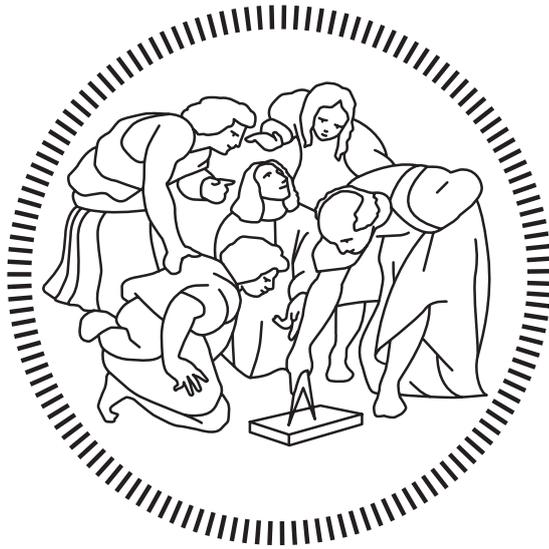


**Politecnico di Milano**

---

SCHOOL OF INDUSTRIAL AND INFORMATION ENGINEERING  
DEPARTMENT OF AEROSPACE SCIENCE AND TECHNOLOGY  
Master of science - Aeronautical Engineering



# Wall turbulence: an exploratory approach to analyse roughness

Supervisor  
**Maurizio QUADRIO**

Co-Supervisor  
**Andrea SCHILLACI**

Candidate  
**Marco NEGRI - 921040**

---

Academic Year 2020 – 2021



# Acknowledgements

La mia Tesi Magistrale conclude un percorso che posso entusiasticamente ritenere di avere vissuto a fondo. In ogni sua sfumatura, in ogni suo profumo. In Università, ho avuto la fortuna di percepire sulla mia pelle le emozioni che devono colorare le giornate di un ventenne. Esco con la testa sulle spalle, dopo esserci entrato con l'ingenuità di un ragazzino. Non riesco ancora ad immaginare cosa mi riserverà il futuro. Adesso, però, voglio pensare al presente. Lo voglio fare ringraziando quelle persone che sono state importanti in questo periodo, che ora posso felicemente etichettare come "passato". È soprattutto grazie a voi se ora taglio questo piccolo, ma enorme traguardo. Con il sorriso a trentadue denti, con il cuore a mille. Voglio dedicare a voi un pensiero, un po' a modo mio. Una sovrapposizione di immagini. Sperando che voi, leggendo queste parole, possiate vedere le emozioni che vorrei trasmettervi, prima ancora di sentirle.

Ringrazio due persone con la quale ho assunto la definitiva consapevolezza delle mie potenzialità e lo spirito per affrontare la mia prossima esperienza professionale. Qualunque essa sia.

Il Professore Maurizio Quadrio. Per aver creduto in me da subito, delineando un progetto che mi rispecchiasse alla perfezione: innovativo; impegnativo; ambizioso; pericolosamente rischioso; terribilmente realizzante. Andrea. Per avermi seguito con la professionalità di un Dottorando e con l'affetto di un Amico. Per avermi sempre sostenuto e difeso. Per avermi reso partecipe del suo mondo di Ricerca.

Ringrazio le persone a me care. Il kernel delle mie energie. La curvatura del mio stato d'animo. Il flusso delle mie emozioni.

La mia famiglia. Per avermi dato la possibilità di coltivare e realizzare il mio sogno più grande. Per avermi protetto dai primi pianti in culla. Per proteggermi, ancora oggi, dai pianti disperati di stanchezza. Per essere le prime persone a credere in me. Nonostante i momenti bui. Le macchie indelebili. Le difficoltà.

I miei migliori amici: Antonio, Daniele e Rachele. Non servono parole per esprimervi quanto voi siate importanti per me. Non riuscirei nemmeno a trovarne. Tante splendide esperienze delineano il nostro rapporto. Tante altre lo arricchiranno in futuro. Non vedo l'ora di viverle tutte.

Gli amici della mia compagnia di casa. Di Lecco. Di Milano. I miei (ormai ex) compagni di allenamento. I miei (ormai ex) compagni di corso. Che sia stata un'ora, un giorno o che siano stati anni. Che sia stato un caffè o una birra al bar. Lungolago

o sui Navigli. Qualche pagaita sul lago. Due chiacchiere tra una lezione e l'altra. Non importa la natura dell'occasione, nè la sua durata. Conta solo che, in quei momenti, eravamo insieme e spensierati. A dimenticarci dei nostri problemi. A condividere una parte di noi. A sognare il futuro. Attimi tanto fugaci quanto significativi che ci hanno permesso di crescere e di diventare le persone che siamo oggi.

Non sono mai stato davvero bravo ad accettare i complimenti. Forse per diffidenza. Forse per paura. Oggi, però, voglio riconoscermi il merito di non avere mai smesso di sognare. Di avere sempre coltivato un briciolo di magia. Di essermi sempre lasciato guidare dai miei obiettivi. Senza mai mollare. Imparando dagli errori. Rialzandomi dalle cadute.

Un calice di vino per brindare. Al futuro che verrà.

# Sommario

Una rappresentazione nuova delle superfici di rugosità, che sfrutta tecniche di analisi di forma statistica e di apprendimento automatico (AA), viene presentata in questo lavoro.

Viene generato un database di geometrie di rugosità che permette di investigare l'effetto di diverse proprietà topografiche. L'altezza superficiale di tali geometrie, scalate in unità viscosa, è scritta come combinazione lineare delle autofunzioni dell'operatore di Laplace-Beltrami (LB), discretizzato su una parete piana di riferimento con le stesse dimensioni fisiche e risoluzione di griglia delle superfici ruvide. I coefficienti di tali equazioni sono poi calcolati con il metodo Least Absolute Shrinkage and Selection Operator (LASSO), che individua i predittori (autofunzioni di LB) rilevanti nella definizione delle differenti geometrie. Infine, si valuta la possibilità di calcolare una correlazione universale di rugosità, che lega l'altezza di rugosità del granello di sabbia equivalente  $k_s$  relativa ad una superficie ruvida alle sue proprietà topografiche, e che rappresenta il principale obiettivo dell'approccio stato dell'arte di affrontare il problema di rugosità.

Questo lavoro propone una rappresentazione fisicamente affidabile della rugosità attraverso un modello che utilizza un numero relativamente piccolo di predittori, rappresentando così un input efficiente per diverse applicazioni di AA, come la predizione di modelli attraverso reti neurali addestrate, che potrebbero rappresentare lo strumento principale di analisi della rugosità negli studi futuri.

**Parole Chiave:** Rugosità; Statistica; Laplace-Beltrami; LASSO; Correlazione; Apprendimento automatico.



# Abstract

A new representation of roughness, which exploits statistical shape analysis and machine learning (ML) techniques, is here presented.

A geometric database is generated to investigate the effect of several topographical features. The surface elevation of these geometries, scaled in viscous units, is written as a linear combination of the eigenfunctions of the Laplace-Beltrami (LB) operator, discretized on a reference smooth wall with the same physical domain sizes and grid resolution of the rough surfaces. Coefficients of these equations are then computed with the Least Absolute Shrinkage and Selection Operator (LASSO) method, which highlights the relevant predictors (LB eigenfunction) in the definition of the different geometries. Lastly, it is assessed the possibility of computing a universal roughness correlation tying the equivalent sand-grain roughness height  $k_s$  related to a rough surface to its topographical properties, which is the main goal of the state-of-art approach to address the roughness problem.

This works proposes a physically reliable representation of roughness by means of a model which uses a relatively small number of predictors, thus representing an efficient input for many ML applications, as model predictions through neural networks, that might represent the main tool to analyse roughness in future works.

**Key Words:** Roughness; Statistics; Laplace-Beltrami; LASSO; Correlation; Machine Learning.



# Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>Sommario</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Contents</b>	<b>1</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Roughness problem . . . . .	1
1.2 Outline . . . . .	3
<b>2 Roughness</b>	<b>5</b>
2.1 State-of-art . . . . .	5
2.2 Surface properties . . . . .	6
2.3 Surface correlations . . . . .	9
<b>3 Method</b>	<b>11</b>
3.1 Database generation . . . . .	11
3.2 Description of the method . . . . .	15
<b>4 Experiment and Results</b>	<b>19</b>
4.1 Application of the method . . . . .	19
4.2 Results . . . . .	22
<b>Conclusions</b>	<b>31</b>
<b>A MATLAB Code to compute LB eigenfunctions</b>	<b>35</b>
<b>B Details of the experiment</b>	<b>39</b>
B.1 Numerical error . . . . .	39
B.2 Parameters tuning . . . . .	40
B.2.1 Effect of $n$ . . . . .	40
B.2.2 Effect of $\lambda$ . . . . .	41
<b>Bibliography</b>	<b>50</b>



# List of Figures

Figure 1.1	<i>Inset</i> : The setup of the roughness problem. <i>Main figure</i> : A close-up view of the rectangular region of roughness topography in the inset. . . . .	2
Figure 2.1	Surface sketches to illustrate surface total plan area $A_t$ and roughness element frontal ( $A_f$ ) and plan ( $A_p$ ) area, as well as clustering and directionality. Empty and filled arrows in sketch $g$ indicate different flow directions. . . . .	7
Figure 3.1	( <i>Left</i> ) Element profile with $n = 0.5$ ; ( <i>Middle</i> ) $n = 1$ (conical); ( <i>Right</i> ) $n = 1.5$ . All the elements have the same $k_p$ and $\gamma$ . . . . .	12
Figure 3.2	From left to right: first seven Laplace-Beltrami eigenfunctions computed on <i>first row</i> : A7088; <i>second row</i> : smooth wall with the same domain dimensions and grid resolution of the rough surface. Colored scale indicates pointwise values of the LB basis. . . . .	16
Figure 4.1	Selected geometry samples. All dimensions are normalized with effective channel half-height $h$ . Colored scale indicates surface elevation. <i>First row</i> : (a) A7088 (reference geometry); <i>second row</i> : (b) B7088, (c) C7088 - geometries with the same $ES$ and $\Delta$ as the reference one but different $S_k$ ; <i>third row</i> : (d) A3588, (e) A0088 - geometries with the same $S_k$ and $ES$ as the reference one but different $\Delta$ ; <i>fourth row</i> : (f) A7060, (g) A7040 - geometries with the same $S_k$ and $\Delta$ as the reference one but different $ES$ . Each sample has the same $k_{rms} = 0.045h \pm 2\%$ and $K_u = 2.61 \pm 5\%$ . . . . .	23
Figure 4.2	Absolute value of LASSO coefficients between surface elevation vectors of A7088 geometries and the first 500 LB eigenfunctions on the reference plane, computed with $\lambda = 1$ . (a) Original geometry; (b) <i>first row</i> : geometries obtained by shuffling the reference A7088 center locations; <i>second row</i> : geometries obtained by shuffling the reference A7088 peak sizes. . . . .	24
Figure 4.3	Absolute value of LASSO coefficients between surface elevation vectors of the twenty-five database geometries and the first 500 LB eigenfunctions on the reference plane, computed with $\lambda = 1$ . . . . .	25

