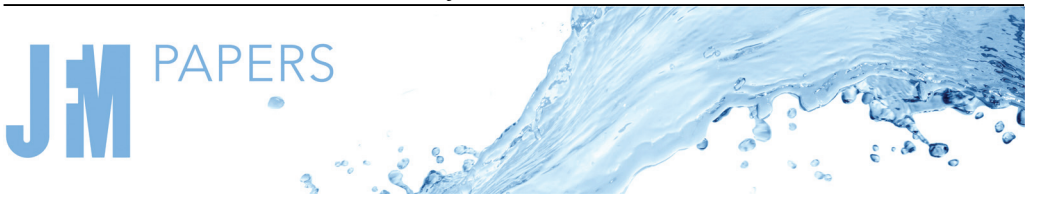
J. Fluid Mech. (2024), *vol*. 998, A3, doi:10.1017/jfm.2024.736



The generalised resolvent-based turbulence estimation with arbitrarily sampled measurements in time

Anjia Ying[1], **Zhigang Li**[1] **and Lin Fu**[1,2,3,4,†]

[1]Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

[2]Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

[3]HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen 518045, PR China

[4]Center for Ocean Research in Hong Kong and Macau (CORE), The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

(Received 21 January 2024; revised 20 July 2024; accepted 6 August 2024)

In this study, the resolvent-based estimation (RBE) is further generalised to cases with arbitrarily sampled measurements in time, where the generalised RBE is denoted as GRBE in this study. Different from the RBE that constructs the transfer function at each frequency, the GRBE minimises the estimation error energy in the physical temporal domain by considering the forcing and noise statistics. The GRBE is validated by estimating the complex Ginzburg–Landau equation and turbulent channel flows with the friction Reynolds number $Re_\tau = 186$, 547 and 934, where the results from the RBE are also included. When the measurements are temporally resolved, the estimation results of the two approaches are equivalent to each other, and both match well with the reference numerical results. For the temporally unresolved cases, the estimation errors from the GRBE are obviously lower than those from the RBE. After validation, the GRBE is applied to investigate the impacts of the abundance of the measured information, including the temporal information and sensor types, on the estimation accuracy. Compared with the mean square error (MSE) in the estimation with temporally resolved measurements, that with measurements at only one snapshot, i.e. without any temporal information, increases by approximately 15 %. On the other hand, it can effectively improve the estimation accuracy by increasing the number of sensor types. With temporally resolved measurements, the relative MSE decreases by 12.3 % when the sensor types increase from

† Email address for correspondence: linfu@ust.hk

$\{\tau_u\}$ to $\{\tau_u, \tau_w, p\}$, where τ_u, τ_w and p are the streamwise shear stress, spanwise shear stress and pressure at the wall. Finally, several existing forcing models are incorporated into the GRBE to investigate their performance in the linear estimation of flow state. The wall-distance-dependent model (W-model) results match well with the optimal linear estimations when the measurements are temporally unresolved. Meanwhile, with the increase of temporal information of the measurement, the estimation errors from the tested W-model and the scale-dependent model (λ-model) both increase, which contradicts the tendency observed in the optimal linear GRBE estimation results. Such a phenomenon highlights the importance of proper modelling of the forcing in the temporal domain for the accuracy of flow state estimation.

Key words: turbulence theory, turbulence modelling

1. Introduction

In experimental and numerical studies, the flow information is usually only partially available due to technical limitations in measuring sensors or the inaccuracy of the numerical models. For numerical studies, the popular wall-modelled large-eddy simulation (WMLES), which resolves only the large-scale motions beyond the local grid scale while modelling the near-wall small-scale motions with the wall models (Larsson *et al.* 2016; Fu *et al.* 2021; Fu, Bose & Moin 2022; Griffin, Fu & Moin 2023), achieves a good balance between the computational cost and accuracy. However, the near-wall dynamics, as an important turbulence property, is usually inaccurately evaluated by the WMLES (Bae *et al.* 2018). On the other hand, the spatial and temporal resolutions of the experimental results are usually restricted by the limitations in the measuring equipment (Nobach & Bodenschatz 2009; Cameron 2011). Current attempts to estimate the missing dynamics of turbulence can be categorised into the data-driven approaches (Marusic, Mathis & Hutchins 2010; Baars, Hutchins & Marusic 2016; Guastoni *et al.* 2021) and the physics-based ones (Chevalier *et al.* 2006; Hwang & Cossu 2010; McKeon & Sharma 2010; Martini *et al.* 2020; Towne, Lozano-Durán & Yang 2020).

The data-driven approaches estimate the flow field with the knowledge obtained from the existing flow data. The well-known attached-eddy model (AEM) (Townsend 1976) and the inner–outer interaction model (IOIM) (Marusic *et al.* 2010; Baars *et al.* 2016) that describe the flow motions in turbulence are proposed based on extensive observation and investigation on the wall-bounded turbulence. According to the IOIM, the large-scale motions (LSMs) and very-large-scale motions (VLSMs) influence the near-wall turbulence through the superposition and modulation effects. Specifically, the superposition effects of the LSMs and VLSMs on near-wall motions can be determined through the transfer kernel estimated from the cross-spectral density (CSD) tensor, which builds up the linear relationship between the flow motions at different wall-normal locations (Gupta *et al.* 2021). It can be demonstrated that the superposition effect defined by the IOIM corresponds to the optimal linear estimation with the knowledge of the spatial cross-spectra of the streamwise velocity. The consistency between the AEM and IOIM has been demonstrated in Cheng & Fu (2022), which is then used to study the isolated attached eddies with a give scale (Cheng, Shyy & Fu 2022). Besides the traditional data-driven approaches that use predefined transfer functions, the deep neural networks and convolutional neural networks (CNNs) are used for the prediction of the flow system in a nonlinear manner (e.g. Lozano-Durán, Bae & Encinar 2020; Guastoni *et al.* 2021). Guastoni *et al.* (2021) predicts the wall-bounded turbulence from the wall

measurements with the CNNs using a linear combination of orthonormal basis functions obtained from proper orthogonal decomposition, which improves the prediction accuracy at large wall-normal distances compared with that from the CNN that directly predicts the turbulent motions.

The physics-driven approaches estimate the flow field based on the Navier–Stokes (NS) equations that govern the flow dynamics. The NS equations can be rearranged into the form of the linearised relationship between the nonlinear forcing (input) and the response of the velocity, pressure and temperature (output) (Hwang & Cossu 2010; McKeon & Sharma 2010). When the linear relationship is built between the forcing and response in the frequency domain, the corresponding linear operator is named the resolvent (McKeon & Sharma 2010). The resolvent analysis usually focuses on uncovering the coherence properties of turbulence through mathematical investigations on the resolvent operator. Through the singular value decomposition of the resolvent operator, the normalised forcing and response modes ordered by their gains are obtained. Specifically, when the gain of the leading mode is much larger than the sequential modes, the corresponding response mode could be considered to dominate the resultant response. The domination of the leading resolvent response mode over the sequential ones is denoted as low-rank behaviour (Pickering *et al.* 2021). Assuming low-rank behaviour, the leading response mode at a given spatio-temporal scale can be regarded as the representative flow pattern at such a scale. Such simplification achieves large success in qualitatively describing the coherent structures of turbulence with only the mean velocity profile obtained from the direct numerical simulation (DNS) (McKeon & Sharma 2010; Moarref *et al.* 2013) or the ordinary differential equation solvers (Hwang & Cossu 2010; Chen *et al.* 2023*a*,*b*). However, the actual response of the flow dynamic system is determined by, not only the response modes and gains, but also the variations of forcing projections on different resolvent forcing modes, which cannot be ignored for the energy-containing flow scales (Morra *et al.* 2019, 2021). Thus, for the flow state estimation in this study, where higher accuracy is needed, the low-rank behaviour is not suitable to directly utilise.

Besides the approaches utilising the low-rank behaviour of the resolvent operator, the modelling of the input forcing is also broadly investigated for the prediction of the flow statistics (Madhusudanan, Illingworth & Marusic 2019; Morra *et al.* 2019; Gupta *et al.* 2021; Ying *et al.* 2023). By modelling a part of the forcing with the eddy-viscosity model (Cess 1958; Reynolds & Hussain 1972) while considering the remaining part of the forcing to be white in space, the predicted accuracy of the energy profile of the response is much improved (Hwang & Cossu 2010; Morra *et al.* 2019). Later, with the cross-spectra predicted by such eddy-viscosity-modelled forcing, Madhusudanan *et al.* (2019) estimate the large-scale structures of the turbulent channel flow via the linear stochastic estimation. Compared with the ones without the eddy-viscosity model, the predicted energy spectra and large-scale structures from the modelled forcing match much better with the DNS results (Madhusudanan *et al.* 2019), which demonstrates the importance of the forcing model in turbulent estimation in cases where the real forcing statistics are unavailable. Gupta *et al.* (2021) further improve the estimation accuracy by introducing additional forcing models that consider the variation of the stochastic forcing profile with heights and flow scales, namely the W-model and λ-model. The W-model and λ-model are demonstrated to further improve the prediction accuracy of the linear coherence spectrum and the large-scale flow structures. Ying *et al.* (2023) propose the resolvent-informed white-noise-based estimation (RWE) approach to model the remaining part of the forcing by modifying the initially white forcing according to the estimated near-wall energy profile based on resolvent-based estimation (RBE) that will be introduced in the following.

With input data of the mean velocity profile and temporally resolved measurement data, the RWE provides an accurate prediction of the turbulence statistics, including the energy spectra, space–time correlation and large-scale structures in the near-wall region.

The above-introduced forcing models focus on the prediction of the flow statistics. On the other hand, it is also very important to estimate the instantaneous flow state with the measured flow state. The predicted flow statistics from the forcing models can be used to inform the estimators of the fluctuation flow state, as will be illustrated in the following. The physics-based approaches for state estimation of the dynamic system include the Kalman-filter-based approaches (Kalman 1960; Hœpffner *et al.* 2005; Chevalier *et al.* 2006; Colburn, Cessna & Bewley 2011), the H_2-optimal control approach (Illingworth, Monty & Marusic 2018) and the resolvent-based approaches (Martini *et al.* 2020; Towne *et al.* 2020; Amaral *et al.* 2021). The Kalman filter (Kalman 1960) includes a set of recursive algorithms that estimate the flow state with the knowledge of the statistics of the stochastic input forcing and measurement noise, which are both assumed to be white in time. To estimate the flow state at the current time step, the causal Kalman filter includes two phases, namely the prediction phase and the update phase. In the prediction phase, the current state is predicted from the state estimation at the previous time step. Then, the prediction is modified by minimising the expected estimation error covariance according to the current noise-containing measurements in the update phase. For non-causal cases, the so-called Kalman smoother further corrects the estimation at a given time step with later measurements. Regardless of whether the stochastic input and measurement noise are Gaussian, the Kalman filter could provide the optimal linear estimation in the minimum mean square error (MSE) sense (Humpherys, Redd & West 2012). When applied in turbulence estimation, it is crucial to appropriately model the temporal evolution of the input forcing to ensure the prediction accuracy of the Kalman filter. The early Kalman-filter-based studies in turbulence estimation (Hœpffner *et al.* 2005; Chevalier *et al.* 2006) assume that the forcing is white in time, which cannot model well the real forcing that is coloured in time (Zare, Jovanović & Georgiou 2017; Zare, Georgiou & Jovanović 2020). Zare *et al.* (2016) propose a stochastic modelling approach that generates a coloured-in-time forcing sequence from the knowledge of the spatial spectra of response along the height, which can potentially enhance the effectiveness of the Kalman filter by introducing that coloured forcing. However, the current stochastic dynamical modelling approach cannot well reproduce the energy distribution of forcing in the frequency domain (Zare *et al.* 2020), which needs further exploration. On the other hand, the H_2-optimal control approach (Doyle *et al.* 1989) constructs a rational controller that minimises the H_2 norm of the transfer matrix that links the external input and the estimation error when this approach is used for causal state estimation. The H_2-optimal control approach assumes that the forcing and measurement noise are both white in time, which is the optimal linear causal estimator when the spatial covariance of the external forcing is known. Illingworth *et al.* (2018) utilise the H_2-optimal control approach in estimating the large-scale structures in turbulent channel flow using the measurements at the logarithmic region.

The RBE (Martini *et al.* 2020; Towne *et al.* 2020; Amaral *et al.* 2021) estimates the flow state in the frequency domain, which can be regarded as the multiple-input and multiple-output (MIMO) Wiener filter (Wiener 1930, 1949). The RBE is initially proposed by Towne *et al.* (2020), which estimates the input forcing of the dynamic system by finding the minimum L^2-norm solution that reproduces the measurements. Here, the L^2 norm is defined as the inner product of a vector. Later, Martini *et al.* (2020) improve the

RBE, which enables the incorporation of the forcing models. The linear estimator of the improved RBE can be constructed by searching for the stationary point of the expected CSD tensor of the estimation error. In this sense, the original RBE by Towne *et al.* (2020) can be regarded as the one that finds the stationary point of the error CSD tensor assuming that the forcing is uniform and uncorrelated in space. With the improved RBE informed by the real forcing statistics, Amaral *et al.* (2021) estimate the flow state of turbulent channel flow with the measurements of wall shear stress and pressure.

Although the RBE has the desirable property of minimising the generic error norm, it has a high requirement on the spatio-temporal resolutions of the measurement. In practice, temporally resolved data could be unavailable due to the high computational and storage costs in numerical studies or the equipment limitations in experiments. For high Reynolds number turbulence, the temporally resolved DNS data could hardly be stored due to the huge storage cost, which hinders the validation and application of the resolvent-based approaches in such cases. Meanwhile, in numerical or experimental studies, some snapshots might become unavailable due to the corruption of the digital storage device or the failure of the measuring equipment. Thus, an estimator that enables arbitrary sampling time intervals is needed for the flow estimations in such cases with unresolved-in-time measurement data. In this study, a generalised version of the RBE will be proposed, which enables physics-based estimation using the measured data with arbitrary time intervals. By incorporating the spatio-temporal forcing statistics or forcing models, the new approach constructs the linear estimator by finding the stationary points of the spatio-temporal covariance tensor of the estimation error. When the real forcing and noise statistics are incorporated into the estimator, the optimal linear estimation can be obtained with the given temporal resolution of the measurements.

The remainder of this article is organised as follows. In § 2, the generalised resolvent-based estimator (GRBE) approach is derived based on the linearised NS equations. In § 3, the newly proposed method is validated for the complex Ginzburg–Landau equation and turbulent channel flows. In § 4, the new approach is utilised to investigate the impacts of different measuring conditions and the incorporation of forcing models on the estimation accuracy. Discussions and concluding remarks are presented in § 5.

2. Methodology

In this section, the linearisation of the incompressible NS equations that govern the dynamics of turbulence is described, followed by the introduction of the resolvent analysis. Then, the derivation and discussions of the newly proposed GRBE will be presented.

2.1. *Linearisation of the incompressible Navier–Stokes equations*

The incompressible NS equations are given by

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = -\boldsymbol{\nabla}p + \frac{1}{Re_\tau}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\boldsymbol{u}^{\mathrm{T}}), \tag{2.1a}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{2.1b}$$

where $Re_\tau = u_\tau h/\nu$ is the friction Reynolds number, u_τ is the friction velocity, h is the half-channel width, ν is the kinetic viscosity and the superscript $^{\mathrm{T}}$ denotes transpose. The linearised NS equations with respect to the fluctuation velocity $\boldsymbol{u}'$ and pressure p' are

expressed by

$$\frac{\partial \boldsymbol{u}'}{\partial t} + \boldsymbol{U}\cdot\nabla\boldsymbol{u}' + \boldsymbol{u}'\cdot\nabla\boldsymbol{U} + \nabla p' - \frac{1}{Re_\tau}\nabla\cdot(\nabla\boldsymbol{u}' + \nabla\boldsymbol{u}'^{\mathrm{T}}) = \boldsymbol{f}_0', \tag{2.2a}$$

$$\nabla\cdot\boldsymbol{u}' = 0, \tag{2.2b}$$

where $'$ denotes the fluctuation variable, U is the mean velocity and the nonlinear forcing term is defined as

$$\boldsymbol{f}_0' = -\left(\boldsymbol{u}'\cdot\nabla\boldsymbol{u}' - \overline{\overline{\boldsymbol{u}'\cdot\nabla\boldsymbol{u}'}}\right) \tag{2.3}$$

with the double overlines denoting the averaged value along the temporal and uniform spatial directions.

On the other hand, the nonlinear term on the right-hand side of (2.2) can be further decomposed into two parts, i.e. the eddy-viscosity terms that are linear functions of $\boldsymbol{u}'$ and the forcing $\boldsymbol{f}'$ (Illingworth *et al.* 2018; Morra *et al.* 2019; Towne *et al.* 2020). The linearised NS equations considering the eddy-viscosity model are expressed by

$$\frac{\partial \boldsymbol{u}'}{\partial t} + \boldsymbol{U}\cdot\nabla\boldsymbol{u}' + \boldsymbol{u}'\cdot\nabla\boldsymbol{U} + \nabla p' - \frac{1}{Re_\tau}\nabla\cdot\left[\frac{\nu_t+\nu}{\nu}(\nabla\boldsymbol{u}' + \nabla\boldsymbol{u}'^{\mathrm{T}})\right] = \boldsymbol{f}', \tag{2.4a}$$

$$\nabla\cdot\boldsymbol{u}' = 0, \tag{2.4b}$$

where the forcing term is defined as

$$\boldsymbol{f}' = \boldsymbol{f}_0' - \frac{1}{Re_\tau}\nabla\cdot\left[\frac{\nu_t}{\nu}(\nabla\boldsymbol{u}' + \nabla\boldsymbol{u}'^{\mathrm{T}})\right]. \tag{2.5}$$

The eddy viscosity ν_t can be calculated from the semi-empirical expression of Cess (1958) as reported by Reynolds & Hussain (1972)

$$\nu_t = \frac{\nu}{2}\left\{1 + \frac{\kappa^2 Re_\tau^2}{9}(2\tilde{y} - \tilde{y}^2)^2(3 - 4\tilde{y} + 2\tilde{y}^2)^2\left[1 - \exp\left(\frac{-Re_\tau(1 - |1-\tilde{y}|)}{A}\right)\right]^2\right\}^{1/2} - \frac{\nu}{2}, \tag{2.6}$$

where $\tilde{y} = y/h$, the constants $\kappa = 0.426$ and $A = 25.4$. For the flow state estimation problems that are focused on in this study, the inclusion of the eddy-viscosity terms provides improved results (Illingworth *et al.* 2018; Amaral *et al.* 2021). Thus, we will also adopt the eddy-viscosity-modelled linearised formulations (2.4) in this study.

