A FREQUENCY DOMAIN ADAPTIVE ALGORITHM FOR WAVE SEPARATION

G. De Sanctis and M. van Walstijn

Sonic Arts Research Centre
School of Electronics, Electrical Engineering and Computer Science
Queen's University Belfast
{gdesanctis01/m.vanwalstijn}@qub.ac.uk

ABSTRACT

We propose a frequency domain adaptive algorithm for wave separation in wind instruments. Forward and backward travelling waves are obtained from the signals acquired by two microphones placed along the tube, while the separation filter is adapted from the information given by a third microphone. Working in the frequency domain has a series of advantages, among which are the ease of design of the propagation filter and its differentiation with respect to its parameters.

Although the adaptive algorithm was developed as a first step for the estimation of playing parameters in wind instruments it can also be used, without any modifications, for other applications such as in-air direction of arrival (DOA) estimation. Preliminary results on these applications will also be presented.

1. INTRODUCTION

One of the most appreciated features offered by physical modelling synthesis is the availability of physically meaningful parameters. The idea of exploring all the timbral possibilities offered by a given model by changing its parameters in an intuitive way based on everyday experience is appealing, and it also allows extrapolation beyond real-world limits. This adherence to reality holds not only for construction parameters (i.e. materials, dimensions, etc.) but also for playing parameters (such as the velocity and the position of a violin's bow with respect to the strings, or the state of the toneholes and the blowing pressure in a woodwind instrument), and as a consequence we need to have some in-depth knowledge of the real instrument in order to fully exploit its model. In general it is not feasible to play a physical model using a standard keyboard controller, because all the playing parameters need to be controlled concurrently and not all their combinations will produce a sound. Therefore, unless we have a specifically designed controller that is able to capture all the information needed from the musician's gestures, we are forced to restrict the space of parameters to a "safe" sub-space that will always produce a sound, sometimes at the price of reducing our possibilities. Unfortunately these custom controllers are often quite different from the original instrument so that the musician may not be comfortable with it and this could affect the performance. To avoid this problem we can use a standard instrument and extract the playing parameters from audio specifically taken by a number of sensors. A typical example of this approach is the "guitar-to-MIDI" converter, which uses a special pickup able to acquire the signal of each string separately. If a live performance is not indispensable, the parameters extraction can be done off-line, so that the requirements in terms of efficiency can be relaxed to advantage precision and accuracy. By working on recorded audio we can also make use of multi-pass algorithms and access the acquired signals at any time instant (e.g. by using non-causal filters).

Our work is based on the latter approach. The focus is on wind instruments and in particular on the clarinet, although the idea could in principle be applied to other types of sound generation mechanisms. We chose the clarinet primarily because of the broad availability of literature on physical models for reed instruments. Wind instruments are characterised by a continuous excitation and require the concurrent estimation of the state of both the exciter (the reed) and the resonator (the bore). Despite this added difficulty, the one-dimensionality of the variables acquired and the well established in-duct measurement techniques also contributed to the choice. The extraction of parameters from audio applied to a reed instrument is a plausible solution, since although the state of the keys could be sensed directly (through suitable sensors applied on each tonehole), the same cannot be easily done on the reed. On the other hand, the extraction of lumped reed parameters from audio seems to be a promising direction, as presented in [2]. In order to independently infer the state of both the reed and the bore we need to "break the loop" by estimating forward and backward travelling waves at a suitable section that separates the two blocks [1]. For this purpose the barrel seems to be the most suitable part of the instrument where the microphones can be placed. It

This work has been supported by the EPSRC, grant number EP/D074983/1.

also has the advantage of being easily replaced, avoiding any irreversible modification on the instrument used, thus ensuring the non-invasiveness of the method.

To the authors' knowledge, the only attempt to tackle a similar problem is the work by Guérard and Boutillion [3]. They were able to separate forward and backward travelling waves in a tube by gathering the pressure, measured at the reference section, and a discrete spatial differentiation of it, computed from the pressure measured at a number of points along the tube. Since the discrete derivative is simply a weighted sum of all the outputs from the microphones, it can be easily done in hardware thus requiring only a dual-channel acquisition. This approach however is not suitable for our purpose. Firstly, the use of a discrete derivative sets a theoretical limit for the precision. It also requires at least five microphones [3] to obtain a good approximation of the pressure's derivative, and it would be difficult to place them consistently in a limited space such as the barrel of a clarinet. More importantly, the method is not adaptive and it cannot be adjusted to construction tolerances or changes in the propagation conditions (temperature and humidity fluctuations due to the musician's breath). Also wall losses are not taken into account. For these reasons we decided to investigate a different approach which only requires two microphones, and it can be made adaptive by adding a third one.

The adaptive scheme developed can also be used in other adaptive filtering applications, when *a priori* knowledge on the filter is available. The in-air DOA estimation constitutes an interesting parallel with the main application and it will be introduced at the end of the paper.

