Improving the Convergence of Spectral Proper Orthogonal Decomposition using Multitaper Estimates

Oliver T. Schmidt

Department of Mechanical and Aerospace Engineering, University of California San Diego, La Jolla, CA 92093, USA

Abstract

The use of multitaper estimates for spectral proper orthogonal decomposition (SPOD) is explored. Multitaper and multitaper-Welch estimators that use discrete prolate spheroidal sequences (DPSS) as orthogonal data windows are compared to the standard SPOD method that exclusively relies on weighted overlapped segment averaging, or Welch’s method, to estimate the cross-spectral density matrix. Two sets of turbulent flow data, one experimental and the other numerical, are used to discuss the choice of resolution bandwidth and the bias-variance tradeoff. Multitaper-Welch estimators that combine both approaches by applying orthogonal tapers to overlapping segments allow for flexible control of resolution, variance, and bias, and outperform the standard method in any of these three aspects if the other two are comparable.

Keywords:

1. Introduction

SPOD is the specialization of the most general form of proper orthogonal decomposition (POD, [1]) to statistically stationary data and is best computed in the frequency domain. An early application of the method can be found in the experimental work of Glauser et al. [2]. The method has recently gained popularity as theoretical connections to dynamic mode decomposition (DMD, [3])...
and resolvent analysis of turbulent mean flows [4] were established by Towne et al. [5]. Another factor is its application to large numerical data that were specifically generated for this purpose [6]. SPOD has been applied to a broad range of canonical [7, 8, 9, 10, 11, 12, 13] and technical [14, 15, 16, 17] turbulent flows, weather and climate data [18], and found diverse applications in aeroacoustics [19, 20], stochastic estimation [21], frequency-time analysis [22], flow field reconstruction [23, 22], and reduced-order modeling [24, 25].

The SPOD modes and their mode energies are, respectively, the eigenvectors and eigenvalues of the weighted cross-spectral density (CSD) matrix that is estimated from the data. The standard approach to estimate the CSD matrix for SPOD is weighted overlapped segment averaging (WOSA), or Welch’s method [26, 27]. A different approach that is routinely used for the estimation of power spectra from time signals is the multitaper (MT), or Thomson’s multitaper method [28]. The multitaper method corresponds to weighted averaging of independent spectral estimates that are obtained from multiple orthogonal data windows. Most notably, it has been demonstrated by Bronez [29] that the multitaper power spectrum estimator outperforms the Welch estimator in terms of leakage, variance, and resolution if two of these three quantities are fixed to determine the third.

The remainder of the paper is organized as follows. The technical background is discussed in §2 and includes a brief summary of the discrete SPOD problem (§2.1) and DPSS (§2.2), as well as a recapitulation of Welch’s method, and the introduction of multitaper estimates for SPOD (§2.3). The two example data sets are presented in §2.4. In §3, SPOD analyses conducted using the basic Welch and single-block multitaper estimates (§3.1) are compared and discussed in the light of parameter selection, frequency resolution, bandwidth, and the variance-bias tradeoff. Next, multitaper-Welch estimates that combine the advantages of both methods are computed for the two examples and discussed in §3.2. Implications and future applications of this work are discussed in §4.
2. Background and Methodology

2.1. Spectral proper orthogonal decomposition (SPOD)

The discrete SPOD of an ensemble of \( N_t \) snapshots of spatio-temporal multivariate data with zero mean,

\[
q_i = q(t_i), \quad i = 1, \ldots, N_t, \tag{1}
\]

here written as a sequence of column vectors, is obtained by computing the eigendecompositions

\[
C_j W \Phi_j = \Phi_j A_j, \quad j = 1, \ldots, N_{\text{fft}}, \quad C_j \in \mathbb{R}^{N_{\text{fft}} \times N_{\text{fft}}}, \tag{2}
\]

of the cross-spectral density (CSD) matrices, \( C_j \). The rank \( M \) of the CSD matrices corresponds to the number of independent realizations of the Fourier transform that are used for their construction. The size of the CSD matrices is determined by the number of discrete frequencies, \( N_{\text{fft}} \), obtained from the discrete Fourier transform (DFT) in the permissible frequency band between zero and the Nyquist frequency, \( f_{\text{Nyq}} = \frac{f_s}{2} \), where \( f_s = 1/\Delta t \) is the sampling frequency. The columns of \( \Phi_j = \left[ \phi_j^{(1)}, \phi_j^{(2)}, \ldots, \phi_j^{(M)} \right] \), i.e., the eigenvectors of \( C_j \), are the SPOD modes and the diagonal entries of \( A_j \), i.e., the eigenvalues \( \lambda_j^{(1)} \geq \lambda_j^{(2)} \geq \cdots \geq \lambda_j^{(M)} \) of \( C_j \), the corresponding mode energies. The weight matrix \( W \) accommodates spatial integration and, for multivariate data, variable-dependent weights. The discrete (sample) CSD matrix is computed as

