{r}_\mathrm{m})/\overline{w}_x(\mathbf{r}_\mathrm{m}). \tag{24}$$

To estimate this parameter, the major challenge lies in the determination of the ideal resolution $\overline{\delta x}_0(\mathbf{r}_\mathrm{m})$ for broadband signals. To do so, the frequency spectrum of the ultrasound image should be taken into account. For each CMP $\mathbf{r}_\mathrm{m}$, $\mathcal{I}(\mathbf{r}_\mathrm{m}, f)$ is an estimation of the desired frequency spectrum [Fig. 2(d)]. Equation (9) can then be extended to broadband signals by using $\mathcal{I}(\mathbf{r}_\mathrm{m}, f)$ as a weighting factor. The expected lateral resolution is then given by:

$$\overline{\delta x}_0(\mathbf{r}_\mathrm{m}) = \left\langle W_{\Delta\mathbf{r}}(\mathbf{r}-\mathbf{r}_\mathrm{m}) \frac{\int_{f_-}^{f_+} df\ \mathcal{I}(\mathbf{r}, f)\ \delta x_0(f, \mathbf{r})}{\int_{f_-}^{f_+} df\ \mathcal{I}(\mathbf{r}, f)} \right\rangle_{\mathbf{r}}. \tag{25}$$

By injecting (25) into (24), the focusing criterion, $F(\mathbf{r}_\mathrm{m})$, can be computed. Fig. 3(b) displays the focusing criterion superimposed onto the conventional B-mode image. The extension of the spatial window $W_{\Delta\mathbf{r}}$ has been set to $7\lambda_c$. While high values of F ($F \sim 1$, blue areas) indicate good image reliability, low values of F ($F < 0.3$, yellow areas) indicate a poor quality of focus and gray areas are associated with a low SNR. Both yellow and grey areas seem to correspond to blurring of the ultrasound image [Fig. 3a]. Indeed, these areas correspond to the situation in which the estimation of the image resolution has failed, meaning that there is no intensity enhancement of the close-diagonal coefficients in $\mathbf{R}_{xx}(z, \Delta f)$. Two complementary reasons can explain this behaviour: Either the single scattering contribution is drowned out by a more dominant background noise (caused by multiple scattering processes and electronic noise), or the aberrations are so intense that the confocal spot spreads over an extended imaging PSF, thereby pushing the single scattering intensity at focus below the noise level. Unsurprisingly, this situation occurs at large depth and in areas where the medium reflectivity is weak. In any case, if a medical diagnosis is desired, the features of the ultrasound image in these areas should be more carefully interpreted.

Fig. 3(b) reveals a poor focusing quality at a shallow depth that could be due to a mismatch between the speed of sound in the skin, $c_\mathrm{skin} \approx 1500 - 1700$ m/s [24], and our wave velocity model ($c_0 = 1580$ m/s). The ultrasound images show different structures that are associated with their own speed of sound: (i) muscles tissues with three different fiber orientations [areas I, II, III on Fig. 3(a)]; (ii) two veins located at $\{x, z\} = 12,5$ mm and $-5,33$ mm; and (iii) the fibula, located at the bottom left of the figure $-12,45$ mm. However, finding the link between quality of focus provided by the focusing criterion and the spatial distribution of the speed of sound throughout the medium is a difficult task that is beyond the scope of this paper.

IX. DISCUSSION

The study presented here – *in-vivo* UMI of a human calf – provides new insights into the construction of the FR matrix and the focusing criterion. Our results are representative of *in-vivo* ultrasound imaging in which the medium under investigation is composed of different kinds of tissues, which can themselves be heterogeneous. As the medium is composed of a mix of unresolved scatterers and specular reflectors, the scattering of ultrasound varies considerable in space, ranging from areas of strong to weak scattering. A direct and important application of this work would thus be to aid aberration correction by employing the parameter F as a virtual guide star for adaptive focusing techniques. Currently, in the literature in this area, this role is performed by the coherence factor C [25–27].

