Study of the energy convergence of the Karhunen–Loève decomposition applied to the large-eddy simulation of a high-Reynolds-number pressure-driven boundary layer

Pieter Bauweraerts and Johan Meyers
KU Leuven, Mechanical Engineering, Celestynelaan 300A, B3001 Leuven, Belgium
(Dated: March 23, 2020)

We study the energy convergence of the Karhunen–Loève decomposition of the turbulent velocity field in a high-Reynolds-number pressure-driven boundary layer as function of the number of modes. An energy-optimal Karhunen–Loève (KL) decomposition is obtained from wall-modelled large-eddy simulations at ‘infinite’ Reynolds number. By explicitly using Fourier modes for the horizontal homogeneous directions, we are able to construct a basis of full rank, and we demonstrate that our results have reached statistical convergence. The KL-dimension, corresponding to the number of modes per unit volume required to capture 90% of the total turbulent kinetic energy, is found to be $2.4 \times 10^3 \, [H^{-3}]$ (with $H$ the boundary layer height). This is significantly higher than current estimates, which are mostly based on the method of snapshots. In our analysis, we carefully correct for the effect of subgrid scales on these estimates. To this end, we identify two classical scaling regimes, corresponding to shear production and inertial range respectively. Finally, the degrees of freedom (DOF) per unit volume necessary to represent 90% of the energy in our pseudo-spectral LES code is also identified, and roughly corresponds $2.6 \times 10^6 \, [H^{-3}]$.

I. INTRODUCTION

Low dimensional models describing the dynamics of the atmospheric boundary layers have important applications, e.g., ranging from dispersion of pollutants, to predicting the power output of a wind turbine, to controlling the turbulence in wind farms for enhanced power production. An assessment of the required dimensionality of a reduced order system can be made by studying the system dynamics typically done via the Kaplan–Yorke (KY) definition of the dimension of the attractor, which is based on the Lyapunov exponents of the dynamical system. This becomes increasingly difficult for large scale systems. Alternatively the Karhunen–Loève (KL) decomposition (also known as proper orthogonal decomposition (POD), empirical orthogonal functions (EOF) or principal component analysis (PCA)) can be used, which generates modes which are well known to be energy optimal, in the sense that off all possible mode sets they capture on average the most energy [1]. The KL-modes are often used to identify structures in turbulent flows, and used as a basis for a reduced order system. An assessment of the number of modes required is the so called KL dimension, introduced into the fluid dynamics community by Sirovich [2]. This is defined as: “the number of actual eigenfunctions required so that the captured energy is at least 90% of the total (as measured by the energy norm) and that no neglected mode, on average, contains more than 1% of the energy contained in the principle eigenfunction mode.” This dimension has previously been found to be of the same order of magnitude as the KY-dimension. The dimensionality of turbulent flows is typically significantly lower than the amount of degrees of freedom (DOF) $\sim (\eta/L)^3 \sim Re^{9/4}$, with $\eta$ the Kolmogorov scale, $L$ a characteristic length and $Re$ the Reynolds number, due to the prevalence of flow field correlations in the large energy containing scales. Moreover, at asymptotically high Reynolds numbers, such as encountered in the ABL, we expect a finite KL-dimension that becomes independent of the Reynolds number.

The KL-modes and corresponding eigenvalues are found as the eigenfunctions, and eigenvalues of the two-point covariance tensor [3]. However, the two-point covariance tensor has $(3N \times 3N)$ elements, where $N$ is the amount of grid points. For a typical high-Reynolds-number turbulent boundary-layer simulation, this is in the order of $10^7 - 10^{10}$ points, which is too big to handle with current computational resources. A popular strategy for transforming this into a tractable problem is the so-called method of snapshots [4], where the duality of the space and time correlation is exploited to transform the problem into a $N_s \times N_s$ problem, with $N_s$ the amount of samples, typically of the order $10^3 - 10^4$. This strategy is applicable to general flow geometries, but the rank is thereby capped by the number of samples, and the eigenvalue spectrum has been shown to converge very slow for large scale problems [5]. An alternative approach is available if the problem exhibits homogeneous directions, which is often the case for the canonical flow cases studied in turbulence. In this case, it is easily shown that the POD-modes correspond to Fourier modes in these

---

*Electronic address: pieter.bauweraerts@kuleuven.be
†Electronic address: johan.meyers@kuleuven.be
directions, and the large scale eigenvalue problem can be replaced by a smaller scale eigenvalue problem per wave number (see e.g. [3]).

The KL-dimension has already been determined for canonical flow cases at low Reynolds/Rayleigh numbers. An overview of studies is given in Table I. The KL-dimension for turbulent channel flow has been determined in Ref. [6], while in Ref. [7] a minimum flow unit was studied and compared to a larger test-case, identifying a linear increase in KL-dimension in correspondence with the extensive property of dimension. Refs. [8, 9] on the other hand found a strong increase in dimensionality with the Reynolds number, while Ref. [10] considered the influence of visco-elasticity. Other flow cases considered are Couette flow [11], Rayleigh–Bénard convection [12, 13], turbulent pipe flow [14], and a turbulent boundary layer [15].