2. WAVE SEPARATION

In our study we will assume that the radius of the bore and the maximum frequency of the signal are small enough so that there is only one propagating mode. We can thus use the solution for a one-dimensional plane wave at the position x

$$\begin{array}{l} P(x) = Ae^{\Gamma x} + Be^{-\Gamma x} = P^+(x) + P^-(x) \\ ZU(x) = Ae^{\Gamma x} - Be^{-\Gamma x} = P^+(x) - P^-(x) \end{array} , \quad (1)$$

where P, U, A and B are respectively the pressure, the volume flow and two suitable pressure values, in phasor form[1], while Z is the characteristic impedance and Γ is the propagation function.

The wave separation problem is thus reduced to finding the forward travelling wave $P^+(x)$ at a suitable reference plane $x = 0$, starting from a number of measures of the pressure taken at different points along the tube. The problem can be formulated entirely in the Kirchhoff domain, i.e. using the (continuous time) variables (P, U) and discretising the result or, alternatively, in the wave domain, starting with a discretised version of the propagation filter (i.e. the filter that relates the pressure measured at two points on the x axis).

2.1. Kirchhoff Domain

From equations (1) we can write

$$P(x) = P_0\, cosh(\Gamma x) + ZU_0\, sinh(\Gamma x), \quad (2)$$

where P_0 and U_0 represent respectively the pressure and the volume flow at $x = 0$. We can substitute $x = d$ and $x = -d$ to find the pressure at two points at a distance d on each side of the reference section,

$$\begin{array}{l} P_1 = P_0 cosh(\Gamma d) + ZU_0 sinh(\Gamma d) \\ P_2 = P_0 cosh(\Gamma d) - ZU_0 sinh(\Gamma d) \end{array} .$$

By taking their sum and difference we obtain

$$P_0 = \frac{P_1 + P_2}{2\, cosh(\Gamma d)}, \quad ZU_0 = \frac{P_1 - P_2}{2\, sinh(\Gamma d)} \ .$$

The forward travelling pressure wave P_0^+ can thus be estimated from the pressures P_1 and P_2 measured by two microphones:

$$P_0^+ = \tfrac{1}{2}(P_0 + ZU_0) \Rightarrow \hat{P}_0^+ = \frac{1}{2}\left(\tfrac{P_1+P_2}{2\, cosh(\Gamma d)} + \tfrac{P_1-P_2}{2\, sinh(\Gamma d)}\right).$$

The above expression can be simplified by replacing $H(\omega) = e^{-\Gamma d}$, $H(\omega)$ being a suitable filter that describes the propagation in a segment of tube of length d:

$$\hat{P}_0^+ = \frac{1}{2}\left(\frac{P_1 + P_2}{1 + H^2}H + \frac{P_1 - P_2}{1 - H^2}H\right) = \frac{P_1 - H^2 \cdot P_2}{1 - H^4}H. \quad (3)$$

By noticing that filtering by H corresponds to a shift by a distance d towards the positive direction, we can remove the term H and shift the reference for the estimated forward travelling wave to the first microphone, $\mathcal{M}_1$. To generalise this result, considering any two microphones $\mathcal{M}_a$ and $\mathcal{M}_b$, and the propagation filter H_{ab}, the estimate for the forward travelling wave referenced to $\mathcal{M}_a$ is defined as

$$\hat{P}_a^+ = \frac{P_a - H_{ab} \cdot P_b}{1 - H_{ab}^2}. \quad (4)$$

2.2. Discrete Wave Domain

If we "open the loop" on the reed side, by treating the wave coming from the mouthpiece as a generic signal $s = p_1^+$, we obtain the scheme shown in Fig. 1, which is similar to a *digital waveguide structure* [4, 5]. The signal s is the forward travelling wave referred at the position of $\mathcal{M}_1$ that we want

[1] We use uppercase letters to indicate variables and functions in the (complex) frequency domain, often dropping the dependence from z or ω which should be clear from the context.

to explicitate with respect to $P_1(z)$ and $P_2(z)$. Referring to the figure, we can write the expressions of P_1 and P_2 as the sum of the forward and backward travelling waves, both filtered versions of S. These are

$$\begin{aligned} P_1(z) &= S(z) + H_{12}^2(z)R(z)S(z) \\ P_2(z) &= H_{12}(z)S(z) + H_{12}(z)R(z)S(z) \end{aligned} \quad ,$$

where $H_{12}(z)$ is the propagation filter from $\mathcal{M}_1$ to $\mathcal{M}_2$ while $R(z)$ is a generic reflection filter that describes the bore of the instrument. To explicitate $S(z)$ we can take the difference between the first and the second, the latter multiplied by H_{12}, in order to eliminate the dependence from $R(z)$:

$$P_1(z) - H_{12}(z)P_2(z) = S(z) - H_{12}^2(z)S(z).$$

We can then explicitate $S(z)$,

$$S(z) = \frac{P_1(z) - H_{12}(z)P_2(z)}{1 - H_{12}^2(z)}, \tag{5}$$

which is a discretised version of (4).