\[
C_j = \frac{1}{M} \sum_{k=1}^{M} \hat{q}_j^{(k)} \left( \hat{q}_j^{(k)} \right)^*, \tag{3a}
\]

\[
= \frac{1}{M} \hat{Q}_j \hat{Q}_j^*, \tag{3b}
\]

where \( \hat{q}_j^{(k)} \) is the \( k \)-th out of \( M \) realizations of the discrete Fourier transform at the \( j \)-th frequency. The matrix representation, equation \( (3b) \), uses the data matrix \( \hat{Q}_j = \left[ \hat{q}_j^{(1)}, \hat{q}_j^{(2)}, \ldots, \hat{q}_j^{(M)} \right] \). The denominator is either \( M \) or \( M - 1 \), depending on whether the true mean, if available, or the sample mean is subtracted from the data, respectively. The desirable mathematical properties of the SPOD, its continuous formulation, and algorithmic implementations are discussed in detail elsewhere [5, 27].
2.2. Discrete Prolate Spheroidal Sequences (DPSS)

![Figure 1: Leading seven DPSS for $N = 4096$ and $b_{\text{win}} = 4$ compared to a standard Hamming window. The DPSS are normalized by the maximum value of $s_1$ for clarity.](image)

Discrete prolate spheroidal sequences (DPSS), or Slepians, are the discrete-time equivalent of prolate spheroidal wave functions [30, 31]. They are the optimal solutions to the spectral concentration problem that seeks orthogonal time sequences whose discrete Fourier transform is maximally localized in a frequency band $[-\delta f, \delta f]$, defined by the resolution half-bandwidth, $\delta f$. Given the window length, $N$, and resolution half-bandwidth, $\delta f$, a common way to find the DPSS is the eigendecomposition

$$
\mathbf{Hs} = \gamma \mathbf{s},
$$

where the entries of the Hermitian matrix $\mathbf{H} = \mathbf{H}(\delta f, N)$ are given by

$$
h_{ij} = \frac{\sin(2\pi \delta f (j - i))}{\pi(j - i)}, \quad i, j = 0, \ldots, N - 1.
$$

The eigenvalues $1 > \gamma_0 > \gamma_1 > \cdots > \gamma_{N-1} > 0$ correspond to the spectral concentrations of energy of the respective tapers, $s_0, s_1, \ldots, s_{N-1}$, in the frequency band $[-\delta f, \delta f]$. Since $\mathbf{H}$ is Hermitian, the DPSS are mutually orthogonal. It can be shown that only $2\delta f N$ (Shannon number) of the spectral concentrations are close to unity. It is hence customary to use only the first $N_{\text{win}} = 2\delta f N - 1$ Sleipian tapers as windows. In terms of the time-halfbandwidth product, $b_{\text{win}} = \delta f N$, this statement becomes

$$
N_{\text{win}} = \lfloor 2b_{\text{win}} \rfloor - 1,
$$

4
and guarantees that the Fourier transforms of all windows are well-concentrated in \([-\delta f, \delta f]\) to minimize spectral leakage to neighboring frequencies. The leading seven DPSS for \(N = 4096\) and \(b_{\text{win}} = 4\) are compared to a standard Hamming window in figure 1.

2.3. Welch and multitaper estimates of the SPOD

The \(N\) independent realizations of the discrete Fourier transform that form the columns of the sample CSD matrix in equation (3a) are obtained in one of two ways. The first option is to obtain well-separated sequences, or blocks, \(\hat{q}_j^{(k)}\) with \(j = 1, \ldots, N_{\text{fft}}\) and \(k = 1, \ldots, N_{\text{blk}}\), one at a time. Denote by \(N_{\text{blk}}\) the total number of blocks, each consisting of \(N_{\text{fft}}\) consecutive snapshots, equally spaced in time by \(\Delta t\). This scenario occurs when data acquisition at a high sampling rate is the bottleneck, as is the case in many experimental settings. The second possibility is that the data is acquired as a single, long time series of \(N_t\) snapshots, \(q_j = q(t_j)\), with \(j = 1, \ldots, N_t\). This is often the case in numerical settings, where the overall runtime is the bottleneck. In this case, the long time series is segmented into \(N_{\text{blk}}\) blocks, \(\hat{q}_j^{(k)} = q_j + (k-1)(N_{\text{fft}}-N_{\text{ovlp}})+1\) with \(j = 1, \ldots, N_{\text{fft}}\) and \(k = 1, \ldots, N_{\text{blk}}\), under the ergodicity hypothesis. Adjacent blocks are allowed to overlap by \(N_{\text{ovlp}}\) snapshots, which results in a total number of \(N_{\text{blk}} = \left\lfloor \frac{N_t-N_{\text{ovlp}}}{N_{\text{fft}}-N_{\text{ovlp}}} \right\rfloor\) blocks. This segmentation is the idea behind the method of Welch [26], that was originally devised for the estimation of power spectra of time signals. In the same work, Welch established from theoretical considerations the best practice of using an overlap of 50%. After segmenting the data, \(N_{\text{blk}}\) realizations of the Fourier transform are obtained from the weighted temporal discrete Fourier transform of each block,