POD studies considering the high-Reynolds atmospheric boundary layer, simulated using large-eddy simulations (LES), are numerous. A non-exhaustive overview is given hereafter. 3D POD of the ABL using snapshot POD approach are performed in Refs. [16–18]. However, the high dimensionality and slow convergence of the snapshot POD, leads to an underestimation of the KL dimension. A comparison of POD in 2D planes, significantly reducing the dimensions, but losing part of the 3D structures, is performed in Refs. [19–24]. In Huang et al. [25] both 1D and 3D POD computations were performed, where homogeneity assumptions are used for the latter. For the 1D case, the convergence of LES and DNS is compared. A slower convergence of the eigenvalues was found for the LES, which was attributed to the Reynolds number (as was demonstrated in Ref. [9]). Other studies only considered a few dominant modes [26, 27]. Note that studies [16–18] considered the impact of wind turbines in an atmospheric boundary layer, such that the Fourier approach is no longer applicable, and only discrete translational symmetry can be used in periodic directions to extend the snapshot base. Next to LES studies, also experimental wind tunnel studies considering 2D POD modes using the spectral approach [28], and using the method of snapshots [29] are found.

In the current study we consider a rough pressure-driven turbulent boundary as a substitute for a neutral atmospheric boundary layer. This approach has been often used for LES studies of the neutral ABL [30–36], and is known to represent statistics in the logarithmic layer very well. Simulations are based on wall-modelled LES, using a wall-stress model, and direct effects of viscosity are neglected, so that all dissipation is handled by the subgrid-scale model, approaching effectively the limit of an ‘infinite’ Reynolds number. Analytical expressions for the modes in the horizontal homogeneous directions are used, such that a complete KL basis and the corresponding eigenvalues can be determined. We focus on the eigenvalue spectrum, and more precisely on the convergence. The structure of the associated dominant modes are already extensively reported in the ABL studies summarized above, and are not repeated in the current study. We further identify the KL dimension, i.e. the number of modes required to represent 90% of the energy, and also identify the number of LES degrees of freedom to capture the same amount of energy. We note that in the context of LES, capturing at least 80-90% of the energy is often considered as a heuristic for quality of the simulation [37, 38].

The paper continues by given a brief overview of the POD methodology and the Fourier approach for the horizontal directions. Next, we describe our case setup, LES model, discretization, and the sampling specifications. Subsequently, the results are presented, considering both the statistical convergence of the results and aspects of dimensionality. Finally, the most important conclusions are discussed.

II. CASE STUDY AND SIMULATION METHODOLOGY

In this case study we consider a pressure driven boundary layer, a summary of the simulation parameters is given in Table I and a snapshot of the flow field is given in Fig. 1. We consider a relatively large domain of $42H \times 12H \times H$ to avoid spurious influence of the periodic boundary conditions on the two-point velocity covariance tensor, which is known to extend up to $10H$ for the streamwise velocity component [39, 40]. The surface roughness length $z_0$ is chosen such that for a BL height of $H = 1000$ m, we get a value of $2 \times 10^{-4}$ m, which is, e.g., typical for offshore conditions.

The code used for this study has been extensively documented in past studies, see e.g. Ref. [41, 42] for further details. In the horizontal directions we use periodic boundary conditions, in the vertical directions we use impermeability in combination with a wall-stress model [43] at the bottom wall and zero stress at the top. As a subgrid-scale model, we use a classical Smagorinsky model [44], combined with wall-damping close to the wall [45]. ABLs occur at very high Re-number, such that the effect of the kinematic viscosity on the resolved flow can be neglected. In this way our flow becomes Re-independent and can be interpreted as the asymptotic behaviour at infinitely/very high Re-number [46]. The horizontal directions use a Fourier spectral discretization, dealiased using the 2/3 rule (see e.g. [47]). For the vertical direction we employ a fourth-order energy conservative scheme [48]. For the time integration we use an explicit fourth order Runge-Kutta method combined with a 0.4 CFL number limit on the time step. In order to speed up the simulations, the equations are solved in a frame of reference moving at approximately half the maximum flow speed $\sim 9.5H/u_*$ in the streamwise direction, allowing for a doubling of the stability time step.
### TABLE I: Comparison of different 3D KL studies.

| Case        | $Re_r$ | $D_{kl}$ | $L_1/H$ | $L_2/H$ | $L_3/H$ | $d_{kl} [H^{-3}]$ | Method | Reference |
|-------------|--------|----------|---------|---------|---------|-------------------|--------|-----------|
| DNS-CH      | 80     | 380      | 1.6$\pi$| 1.6$\pi$| 2       | 15.40             | F − SC | Ball et al. [5] |
| DNS-CH      | 110    | 13452    | 5$\pi$  | 2$\pi$  | 2       | 68.16             | S − SC | Iwamoto et al. [9] |
| DNS-CH      | 125    | 4186     | 5$\pi$  | 5$\pi$  | 2       | 83.72             | F − SC | Sirovich et al. [8] |
| DNS-CH      | 136    | 658      | $\pi$   | 0.3$\pi$| 2       | 111.12            | F − SC | Webber et al. [7] |
| DNS-CH      | 180    | 18920    | 9$\pi$  | 4.5$\pi$| 2       | 233.6             | F − SC | Housiadas et al. [10] |
| DNS-CH      | 300    | 36520    | 2.5$\pi$| $\pi$   | 2       | 74.04             | F − SC | Iwamoto et al. [9] |

