Robustness and reliability of state-space, frame-based modeling for thermoacoustics

Mathieu Cances, Luc Giraud, Michael Bauerheim, Laurent Y.M. Gicquel, Franck Nicoud

Robustness and Reliability of state-space, frame-based modeling for Thermoacoustics.

Mathieu Cances[a,f,*], Luc Giraud[d], Michael Bauerheim[e], Laurent Gicquel[a], Franck Nicoud[b,c]

[a]*CERFACS, 42 Av. Gaspard Coriolis, 31100 Toulouse, France*
[b]*IMAG, University of Montpellier, CNRS, Montpellier, France*
[c]*Institut Universitaire de France (IUF), Paris, France*
[d]*Inria, joint Inria, Airbus CR & T and Cerfacs, France*
[e]*ISAE-Supaero, 10 Avenue Edouard-Belin, Toulouse Cedex 4 31055, France*
[f]*Safran Aircraft Engines, 77550 Moissy-Cramayel, France*

Abstract

The Galerkin modal expansion is a well-known method used to develop reduced order models for thermoacoustics. A known issue is the appearance of Gibbs-type oscillations on velocity fluctuations at the interface between subdomains and at boundary conditions. Recent work of Laurent *et al.*, Comb. and Flame, **206**, (2019) and Laurent, Badhe and Nicoud, J. of Comp. Physics, **428** (2021) have shown that it is possible to overcome this issue by using an over-completed frame, instead of a Galerkin modal basis. However, the low-order modeling based on this frame modal expansion may generate spurious modes. In this paper, the origin of these non-physical modes is identified and a method is proposed to automatically remove them from the outcome. By preventing any interaction between the physical and non-physical components, the proposed methodology drastically improves the robustness and reliability of the frame modal expansion modeling for thermoacoustics.

Keywords: Thermoacoustic instabilities, Low-Order model, Modal expansion, State-space, Singular Value Decomposition

Introduction

Any combustor may be prone to thermoacoustic instabilities, which correspond to the constructive coupling between the acoustic modes of the geometry and the unsteady heat release rate related to the flame [1]. Pressure fluctuations due to these instabilities can reach values of the order of a large fraction of the mean pressure, leading to serious consequences like the flashback of the flame and the destruction of the combustor. While Large Eddy Simulation (LES) is a well-known method that can accurately predict thermoacoustic instabilities [2, 3], it is not suitable for the engine design step due to its high computational cost. Thus, developing reliable and computationally efficient low-order models (LOMs) for thermoacoustic instabilities is of great importance. In this context, it is common to consider the evolution of small amplitude velocity/pressure fluctuations superimposed to a stationary and zero Mach number baseline flow. The thermoacoustic system is represented by a Helmholtz equation forced by a combustion term [4], which can be numerically solved using Helmholtz solvers, thus avoiding expensive LES calculations (see Fig. 1). The forcing term, which represents the effect of the unsteady rate of heat release on the pressure fluctuations, must be modeled as a function of a purely acoustic quantity (either pressure or acoustic velocity) [5, 6, 7]. This dependency is usually expressed through a flame transfer function (FTF) depending on the frequency. The FTF can also depend on the amplitude of the velocity oscillations in which case it is called a FDF (Flame Describing Function) [8, 9]. FTFs and FDFs required for any Helmholtz solver can be obtained analytically in simple cases [10, 11], numerically [3, 12] or experimentally [13]. Regardless of the method used to obtain the flame dynamics, once acquired, employing a Helmholtz solver allows for the computation of thermoacoustic instabilities.

Even for Helmholtz solvers, optimizing numerical methods and reducing computational cost is crucial. In the case of simple 1D-configurations, the problem can be simplified by making assumptions on the geometry, leading to pseudo-analytical models that are computationally free (see bottom right of Fig. 1). A canonical example of such a method is the ATACAMAC approach [14], where a 1.5-dimensional model is developed to compute azimuthal thermoacoustic modes in an annular chamber fed by an annular plenum through N burners. For more complex 3D geometries, the coupled Helmholtz equation is not reducible to a simple set of equations. The Finite Element Method (FEM) must be employed to solve it on a discretized mesh of a complex geometry. The mesh describes each element of the system and the minimum cell size is defined through the minimum wavelength of the modes of interest. This typically leads to a number of vertices $N_v \sim 10^6$ on which the coupled Helmholtz equation is solved.

⋆PhD student, CERFACS, Safran Aircraft Engines
*Tel.: +33-789625276;
Email address: cances@cerfacs.fr (Mathieu Cances)

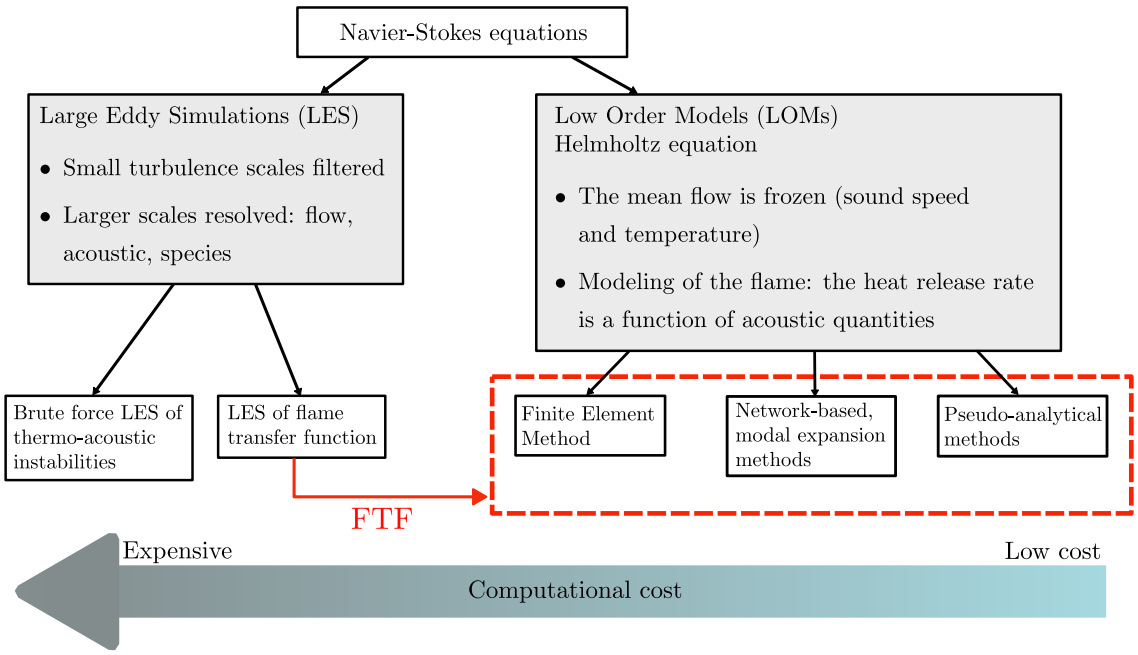

Figure 1: Numerical methods for thermoacoustic instabilities. On the left, the LES method. On the right, the class of Low-Order Models, which require a Flame Transfer Function (FTF) as an input. Adapted from Poinsot [1]

Numerically, the solver has to compute the solutions of a non-linear eigenvalue system $A(\omega)P = -\omega^2 P$, where the A matrix is of size $N_v \times N_v$ and depends on the searched frequency ω. This dependency requires an iterative process to solve the problem, in which an initial guess ω_0 must be provided to the solver by the user. Algorithms based on this method return the converged frequency ω corresponding to the thermoacoustic mode with the eigenfrequency closest to ω_0 [15]. At each iteration the A matrix is eigen-solved and at the end, only one thermoacoustic mode is computed. Although considerably reduced compared to LES, the calculation time of such a method is too large to be used in parametric studies, where different physical parameters are prone to be modified, and several modes have to be computed. Note also that there is no garanty that all the thermoacoustic modes of interest will be found by following this approach, as it is often the case when dealing with non-linear eigenvalue problems.

Intrinsic thermoacoustic modes (ITA) are usually missed when using this approach [16, 17]. Buschmann *et al.* [18] proposed an integral method based on Beyn's algorithms that is guaranteed to find all modes in a given contour, but this method is not studied in this paper.

The Galerkin modal expansion [19] can be used to reduce the problem size, in which the pressure field is no more decomposed on each vertex of the mesh as in FEM, but on the pure acoustic modes of the system. This decomposition corresponds to a spectral discretization of the domain instead of a spatial discretization as in Finite Element Method.

This article focuses on a particular low-order model based on an extension of the Galerkin method, called the *frame modal expansion*, recently adapted for thermoacoustics by Laurent *et al.* [20]. The frame modal expansion provides a better description of boundary conditions (BCs) or jump conditions compared to the modal decomposition using Galerkin's method, where Gibbs phenomena can appear at the boundaries or domain junctions (see Fig. 3 in [20]).

This work employs state-space formalism to decompose a global system into a network of interconnected subdomains. State-space models have been extensively used in thermoacoustics; for instance, Bothien *et al.* [21] developed a state-space model to compute the thermoacoustic modes of a full-scale annular gas turbine, and Orchini *et al.* [22] utilized state-space models to incorporate the influence of mean flow on the acoustic field.

By using the spectral discretization of the pressure decomposition and the state-space mathematical formalism, a linear eigenvalue problem $MP = \omega P$ can be assembled for representing the entire thermoacoustic system. Even if this eigenvalue problem seems close to the FEM one, there are actually two main differences. First, M does not depend on ω in this case, so the resolution of the eigen-problem gives all the thermoacoustic modes of the system in one step, and no iterative process is required. Then, the size of M here is typically 10^3 instead of 10^6 for FEM, so the associated eigen-problem can be solved by a standard desktop computer in a few seconds, effectively opening up the possibility of conducting extensive parametric studies. The frame-based modeling shows accurate results on reference configurations such as 1-dimensional burners [20] or 3D multiperforated liner [23] in terms of frequency of oscillation, growth rate and mode shape of the thermoacoustic modes. Despite its efficiency, this method also generates non-physical modes as output, and the crucial matter of discerning physical from non-physical modes has not been discussed in the previously cited papers.

The objective of this paper is to unravel the origin of the spurious modes generated by the frame-based projection method, as well as to propose a systematic strategy to classify the computed modes as physical/spurious. The formalism is detailed in Section 1 while some limits of the frame-based expansion approach are discussed in Section 2. Methodologies to improve the quality/robustness of the computed modes and to identify spurious solutions are described in Sections 3 and 4, respectively. The

potential of the approach is illustrated by considering the case of an experimental burner in Section 5

1. Formalism

1.1. General Helmholtz equation

The spatio-temporal evolution of thermoacoustic fluctuations are described by the generalized Helmholtz equation, obtained by combining the perfect-gas law and the Linearized Euler Equations (LEEs) at zero Mach number [4]:

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho_0} \boldsymbol{\nabla} p' \right) - \frac{1}{\gamma p_0} \frac{\partial^2 p'}{\partial t^2} = - \frac{\gamma - 1}{\gamma p_0} \frac{\partial \omega'_T}{\partial t} \tag{1}$$

In this equation, p_0 is the homogeneous static pressure, ρ_0 and γ stand for the baseline density and heat capacity ratio, respectively. Moreover, $p' = p'(\mathbf{x}, t)$ is the field of pressure fluctuations, while ω'_T is the fluctuating rate of heat release (in $W.m^{-3}$). The right-hand-side term represents the forcing of the acoustics by the unsteady combustion. When dealing with linear acoustics, it is convenient to introduce $\hat{p}(\mathbf{x}, \omega)$, the Fourier transform of the acoustic field $p'(\mathbf{x}, t)$:

$$\begin{aligned} \hat{p}(\mathbf{x}, \omega) &= \int_{-\infty}^{+\infty} p'(\mathbf{x}, t) e^{-j\omega t} \mathrm{d}t \\ p'(\mathbf{x}, t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{p}(\mathbf{x}, \omega) e^{j\omega t} \mathrm{d}\omega \end{aligned} \tag{2}$$

Applying this transformation to Eq. 1 leads to the frequency domain Helmholtz equation, traducing the coupling between combustion and acoustics.

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho_0} \boldsymbol{\nabla} \hat{p} \right) + \frac{\omega^2}{\gamma p_0} \hat{p} = -j\omega \frac{\gamma - 1}{\gamma p_0} \hat{\omega_T} \tag{3}$$

In order to build a reduced-order model of this equation, the source term representing the flame is modeled using a Flame Transfer Function (FTF) [24]. Eq. 3 becomes an eigenvalue problem, and the quantified solutions $(\omega_n, \hat{p}_n)_{n \in \mathbf{N}}$ are the thermoacoustic modes.

1.2. State-space formalism

The present work makes use of the state-space formalism and a "divide and conquer" strategy. The idea is first to decompose the thermoacoustic system (e.g. a gas turbine), into a set of subdomains, individually described by their own dynamics. Inputs and outputs quantities are also defined for each subdomain, which allows them to be connected to each other. The state-space formalism adapted for connecting subdomains will be detailed further in the paper. To illustrate how a combustion chamber can be divided in this way, Fig. 2 shows an example of such a network decomposition.

On this sketch, several subdomains are present:

- Ω_1 is a 3D subdomain with hot gases.
- Ω_2 is a 3D subdomain with cold gases.
- H is a 3D subdomain associated with a heat source in the Ω_1 subdomain.
- Multiple 2D subdomains noted C, which represent different connection elements, such as a liner-type pressure jump condition between the two geometrical subdomains Ω_1 and Ω_2, or impedance boundary conditions. They are all noted C because they all follow the same formalism specific to 2D surfaces.

All above subdomains can be sorted in 3 different classes.

