

# Data-driven closure of the harmonic-balanced Navier-Stokes equations in the frequency domain

By G. Rigas<sup>†</sup>, P. J. Schmid<sup>‡</sup>

The Fourier-Galerkin method is employed to calculate the multifrequency and multiscale asymptotic nonlinear flow response in the frequency domain, by expanding the solution as Fourier series. The resulting equations are known as the harmonic-balanced Navier-Stokes (HBNS) equations. Although near the threshold of transition a small number of harmonics suffice to achieve convergence, further away the computational cost becomes intractable because energy is transferred to higher harmonics, which can no longer be neglected. In this study, we propose a data-driven framework to model the residual (nonlocal closure) terms for the frequency-truncated HBNS equations. By splitting the sought solution into low-frequency (resolved) and high-frequency (unresolved) harmonics, we systematically express the low-frequency residual as a function of the resolved frequency harmonics only. A consistent deep learning architecture, which parameterizes the residual function, is designed and trained using high-fidelity results near the thresholds of transition for two-dimensional (2D) cylinder flow. We show that our proposed framework achieves low generalization error by predicting accurately the coarse-grained residual for unseen Reynolds numbers, and significantly reduces the computational cost by solving accurately for the coarse-grained dynamics.

---

## 1. Introduction

Coherent structures in turbulent flows can be directly linked to instabilities in the transitional regime or to instabilities of the turbulent mean flow. For small instability amplitudes, and typically close to the transitional thresholds of instability, linear and weakly nonlinear analyses of the governing Navier-Stokes equations accurately predict the evolution of perturbations. For the saturated regime, fully nonlinear approaches are required to predict the evolution of multiscale and multifrequency coherent structures. However, at realistic turbulent regimes, direct numerical simulation (DNS) of the Navier-Stokes equations is intractable, despite the advent of exascale computing. Thus, modelling is required, yielding typically lower-fidelity models by parameterizing the small-scale interactions with the large-scale structures in subgrid models. In recent years, super-resolution tasks for chaotic and turbulent flows have been effectively addressed using physics-informed machine learning techniques by learning functions or operators that accurately reconstruct or correct coarse-grained solutions (see, e.g., Bar-Sinai *et al.* (2019); Kelshaw *et al.* (2022)).

At transitional regimes, simulating the Navier-Stokes equations in the frequency domain (Rigas *et al.* 2021) is an efficient and complementary alternative to time-domain methods (Kerswell 2018). Periodic solutions of the Navier-Stokes equations are sought in the frequency domain using the harmonic balance method, which is a version of Galerkin's

<sup>†</sup> Department of Aeronautics, Imperial College London, United Kingdom

<sup>‡</sup> Department of Mechanical Engineering, King Abdullah University of Science and Technology, Saudi Arabia

weighted residual method. The periodic solution is approximated by a truncated Fourier series, and the governing HBNS equations are obtained by requiring that the Fourier coefficients of the residual term vanish.

Retaining only one fundamental harmonic (and the mean flow) of the Fourier series expansion is known as quasi-Linear modelling (Mantič-Lugo & Gallaire 2016; Marston & Tobias 2023), where a self-consistent solution is obtained by solving the coupled nonlinear mean flow and the linear perturbation equation. The methodology can be generalized to an arbitrary number of frequency harmonics (Rigas *et al.* 2021; Bengana & Tuckerman 2021; Sierra-Ausin *et al.* 2022), which is necessary to capture the generation of multiple harmonics and spectral energy transfer. Although at early transitional regimes a few harmonics suffice to obtain a converged solution, during and after the transition to chaotic or turbulent regimes, a large number of harmonics is required.

In this study, we develop a methodology to learn closures for unsteady flows in the frequency domain by data-driven modeling of the residual HBNS terms for the coarse-grained dynamics. The remainder of this report is structured as follows. In Section 2, the derivation of the HBNS equations is outlined. In Section 3, we present the data-driven framework for modelling the residual terms due to the frequency (spectral) truncation of the HBNS equations. In Section 4, we apply the proposed framework to the 2D flow past a cylinder. Conclusions are drawn in Section 5.