| LES-ABL     | $\infty$ | −        | 30      | 30      | 1       | −                | F − SC | Keith Wilson [27] |
| LES-ABL     | $\infty$ | −        | 8       | 8       | 1       | −                | F − SC | Esau [26] |
| LES-ABL     | $\infty$ | $10^3$   | 2$\pi$  | $\pi$  | 1       | 50.66            | TC     | Ali et al. [17] |
| LES-PDBL    | $\infty$ | −        | $2\pi$  | $2\pi$  | 1       | −                | F − SC | Huang et al. [25] |
| LES-PDBL    | $\infty$ | $4 \times 10^3$ | $\pi$  | $\pi$  | 1       | 405              | TC     | VerHulst and Meneveau [16] |
| LES-PDBL    | $\infty$ | $3 \times 10^3$ | $4\pi$  | $2\pi$  | 1       | 38               | TC     | Zhang and Stevens [18] |
| LES-PDBL    | $\infty$ | $9.77 \times 10^6$ | 42     | 12      | 1       | 240000           | F − SC | Current manuscript |

*The reported values did not take into account the degeneracy of the modes to calculate the dimensionality, therefore the reported dimension was multiplied by 4 as an approximation/upper boundary.*

### TABLE II: Summary of the simulation grid setup and simulation parameters.

| Domain size | $L_1 \times L_2 \times L_3$ | $42H \times 12H \times H$ |
|-------------|-----------------------------|---------------------------|
| Grid size   | $N_1 \times N_2 \times N_3$ | $2800 \times 800 \times 200$ |
| Cell size   | $\Delta x_1 \times \Delta x_2 \times \Delta x_3$ | $0.015H \times 0.015H \times 0.005H$ |

Roughness length $z_0/H$ $2 \times 10^{-7}$

### III. PROPER ORTHOGONAL DECOMPOSITION

A brief overview is given of the POD method. For a more elaborate discussion we refer to Ref. [3]. The POD-mode $\phi$ and its corresponding eigenvalue $\lambda$ is given by respectively the eigenvectors and eigenvalues of the following

![FIG. 1: The top and bottom figure respectively represent a $x$-$y$ and $x$-$z$ section of an instantaneous streamwise velocity field.](image-url)
Fredholm eigenvalue problem

\[ \frac{1}{|\Omega|} \int_{\Omega} R_{ij}(\mathbf{x}, \mathbf{\bar{x}}) \phi_i(\mathbf{\bar{x}}) \, d\mathbf{\bar{x}} = \lambda \phi_j(\mathbf{x}), \]

with \( R_{ij}(\mathbf{x}, \mathbf{\bar{x}}) = \langle \Bar{u}_j^i(\mathbf{x}) \Bar{u}_j^j(\mathbf{\bar{x}}) \rangle \) the two-point covariance tensor, where \( \Bar{u}_j^i = \Bar{u} - \langle \Bar{u} \rangle \) is the filtered velocity fluctuation with respect to the mean in direction \( \mathbf{e}_i \). The eigenvalues are numbered and ordered in decreasing order \( \lambda_k \geq \lambda_{k+1} \). The eigenvectors are orthogonal and can be normalized such that \( \langle \phi_i(\mathbf{x}), \phi_j(\mathbf{\bar{x}}) \rangle = \delta_{ij} \), where the inner product \( \langle \cdot, \cdot \rangle_{\Omega} \) is used, defined as \( \langle \phi, \psi \rangle_{\Omega} = |\Omega|^{-1} \int_{\Omega} \phi^*(\mathbf{x}) \psi(\mathbf{x}) \, d\mathbf{x} \). The average turbulent kinetic energy per unit volume captured by the first \( n \) modes, \( \lambda_n \), is directly related to the sum of the eigenvalues \( \lambda_n = \sum_{q=1}^{n} \lambda_q/2 \). It is easily shown that for the horizontal homogeneous directions the eigenvectors are Fourier modes, such that we have eigenmodes of the form \( \sqrt{2} \cos \left( k_1 x_1 + k_2 x_2 + \angle \phi_i^j(k)(x_3) \right) \phi_i^j(k)(x_3) \) and \( \sqrt{2} \sin \left( k_1 x_1 + k_2 x_2 + \angle \phi_i^j(k)(x_3) \right) \phi_i^j(k)(x_3) \), where the vertical modes \( \phi_i^j(k)(x_3) \) are found by solving the following Fredholm eigenvalue problem

\[ \frac{1}{H} \int_0^H \tilde{R}_{ij}(\mathbf{k}, x_3, \mathbf{\bar{x}}_3) \phi_i^j(\mathbf{\bar{x}}_3) \, d\mathbf{\bar{x}}_3 = \lambda_k \phi_i^k(\mathbf{\bar{x}}_3), \]

