ORIGINAL ARTICLE

Brandon C. Y. Yeung · Oliver T. Schmidt

Adaptive spectral proper orthogonal decomposition of broadband-tonal flows

Abstract An adaptive algorithm for spectral proper orthogonal decomposition (SPOD) of mixed broadband-tonal turbulent flows is developed. Sharp peak resolution at tonal frequencies is achieved by locally minimizing bias of the spectrum. Smooth spectrum estimates of broadband regions are achieved by locally reducing variance of the spectrum. The method utilizes multitaper estimation with sine tapers. An iterative criterion based on modal convergence is introduced to enable the SPOD to adapt to spectral features. For tonal flows, the adaptivity is controlled by a single user input; for broadband flows, a constant number of sine tapers is recommended without adaptivity. The discrete version of Parseval's theorem for SPOD is stated. Proper normalization of the tapers ensures that Parseval's theorem is satisfied in expectation. Drastic savings in computational complexity and memory usage are facilitated by two aspects: (i) sine tapers, which permit *post hoc* windowing of a single Fourier transform; and (ii) time-domain lossless compression using a QR or eigenvalue decomposition. Sine-taper SPOD is demonstrated on time-resolved particle image velocimetry (TR-PIV) data from an open cavity flow (Zhang et al. in Exp Fluids 61(226):1–12, https://doi.org/10.1007/s00348-020-03057-8, 2020) and high-fidelity large-eddy simulation (LES) data from a round jet (Brès et al. in J. Fluid Mech. 851:83–124, https://doi.org/10.1017/jfm.2018.476, 2018), with and without adaptivity. For the tonal cavity flow, the adaptive algorithm outperforms Slepian-based multitaper SPOD in terms of variance and local bias of the spectrum, mode convergence, and memory usage. The tonal frequencies associated with the Rossiter instability are accurately identified. For both the tonal cavity and the broadband jet flows, results comparable to or better than those from standard SPOD based on Welch's overlapped segment averaging are obtained with up to 75% fewer snapshots, including similar convergence of the Rossiter modes and Kelvin-Helmholtz wavepacket structures for the cavity and jet examples, respectively. Drawing from these examples, we establish best practices.

Keywords SPOD · Sine taper · Adaptive · Broadband-tonal turbulence · PIV · LES

1 Introduction

Owing to its stochastic nature, turbulence has been studied using a variety of data-driven statistical approaches. Among them are dynamic mode decomposition (DMD, [16]) and proper orthogonal decomposition (POD, [7]), the latter of which forms the subject of this paper. POD in the most general form seeks a set of orthogonal modes

---

B. C. Y. Yeung· O. T. Schmidt (✉)
Department of Mechanical and Aerospace Engineering, Jacobs School of Engineering, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0411, USA
E-mail: oschmidt@ucsd.edu

B. C. Y. Yeung
E-mail: byeung@ucsd.edu

that optimally represent the second-order space-time statistics of turbulent flows. For wide-sense stationary flows, space-time POD was later specialized by adding a temporal Fourier transform of the data, before a space-only POD is performed per frequency (see [5] for an early example). This procedure is termed spectral POD (SPOD, [12]). Since SPOD modes are monochromatic and coherent in both space and time, they can be interpreted as coherent structures. Convergence of SPOD modes and their modal energies often demands a large number of realizations of the Fourier transform. Collecting sufficient samples can be experimentally, numerically, and literally costly.

In the standard implementation of SPOD [18, 20, 27], the samples are obtained using Welch estimation [29]. The time-domain data are first segmented into overlapping blocks, each of which is then Fourier transformed. The block size is inversely related to the frequency resolution of the spectrum. For the sample size to be large, each block must be short, which lowers resolution and increases bias. Conversely, long blocks raise resolution, but make fewer samples available, increasing the variance of the estimate. More precise control of this trade-off can be achieved with Thomson's multitaper estimator [26], recently adopted for SPOD by [17]. Rather than segmenting the data, Thomson's multitaper windows them using a family of mutually orthogonal tapers, called discrete prolate spheroidal sequences (DPSS, also known as Slepians, [24, 25]), that render the Fourier realizations approximately uncorrelated with each other. As each Fourier transform spans the entire data, spectral resolution is maximized. Independent of resolution, the variance and bias of the estimate are controlled via the number and bandwidth of the tapers. The DPSS are the family of tapers that maximize spectral concentration. Thomson also proposed adaptive weighting of the windowed Fourier transforms to offer additional protection against spectral leakage [26]. However, Thomson's multitaper estimator offers scant protection from local bias [14], a phenomenon that causes discrete tones in the spectrum to become smeared or flattened, though improvements were later made by [13]. In the context of SPOD, [17] found that Thomson's multitaper also incurs dramatically higher computational cost and memory usage relative to the Welch estimator and, as a compromise, suggested a hybrid multitaper-Welch approach.

Our overarching goal is to formulate an SPOD algorithm that combines strong variance reduction with low bias, and that is computationally efficient. Such an algorithm would be well-suited to estimating spectra that include tones due to hydrodynamic instabilities, resonances, and exogenous inputs. It would also be able to handle large data, such as those produced by high-fidelity simulations. To this end, we propose an adaptive multitaper SPOD that uses sine functions as tapers, with the aim of achieving better local bias protection and lower memory usage than DPSS-taper SPOD, while ensuring reasonable convergence of modes at all frequencies. Introduced by [14] for one-dimensional signal processing, sine tapers optimally minimize local bias. The properties of sine functions also allow an arbitrary number of sine-tapered Fourier realizations to be assembled from a single Fourier transform. The reduction in computational effort this affords may be immaterial to one-dimensional signal processing, but is crucial in the context of SPOD. To further mitigate computational complexity and peak memory consumption, we demonstrate lossless compression of the time-domain data. When a spectrum contains tones, frequency-dependent estimation parameters enable the bias and variance to adapt to the spectrum. Riedel and Sidorenko [14] derived an expression for the optimum number of sine tapers, in an expected error sense, to apply as a function of frequency, with the caveat that the expression depends on the unobservable true spectrum. Barbour and Parker [1] substituted the true spectrum with an estimated spectrum, then determined the taper number by iteratively updating the estimate. Hansson [6] obtained frequency-dependent weights for tapered Fourier transforms, including those computed using sine tapers, through an iterative procedure that also requires knowledge of the true spectrum. The application of sine tapers to spectral modal decomposition offers a novel pathway to realizing adaptivity. Unlike signal processing, SPOD accounts for spatial coherence. It requires mode shapes to be well-converged so that they may be employed for physical discovery and reduced-order modeling (ROM). Convergence of the mode shapes is in turn coupled to that of the spectrum. We propose an adaptive algorithm that protects tones in the spectrum from bias by leveraging the spatial information inherent in SPOD. Using mode convergence, the algorithm determines a taper number for each frequency such that a smaller number is applied to tones than to broadband turbulence. We demonstrate adaptive sine-taper SPOD on time-resolved particle image velocimetry (TR-PIV) data of an open cavity flow [31], which contain noise and mix tones with broadband turbulence, and show that the judicious selection of a single parameter, the similarity tolerance, suffices to generate a spectrum with low bin-to-bin variance and sharp tones, and smooth modes at a broad range of frequencies. Furthermore, we apply non-adaptive sine-taper SPOD to high-fidelity large-eddy simulation (LES) data of a turbulent round jet [3] to discuss the advantages of the technique for non-tonal, broadband flows.

The remainder of the paper is organized as follows. The standard SPOD algorithm is briefly summarized in Sect. 2.1. Section 2.2 introduces multitaper SPOD with sine tapers. Acceleration of SPOD via lossless

Table 1 Parameters of the cavity PIV and jet LES data

Case	M	$\mathbf{q}$	Δt	N_t	N_x	N_y or N_r
Cavity PIV	0.6	$[\mathbf{u}, \mathbf{v}]^T$	6.25×10^{-5} s	16,000	156	55
Jet LES	0.9	$[\hat{\boldsymbol{\rho}}, \hat{\mathbf{u}}_x, \hat{\mathbf{u}}_r, \hat{\mathbf{u}}_\theta, \hat{\mathbf{T}}]^T$	$0.18 D_j/u_j$	10,000	656	138

For the cavity, the time step is expressed in seconds (see Sect. 3). For the jet, the time step is non-dimensionalized by the nozzle diameter, D_j, and nozzle exit velocity, u_j (see Sect. 4). Only the axisymmetric component of the jet data is used
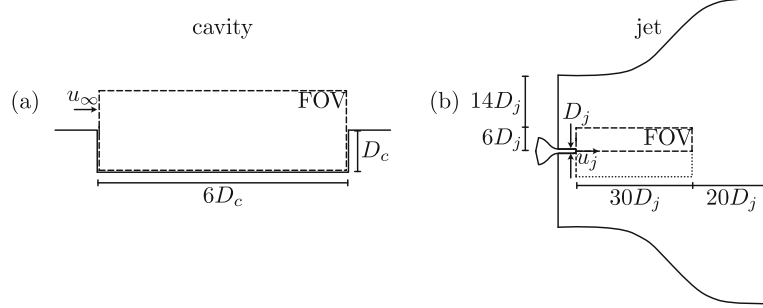

Fig. 1 2D representations of the geometries of the cavity PIV and jet LES data. The field of view (FOV) of each case is indicated by dashed lines. The cavity has a depth of D_c and a freestream velocity of u_∞. The jet has a nozzle diameter of D_j and a nozzle exit velocity of u_j. For the jet case, since only the axisymmetric component is considered, the region marked by dotted lines is redundant
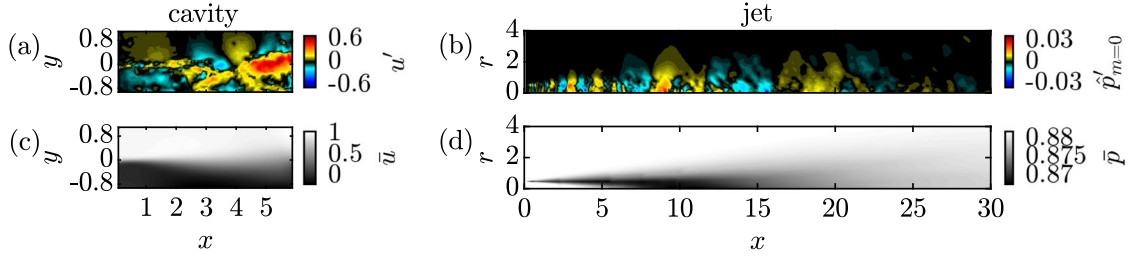

Fig. 2 Instantaneous fluctuations (**a**, **b**) and mean (**c**, **d**) of: **a**, **c** streamwise velocity component of the cavity PIV data; **b**, **d** axisymmetric pressure component of the jet LES data. For the remainder of the paper, for clarity of visualizations, we will focus on subdomains but have confirmed that all conclusions hold for the full domain

compression is developed in Sect. 2.3. Section 2.4 introduces the adaptive algorithm. Applications of sine-taper SPOD to the cavity and jet flows, with and without adaptivity, are presented in Sects. 3 and 4 respectively and discussed in Sect. 5. The parameters of the cavity and jet data are summarized in Table 1. Their geometries are illustrated in Fig. 1, and their flow fields are visualized in Fig. 2. Appendix A reports computational performance data. A strategy to tackle excessive time-to-solution is outlined in Appendix B.

2 Adaptive multitaper SPOD

2.1 SPOD

We begin with an outline of the SPOD algorithm. Given N_t snapshots $\mathbf{q}_{t_j} \in \mathbb{C}^{N_{\text{dof}}}$, $j = 1, \ldots, N_t$, of a statistically stationary flow sampled at N_{dof} points, where N_{dof} is the number of degrees of freedom (the product of the number of spatial locations and the number of variables) recorded, the Reynolds decomposition

$$\mathbf{q}_{t_j} = \bar{\mathbf{q}} + \mathbf{q}'_{t_j} \tag{1}$$

partitions $\mathbf{q}$ into the long-time mean, $\bar{\mathbf{q}}$, and the fluctuations, $\mathbf{q}'$. In general the snapshots may be complex, for example after Fourier transforms in one or more homogeneous directions. The fluctuations are assembled in the data matrix

$$\mathbf{Q}' = [\mathbf{q}'_{t_1}, \ldots, \mathbf{q}'_{t_{N_t}}] \in \mathbb{C}^{N_{\text{dof}} \times N_t}. \tag{2}$$

In the following, we will consider fluctuation quantities only and drop the $(\cdot)'$ notation.

