Enhancing Multichannel Laser-Doppler Vibrometry Signals with Application to (Carotid-Femoral) Pulse Transit Time Estimation

Simeon Beeckman[1], Yanlu Li[3], Soren Aasmul[4], Roel Baets[3], Pierre Boutouyrie[5], Patrick Segers[1] and Nilesh Madhu[2]

***Abstract*— Pulse-wave velocity (PWV) can be used to quantify arterial stiffness, allowing for a diagnosis of this condition. Multi-beam laser-doppler vibrometry offers a cheap, non-invasive and user-friendly alternative to measuring PWV, and its feasibility has been previously demonstrated in the H2020 project CARDIS. The two handpieces of the prototype CARDIS device measure skin displacement above main arteries at two different sites, yielding an estimate of the pulse-transit time (PTT) and, consequently, PWV. The presence of multiple beams (channels) on each handpiece can be used to enhance the underlying signal, improving the quality of the signal for PTT estimation and further analysis. We propose two methods for multi-channel LDV data processing: beamforming and beamforming-driven ICA. Beamforming is done by an SNR-weighted linear combination of the time-aligned channels, where the SNR is blindly estimated from the signal statistics. ICA uses the beamformer to resolve its inherent permutation and scale ambiguities. Both methods yield a single enhanced signal at each handpiece, where spurious peaks in the individual channels as well as stochastic noise are well suppressed in the output. Using the enhanced signals yields individual PTT estimates with a low spread compared to the baseline approach. While the enhancement is introduced in the context of PTT estimation, the approaches can be used to enhance signals in other biomedical applications of multi-channel LDV as well.**

I. INTRODUCTION

The large arteries, and specifically the aorta, play a central role in the blood circulation. Their structure allows the vessel wall to distend during heart contraction, storing elastic energy, which is used during the consequent relaxation to drive the blood flow [1]. This way, a near-continuous flow is assured at the smaller arteries that provide organ perfusion. This 'buffer' function [2] deteriorates significantly when these large arteries stiffen, leading to poorer blood flow and consequent organ damage – especially in low resistance organs such as the heart itself [3], [4], [5].

A metric that allows for quantifying arterial stiffening is the pulse-wave velocity (PWV) [5], [6]. Especially carotid-femoral PWV has been studied and showed a significant relationship with arterial stiffness. The speed of the pulse-wave induced by heart-contraction increases with arterial stiffness. If we measure the arrival time of the pulse-wave at two points on the arterial tract, separated by a distance dx, it holds that $\text{PWV} = \frac{dx}{dt}$, where dt is the delay in the pulse arrival time across the two points. We refer to dt as the pulse-transit time (PTT).

Aside from the current state-of-the-art measurement methods, such as applanation tonometry and ultrasound [7], [6], an alternative approach based on laser-doppler vibrometry (LDV) [8], [9], [10] has also been applied to the measuring of PWV. This cheaper, more user-friendly method measures skin displacement above large arteries such as the carotid and femoral arteries, based on which the pulse arrival time and PTT can be estimated. A feasibility study was previously conducted using industrial LDV devices [8]. In the scope of the H2020 CARDIS project, a first prototype was constructed with two handpieces, with which skin displacement could be simultaneously measured at two different locations. Each handpiece captured data using six laser *beams* with a wavelength of 1550nm, where the inter-beam distance was 5mm. The displacement signals obtained were differentiated twice, yielding acceleration. The time-point of the arrival of the pulse wave was then detected via the point of maximum acceleration.

Several datasets were obtained with this prototype [11]. However, despite the availability of multi-channel data on both handpieces, PTT and PWV estimates were based on a laborious combination of the individual channels across both handpieces.

Since the underlying signal captured by all the channels in a single handpiece is the same, *linearly* combining the channels could yield an output with a better signal-to-noise ratio. Such methods of combining signals from spatially separated sensors to enhance a desired (target) signal fall under the umbrella of *beamforming* [12], [13]. A beamformer is essentially a spatial filter that, like its temporal/spectral counterpart, combines signals captured by spatially distributed sensors to extract signals coming from desired *directions* while suppressing interference and noise from unwanted ones. Since the original problem in the time-domain is a convolutive mixing, beamforming is typically applied in the transform (Fourier) domain, where the signal model reduces to an *instantaneous* mixture with complex coefficients at each frequency. This simplifies the subsequent enhancement as it simplifies to a straightforward matrix multiplication instead of a deconvolution. Such methods

*Work done within the scope of the H2020-funded project: InSiDe.

[1]Simeon Beeckman and Patrick Segers are with IBiTech-bioMMeda, Ghent University. simeon.beeckman@ugent.be

[2]Nilesh Madhu is with IDLab, Ghent University - imec. nilesh.madhu@ugent.be

[3]Yanlu Li and Roel Baets are with Photonics Research Group & Center for Nano- and Biophotonics, Ghent University-imec, Ghent, Belgium. yanlu.li@ugent.be & roel.baets@ugent.be

[4]Soren Aasmul is with Medtronic Bakken Research Center, Maastricht, The Netherlands soren.aasmul@medtronic.com

[5]Pierre Boutouyrie is with NSERM U970, Université de Paris, Assistance Publique Hôpitaux de Paris, Paris, France pierre.boutouyrie@aphp.fr

have also been applied to electroencephalography (EEG) data e.g. [14], where an instantaneous mixing model is directly applicable in time domain.

