MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Fast computational aeroacoustics using random distributions of Kirchhoff's spinning vortices with application to sibilant sound generation.

Pont, Arnau,
Guasch, Oriol[1]**,**
Arnela, Marc

GTM Grup de recerca en Tecnologies Mèdia, La Salle - Universitat Ramon Llull
C/ Quatre Camins 30, 08022 Barcelona, Catalonia

ABSTRACT

Hybrid approaches to low Mach number computational aeroacoustics (CAA) for three-dimensional problems have a high computational cost and often demand resorting to supercomputing facilities. The bottleneck concerns the first step of the process, in which a computational fluid dynamics (CFD) simulation is carried out to solve the incompressible Navier-Stokes equations, to obtain the source term for the acoustic wave equation. In this work we suggest that, for some problems in which average results are only needed, it may be possible to avoid the CFD simulation and approximate the flow noise sources by means of a random distribution of Kirchhoff's spinning vortices. In this way, one simply needs to solve an acoustic linear wave operator to solve the aeroacoustics problem. We have applied such methodology to simulate the generation of the sibilant sound /s/ on a realistic geometry for which CAA and experimental data exist. After validation, the versatility of the proposed approach is tested on a simplified geometry, which may be useful to synthesize more complex sound in the future, like syllables. Implementation details of the vortex distribution in a stabilized finite element (FEM) code are also discussed.

Keywords: Fricatives, Finite Element Method, Quadrupoles
I-INCE Classification of Subject Number: 21, 26

1. INTRODUCTION

To generate a sibilant sound like /s/, we press the tongue against the hard palate, which accelerates the flow emanating from the glottis and directs it towards the small gap between the incisors. There, a rapid turbulent jet develops that crosses the cavity between the lower incisors and the lower lips. The strong turbulent eddies in that region,

[1]oriol.guasch@salle.url.edu

as well as their diffraction by the upper incisors, become a source of aerodynamic sound, which we perceive as an /s/ (see e.g., [1–4]).

The physical mechanisms of sibilants' generation has been an object of study for decades (see e.g., [5] for models and references). In recent years, the advent of supercomputer facilities has allowed to perform large-scale simulations based on hybrid computational aeroacoustics (CAA) methods. In this sense, [1] presents a very detailed analysis of the influence of the geometrical parameter variation on the generation of /s/ and /ʃ/ in a simplified rectangular three-dimensional (3D) vocal tract, while [2] provides a qualitative validation of the acoustic output of an /s/ in a realistic geometry.

However, the extreme computational cost of CAA makes the generation of audible sibilant sequences still unaffordable. The bottleneck in the computational cost resides in the solution of the non-linear incompressible Navier-Stokes equations, which is carried out in the first step of hybrid CAA to obtain the acoustic source terms (typically the double divergence of Lighthill's tensor). In this work, we suggest to skip that step by replacing the source term with a tailored, random, distribution of Kirchhoff's spinning vortices [6, 7]. In this fashion, only a linear wave operator needs to be solved. The use of simplified tuned models for /s/ generation is of common practice in 1D approaches to voice synthesis [8,9]. Nonetheless, our goal here is to develop a physics-based model that could emulate the turbulent eddies in the source region of /s/ generation. We validate that model against existing CAA results and also prove its usefulness in simplified vocal tract geometries, which could open the door to synthesize more complex sounds as syllable /sa/.

This paper is structured as follows. Sec. 2 is devoted to presenting the underlying equations, with special attention to the wave equation in mixed form [10], as well as to describe the proposed random distribution of Kirchhoff's vortices that should match the large eddy simulations (LES) in [2]. Next, in Sec. 3 we introduce the stabilized FEM formulation used to discretize the equations, placing primary emphasis on the treatment of the suggested vortex distribution. The results from numerical simulations are shown in Sec. 4. The validation carried out in [4] is briefly reviewed and the outcomes of the simulations on a simplified vocal tract are presented. Conclusions are drawn in Sec. 5.

2. PROBLEM STATEMENT

2.2.1. The standard hybrid CAA approach

The most common approach for simulating low Mach number flow noise resorts to hybrid CAA exploiting Lighthill's acoustic analogy. A CFD simulation of the incompressible Navier-Stokes equations is carried out in the first step of method to compute the acoustic source term, which takes most of the computational cost. The source term, namely the double divergence of Lighthill's tensor $\boldsymbol{T}$ [11], is then plugged into a linear wave operator that propagates sound to the far-field, namely

$$c_0^{-2}\partial_{tt}p - \nabla^2 p = (\nabla\otimes\nabla) : \boldsymbol{T} \approx \rho_0\,(\nabla\otimes\nabla) : (\boldsymbol{u}^0\otimes\boldsymbol{u}^0) \quad \text{in } \Omega,\ t>0. \tag{1}$$

In Equation 1, p denotes the acoustic pressure, ρ_0 the air density, $\boldsymbol{u}^0$ the incompressible flow velocity and c_0 the speed of sound. Note that we have made use of the standard approximation of the Lighthill's tensor for low Mach numbers. For the sake of conciseness, boundary and initial conditions have not been specified.