Here, to highlight the benefit of our matrix approach with respect to the state-of-the-art, we build a map of the standard coherence factor C for the ultrasound image of the human calf. C is equal to the ratio of the coherent intensity to the incoherent intensity of the realigned reflected wave-fronts recorded by the probe for each input focusing beam [28]. Just as with the calculation of the CMP intensity profile (14), the raw coherence factor $C(\mathbf{r}_\mathrm{in})$ is then spatially averaged over overlapping spatial windows to smooth the fluctuations due to the random reflectivity. The result is mapped onto the ultrasound image in Fig. 3c. Compared to the $F-$map (Fig. 3b), the coherence factor C provides a weakly con-

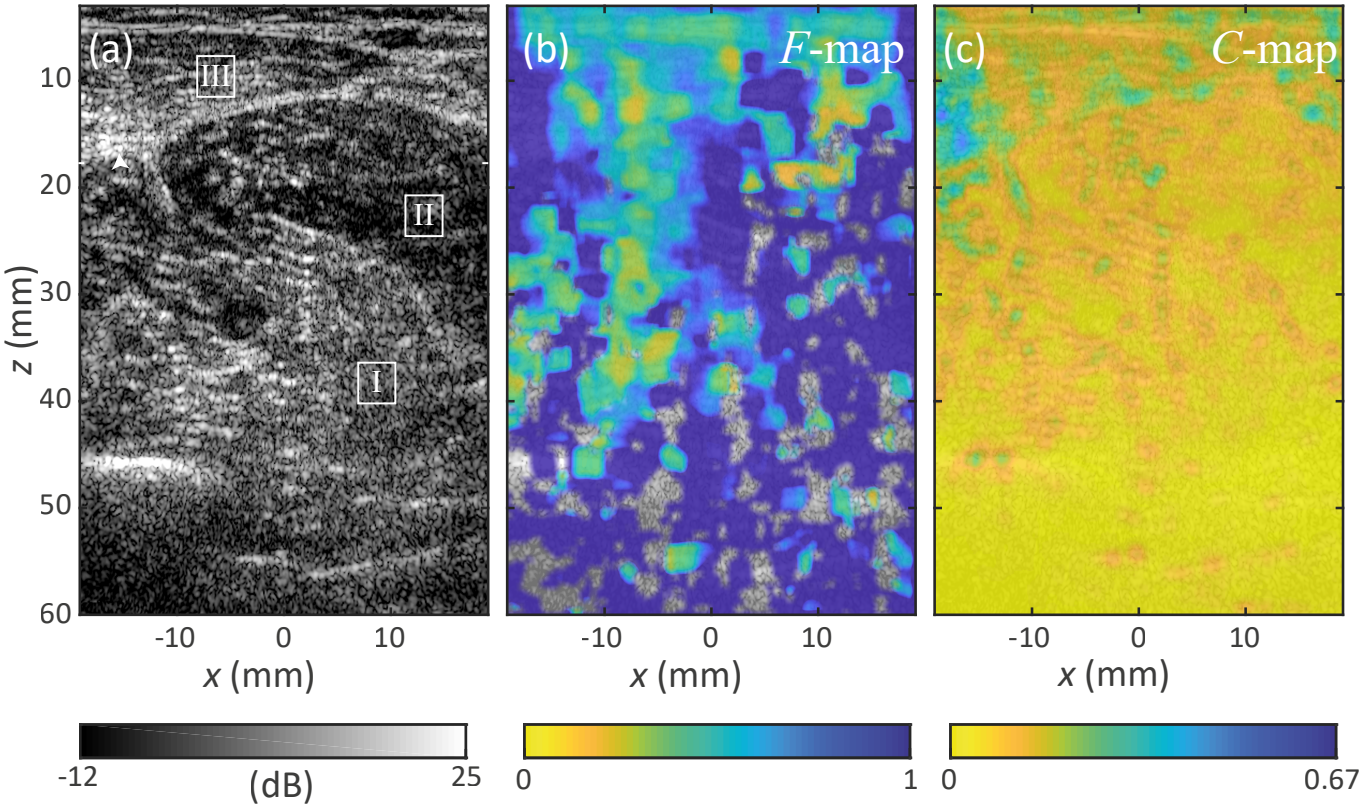

FIG. 3. (a) Conventional ultrasound image extracted from the diagonal of the broadband FR matrix $\mathbf{R_{rr}}$. (b) Focusing factor F. (c) Coherence factor C. Blue and yellow areas correspond to a high and low quality of focus respectively. In (b), gray areas highlights location where the estimation of the focusing criterion have failed, implying a very low quality of focus and/or a weak SNR.