Unlike conventional beamforming or the methods applied to EEG signal processing, however, the signal model for LDV captures cannot be straightforwardly transformed into an instantaneous mixing model in the frequency or the time domain. In this contribution, therefore, we investigate how to adapt and apply beamforming to multichannel LDV data. We also show how to extend the framework to incorporate more sophisticated approaches such as independent component analysis (ICA). While the approach will be demonstrated in the context of carotid-femoral PTT estimation, the underlying signal model and concepts can be applied to a wide range of applications using multichannel LDV - which is being explored in the scope of the follow-up H2020 project InSiDe.

The paper is structured as follows: we first present the CARDIS data and the 'brute-force' baseline PTT estimation method. Next, we present the signal model for the multichannel LDV data. Based on this model we first adapt the beamforming strategy to perform a signal-to-noise ratio (SNR)-weighted averaging of the channels. Lastly, we show how, with the adapted signal model, ICA can also be applied and how the permutation and scaling ambiguity associated with ICA can be resolved. The approaches are then compared in terms of the quality of the PTT estimates obtained from the enhanced signals. It will be demonstrated that individual PTT estimates obtained on the beamformed or ICA processed signals are more reliable (lower variance) than that obtained from the baseline method. Further, the SNR estimates computed in the course of the beamforming provide valuable, additional information regarding the reliability and quality of different segments of the signal - opening up new possibilites for such signal analysis.

II. Materials & Baseline Method

A. The CARDIS database

The CARDIS device consists of two handpieces (HP) - which we will refer to as HP1 and HP2. While the reader is referred to [15], [16] for the details, the prototype is illustrated in Fig. 1, along with typically captured signal traces, for convenience. Note the six channels per handpiece, which capture the skin displacement. To guarantee that sufficient light is reflected by the skin back to the device, application of retro-reflective tape on the measurement site was required. Finally, to ensure stability of positioning and an optimized focus distance during the measurements, a spacer is included in the build.

Data can be gathered simultaneously with both handpieces, each of which is located at a separate measurement site – such as above the carotid and femoral artery to get a carotid-femoral PWV estimation, or on measurement sites that are $25-50$mm apart, to measure local pulse wave propagation in e.g. the carotid artery. Note that while the developed approaches are illustrated on the carotid-femoral setup, they are equally applicable to other setups as well.

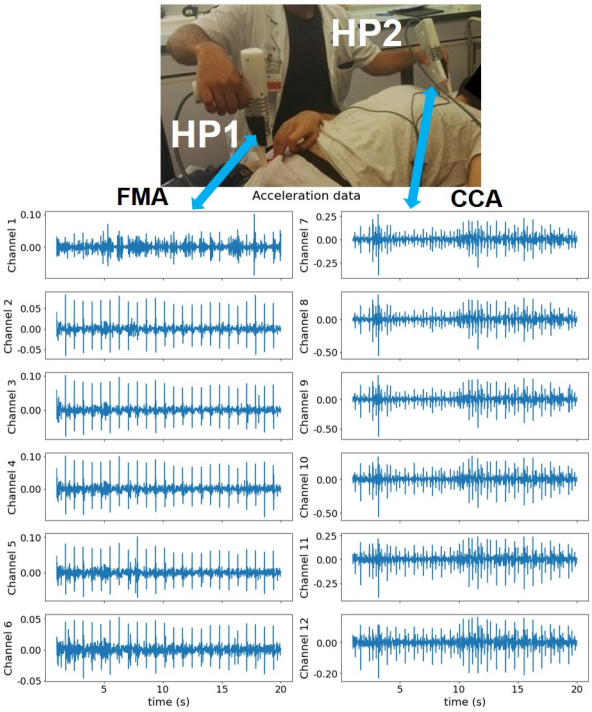

Fig. 1. The CARDIS device being used during a carotid-femoral measurement. Handpiece 1 measures skin displacement above the femoral artery and handpiece 2 above the carotid artery. 20-second recordings for the 12 beams are displayed. Channel 1-6 correspond with handpiece 1 and channel 7-12 to handpiece 2.

The carotid-femoral LDV-data used in this analysis were acquired in a clinical feasibility study at the Hôpital Européen Georges Pompidou (HEGP) in Paris, France. Data was gathered from 100 subjects with varying ages (19-85), sex, BMI, and history with cardiovascular risk-factors and illnesses (from mild to stage three hypertension) [11]. For every subject, four to five sets of measurements were conducted, resulting in 410 datasets for the carotid-femoral database. Each measurement contained six LDV signals per handpiece, yielding skin-displacement data sampled at $f_s = 10$kHz. This was passed through a linear-phase low-pass filter with a cut-off frequency of 30 Hz to suppress high-frequency noise. Filtering was applied in a zero-phase manner. The signal was then differentiated twice to yield acceleration data – which is a more robust feature as the differentiation removes any drift in the displacement signal.

B. Brute-force PTT estimation

To validate the applicability of LDV for PWV estimation, the carotid-femoral PTT estimate was calculated and compared with ground truth data, which was obtained using applanation tonometry. As mentioned previously, PTT is defined as the time delay between the arrival of the pulse wave at the carotid artery and the femoral artery. It should be noted that, since the arterial pathway from the heart to the carotid artery is shorter than that from heart to femoral, the pulse wave should always be detected first at the carotid. The time of arrival of the pulse at each measurement point is taken as the instant of maximum acceleration for a given heartbeat. Demarcation of the pulse arrival time and the PTT

computation is illustrated in Fig. 2 where one channel each from HP1 and HP2 are considered. By pooling the estimates across the 36 possible channel combinations between HP1 and HP2, more robust PTT estimates can be obtained.