The computational cost of the second step of the hybrid CAA (i.e. solving Equation 1) is much more affordable than the first one because the wave equation is linear and,

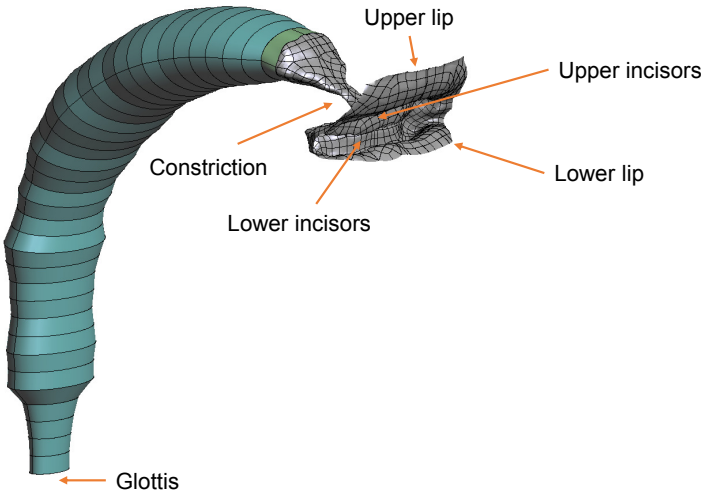

Figure 1: Realistic 3D vocal tract for sibilant /s/.

moreover, one does not require the very fine meshes needed to resolve turbulent flows. As mentioned in the Introduction, the main focus of this work consists in finding an alternative procedure to derive the acoustic sources involved in the generation of sibilant sound /s/. This would prevent the use of supercomputer facilities as in [1, 2] to produce that sound.

2.2.2. Random distribution of Kirchhoff's spinning vortices

Given that the acoustic source term in the right hand side of Equation 1 consists in a distribution of quadrupoles, it has been recently proposed in [4] that it could be well approximated by means of a random distribution of Kirchhoff's spinning vortices (see e.g., [6, 7]). The idea is to replace Lighthill's tensor, $\boldsymbol{T}$, with the expression,

$$\boldsymbol{T} \approx \sum_{n=1}^{N} \boldsymbol{T}_n(\boldsymbol{Y}_n, t)\delta(X_n), \tag{2}$$

where N stands for the number of vortices and δ for the Dirac generalized function; X_n indicating the location of the n-th vortex. $\boldsymbol{T}_n$ designates the vortex strength, which depends on the parameter vector $\boldsymbol{Y}_n = (Y_1 \dots Y_P)^\top$, P being the number of properties needed to describe $\boldsymbol{T}_n$. A Kirchhoff vortex is a two-dimensional rotating eddy with elliptic cross-section and uniform vorticity modulus Ω_n, which leads to an angular frequency $\Omega_n/4$. The vortex radiates sound with a quadrupolar sound radiation pattern at frequency $f_n = \Omega_n/4\pi$.

To reproduce the turbulence regime encountered at the sibilant generation region, between the inter-dental gap and the lower lips, the distribution in Equation 2 is considered to be random, with a uniform probability density function assigned to X_n. Each vortex is also assigned a random lifetime and the P parameters in $\boldsymbol{Y}$, are also identified as random variables. For a vortex spinning in the yz sagittal plane of the vocal

tract geometry in Fig. 1, the expression for $\boldsymbol{T}_n$ reads

$$\boldsymbol{T}_n(t) = 2\pi^3 f_n^2 \epsilon_n a_n^4 \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi f_n t) & \sin(2\pi f_n t) \\ 0 & \sin(2\pi f_n t) & -\cos(2\pi f_n t) \end{pmatrix} \left[\sin^2(2\pi\beta_n t) - \sin^2(2\pi\beta_n t - 1/2)\right], \tag{3}$$

with $a_n(1 \pm \epsilon_n)$ being the semi-major and minor axes of an ellipse with eccentricity ϵ_n. The eccentricity has been fixed constant for simplicity in the forthcoming expressions. The remaining elements in Equation 3 are the frequency, the rotation matrix and a Hann windowing which determines the finite lifetime of the vortex in the semiperiod $[0, \beta_n^{-1}/2]$. The parameter vector of random variables in Equation 2 is therefore identified with $\boldsymbol{Y}_n = (a_n, f_n, \beta_n)^\top$.

The next task is to assign proper sampling spaces to the variables in $\boldsymbol{Y}_n$, such that the distribution in Equation 2 could match the outputs from the CAA simulations on realistic geometries, like the one in Fig. 1 (see [2]). For instance, the parameters in $\boldsymbol{Y}_n$ should guarantee that the energy distribution of the vortices is that obtained from a CFD simulation of the turbulent flow on the vocal tract (see Fig. 2, derived from [2]). In that figure, three different regions can be distinguished, the last one corresponding to the celebrated Kolmogorov inertial range. To reproduce them, it is found that a_n should behave as follows,

$$a_n = \begin{cases} O(f_n^{-3/5}) & \text{if } 0.4 \text{ kHz} < f_n \leq 2.3 \text{ kHz}, \\ O(f_n^{-0.93}) & \text{if } 2.3 \text{ kHz} < f_n \leq 12 \text{ kHz}, \\ O(f_n^{-11/6}) & \text{if } f_n > 12 \text{ kHz}. \end{cases} \tag{4}$$

