Brass player's mask parameters obtained by inverse method

Sylvain Maugeais, Gilbert Joel

Brass player's mask parameters obtained by inverse method

S. Maugeais[†] and J. Gilbert[⋆]

[†]Sylvain.Maugeais@univ-lemans.fr, Laboratoire Manceau de Mathématiques
[⋆]Laboratoire d'Acoustique de l'Université du Maine, UMR CNRS 6613, Le Mans, France

Abstract

An optimization method is proposed to find mask parameters of a brass player coming from a one degree of freedom lip model, with only constant mouth pressure and periodic mouthpiece pressure as input data, and a cost function relying on the waveform and the frequency of the signal. It delivers a set of parameters called $\mathscr{C}$-admissible, which is a subset of all mask parameters that allow the inverse problem to be well defined up to an acceptable precision.
Values for the mask parameters are found that give a good aproximation of real signals, with an error on the playing frequency of less than 5 cents for some notes. The evolution of the mask parameters is assessed during recordings with real musicians playing bend notes and their effects on the playing frequency are compared to the theoretical change on a model.

1 Introduction

Models of brass instruments have been used for a long time to understand the physics (cf. [6]) and synthesize their sound (cf. [16]). However, one remaining difficulty is the calibration of these models as many of the constants appearing in the mathematical equations are difficult to measure, in particular when they involve human body parts such as the lips of a brass player.

A good calibration of these parameters is useful not only for sound synthesis, where a realistic sound is the main goal, but also for theoretical purposes, as the equations governing models of brass instruments are nonlinear, and are therefore very sensitive to changes in the constants of the model.

In the literature, many models exist for the lips, with different degrees of complexity, either as a one degree of freedom oscillator (which is the most common), or a two degrees of freedom oscillator taking into account different polarities (cf. [3]), and models trying to come closer to the geometry of the opening section of vibrating lips (cf [6], section 5.1.2). However, the more complex the model, the more constants there are that need to be calibrated, and the higher the uncertainty.

The question of calibration of the parameters is not restricted to lip models of brass players, and a related problem is that of the embouchure of reed instruments which may seem easier as measures can be undertaken directly on the reed. The study of the reed parameters is the source of a large literature (see for example [2, 27, 21, 8, 17, 7, 26]) which can be used as a source for methods dedicated to the lips.

Concerning brass instruments, the body of literature is more reduced, although many articles deserve to be cited and will serve as reference for the present work (see [28] for a table summarizing known values). Most notably, [13] who was one of the first to give a complete set of parameters, [10] who built an asymptotic state observer, [29] who used simulated annealing with a cost function depending on playing frequency only, [3] with data coming from high speed camera and [15] using bifurcation diagrams.

The present article aims to provide a new method to identify embouchure parameters that can be used with very little apparatus on actual musicians, and that can follow their evolution while playing. It introduces a new cost function which is a combination of [29] (for the frequency part) and [7] (without the displacement), together with some penalization (see section 2.4). An optimization algorithm is used to minimize this cost function on recordings with actual musicians and the results are discussed (cf. section 3).

2 Method

2.1 Model

To reduce the amount of parameters that must be calibrated, the model chosen in the present article is the simplest one described by (2.1), with only three equations (cf. [20]), that proved to replicate many of the properties of brass instruments (see for example [19]). It relates the mouth pressure p_m to the mouthpiece pressure p and the opening h through a spring-mass-dashpot equation describing the lips, a valve effect computing the flow u through the lip from the difference of pressure, and the expression of the input impedance:

$$\begin{cases} \ddot{h} + \frac{\omega_\ell}{Q_\ell}\dot{h} + \omega_\ell^2(h-H) = \frac{p_m - p}{\mu} \\ \dot{p}_n = Z_c C_n u + s_n p_n \\ p = 2\Re\left(\sum_{n=1}^{N} p_n\right) \\ u = w h^+ \mathrm{sgn}(p_m - p)\sqrt{\frac{2|p_m - p|}{\rho}} \end{cases} \tag{2.1}$$

where $h^+ = max(h, 0)$, w is the width of the lips, and Z_c the characteristic impedance.

Here, the impedance is decomposed using modal analysis as a sum of simple fractions

$$Z(\omega) = Z_c \left(\sum_n \frac{C_n}{j\omega - s_n} + \frac{C_n^*}{j\omega - s_n^*} \right) \qquad (2.2)$$

and p_n being complex valued.

The variables p_n form a decomposition of the mouthpiece pressure using the special form of the impedance. They are not really modal coordinates, as they are not naturally orthogonal for some scalar product, but give a convenient way to solve the problem.

The input impedance is measured through an impedance bridge and is therefore known. It is decomposed into formula (2.2) using the Rational Fraction Polynomial method (see [14], section 4.4.3), giving the values for s_n and C_n which are therefore fixed for a given instrument.

The value of C_n and s_n computed for the Bb trombone in first position and used in the measures (basse trombone Courtois and mouthpiece Holton) are given in table 5.

The mask parameters are the remaining constants appearing in equations (2.1), and are the lip angular resonance frequency $\omega_\ell = 2\pi F_\ell$, the quality factor Q_ℓ, its surface density μ and its opening at rest H.

The set of equations 2.1 can be rewritten as a real valued ordinary differential equation (cf. [20]) $\dot{y} = f(y)$ of order 1 in dimension $2N + 2$ with a state variable

$$y = [h, \dot{h}, \Re(p_1), \mathrm{Im}\,(p_1) \ldots, \Re(p_N), \mathrm{Im}\,(p_N)]. \qquad (2.3)$$

Such a formulation is very convenient both for time numerical simulation, and for the study of bifurcation diagrams, as computed by the software auto-07p [12] (see section 2.9).

The time simulations performed in this project are based on the Runge-Kutta 4 algorithm and are coded in C language as a large number of them are performed. The sampling rate is set to $44100Hz$ as it gives a sufficient precision for the results (comparable to measured data), ensures the stability and convergence of the numerical scheme, and is sufficiently fast (0.18s for a one second signal).

The time simulations have to be performed on a time range long enough so that stationary regime is attained. In practice, this means signals of up to 4 seconds have to be simulated as transient regime can be quite long for some mask parameters.

For a set of mask parameters

$$\mathfrak{M} = (\omega_\ell, Q_\ell, H, \mu), \qquad (2.4)$$

we write $p_{\mathfrak{M}}$ for the mouthpiece pressure obtained by solving the system (2.1) with zero initial value $y(0) = [0, \cdots, 0]$ (cf. equation 2.3). All the other parameters of the model, including p_m, are fixed and constant during one optimization.

2.2 Experimental protocol

The goal of the project is to get access to as many mask parameters as possible with as little apparatus as necessary so as not to hinder the musician's playing. We therefore focused on only two piezoresistive pressure sensors (Endevco 8507C-5): one in the mouth of the musician (or artificial mouth), and one in the mouthpiece. Simultaneous recordings give access at each time step to the quantities p_m and p. Both signals are sampled at 44100Hz.

As only one period of p is needed by the algorithm below, the signal can be broken into pieces to see the evolution of parameters (see section 2.4). Once a period is chosen, p_m is averaged on the same time period to have a constant value to feed the numerical simulation.

Three sets of measures were performed with two experienced amateur trombone players, one of them being recorded twice, labeled A1, A2 and B in the rest of this article. For each session, the musician was asked to play 6 notes on a Bb bass trombone in first position (Bb2, F3, Bb3, D4, F4, Bb4), together with a bend on F4 (first down, then up), and a crescendo on F4.

2.3 Extraction of a signal's period

For a given sampled signal p, either simulated (as in equation (2.1)) or measured (see section 2.2), the determination of a period is critical for the method and the first step to perform it is the identification of the periodic regime and its frequency. This is done using a python implementation [23] of the Yin algorithm (see [11]).

The Yin algorithm produces an estimator called the harmonic rate, which is a real number between 0 and 1, that gives a quantification of how periodic the signal is: A harmonic rate very small (ideally 0) meaning that the signal is close to being periodic. In practice, we consider the signal to be periodic when the harmonic rate is lower than 10^{-3}, and extract the part of the periodic regime with the smallest harmonic rate.