\[
\hat{q}_j^{(k)} = \sum_{i=0}^{N_{\text{fft}}-1} w[i]q_i^{(k)} e^{\frac{2\pi i}{N_{\text{fft}}}ij}, \quad k = 1, \ldots, N_{\text{blk}},
\]

where \(q_i^{(k)}\) is the \(k\)-th realization of the Fourier transform at the \(j\)-th discrete frequency. It is common practice to use a data window, \(w\), to prevent spectral leakage due to the non-periodicity of the samples. A popular choice is the
A symmetric Hamming window,

$$w[i] = 0.54 - 0.46 \cos \left( \frac{2\pi i}{N_{\text{fft}} - 1} \right), \quad i = 0, \ldots, N_{\text{fft}} - 1.$$  (8)

Given the ensemble of realizations, equation (9), the CSD matrix is formed according to equation (3a), and its eigendecomposition, equation (2), concludes the standard SPOD algorithm.

Inspired by the method of Thomson [28], also originally intended for the estimation of power spectra, we propose to use a set of $N_{\text{win}}$ orthogonal DPSS (see §2.2) as data windows for the SPOD. We restrict this set to the first $2b_{\text{win}} - 1$ sequences (equation (6)) to ensure that all data windows are well-concentrated in $[-\delta f, \delta f]$. This minimizes spectral leakage and avoids the necessity to use different weights for different windows. Since the DPSS are orthogonal, the different realizations are independent. By combining segmenting and orthogonal windowing, a total number of $N_{\text{blk}}N_{\text{win}}$ realizations,

$$\hat{q}_{j}^{(k,l)} = \sum_{i=0}^{N_{\text{win}}-1} s_l[i] q_l^{(k)} e^{i \frac{2\pi}{N_{\text{win}}} ij}, \quad k = 1, \ldots, N_{\text{blk}}, \quad l = 0, \ldots, N_{\text{win}} - 1,$$  (9)

is obtained. Here, $s_l[i]$ is the $i$-th element of the $l$-th DPSS, $s_l$, computed from equation (4). The first seven DPSS for $N_{\text{fft}} = 4096$ and $b_{\text{win}} = 4$ are shown in figure 1 and compared to the Hamming window, equation (8). The corresponding sample CSD matrix is

$$C_j = \frac{1}{N_{\text{blk}}N_{\text{win}}} \sum_{k=1}^{N_{\text{blk}}} \sum_{l=1}^{N_{\text{win}}} \hat{q}_{j}^{(k,l)} \left( \hat{q}_{j}^{(k,l)} \right)^*.$$  (10)

Its matrix representation is also given by equation (3b), but for $M = N_{\text{blk}}N_{\text{win}}$ and $\hat{Q}_j = [\hat{q}_j^{(1,1)}, \hat{q}_j^{(1,2)}, \ldots, \hat{q}_j^{(N_{\text{blk}}-1,N_{\text{win}})}, \hat{q}_j^{(N_{\text{blk}},N_{\text{win}})}]$.

The standard Welch SPOD method is recovered if a single data window is applied to multiple blocks, ($N_{\text{win}} = 1$, $N_{\text{blk}} > 1$); a multitaper SPOD method if multiple orthogonal windows are applied to a single segment ($N_{\text{win}} > 1$, $N_{\text{blk}} = 1$); and a hybrid multitaper-Welch SPOD method if blocking and orthogonal windowing are combined ($N_{\text{win}} > 1$, $N_{\text{blk}} > 1$).
2.4. Data

| Example        | Variables | $N_x$ | $N_{y,r}$ | $N_t$ | $\Delta t$ |
|----------------|-----------|-------|-----------|-------|------------|
| Jet LES\cite{32} | $p$       | 261   | 58        | 10,000| 0.2        |
| Cavity PIV\cite{33} | $u, v$   | 78    | 28        | 16000 | 6.24 $\cdot$ 10$^{-5}$ |

Table 1: Database and spectral estimation parameters. The jet LES data are non-dimensionalized by the jet diameter and velocity, whereas the cavity PIV data are reported in SI units.

Two sets of data are taken as examples. The first data set consists of 10000 snapshots of the pressure field from the wall-modeled large-eddy simulation (LES) of a turbulent $M = 0.9$ jet by Brèş et al.\cite{32}. This data has been studied in detail using SPOD by Schmidt et al.\cite{6}. The second data set consists of 16000 snapshots of the streamwise and wall-normal velocity fields of the turbulent $M = 0.6$ flow over an open cavity acquired using time-resolved particle image velocimetry (TR-PIV) by Zhang et al.\cite{33}. The open cavity has a length-to-depth ratio of $L/D = 6$ and a width-to-depth ratio of $W/D = 3.85$. More details on the experimental setup can be found in Zhang et al.\cite{34}, and an analysis of the data, including using SPOD, in Zhang et al.\cite{33}. The spatial resolutions of both databases were reduced for this study.