Figure 4.4	Absolute value of correlation coefficients between LASSO coefficients of the A7088 collection and their ( <i>first row</i> ) $k_s/k_{rms}$ ; ( <i>second row</i> ) $k_s/k_p$ ; ( <i>third row</i> ) $k_s/k_t$ ; ( <i>fourth row</i> ) $k_s/k_z$ vector. . . . .	28
Figure 4.5	Correlation matrix between LASSO coefficients of all the different geometries and their ( <i>first row</i> ) $k_s/k_{rms}$ ; ( <i>second row</i> ) $k_s/k_p$ ; ( <i>third row</i> ) $k_s/k_t$ ; ( <i>fourth row</i> ) $k_s/k_z$ vectors. Colored scale indicates correlation coefficient values. . . . .	29
Figure 4.6	Rough surfaces characterized by ( <i>left</i> ) $\Delta U^+ = 7.60$ ; ( <i>middle</i> ) $\Delta U^+ = 7.67$ ; ( <i>right</i> ) $\Delta U^+ = 7.99$ . All dimensions are normalized with effective channel half-height $h$ . Colored scale indicates surface elevation. . . . .	30
Figure B.1	Absolute value of LASSO coefficients between surface elevation vectors of A7088 geometries and the first 500 LB eigenfunctions on the reference plane, computed with $\lambda = 1$ . ( <i>a</i> ) Original geometry; ( <i>b</i> ) <i>first row</i> : geometries obtained by shuffling the reference A7088 center locations; <i>second row</i> : geometries obtained by shuffling the reference A7088 peak sizes. Blue points are computed when <i>case 1</i> (explained in section B.1) is considered; orange points are computed when <i>case 2</i> is considered. . . . .	40
Figure B.2	Correlation matrix between LASSO coefficients of all the different geometries and their ( <i>first row</i> ) $k_s/k_{rms}$ ; ( <i>second row</i> ) $k_s/k_p$ ; ( <i>third row</i> ) $k_s/k_t$ ; ( <i>fourth row</i> ) $k_s/k_z$ vectors. Coefficients are computed with $\lambda = 1$ and $n = 300$ . Colored scale indicates correlation coefficient values. . . . .	43
Figure B.3	Correlation matrix between LASSO coefficients of all the different geometries and their ( <i>first row</i> ) $k_s/k_{rms}$ ; ( <i>second row</i> ) $k_s/k_p$ ; ( <i>third row</i> ) $k_s/k_t$ ; ( <i>fourth row</i> ) $k_s/k_z$ vectors. Coefficients are computed with $\lambda = 1$ and $n = 1000$ . Colored scale indicates correlation coefficient values. . . . .	44
Figure B.4	Correlation matrix between LASSO coefficients of all the different geometries and their ( <i>first row</i> ) $k_s/k_{rms}$ ; ( <i>second row</i> ) $k_s/k_p$ ; ( <i>third row</i> ) $k_s/k_t$ ; ( <i>fourth row</i> ) $k_s/k_z$ vectors. Coefficients are computed with $\lambda = 0.1$ and $n = 500$ . Colored scale indicates correlation coefficient values. . . . .	45
Figure B.5	Correlation matrix between LASSO coefficients of all the different geometries and their ( <i>first row</i> ) $k_s/k_{rms}$ ; ( <i>second row</i> ) $k_s/k_p$ ; ( <i>third row</i> ) $k_s/k_t$ ; ( <i>fourth row</i> ) $k_s/k_z$ vectors. Coefficients are computed with $\lambda = 2$ and $n = 500$ . Colored scale indicates correlation coefficient values. . . . .	46

# Chapter 1

## Introduction

The roughness problem formulation is presented in section 1.1 to understand the relationship between roughness topography and drag behaviour, while section 1.2 reports the outline of the paper.

### 1.1 Roughness problem

In fluid mechanics, whether a surface is smooth or rough is judged by the flow occurring next to it. For turbulent flows, a surface must be considered rough if its topographical features are large enough to disrupt the smallest eddies near the wall, thus altering the transfer of mass, momentum and heat [6]. Each roughness topography affects flow differently, making prediction of performance-critical quantities (e.g. drag and heat transfer) insufficiently reliable, with uncertainties costing billions of dollars per year. As such, roughness is still an active research area [17]. The main challenge has been the large number of relevant topographies and the cost associated with testing each of them to identify key roughness parameters and quantify their influence. However, recent advances in computations and experiments have allowed to reach unprecedented levels of detail and accuracy. Specifically, these studies have systematically highlighted the effects of different topographical properties and explored the far reaches of the Moody chart, which has been - and it still is - the most widely used tool for the prediction of skin friction in flows over rough walls [9].

The roughness problem setup is depicted in figure 1.1. The desired output is the wall shear stress  $\tau_w$  [ $N/m^2$ ] or the wall heat flux  $q_w$  [ $W/m^2$ ], given any roughness topography described by a physical scale  $k$ , here chosen to be the maximum peak-to-trough roughness height  $k_t$ ; flow at wall-normal distance  $y$  characterized by mean velocity  $U(y)$ , mean temperature  $\Theta(y)$  and boundary-layer thickness  $\delta$ ; fluid properties such as kinematic viscosity  $\nu$  [ $m^2/s$ ], thermal diffusivity  $\alpha$  [ $1/K$ ], density  $\rho$  [ $kg/m^3$ ] and specific heat at constant pressure  $c_p$  [6]. Each unit of measure is expressed in the SI reference. In dimensionless form, introducing the subscript  $+$  to denote viscous-friction scaling, the problem can be stated as finding  $\tau_w/(\rho U^2/2) = 2/(U^{+2})$  and  $q_w/(\rho c_p U \Theta) = 1/(U^+ \Theta^+)$  as function of  $y^+$ ,  $k^+$  and Pr. In this case, the friction-

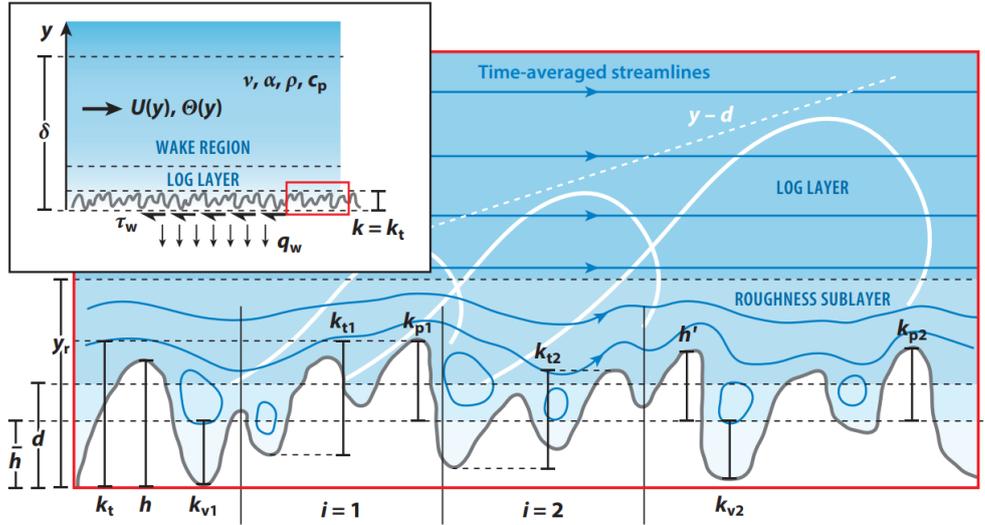


Figure 1.1. *Inset*: The setup of the roughness problem. *Main figure*: A close-up view of the rectangular region of roughness topography in the inset.

scaled mean velocity is  $U^+ \equiv U/u_\tau$  and the friction-scaled mean temperature is  $\Theta^+ \equiv \Theta/\theta_\tau$ , with  $u_\tau \equiv (\tau_w/\rho)^{1/2}$  and  $\theta_\tau \equiv (q_w/\rho c_p u_\tau)$  being the friction velocity and the friction temperature, respectively. Lastly,  $y^+ \equiv y u_\tau/\nu$  is the viscous-scaled wall distance,  $k^+ \equiv k u_\tau/\nu$  is the roughness Reynolds number and  $Pr \equiv \nu/\alpha$  is the Prandtl number. The basic formulation here presented is preserved if the problem is posed, for pipe or channel flows of cross section  $\int dA$ , using the bulk velocity  $U_b \equiv \int U dA$ , mixed mean temperature  $\Theta_m \equiv \int U \Theta dA / \int U dA$  and pipe radius of half channel  $\delta$ , or, for boundary layers, using the freestream velocity  $U_\delta \equiv U(y = \delta)$ , temperature  $\Theta_\delta \equiv \Theta(y = \delta)$  and boundary-layer thickness  $\delta$  [6].

In the close-up view,  $h$  describes the roughness height function. Its average position is located at  $y = \bar{h}$ , from which the variation  $h'$  is measured. The roughness sublayer ( $y < y_r$ ) is a region near the wall in which the flow is influenced by the local roughness topography. In the region above it (log layer), the time-averaged flow (*blue streamlines*) is spatially homogeneous, meaning that smooth- and rough-wall turbulence behave similarly, accordingly to the *Outer-Layer Similarity* [31] theory. Specifically, this important assumption, on which all predictive models rely, declares that friction-scaled turbulent relative motions in the outer layer are independent of the surface condition at sufficiently high Reynolds numbers and large scale separations. Lastly, a key parameter to be considered is the wall offset  $d$ , which is a wall-normal coordinate depending on both the roughness topography and  $k^+$ . The outer turbulent flow does not perceive its origin to be at  $y = 0$  but at  $y = d$ . Specifically, for drag reduction roughness (e.g. riblets), the virtual origin of the flow velocity profile is known as the "protrusion height" [22]. This quantity, if properly dimensionless for the period of the corrugation, is a purely geometric parameter which depends only on the shape of the wall corrugations.

To address the roughness problem, many works in literature, e.g. [7–9, 24], have tried

to tie either the roughness function  $\Delta U^+$ , i.e. the shift at matched  $y^+$  of the velocity profile in the log layer relative to that of the smooth wall, or  $k_s$ , a hydraulic scale - described in Chapter 2 - which can be directly computed from  $\Delta U^+$  in the fully-rough regime, to some roughness topographical properties. However, "an all-encompassing equation relating  $k_s$  to topography remains elusive" [6]. Furthermore, advances in this framework are only linked to an improved appreciation of key topographical parameters. The main purpose of this work is thus to analyse roughness from a different point of view, proposing a method which uses statistical shape analysis and machine learning (ML) techniques. Specifically, a powerful tool to study surfaces, which has applications across multiple domains including statistical shape analysis [1, 13], shape correspondences [14] and co-segmentation [37], is exploited. This method allows to compare rough surfaces independently from their topographical properties, and offers a representation of these geometries which is suitable for several ML applications, e.g. feature extraction and model prediction, in which many recent works, e.g. [18, 21], are acquiring a growing interest.

## 1.2 Outline

The layout of this paper is outlined as follows. Chapter 2 describes the state-of-art approach to address the roughness problem, underlying its underpinnings and limits. Chapter 3 reports an innovative representation of roughness. In particular, section 3.1 describes the database generation procedure, while section 3.2 proposes the method on which this work relies. The application of this method to the computed dataset and the results obtained accordingly are then presented in Chapter 4 . Lastly, final conclusions are exposed, along with the main possible future developments of this work.



# Chapter 2

## Roughness

The state-of-art approach to address the roughness problem is described in section 2.1. Consequently, the main topographical features, which this method aims to correlate with the roughness function  $\Delta U^+$  related to a rough surface, are summarized in section 2.2. Lastly, some of the most widely used roughness correlations are briefly discussed in section 2.3.