Taking the Fourier transformation to (2.4) in the homogeneous spatial directions, the linearised NS equations in each spatial scale $\boldsymbol{k}_s$ are obtained. For instance, in the fully developed turbulent channel flow, the Fourier transformation is taken in the streamwise (x) and spanwise (z) directions, where $\boldsymbol{k}_s = (k_x, k_z)$. The linearised NS equations are written in a discretised state-space form with N_y points in the wall-normal direction as (Luhar, Sharma & McKeon 2014*a*,*b*; Amaral *et al.* 2021)

$$\boldsymbol{M}\frac{\partial \hat{\boldsymbol{q}}_{\boldsymbol{k}_s}(t)}{\partial t} = \boldsymbol{A}_{\boldsymbol{k}_s}\hat{\boldsymbol{q}}_{\boldsymbol{k}_s}(t) + \boldsymbol{B}\hat{\boldsymbol{f}}_{\boldsymbol{k}_s}(t), \tag{2.7a}$$

$$\hat{\boldsymbol{m}}_{\boldsymbol{k}_s}(t) = \boldsymbol{C}\hat{\boldsymbol{q}}_{\boldsymbol{k}_s}(t) + \hat{\boldsymbol{n}}_{\boldsymbol{k}_s}(t), \tag{2.7b}$$

where $\hat{}$ denotes the variable after Fourier transformation, the state $\hat{\boldsymbol{q}}_{\boldsymbol{k}_s}(t) = [\hat{u}, \hat{v}, \hat{w}, \hat{p}]^{\mathrm{T}}$ and the operators are defined as

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{I}_{3N_y} & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{A}_{\boldsymbol{k}_s} = \begin{bmatrix} \mathcal{L}_{\boldsymbol{k}_s} & -\nabla^{\mathrm{T}} \\ \nabla & 0 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{I}_{3N_y} \\ 0 \end{bmatrix}, \tag{2.8a–c}$$

where $\boldsymbol{C}$ is the spatial observation tensor, $\hat{\boldsymbol{n}}_{\boldsymbol{k}_s}(t)$ is the measurement noise, ∇ is the Hamiltonian operator and $\mathcal{L}_{\boldsymbol{k}_s}$ is the spatial linearised NS operator. The expressions of $\mathcal{L}_{\boldsymbol{k}_s}$, ∇ and $\boldsymbol{C}$ are provided in Appendix A. Further taking Fourier transformations along the statistically homogeneous temporal direction, the linearised NS equations at each spatio-temporal scale $\boldsymbol{k} = (\boldsymbol{k}_s, \omega)$ are obtained, as expressed by

$$\hat{\boldsymbol{q}}_{\boldsymbol{k}} = \boldsymbol{H}_{\boldsymbol{k}} \cdot \hat{\boldsymbol{f}}_{\boldsymbol{k}}, \tag{2.9}$$

where $\boldsymbol{H}_{\boldsymbol{k}}$ is the resolvent operator that is expressed as

$$\boldsymbol{H}_{\boldsymbol{k}} = \left(-\mathrm{i}\omega \boldsymbol{M} - \boldsymbol{A}_{\boldsymbol{k}_s}\right)^{-1}, \tag{2.10}$$

with $\mathrm{i} = \sqrt{-1}$. From (2.9), $\hat{\boldsymbol{q}}_{\boldsymbol{k}}$ is regarded as the response of the input forcing $\hat{\boldsymbol{f}}_{\boldsymbol{k}}$ through the resolvent operator $\boldsymbol{H}_{\boldsymbol{k}}$. Taking singular value decomposition of the transfer function $\boldsymbol{H}_{\boldsymbol{k}}$ yields the resolvent input and output modes ordered by the gains, where the low-order modes can be utilised to investigate the predominant coherent structures assuming the low-rank behaviour, see, e.g. Moarref *et al.* (2013) and Sharma & McKeon (2013). Meanwhile, only the linearised relationship (2.9) built by the resolvent analysis will be utilised in this study.

Based on the resolvent formulation (2.9), the RBE that estimates the flow state from limited measurements was proposed by Towne *et al.* (2020), and is further improved (Martini *et al.* 2020) by considering the forcing statistics. In the following, the RBE approach is further generalised to cases where the measurements are arbitrarily sampled.

2.2. *Generalised RBE in cases with arbitrary temporal measurements*

In this section, the case where the information of the turbulence field is partially available from measurements that are noise contaminated and arbitrarily sampled in time is considered. Before the estimation, Fourier transformations are conducted in statistically uniform spatial directions of the flow field, such as the streamwise and spanwise directions in fully developed turbulent channel flows, leading to the decomposed measurements at separated spatial scales $\boldsymbol{k}_s$. Meanwhile, since the flow signals can be unresolved in time, the Fourier transformation will not be conducted along the temporal direction. In this section, the subscript $_{\boldsymbol{k}_s}$ denoting the spatial scale of a flow quantity is omitted for brevity.

Consider the MIMO estimation case where the measured data are sampled at $N_{\check{T}}$ time instances with inadequate resolution. Meanwhile, to fully resolve the temporal information, N_T snapshots are actually needed ($N_T > N_{\check{T}}$). The amounts of response and observation of different types or at different locations at a given spatial scale $\boldsymbol{k}_s$ are denoted as N_q and N_m, respectively. Without loss of generality, a sequence of measured data $m(j)$, where $j \in [1, N_m N_{\check{T}}]$, is taken into consideration. The linear estimation of the flow state $q(i)$, where $i \in [1, N_q N_{\check{T}}]$ is based on all the available measurements $m(j)$, can be expressed by the convolution operation as

$$q_e(i) = \sum_{j=1}^{N_m N_{\check{T}}} \hbar(i,j) m(j), \tag{2.11}$$

where the subscript $_e$ denotes the estimated quantity and $\hbar(i,j)$ is the convolution kernel. Equation (2.11) can also be expressed in matrix form by

$$\boldsymbol{q}_e = \boldsymbol{h} \cdot \boldsymbol{m}, \tag{2.12}$$

where $\boldsymbol{h}$ is a two-dimensional matrix consisting of $\hbar(i,j)$ at the (i,j) location.

Denoting the estimation error $\varepsilon(i) = q(i) - q_e(i)$, the cost function $\mathcal{J}(i)$ is defined as the MSE that is expressed by

$$\mathcal{J}(i) = \left\langle |\varepsilon(i)|^2 \right\rangle = \left\langle \left[q(i) - \sum_{j=1}^{N_m N_{\check{T}}} \hbar(i,j) m(j) \right] \cdot \overline{\left[q(i) - \sum_{k=1}^{N_m N_{\check{T}}} \hbar(i,k) m(k) \right]} \right\rangle, \tag{2.13}$$

where the overline denotes the complex conjugate and $\langle\rangle$ denotes ensemble average. To minimise the MSE of the estimated signal with optimised transfer kernel $\hbar(i,j) = \hbar^R(i,j) + \mathrm{i}\hbar^I(i,j)$, the partial derivatives of the cost function $\mathcal{J}(i)$ to any $\hbar^R(i,j)$ or $\hbar^I(i,j)$ with $j \in [1, N_m N_{\check{T}}]$ should be zero, i.e.

$$\left.\begin{aligned} \frac{\partial \mathcal{J}(i)}{\partial \hbar^{R(I)}(i,j)} &= \left\langle -2 \left[q(i) - \sum_{j=1}^{N_m N_{\check{T}}} \hbar(i,k) m(k) \right] \overline{m(j)} \right\rangle_{R(I)} = 0, \\ \Rightarrow \quad \left\langle q(i)\overline{m(j)} \right\rangle &- \sum_{j=1}^{N_m N_{\check{T}}} \hbar(i,k) \left\langle m(k)\overline{m(j)} \right\rangle_{R(I)} = 0. \end{aligned}\right\} \tag{2.14}$$

The optimal transfer kernel can be obtained by solving the system of linear equations (2.14). The solution can be expressed as

$$\boldsymbol{h} = \left\langle \boldsymbol{q}\boldsymbol{m}^* \right\rangle \cdot \left\langle \boldsymbol{m}\boldsymbol{m}^* \right\rangle^{-1}, \tag{2.15}$$

where the superscript * denotes the Hermitian transpose.

Equation (2.15) provides the expression of the optimal linear transfer kernel in terms of the covariance matrices of the flow state and measured signal. However, since such spatio-temporal covariance matrices are not directly available, the expressions of the optimal transfer kernel as a function of the forcing CSD tensors, which can be provided by the forcing models, should be further derived.

Considering the resolvent formulation, the expressions of $\boldsymbol{q}$ and $\boldsymbol{m}$ in the physical space as functions of the forcing $\hat{\boldsymbol{f}}_j$ at ω_j, where $j \in [1, N_T]$, in the Fourier space are expressed as

$$\boldsymbol{q} = \frac{1}{N_T} \sum_{j=1}^{N_T} \check{\zeta}_{\boldsymbol{q},j} \cdot \hat{\boldsymbol{q}}_j = \frac{1}{N_T} \sum_{j=1}^{N_T} \check{\zeta}_{\boldsymbol{q},j} \cdot \left(\boldsymbol{H}_j \cdot \hat{\boldsymbol{f}}_j \right), \tag{2.16a}$$

$$\boldsymbol{m} = \frac{1}{N_T} \sum_{j=1}^{N_T} \check{\zeta}_{\boldsymbol{m},j} \cdot \hat{\boldsymbol{m}}_j = \frac{1}{N_T} \sum_{j=1}^{N_T} \check{\zeta}_{\boldsymbol{m},j} \cdot \left(\boldsymbol{C} \cdot \boldsymbol{H}_j \cdot \hat{\boldsymbol{f}}_j + \hat{\boldsymbol{n}}_j \right), \tag{2.16b}$$

where $\hat{\boldsymbol{q}}_j$, $\hat{\boldsymbol{m}}_j$ and $\hat{\boldsymbol{n}}_j$ are the Fourier transforms of response, measurement and sensor noise at ω_j. Here, $\check{\zeta}_{\boldsymbol{q}(\boldsymbol{m}),j} = \mathrm{diag}[\overbrace{\check{\xi}_j, \ldots, \check{\xi}_j}^{N_{q(m)}}]$, $\check{\xi}_j = \mathrm{e}^{\mathrm{i}\omega_j \check{\boldsymbol{t}}}$, $\check{\boldsymbol{t}} \in \mathbb{R}^{N_{\check{T}}}$ is the vector consisting of the sampled time instances.

Algorithm 1 GRBE

Input: the CSD tensors $\boldsymbol{S}_{\boldsymbol{ff}}$, $\boldsymbol{S}_{\boldsymbol{nn}}$ and the resolvent operator $\boldsymbol{H}_j$ in a wide-enough frequency range such that the spectrum aliasing is negligible.

1: Compute the Fourier operators $\check{\boldsymbol{\zeta}}_{\boldsymbol{q}}$ and $\check{\boldsymbol{\zeta}}_{\boldsymbol{m}}$ according to the sensor types and sampling time instances $\check{\boldsymbol{t}}$.

2: Compute the transfer kernel $\boldsymbol{h}$ with (2.18).

3: Obtain the estimations $\boldsymbol{q}_e$ by convolving the measurements with (2.12).

Output: the estimated flow state $\boldsymbol{q}_e$.

Using (2.16*a*,*b*), the spatio-temporal covariance matrices can be predicted with the forcing models by

$$\langle \boldsymbol{q}\boldsymbol{m}^* \rangle_p = \sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{q},j} \boldsymbol{S}_{\boldsymbol{qm},j} \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j} = \sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{q},j} \boldsymbol{H}_j \boldsymbol{S}_{\boldsymbol{ff},j} \boldsymbol{H}^*_{\boldsymbol{m},j} \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j}, \tag{2.17a}$$

$$\langle \boldsymbol{m}\boldsymbol{m}^* \rangle_p = \sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{m},j} \boldsymbol{S}_{\boldsymbol{mm},j} \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j} = \sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{m},j} \left(\boldsymbol{H}_{\boldsymbol{m},j} \boldsymbol{S}_{\boldsymbol{ff},j} \boldsymbol{H}^*_{\boldsymbol{m},j} + \boldsymbol{S}_{\boldsymbol{nn},j} \right) \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j}, \tag{2.17b}$$

where $\boldsymbol{S}_{\boldsymbol{qm},j} = \langle \hat{\boldsymbol{q}}_j \hat{\boldsymbol{m}}^*_j \rangle$, $\boldsymbol{S}_{\boldsymbol{mm},j} = \langle \hat{\boldsymbol{m}}_j \hat{\boldsymbol{m}}^*_j \rangle$, $\boldsymbol{S}_{\boldsymbol{ff},j} = \langle \hat{\boldsymbol{f}}_j \hat{\boldsymbol{f}}^*_j \rangle$ and $\boldsymbol{S}_{\boldsymbol{nn},j} = \langle \hat{\boldsymbol{n}}_j \hat{\boldsymbol{n}}^*_j \rangle$. Substituting (2.17*a*,*b*) into (2.15), the optimal transfer kernel that considers the forcing and noise statistics is expressed by

$$\boldsymbol{h} = \langle \boldsymbol{q}\boldsymbol{m}^* \rangle_p \cdot \langle \boldsymbol{m}\boldsymbol{m}^* \rangle_p^{-1} \tag{2.18a}$$

$$= \left(\sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{q},j} \boldsymbol{S}_{\boldsymbol{qm},j} \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j} \right) \cdot \left(\sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{m},j} \boldsymbol{S}_{\boldsymbol{mm},j} \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j} \right)^{-1} \tag{2.18b}$$

$$= \left[\sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{q},j} \boldsymbol{H}_j \boldsymbol{S}_{\boldsymbol{ff},j} \boldsymbol{H}^*_{\boldsymbol{m},j} \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j} \right] \cdot \left[\sum_j \check{\boldsymbol{\zeta}}_{\boldsymbol{m},j} \left(\boldsymbol{H}_{\boldsymbol{m},j} \boldsymbol{S}_{\boldsymbol{ff},j} \boldsymbol{H}^*_{\boldsymbol{m},j} + \boldsymbol{S}_{\boldsymbol{nn},j} \right) \check{\boldsymbol{\zeta}}^*_{\boldsymbol{m},j} \right]^{-1}. \tag{2.18c}$$

From the above derivations, the transfer kernel $\boldsymbol{h}$ is the optimal linear estimator of the response $\boldsymbol{q}$ when the real statistics $\boldsymbol{S}_{\boldsymbol{ff},j}$ and $\boldsymbol{S}_{\boldsymbol{nn},j}$ are incorporated, which can be obtained from another existing database. More importantly, the transfer kernel defined in (2.18) also constructs an open framework that enables the incorporation of the spatio-temporal forcing models that predict the forcing CSD tensors all over the frequency domain at each spatial scale $\boldsymbol{k}_s$ when the real statistics of forcing are unknown. The current models that provide spatio-temporal prediction of the forcing include the baseline model (B-model) (Hwang & Cossu 2010; Madhusudanan *et al.* 2019), W-model and λ-model (Gupta *et al.* 2021), which assume that the forcing is white in time and uncorrelated in space with certain forms of spatial energy distributions along the wall-normal height. These forcing models are reviewed in Appendix B. The new approach generalises the RBE to the cases with arbitrary sampled measurements in time, and thus it is denoted as the GRBE in this study. The procedure of the GRBE is provided in Algorithm 1.

2.3. *Expected estimation error from the GRBE*

In § 2.2, the theoretically optimal transfer kernel of arbitrarily sampled measurements is derived in (2.18). On the other hand, the expected statistics of the estimation error can also

be theoretically derived. Recall (2.13) that describes the linear relationship between the observations and estimations, the matrix form holds

$$\begin{aligned}\langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{*}\rangle &= \langle (\boldsymbol{q}-\boldsymbol{h}\cdot\boldsymbol{m})\cdot(\boldsymbol{q}-\boldsymbol{h}\cdot\boldsymbol{m})^{*}\rangle \\ &= \langle \boldsymbol{q}\boldsymbol{q}^{*}\rangle + \boldsymbol{h}\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle\cdot\boldsymbol{h}^{*} - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\boldsymbol{h}^{*} - \boldsymbol{h}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle . \end{aligned} \tag{2.19}$$

When the actual forcing and noise statistics are incorporated into the transfer kernel $\boldsymbol{h}$, the minimum expected estimation error can be obtained, as calculated by

$$\begin{aligned}\langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{*}\rangle_{OPT} &= \langle \boldsymbol{q}\boldsymbol{q}^{*}\rangle + \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle \\ &\quad - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle \\ &= \langle \boldsymbol{q}\boldsymbol{q}^{*}\rangle - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1}\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle , \end{aligned} \tag{2.20}$$

where the subscript '$_{OPT}$' denotes the variable from the optimal linear estimation. Considering that $\langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1}\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle = \langle \boldsymbol{q}_e\boldsymbol{q}_e^{*}\rangle$ according to (2.12) and (2.18*a*), (2.20) indicates the orthogonality principle for the optimal linear estimation, i.e. $\langle \boldsymbol{q}_e\cdot\boldsymbol{\varepsilon}^{*}\rangle = \boldsymbol{O}$. Considering (2.16), the covariance matrix of the estimated error is further expressed by

$$\begin{aligned}\langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{*}\rangle_{OPT} &= \left[\sum_j \check{\boldsymbol{\zeta}}_{q,j}\boldsymbol{H}_j\boldsymbol{S}_{ff,j}\boldsymbol{H}_j^{*}\check{\boldsymbol{\zeta}}_{q,j}^{*}\right] - \left[\sum_j \check{\boldsymbol{\zeta}}_{q,j}\boldsymbol{H}_j\boldsymbol{S}_{ff,j}\boldsymbol{H}_{m,j}^{*}\check{\boldsymbol{\zeta}}_{m,j}^{*}\right] \\ &\quad\cdot\left[\sum_j \check{\boldsymbol{\zeta}}_{m,j}\left(\boldsymbol{H}_{m,j}\boldsymbol{S}_{ff,j}\boldsymbol{H}_{m,j}^{*}+\boldsymbol{S}_{nn,j}\right)\check{\boldsymbol{\zeta}}_{m,j}^{*}\right]^{-1}\cdot\left[\sum_j \check{\boldsymbol{\zeta}}_{q,j}\boldsymbol{H}_j\boldsymbol{S}_{ff,j}\boldsymbol{H}_{m,j}^{*}\check{\boldsymbol{\zeta}}_{m,j}^{*}\right]. \end{aligned} \tag{2.21}$$

On the other hand, when the forcing models rather than the real flow statistics are incorporated in the GRBE estimator, the orthogonal principle as in (2.20) for the optimal linear estimation no longer holds. Instead, the expected error energy from the modelled GRBE estimator should be directly derived from (2.19) by considering (2.18*a*), i.e.

$$\begin{aligned}\langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{*}\rangle_{MOD} &= \langle \boldsymbol{q}\boldsymbol{q}^{*}\rangle + \boldsymbol{h}_{q,MOD}\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle\cdot\boldsymbol{h}_{q,MOD}^{*} - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\boldsymbol{h}_{q,MOD}^{*} - \boldsymbol{h}_{q,MOD}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle \\ &= \langle \boldsymbol{q}\boldsymbol{q}^{*}\rangle + \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle_p\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle_p^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle_p^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle_p \\ &\quad - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle_p^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle_p - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle_p\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle_p^{-1}\cdot\langle \boldsymbol{m}\boldsymbol{q}^{*}\rangle , \end{aligned} \tag{2.22}$$

where $\langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle_p$ and $\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle_p$ are defined in (2.17) and the subscript '$_{MOD}$' denotes the variable corresponding to the linear estimation informed by the modelled forcing. Substituting the estimation error from optimal linear estimation in (2.20) to (2.22), the estimation error from the modelled forcing can be expressed as

$$\begin{aligned}\langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{*}\rangle_{MOD} &= \langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{*}\rangle_{OPT} + \left[\langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1} - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle_p\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle_p^{-1}\right] \\ &\quad\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle\cdot\left[\langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle^{-1} - \langle \boldsymbol{q}\boldsymbol{m}^{*}\rangle_p\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle_p^{-1}\right]^{*} \\ &= \langle \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{*}\rangle_{OPT} + \left(\boldsymbol{h}_{q,OPT}-\boldsymbol{h}_{q,MOD}\right)\cdot\langle \boldsymbol{m}\boldsymbol{m}^{*}\rangle\cdot\left(\boldsymbol{h}_{q,OPT}-\boldsymbol{h}_{q,MOD}\right)^{*}, \end{aligned} \tag{2.23}$$

where $\boldsymbol{h}_{\boldsymbol{q},OPT} = \langle \boldsymbol{q}\boldsymbol{m}^* \rangle \cdot \langle \boldsymbol{m}\boldsymbol{m}^* \rangle^{-1}$. On the right-hand side of (2.23), the extra term $[(\boldsymbol{h}_{\boldsymbol{q},OPT} - \boldsymbol{h}_{\boldsymbol{q},MOD}) \cdot \langle \boldsymbol{m}\boldsymbol{m}^* \rangle \cdot (\boldsymbol{h}_{\boldsymbol{q},OPT} - \boldsymbol{h}_{\boldsymbol{q},MOD})^*]$ is positive semi-definite. Thus, the estimation error energy from the modelled forcing will be minimised when $\boldsymbol{h}_{\boldsymbol{q},MOD} = \boldsymbol{h}_{\boldsymbol{q},OPT}$.

2.4. *Discussions on the relationship between the GRBE and RBE*

The GRBE fundamentally differs from the RBE (Martini *et al.* 2020) in their different optimisation targets when constructing the transfer kernel. The RBE minimises the expected error energy at each frequency, while the GRBE directly minimises the expected error energy in physical space. Although these two approaches provide the same results when the measurements are temporally resolved, their results are different when the measurement information is unresolved in time. We will elaborate the discussions between the GRBE and RBE in the following.