At the core of SPOD is the eigendecomposition of the weighted cross-spectral density (CSD) matrix, $\mathbf{S}_{f_k} = \hat{\mathbf{X}}_{f_k}\hat{\mathbf{X}}_{f_k}^*$, where $\hat{\mathbf{X}}_{f_k} = \sqrt{\Delta t/N}\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{Q}}_{f_k}$, Δt is the time step between snapshots, $\mathbf{W}$ is a diagonal weight matrix, and the columns of $\hat{\mathbf{Q}}_{f_k}$ comprise N realizations of the temporal Fourier transform of $\mathbf{Q}$ at frequency f_k. At each frequency, the eigendecomposition

$$\mathbf{S}_{f_k}\mathbf{\Theta}_{f_k} = \mathbf{\Theta}_{f_k}\mathbf{\Lambda}_{f_k} \tag{3}$$

provides the SPOD modal energies, $\mathrm{diag}(\mathbf{\Lambda}_{f_k})$, and the modes, $\mathbf{\Phi}_{f_k} = \mathbf{W}^{-\frac{1}{2}}\mathbf{\Theta}_{f_k}$. Inclusion of the weight ensures that the modes are optimal in the spatial inner product

$$\langle \hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2 \rangle_{\mathbf{x}} = \hat{\mathbf{q}}_2^*\mathbf{W}\hat{\mathbf{q}}_1. \tag{4}$$

If $N_{\mathrm{dof}} > N$, as is typical in fluid flow applications, the SPOD is more efficiently computed using the method-of-snapshots [23]. Rather than explicitly constructing then diagonalizing the $N_{\mathrm{dof}} \times N_{\mathrm{dof}}$ CSD matrix, we solve the $N \times N$ eigendecomposition,

$$\hat{\mathbf{X}}_{f_k}^*\hat{\mathbf{X}}_{f_k}\mathbf{\Psi}_{f_k} = \mathbf{\Psi}_{f_k}\mathbf{\Lambda}_{f_k}, \tag{5}$$

to obtain the same nonzero eigenvalues, $\mathrm{diag}(\mathbf{\Lambda}_{f_k})$, and modes,

$$\mathbf{\Phi}_{f_k} = \mathbf{W}^{-\frac{1}{2}}\hat{\mathbf{X}}_{f_k}\mathbf{\Psi}_{f_k}\mathbf{\Lambda}_{f_k}^{-\frac{1}{2}}, \tag{6}$$

as before. For details on the SPOD algorithm, we refer the reader to [18,19,27].

2.2 Multitaper estimation

Estimation of the CSD matrix requires assembling $N > 1$ independent realizations of the Fourier mode (coefficient), $\hat{\mathbf{q}}_{f_k}^{(n)}, n = 1, \ldots, N$, at each frequency. The classic Welch method [29] produces these realizations by segmenting a single time series into $N = N_{\mathrm{blk}}$ partially overlapping blocks of length N_{fft}, before taking the windowed fast Fourier transform (FFT) of each block. The resulting spectral estimates have a normalized half-bandwidth of $\delta f = 1/N_{\mathrm{fft}}$. For this estimator, the variance-bias trade-off necessitates a compromise between maximizing N, which reduces variance in the estimates, and maximizing N_{fft}, which improves frequency resolution. In contrast, the multitaper estimator [26] produces N independent realizations using $N = N_{\mathrm{win}}$ mutually orthogonal windows, or tapers, $\mathbf{v}^{(n)} = [v_1^{(n)}, \ldots, v_{N_{\mathrm{fft}}}^{(n)}]$, then performing and storing an FFT for each window:

$$\hat{\mathbf{q}}_{f_k}^{(n)} = \sum_{j=0}^{N_{\mathrm{fft}}-1} v_{j+1}^{(n)}\mathbf{q}_{t_{j+1}}\mathrm{e}^{-\mathrm{i}2\pi jk/N_{\mathrm{fft}}}, \quad k = 0, \ldots, N_{\mathrm{fft}} - 1, \; n = 1, \ldots, N_{\mathrm{win}}. \tag{7}$$

The choice of tapers in the first application of multitaper estimation to SPOD [17] were the DPSS. These uniquely maximize the spectral concentration of the estimates within a given bandwidth and, lacking an analytical form, must be computed from a Hermitian eigendecomposition. Determining an optimal choice of N_{win} forms the subject of Sect. 2.4.

In controlling the variance-bias trade-off a different objective is to minimize the bias due to local variation of the true spectrum within the finite bandwidth of the estimator. By expanding the true spectrum in a Taylor series, it can be shown (see e.g. [26]) that this local bias is dominated by a term proportional to $\int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} |\hat{v}(f)|^2 f^2 \mathrm{d}f$, where $\hat{v}$ is the discrete-time Fourier transform of the taper, $\mathbf{v}$. The tapers that uniquely minimize this term are referred to as the minimum bias tapers [14]. Like the DPSS, they are the solutions to a Hermitian eigenvalue problem (EVP), but can be closely approximated by the set of scaled sinusoids [14],

$$v_j^{(n)} = \sqrt{\frac{2}{N_{\mathrm{fft}} + 1}} \sin \frac{\pi n j}{N_{\mathrm{fft}} + 1}, \quad j = 1, \ldots, N_{\mathrm{fft}}, \quad n = 1, \ldots, N_{\mathrm{win}}, \tag{8}$$

where j is the discrete time index, and n the taper index. The $\sqrt{2/(N_{\mathrm{fft}} + 1)}$ factor normalizes each sine taper such that $\sum_j (v_j^{(n)})^2 = 1$. This ensures that the FFT in Eq. (7) satisfies Parseval's theorem [11], which

also relates SPOD eigenvalues to average power. With this normalization, the discrete version of Parseval's theorem for SPOD guarantees that the eigenvalues at each frequency sum to the spatially-integrated power spectral density (PSD); the integral of the latter over frequency recovers the average power of the zero-mean stationary process. For discrete time and frequency, this relationship can be expressed as

$$\Delta f \sum_{k=0}^{N_{\mathrm{fft}}-1} \sum_{i=1}^{N_{\mathrm{win}}} E\{\lambda_{i,f_k}\} = \|\boldsymbol{\sigma}\|_{\mathbf{x}}^2 = \frac{1}{N_t} \sum_{j=1}^{N_t} E\left\{\left\|\mathbf{q}_{t_j}\right\|_{\mathbf{x}}^2\right\} \tag{9}$$

(Parseval's theorem for SPOD),

where $\Delta f = 1/(N_{\mathrm{fft}}\Delta t)$ is the frequency bin width, λ the estimated eigenvalues, $E\{\cdot\}$ the expectation operator, $\boldsymbol{\sigma}$ the pointwise true standard deviation, and $\|\mathbf{q}\|_{\mathbf{x}}^2 = \langle \mathbf{q}, \mathbf{q} \rangle_{\mathbf{x}}$. Given a finite number of samples—as is always the case when dealing with data—the identity can instead be approximated as

$$\Delta f \sum_{k=0}^{N_{\mathrm{fft}}-1} \sum_{i=1}^{N_{\mathrm{win}}} \lambda_{i,f_k} \approx \frac{1}{N_t} \sum_{j=1}^{N_t} \left\|\mathbf{q}_{t_j}\right\|_{\mathbf{x}}^2, \tag{10}$$

which is exact only for the rectangular window, with $v_j = 1/\sqrt{N_{\mathrm{fft}}}$. Conservation of power, i.e., properties (9) and (10), simplify comparisons between eigenvalues computed using different spectral estimation parameters.

Relative to DPSS tapers, which by construction are optimal in spectral concentration, sine tapers have slightly lower spectral concentration [14]. With DPSS tapers, customarily the user supplies the normalized half-bandwidth, δf, from which the number of usable tapers is derived as $N_{\mathrm{win}} = \lfloor 2\delta f N_{\mathrm{fft}} \rfloor - 1$ [26]. With sine tapers, the user instead directly supplies N_{win}, which translates to a half-bandwidth of $\delta f = (N_{\mathrm{win}} + 1)/[2(N_{\mathrm{fft}}+1)]$ [28]. In DPSS-taper SPOD [17], the Fourier realizations are uniformly weighted with weight $\mu_n = 1/N_{\mathrm{win}}$, $n = 1, \ldots, N_{\mathrm{win}}$. In sine-taper SPOD, we follow the recommendation of [14] and apply nonuniform weighting to each Fourier realization. The weights are given by

$$\mu_n = \frac{6}{4N_{\mathrm{win}}^3 + 3N_{\mathrm{win}}^2 - N_{\mathrm{win}}}[N_{\mathrm{win}}^2 - (n-1)^2], \quad n = 1, \ldots, N_{\mathrm{win}}, \tag{11}$$

and satisfy $\sum_{n=1}^{N_{\mathrm{win}}} \mu_n = 1$. They penalize higher-order tapers that possess larger bandwidths. With these weights, we now define the modified matrix of Fourier realizations as

$$\hat{\mathbf{X}}_{f_k} = \sqrt{\Delta t}\,\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{Q}}_{f_k}\mathrm{diag}([\mu_1, \ldots, \mu_{N_{\mathrm{win}}}])^{\frac{1}{2}}. \tag{12}$$

The simple analytical form of the sine tapers in Eq. (8) affords a significant computational simplification [14] via the identity

$$\hat{\mathbf{q}}_{f_k}^{(n)} = \sqrt{\frac{2}{N_{\mathrm{fft}}+1}}\frac{\hat{\mathbf{q}}_{f_{k-n\Delta k}} - \hat{\mathbf{q}}_{f_{k+n\Delta k}}}{2\mathrm{i}}, \quad n = 1, \ldots, N_{\mathrm{win}}, \tag{13}$$

where $\hat{\mathbf{q}}_{f_{k-n\Delta k}}$ and $\hat{\mathbf{q}}_{f_{k+n\Delta k}}$ are the unwindowed Fourier modes at frequency indices $k - n\Delta k$ and $k + n\Delta k$, and $\Delta k = N_{\mathrm{fft}}/[2(N_{\mathrm{fft}}+1)] \approx 0.5$ for $N_{\mathrm{fft}} \gg 1$.[1] The unwindowed modes, $\hat{\mathbf{q}}_{f_{k\pm n\Delta k}}$, for both even and odd n are obtained by zero-padding the snapshots in the time domain (see e.g. [9]) to length $2N_{\mathrm{fft}}$, then performing the FFT:

$$\hat{\mathbf{q}}_{f_k} = \sum_{j=0}^{2N_{\mathrm{fft}}-1} \mathbf{q}_{t_{j+1}}\mathrm{e}^{-\mathrm{i}2\pi jk/(2N_{\mathrm{fft}})}, \quad k = 0, \ldots, 2N_{\mathrm{fft}} - 1, \tag{14}$$

where $\mathbf{q}_{t_{j+1}} = \mathbf{0}$ for $j = N_{\mathrm{fft}}, \ldots, 2N_{\mathrm{fft}} - 1$. Since only one FFT needs to be computed and stored, regardless of the number of tapers, a significant computational advantage is attained. Thereafter the windowed Fourier realizations, $\hat{\mathbf{q}}_{f_k}^{(n)}$, can be recovered using Eq. (13) by sampling from $\hat{\mathbf{q}}_{f_k}$ in Eq. (14). In [17], memory considerations motivated a hybrid multitaper-Welch SPOD, with a user-specified $N_{\mathrm{fft}} < N_t$ and $N_{\mathrm{blk}} > 1$. The central difference identity in Eq. (13) largely neutralizes such concerns, as the scaling experiments in Appendix A attest. For this work, we set $N_{\mathrm{fft}} = N_t$, both maximizing frequency resolution and simplifying user inputs. We will revisit this decision in Sects. 3.1 and 5.

---

[1] In our experience the quality of this approximation only becomes relevant for values of N_{fft} that are too low to be suitable for multitaper SPOD; for all practical purposes, it is appropriate to work with $\Delta k = 1/2$.