Figure 2: Instantaneous flow field visualizations: (a) fluctuating pressure of the turbulent jet LES; (b) streamwise fluctuating velocity of the open cavity PIV data.
Instantaneous flow field visualizations of both data sets are shown in figure 2, and relevant parameters are reported in table 1. The turbulent nature of both flows is apparent from the visualizations. The turbulent jet data are representative of turbulent flows with broadband turbulence spectra. The open cavity data are representative of turbulent flows with tonal peaks and an underlying broadband spectrum. The two sets also differ in that measurement noise is present only in the experimental cavity flow data.

3. Results

For the remainder of this paper, we follow best spectral estimation practices [26], and set $N_{\text{ovlp}}$ to 50% of $N_{\text{fft}}$ for all Welch estimates. To compare multiple spectra for different parameters, we first focus on the leading eigenvalues, $\lambda^{(1)}_j$. The variance of the remaining eigenvalues is similar to that of the leading eigenvalue. We do, however, inspect the statistical convergence of the leading two SPOD modes, $\phi^{(1)}_j$ and $\phi^{(2)}_j$, and note that the convergence of the second mode is indicative of the convergence of the remaining modes. The leading SPOD modes play a special role in that they often capture physical mechanisms like resonances or hydrodynamic instabilities that dominate the dynamics at a given frequency. The presence of such mechanisms leads to a rapid convergence of the leading mode and separation of the leading two eigenvalues that represent the mode energies. The latter phenomenon is often referred to as ‘low-rank behavior’, which should be understood in a physical rather than strictly mathematical sense ($C_j \in \mathbb{R}^{N \times N}$ is positive semi-definite in general and has full rank only if constructed from $M = N$ independent realizations $q_j^{(k)}$, $k = 1, \ldots, N$). We begin by comparing the basic Welch and multitaper estimates of the SPOD for the open cavity data in §3.1. In this context, we discuss the bias-variance tradeoff and some subtleties of the estimation process that are best explained by example. A comparative study of both test databases using multitaper-Welch estimates is presented in §3.2.
3.1. Standard Welch and single-block multitaper estimates of the SPOD

Figure 3: Leading SPOD mode spectra of the open cavity PIV data using different estimators: (a) Welch; (b) multitaper. The leading SPOD modes at $f = 421.9$ and $f = 938.5$ are shown in figure 4 below.

Figure 3 compares SPOD eigenvalue spectra for the open cavity flow PIV data obtained using standard Welch and single-segment multitaper estimates. The free parameter that is varied for the Welch estimates in figure 3(a) is the number of samples per block, $N_{\text{fft}}$. For a fixed overlap of 50%, this also fixes the number of blocks, $N_{\text{blk}}$. The multitaper estimates in figure 3(b) are obtained for the entire dataset of $N_{\text{fft}} = N_t = 16000$ snapshots, and the time-halfbandwidth product, $b_{\text{win}}$, was varied. The number of tapers used for each estimate, $N_{\text{win}}$, then directly follows from equation (6). The modes corresponding to two particular frequencies of $f = 938.5$ and $f = 421.9$ that are labeled on the frequency axis will be discussed later. The higher frequency corresponds to the highest peak in the spectrum and is associated with the first Rossiter mode [35], a resonant hydrodynamic instability. The lower frequency corresponds to the much broader peak in the region of high variance that appears close to the Rossiter frequency on the logarithmic scale.