with \( \tilde{R}_{ij}(\mathbf{k}, x_3, \mathbf{\bar{x}}_3) = \langle \Bar{u}_j^i(\mathbf{k}, x_3) \Bar{u}_j^j(\mathbf{\bar{x}}_3) \rangle \) the horizontal-spectrum vertical-covariance tensor. Here, the horizontal wave vector \( \mathbf{k} = [k_1, k_2] \) has been introduced for sake of conciseness. The numbering of the POD modes per wave number (also called quantum number) is also done in decreasing eigenvalue order \( \lambda_{(k,l)} \geq \lambda_{(k,l+1)} \). The corresponding vertical modes are in analogy to the 3D modes normalized such that \( \langle \phi_i^{(m)}(\mathbf{k}), \phi_j^{(n)}(\mathbf{k'}) \rangle_{x_3} = \delta_{mn} \delta_{ij} \). Due to the periodic horizontal boundary conditions, only integer multiples of the base wave numbers \( k_1^* = 2\pi/L_1 \) and \( k_2^* = 2\pi/L_2 \) for respectively the streamwise and spanwise direction, have a non-zero contribution. The highest allowable wavenumbers are determined by the grid cut-off wave numbers, i.e. \( \pi/\Delta_1 \) and \( \pi/\Delta_2 \). Moreover, we discretize the integral in Eq. 1 using the midpoint rule at the vertical grid locations in of our LES grid. In this way Eq. [1] is reduced to a finite-sized eigenvalue problem. By bringing the eigenvalues of the different wave numbers together and reordering we obtain the set of ordered eigenvalues of the original 3D problem.

The flow field is sampled every \( 3.16 \times 10^{-3}H/u_* \) and the symmetry of the equations in the spanwise directions is used to double sample size. Note that the samples are for most wave numbers well within their integral time scale, and therefore correlated such that the effective sample size will be smaller and dependent on the considered wave numbers. A total of 8200 samples is generated, leading to an averaging time of \( 12.97H/u_* \) time units.

IV. RESULTS

A. Sampling time: convergence of the results

First we study the convergence of the eigenvalues as a function of the amount of snapshots which were used to compute the two-point covariance tensor. Figure 2 shows the eigenvalues of the different POD modes, which are...
FIG. 3: (a) Eigenvalues $\lambda_n/u_s^2$ (b) premultiplied eigenvalues $n\lambda_n/u_s^2$ as a function of the index number $n$. The black trend lines indicate the $n^{-1}$ and $n^{-11/9}$ scaling. The figure is suggestively subdivided in an inactive range (I), a shear production range (II) and an inertial range (III).

### TABLE III: Summary of the properties of the most energetic modes $\hat{\phi}_i^{(k,m)}$. The wave numbers are normalized by $k_i^* = 2\pi/L_i$. The degeneracy denotes the multiplicity of the eigenvalues.

| $n$ | $k_1/k_i^*$ | $k_2/k_i^*$ | $m$ | $\lambda_n/u_s^2$ | Degeneracy |
|-----|-------------|-------------|-----|-------------------|------------|
| 1-4 | ±1          | ±6          | 0   | 0.02745          | 4          |
| 5-6 | 0           | ±6          | 0   | 0.02132          | 2          |
| 7-10| ±1          | ±5          | 0   | 0.01190          | 4          |
| 11-12| 0          | ±4          | 0   | 0.01186         | 2          |
| 13-16| ±1          | ±2          | 0   | 0.01074         | 4          |

ordered such that $\lambda_k \geq \lambda_{k+1}$ regardless of the originating wave number, and the average fraction of unresolved energy $\mathcal{E}_n^\Delta = 1 - K_n/K_\Delta$, where $K_\Delta$ is the resolved kinetic energy $K_\Delta = |\Omega|^{-1} \int_\Omega \overline{\langle u'_i u'_j \rangle} / 2 \, dx$. Note that our POD modes span the whole space of filtered solenoidal fields, such that $\mathcal{E}_n^\Delta = 0$, with $N_m = N_1 N_2 (2N_3 - 1) + 1$. The difference between $N_m$ and the number of degrees of freedom $N_1 N_2 (3N_3 - 1) - 1$ due to staggered vertical velocity in the vertical direction) results from the amount of independent continuity constraints $N_1 N_2 N_3 - 1$. The possible effect of the unresolved scales in the LES is further discussed in §IVC.

In Figure 2, it is observed that starting from 2048 samples the shape of the different curves becomes almost independent of the amount of samples. The eigenvalue curves show that the amount of non-zero eigenvalues grows linearly with the amount of samples, until an abrupt change occurs at around 2/3 of the total amount of eigenvalues. The linear behaviour is caused by the rank of the spectral correlation matrix being limited by the amount of samples $N_s$. The abrupt change to zero is caused by the rank deficiency of the two-point correlation tensor, due to samples being restricted to divergence free flow fields, because $\partial R_{ij}(x,x')/\partial x_i = 0$. The scaling of the eigenvalues and the residual energy behaviour is further discussed in more detail below.

The elements of the tensor $\hat{R}_{ij}^T = \langle \tilde{u}'_i(k,x_3) \tilde{u}'_j(k,x_3) \rangle_T$, have a variance, which for sufficiently long time averaging horizon $T$, is of the form $\text{Var}(\hat{R}_{ij}^T) \sim \text{Var}(\hat{R}_{ij}^0) T/T$, here $T$ is the integral time scale of $\hat{R}_{ij}^0$, defined as $T = \int \rho(t) \, dt$, with $\rho(t)$ the autocorrelation of $\hat{R}_{ij}^0$. The exact proportionality factor depends on the probability distribution of $\hat{R}_{ij}^0$ and is 2 for a normal distribution [49]. If we assume the eigenvalues close enough to their true value, a first order Taylor series can be used to approximate the propagation of uncertainty from matrix elements to eigenvalues [50], and the variance of the eigenvalues can be represented as a linear combination of the variance of the spectral covariance tensor, such that also here a $T^{-1}$ convergence can be expected.