The first two slopes have been obtained heuristically, whereas the last one has been adapted to match with Kolmogorov's energy cascade. In that range the kinetic energy is characterized by $\boldsymbol{v}_n^2 \sim k_n^{-5/3}$, where $\boldsymbol{v}_n$ represents the vortex tangential speed. Because $k_n = 2\pi f_n / c_0$, $\lambda_n \sim O(l_n/M)$ and $\boldsymbol{v}_n^2 = 4\pi^2 a_n^2 f_n^2$, the average size of the vortices shall

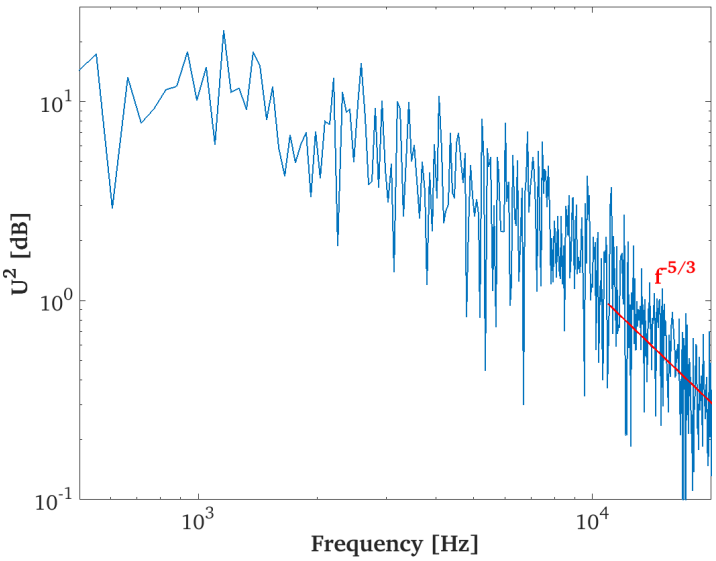

Figure 2: Energy spectrum of the flow velocity obtained with an incompressible LES at a point in the noise generation region.

behave as $a_n \sim O(f_n^{-11/6})$. In order to avoid unrealistic sizes, f_n has been limited to values above 400 Hz (see [4] for full details).

As regards the frequency, f_n, it has been assigned a uniform probability distribution to cover the whole audible range. Besides, and as said before, the parameter β_n is introduced to assign the vortices a finite life time. This avoids tracking them, which is complex in an Eulerian flow description. The arrival and departure of a vortex at a given location X_n is represented by a $\beta_n^{-1}/2$ Hann windowing, with $\beta_n \in [0.4, 2.3]$ kHz. This provides a smooth activation and deactivation of T_n. A minimum value of $\beta_n = 0.4$ kHz has been set to avoid excessive vortex lifetimes, while an upper limit of $\beta_n = 2.3$ kHz has been imposed consistent with the low frequency range of Equation 4. Those choices reinforce the dominance of the low frequency scales found in Fig. 2. However, this model interacts with the spectrum assigned to the vortices in Equation 2. Consequently, the heuristic process that has led to the correlation for a_n in the [0.4, 2.3] kHz range has considered the windowing effect.

The random vortices are assigned to $\sim$ 6% of the points in the source region. This relatively low density avoids excessive shadowing between them and reproduces well the results from the CAA simulations in [2].

2.2.3. Weak formulation of the problem

As mentioned in the Introduction, the formulation has been implemented in a finite element code which involves dealing with the weak form of the problem. That corresponding to Equation 1 can be derived as usual, first multiplying the equation by a pressure test function, q, and then integrating over the computational domain Ω. The mathematical problem becomes that of finding $p(\cdot, t)$ for all $t > 0$ such that,

$$\begin{aligned} c_0^{-2}(\partial_{tt}^2 p, q) + (\nabla p, \nabla q) - c_0^{-1}(\partial_t p, q)_{\Gamma_I \cup \Gamma_\infty} + \mu_W c_0^{-1}(\partial_t p, q)_{\Gamma_W} \\ = \sum_{n=1}^{N} \left[-(\nabla q, \nabla \cdot [\boldsymbol{T}_n' \delta(X_n)]) + (q, \nabla \cdot [\boldsymbol{T}_n' \delta(X_n)] \cdot \boldsymbol{n})_{\partial\Omega} \right] \end{aligned} \tag{5}$$

for all q. For the sake of brevity and simplicity, the description of the functional spaces where the solution and test functions belong will be omitted in this work. Equation 5 is to be supplemented with appropriate initial conditions.