## 2. Harmonic-balanced Navier-Stokes equations

We consider the forced three-dimensional (3D) incompressible Navier-Stokes equations

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \nu \Delta \mathbf{u} + \mathbf{f}(\mathbf{x}, t), \quad (2.1a)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2.1b)$$

where  $\mathbf{f}$  is a volumetric time-dependent momentum forcing. We discretize the inhomogeneous spatial directions  $x$  and  $y$  with finite elements using FreeFem++ (Hecht 2012). The semi-discretized equations can be written in compact form, introducing the state vector  $\mathbf{w} = [\mathbf{u}, p] = [u, v, w, p]$ , as

$$\mathbf{M} \partial_t \mathbf{w} + \mathbf{L} \mathbf{w} + \frac{1}{2} \mathbf{N}(\mathbf{w}, \mathbf{w}) = \mathbf{M} \mathbf{P} \mathbf{f}(z, t), \quad (2.2)$$

where  $\mathbf{P}$  is the prolongation matrix mapping the  $[u, v, w]$  velocity vector into the  $[u, v, w, 0]$  velocity-pressure vector. The matrices  $\mathbf{M}$  and  $\mathbf{L}$  and the bilinear operator  $\mathbf{N}$  are defined as

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}' & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} -\nu \Delta(\cdot) & \nabla(\cdot) \\ \nabla \cdot (\cdot) & 0 \end{pmatrix}, \quad \mathbf{N}(\mathbf{w}_1, \mathbf{w}_2) = \begin{pmatrix} \mathbf{u}_1 \cdot \nabla \mathbf{u}_2 + \mathbf{u}_2 \cdot \nabla \mathbf{u}_1 \\ 0 \end{pmatrix},$$

where  $\mathbf{M}$  and  $\mathbf{M}'$  are the mass matrices associated with the spatial finite-element discretization,  $\mathbf{L}$  denotes the Stokes operator, and  $\mathbf{N}$  represents the symmetrized nonlinear convection operator.

For 3D flows, the spanwise direction can be discretized, i.e. by using a finite element or Fourier basis. In the latter case, a harmonic balance can also be performed in  $z$  to limit the number of retained spanwise harmonics. Without loss of generality, here we consider linearly globally unstable 2D flows, leading to self-sustained dynamics in the absence of forcing  $\mathbf{f}$ . In this case, we can obtain  $N$ -periodic oscillatory dynamics, without modeling or optimizing for the volumetric forcing by setting  $\mathbf{f} = 0$ .

2.1. Harmonic balance in  $\omega$ 

We consider solutions approximated by  $N$  harmonics in the frequency domain

$$\mathbf{w}(t) = \sum_{-N \leq n \leq N} e^{in\omega t} \hat{\mathbf{w}}_n. \quad (2.3)$$

Using the above expansion, the Navier-Stokes equations can be written in the frequency domain applying the Fourier-Galerkin method (harmonic balance) as

$$\mathbf{R}(\hat{\mathbf{w}}) \equiv [in\omega\mathbf{M} + \mathbf{L} + \gamma_0^n \mathbf{N}(\hat{\mathbf{w}}_0, \cdot)] \hat{\mathbf{w}}_n + \sum_{\substack{n=n_1+n_2 \\ -N \leq n_1 \leq n_2 \leq N \\ n_1, n_2 \neq 0}} \gamma_{n_1}^{n_2} \mathbf{N}(\hat{\mathbf{w}}_{n_1}, \hat{\mathbf{w}}_{n_2}) = 0. \quad (2.4)$$

The above system of equations, known as the HBNS equations, is nonlinearly coupled through triadic interactions and self-consistent (closed). The coefficients  $\gamma_{n_1}^{n_2}$  are 0.5, if  $n_1 = n_2$ , and 1 otherwise. The solution vector  $\hat{\mathbf{w}}$  has  $(2N + 1)$  complex unknowns

$$\hat{\mathbf{w}} = \left( \hat{\mathbf{w}}_0, \hat{\mathbf{w}}_1, \hat{\mathbf{w}}_{-1}, \dots, \hat{\mathbf{w}}_N, \hat{\mathbf{w}}_{-N} \right)^T. \quad (2.5)$$

Since  $\mathbf{w}$  is real, the complex harmonics satisfy  $\hat{\mathbf{w}}_n = \overline{\hat{\mathbf{w}}_{-n}}$ , where the overbar denotes complex conjugate, and the mean flow  $\hat{\mathbf{w}}_0$  is real. The above system corresponds to an extension of the quasi-linear self-consistent model (Mantič-Lugo & Gallaire 2016) to an arbitrary number of truncated harmonics (Rigas *et al.* 2021; Bengana & Tuckerman 2021; Sierra-Ausin *et al.* 2022).