Yin also gives an estimation of the instantaneous frequency F at each point. From this, it is already possible to extract a waveform $p^\dagger$ of duration exactly one period in the periodic regime. However, as this waveform has to be compared to another one coming from a reference signal, a phase condition has to be fixed. We therefore demand that all the waveforms

$$\begin{cases} \bullet & \text{begin by crossing 0,} \\ \bullet & \text{in an increasing way.} \end{cases} \qquad (2.5)$$

This is always possible in practice as p has a mean value equal to 0. The normalization is achieved by considering the waveform $p^\dagger$ and shifting it to the left until the first point satisfies the phase condition, giving rise to a new waveform $\widetilde{p}$ which is used as a reference.

It should be noted that on a general signal, the phase condition may not be sufficient to uniquely determine $\widetilde{p}$. However, for the signals obtained either numerically or experimentally, this condition proved to be sufficient.

For a set of mask parameters $\mathfrak{M}$, we also write $F_{\mathfrak{M}}$ and $\widetilde{p}_{\mathfrak{M}}$ for the frequency and normalized waveform of the signal $p_{\mathfrak{M}}$ obtained in section 2.1.

2.4 Definition of the cost function

The goal of the cost function is to try to compare two periodic signals p_{ref} and p, from which frequencies F_{ref}, F and waveforms $\widetilde{p}_{ref}, \widetilde{p}$ are extracted. The signal $\widetilde{p}_{ref}$ can be either a recorded signal, or a simulated signal obtained from known mask parameters (for test purposes), and is the reference against which the model outputs are compared. Although in theory it should be enough to compare $\widetilde{p}_{ref}$ and $\widetilde{p}$, it puts too much emphasis on the waveform itself, and too little on the frequencies. As both timbre and intonation are important for the applications, it is necessary to add an extra weight to the difference in frequencies.

The preliminary cost function $\mathscr{C}_\circ$ is therefore

$$\mathscr{C}_\circ(p_{ref}, p) = \frac{1}{\|\widetilde{p}_{ref}\|_2^2} \int_0^{\min\left(\frac{1}{F_{ref}}, \frac{1}{F}\right)} \left(\widetilde{p}_{ref}(t) - \widetilde{p}(t)\right)^2 dt + \alpha_F \left(1200 \log_2 \left(F_{ref}/F\right)\right)^2 \quad (2.6)$$

with

$$\|\widetilde{p}_{ref}\|_2^2 = \int_0^{1/F_{ref}} \widetilde{p}_{ref}(t)^2 dt.$$

The first term of the sum is the square of the *relative RMS difference*, and the second one the square of the *relative frequency difference*.

The choice of the constant α_F guides the optimization procedure either toward a better approximation of the waveform (small α_F) or toward a better approximation of the frequency (big α_F).

In our case, the choice of $\alpha_F = 0.02$, obtained by trial and error, leads to good results during optimization for trombone sounds, in that intonation (errors around 10 cents) and waveforms (errors around 30%) are respected. In particular it means that the difference in cents between two signals is $\leq \sqrt{\mathscr{C}_\circ(p_{ref}, p)/\alpha_F}$.

2.5 Penalization

As is already well known (cf. [17]) the inversion problem is not well defined and it is actually easy to find multiple sets of mask parameters which give signals with very similar waveforms (see table (1))

	$\mathfrak{M}_{ref}$	$\mathfrak{M}$	Difference
F_ℓ (Hz)	177.70	184.05	60 cents
Q_ℓ	3.82	4.79	20%
μ (kg.m^{-2})	1.28	1.42	10.9%
H (m)	1.2×10^{-5}	1.9×10^{-4}	1483%

Table 1: Values of two different sets of mask parameters giving almost identical signals: relative RMS difference is 0.9% and difference in frequencies is 1.41 cents. Mouth pressure is equal for both simulations and fixed at 2500Pa.

This means in particular that the cost function lacks convexity. One typical solution to remedy this problem is to convexify the cost function using Tikhonov regularization, which amounts to adding quadratic terms with respect to some mask parameters. More precisely, we define the complete cost function $\mathscr{C}$

$$\mathscr{C}(p_{ref}, \mathfrak{M}) = \mathscr{C}_\circ(p_{ref}, p_{\mathfrak{M}}) + \beta_Q Q_\ell^2 + \beta_H H^2 \quad (2.7)$$

The choice on the specific penalization has been made on Q_ℓ and H because it proved to be

- sufficient to have a well defined solution up to a sufficiently good precision (cf. table 2),
- necessary to remove very different solutions (cf. section 2.8).

It should be noted that this particular choice of penalization, instead of a more general form like $(H - H_0)^2$ for a reference H_0 which should be fixed for the whole optimization, implies that the optimization procedure will favor solutions with the smallest quality factor and lip opening. This was chosen for lack of a good candidate for H_0.

This particular choice of penalization proved to give results close to those in the literature, except for H (cf. section 3).

The method for fixing the values of β_Q and β_H is done so that a typical value of the penalization should be of the same magnitude as $\mathscr{C}_\circ(p_{ref}, p)$. As we expect $\mathscr{C}_\circ(p_{ref}, p)$ to be about 0.3 (cf. section 3), that $Q \cong 7$ and $H \cong 10^{-4} m$ (chosen among the known values in [28]), we took $\beta_Q = 5 \times 10^{-3}$ and $\beta_H = 3 \times 10^7 m^{-2}$.

2.6 Continuity, optimization algorithm

The algorithm chosen to find the minimum of the cost function is the dual annealing optimization (cf. [30]) which is a stochastic algorithm that requires neither the cost function to be regular, nor the minimum to be unique.

Indeed, the cost function in this article is not continuous: a small variation of the mask parameters can lead to completely dissimilar solutions. For example the trombone player can obtain the different notes by only varying its lip resonance frequency: the variation of playing frequency with lip resonance frequency clearly has jumps (cf. [20] figure 4).

Another problem of the cost function is that it has many local minima. Although we do not have a mathematical or musical reason for this, it is clearly seen during the optimization process using dual annealing as it performs local searches before jumping to other locations (see table 3 where each line represents a local minimum).

2.7 Limitations of dual annealing

The dual annealing algorithm is known to give a solution for some very slow (logarithmically) decreasing temperatures (cf. [30]), but that kind of evolution of the temperature implies a very slow convergence. In practice, a faster decreasing temperature is used, but the convergence is not assured.

Moreover, as this algorithm is of stochastic nature, there is no simple criterion to stop it, and an arbitrary condition

has to be chosen: the choice was made to bound by 1000 the number of calls to the cost function. With this choice, a typical run lasts about 1 hour on a desktop computer.

The probabilistic nature of this algorithm also means that two runs of the same algorithm, with the same initialization data (except for the seed of the random generator) give different solutions that can be quite different, as the global minimum might not have been reached in one run. It is therefore often necessary to launch the algorithm iteratively until a new run does not produce a solution with lower cost function.

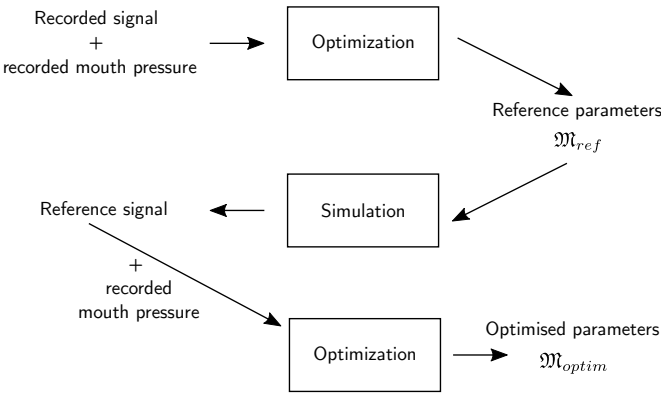

Figure 1: Block diagram explaining how to get a $\mathscr{C}$-admissible set of mask parameters from a recorded signal and use it to assess the precision of the algorithm.