### 2.1 State-of-art

For fully-developed turbulent wall-bounded flows over smooth and isothermal walls, the viscous sublayer is scaled by  $\nu/u_\tau$ , and when the roughness height  $k$  is small relative to this scale ( $k^+ \ll 1$ ), roughness is submerged below the viscous sublayer and thus the surface appears smooth to the flow. Then, if the outer Reynolds number is sufficiently large ( $\delta \gg \nu/u_\tau$ ), an inertial range where viscosity does not evidently matter ( $y \gg \nu/u_\tau$ ) emerges near the wall which is independent from the outer-flow geometry [6]. In this region, the velocity profile can be computed through the log law equation:

$$U_S^+ = (1/k) \ln y^+ + A \quad (2.1)$$

where the subscript  $S$  stands for smooth conditions,  $k$  ( $\approx 0.4$ ) is the von Kármán constant and  $A$  ( $\approx 5$ ) indicates the log-law intercept, universal for smooth walls ( $k^+ \ll 1$ ). A similar relation can be obtained even if the surface is rough such that  $k^+$  is no longer small, but the log-law intercept becomes function of both roughness topography and roughness Reynolds number:

$$U_R^+ = (1/k) \ln (y/k) + B(k^+) \quad (2.2)$$

where the subscript  $R$  stands for rough conditions. If viscous effects can be neglected, then for  $k^+ \gg 1$  the intercept  $B$  approaches a finite value  $B(\infty)$  which is independent of roughness Reynolds number but depends only on roughness topography.

Equation 2.2 suggests that the unknown function  $B(k^+)$ , which can be computed by subtracting 2.1 to 2.2, determines the effect of a rough surface on the flow occurring

next to it. It is thus obtained a relation for the roughness function  $\Delta U^+$ :

$$\Delta U^+(k^+) = U_S^+ - U_R^+ = (1/k) \ln k^+ + A - B(k^+) \quad (2.3)$$

$\Delta U^+$  describes the momentum deficit resulting from surface roughness and thus represent a measure of drag penalty relative to a smooth wall [12]. Surfaces with  $\Delta U^+ = 0$  are hydrodynamically smooth since viscosity damps out the perturbations caused by roughness, whereas drag increases or reduces for  $\Delta U^+ > 0$  and  $\Delta U^+ < 0$ , respectively.

A roughness Reynolds number  $k^+$  must be set to solve equation 2.3. As such, the equivalent sand-grain roughness  $k_s$  on a surface, inspired by the pioneering studies of Nikuradse [27], is introduced. This parameter is defined as the size of uniform, close-packed sand grains on a hypothetical surface that would produce the same friction factor as the surface of interest if exposed to the same flow in the fully-rough regime. Based on this hydraulic scale, the roughness function for all surfaces can be written as:

$$\Delta U^+(k_s^+) = (1/k) \ln k_s^+ + A - B_s(k_s^+) \quad (2.4)$$

where  $B_s(k_s^+)$  is the unique log-law intercept of  $U_R^+$  for uniform sand grains. In the fully-rough regime, it assumes the limit value of  $B_s(\infty) = 8.5$ .

Lastly, by adding a wake function  $W$  of strength  $\Pi$ , representing the departure of the mean velocity profile from the log law in the outer layer (for wall-bounded turbulent flows), it is possible to write the following expression for the mean velocity profile, valid from the log region up to  $y = \delta$ :

$$U^+ = (1/k) \ln y/k_s + B_s(k_s^+) + (\Pi/k)W(y/\delta) \quad (2.5)$$

At the edge of the boundary layer,  $U^+ = U_\delta^+ = \sqrt{2/C_f}$ , thus equation 2.5 reduces to an implicit relation that can be solved to make full-scale predictions of skin friction coefficient  $C_f(k_s/\delta, Re_\delta)$ , known  $B_s$  as function of  $k_s^+$  and  $Re_\delta = (\delta \cdot U_\delta)/\nu$ .

## 2.2 Surface properties

The critical point in making full-scale drag predictions regards the possibility to assign  $k_s$ , based only on topographical features:

$$k_s = f(\text{measured topographical properties}) \quad (2.6)$$

The ultimate challenge is to identify what is the bare-minimum set of invaluable topographical properties, which should be clearly defined and measurable, required to estimate the drag on a surface.

Although it is very difficult to characterize roughness from the wide range of geometric morphologies, a correlation between surface topography and wall drag is likely to require measures of (1) roughness height  $k$ ; (2) frontal solidity  $\lambda_f$ , effective slope  $ES$  or roughness density parameter  $\Lambda_s$ ; (3) plan solidity  $\lambda_p$  or skewness  $S_k$  [6]. Finally, clustering and directionality should also be accounted.

- **Roughness height**

Some measure of roughness height is needed to relate a surface topography to  $k_s$ . In literature, the most adopted choices of this parameter include (a) average roughness height  $k_a \equiv 1/A_t \int |h'| dA$ , where  $A_t$  is the total plan area; (b) root-mean-square roughness height  $k_{rms} \equiv \sqrt{1/A_t \int h'^2 dA}$ , i.e. standard deviation of roughness elevation; (c) maximum peak-to-trough elevation  $k_t$ . The former two, involving area integrals, are less corrupted by extreme asperities and thus represent more reliable measures of mean surface elevation. To partially address the issue concerning the latter,  $k_t$  is often replaced by a subsample-average peak-to-trough roughness amplitude  $k_z \equiv 1/N \sum_{i=1}^N (k_{pi} + k_{vi})$ , where  $k_{pi} + k_{vi}$  is the highest peak-to-trough height of the entire  $i_{th}$  sample.

- **Frontal projected area parameter**

A parameter regarding the frontal projected area of roughness elements  $A_f$  is probably required. Frontal solidity  $\lambda_f$ , for an array of wall-mounted roughness elements as depicted in Figure 2.1 [6], is simply computed as  $\lambda_f \equiv A_f/A_t$ . It

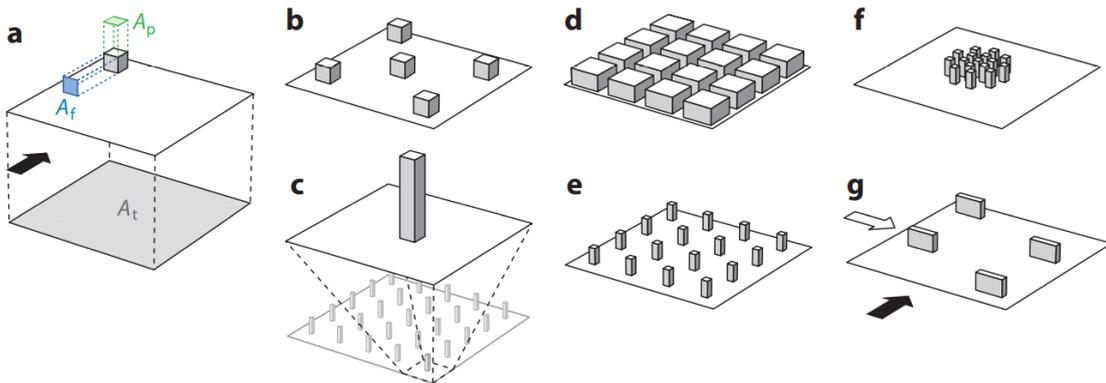


Figure 2.1. Surface sketches to illustrate surface total plan area  $A_t$  and roughness element frontal ( $A_f$ ) and plan ( $A_p$ ) area, as well as clustering and directionality. Empty and filled arrows in sketch  $g$  indicate different flow directions.

measures the available area exposed to pressure drag. There are increasing expanses of smooth-like conditions between isolated roughness elements ( $\lambda_f \rightarrow 0$ ); similarly, very densely packed roughness elements are sheltered due to proximity to neighbors. Between these regimes, there is a range, typically  $0.1 < \lambda_f < 0.3$ , where drag is maximum [17].

Ref. [35] proved that frontal solidity is equivalent to half of the mean absolute streamwise gradient of the rough surface,  $ES \equiv 1/A_t \int |\frac{\partial h'}{\partial x}| dA$ , suggesting that low  $\lambda_f$  could be considered as either sparsely packed (sketch **b** in figure 2.1) or long-wavelength shallow roughness, as well as high  $\lambda_f$  could be thought of as either densely packed (sketch **d** in figure 2.1) or short-wavelength steep roughness.

Another choice for a roughness density parameter widely used in literature is

$\Lambda_f \equiv \lambda_f^{-1}(A_f/A_w)$ , where  $A_w$  is the total windward wetted surface area. Accordingly to [2], the local value of  $\Lambda_f$  is related to the local surface mean angle  $\alpha_h \equiv \arctan(\frac{\partial h'}{\partial x})$  for small  $\alpha_h$ , and argued that the latter is a more appropriate parameter to use in a roughness correlation as it is easier to measure using available techniques. Since  $|\alpha_h|$  is approximately  $ES$  locally, only one of these variables is sufficient for correlations.

- **Plan area parameter**

A plan area parameter is indispensable to distinguish between geometries with different packing densities and element aspect ratios but with matched  $\lambda_f$  and height, like those reported in sketches (b,c) in figure 2.1. When combined with  $\lambda_f$ , plan solidity  $\lambda_p \equiv A_p/A_t$  gives indeed an indication about the aspect ratio of the roughness elements.  $\lambda_p \rightarrow 0$  indicates sparsely packed arrangements, whereas  $\lambda_p \rightarrow 1$  imply dense roughness. For wall-mounted cubes,  $\lambda_p$  is related to surface skewness  $S_k \equiv (1/A_t \int h'^3 dA)/k_{rms}^3$ , a measure of the asymmetry in surface elevation distribution. However, this relationship does not hold in general and  $S_k$ , unlike  $\lambda_p$ , is easy to calculate for irregular surfaces [9], thus representing a more suitable parameter for converting from topography to  $k_s$ .

- **Clustering and directionality**

All the topographical parameters described until now fail to capture clustering and directionality. For examples, sketches (e,f) and (f,g) in figure 2.1 show a comparison between surfaces with the same ( $k_a, \lambda_f, \lambda_p$ ) and all higher-order moments (due to their identical elevation distributions), yet different arrangements and levels of directionality, respectively. In both cases, different outcomes on the flow are expected.

Ref. [34] analysed surfaces with identical topographical properties, but varying degrees of clustering, discovering a decrease in  $\Delta U^+$  with increased clustering. Similarly, ref. [9] realised that, for geometries with matched  $k_{rms}$ , monodisperse roughness elements lead to a larger  $k_s$  than polydisperse ones. The concept of directionality is instead due to either surface shape or alignment of roughness elements. Sketch f in figure 2.1 shows an isotropic roughness, in the sense that drag is relatively invariant to flow direction, while the anisotropic surface in sketch g has aligned elements, thus the pressure drag is expected to be very different if the flow came from the direction of the open and closed arrow. In particular, ref. [3] shown that spanwise-aligned surface patterns tend to have a higher drag penalty than streamwise alignments. Generally, random roughness is isotropic, while a regular surface is often anisotropic, with high directionality leading to atypical drag behaviours [6].

## 2.3 Surface correlations

There are in general two types of correlation in literature: those directly predicting  $k_s$  and those predicting  $\Delta U^+$  [38].