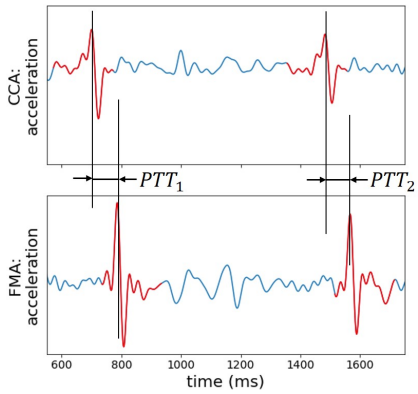

Fig. 2. Example matching of simultaneously measured carotid and femoral LDV-traces. Two corresponding heartbeats from which pulse-transit times can be calculated are indicated in red. Y-axis is not labelled because the scale is arbitrary.

As the above example illustrates, computing the PTT consists of first identifying the heart cycles in the acceleration signals of the carotid-femoral channel-pair considered, followed by identifying the pulse arrival time in each channel and for each cycle. However, the problem is that either or both channels may not pick up the pulse in each beat - because of insufficient reflection of the beam or movement of the handpiece or insufficient skin displacement. Further, as the heart rate can also change during a measurement, an online estimation of the heart rate (and, thus, the location of a cycle) becomes necessary. Thus, the following procedure was adopted for estimating the PTT from a given recording.

First, a beat-detection algorithm based on *template matching* [17], [18] was applied to each of the 12 traces (we term the acceleration signal of a channel as a *trace*). Template matching is a pattern recognition technique which essentially consists of computing the normalised cross-correlation between a so-called template (here, the characteristic waveform of the acceleration signal when the pulse traverses the measurement point) and the trace at different time-lags. For time-lags where the segment of the trace 'matches' well with the template, the normalised cross-correlation will be high (ideally ≈ 1) – indicating the presence of the pulse-wave. An empirically selected threshold of 0.7 was selected and at time instants where this value was exceeded, it was be assumed that a beat was detected. As the acceleration signal demonstrates different characteristic wave pattern at the carotid (HP1) and femoral (HP2) measurement points, separate templates were used each case, and were obtained offline by an ensemble average across all recordings. The length of the carotid template was 200ms and that of the femoral was 500ms (please see [19] for details).

Next, *matching* beat-pairs were identified using the carotid as reference (typically more carotid beats are detected compared to femoral). When a pulse is detected on *any* of the carotid channels, all femoral channels are examined for pulses within a time-window of 200ms of this pulse (two matching beat-pairs are indicated in Fig. 2). The range of 200ms was chosen to accommodate a wide range of PWVs while minimising the risk that a *subsequent* beat on the femoral is matched (i.e., carotid detects beat P and is matched to beat $P+1$ on the femoral). Since a 200ms beat period corresponds to the unlikely heart rate of 300bpm, such confusion is avoided. Finally, PTTs were calculated from *each* matching beat pair in the 36 trace combinations. The *median* value of these was taken as the final PTT estimate.

In addition to being laborious (template matching applied to 12 channels, exhaustive search for a matching beat in the femoral channels for each beat found on a carotid channel,...) this approach implicitly includes data from noisier channels as well. With no way of indicating which estimate comes from a good or from a bad channel, this leads to a wide spread of the results. Hence we investigate the linear combination of all channels at a handpiece to generate a single output signal with improved SNR. This should reduce the spread of the PTT estimates and offer potential to reduce the computational complexity as well.

III. Beamforming for PTT estimation

A. Signal model

For any handpiece, the acceleration signal at a channel $m \in \{1,2,\ldots,6\}$ is modelled as:

$$x_m(n) = s_m(n) + v_m(n)\,, \tag{1}$$

where $s_m(n)$ is the underlying target signal *in channel m* and $v_m(n)$ is the noise in that channel. We further assume:

$$s_m(n) = \alpha_m s(n-\tau_m)\,, \tag{2}$$

i.e., the target signal is received at each channel with a channel-dependent scale factor α_m and delay τ_m.

The channel-dependent delay in (1) makes it difficult to form the enhanced output ($y(n)$) by a weighted linear combination of the form:

$$y(n) = \sum_m w_m x_m(n)\,, \tag{3}$$

while ensuring a constructive addition of the desired component $s(n)$. The general solution requires the estimation of optimal individual *filters* $w_m(n)$ for each channel – which is not a straightforward problem (see, e.g., [20]). However, the model of (2) allows for a simpler alternative: by time-aligning the target components $s_m(n)$, (1) reduces to:

$$x'_m(n) = \alpha_m s(n) + v'_m(n)\,, \tag{4}$$

allowing for the application of (3). We address, next, the blind estimation of the time-delays τ_m and the optimal weighting factor w_m. Since the underlying signal $s(n)$ is essentially unknown, a blind estimation of τ_m and α_m is always subject to an offset and a scale ambiguity respectively - i.e., any solution of the kind $\widehat{\tau_m} = \tau_m + T$ and $\widehat{\alpha_m} = \mathscr{C}\alpha_m$, where T and $\mathscr{C}$ are constants, is acceptable – the constant values being subsumed into the definition of $s(n)$. The ambiguities vanish, however, if we estimate the scale and

delay with respect to the signal component in a *reference* channel m_{ref}. Thus, reinterpreting (1) and (2) with respect to a reference channel m_{ref} we obtain:

$$\begin{aligned} x_m(n) &= \alpha_m s(n-\tau_m) + v_m(n) \quad \forall m \neq m_{\text{ref}} \\ x_{m_{\text{ref}}}(n) &= s(n) + v_{m_{\text{ref}}}(n) \end{aligned} \tag{5}$$

where α_m and τ_m are now with respect to the signal component in the reference channel.