2.8 Precision of the algorithm

For a given set of mask parameters $\mathfrak{M}_{ref}$, the goal is to find an algorithm that captures an approximation of $\mathfrak{M}_{ref}$ with only the knowledge of $\widetilde{p}_{ref}$ and F_{ref}. The algorithm we propose here is not able to do that without a prior assumption on $\mathfrak{M}_{ref}$, namely

$(\star)$ $\mathfrak{M}_{ref}$ must be a minimum of the cost function $\mathfrak{M} \mapsto \mathscr{C}(p_{ref}, \mathfrak{M})$

We say that a set of mask parameters satisfying this hypothesis is *$\mathscr{C}$-admissible*. A random set of mask parameters is not $\mathscr{C}$-admissible in general . Indeed, two sets $\mathfrak{M}_1$ and $\mathfrak{M}_2$ can give very similar waveforms (cf. table 1), so that $\mathscr{C}_\circ(p_{ref}, p_{\mathfrak{M}_1}) = \mathscr{C}_\circ(p_{ref}, p_{\mathfrak{M}_2})$, but have different quality factor or lip opening, and implying for example $\mathscr{C}(p_{ref}, \mathfrak{M}_1) < \mathscr{C}(p_{ref}, \mathfrak{M}_2)$. In that case, $\mathfrak{M}_2$ cannot be $\mathscr{C}$-admissible.

Moreover, this definition highly depends on the choices made for the definition of $\mathscr{C}$, be it α_F or the choice of penalizations. A set of mask parameters may be $\mathscr{C}$-minimal for one choice, but no longer for another one!

The assumption in this article is that for every "realistic" signal (i.e. coming from the recording of a trombone), there is *only one* mask parameter that is $\mathscr{C}$-admissible. It is not at all clear that this is true, and this is even known to be false if the penalizations are not added (see section 2.5). Taking this assumption for granted, a set of mask parameters obtained by minimization of the cost function is automatically $\mathscr{C}$-admissible.

To find a suitable $\mathscr{C}$-admissible set of mask parameters and to stay close to an actual trombone signal, so that the robustness of the optimization procedure can be tested, we applied the optimization procedure to a reference signal on a recorded D4 (cf. section 2.2), cf. figure 1.

The obtained values are given in the second column of table (2), denoted by $\mathfrak{M}_{ref}$. The set of mask parameters used for the initialization of the dual annealing algorithm is given in the third column, and the result of the optimization algorithm is in the last one. The search space is given in its caption.

	$\mathfrak{M}_{ref}$	Init.	$\mathfrak{M}_{optim}$ (error)
F_ℓ (Hz)	177.70	160	178.02 (3.1 cents)
Q_ℓ	3.82	5	3.64 (4.7%)
μ $(kg.m^{-2})$	1.28	3	1.30 (1.5%)
H (mm)	1.2×10^{-2}	0.4	2.48×10^{-2} (105%)

Table 2: Data for optimization: Mouth pressure $p_m = 1656Pa$ and width of the lips $w = 12.10^{-3}m$. Search space is $F_\ell \in [150, 200]$, $Q_\ell \in [0.1, 6]$, $\mu \in [0.1, 3]$, $H \in [10^{-5}, 10^{-3}]$. Error on playing frequency : 6cents

The resulting waveforms for the two sets of mask parameters are indistinguishable, with a relative RMS error of only 3%, and an error on the frequencies of about 6 cents.

The results of optimization are acceptable (much lower than the dispersion of the values found in the literature) and representative of the errors we found on other simulations (cf. section 3) except for the opening at rest H, but its value is so small that it is hard to give a physical interpretation (see section 3.2).

The difference between $\mathfrak{M}_{ref}$ and $\mathfrak{M}_{optim}$ may be explained by the fact that we had to stop the algorithm at one point (cf. section 2.7) or because $\mathscr{C}$ is insufficiently convexified.

2.9 Continuation

During the analysis of the bend (cf. section 3.4), the continuation software auto-07p ([12]) is used to follow the evolution of the playing frequency with different parameters.

The continuation is first initialized for a reference set of mask parameters $\mathfrak{M}_1$ which is chosen for each musician to be the point which minimizes the cost function $\mathscr{C}_\circ$ among all the optimized values of mask parameters, so as to be as close as possible to the actual recording.

The bifurcation diagram is built using the dependency on p_m up to the recorded value of the reference signal, so that auto-07p is now precisely set to the signal $p_{\mathfrak{M}_1}$.

Then a continuation curve along one of the mask parameters (either Q_ℓ, F_ℓ, H or μ), δ is computed, all other physical variables being fixed, and the playing frequency is drawn.

3 Results

The results obtained for the different musicians are presented in the following subsections, but first it is interesting to look at the mouth pressure as a function of the playing frequency, cf figure 2. Indeed, although all these data are directly recorded, and not optimized, we can clearly see differences between the two recording sessions of musician A, where the second session has a larger mouth pressure, which could translate into perceptible differences within the optimized data of a single musician.

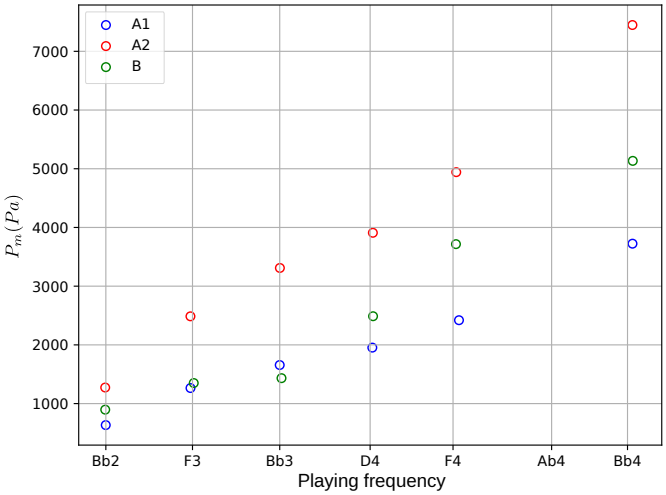

Figure 2: Measured mouth pressure P_m averaged over one period as a function of the playing frequency for all three musicians and the six notes.

3.1 Errors on sustained notes

Both RMS and frequency errors obtained at the end of optimization are presented in figures 3 and 5. The RMS error can be quite large for some notes (up to 40%), which is not surprising as the model is one of the simplest and many physical details are neglected. As the optimization looks for a best fit among all mask parameters, this means the model should be complexified to take into account more of the physics of the instrument if precision on timbre and playing frequency have to be maintained, provided the dual annealing algorithm give results close enough to the global minimum.

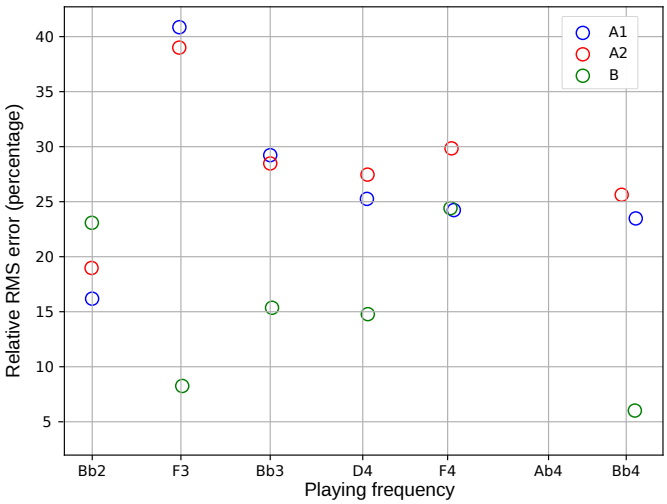

Figure 3: RMS error of **the signals obtained from optimized mask parameters,** relative to the measured signal for 6 different notes and the three musicians.

For reference, a typical waveform is shown in figure (4) where the reference signal is given in green, the reconstructed signal is in orange and the difference between them is in dotted red. The relative RMS error for this particular signal is 0.28. Although many properties are well approximated, the higher harmonics of the signal are clearly not in agreement with the experimental signal. This gives a typical value that can be expected for the RMS error.