For example, for  $N = 4$ , we obtain

$$\begin{aligned} \left[ \mathbf{L} + \frac{1}{2} \mathbf{N}(\hat{\mathbf{w}}_0, \cdot) \right] \hat{\mathbf{w}}_0 + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_1) + \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_2) + \mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_3) + \mathbf{N}(\overline{\hat{\mathbf{w}}_4}, \hat{\mathbf{w}}_4) &= 0, \\ [i\omega\mathbf{M} + \mathbf{L} + \mathbf{N}(\hat{\mathbf{w}}_0, \cdot)] \hat{\mathbf{w}}_1 + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_2) + \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_3) + \mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_4) &= 0, \\ [2i\omega\mathbf{M} + \mathbf{L} + \mathbf{N}(\hat{\mathbf{w}}_0, \cdot)] \hat{\mathbf{w}}_2 + \frac{1}{2} \mathbf{N}(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_1) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_3) + \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_4) &= 0, \\ [3i\omega\mathbf{M} + \mathbf{L} + \mathbf{N}(\hat{\mathbf{w}}_0, \cdot)] \hat{\mathbf{w}}_3 + \mathbf{N}(\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2) + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_4) &= 0, \\ [4i\omega\mathbf{M} + \mathbf{L} + \mathbf{N}(\hat{\mathbf{w}}_0, \cdot)] \hat{\mathbf{w}}_4 + \frac{1}{2} \mathbf{N}(\hat{\mathbf{w}}_2, \hat{\mathbf{w}}_2) + \mathbf{N}(\hat{\mathbf{w}}_3, \hat{\mathbf{w}}_1) &= 0. \end{aligned}$$

A solution to the nonlinear system of equations, spectrally truncated at a finite number of harmonics  $N$ , is obtained with a Newton method. An initial guess  $\hat{\mathbf{w}}_i$  may be improved according to  $\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i - \delta\hat{\mathbf{w}}_i$  by solving

$$\mathbf{A}\delta\hat{\mathbf{w}}_i = \mathbf{R}(\hat{\mathbf{w}}_i),$$

where  $\mathbf{A} = \partial\mathbf{R}/\partial\hat{\mathbf{w}}$  is the Jacobian of the operator  $\mathbf{R}$ , known as the finite-dimensional block Hill matrix. More details regarding efficient iterative methods to solve this system using a preconditioned generalized minimal residual method are described by Rigas *et al.* (2021) and Sierra-Ausin *et al.* (2022).

In the case of flow past a cylinder with the threshold of the first Hopf bifurcation at  $R = 47$ , the flow is approximated by a fundamental frequency and multiple higher harmonics. The HBNS system of equations (Eq. (2.4)) is augmented with a scalar equation for determining the unknown fundamental frequency of the limit-cycle oscillations by fixing the phase of the cycle. Several choices are possible; here, we follow Fabre *et al.* (2018) and set the lift force at  $t = 0$  to be a maximum.

### 3. Data-driven modeling of super-spectral scales

We introduce the following decomposition into low-frequency,  $\mathbf{w}_l$ , and high-frequency,  $\mathbf{w}_h$ , components such that

$$\mathbf{w} = \mathbf{w}_l + \mathbf{w}_h. \quad (3.1)$$

The sharp bandpass filters we chose are projection operators, denoted by  $\mathcal{P}_l$  and  $\mathcal{P}_h$ , that separate the solution into low- and high-frequency components at  $\Lambda\omega$ , with  $\Lambda < N$ ,

$$\mathbf{w}_l(t) = \sum_{|l| \leq \Lambda} e^{il\omega t} \hat{\mathbf{w}}_l, \quad (3.2a)$$

$$\mathbf{w}_h(t) = \sum_{|h| > \Lambda} e^{ih\omega t} \hat{\mathbf{w}}_h, \quad (3.2b)$$

such that  $\mathcal{P}_l[\mathbf{w}] \equiv \mathbf{w}_l$  and  $\mathcal{P}_h[\mathbf{w}] \equiv \mathbf{w} - \mathbf{w}_l = \mathbf{w}_h$ . The HBNS equations (Eq. (2.4)) are projected onto the low- and high-frequency subspaces