trasted image of the focusing quality. To understand this difference, the analytical expression of C can be derived under the isoplanatic hypothesis (11). While the CMP intensity enables a full retrieval of the spatial convolution between the incoherent input and output PSFs (16), the coherence factor C only probes this quantity at $\Delta x = 0$ (See Supplementary Material, Section S3).:

$$C(\mathbf{r}) = \left[|\overline{H}_{\rm in}|^2 \overset{\Delta x}{\circledast} |\overline{H}_{\rm out}|^2 \right] (\Delta x = 0, \mathbf{r}). \qquad (26)$$

In the speckle regime, the coherence factor C ranges from 0, for strong aberrations, to 2/3 in the ideal case [4]. However, the presence of multiple scattering and noise can strongly hamper its measurement. As highlighted by Fig. 2(b_2), the incoherent input-output PSF, $|\overline{H}_{\rm in}|^2 \overset{\Delta x}{\circledast} |\overline{H}_{\rm out}|^2$, exhibits a noise background that is far from being negligible. Unlike the C-factor, the CMP intensity enables a clear discrimination between the single scattering contribution and the noise components. Its transverse width w_x is sensitive only to aberrations, since it is estimated after removing the incoherent background. This crucial feature accounts for the difference in behavior between F and C in Fig. 3, especially at large depths. While F shows a close to ideal value of 1 beyond $z = 50$ mm, C tends to 0 because of a predominant noise background. This result demonstrates one of the benefits of UMI compared to standard ultrasound imaging: Probing the focusing quality in the focused basis drastically improves the robustness to multiple scattering and noise compared to a direct cross-correlation of back-scattered echoes in the transducer basis.

In the second paper of the series [16], we present a matrix approach for aberration correction in *in vivo* ultrasound imaging. An important aspect of that study is the use of F to map the resolution of the image at each step of the aberration correction process.

X. CONCLUSION

In summary, we successfully applied the concept of the FR matrix to the case of *in vivo* ultrasound imaging of a human calf. Thanks to the intuitive concept of virtual transducers, the FR matrix provides a wealth of information on the medium that goes well beyond a single confocal image. By splitting the location of the transmitted and received focal spots, the local resolution of the ultrasound image can be assessed at any pixel. By performing a time-frequency analysis of the reflection

matrix, the contributions of single and multiple scattering and their impact on the resolution and contrast were carefully investigated. This time-frequency study of the FR matrix paves the way towards a quantitative characterization of soft tissues by measuring parameters such as the attenuation coefficient or the scattering mean free path. In the accompanying paper [16], the FR matrix will be used as a key building block of UMI for a local aberration correction. Relatedly, a focusing criterion was defined from the FR matrix in order to quantify the impact of aberrations on each pixel of the ultrasound image. Compared to the coherence factor generally used in the literature [4, 5], our focusing parameter is much more robust to noise and multiple scattering. Our focusing parameter is thus promising for use in medical imaging as a reliability index of the ultrasound image. It can also be used as a guide star for adaptive focusing techniques, or as a local aberration indicator for UMI [16].

ACKNOWLEDGMENT

The authors wish to thank Victor Barolle, Amaury Badon and Thibaud Blondel whose own research works in optics and seismology inspired this study. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 819261); the Labex WIFI (Laboratory of Excellence within the French Program Investments for the Future, ANR-10-LABX-24 and ANR-10-IDEX-0001-02 PSL*). W.L. acknowledges financial support from the Supersonic Imagine company. L.C. acknowledges financial support from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 744840.

Supplementary Information

This document provides further information on: (i) the spatial aliasing of the focused reflection matrix; (ii) the theoretical derivation of the incoherent input-output PSF in the broadband regime; (iii) the theoretical expression of the coherence factor C.