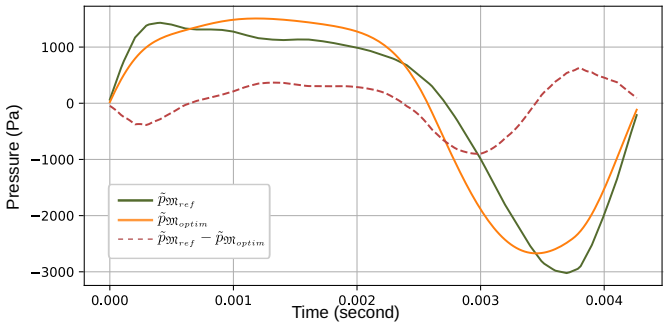

Figure 4: Typical waveform for the recorded signal (green) and for the signal obtained from optimized mask parameters (orange) as played by musician A2 on a Bb3. The difference between signals is drawn in dashed red. The relative RMS error is 0.28

Concerning repeatability, estimations of the mask parameters of musician A are coherent and give almost the same results for both the RMS error and the frequency error (except for the playing frequency of the note Bb4). However, the errors differ largely between both players, player B mainly getting the lowest error. This may suggest that both musicians use different techniques, and that player B is closer to the simple model (2.1).

Note in particular the difference between errors for the note F3, where musician A has the largest RMS error of all (cf. figure 4), and musician B has one of the lowest.

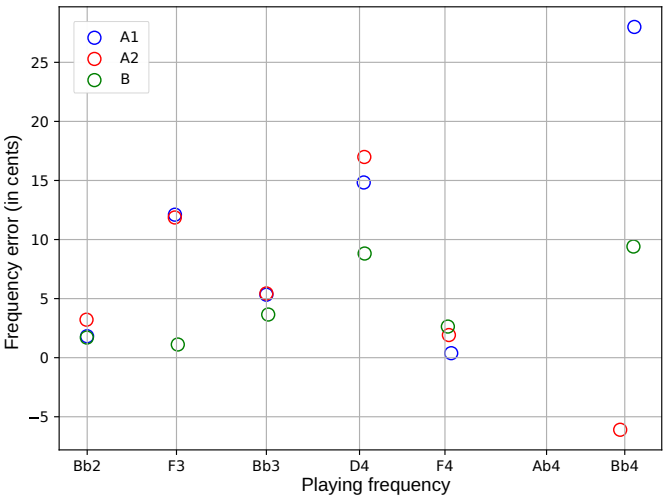

Figure 5: Frequency error in cents of the signals obtained from optimized mask parameters, relative to the recorded signal for 6 different notes and the three musicians

3.2 Discussion on sustained notes

Lip resonance frequency

The lip resonance frequency as a function of the playing frequency is shown in figure 6 for all three musicians. As for any outward model, the frequency of the lips is lower than the playing frequency (cf. [5]), which is clearly seen in this figure as the circles are below the line $F_{play} = F_\ell$.

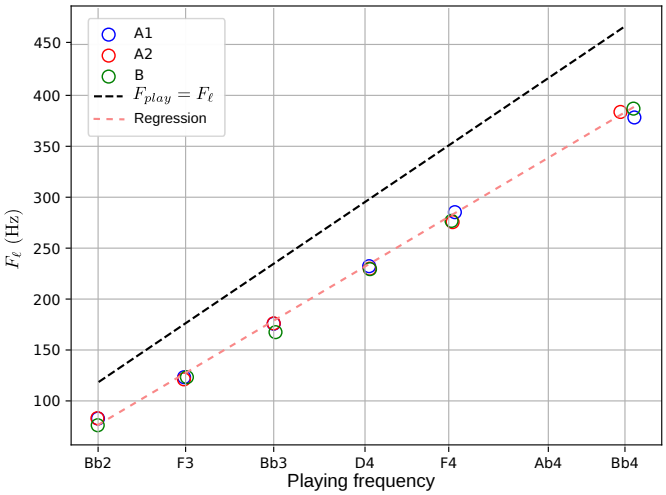

Figure 6: Variation of F_ℓ as a function of F_{play} for all three musicians in circles, together with the diagonal $F_{play} = F_\ell$ in dashed black, and the regression line in dashed red.

Note that for a given frequency, there is little dispersion from player A (A1 or A2) to player B. Moreover the regression line

$$F_\ell = 0.9366 F_{play} - 31.57 \tag{3.1}$$

gives a good fit with $R^2 = 0.996$ and could be used as a first estimation of the playing frequency using only the lip frequency.

Quality factor

The estimated quality factors for all three players are displayed in figure 7. As in the case of lip frequency (cf. 3.2) results are very close for all three players, and also for all notes, being between 2 and 5.

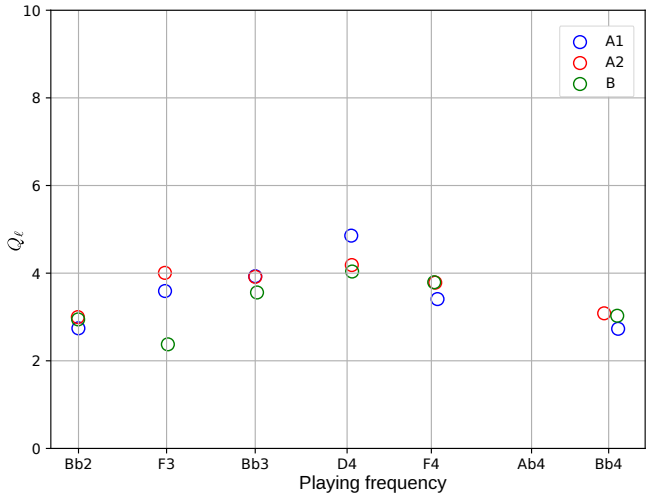

Figure 7: Variation of Q_ℓ as a function of F_{play} for all three musicians in circles.

Compared to the literature, they are however smaller than the measured values of [9] (between 9 and 10.5) but comparable to the estimation of [18] (around 5), [22] (between 1.2 and 1.8), [24] (around 3.7), [25] (around 2.88) and [1] (between 0.5 and 3). Except for the first reference, this justifies the penalization on Q_ℓ, which tends to favor the smallest possible values.

The values obtained for the two recordings of player A are always very close, this may mean that it depends very little on the loudness.

Surface density

The optimized values of μ^{-1} are given in figure 8. Except for the 4 highest values, they are comparable to [25]. However, they are overestimated compared to other values found in the literature ([13, 9, 18, 24]) where they are between 0.03 and $0.2 m^2.kg^{-1}$.

We can see that the data for A1 is systematically higher than that of A2, indicating a possible dependency on the mouth pressure and the loudness.

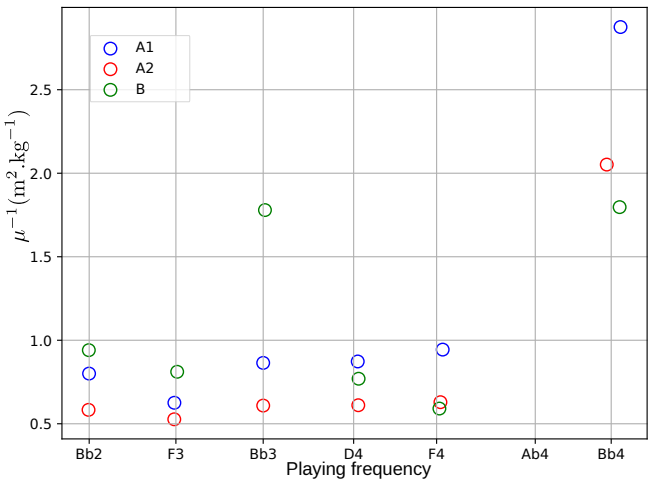

Figure 8: Variation of μ^{-1} as a function of F_{play} for all three musicians.

Opening at rest

The values of opening at rest obtained by optimization are given in figure 9. They seem very small compared to what was obtained by other authors, up to a factor 10: a typical value obtained by optimization is around $2.10^{-5}m$ (see figure 9) whereas [9] and [24] have a typical value of $5.10^{-4}m$, [18] $2.10^{-4}m$, and [1] $1.10 \times 10^{-3}m$, all with comparable lip's width.