For transitionally-rough flows, the latter group is more suitable since  $k_s$  is essentially defined only when the flow is fully-rough. Ref. [35] proposed a correlation to predict  $\Delta U^+$  based on a roughness parameter  $\lambda_T$ :

$$\Delta U^+ = \alpha_T \lambda_T + \beta_T; \quad (2.7)$$

$$\lambda_T = \ln\left(\frac{A_f}{A_t}\right) \left[1 + 0.09 \ln\left(\frac{L_x^{corr}}{k_z}\right) \left(\frac{4k_{rms}}{k_z}\right)^{-0.44} e^{-0.074S_k}\right]$$

where  $\alpha_T = 1.4699$  and  $\beta_T = 8.0394$  are empirical constants, while  $L_x^{corr}$  refers to the roughness length scale where the streamwise auto-correlation drops under 0.2.

In ref. [5], a correlation based on 3D sinusoidal roughness data in both transitionally- (eq. 2.8) and fully-rough (2.9) regimes is developed:

$$\Delta U^+ = \frac{1}{k} \log(k_a^+) + 1.12 \log(ES) + 1.47; \quad (2.8)$$

$$k_s/k_a = 7.3k_a ES^{0.45} \quad (2.9)$$

For fully-rough flows, several predictive correlations for  $k_s/k$  have been formulated using different measures for  $k$  and different choices of the parameters explained in section 2.2. Ref. [7] outlined a correlation based on [8]:

$$k_s/k_{rms} = \begin{cases} 2.48(1 + S_k)^{2.24} & S_k > 0 \\ 2.11 & S_k = 0 \\ 2.73(2 + S_k)^{-0.45} & S_k < 0 \end{cases} \quad (2.10)$$

Equation 2.10 was obtained using a wide range of roughness topographies, but it does not include any slope measure, as suggested in the bare-minimum set in section 2.2. This correlation was indeed predominantly written for surfaces with a relatively narrow range of  $ES$  in which  $k_s/k_{rms}$  is invariant [17], and the authors themselves suggested to supplement it by a density parameter for sparse roughness.

Ref. [9] understood that a correlation based on surface height skewness and effective slope can satisfactorily predict  $k_s$  normalized with the average maximum peak-to-trough roughness height. They studied surfaces with both regular and irregular arrangements and size distributions of roughness elements and proposed the following formulation:

$$k_s/k_z = (0.67S_k^2 + 0.93S_k + 1.3)[1.07(1 - e^{-3.5ES})] \quad (2.11)$$

Both equations 2.10 and 2.11 struggle with low- $ES$  surfaces, but they both could be extended to this range if long-wavelength contributions are filtered out prior to

computing statistics [4].

In general, existing roughness correlations are developed based on a limited number of data points covering a relatively narrow region of the parameter space, and their use beyond these respective regions involves a dangerous degree of extrapolation [6]. For example, data from [23] suggest that, at fixed  $S_k$ ,  $k_s/k$  increases with  $ES$  for  $ES < 0.3 - 0.6$ , beyond which it eventually decreases in the dense fully-rough regime, depending on the roughness topography. However, none of these correlations capture this behaviour. Moreover, ref. [38] conducted Direct Numerical Simulation (DNS) to study turbulent flows over irregular rough surfaces and assessed the roughness correlations presented in this section, using their database. They realised that, although none of the assessed correlations shows a dramatic loss of accuracy when used outside of the parameter space of its original fitting data, the most successful correlation can reproduce the values of  $k_s$  from DNS only within a  $\pm 30\%$  error, while none of them shows a better predictive accuracy. Lastly, there are some challenging topographies (e.g. dense, non-homogeneous, multiscale and wavy roughness) which could not behave accordingly to the framework described in section 1.1 [6], making results from the state-of-art approach to address the roughness problem completely unreliable.

# Chapter 3

## Method

An innovative representation of roughness, which does not try to tie  $k_s$  related to a rough surface to its topographical features, is here presented. Specifically, section 3.1 describes the database generation procedure, while section 3.2 proposes the method which will be applied to the computed surfaces.

### 3.1 Database generation

The database generation procedure follows the guideline proposed in Chapter 3 of [9], with some differences here discussed. Specifically, this work proposes a process which allows systematic variations of moments of surface height probability density function (PDF), surface slope, center locations and size distribution of roughness peaks. These surfaces have indeed a similar root-mean-square roughness height ( $k_{rms} \approx 0.045h$ ) and are characterized by several topographical parameters, including effective slope  $ES$ , skewness  $S_k$  and kurtosis  $K_u \equiv (1/A_t \int h'^4 dA)/k_{rms}^4$ , in a relatively wide range ( $0.3 \leq ES \leq 0.88$ ,  $-0.34 \leq S_k \leq 0.66$ ,  $1.9 \leq K_u \leq 2.61$ ). In this case,  $h$  is the half-height of the channel used for carrying out the Direct Numerical Simulations (DNS) on the rough geometries. To focus more attention on real roughness rather than simply regular roughness element arrays, in addition to surface moments and effective slope, a parameter  $\Delta$  is introduced to indicate the distribution of roughness element sizes. In the special case of  $\Delta = 0$ , all elements are identical, while positive variations of this parameters at constant values of the others allow to investigate whether and how a roughness generated by identical elements differs from a generic roughness. It is worth mentioning that these geometries are all statistically homogeneous, i.e. the statistical properties do not depend on the origin of the coordinate system.

The goal of the roughness generation procedure is to generate a surface height function  $k(x, z)$ , with  $x$  and  $z$  being the streamwise and spanwise coordinates, respectively, with prescribed statistical properties. To achieve it, axisymmetric roughness elements with imposed shape and random size are distributed randomly on a reference flat surface. The  $i$ -th element has a height of  $k_p^{(i)}$ , a base diameter  $D^{(i)}$  and is truncated on top at  $d^{(i)} = 0.1 \cdot D^{(i)}$ , accordingly to the original paper. The ratio of height to

base diameter  $\gamma$  is constant between all elements. Index  $i$  varies from 1 to  $N$ , with  $N$  being the total number of the elements.  $N$  is determined by dividing the whole reference surface area ( $L_x \cdot L_z$ , with  $L_x$  and  $L_z$  being the streamwise and spanwise geometry lengths, respectively) by the average area occupied by one element - a circle of radius  $\beta \cdot D^{(i)}$ . The profile of the  $i$ -th element is determined by the function:

$$Y = k_p^{(i)} [1 - ((r - d^{(i)}/2)/(D^{(i)}/2 - d^{(i)}/2))^n] \quad (3.1)$$

with  $Y$  and  $r$  being wall-normal and radial locations of the element wall surface, respectively. The size of  $i$ -th element  $k_p^{(i)}$  is determined using a random function with normal distribution of mean value  $k_p$  and standard deviation  $(\Delta/2.33) \cdot k_p$ , with  $\Delta$  being explicitly prescribed in the beginning of the roughness generation procedure. Exponent  $n$  determines the element shapes, as shown in figure 3.1. If initial guesses

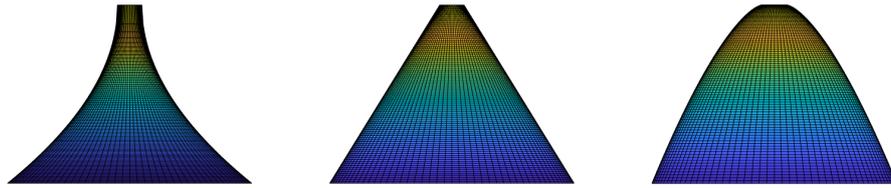


Figure 3.1. (Left) Element profile with  $n = 0.5$ ; (Middle)  $n = 1$  (conical); (Right)  $n = 1.5$ . All the elements have the same  $k_p$  and  $\gamma$ .

for the four input parameters ( $k_p, \gamma, \beta, n$ ) are made, the number of elements can be computed as  $N = \frac{L_x \cdot L_z}{\pi \cdot (\beta \cdot \gamma \cdot k_p)^2 / 4}$ . At this point, each roughness element is distributed individually on a flat surface of sizes  $(L_x, L_z) = (3h, 3h)$  using the MATLAB tool *griddata*. The  $x$  and  $z$  position of the element centers is computed using a random function with uniform distribution in the range of  $(0, 3h)$ . Specifically, function *griddata* interpolates the roughness element height into a mesh grid built on the reference plate, with resolution equal to that considered in the original paper (432 points along both streamwise and spanwise directions). Due to the randomness of the center locations, some elements may intersect or even incorporate others. For this reason, only the highest roughness size registered in each grid point is considered. This passage differs from what is done in [9], where the Immersed Boundary Method (IBM) based on [11] is used to implement roughness samples - separately generated in MATLAB - on the wall. Furthermore, each sample considered in the original work

has a size of  $(L_x, L_z) = (8H, 4H)$ . The choice of using smaller domain sizes is due to the fact that, in this case, a smaller number  $N$  of elements is required to match the prescribed statistical values, thus reducing the computational time needed to generate the surface. However,  $N$  must be sufficiently large in order to obtain a statistically homogeneous geometry.

From this computed point cloud, a rough surface, corresponding to the initial guessed parameters, is generated. For this surface, the statistical properties of interest  $(k_{rms}, S_k, K_u, ES)$  can be numerically measured. These values are also computed for the same geometries when rotated by 90 and 180 degrees, to check the homogeneity constraint. The same values, within an uncertainty interval, are obtained, thus confirming that  $L_x = L_z = 3h$  is an appropriate domain size. Similarly to the original paper, if  $S_k$  and  $K_u$  do not match the prescribed values,  $n$  and  $\beta$  are used to adjust them iteratively. Then,  $k_p$  and  $\gamma$  can be used to scale  $k_{rms}$  and  $ES$  to their imposed values. A change in  $k_p$  and  $\gamma$ , indeed, only stretches the surface in vertical and horizontal directions, respectively; hence, they do not affect other surface statistics [9]. However, if these two parameters change significantly from their initially guessed values, the computed surface might be remarkably different from the original one. This is due to some odd interpolations of *griddata* which can happen, for example, when roughness elements, separated before the iteration, intersect, producing a new element with a strange shape. In this case, *griddata* might locally register roughness heights pretty different from that in output from the first iteration. It is thus necessary to check the value of each statistical property at the end of the second iteration and keep iterating the input parameters until  $(k_{rms}, S_k, K_u, ES)$  match all the corresponding prescribed values within a given tolerance.