Unfortunately, the inclusion of the source term of Equation 2 into Equation 5 leads to the presence of the derivative of a delta function in the right hand side. In principle, this could be avoided by making use of the wave equation in mixed form [10, 12], which involves the acoustic particle velocity, $\boldsymbol{u}$, along with the acoustic pressure. The mixed formulation of Equation 1 reads,

$$\frac{1}{\rho_0 c_0^2} \partial_t p + \nabla \cdot \boldsymbol{u} = 0, \tag{6}$$

$$\rho_0 \partial_t \boldsymbol{u} + \nabla p = -\sum_{n=1}^{N} \nabla \cdot [\boldsymbol{T}_n' \delta(X_n)] \,. \tag{7}$$

The weak for of the above expressions is now found multiplying Equation 6 by a pressure test function, q, and Equation 7 by a velocity test function, $\boldsymbol{v}$, so that once having

integrated over the computational domain Ω we are left with,

$$\frac{1}{\rho_0 c_0^2}(\partial_t p, q) - (\boldsymbol{u}, \nabla q) + \frac{1}{\rho_0 c_0}(p, q)_{\Gamma_I \cup \Gamma_\infty} - \frac{\mu_W}{\rho_0 c_0}(p, q)_{\Gamma_W} = 0,$$

$$\rho_0(\partial_t \boldsymbol{u}, \boldsymbol{v}) + (\nabla p, \boldsymbol{v}) = \sum_{n=1}^{N}(\boldsymbol{T}_n' \delta(X_n), \nabla \boldsymbol{v}), \qquad (8)$$

for all q and $\boldsymbol{v}$. Note that no Dirac's delta function derivative appears in Equation 8 and that the boundary source term stemming from the integration by parts, namely $(\boldsymbol{T}_n' \delta(X_n) \cdot \boldsymbol{n}, \boldsymbol{v})_{\partial\Omega}$, disappears because the test functions vanish on the domain borders. Working with the wave equation in mixed form has another advantage, it is the natural choice if one wanted to generate dynamic sounds like a diphthong (see [13]) or a syllable like /sa/.

3. NUMERICAL DISCRETIZATION

3.3.1. Spatial discretization

The discretization of Equation 8 has been carried out with the method of lines. A second order backwards finite difference scheme has been implemented for the time discretization, while finite elements have been used for the spatial one. The former is standard so let us focus on the latter because it has some implications concerning the source term in Equation 2. First we note that unlike the wave propagation problem in Equation 1, the mixed problem of Equation 8 is constrained by an inf-sup compatibility condition for p and $\boldsymbol{u}$, which prevents using equal interpolation orders for them. To circumvent that condition, one can resort to stabilization methods, like the variational multiscale (VMS) ones [14, 15], which have been already applied to the wave equation in mixed form in [10, 13, 16].

According to those works, given a finite element partition of the computational domain Ω, with n_{el} elements and n_p nodes, the stabilized FEM approach to Equation 8 consists in finding $p_h(\cdot, t)$ and $\boldsymbol{u}_h(\cdot, t)$ for all $t > 0$, such that

$$\frac{1}{\rho_0 c_0^2}(\partial_t p_h, q_h) - (\boldsymbol{u}_h, \nabla q_h) + \frac{1}{\rho_0 c_0}(p_h, q_h)_{\Gamma_I \cup \Gamma_\infty} - \frac{\mu_W}{\rho_0 c_0}(p_h, q_h)_{\Gamma_W} + \sum_{n_{el}}(\tau_p \nabla \cdot \boldsymbol{u}_h, \nabla \cdot \boldsymbol{v}_h)_{\Omega_e} = 0,$$

$$\rho_0(\partial_t \boldsymbol{u}_h, \boldsymbol{v}_h) + (\nabla p_h, \boldsymbol{v}_h) + \sum_{n_{el}}(\tau_{\boldsymbol{u}} \nabla p_h + \tau_{\boldsymbol{u}} \sum_{n=1}^{N} \nabla \cdot [\boldsymbol{T}_n' \delta(X_n)], \nabla q_h)_{\Omega_e} = \sum_{n=1}^{N}(\boldsymbol{T}_n' \delta(X_n), \nabla \boldsymbol{v}_h), \qquad (9)$$

where τ_p and $\tau_{\boldsymbol{u}}$ are stabilization parameters whose expressions can be found e.g., in [10, 12, 13]. Note that with the mixed formulation we got rid off the delta function derivative, but that recurs in the stabilization term of the second line in Equation 9. We shall next see how to deal with that term, as well as with the force one in the right hand side of the same equation.

3.3.2. Discretized distribution of Kirchhoff's vortices

The trick to deal with the terms involving delta functions in Equation 9 consists in properly smoothing them. Let us refer to the finite element shape functions as N^a, with a running from 1 to n_p and designating a node located at $\boldsymbol{x}^a$. Both, the unknowns and

test functions of the problem can be expanded in terms of the shape function basis. For example, the i-th component of the velocity can be expressed as $u_{hi} = \sum_a N^a(\boldsymbol{x}) U_i^a$, with U_i^a standing for the nodal value $u_{hi}(\boldsymbol{x}^a)$.