2.4.1. *Temporally resolved measurements*

The RBE transfer function that minimises the expected error energy at a given frequency is given by (Martini *et al.* 2020)

$$\hat{\boldsymbol{T}} = \boldsymbol{H}\boldsymbol{S}_{\boldsymbol{f}\boldsymbol{f}}\boldsymbol{H}_{\boldsymbol{m}}^* \left(\boldsymbol{H}_{\boldsymbol{m}}\boldsymbol{S}_{\boldsymbol{f}\boldsymbol{f}}\boldsymbol{H}_{\boldsymbol{m}}^* + \boldsymbol{S}_{\boldsymbol{n}\boldsymbol{n}}\right)^{-1}. \tag{2.24}$$

The estimated $\hat{\boldsymbol{q}}_e$ is obtained with

$$\hat{\boldsymbol{q}}_{e,j} = \hat{\boldsymbol{T}}_j \cdot \hat{\boldsymbol{m}}_j. \tag{2.25}$$

When the measurements are temporally resolved, the RBE approach can be applied in the time domain through inverse Fourier transform (Martini *et al.* 2020; Amaral *et al.* 2021) i.e.

$$\boldsymbol{h}^{RBE}(t,\tau) = \mathcal{F}^{-1}\left(\hat{\boldsymbol{T}}(\omega)\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\boldsymbol{T}}(\omega)\exp(\mathrm{i}\omega(t-\tau))\,\mathrm{d}\omega, \tag{2.26}$$

and the flow state $\boldsymbol{q}$ can be estimated by performing a convolution of the transfer kernel $\boldsymbol{h}^{RBE}$ and the measurements $\boldsymbol{m}$,

$$\boldsymbol{q}_e(t) = \frac{1}{T}\int_{-\infty}^{\infty} \boldsymbol{h}^{RBE}(t,\tau)\boldsymbol{m}(\tau)\,\mathrm{d}\tau. \tag{2.27}$$

Given temporally resolved information, the discrete form of the convolution operation of the RBE is expressed by

$$\boldsymbol{q}_e = \boldsymbol{h}^{RBE} \cdot \boldsymbol{m}, \tag{2.28}$$

with

$$\begin{aligned}\boldsymbol{h}^{RBE} &= \sum_j \zeta_{\boldsymbol{q},j}\hat{\boldsymbol{T}}(\omega_j)\zeta_{\boldsymbol{m},j}^* \\ &= \sum_j \zeta_{\boldsymbol{q},j}\boldsymbol{H}\boldsymbol{S}_{\boldsymbol{f}\boldsymbol{f}}\boldsymbol{H}_{\boldsymbol{m}}^* \left(\boldsymbol{H}_{\boldsymbol{m}}\boldsymbol{S}_{\boldsymbol{f}\boldsymbol{f}}\boldsymbol{H}_{\boldsymbol{m}}^* + \boldsymbol{S}_{\boldsymbol{n}\boldsymbol{n}}\right)^{-1}\zeta_{\boldsymbol{m},j}^*,\end{aligned} \tag{2.29}$$

where $\zeta_{\boldsymbol{q}(\boldsymbol{m}),j} = \mathrm{diag}[\overbrace{\xi_j,\ldots,\xi_j}^{N_{q(m)}}]$, $\xi_j = \mathrm{e}^{\mathrm{i}\omega_j \boldsymbol{t}}$, $\boldsymbol{t} \in \mathbb{R}^{N_T}$ is the vector consisting of fine enough time steps that are evenly distributed. Comparing the RBE transfer kernel (2.29) with the

Algorithm 2 RBE-F

Input: the CSD tensors $\boldsymbol{S}_{ff}$, $\boldsymbol{S}_{nn}$ and the resolvent operator $\boldsymbol{H}_j$ in the frequency range resolved by the sampled data.

1: Obtain $\hat{\boldsymbol{m}}_j$ in the frequency domain through DFT to the measurements in time.
2: Compute the transfer function $\hat{\boldsymbol{T}}_j$ with (2.24).
3: Obtain the estimations $\hat{\boldsymbol{q}}_e$ with (2.25).
4: Convert the estimated flow state to the physical space with inverse DFT.

Output: the estimated flow state $\boldsymbol{q}_e$.

GRBE one (2.18*c*), it can be demonstrated that these two transfer kernels are the same only when $\check{\zeta}_{\boldsymbol{q}(\boldsymbol{m}),j} = \zeta_{\boldsymbol{q}(\boldsymbol{m}),j}$.

It should be noted that the above relationship between the GRBE and RBE is only valid when the measurement information is resolved in time. In the other cases, these two approaches are different due to their essential difference in constructing the transfer kernel. In the next subsection, the implementations of the RBE in temporally unresolved cases will be provided to facilitate further comparisons between the GRBE and RBE.

2.4.2. *Implementations of the RBE in temporally unresolved cases*

When the sampling time steps are not fine enough to resolve the flow data, the convolution operation (2.27) of RBE cannot be well approximated by the discrete counterpart (2.28) with the temporally unresolved measurements, even if the spatio-temporal statistics of the flow state are available. Despite this, if the RBE is to be deployed in such case, there are two different choices of directly applying the RBE formulation in the frequency or temporal domain, i.e.

(i) To conduct the discrete Fourier transformation (DFT) of the measurements with the sparse-in-time signal, then estimate the flow state via (2.25) in the frequency domain. Since this approach estimates the flow state with a discrete-in-frequency resolvent-based transfer function, it is denoted as RBE-F in the following section. The procedure of RBE-F is provided in Algorithm 2.

(ii) To deploy the discrete RBE convolution formulation (2.28) in estimating the sparse-in-time flow state. To discretise the RBE transfer kernel in sparse time instances, it is calculated by

$$\boldsymbol{h}^{RBE}(i,j) = \int_{\Delta t_{j-}}^{\Delta t_{j+}} w(\tau)\boldsymbol{h}^{RBE}(t_i, t_j+\tau)\,\mathrm{d}\tau, \tag{2.30}$$

where $\Delta t_{j-} = (t_{j-1} - t_j)/2$, $\Delta t_{j+} = (t_{j+1} - t_j)/2$ and $w(\tau)$ is the weighting function for integration, which is specified as $w(\tau) = 1$ in this study. Since this approach utilises a discrete-in-time resolvent-based transfer kernel, it is denoted as RBE-T in the following section. The procedure of the RBE-T is provided in Algorithm 3.

From the above discussions, both RBE-F and RBE-T, although seemingly straightforward, utilise different assumptions when applied in temporally unresolved cases. Investigations on the sources of estimation errors of such two variants of the RBE are provided in Appendix C. For the validations of the GRBE in § 3, only the RBE-T approach will be included, since RBE-T can be directly applied in the cases with uneven samplings in time.

Besides the non-causal estimation applied in post-processing of the measured data, the causal estimation that utilises only the measurements up to the current time step in

Algorithm 3 RBE-T

1: Initialise: prepare $\boldsymbol{S_{ff}}$, $\boldsymbol{S_{nn}}$ and the resolvent operator $\boldsymbol{H}_j$ in a wide-enough frequency range such that the spectrum aliasing is negligible.
2: Compute the Fourier operators $\check{\zeta}_{\boldsymbol{q}}$ and $\check{\zeta}_{\boldsymbol{m}}$ according to the sensor types and sampling time instances $\check{\boldsymbol{t}}$.
3: Compute the transfer function $\hat{\boldsymbol{T}}_j$ with (2.24).
4: Compute the transfer kernel $\boldsymbol{h}^{RBE}$ with (2.30).
5: Obtain the estimations $\boldsymbol{q}_e$ by convolving the measurements with (2.28).

Output: the estimated flow state $\boldsymbol{q}_e$.

real-time estimation is also an interesting topic. Since the GRBE does not impose any restrictions on the temporal resolution of the measurements, it is naturally capable of providing an optimal causal estimation of the flow state in real-time estimation when informed by the actual forcing colour. On the other hand, as described in Martini *et al.* (2020), the RBE can also be applied for causal estimation by truncating the resolvent-based kernel (2.26) by setting $\boldsymbol{h}^{RBE}(t, \tau) = 0$ when $\tau > t$, which is essentially different from the causal GRBE approach. The discussions on the RBE and GRBE results in causal estimation are provided in § 3.1.2.

3. Validations

To validate the new estimation framework, two cases, i.e. the complex Ginzburg–Landau equation and turbulent channel flows with friction Reynolds numbers $Re_\tau = 186$, 547 and 934, will be used in this section. The RBE (Martini *et al.* 2020; Towne *et al.* 2020) and the H_2-optimal estimation (Illingworth *et al.* 2018) are also included in this section for comparison. To extend the RBE in temporally unresolved cases, the RBE-T defined in § 2.4.2 is utilised. For brevity, we will still use the term 'RBE' to refer to the RBE-T approach in this section unless otherwise specified. For reference, the details of the H_2-optimal estimation method are provided in Appendix D. In all the cases tested in this section, the measurement noise is neglected for brevity.

3.1. *The complex Ginzburg–Landau equation*

The complex Ginzburg–Landau equation is a one-dimensional model that mimics a real turbulence system (Bagheri, Brandt & Henningson 2009; Towne, Schmidt & Colonius 2018; Towne *et al.* 2020) and facilitates the analysis of the properties of the linear operator. The equation can be written in the same form as (2.7*a*), with the linear operators $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{M}$ expressed by

$$\left.\begin{aligned} \boldsymbol{A} &= -\nu\frac{\partial}{\partial x} + \gamma\frac{\partial^2}{\partial x^2} + \mu(x), \\ \boldsymbol{M} &= \boldsymbol{B} = \boldsymbol{I}. \end{aligned}\right\} \tag{3.1}$$

The parameters in (3.1) are set as $\mu(x) = (\mu_0 - c_u^2) + \mu_2 x^2/2$, $\nu = 2 + 0.2\mathrm{i}$, $\gamma = 1 - \mathrm{i}$, $\nu_0 = 0.3$, $c_u = 0.2$ and $\mu_2 = -0.01$, which are identical to those in Towne *et al.* (2018). The spatial domain is discretised with second-order upwind scheme at evenly distributed grid nodes with $N_x = 221$ in a computational domain $x \in [-85, 85]$, where the grid size $\Delta x = 0.77$. Following Towne *et al.* (2018), the input forcing that excites the system is realised by setting the value at each discrete time step (with a constant step size of 0.05)

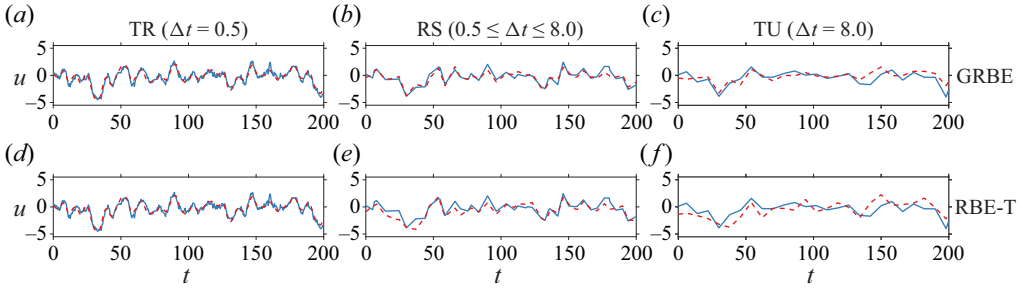


Figure 1. Estimation of the flow states at $x = 10$ using the measurements at $x = 0$ with the GRBE (*a–c*) and RBE (*d–f*) in cases TR (*a*,*d*), RS (*b*,*e*) and TU (*c*,*f*). The solid and dashed curves denote the reference numerical results and the estimated ones, respectively.

and grid node to be a random complex number with uniformly distributed phase from 0 to 2π and normally distributed amplitude with unit variance. The forcing is then low-pass filtered using a tenth-order finite-impulse-response filter with a cutoff frequency equal to 60 % of the Nyquist frequency ω_s equal to 20π. The spatial scope of forcing is limited in the interior of the computational domain with an exponential envelope of $\exp[-(x/L)^p]$, where $L = 60$ and $p = 10$. The equations are integrated in time with the Crank–Nicolson scheme (Crank & Nicolson 1947). After the computational result achieves a statistical stationary state, the flow data at the measurement location are sampled for estimation.

3.1.1. *Non-causal estimation results*

In this section, three cases in terms of different sampling time intervals are considered in this section, namely the temporally resolved case (TR), the randomly sampled case (RS) and the temporally unresolved case (TU), with $\Delta t = 0.5$, $\Delta t = \mathrm{rand}(0.5, 8.0)$ that is uniformly distributed between 0.5 and 8.0, and $\Delta t = 8.0$, respectively. The same sampling period of $T = 200$ is applied for each case. To construct the estimator of the GRBE in (2.18) and RBE in (2.24), the incorporated CSD tensor of the forcing is set to be $\boldsymbol{S}_{\boldsymbol{ff}} = \boldsymbol{I}$ at each frequency.

In the TR case with $\Delta t = 0.5$, the temporal information of the measured signals is resolved in time. In this case, the flow information at the energy-containing temporal scales can be well recovered by the Fourier transform. According to the derivations in § 2.4, the RBE results should be the same as the GRBE in this case. Figure 1(*a*,*d*) depicts the estimation of the flow data at $x = 10$ using the measurements at $x = 0$. It is observed that the prediction results of the RBE are the same as those of the GRBE, which match pretty well with the reference numerical results. To quantitatively analyse the dependence of estimation accuracy on the distance between the measured and predicted signals, the relative MSE at the ith location is defined as

$$\mathrm{Err} = \frac{\left\langle \sum_{j=1}^{N_{\check{T}}} |\varepsilon(i,j)|^2 \right\rangle}{\left\langle \sum_{j=1}^{N_{\check{T}}} |u(i,j)|^2 \right\rangle}, \tag{3.2}$$

where $\varepsilon(i,j) = u(i,j) - u_e(i,j)$ is the error at the ith location and jth time instance. Figure 2(*a*) shows the relative MSEs of the estimations at different spatial locations in the TR case. The relative MSEs from the RBE and GRBE overlap with each other at all of the considered locations.

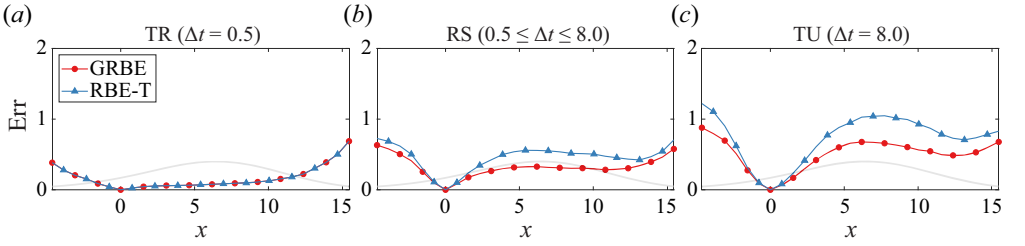


Figure 2. Relative MSEs of the estimations with measurements at $x = 0$ in cases TR (a), RS (b) and TU (c). The translucent grey curves denote the spatial distribution of the relative fluctuation energy from the reference numerical simulation.

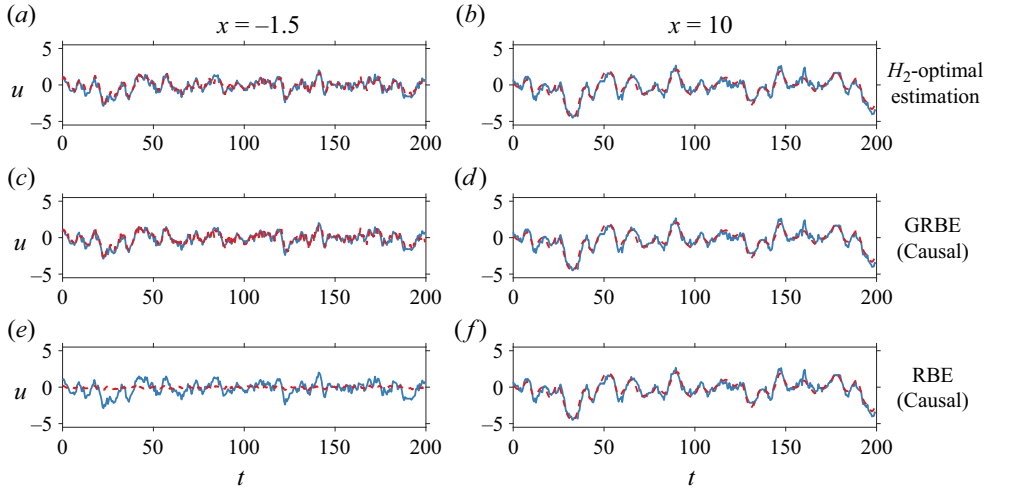


Figure 3. Causal estimation of the flow states at $x = -1.5$ (a,c,e) and $x = 10$ (b,d,f) using the measurements at $x = 0$ with the H_2-optimal estimation (a,b), GRBE (c,d) and RBE (e,f). The solid and dashed curves denote the reference numerical results and the estimated ones, respectively.

In the RS case, the sampling time steps are set to be uniformly distributed in the range of $0.5 \leqslant t \leqslant 8.0$. From figure 1($b$,$e$), the GRBE results match well with the reference numerical results, while those from the RBE deviate obviously from the reference data. In figure 2(b), the relative MSEs of the RBE increase rapidly when moving away from the measurement location, while those of the GRBE are much lower at corresponding locations. Similar phenomena are observed in the TU case, where the estimation accuracy of the GRBE is also much lower than the RBE.

3.1.2. *Causal estimation results*

In Martini *et al.* (2020), the casual RBE estimator is proposed by truncating the resolvent-based kernel (2.26), which can be explicitly expressed in (2.30) with the weighting function $w(\tau) = 0$ when $\tau > t$, where t is the estimation time step. On the other hand, the causal GRBE estimator can be realised by setting $\check{\boldsymbol{t}}$ containing only the time instances with $\tau = t$ and $\tau \leqslant t$ when constructing the Fourier bases $\check{\zeta}_{\boldsymbol{q},j}$ and $\check{\zeta}_{\boldsymbol{m},j}$ in (2.18), respectively.

Figures 3 and 4 show the estimated instantaneous flow states and relative MSEs of the one-dimensional Ginzburg–Landau equation with measurements located at $x = 0$.

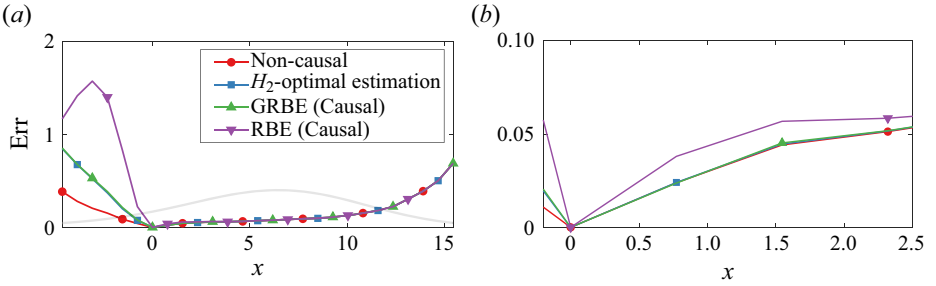


Figure 4. Relative MSEs of the causal estimations with measurements at $x = 0$ (a), with zoomed results near $x = 0$ (b). The translucent grey curves denote the spatial distribution of the relative fluctuation energy from the reference numerical simulation.

Re_τ	Nx	Nz	Ny	L_x/h	L_z/h	L_y/h	$(\bar{u}_c\delta t)/h$	δt^+	$(u_\tau T)/h$
186	768	512	97	12π	4π	2	0.2	2.13	13.74
547	1536	1536	257	8π	4π	2	0.18	4.81	8.80
934	3072	2304	385	8π	3π	2	0.12	5.09	5.23

Table 1. Parameters of the incompressible channel DNS set-ups.

The sampling time interval is set as $\Delta t = 0.5$, which is the same as that in the TR case. Since the H_2-optimal estimation provides optimal causal results with white-in-time forcing (Doyle *et al.* 1989; Zhou & Doyle 1998), which is met in this case, its results are used for the validation of the causal GRBE herein. From figures 3 and 4, the causal GRBE results overlap with the H_2-optimal estimation results at all of the tested locations. On the other hand, the estimation errors from the causal RBE results are obviously larger than those from the causal GRBE results at $x < 0$. Even in the downstream region with $x \geqslant 0$, the estimation errors of the causal RBE estimator are still subtly larger than those of the causal GRBE estimator, as in figure 4(b). Such observations are consistent with those in Martini *et al.* (2020), where the causal RBE results are found to underestimate the flow state compared with the causal Kalman filter results.

3.2. *Estimation of the turbulent channel flows*

In this section, validations of the GRBE in estimating the turbulent channel flows with $Re_\tau = 186$, 547 and 934 will be conducted, where the RBE results are also presented.