- Geometrical subdomains, noted Ω, can be defined by their respective pure acoustic modes, obtained as solutions of the Helmholtz equation (Eq. 3) without combustion forcing (right-hand-side term set to 0) and associated with open ($\hat{p} = 0$) or closed ($\nabla_n \hat{p} = 0$) boundary conditions, or to a mixture of these two types of conditions. ∇_n stands for the gradient operator projected onto the normal direction of the boundary. Only objects of this class are modeled by the frame modal expansion detailed in Section 1.3.
- Connection subdomains, noted C, are defined by their jump relations or boundary conditions. They can represent multiperforated liners, impedance condition or any type of condition applied to an interface or a frontier of a geometrical subdomain.
- Flame subdomains, noted H, and defined by a relation modeling the flame response to the acoustics. For instance, a flame transfer function links the unsteady heat release rate of the flame to the velocity fluctuations upstream of the flame.

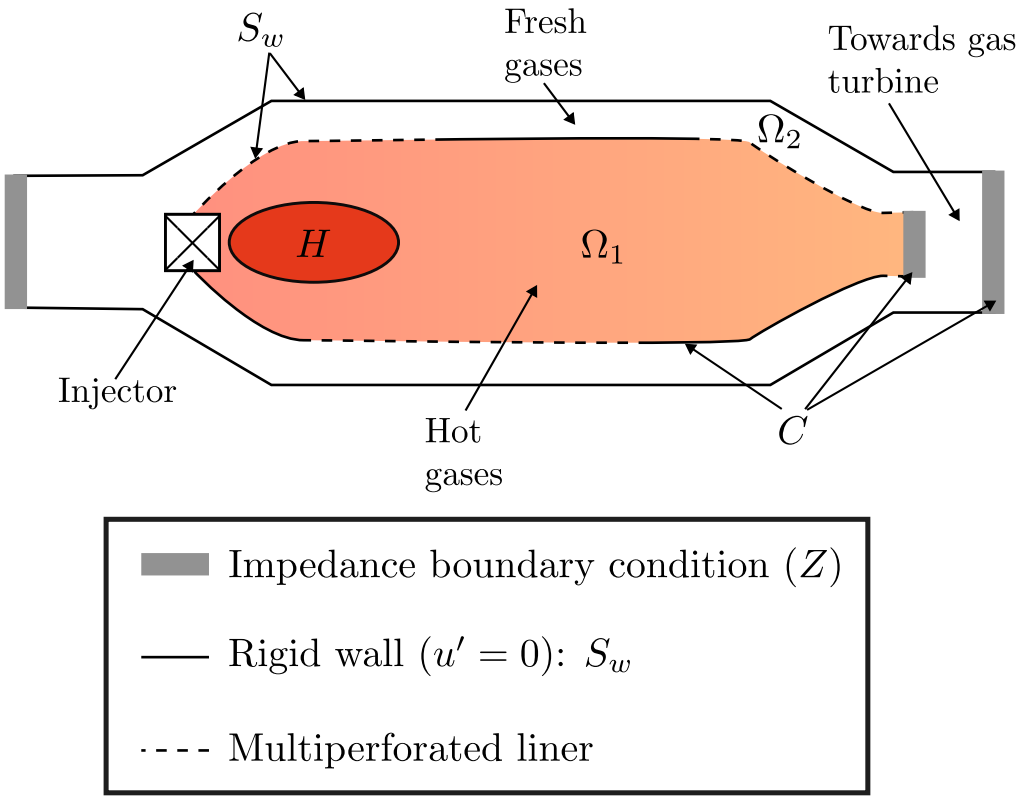

Figure 2: Schematic representation of a combustion chamber decomposition from state-space point of view.

Once the network decomposition is chosen, the coupling between subdomains remains to be formulated. This operation is done by associating a continuous time-invariant state-space to each subdomain [25]. A state-space for any subdomain (i) is a system of two matrix equations, for which $\boldsymbol{A}_i$, $\boldsymbol{B}_i$ and $\boldsymbol{C}_i$ are time independent:

$$\left\{ \begin{array}{rcl} \dot{\boldsymbol{X}}_i(t) & = & \boldsymbol{A}_i \boldsymbol{X}_i(t) + \boldsymbol{B}_i \boldsymbol{U}_i(t) \\ \boldsymbol{Y}_i(t) & = & \boldsymbol{C}_i \boldsymbol{X}_i(t) + \boldsymbol{D}_i \boldsymbol{U}_i(t) \end{array} \right. \tag{4}$$

The first line of Eq. 4 is the evolution equation, which describes the dynamics of the subdomain (i) under the forcing of other subdomains present in the network. This forcing is the product of the input matrix $\boldsymbol{B}_i$ and the input vector $\boldsymbol{U}_i(t)$. The dynamic variables of interest related to subdomain (i) are stored in the state vector $\boldsymbol{X}_i$. The second line of Eq. 4 serves to define an output vector $\boldsymbol{Y}_i$ for the subdomain (i), allowing to assemble it to other subdomains, by connecting their respective inputs/outputs (U / Y). In the state-spaces presented here, all the feedthrough matrices $\boldsymbol{D}_i$ are zero, so they are omitted from the rest of the text. The inputs/outputs of each type of subdomains are summarized in Table 1, where the acoustic quantities u_n, φ and p are respectively the velocity normal to the boundary, the acoustic potential and the pressure measured at the edge of the connected subdomain.

Table 1: Input/output vectors associated with the three different subdomain classes. Q_f is the fluctuating heat release rate representing the flame. φ is the acoustic potential defined as $p' = -\rho_0 \partial\varphi/\partial t$. The superscript T is the transpose operator.

Subdomains	Geometrical (Ω)	Flame (H)	Connection (C)
Input vectors $\boldsymbol{U}$	$\boldsymbol{U}_\Omega^T = [u_n \;\; \varphi \;\; Q_f]$	$\boldsymbol{U}_H^T = [p \;\; u_n]$	$\boldsymbol{U}_C^T = [p \;\; u_n]$
Output vectors $\boldsymbol{Y}$	$\boldsymbol{Y}_\Omega^T = [p \;\; u_n]$	$\boldsymbol{Y}_H^T = [Q_f]$	$\boldsymbol{Y}_C^T = [u_n \;\; \varphi]$

All the individual state-spaces can then be assembled by applying recursively the Redheffer Star Product [26], which consists in relating the input and output vectors of subdomains that are connected. The result of this inter-connection process is a global state-space representing the complete thermoacoustic system, which reads:

$$\dot{\boldsymbol{X}}_g(t) = \boldsymbol{A}_g \boldsymbol{X}_g(t) \tag{5}$$

Once this global system is assembled, all physical information present in matrices $\boldsymbol{A}_i$, $\boldsymbol{B}_i$ and $\boldsymbol{C}_i$ of each subdomain contribute to the global state matrix $\boldsymbol{A}_g$, while all the input/output vectors $\boldsymbol{U}_i$ and $\boldsymbol{Y}_i$ vanish. The global state vector $\boldsymbol{X}_g$ is merely the

concatenation of each subdomain state vector $\boldsymbol{X}_i$. Note that Eq. 5 can be solved in time, or the eigen solutions of dynamic matrix $\boldsymbol{A}_g$ can be computed to obtain the set of thermoacoustic modes of the global system. The eigenvalues/eigenvectors of $\boldsymbol{A}_g$ give the pulsations/shapes of all the thermoacoustic modes of the global system. If ω_n is a complex eigenvalue of $\boldsymbol{A}_g$, $\Re(\omega_n)$ and $\Im(\omega_n)$ respectively give the growth rate and the pulsation of mode n. The associated eigenvector $\boldsymbol{V}_n$ contains all the information to compute the pressure and velocity fields corresponding to the mode n - see Section 1.3. Note that $\boldsymbol{A}_g$ is a square matrix, typically of size $\sim$ 1000 or much less, as it corresponds to the sum of the size of each connected subdomain ($\boldsymbol{X}_g$ is the concatenation of each individual $\boldsymbol{X}_i$).

1.3. Frame modal expansion - Why Galerkin modal expansion is not suitable ?

In order to solve Eq. 3, the "divide and conquer" strategy defined earlier is used. As it is illustrated in Fig. 2, a global system is decomposed into different subdomains, which individually represent a particular physics. This section is dedicated to any element belonging to the geometrical subdomain class, which is described by its pure acoustic modes. A classical method for describing an acoustic domain relies on the Galerkin modal expansion, introduced by Morse in 1968 [27]. In that case, the acoustic pressure p' on a subdomain Ω_i is decomposed on a known modal basis $(\phi_n)_{n\in\mathbf{N}}$. To simplify the notation, and since we are focusing on a single sub-domain, the index i referring to subdomain Ω_i is omitted, so that:

$$p'(\mathbf{x},t) = \sum_{n}^{\infty} \dot{\Gamma}_n(t)\phi_n(\mathbf{x}) \tag{6}$$

$\dot{\Gamma}_n$ corresponds to the temporal derivative of the modal amplitude Γ_n, introduced here for convenience in the further equations. $\phi_n(\mathbf{x})$ are solutions of the Helmholtz equation (Eq. 7), without sources and with the Neumann boundary condition, that is $\nabla_n\phi_n = 0$. As it is numerically not possible to build a set of infinite number of vectors, an arbitrary value N is usually prescribed to limit the size of the modal basis which becomes $(\phi_n)_{n\leq N}$ and for which,

$$\boldsymbol{\nabla}\cdot\left(\frac{1}{\rho_0}\boldsymbol{\nabla}\phi_n\right) + \frac{\omega_n}{\gamma p_0}\phi_n = 0 \tag{7}$$

The problem is hence reduced to finding the N amplitudes $\Gamma_n(t)$ corresponding to each component $\phi_n(\mathbf{x})$ in the subdomain Ω_i. The inherent issue of such a Low-Order Model (LOM) using the Galerkin modal expansion is the proper definition of the boundary conditions of the problem. By construction, the acoustic field is always decomposed onto a basis satisfying a single type of boundary condition (homogeneous Neumann modeling a wall, or Dirichlet modeling an open atmosphere) at the frontier of each subdomain. Let us consider the example of a 1D closed duct connected to an impedance block representing a complex boundary condition (Fig. 3) to further illustrate this issue.

Typically, the network decomposition of this simple academic case is based on two elements:

- One geometrical subdomain Ω representing the 1D duct for which the pressure field is decomposed onto the corresponding basis.
- One connection subdomain C representing the impedance boundary condition $\hat{p} = \rho_0 c_0 Z\hat{u}_n$, where Z is the value of the impedance and $\hat{u}'_n$ the velocity normal to the surface.

If the Galerkin decomposition is used for the duct, and for instance the modal basis satisfies the Neumann boundary condition at the impedance interface, the gradient of the computed acoustic pressure is necessary 0 because the vectors $\phi_n(\mathbf{x})$ constituting the basis are all satisfying $\nabla_n\phi_n = 0$. So, regardless of the modal amplitudes Γ_n, the pressure gradient (and acoustic velocity) will be zero at the impedance location, which is non-physical. The same issue is also present if the basis satisfies Dirichlet boundary condition, which will impose a node for the acoustic pressure field at the impedance location. To remedy this inconsistency, the frame modal decomposition has been proposed and developed for thermoacoustics by Laurent *et al.* [20]. As the value of the pressure field is unknown at the interface, the Galerkin basis is over-completed with another basis, satisfying a different boundary condition at that interface. In the duct, the pressure field is hence decomposed onto a frame, so that:

$$\begin{aligned} p'(\mathbf{x},t) &= \textstyle\sum_{n=1}^{N} \dot{\Gamma}_n(t)\phi_n(\mathbf{x}) \\ \text{with } (\phi_n)_{n\leq N} &= (\phi^o_m)_{m\leq N/2} \cup (\phi^w_m)_{m\leq N/2} \end{aligned} \tag{8}$$

If N is not even, the frame can be defined as $(\phi_n)_{n\leq N} = (\phi^o_m)_{m\leq(N+1)/2} \cup (\phi^w_m)_{m\leq(N-1)/2}$ or vice-versa. Note that the set of vectors $(\phi_n(\mathbf{x}))_{n\leq N}$ introduced in Eq. 8 is no more an orthogonal basis but a frame, defined as the concatenation of two orthogonal basis $(\phi^o_n(\mathbf{x}))_{n\leq N/2}$ and $(\phi^w_n(\mathbf{x}))_{n\leq N/2}$ which correspond to the solutions of the Helmholtz equation (Eq. 7) with two different types of BCs at the impedance interface: Dirichlet and Neumann (see Fig. 3).