### B. Eigenvalue spectrum and KL-dimension

In this section we take a more in depth look at the converged eigenvalue spectrum of the flow. Figure 3 (a) shows the eigenvalues as a function of the index number, which is a decreasing function due to the ordering. Similar, to the
classical boundary layer spectrum, three different regions seem to exist. A first region of the very large scale motions also known as the inactive range due to lack of contribution to shear stresses. A second region seems to adhere to a $\lambda \sim n^{-1}$ spectrum, seemingly related to the well known $k^{-1}$ scaling of the streamwise energy spectrum in turbulent boundary layers in the so-called shear production range \[51\], and finally, the inertial region where $\lambda \sim n^{-11/9}$ as was demonstrated by Ref. \[52\] and later proven more rigorously in Ref. \[53\]. It is clearly seen, certainly for the first modes, that the eigenvalues are not unique. This is due to the symmetries that exist in $\hat{R}_{ij}$. Depending on whether the originating wave vector is zero, has a single zero component or both non-zero, the eigenvalues are unique or have a multiplicity of 2 or 4. This is further illustrated in Table III, where the wave numbers, quantum numbers and eigenvalues are summarized of the most energetic modes. It is interesting to note that the wave numbers correspond to very long structures in the streamwise direction, in accordance with the subdivision on Fig. 3. Figure 3 (b) on the other hand shows the premultiplied spectrum $n\lambda_n/u_s^2$, giving a visual interpretation of the distribution of the energy over the different mode numbers.

The KL-dimension is the amount of POD methods necessary to capture 90% of energy on average and can be determined from the unresolved energy (see Fig. 2 (b)). The influence of the amount of samples on the KL-dimension is shown in Fig. 4. After an initial monotonous increase up to around 4000 samples (corresponding to a total averaging time of $6.3H/u_*$), $D_{KL}^\Delta$ reaches a steady state value of $9.77 \times 10^6$. The superscript $\Delta$ is added in the notation to indicate that results are based on a filtered velocity field $\tilde{u}$ with equivalent filter width $\Delta$. The KL-dimension is known to be an extensive property, and for sake of comparison, better expressed per unit volume (normalized by the BL-height), i.e. $d_{KL}^\Delta = D_{KL}^\Delta/|\Omega| = 1.9 \times 10^4 [H^{-3}]$. The averaging time used here is significantly less than typical used values for snapshot POD studies of the ABL or PDBL. The KL-dimension on the other hand is 2 to 3 orders of magnitude bigger than the numbers found by previous studies using snapshot POD of ABL (see Table \[I\]). The difference can be explained by the slow convergence of the method of snapshots for high dimensional systems \[5\]. The substantial increase of KL-dimension with Reynolds number was already demonstrated in Ref. \[9\], for channel flows of $Re_\tau = 180$ and $Re_\tau = 300$, also see Table \[I\]. In Ref. \[25\] a similar increase was found in the comparison of 1D vertical POD between DNS and LES.

C. Estimation of the effect of subgrid-scale energy

The slow decrease in the energy of the KL modes with increasing mode number indicates that the KL-dimension is sensitive to the fraction of unresolved energy. In this section we estimate this unresolved energy based on the power law scaling found in the inertial zone. We start by introducing the volume averaged total turbulent kinetic energy, which is, in analogy to $K_\Delta$, defined as $K \equiv |\Omega|^{-1} \int_{\Omega} (u'_i u'_j)/2 \mathrm{d}x$. Assuming that the LES filter cut-off is in the inertial range, it is easily shown, by integrating a $n^{-11/9}$ spectrum from $n$ to $\infty$, that the residual kinetic energy $K - K_n$ scales as $K - K_n \sim n^{-2/9}$, such that the normalized residual $E_n$ can be expressed as

$$E_n \equiv 1 - \frac{K_n}{K} = C_{KL} \left( \frac{n}{|\Omega'|} \right)^{-2/9},$$