3.2.1. *Descriptions of the DNS database*

The code that is used to generate the widely used open-source DNS database (Del Alamo & Jiménez 2003; Hoyas & Jiménez 2008) for incompressible turbulent channel flows with $Re_\tau = 186$, 547 and 934 is applied to compute the DNS results in this study. Details of the DNS set-ups are listed in table 1. To provide temporally resolved results, the sampling time intervals $\delta t^+ = \delta t{\cdot}u_\tau^2/\nu$ are set as 2.13, 4.81 and 5.09 for cases with $Re_\tau = 186$, 547 and 934, respectively. The normalised total simulation time $(u_\tau{\cdot}T)/h$ is larger than 5.0 in each case to obtain statistically converged results. To process the DNS datasets, the computational domain is divided into blocks with a spatial size of $L_x/h = 4\pi$, $L_z/h = \pi$ and $L_y/h = 2$ in each case. Considering that the fully developed channel flow is

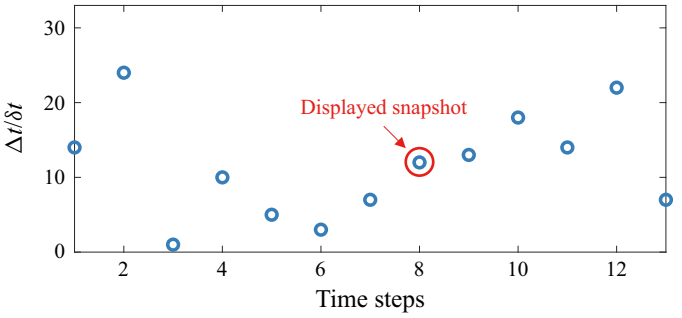


Figure 5. Sampling time intervals in the RS case of turbulent channel flow with $Re_\tau = 934$.

statistically symmetric about the centreline $y = h$, the DNS data that are mirrored about the centreline are treated as another set of blocks besides the original one. The time periods of the blocks are set as $240\delta t$, $160\delta t$ and $160\delta t$ for cases with $Re_\tau = 186$, 547 and 934, respectively, with 75 % overlap in the temporal direction. The rectangular window function is utilised for spectral analysis. With the above set-ups for data processing, the numbers of blocks are 408, 352 and 252 for cases with $Re_\tau = 186$, 547 and 934, respectively. The DNS datasets utilised in this study have been validated in Ying *et al.* (2023) by comparisons with the time-averaged and r.m.s. profiles of the velocities from the open-source DNS database (Del Alamo & Jiménez 2003; Hoyas & Jiménez 2008). In the following, the DNS results with $Re_\tau = 934$ will be mainly discussed, supplemented by the investigations on the estimation results with $Re_\tau = 186$ and 547.

3.2.2. *Comparisons of the estimation results from the GRBE and RBE*

Like the case settings in the validations with the Ginzburg–Landau equation, three cases are set with different temporal resolutions, namely the TR, RS and TU cases. In the TR case, the sampled data are temporally resolved with $\Delta t^+ = \delta t^+$, which corresponds to the case where the measured data are finely sampled in time and well restored. In the RS case, the sampling time intervals are uniformly distributed between $\Delta t^+ = \delta t^+$ and $32\delta t^+$, as shown in figure 5. In this case, the temporal information is partially utilised, which corresponds to the case where the initially time-resolved measured data are missing arbitrarily. As for the TU case, the sampling interval is set as $\Delta t^+ = 32\delta t^+$. Besides the sampling temporal intervals, the types of observed quantities may also influence the estimation accuracy. Meanwhile, in this section, we only show the estimation results from the observations of both the wall shear stress and pressure at the wall for brevity. On the other hand, detailed discussions on the impacts of choices of measurement types and the inclusion of forcing models on the estimation accuracy of the GRBE will be presented in the following § 4. The estimators of both RBE and GRBE are informed by the real flow statistics. With the above settings, the GRBE and RBE estimators are informed by abundant information, including all the considered wall quantities, i.e. wall shear stress and pressure, and exact flow statistics. Such choice enables us to adequately compare the performance of the two methods with given sampling time intervals of the measurement while excluding the influence from other factors.

Intuitive insights into the estimation results can be found in the instantaneous flow field at $y^+ = 100$ with $Re_\tau = 934$ shown in figure 6. It is observed that the estimation results from the GRBE and RBE are the same in the TR case. The estimated results from both approaches recover well the representative flow structures in the DNS results. With the

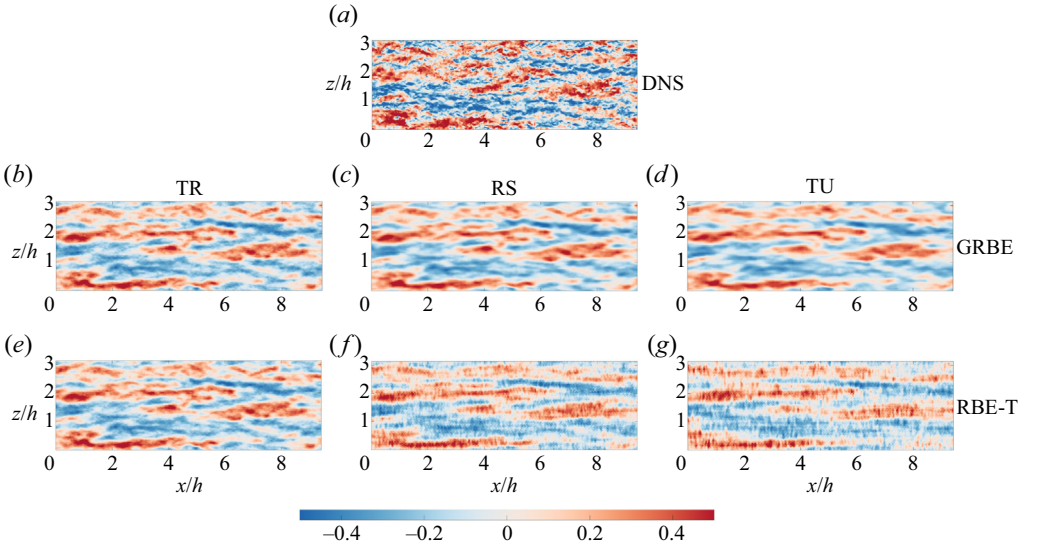


Figure 6. Instantaneous flow state at $y^+ = 100$ from the DNS results (*a*), the GRBE results (*b–d*) and the RBE results (*e–g*). The sampling time intervals are set as temporally resolved (in the TR case) (*b*,*e*), randomly sampled (in the RS case) (*c*,*f*) and temporally unresolved (in the TU case) (*d*,*g*). The values shown in the figures are normalised by the maximum velocity fluctuation value in the DNS result.

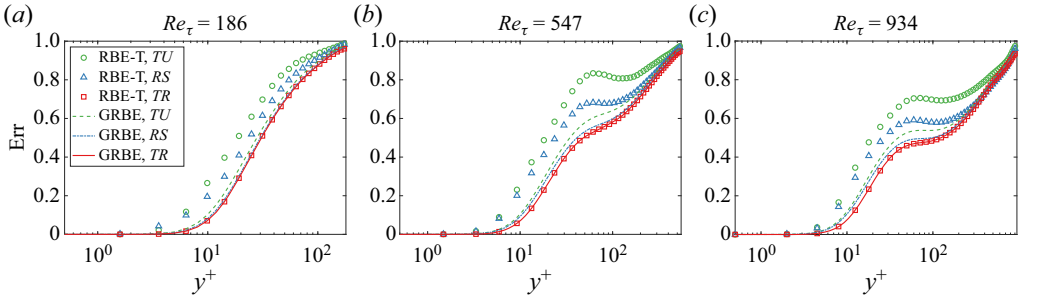


Figure 7. Comparison of the relative MSEs from the RBE and GRBE with different measurement sampling intervals with $Re_\tau = 186$ (*a*), 547 (*b*) and 934 (*c*).

increase of the sampling intervals, the estimation accuracy of the RBE rapidly deteriorates, where some non-physical small-scale spanwise streaks are observed, as in figure 6(*f*,*g*) in the RS and TU cases. In the meantime, although the flow structures estimated by the GRBE also become blurred a little in cases RS and TU, which is attributed to the decrease of measurement information, they still provide credible estimations of the large-scale structures.

To quantitatively investigate the estimation accuracy of the GRBE and RBE results, the relative MSE for the estimated turbulent channel flow is defined as

$$\mathrm{Err}(y) = \frac{\left\langle \sum_{k_x} \sum_{k_z} \sum_t \hat{\varepsilon}(k_x, k_z, y, t) \bar{\hat{\varepsilon}}(k_x, k_z, y, t) \right\rangle}{\left\langle \sum_{k_x} \sum_{k_z} \sum_t \hat{u}(k_x, k_z, y, t) \bar{\hat{u}}(k_x, k_z, y, t) \right\rangle}. \tag{3.3}$$

The profiles of the relative MSEs from the GRBE and RBE with different Re_τ and sampling time intervals are shown in figure 7. In the TR case where $\Delta t = \delta t$, the curves

denoting the estimation error from the GRBE and RBE overlap with each other, which again demonstrates that the GRBE is equivalent to RBE when the sampling time is resolved in time, as pointed out in § 2.4. When the sampling intervals are randomly distributed in the RS case, the estimation errors of the GRBE along the height slightly increase compared with those in the TR case due to the reduction of measurement information. Meanwhile, the estimation errors of the RBE in the RS case are much larger in comparison with those in the TR case. When the sampling time interval increases to $32\delta t^+$ in the TU case, the estimation errors of the RBE results become even larger, while those from the GRBE maintain a relatively low magnitude. For instance, compared with the TR cases, the relative MSE from the RBE at $y^+ = 100$ with $Re_\tau = 934$ increases by 43.7 % in the TU case, while that from the GRBE increases by only 10.1 %.

From the above comparisons between the estimation results from the RBE and GRBE, the GRBE is capable of providing a satisfying estimation for the turbulent channel flows with different sampling time intervals when the measurements are wall shear stress and pressure with real flow statistics informing the estimator.

4. Applications

In this section, the validated GRBE will continue to be applied in the estimations of turbulent channel flows to systematically investigate the characteristics of the GRBE in different application scenarios. In general, there are several related problems to be further uncovered, including (i) the influence of the abundance of the measured information on the estimation accuracy, including the amount of temporal information and sensor types, (ii) the Reynolds number effects and (iii) the performance of the forcing models when incorporated into the GRBE estimator. Discussions on the impacts of measurement abundance and Reyolds number effects are provided in § 4.1, while those on the impacts of the forcing models are presented in § 4.2.

4.1. *Impacts of the abundance of measurement information on the estimation accuracy*

The impact of different observation types, including wall shear stress and pressure, is preliminarily studied in Amaral *et al.* (2021), leading to the conclusion that the estimation accuracy is improved with the increasing amount of the measured quantity types. Meanwhile, the problem becomes more complex when the abundance degree of temporal information is introduced. The variations of estimation accuracy with the measurement abundance at different flow scales also await further examination.

To investigate the impact of the abundance of temporal measurement information, three cases with different sampling temporal intervals are set, namely TU for the temporally unresolved case, RS for the randomly sampled case and TR for the temporally resolved case. Note that the sampling intervals in the TU case in this section are further enlarged to be $240\delta t$, $160\delta t$ and $160\delta t$ for turbulent channel flows with $Re_\tau = 186$, 547 and 934, respectively, which are all the same as the lengths of the DNS data in temporal direction for each block with corresponding Re_τ, as presented in § 3.2.1. This means that the flow information at only one snapshot is used for each TU case. On the other hand, to investigate the impacts of the sensor types on the estimation accuracy, different combinations of wall quantities, including the streamwise shear stress τ_u, spanwise shear stress τ_w and pressure p, are also considered in this study. All of the to-be-tested combinations of the sampling intervals and sensor types in this study are summarised in table 2. In this section, the estimations with wall measurements informed by the real flow statistics are conducted.

	τ_u	τ_w	p	τ_u, τ_w	τ_u, τ_w, p
Temporally unresolved (TU)	TU_u	TU_w	TU_p	$\mathrm{TU}_{u,w}$	$\mathrm{TU}_{u,w,p}$
Randomly sampled (RS)	RS_u	RS_w	RS_p	$\mathrm{RS}_{u,w}$	$\mathrm{RS}_{u,w,p}$
Temporally resolved (TR)	TR_u	TR_w	TR_p	$\mathrm{TR}_{u,w}$	$\mathrm{TR}_{u,w,p}$

Table 2. Nomenclature of the cases testing the impacts of the sampling time intervals and sensor types on estimating the turbulent channel flow in this study.

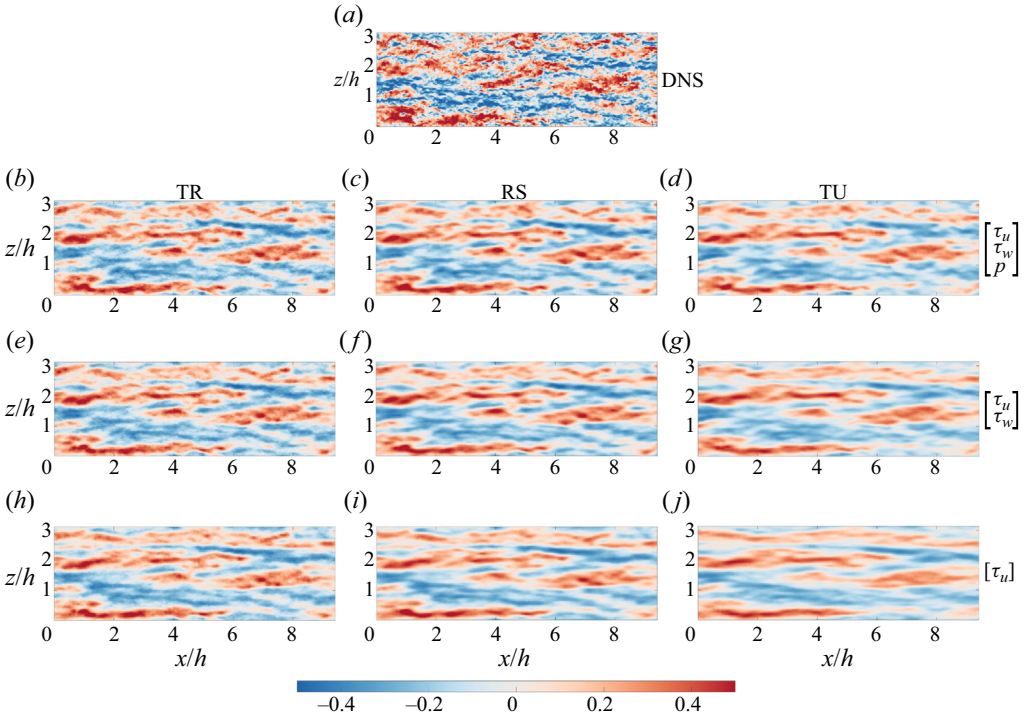


Figure 8. Estimated instantaneous flow state at $y^+ = 100$ with measurements at the wall from the GRBE. The DNS result (*a*) is presented as a reference. The measured quantities are set as (τ_u, τ_w, p) (*b–d*), (τ_u, τ_w) (*e–g*) and τ_u (*h–j*). The sampling time intervals are set as temporally resolved (*b*,*e*,*h*), randomly sampled (*c*,*f*,*i*) and temporally unresolved (*d*,*g*,*j*). The values shown in the figures are normalised by the maximum velocity fluctuation value in the DNS result.

Instantaneous snapshots of the DNS result and the estimations at $y^+ = 100$ with $Re_\tau = 934$ are depicted in figure 8. Different types of sensors with various extents of temporal information are considered. In case $\mathrm{TR}_{u,w,p}$ where the measured flow information is the most abundant, the GRBE estimator provides the most accurate results, where the flow structures recover well the representative ones in the DNS results. With the decrease of either the sensor types or temporal information, the estimated small flow structures become blurred. In case TU_u with the least measured information, only the large-scale structures with long streaky structures could be observed. Interestingly, it is found that the flow structures are still well estimated in case $\mathrm{TU}_{u,w,p}$ where the flow information at only one snapshot is used for the estimation. Compared with case TR_u with measurements of

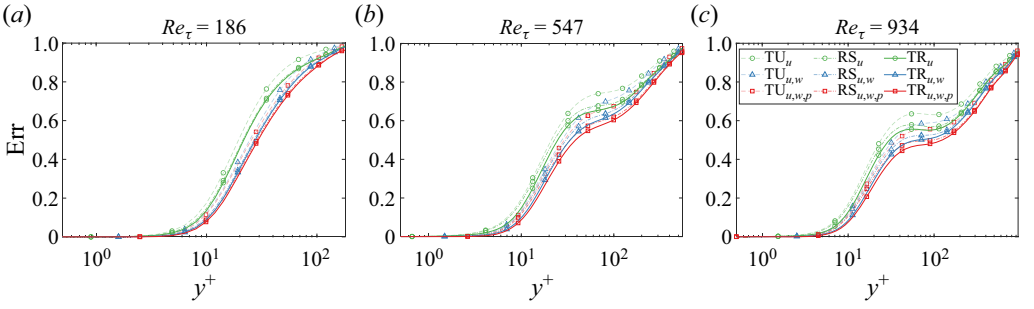


Figure 9. Profiles of the relative MSEs of the streamwise velocity fluctuations from the GRBE in turbulent channel flows with $Re_\tau = 186$ (*a*), 547 (*b*) and 934 (*c*).

$N_x \times N_z \times N_T$, the results from case $\mathrm{TU}_{u,w,p}$ with $3 \times N_x \times N_z$ provide comparable flow information with a much lower requirement of measured data.

Next, we further investigate the impacts of the measurement abundance through the profiles of the relative MSEs. Figure 9 shows the profiles of the relative MSEs along the height. For all the tested Re_τ, it is observed that the estimation errors generally increase with the enlarged distance from the measurement location, i.e. the wall. At each height, the results in TR cases are the most accurate among those with the same sensor types. Meanwhile, case $\mathrm{TU}_{u,w,p}$ that only needs the measured data at one instant has an even smaller error compared with case TR_u that needs temporally resolved measurements when $y^+ < 100$. The Reynolds number effects on the GRBE accuracy can also be observed from figure 9. At a given height y^+, the estimation errors decrease with the enlargement of Re_τ for each case. For instance, at $y^+ = 100$, the estimation errors in case $\mathrm{TR}_{u,w,p}$ are equal to 0.88, 0.60 and 0.48 when $Re_\tau = 186$, 547 and 934, respectively. In the meantime, an approximate plateau of the relative MSEs in the buffer zone is observed when $Re_\tau = 547$ and 934, which is more obvious with the higher Re_τ. When $Re_\tau = 934$, the estimation errors even slightly decrease with the increase of height for cases with measurements of τ_u.

To further excavate the underlying mechanism of the differences in the estimation accuracy with different Re_τ, the relative estimation error energy as a function of y and the flow scale λ_x is defined by

$$\mathrm{Err}_x(\lambda_x, y) = \frac{\left\langle \sum_{k_z} \sum_t \hat{\varepsilon}(k_x, k_z, y, t)\bar{\hat{\varepsilon}}(k_x, k_z, y, t) \right\rangle}{\left\langle \sum_{k_z} \sum_t \hat{u}(k_x, k_z, y, t)\bar{\hat{u}}(k_x, k_z, y, t) \right\rangle}, \tag{4.1}$$

where $\lambda_x = 2\pi/k_x$. As shown in figure 10, with the increasing flow scale, the estimation errors at a given height decrease monotonically. In the meantime, the plateaus in the estimation errors that are observed in the overall error energy profiles in figure 9(*b*,*c*) are much more obvious for the VLSMs ($\lambda_x \geqslant 3h$) (Adrian, Meinhart & Tomkins 2000; Lee & Sung 2013; Lee *et al.* 2014) in the buffer zone. Especially, for the VLSMs with $\lambda_x = 4\pi$, the values of the estimation errors are around 0.18 at the plateaus at $y^+ = 40 \sim 100$ in case $\mathrm{TR}_{u,w,p}$ with $Re_\tau = 934$. Meanwhile, for the motions with small scales of $\lambda_x \leqslant 3h$, the plateaus are not found in the buffer zone, which indicates that the nonlinear effects are stronger for such small-scale motions.