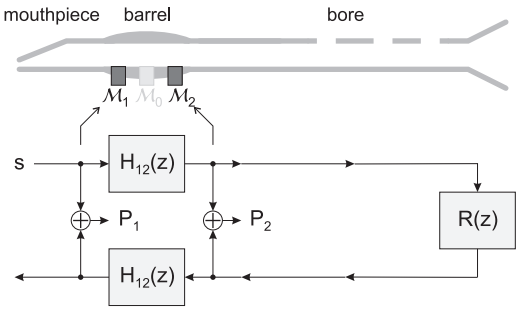

Figure 1: Model for the signals acquired and placement of the microphones on the barrel of the instrument. The third microphone used for the adaptation ($\mathcal{M}_0$) is also shown.

2.3. Wall Losses

Although the wall losses in the short segment of tube between two microphones are very small, taking them into account will improve the estimation. As we will show with an example in the next Subsection, if H is a pure delay then the filter in (5) will have poles on the unit circle at periodic frequency values. To move these poles inside the unit circle and reduce the peaks at these *singular frequencies*, $H(z)$ must include losses. We will use Pierce's approximation for the propagation function [7], formulated as

$$\Gamma(\omega) \cdot d = \left(j\omega\frac{1}{c} + g \cdot \frac{1}{c}\sqrt{j\omega} \right) \cdot d = j\omega\tau + g \cdot \tau\sqrt{j\omega},$$

where c is the speed of sound and g is a positive constant that embeds all the losses per unit length.

Although a time-domain formulation for the propagation filter $H(\omega) = e^{-\Gamma(\omega)\cdot d}$ exists, this is not well sampled using common audio sample rates. A number of techniques have been proposed to design a band limited version of this filter (see for example [8]), but none of them seem to be suitable for an adaptive algorithm implementation. This is one of the main reasons that led us to tackle the problem in the frequency domain, using a sampled version of the analog filter

$$H(\omega) = e^{-\tau(j\omega + g\cdot\sqrt{j\omega})}. \tag{6}$$

2.4. Wave Separation - An Example

As an introductory example we present here a simulation of wave separation using (5). The data was generated using the scheme of Fig. 1, where the input signal used was a square wave of unitary amplitude, and H was designed for $d = 20\,mm$ and $g = 0.765\sqrt{Hz}$ (the latter calculated for an inner tube diameter of $15\,mm$, at a temperature of $20\,^{\circ}$C). White noise was also independently added to the generated signals, in order to have a SNR of $20\,dB$ with respect to the input signal s.

As shown in Fig. 2, the separation error cannot be appreciated by comparing the original signal and its estimate in the time domain, while looking at the error in the frequency domain we can see that it is due to resonances in the frequency response of the separation filter. These occur when d is a multiple of half the wavelength of the signal, so that the microphones' measures are in-phase or in phase opposition, thus providing the same information [6]. This is a physical limitation of the system when fed with periodic signals and it does not depend on the separation method chosen, although these effects can be attenuated by including losses, as mentioned above. Analytically, this phenomenon can be understood by studying the frequency response from each noise source to the output. It is considered noise any additive signal on P_1 that is not present, delayed by the propagation function, in P_2, or vice versa. As a result, the noise is filtered by the transfer function from P_1 to the output, $\frac{1}{1-H^2}$, or the one from P_2 to the output, $\frac{H}{1-H^2}$, both having practically the same spectral characteristics.

3. THE ADAPTIVE ALGORITHM

From the above example it can be deduced that $H(z)$ must describe the propagation from the first microphone to the second as well as possible, or the estimation error will rapidly increase. It follows that the filter must track changes in the propagation due to variations in temperature and humidity of the air, as these can change considerably when the player is blowing air into the instrument. To gather a measure of the filter's accuracy we can place a third microphone $\mathcal{M}_0$ and compare its output with the pressure at the

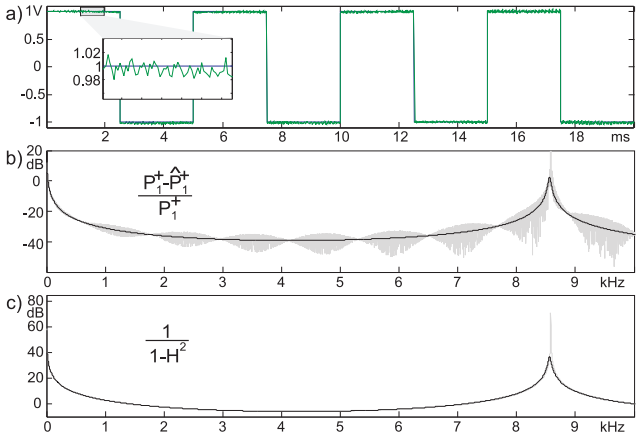

Figure 2: Separation error: a) in the time domain; b) in the frequency domain; c) frequency response of the denominator of (5). (b) and (c) also show the same plots when no losses are taken into account (shown in light gray)

same position, estimated from P_1 and P_2. We can then adjust $H(z)$ in order to minimise this difference.