Welch’s method is asymptotically unbiased for stationary data and converges
to the true spectral density as $N_{\text{blk}}, N_{\text{fft}} \to \infty$. For limited data, however, care has to be taken regarding the balance of bias and variance. This becomes clear in figure 3(a). For $N_{\text{fft}} = 2^{12} = 4096$ (black line), the data are segmented into only six blocks, and the noisy appearance of the SPOD spectrum indicates the high variance of the estimate. As $N_{\text{fft}}$ is decreased, the number of blocks increases, and the variance is significantly reduced. Observe, however, that the distinct spectral peaks at the Rossiter and some higher frequencies vanish simultaneously. This is evidence of bias due to the decreased frequency resolution, $\Delta f_{\text{Welch}} = \frac{1}{N_{\text{fft}} \Delta t} = \frac{f_s}{N_{\text{fft}}}$, with decreasing $N_{\text{fft}}$ and the associated spectral leakage of the under-resolved features. Similar to Welch’s method, multitaper estimates of the spectral density based on DPSS are asymptotically unbiased and their mean square error converges to zero as $N \to \infty$ (but at the same time $\delta f \to 0$ with $b_{\text{win}} \to \infty$, and $N_{\text{win}} \leq 2b_{\text{win}}$ but $N_{\text{win}} \to \infty$; see [36] for details). As $b_{\text{win}}$ is increased in figure 3(b), the same qualitative trends are observed as for the Welch estimator, i.e., the variance decreases as the bias increases. The multitaper estimate retains the same frequency resolution, or bin size, of $\Delta f = \frac{1}{\Delta t N_{\text{fft}}} = 1$, whereas for the Welch estimate, it increases from $\Delta f_{\text{Welch}} = \frac{1}{\Delta t N_{\text{fft}}} = 3.9$ for $N_{\text{fft}} = 4096$ to $\Delta f_{\text{Welch}} = 250$ for $N_{\text{fft}} = 64$. For the multitaper estimator at a fixed $N$, the time-halfbandwidth product, $b_{\text{win}}$, determines the frequency bandwidth, $\Delta f_{\text{DPSS}} = f_s \delta f$, over which the DPSS tapers yield an integrated average. Welch and multitaper estimates may be compared if they average over a similar frequency band, this is if $\Delta f_{\text{Welch}} \approx \Delta f_{\text{DPSS}}$. For the Welch estimate with $N_{\text{fft}} = 512$, for example, the frequency resolution is reduced by a factor of $\frac{N_{\text{fft}}}{N_{\text{fft}}} = 31.25$. Because $\Delta f = 1$ (16000 samples recorded at 16 kHz over 1 sec) in the experiment, the same value, $\Delta f_{\text{Welch}} = 31.25$, is obtained for the frequency bin size. Since $b_{\text{win}}$ determines the resolution bandwidth in multiples of $\Delta f$, we consider $b_{\text{win}} = 31$, and may confirm that $\Delta f_{\text{DPSS}} = 31$ (again, because $\Delta f = 1$). We observe that the Welch estimate for $N_{\text{fft}} = 512$ (blue line in figure 3(a)) and the multitaper estimate for $b_{\text{win}} = 31$ (gray line in figure 3(b)), are indeed qualitatively similar in terms of their variance and resolution of peaks. If the variance is further reduced (cyan lines in figure 3(a,b)), then the
bias is further increased, and spectral peaks are increasingly under-resolved for
the Welch estimator, and flattened and broadened for the multitaper estimator.
An advantage of multitaper SPOD for comparable bias and variance is that it
resolves lower frequencies down to $\Delta f$. A disadvantage is that it becomes com-
putationally expensive for large $N_{\text{win}}$, as it requires parallel processing of $N_{\text{win}}$
realizations of the Fourier transform of the full data.

Figure 4: Real part of the leading SPOD modes of the open cavity PIV data using different
estimators: (a,b) Welch for varying $N_{\text{fft}}, N_{\text{blk}}$; (c,d) multitaper for varying $b_{\text{win}}, N_{\text{win}}$. The
leading SPOD modes at $f = 421.9$ and $f = 938.5$ are shown in the left and right columns,
respectively (labelled in figure 3 above).

Figure 4 shows SPOD modes obtained using the Welch (top) and multitaper
(bottom) estimators at the two different frequencies, respectively representing
the low-frequency region with high variance (left) and the tonal peak frequency
associated with the first Rossiter mode (right). The phases of the complex
modes have been approximately aligned for comparability. At $f = 938.5$, the Rossiter mode in figure 4(b,d) is extracted by both methods, and well-converged for all parameters. This is typical for hydrodynamic instabilities and resonant mechanisms that are both energetic and evolve coherently over large regions in space and time. At the lower frequency of $f = 421.9$, no such mechanism is present, and both the Welch estimate for $N_{blk} = 6$ and the multitaper estimate for $N_{win} = 3$ appear noisy. Well-converged modes are obtained when $N_{fft}$ is decreased (and therefore $N_{blk}$ increased), or $b_{win}$ (and therefore $N_{win}$) are increased for the multitaper estimate.

3.2. Multitaper-Welch estimates of the SPOD

As demonstrated in §3.1 above, the multitaper method can, in principle, obtain low-variance estimates without segmentation if a large number of data windows is used. Doing so, however, quickly becomes computationally prohibitive. Furthermore, the estimate at the lowest resolvable frequency relies on a single realization of the flow process if $N = N_t$, and is therefore not statistically representative. These issues are addressed by applying multiple orthogonal windows to overlapping segments, that is, the multitaper-Welch estimate given by equation (10) with $N_{blk}, N_{win} > 1$. Our results from §3.1 suggest that Welch and multitaper SPOD yield comparably well-converged modes and spectra for similar bandwidths, $\Delta f_{Welch} \approx \Delta f_{DPSS}$, leaving computational cost as the only advantage of the Welch estimator.