The smoothing of the force term of Equation 9 takes place in two stages. First, a vortex is no longer considered to be a point source at X_n, but to occupy an entire finite element. This allows us to replace the problematic delta function $\delta(X_n)$ by a sort of Heaviside function Π_K, such that $\Pi_K(X_n) = 1$ if $X_n \in \text{int}(K)$, K being a particular element domain, and zero elsewhere. Then, the contribution of a single vortex becomes

$$(\boldsymbol{T}_n' \delta(X_n), \nabla \boldsymbol{v}_h) \simeq (\boldsymbol{T}_n' \Pi_K(X_n), \nabla \boldsymbol{v}_h) = \sum_a V_i^a \int_\Omega \boldsymbol{T}_n'|_{ij} \Pi_K(X_n) \partial_j N^a d\Omega$$
$$= \sum_a V_i^a \sum_{n_e} \int_{\Omega_e} \boldsymbol{T}_n'|_{ij} \Pi_K(X_n) \partial_j N^a d\Omega_e = \sum_a V_i^a \boldsymbol{T}_n'|_{ij} \int_K \partial_j N^a dK, \quad (10)$$

where the summation convention over spatial repeated indexes is assumed. Unfortunately, numerical tests revealed that Equation 10 still suffers from numerical instabilities, so an additional smoothing was applied by L_2 projecting $\boldsymbol{T}_n' \Pi_K(X_n)$ onto the finite element space. That is found as,

$$\mathcal{P}(\psi_h) = \sum_{a=1}^{n_u} N^a(\boldsymbol{x}) P^a, \quad (11)$$

where the coefficients P^a are given by the solution of the linear system

$$\sum_a M^{ba} P^a = \int_\Omega N^b \psi_h d\Omega, \quad b = 1, \ldots, n_p \quad (12)$$

$$M^{ba} := \int_\Omega N^b N^a d\Omega. \quad (13)$$

If we transform $\boldsymbol{M}$ in (13) into a lumped diagonal matrix $\text{diag}(\mathcal{M}_{11}, \ldots, \mathcal{M}_{n_p n_p})$ using a nodal quadrature rule, the following approximation of $(\boldsymbol{T}_n' \delta(X_n), \nabla \boldsymbol{v}_h)$ can replace that in Equation 10,

$$(\boldsymbol{T}_n' \delta(X_n), \nabla \boldsymbol{v}_h) \simeq (\mathcal{P}[\boldsymbol{T}_n' \Pi_K(X_n)], \nabla \boldsymbol{v}_h) = \sum_a \sum_b V_i^a \Big[\int_\Omega \partial_j N^a N^b d\Omega \Big] P^b$$
$$= \sum_a \sum_b V_i^a \Big[\int_\Omega \partial_j N^a N^b d\Omega \Big] \mathcal{M}_{bb}^{-1} \Big[\int_\Omega \boldsymbol{T}_n'|_{ij} \Pi_K(X_n) N^b d\Omega \Big]$$
$$= \sum_a \sum_b V_i^a \Big[\int_\Omega \partial_j N^a N^b d\Omega \Big] \mathcal{M}_{bb}^{-1} \Big[\boldsymbol{T}_n'|_{ij} \int_K N^b dK \Big]. \quad (14)$$

In what concerns smoothing the stabilization term $(\tau_{\boldsymbol{u}} \nabla \cdot [\boldsymbol{T}_n' \delta(X_n)], \nabla q_h)_{\Omega_e}$ in Equation 9, we have proceed with a first projection onto the finite element space $(\tau_{\boldsymbol{u}} \mathcal{P}\{\nabla \cdot [\boldsymbol{T}_n' \Pi_K(X_n)]\}, \nabla q_h)_{\Omega_e}$, followed by an integration by parts restricting the source to the interior of the elements, which provides the nodal vector values

$$\boldsymbol{P}^b = \mathcal{M}_{bb}^{-1} \int_\Omega N^b \nabla \cdot [\boldsymbol{T}_n' \Pi_K(X_n)] d\Omega = -\mathcal{M}_{bb}^{-1} \int_K \nabla N^b \cdot \boldsymbol{T}_n' dK. \quad (15)$$

Therefore, the final expression for the stabilization term $(\tau_{\boldsymbol{u}} \mathcal{P}\{\nabla \cdot [\boldsymbol{T}_n' \Pi_K(X_n)]\}, \nabla q_h)_{\Omega_e}$ becomes

$$(\tau_{\boldsymbol{u}} \mathcal{P}\{\nabla \cdot [\boldsymbol{T}_n' \Pi_K(X_n)]\}, \nabla q_h)_{\Omega_e} = -\sum_a \sum_b Q^a \Big[\int_\Omega \partial_j N^a N^b d\Omega \Big] \mathcal{M}_{bb}^{-1} \Big[\boldsymbol{T}_n'|_{ij} \int_K \partial_i N^b dK \Big]. \quad (16)$$

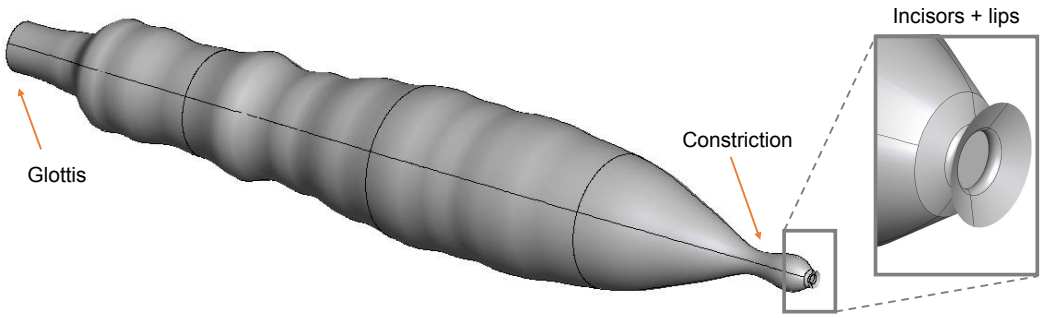

Figure 3: Simplified vocal tract geometry for sibilant /s/.