Note also that for this simple case, an analytical solution of the system exists but it is not used, the goal is to propose a methodology for representing any geometrical subdomain under the state-space framework. Generally, for 1D, 2D or 3D cases, the frame representation is applied to each geometrical subdomain Ω which has an interface C (see Fig. 2) where the behaviour of pressure is non-trivial (neither Neumman nor Dirichlet). These interfaces can be either a jump relation between

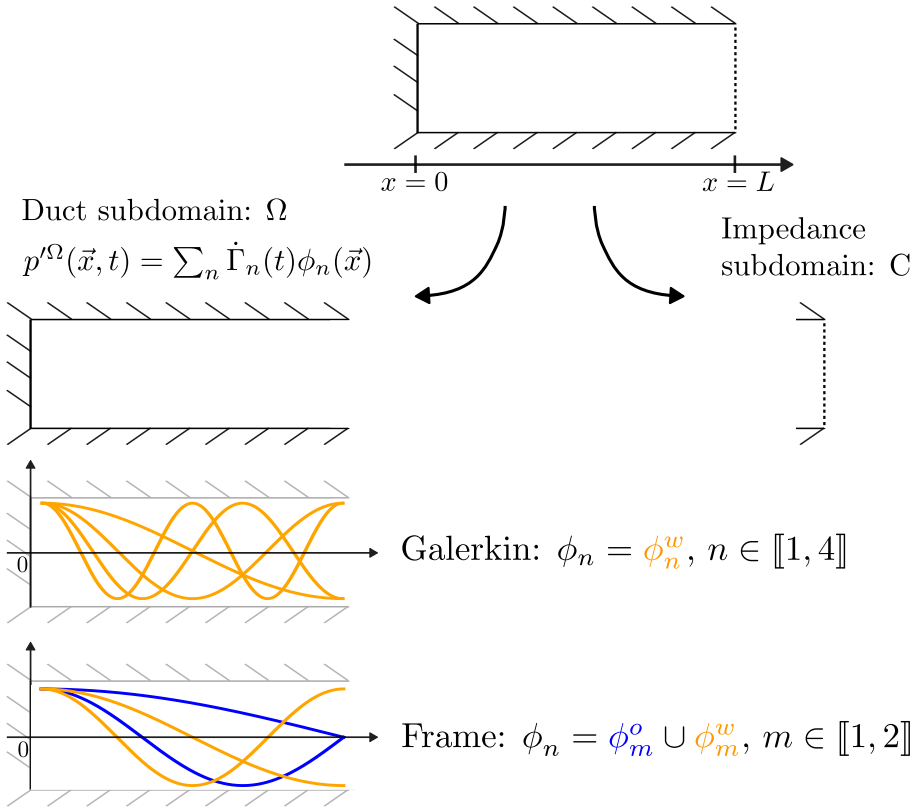

Figure 3: State-space decomposition of an academic configuration: A 1-dimensional duct connected to an impedance block defining the boundary condition $p' = \rho_0 c_0 Z u'_n$, where Z is the value of the impedance. The difference between the frame and the Galerkin modal decompositions is also illustrated. In the duct geometrical subdomain, the frame is built as the concatenation of two orthogonal basis ϕ^w and ϕ^o, while the Galerkin basis is only composed by one orthogonal basis ϕ^w.

acoustic quantities as multiperforated liners, or complex boundary conditions as impedances. Using the frame instead of a simple orthogonal basis allows the pressure field to be well-defined at the interfaces/boundaries of each geometrical subdomains. On the contrary, if an orthogonal basis is used, a Gibbs phenomenon usually appears at the interface/boundary (Fig. 3 in [20]). Depending on the boundary condition used to define the Galerkin modal basis, the Gibbs phenomenon appears either on the pressure fluctuations (Dirichlet BC) or on the velocity fluctuations (Neumann BC)

2. Frame conditioning - limits of the frame modal expansion

To understand the limits of the frame modal expansion, the test case presented in Fig. 3 is considered. The configuration consists in a 1-dimensional tube of length $L = 1$ m with zero velocity at one side ($x = 0$) and finite-valued impedance $Z = 0.5$ at the other boundary ($x = L$). The value of the sound speed is set to $c_0 = 390.3\ m.s^{-1}$. The system is decomposed into two subdomains: the duct Ω and the impedance block C. A state-space (Eq. 4) is then associated with each subdomain. All this section focuses on the geometrical subdomain: the duct.

2.1. State-space representation of an acoustic geometrical subdomain

The fact that pressure fluctuations decomposed using Eq. 8 are also solution of the thermoacoustic problem expressed by Eq. 3 must now be translated into the state-space language. Since the unknowns of the problem are the modal amplitude coefficients $\Gamma_n(t)$, these become the state variables of the state-space set of equations. The equation governing the time evolution of each Γ_n can hence be found by injecting the pressure frame decomposition (Eq. 8) into Eq. 3. After some algebra [20], the differential equation of the modal amplitudes in the duct reads,

$$\begin{cases} p'(\mathbf{x},t) &= \sum_n \dot{\Gamma}_n(t)\phi_n(x) \\ \ddot{\Gamma}_n(t) &= -\omega_n^2\Gamma_n(t) + \rho_0 c_0^2 \mathbf{\Lambda^{-1}} \boldsymbol{\mathcal{I}}_{\boldsymbol{\phi}}(t) + \mathbf{\Lambda^{-1}} \boldsymbol{\mathcal{F}}_{\boldsymbol{\phi}}(t) \end{cases} \tag{9}$$

The last term on right-hand side of the second line of Eq. 9 represents the external forcing contribution to the subdomain due to the impedance block C. This connection is managed by the term $\boldsymbol{\mathcal{I}}_{\boldsymbol{\phi}}(t)$, which involves the projection of the frame $\boldsymbol{\phi}$ onto the connection interface C, the acoustic potential and the normal velocity at this interface. Its explicit expression, which is not useful for the present analysis, is provided in [23]. ρ_0 and c_0 are the mean density and sound speed at the location where the connection takes place. $\mathbf{\Lambda^{-1}}$ is the inverse of the Gram matrix, defined as the inner product between frame components: $\mathbf{\Lambda} = \boldsymbol{\phi}^T\boldsymbol{\phi}$ where

$\boldsymbol{\phi}$ is the row vector $\boldsymbol{\phi} = [\phi_1, \cdots, \phi_N]$. The term $\mathcal{F}_{\boldsymbol{\phi}}(t)$ represents the volumic source in the case where a flame is present in the system. Its explicit expression is detailed in Section 4. Still, $\mathbf{\Lambda^{-1}}$ can be difficult to compute numerically due to poor numerical conditioning. This issue would arise as soon as a vector ϕ_n of the frame $\boldsymbol{\phi}$ is close to a linear combination of other vectors of the frame: $\phi_n \simeq \sum_{m\neq n} \alpha_m \phi_m$. This situation which is in fact more and more likely when the frame size increases as discussed in the next section.

2.2. Conditioning issues

In the case depicted in Fig. 3, the explicit expression of the modal frame $\boldsymbol{\phi}$ corresponding to the duct is the following:

$$\begin{aligned} \boldsymbol{\phi} &= (\phi_n(\mathbf{x}))_{n\leq N} \\ \boldsymbol{\phi} &= \underbrace{\left(\cos\frac{m\pi x}{L}\right)_{m\leq N/2}}_{\phi^w_m} \cup \underbrace{\left(\cos\frac{(2m+1)\pi x}{2L}\right)_{m\leq N/2}}_{\phi^o_m} \end{aligned} \tag{10}$$

This frame $\boldsymbol{\phi}$ is associated with the following eigen-frequencies:

$$\begin{aligned} \boldsymbol{f} &= (f_n)_{n\leq N} \\ \boldsymbol{f} &= \underbrace{\left(\frac{mc_0}{2L}\right)_{m\leq N/2}}_{f^w_m} \cup \underbrace{\left(\frac{(2m+1)c_0}{4L}\right)_{m\leq N/2}}_{f^o_m} \end{aligned} \tag{11}$$

The maximum value reached by $\boldsymbol{f}$ in Eq. 11, $f_{max} = (N+1)c_0/4L$, defines an upper limit for the computable frequencies of the whole system as no mode with a frequency higher than f_{max} can be accurately reconstructed in the duct by a linear combination of modes with lower frequencies. This reasoning can be extended for cases which involve several geometrical subdomains and, to be more precise, the global frequency upper limit is the minimum of all the f_{max} defined in each geometrical subdomains.

Once the frame is defined, the second line of Eq. 9 can be written as a state-space representation (Eq. 4), where the state vector associated with the duct, $\boldsymbol{X_d}$, is the concatenation of all the unknown modal amplitudes $\Gamma_n(t)$ and their temporal derivative $\dot{\Gamma}_n(t)$. The explicit expression of state-space matrices for a generic geometrical subdomain can be found in Appendix A of [20]. Another state-space is associated with the impedance block C. Then the Redheffer star product allows to assemble a global system of the form $\dot{\boldsymbol{X}}_{\boldsymbol{g}} = \boldsymbol{A_g X_g}$ by connecting the respective inputs/outputs of each individual state-space. The schematic in Fig. 4 illustrates this point.

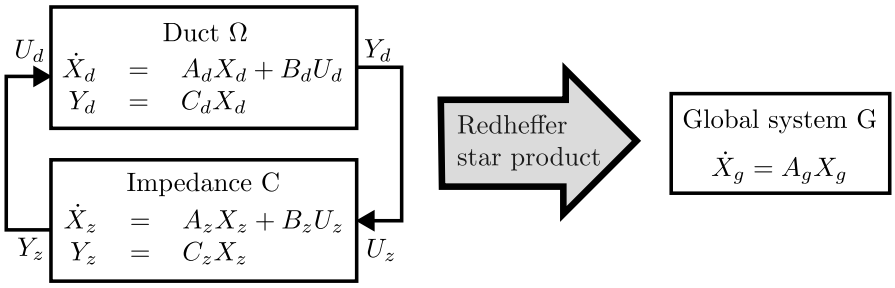

Figure 4: Block diagram of the state-space representation of the configuration depicted Fig. 3, in which 2 blocks are involved (i.e. 2 subdomains): the duct Ω and the impedance C.

Note that the final assembled equation $\dot{\boldsymbol{X}}_{\boldsymbol{g}} = \boldsymbol{A_g X_g}$ is also a State-space but without any input/output. For this case it reads:

$$\underbrace{\begin{bmatrix}\dot{\boldsymbol{X}}_{\boldsymbol{d}} \\ \dot{\boldsymbol{X}}_{\boldsymbol{z}}\end{bmatrix}}_{\dot{\boldsymbol{X}}_{\boldsymbol{g}}} = \underbrace{\begin{bmatrix}\boldsymbol{A_d} & \boldsymbol{B_d C_z} \\ \boldsymbol{B_z C_d} & \boldsymbol{A_z}\end{bmatrix}}_{\boldsymbol{A_g}} \underbrace{\begin{bmatrix}\boldsymbol{X_d} \\ \boldsymbol{X_z}\end{bmatrix}}_{\boldsymbol{X_g}} \tag{12}$$

The global system is solved for 3 different values of the frame size, namely $N = 12$, 20 and 26, increasing the range of solution frequencies (i.e increasing f_{max}). Figure 5 shows the frequency and the growth rate of the modes between 0 and f_{max} returned by the frame expansion method for each frame size. They are compared to the analytical complex frequencies $s^A_n = f^A_n + i\sigma^A_n$, solutions of the dispersion relation associated with the configuration, given by:

$$iZ\tan\left(\frac{2\pi s^A_n}{c_0}L\right) = 1 \tag{13}$$

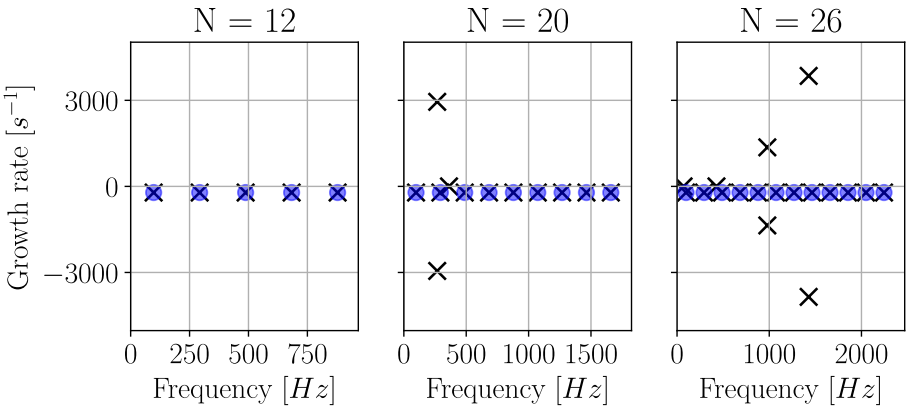

Figure 5: Modes of the configuration depicted in Fig. 3 displayed in the frequency plane. Crosses: Numerical results from the frame modal expansion method, with different values of the frame size: $N = 12$ (left), $N = 20$ (middle), $N = 26$ (right). Blue circles: analytical solutions (Eq. 13). Only modes with frequency between 0 and $f_{max} = (N + 1)c_0/4L$ are plotted.

The agreement is very good for $N = 12$ since all the five first analytical modes are well represented by the frame-based numerical procedure, both for the oscillation frequency and for the growth rate. The agreement is still good for $N = 20$ and $N = 26$ but some spurious modes appear (see the $\times$ symbols that do not match the reference circles for $N = 20$ and for $N = 26$). The generation of these non-physical components is due to numerical rounding errors in the inversion process of the Gram matrix $\mathbf{\Lambda}$. As the inverse of the Gram matrix is involved in the computation of modal amplitudes Γ_n (see Eq. 9), numerical errors on $\mathbf{\Lambda}^{-1}$ propagate, resulting in some spurious modes in the output of the method. To support this claim, Figure 6 displays the singular values $s(\mathbf{\Lambda})$ of the Gram matrix for each frame size $N = 12$, 20 and 26. The condition number of the matrix, $\kappa(\mathbf{\Lambda})$, allows to quantify the conditioning of the geometrical subdomain (the duct), and is defined as the ratio between the maximum and the minimum of singular values:

$$\kappa(\mathbf{\Lambda}) = \frac{\max s(\mathbf{\Lambda})}{\min s(\mathbf{\Lambda})} \tag{14}$$

The presence of spurious modes in the output is expected if $1/\kappa(\mathbf{\Lambda})$ get close to the machine precision, about 10^{-16} for double precision coding. From Figure 6, this condition is met for $N = 20$ and 26 where a plateau-like behaviour is present for the lowest singular values (around 10^{-16}), which is responsible for the poor numerical conditioning of the system. However this plateau is not reached for $N = 12$, in other words, $\mathbf{\Lambda^{-1}}$ is computed accurately for $N = 12$ only, which is consistent with the results in Fig. 5.