Taking these considerations into account, we chose the highest $N_{fft}$ possible that results in the lowest number of blocks that still gives us confidence in the statistical convergence at low frequencies. That segment length is $N_{fft} = 4096$, resulting in $N_{blk} = 3$ and $N_{blk} = 6$ for the jet and cavity data, respectively. We then vary the bandwidth parameter, $b_{win}$, and therefore the number of Slepian tapers, $N_{win}$.
Figure 5 shows the SPOD spectra for the leading modes and both cases. The bandwidth parameter is varied from $b_{\text{win}} = 1$ to 25, resulting in a number of windows between $N_{\text{win}} = 1$ (Welch estimator) and 49. The total number of realizations, $N_{\text{blk}} N_{\text{fft}}$, is between 3 and 147 for the jet data and between 6 and 294 for the cavity data. A similar decrease in variance is observed for both data as the bandwidth is increased. This decrease is linked to an increase in bias that leads to the loss of certain features, such as several broader peaks at low frequencies in the jet data (still resolved for $b_{\text{win}} = 9$, green line) and several double peaks at high frequencies in the cavity data (still clearly visible for $b_{\text{win}} = 17$, gray line). Under the caveat that the bias cannot be quantified because the true spectrum is unknown, these observations reflect the usual variance-bias tradeoff.
Figure 6: Two leading SPOD modes for the jet LES data using multitaper-Welch estimates of different bandwidths: (a) first mode for $f = 0.055$; (b) first mode for $f = 0.33$; (c) second mode for $f = 0.055$; (d) second mode for $f = 0.33$. The corresponding leading SPOD energy spectra are reported in figure 5.

The first and second SPOD modes of the jet data at the two representative frequencies are shown in figure 6. Modes obtained for a single, an intermediate, and the highest number of tapers are compared. As expected, better-converged modes are obtained for higher bandwidths. Take as an example the leading SPOD mode for $f = 0.33$ shown in (b). For $b_{\text{win}} = 1$ and a single taper, the mode appears noisy, particularly in the downstream region for $x \gtrsim 10$. For the two higher bandwidths, the waveform is much more compact and smooth, and very similar for $b_{\text{win}} = 9$ and 25. This is indirect evidence of the statistical convergence of the mode. For the second (first subdominant) SPOD mode shown in figure (d), continual improvement is observed as the bandwidth is increased and a double-lobed structure emerges for the highest value of $b_{\text{win}} = 25$. 


Figure 7: Same as figure 6 but for the cavity PIV data: (a) first mode for \( f = 421.9 \); (b) first mode for \( f = 938.5 \); (c) second mode for \( f = 421.9 \); (d) second mode for \( f = 938.5 \).

Similar observations are made for the SPOD modes of the cavity PIV data shown in figure 7. The improvement of mode convergence with increasing bandwidth is apparent, in particular for the second SPOD modes in figure 7(c,d). A clear spatial pattern that is physically interpretable and quantifiable in terms of, for example, streamwise and wall-normal wavenumbers, is only revealed for the two higher bandwidths. A notable exception, as discussed earlier in the context of figure 4, is the Rossiter mode shown in figure 7(b) at a frequency of \( f = 938.5 \). It is associated with the overall largest SPOD eigenvalue and is so prevalent that it can be spotted in the instantaneous velocity fields in figure 7(b).
Finally, we investigate in figure 8 the full SPOD spectrum for the cavity PIV data and the relative energy content of the leading and higher SPOD modes. Four values of $b_{\text{win}}$ between 1, in figure 8(a), and the maximum of 25, in 8(d), are considered. It is, as expected, observed that the variance of the higher mode spectra decreases in a very similar manner to that of the leading mode. The normalized cumulative energy, that is, the cumulative sum of the eigenvalues from the highest to the lowest at each frequency, is indicated by the color. The number of modes required to capture 90% of the energy at any given frequency corresponds to the number of all colored (non-gray) lines. For the two highest peaks and the lowest bandwidth of $b_{\text{win}} = 1$ shown in 8(a), the first modes alone account for more than 90% of the energy. As the bandwidth is increased,
the peaks flatten, and this number reduces to 60%-80%. At the same time, the number of modes required to reach 90% stabilizes, and the main effect of increasing \( b_{\text{win}} \) is, apart from the discussed effects on bias and variance, that more eigenvalues of lower energy are added to the bottom of the spectrum. This, again, is an indicator of statistical convergence.

4. Discussion

In the light of the above comparative study, we start by assessing the performance of multitaper and multitaper-Welch estimates of the SPOD compared to the standard Welch approach. By comparing SPOD spectra obtained using Welch-only and multitaper-only estimators in figure 3, we have established that both methods are comparable in terms of variance, bias, and mode convergence if the frequency resolution of the Welch estimate is similar to the bandwidth of the multitaper estimate. This leaves the multitaper estimate with the advantage of a higher resolution and the associated ability to resolve lower frequencies. Hybrid multitaper-Welch estimates offer the flexibility to control to a large degree the resolution, bias, and variance. In the limit of unit bandwidth and a single taper, the classical Welch estimator is recovered. This allows us to assess the performance advantage of multitaper-Welch estimates over Welch estimates directly from figures 5-8 that all contain this limiting case. All comparisons show that additional tapers reduce the variance of the basic Welch estimate and lead to better-converged modes. Similar to the findings of Bronez [29] on the estimation of power spectra from time signals, our results suggest that the multitaper-Welch approach generally outperforms the Welch-only approach in terms of variance, bias, or resolution if the other two quantities are comparable. This performance advantage comes at the additional computational cost associated with processing a factor equal to the number of tapers of additional Fourier transforms.