with $C_{KL}$ and $K$ parameters that need to be further identified, and introducing the normalized volume $|\Omega'| = |\Omega|/H^3$ for notational convenience. Equation \[3\] is expected to hold for $x_3 \gg \Delta$, but is not valid close to the wall.
unresolved energy close to the wall is estimated in two steps: first the energy below the logarithmic region (i.e. the roughness sublayer), and secondly the contribution of the logarithmic region. The roughness sublayer is very narrow compared to the BL height \((z_0/H \ll 1)\), and although there is a peak of turbulent kinetic energy, its total contribution is therefore negligible. Similar considerations hold for smooth walls, for which, e.g., the peak of TKE scales with \(\langle u'_i u'_i \rangle_{\text{max}} \sim u_*^2 \log \Delta \) for smooth walls (see e.g. [54]) and becomes lower with increasing wall roughness \([55]\), while the width below the log region scales with \(\tau_{\text{log}} \sim H \log Re\), which equals e.g. \(5 \times 10^{-3}\) for \(Re\_c = 10^7\) a typical value in the atmospheric boundary layer. To estimate the contribution of the unresolved energy in the logarithmic region, we employ Townsend’s similarity hypothesis for the velocity fluctuation \([54]\), i.e. \(\langle u'_i u'_i \rangle / u_*^2 = B - A \log(x_3/H)\), usually expressed per velocity component \(\langle u'^2_i / u_*^2 \rangle = B_i - A_i \log(x_3/H)\), with \(B = B_1 + B_2 + B_3\) and \(A = A_1 + A_2\). Integrating from 0 to \(\Delta\) gives \(\Delta (A + B + B \log(\Delta/H))\), such that the energy fraction is \(\mathcal{O}(\Delta/H)\), which is typically \(\mathcal{O}(10^{-2})\) and therefore contributions to \(\mathcal{E}_n\) are expected to be of similar magnitude, and are further neglected.

We find the parameters \(C_{\text{KL}}\) and \(K\) in Eq. (3) by a least squares fit using the data of \(K_n\) from [IVB] in the range \(n/|\Omega'|\) from 10 to 10^3, resulting in \(C_{\text{KL}} = 1.57\) and \(K = 2.56 u_*^2\). In Fig. 5 we show the result of this fitting. We find from the asymptotic behaviour of \(\mathcal{E}_n\) (dashed orange curve) to hign \(n\), that still a significant portion of the energy \(\mathcal{E}_\Delta \approx 1 - K\Delta/K = \mathcal{E}_{N_n} = 8.3\%\) is unresolved. This is higher than, e.g., reported in [38, 40], and therefore we further verify this number based on an alternative method. To this end, the unresolved energy fraction is estimated from a height-dependent Kolmogorov energy spectrum \(E(k, z) = C_K k^{2/3} k^{-5/3}\), with \(C_K \approx 1.6\) the Kolmogorov constant and \(\epsilon\) the local dissipation of turbulent kinetic energy. For the dissipation \(\epsilon\) we use the usual hypothesis that local production equals dissipation such that \(\epsilon \approx \langle u_1 u_3 \rangle \partial \langle u_1 \rangle / \partial x_3 \approx \kappa^{-1} u_3^2 (1 - x_3/H) / x_3\), where the log-law is used for the mean velocity profile \(\langle u_1 \rangle / u_* = \kappa^{-1} \log(x_3/z_0)\), with \(\kappa \approx 0.4\) the Von Kármán constant. The viscous contribution of the shear stresses is ignored, such that \(\langle u_1 u_3 \rangle = u_*(1 - x_3/H)\). An estimate of the unresolved kinetic energy \(K - K_\Delta\) is obtained by integrating the energy spectrum \(E\) over the domain \(\Omega\) and from the cut-off wave number \(k_\Delta\) in wave number space, this gives

\[
K - K_\Delta \approx \frac{\pi/\Delta}{u_*^2} \int_{|\Omega|} \int_{k_\Delta} E(k, z) \, dk \, dz = \frac{2\pi}{\sqrt{3}} C_K (\kappa k_\Delta H)^{-2/3}, \tag{4}
\]

with \(k_\Delta = \pi/\Delta\), with \(\Delta = (\Delta_1 \Delta_2 \Delta_3)^{1/3}\) the characteristic grid spacing. The only input we need is \(K_\Delta\), the resolved kinetic energy from a single simulation, with its corresponding wave number \(k_\Delta\). In this way, we find that \((K - K_\Delta) = 0.237 u_*^2\), by using \(K_\Delta = 2.29 u_*^2\) we find that \(K = 2.52 u_*^2\) and \(\mathcal{E}_\Delta = 9.3\%\), which is remarkably close to the values found by the asymptotic behaviour of the grid and the POD-modes.

Finally, we compare the variance of the streamwise velocity component in our simulations with experimental data. Here we compare with measurements at the SLTEST site [56, 58], using \(u_* = 0.1884\) at \(H = 60\text{ m}\) from Ref. [56]. It is observed that the LES data (blue dots) are consistently lower than the measurement data (black dots). Corrected LES data for the unresolved energy (orange dots) are also shown in the figure, based on \(\langle u_1 u_1 \rangle - \langle \tilde{u}_1 \tilde{u}_1 \rangle \approx C_K k^{2/3} k_\Delta^{-2/3}\) and