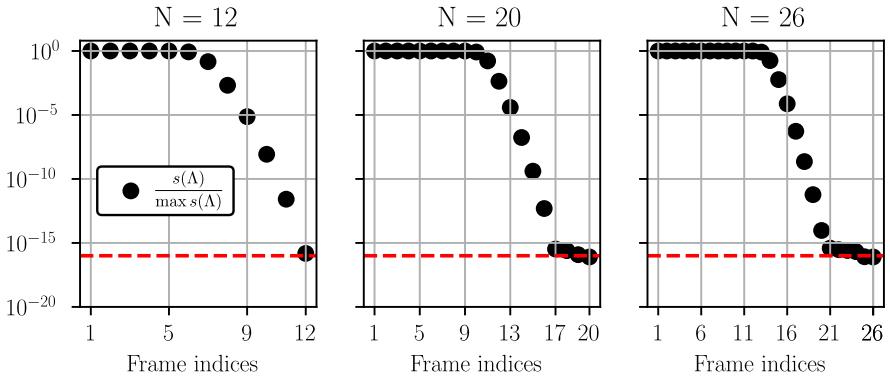

Figure 6: Singular values distribution of the Gram matrix normalized by the maximum of singular values, for $N = 12$ (left), $N = 20$ (middle) and $N = 26$ (right). The red horizontal dashed lines correspond to 10^{-16}, an indication of the machine precision.

It is important to note that the frame size N is depending on the upper bound of the range of frequencies of interest in the global system f_{up}. The frame size N imposes a maximum value $f_{max}(N)$ for the frequencies available in the frame, which has to be greater than f_{up}. For the case considered in Fig. 3, the frame size is increased only for illustrating ill-conditioning issues. In the case where high frequency modes (screech) are of interest, the value of N must be selected large enough for reaching the target frequency f_{up}, and ill-conditioning issues can occur.

3. Singular Value Decomposition of the frame

To improve the conditioning issue discussed in Section 2.2, the proposed methodology is first to transform the frame into an orthonormal basis using the Singular Value Decomposition (SVD), and then to truncate the frame by removing the vectors

associated with the smallest singular values.

3.1. SVD of the frame

In each geometrical subdomain Ω_i (i.e only the duct in the case studied), the SVD decomposition is applied to the frame $\boldsymbol{\phi}$ as follows. The subscript i is omitted.

$$\boldsymbol{\phi} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T \tag{15}$$

$\boldsymbol{U}$ is a row vector $\boldsymbol{U} = [U_1(\mathbf{x}) \ \cdots \ U_N(\mathbf{x})]$ of the same shape as $\boldsymbol{\phi}$ and which contains a set of orthonormal vectors spanning the same space as $\boldsymbol{\phi}$. $\boldsymbol{\Sigma}$ is the diagonal matrix of singular values, sorted in descending order. The unitary matrix $\boldsymbol{V}$ contains the right singular vectors of $\boldsymbol{\phi}$. Doing this SVD requires a metric adaptation detailed in Appendix 8.1. Once this decomposition is done, the new state-space based on the $\boldsymbol{U}$ basis has to be derived. For this purpose, the unknowns are now noted $\dot{\Upsilon}_n$ (see Eq. 16). The index i referring to any geometrical subdomain Ω_i is omitted again, and the pressure fluctuations is decomposed onto the $\boldsymbol{U}$ basis instead of the frame $\boldsymbol{\phi}$:

$$p'(\mathbf{x}, t) = \sum_{n=1}^{N} \dot{\Upsilon}_n(t) U_n(\mathbf{x}) \tag{16}$$

The next step is to derive the modified Helmholtz equation which solutions are the components of $\boldsymbol{U}$. This is achieved by remarking that the classical Helmholtz equation can be written in the following vector form:

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho_0} \boldsymbol{\nabla} \boldsymbol{\phi} \right) + \frac{1}{\gamma p_0} \boldsymbol{\phi} \boldsymbol{\Omega}^{\mathbf{2}} = \mathbf{0} \tag{17}$$

where $\boldsymbol{\Omega}$ is the diagonal of eigenfrequencies, $\boldsymbol{\Omega} = diag(\omega_1, \cdots, \omega_N)$. By replacing $\boldsymbol{\phi} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ and then right-multiplying the equation by $(\boldsymbol{\Sigma}\boldsymbol{V}^T)^{-1} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}$, the following modified Helmholtz equation is obtained:

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho_0} \boldsymbol{\nabla} \boldsymbol{U} \right) + \frac{1}{\gamma p_0} \boldsymbol{U} \left[\boldsymbol{\Sigma} \boldsymbol{V}^T \boldsymbol{\Omega}^{\mathbf{2}} \boldsymbol{V} \boldsymbol{\Sigma}^{-1} \right] = \mathbf{0} \tag{18}$$

The coupling between U_n components then reads:

$$\boldsymbol{\nabla} \cdot \left(\frac{1}{\rho_0} \boldsymbol{\nabla} U_n(\mathbf{x}) \right) + \frac{1}{\gamma p_0} \sum_{k=1}^{N} [\boldsymbol{\Sigma} \boldsymbol{V}^T \boldsymbol{\Omega}^2 \boldsymbol{V} \boldsymbol{\Sigma}^{-1}]_{k,n} U_k(\mathbf{x}) = 0 \tag{19}$$

No boundary condition needs to be defined for this equation since it is never solved in practice; instead the $\boldsymbol{U}$ basis is computed thanks to the orthonormalization process $\boldsymbol{\phi} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$. This equation does not represent any physical phenomenon, but is needed for developing the new state-space of geometrical subdomains. Note that there are similitudes between this modified Helmholtz equation for U_n and the classical Helmholtz equation for ϕ_n. For example, the set of equivalent pulsations $[\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{\Omega}^{\mathbf{2}}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}]_{k,n}$, with $k \in [\![1, N]\!]$ can be associated with the mode U_n as ω_n^2 is associated with the mode ϕ_n. The new state-space of any geometrical subdomain Ω_i can now be derived from Eq. 19 and writes:

$$\begin{cases} p'(\mathbf{x}, t) &= \sum_n^N \dot{\Upsilon}_n(t) U_n(x) \\ \ddot{\Upsilon}_n(t) &= -\sum_k^N [\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{\Omega}^{\mathbf{2}}\boldsymbol{V}\boldsymbol{\Sigma}^{-1}]_{n,k} \Upsilon_k(t) \\ &+ \rho_0 c_0^2 \boldsymbol{\mathcal{I}}_{\boldsymbol{U}}(t) \\ &+ \boldsymbol{\mathcal{F}}_{\boldsymbol{U}}(t) \end{cases} \tag{20}$$

Some similitudes can be highlighted between this modified state-space (Eq. 20) and the original one (Eq. 9). The terms $\boldsymbol{\mathcal{I}}_{\boldsymbol{U}}(t)$ and $\boldsymbol{\mathcal{F}}_{\boldsymbol{U}}(t)$ are the counterpart of the $\boldsymbol{\Lambda}^{-1}\boldsymbol{\mathcal{I}}_{\boldsymbol{\phi}}$ and $\boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{\mathcal{F}}_{\boldsymbol{\phi}}(t)$ terms in Eq. 9 and respectively represent the external forcing contributions from the other subdomains and the volumic source induced by the flame. Note that the $\boldsymbol{\Lambda}^{-1}$ dependency vanishes since $\boldsymbol{U}$ is an orthogonal basis, and the projection terms $\boldsymbol{\mathcal{I}}_{\boldsymbol{U}}$ and $\boldsymbol{\mathcal{F}}_{\boldsymbol{U}}$ are now depending on U_n vectors instead of ϕ_n. The numerical illposedness of the SVD-based state-space (Eq. 20) is related to the conditioning of $\boldsymbol{\Sigma}$ (see the term $\boldsymbol{\Sigma}^{-1}$ in Eq. 20) instead of $\boldsymbol{\Lambda}$ for the original state-space formulation (Eq. 9). Since $\boldsymbol{\Lambda} = \boldsymbol{V}\boldsymbol{\Sigma}^2\boldsymbol{V}^T$, the singular values of $\boldsymbol{\Sigma}$ are the square root of those of $\boldsymbol{\Lambda}$, so $\kappa(\boldsymbol{\Sigma}) = \sqrt{\kappa(\boldsymbol{\Lambda})}$. Thus the state-space formulation in Eq. 20 should be better conditioned than the one using the original frame modal expansion. Still, one can expect that numerical quality issues will remain in cases where $\kappa(\boldsymbol{\Sigma}) \geq 10^{16}$, that is to say $\kappa(\boldsymbol{\Lambda}) \geq 10^{32}$. To illustrate this point, the configuration described in Fig. 3 is solved using the SVD-based state-space (Eq. 20) for the same values of frame size (N = 12, 20 and 26) as in Figs. 5 and 6. The results in terms of frequencies are presented in top row of Fig. 7.

The agreement between physical modes and the reference frequencies remains good for all the values of the frame size N. For N = 12, both the direct frame (Eq. 9) and the SVD-based approach (Eq. 20) give the correct results, without spurious modes, as the global system is well conditioned in both cases. For N = 20 the conditioning of the problem is $\kappa(\boldsymbol{\Sigma}) \simeq 1.2 \times 10^{14}$ using

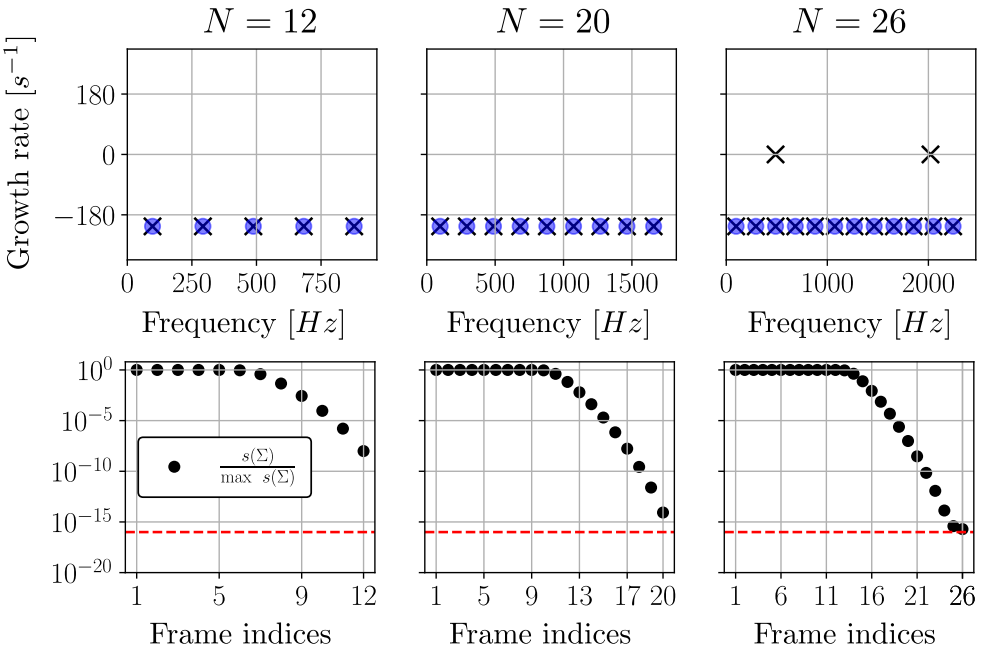

Figure 7: Top row: Modes of the configuration depicted in Fig. 3 displayed in the frequency plane. Crosses: Numerical results from the SVD-based modal expansion method, with different values of the frame size: $N = 12$ (left), $N = 20$ (middle), $N = 26$ (right). Blue circles: analytical solutions (Eq. 13). Only modes with frequency between 0 and $f_{max} = (N+1)c_0/4L$ are plotted. Bottom row: Singular values distribution of the Σ matrix normalized by the maximum of singular values, for $N = 12$ (left), $N = 20$ (middle) and $N = 26$ (right). The red horizontal dashed lines correspond to 10^{-16}, an indication of the machine precision.

the SVD-based framework and $\kappa(\mathbf{\Lambda}) \simeq 1.4 \times 10^{28}$ for the original formulation (Eq. 9). Consistently, the three spurious modes observed previously are now removed (compare Fig. 5 and top row of Fig. 7 for $N = 20$). Note that, comparing Fig. 6 and the bottom row of Fig. 7, the theoretical relation between the sets of singular values $s(\mathbf{\Sigma}) = \sqrt{s(\mathbf{\Lambda})}$, is not met for the lowest singular values for $N = 20$ and $N = 26$. The computation of singular values itself is actually sensitive to the conditioning, and singular values close to/lower than $10^{-16} \times \max(s)$ cannot be computed accurately. The important point to understand is the following: if a plateau-like behaviour is present for the lowest singular values, then all singular values located at or below this plateau are considered as numerical zeros, leading to an ill-conditioned system, and consequently, the apparition of spurious modes. As it is illustrated in Fig. 7 for $N = 26$, the two lowest singular values have reached this plateau around 6×10^{-16} and two spurious modes are still present in the output. The idea to remove spurious modes in cases where even the SVD is ill-conditioned is to truncate the $\boldsymbol{U}$ basis; this approach is considered in the following section.