We hence recommend using the multitaper-Welch estimator to compute the SPOD and the following best practice: for given data consisting of \( N_t \) snap-
shots, first select $N_{\text{fft}}$ to obtain the smallest small number of blocks that still ensures that the dynamics at the lowest frequencies are statistically adequately represented. Then choose either the minimum bandwidth necessary to reduce the variance (or increase the mode convergence) to the desired level, or select the maximum bandwidth acceptable in terms of bias (or any value in between the two). In settings where well-separated blocks of data are obtained individually, for example, in carefully designed experiments like those by Citriniti and George [37], multitaper estimates from each block can be leveraged to reduce the number of required blocks, or reduce the variance and improve the mode convergence for an existing set of data.

While SPOD computes modes that optimally represent the data in terms of its second-order statistics, i.e., variance or energy, multitaper estimates are equally applicable to estimate higher-order statistics. They can hence be leveraged to improve estimates of the bispectral mode decomposition [38]. SPOD can quickly become computationally intractable for very large data. This problem can be addressed by updating algorithms that converge the SPOD from streaming data that become available one snapshot at a time. While impossible for the multitaper-only estimator, the method proposed by Schmidt and Towne [39] can readily be extended to accommodate multiple data windows and compute multitaper-Welch estimates on the fly.

The multitaper-Welch estimator using DPSS was implemented in the existing open-source Matlab code SPOD [5]. The second argument of the function can be used to specify $N_{\text{fft}}$ ($N_{\text{FFT}}$) and $b_{\text{win}}$ ($BW$) using the syntax $[L,P,F] = \text{SPOD}(X,[N_{\text{FFT}} BW],\ldots)$. The number of DPSS tapers is then determined by equation (6).

Acknowledgements. OTS gratefully acknowledges support from Office of Naval Research grant N00014-20-1-2311 and NSF grant CBET 2046311. I would like to thank Lou Cattafesta and Yang Zhang for providing the TR-PIV data, Guillaume Brès for generating the turbulent jet database, and Eduardo Martini for the constructive feedback. The TR-PIV data was created with support from
References

1. Lumley JL. Stochastic tools in turbulence. New York: Academic Press; 1970.

2. Glauser MN, Leib SJ, George WK. Coherent structures in the axisymmetric turbulent jet mixing layer. *Turbulent shear flows* 1987;5:134–45.

3. Schmid PJ. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics* 2010;656:5–28. doi:10.1017/S0022112010001217.

4. McKeon BJ, Sharma AS. A critical-layer framework for turbulent pipe flow. *Journal of Fluid Mechanics* 2010;658:336–382. doi:10.1017/S002211201000176X.

5. Towne A, Schmidt OT, Colonius T. Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *Journal of Fluid Mechanics* 2018;847:821–867. doi:10.1017/jfm.2018.283.

6. Schmidt OT, Towne A, Rigas G, Colonius T, Brès GA. Spectral analysis of jet turbulence. *Journal of Fluid Mechanics* 2018;855:953–982. doi:10.1017/jfm.2018.675.

7. Arndt REA, Long DF, Glauser MN. The proper orthogonal decomposition of pressure fluctuations surrounding a turbulent jet. *Journal of Fluid Mechanics* 1997;340:1–33.

8. Citriniti JH, George WK. Reconstruction of the global velocity field in the axisymmetric mixing layer utilizing the proper orthogonal decomposition. *Journal of Fluid Mechanics* 2000;418:137–66.
9. Nidhan S, Chongsiripinyo K, Schmidt OT, Sarkar S. Spectral proper orthogonal decomposition analysis of the turbulent wake of a disk at Re=50000. *Physical Review Fluids* 2020;5(12):124606. doi:10.1103/PhysRevFluids.5.124606.

10. Gordeyev SV, Thomas FO. Coherent structure in the turbulent planar jet. part 1. extraction of proper orthogonal decomposition eigenmodes and their self-similarity. *Journal of Fluid Mechanics* 2000;414:145–94.

11. Gudmundsson K, Colonius T. Instability wave models for the near-field fluctuations of turbulent jets. *Journal of Fluid Mechanics* 2011;689:97–128.

12. Hellström LHO, Smits AJ. The energetic motions in turbulent pipe flow. *Physics of Fluids* 2014;26(12):125102.

13. Tutkun M, George WK. Lumley decomposition of turbulent boundary layer at high reynolds numbers. *Physics of Fluids* 2017;29(2):020707.

14. Araya DB, Colonius T, Dabiri JO. Transition to bluff-body dynamics in the wake of vertical-axis wind turbines. *Journal of Fluid Mechanics* 2017;813:346–81.