2.3 Lossless compression

For large data, even with the techniques of Sects. 2.1 and 2.2, sine-taper SPOD can be prohibitively slow. To accelerate the SPOD, we propose to losslessly compress the weighted data matrix, assumed to be full-rank, using its QR factorization

$$\mathbf{X} = \mathbf{W}^{\frac{1}{2}}\mathbf{Q} = \mathbf{U}\widetilde{\mathbf{X}}, \tag{15}$$

where $\mathbf{U} \in \mathbb{C}^{N_{\text{dof}} \times N_t}$ is an orthonormal basis for the column space of $\mathbf{X}$, and $\widetilde{\mathbf{X}} \in \mathbb{C}^{N_t \times N_t}$ is the compressed data matrix. For data whose $N_{\text{dof}} \gg N_t$, it can be more efficient to orthogonalize $\mathbf{X}$ using the method-of-snapshots via the EVP

$$\mathbf{X}^*\mathbf{X}\mathbf{V} = \mathbf{V}\mathbf{L}, \tag{16}$$

where $\mathbf{V} \in \mathbb{C}^{N_t \times N_t}$ contains the eigenvectors, and $\mathbf{L} \in \mathbb{R}^{N_t \times N_t}$ contains the eigenvalues on its diagonal. The orthonormal basis and compressed data matrix are recovered as

$$\mathbf{U} = \mathbf{X}\mathbf{V}\mathbf{L}^{-\frac{1}{2}} \quad \text{and} \quad \widetilde{\mathbf{X}} = \mathbf{L}^{\frac{1}{2}}\mathbf{V}^*, \tag{17}$$

respectively. In practice, it is advantageous to replace the QR in Eq. (15) with the method-of-snapshots in Eqs. (16) and (17) above a threshold of $N_{\text{dof}} \gtrsim 8N_t$. As an added benefit, compression of the snapshot matrix further reduces the memory consumption of the FFT in Eq. (14). It can also be deduced from the dimensions of $\widetilde{\mathbf{X}}$ in Eqs. (15) and (17) that compression renders the overall wall time of the SPOD largely independent of N_{dof}; this is exemplified in Appendix A.

The remainder of sine-taper SPOD proceeds almost unchanged from the steps in Sects. 2.1 and 2.2. We assemble the compressed Fourier realizations, $\hat{\widetilde{\mathbf{X}}}_{f_k}$, at frequency f_k according to Eqs. (14), (13), and (12), then solve the $N_{\text{win}} \times N_{\text{win}}$ EVP

$$\hat{\widetilde{\mathbf{X}}}_{f_k}^* \hat{\widetilde{\mathbf{X}}}_{f_k} \mathbf{\Psi}_{f_k} = \mathbf{\Psi}_{f_k} \mathbf{\Lambda}_{f_k}. \tag{18}$$

Since $\mathbf{U}$ is a complete orthonormal basis,

$$\hat{\mathbf{X}}_{f_k}^* \hat{\mathbf{X}}_{f_k} = \hat{\widetilde{\mathbf{X}}}_{f_k}^* \hat{\widetilde{\mathbf{X}}}_{f_k}, \tag{19}$$

i.e., the compression is lossless. The uncompressed EVP in Eq. (5) and its compressed counterpart in Eq. (18) thus yield identical SPOD modal energy and expansion coefficients, $\mathbf{\Lambda}_{f_k}$ and $\mathbf{\Psi}_{f_k}$, respectively. The SPOD modes are recovered from the compressed data as

$$\mathbf{\Phi}_{f_k} = \mathbf{W}^{-\frac{1}{2}}\mathbf{U}\widetilde{\mathbf{\Theta}}_{f_k}, \quad \text{with} \quad \widetilde{\mathbf{\Theta}}_{f_k} = \hat{\widetilde{\mathbf{X}}}_{f_k} \mathbf{\Psi}_{f_k} \mathbf{\Lambda}_{f_k}^{-\frac{1}{2}}. \tag{20}$$

If the computational effort remains excessive, the SVD version of the lossless compression above may readily be used for lossy compression through a rank-d truncation, with $d < N_t$. The optimality of the truncation is guaranteed by the properties of SVD. We accomplish this by retaining the first d rows of $\widetilde{\mathbf{X}}$ in Eq. (17). Appendix B offers an example in which compression is achieved by retaining the first $d = 1000$ rows, and that shows an 87% reduction in computation time, while the spectrum and modes remain visually indistinguishable from the losslessly compressed results.

Though we have proposed the application of spatial compression in the context of sine-taper SPOD, compression is agnostic to the choice of spectral estimator, and could be applied to DPSS-taper SPOD, as well as to standard Welch SPOD.

2.4 Adaptivity

Two distinctive features of sine tapers enable the bandwidth to vary with frequency [14]. They are: (i) the bandwidth depends solely on N_{win}; and (ii) an arbitrary number of Fourier realizations can be sampled from one precomputed FFT. For any frequency f_k, the bandwidth can be locally broadened or narrowed by raising or lowering the number of tapers, N_{win, f_k}. In principle, N_{win, f_k} can vary arbitrarily to locally achieve the desired balance between variance and bias. In practice, however, bin-to-bin discontinuities in N_{win, f_k} may lead to undesired artifacts. To mitigate such artifacts, a best practice proposed by [1] limits the bin-to-bin difference in N_{win, f_k} to one, that is,

$$\left| N_{\text{win}, f_k} - N_{\text{win}, f_{k\pm1}} \right| \leq 1. \tag{21}$$

Despite this constraint, for temporally finely-resolved data and $N_{\text{fft}} = N_t$, the resulting high frequency resolution of the spectrum permits N_{win, f_k} to vary rapidly over a small frequency interval—a desirable trait.

To determine N_{win, f_k}, we exploit the fact that SPOD modes must be well-converged in order to be utilized successfully for physical discovery and ROM. We therefore propose to use the convergence of the leading SPOD mode, $\boldsymbol{\phi}_{1, f_k}$, between iterations to gauge the N_{win, f_k} required at each frequency. At the ith iteration, we compute the sine-taper SPOD with $N_{\text{win}}^{(i)}$ tapers, then evaluate the similarity [21] between the leading modes at the ith and $(i-1)$th iterations:

$$\alpha_{f_k}^{(i)} = \left| \left(\boldsymbol{\phi}_{1, f_k}^{(i)} \right)^* \mathbf{W} \boldsymbol{\phi}_{1, f_k}^{(i-1)} \right|. \tag{22}$$

The weight, $\mathbf{W}$, is the same as that in the inner product of Eq. 4, thus ensuring the similarity measure is consistent with the energy norm of the SPOD. Since SPOD modes are normalized to unit energy, by the Cauchy-Schwarz inequality, $\alpha_{f_k}^{(i)} \leq 1$. This implies that $\alpha_{f_k}^{(i)} \to 1$ as $\boldsymbol{\phi}_{1, f_k}^{(i)} \to \boldsymbol{\phi}_{1, f_k}^{(i-1)}$. Using Eq. (20), the similarity can be computed entirely in the compressed subspace as

$$\alpha_{f_k}^{(i)} = \left| \left(\widetilde{\boldsymbol{\theta}}_{1, f_k}^{(i)} \right)^* \widetilde{\boldsymbol{\theta}}_{1, f_k}^{(i-1)} \right|. \tag{23}$$

We consider $\boldsymbol{\phi}_{1, f_k}^{(i)}$, or equivalently $\widetilde{\boldsymbol{\theta}}_{1, f_k}^{(i)}$, converged if $\alpha_{f_k}^{(i)}$ satisfies

$$1 - \alpha_{f_k}^{(i)} \leq tol, \tag{24}$$

where tol is the similarity tolerance.

One of our goals is to protect tones from local bias, which requires N_{win, f_k} to be minimized. We achieve this by initializing the iterations with a small N_{win}, then incrementally increasing it. In each subsequent iteration, we fix the taper number for all frequencies that either meet the convergence criterion (24), or cannot accept more tapers due to constraint (21). The frequencies that converge most rapidly are thus windowed using small numbers of tapers, with narrow bandwidths. Such frequencies are associated with tones or dominant hydrodynamic instabilities characterized by low-rankness [20], i.e., a large separation between the leading and suboptimal SPOD eigenvalues. Where the spectrum is broadband, the prevalent flow structures are expected to vary smoothly as a function of frequency. We leverage this for statistical convergence by widening the bandwidth in regions of broadband turbulence.

Our proposed adaptive SPOD algorithm is described in more detail in Algorithm 1, in which the relevant equation numbers are given in parentheses. For simplicity, Algorithm 1 neglects the lossless compression we introduced in Sect. 2.3, but can be readily modified to include it. Figure 3 depicts an example of the taper numbers iteratively generated during the adaptive SPOD of the cavity data. The algorithm is initialized at iteration $i = 1$ with three tapers for all frequencies. Each subsequent iteration adds one taper to all or part of the spectrum. As a result, the taper numbers and iteration counts are offset by two. Each of the gray contours represents the tapers accumulated during the preceding 50 iterations. Where N_{win} plateaus, it has yet to be fixed, and the corresponding leading SPOD mode is not converged. For comparison, the PSD, which following Eq. (9) is the sum of all eigenvalues at each frequency, is superimposed. The general trend displayed by the contours is that N_{win} becomes fixed faster at lower frequencies than at higher frequencies. This trend is a consequence of low-rankness, which leads to rapid modal convergence because it indicates specific physical mechanisms are dominant. In the cavity flow, the low frequencies exhibit low-rankness, whereas the high frequencies consist mainly of broadband turbulence. The former thus converge more quickly. This distinction becomes particularly clear at the second, third, and fourth Rossiter frequencies, denoted by f_{R_2}, f_{R_3}, and f_{R_4} respectively, and marked by dotted cyan lines. These discrete tones in the PSD coincide with prominent dips in N_{win}, and are the first frequencies to achieve convergence. We will examine this phenomenon in more detail in Sect. 3.

3 Application to mixed tonal and broadband flows: example of an open cavity

We first demonstrate sine-taper SPOD on the turbulent open cavity flow from [31]. The cavity flow has a Mach number of $M = u_\infty / c_\infty = 0.6$, where u_∞ and c_∞ are the streamwise velocity and speed of sound in the freestream. TR-PIV was used to capture the streamwise and wall-normal velocities, u and v, at a time interval of $\Delta t = 6.25 \times 10^{-5}$ s, for a total of $N_t = 16{,}000$ snapshots. The velocities have been non-dimensionalized

---

Algorithm 1: Adaptive SPOD

Input: $\mathbf{Q}$, $\mathbf{W}$, and tol.
Output: $\mathbf{\Lambda}_{f_k}$, $\mathbf{\Phi}_{f_k}$, and N_{win, f_k} for $k = 0, \ldots, N_t - 1$.
Obtain the unwindowed Fourier modes, $\hat{\mathbf{q}}_{f_k}$ (14).

Initialize $i \leftarrow 0$ and $N_{\text{win}}^{(i)} \leftarrow 2$.
while *not all* N_{win, f_k} *fixed* **do**
    $i \leftarrow i + 1$.
    $N_{\text{win}}^{(i)} \leftarrow N_{\text{win}}^{(i-1)} + 1$.
    for *all k for which* N_{win, f_k} *not fixed* **do**
        Assemble $\hat{\mathbf{X}}_{f_k}^{(i)}$ for $N_{\text{win}}^{(i)}$ tapers from $\hat{\mathbf{q}}_{f_k}$ (13), and solve $(\hat{\mathbf{X}}_{f_k}^{(i)})^* \hat{\mathbf{X}}_{f_k}^{(i)} \mathbf{\Psi}_{f_k}^{(i)} = \mathbf{\Psi}_{f_k}^{(i)} \mathbf{\Lambda}_{f_k}^{(i)}$ for $\mathbf{\Lambda}_{f_k}^{(i)}$ and $\mathbf{\Psi}_{f_k}^{(i)}$ (5).   `% EVP`
        $\mathbf{\Phi}_{f_k}^{(i)} \leftarrow \mathbf{W}^{-\frac{1}{2}} \hat{\mathbf{X}}_{f_k}^{(i)} \mathbf{\Psi}_{f_k}^{(i)} (\mathbf{\Lambda}_{f_k}^{(i)})^{-\frac{1}{2}}$ (6).   `% SPOD modes`
        if $N_{\text{win}}^{(i)} > 3$ **then**
            $\alpha_{f_k}^{(i)} \leftarrow \left| (\boldsymbol{\phi}_{1, f_k}^{(i)})^* \mathbf{W} \boldsymbol{\phi}_{1, f_k}^{(i-1)} \right|$ (22).   `% similarity`
            if $1 - \alpha_{f_k}^{(i)} \le tol$ **then**   `% converged`
                Fix $\mathbf{\Lambda}_{f_k} \leftarrow \mathbf{\Lambda}_{f_k}^{(i)}$, $\mathbf{\Phi}_{f_k} \leftarrow \mathbf{\Phi}_{f_k}^{(i)}$, and $N_{\text{win}, f_k} \leftarrow N_{\text{win}}^{(i)}$.
            else   `% not converged`
                if $N_{\text{win}, f_{k-1}}$ *is fixed* **and** $N_{\text{win}}^{(i)} - N_{\text{win}, f_{k-1}} > 0$ **then**
                    `% constrained by left neighboring frequency`
                    Fix $\mathbf{\Lambda}_{f_k} \leftarrow \mathbf{\Lambda}_{f_k}^{(i)}$, $\mathbf{\Phi}_{f_k} \leftarrow \mathbf{\Phi}_{f_k}^{(i)}$, and $N_{\text{win}, f_k} \leftarrow N_{\text{win}}^{(i)}$.
                else if $N_{\text{win}, f_{k+1}}$ *is fixed* **and** $N_{\text{win}}^{(i)} - N_{\text{win}, f_{k+1}} > 0$ **then**
                    `% constrained by right neighboring frequency`
                    Fix $\mathbf{\Lambda}_{f_k} \leftarrow \mathbf{\Lambda}_{f_k}^{(i)}$, $\mathbf{\Phi}_{f_k} \leftarrow \mathbf{\Phi}_{f_k}^{(i)}$, and $N_{\text{win}, f_k} \leftarrow N_{\text{win}}^{(i)}$.
                end
            end
        end
    end
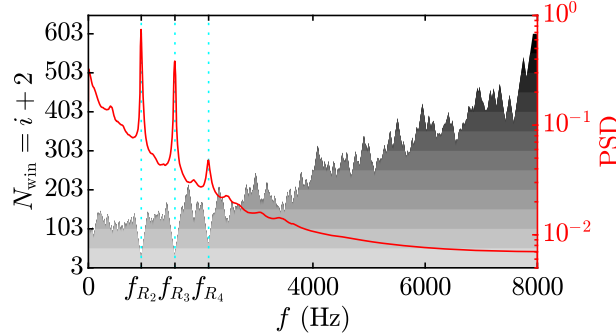**end**