$$in\omega \hat{\mathbf{w}}_l = \mathcal{L}[\hat{\mathbf{w}}_l] + \mathcal{P}_l(\mathcal{N}[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_l] + \mathcal{N}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h] + \mathcal{N}[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_h] + \mathcal{N}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_l]), \quad (3.3a)$$

$$im\omega \hat{\mathbf{w}}_h = \mathcal{L}[\hat{\mathbf{w}}_h] + \mathcal{P}_h(\mathcal{N}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_h] + \mathcal{N}[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_h] + \mathcal{N}[\hat{\mathbf{w}}_h, \hat{\mathbf{w}}_l] + \mathcal{N}[\hat{\mathbf{w}}_l, \hat{\mathbf{w}}_l]), \quad (3.3b)$$

where  $\mathcal{L}$  and  $\mathcal{N}$  correspond to the linear and nonlinear operators, respectively. No approximation has been introduced up to this point, and the above set of equations corresponds to a separation of the dynamics into low- and high-frequency motions.

#### 3.1. Data-driven modeling of the super-spectral residual: a linear example

In the present study, instead of making approximations in the operator governing the high-frequency dynamics (i.e. GQL; see Appendix A), we apply a data-driven approach to learn the influence of the high-frequency dynamics on the low-frequency dynamics without any approximation. The proposed approach can be regarded as the equivalent for spatially coarse-graining dynamical systems using the Mori–Zwanzig formalism (Parish & Duraisamy 2017), but applied in time (i.e., time coarse-graining or, equivalently, frequency truncation).

We first consider a demonstrative example to introduce the reader to the proposed formalism. Consider a two-state linear system, written in the frequency domain, given in matrix form by

$$\begin{bmatrix} A_{ll} & A_{lh} \\ A_{hl} & A_{hh} \end{bmatrix} \begin{bmatrix} u_l \\ u_h \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Our goal is to create a reduced-order model that depends only on  $u_l$ ; that is

$$A_{ll}u_l + R(u_l) = 0. \quad (3.4)$$

For an accurate prediction of  $u_l$ , the function  $R(u_l)$  has to be determined. This function represents the effect of the unresolved variable  $u_h$  on the resolved variable  $u_l$ . In physical space, this is analogous to modeling subgrid-scale effects in large-eddy simulations (Parish & Duraisamy 2017).

For the simple linear system introduced above,  $R(u_l)$  can be exactly determined by solving for  $u_h$  given  $u_l$ . Through this process, the two-component system can be cast as

a one-component system that has the closed form

$$A_l u_l - \underbrace{A_{lh} A_{hh}^{-1} A_{lh}}_{\substack{u_h \\ \text{residual } R(u_l)}} u_l = 0. \quad (3.5)$$

Note, however, that one has to solve for the unresolved variable  $u_h$ . The one-dimensional representation of the original system offers no decrease in computational complexity. Instead, it provides a starting point for the construction of closure models, since the effect of the unresolved variable  $u_h$  has been expressed purely in terms of the resolved variable  $u_l$ .

### 3.2. Data-driven modeling of the super-spectral residual: HBNS equations

The above approach can be generalized to nonlinear functions (or operators). In the case where the dynamics are governed by the HBNS equations, the low-frequency dynamics are dictated by

$$\underbrace{-in\omega \mathbf{q}_l + \mathcal{L}[\mathbf{q}_l] + \mathcal{P}_l(\mathcal{N}[\mathbf{q}_l, \mathbf{q}_l])}_{HBNS_l(\mathbf{w}_1)} + \underbrace{\mathcal{P}_l(\mathcal{N}[\mathbf{q}_h, \mathbf{q}_h] + \mathcal{N}[\mathbf{q}_l, \mathbf{q}_h] + \mathcal{N}[\mathbf{q}_h, \mathbf{q}_l])}_{F(\mathbf{w}_1, \mathbf{w}_h)} = 0. \quad (3.6)$$

After eliminating the nonlinear high-frequency dynamics by replacing  $\mathbf{w}_h$  from the solution of Eq. (3.3b), the residual term  $F$  can be expressed only as a (nonlinear) function of  $\mathbf{w}_l$

$$HBNS_l(\mathbf{w}_1) + R_\theta(\mathbf{w}_1) = 0. \quad (3.7)$$

The term  $R_\theta(\mathbf{w}_1)$  corresponds to the exact residual of the frequency-truncated (temporally coarse-grained) HBNS equations.