15. Abreu LI, Cavalieri AVG, Wolf W. Coherent hydrodynamic waves and trailing-edge noise. In: 23rd AIAA/CEAS aeroacoustics conference; vol. AIAA 2017-3173. 2017;doi:10.2514/6.2017-3173.

16. He X, Fang Z, Rigas G, Vahdati M. Spectral proper orthogonal decomposition of compressor tip leakage flow. *Physics of Fluids* 2021;33(10):105105.

17. Li XB, Chen G, Liang XF, Liu DR, Xiong XH. Research on spectral estimation parameters for application of spectral proper orthogonal decomposition in train wake flows. *Physics of Fluids* 2021;33(12):125103.

18. Schmidt OT, Mengaldo G, Balsamo G, Wedi NP. Spectral empirical orthogonal function analysis of weather and climate data.
19. Sanjose M, Towne A, Jaiswal P, Moreau S, Lele S, Mann A. Modal analysis of the laminar boundary layer instability and tonal noise of an airfoil at reynolds number 150,000. *International Journal of Aeroacoustics* 2019;18(2-3):317–50.

20. Nekkanti A, Schmidt OT. Modal analysis of acoustic directivity in turbulent jets. *AIAA Journal* 2020;0(0):1–12. URL: https://doi.org/10.2514/1.J059425 doi:10.2514/1.J059425 arXiv:https://doi.org/10.2514/1.J059425

21. Baars WJ, Tinney CE. Proper orthogonal decomposition-based spectral higher-order stochastic estimation. *Physics of Fluids* 2014;26(5):055112.

22. Nekkanti A, Schmidt OT. Frequency–time analysis, low-rank reconstruction and denoising of turbulent flows using spod. *Journal of Fluid Mechanics* 2021;926:A26. doi:10.1017/jfm.2021.681

23. Ghate AS, Towne A, Lele SK. Broadband reconstruction of inhomogeneous turbulence using spectral proper orthogonal decomposition and gabor modes. *Journal of Fluid Mechanics* 2020;888:R1. doi:10.1017/jfm.2020.78

24. Chu T, Schmidt OT. A stochastic spod-galerkin model for broadband turbulent flows. *Theoretical and Computational Fluid Dynamics* 2021;URL: https://doi.org/10.1007/s00162-021-0588-6 doi:10.1007/s00162-021-0588-6

25. Towne A. Space-time galerkin projection via spectral proper orthogonal decomposition and resolvent modes. vol. AIAA 2021-1676. 2021:1676.
26. Welch P. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics* 1967;15(2):70–3.

27. Schmidt OT, Colonius T. Guide to spectral proper orthogonal decomposition. *AIAA Journal* 2020;58(3):1023–33. doi:10.2514/1.J058809

28. Thomson DJ. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE* 1982;70(9):1055–96.

29. Bronez TP. On the performance advantage of multitaper spectral analysis. *IEEE Transactions on Signal Processing* 1992;40(12):2941–6.

30. Slepian D, Pollak HO. Prolate spheroidal wave functions, fourier analysis and uncertainty—i. *Bell System Technical Journal* 1961;40(1):43–63.

31. Slepian D. Prolate spheroidal wave functions, fourier analysis, and uncertainty—v: The discrete case. *Bell System Technical Journal* 1978;57(5):1371–430.

32. Brès G, Jordan P, Le Rallic M, Jaunet V, Cavalieri AVG, Towne A, Lele S, Colonius T, Schmidt OT. Importance of the nozzle-exit boundary-layer state in subsonic turbulent jets. *Journal of Fluid Mechanics* 2018;851:83–124. doi:10.1017/jfm.2018.476

33. Zhang Y, Cattafesta L, Ukeiley L. A spectral analysis modal method applied to cavity flow oscillations. In: TSFP11, Southampton, UK. 2019:

34. Zhang Y, Cattafesta L, Ukeiley L. Identification of coherent structures in cavity flows using stochastic estimation and dynamic mode decomposition. In: TSFP10, Chicago, USA. 2017: 3.

35. Rossiter JE. Wind tunnel experiments on the flow over rectangular cavities at subsonic and transonic speeds. *RAE Technical Report No 64037* 1964:

36. Lii KS, Rosenblatt M. Prolate spheroidal spectral estimates. *Statistics & Probability Letters* 2008;78(11):1339–48.
37. Citriniti JH, George WK. Reconstruction of the global velocity field in the axisymmetric mixing layer utilizing the proper orthogonal decomposition. *Journal of Fluid Mechanics* 2000;418:137–66.

38. Schmidt OT. Bispectral mode decomposition of nonlinear flows. *Nonlinear Dynamics* 2020;(102(4)):2479–501. doi:10.1007/s11071-020-06037-z.

39. Schmidt OT, Towne A. An efficient streaming algorithm for spectral proper orthogonal decomposition. *Computer Physics Communications* 2018;URL: http://www.sciencedirect.com/science/article/pii/S0010465518304016 doi:https://doi.org/10.1016/j.cpc.2018.11.009.