4. RESULTS

As regards the numerical simulations, two aspects will be considered. The first one concerns the validation of the proposed Kirchhoff's vortex random model in a realistic geometry like that in Fig. 1. This task has been very recently carried out in detail in [4], through comparison with the CAA results presented in [2] and the experimental data in [17]. It will be shortly reviewed below. The second aspect we want to address concerns checking whether the vortex model could be applied to further geometries than the realistic one used to adjust the model. To that purpose the simplified axisymmetric vocal tract geometry in Fig. 3 will be considered. Could the model perform well in such simple geometries that would enormously facilitate future steps like generating syllables using them (e.g., /sa/).

4.4.1. Validation of the Kirchhoff vortex random model as a source for fricative /s/

The realistic oral cavity geometry used to validate the Kirchhoff's vortex model in [4] with the CAA in [2] was extracted from [18] and [19]. The random model of spinning vortices was prescribed in the space between the incisors and the lower lip, where most turbulent eddies get generated according to [2]. The kinetic energy of various realizations of the random model were first verified in [4] to satisfy the three slope behavior of Fig. 2, obtained from the CFD step of the CAA in [2]. The Kirchhoff's vortex model therefore succeed in emulating the physical patterns of flow turbulence (we do not reproduce those results here though).

A second, and critical, validation feature was that of checking whether the model was capable of reproducing the characteristic spectrum of phoneme /s/. In this regard, the Welch power spectrum at a point located outside the mouth was computed in [4] from the average of 20 random realizations of the model. That average spectrum is reproduced in Fig. 4 (dashed black line) and compared to that from the CAA in [2] (dotted red line) and the experimental results in [17] (dotted blue line). The resemblance of the three curves is apparent, showing a strong dip near 2 kHz, followed by a peak close to $\sim$ 4 kHz and a high energy concentration between 8 and 10 kHz.

We would like to emphasize that one of the most relevant outcomes of the proposed method concerns the reduction of the computational mesh size. The mesh in the simulations has been dimensioned to capture the smallest audible wavelength involved in sibilant /s/ generation, namely $\lambda = 0.0175$ m, corresponding to $\sim$ 12 kHz. For a proper resolution, the bulk element size has been set to $h = 0.002$ m, with a finer remeshing at the occlusion because of the complexity of the geometry. As a result, the $\sim 4.5 \times 10^7$

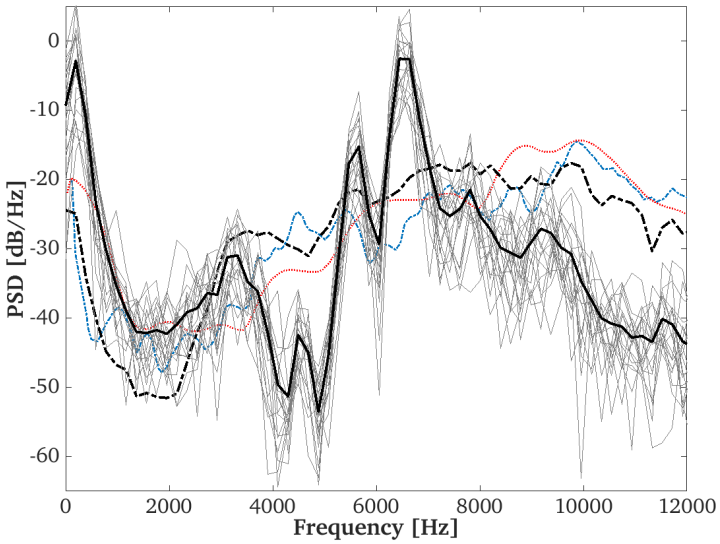

Figure 4: Welch power spectrum at P2 for different realizations of sibilant /s/ using a simplified geometry (continuous grey lines) with its corresponding average value (continuous black line). Equivalent spectra using a realistic vocal tract geometry: Kirchhoff's vortex model (dashed black line), CAA computation (dotted red line) and experimental values (dotted blue line).

elements of the CAA model in [2] get reduced to only $\sim 7 \times 10^5$ linear P1/P1 elements. In regard to the parameters in Equation 9, a time step $\delta t = 2.5 \times 10^{-6}$ s has been used, together with an air density $\rho_0 = 1.2$ kg/m^3 and a sound speed $c_0 = 350$ m/s.