3.2. Truncated SVD - Removing multicollinearity

Even if the $\boldsymbol{U}$ basis is better conditioned than the frame $\boldsymbol{\phi}$, the SVD process is just a change of basis and we have no control on the value of the conditioning. The previous section shows that even the $\boldsymbol{U}$ basis can be ill-conditioned (case $N = 26$), leading to the appearance of spurious modes in the output when $\kappa(\mathbf{\Sigma}) \sim 10^{16}$ or more. Remember that in each geometrical subdomain, the frame $\boldsymbol{\phi}$ is built from two concatenated bases. If a vector ϕ_n of the first basis can numerically be approximated by a linear combination of modes in the second basis ($\phi_n \simeq \sum_{m \neq n} \alpha_m \phi_m$) then this ϕ_n leads to a small singular value, which increases $\kappa(\mathbf{\Sigma})$. This phenomenon is commonly called multicollinearity [28] and can be removed by using the truncated SVD:

$$\boldsymbol{\phi}_t = \boldsymbol{U}_t \mathbf{\Sigma}_t \boldsymbol{V}_t^T \tag{21}$$

The truncated SVD serves to approximate the space spanned by the initial $\boldsymbol{\phi}$ basis with another basis $\boldsymbol{\phi_t}$, built from a truncated set of vectors $\boldsymbol{U}_t$ and truncated matrices $\mathbf{\Sigma}_t$ and $\boldsymbol{V}_t^T$. The difference between the classical SVD (Eq. 15) and the truncated SVD (Eq. 21) is all about the size of matrices $\boldsymbol{U}$, $\mathbf{\Sigma}$ and $\boldsymbol{V}$. The matrix $\mathbf{\Sigma}$ in Eq. 15 is the diagonal matrix of singular values of $\boldsymbol{\phi}$, sorted in descending order such that $\mathbf{\Sigma} = diag(\sigma_1, \cdots, \sigma_N)$ with $\sigma_1 > \cdots > \sigma_N$, where N is the frame size. In Eq. 21, the matrix $\mathbf{\Sigma}_t$ is a reduction of $\mathbf{\Sigma}$, where singular values lower than $\varepsilon_t \times \max(s(\mathbf{\Sigma}))$ are removed, where ε_t is the truncation threshold. The truncated singular matrix is then $\mathbf{\Sigma}_t = diag(\sigma_1, \cdots, \sigma_{N_t})$, where $N_t < N$ is the number of remaining singular values in $\mathbf{\Sigma}_t$. $\boldsymbol{U}_t$ is a row vector of size N_t which contains the first N_t orthonormalized vectors, corresponding to the N_t largest singular values. $\boldsymbol{V}_t^T$ is of size $N_t \times N$ in order to have $\boldsymbol{\phi}_t$ of the same size as $\boldsymbol{\phi}$. As spurious modes are still present for the case $N = 26$, only this case is studied in this part. The natural choice for the threshold of the truncated SVD is the machine precision $\varepsilon_t = 10^{-16}$. However, as explained earlier, the lowest singular values are difficult to compute accurately so the threshold is set to $\varepsilon_t = 10^{-15}$, just above the plateau of the two lowest singular values (see the left plot in Fig. 8). To study the robustness of the method, six different values of the truncation threshold have been tested from $\varepsilon_t = 10^{-15}$ to $\varepsilon_t = 10^{-1}$, thus decreasing the size of the reduced frame from

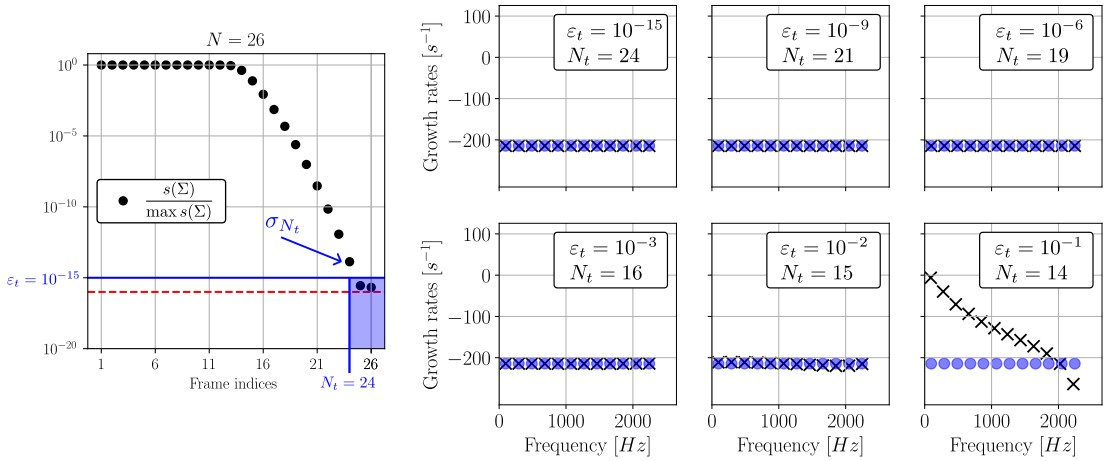

Figure 8: Left: Singular values distribution of the Σ matrix normalized by the maximum of singular values, for $N = 26$. The red horizontal dashed lines correspond to 10^{-16}, an indication of the machine precision. In this plot, the truncation threshold is set to $\varepsilon_t = 10^{-15}$ just above the plateau, leading to $N_t = 24$ remaining vectors in the U-basis. Right: Modes of the configuration depicted in Fig. 3 displayed in the frequency plane, for six different values of the truncation threshold ε_t. Crosses: Numerical results from the truncated-SVD-based modal expansion method, with the initial frame size $N = 26$. Blue circles: analytical solutions (Eq. 13). Only modes with frequency between 0 and $f_{max} = (N + 1)c_0/4L$ are plotted.

$N_t = 24$ to $N_t = 14$. The results are presented in the right hand side part of Fig. 8, where the two spurious modes have vanished from the global output, because the two multicollinear modes have been removed from the frame.

The truncation threshold values do not impact the calculated growth rate in comparison with the analytical solutions up until $\varepsilon_t = 10^{-3}$. While for $\varepsilon_t = 10^{-2}$ the results remain stable with a small shift occurring at higher frequency, for a truncation threshold fixed at $\varepsilon_t = 10^{-1}$, the method strongly deviates from the analytical solutions. All in all, the results clearly show the robustness of the proposed methodology, emphasized by the large spectrum of ε_t values that is user-defined.

Once the truncated SVD is performed, the distance between the spaces spanned by either $\boldsymbol{\phi}_t$ or $\boldsymbol{\phi}$ can be quantified thanks to the Eckart-Young Theorem (see [29, Theorem 2.4.8, p. 79]) which states:

$$||\boldsymbol{\phi} - \boldsymbol{\phi}_t||_2 = \sigma_{N_t+1} \tag{22}$$

where $||\cdot||_2$ denotes here the 2-norm of a matrix that is:

$$||\boldsymbol{\phi}||_2 = \max_{||x||_2=1} ||\boldsymbol{\phi}x||_2 \tag{23}$$

For instance, with $\varepsilon_t = 10^{-15}$, two modes have been removed from the frame so $||\boldsymbol{\phi} - \boldsymbol{\phi}_t||_2 = \sigma_{25} \simeq 2.8 \times 10^{-16}$. The truncation threshold parameter allows the user to control the precision of the search space on which the pressure field is decomposed.

4. Flames low-order modeling

To understand how the flames are modeled under the network framework presented in this paper, a flame is added to the 1D duct case depicted in Fig. 3, at the position $x_f = 0.23\ m$. The truncation threshold of the SVD is set to $\varepsilon_t = 10^{-15}$ for all this section.

4.1. Modeling

The Helmholtz equation (Eq. 3) is an eigenproblem, for which solutions are the thermoacoustic modes, if the source term is modeled by a Flame Transfer Function (FTF). The fluctuating heat release rate is decomposed as a product of a mean flame shape and a FTF such as:

$$\hat{\omega}_T = H(\boldsymbol{x})Q(t) \tag{24}$$

$H(\boldsymbol{x})$ is the mean flame shape (in m^{-3}) normalized over the whole volume V: $\int_V H(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1$. In this example, the flame shape is defined as a normalized gaussian centered at x_f and of thickness $\delta_f = 0.02\ m$:

$$H(\boldsymbol{x}) = \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{(x-x_f)^2}{\delta_f^2}} \tag{25}$$

The term $Q(t)$ (in W) is modeled using the adimensional form of the $n - \tau$ FTF:

$$Q(t) = \frac{\bar{Q}}{\bar{u}} N e^{-j\omega\tau} u'_n(\boldsymbol{x_r}, t) \tag{26}$$

$u'_n(\boldsymbol{x}_r, t)$ is the reference velocity fluctuation in the direction of the flame, evaluated at the reference point $x_r = 0.21\ m$, just upstream the flame. The mean flame power is set to $\bar{Q} = 40\ W$ and the mean flow speed is set to $\bar{u} = 1\ m.s^{-1}$. The gain N and the delay τ of the FTF are *a priori* depending on the frequency ω but they are set constant for simplicity: $N = 1$ and $\tau = 1\ ms$

Following these notations, the expression of the term $\boldsymbol{\mathcal{F}_U}(t)$ in the state-space of the 1D duct (Eq. 20) writes:

$$\begin{cases} p'(\mathbf{x}, t) &= \sum_n^N \dot{\Upsilon}_n(t) U_n(x) \\ \ddot{\Upsilon}_n(t) &= -\sum_k^N [\boldsymbol{\Sigma V}^T \boldsymbol{\Omega^2 V \Sigma}^{-1}]_{n,k} \Upsilon_k(t) \\ &+ \rho_0 c_0^2 \boldsymbol{\mathcal{I}_U}(t) \\ &+ (\gamma - 1)\langle H|U_n\rangle Q(t) \end{cases} \tag{27}$$

The term $(\gamma - 1)\langle H|U_n\rangle Q(t)$ traduces the influence of the flame onto each modal amplitude Υ_n of the unknown pressure field. γ is the heat capacity ratio, and fixed equal to $\gamma = 1.4$. The term $\langle H|U_n\rangle$ is the projection of the normalized flame shape onto the frame mode U_n, defined as: $\langle H|U_n\rangle = \int_V H(\boldsymbol{x}) U_n(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$. The flame shape is so approximated by its projection on the space spanned by the U-basis. The error between the real flame shape $H(\boldsymbol{x})$ and its projection is computed using the classical 2-norm of a vector as follows:

$$\epsilon_{fs} = \frac{\left\| H(\boldsymbol{x}) - \sum_n^{N_t} \langle H|U_n\rangle U_n(\boldsymbol{x}) \right\|_2}{\|H(\boldsymbol{x})\|_2} \tag{28}$$

Here, $\|\cdot\|$ denote the 2-norm of the vector. As highlighted in Fig. 9, the frame size $N = 26$ used in Section 3 is too low to have a correct representation of the flame shape through its projection onto the U-basis. The projection error is $\epsilon_{fs}(N = 26) = 2.7 \times 10^{-1}$. The frame size is then increased to $N = 100$ to have a better approximation of the flame shape: $\epsilon_{fs}(N = 100) = 5.8 \times 10^{-4}$

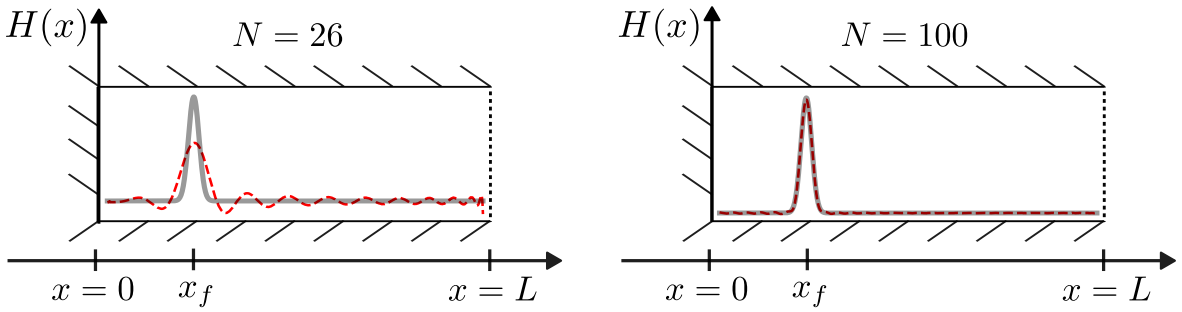

Figure 9: The flame added to the configuration depicted in Fig. 3. Gray plots: real flame shape $H(x)$ defined in Eq. 25. Red dashed plots: reconstruction of the flame shape through its projection onto the U-basis. Left: case $N = 26$. Right: case $N = 100$.

Once the flame shape is correctly projected onto the U-basis, the FTF (Eq. 24) must be written as a state-space formulation (Eq. 4). To do so, the term $Ne^{-j\omega\tau}$ in the definition of $Q(t)$ is fitted using the Vector Fitting algorithm [30, 31]. This algorithm is used to fit the input dataset $(\omega, Ne^{-j\omega\tau})$ as a sum of Pole Base Functions (PBFs), defined as follows:

$$Ne^{-j\omega\tau} = R_0 + \sum_{k=1}^{N_{PBF}} \frac{R_k}{j\omega - p_k} \tag{29}$$

In this expression, N_{PBF} is the order of the fit (i.e. number of PBFs), R_0 is a constant and p_k are the poles, with their associated residues R_k. A maximum frequency must be provided to the fitting algorithm, and is set to 2100 Hz. The order of the fit N_{PBF} is set to 20.

4.2. Direct results

The state space realisation of the flame block is explicited in Appendix 8.2. Figure 10 shows the studied 1D duct configuration, with the flame and the impedance boundary condition, and the corresponding block diagram.

The direct results of the presented modal expansion method are compared to the pseudo analytical solutions, provided by the approximation of a 1D duct and an infinitely thin flame. Results in terms of frequencies and growth rates are depicted in Fig. 11.