We have seen that the propagation filter is specified in the frequency domain and a time-domain version cannot be easily obtained. There are also a number of other reasons why the frequency domain should be preferred. The filter's parameters vary slowly with time so that it makes sense to work with blocks of data and average the estimate, to have a better performance against the noise. However, the classic block LMS [9] scheme would not exploit all the *a priori* information about the filter, since, as we have seen, this is defined except for the two parameters τ and g. We used a technique similar to the one devised in [12] and further developed in [13], where the update is computed on the parameters which are then used to design the filter for the next iteration. Once a suitable objective function has been defined, its gradient (or even a higher order derivative) provides the update direction for the parameters, and can be easily computed in the frequency domain. In the time domain this would not be so immediate as we would need an appropriate interpolation filter to account for fractional delays [13].

Since the whole process is carried out off-line, the adaptation can be seen as an optimisation problem, where the aim is to find the optimum set of parameters that minimise the objective function within the current data block. If the size of the block is sufficiently large (condition easily verified in practice) we also take advantage of the convolution theorem, having to compute only vector multiplications instead of convolutions [11, 14], thus increasing the efficiency. Considering that the optimization is carried out within the same block, we only need one FFT and one IFFT per block, leading to an even more efficient solution[1].

3.1. Objective Function for Equi-spaced Microphones

If we place the third microphone at the mid-point between $\mathcal{M}_1$ and $\mathcal{M}_2$, then the propagation filters (from $\mathcal{M}_1$ to $\mathcal{M}_0$ and from $\mathcal{M}_0$ to $\mathcal{M}_2$) will be identical, i.e. $H_1 = H_2 = H$. By recalling that $\hat{P}_0 = \frac{P_1+P_2}{1+H^2}H$ (see Eq. (3)) we can write an expression for the error function, or the difference between H and the actual propagation function between any two adjacent microphones:

$$E = (1+H^2)(\hat{P}_0 - P_0) = H\cdot(P_1+P_2) - (1+H^2)\cdot P_0. \quad (7)$$

In the above equation, E is a complex function of ω while its time-domain counterpart e, being a combination of real functions, is real. However, in the discrete case, e will differ from its version computed entirely in the time-domain because of the effect of the circular convolution. As a general rule, to ensure that the objective function will have a "sharp" minimum when H equals the optimum filter, it is necessary to apply an IFFT to E, null those samples that are heavily corrupted by the circular convolution (i.e. those which would not become null even when H equals the optimum filter) and then apply an FFT to return to the frequency domain (see also [14, 15]).

The objective function in a least square sense is obtained, in the discrete frequency domain, as

$$F = \sum |W \cdot E|^2 = \sum (W \cdot E)\,(W \cdot E)^*, \quad (8)$$

where $W(\omega)$ is a suitable weighting function and the summation is extended to all the frequency samples, so that F is a real function of the filter's parameters.

3.2. Extension to Arbitrary Microphones Spacing

Due to mechanical tolerances, especially in the coupling between microphones and tube, and because of the fact that the "acoustic center" of the microphones is not guaranteed to be on the main axis of the sensor, the distances between the three microphones can be sensibly different from their nominal value. In addition, we may want to place the three microphones at generic positions that are not necessarily equi-spaced. To do so we need a generalization of the expression (7) that is able to account for different distances, or $H_1 \neq H_2$. Instead of estimating the pressure at the central microphone, we can write the error by starting from the difference between two estimates for the forward travelling wave at the same point, obtained from two different groups of microphones. Using Eq. (4), replacing P_a, P_b with P_1,

[1] This applies only if the IFFT/FFT necessary to null the effect of the circular convolution can be avoided [14, 15].

P_0 for the first estimate and P_1, P_2 for the second estimate and taking their difference we have

$$P_1^+\big\rfloor_{10} - P_1^+\big\rfloor_{12} = \frac{P_1 - H_1 \cdot P_0}{1 - H_1^2} - \frac{P_1 - H_1 H_2 \cdot P_2}{1 - H_1^2 H_2^2},$$

and, after multiplication by $\left(1 - H_1^2\right)\left(1 - H_1^2 H_2^2\right)$ we obtain the new expression for the error:

$$E = (1-H_2^2)H_1 \cdot P_1 + (1-H_1^2)H_2 \cdot P_2 - (1-H_1^2 H_2^2) \cdot P_0. \tag{9}$$

Although obtained in a different way, Eq. (9) is consistent with (7), as by imposing $H_1 = H_2 = H$ we re-obtain the expression for the error in the equi-spaced case. It can also be verified that (9) is not affected by a different choice of microphone groupings or a shift of the reference plane.