Fig. 3 The taper numbers, N_{win}, of the cavity flow generated iteratively by the adaptive SPOD algorithm, shown for every 50 iterations starting from $i = 1$. Darker contours denote higher iteration counts. The volume-integrated PSD, in red, is superimposed. The second, third, and fourth Rossiter frequencies are marked by dotted cyan lines (Colour figure online)

by u_∞, and lengths by the cavity depth, D_c. For consistency and ease of comparison with [31], frequencies for the cavity example are reported in units of Hz. The field of view is resolved by $N_x \times N_y = 156 \times 55$ points. Prior to carrying out the SPOD, the velocity components are arranged in the state vector $\mathbf{q} = [\mathbf{u}, \mathbf{v}]^{\text{T}}$. The weight matrix is the identity, that is, $\mathbf{W} = \mathbf{I}$. Since $\mathbf{q}$ contains the two Cartesian velocity components and the grid is uniform, the norm corresponds to the incompressible turbulent kinetic energy. These parameters are summarized in Table 1.

Instantaneous streamwise velocity fluctuations, u', of the cavity are visualized in Fig. 2a in the introduction. The temporal mean field is shown in Fig. 2c. The fluctuations of the cavity display structures that span a wide range of spatial scales, which is characteristic of broadband turbulent flows. It is well-known that open cavity flows also contain tones associated with the Rossiter mechanism (see [15] for a review). Since we designed

the adaptive version of sine-taper SPOD specifically for spectra with mixed tonal and broadband content, in the following we will focus primarily on the cavity data. For non-tonal flows, on the other hand, adaptivity is not necessary. For the broadband jet data in Sect. 4 we will thus demonstrate the non-adaptive version of sine-taper SPOD.

3.1 Preamble: non-adaptive sine-taper SPOD with constant N_{win}

We begin with a comparison between the state-of-the-art DPSS-taper SPOD and our proposed sine-taper SPOD using a frequency-independent N_{win} in the left and right columns of Fig. 4 respectively. In the standard implementation of DPSS-taper SPOD, Fourier realizations receive uniform weights, $\mu = 1/N_{\text{win}}$. The nonuniform weights in Eq. 11 are used for sine tapers. The data are not segmented, i.e., $N_{\text{fft}} = N_t = 16{,}000$. This demands a potentially large N_{win} for convergence. We reiterate that the need for segmenting is eliminated by the computational savings brought about by sine tapers. As Appendix A exemplifies, the memory consumption of sine-taper SPOD is independent of N_{win}, whereas that of the DPSS version grows linearly with it. This makes DPSS-taper SPOD without segmenting inconvenient for the small cavity data, but prohibitively expensive for the large jet data in Sect. 4.

The eigenvalue spectra in Fig. 4a, b, i, and j fade from black to white with increasing mode number. Each row shares the same N_{win}, giving both taper families approximately equal bandwidths. In Fig. 4a–h,
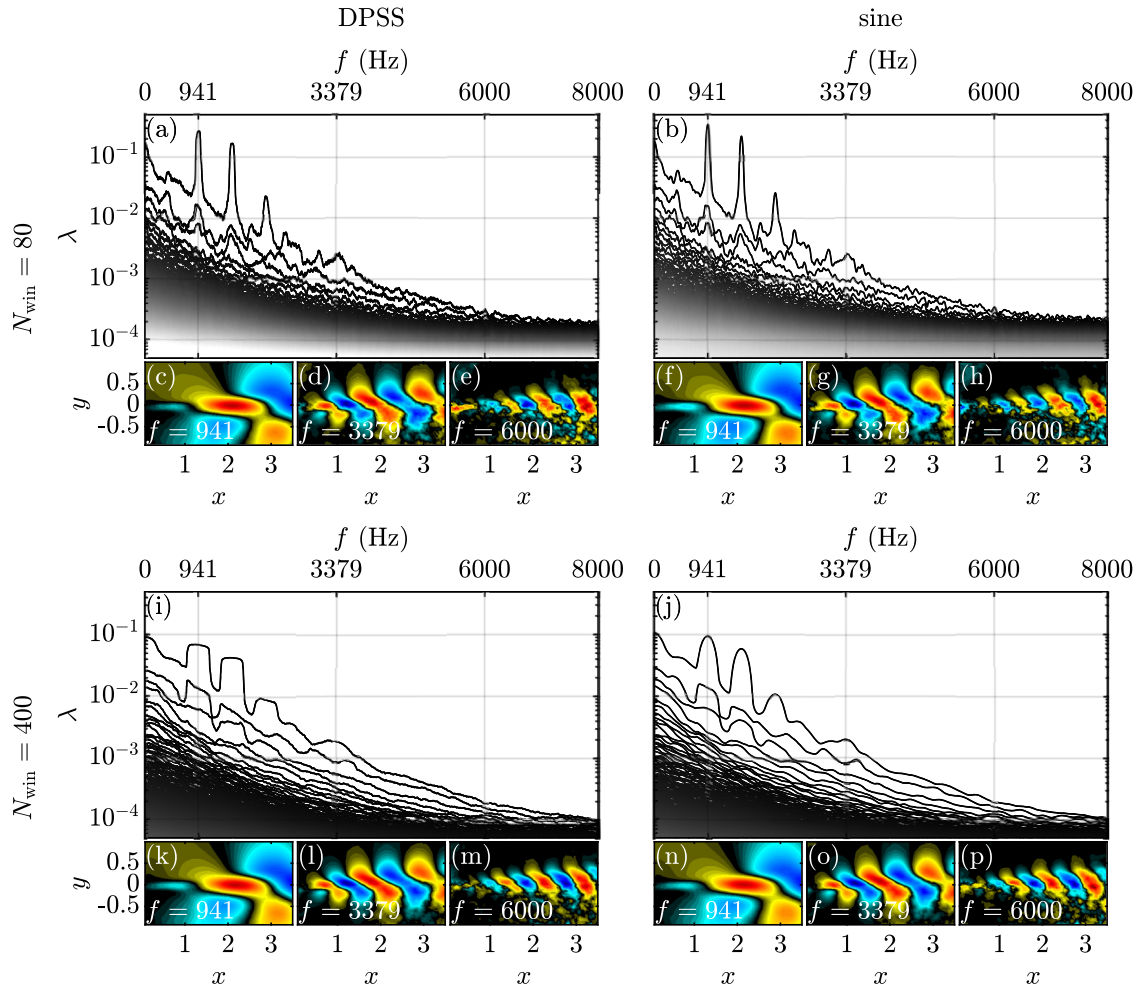

Fig. 4 SPOD eigenvalue spectra and leading modes of the cavity, computed with: **a**, **c**, **d**, **e**, **i**, **k**, **l**, **m** DPSS; **b**, **f**, **g**, **h**, **j**, **n**, **o**, **p** sine tapers. 80 (**a–h**) and 400 (**i–p**) tapers are used; for DPSS, these correspond to setting $\delta f\, N_{\text{fft}} = 40.5$ and 200.5, respectively. The real part of the streamwise velocity component of each mode, $\left|\boldsymbol{\phi}_u\right|$, is shown. The colors saturate at $\left|\boldsymbol{\phi}_u\right| / \max \left|\boldsymbol{\phi}_u\right| = \pm 1$

$N_{\text{win}} = 80$. Owing to their matching bandwidths, the eigenvalues in Fig. 4a and b display similar spectral shapes. In particular, the discrete peaks corresponding to Rossiter tones are precisely localized in frequency, and are visible in at least the three leading eigenvalues. The two spectra possess similar energies in the broadband region, $f \gtrsim 4000$ Hz. However, the sine-tapered spectrum in Fig. 4b is distinguished by sharper peaks and higher energies at the tonal frequencies, e.g., $f = 941$ Hz, which indicates lower local bias. Bin-to-bin variance, which can obscure important features, especially in shorter data, affects both spectra but is more severe in the DPSS-tapered spectrum. Below each spectrum, representative modes at three frequencies, $f = 941, 3379$, and 6000 Hz, are shown. Each mode, or eigenvector, is determined only up to an arbitrary phase. Without loss of generality, we approximately align the phases of the modes at the same frequency to faciliate comparison. At each frequency, DPSS- and sine-tapered modes are visually nearly indistinguishable. The modes at the low frequency, $f = 941$ Hz, a Rossiter tone, have very smooth waveforms that indicate they are fully converged. This is expected since the underlying Rossiter instability ensures the first and second eigenvalues, λ_1 and λ_2, are separated by a large gap in energy.[2] For DPSS, $\lambda_1 - \lambda_2 \approx 0.25$; for sine, $\lambda_1 - \lambda_2 \approx 0.32$. Such low-rank behavior suggests that the structures are prevalent and deterministic, therefore rapidly converged. The leading eigenvalue at the intermediate frequency, $f = 3379$ Hz, also reveals a peak. However, at this frequency, $\lambda_1 - \lambda_2 \approx 0.0017$ and 0.0016 for DPSS and sine tapers, respectively. Consistent with the smaller eigenvalue separation, the modes are less well-converged, though they continue to display the clear streamwise and wall-normal wavenumbers that are characteristic of Kelvin-Helmholtz (KH) wavepackets. At the high frequency, $f = 6000$ Hz, the two leading eigenvalues possess similar energies. The mode at this frequency exemplifies the slow rate of convergence in the broadband region. The KH wavepackets, while visible, appear distorted and lack a clearly delineated envelope. Overall, these results are consistent with the spectrum and modes obtained by [31] using the Welch method.

If N_{win} is increased to 400, the leading modes at the Rossiter frequency, $f = 941$ Hz, in Fig. 4k and n remain visually unchanged, since they are already converged with $N_{\text{win}} = 80$. The modes at the intermediate and high frequencies, $f = 3379$ and 6000 Hz, on the other hand, have become smoother and more compact, though their convergence may be further improved with an even greater N_{win}. However, raising N_{win} further would be inadvisable, as the obvious and undesired broadening of the tones in the spectra in Fig. 4i and j demonstrate. As expected, with $N_{\text{win}} = 400$ the leading eigenvalues of both spectra have lower energy at all frequencies in comparison with $N_{\text{win}} = 80$. This phenomenon is related to the energy being distributed over a much higher number of modes. The sine-tapered spectrum in Fig. 4j continues to preserve more energy at the tonal frequencies than the DPSS-tapered spectrum does in Fig. 4i. In both spectra, however, the peaks are drastically flattened as the larger N_{win} effectively averages over a larger bandwidth. Less energetic, closely-spaced peaks are merged. At the peaks, the DPSS spectrum takes on a characteristic boxy shape [13], whereas the sine spectrum retains higher curvature. This means that despite the excessive broadening, the frequencies of the most dominant Rossiter tones can still be accurately estimated from the sine spectrum. In the DPSS spectrum, the tonal frequencies are much less distinct, and have to be estimated as the midpoint of each of the flattened peaks. From these observations we draw two conclusions: (i) DPSS- and sine-taper SPOD demonstrate similar variance reduction, with a slight advantage to sine tapers; however, the latter excels at resolving tones; and (ii) in the broadband region a higher N_{win} is desirable because it leads to lower variance and more converged modes; at tonal frequencies, however, it is not, since well-converged modes and accurately-resolved peaks are obtained with a low N_{win}. These conclusions motivate the use of the adaptive algorithm for mixed tonal and broadband flows.