4.4.2. Generation of sibilant fricative /s/ using a full length simplified vocal tract

As mentioned before, it would be very convenient that the proposed vortex model could also work well for simplified geometries. This is so because working with realistic geometries has one important disadvantage: the difficulty to extend them in an Arbitrary Lagrangian Eulerian reference (ALE) for dynamic phonation problems, such as the generation of diphthongs and syllables, since all the nodes of the vocal tract boundary nodes need to follow a compatible deformation mapping from the initial to the final shape. This geometrical compliance becomes much easier for simplified geometries with circular or elliptic cross sections, where some tricks can be exploited to avoid costly remeshing strategies [13, 20]. For this reason, the realistic-like geometry in Fig. 1 is replaced with that in Fig. 3. The latter has been obtained using the 1D area functions in [21] to generate an axisymmetric 3D vocal tract with circular cross-sections, which has been somewhat tuned with an additional extension beyond the constriction to consider the incisors. The challenge now consists in reproducing an acoustic output for /s/ similar to the ones from the previous section, depicted in Fig. 4.

It is to be mentioned that the drastic geometric simplifications have required tuning one parameter of the model: the vortex density. In the previous section that was set to 6% but now has been increased to 35%. This makes sense if one considers that the region where the source is now prescribed is much smaller than that from the realistic vocal tract, due to the symmetry of the occlusion. Proceeding like in [4], we have ran several realizations of the model and computed the corresponding Welch power spectra at the same point in the outer domain as for the realistic geometry. Those spectra, together with their average

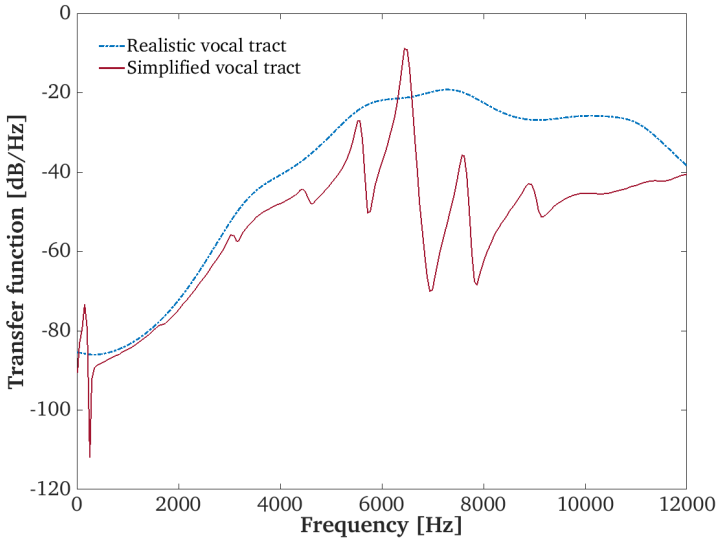

Figure 5: Impulsional transfer function of the realistic and the simplified geometries.

value, are also plotted in Fig. 4. As observed from the figure, the general trends of the realistic /s/ are recovered except for two big differences. The first and most remarkable one is the large dip in the $4-5$ kHz region, followed by the strong peaks at 6 and 8 kHz. The explanation for that behavior comes from the difference between the impulsional transfer function of the simplified vocal tract geometry and that of the realistic one, which have been plotted in Fig. 5. While the transfer function of the realistic vocal tract increases monotonically up to 8 kHz, that from the simplified geometry becomes smaller beyond 4 kHz and exhibits a series of dips and peaks which coincide with those observed in Fig. 4. For instance, the two resonances around 6 kHz in Fig. 4 get clearly reflected in Fig. 5. The high frequency peaks in the transfer function of the simplified geometry might be caused by an excessive length of the appended extension and/or the shape of the vocal tract exit. By modifying such elements, the peaks in the transfer function may move to much higher frequencies and better resemble that of the realistic vocal tract.

The other big difference between the spectra of the realistic and simplified vocal tracts concerns the high frequency content. As observed, when the Kirchhoff's vortex model is applied to the simplified vocal tract it is unable to generate enough acoustic pressure at high frequencies. This is partly justified by the lower values of the transfer function at this range but also by the circular symmetry of the vocal tract geometry. Probably, the occlusion of a vocal tract with elliptical cross-sections (see e.g., [22, 23]) would be closer to the realistic one and produce better results.

5. CONCLUSIONS

In this paper, we have suggested to construct an acoustic source term based on a random distribution of Kirchhoff's spinning vortices, to avoid long CAA computations in order to generate sibilant sounds. The new source model has been validated against CAA computations and experimental results in realistic 3D vocal tract geometries, showing great potential to produce sibilants at a low computational cost.

Besides, the vortex model has been also tested in a simplified vocal tract geometry to check whether it could be exploited to generate more complex sounds, like syllables. The

results revealed some remarkable discrepancies with those obtained from a realistic vocal tract. An analysis of the realistic and simplified vocal tract transfer functions has revealed that those discrepancies may be probably attributed to inaccuracies in the simplified vocal tract, rather than to the performance of the Kirchhoff's vortex model. Future efforts will be placed on improving the simplified vocal tracts so that their response could better resemble that of realistic ones. Changing the shape, location and size of the constrictions, and replacing circular cross-sections with elliptical ones could probably improve the herein reported results.