In this study, we parameterize the residual closure term using a deep neural network (DNN). The unknown parameters of the network,  $\theta$ , are learned using data from high-fidelity simulations (i.e., a high number of harmonics  $N$  to ensure convergence to DNS precision). The high-fidelity data are truncated to  $\Lambda$  harmonics to obtain  $(\mathbf{w}_1)$ , and then the residual is obtained by evaluating the coarse-grained operator using the high-fidelity truncated solution,  $R = -HBNS_l(\mathbf{w}_1)$ .

## 4. Results

We demonstrate the proposed framework on the flow around a 2D circular cylinder. The HBNS equations are discretized in the streamwise  $x$  and vertical  $y$  directions in FreeFem++. The computational domain is a rectangle ( $[-10, 30]$  in  $x$ ,  $[-10, 10]$  in  $y$ ), and the cylinder center is placed at  $(x, y) = (0, 0)$ . Equations and results are nondimensionalized by the cylinder diameter  $D$  and the freestream velocity  $U_\infty$ , and the Reynolds number is defined as  $Re = U_\infty D / \nu$ .

### 4.1. 2D limit-cycle solutions with HBNS equations

Above  $Re = 47$ , the flow is linearly unstable. A stable limit-cycle solution with periodic vortex shedding is obtained by solving the HBNS equations using a finite number of frequency harmonics  $N$ . The HBNS equations are initialized with the linearly unstable eigenmode and the base flow.

Using one frequency harmonic, we recover the results obtained by Mantić-Lugo &

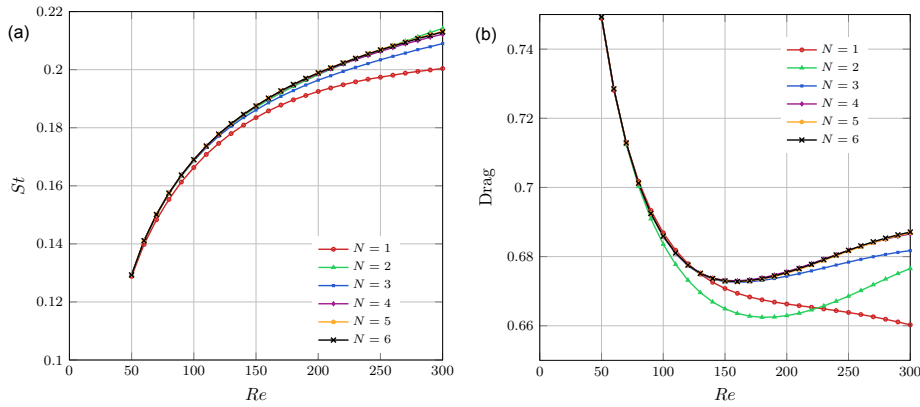


Figure 1: Strouhal number (a) and drag force (b) as a function of the Reynolds number. Truncation in the number of frequency harmonics  $N$ .

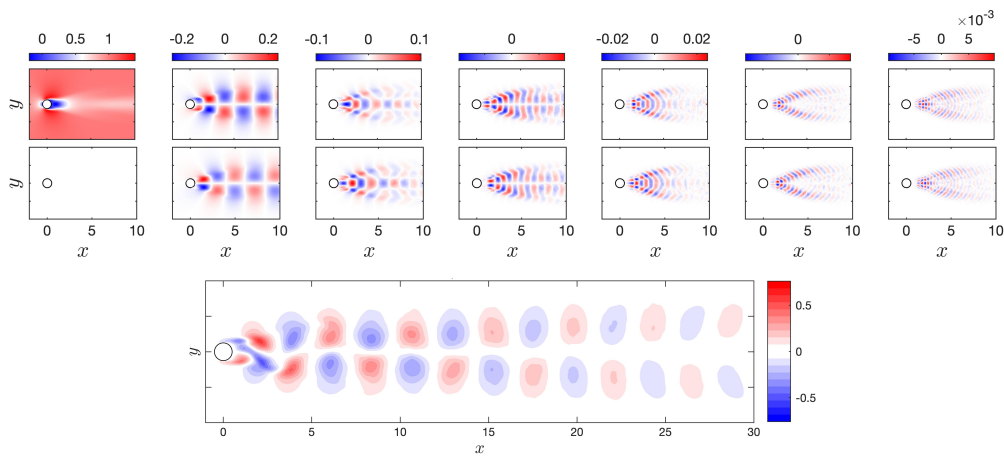


Figure 2: Streamwise velocity components at  $Re = 200$ . Real and imaginary harmonic components are shown for  $N = 6$  (top). Total perturbation field in physical space at  $t = 0$  (bottom).