S1. SPATIAL ALIASING IN FOCUSED REFLECTION MATRIX

This Supplementary Section provides a theoretical description of the maximal lateral distance $\Delta x_{\max}$ between the two virtual transducers. Performing a change of basis between the focused basis and the plane wave basis can be interpreted as a spatial Fourier transform. As such, to avoid aliasing, the Shannon criterion should be respected. As stated in Sec.III of the accompanying paper, the FR matrix is the result of the convolution of the input and output focusing matrices with the medium reflectivity (6). In the absence of aberration, $\mathbf{H}_{\rm in}$ can be written:

$$\mathbf{H}_{\rm in} = \mathbf{P_0} \times \mathbf{P_0}^{\dagger}, \tag{S1}$$

and in term of coefficients:

$$H(\mathbf{r}_{\rm in}, \mathbf{r}, f) = \sum_{\theta} \exp\left[i \frac{2\pi}{\lambda} \ (x_{in} - x) \ \sin(\theta) \right]. \tag{S2}$$

This equation is a spatial Fourier transform that requires the following condition:

$$\Delta x_{\max} = \frac{\lambda}{2\delta\theta_{\rm in}}, \tag{S3}$$

with $\Delta x_{\max}$ the maximal transverse distance between the two virtual transducers and $\delta\theta_{\rm in}$ the angle step fixed by the ultrasound sequence used to insonify the medium. A similar criterion also exists in reception for $\mathbf{H}_{\rm out}$; however, ultrasound probes are generally designed with a pitch $p \propto \lambda$ or $p \propto \lambda/2$. The transmitcondition is then generally the limiting one for plane wave imaging modes.

S2. DERIVATION OF THE INCOHERENT INPUT-OUTPUT PSF IN THE BROADBAND REGIME

In the broadband regime, a theoretical expression for the reflection matrix can be derived by injecting (8) and (19) into (18):

$$\begin{aligned}\overline{R}(\mathbf{r}_{\rm out}, \mathbf{r}_{\rm in}, \Delta f) = \frac{1}{\Delta f} \int_{f_-}^{f_+} df \\ \int d\mathbf{r} \ \gamma(\mathbf{r}) \ \overline{H}_{\rm out}(\mathbf{r} - \mathbf{r}_{\rm out}, \mathbf{r}_{\rm m}, f) \\ \overline{H}_{\rm in}(\mathbf{r} - \mathbf{r}_{\rm in}, \mathbf{r}_{\rm m}, f) \ e^{j2\pi f(2z - z_{\rm in} - z_{\rm out})/c}\end{aligned} \tag{S4}$$

If, in first approximation, we assume that the PSF envelopes $\overline{H}_{\rm in/out}$ do not depend on frequency, (20) is obtained. In the following, this approximation will not be made.

To express the CMP intensity profile, the change of variable $(\mathbf{r}_{\rm in}, \mathbf{r}_{\rm out}) \rightarrow (\mathbf{r}_{\rm m}, \Delta\mathbf{r})$ should first be made. It yields:

$$\begin{aligned}\overline{R}(\mathbf{r}_{\rm m} + \Delta\mathbf{r}/2, \mathbf{r}_{\rm m} - \Delta\mathbf{r}/2, \Delta f) = \\ \frac{1}{\Delta f} \int_{f_-}^{f_+} df \ \int d\mathbf{r}' \ \overline{H}_{\rm out}(\mathbf{r}' - \Delta\mathbf{r}/2, f) \ \gamma(\mathbf{r}' + \mathbf{r}_{\rm m}) \\ \overline{H}_{\rm in}(\mathbf{r}' + \Delta\mathbf{r}/2, f) \ e^{j4\pi f z'/c}\end{aligned} \tag{S5}$$

where the change of variable $\mathbf{r}' = \mathbf{r} - \mathbf{r}_{\rm m}$ has been made in the integrand. Injecting (S5) into (10) leads to the following expression for the CMP intensity profile:

$$\begin{aligned}\overline{I}(\mathbf{r}_{\rm m}, \Delta\mathbf{r}) = \frac{1}{\Delta f^2} \int_{f_-}^{f_+} df \int_{f_-}^{f_+} df' \int d\mathbf{r}_1 \int d\mathbf{r}_2 \\ e^{j4\pi f z_1/c} \ e^{-j4\pi f' z_2/c} \ \gamma(\mathbf{r}_1 + \mathbf{r}_{\rm m}) \ \gamma^*(\mathbf{r}_2 + \mathbf{r}_{\rm m}) \\ \overline{H}_{\rm out}\left(\mathbf{r}_1 - \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f\right) \ \overline{H}^*_{\rm out}\left(\mathbf{r}_2 - \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f'\right) \\ \overline{H}_{\rm in}\left(\mathbf{r}_1 + \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f\right) \ \overline{H}^*_{\rm in}\left(\mathbf{r}_2 + \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f'\right).\end{aligned} \tag{S6}$$