3.3. Optimisation

Once an objective function is defined, the optimum filter is given by the set $\mathbf{h}$ of its parameters that minimises it:

$$\min_{\mathbf{h}} F(\mathbf{h}). \tag{10}$$

As we have seen, for generic propagation filters $H_1 \neq H_2$, $\mathbf{h}$ includes τ_1, τ_2 and g (the latter being the same for H_1 and H_2, as it is independent from the travelled distance). However, it can easily be verified from (9) that although the estimate for these parameters is insensitive to a global scaling of the pressure signals acquired, it is affected to a mismatch in their individual scale, perhaps due to microphones having different sensitivities. Therefore $\mathbf{h}$ must also include two of the three scale factors (the third would depend on the other two and can be assumed unitary), for example a_1 and a_2, relative to $\mathcal{M}_1$ and $\mathcal{M}_2$.

To take into account the effects of different sensitivities, the signals from the three microphones are now defined as $S_1 = a_1 P_1$, $S_0 = P_0$, and $S_2 = a_2 P_2$, which can be substituted in the error function (9). This, after multiplication by $a_1 a_2$ is redefined as

$$\begin{aligned} E = H_1(1 - H_2^2) \cdot a_2 S_1 + H_2(1 - H_1^2) \cdot a_1 S_2 + \\ -(1 - H_1^2 H_2^2) a_1 a_2 S_0. \end{aligned} \tag{11}$$

The nonlinear optimisation (minimisation) problem stated in (10) is solved using an iterative method. Since the optimisation is now a function of five parameters (a number that is likely to increase in the future), it would not be advisable to adopt a search technique such as the simplex method. Gradient-based methods are much faster since the analytical differentiation of F with respect to $\mathbf{h}$ can be computed quite straightforwardly. This may be a tedious task due to the complexity of F but it can be simplified noting that

$$\frac{\partial F}{\partial \mathbf{h}} = \frac{\partial (E^* E)}{\partial \mathbf{h}} = 2\Re\left\{\frac{\partial E}{\partial \mathbf{h}} \cdot E^*\right\}.$$

This will also increase the efficiency of the implementation. For example, the derivative of F with respect to τ_1 is

$$\begin{aligned} \frac{\partial F}{\partial \tau_1} = 2\Re\left\{E^* \cdot \frac{\partial H_1}{\partial \tau_1} \cdot \frac{\partial E}{\partial H_1}\right\} = 2\Re\left\{E^*(-j\omega) \cdot \right. \\ \left. \cdot \left((1 - H_2^2) H_1 a_2 S_1 - 2H_1^2 H_2 a_1 S_2 + 2H_1^2 H_2^2 a_1 a_2 S_0\right)\right\} \end{aligned}$$

while the differentiation with respect to the other parameters can be done following the same procedure.

By iterating this mechanism it is also possible to calculate the second order derivatives and thus the Hessian matrix of F in order to solve (10) using a Newton's method.

4. RESULTS

In this section we present the results from a number of tests carried out to study the convergence of the algorithm. Measurements were taken on a purposely built apparatus, essentially composed by a tube (the bore of the instrument) connected to a compression driver (acting as the mouthpiece) as schematised in Fig. 3. Three microphones (Endevco model 8507C-1) were placed flush with the inner wall of the tube, equi-spaced by a distance of $20\,mm$. The acquisitions were compared with simulations obtained using the scheme of Fig. 4, adjusted with the same parameters of the experimental setup. The input signals used were a square wave ($f = 200 Hz$) and a maximum length sequence (MLS).

Fig. 5 represents the basin of convergence for (τ_1, τ_2) when all the other parameters are set to their optimum value. The plots were obtained using a data window of 8192 samples. We notice that the convergence is guaranteed over a wide area around the solution and it can be improved by reducing the limits of the summation in (8) to the range of frequencies where the most part of the signal's energy is concentrated, so that any disturbance outside the band of interest is not taken into account. Also, the shape of the surface is not affected by additive noise applied to the sensors.

Similar plots can be obtained for (a_1, a_2) and g, although we found that the latter is quite sensitive to (differences in) the frequency response of the three microphones. This happens because g converges to the value that best explains a non-flat frequency response of the sensors. The problem is even more accentuated when the losses for H_1 and H_2 are optimised independently (i.e. by replacing g with g_1 and g_2).

5. OTHER APPLICATIONS

The adaptive algorithm developed can also be effectively used in other applications where the filter estimated is easily described in the frequency domain as a function of a few parameters. A special case is represented by the pure fractional delay, which has a very simple Fourier transform that is easily differentiable with respect to its parameter.