The frequency and growth rate of each physical modes are correct, the little shift between black crosses and blue dots in Fig. 11 is due to the modeling of the flame, which has a zero thickness when using the pseudo-analytical tool instead of $\delta_f = 0.02\ m$ in the presented method. However, three non-physical modes appear in the output (see the cross points that do not match the blue circles in Fig. 11). They correspond to the complex poles of the fit p_k which have not converged towards a physical mode.

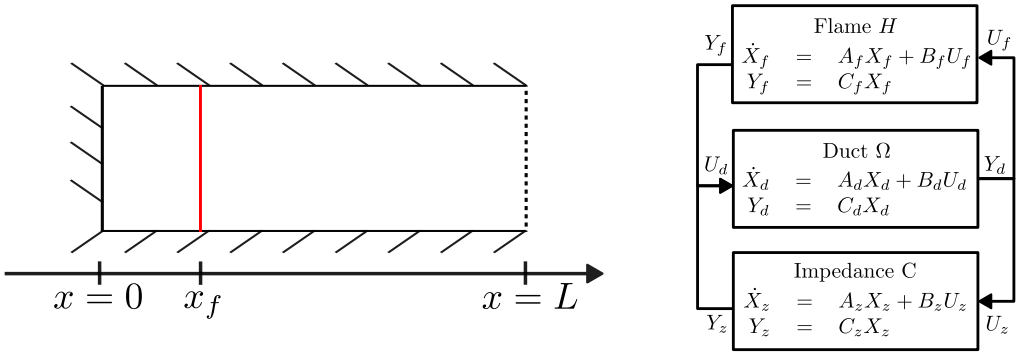

Figure 10: Left: studied academic configuration, composed with a 1D duct of length $L = 1\ m$, an impedance boundary condition on the right side and a flame positioned at $x_f = 0.23\ m$ of thickness $\delta_f = 0.02\ m$. Right: the corresponding block diagram.

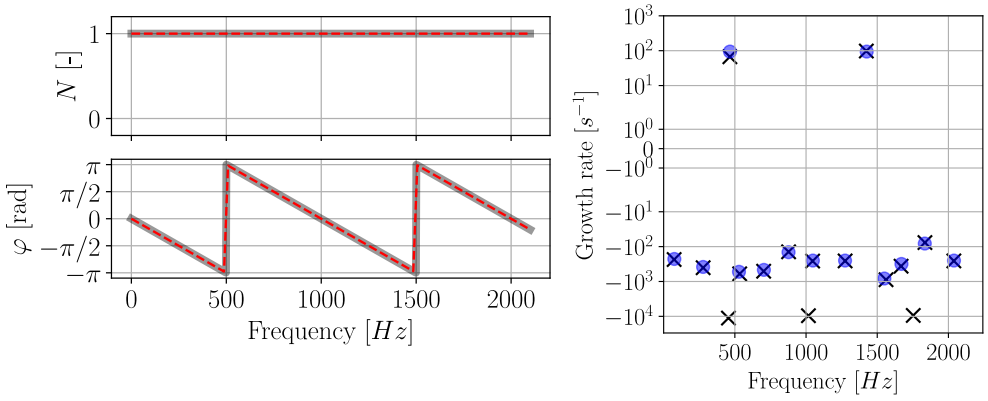

Figure 11: Left: Gain N and phase $\varphi = \omega\tau$ of the FTF. Gray plot: constant $N - \tau$ FTF with $N = 1$ and $\tau = 0.001\ s$. Red dashed plot: the fit reconstructed using Eq. 29. Right: direct results of the modal expansion method (black crosses), and pseudo analytical solutions (blue dots).

4.3. Identification of spurious modes due to the fit

The criterion to identify these spurious modes is based on the idea: only the physical modes are independent of the order of the fit N_{PBF}. The strategy is to fit the same FTF with a different order, and compare both outputs of the method using these two fits. The modes with the highest shift are then considered as non-physical. For this case, a second fit is done with $N_{PBF} = 25$. The shift $\epsilon_s(n)$ associated with each mode n is defined as follow. Let $f_n^{(20)}$, $\sigma_n^{(20)}$ the frequency and growth rate of the mode n in the family of modes computed using $N_{PBF} = 20$ and $f_m^{(25)}$, $\sigma_m^{(25)}$ the frequency and growth rate of the mode m in the family of modes computed using $N_{PBF} = 25$. It is assumed that the mode m is the closest mode (in the frequency plane) to the mode n among all the returned modes $(f_k^{(25)}, \sigma_k^{(25)})_k$.

$$\epsilon_s(n) = \frac{\sqrt{\left(f_n^{(20)} - f_m^{(25)}\right)^2 + \left(\sigma_n^{(20)} - \sigma_m^{(25)}\right)^2}}{\sqrt{\left(f_n^{(20)}\right)^2 + \left(\sigma_n^{(20)}\right)^2}} \tag{30}$$

On the left part of Fig. 12, this shift is the relative cartesian distance between a black cross and the closest red star. The value of $\ln(\epsilon_s(n))$ is depicted on the right part of Fig. 12. The one dimensional K-means algorithm is applied to the dataset $\ln(\epsilon_s(n))$ to separate the two clusters. The cluster with the lowest center value is colored in green and corresponds to the physical modes (smaller shift), and the other one is colored in orange.

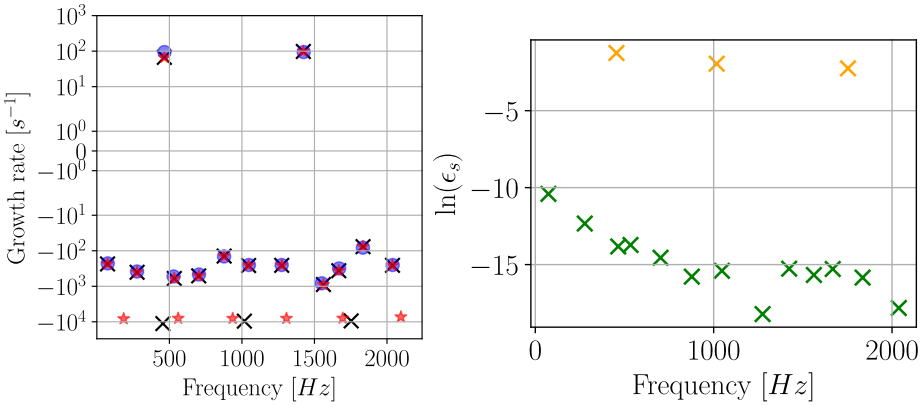

Figure 12: Left: direct results of the modal expansion method using either $N_{PBF} = 20$ (black crosses) or $N_{PBF} = 25$ (red stars), and pseudo analytical solutions (blue dots). Right: $\ln(\epsilon_s)$ computed for each modes (each black crosses on the left figure). The coloring is based on the K-means algorithm with two clusters.

In this academic configuration, the two clusters of physical and non-physical modes have been well separated, allowing to clearly identify the physical ones (green crosses in right part of Fig. 12). The capability of the overall methodology to handle a more complex case is investigated in the remaining of the paper.

5. Experimental configuration

The methodology to make the frame-based projection approach more reliable and predictive is tested on an experimental configuration studied at the Institut de Mecanique des Fluides de Toulouse (IMFT), named MIRADAS. The configuration setup is detailed in Fig. 1 and Fig. 2 in [32]. Oztarlik *et al.* have notably measured the Flame Transfer Function for different operating points (see Fig. 14 in [32]). In this study, the operating point named "Ref" is studied, which presents an unstable mode around $f = 591$ Hz (see Tab. 2 in [32]). The FTF and its fit with the Vector Fitting algorithm is depicted in the left part of Fig. 13. The reflection coefficient $R(\omega)$ at the output of the combustion chamber was also measured by Oztarlik, given in Fig. 4.7 in [33]. As this reflection coefficient is depending on ω, the outlet impedance $Z(\omega) = (1 + R)/(1 - R)$ must also be fitted with the Vector Fitting algorithm (see right part of Fig. 13)

The mean physical fields $\gamma(\boldsymbol{x})$, $\rho_0(\boldsymbol{x})$ and the value of p_0 used to compute the frame modal basis (see Eq. 7) are determined by the LES solver AVBP [34] developed at CERFACS and interpolated on the mesh used for the presented thermoacoustic study (see Fig. 14). The value of the mean pressure is $p_0 = 101400\ Pa$.

As an impedance boundary condition has to be applied at the outlet of the combustion chamber, the two orthogonal basis constituting the frame are computed by applying either a Dirichlet or a Neumann boundary condition at the outlet. The modes are computed with the FEniCS software [35, 36], using Krylov-Schur method. The number of modes in each basis is set to 30, leading to a total of 60 modes in the initial frame. The solver tolerance is set to 10^{-7}. As detailed in Appendix 8.3, the frame truncation threshold should theoretically be set close to the solver tolerance. To ensure the quality of the truncated frame, the truncation threshold is set just above the solver tolerance: $\varepsilon_t = 10^{-6}$ (see the singular value distribution in Fig. 15). All the modes associated with a singular value lower than the threshold, totaling 23 modes, are removed. This process effectively eliminates 23 non-physical modes from the overall output of the method.

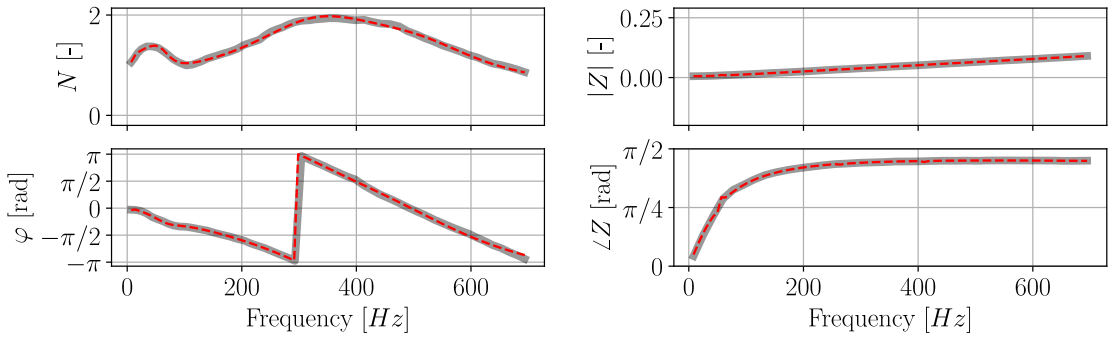

Figure 13: Left: Modulus and phase of the flame Transfer Function associated with the "Ref" operating point in Oztarlik *et al* [33]. Right: Modulus and phase of the impedance boundary condition at the combustion chamber outlet. Gray lines: experimental measurements from [32, 33]. Red dashed lines: the fits reconstructed using Eq. 29 and the Vector Fitting algorithm [31]. The number of Pole Base Function is $N_{PBF}^{fl} = 20$ for the FTF fit and $N_{PBF}^{imp} = 11$ for the impedance fit.

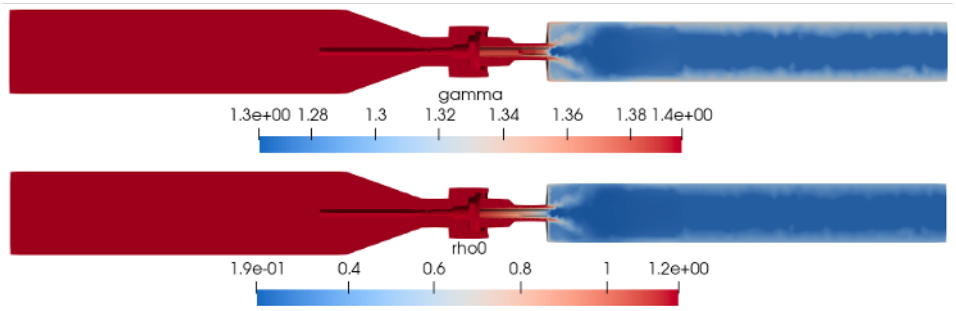

Figure 14: Mean fields γ and ρ_0 used to compute the frame modal basis, determined using Large Eddy Simulation.

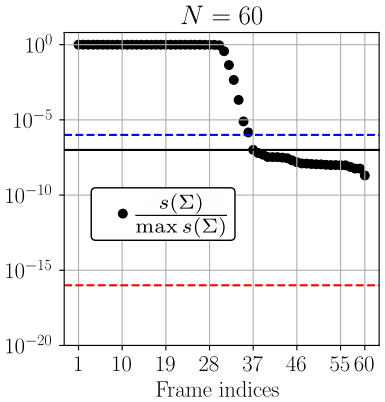

Figure 15: Singular value distribution of the Σ matrix normalized by their maximum, associated with the MIRADAS configuration. The red dashed line corresponds to 10^{-16}, an indication of the machine precision. The blue dashed line corresponds to the truncation threshold set to $\varepsilon_t = 10^{-6}$, just above the solver tolerance at 10^{-7} (horizontal black line).

The presented method is compared to the well-established thermoacoustic code AVSP developed at CERFACS [37]. As two different fit are used in this configuration (impedance and flame), the criterion to identify the non-physical modes uses a cross-comparison based on the Eq. 30. A total of six simulations were performed, using two different values for the impedance fit order: $N_{PBF}^{imp} = 8$ and 11, and three different values for the flame fit order: $N_{PBF}^{fl} = 6$, 20 and 26. The FTF and the impedance are still well-fitted using these orders. As the criterion to identify the spurious modes is based on a frequency shift, orders in the fits are chosen significantly different, so that the poles are very distinct. The case $N_{PBF}^{fl} = 20$ and $N_{PBF}^{imp} = 11$ is defined as the main case and five sets of errors $\epsilon_s(n)$ (see Eq. 30) are computed using the five other simulations. As the cost of one fit followed by one simulation is about 2 seconds, we are not limited by the number of simulations in terms of computational cost. These sets of errors are added to define a global error: $\epsilon_g(n) = \sum_{i=1}^{5} \epsilon_s^i(n)$, where the index i refers to the five different cases summarized in Tab. 2 . The logarithm of $\epsilon_g(n)$ is plotted in Fig. 16.