In the speckle regime (12) and upon spatial averaging (14), an expression for the mean broadband CMP intensity profile can be derived:

$$\begin{aligned}\overline{I}_{\rm av}(\mathbf{r}_{\rm m}, \Delta\mathbf{r}, \Delta f) = \\ \frac{\langle|\gamma|^2\rangle}{\Delta f^2} \int_{f_-}^{f_+} df \int_{f_-}^{f_+} df' \int d\mathbf{r} \ e^{j4\pi(f - f') \ z/c} \\ \left\langle \overline{H}_{\rm out}\left(\mathbf{r} - \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f\right) \ \overline{H}^*_{\rm out}\left(\mathbf{r} - \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f'\right) \right\rangle \\ \left\langle \overline{H}_{\rm in}\left(\mathbf{r} + \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f\right) \ \overline{H}^*_{\rm in}\left(\mathbf{r} + \Delta\mathbf{r}/2, \mathbf{r}_{\rm m}, f'\right) \right\rangle.\end{aligned} \tag{S7}$$

By applying the following change of variable, $f_0 = (f + f')/2$ and $f'' = (f - f')/2$, the last equation can be

rewritten as follows:

$$\overline{I}_{\rm av}(\mathbf{r}_{\rm m},\Delta\mathbf{r},\Delta f)=\frac{\langle|\gamma|^2\rangle}{\Delta f^2}\int_{f_-}^{f_+}df_0 \int_{-\Delta f/2}^{\Delta f/2}df''\eta_{\rm out}(f'')\eta_{\rm in}(f'')\ e^{j8\pi f''z/c} \int d\mathbf{r}\overline{H}_{\rm out}\left(\mathbf{r}-\Delta\mathbf{r},\mathbf{r}_{\rm m},f_0\right)|^2|\overline{H}_{\rm in}\left(\mathbf{r}+\Delta\mathbf{r}/2,\mathbf{r}_{\rm m},f_0\right)|^2, \tag{S8}$$

where $\eta_{\rm in/out}(f'')$ is the frequency correlation function of $\overline{H}_{\rm in/out}$:

$$\eta_{\rm in/out}(f'')=\frac{\left\langle\overline{H}_{\rm in/out}\left(\mathbf{r},\mathbf{r}_{\rm m},f_0+f''\right)\overline{H}^*_{\rm in/out}\left(\mathbf{r},\mathbf{r}_{\rm m},f_0-f''\right)\right\rangle}{|\overline{H}_{\rm in/out}\left(\mathbf{r},\mathbf{r}_{\rm m},f_0\right)|^2}. \tag{S9}$$

Here we have assumed implicitly that the distribution of the heterogeneities in the medium is statistically invariant both in the spatial domain and over the frequency bandwidth. This implies a normalized frequency correlation function that does not depend on $\mathbf{r}_{\rm m}$, $\mathbf{r}$ or f_0. A further assumption is to consider the aberrations as relatively coherent over the frequency bandwidth, such that $\eta_{\rm in/out}$ is a slow function of f'' over the frequency bandwidth Δf: $\eta_{\rm in/out}\simeq 1$. Under this approximation, (S9) then simplifies to:

$$\overline{I}_{\rm av}(\mathbf{r}_{\rm m},\Delta\mathbf{r},\Delta f)=\frac{\langle|\gamma|^2\rangle}{\Delta f}\int_{f_-}^{f_+}df_0\int d\mathbf{r}\,\mathrm{sinc}\left(4\pi z\Delta f/c\right) |\overline{H}_{\rm out}\left(\mathbf{r}-\Delta\mathbf{r}/2,\mathbf{r}_{\rm m},f_0\right)|^2\ |\overline{H}_{\rm in}\left(\mathbf{r}+\Delta\mathbf{r}/2,\mathbf{r}_{\rm m},f_0\right)|^2$$