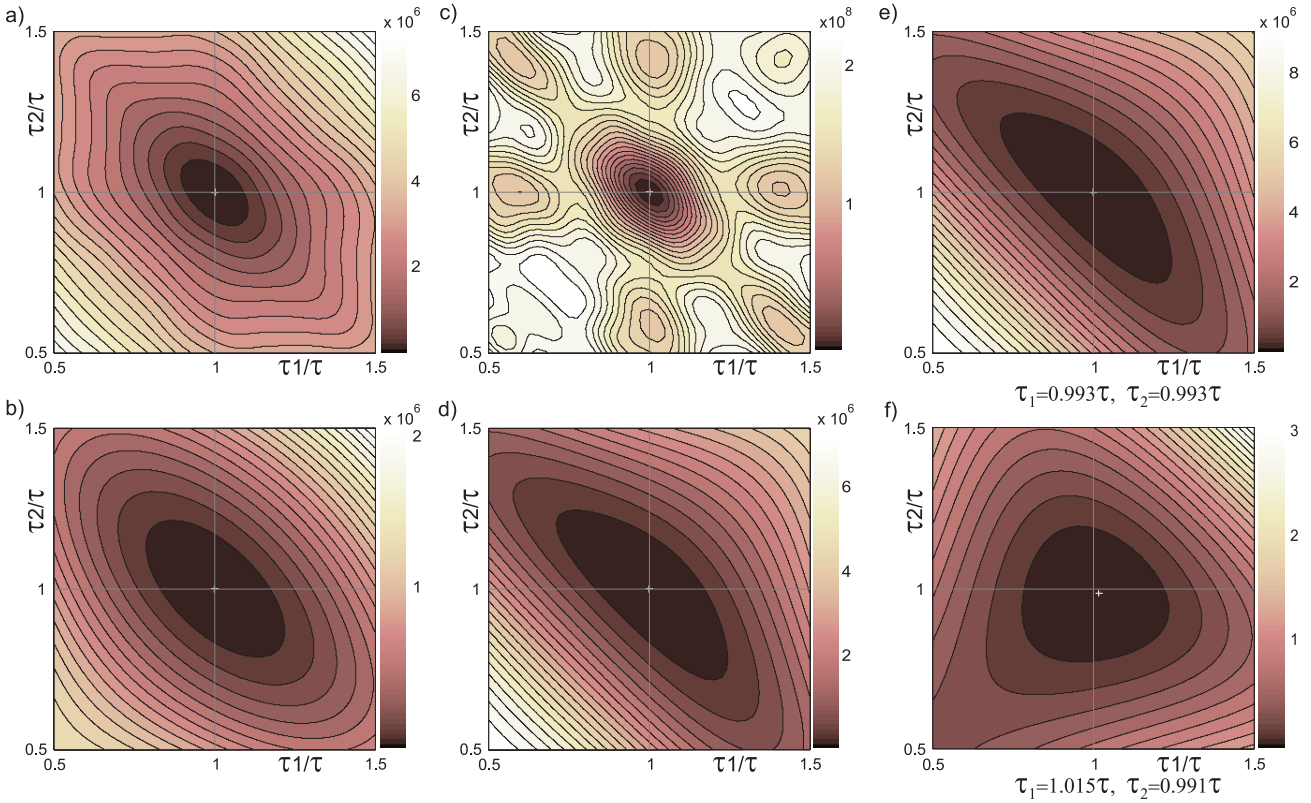

Figure 5: Contour plot of the objective function when τ_1 and τ_2 vary from 0.5 to 1.5 times their nominal value. The solution (τ_1, τ_2) found is marked with a cross in each plot: a) simulation using a 200 Hz square wave as input; b) same as (a) but limiting the summation in Eq.(8) from 100 Hz to 6 kHz; c) and d) same as (a) and (b) respectively, but using a maximum length sequence as input signal; e) same as (d), with additive white Gaussian noise added independently to each microphone to have a SNR of 20dB with respect to the input signal; f) Measurement taken on the apparatus. Estimates for a single realisation are also shown in (e) and (f).

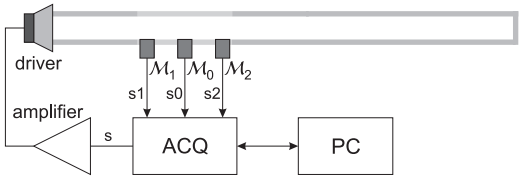

Figure 3: Experimental setup used to study the convergence of the algorithm.

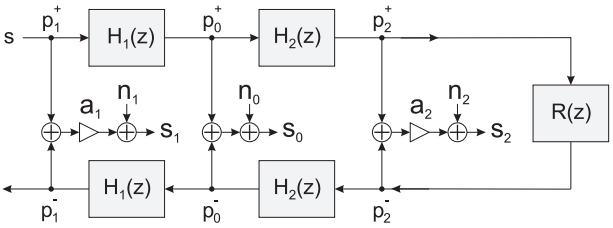

Figure 4: Scheme used for the simulation of the microphones' signals starting from a generic sequence s.

As an example of application using delay filters we will briefly introduce in this section two different DOA estimation schemes using microphone arrays. They are particularly interesting examples as they also make use of the same error functions developed for the original application.