6. ACKNOWLEDGEMENTS

This research has been supported by the Agencia Estatal de Investigación (AEI) and FEDER, EU, through project GENIOVOX TEC2016-81107-P. The second author would also like to thank l'Obra Social de la Caixa and the Universitat Ramon Llull for their support under grant 2018-URL-IR2nQ-031. The authors also gratefully acknowledge the International Center for Numerical Methods in Engineering for their support with the computational code FEMUSS and the cluster HPC0.

7. REFERENCES

[1] T. Yoshinaga, K. Nozaki, and S. Wada. Experimental and numerical investigation of the sound generation mechanisms of sibilant fricatives using a simplified vocal tract model. *Phys. Fluids*, 30(3):035104, 2018.

[2] A. Pont, O. Guasch, J Baiges, R. Codina, and A. Van Hirtum. Computational aeroacoustics to identify sound sources in the generation of sibilant /s/. *Int. J. Numer. Meth. Biomed. Eng.*, 35(1):e3153, 2019.

[3] O. Guasch, A. Pont, J. Baiges, and R. Codina. Concurrent finite element simulation of quadrupolar and dipolar flow noise in low mach number aeroacoustics. *Comput. Fluids*, 133:129–139, 2016.

[4] A. Pont, O. Guasch, and M. Arnela. Finite element generation of sibilants /s/ and /z/ using random distributions of kirchhoff's vortices. *Submitted*, 2019.

[5] M.S. Howe and R.S. McGowan. Aeroacoustics of [s]. *Proc. R. Soc. A*, 461:1005–1028, 2005.

[6] M.S. Howe. Contributions to the theory of aerodynamic sound, with application to excess jet noise and the theory of the flute. *J. Fluid Mech.*, 71(4):625–673, 1975.

[7] M.S. Howe. *Acoustics of fluid-structure interactions*. Cambridge university press, 1998.

[8] S. Narayanan and A. Alwan. Noise source models for fricative consonants. *IEEE Trans. Audio Speech Lang. Process.*, 8(3):328–344, 2000.

[9] B.H. Story. Phrase-level speech simulation with an airway modulation model of speech production. *Comput. Speech Lang.*, 27(4):989–1010, 2013.

[10] R. Codina. Finite element approximation of the hyperbolic wave equation in mixed form. *Comput. Methods Appl. Mech. Engrg.*, 197(13–16):1305–1322, 2008.

[11] M.J. Lighthill. On sound generated aerodynamically i. general theory. *Proc. R. Soc. Lond. A*, 211(1107):564–587, 1952.

[12] S. Badia, R. Codina, and H. Espinoza. Stability, convergence and accuracy of stabilized finite elements methods for the wave equation in mixed form. *SIAM J. Numer. Anal.*, 52:1729–1752, 2014.

[13] O. Guasch, M. Arnela, R. Codina, and H. Espinoza. A stabilized finite element method for the mixed wave equation in an ALE framework with application to diphthong production. *Acta Acust. united Ac.*, 102(1):94–106, 2016.

[14] T.J.R. Hughes. Multiscale phenomena: Green's function, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized formulations. *Comput. Methods Appl. Mech. Engrg.*, 127:387–401, 1995.

[15] T.J.R. Hughes, G. Feijóo, L. Mazzei, and J. Quincy. The variational multiscale method, a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166:3–24, 1998.

[16] R. Codina, J.M. González-Ondina, G. Díaz-Hernández, and J. Principe. Finite element approximation of the modified Boussinesq equations using a stabilized formulation. *Int. J. Numer. Meth. Fluids*, 57(9):1249–1268, 2008.

[17] K. Nozaki, T. Yoshinaga, and S. Wada. Sibilant/s/simulator based on computed tomography images and dental casts. *J. Dent. Res.*, 93(2):207–211, 2014.

[18] A. Van Hirtum, Y. Fujiso, and K. Nozaki. The role of initial flow conditions for sibilant fricative production. *J. Acoust. Soc. Am.*, 136(6):2922–2925, 2014.

[19] Y. Fujiso, K. Nozaki, and A. Van Hirtum. Towards sibilant physical speech screening using oral tract volume reconstruction: Some preliminary observations. *Appl. Acoust.*, 96:101–107, 2015.

[20] M. Arnela and O. Guasch. Finite element synthesis of diphthongs using tuned two-dimensional vocal tracts. *IEEE/ACM Trans. Audio Speech Lang. Process.*, 25(10):2013–2023, 2017.

[21] B.H. Story. A parametric model of the vocal tract area function for vowel and consonant simulation. *J. Acoust. Soc. Am.*, 117(5):3231–3254, 2005.

[22] M. Arnela and O. Guasch. Finite element computation of elliptical vocal tract impedances using the two-microphone transfer function method. *J. Acoust. Soc. Am.*, 133(6):4197–4209, 2013.

[23] M. Arnela, S. Dabbaghchian, R. Blandin, O. Guasch, O. Engwall, A. Van Hirtum, and X. Pelorson. Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds. *J. Acoust. Soc. Am.*, 140(3):1707–1718, 2016.