In the following, deeper investigations on the GRBE estimator are conducted by studying the two-dimensional energy spectra of the estimation and the corresponding energy distributions of the errors as functions of λ_x and λ_z. Similar to (4.1), the

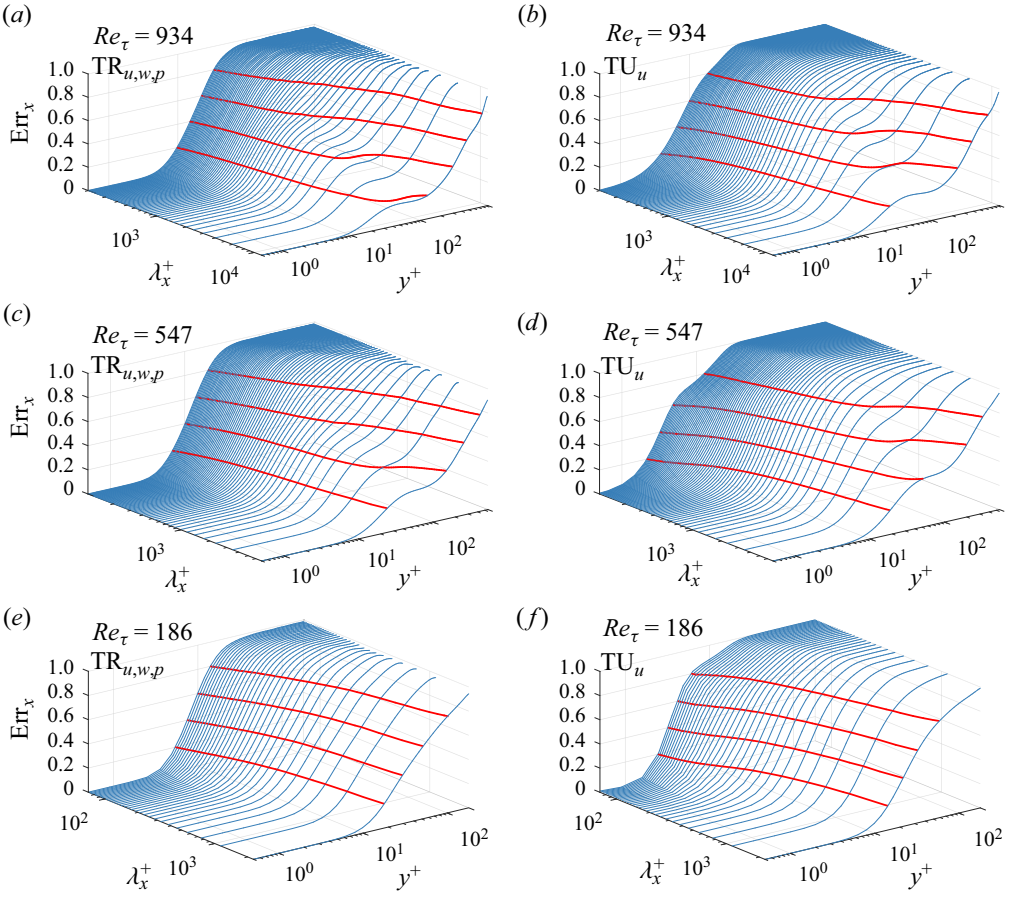


Figure 10. Relative MSEs of the streamwise velocity fluctuations from the GRBE in turbulent channel flows as functions of the flow scale λ_x^+ and wall-normal distance y^+ in cases $\mathrm{TR}_{u,v,w}$ (*a*,*c*,*e*) and TU_u (*b*,*d*,*f*) with $Re_\tau = 934$ (*a*,*b*), 547 (*c*,*d*), and 186 (*e*,*f*). The blue curves denote the profiles of the relative MSEs for flow motions with given λ_x^+, while the red ones denote the contours of $\mathrm{Err}_x = 0.2$, 0.4, 0.6 and 0.8, respectively.

two-dimensional relative energy distribution of the estimation error at a given height y_o is defined by

$$\mathrm{Err}_{xz}(\lambda_x, \lambda_z) = \frac{\left\langle \sum_t \hat{\varepsilon}(k_x, k_z, y_o, t)\bar{\hat{\varepsilon}}(k_x, k_z, y_o, t) \right\rangle}{\left\langle \sum_t \hat{u}(k_x, k_z, y_o, t)\bar{\hat{u}}(k_x, k_z, y_o, t) \right\rangle}, \tag{4.2}$$

where $\lambda_x = 2\pi/k_x$. The premultiplied spectra of estimated velocity fluctuation energy and the corresponding relative MSEs of u at $y^+ = 40$ and 100 are shown in figures 11 and 12, respectively. To quantify the spectral region with relatively high estimation accuracy, the contours of $\mathrm{Err}_{xz}(\lambda_x, \lambda_z) = 0.25$ and 0.75 in the corresponding figures are highlighted with white solid and dashed curves, respectively. Meanwhile, the scope of the self-similar attached eddies, i.e. $\lambda_x \geqslant 2\lambda_z$, $\lambda_x \geqslant 10y$, and $\lambda_z \geqslant 2y$, as suggested by Lozano-Durán, Flores & Jiménez (2012), Hwang (2015) and Cheng *et al.* (2019), are also highlighted with black dot-dashed lines.

From figures 11 and 12, estimated premultiplied energy spectra in the dashed curves match relatively well with the DNS results within the spectral region corresponding to

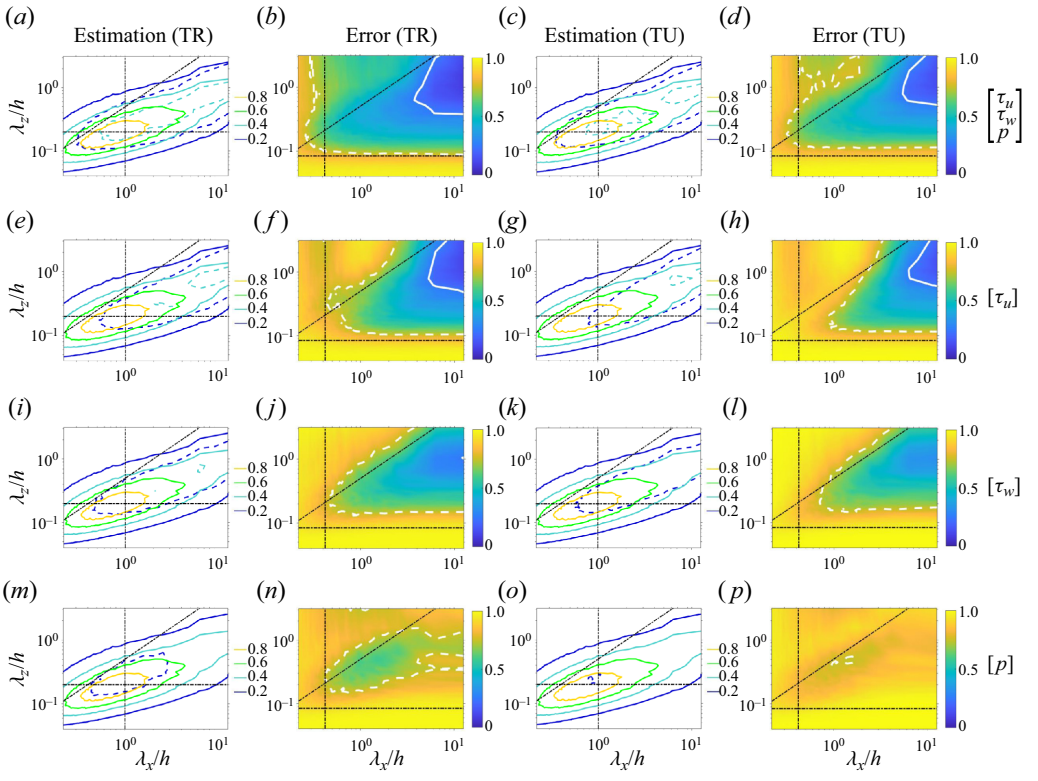


Figure 11. Premultiplied energy spectra (1st and 3rd panels) and the relative error energies (2nd and 4th panels) of u at $y^+ = 40$ using the measurements at the wall. The sensor types are τ_u, τ_w, and p (*a–d*), τ_u (*e–h*), τ_w (*i–l*) and p (*m–p*). The black dot-dashed lines in each figure denote the range of the attached eddies. In the 1st and 3rd panels, the solid and dashed curves denote the contours of the energy spectra from DNS and estimations, respectively. In the 2nd and 4th panels, the contours of $\mathrm{Err}_{xz} = 0.25$ and 0.75 are highlighted with white solid and dashed curves, respectively.

the attached eddies when the sensor types are $\{\tau_u, \tau_w, p\}$ or $\{\tau_u\}$. Correspondingly, the relative energies of the estimation errors are generally smaller for the flow portions with larger wavelengths. On the other hand, the spectral regions where $\mathrm{Err}_{xz} \leqslant 0.75$ at $y^+ = 40$ are much larger than those at $y^+ = 100$, which means that the estimation errors of the small-scale structures are significantly enlarged. Meanwhile, the extents of the regions where $\mathrm{Err}_{xz} \leqslant 0.25$ for the large-scale flows do not shrink so obviously when the estimation layers lifts from $y^+ = 40$ to $y^+ = 100$. This phenomenon is consistent with the plateaus for the VLSMs existing at the buffer zones observed in figure 10.

The sensor types are found to have significant impacts on the energy distributions of the estimation errors in figures 11 and 12. Comparing the results in cases $\mathrm{TR}_{u,w,p}$ and TR_u, obvious increases of the estimation errors are observed in the spectral region of $10y < \lambda_x < 2\lambda_z$. At $(\lambda_x/h, \lambda_z/h) = (0.36\pi, \pi)$ on $y^+ = 40$, the value of the error energy is 0.52 and 0.94 in cases $\mathrm{TR}_{u,w,p}$ and TR_u, respectively, with an increasing rate of 81.5 %. On the other hand, at $(\lambda_x/h, \lambda_z/h) = (4\pi, 0.5\pi)$ inside the scope of the attached eddies, the error energy increases by 20.0 % in case TR_u in comparison with that in case $\mathrm{TR}_{u,w,p}$, which is not so dramatic as that at $(\lambda_x/h, \lambda_z/h) = (0.36\pi, \pi)$. Besides the estimation results with measurements of $\{\tau_u, \tau_w, p\}$ and $\{\tau_u\}$, those with $\{\tau_w\}$ and $\{p\}$ are also shown in figures 11 and 12 to study the characteristics of the GRBE in cross-quantity estimation.

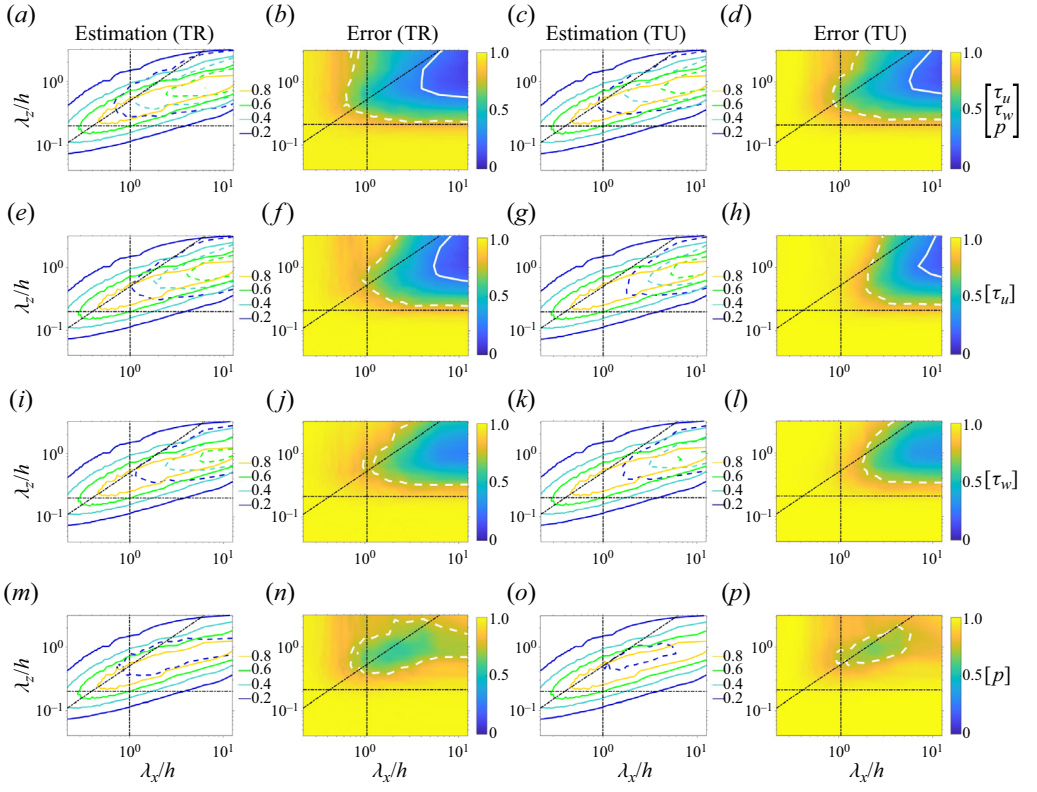


Figure 12. Same as figure 11, but at $y^+ = 100$.

In case TR_w, the region where $\mathrm{Err}_{xz}(\lambda_x, \lambda_z) \leqslant 0.75$ that is surrounded by the white dashed lines is comparable to that in case TR_u, which roughly matches the range of the attached eddies. However, for large-scale flow motions, the estimation errors in case TR_w are generally larger than 0.25, which is larger than those in TR_u. Thus, although the GRBE is capable of providing fair estimations of the small-scale structures of u with measurements of τ_w, the large-scale structures cannot be well estimated from cross-quantity estimations. In case TR_p with wall pressure measurements, the estimation errors are obviously lower than those in the TR_u case for the small-scale flows, which is beyond the scope of the attached eddies. For $(\lambda_x/h, \lambda_z/h) = (0.36\pi, \pi)$ at $y^+ = 40$, the error energy in case TR_p is 0.75, which is 20 % lower than that in case TR_u. Meanwhile, for the estimations of large-scale structures, the GRBE in the TR_p case is not as good as those in other cases.

Besides the sensor types, the abundance of the measurement temporal information is indeed another factor that influences the performance of the GRBE. Compared with cases $\mathrm{TR}_{u,w,p}$ and $\mathrm{TU}_{u,w,p}$, the reduction of the measured information in time reduces the estimation accuracy for all the considered spatial scales. Unlike the sensor types, which have significant influences on the patterns of the energy distributions of errors, the variations of the temporal information of measurements do not obviously change the error energy distributions. For instance, at $(\lambda_x, \lambda_z) = (0.36\pi, \pi)$ and $(4\pi, 0.5\pi)$ on $y^+ = 40$ that are outside and inside the scopes of attached eddies, the value of the relative error energy increases by 46.2 % and 53.9 % in cases $\mathrm{TU}_{u,w,p}$ compared with those in the $\mathrm{TR}_{u,w}$ case, respectively, which are relatively close to each other.

To find out the performance of the GRBE with the variations of sampling intervals, we gradually increase the sampling time interval. The variations of the relative MSEs with the

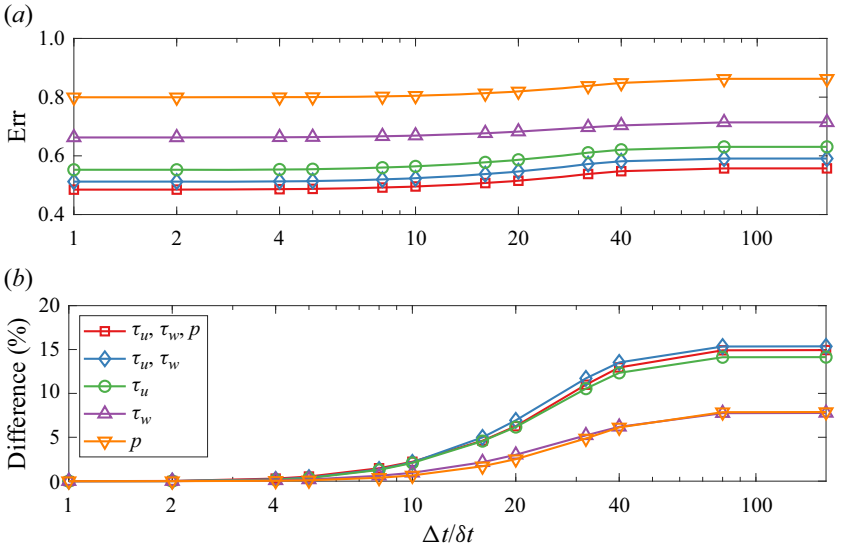


Figure 13. Variations of the estimation error energies with measurement time intervals. (*a*) Values of the relative MSEs. (*b*) Variations of the relative MSEs compared with those in the TR cases with the corresponding measured quantities.

enlargement of the sampling intervals are shown in figure 13(*a*), while the increasing rate of the relative MSEs compared with those in the TR cases are shown in figure 13(*b*). For each combination of the sensor types, the estimation error does not obviously increase when $\Delta t/\delta t \leqslant 5$. The relative deviations compared with the TR cases are no greater than 5 % until $\Delta t/\delta t$ increases to 16 ($\Delta t^+ = 81.41$). When $16 \leqslant \Delta t/\delta t \leqslant 80$, a relatively obvious increase in the estimation error with the enlarging sampling interval is observed. When the sampling intervals increase over $80\delta t$, the estimation errors become almost unchanged with a further increase of Δt. For the estimations with measurements including τ_u, the largest relative MSEs in the case with $\Delta t = 160\delta t$ are around 15 % larger than the temporally resolved cases. Meanwhile, the errors of the cross-quantity estimation with measurements of $\{\tau_w\}$ and $\{p\}$ increase by around 8 % in TU cases compared with the TR ones. The relatively small increase in the cross-quantity estimation errors should be attributed to the fact that their magnitudes are already large in the TR cases, as shown in figure 13(*a*). In the meantime, it is noticed that the estimation error in case $\mathrm{TU}_{u,w,p}$ is almost the same as that in the TR_u case with a discrepancy of less than 1 %.

To further investigate the impact of sampling temporal intervals on the estimation accuracy, the increasing rates of the error energies compared with those in the TR cases as functions of measurement time intervals Δt and the streamwise wavelength λ_x are depicted in figure 14. To quantify the impact of the increasing sampling temporal intervals on the estimation accuracy of flow motions with different scales, the threshold sampling interval Δt_T is defined herein as the maximum sampling interval with which the increasing rate of Err_x compared with the TR cases is smaller than 5 %. For the attached eddies with $\lambda_x \geqslant 10y$, the threshold temporal interval Δt_T is generally larger than $20\delta t$, where $\Delta t_T^+ \geqslant 101.76$. On the other hand, when $3y \leqslant \lambda_x < 10y$, Δt_T decreases with decreasing wavelength, which indicates that finer temporal resolution is required to provide effective information of the smaller flow scales. When $\lambda_x = 3y$ with the sensor of $\{\tau_u, \tau_w, p\}$, Δt_T decreases to the minimum value of approximately $10\delta t$. When the flow scales further decrease to $\lambda_x < 3y$, Δt_T dramatically increases instead, which is attributed to the fact

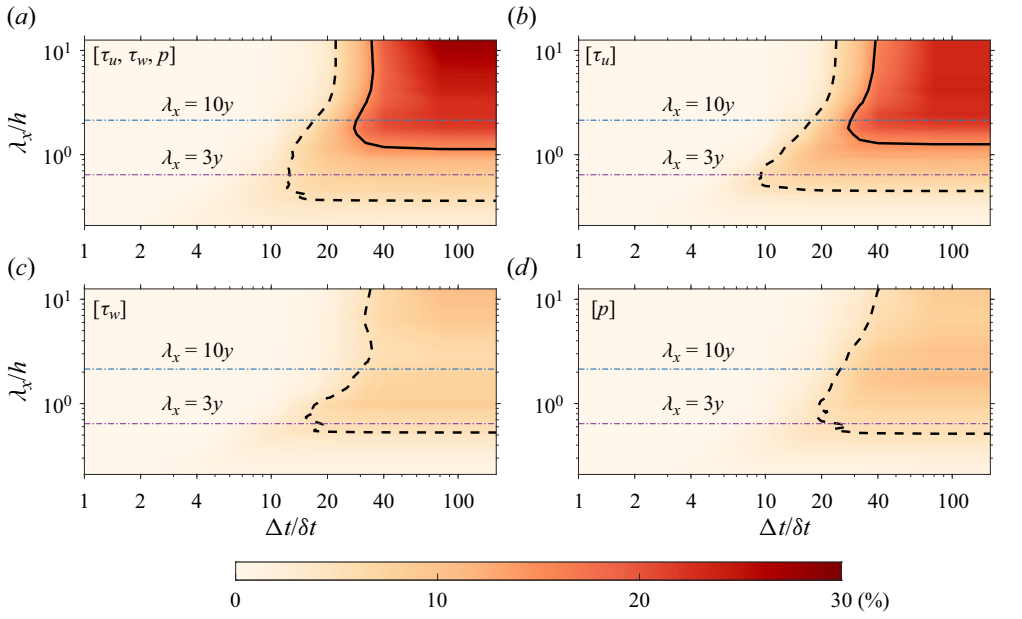


Figure 14. The increasing rate of the estimation error energies compared with those in the TR case as functions of sampling time intervals Δt and streamwise wavelengths λ_x at $y^+ = 100$ with wall measurements of $\{\tau_u, \tau_w, p\}$ (*a*), $\{\tau_u\}$ (*b*), $\{\tau_w\}$ (*c*) and $\{p\}$ (*d*). The black dashed and solid lines denote the contours of 5 % and 15 %, respectively.

that the estimation errors with $\lambda_x < 3y$ are already near 1.0 in the TR cases, as shown in figure 10.