In general, the reference channel can be arbitrarily chosen for each handpiece – the middle channels ($m = 3,4$) being a logical choice. Alternatively, with the help of overall-quality estimated per channel over the whole recording – as in [19] – the 'best' channel (on average) may be chosen. In the following we denote this generically as m_{ref}.

B. Time alignment of the $s_m(n)$

As the signal component in channel m can either be advanced or delayed with respect to the signal component at channel m_{ref}, the reference signal $x_{m_{\text{ref}}}(n)$ is first shifted by a group delay of $D > \max\{\tau_m\}$ samples. This ensures that all other channels are shifted *causally* for time alignment. Thus, the reference signal becomes:

$$\tilde{x}_{m_{\text{ref}}}(n) = s(n-D) + \tilde{v}_{m_{\text{ref}}}(n)\,. \tag{6}$$

It is easy to see that delaying each $x_m(n)$ $(m \neq m_{\text{ref}})$ by $T_m = D - \tau_m$ ensures that the signal component is aligned with the reference. By cross-correlating $x_m(n)$ with $\tilde{x}_{m_{\text{ref}}}(n)$ an estimate the time-delay can be obtained [21], [22]. The *integer* part of the delay corresponds to the time-lag at which the cross-correlation peak is observed. The delay estimate can be further refined to account for *fractional* shifts. This can be done by a simple three-point parabolic fit around the observed cross-correlation peak, or by more sophisticated methods (see, e.g., [23]). Having estimated T_m, generic delay filters $h_m(n)$ can be obtained by (truncated) sinc functions [24] of order $L \geq 2D$:

$$h_m(n) = \text{sinc}\big(n - T_m\big)\,. \tag{7}$$

This formulation of the delay filter allows us to account for fractional sample shifts. It is easy to see that the filter reduces to a shifted Kronecker's delta function when T_m is an integer – as we expect.

C. Compensating for α_m

By appropriately compensating α_m, we can express the signal at *each* channel as:

$$\tilde{x}_m(n) = s(n-D) + \tilde{v}_m(n)\,, \tag{8}$$

which, as we subsequently show, allows for an intuitive definition of the weighting factors w_m in (3). We exploit the observation that $s(n)$ is characterised by high-energy regions corresponding to the traversal of a pulse. Thus, comparing energies of $x_m(n)$ to $x_{m_{\text{ref}}}(n)$ (or their time-aligned versions) at these regions can yield a reasonable idea of α_m.

While identification of high-energy regions can be done using template matching (as in Sec. II-B), this requires extra computation and the availability of templates. Therefore, we present a more generic method based on the *statistics* of the signal energy – so the approach is also applicable where prior knowledge of templates are unavailable.

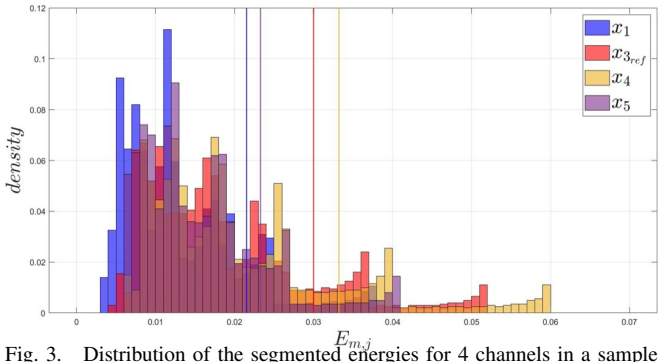

Fig. 3. Distribution of the segmented energies for 4 channels in a sample dataset. The 85$^{\text{th}}$ percentiles are displayed by vertical lines, and in the colour chosen for the corresponding channel.

We segment $x_{m_{\text{ref}}}(n)$ and $x_m(n)$ (or their time-aligned versions) into J non-overlapping segments of length N samples, where N is roughly the length of a pulse waveform ($\approx$ 200ms). The energy $E_{m,j}$ (resp. $E_{m_{\text{ref}},j}$) is obtained as:

$$E_{m,j} = \sum_{n=0}^{N-1} \big(x_m(n + jN)\big)^2\,. \tag{9}$$

Subsequently, as we are interested in the high-energy segments, the 85$^{\text{th}}$ percentile of the energies was calculated. This threshold was empirically found to be a good balance between accounting for the variation of the received pulse energy in the trace and not being biased by the low-energy segments which mainly contain channel-dependent noise. Based on this, α_m can be estimated as:

$$\widehat{\alpha}_m = \sqrt{\frac{\text{perc}\big(\{E_{m,j}\}, 85\big)}{\text{perc}\big(\{E_{m_{\text{ref}},j}\}, 85\big)}} \tag{10}$$

The distribution of the energies are illustrated for an example dataset in Fig. 3 and the 85$^{\text{th}}$ percentiles are shown as vertical lines in the colour of the corresponding distribution.

Based on $\widehat{\alpha}_m$ above and $h_m(n)$ from (7), we modify the signal of each channel as:

$$\tilde{x}_m(n) = \frac{1}{\widehat{\alpha}_m} h_m(n) * x_m(n) \;\approx s(n-D) + \tilde{v}_m(n)\,, \tag{11}$$

where $*$ represents the discrete-time convolution operator.

D. Estimating w_m and the enhanced signal $y(n)$

Using (5) and (11) and stacking the signals into an $M-$dimensional column vector for each time-instant yields:

$$\widetilde{\mathbf{x}}(n) = \mathbf{1}\, s(n-D) + \widetilde{\mathbf{v}}(n)\,, \tag{12}$$

where $\widetilde{\mathbf{x}}(n) = \big[\tilde{x}_1(n),\quad \tilde{x}_2(n),\quad \ldots,\quad \tilde{x}_M(n)\big]^T$, $\mathbf{1}$ is an $M \times 1$ vector of ones and $\widetilde{\mathbf{v}}(n)$ is similarly defined to $\widetilde{\mathbf{x}}(n)$. This model now allows the computation of the enhanced signal as in (3) and an intuitive way to define the weights would be to make them proportional to the SNR at each channel.