Following this procedure, twenty-five geometries are reproduced. Each sample is named accordingly to the *Xaabb* convention. Letter  $X$  can take values of A, B, C and D, each corresponding to one  $(S_k, K_u)$  pair; A, B and C all have  $K_u = 2.61 \pm 5\%$  and  $S_k = 0.21 \pm 5\%$ ,  $-0.34 \pm 3\%$  and  $0.66 \pm 2\%$ , respectively. D has the same  $S_k$  of A but  $K_u = 1.9 \pm 2\%$ . The two-digit number  $aa$  is equal to  $\Delta \cdot 100$  and takes either of the values 00 (uniform peak size), 15, 35 or 70. Lastly, the two-digit number  $bb$  corresponds to  $ES \cdot 100$  and ranges from  $30 \pm 1\%$  (least steep surface) to  $88 \pm 1\%$  (steepest surface). A summary of all cases and their statistical properties is presented in table 3.1. Along with the statistical values, the roughness height measurements described in section 2.2 are reported for each geometry. Specifically, the average roughness height, here denoted as  $k_p$ , is simply the mean of the normal distribution governing the element peak sizes, while the average peak-to-trough roughness amplitude  $k_z$  is computed considering nine  $(1h \cdot 1h)$  subsamples, similarly to the original paper. These values match those presented in table 1 of [9] within a tolerance of 1% and 10% for statistical properties and roughness height scales, respectively. From that table, for each rough surface the roughness function  $\Delta U^+$  is also taken, along with the prescribed  $Re_\tau \equiv \frac{u_\tau h}{\nu}$  on which the DNS is carried out to compute that value of  $\Delta U^+$ . In this case, the friction-Reynolds number  $Re_\tau \equiv \frac{u_\tau h}{\nu}$  is defined considering the half-height of the

Table 3.1. Surface parameters and main flow properties of all cases

Case	$k_{rms}/h$	$S_k$	$K_u$	$ES$	$\Delta$	$k_p/h$	$k_t/h$	$k_z/h$	$Re_\tau$	$\Delta U^+$	$k_s^+$
A7088	0.0455	0.210	2.619	0.876	0.7	0.128	0.255	0.223	498	8.92	162.748
A7060	0.0452	0.209	2.620	0.602	0.7	0.127	0.242	0.214	496	8.67	146.892
A7040	0.0448	0.217	2.620	0.404	0.7	0.127	0.226	0.204	496	8.19	120.651
A7030	0.0454	0.219	2.620	0.303	0.7	0.128	0.235	0.196	496	7.67	97.485
A3588	0.0457	0.211	2.572	0.883	0.35	0.168	0.235	0.225	499	9.22	184.049
A1588	0.0454	0.209	2.551	0.880	0.15	0.186	0.221	0.214	497	9.45	202.249
A0088	0.0445	0.214	2.580	0.887	0.00	0.210	0.210	0.204	499	9.60	215.078
A0060	0.0450	0.211	2.620	0.602	0.00	0.211	0.211	0.200	499	9.54	209.852
A0040	0.0452	0.202	2.620	0.398	0.00	0.212	0.212	0.190	498	9.15	178.842
B7088	0.0455	-0.331	2.619	0.879	0.7	0.115	0.227	0.207	501	7.60	94.727
B7060	0.0447	-0.350	2.621	0.606	0.7	0.113	0.207	0.195	499	7.38	86.557
B7040	0.0451	-0.340	2.620	0.401	0.7	0.113	0.208	0.191	499	6.89	70.803
B7030	0.0449	-0.350	2.621	0.302	0.7	0.113	0.207	0.180	498	6.37	57.208
B3588	0.0450	-0.330	2.575	0.883	0.35	0.151	0.234	0.213	499	8.00	111.609
B1588	0.0456	-0.330	2.639	0.872	0.15	0.215	0.255	0.239	497	8.21	121.644
B0088	0.0443	-0.330	2.593	0.875	0.00	0.230	0.230	0.212	498	8.35	128.831
C7088	0.0448	0.660	2.642	0.888	0.7	0.120	0.238	0.214	502	9.49	205.593
C7060	0.0455	0.682	2.619	0.600	0.7	0.120	0.221	0.197	498	9.30	190.186
C7040	0.0453	0.652	2.624	0.400	0.7	0.118	0.226	0.193	499	8.37	129.892
C7030	0.0447	0.670	2.620	0.299	0.7	0.123	0.204	0.175	497	8.28	125.186
C3588	0.0449	0.670	2.601	0.879	0.35	0.153	0.223	0.208	497	9.68	222.249
C1588	0.0445	0.670	2.591	0.881	0.15	0.169	0.205	0.195	497	9.90	243.228
C0088	0.0447	0.672	2.622	0.884	0.00	0.175	0.175	0.175	496	10.09	262.933
D7088	0.0448	0.210	1.920	0.879	0.7	0.100	0.185	0.172	498	7.99	111.152
D0088	0.0445	0.220	1.881	0.887	0.00	0.149	0.149	0.149	501	8.76	152.414

computational channel box  $h$  as the reference height. It is indeed supposed that  $\Delta U^+$  depends only on  $(k_{rms}/h, S_k, K_u, ES, \Delta)$  values on a surface if the latter is isotropic, like those presented in both this work and the original one. As such, although the samples here reproduced are geometrically different from those described in the original work, due to their matched statistical properties, it is possible to transfer to the former surfaces the  $\Delta U^+$  values computed, carrying out a DNS, on the latter ones. Lastly, the equivalent sand-grain roughness height (in viscous units)  $k_s^+$  is directly computed from  $\Delta U^+$  through the equation 2.4.

It is worth noticing that, in ref. [9], thirty-four samples are generated, while this work considers only twenty-five of them. However, the nine residual geometries include:

- Two samples (D0088s,D0088a) characterized by elements which are distributed in regular arrays with staggered and aligned arrangements, respectively. This work does not intend to analyse non-isotropic roughness.
- Four samples (A7020,B7020,C7020,A0020) with  $ES = 0.2$  and one (A0030) with a low value of both  $\Delta$  and  $ES$ . The database generation procedure highlights the need to decrease the roughness element number  $N$  to decrease either  $ES$  or  $\Delta$ , at fixed values of the other statistics. These surfaces are not reproduced in this work because an insufficient value of  $N$  would be required to match their statistical properties with the prescribed domain sizes, thus violating the statistical homogeneity constraint.
- Two samples (A3560,A3540) which do not add any significant information. To

safe computational time, since the twenty-five geometries considered in this work allow to investigate the effect of each topographical property of interest in a sufficiently large range, these two surfaces are not reproduced.

## 3.2 Description of the method

A powerful tool to analyse and compare surfaces, which has applications across multiple domains in computer graphics as introduced in Chapter 1, concerns the discretization of the Laplace-Beltrami (LB) operator [32].

Let consider a functional space on each shape of interest (e.g. the Hilbert space  $L_2$  of two square-integrable real-valued functions) such that a function defined on that surface can be written as a linear combination of the space bases. There are many possible choices of these bases, some of which can lead to a significant reduction in representation complexity. Ref. [28] realised that the two most important characteristics for choosing a basis might be compactness and stability. The former implies that the most natural functions on a shape should be well approximated by using a small number of basis elements, whereas the latter means that the space of functions spanned by all linear combinations of basis functions must be stable under small shape deformations.

A natural choice of basis in literature is the LB eigenfunctions. Specifically, let  $f$  be a  $C^2$  real-valued function defined on a manifold  $M$  with Riemann metric. The LB operator  $\Delta$  is the divergence of the gradient on manifold  $M$ :

$$\Delta f \equiv \nabla \cdot \nabla(f) \quad (3.2)$$

Since the LB operator is self-adjoint and semi-positive definite [33], it admits an orthonormal eigensystem  $(\lambda_i, \phi_i)$ , which is a basis of the space of two square-integrable functions, with  $\Delta\phi_i = -\lambda_i\phi_i$ . Furthermore, since  $\lambda_0 \leq \lambda_1 \leq \dots, \lambda_i \leq \lambda_{i+1} \leq \dots \leq +\infty$ , LB eigenfunctions  $\phi_i$  are ordered by eigenvalues from "low frequency" to "higher frequency", providing a natural multi-scale way to approximate functions between surfaces. The most common discretization of the LB operator is the standard cotangent-weight scheme [26, 29]. This method computes the first  $n$  LB eigenfunctions on a shape, discretized with a 3D triangular mesh, by solving the linear system:

$$L = A^{-1}W \quad (3.3)$$

where  $A$  is a diagonal matrix of lumped area weights (mass) and  $W$  is a sparse matrix of cotangent weights (stiffness). In this case,  $L$  is a matrix of dimension  $(n_{grid\ points}, n)$ , where  $L(i, j)$  contains the pointwise value of the  $j$ -th LB eigenfunction in the grid-point  $(x_i, z_i, k(x_i, z_i))$ . The matrix equation 3.3 can be efficiently solved with a sparse matrix eigensolver implemented in MATLAB. Its code is reported and briefly discussed in Appendix A. Typically,  $n$  is fairly small [15], giving rise to compact functional representation of shapes.

Figure 3.2 proposes the plot of the first seven LB eigenfunctions computed on one

database roughness, named A7088, and a flat plate with identical domain sizes and grid resolution. These functions are defined pointwise on a mesh grid. As such, if a surface is discretized with a grid of  $n_{grid-points}$  points, then a LB eigenfunction  $\phi$  is a  $(1, n_{grid-points})$  vector, where  $\phi_i$  takes the pointwise value of that function in the grid-point  $(x_i, z_i)$ . It is thus possible to color all grid-points of a surface mesh with their corresponding pointwise values of one LB eigenfunction. The first eigenfunction is always constant, i.e. all grid-points are identically colored. A critical point in the use of the Laplace-Beltrami operator is that individual eigenfunctions are known to be unstable under perturbations, suffering from well-known phenomena such as sign flipping and eigenfunction order changes [28]. For example, it is clear that eigenfunctions of index "2" and "3" in figure 3.2 are inverted in the two geometries, while the fifth eigenfunctions have flipped sign. However, the space of functions spanned by the first  $n$  eigenfunctions of the LB operator is shown to be stable under near-isometries as long as the  $n^{th}$  and the  $(n + 1)^{th}$  eigenvalues are well separated [19]. In this case, it is thus possible to transfer functions from a surface (source) to another surface (target), if written as a linear combination of the LB eigenfunctions computed on the corresponding surface.

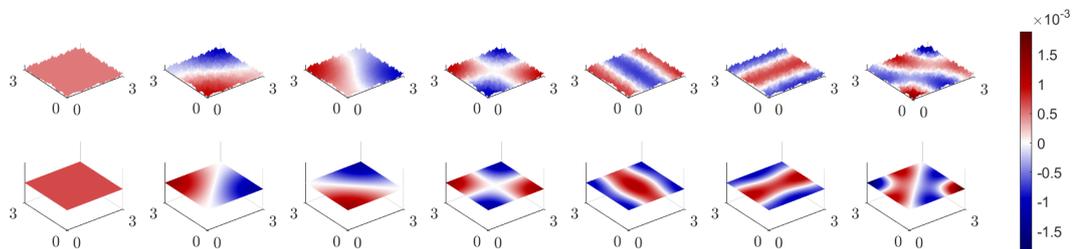


Figure 3.2. From left to right: first seven Laplace-Beltrami eigenfunctions computed on *first row*: A7088; *second row*: smooth wall with the same domain dimensions and grid resolution of the rough surface. Colored scale indicates pointwise values of the LB basis.