Next we systematically and quantitatively investigate the effect of N_{win} on tonal frequency estimation. The contours in Fig. 5 show the evolution of the leading SPOD eigenvalues as N_{win} is increased from five to 400 for DPSS (5a) and sine (5b) tapers. Darker shades of gray represent larger eigenvalues. In both spectra, thin striations in the region $N_{\text{win}} \lesssim 50$ indicate high bin-to-bin variance. For higher N_{win}, the striations are smoothed and replaced by bands that gradually widen with bandwidth. Dark bands centered at f_{R_2}, f_{R_3}, and f_{R_4} correspond to the Rossiter frequencies. Accurately locating these frequencies becomes more challenging as the bands are broadened, especially with DPSS tapers. As Fig. 4i showed, when N_{win} is high, DPSS tapers flatten peaks, making the peak frequencies ambiguous. In Fig. 5, for each N_{win} we overlay the peak frequency within each dark band,

---

[2] A large separation between λ_1 and λ_2 is sufficient but not necessary for rapid modal convergence. A counterexample is a flow in which two different physical mechanisms are active at the same frequency, leading to a large separation between λ_2 and λ_3 instead. This may occur if, e.g., statistical symmetries are present. An example can be found in [30].
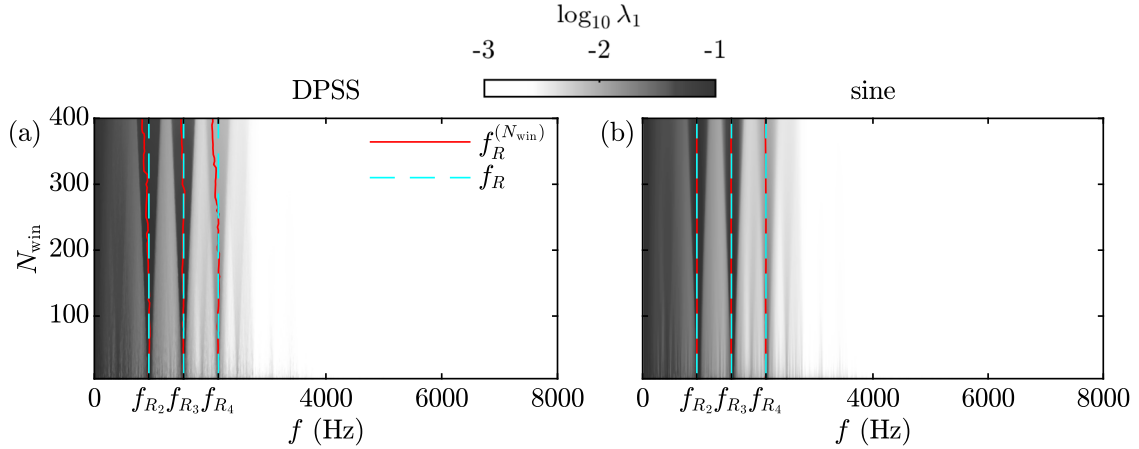
Fig. 5 Leading SPOD eigenvalues of the cavity computed with: **a** DPSS; **b** sine tapers. The solid red lines correspond to the peak frequencies of the three most prominent Rossiter tones for each N_{win}. For comparison, the dashed cyan lines mark the actual tonal frequencies, f_{R_2}, f_{R_3}, and f_{R_4}, previously determined in Fig. 3. **a**, **b** share the same contour levels (Colour figure online)
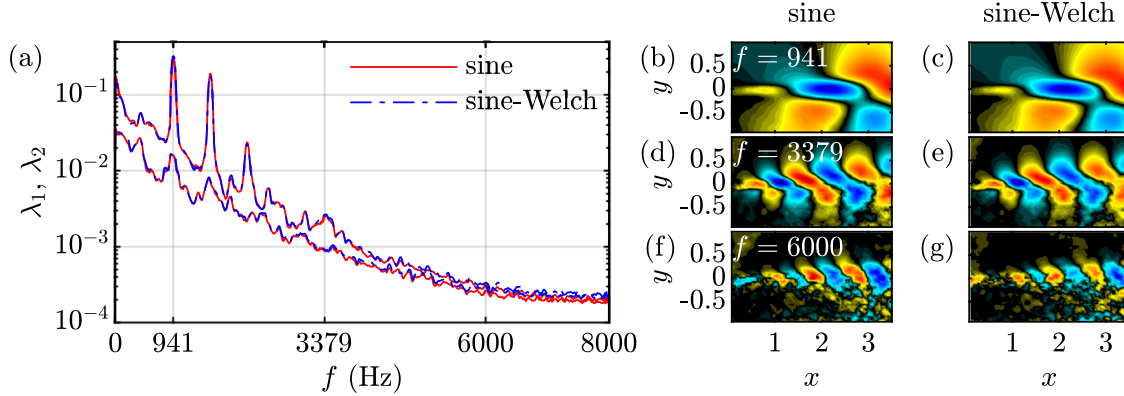

Fig. 6 First two SPOD eigenvalues, λ_1 and λ_2, and leading modes of the cavity at $f = 941$ (**b**, **c**), 3379 (**d**, **e**), and 6000 (**f**, **g**) Hz. The solid red spectrum in (**a**) and modes in (**b**, **d**, **f**) use multitaper estimation, with a block size of $N_{\text{fft}} = N_t$ and 97 sine tapers. The dash-dotted blue spectrum in (**a**) and modes in (**c**, **e**, **g**) use multitaper-Welch estimation, with a block size of $N_{\text{fft}} = 4096$, 50% overlap, and 24 sine tapers

$$f_R^{(N_{\text{win}})} = \arg\max_{|f_k - f_R| < \frac{\delta f}{\Delta t}} \lambda_{1, f_k}^{(N_{\text{win}})}, \tag{25}$$

as solid red lines. The dashed cyan lines mark the actual Rossiter frequencies determined in Fig. 3. In Fig. 5a, for $N_{\text{win}} \lesssim 200$, the peak frequencies $f_R^{(N_{\text{win}})}$ of the DPSS spectrum initially stay close to—but meander around—the reference frequencies f_R. The meandering is a consequence of high bin-to-bin variance. For $N_{\text{win}} \gtrsim 200$, $f_R^{(N_{\text{win}})}$ becomes biased towards low frequencies. The leftward bias grows with N_{win}, as a larger bandwidth increasingly merges the discrete tones with the underlying broadband spectrum, which decays with frequency. Although the meandering and frequency bias are both observed in the sine spectrum in Fig. 5b, they are significantly subtler here than they are in Fig. 5a. The tones in Fig. 5b are tracked smoothly and more accurately for all N_{win}, achieving variance reduction without compromising tonal frequency estimation. These observations apply to peaks that are distant from each other in frequency, but are likely to break down otherwise. If neighboring peaks merge, as some have in Fig. 4j for $N_{\text{win}} = 400$, their frequencies will not be resolved even by sine tapers. Nonetheless, the favorable performance of sine-taper SPOD demonstrates that minimizing local bias should be prioritized over maximizing spectral concentration when the spectrum comprises both discrete and broadband components.

When memory requirements become prohibitive for long data records, [17] suggested using a hybrid multitaper-Welch SPOD, combining segmenting with multitaper SPOD. As discussed in Sect. 2.2, sine-taper SPOD with compression diminishes the need for segmenting. Figure 6 compares the sine-taper and sine-Welch

approaches. For sine-taper SPOD, the full-length data are windowed using 97 sine tapers. For sine-Welch SPOD, the snapshots are segmented into six blocks of length $N_{fft} = 4096$ each, with 50% overlap between adjacent blocks. Each block is then windowed using 24 sine tapers. These sets of parameters result in similar bandwidths, δf, which in turn produce spectral estimates with comparable variance and bias. The leading two eigenvalues, λ_1 and λ_2, in Fig. 6a and leading modes in Fig. 6b–g produced by the sine-taper and sine-Welch methods match each other closely. The consistency between the two approaches shows that either allows for good variance and bias control, provided that spectral estimation parameters are chosen judiciously. For the sine-Welch estimator, N_{fft} and overlap must be selected in addition to N_{win}; for the sine estimator, only N_{win} needs to be selected. Its maximal frequency resolution and small number of parameters make sine-taper SPOD without segmenting advantageous in most cases, in particular for the proposed adaptive algorithm.

3.2 Adaptive sine-taper SPOD

As previously seen in Fig. 3, while sine-taper SPOD improves upon the existing DPSS-taper algorithm, for the cavity flow it remains unable to provide satisfactory spectral and modal estimates at all frequencies using a single N_{win}. Recall that in the right column of Fig. 4, $N_{win} = 80$ generates a spectrum with precisely localized tones but high bin-to-bin variance, and poorly converged modes at high frequencies. A fivefold increase in N_{win} reduces variance and enhances convergence at the expense of unacceptable local bias. Adaptive sine-taper SPOD, introduced in Sect. 2.4, grants greater control of the variance-bias trade-off by allowing N_{win} to vary with frequency. It determines a frequency-local N_{win, f_k} for each frequency based on the rate of convergence of the leading mode. The contours in Fig. 7 display the similarity, $\alpha_{f_k}^{(N_{win})}$, defined in Eq. (22), between the leading modes obtained using N_{win} and $N_{win} - 1$ tapers. As each mode converges, $\alpha \to 1$ and $\log_{10}(1 - \alpha) \to -\infty$, producing darker contours. Overall the contours darken with increasing N_{win}, confirming that convergence improves with the addition of tapers. Meanwhile, the contours lighten with increasing frequency in the region $f \gtrsim 3000$ Hz characterized by broadband turbulence. The darkest wedges, where α is closest to one, are centered around the Rossiter frequencies. The fastest converging frequencies are thus the ones which exhibit low-rank behavior in Fig. 4; conversely, convergence slows at high frequencies, where the leading and suboptimal eigenvalues cluster around each other. The bin-to-bin variance of the similarity, α, underscores the need to enforce constraint (21) on N_{win, f_k} explicitly.

To study the effects of the similarity tolerance, tol, Fig. 8 shows the adaptive SPOD eigenvalue spectra, leading modes, and taper numbers of the cavity flow for $tol = 10^{-5}, 10^{-6}$, and 10^{-7}. For each spectrum, at each frequency only the first $\min(N_{win, f_k})$ eigenvalues out of a total of N_{win, f_k} are displayed. For $tol = 10^{-5}$, the Rossiter tones are sharp and well-localized in frequency, and the bin-to-bin variance is high. The taper number fluctuates between approximately 10 and 200. The Rossiter tones coincide with local minima in N_{win, f_k}. In particular, at the second and third Rossiter frequencies, $f = 941$ and 1538 Hz, $N_{win, f_k} \approx 10$. Among the modes in Fig. 8c, f, i, and l, only the one at $f = 941$ Hz is smooth. The mode at $f = 6000$ Hz contains KH wavepackets that appear disjointed and distorted. When we tighten the tolerance to 10^{-6}, the bin-to-bin variance of the spectrum is greatly reduced at all frequencies. Unlike in Fig. 4j, however, here the tones remain sharp and individually resolved. This low local bias is enabled by large variations in N_{win, f_k}. While N_{win, f_k}
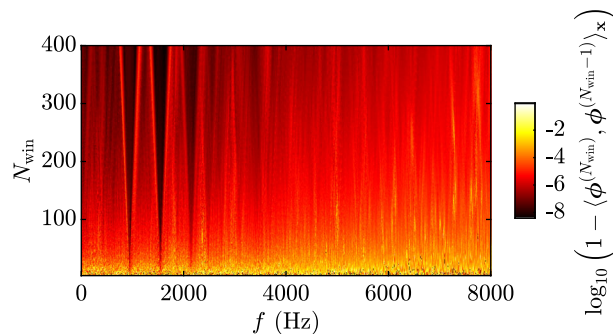