From the above discussions, it can be concluded that the temporal information has limited impacts on the estimation accuracy, with a maximum reduction of around 13.0 % of the estimation error in the TR case compared with the TU cases at $y^+ = 100$ with measurements of wall shear stress and pressure. When the sampling time intervals $\Delta t = 16\delta t$ ($\Delta t^+ = 81.41$), the increasing rates of the estimation errors are around 5 % compared with the TR results. Such findings indicate that the measurements with sparse samplings can be effectively used to provide reasonable estimations of the flow state with the GRBE. On the other hand, the inclusion of more types of observations can be efficient in improving the estimation accuracy. In the TU cases without temporal measurement information, the estimation error reduces by 12.3 % when the sensors for τ_w and p are added besides that for τ_u. Consequently, the relative MSEs at $y^+ = 100$ in the $\mathrm{TU}_{u,w,p}$ case with the most sensor types but least temporal information are almost the same as those in the TR_u case with the least sensor type but most temporal information.

Besides the above conclusions, we would also like to discuss the tolerated estimation error from the GRBE when applied in real engineering scenarios. Generally speaking, the feasibility of the GRBE depends on specific requirements on the estimation accuracy for a given range of flow scales. For instance, in flow control problems, the combination of sensors and actuators that are both limited to one wall-normal location provides a comparable control performance to the case where the limited-in-space actuators are informed by the entire flow information (Oehler & Illingworth 2021). This is due to the fact that the linear actuators are also limited by the flow coherence like the linear estimator, which means that the energy-containing LSMs could be better controlled. Since the GRBE provides reliable estimations of the large-scale structures in the inner layer, it can already well excavate the potentials of the linear actuator in controlling the large-scale

turbulent flows. On the other hand, for the estimation problem of small-scale structures outside the scale range of the attached eddies, which is unresolved yet for all the current linear and nonlinear estimators (e.g. Guastoni *et al.* 2021) to the best of the authors' knowledge, the GRBE is also not able to provide reliable results. Meanwhile, since GRBE has been validated to be the optimal linear estimator, the comprehensive discussions in this section are meaningful to evaluate the reliability and quantify the uncertainties of the linear estimation results in engineering applications.

4.2. *Performance of the forcing models in flow state estimation*

In the above, the performance of the GRBE estimators informed by the real flow statistics under different cases is discussed. Thus, no matter how the performance varies in different cases, they are already the optimal linear estimators with given measurements. In practice, the flow statistics are not always available when estimating the flow state, where the forcing models that predict the spatio-temporal statistical properties of turbulent flows are needed to inform the GRBE estimator. Thus, it is important to clarify their performance when informing the linear estimator of GRBE. In the following, three existing forcing models that provide spatio-temporal descriptions of the forcing statistics, including the B-model (Hwang & Cossu 2010; Madhusudanan *et al.* 2019), W-model and λ-model (Gupta *et al.* 2021), will be applied to inform the GRBE estimator. Mathematical descriptions of these three models are provided in Appendix B. Estimation results with the wall measurements will be used for the discussions in this section.

Figure 15 shows the estimation errors of different forcing models with different sensor types along the wall-normal direction. Moreover, to provide intuitive results, figures 16, 17 and 18 show the instantaneous flow states estimated by GRBE estimators that are informed by the B-model, W-model and λ-model, respectively. In the TU case, the W-model-informed GRBE provides the most accurate results, which are closer to the optimal linear estimated results than those from the other models. In case TU_u, the estimation error in the W-model-informed result at $y^+ = 100$ is 13.3 % larger than that from the optimal linear estimations, as in figure 15(*b*). From figure 17(*d*,*g*,*j*), it is also observed that the estimated large-scale structures from the W-model match well with the DNS results in the TU cases. Such good agreement between the W-model-informed result and the optimal linear result in case TU_u indicates that the W-model can effectively predict the spatial coherence of the streamwise velocity between the wall and the higher region. Compared with the estimation errors from the W-model and λ-model, those from the B-model are obviously larger in the near-wall region of $y^+ \leqslant 10$. Meanwhile, with the increase of estimation height, the errors from the λ-model rapidly increase and exceed those from the B-model where $y^+ \geqslant 15$. On the other hand, the estimation error from each forcing model does not obviously decrease with the increase of sensor types. When converting the sensor types from $\{\tau_u\}$ to $\{\tau_u, \tau_w, p\}$, the estimation errors at $y^+ = 100$ decrease by 0.7 %, 3.9 % and 8.0 % for the B-model, W-model and λ-model, respectively, which are lower than the value of 13.3 % by the optimal linear estimator. The relatively low efficiency of the forcing-model-informed GRBE estimator to improve the estimation accuracy with an increasing number of measured variables is attributed to the fact that all the tested forcing models cannot effectively predict the covariance between different quantities.

With the increase of the temporal information of measurements, it is interesting to find that the estimation errors from the W-model- and λ-model-informed estimators increase accordingly, the tendencies of which are opposite to those of the optimal linear estimator. In the $\mathrm{TR}_{u,w,p}$ case, where the estimation error is minimised in the optimal linear estimator,

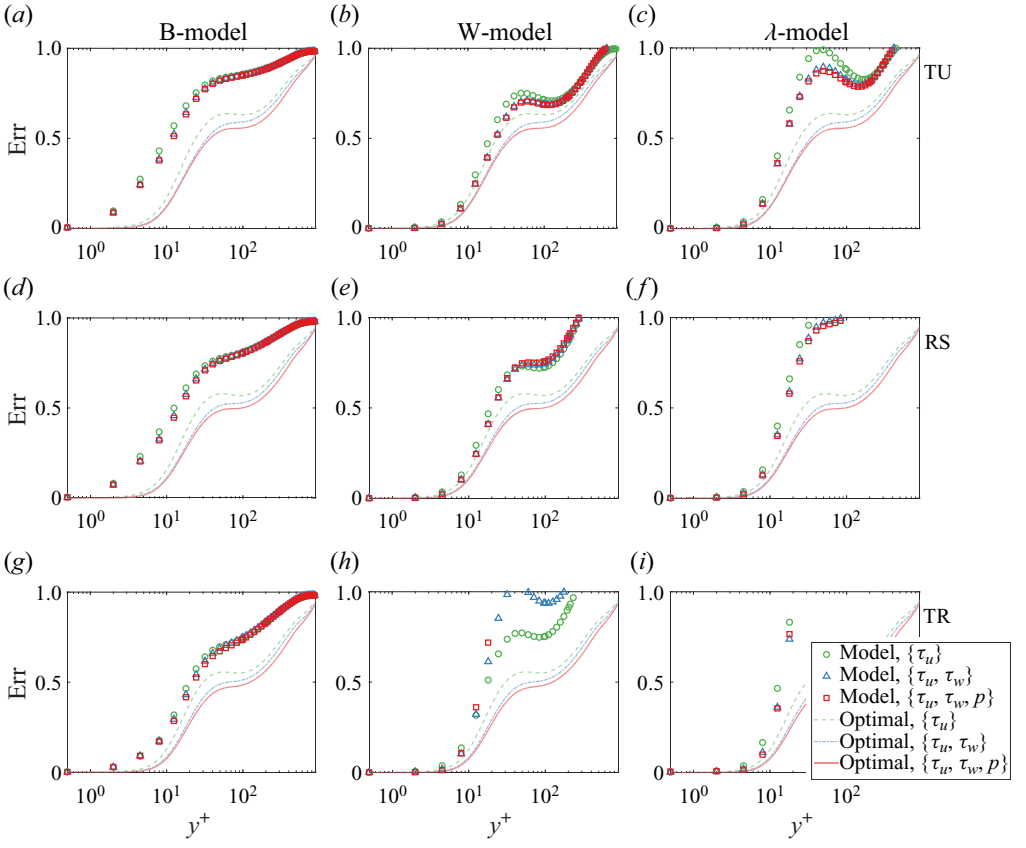


Figure 15. Variations of the estimation errors from the B-model (*a*,*d*,*g*), W-model (*b*,*e*,*h*) and λ-model (*c*,*f*,*i*) with the wall-normal height in the TU (*a–c*), RS (*d–f*) and TR (*g–i*) cases. The scattered symbols and translucent curves denote the estimation results from the GRBE estimator incorporating the forcing models and real flow statistics, respectively.

it is instead maximised for the W-model results, which are 48.9 % larger than that in case $\mathrm{TU}_{u,w,p}$. A similar non-physical phenomenon is observed in the λ-model results. From figures 17(*a*) and 18(*a*), the fluctuation energy of small-scale structures is incorrectly amplified by the W- and λ-models in the TR cases compared with the estimated results in the TU cases. Such non-physical increase of estimation error with the increase of temporal measurement information and sensor types indicates that the W-model and λ-model cannot effectively predict the temporal properties, although the W-model works well in predicting the spatial coherence of the streamwise velocity. On the other hand, the estimation errors of the B-model results decrease monotonically with the increase of the measurement abundance. Compared with the relative MSE in case $\mathrm{TU}_{u,w,p}$ at $y^+ = 100$, that in case $\mathrm{TR}_{u,w,p}$ decreases by 13.2 % from the B-model results.

In summary, the tested forcing models perform differently in cases with different measurement abundance. In case TU_u with the least measurement information, the W-model performs the best. When the sensor types increase from $\{\tau_u\}$ to $\{\tau_u, \tau_w, p\}$, the estimation accuracy of the tested forcing models is not obviously improved. Moreover, with the increase of temporal information, the estimation errors from the W-model and λ-model non-physically increase. Such tendencies are opposite to those in the optimal

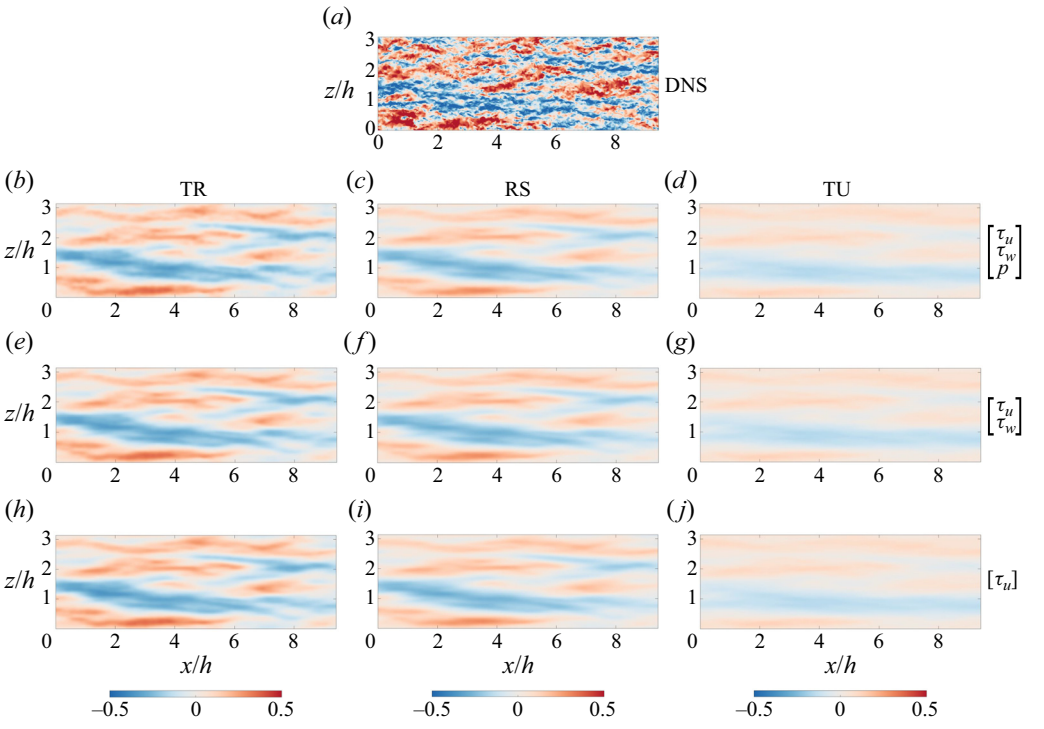


Figure 16. Estimations of the flow state with the B-model-informed estimator. The depicted figures are reference DNS result (*a*); estimations with measurements of wall shear stress and pressure (*b–d*), wall shear stress (*e–g*) and streamwise wall shear stress (*h–j*). The measurements are temporally resolved (*b*,*e*,*h*), randomly sampled (*c*,*f*,*i*) and temporally unresolved (*d*,*g*,*j*). The values shown in the figures are normalised by the maximum velocity fluctuation value in the DNS result.

linear estimation results, indicating that the temporal properties predicted by the W-model and λ-model cannot effectively reflect these of the real turbulence. A forcing model that properly models the forcing in the temporal domain will effectively improve the linear estimation accuracy of the GRBE without the requirement of the real flow statistics, which is to be explored in the future.

5. Conclusions

In this study, the RBE that estimates the turbulence field with temporally resolved measurement data is generalised so that credible linear estimations of the flow state can be obtained when the measurement data are arbitrarily sampled in time. The generalised version of RBE is named the GRBE herein. Unlike the RBE that minimises the expected error energy at each frequency, the transfer kernel of the GRBE is constructed with the target of directly minimising the expected error energy in physical space, by means of finding the stationary point of the MSE as a quadratic function of the transfer kernel. The flow statistics, including the CSD tensors of the forcing and measurement noise, are considered when constructing the GRBE estimator. In cases where the flow statistics are unavailable, the GRBE enables the incorporation of the forcing models that provide spatio-temporal predictions of the flow statistics.

Validations of the GRBE approach include the one-dimensional complex Ginzburg–Landau equation and turbulent channel flows with $Re_\tau = 186, 547$ and 934. The estimation

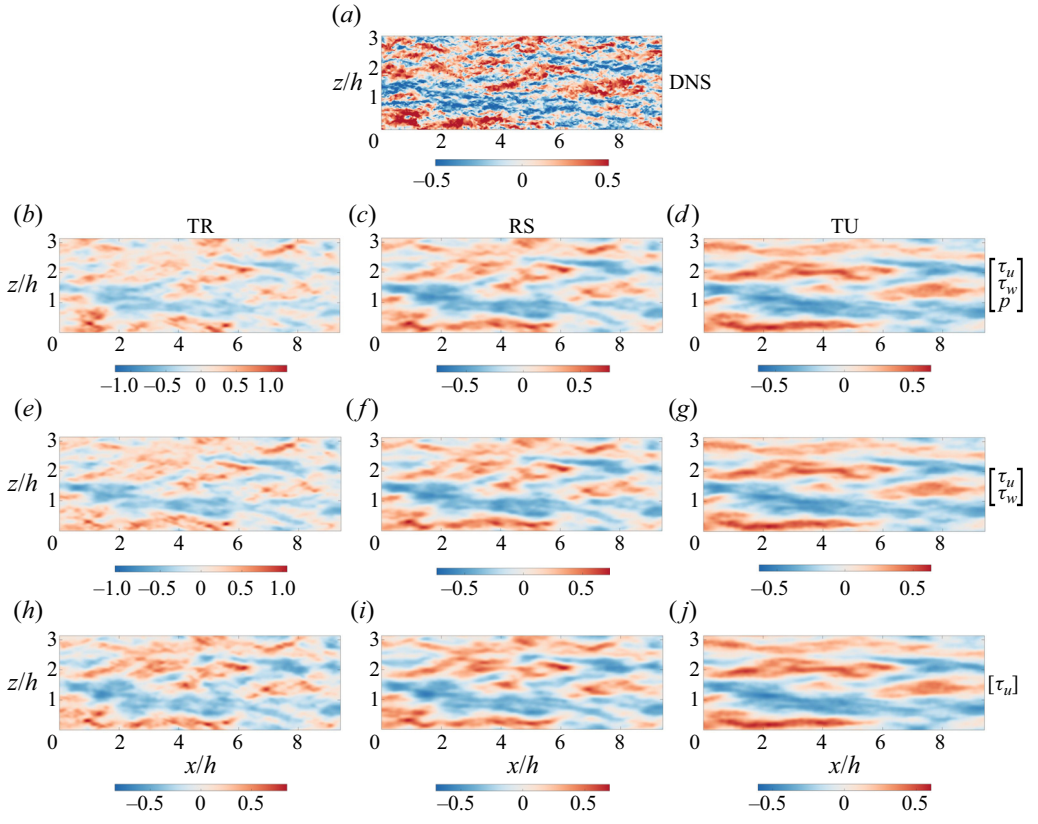


Figure 17. Same as figure 16, but with the W-model-informed estimator.

results from the original RBE (Martini *et al.* 2020; Towne *et al.* 2020) and the H_2-optimal estimation (Doyle *et al.* 1989; Illingworth *et al.* 2018) are also included in the case of the complex Ginzburg–Landau equation. When the measurement information is resolved in the temporal direction, the GRBE results are equivalent to those from the RBE. As the sampling time intervals arbitrarily increase, the relative MSEs of the GRBE results still keep relatively low magnitudes, while those of the RBE are obviously enlarged when the sampling time intervals increase. The GRBE also provides the optimal results in causal estimations by comparing with the H_2-optimal estimation results. In the validation cases with the turbulent channel flows with $Re_\tau = 186$, 547 and 934, the measured signals are set to be the shear stress and pressure at the wall, with the real flow statistics incorporated in the GRBE and RBE estimators. The results of the GRBE and RBE are the same when $\Delta t^+ = \delta t^+ = 5.09$, i.e. the measurements are well resolved in time. When the sampling time intervals Δt^+ increase to $32\delta t^+$, the estimation errors of the RBE are obviously enlarged, while those from the GRBE maintain relatively low magnitudes.

After validations, the GRBE is further utilised to comprehensively investigate its performance in different measuring conditions, where the factors influencing the estimation accuracy include the abundance of temporal measurement information and sensor types, as well as the Reynolds number effects. With the increase of Re_τ, the relative MSE at a given height decreases accordingly, where the plateaus of the relative MSE magnitudes are observed in the buffer zone when $Re_\tau = 547$ and 934. Investigations on the estimation errors with different streamwise scales λ_x reveal that the plateaus

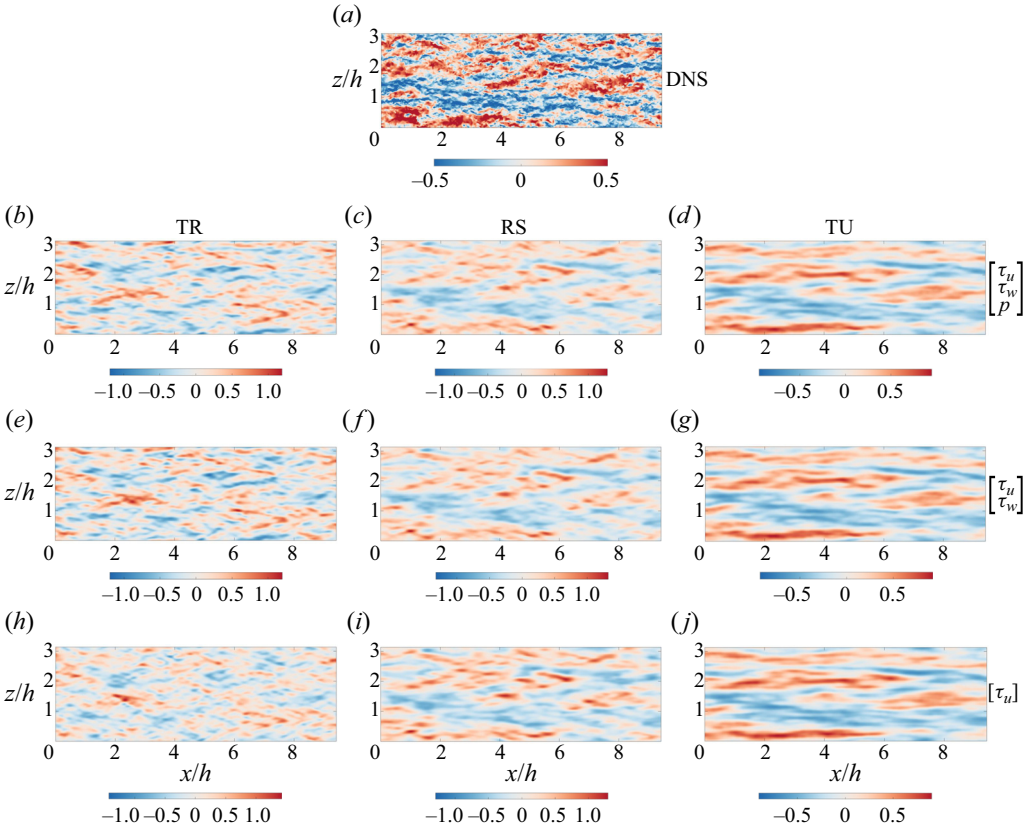


Figure 18. Same as figure 16, but with the λ-model-informed estimator.

mainly appear in the VLSMs, which are more energetic when Re_τ increases. By increasing the measurement types, the estimation accuracy of the turbulence state can be effectively improved. For instance, the relative MSE reduces by 12.3 % when the additional measurements of τ_w and p are added to the original estimations where τ_u is measured in the temporally resolved case. On the other hand, increasing the temporal information of the measurements can also reduce the estimation error. When measuring the wall shear stress and pressure, the relative MSE at $y^+ = 100$ with $Re_\tau = 934$ decreases by 13.0 % in the TR case compared with the TU case with only spatial measurements.