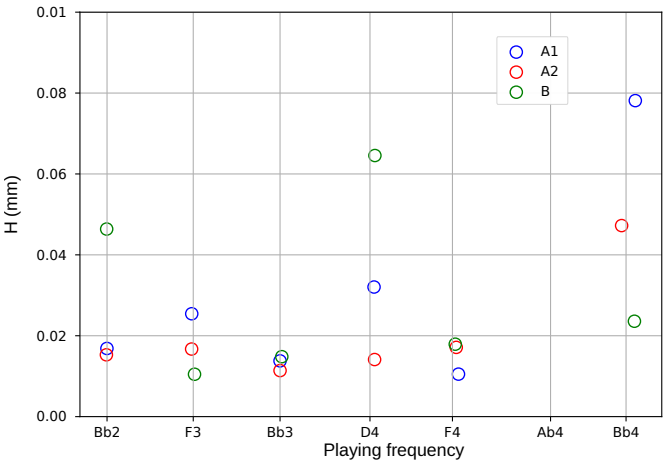

Figure 9: Variation of H as a function of F_{play} for the three musicians.

However, when shifting from opening at rest to mean opening (cf. figure 10) using formula (A.3), which takes also into account the mouth pressure, the lip frequency and the lip surface density, the results are comparable to those of [4], which are between $0.6mm$ and $2mm$. It should be noted that in this case, the value of opening at rest is negligible in the formula (A.3) in appendix A.

Just as for μ^{-1} there appears to be a correlation between loudness and mean opening. It is not only expected, but actually obvious from the formula (A.3) where the mouth pressure appears.

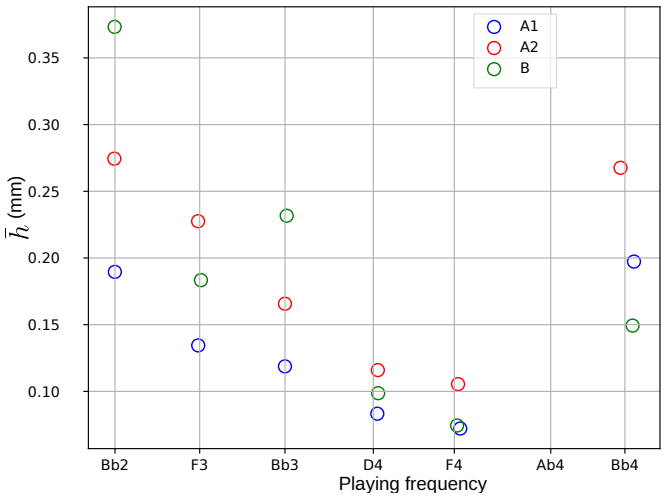

Figure 10: Mean opening of the lips as a function of F_{play} for the three musicians.

3.3 Optimization for bent notes

The musicians were instructed to perform pitch bends on F4: without moving the slide, the player used embouchure adjustments to vary the pitch, first below its normal value, then above, then below, and then back to F4. For each recording of approximately 10 seconds, the signals are cut into chunks of 0.2 seconds, and the optimization procedure is applied independently on each chunk. The note F4 has been chosen because it is one of the most comfortable to bend for the musician. The RMS and frequency errors can be found in figures (11) and (12).

The RMS error is around 0.25, except for a group of notes played by musician B with low frequencies, which may be related to a particular technique used by this musician.

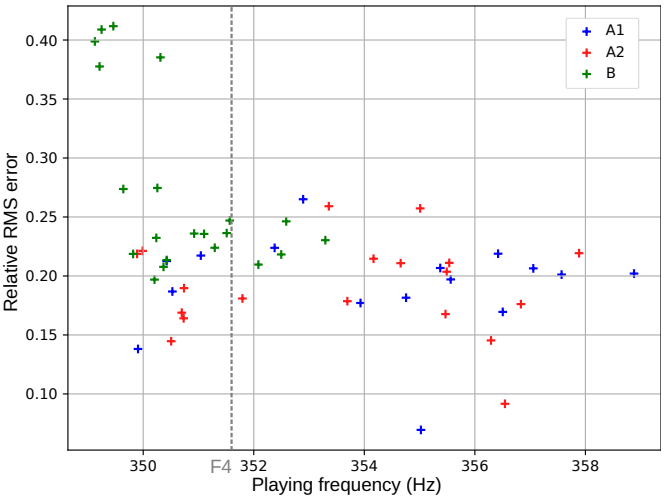

Figure 11: RMS error for bent note on F4 for all three musicians. The dashed line represents the average frequency of the actual played F4 with no bend.

The frequency error is quite low except for the lowest notes, and is in agreement with the very low frequency error found for F4 in figure (5). This suggests that the

model is not able to predict precisely what the musician is doing for the lowest frequencies of the bend.

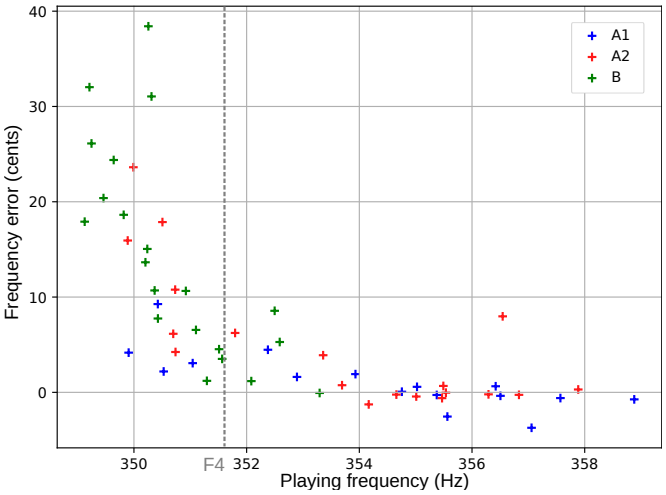

Figure 12: Frequency error for bent note on F4 for all three musicians. The dashed line represents the average frequency of the actual played F4 with no bend.

Indeed, there seems to be two regions for the errors, with a change at around $352Hz$, as if the optimization process could not find parameters that fit the playing frequency below that value. In practice, musicians often use special techniques to bend to very low notes, such as using the vocal tract. This technique is clearly not taken into account in the model, and it seems the algorithm indicates its own limits.

3.4 Discussion on bent notes

During bending, the musician varies many parameters. This makes it quite difficult to see the influence of any of them. In the following diagrams, the evolution of playing frequency is shown with respect to the mask parameters. To put it into perspective, the theoretical evolution with respect to only the considered parameter (the other parameters being kept constant) is also computed using auto-07p (see section 2.9). The mask parameters used to initialize the continuation are those with smallest cost function among all the optimized values for this recording, to ensure that the model is as close to the measures as possible.

Quality factor Q_ℓ

The results of the optimization for bent notes is presented in figure 13 for the quality factor. The values obtained are within the same range as in figure 7.

One striking feature is the proximity from the measures of musician B, and the results of continuation obtained by auto-07p. It seems like the playing frequency is completely predicted by the evolution of the quality factor. However, the precision of the fit must be put into perspective with the rather large errors in the optimization (see figures 11 and 12).

The fit is not so good with musician A, although the results of the continuation go in the right direction.

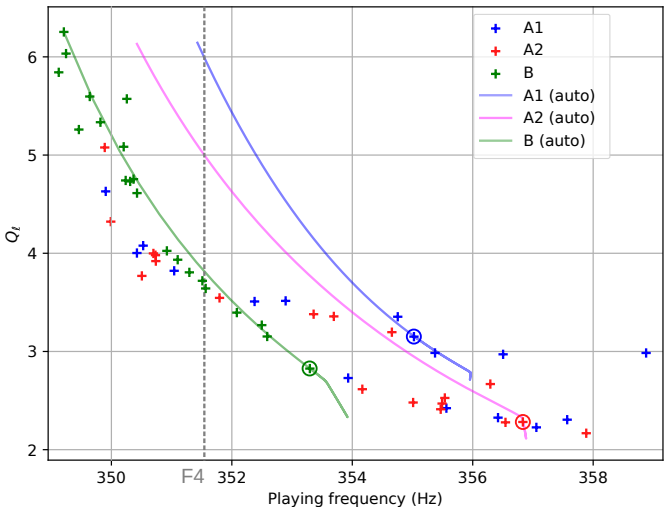

Figure 13: Variation of Q_ℓ as a function of F_{play}. Crosses represent the measures for all three musicians, circles the initialization parameter for continuation (see section 2.9), and lines the continuation obtained with auto-07p. The dashed line represents the average frequency of the actual played F4 with no bend.