However, computing a *single* set of weights over the whole recording would not be optimal, since the quality of the underlying received signal is time-variant in each channel (as can be seen in Fig. 1). Thus, we propose to segment $\tilde{x}_m(n)\ \forall m$ into $\widetilde{J}$ *overlapping* segments of length $\widetilde{N} \approx 1$s, and derive optimal weights for each segment independently. A 50% overlap is considered to avoid edge effects during the subsequent weighted combination.

To estimate the SNR of each channel m in segment $\widetilde{j}$, we expand upon the idea in Sec. III-C. We first partition segment $\widetilde{j}$ into $\widetilde{K} = 2000$ sub-segments of $\widetilde{N}/\widetilde{K} = 5$ samples and compute the energy $E_{m,\widetilde{j},\widetilde{k}}$ of every sub-segment $\widetilde{k}$. Indices $(m,\widetilde{k})$ where $E_{m,\widetilde{j},\widetilde{k}} > 85^{th}$ percentile (for channel m and segment $\widetilde{j}$) are then extracted. This is done separately for all channels. Next, we compare the selected $\widetilde{k}$ across all M channels and only retain those sub-segment indices that occur in a *majority* of the channels. The signal energy in segment $\widetilde{j}$ for each channel (denoted as $P_s(m,\widetilde{j})$) is finally computed as the *average* of the energies $E_{m,\widetilde{j},\widetilde{k}}$ in the *retained* sub-segments $\widetilde{k}$. Since the desired signal $s(n)$ is time-aligned across the channels, when the energy of a particular sub-segment $\widetilde{k}$ lies beyond the 85^{th} percentile in multiple channels, it is likely that this is not a spurious energy peak, but that occasioned by $s(n)$. This selection thus avoids noise sub-segments in a channel from biasing the signal energy estimate.

A similar procedure is applied to estimate the noise energy $P_v(m,\widetilde{j})$, but here we select segment indices where $E_{m,\widetilde{j},\widetilde{k}} < 15^{th}$ percentile. Comparison of sub-segment indices across channels is not required in this case.

The final SNR and the resultant combination weights $w_m(\widetilde{j})$ are then obtained, for segment $\widetilde{j}$, as:

$$\mathrm{SNR}(m,\widetilde{j}) = \frac{P_s(m,\widetilde{j})}{P_v(m,\widetilde{j})} \tag{13}$$

$$w_m(\widetilde{j}) = \frac{\mathrm{SNR}(m,\widetilde{j})}{\sum_{m'=1}^{M} \mathrm{SNR}(m',\widetilde{j})}\,. \tag{14}$$

The beamformed signal of segment $\widetilde{j}$ is subsequently computed using (3) as:

$$y_{\mathrm{BF}}(n,\widetilde{j}) = \mathbf{w}^T(\widetilde{j})\mathbf{x}(n,\widetilde{j})\,. \tag{15}$$

The enhanced signal $y_{\mathrm{BF}}(n)$ is reconstructed from the segments $y_{\mathrm{BF}}(n,\widetilde{j})$ by the overlap-add method, after the application of an $\widetilde{N}$-point von Hann window. This tapered window reduces edge effects at the boundaries of the overlapping segments and, at 50% overlap, allows perfect reconstruction during overlap-add.

E. Beamformer-driven ICA

Given an *instantaneous* multi-channel mixture of the form:

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n)\,, \tag{16}$$

where $\mathbf{x}(n)$ is an $M-$dimensional vector of *observations* at time-instant n and $\mathbf{s}(n)$ is a $Q(\leq M)$-dimensional vector of underlying source activity. $\mathbf{A}$ is an $M \times Q$ *mixing* matrix yielding a linear combination of the sources in the observed signal. Under the assumption that the $s_q(n)$ are statistically *independent*, ICA [20], [25] can be applied to yield a $Q \times M$ *demixing* matrix $\mathbf{W}$. Applied to $\mathbf{x}(n)$ this yields: $\mathbf{y}(n) = \mathbf{W}\mathbf{x}(n)$, with statistically independent $y_{q'}(n)$. For non-Gaussian $s_q(n)$, such maximisation of statistical independence in the output implies that $y_{q'}(n)$ recover $s_q(n)$, subject to a *scale* ambiguity and arbitrary order of outputs (*permutation* ambiguity). Robustly resolving these ambiguities is an enduring challenge, usually requiring extra knowledge!

Our signal model after *time-alignment* may be written as (slightly abusing the notation in (12)):

$$\widetilde{\mathbf{x}}(n) = \boldsymbol{\alpha} s(n-D) + \widetilde{\mathbf{v}}(n)\,, \tag{17}$$

where $\boldsymbol{\alpha} = \left[\alpha_1, \alpha_2, \ldots, \alpha_M\right]^T$. This can be straightforwardly massaged into the standard ICA form of (16):

$$\widetilde{\mathbf{x}}(n) = \mathbf{A}\widetilde{\mathbf{s}}(n)\,, \tag{18}$$

with the noise subsumed into the definition of $\widetilde{\mathbf{s}}(n)$. Application of ICA then gives outputs $\mathbf{y}_{\mathrm{ICA}}(n)$, where $y_{\mathrm{ICA},q}(n)$ corresponds to $s(n-D)$ for some output channel q, and subject to an unknown scale. A simple correlation of each $y_{\mathrm{ICA},q}(n)$ with $y_{\mathrm{BF}}(n)$ can then be used to identify the correct output channel q'– the one with the maximum correlation with $y_{\mathrm{BF}}(n)$. It is straightforward, then, to infer the scale as:

$$\alpha_{\mathrm{ICA}} = \frac{\mathbb{E}\left\{y_{\mathrm{BF}}^2(n)\right\}}{\mathbb{E}\left\{y_{\mathrm{BF}}(n) y_{\mathrm{ICA},q'}(n)\right\}}\,, \tag{19}$$

where $\mathbb{E}\{\cdot\}$ is the expectation operator.