The sinc function, which is a function of z, displays a characteristic width equal to $\overline{\delta z}_0/2$, corresponding to half the axial resolution of the broadband ultrasound image. As $\overline{\delta z}_0$ is much smaller than δz, the z-extension of the input and output focal spots $H_{\rm in}$ and $H_{\rm out}$, the following approximation can be made: $\mathrm{sinc}\left(4\pi\Delta f z/c\right)\simeq\pi\overline{\delta z}_0\delta(z)/2$. It leads to the following expression of the CMP intensity profile:

$$\overline{I}_{\rm av}(\mathbf{r}_{\rm m},\Delta\mathbf{r},\Delta f)=\frac{A}{\Delta f}\int_{f_-}^{f_+}df_0\int dx\ |\overline{H}_{\rm out}\left(\{x-\Delta x/2,-\Delta z/2\},\mathbf{r}_{\rm m},f_0\right)|^2\ |\overline{H}_{\rm in}\left(\{x+\Delta x/2,\Delta z/2\},\mathbf{r}_{\rm m},f_0\right)|^2$$

with $A=\pi\overline{\delta z}_0\langle|\gamma|^2\rangle/2$. At last, given the elongated shape of the input focal spot along the $z-$direction, its energy distribution along z can be assumed as symmetric with respect to the focal plane: $|\overline{H}_{\rm out}\left(\{x-\Delta x/2,-\Delta z/2\},\mathbf{r}_{\rm m},f_0\right)|^2=|\overline{H}_{\rm out}\left(\{x-\Delta x/2,\Delta z/2\},\mathbf{r}_{\rm m},f_0\right)|^2$. It leads to the expression of the CMP intensity profile provided in (22) of the accompanying paper.

S3. THEORETICAL EXPRESSION OF THE COHERENCE FACTOR C

The coherence factor is a measure of the spatial correlation of the field back-scattered by a distribution of scatterers in response to a focused insonification. It corresponds to the ratio between the coherent intensity of the back-scattered echoes recorded by the array and their incoherent intensity. Under a matrix formalism, it can be expressed using the distortion matrix $\mathbf{D}$ introduced in Ref. [11]. When defined in the transducer basis [16], each column of $\mathbf{D}_{u\mathbf{r}}$ contains the realigned wave-fronts $D(u_{\rm out},\mathbf{r}_{\rm in})$ measured by each transducer at the ballistic time for each input focusing beam at $\mathbf{r}_{\rm in}$. A coherence factor can then be deduced as follows:

$$C(\mathbf{r}_{\rm in})=\frac{\left|\langle D(u_{\rm out},\mathbf{r}_{\rm in})\rangle_{u_{\rm out}}\right|^2}{\left\langle|D(u_{\rm out},\mathbf{r}_{\rm in})|^2\right\rangle_{u_{\rm out}}} \tag{S10}$$

To smooth out the fluctuations of this parameter induced by the random reflectivity of the medium, a spatial average over a few resolution cells should be performed, such that

$$C_{\rm av}(\mathbf{r})=\langle W_L(\mathbf{r}_{\rm in}-\mathbf{r})C(\mathbf{r}_{\rm in})\rangle_{\mathbf{r}_{\rm in}} \tag{S11}$$

with $W_L(\mathbf{r})$ the spatial window defined in (15).

Under an isoplanatic hypothesis, the coherence factor can be expressed as follows [4, 28, 29]

$$C(\mathbf{r})=\int\rho_{\rm out}(\Delta u,\mathbf{r})\rho_{\rm in}(\Delta u,\mathbf{r})d\Delta u \tag{S12}$$

where $\rho_{\rm in/out}$ is the autocorrelation function of the aberration transmittance $\tilde{H}_{\rm in/out}$ in the transducer basis, such that

$$\rho_{\rm in/out}(\Delta u,\mathbf{r})=\left[\tilde{H}_{\rm in/out}\overset{u}{*}\tilde{H}_{\rm in/out}\right](\Delta u,\mathbf{r}) \tag{S13}$$

where the symbol $*$ stands for the correlation product. In the Fresnel approximation, $\tilde{H}_{\rm in/out}$ is the Fourier transform of the input/output PSF $\overline{H}_{\rm in/out}$ [16]:

$$\tilde{H}_{\rm in/out}(u,\mathbf{r})=\int\overline{H}_{\rm in/out}(\Delta x,\mathbf{r})e^{-ik_cu\Delta x/(2z)}d\Delta x \tag{S14}$$

where k_c is the wave number at the central frequency. The coherence factor can then be shown to probe the incoherent input-output PSF at $\Delta x=0$:

$$C(\mathbf{r})=\left[|\overline{H}_{\rm in}|^2\overset{\Delta x}{\circledast}|\overline{H}_{\rm out}|^2\right](\Delta x=0,\mathbf{r}). \tag{S15}$$

[1] L. M. Hinkelman, T. L. Szabo, and R. C. Waag, J. Acoust. Soc. Am. **101**, 2365 (1997).