Theoretically, all real-valued functions  $f$  definable on a surface can be exactly recovered if an infinite number of the surface space bases is considered:  $f = \sum_{i=1}^{\infty} \alpha_i \phi_i$ . In practice, the series must be truncated at the  $n$ -th basis:  $f \approx \sum_{i=1}^n \alpha_i \phi_i$ . Coefficients  $\{\alpha_i\}$  can be simply computed through the Ordinary Least Squares (OLS) approach, which produces an estimation  $\{\hat{\alpha}_i\}$  in such a way that the sum of squares of residuals

is as small as possible, i.e. minimizes the following loss function:

$$L_{OLS}(\hat{\alpha}_i) = |f - \sum_{i=1}^n \hat{\alpha}_i \phi_i|^2 \quad (3.4)$$

The OLS method does not bother to use as few predictors (LB eigenfunctions) as possible, since it only aims to minimize  $L_{OLS}$ . However, there are many applications in which a narrow number of predictors might be dominant, i.e. their coefficients have a much higher absolute value, over the others. Moreover, many predictor variables might be highly correlated with each other. In this scenario, it is well known that the OLS estimator, despite being unbiased, has a huge variance [39]. To address this issue, the general solution is to reduce variance at the cost of introducing some bias. This approach is called regularization. In the case of interest, i.e. linear regression, it translates in reducing the model complexity by reducing the number of predictors. There are three main linear regression approaches:

- **Ridge regression**

The Ridge Regression [25] loss function is obtained by adding a term which penalizes the size of parameter estimates in equation 3.4, in order to shrink them towards zero:

$$L_{Ridge}(\hat{\alpha}_i) = |f - \sum_{i=1}^n \hat{\alpha}_i \phi_i|^2 + \lambda \sum_{j=1}^m \hat{\alpha}_j^2 \quad (3.5)$$

where  $\lambda$  is the regularization penalty. This method does not enforce coefficients to be exactly zero, yet penalizes them if they are too far from zero, compelling them to be small. As such, model complexity is decreased while keeping all variables in the model.

- **LASSO regression**

The Least Absolute Shrinkage and Selection Operator (LASSO) [30] adds a penalty on the absolute value of the coefficients. Consequently, for high values of  $\lambda$ , many coefficients are exactly zero, which is never the case of ridge regression. LASSO expression yields:

$$L_{LASSO}(\hat{\alpha}_i) = |f - \sum_{i=1}^n \hat{\alpha}_i \phi_i|^2 + \lambda \sum_{j=1}^m |\hat{\alpha}_j| \quad (3.6)$$

- **Elastic Net**

Elastic Net [39] combines the penalties of Ridge and LASSO regressions to get the best of both. As a matter of fact, Ridge regression works well when most of predictors impact the response, while LASSO is more suitable when there is a small number of significant parameters and the others are close to zero. Since in many applications the true value of these parameters is unknown, it is often

suggested to combine the previous two approaches. The Elastic Net formulation is:

$$L_{ENet}(\hat{\alpha}_i) = \frac{|f - \sum_{i=1}^n \hat{\alpha}_i \phi_i|^2}{2n} + \lambda \left( \frac{1-\gamma}{2} \sum_{j=1}^m \hat{\alpha}_j^2 + \frac{\gamma}{2} \sum_{j=1}^m |\hat{\alpha}_j| \right) \quad (3.7)$$

where  $\gamma$  is the mixing parameter between Ridge ( $\gamma = 0$ ) and LASSO ( $\gamma = 1$ ). In this case, both  $(\lambda, \gamma)$  values must be tuned.

The most critical point in each regression method is the choice of the value of  $\lambda$ . There are two commonly adopted ways to solve this problem. A more traditional method chooses  $\lambda$  such that some information criterion, e.g. Akaike (AIC) or Bayesian (BIC), is the smallest [36]. On the other hand, a original, machine-learning approach performs cross-validation and select the value of  $\lambda$  that minimizes the cross-validated sum of squared-residuals [10]. Section B.2 in Appendix B reports the algorithm to find this value. However, these are not the only ways to set this parameter and, generally, the most suitable choice is really problem-dependent.

# Chapter 4

## Experiment and Results

The application of the method proposed in this work to the computed geometric database is explained in section 4.1, while its results are presented and discussed in section 4.2.

### 4.1 Application of the method

The roughness representation here presented is based on the simple idea that each rough surface can be thought as a specific deformation of a smooth wall. Accordingly, it is not necessary to identify and compute the key topographical properties of the different roughness geometries, since all of them are directly transferred to the reference flat plate. To achieve it, the guideline discussed in section 3.2 is followed. Specifically, the surface elevation of the rough geometries is written as a linear combination of the Laplace-Beltrami (LB) eigenfunctions computed on the smooth wall. Coefficients of those linear systems are then computed with the method LASSO. Each step to be carried out to outline the desired statistical representation of roughness is now described.

Firstly, the roughness database generated in section 3.1 must be enlarged. Otherwise, it cannot be considered statistically significant. As such, ten additional copies of these surfaces are reproduced: five of them are obtained by shuffling element center positions of the corresponding original geometry; the other five are computed with a shuffle on its element height sizes. These shuffles do not change surface statistical properties since the positioning of the elements, as well as their size distribution, do not affect the PDF of  $k$  and its derivatives [9]. However, due to the odd behaviours of the interpolating function *griddata* explained in section 3.1, some little adjustments on the input parameters of the database generation procedure might be needed to adequately match the prescribed statistics. In the particular case of uniform peak sizes ( $\Delta = 0$ ), the shuffle on height distribution becomes meaningless since the surface elements are all identical. The ten copies of these geometries are thus computed by shuffling their element center positions.

Twenty-five shape collections are obtained, each of them containing one geometry

with its ten statistical copies. To fairly compare the different rough surfaces, a flat plate with the same domain sizes and grid resolution is generated. Rough geometries are indeed characterized by a vector  $k$  of dimension  $(1, n_{grid-points})$ , representing their surface elevation - relative to the mid plane - in each grid point. Due to identical lengths and grid resolution, vectors  $k$  can be thought as the discretization of a real-valued function defined on the reference smooth wall, which describes the grid points deformation. Therefore, if the Laplace-Beltrami operator is discretized on the flat plate, the vector  $k$  of each rough surface can be written as a linear combination of the first  $n$  LB eigenfunctions, after truncating the series. Nevertheless, the first LB eigenfunction on any surface is always constant and thus not included in this procedure. One can write:

$$k_{roughness} \approx \sum_{i=2}^n \alpha_i \phi_{i_{plane}} \quad (4.1)$$

The rough surfaces and the smooth wall are scaled in viscous units prior to solving equation 4.1. This step is needed to account the (slightly) different values of  $Re_\tau \equiv \frac{u_\tau h}{\nu}$  on which the DNS, in the work [9], is carried out to compute  $\Delta U^+$  on each rough surface. In this case, to define the friction-Reynolds number  $Re_\tau$ , the "effective channel half-height"  $h$ , i.e. the half-height of a smooth wall with the same cross-sectional area of the computational box used to execute DNS, is chosen as the reference length scale. The domain sizes  $(L_x, L_z, k)$  are thus multiplied by the prescribed  $Re_\tau$  reported in table 3.1. The nominal value of  $Re_\tau = 500$  is considered for the flat plate. Unfortunately, this passage leads to a (small) numerical error. All the rough surfaces and the reference plane have indeed identical physical domain sizes, thus different lengths in viscous scaling if  $Re_\tau \neq 500$  is considered. Since the LB eigenfunctions are defined pointwise on the surface grid and differ if the surface dimensions change, the reconstructed  $k$  is slightly different from the original one. However, each geometry has  $Re_\tau \approx 500$  (ranging from 496 to 502), thus this error is expected to be negligible. For completeness, section B.1 in Appendix B proposes a brief discussion of this topic. Coefficients  $\alpha_i$  must be computed to solve equation 4.1. Firstly, the number  $n$  of LB eigenfunctions beyond which the series is truncated must be set carefully. Dense roughness topographies with very different peak sizes ( $ES = 0.6-0.88$ ,  $\Delta = 0.7$ ) might have very small peaks represented more efficiently by high frequency eigenfunctions, suggesting that  $n$  should not be too small. On the other hand, bases at very high frequency could only be noisy if compared to the others. For this reason, it is decided to set  $n = 500$ . In section B.2.1 of Appendix B, the same study is conducted considering  $n = 300$  and  $n = 1000$ . Subsequently, a method from those presented in section 3.2 must be chosen to compute  $\alpha_i$ . There is a large number of predictors (LB eigenfunctions) and thus, despite being uncorrelated by definition (space bases), the OLS method is not a good choice. The LASSO approach is instead used. LASSO is more suitable than Ridge regression since the former works better when considering a problem with a lot of predictors, yet not all of them are expected to be important (the case of interest), while Ridge should be preferred when there are few predictors, but

all of them are considered relevant to predictions. Moreover, since the features are known, there is no need to imply a Elastic Net approach, that would also require the setting of the additional parameter  $\gamma$ . In this work, it is used  $\lambda = 1$  as regularization penalty value. As such, the two terms in equation 3.6, describing the least-squared minimization and the model dimension reduction, respectively, have the same influence. Appendix B.2.2 proposes the same study (with fixed  $n$ ), when considering  $\lambda = 0.1$ , which is very close to the solution of the machine-learning approach introduced in section 3.2, and  $\lambda = 2$ , to understand how this parameter setting influences the final results.

At this point, the equation 4.1 is solved, for all the eleven geometries of each shape collection, using the MATLAB function *lasso*. Although the rough surfaces belonging to the same collection are statistically equivalent, their LASSO coefficients will be slightly different in both the highlighted predictors (coefficients indices) and their weights (coefficients values). This is due to the fact that these surfaces are not geometrically identical and thus will excite the LB eigenfunctions, which are defined in the space domain, in a different manner. However, it is possible to outline a statistical representation of a specific roughness, starting from its LASSO coefficients. Specifically, each shape collection is now described as a matrix of dimensions  $(11 \cdot n)$ , where the  $i$ -th row contains the  $n$  LASSO coefficients of the  $i$ -th surface of that collection. For these surfaces, the four roughness height measures described in section 3.1 are then computed (in viscous scaling). Their values are not reported for simplicity, but they all match those of the corresponding original geometry within a tolerance of  $\pm 5\%$ . These scales are used to normalize the equivalent sand-grain height  $k_s^+$ . Accordingly to what is stated in section 3.2, a specific rough surface and its statistical copies can indeed be labelled with the same value of  $\Delta U^+$ , i.e. the same  $k_s^+$ . Four additional vectors  $(k_s/k_{rms}, k_s/k_p, k_s/k_t, k_s/k_z)$  of dimensions  $(11 \cdot 1)$ , each of them containing a given normalization of  $k_s^+$  related to the eleven surfaces of that collection, are thus obtained. The desired representation of a specific roughness can be outlined computing the linear correlation between one of these four vectors and the  $(11 \cdot n)$  matrix of the LASSO coefficients related to the shape collection describing that rough surface. This step can be efficiently done by means of the MATLAB function *corr*. It is worth noticing that this method tie  $k_s$  related to a rough surface to the Laplace-Beltrami eigenfunctions computed on a flat plate, rather than the topographical properties of that roughness, as done by the state-of-art approach. As such, this roughness representation is independent of the surface topographical features, allowing to compare different surfaces in a simple and efficient way.

Lastly, the twenty-five different rough surfaces considered in this work must be compared. Since each of them is now represented by a  $(1 \cdot n)$  vector of correlation coefficients, it is simply needed to build a  $(n_{shapes} \cdot n)$  matrix, where each row contains the correlation coefficients related to a specific roughness. To assess the possibility of developing a universal roughness correlation for  $\Delta U^+$ , which is the main target of the start-of-art method to address the roughness problem, the matrix rows are sorted into

increasing  $\Delta U^+$  order, i.e. the first row corresponds to the surface with lowest value of  $\Delta U^+$ , while the last describes the roughness with the highest one. In this way, it is possible to understand if surfaces with similar roughness function values (adjacent matrix rows) are characterized by similar pattern of the correlation coefficients.

## 4.2 Results

The results of the experiment described in the previous section are here presented and discussed.