Finally, the current spatio-temporal forcing models, including the B-model, W-model and λ-model, are incorporated into the GRBE estimator to investigate the performances of these existing models when conducting the linear estimation. The estimated results with wall measurements are investigated to test the abilities of the existing forcing models in turbulence state estimation. The estimation errors from the W-model-informed estimator are relatively low in the TU cases, where the estimation error at $y^+ = 100$ in case TU_u is only 13.3 % larger than that from the optimal linear estimations. This indicates that the W-model-predicted spatial statistics of turbulence reflect well the DNS results. However, for the TR cases, the estimation errors in the results from the W-model and λ-model non-physically increase, especially for the estimations with multiple types of observations. Thus, given a forcing model that provides a fair prediction of the spatial covariance of turbulence, its prediction accuracy on the temporal information is not guaranteed, and it

might induce huge errors for state estimation. This highlights the importance of proper modelling strategies of forcing in the temporal domain.

When applied to experiments or engineering scenarios, the flow patterns and sensor locations may not be the same as those in this study. In such cases, the resolvent operator and, correspondingly, the GRBE transfer kernel, should be tuned to match the specific flow configuration and sensor layouts. More importantly, an appropriate forcing model should be incorporated in the GRBE framework such that the linear relationships between the measurement and estimations are properly evaluated in the given application scenario. With continuous studies on the colour of forcing in the frequency domain (e.g. Zare *et al.* 2016, 2020; Morra *et al.* 2021; Nogueira *et al.* 2021), an advanced forcing model that provides a more accurate description of the spatio-temporal properties of turbulence is promising to be proposed in the future, which will make the GRBE a more powerful tool in estimating the turbulence.

Funding. L.F. acknowledges the fund from the Research Grants Council (RGC) of the Government of Hong Kong Special Administrative Region (HKSAR) with RGC/ECS Project (no. 26200222), RGC/GRF Project (no. 16201023) and RGC/STG Project (no. STG2/E-605/23-N), the fund from Guangdong Basic and Applied Basic Research Foundation (no. 2024A1515011798), the fund from Guangdong Province Science and Technology Plan Project (no. 2023A0505030005), the fund from Center for Ocean Research in Hong Kong and Macau, a joint research center between Laoshan Laboratory and HKUST and the fund from the Project of Hetao Shenzhen–Hong Kong Science and Technology Innovation Cooperation Zone (no. HZQB-KCZYB-2020083).

Declaration of interests. The authors report no conflict of interest.

Data availability. The data that support the findings of this study are available on request from the corresponding author, L.F.

Author ORCIDs.
Anjia Ying https://orcid.org/0000-0002-3358-2786;
Zhigang Li https://orcid.org/0000-0001-7164-1049;
Lin Fu https://orcid.org/0000-0001-8979-8415.

Appendix A. The spatial linearised Navier–Stokes operator

The spatial linearised NS operator $\mathcal{L}_{\boldsymbol{k}_s} \in \mathbb{C}^{3N_y \times 3N_y}$ in (2.8*a–c*) for the momentum equations is calculated as

$$\mathcal{L}_{\boldsymbol{k}_s} = \boldsymbol{L}_0 + \mathrm{i}k_x \boldsymbol{L}_x + \boldsymbol{L}_y \mathcal{D} + \mathrm{i}k_z \boldsymbol{L}_z + \boldsymbol{L}_2(-k_x^2 + \mathcal{D}^2 - k_z^2), \tag{A1}$$

with

$$\left.\begin{aligned} \boldsymbol{L}_x &= \begin{bmatrix} U & -v_T' & 0 \\ 0 & U & 0 \\ 0 & 0 & U \end{bmatrix}, \quad \boldsymbol{L}_y = \begin{bmatrix} -v_T' & 0 & 0 \\ 0 & -2v_T' & 0 \\ 0 & 0 & -v_T' \end{bmatrix}, \\ \boldsymbol{L}_z &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -v_T' & 0 & 0 \end{bmatrix}, \quad \boldsymbol{L}_0 = \begin{bmatrix} 0 & u' & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ \boldsymbol{L}_2 &= \begin{bmatrix} -v_T & 0 & 0 \\ 0 & -v_T & 0 \\ 0 & 0 & -v_T \end{bmatrix}, \end{aligned}\right\} \tag{A2}$$

where U is the mean streamwise velocity profile, $v_T = (1/Re_\tau)(\nu_T/\nu)$, $\mathcal{D}$ and $'$ denote $\partial/\partial y$. The Hamiltonian operator $\boldsymbol{\nabla} \in \mathbb{C}^{N_y \times 3N_y}$ in (2.8*a–c*) is calculated as

$$\boldsymbol{\nabla} = \begin{bmatrix} \mathrm{i}k_x \boldsymbol{I}_{N_y} & \mathcal{D} & \mathrm{i}k_z \boldsymbol{I}_{N_y} \end{bmatrix}. \tag{A3}$$

The observation matrix $\boldsymbol{C}$ in (2.7*b*) is determined by the location of the measured variable. For instance, denoting $(u_1, u_2, u_3) = (u, v, w)$, if the measured variable is the velocity u_i $(i = 1, 2, 3)$ at the nth node, then the observation matrix $\boldsymbol{C} \in \mathbb{C}^{1\times 4N_y}$ should be expressed as

$$\boldsymbol{C} = \left[\boldsymbol{O}_{1\times[(i-1)N_y+n-1]} \quad 1 \quad \boldsymbol{O}_{1\times[4N_y-(i-1)N_y-n]}\right]. \tag{A4}$$

Appendix B. The spatio-temporal forcing models

The concept of the spatio-temporal forcing model defined in this study refers to the method that provides the statistical description of forcing in both the wall-normal and temporal directions at each spatial scale $\boldsymbol{k}_s$. Such spatio-temporal forcing models can be incorporated in the GRBE transfer function $\boldsymbol{H}$ as in (2.18) when the measurements are unresolved in time. The existing methods that belong to the spatio-temporal forcing models include the B-model (Hwang & Cossu 2010), W-model and λ-model (Gupta *et al.* 2021). All these three models assume that the forcing is white in time and uncorrelated along the wall-normal direction. This means that the CSD tensors of the forcing at a given spatial scale $\boldsymbol{k}_s$ are diagonal and unchanged all over the frequencies ω. The B-model specifies the forcing energy $E_{(\boldsymbol{k}_s,\omega)}$ to be unchanged along the height, i.e.

$$E_{(\boldsymbol{k}_s,\omega,y)} = \sigma_{\boldsymbol{k}_s}, \quad \forall \omega, y. \tag{B1}$$

With such a precise description of the forcing, the B-model is widely applied in relevant studies (Hwang & Cossu 2010; Madhusudanan *et al.* 2019; Gupta *et al.* 2021).

Extended from the work of Jovanović & Bamieh (2005) in laminar flow, the W-model (Gupta *et al.* 2021) assumes that the vertical profile of forcing energy is proportional to that of the eddy viscosity, i.e.

$$E_{(\boldsymbol{k}_s,\omega,y)} = \sigma_{\boldsymbol{k}_s} \nu_t^2(y), \quad \forall \omega, y. \tag{B2}$$

Considering that the nonlinear interactions of turbulence are scale dependent (Cho, Hwang & Choi 2018), the λ-model (Gupta *et al.* 2021) is further proposed with the modified eddy viscosity, i.e.

$$\nu_{t,\boldsymbol{k}_s}(y) = \frac{\lambda}{\lambda + \lambda_m(y)} \nu_t(y), \tag{B3}$$

where $\lambda = 2\pi/(k_x^2 + k_z^2)^{0.5}$, and $\lambda_m(y) = 50/Re_\tau + (2 - 50/Re_\tau)\tanh(6(1 - |1 - \tilde{y}|))$. The forcing energy profile is thereby calculated by

$$E_{(\boldsymbol{k}_s,\omega,y)} = \sigma_{\boldsymbol{k}_s} \nu_{t,\boldsymbol{k}_s}^2(y), \quad \forall \omega, y. \tag{B4}$$

Appendix C. Analyses of the estimation errors from the RBE in temporally unresolved cases

In this section, the estimation errors from the RBE-F and RBE-T approaches in the TU cases will be analysed. The estimation errors on different frequency bands will be separately investigated to study how the estimation errors are induced from the RBE approaches due to spectrum aliasing. The flow motions on each frequency band are firstly

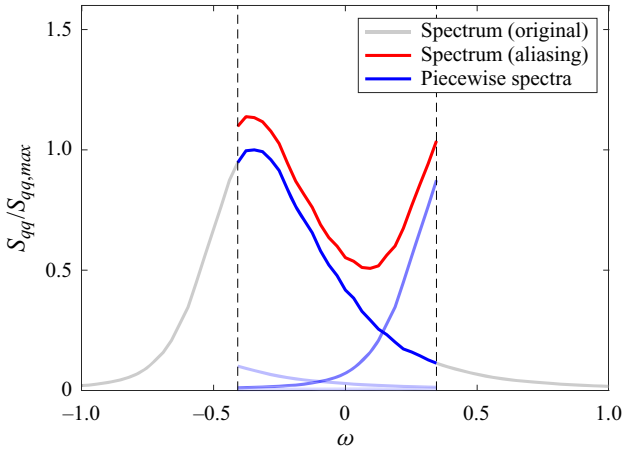


Figure 19. Sketch of spectrum aliasing due to down-sampling of the measurement signal at $x = 10$ of the Ginzburg–Landau equation in the TU case. Grey curve: the original spectrum without aliasing; red curve: the measured spectrum with aliasing; blue curves: the piecewise spectra, with colour shades indicating the numbers of the frequency bands.

extracted from the original flow signal with high temporal resolution, then separately used for estimation in the TU case after down-sampling. To facilitate the following discussions, we denote Pi to be the frequency band with $\omega \in [(2n_i - 1)\omega_N, (2n_i + 1)\omega_N]$, where $n_i = 0, -1, 1, -2, 2, \ldots$ for $i = 1, 2, 3, 4, 5, \ldots$, respectively, and ω_N is the Nyquist frequency. Two cases, including the one-dimensional Ginzburg–Landau equation and turbulent channel flow with $Re_\tau = 934$ with temporally unresolved measurements, will be included in the discussions in this section. Settings of the sampling intervals in the TU cases can be found in §§ 3.1 and 3.2 for the Ginzburg–Landau equation and turbulent channel flows, respectively.

C.1. *The one-dimensional Ginzburg–Landau equation*

Firstly, the spectrum aliasing due to the under-sampling of the measurement data is investigated, which induces the estimation errors in the RBE-F results. Since the RBE-F uses the Fourier series directly from the DFT to conduct the estimations, the flow motions with $\omega = \omega_{P1} + 2n_i\omega_N$ are identified to have a frequency of ω_{P1}, where $|\omega_{P1}| \leqslant \omega_N$ (Karban *et al.* 2022). As a result, the flow motions with $\omega = \omega_{P1} + 2n_i\omega_N$ ($n_i \neq 0$) are incorrectly estimated by RBE-F with the transfer function for ω_{P1}. As shown in figure 19, the under-sampled spectrum in the TU case obviously differs from the original spectra. In figure 20(a), the relative MSE in the RBE-F result corresponding to frequency band P1 with lowest frequencies is the same as that from the optimal GRBE/RBE results in the TR case, and both are lower than the GRBE results in the TU case. This is due to the fact that the RBE-F constructs the transfer function according to P1, while the GRBE comprehensively considers the overall statistics across all the scales. Meanwhile, for the frequency bands with higher frequencies, the relative MSEs from the RBE-F rapidly increases to be larger than unity, which means that the optimal transfer function for band P1 is not effective for the other frequency bands with different frequencies.

On the other hand, the RBE-T uses the discretised temporal transfer kernel by integration of the transfer kernel in the temporally resolved case, as shown in (2.30). From figure 20(a),the relative MSE of RBE-T at frequency band P1 is the largest one among

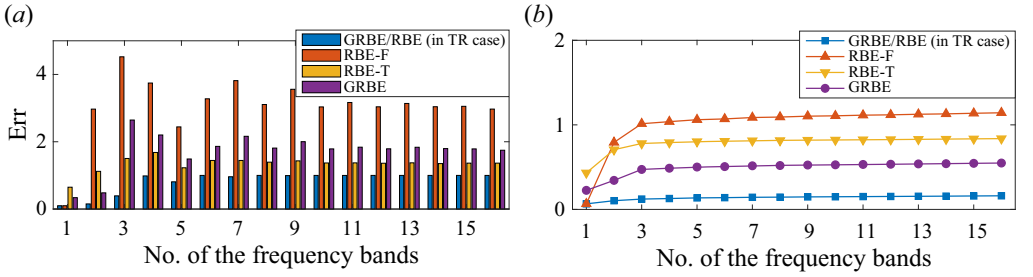


Figure 20. Relative MSEs at $x = 10$ for different frequency bands of the Ginzburg–Landau equation in the TU case. (*a*) The estimation error at separated frequency bands. (*b*) The accumulated estimation error by sequentially increasing the considered numbers of frequency bands.

those from the other methods. For the sequential frequency bands, the estimation errors of the RBE-T approach appear to be lower than those of the RBE-F approach. Meanwhile, both the RBE-F and RBE-T results have larger estimation errors than the GRBE approach when considering the summation of flow motions at all the frequency bands.

Appendix D. The H_2-optimal estimation

The H_2-optimal estimation (Illingworth *et al.* 2018) originates from the H_2-optimal control problem (Zhou & Doyle 1998), where the term 'H_2' is from the H_2 norm that is used to define the cost function in this problem. The H_2 norm of a stable matrix $M(s)$ is defined by

$$\|M\|_2 := \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{Tr}\left[M(\mathrm{i}\omega) M(\mathrm{i}\omega)^* \right] \mathrm{d}\omega \right)^{1/2}. \tag{D1}$$

Now consider a dynamic system with the state evolution equation expressed by

$$\left.\begin{aligned} \frac{\partial \boldsymbol{u}(t)}{\partial t} &= A\boldsymbol{u}(t) + B_1 \boldsymbol{w}(t) + B_2 \boldsymbol{u}_e(t), \\ \boldsymbol{e}(t) &= C_1 \boldsymbol{u}(t) + D_{12} \boldsymbol{u}_e(t), \\ \boldsymbol{y}(t) &= C_2 \boldsymbol{u}(t) + D_{21} \boldsymbol{w}(t), \end{aligned}\right\} \tag{D2}$$

where $\boldsymbol{u}$ is the state variable; $\boldsymbol{u}_e$ is the control input, which corresponds to the estimation for the optimal estimation problems; $\boldsymbol{w}$ is the spatio-temporally white external input, which includes the effects of disturbances, i.e. nonlinear forcing and sensor noise. Note that the cross-spectra of nonlinear forcing and the sensor noise are defined with $\boldsymbol{w}$ along with the matrices of $\boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{D}}$ introduced below. Here, $\boldsymbol{e}$ is the output of the closed-loop system, which corresponds to the estimation error for the optimal estimation problems and $\boldsymbol{y}$ is the measurement. A diagram of the closed-loop system is shown in figure 21, where $G(s)$ denotes the state-space representation of

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = C\left(s\boldsymbol{I} - A\right)^{-1} B + D, \tag{D3}$$

with

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \quad D = \begin{bmatrix} \boldsymbol{O} & D_{12} \\ D_{21} & \boldsymbol{O} \end{bmatrix}. \tag{D4a–c}$$

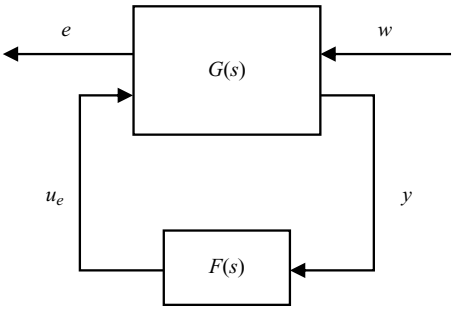


Figure 21. Diagram of the closed-loop system.

The transfer matrix $G(s)$ actually defines the linear relationship between $[\boldsymbol{e}, \boldsymbol{y}]^{\mathrm{T}}$ and $[\boldsymbol{w}, \boldsymbol{u}_e]^{\mathrm{T}}$ after taking Laplace transforms along the temporal direction, i.e.

$$\begin{bmatrix} \hat{\boldsymbol{e}}(s) \\ \hat{\boldsymbol{y}}(s) \end{bmatrix} = G(s) \begin{bmatrix} \hat{\boldsymbol{w}}(s) \\ \hat{\boldsymbol{u}}_e(s) \end{bmatrix}. \tag{D5}$$

For output estimation problems, with respect to the parameters in (2.7), the linear operators and the external input $\boldsymbol{w}$ in (D2) are given by

$$\left.\begin{aligned} A = \boldsymbol{A}, \quad B_1 = [\boldsymbol{\mathcal{B}} \quad \boldsymbol{O}], \quad B_2 = [\boldsymbol{O}], \quad C_1 = [\boldsymbol{I}], \quad C_2 = [\boldsymbol{C}], \\ D_{12} = [-\boldsymbol{I}], \quad D_{21} = [\boldsymbol{O} \quad \boldsymbol{\mathcal{D}}], \quad \boldsymbol{w} = \begin{bmatrix} \tilde{\boldsymbol{f}} \\ \tilde{\boldsymbol{n}} \end{bmatrix}, \end{aligned}\right\} \tag{D6}$$

where $\boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{D}}$ are obtained from the Cholesky decompositions of the CSD tensors of forcing and measurement noise, respectively, i.e. $\boldsymbol{S}_{\boldsymbol{ff}} = \boldsymbol{\mathcal{B}} \cdot \boldsymbol{\mathcal{B}}^*$, $\boldsymbol{S}_{\boldsymbol{nn}} = \boldsymbol{\mathcal{D}} \cdot \boldsymbol{\mathcal{D}}^*$.

Equation (D5) describes the mathematical definitions of the estimation error $\hat{\boldsymbol{e}}(s)$ and observation $\hat{\boldsymbol{y}}(s)$ as functions of the external input $\hat{\boldsymbol{w}}(s)$ and the state estimation $\hat{\boldsymbol{u}}_e(s)$ through the linear operator $G(s)$. Considering (D3)–(D6), (D5) is equivalent to

$$\left.\begin{aligned} \hat{\boldsymbol{e}}(s) &= \hat{\boldsymbol{u}}(s) - \hat{\boldsymbol{u}}_e(s), \\ \hat{\boldsymbol{y}}(s) &= \boldsymbol{C}\hat{\boldsymbol{u}}(s) + \boldsymbol{\mathcal{D}}\hat{\tilde{\boldsymbol{n}}}(s), \end{aligned}\right\} \tag{D7}$$

where $\hat{\boldsymbol{u}}(s) = (s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{\mathcal{B}}\hat{\tilde{\boldsymbol{f}}}(s)$ is the response generated by the stochastic forcing that is defined as $\boldsymbol{\mathcal{B}}\hat{\tilde{\boldsymbol{f}}}(s)$, and $\boldsymbol{\mathcal{D}}\hat{\tilde{\boldsymbol{n}}}(s)$ defines the sensor noise.