[2] F. A. Duck, Physical properties of tissue: A comprehensive reference book , 73 (1990).

[3] L. M. Hinkelman, T. D. Mast, L. A. Metlay, and R. C. Waag, J. Acoust. Soc. Am. **104**, 3635 (1998).

[4] R. Mallart and M. Fink, J. Acoust. Soc. Am. **96**, 3721 (1994).

[5] K. W. Hollman, K. W. Rigby, and M. O'Donnell, in *1999 IEEE Ultrasonics Symposium. Proceedings. International Symposium*, Vol. 2 (1999) pp. 1257–1260.

[6] W. Lambert, L. A. Cobus, M. Couade, M. Fink, and A. Aubry, Phys. Rev. X **10**, 021048 (2020).

[7] R. Ali and J. J. Dahl, in *2018 IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2018) pp. 1–4.

[8] M. Jaeger, E. Robinson, H. Günhan Akarçay, and M. Frenz, Phys. Med. Biol. **60**, 4497 (2015).

[9] R. Rau, D. Schweizer, V. Vishnevskiy, and O. Goksel, in *IEEE Int. Ultrason. Symp.* (IEEE, Glasgow, 2019) pp. 2003–2006.

[10] H. Bendjador, T. Deffieux, and M. Tanter, IEEE Trans. Med. Imag. **39**, 3100 (2020).

[11] W. Lambert, L. A. Cobus, T. Frappart, M. Fink, and A. Aubry, Proc. Natl. Acad. Sci. USA **117**, 14645 (2020).

[12] A. Badon, D. Li, G. Lerosey, A. C. Boccara, M. Fink, and A. Aubry, Sci. Adv. **2**, e1600370 (2016).

[13] A. Badon, V. Barolle, K. Irsch, A. Boccara, M. Fink, and A. Aubry, Sci. Adv. **6**, eaay7170 (2020).

[14] T. Blondel, J. Chaput, A. Derode, M. Campillo, and A. Aubry, J. Geophys. Res.: Solid Earth **123**, 10936 (2018).

[15] R. Touma, R. Blondel, A. Derode, M. Campillo, and A. Aubry, arXiv: 2008.01608 (2020).

[16] W. Lambert, L. C. Cobus, M. Fink, and A. Aubry, arXiv: 2103.02036 (2021).

[17] A. Aubry and A. Derode, J. Appl. Phys. **106**, 044903 (2009).

[18] A. Aubry, A. Derode, and F. Padilla, Appl. Phys. Lett. **92**, 124101 (2008).

[19] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink, IEEE Trans. Ultrason., Ferroelectr., Freq. Control **56**, 489 (2009).

[20] K. Watanabe, *Integral transform techniques for Green's functions* (Springer, Cham, Switzerland, 2014) Chap. 2.

[21] C. Passmann and H. Ermert, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 545 (1996).

[22] M.-H. Bae and M.-K. Jeong, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 1510 (2000).

[23] M. Born and E. Wolf, *Principles of optics (Seventh edition)* (Cambridge University Press, Cambridge, 2003).

[24] C. Moran, N. Bush, and J. Bamber, Ultrasound Medicine Biol. **21**, 1177 (1995).

[25] G. Montaldo, M. Tanter, and M. Fink, Phys. Rev. Lett. **106**, 054301 (2011).

[26] M. A. Lediju, G. E. Trahey, B. C. Byram, and J. J. Dahl, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **58**, 1377 (2011).

[27] J. J. Dahl, D. Hyun, Y. Li, M. Jakovljevic, M. A. Bell, W. J. Long, N. Bottenus, V. Kakkad, and G. E. Trahey, in *2017 IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2017) pp. 1–10.

[28] J.-L. Robert and M. Fink, J. Acoust. Soc. Am. **123**, 866 (2008).

[29] J.-L. Robert, *Evaluation of Green's functions in complex media by decomposition of the Time Reversal Operator: Application to Medical Imaging and aberration correction*, Ph.D. thesis, Universite Paris 7 - Denis Diderot (2007).