Fig. 7 Similarity between the leading SPOD mode of the cavity at each frequency estimated using N_{win} sine tapers and the corresponding mode estimated using $N_{win} - 1$ tapers, for $N_{fft} = N_t$. Dark red indicates high mode similarity (Colour figure online)
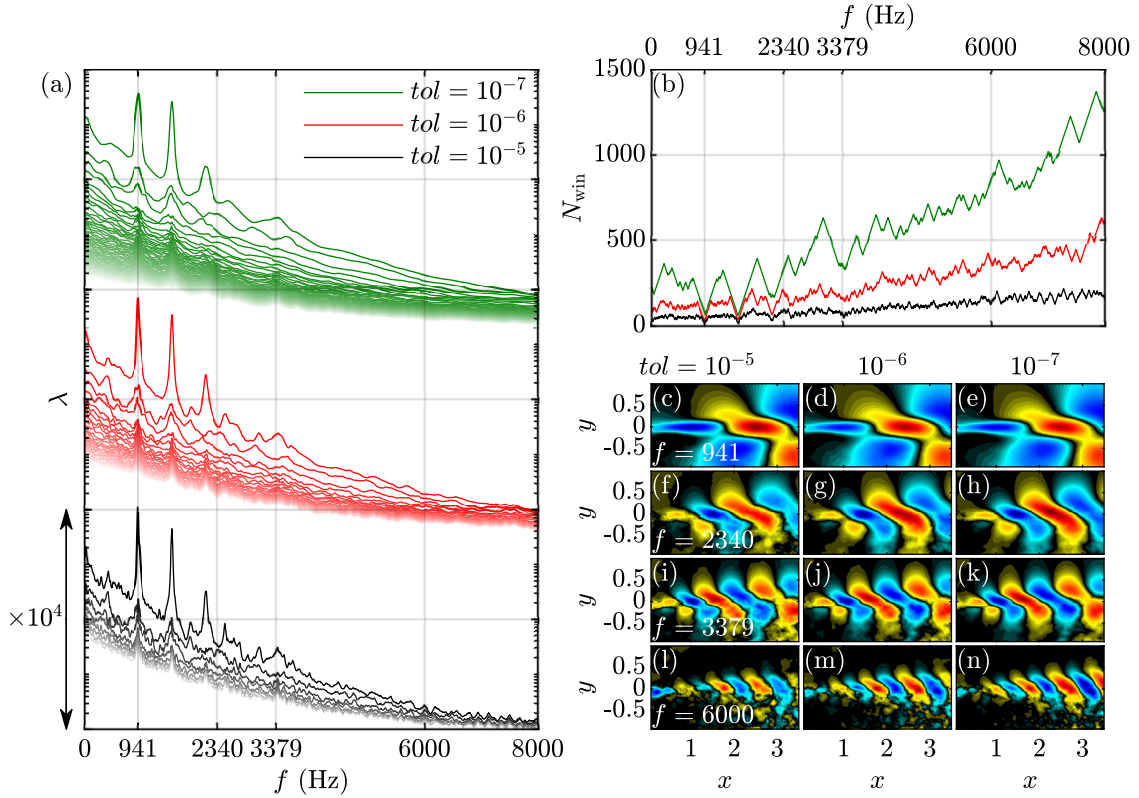
Fig. 8 Adaptive SPOD eigenvalue spectra of the cavity for different similarity tolerances (**a**), the corresponding frequency-local taper numbers (**b**), and the leading modes at $f = 941$ (**c–e**), 2340 (**f–h**), 3379 (**i–k**), and 6000 (**l–n**) Hz, for $tol = 10^{-5}$ (black), 10^{-6} (red), and 10^{-7} (green). For ease of comparison, the red and green spectra have been shifted. The first $\min(N_{\mathrm{win}, f_k})$ adaptive SPOD eigenvalues are shown in (**a**) (Colour figure online)

is generally much higher for $tol = 10^{-6}$ than for 10^{-5}, it dips to around 30 at the most prominent tones. The modes have also become smoother. This is most apparent at $f = 6000$ Hz, where the wavepackets acquire clearly-defined streamwise and wall-normal wavenumbers. Reducing the tolerance to 10^{-7} necessitates still higher N_{win, f_k}. The bin-to-bin variance of the spectrum is now largely eliminated at the expense of more severe leakage. Some of the closely-spaced peaks have merged with their neighbors, though the Rossiter tones remain sharp and their broadening is limited. The modes again show smoother waveforms, free of distortion or artifacts.

Of the three similarity tolerances, we deem $tol = 10^{-6}$ to result in both low bias and low variance. From Fig. 4 we concluded that with a frequency-independent N_{win}, the variance-bias trade-off cannot be adequately controlled in all regions of a tonal spectrum. Compared to non-adaptive SPOD, the adaptive estimator combines the low bias of a small N_{win} to resolve discrete tones with the low variance of a large N_{win} to estimate the broadband portions of the spectrum. The tones in the $tol = 10^{-6}$ spectrum in Fig. 8a experience a similar degree of broadening to those in the $N_{\mathrm{win}} = 80$ spectrum in Fig. 4b, though the former attain higher energies. The bin-to-bin variance of the $tol = 10^{-6}$ spectrum, on the other hand, is nearer that of the $N_{\mathrm{win}} = 400$ spectrum in Fig. 4j. The modes computed using $tol = 10^{-6}$ in Fig. 8d, j, and m resemble those using $N_{\mathrm{win}} = 400$ in Fig. 4n, o, and p respectively. In summary, with the right similarity tolerance, the proposed adaptive algorithm simultaneously achieves accurate resolution of tonal peaks and smooth estimation of broadband spectra.

Next we demonstrate that adaptive SPOD only requires a fraction of the data to produce estimates comparable to those of standard Welch SPOD. Figure 9a shows the Welch SPOD spectrum obtained by [31], the work in which the cavity data were originally published, using the full data, with parameters $N_{\mathrm{fft}} = 1280$, 75% overlap, and the Hamming window. In contrast, the adaptive SPOD spectrum shown in Fig. 9b uses only the first 4000 snapshots, i.e., 25% of the original data. The similarity tolerance is adjusted such that the tones in both spectra are approximately equally resolved. We have confirmed that the differences in magnitudes of the tones between the two spectra are associated with the reduced test data segment, not with the estimators. The
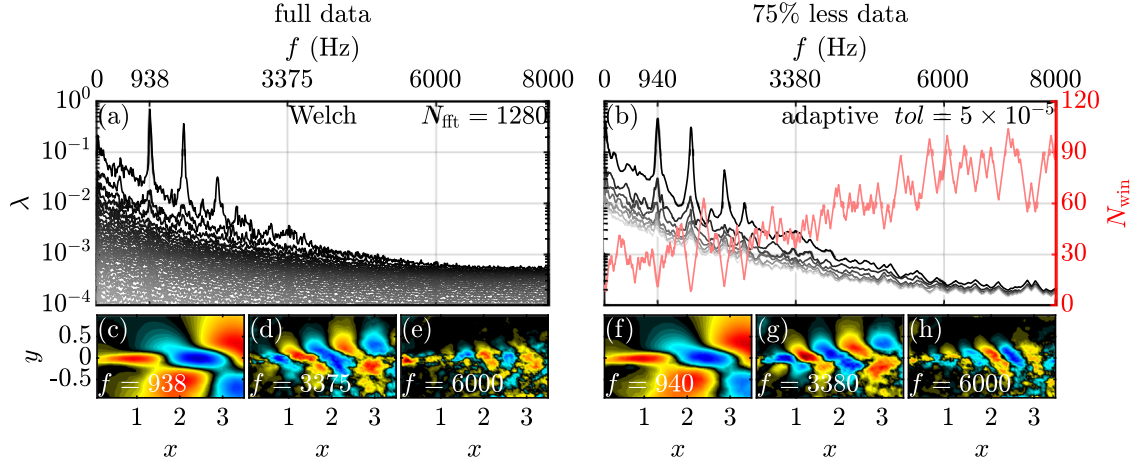
Fig. 9 SPOD eigenvalues and leading modes of the cavity computed from: **a**, **c–e** all 16,000 snapshots using the Welch method; **b**, **f–h** the first 4000 snapshots, or 75% less data, using adaptive sine-taper SPOD. The Welch SPOD uses $N_{\text{fft}} = 1280$ with 75% overlap and the Hamming window—the same parameters as in [31]. The adaptive SPOD uses $tol = 5 \times 10^{-5}$. For the adaptive case, N_{win} is overlaid

adaptive spectrum exhibits less bin-to-bin variance than the Welch spectrum does. The modes at $f \approx 940$ Hz in Fig. 9c and f are comparable, whereas those at $f \approx 3380$ and 6000 Hz in Fig. 9d, e, g, and h appear slightly smoother and more distinct with the adaptive estimator. As Fig. 8 shows, given more data adaptive SPOD will perform even better. Nevertheless, its efficient usage of scarce data is a major advantage over standard SPOD.

4 Application to broadband flows: example of a turbulent jet

Having demonstrated sine-taper SPOD on the open cavity flow, we next examine its application to the round jet flow from [3], first studied using standard Welch SPOD by [20]. The jet is an example of a broadband turbulent flow, for which we explicitly highlight that the adaptive algorithm is not recommended. Instead we recommend using a constant number of sine tapers to benefit from the increased frequency resolution. The jet flow has a Mach number of $M = u_j/c_j = 0.9$, where u_j and c_j are the centerline velocity and sound speed at the nozzle exit, respectively. Large-eddy simulation (LES) using the solver 'Charles' [2] produced $N_t = 10{,}000$ snapshots at an interval of $\Delta t = 0.18$. Each snapshot contains the density, ρ, axial, radial, and azimuthal velocities, u_x, u_r, and u_θ, and temperature, T. The density, velocities, and temperature have been non-dimensionalized by their values on the centerline at the nozzle exit, ρ_j, u_j, and T_j, respectively. Lengths are non-dimensionalized by the nozzle diameter, D_j, and time by D_j/u_j. Frequencies for the jet example are thus expressed as Strouhal numbers, $f D_j/u_j$. For brevity, we denote them by f, which should be understood to be dimensionless. The unstructured LES data are interpolated onto a nonuniform cylindrical grid of size $N_x \times N_r \times N_\theta = 656 \times 138 \times 128$. The rotational symmetry of the nozzle permits independent consideration of each azimuthal wavenumber component. Only the axisymmetric component, with wavenumber $m = 0$, will be examined in this work. The five variables are assembled in the state vector $\mathbf{q} = [\hat{\rho}, \hat{\mathbf{u}}_x, \hat{\mathbf{u}}_r, \hat{\mathbf{u}}_\theta, \hat{\mathbf{T}}]^{\text{T}}$. The weight matrix $\mathbf{W}$ discretizes the inner product

$$\langle \boldsymbol{q}_1, \boldsymbol{q}_2 \rangle_{\boldsymbol{x}} = \iiint_\Omega \boldsymbol{q}_2^* \text{diag}\left(\left[\frac{\bar{T}}{\gamma \bar{\rho} M^2}, \bar{\rho}, \bar{\rho}, \bar{\rho}, \frac{\bar{\rho}}{\gamma(\gamma-1)\bar{T} M^2}\right]\right) \boldsymbol{q}_1 r \mathrm{d}x \mathrm{d}r \mathrm{d}\theta, \tag{26}$$

where Ω denotes the domain and γ is the adiabatic constant. Thus the eigenvalues are the volume-integrated compressible energy [4] of each mode. The parameters are summarized in Table 1. We note that the higher spatial resolution and higher number of variables of the jet database make it an order of magnitude larger than the cavity database.