Gallaire (2016) and Fabre *et al.* (2018). This corresponds to the predictions of the quasi-linear model, where the mean and perturbation fields are obtained in a self-consistent way. In the above case where  $N = 1$ , the perturbation equation is linear. Close to and above the critical Reynolds number of the Hopf bifurcation, the higher harmonics are weak, and this approximation yields convergent and hence accurate results. The convergence study is shown in Figure 1, which shows the Strouhal number,  $St = fD/U_\infty$ , of the fundamental limit-cycle frequency and the drag force. However, departing from the threshold of the bifurcation, energy is transferred to the higher harmonics, requiring more harmonics to obtain converged results. Convergence of the HBNS solution (to a fully resolved DNS) is achieved with  $N = 4$  frequency harmonics over the Reynolds number range  $[50, 300]$ .

Figure 2 shows the (real) mean flow and the first six (complex) harmonic components for the streamwise velocity component  $u$ . Higher harmonics are characterized by higher frequency, shorter wavelengths, and lower amplitudes. Figure 2 also shows the total  $u'$

perturbation field in the time domain for  $t = 0$ , reconstructed from the frequency domain solution according to

$$u'(t) = \sum_{|n|=1}^4 e^{in\omega t} \hat{\mathbf{u}}_n. \quad (4.1)$$

#### 4.2. HBNS truncation and residual

For the subsequent analysis we will consider the  $N = 4$  solution as the ground truth (converged DNS solution), and we will identify data-driven closures of the HBNS equations to achieve the same level of accuracy, but using a lower (truncated) number of harmonics.

Here, we will model the residual term for the severely truncated case, where we keep only only harmonic (quasi-Linear model with  $N = 1$ )

$$\begin{aligned} \left[ \mathbf{L} + \frac{1}{2} \mathbf{N}(\hat{\mathbf{w}}_0, \cdot) \right] \hat{\mathbf{w}}_0 + \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_1) &= -\mathbf{R}_0, \\ [i\omega \mathbf{M} + \mathbf{L} + \mathbf{N}(\hat{\mathbf{w}}_0, \cdot)] \hat{\mathbf{w}}_1 &= -\mathbf{R}_1, \end{aligned}$$

where

$$\begin{aligned} \mathbf{R}_0 &= \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_2) + \mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_3) + \mathbf{N}(\overline{\hat{\mathbf{w}}_4}, \hat{\mathbf{w}}_4) + \dots, \\ \mathbf{R}_1 &= \mathbf{N}(\overline{\hat{\mathbf{w}}_1}, \hat{\mathbf{w}}_2) + \mathbf{N}(\overline{\hat{\mathbf{w}}_2}, \hat{\mathbf{w}}_3) + \mathbf{N}(\overline{\hat{\mathbf{w}}_3}, \hat{\mathbf{w}}_4) + \dots, \end{aligned}$$

The residual terms  $\mathbf{R} = [\mathbf{R}_0, \mathbf{R}_1]$  will be modeled here using only known low-frequency harmonics,  $\hat{\mathbf{w}}_1 = [\hat{\mathbf{w}}_0, \hat{\mathbf{w}}_1]$ . It is evident that these residual terms are functions of the higher unknown harmonics (see Figure 3(a)). However, our proposed framework expresses them only as a function of the low harmonics by decoding the low-frequency space to the high-frequency space and then forming the residual term. Equivalently, the decoder is learning the solution of the neglected high-frequency HBNS equations, which is taken as a function of the low-frequency solution. Instead of a two-step process, we can learn the input-output mapping with a single nonlinear mapping  $R_\theta(\mathbf{w}_1)$ .

#### 4.3. Data-driven modeling of super-spectral scales

The residual nonlinear mapping  $R_\theta(\mathbf{w}_1)$  is parameterized as a DNN. We use a fully connected network (also known as a multilayer perceptron) composed of a decoder (expanding neural network) and an encoder (contracting neural network). The input to the neural network is the resolved low-frequency velocity harmonics, and the output is the low-frequency residual. Since we are considering the coarse-grained HBNS equations with  $N = 1$ , the input has six real pointwise components,  $\mathbf{w}_1 = [\mathbf{w}_0, \Re(\mathbf{w}_1), \Im(\mathbf{w}_1)]$ , and the residual output has the same size, sampled at a specific  $(x, y)$  location. The input to the residual neural network for the closure terms contains only localized information at a single point, and of future interest is the exploration of convolutional neural networks or stencil inputs (Maulik *et al.* 2019).