5.1. DOA estimation with 3 microphones in line

Fig. 6 shows a linear array composed by three equi-spaced microphones. The DOA estimation problem consists in finding the angle between the propagation direction of a plane wave and the direction normal to the array. This is equivalent to the estimation of the time taken by a signal to travel a distance $d_1 = d_2$, function of the angle φ. If we assume $\tau = d/c$ known and constant, we can obtain the DOA by estimating $\tau_1 = \tau_2$. Instead of using the classic delay-and-sum technique we can apply the separation filter (4), where in this case $H = exp(-j\omega\tau)$, and use the error

function for equi-spaced microphones (7). We can directly minimise F as a function of the angle of arrival by nulling

$$\frac{\partial F}{\partial \varphi} = 2\Re\left\{\sum_{\omega} j\omega\tau_0\, cos\varphi H \cdot E^*\left(P_1 + P_2 - H \cdot P_0\right)\right\}.$$

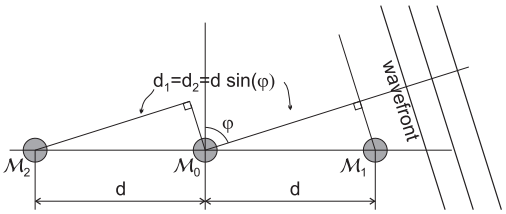

Figure 6: Linear array with 3 microphones.

5.2. DOA estimation with 3 microphones on a triangle

It is reasonable to expect that the precision of the estimate using a linear array will depend on the angle. This is because τ changes very little around $\varphi = \pm 90°$, while around $0°$ its variation is much more pronounced. In order to have a more uniform estimation error with respect to the DOA using the same number of sensors, we can use a triangular array, which can be seen as a circular array of three microphones. Depending on the DOA, the *time of flight* between two microphones will be proportional to the distance between their projection along the propagation direction, as shown in Fig. 7. Therefore, the situation is equivalent to a linear array where the spacing between its sensors is a function of the DOA. We can thus use the more general expression (9) to estimate φ, replacing $H_1 = exp(-j\omega\tau_1)$, $H_2 = exp(-j\omega\tau_2)$ and differentiating the objectve function with respect to the angle:

$$\frac{\partial F}{\partial \varphi} = 2\Re\left\{\sum_{\omega} j\omega E^* \cdot \left[\left(T_1 - \left(T_1 + 2T_2\right) H_2^2\right) H_1 P_1 + \right.\right.$$
$$\left.\left. + \left(T_2 - \left(T_2 + 2T_1\right) H_1^2\right) H_2 P_2 + 2\left(T_1 + T_2\right) H_1^2 H_2^2\right]\right\};$$

$$T_1 = \tau_0 \cdot cos(\pi/6 - \varphi),\; T_2 = \tau_0 \cdot cos(\pi/6 + \varphi)\,.$$

Fig. 8 shows simulated and measured objective functions obtained using a triangular array with $d = 50\,mm$. The functions are normalised to allow an easier comparison. The source signal was a sine wave and results from four different frequencies ($400Hz$, $800Hz$, $1600Hz$ and $2400Hz$) were collected. Measured signals were acquired in a semi-reflective room in order to reduce phase alterations and consequent interferences on the estimated DOAs due to reflections. Measured and simulated curves have the same trend, although with the measured data the estimated angles vary slightly with the frequency, as highlighted by the normalised

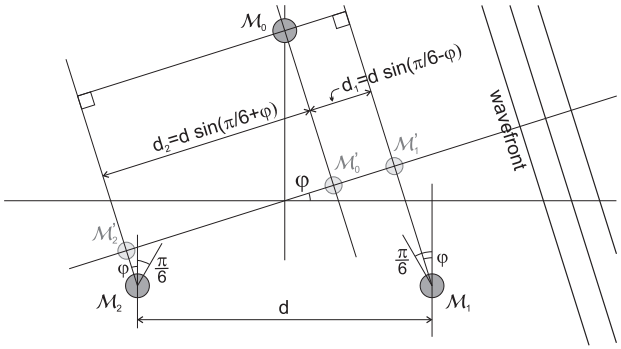

Figure 7: Triangular (circular) array. The smaller circles in light gray represent the "virtual" linear array seen by the wavefront.

inverse plot ($F_{inv} = 1/(k\,F + 1)$, $k = 1000$). This is due to an imprecise relative positioning of the microphones used for the acquisitions.

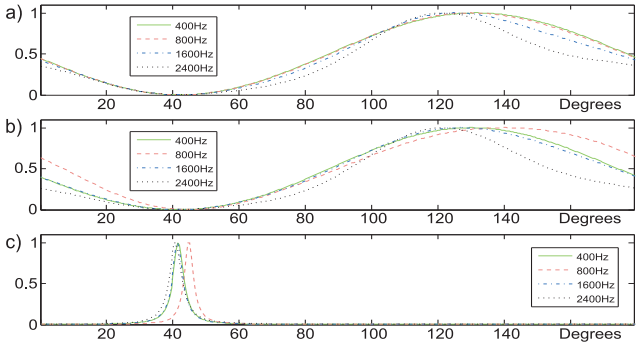

Figure 8: Normalised objective functions for a triangular array and DOA of $42°$: a) simulated; b) measured; c) normalised inverse for the measured data.