Now define the controller $F(s)$ and transfer matrix $T_{\boldsymbol{ew}}(s)$, i.e.

$$\left.\begin{aligned} \hat{\boldsymbol{u}}_e(s) &= F(s)\hat{\boldsymbol{y}}(s), \\ \hat{\boldsymbol{e}}(s) &= T_{\boldsymbol{ew}}(s)\hat{\boldsymbol{w}}(s), \end{aligned}\right\} \tag{D8}$$

where the explicit expression of $T_{\boldsymbol{ew}}$ can be derived to be

$$T_{\boldsymbol{ew}}(s) = \left[(\boldsymbol{I} - F(s)\boldsymbol{C}) (s\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{\mathcal{B}}, \quad -F(s)\boldsymbol{\mathcal{D}} \right], \tag{D9}$$

according to (D5). The goal of the H_2 control problem includes finding a rational controller $F(s)$ that minimises the H_2 norm of the transfer matrix $T_{\boldsymbol{ew}}$ that links the external input $\boldsymbol{w}$ and the estimation error $\boldsymbol{e}$ (Zhou & Doyle 1998), which directly builds up the foundation of the H_2-optimal estimation in Illingworth *et al.* (2018). Details of the mechanism and

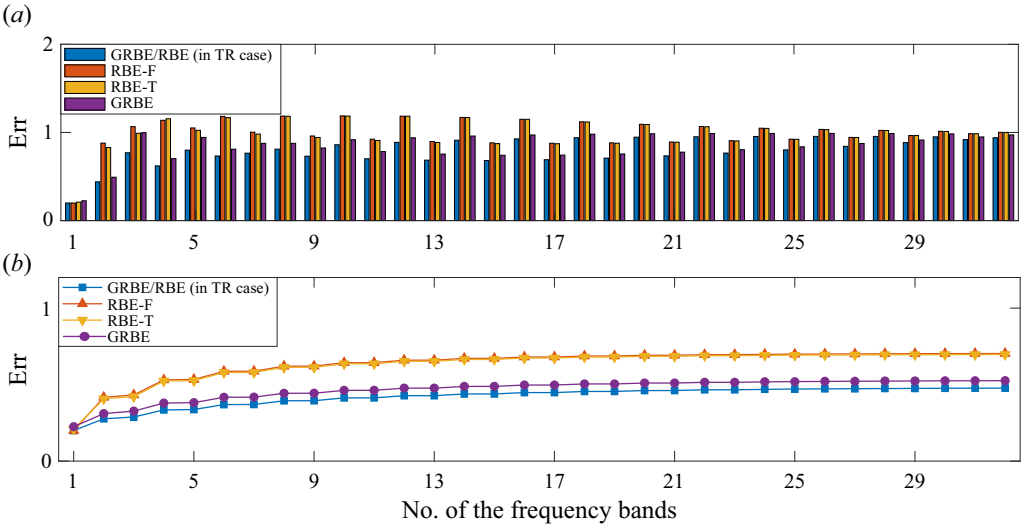


Figure 22. Relative MSEs at $y^+ = 100$ for different frequency bands of the turbulent channel flow with $Re_\tau = 934$ in the TU case. (*a*) The estimation error at separated frequency bands. (*b*) The accumulated estimation error by sequentially increasing the considered numbers of frequency bands.

methodology for solving the H_2 optimisation problem can be found in Doyle *et al.* (1989) or Zhou & Doyle (1998). The algorithm for solving the H_2-optimal problem has been integrated in MATLAB and can be solved via the function `h2syn`, i.e.

$$[F, T_{\boldsymbol{ew}}, \gamma] = \texttt{h2syn}(G, n_{\boldsymbol{y}}, n_{\boldsymbol{u_e}}), \tag{D10}$$

where G is assigned in MATLAB via $G = \texttt{ss}(A, B, C, D)$, $n_{\boldsymbol{y}}$ and $n_{\boldsymbol{u_e}}$ are the dimensions of the measurement $\boldsymbol{y}$ and estimation $\boldsymbol{u}_e$, respectively, and γ is the resultant minimum H_2 norm of $T_{\boldsymbol{ew}}$.

According to Doyle *et al.* (1989) and Zhou & Doyle (1998), the H_2-optimal approach provides causal estimations, and utilises the measurement information at $\tau \leqslant t$ when estimating the flow state at time t. From the definition of the H_2 norm in (D1), the expected estimation error energy is minimised only when the forcing is white in time, i.e. the CSD tensors of the forcing at a given spatial scale $\boldsymbol{k}_s$ are unchanged for all the temporal frequencies.

D.1. *The turbulent channel flow with* $Re_\tau = 934$

The distributions of the estimation errors at $y^+ = 100$ with $Re_\tau = 934$ in the TU case for different frequency bands are shown in figure 22. As in figure 22(*a*), the RBE-F results have the smallest estimation error at P1, since the RBE-F utilises the optimal transfer kernels for P1 to estimate the flow state at all the sequential frequency bands. From the second frequency band P2 on, the relative MSEs of the RBE-F results become larger than those from the GRBE. Meanwhile, the RBE-T results nearly overlap with the RBE-F results in almost all the considered frequency bands. On the other hand, the GRBE results in the TU and TR cases are pretty close to each other for the turbulence estimation.

In figure 23, the instantaneous flow states at $y^+ = 100$ within different frequency bands are depicted. For the first frequency band P1 with the lowest frequencies, all the tested methods provide reasonable results. For the sequential frequency bands P2 and P8, the patterns of the flow structures become different in the GRBE and RBE results.

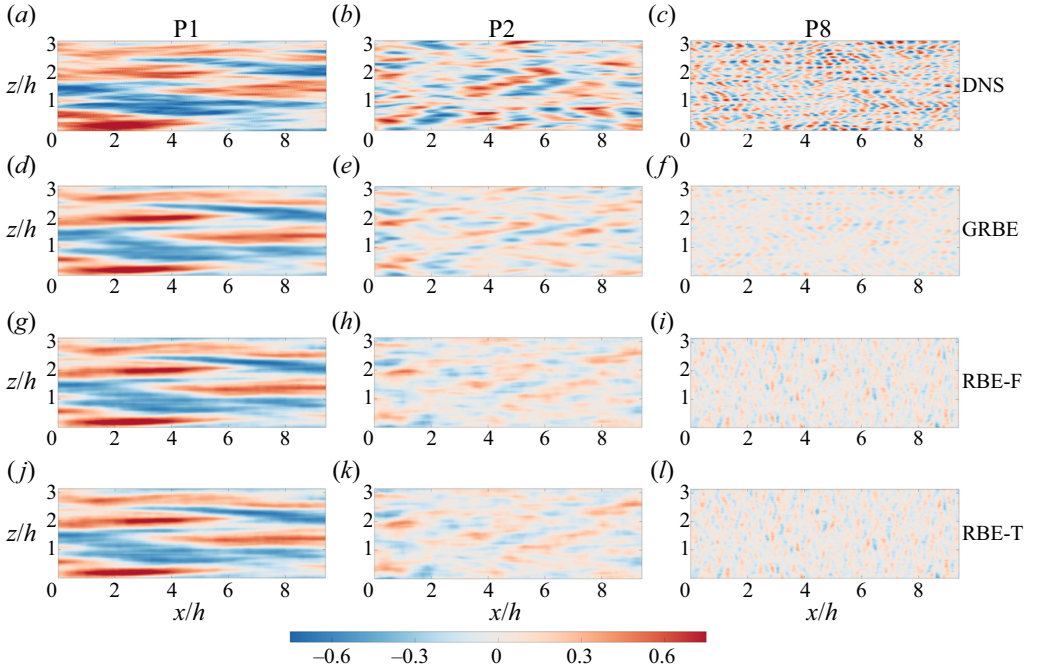


Figure 23. Instantaneous flow state at $y^+ = 100$ in the TU case corresponding to the frequency bands of P1 (*a*,*d*,*g*,*j*), P2 (*b*,*e*,*h*,*k*) and P8 (*c*,*f*,*i*,*l*). The results are from DNS (*a–c*), GRBE (*d–f*), RBE-F (*g–i*) and RBE-T (*j–l*).

Although the estimated flow energies are lower than those from the DNS, the flow patterns of the GRBE results in the sequential frequency bands keep consistent with the DNS ones. Meanwhile, the flow patterns estimated by the RBE-F and RBE-T become different from the DNS results in the sequential frequency bands. Specifically, for P8 with higher frequencies, flow structures that are elongated in the spanwise direction can be observed in the RBE results, which are considered to contribute to the non-physical small-scale spanwise streaks in the RBE-T results shown in figure 6.

References

Adrian, R.J., Meinhart, C.D. & Tomkins, C.D. 2000 Vortex organization in the outer region of the turbulent boundary layer. *J. Fluid Mech.* **422**, 1–54.

Amaral, F.R., Cavalieri, A.V.G., Martini, E., Jordan, P. & Towne, A. 2021 Resolvent-based estimation of turbulent channel flow using wall measurements. *J. Fluid Mech.* **927**, A17.

Baars, W.J., Hutchins, N. & Marusic, I. 2016 Spectral stochastic estimation of high-Reynolds-number wall-bounded turbulence for a refined inner-outer interaction model. *Phys. Rev. Fluids* **1** (5), 054406.

Bae, H.J., Lozano-Durán, A., Bose, S.T. & Moin, P. 2018 Turbulence intensities in large-eddy simulation of wall-bounded flows. *Phys. Rev. Fluids* **3**, 014610.

Bagheri, S., Brandt, L. & Henningson, D.S. 2009 Input–output analysis, model reduction and control of the flat-plate boundary layer. *J. Fluid Mech.* **620**, 263–298.

Cameron, S.M. 2011 PIV algorithms for open-channel turbulence research: accuracy, resolution and limitations. *J. Hydro-Environ. Res.* **5** (4), 247–262.

Cess, R.D. 1958 A survey of the literature on heat transfer in turbulent tube flow. *Tech. Rep.* 8–0529–R24.

Chen, X., Cheng, C., Fu, L. & Gan, J. 2023*a* Linear response analysis of supersonic turbulent channel flows with a large parameter space. *J. Fluid Mech.* **962**, A7.

Chen, X., Cheng, C., Gan, J. & Fu, L. 2023*b* Study of the linear models in estimating coherent velocity and temperature structures for compressible turbulent channel flows. *J. Fluid Mech.* **973**, A36.

CHENG, C. & FU, L. 2022 Consistency between the attached-eddy model and the inner–outer interaction model: a study of streamwise wall-shear stress fluctuations in a turbulent channel flow. *J. Fluid Mech.* **942**, R9.

CHENG, C., LI, W., LOZANO-DURÁN, A. & LIU, H. 2019 Identity of attached eddies in turbulent channel flows with bidimensional empirical mode decomposition. *J. Fluid Mech.* **870**, 1037–1071.

CHENG, C., SHYY, W. & FU, L. 2022 Streamwise inclination angle of wall-attached eddies in turbulent channel flows. *J. Fluid Mech.* **946**, A49.

CHEVALIER, M., HŒPFFNER, J., BEWLEY, T.R. & HENNINGSON, D.S. 2006 State estimation in wall-bounded flow systems. Part 2. Turbulent flows. *J. Fluid Mech.* **552**, 167–187.

CHO, M., HWANG, Y. & CHOI, H. 2018 Scale interactions and spectral energy transfer in turbulent channel flow. *J. Fluid Mech.* **854**, 474–504.

COLBURN, C.H., CESSNA, J.B. & BEWLEY, T.R. 2011 State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter. *J. Fluid Mech.* **682**, 289–303.

CRANK, J. & NICOLSON, P. 1947 A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. In *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 43, pp. 50–67. Cambridge University Press.

DEL ALAMO, J.C. & JIMÉNEZ, J. 2003 Spectra of the very large anisotropic scales in turbulent channels. *Phys. Fluids* **15** (6), L41–L44.

DOYLE, J.C., GLOVER, K., KHARGONEKAR, P.P. & FRANCIS, B.A. 1989 State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems. *IEEE Trans. Automat. Control* **34** (8), 831–847.

FU, L., BOSE, S. & MOIN, P. 2022 Prediction of aerothermal characteristics of a generic hypersonic inlet flow. *Theor. Comput. Fluid Dyn.* **36** (2), 345–368.

FU, L., KARP, M., BOSE, S.T., MOIN, P. & URZAY, J. 2021 Shock-induced heating and transition to turbulence in a hypersonic boundary layer. *J. Fluid Mech.* **909**, A8.

GRIFFIN, K.P., FU, L. & MOIN, P. 2023 Near-wall model for compressible turbulent boundary layers based on an inverse velocity transformation. *J. Fluid Mech.* **970**, A36.

GUASTONI, L., GÜEMES, A., IANIRO, A., DISCETTI, S., SCHLATTER, P., AZIZPOUR, H. & VINUESA, R. 2021 Convolutional-network models to predict wall-bounded turbulence from wall quantities. *J. Fluid Mech.* **928**, A27.

GUPTA, V., MADHUSUDANAN, A., WAN, M., ILLINGWORTH, S.J. & JUNIPER, M.P. 2021 Linear-model-based estimation in wall turbulence: improved stochastic forcing and eddy viscosity terms. *J. Fluid Mech.* **925**, A18.

HŒPFFNER, J., CHEVALIER, M., BEWLEY, T.R. & HENNINGSON, D.S. 2005 State estimation in wall-bounded flow systems. Part 1. Perturbed laminar flows. *J. Fluid Mech.* **534**, 263–294.

HOYAS, S. & JIMÉNEZ, J. 2008 Reynolds number effects on the Reynolds-stress budgets in turbulent channels. *Phys. Fluids* **20** (10), 101511.

HUMPHERYS, J., REDD, P. & WEST, J. 2012 A fresh look at the Kalman filter. *SIAM Rev.* **54** (4), 801–823.

HWANG, Y. 2015 Statistical structure of self-sustaining attached eddies in turbulent channel flow. *J. Fluid Mech.* **767**, 254–289.

HWANG, Y. & COSSU, C. 2010 Linear non-normal energy amplification of harmonic and stochastic forcing in the turbulent channel flow. *J. Fluid Mech.* **664**, 51–73.

ILLINGWORTH, S.J., MONTY, J.P. & MARUSIC, I. 2018 Estimating large-scale structures in wall turbulence using linear models. *J. Fluid Mech.* **842**, 146–162.

JOVANOVIĆ, M.R. & BAMIEH, B. 2005 Componentwise energy amplification in channel flows. *J. Fluid Mech.* **534**, 145–183.

KALMAN, R.E. 1960 A new approach to linear filtering and prediction problems. *J. Basic Engng* **82** (1), 35–45.

KARBAN, U., MARTINI, E., JORDAN, P., BRÈS, G.A. & TOWNE, A. 2022 Solutions to aliasing in time-resolved flow data. *Theor. Comput. Fluid Dyn.* **36** (6), 887–914.

LARSSON, J., KAWAI, S., BODART, J. & BERMEJO-MORENO, I. 2016 Large eddy simulation with modeled wall-stress: recent progress and future directions. *Mech. Engng Rev.* **3** (1), 15–00418.

LEE, J., LEE, J.H., CHOI, J.-I. & SUNG, H.J. 2014 Spatial organization of large-and very-large-scale motions in a turbulent channel flow. *J. Fluid Mech.* **749**, 818–840.

LEE, J.H. & SUNG, H.J. 2013 Comparison of very-large-scale motions of turbulent pipe and boundary layer simulations. *Phys. Fluids* **25**, 045103.

LOZANO-DURÁN, A., BAE, H.J. & ENCINAR, M.P. 2020 Causality of energy-containing eddies in wall turbulence. *J. Fluid Mech.* **882**, A2.

LOZANO-DURÁN, A., FLORES, O. & JIMÉNEZ, J. 2012 The three-dimensional structure of momentum transfer in turbulent channels. *J. Fluid Mech.* **694**, 100–130.

LUHAR, M., SHARMA, A.S. & MCKEON, B.J. 2014*a* On the structure and origin of pressure fluctuations in wall turbulence: predictions based on the resolvent analysis. *J. Fluid Mech.* **751**, 38–70.
LUHAR, M., SHARMA, A.S. & MCKEON, B.J. 2014*b* Opposition control within the resolvent analysis framework. *J. Fluid Mech.* **749**, 597–626.
MADHUSUDANAN, A., ILLINGWORTH, S.J. & MARUSIC, I. 2019 Coherent large-scale structures from the linearized Navier–Stokes equations. *J. Fluid Mech.* **873**, 89–109.
MARTINI, E., CAVALIERI, A.V.G., JORDAN, P., TOWNE, A. & LESSHAFFT, L. 2020 Resolvent-based optimal estimation of transitional and turbulent flows. *J. Fluid Mech.* **900**, A2.
MARUSIC, I., MATHIS, R. & HUTCHINS, N. 2010 Predictive model for wall-bounded turbulent flow. *Science* **329** (5988), 193–196.
MCKEON, B.J. & SHARMA, A.S. 2010 A critical-layer framework for turbulent pipe flow. *J. Fluid Mech.* **658**, 336–382.
MOARREF, R., SHARMA, A.S., TROPP, J.A. & MCKEON, B.J. 2013 Model-based scaling of the streamwise energy density in high-Reynolds-number turbulent channels. *J. Fluid Mech.* **734**, 275–316.
MORRA, P., NOGUEIRA, P.A., CAVALIERI, A.V.G. & HENNINGSON, D.S 2021 The colour of forcing statistics in resolvent analyses of turbulent channel flows. *J. Fluid Mech.* **907**, A24.
MORRA, P., SEMERARO, O., HENNINGSON, D.S. & COSSU, C. 2019 On the relevance of Reynolds stresses in resolvent analyses of turbulent wall-bounded flows. *J. Fluid Mech.* **867**, 969–984.
NOBACH, H. & BODENSCHATZ, E. 2009 Limitations of accuracy in PIV due to individual variations of particle image intensities. *Exp. Fluids* **47**, 27–38.
NOGUEIRA, P.A.S., MORRA, P., MARTINI, E., CAVALIERI, A.V.G. & HENNINGSON, D.S. 2021 Forcing statistics in resolvent analysis: application in minimal turbulent couette flow. *J. Fluid Mech.* **908**, A32.
OEHLER, S.F. & ILLINGWORTH, S.J. 2021 Linear control of coherent structures in wall-bounded turbulence at $Re_\tau = 2000$. *Intl J. Heat Fluid Flow* **87**, 108735.
PICKERING, E., RIGAS, G., SCHMIDT, O.T., SIPP, D. & COLONIUS, T. 2021 Optimal eddy viscosity for resolvent-based models of coherent structures in turbulent jets. *J. Fluid Mech.* **917**, A29.
REYNOLDS, W.C. & HUSSAIN, A.K.M.F. 1972 The mechanics of an organized wave in turbulent shear flow. Part 3. Theoretical models and comparisons with experiments. *J. Fluid Mech.* **54** (2), 263–288.
SHARMA, A.S. & MCKEON, B.J. 2013 On coherent structure in wall turbulence. *J. Fluid Mech.* **728**, 196–238.
TOWNE, A., LOZANO-DURÁN, A. & YANG, X. 2020 Resolvent-based estimation of space–time flow statistics. *J. Fluid Mech.* **883**, A17.
TOWNE, A., SCHMIDT, O.T. & COLONIUS, T. 2018 Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *J. Fluid Mech.* **847**, 821–867.
TOWNSEND, A.A. 1976 *The Structure of Turbulent Shear Flow*, 2nd edn. Cambridge University Press.
WIENER, N. 1930 Generalized harmonic analysis. *Acta Mathematica* **55** (1), 117–258.
WIENER, N. 1949 *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*. MIT.
YING, A., LIANG, T., LI, Z. & FU, L. 2023 A resolvent-based prediction framework for incompressible turbulent channel flow with limited measurements. *J. Fluid Mech.* **976**, A31.
ZARE, A., CHEN, Y., JOVANOVIĆ, M.R. & GEORGIOU, T.T. 2016 Low-complexity modeling of partially available second-order statistics: theory and an efficient matrix completion algorithm. *IEEE Trans. Autom. Control* **62** (3), 1368–1383.
ZARE, A., GEORGIOU, T.T. & JOVANOVIĆ, M.R. 2020 Stochastic dynamical modeling of turbulent flows. *Annu. Rev. Control Robot. Auton. Syst.* **3**, 195–219.
ZARE, A., JOVANOVIĆ, M.R. & GEORGIOU, T.T. 2017 Colour of turbulence. *J. Fluid Mech.* **812**, 636–680.
ZHOU, K. & DOYLE, J.C. 1998 *Essentials of Robust Control*, vol. 104. Prentice Hall.