The training data are generated from running fully resolved HBNS calculations with  $N = 4$  at  $Re = [100, 150, 200]$ . The decoder-encoder architecture we use in this study has five hidden layers, with  $[20, 40, 60, 40, 20]$  nodes for each layer and a ReLU activation function.

Figure 3(b) shows the predictions of the residual at  $Re = 300$ , using as an input the low-frequency (coarse-grained components). This is a case that has not been used for training, and it will test the capabilities of the proposed approach for extrapolation.

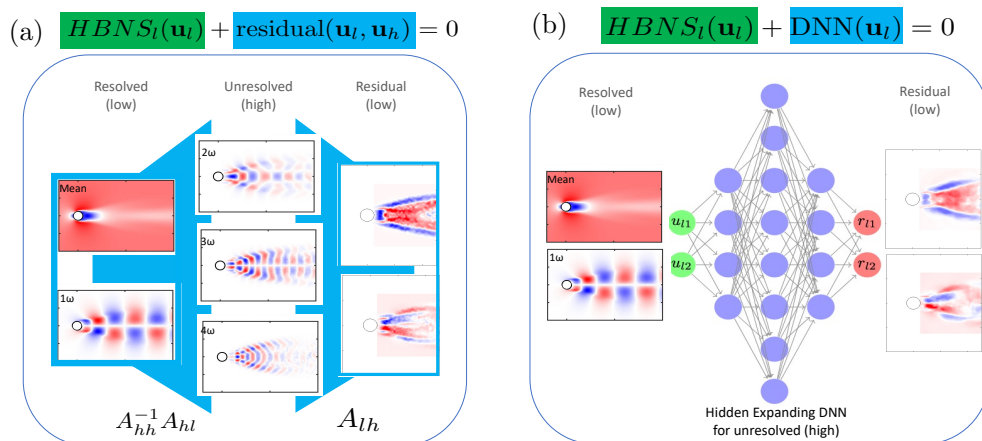


Figure 3: Data-driven modeling and identification of the super-spectral (high/unresolved) scales, from sub-spectral (low/resolved) scales. (a) Operator structure of the residual terms. (b) Deep neural network structure for data-driven identification.

Compared with the exact residual, shown in Figure 3(a), which was calculated solving the HBNS equations with  $N = 4$ , excellent agreement is observed. The results show the capability of using only a low number of frequency harmonics (here  $N = 1$ ), and achieving prediction errors similar to those obtained from fully resolved calculations ( $N = 4$ ).

## 5. Conclusions

In this study, we demonstrated a data-driven framework to model the residual (non-local closure) terms for the frequency-truncated HBNS equations. By splitting the sought solution into low-frequency (resolved) and high-frequency (unresolved) harmonics, we systematically express the low-frequency residual as a function of the resolved frequency harmonics only. A consistent deep learning architecture, which parameterizes the residual function, was designed and trained using high-fidelity results above the threshold of transition for 2D cylinder flow. We showed that our proposed framework achieves low generalization errors by predicting accurately the coarse-grained residual for unseen Reynolds numbers, and significantly reduces the computational cost by solving accurately for the coarse-grained dynamics.

Future work will evaluate the proposed framework in the fully turbulent regime and will compare it to linearized Navier-Stokes models and common eddy-viscosity closures (Pickering *et al.* 2021).

### Acknowledgments

G. R. thanks the participants of the CTR summer program for inspiring discussions. Special thanks are extended to Luca Magri and Anh Khoa Doan for stimulating discussions and their help regarding the machine learning implementation, and to Petros Ioannou and Brian Farrell for their constructive comments.