Simulation index (i)	N_{PBF}^{imp}	N_{PBF}^{fl}
Main	11	20
1	11	6
2	11	26
3	8	6
4	8	20
5	8	26

Table 2: Value of the fit order in the six different simulations performed used to identify non-physical components in the output.

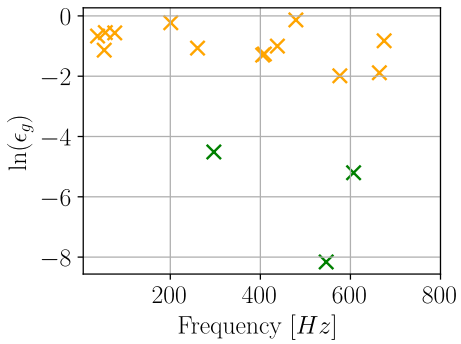

Figure 16: Log of the global error ϵ_g, computed as the sum of five sets of errors ϵ_s. The coloring is based on the K-means algorithm with two clusters.

The frequency and growth rate of the three modes with the lowest global errors $\epsilon_g(n)$ (green crosses in Fig. 16) are plotted in the frequency plane in the right part of Fig. 17.

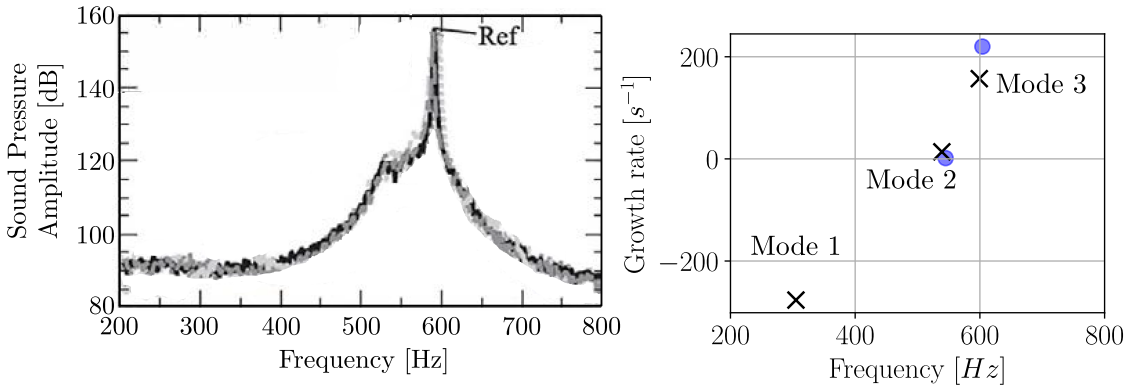

Figure 17: Left: Experimental measurements of the pressure amplitude (adapted from Oztarlik [32]). Right: Numerical results using the truncated SVD modal expansion method (black crosses) and the FEM solver AVSP (blue dots). Modes are indexed in increasing order of frequency.

The presented method found an unstable mode (mode 3) at $f_3 = 606$ Hz, $\sigma_3 = 25\ s^{-1}$ which is consistent with the mode measured in the experiment at 591 Hz. It corresponds to the coupling between the flame and the second pure acoustic mode (at 602 Hz).

The second mode at $f_2 = 546$ Hz is computed as a marginally unstable mode (low growth rate: $\sigma_2 = 1.2\ s^{-1}$), and corresponds to a small coupling between the flame and the first pure acoustic mode of the geometry, which is at 558 Hz. The first mode at $f_1 = 297$ Hz is only found by the presented method.

The Finite Element solver AVSP requires an initial guess f_0 to compute one by one the thermoacoustic modes. By setting the initial guess at f_1, AVSP converges to the mode 2. As the phase difference of the mode 1 between the velocity upstream and downstream of the flame is $-\pi$ (see Fig. 18), the mode can be classified as an Intrinsic Thermoacoustic (ITA) mode [38].

Furthermore, as the basin of attraction of ITA modes is very small when using iterative algorithms to solve the non-linear eigenvalue problem [16], AVSP failed to find this mode.

The Contour Integration method [39, 18] seems to be a good alternative technique to overcome the initial guess issue, but it is not studied in this paper.

Fortunately, the presented modal expansion method do not require any initial guess, and all the thermoacoustic modes are computed in one step, as the eigenproblem which defines the system is linear.

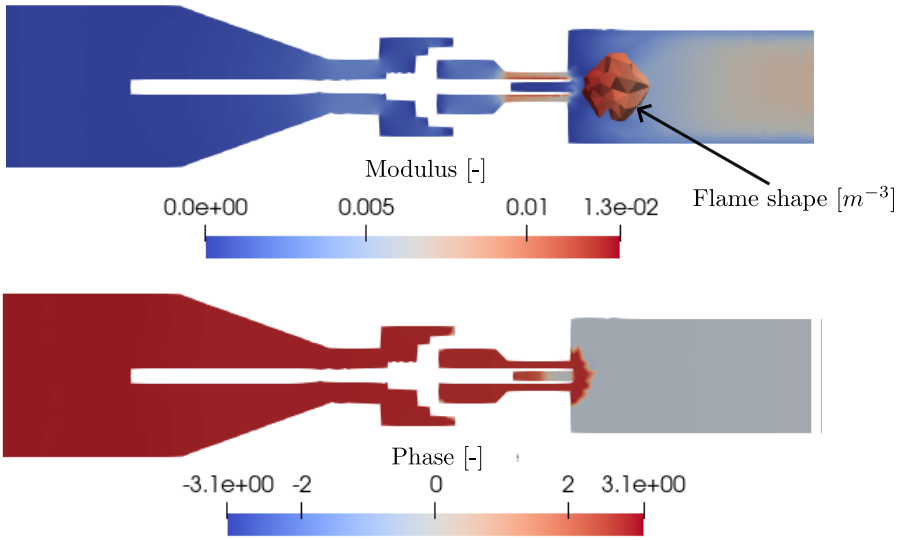

Figure 18: Top: Modulus of velocity fluctuations associated with the mode 1, normalized by its L2-norm. Flame shape $H(\boldsymbol{x})$ used in the simulation that satisfies $\int_V H(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1$. Bottom: Phase of the mode 1.

To give an order of magnitude of computation times in this configuration ($\sim$ 30000 degrees of freedom), the non-linear eigensolver AVSP takes about 30min - 1h to compute the two thermoacoustic modes without the ITA mode. On the other hand, the presented frame modal expansion method requires to build the initial frame. So 60 pure acoustic modes are computed using the Finite Element Method ($\sim$ 5 *min*), and the global thermoacoustic system is assembled using the state-space formalism. That leads to a linear eigen-system of size $\sim$ 100, which is solved by a direct solver in 1 second. The spurious mode identification process which identifies the three physical modes requires 5 other simulation of 1 second each, so that takes about 10 seconds to be performed. Moreover, a small change in the FTF or in the boundary condition is free to compute with the presented method, as the 60 modes constituting the frame are already computed.

6. Conclusion

This article improves the formalism recently proposed by Laurent *et al.* [23, 20] for representing any geometrical subdomains using frame modal expansion. It is shown that the generation of spurious modes is related to the ill-conditioning of the Gram matrix associated with the frame. A strategy based on the truncated SVD is proposed and successfully tested on an academic and an experimental configuration. The inverse of ill-conditioned Gram matrix $\boldsymbol{\Lambda}$ is no longer something to compute. The system using the truncated SVD is way better conditioned than using the direct frame for two reasons:

- The condition number is lower by construction, $\kappa(\boldsymbol{\Sigma}) = \sqrt{\kappa(\boldsymbol{\Lambda})}$.
- The frame size is reduced, by removing the frame multicollinearities.

The $\boldsymbol{U}$ basis is an orthonormal basis constructed as a linear combination of vectors of the frame, so it does not represent any particular fluctuations. Still, a modified Helmholtz equation that couples the $\boldsymbol{U}$ components is required to derive the state-space formulation associated with the SVD-based modal expansion. However, this equation is not derived from physical laws, it is just verified by U_n vectors. If the frame is built analytically, the value of the threshold ε_t in the truncated SVD has to be set just above the machine precision.

For more complex 3D geometries, the frame cannot be built analytically as for the MIRADAS configuration presented in this paper, so it must be computed from the outcome of an homogeneous Helmholtz solver. The numerical quality of the frame itself depends on the mesh refinement, therefore the truncation threshold does as well. It must be set higher than the plateau-like behaviour in the singular value distribution of the frame.

This article also provides an automatic method to identify non-physical output components due to the fit of the Flame Transfer Function or the impedance, based on a frequency comparison and using the K-means algorithm to separate the two clusters. This method remains sensitive to the studied system and even if it gives correct results on the 1D duct and MIRADAS cases presented, it must be tested/validated on other complex configurations.

Given the general nature of the developments proposed, the truncated SVD based modal expansion method is expected to be effective in reducing numerical quality problems in any thermoacoustic case, including complex geometries inspired by industrial systems.

7. Acknowledgments

The authors would like to acknowledge the financial support provided by the Convention Industrielle de Formation par la Recherche (CIFRE) program (n°2021/1380), managed by the Association Nationale de Recherche et Technologie (ANRT), in collaboration with Safran Aircraft Engines.

References

[1] T. Poinsot, Prediction and control of combustion instabilities in real engines, Proceedings of the Combustion Institute 36 (2017) 1–28.

[2] A. Urbano, L. Selle, G. Staffelbach, B. Cuenot, T. Schmitt, S. Ducruix, S. Candel, Exploration of combustion instability triggering using Large Eddy Simulation of a multiple injector liquid rocket engine, Combustion and Flame 169 (2016) 129–140.

[3] L. Y. M. Gicquel, G. Staffelbach, T. Poinsot, Large Eddy Simulations of gaseous flames in gas turbine combustion chambers, Progress in Energy and Combustion Science 38 (2012) 782–817.

[4] T. Poinsot, D. Veynante, Theoretical and numerical combustion, R. T. Edwards, Inc., 2005.

[5] L. Crocco, S. I. Cheng, Theory of combustion instability in liquid propellant rocket motors, volume Agardograph No 8, Butterworths Science, 1956.

[6] C. O. Paschereit, B. Schuermans, W. Polifke, O. Mattson, Measurement of Transfer Matrices and Source Terms of Premixed Flames, Journal of Engineering for Gas Turbines and Power 124 (2002) 239–247.

[7] K. Truffin, T. Poinsot, Comparison and extension of methods for acoustic identification of burners, Combustion and Flame 142 (2005) 388–400.

[8] N. Noiray, D. Durox, T. Schuller, S. Candel, A unified framework for nonlinear combustion instability analysis based on the describing function, Journal of Fluid Mechanics 615 (2008) 139–167.

[9] P. Palies, D. Durox, T. Schuller, S. Candel, Nonlinear combustion instability analysis based on the flame describing function applied to turbulent premixed swirling flames, Combustion and Flame 158 (2011) 1980–1991.

[10] A. Cuquel, D. Durox, T. Schuller, Impact of flame base dynamics on the non-linear frequency response of conical flames, Comptes Rendus Mécanique 341 (2013) 171–180.

[11] K. S. Kedia, H. M. Altay, A. F. Ghoniem, Impact of flame-wall interaction on premixed flame dynamics and transfer function characteristics, Proceedings of the Combustion Institute 33 (2011) 1113–1120.

[12] C. Fureby, LES of a Multi-burner Annular Gas Turbine Combustor, Flow, Turbulence and Combustion 84 (2010) 543–564.

[13] S. Ducruix, D. Durox, S. Candel, Theoretical and experimental determinations of the transfer function of a laminar premixed flame, Proceedings of the Combustion Institute 28 (2000) 765–773.

[14] M. Bauerheim, J.-F. Parmentier, P. Salas, F. Nicoud, T. Poinsot, An analytical model for azimuthal thermoacoustic modes in an annular chamber fed by an annular plenum, 161 (2014) 1374 – 1389.

[15] F. Nicoud, L. Benoit, C. Sensiau, T. Poinsot, Acoustic Modes in Combustors with Complex Impedances and Multidimensional Active Flames, AIAA Journal 45 (2007) 426–441.

[16] G. A. Mensah, P. E. Buschmann, A. Orchini, Iterative solvers for the thermoacoustic nonlinear eigenvalue problem and their convergence properties, International Journal of Spray and Combustion Dynamics 14 (2022) 30–41. Publisher: SAGE Publications Ltd STM.

[17] C. F. Silva, Intrinsic thermoacoustic instabilities, Progress in Energy and Combustion Science 95 (2023) 101065.

[18] P. E. Buschmann, G. A. Mensah, F. Nicoud, J. P. Moeck, Solution of Thermoacoustic Eigenvalue Problems With a Nonlinear Method, Journal of Engineering for Gas Turbines and Power 142 (2020) 031022.

[19] M. J. Gander, G. Wanner, From Euler, Ritz, and Galerkin to Modern Computing, SIAM Review 54 (2012) 627–666.

[20] C. Laurent, M. Bauerheim, T. Poinsot, F. Nicoud, A novel modal expansion method for low-order modeling of thermoacoustic instabilities in complex geometries, Combustion and Flame 206 (2019) 334–348.

[21] M. R. Bothien, N. Noiray, B. Schuermans, Analysis of Azimuthal Thermo-acoustic Modes in Annular Gas Turbine Combustion Chambers, Journal of Engineering for Gas Turbines and Power 137 (2015) 061505.