\[
\mathcal{E}_\Delta = 9.3\%.
\]
unresolved energy fraction $E$ is respectively given by the resolved kinetic energy $K$, and is compared to POD modes, and it is found that for capturing the same amount of energy up to twice as many Fourier–Chebychev modes were ranked in decreasing energy order, which was determined using the full resolution simulation. Here, we investigate the effectiveness of LES compared to POD for a grid that is uniformly distributed and coarsened/refined in all directions – requiring no a priori ranking of modes. We use the data from Ref. [59], where a detailed LES error analysis was made of the LES error based on four different grid resolutions, each time refined by a factor of two in all directions. The finest grid used cell sizes $0.030H \times 0.015H \times 0.005H$, which is similar to the set-up used in the current work. The code used for the simulations was the same code as used in this study, for further simulation details we refer the reader to the corresponding paper. The number of degrees of freedom of a simulation is $n = 3N_1N_2N_3$, where the factor three stems from the different velocity components. The number of DOF per unit of normalized volume $n_\Delta/|\Omega|$ gives for the different cases $n_\Delta/|\Omega| = 3.75 \times 10^3$, $3.0 \times 10^4$, $2.4 \times 10^5$, $1.92 \times 10^6$. The resolved kinetic energy $K_\Delta$ is found by restricting the turbulent flow field from the finest grid to a different grid and then interpolating, and is respectively given by $K_\Delta = 1.42 u_1^2$, $1.82 u_2^2$, $2.08 u_3^2$, $2.25 u_4^2$. From Eq. [4] we find that the unresolved energy fraction $\mathcal{E}$ scales as $\mathcal{E}_\Delta \sim k_\Delta^{-2/3}$. Further using $k_\Delta \sim n^{1/3}$, we find (similar to the analysis above),

$$\mathcal{E}_\Delta \triangleq 1 - \frac{K_\Delta}{K} = C_\Delta \left( \frac{n_\Delta}{|\Omega|} \right)^{-2/9},$$

with $C_\Delta = 2.72$ and $K = 2.52 u_1^2$ obtained from a least squares fit of the LES data from Ref. [59]. This fit is visualized in blue in Fig. [4]. We find an unresolved energy fraction for the finest mesh of $\mathcal{E}_\Delta = 10.7\%$, which is similar to the unresolved fraction of the POD study above. This is expected since the meshes are quite similar.

FIG. 6: Variance of streamwise velocity component. (Blue, ◦) represents $\langle \tilde{u}_1 \tilde{u}_1 \rangle$, (orange, □) represents a correction for the unresolved energy $\langle u_1 u_1 \rangle - \langle \tilde{u}_1 \tilde{u}_1 \rangle$, (black, △) neutral ABL measurement data from Ref. [56]. The error bars indicate the 10% uncertainty intervals on the friction velocity $u_\ast$. Using $\mathcal{E} \triangleq 0.1$ gives a KL-dimension $d_{KL} = 2.41 \times 10^5$, which is more than a factor 10 larger than the value determined based on the filtered velocity field.

D. Comparison of POD representation to LES discretization

It is interesting to compare the effectivity of POD-modes for dimension reduction compared to simply using the pseudo-spectral discretization used in our LES. A similar performance analysis has been performed in Ref. [6], for a low Reynolds number channel flow. In this paper the performance of a Fourier–Chebychev basis, ranked in decreasing energy order, is compared to POD modes, and it is found that for capturing the same amount of energy up to twice as many Fourier–Chebychev modes where required compared to POD modes [6]. However, the Fourier–Chebychev modes were ranked in decreasing energy order, which was determined using the full resolution simulation.

In summary, we have found that the fraction of unresolved energy of the reference simulation is estimated at $\mathcal{E}_\Delta = 8.4\%$, and therefore, the KL dimension $d_{KL}^\Delta$, obtained in IV.B is an underestimation. The relation between the resolved $\mathcal{E}_n$ and unresolved fraction $\mathcal{E}_n^\Delta$ is found as $\mathcal{E}_n = (1 - \mathcal{E}_\Delta)\mathcal{E}_n^\Delta + \mathcal{E}_\Delta$. Inverting Eq. [3] leads to $n/|\Omega| = (\mathcal{E}/C_{KL})^{-9/2}$, which yields an expression for the number of modes required to express a specified unresolved fraction of energy $\mathcal{E}_\Delta$. Using $\mathcal{E}^M = 0.1$ gives a KL-dimension $d_{KL} = 2.41 \times 10^5$, which is more than a factor 10 larger than the value determined based on the filtered velocity field.
Note that this grid is far from optimized, constant grid spacing is used in the vertical direction and an ad-hoc aspect ratio has been used. The required degrees of freedom for regular grid cells to POD modes to attain the same accuracy is obtained by equating Eqs. 5 and 3 leading to \( n_\Delta / n_{KL} = (C_\Delta / C_{KL})^{9/2} = 11.8 \). We remark that this result is expected to both depend on the case as the discretization. Nonetheless it gives an interesting order of magnitude, indicating that for ABL flows, our pseudo-spectral discretization in combination with fourth-order FV in the vertical direction is quite effective when compared to POD. Obviously, this should not surprise, since in the horizontal directions, the Fourier-modes used in the pseudo-spectral method effectively correspond to the POD modes.

Finally, the constant \( C_\Delta \) can also be found directly from the simple model elaborated in the previous section for \( K - K_\Delta \). The amount of grid cells is directly related to the cut off wave number, \( k_\Delta H = \pi (n/|\Omega'|)^{1/3} \). In this way, a proportionality factor can be determined analytically, with some algebra this is found as

\[
(C_\Delta^M)^{-1} = \frac{\sqrt{3}}{2\pi^{1/3}} C_R^{-1} K_\Delta^{2/3} K_\Delta / u_*^2 + \left( \frac{n_\Delta}{|\Omega'|} \right)^{-2/9},
\]

where the second part of the expression becomes negligible when a fine reference grid (high \( n_\Delta \)) is used for determining \( K_\Delta \). This leads to a value \( C_\Delta^M = 2.03 \), which is in reasonable agreement 2.72 obtained from a direct fit on the LES data.