Figure 4.1 shows selected rough surfaces computed in the database generation procedure, described in section 3.1, to illustrate how variations of roughness parameters affect the surface geometry. Figures 4.1b, 4.1c report two geometries with statistical properties identical to those of the reference roughness (A0778), yet different skewness value. These surfaces are characterized by roughness elements with very different shapes, thus different outcomes on the flow, i.e.  $\Delta U^+$  values, are expected. Ref. [16] confirmed this theory, noting that peaked surfaces ( $S_k > 0$ ) produce much more drag than pitted surfaces ( $S_k < 0$ ). In general, many works in literature, e.g. [7, 20], showed  $k_s/k_{rms}$  increasing with  $S_k$ . Regarding the parameter  $\Delta$  (figures 4.1d, 4.1e), Ref. [9] realised that, at constant values of surface moments and ES, a rough surface with more uniform distribution of peak sizes (smaller  $\Delta$ ) causes a higher skin friction. They gave a qualitative explanation of this concept stating that, with a decrease in  $\Delta$ , roughness geometry needs to reshape in a such way that its elements are sharper on top and flatter near the root, bringing to a similar behaviour than those of increasing  $S_k$ . Lastly, it is clear how lower-ES surfaces (figures 4.1f, 4.1g) are generated with a smaller number of roughness elements. These geometries have a smaller value of  $\Delta U^+$ , due to a more gradual transition from hydraulically-smooth to fully-rough regime [7]. Figure 4.2 reports the coefficients, computed with the LASSO method, of the linear system 4.1, relating the surface elevation vector  $k$  of the rough surface A7088 (figure 4.2a) and its ten statistical copies (figure 4.2b) to the first 500 LB eigenfunctions computed on the reference flat plate, considering  $\lambda = 1$ , as described in section 4.1. As highlighted in figure 4.2a, LASSO cancels the effect of many eigenfunctions but retains some of them at very high frequency, suggesting that truncating the series at much lower values of  $n$  could neglect the presence of some important features. On the other hand, exploring very high frequencies could drastically increase the model complexity while adding very few information, or even only noise. A general decrease of the absolute value of the LASSO coefficient with increasing frequency (eigenfunction index) is indeed outlined. Figure 4.2b underlines that both center positioning (first row) and size distribution (second row) influence the importance of predictors in reconstructing the roughness elevation  $k$ , since these LASSO coefficients (slightly) differ in both their absolute values and the highlighted functions. This is due to the (small) geometrical differences between surfaces with matched statistical properties. However, these coefficients seem to follow a general trend and thus, if a

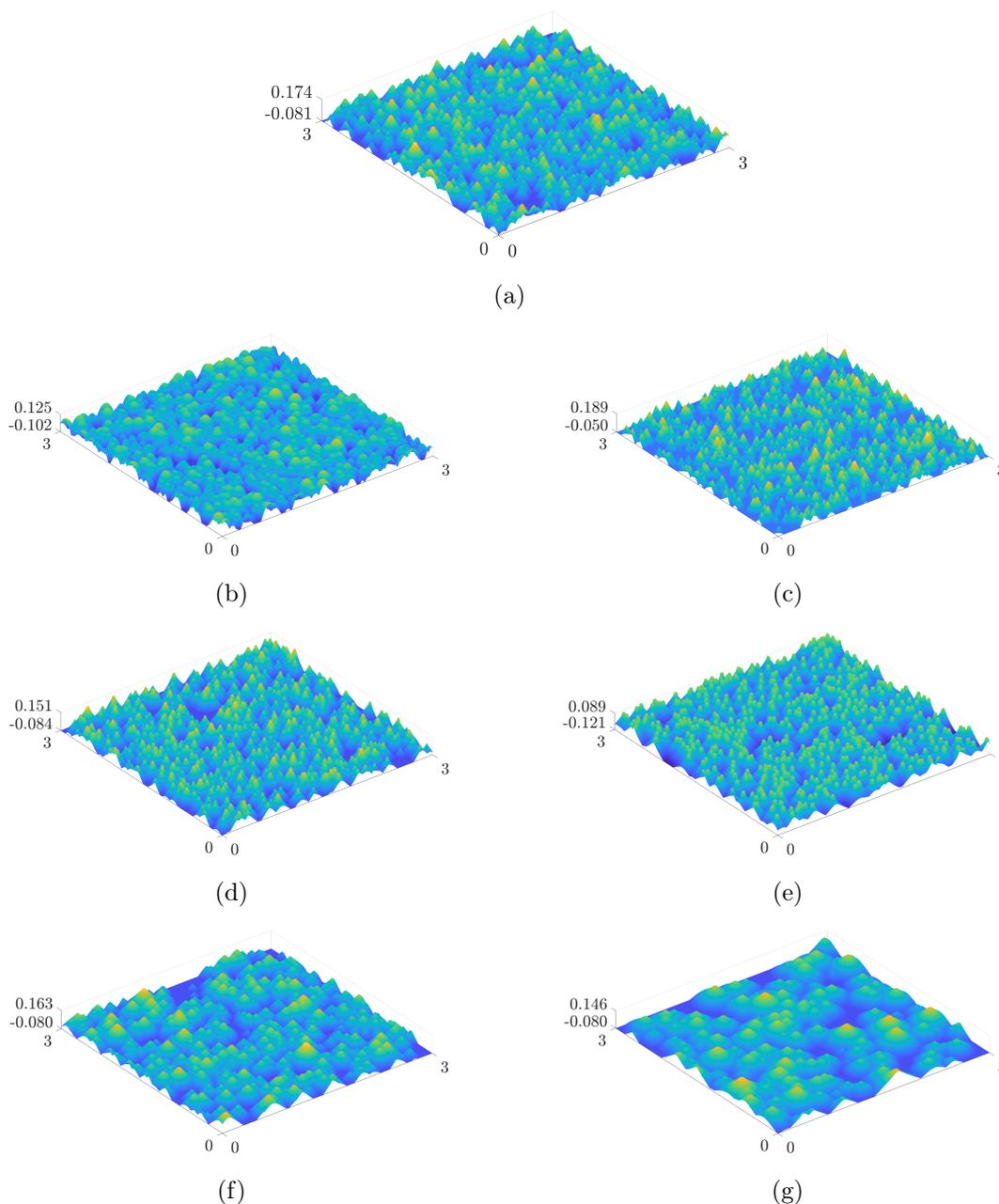
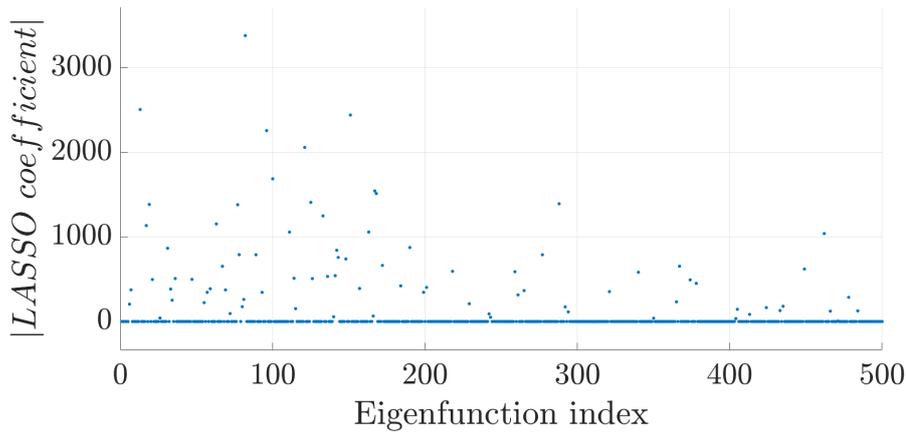
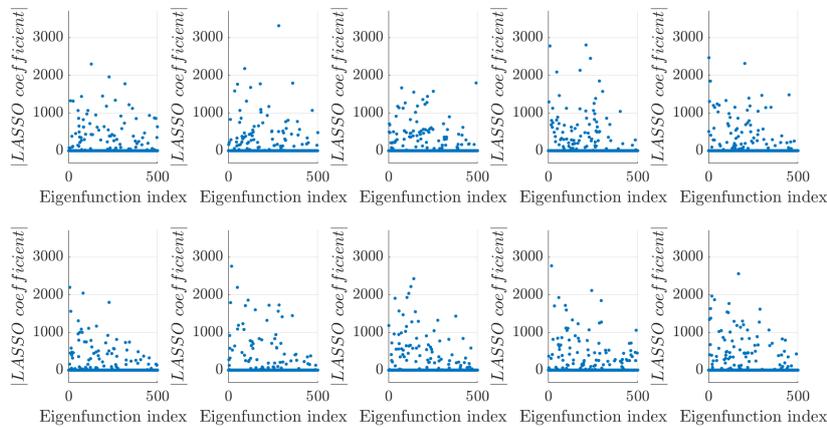


Figure 4.1. Selected geometry samples. All dimensions are normalized with effective channel half-height  $h$ . Colored scale indicates surface elevation. *First row*: (a) A7088 (reference geometry); *second row*: (b) B7088, (c) C7088 - geometries with the same  $ES$  and  $\Delta$  as the reference one but different  $S_k$ ; *third row*: (d) A3588, (e) A0088 - geometries with the same  $S_k$  and  $ES$  as the reference one but different  $\Delta$ ; *fourth row*: (f) A7060, (g) A7040 - geometries with the same  $S_k$  and  $\Delta$  as the reference one but different  $ES$ . Each sample has the same  $k_{rms} = 0.045h \pm 2\%$  and  $K_u = 2.61 \pm 5\%$ .

sufficiently large dataset is taken for each shape collection, it is possible to outline a statistically significant description of the roughness represented by that collection.



(a)



(b)

Figure 4.2. Absolute value of LASSO coefficients between surface elevation vectors of A7088 geometries and the first 500 LB eigenfunctions on the reference plane, computed with  $\lambda = 1$ . (a) Original geometry; (b) *first row*: geometries obtained by shuffling the reference A7088 center locations; *second row*: geometries obtained by shuffling the reference A7088 peak sizes.

Figure 4.3 proposes a comparison between the different geometries analysed in this work. Accordingly to the Xaabb convention described in section 3.1, each row corresponds to a specific letter X, i.e. a  $(S_k, K_u)$  pair, while each column represent a different combination of the aabb numbers, i.e. a  $(\Delta, ES)$  pair. Surfaces with the same statistical moments yet different effective slope  $ES$  show that, for decreasing values of  $ES$ , LASSO retains few eigenfunctions and at lower frequencies, yet their influence (i.e. the absolute value of their LASSO coefficient) is more relevant in the reconstruction of  $k$  than that of the same predictors in the case of higher  $ES$  values. A possible explanation of this result is that, for decreasing values of  $ES$ , the surface becomes more wavy, i.e. the ratio of height to base diameter of the roughness elements increase. Therefore, there is no need to exploit eigenfunctions at very high

spatial frequency to catch these "large" peaks. Furthermore, the generation procedure highlights that, to decrease  $ES$  while keeping the surface dimensions fixed, the total number of elements must decrease. Less eigenfunctions are needed to represent less elements. Conversely, the other statistical parameters do not exert an influence on LASSO coefficients as clear as the effective slope and thus a further investigation is required.