Lip resonance frequency F_ℓ

The results of optimization for the lip resonance frequency are given in figure 14, and are quite difficult to interpret. Even more than in figure 12, there seem to be two regions, one before $352Hz$, and one after.

Above 352Hz, the estimation of lip frequency does not give a clear tendency. Although we could expect the lip frequency to increase with playing frequency, just as in figure 6, this is not what appears in the figure. This suggests that the lip frequency is only a coarse tuner, and the quality factor is actually the fine tuner.

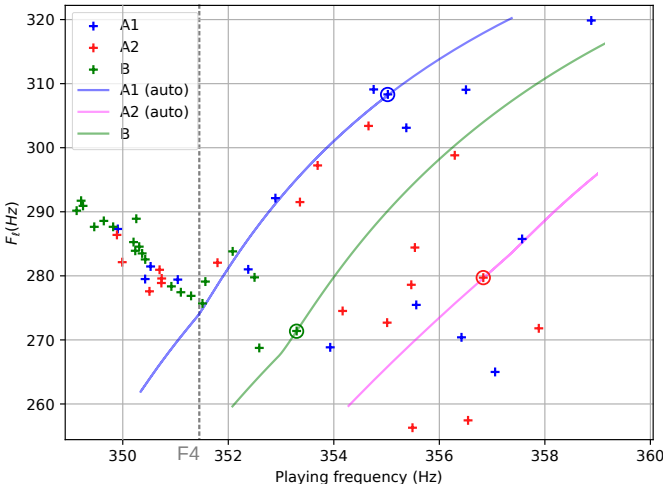

Figure 14: Variation of F_ℓ as a function of F_{play}. Crosses represent the measures for all three musicians, circles the initialization parameter for continuation (see section 2.9), and lines the continuation obtained with auto-07p. The dashed line represents the average frequency of the actual played F4 with no bend.

Lip surface density μ

The results of the optimization for bent notes is presented in figure 15 for the lip surface density. The values are compatible with those in figure 8, and the evolution of playing frequency relatively to μ is compatible with the theoretical one obtained by continuation.

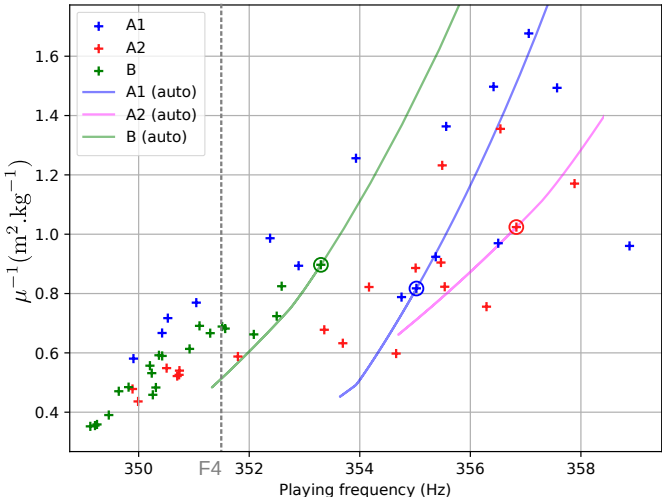

Figure 15: Variation of μ^{-1} as a function of F_{play}. Crosses represent the measures for all three musicians, circles the initialization parameter for continuation (see section 2.9), and lines the continuation obtained with auto-07p. The dashed line represents the average frequency of the actual played F4 with no bend.

Opening at rest H

The results of the optimization for bent notes is presented in figure 16 for the opening at rest. The values obtained are within the same range as in figure 9.

As explained in section 3.2, the values obtained for H are very small, and therefore not very well defined (see error term in table 2).

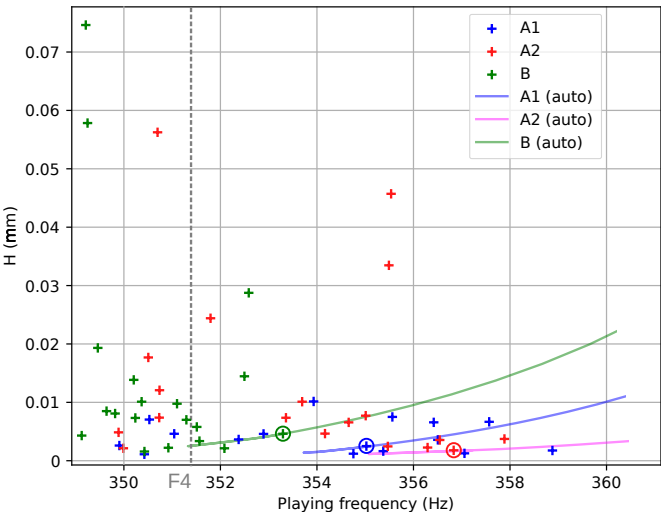

Figure 16: Variation of H as a function of F_{play}. Crosses represent the measures for all three musicians, circles the initialization parameter for continuation (see section 2.9), and lines the continuation obtained with auto-07p. The dashed line represents the average frequency of the actual played F4 with no bend.

The mean opening obtained from other optimized values and formula (A.3) is given in figure 17, and a clear tendency can be observed above $350Hz$: the mean opening increases with the playing frequency for all musicians.

Below $350Hz$ the tendency is not so clear. Moreover, one must be careful with interpretation as there may be other phenomena involved than those directly modeled (cf. section 3.3).

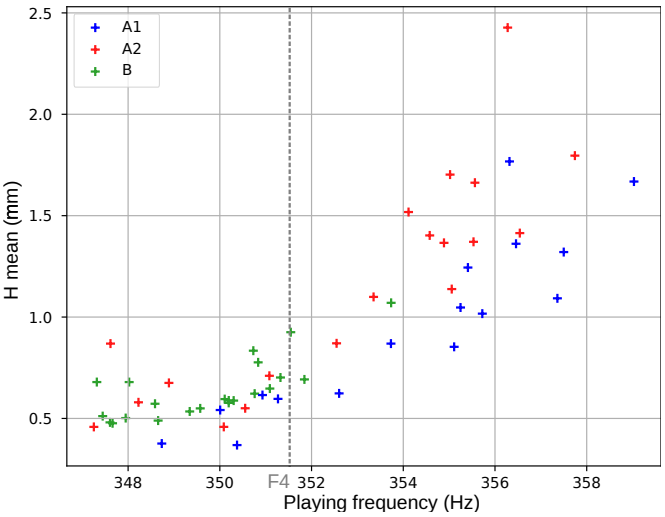

Figure 17: Variation of H_{mean} as a function of F_{play}. Crosses represent the measures for all three musicians, circles the initialization parameter for continuation (see section 2.9), and lines the continuation obtained with auto-07p. The dashed line represents the average frequency of the actual played F4 with no bend.

3.5 Optimization for a crescendo

The same procedure as in section 3.3 was used for the recordings of a crescendo on F4 for all three musicians.

The relative RMS error is presented in figure 18, and indicates that the higher the mouth pressure, the higher the RMS error. This shows that the simple model (2.1) is good at reproducing the timbre for low pressure, but not so much for higher pressures. This may be due to the nonlinear propagation along the length of the trombone.

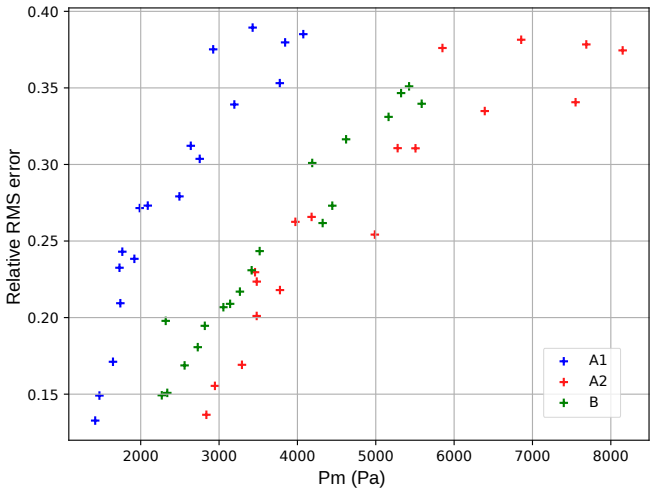

Figure 18: Variation of RMS error as a function of P_m for all three players.