The broadband nature of the SPOD spectrum of the jet [20] favors a constant rather than adaptive N_{win}. In general, however, we may not be able to exploit a priori knowledge of a particular flow. Without such foresight, distinguishing a tonal spectrum from a noisy estimate of a broadband spectrum may not be a trivial task. Adaptive sine-taper SPOD, while not designed for broadband flows, nevertheless must be robust in such
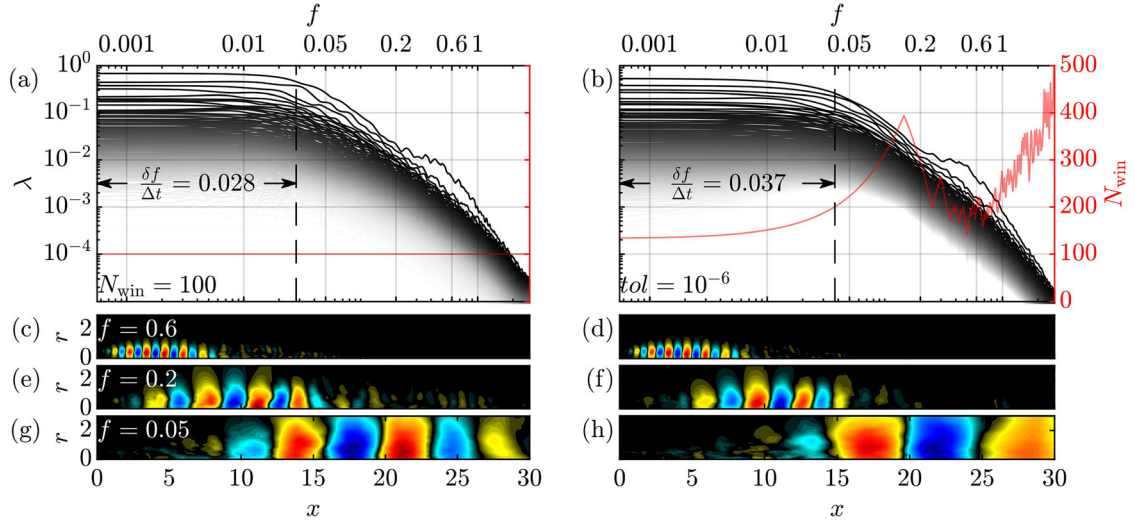
Fig. 10 SPOD eigenvalues (black) and N_{win} (red) of the jet flow computed using: **a** $N_{\mathrm{win}} = 100$; **b** $tol = 10^{-6}$, along with their respective leading modes at frequencies $f = 0.6$ (**c**, **d**), 0.2 (**e**, **f**), and 0.05 (**g**, **h**). In (**b**), only the first min N_{win, f_k} eigenvalues are shown. Dashed lines mark the half-bandwidth, $\delta f / \Delta t$, of the $f = 0$ bin in (**a**, **b**)

cases. For the jet, Fig. 10b shows the adaptive SPOD eigenvalues and taper numbers computed using $tol = 10^{-6}$ and $N_{\mathrm{fft}} = N_t = 10{,}000$. Compared to the spectrum produced by [20] using Welch SPOD (reproduced in Fig. 11a), the adaptive spectrum possesses a similar shape. Notably, the adaptivity in the high-frequency broadband regions does not generate numerical artifacts. Unlike the cavity flow in Fig. 8b, here N_{win, f_k} is uninterrupted by deep valleys that signify dramatic reductions in the local bandwidth. Consistent with the absence of tones, the taper number remains moderately high: $N_{\mathrm{win}, f_k} \gtrsim 100$ over the entire frequency range. N_{win, f_k} does experience a spike in the range $0.04 \lesssim f \lesssim 0.3$, which is associated with the KH mechanism becoming dominant for $f \gtrsim 0.2$ (see [20]). The slowing of mode convergence when the eigenvalue separation is small is thus common to both the jet and the cavity.

Below $f \approx 0.05$, the spectrum gradually flattens into a plateau. The half-width of this plateau is comparable to the half-bandwidths, $\delta f / \Delta t$, of the low frequencies. As an example, the half-bandwidth of the $f = 0$ bin, $\delta f / \Delta t = 0.037$, is indicated in Fig. 10. The shape of the plateau is due to the main lobe of the mean spectral window (i.e., mean PSD) of DPSS and sine tapers [14,26,28], which becomes nearly rectangular with width equal to the bandwidth as N_{win} grows. In effect, the plateau—and therefore the bandwidth—determines the lowest nonzero frequency resolvable by the spectrum.

In Fig. 10a, the spectrum windowed using $N_{\mathrm{win}} = 100$ tapers is shown for reference, and is consistent with the adaptive spectrum in Fig. 10b. The leading modes at $f = 0.6$ in Fig. 10c, d, as well as those at $f = 0.2$ in Fig. 10e, f, are also well-matched, though the adaptively tapered mode at $f = 0.2$ has a smoother, more compact waveform. At $f = 0.05$, however, the adaptive mode in Fig. 10h significantly underestimates the wavenumber, k, of the structures. This discrepancy occurs only for $0.04 \lesssim f \lesssim 0.2$, where N_{win, f_k} spikes, and suggests that in the low-frequency region averaging over a large bandwidth may bias modes towards lower wavenumbers, or equivalently lower frequencies. This phenomenon will be discussed in Sect. 5. We conclude that the $N_{\mathrm{win}} = 100$ estimate is better, and recommend non-adaptive sine-taper SPOD for broadband flows to take advantage of the high frequency resolution of the method. An N_{win} of 100 is a good starting point, but should be adjusted to balance variance and bias.

Analogous to Fig. 9, Fig. 11 compares the Welch SPOD obtained by [20] using the full jet data, and non-adaptive sine-taper SPOD with 75% of the data discarded. For the latter, 100 tapers are used. The sine-taper spectrum displays lower variance than the Welch spectrum does. Both algorithms produce modes of comparable smoothness. We conclude that the non-adaptive sine-taper algorithm also requires a fraction of the snapshots that the standard Welch algorithm demands for broadband flows, just as the adaptive version does for tonal flows.
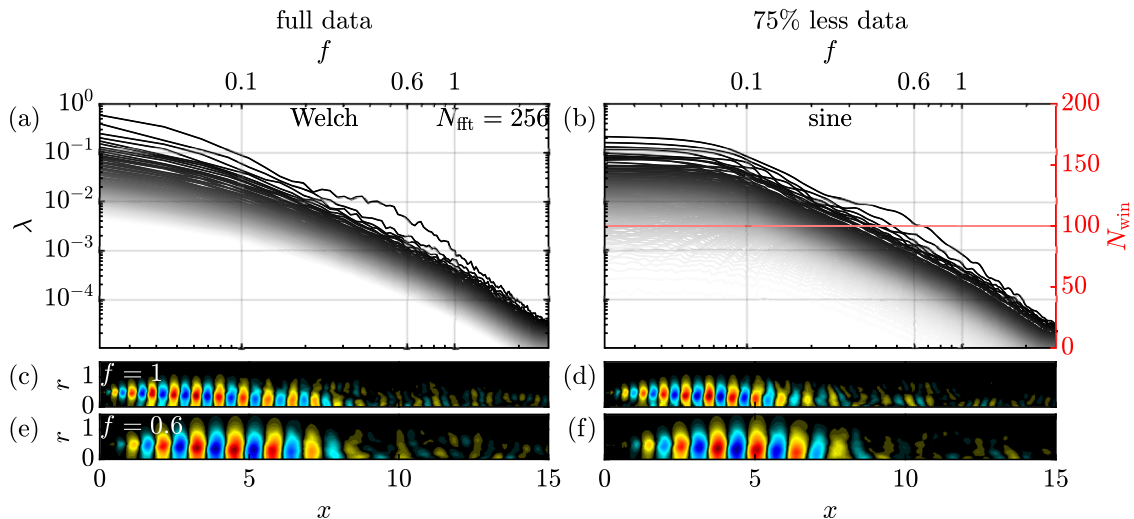
Fig. 11 SPOD eigenvalues and leading modes of the jet computed from: **a**, **c**, **e** all 10,000 snapshots using the Welch method; **b**, **d**, **f** the first 2500 snapshots, or 75% less data, using non-adaptive sine-taper SPOD. The Welch SPOD uses $N_{\text{fft}} = 256$ with 50% overlap and the Hamming window—the same parameters as in [20]. The sine-taper SPOD uses $N_{\text{win}} = 100$

5 Discussion

Distinct from spectral estimation of one-dimensional signals, SPOD requires the estimation of both modes and their corresponding energies. The latter, in particular, must be well-converged to aid in physical discovery and reduced-order modeling efforts. An accurate estimate of SPOD modes and modal energies hinges on satisfying two competing objectives: on one hand, minimizing the variance of the modes and energies, and on the other, minimizing their bias. With standard Welch-SPOD, the user attempts to balance both objectives by selecting an N_{fft}. The recently-introduced multitaper SPOD of [17], which uses the family of DPSS tapers, affords finer-grained control of this balance by adding N_{win} as an input. Its shortcomings include excessive bias when estimating tonal spectra, and high computational cost. In this work, we introduce sine tapers to multitaper SPOD. By construction, sine tapers minimize local bias. They also do not require the explicit computation and storage of N_{win} Fourier transforms. Instead, sine-tapered Fourier realizations can be recovered from the central differences between pairs of Fourier modes sampled from a single FFT. In addition, sine tapers enable N_{win} to vary with frequency and adapt to rapid variations of the spectrum within a given bandwidth, making sine-taper SPOD well-suited to tonal spectra.

To further protect tones in the spectrum from bias, we propose an iterative procedure that uses modal convergence to determine the taper number, N_{win, f_k}, to apply to each frequency. Starting from a small taper number, we incrementally add tapers to all frequencies, but fix N_{win, f_k} for any frequencies at which the leading SPOD mode has converged to tolerance. For the remaining frequencies, the process is repeated until all leading modes are converged. The frequencies that converge faster tend to be characterized by tones or other low-rank dynamics. As a result, the bandwidths at these frequencies become narrower than those in broadband regions. The similarity tolerance is the only parameter that the user need select. For large data, SPOD, especially multitaper SPOD, may become intractable. To further alleviate high computational complexity and memory usage, we losslessly compress the data in the time domain via a QR or eigenvalue decomposition. To a large extent, the combination of sine tapers and lossless compression frees the user from limitations on estimation parameters due to a lack of local computational resources. For the TR-PIV data of the cavity flow, compared to DPSS-taper SPOD with a constant N_{win}, we find adaptive sine-taper SPOD enhances the resolvability of tones, reduces bin-to-bin variance of the spectrum, and improves mode convergence. For the same similarity tolerance, we observe similar bin-to-bin variance and mode smoothness for the turbulent jet data. After testing on an additional TR-PIV database [22] and a direct numerical simulation (DNS) database (with similar parameters to [8]), we feel confident in recommending $tol = 10^{-6}$ as best practice. While it is always beneficial to obtain more data, the adaptive SPOD of tonal flows produces results similar to those of standard Welch SPOD with up to 75% less data.

In Sect. 3.1 we compared sine-taper SPOD with and without segmenting. While segmenting can provide comparable results, avoiding it offers greater simplicity and maximizes frequency resolution. If the data record

is very long, however, allowing N_{fft} to grow with N_t may become untenable. This is easily prevented by capping N_{fft} at some fixed value. Care must also be exercised when interpreting the total power (i.e., area under the curve) of an adaptively tapered spectrum. For the adaptive version of Thomson's multitaper estimator applied to one-dimensional signals [26], it can be shown [10,11] that Parseval's Theorem in Eq. (9) is no longer guaranteed to be satisfied, even in expected value. For our adaptive sine-taper SPOD, we also observe this empirically. In practice this effect is modest; the total power of the $tol = 10^{-6}$ spectrum of the cavity in Fig. 8a, for instance, exceeds the sample variance only by 2%. The jet example in Fig. 10 highlights a different limitation of multitaper SPOD: when N_{win} is large, modes may be biased in wavenumber and frequency. This bias is common to DPSS and sine tapers, but most severe for the adaptive algorithm. It arises due to partial correlation between modes at different frequencies within a given bandwidth. The bias becomes more pronounced at small f, since a given bandwidth encompasses structures of more disparate wavenumbers at low frequencies than it does at high frequencies. Because the underlying spectrum decays with frequency, low-frequency structures are energetically dominant, biasing the estimated modes towards $f = 0$ and $k = 0$. This is another manifestation of the variance-bias trade-off. The bias can be alleviated with a variable tol that decreases with frequency, sacrificing convergence for small f. In designing the adaptive algorithm for tonal flows, we have prioritized protecting tones and controlling variance over extracting low-frequency modes. For the most accurate estimation of broadband flows, adaptivity should be switched off entirely. Instead we recommend a constant N_{win} to obtain spectral estimates that retain the advantages of non-adaptive sine-taper SPOD over the DPSS and Welch estimators.

Availability of data and materials Data is available on request from the authors. Code is openly available in a public repository.

Declarations

Ethical approval Not applicable.

Appendix A: Wall time and memory usage

Figure 12 reports the wall-clock time and peak memory usage for the cavity and jet data as N_{win} is varied, using non-adaptive multitaper SPOD. Tests were conducted using MATLAB on a workstation with 48 Intel Xeon cores and 500 GB of memory. The most time-consuming operations, including QR, FFT, and EVD, were performed in single precision. At each frequency, only the three leading SPOD modes were stored. For DPSS-taper SPOD, the largest N_{win} tested for the jet was limited by available memory to 24. For sine-taper SPOD, lossless compression was achieved using QR and EVD for the cavity and jet, respectively.