### Appendix A. Generalised quasi-linear approximation

The generalised quasi-linear (GQL) approximation (Marston *et al.* 2016) corresponds to linearization of the high-frequency dynamics around the low-frequency dynamics. The GQL system, applied to the HBNS equations, reads

$$in\omega\mathbf{u}_l = \mathcal{L}[\mathbf{u}_l] + \mathcal{P}_l(\mathcal{N}[\mathbf{w}_l, \mathbf{w}_l] + \mathcal{N}[\mathbf{w}_h, \mathbf{w}_h]), \quad (\text{A } 1a)$$

$$im\omega\mathbf{w}_h = \mathcal{L}[\mathbf{w}_h] + \mathcal{P}_h(\mathcal{N}[\mathbf{w}_h, \mathbf{w}_l] + \mathcal{N}[\mathbf{w}_l, \mathbf{w}_h]). \quad (\text{A } 1b)$$

The terms  $\mathcal{P}_h(\mathcal{N}[\mathbf{w}_h, \mathbf{w}_h])$  and  $\mathcal{P}_h(\mathcal{N}[\mathbf{w}_l, \mathbf{w}_l])$  have been omitted due to the linearization. The terms  $\mathcal{P}_l(\mathcal{N}[\mathbf{w}_l, \mathbf{w}_h])$  and  $\mathcal{P}_l(\mathcal{N}[\mathbf{w}_h, \mathbf{w}_l])$  are set to zero such that the GQL system conserves energy.

In this study, we applied the GQL approximation in the frequency domain, dictating the asymptotic dynamics of the flow. Although the application of the GQL in the time domain has yielded significant improvements in the prediction of the dynamics in other applications (in the turbulent regime, for example, Hernández *et al.* (2022)), in the present application we found that the high-frequency subspace admits only the trivial (null) solution due to the stable nature of the operator dictating the high-frequency dynamics.

### REFERENCES

- BAR-SINAI, Y., HOYER, S., HICKEY, J. & BRENNER, M. P. 2019 Learning data-driven discretizations for partial differential equations. *P. Natl. Acad. Sci. USA* **116** (31), 15344–15349.
- BENGANA, Y. & TUCKERMAN, L. S. 2021 Frequency prediction from exact or self-consistent mean flows. *Phys. Rev. Fluids* **6**, 063901.
- FABRE, D., CITRO, V., FERREIRA-SABINO, D., BONNEFIS, P., SIERRA, J., GIANNETTI, F. & PIGOU, M. 2018 A practical review on linear and nonlinear global approaches to flow instabilities. *Appl. Mech. Rev.* **70**, 060802.
- HECHT, F. 2012 New development in FreeFem++. *J. Numer. Math.* **20**, 251–265.
- HERNÁNDEZ, C. G., YANG, Q. & HWANG, Y. 2022 Generalised quasilinear approximations of turbulent channel flow. Part 1. Streamwise nonlinear energy transfer. *J. Fluid Mech.* **936**, A33.
- KELSHAW, D., RIGAS, G. & MAGRI, L. 2022 Physics-informed CNNs for super-resolution of sparse observations on dynamical systems. *arXiv:2210.17319 [physics.flu-dyn]*.
- KERSWELL, R. R. 2018 Nonlinear nonmodal stability theory. *Ann. Rev. Fluid Mech.* **50**, 319–345.
- MANTIČ-LUGO, V. & GALLAIRE, F. 2016 Self-consistent model for the saturation mechanism of the response to harmonic forcing in the backward-facing step flow. *J. Fluid Mech.* **793**, 777–797.
- MARSTON, J. B., CHINI, G. P. & TOBIAS, S. M. 2016 Generalized quasilinear approximation: application to zonal jets. *Phys. Rev. Lett.* **116**, 214501.
- MARSTON, J. B. & TOBIAS, S. M. 2023 Recent developments in theories of inhomogeneous and anisotropic turbulence. *Ann. Rev. Fluid Mech.* **55** (1).
- MAULIK, R., SAN, O., RASHEED, A. & VEDULA, P. 2019 Subgrid modelling for two-dimensional turbulence using neural networks. *J. Fluid Mech.* **858**, 122–144.
- PARISH, E. J. & DURAISAMY, K. 2017 A dynamic subgrid scale model for large eddy simulations based on the Mori–Zwanzig formalism. *J. Comput. Phys.* **349**, 154–175.

- 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.
- RIGAS, G., SIPP, D. & COLONIUS, T. 2021 Nonlinear input/output analysis: application to boundary layer transition. *J. Fluid Mech.* **911**, A15.
- SIERRA-AUSIN, J., CITRO, V., GIANNETTI, F. & FABRE, D. 2022 Efficient computation of time-periodic compressible flows with spectral techniques. *Comput. Methods Appl. Mech. Eng.* **393**, 114736.