Consistent with Sec. III-D, ICA is similarly applied segment-wise and the final output obtained by overlap-add.

IV. Experimental results and discussion

We start with a qualitative evaluation on a sample dataset: Fig. 4 depicts the captured signal on four channels of a handpiece, after scale compensation and time-alignment. To illustrate the weight computation of (14), three segments $(\widetilde{j})$ are highlighted in each channel, and the corresponding weights $(w_m(\widetilde{j}))$ assigned to those segments are indicated. As can be seen, when using these four channels for beamforming, our blind SNR-estimation approach performs correctly, assigning a very low weight to channel 1 (essentially noise) and giving more-or-less similar weights to the other three channels (of similar quality). Also evident is the time-varying SNR in each channel (indicated by the changing weights) – indirectly validating the wisdom of segment-wise processing. This is further highlighted in Fig. 5 which shows, for the same sample dataset as in Fig. 4, a different set of channels as well as the enhanced signals ($y_{\mathrm{BF}}(n)$ and $y_{\mathrm{ICA}}(n)$). Channel 1 is omitted for space reasons (and, as Fig 4 shows, this channel will not contribute to the output). It is instructive to note that despite all channels being of relatively good quality in Fig. 5, spurious peaks occur in individual channels – which are effectively removed in the enhanced signals.

Lastly, we test the benefit of the proposed beamforming and ICA-based enhancement for PTT estimation on a subset

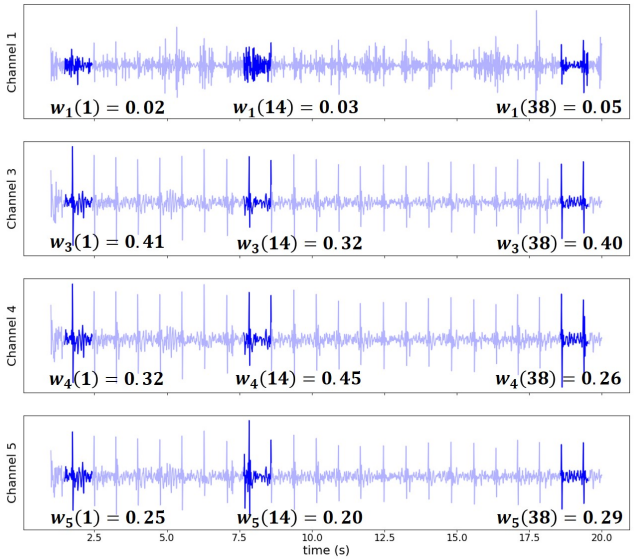

Fig. 4. Weights w_m computed for four input channels. Three representative segments are highlighted. The correlation between the segment weights and the SNR is evident – indicating a good (blind) SNR estimation.

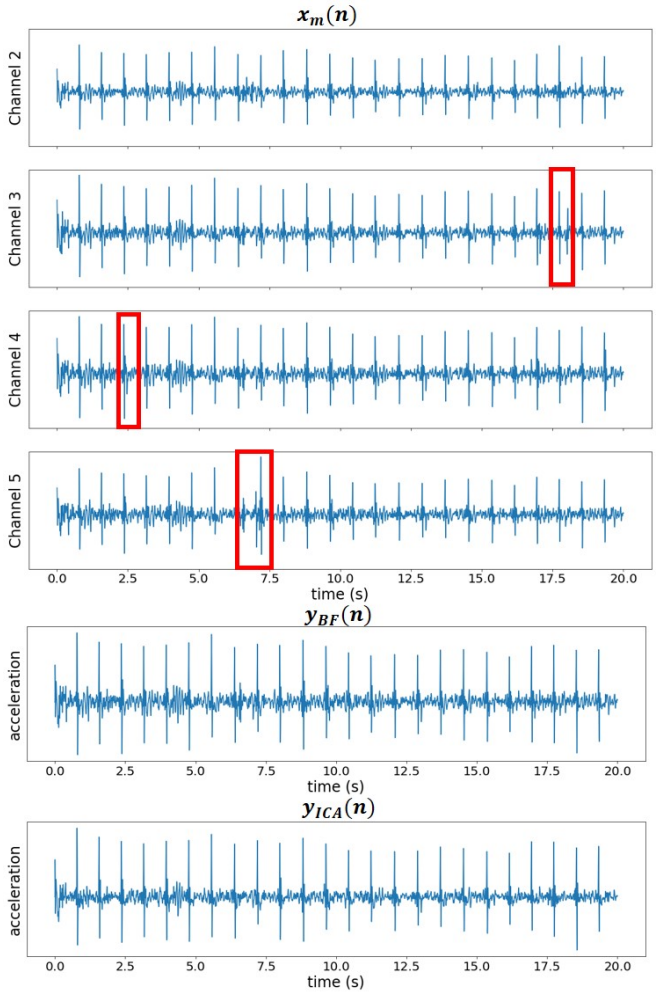

Fig. 5. A selection of four input signals from the femoral recording, along with the beamformer & ICA enhanced outputs, for an example dataset. The segment-wise operation with time-varying weights effectively suppresses spurious pulses in individual channels (highlighted). The ICA also yields a cleaner signal compared to the beamformer.