The wall time and peak memory consumption of DPSS-taper SPOD grow linearly with N_{win}. For sine-taper SPOD, the wall time is dominated by compression of the snapshots and expansion of the SPOD modes; memory usage is dominated by storage of the snapshots and the modes. As a result, the wall time of sine-taper SPOD depends only weakly on N_{win}, while its memory usage stays constant. For the same reason, the memory usage of adaptive sine-taper SPOD is comparable to that of the non-adaptive algorithm. The adaptive algorithm is considerably slower, however. With $tol = 10^{-6}$, adaptive SPOD of the cavity and jet examples take approximately 18 and eight hours, respectively, to complete. Although the DPSS-taper and Welch algorithms may also benefit from compression, they require computing and storing an FFT for every taper; the sine-taper algorithm, on the other hand, uses the central difference identity in Eq. (13), thus requiring only a single FFT. For cases where $N_{\text{dof}} \gg N_t$, e.g., the jet, lossless compression and the central difference both contribute to efficiency; for small N_{dof}, e.g., the cavity, the central difference is responsible for most of the advantage. Comparing Fig. 12a and b also reveals that for sine-taper SPOD the wall times of the cavity and jet cases are of the same order of magnitude, despite a 26-fold difference in N_{dof}. Clearly the wall time does not scale with
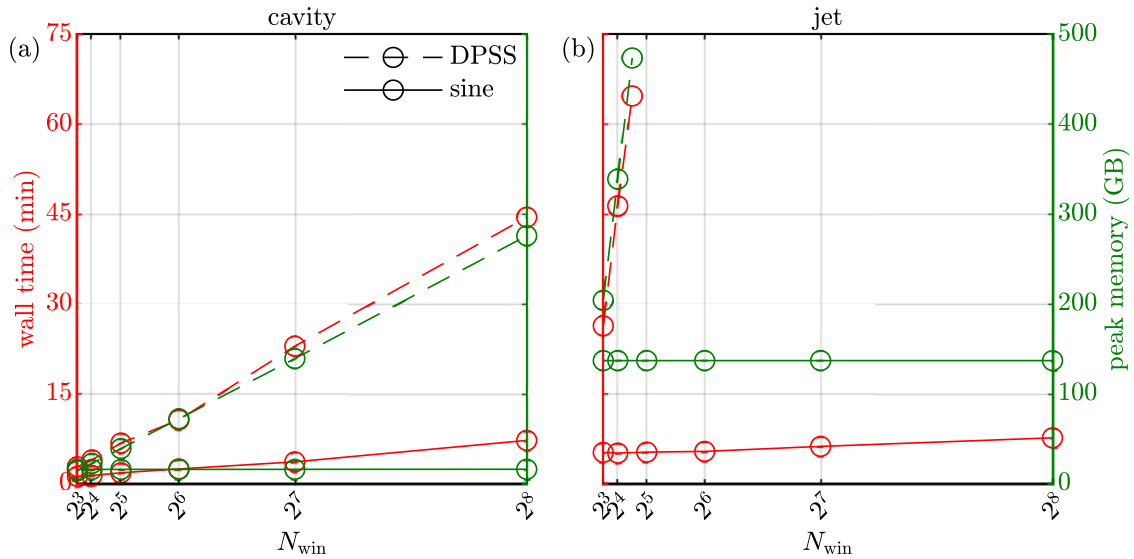
Fig. 12 Wall-clock time (red) and peak memory usage (green) of DPSS-taper (dashed line) and sine-taper (solid line) SPOD: **a** cavity; **b** jet (Colour figure online)

grid size. Together with reduced memory usage, we anticipate this feature will be crucial to the analysis of three-dimensional flow data that lack spatial homogeneities.

Appendix B: Lossy compression

As outlined in Sect. 2.3, if the computational effort required by adaptive SPOD is excessive, SVD truncation can achieve further acceleration at the expense of convergence at high frequencies, which tend to have low energy. However, once N_{win, f_k} has been determined adaptively, we recompute the SPOD using the losslessly-compressed data. The final spectrum and modes are thus dictated solely by N_{win, f_k}. For the SPOD in Fig. 13 computed with $tol = 10^{-6}$, we retain the first $d = 1000$ singular values and vectors of the weighted data matrix $\mathbf{X}$. This represents an additional 16-fold compression relative to lossless compression, and drops the wall time from 18 to 2.4 hours. The spectrum, modes, and N_{win, f_k} display excellent agreement with those in Fig. 8 for the same value of tol.
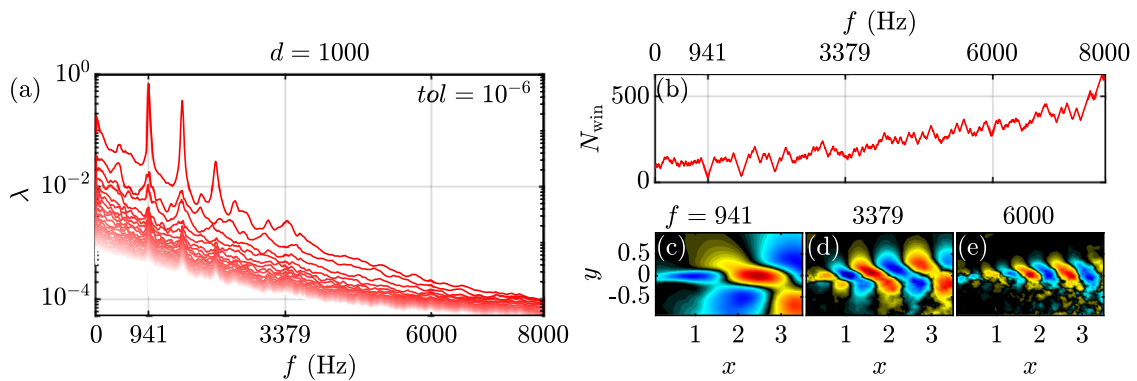

Fig. 13 Same as Fig. 8 for $tol = 10^{-6}$ (red curves), but computed with lossy compression, retaining the first $d = 1000$ singular values and vectors of the weighted data matrix (Colour figure online)

References

1. Barbour, A.J., Parker, R.L.: psd: Adaptive, sine multitaper power spectral density estimation for R. Comput. Geosci. **63**, 1–8 (2014). https://doi.org/10.1016/j.cageo.2013.09.015
2. Brès, G.A., Ham, F.E., Nichols, J.W., Lele, S.K.: Unstructured large-eddy simulations of supersonic jets. AIAA J. **55**(4), 1164–1184 (2017). https://doi.org/10.2514/1.J055084
3. Brès, G.A., Jordan, P., Jaunet, V., Rallic, M.L., Cavalieri, A.V.G., Towne, A., Lele, S.K., Colonius, T., Schmidt, O.T.: Importance of the nozzle-exit boundary-layer state in subsonic turbulent jets. J. Fluid Mech. **851**, 83–124 (2018). https://doi.org/10.1017/jfm.2018.476
4. Chu, B.T.: On the energy transfer to small disturbances in fluid flow (part I). Acta Mech. **1**, 215–234 (1965). https://doi.org/10.1007/BF01387235
5. Glauser, M.N., Leib, S.J., George, W.K.: Coherent structures in the axisymmetric turbulent jet mixing layer. Turbul. Shear Flows **5**, 134–145 (1987). https://doi.org/10.1007/978-3-642-71435-1_13
6. Hansson, M.: Optimized weighted averaging of peak matched multiple window spectrum estimates. IEEE Trans. Signal Proc. **47**(4), 1141–1146 (1999). https://doi.org/10.1109/78.752613
7. Lumley, J.L.: Stochastic Tools in Turbulence. Academic Press, New York (1970)
8. Mathias, M.S., Medeiros, M.A.F.: Interaction between Rossiter and Görtler modes in the compressible flow in an open cavity. In: AIAA Aviation 2020 Forum, vol. AIAA 2020-3074 (2020). https://doi.org/10.2514/6.2020-3074
9. Oppenheim, A.V., Schafer, R.W.: Discrete-Time Signal Processing. Pearson, Upper Saddle River (2010)
10. Park, J., Lindberg, C.R., Vernon, F.L., III.: Multitaper spectral analysis of high-frequency seismograms. J. Geophys. Res. **92**(B12), 12675–12684 (1987). https://doi.org/10.1029/JB092iB12p12675
11. Percival, D.B., Walden, A.T.: Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques. Cambridge University Press, Cambridge (1993)
12. Picard, C., Delville, J.: Pressure velocity coupling in a subsonic round jet. Int. J. Heat Fluid Flow **21**(3), 359–364 (2000). https://doi.org/10.1016/S0142-727X(00)00021-7
13. Prieto, G.A., Parker, R.L., Thomson, D.J., Vernon, F.L., Graham, R.L.: Reducing the bias of multitaper spectrum estimates. Geophys. J. Int. **171**(3), 1269–1281 (2007). https://doi.org/10.1111/j.1365-246X.2007.03592.x
14. Riedel, K.S., Sidorenko, A.: Minimum bias multiple taper spectral estimation. IEEE Trans. Signal Proc. **43**(1), 188–195 (1995). https://doi.org/10.1109/78.365298
15. Rowley, C.W., Williams, D.R.: Dynamics and control of high-Reynolds-number flow over open cavities. Annu. Rev. Fluid Mech. **38**, 251–276 (2006). https://doi.org/10.1146/annurev.fluid.38.050304.092057
16. Schmid, P.J.: Dynamic mode decomposition of numerical and experimental data. J. Fluid Mech. **656**, 5–28 (2010). https://doi.org/10.1017/S0022112010001217
17. Schmidt, O.T.: Spectral proper orthogonal decomposition using multitaper estimates. Theor. Comput. Fluid Dyn. **36**, 741–754 (2022). https://doi.org/10.1007/s00162-022-00626-x
18. Schmidt, O.T., Colonius, T.: Guide to spectral proper orthogonal decomposition. AIAA J. **58**(3), 1023–1033 (2020). https://doi.org/10.2514/1.J058809
19. Schmidt, O.T., Towne, A.: An efficient streaming algorithm for spectral proper orthogonal decomposition. Comput. Phys. Commun. **237**, 98–109 (2019). https://doi.org/10.1016/j.cpc.2018.11.009
20. Schmidt, O.T., Towne, A., Rigas, G., Colonius, T., Brès, G.A.: Spectral analysis of jet turbulence. J. Fluid Mech. **855**, 953–982 (2018). https://doi.org/10.1017/jfm.2018.675
21. Semeraro, O., Jaunet, V., Jordan, P., Cavalieri, A.V.G., Lesshafft, L.: Stochastic and harmonic optimal forcing in subsonic jets. In: 22nd AIAA/CEAS Aeroacoustics Conference, vol. AIAA 2016-2935 (2016). https://doi.org/10.2514/6.2016-2935
22. Seo, J., Zhang, Y., Mittal, R., Cattafesta, L.N., III.: Vortex-induced sound prediction of slat noise from time-resolved particle image velocimetry data. Exp. Fluids **64**(99), 1–13 (2023). https://doi.org/10.1007/s00348-023-03636-5
23. Sirovich, L.: Turbulence and the dynamics of coherent structures part I: coherent structures. Quart. Appl. Math. **45**(3), 561–571 (1987). https://doi.org/10.1090/qam/910462
24. Slepian, D.: Prolate spheroidal wave functions, Fourier analysis, and uncertainty–v: the discrete case. Bell Syst. Tech. J. **57**(5), 1371–1430 (1978). https://doi.org/10.1002/j.1538-7305.1978.tb02104.x
25. Slepian, D., Pollak, H.O.: Prolate spheroidal wave functions, Fourier analysis and uncertainty—I. Bell Syst. Tech. J. **40**(1), 43–63 (1961). https://doi.org/10.1002/j.1538-7305.1961.tb03976.x
26. Thomson, D.J.: Spectrum estimation and harmonic analysis. Proc. IEEE **70**(9), 1055–1096 (1982). https://doi.org/10.1109/PROC.1982.12433
27. Towne, A., Schmidt, O.T., Colonius, T.: Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. J. Fluid Mech. **847**, 821–867 (2018). https://doi.org/10.1017/jfm.2018.283
28. Walden, A.T., McCoy, E.J., Percival, D.B.: The effective bandwidth of a multitaper spectral estimator. Biometrika **82**(1), 201–214 (1995). https://doi.org/10.1093/biomet/82.1.201
29. Welch, P.D.: The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans. Audio Electroacoust. **15**(2), 70–73 (1967). https://doi.org/10.1109/TAU.1967.1161901
30. Yeung, B.C.Y., Schmidt, O.T., Brès, G.A.: Three-dimensional spectral POD of supersonic twin-rectangular jet flow. In: AIAA Aviation 2022 Forum, vol. AIAA 2022-3345 (2022). https://doi.org/10.2514/6.2022-3345
31. Zhang, Y., Cattafesta, L.N., III., Ukeiley, L.: Spectral analysis modal methods (SAMMs) using non-time-resolved PIV. Exp. Fluids **61**(226), 1–12 (2020). https://doi.org/10.1007/s00348-020-03057-8