The error in frequency is presented in figure 19 and is compatible with that in figure 5. It proves that the model (2.1) is actually quite good at reproducing the playing frequency, whatever the dynamic of the playing. This indicates that the limits of the model are not so much on the frequency, but more on timbre.

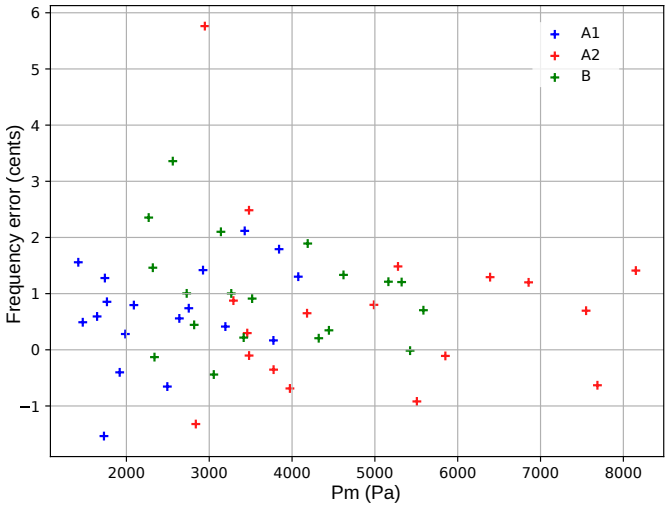

Figure 19: Variation of frequency error as a function of P_m for all three players.

The waveform for two different dynamics are shown in figure 20, both for the measured signal, together with the reconstructed signal from optimized mask parameters. The difference in timbre is clearly seen for the forte recording.

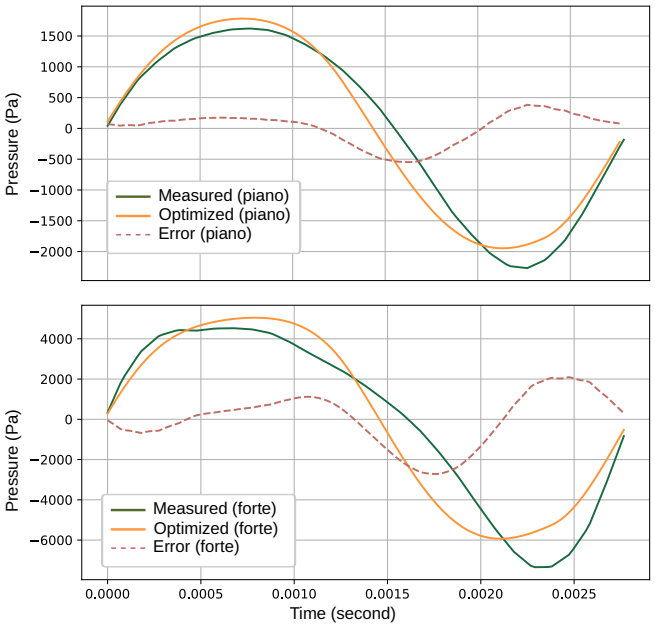

Figure 20: Comparison of wave functions for measured signal vs. optimized signal for musician B. Top for piano (relative RMS error : 20%, playing frequency error 1.5 cents)), and bottom for forte (relative RMS error : 34%, playing frequency error 0.7 cents))

The details of the figures obtained through optimization during a crescendo can be found at Comparison of measured vs. simulated sounds with these parameters can be found at `http://perso.univ-lemans.fr/~smauge/mask`.

4 Conclusion

In this article, a new method is proposed to estimate the mask parameters of a brass musician within a set of acceptable parameters (so called $\mathscr{C}$-admissible). This approach is used on recordings of actual musicians during playing, and is able to deliver a coherent set of parameters (except maybe for the opening at rest), in that they are not too far from existing results in the literature, and their values evolve in a way that is compatible with theory during the playing.

The values obtained prove that a simple model is already capable of reproducing a playing frequency close to that played by an actual musician, with a dynamic and waveform that are similar to the measured ones. This may prove useful for instrument making, although more research should be done to assess the robustness of the method, and investigate the variability of the mask parameters from player to player. Moreover, it can give new leads to a better understanding of intonation, such as the almost linear relation between playing and lip frequencies, or the role of quality factor as a fine tuner.

Furthermore, the system seems to be able to detect when a particular technique is used, for example on the lowest part of bent notes where the vocal tract is used by the musician, as the difference between the techniques is clearly seen on the cost function.

Acknowledgment

The authors would like to thank Christophe Vergez for many discussions and very helpful comments on the manuscript.

In memoriam

This article is dedicated to the memory of Joël Gilbert (1963-2022), who was at the origin of this project and a constant source of ideas.

References

[1] Seiji Adachi and Masa-aki Sato. Time-domain simulation of sound production in the brass instrument. *The Journal of the Acoustical Society of America*, 97(6):3850–3861, June 1995. Publisher: Acoustical Society of America.

[2] Federico Avanzini and Maarten van Walstijn. Modelling the Mechanical Response of the Reed-Mouthpiece-Lip System of a Clarinet. Part I. A One-Dimensional Distributed Model. *Acta Acustica united with Acustica*, 90(3):537–547, May 2004.

[3] Henri Boutin, John Smith, and Joe Wolfe. Trombone lip mechanics with inertive and compliant loads ("lipping up and down"). *The Journal of the Acoustical Society of America*, 147(6):4133–4144, June 2020. Publisher: Acoustical Society of America.

[4] Seona Bromage, Murray Campbell, and Joël Gilbert. Open Areas of Vibrating Lips in Trombone Playing. *Acta Acustica united with Acustica*, 96(4):603–613, July 2010.

[5] Murray Campbell. Brass Instruments As We Know Them Today. *Acta Acustica united with Acustica*, 90(4):600–610, July 2004.

[6] Murray Campbell, Joël Gilbert, and Arnold Myers. *The Science of Brass Instruments*.

[7] Vasileios Chatziioannou, Sebastian Schmutzhard, Montserrat Pàmies-Vilà, and Alex Hofmann. Investigating Clarinet Articulation Using a Physical Model and an Artificial Blowing Machine. *Acta Acustica united with Acustica*, 105(4):682–694, July 2019.

[8] Vasileios Chatziioannou and Maarten van Walstijn. Estimation of Clarinet Reed Parameters by Inverse Modelling. *Acta Acustica united with Acustica*, 98(4):629–639, July 2012.

[9] J. S. Cullen, J. Gilbert, and D. M. Campbell. Brass Instruments: Linear Stability Analysis and Experiments with an Artificial Mouth. *Acta Acustica united with Acustica*, 86(4):704–724, July 2000.

[10] Brigitte d'Andréa Novel, Jean-Michel Coron, and Thomas Hélie. Asymptotic State Observers for a Simplified Brass Instrument Model. *Acta Acustica united with Acustica*, 96(4):733–742, July 2010.

[11] Alain de Cheveigné and Hideki Kawahara. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930, April 2002. Publisher: Acoustical Society of America.

[12] E.J. Doedel. Auto-07p, continuation and bifurcation software for ordinary differential equations. Ver. 0.9.3 https://github.com/auto-07p/auto-07p.

[13] S. J. Elliott and J. M. Bowsher. Regeneration in brass wind instruments. *Journal of Sound and Vibration*, 83(2):181–217, July 1982.

[14] D.J. Ewins. *Modal Testing: Theory, Practice and Application, 2nd Edition | Wiley.*

[15] Vincent Fréour, Louis Guillot, Hideyuki Masuda, Christophe Vergez, and Bruno Cochelin. Parameter identification of a physical model of brass instruments by constrained continuation. *Acta Acustica*, 6:9, 2022. Publisher: EDP Sciences.