of the CARDIS carotid-femoral database. Only those datasets were taken that had at least one good quality channel per handpiece [19] (to allow reasonable PTT estimation with the baseline). A total of 54 datasets passed this threshold. For these datasets, the $y_{\mathrm{BF}}(n)$ and $y_{\mathrm{ICA}}(n)$ are obtained for each handpiece. Following this, PTT is estimated on these signals, in a similar manner as described in Sec. II-B. These estimates were compared with those generated by the baseline ('brute-force') and the ground truth. Fig. 6 shows the *error* of each

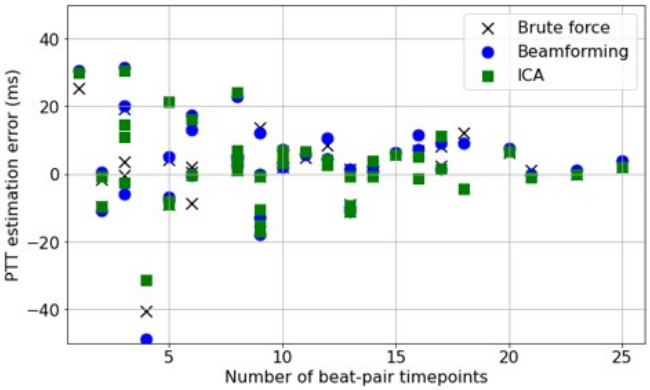

Fig. 6. PTT estimation error for the different methods, compared to the ground-truth. The error is shown in function of the number of individual estimates obtained from matching beat-pairs. As the number of individual estimates increase, their median is closer to the ground truth.

method compared to the ground-truth as a function of the number of PTT estimates (i.e., the number of matching beat-pairs across which the PTT estimate is computed). For all methods the accuracy of the final PTT estimate improves with increasing number of timepoints at which a pulse-transit time estimation could be made. While all methods deviate somewhat from the applanation-tonometry reference values, between brute force, beamforming and ICA, results were mostly similar. This is not wholly unexpected: $y_{\mathrm{BF}}(n)$ and $y_{\mathrm{ICA}}(n)$ are obtained by linearly combining the different channels. Inherently, the brute-force search across all beat-pairs and all channels performs such a linear combination. As the *final* estimate is obtained from the median value in all cases, the results are expected to be rather homogeneous.

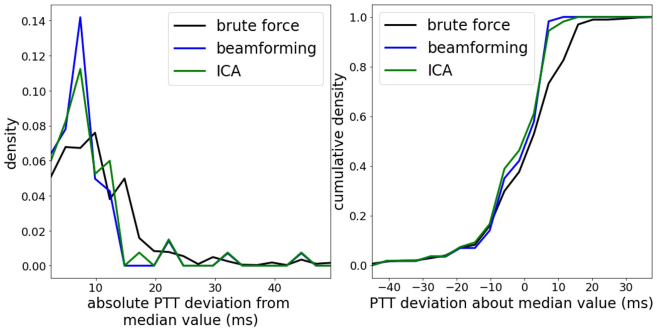

Fig. 7. PTT-estimation distributions calculated via brute force, beamforming and ICA methods for ten example datasets. The median value was subtracted from the estimations.

It is more instructive, therefore, to consider the distribution of the *individual* PTT estimates in all cases. This is depicted

in Fig. 7, where the left sub-plot shows the distribution of the PTT-deviation about the median value and the right sub-plot shows the cumulative density. Both plots indicate a clear reduction of variance in the PTT estimates obtained after beamforming and ICA enhancement, whereas the brute-force method has a larger variance and several outliers. Between beamforming and ICA, for the purpose of PTT estimation, the performance is comparable, with the ICA being *marginally* better.

V. Conclusions

We proposed two methods for the enhancement of multi-channel LDV signals, as applied to the task of pulse-wave velocity estimation. Data from a pilot study indicate that the desired signal component in each channel can be modelled as a scaled and time-shifted version of the true, underlying signal. Exploiting the fact that this underlying signal is characterised by high-energy peaks at instants of pulse traversal, an analysis of statistics of the short-term signal energy allows for an estimation of the scale factor, and correlation analysis yields the necessary time-shift for the signal alignment across the channels. By compensating the scale and time-aligning the signals across all channels of a handpiece, an SNR-weighted linear combination yields the beamformed signal. The SNR in each channel is *blindly* estimated, based on percentile statistics of the short-term signal energies and by cross-validating across channels – increasing the robustness of the estimates.

Experiments demonstrate that the weights assigned during beamforming reflect the signal quality in the channels – validating the SNR estimation. Further, the beamformed output shows a cleaner signal with spurious pulses in the individual channels being well suppressed. Because the ICA does not require an explicit compensation of the individual channel scale factors and is free to derive the optimal weighting in terms of maximising statistical independence, it yields sharper, better formed outputs compared to the beamformer.

Regarding the reliability of individual PTT estimations: the distribution of the estimates has low variance when estimated on the beamformed or ICA-enhanced outputs. In comparison, the estimates from the brute-force method exhibit a larger variance – indicating the influence of noisy and unreliable beat-pairs.

While we have demonstrated the benefit of the proposed approaches in the context of PWV estimation, the underlying ideas for blind SNR estimation and delay- and scale compensation can be more broadly applied to analyse the quality of, or enhance the signals from multidimensional LDV data in biomedical applications.