V. CONCLUSIONS

We performed a Karhunen–Loève decomposition for a LES of a PDBL. We find regimes for which the eigenvalues demonstrate a -1 and -11/9 scaling behaviour, related to the shear production and inertial range. We conclude that to resolve 90% of the TKE on average – the so called KL-dimension, \( 2.41 \times 10^9 H^{-3} \) modes per unit volume are needed, which is significantly higher than indicated by previous studies. We also investigate the DOF required to both depend on the case as the discretization. Nonetheless it gives an interesting order of magnitude, indicating that approximately 11 times more DOF are needed that POD modes.

[1] K. Pearson, On lines and planes of closest fit to systems of points in space, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2, 559 (1901).
[2] L. Sirovich, Chaotic dynamics of coherent structures, Physica D: Nonlinear Phenomena 37, 126 (1989).
[3] G. Berkooz, P. Holmes, and J. L. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, Annual Review of Fluid Mechanics 25, 539 (1993).
[4] L. Sirovich, Turbulence and the dynamics of coherent structures. I. Coherent structures, Quarterly of Applied Mathematics 45, 561 (1987).
[5] A. Duggleby, K. S. Ball, and M. Schwaenen, Structure and dynamics of low Reynolds number turbulent pipe flow, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 367, 473 (2008).
[6] K. Ball, L. Sirovich, and L. Keefe, Dynamical eigenfunction decomposition of turbulent channel flow, International Journal for Numerical Methods in Fluids 12, 585 (1991).
[7] G. A. Webber, R. Handler, and L. Sirovich, The Karhunen–-Loeve decomposition of minimal channel flow, Physics of Fluids 9, 1054 (1997).
[8] L. Sirovich, K. Ball, and R. Handler, Propagating structures in wall-bounded turbulent flows, Theoretical and Computational Fluid Dynamics 2, 307 (1991).
[9] K. Iwamoto, Y. Suzuki, and N. Kasagi, Reynolds number effect on wall turbulence: toward effective feedback control, International journal of heat and fluid flow 23, 678 (2002).
[10] K. D. Houisiadas, A. N. Beris, and R. A. Handler, Viscoelastic effects on higher order statistics and on coherent structures in turbulent channel flow, Physics of Fluids 17, 035106 (2005).
[11] T. Smith, J. Moehlis, and P. Holmes, Low-dimensional models for turbulent plane Couette flow in a minimal flow unit, Journal of Fluid Mechanics 538, 697 (2005).
[12] S. Ciliberto and B. Nicolaenko, Estimating the number of degrees of freedom in spatially extended systems, EPL (Europhysics Letters) 14, 303 (1991).
[13] L. Sirovich and A. E. Deane, A computational study of Rayleigh–Bénard convection. Part 2. Dimension considerations, Journal of Fluid Mechanics 222, 251 (1991).
[14] A. Duggleby, K. S. Ball, M. R. Paul, and P. F. Fischer, Dynamical eigenfunction decomposition of turbulent pipe flow, Journal of Turbulence , N43 (2007).
[15] J. Cardillo, Y. Chen, G. Araya, J. Newman, K. Jansen, and L. Castillo, DNS of a turbulent boundary layer with surface roughness, Journal of Fluid Mechanics 729, 603 (2013).
[16] C. VerHulst and C. Meneveau, Large eddy simulation study of the kinetic energy entrainment by energetic turbulent flow structures in large wind farms, Physics of Fluids 26, 025113 (2014).
for molecular dynamics, The Journal of Chemical Physics 126, 244101 (2007).
[51] A. Perry, S. Henbest, and M. Chong, A theoretical and experimental study of wall turbulence, Journal of Fluid Mechanics 165, 163 (1986).
[52] B. Knight and L. Sirovich, Kolmogorov inertial range for inhomogeneous turbulent flows, Physical Review Letters 65, 1356 (1990).
[53] R. D. Moser, Kolmogorov inertial range spectra for inhomogeneous turbulence, Physics of Fluids 6, 794 (1994).
[54] C. Meneveau and I. Marusic, Generalized logarithmic law for high-order moments in turbulent boundary layers, Journal of Fluid Mechanics 719 (2013).
[55] J. Jiménez, Turbulent flows over rough walls, Annu. Rev. Fluid Mech. 36, 173 (2004).
[56] I. Marusic, J. P. Monty, M. Hultmark, and A. J. Smits, On the logarithmic region in wall turbulence, Journal of Fluid Mechanics 716 (2013).
[57] A. Townsend, The structure of turbulent shear flow, Cambridge and New York, Cambridge University Press, 1976. 438 p. (1976).
[58] N. Hutchins, K. Chauhan, I. Marusic, J. Monty, and J. Klewicki, Towards reconciling the large-scale structure of turbulent boundary layers in the atmosphere and laboratory, Boundary-Layer Meteorology 145, 273 (2012).
[59] P. Bauweraerts and J. Meyers, On the feasibility of using large-eddy simulations for real-time turbulent-flow forecasting in the atmospheric boundary layer, Boundary-Layer Meteorology 171, 213 (2019).