[16] R. L. Harrison-Harsley. *Physical Modelling of Brass Instruments using Finite-Difference Time-Domain Methods*. PhD thesis, University of Edinburgh, 2018.

[17] T. Helie, C. Vergez, J. Levine, and X. Rodet. Inversion of a physical model of a trumpet. In *Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No.99CH36304)*, volume 3, pages 2593–2598 vol.3, December 1999. ISSN: 0191-2216.

[18] I. Lopez, A. Hirschberg, A. Van Hirtum, N. Ruty, and X. Pelorson. Physical Modeling of Buzzing Artificial Lips: The Effect of Acoustical Feedback. *Acta Acustica united with Acustica*, 92(6):1047–1059, November 2006.

[19] Rémi Mattéoli, Joël Gilbert, Soizic Terrien, Jean-Pierre Dalmont, Christophe Vergez, Sylvain Maugeais, and Emmanuel Brasseur. Diversity of ghost notes in tubas, euphoniums and saxhorns. *Acta Acustica*, 6:32, 2022. Publisher: EDP Sciences.

[20] Rémi Mattéoli, Joël Gilbert, Christophe Vergez, Jean-Pierre Dalmont, Sylvain Maugeais, Soizic Terrien, and Frédéric Ablitzer. Minimal blowing pressure allowing periodic oscillations in a model of bass brass instruments. *Acta Acustica*, 5:57, 2021. Publisher: EDP Sciences.

[21] Alberto Muñoz Arancón, Bruno Gazengel, Jean-Pierre Dalmont, and Ewen Conan. Estimation of saxophone reed parameters during playing. *The Journal of the Acoustical Society of America*, 139(5):2754–2765, May 2016. Publisher: Acoustical Society of America.

[22] Michael James Newton. Experimental Mechanical and Fluid Mechanical Investigations of the Brass Instrument Lip-reed and the Human Vocal Folds. 2009. Accepted: 2009-10-27T16:06:35Z Publisher: The University of Edinburgh.

[23] Patrice Guyot. Fast python implementation of the yin algorithm. (Version v1.1.1). Zenodo. http://doi.org/10.5281/zenodo.1220947.

[24] Orlando Richards. Investigation of the lip reed using computational modelling and experimental studies with an artificial mouth. 2003. Accepted: 2015-11-04T16:04:16Z Publisher: The University of Edinburgh.

[25] X. Rodet and C. Vergez. Physical models of trumpet-like instruments. detailed behavior and model improvements. *In Proceedings of ICMC (Hong-Kong)*, pages 448–453, 1996.

[26] Tamara Smyth and Jonathan S. Abel. Toward an estimation of the clarinet reed pulse from instrument performance. *The Journal of the Acoustical Society of America*, 131(6):4799–4810, June 2012. Publisher: Acoustical Society of America.

[27] Maarten van Walstijn and Federico Avanzini. Modelling the Mechanical Response of the Reed-Mouthpiece-Lip System of a Clarinet. Part II: A Lumped Model Approximation. *Acta Acustica united with Acustica*, 93(3):435–446, May 2007.

[28] Lionel Velut, Christophe Vergez, Joël Gilbert, and Mithra Djahanbani. How Well Can Linear Stability Analysis Predict the Behaviour of an Outward-Striking Valve Brass Instrument Model? *Acta Acustica united with Acustica*, 103(1):132–148, January 2017.

[29] Christophe Vergez and Patrice Tisserand. The BRASS Project, from Physical Models to Virtual Musical Instruments: Playability Issues. In Richard Kronland-Martinet, Thierry Voinier, and Sølvi Ystad, editors, *Computer Music Modeling and Retrieval*, Lecture Notes in Computer Science, pages 24–33, Berlin, Heidelberg, 2006. Springer.

[30] Y Xiang, D. Y Sun, W Fan, and X. G Gong. Generalized simulated annealing algorithm and its application to the Thomson model. *Physics Letters A*, 233(3):216–220, August 1997.

A Mean opening

Suppose that p and h are of period T, write $^-$ for the mean value of any T-periodic function. Applying it on equation

$$h'' + \frac{\omega_\ell}{Q_\ell} h' + \omega_\ell^2 (h - H) = \frac{p_m - p}{\mu} \tag{A.1}$$

gives

$$\overline{h''} + \frac{\omega_\ell}{Q_\ell} \overline{h'} + \omega_\ell^2 (\overline{h} - H) = \frac{p_m - \overline{p}}{\mu} \tag{A.2}$$

As h is T-periodic, $\overline{h'} = \overline{h''} = 0$, and as $\overline{p} = 0$ we get $\omega_\ell^2 (\overline{h} - H) = \frac{p_m}{\mu}$. So that

$$\overline{h} = H + \frac{p_m}{\mu \omega_\ell^2} \tag{A.3}$$

B Tables

$F_\ell(Hz)$	Q_ℓ	$\mu(kg.m^{-2})$	$H(mm)$	Cost	Frequency (Hz)
184.38090	4.27669	1.27669	0.09	0.11377	233.18
184.33001	4.54898	1.04458	0.08	0.10951	233.18
180.35897	4.34470	1.14597	0.11	0.10177	233.04
178.16236	4.37235	1.17461	0.15	0.09876	232.95
176.77990	4.35483	1.02964	0.09	0.09696	232.98
176.77899	4.33623	1.03454	0.11	0.09614	233.01
175.21228	4.10074	1.10552	0.11	0.09348	233.04
175.21124	4.10265	1.09939	0.13	0.09318	233.08
175.21124	4.10265	1.09939	0.03	0.09000	232.93
175.21816	4.08641	1.05548	0.01	0.08922	232.97
174.56073	4.13190	1.07071	0.03	0.08912	232.91
174.56073	4.13190	1.07071	0.03	0.08909	232.91
174.56073	4.15607	1.07071	0.03	0.08718	232.71
174.56169	4.15607	1.07071	0.03	0.08613	232.47
174.56169	4.15610	1.07071	0.03	0.08608	232.43
174.56160	4.15610	1.07071	0.03	0.08608	232.40
174.56160	4.15610	1.07015	0.03	0.08600	232.56
174.56098	4.15610	1.07015	0.03	0.08593	232.52
174.55610	4.15610	1.07015	0.03	0.08587	232.47
174.55610	4.15539	1.07015	0.03	0.08548	232.49

Table 3: Optimization steps. Each line represents the result of a gradient descent performed by the dual annealing algorithm

Note	$F_\ell(Hz)$	Q_ℓ	$\mu(kg.m^{-2})$	$H(mm)$	$P_m(Pa)$	$F_{play}(Hz)$
Bb2	78.78	3.18	1.33	0.32	896	116.12
Bb3	167.71	3.54	0.57	0.0114	1433	234.26
Bb4	392.90	3.42	0.62	0.452	5134	469.69
D4	232.11	4.56	1.32	0.38	2488	295.62
F3	123.77	2.40	1.22	0.0139	1350	175.45
F4	277.53	3.79	1.68	0.0132	3716	351.19

Table 4: Values obtained by optimization for different notes played by musician B (mezzo forte) for a width $w = 0.012m$.
Comparison of measured vs. simulated sounds with these parameters can be found at `http://perso.univ-lemans.fr/~smauge/mask/#sounds`

s_n	C_n
$-12.24 + 237.80i$	$237.83 + 12.23i$
$-16.01 + 700.45i$	$210.28 + 4.81i$
$-21.08 + 1063.24i$	$222.43 + 4.41i$
$-24.09 + 1438.12i$	$210.97 + 3.53i$
$-26.14 + 1838.74i$	$237.99 + 3.38i$
$-30.00 + 2176.32i$	$268.08 + 3.70i$
$-32.85 + 2532.56i$	$159.47 + 2.07i$
$-35.80 + 2920.16i$	$150.19 + 1.84i$
$-40.08 + 3307.88i$	$175.93 + 2.13i$
$-45.82 + 3707.10i$	$99.70 + 1.23i$
$-276.05 + 4886.60i$	$1036.27 + 58.54i$

Table 5: Values of the coefficients in equation for the impedance decomposition for the Bb trombone in first position 2.2