References

[1] H. Wolinsky and S. Glagov, "A lamellar unit of aortic medial structure and function in mammals," *Circulation Research*, vol. 20, pp. 99–111, 1967.

[2] N. Westerhof, J.-W. Lankhaar, and B. E. Westerhof, "The arterial windkessel," *Medical & biological engineering & computing*, vol. 47, no. 2, pp. 131—141, February 2009.

[3] J. A. Chirinos, P. Segers, T. Hughes, and R. Townsend, "Large-artery stiffness in health and disease: JACC state-of-the-art review," *Journal American College of Cardiology*, vol. 74, no. 9, pp. 1237–1263, 2019.

[4] C. Vlachopoulos, K. Aznaouridis, and C. Stefanadis, "Prediction of cardiovascular events and all-cause mortality with arterial stiffness: A systematic review and meta-analysis," *Journal of the American College of Cardiology*, vol. 55, no. 13, pp. 1318–1327, 2010.

[5] S Laurent, J. Cockcroft, L. Van Bortel, P Boutouyrie, C. Giannattasio, D. Hayoz, B. Pannier, C. Vlachopoulos, I. Wilkinson, and H. Struijker-Boudier, "Expert consensus document on arterial stiffness: methodological issues and clinical applications," *European Heart Journal*, vol. 27, no. 21, pp. 2588–2605, 2006.

[6] P. Segers, E. R. Rietzschel, and J. A. Chirinos, "How to measure arterial stiffness in humans," *Arteriosclerosis, thrombosis, and vascular biology*, vol. 40, no. 5, pp. 1034–1043, 2020.

[7] T. Pereira, C. Correia, and J. Cardoso, "Novel methods for pulse wave velocity measurement," *Journal of medical and biological engineering*, vol. 35, no. 5, pp. 555–565, 2015.

[8] M. De Melis, U. Morbiducci, L. Scalise, E. P. Tomasini, D. Delbeke, R. Baets, L. M. Van Bortel, and P. Segers, "A noncontact approach for the evaluation of large artery stiffness: a preliminary study," *American journal of hypertension*, vol. 21, no. 12, pp. 1280–1283, 2008.

[9] A. Campo and J. Dirckx, "Dual-beam laser doppler vibrometer for measurement of pulse wave velocity in elastic vessels," in *22nd Congress Intl. Commission for Optics: Light for the Development of the World*. International Society for Optics and Photonics, 2011, vol. 8011, p. 80118Y.

[10] A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, P.-H. Lai, and J. W. Rohrbaugh, "Hidden state models for noncontact measurements of the carotid pulse using a laser doppler vibrometer," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 3, pp. 744–753, Mar. 2012.

[11] L Marais, H Khettab, Y. Li, P. Segers, R. Baets, K. Reesink, S. Aasmul, M. De Melis, and P. Boutouyrie, "Measurement of aortic stiffness by laser doppler vibrometry: The CARDIS study," *Journal of Hypertension*, vol. 37, pp. e88, 2019.

[12] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory. Part 4. Optimum Array Processing*, John Wiley & Sons, New York, 2002.

[14] A. Ivannikov, T. Kärkkäinen, T. Ristaniemi, and H. Lyytinen, "Spatially weighted averaging for ERP denoising in EEG data," in *Proc. Intl. Symp. Communication, Control, Signal Proc. (ISCCSP)*, 2010.

[15] Y. Li, L. Marais, H. Khettab, Z. Quan, S. Aasmul, R. Leinders, R. Schüler, P. E Morrissey, S. Greenwald, and P. Segers *,et al*, "Silicon photonics-based laser doppler vibrometer array for carotid-femoral pulse wave velocity (PWV) measurement," *Biomedical Optics Express*, vol. 11, no. 7, pp. 3913–3926, 2020.

[16] Y. Li, P. Segers, J. Dirckx, and R. Baets, "On-chip laser doppler vibrometer for arterial pulse wave velocity measurement," *Biomedical optics express*, vol. 4, no. 7, pp. 1229–1235, 2013.

[17] C. Won-Du and I. Chang-Hwan, "Enhanced Template Matching Using Dynamic Positional Warping for Identification of Specific Patterns in Electroencephalogram," *Journal of Applied Mathematics*, Apr 2014.

[18] C. Jiun-Hung, C. Chu-Song, and C. Yong-Sheng, "Fast algorithm for robust template matching with M-estimators," *IEEE Trans. Signal Processing*, vol. 51, no. 1, Jan 2003.

[19] S. Seoni, S. Beeckman, Y. Li, S. Aasmul, U. Morbiducci, R. Baets, P. Boutouyrie, F. Molinari, N. Madhu, and P. Segers, "Template matching and matrix profile for signal quality assessment of carotid and femoral laser doppler vibrometer signals," *Frontiers in Physiology*, vol. 12, 2022.

[20] A. Hyvärinen, E. Oja, and J. Karhunen, *Independent Component Analysis*, John Wiley & Sons, New York, 2001.

[21] R. M. Gray and L. D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004.

[22] N. Madhu, *Acoustic source localization: Algorithms, applications and extensions to source separation*, Ph.D. thesis, Ruhr-Universität Bochum, 2009.

[23] D. Ventzas and N. Petrellis, "Peak searching algorithms and applications," *Proceedings of the IASTED International Conference on Signal and Image Processing and Applications, SIPA 2011*, 08 2011.

[24] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay - tools for fractional delay filter design," *IEEE Signal Processing Magazine*, vol. 13, no. 1, Jan. 1996.

[25] R. Ranta and N. Madhu, "Reference estimation in EEG: Analysis of equivalent approaches," *IEEE Signal Processing Letters*, vol. 19, no. 1, pp. 12–15, 2012.