6. CONCLUSION AND FUTURE WORK

We developed an adaptive algorithm for wave separation in tubes which represents a fundamental step towards the independent characterisation of exciter and resonator in wind instruments for physical modelling parameters extraction. Forward and backward travelling waves correspond to the input and output of the bore's reflectance filter and allow the estimation of pressure and volume flow in proximity of the reed. The time-domain design of the propagation filter as a function of physical parameters is a difficult task and the techniques found in literature are not suitable for an efficient adaptive filtering algorithm implementation. The obvious choice was thus to work entirely in the frequency domain. This also allowed a closed-form computation of the objective function's first-order derivatives, enabling the use

of gradient based optimisation algorithms. Since second-order derivatives are just as easily obtained, Newton methods could also be used. The only weak point is represented by the effects of the circular convolution which must be taken into account, as they may reduce the precision of the solution.

Current research is addressing an extension of the model to describe a mismatch in the frequency response of the microphones, as this affects the estimation of g and, to a much smaller extent, τ_1 and τ_2. The convergence of g is in fact quite slow and less robust compared to the other parameters and although its influence on H is modest, a further study of its behaviour may increase the precision of the estimate.

We found that the developed algorithm can be used in other adaptive filtering applications, when the filter is a non-linear function of a few parameters as in (6). A special case of this class of filters is the fractional delay. As an example of delay estimation, two DOA estimation schemes using linear and triangular (or circular) arrays have been introduced because of their equivalence to the main application, respectively to the equi-spaced microphones case and to the more general situation. This similarity also provides an alternative perspective of the original problem. However it is interesting to note that in the DOA estimation we rely on the nominal value of $\tau = d/c$ and interpret the estimated delay as the projection of the propagation delay onto the axis of the array to estimate the angle of arrival. In the main application this angle is known (it is a one-dimensional problem) and the error is used to refine the value of τ.

Further developments of the algorithm applied to arrays of microphones will include a more in depth study of the DOA estimation problem and how it compares with existing techniques, especially in reflective environments or when a small number of microphones is available. A variation of the wave separation algorithm could also be applied to the signals acquired by the array for beamforming.

7. ACKNOWLEDGEMENTS

G. De Sanctis would like to thank Dr F. Antonacci at *Politecnico di Milano* for his interesting comments on the DOA estimation application and for providing the acquisitions used in Section 5.2.

8. REFERENCES

[1] M. van Walstijn and G. De Sanctis, "Towards Physics-Based Re-Synthesis of Woodwind Tones", International Congress on Acoustics 2007, Madrid, Spain

[2] V. Chatziioannou, M. van Walstijn, "Inverting the Clarinet", Proc. of the 12th Int. Conference on Digital Audio Effects (DAFx-09), Como, Italy, September 1-4, 2009.

[3] J. Guérard and X. Boutillion, "Real time acoustic travelling waves separation in a tube", In Proc. 1997 Int. Symposium on Musical Acoustics, 1997

[4] J. O. Smith, "Physical Modeling using Digital Waveguides", Computer Music Journal, vol. 16, no. 4, pp. 74, Winter 1992

[5] M. Karjalainen, V. Välimäki and T. Tolonen, "Plucked-string models: From the karplus-strong algorithm to digital waveguides and beyond", Computer Music Journal, vol. 22, no. 3, pp. 17–32, 1998

[6] Y. B. Jim and Y. H. Kim, "A measurement method of the flow rate in a pipe using a microphone array", Journal of the Acoustical Society of America, vol. 112, no. 2, pp. 856-864, 2002

[7] A. D. Pierce, "Acoustics - An Introduction to its Physical Principles and Applications", Acoustical Society of America, New York, 1994

[8] J. Abel, T. Smyth, and J. O. Smith, "A simple, accurate wall loss filter for acoustic tubes", Proc. of the 6th Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 2003, pp. 53-57

[9] B. Farhang-Boroujeny, "Adaptive Filters: Theory and Applications", Chichester, New York: Wiley, 1998, pp. 529

[10] S. S. Haykin, "Adaptive Filter Theory", 4th ed. Upper Saddle River, N.J.: Prentice Hall, 2002, pp. 920

[11] A. V. Oppenheim, R. W. Schafer and J. R. Buck, "Discrete-Time Signal Processing", 2nd ed. Upper Saddle River, N.J.: Prentice Hall, 1999, pp. 870

[12] D. Etter and S. Stearns, "Adaptive estimation of time delays in sampled data systems", Acoustics, Speech and Signal Processing, IEEE Transactions on , vol.29, no.3, pp. 582-587, Jun 1981

[13] H. C. So, P. C. Ching and Y. T. Chan, "A new algorithm for explicit adaptation of time delay", Signal Processing, IEEE Transactions on , vol.42, no.7, pp.1816-1820, Jul 1994

[14] E. Ferrara, "Fast implementations of LMS adaptive filters", IEEE Tr. on Acoustics, Speech and Signal Processing, vol.28, no.4, pp. 474-475, Aug 1980

[15] D. Mansour and A. Gray, "Unconstrained frequency-domain adaptive filter", IEEE Tr. on Acoustics, Speech and Signal Processing, vol.30, no.5, pp. 726-734, Oct 1982