[22] A. Orchini, S. Illingworth, M. Juniper, Frequency domain and time domain analysis of thermoacoustic oscillations with wave-based acoustics, Journal of Fluid Mechanics 775 (2015) 387–414.

[23] C. Laurent, A. Badhe, F. Nicoud, Representing the geometrical complexity of liners and boundaries in low-order modeling for thermoacoustic instabilities, Journal of Computational Physics 428 (2021) 110077.

[24] L. Crocco, Aspects of combustion instability in liquid propellant rocket motors. Part I., 21 (1951) 163–178.

[25] T.Emmert, M.Meindl, S.Jaensch, Linear state space interconnect modeling of acoustic systems, Acta Acustica united with Acustica (2016).

[26] B. Schuermans, V. Bellucci, C. O. Paschereit, Thermoacoustic Modeling and Control of Multi Burner Combustion Systems, in: Volume 2: Turbo Expo 2003, ASMEDC, Atlanta, Georgia, USA, 2003, pp. 509–519. URL: https://asmedigitalcollection.asme.org/GT/proceedings/GT2003/36851/509/296081. doi:10.1115/GT2003-38688.

[27] P. M. Morse, K. U. Ingard, Theoretical acoustics, volume 332, Princeton University Press, 1968.

[28] F. Öztürk, F. Akdeniz, Ill-conditioning and multicollinearity, Linear Algebra and its Applications 321 (2000) 295–305.

[29] G. Golub, C. Van Loan, Matrix Computations, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, 2013. URL: https://books.google.fr/books?id=5U-l8U3P-VUC.

[30] P. Triverio, Robust Causality Check for Sampled Scattering Parameters via a Filtered Fourier Transform, IEEE Microwave and Wireless Components Letters 24 (2014) 72–74.

[31] P. Triverio, Vector Fitting, 2019. URL: http://arxiv.org/abs/1908.08977, arXiv:1908.08977 [physics].

[32] G. Oztarlik, L. Selle, T. Poinsot, T. Schuller, Suppression of instabilities of swirled premixed flames with minimal secondary hydrogen injection, Combustion and Flame 214 (2020) 266–276.

[33] G. Oztarlik, Numerical and Experimental Investigations of Combustion Instabilities of Swirled Premixed Methane-Air Flames With Hydrogen Addition (2020).

[34] T. Schønfeld, M. Rudgyard, Steady and unsteady flows simulations using the hybrid flow solver avbp, 37 (1999) 1378–1385.

[35] H. P. Langtangen, A. Logg, Solving PDEs in Python, Springer International Publishing, Cham, 2016. URL: http://link.springer.com/10.1007/978-3-319-52462-7. doi:10.1007/978-3-319-52462-7.

[36] I. A. Baratta, J. P. Dean, J. S. Dokken, M. Habera, J. S. Hale, C. N. Richardson, M. E. Rognes, M. W. Scroggs, N. Sime, G. N. Wells, DOLFINx: The next generation FEniCS problem solving environment, 2023. URL: https://zenodo.org/records/10447666. doi:10.5281/zenodo.10447666.

[37] L. Benoit, F. Nicoud, Numerical assessment of thermo-acoustic instabilities in gas turbines, 47 (2005) 849–855.

[38] K. J. Yong, C. F. Silva, G. J. J. Fournier, W. Polifke, Categorization of Thermoacoustic Modes in an Ideal Resonator with Phasor Diagrams, 2022. URL: http://arxiv.org/abs/2211.03842. doi:10.48550/arXiv.2211.03842, arXiv:2211.03842 [physics].

[39] M. Merk, P. E. Buschmann, J. P. Moeck, W. Polifke, The Nonlinear Thermoacoustic Eigenvalue Problem and Its Rational Approximations: Assessment of Solution Strategies, Journal of Engineering for Gas Turbines and Power 145 (2023) 021028.

[40] Y. Saad, Numerical methods for large eigenvalue problems: revised edition, SIAM, 2011.

8. Appendix

8.1. Metric adaptation

The initial frame $\boldsymbol{\phi}(\vec{x})$ and the orthonormalized basis $\boldsymbol{U}(\vec{x})$ are row-vectors defined as follows: $\boldsymbol{\phi}(\vec{x}) = [\phi_1(\vec{x}) \ \cdots \ \phi_N(\vec{x})]$ and $\boldsymbol{U}(\vec{x}) = [U_1(\vec{x}) \ \cdots \ U_N(\vec{x})]$. In this notation, each component $\phi_i(\vec{x})$ or $U_i(\vec{x})$ is a vector of size N_v, the number of vertices of the mesh. The SVD orthonormalization process must be done with respect to the inner product: $\langle \phi_n | \phi_m \rangle = \int_V \phi_n(\vec{x}) \phi_m(\vec{x}) \mathrm{d}V$, involving the volume integral of the component-wise product of the two vectors ϕ_n and ϕ_m. From a numerical point of view, this inner product must be weighted by the volume of each vertices of the mesh as follows:

$$\langle \phi_n | \phi_m \rangle = \phi_n^T \boldsymbol{M} \phi_m \tag{31}$$

where $\boldsymbol{M}$ is the mass matrix associated with the mesh. This quantity is inherent to each mesh, and is calculated as the inner product between each basis function spaces of the mesh. If the SVD is directly done on $\boldsymbol{\phi}$ without any modification, the algorithm will orthonormalize the frame with respect to the numerical inner product $\sum_{i=1}^{N_v} \phi_n(x_i)\phi_m(x_i)$ without considering the volume of each vertex $\mathrm{d}V_i$. Thus, a metric adaptation is applied on $\boldsymbol{\phi}$ before computing the SVD. The metric-adapted frame $\tilde{\boldsymbol{\phi}}$ is:

$$\tilde{\boldsymbol{\phi}} = \boldsymbol{C}\boldsymbol{\phi} \tag{32}$$

where $\boldsymbol{C}$ is the lower triangulate matrix from the Cholesky decomposition of the mass matrix such as: $\boldsymbol{M} = \boldsymbol{C}^T\boldsymbol{C}$. The SVD is then done on this metric-corrected frame $\tilde{\boldsymbol{\phi}}$ to orthonormalize the frame with respect to the correct numerical inner product defined in Eq. 31. It writes:

$$\tilde{\boldsymbol{\phi}} = \tilde{\boldsymbol{U}}\boldsymbol{\Sigma}\boldsymbol{V}^T \tag{33}$$

Finally, the $\boldsymbol{U}$ basis required is obtained with $\boldsymbol{U} = \boldsymbol{C}^{-1}\tilde{\boldsymbol{U}}$.

8.2. State space realisation of the flame block

Under the Vector Fitting formalism, the FTF $Q(t)$ writes:

$$Q(t) = \frac{\bar{Q}}{\bar{u}}\left(R_0 + \sum_k^{N_{PBF}} \frac{R_k}{j\omega - p_k}\right) u'_n(\boldsymbol{x_r}, t) \tag{34}$$

and can be recast as a state-space formulation as follows:

$$\frac{d}{dt}\underbrace{\begin{bmatrix} z_1 \\ \vdots \\ z_{N_{PBF}} \\ x(t) \end{bmatrix}}_{X_f} = \underbrace{\begin{bmatrix} p_1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & 0 \\ 0 & \cdots & p_{N_{PBF}} & 0 \\ 0 & \cdots & 0 & -\Omega_0 \end{bmatrix}}_{A_f} \begin{bmatrix} z_1 \\ \vdots \\ z_{N_{PBF}} \\ x(t) \end{bmatrix} + \underbrace{\begin{bmatrix} R_1 \\ \vdots \\ R_{N_{PBF}} \\ \Omega_0 \end{bmatrix}}_{B_f} \underbrace{\left[u'_n(\boldsymbol{x_r}, t)\right]}_{U_f} \tag{35}$$

$$\underbrace{\left[Q(t)\right]}_{Y_f} = \frac{\bar{Q}}{\bar{u}} \underbrace{\begin{bmatrix} 1 & \cdots & 1 & R_0 \end{bmatrix}}_{C_f} \begin{bmatrix} z_1 \\ \vdots \\ z_{N_{PBF}} \\ x(t) \end{bmatrix} \tag{36}$$

where the variable z_k is introduced for convenience: $z_k = R_k u'_n(\boldsymbol{x_r}, t)/(j\omega - p_k)$.

Ω_0 is set to a high value to built a time-derivator in the last line, forcing the variable $x(t)$ to exponentially converge to $u'_n(\boldsymbol{x_r}, t)$. This trick allows to put a variable from the input vector U into the state vector X (here $u'_n(\boldsymbol{x_r}, t)$).

8.3. Short note on the threshold selection for the eigensolver and truncated SVD

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix.

8.3.1. Residual versus forward error on eigenvectors

The purpose of this section is to describe a bound on the norm of the difference between an exact unitary eigenvector ϕ of A and the one computed numerically. This bound includes the 2-norm of the residual $r = (A - \hat{\lambda}I)\hat{\phi}$ associated with the corresponding eigenpair ($\hat{\lambda} = \hat{\phi}^T A\hat{\phi}, \hat{\phi}$) and the distance with the closest eigenvalue of $\hat{\lambda}$.

Theorem 1. *(See e.g., [40, Theorem 3.9, p. 63]).*
Let $\hat{\phi}$ be an approximate eigenvector of unit norm of A, $\hat{\lambda} = \hat{\phi}^T A\hat{\phi}$ its associated approximate eigenvalue and $r = (A - \hat{\lambda}I)\hat{\phi}$. Let λ be the eigenvalue of A closest to $\hat{\lambda}$ and δ the distance from $\hat{\lambda}$ to the rest of the spectrum, i.e., $\delta = \min_i\{|\lambda_i - \hat{\lambda}|, \lambda_i \neq \lambda\}$. The eigenvalue for which the minimum distance δ is reached is noted λ_j, i.e $\delta = |\tilde{\lambda} - \lambda_j|$. If ϕ is the eigenvector of A associated with λ we have

$$\sin\theta(\hat{\phi}, \phi) \leq \frac{||r||_2}{\delta}. \tag{37}$$

Let us write $\hat{\phi} = \cos\theta\phi + \sin\theta z$ where z is a unit vector orthogonal to u. If $\theta \ll 1$, so that $\cos\theta \approx 1$ and $\sin\theta \approx \theta$ then $\hat{\phi} - \phi \approx \theta z$ so that

$$||\hat{\phi} - \phi|| \leq \frac{||r||_2}{\delta}.$$

Notice that the assumption $\theta \ll 1$ can be discarded to get an exact expression of

$$||\hat{\phi} - \phi|| = \tan\theta(1 + \sin^2\theta)^{\frac{1}{2}} \tag{38}$$

that is not much useful as we shall try to compute the eigenpairs accurately enough using a stopping criterion based on the residual (possibly scaled).

In practice the closest eigenvalue to $\hat{\lambda}$ is not known so that δ cannot be computed, but a lower bound can be derived that only involves the computed eigenvalues and the residual norm [40, p. 64]:

$$\delta = |\hat{\lambda} - \lambda_j| \geq |\hat{\lambda} - \hat{\lambda}_j| - \|r_j\|_2.$$

This result indicates that the computed eigenvectors are at a distance that depends on the associated residual norm and how close are the computed eigenvalues of A.

For a set of computed eigenvectors (frame) $\hat{\Phi} = [\hat{\phi}_1, ..., \hat{\phi}_m]$ we will give in the next section the norm of the difference between $\hat{\phi}_i$ and its approximation extracted from its truncated SVD.

8.3.2. Error associated with a computed eigenvector $\hat{\phi}$ versus truncated SVD of $\hat{\Phi}$

Let $\hat{\Phi} = U\Sigma V^T$ be the SVD of $\hat{\Phi}$, where $U = [u_1, ..u_r] \in \mathbb{R}^{n\times r}$ are the left singular vectors that form an orthonormal basis of $range(\hat{\Phi})$, Σ is a diagonal matrix where the diagonal entries are the non-zero singular values σ_i is a decreasing order and $V = [v_1, ..., v_r] \in \mathbb{R}^{m\times r}$ the right singular vectors.

For $k \leq m$, we define the best rank k approximation of $\hat{\Phi}$ as $\hat{\Phi}_k = U_k\Sigma_k V_k^T$ with $U_k = [u_1, ..., u_k]$, ...

Theorem 2. *Eckart-Young Theorem (See e.g., [29, Theorem 2.4.8, p. 79]).*

$$\min_{rank(\Psi)=k} \|\hat{\Phi} - \Psi\|_2 = \|\hat{\Phi} - \hat{\Phi}_k\|_2 = \sigma_{k+1}$$

Coming back to the definition of the 2-norm of a matrix that is

$$\|\hat{\Phi}\|_2 = \max_{\|x\|_2=1} \|\hat{\Phi}x\|_2$$

and considering the canonical basis vectors we have

$$\|(\hat{\Phi} - \hat{\Phi}_k)e_i\|_2 = \|\hat{\phi}_i - \hat{\Phi}_k e_i\|_2 \leq \sigma_{k+1}.$$

This shows that the best approximation of any vector $\hat{\phi}$ in the space spanned by U_k is at most at a distance (in 2-norm) of σ_{k+1}, that is

$$\min_{\tilde{\phi}\in span(U_k)} \|\hat{\phi}_i - \tilde{\phi}\| \leq \sigma_{k+1} \tag{39}$$

The bounds (38) and (39) indicate that the stopping criterion threshold for the eigensolvers used to compute the $\hat{\phi}$ and the threshold used to define the truncated SVD should be chosen consistently. In practice, the truncation threshold is chosen to be ten times higher than the solver tolerance.