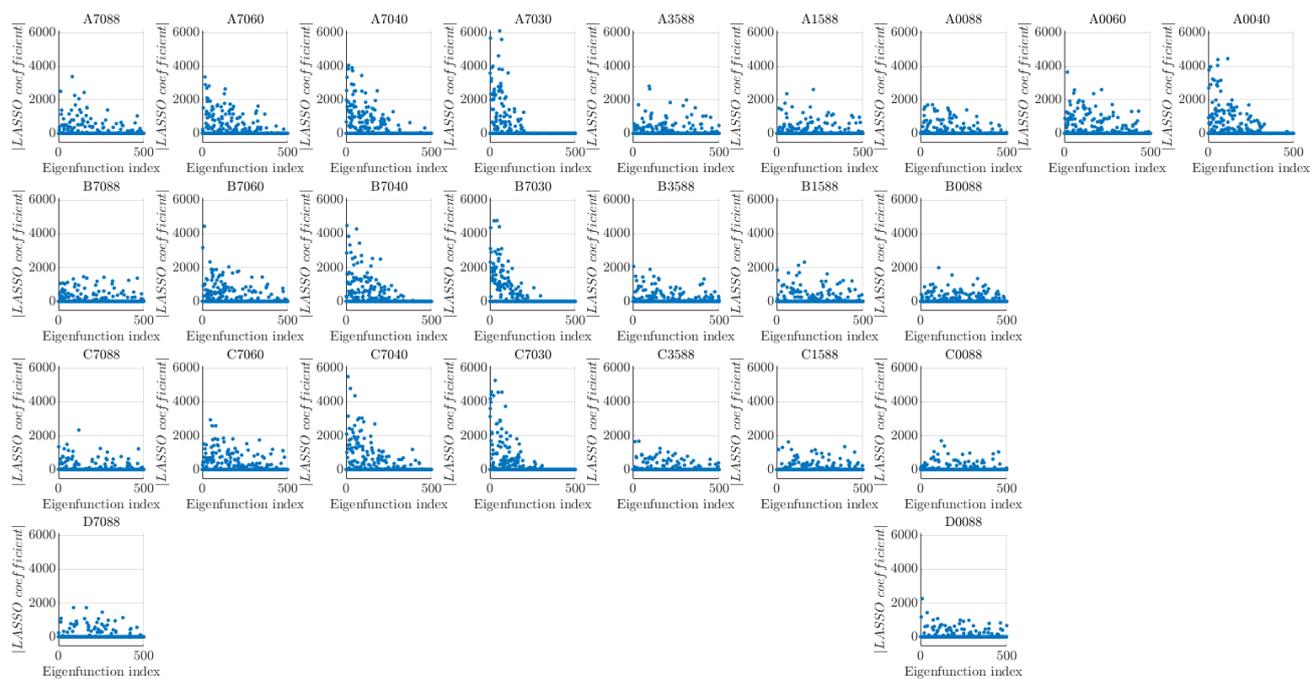


Figure 4.3. Absolute value of LASSO coefficients between surface elevation vectors of the twenty-five database geometries and the first 500 LB eigenfunctions on the reference plane, computed with  $\lambda = 1$ .

Figure 4.4 shows, for the shape collection representing roughness A7088, the absolute value of the linear correlation coefficients between LASSO coefficients and the four vectors regarding the different normalizations of  $k_s$ , described in section 4.1, computed for that collection. In this case,  $k_{rms}$  seems to be the most suitable choice to scale  $k_s$ . Indeed, as depicted in the first plot of figure 4.4, many predictors have a very low correlation coefficient ( $\leq 0.4$ ) in absolute value, yet few of them have a very high one ( $\approx 0.9$ ). This is the best possible scenario when developing a model complexity reduction method, since only few features are supposed to be influential to the desired output. This result can be explained by the nature of this particular geometry, which has peaks with very different sizes, thus a statistical scale, which can "average" these heights, is an appropriate choice. However, this method is applied to a large dataset including rough surfaces with very different geometrical features, as highlighted in

figure 4.1. As such, it is not guaranteed that the correlation based on a particular roughness height scale is the best choice for all of those topographies.

The comparison between correlation coefficients of all the different geometries is needed. Figure 4.5 reports, for each of these geometries, the same results of those presented in figure 4.4 for the A7088. Specifically, each matrix corresponds to a specific normalization of  $k_s$ , while each matrix row represents the specific rough surface labelling that row on the left. Matrix rows are sorted into increasing  $\Delta U^+$  order, as discussed in section 4.1. Some important features that might be useful to efficiently cluster different roughness topographies are outlined. A good choice of the roughness height measurement for a shape collection is such that the correlation based on this scale retains few predictors with high absolute value of their coefficients, yet the others are either negligible or null. This constraint translates to few matrix elements in figure 4.5 with colored scale approaching to  $\pm 1$  (red or blue), while the others are faded or white. Accordingly, the four figures highlight that geometries with uniform peak sizes ( $\Delta = 0$ ) are well represented by geometrical scales  $k_p$ ,  $k_t$  (the same, in this case) and  $k_z$  (slightly different due to the subsampling). This is due to the fact that these particular surfaces are generated with roughness elements that are all identical, thus a geometrical scale directly representing the height of these elements is an appropriate choice. Moving towards multiscale roughness, i.e. increasing  $\Delta$ ,  $k_{rms}$  becomes the most suitable choice. These geometries are indeed characterized by elements with several sizes and some of them might be small enough to be hydrodynamically irrelevant. These small scales can contaminate the peak height statistics, thus a statistical scale that can "average" the different heights, filtering extreme asperities, is appropriate. A similar statement can be claimed for low- and high-ES surfaces. As explained earlier, the total number of roughness elements must be decreased to achieve a smaller  $ES$  at fixed values of the other statistical properties. As such, low-ES surfaces are characterized by few roughness elements and, consequently, can be well represented by a geometrical scale. On the other hand, high-ES surfaces are built using a large number of elements, thus it is preferable to use a statistical scale, e.g.  $k_{rms}$ . These considerations demonstrate the efficacy of this method in representing roughness in a reliable way. On the other hand, it is clear how none of these correlations, based on the roughness height measures commonly adopted in the state-of-art approaches, are such that surfaces with similar values of the roughness function have also similar correlation coefficients. In particular, all of these matrices highlight very different coefficients patterns between low- and high-ES surfaces. The entire spectrum of low-ES surfaces is included into the first 300-350 (depending on which correlation is considered) LB eigenfunctions computed on the reference smooth wall, since almost all the matrix entries beyond that indices are null. In the case of dense roughness topographies, many predictors at very high frequencies are instead characterized by very high correlation coefficients, suggesting that if a larger value of  $n$  was considered, some important features might have been underlined by this method. Actually, figure B.3 proposes the same work with  $n = 1000$  and

realises that only few predictors at index larger than 500 have a relevant correlation coefficient, thus it is not worthy to double-size the model to catch them.

Accordingly to the results presented in figure 4.5, a universal correlation tying  $\Delta U^+$  to any roughness topography is not feasible, due to the different coefficient patterns between surfaces with similar values of  $\Delta U^+$ . However, the roughness representation presented in this work allows to understand the limit of that approach. Figure 4.6 aligns three geometries (B7088, A7030, D7088) with very different topographical properties, yet similar values of  $\Delta U^+$ . The flows occurring next to these surfaces are characterized by a different intrinsic physics, as discussed at the beginning of this section, due to variations of the statistical properties. These effects, if combined to each other, can somehow bring to similar values of the roughness function, but the method presented in this work, which proposes a statistical description of roughness based only on its surface elevation, cannot catch these flow patterns. The key point of this analysis is that even all the existing roughness correlations, which are based only on the topographical properties of the rough surfaces described in section 2.2, unavoidably fall into the same issue for the same reason.

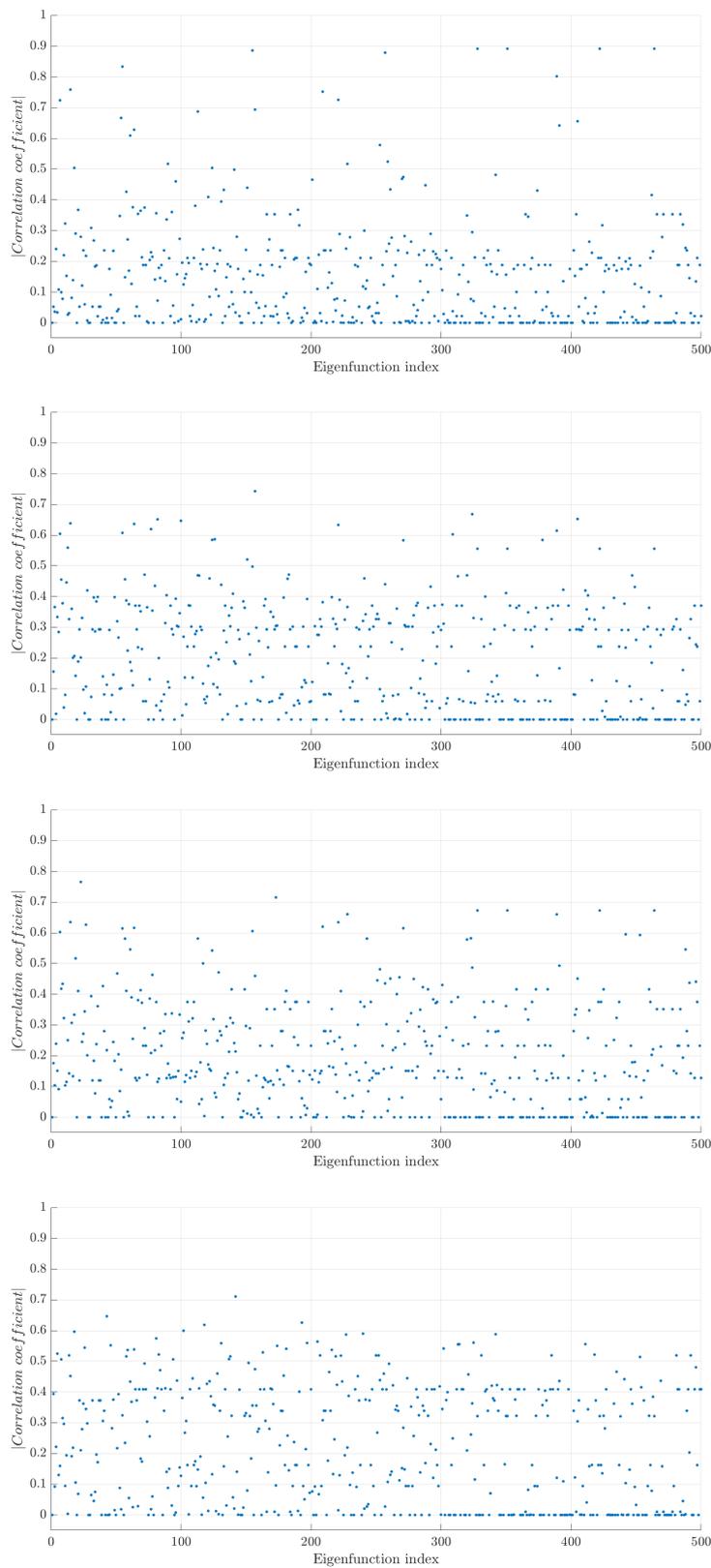


Figure 4.4. Absolute value of correlation coefficients between LASSO coefficients of the A7088 collection and their (*first row*)  $k_s/k_{rms}$ ; (*second row*)  $k_s/k_p$ ; (*third row*)  $k_s/k_t$ ; (*fourth row*)  $k_s/k_z$  vector.

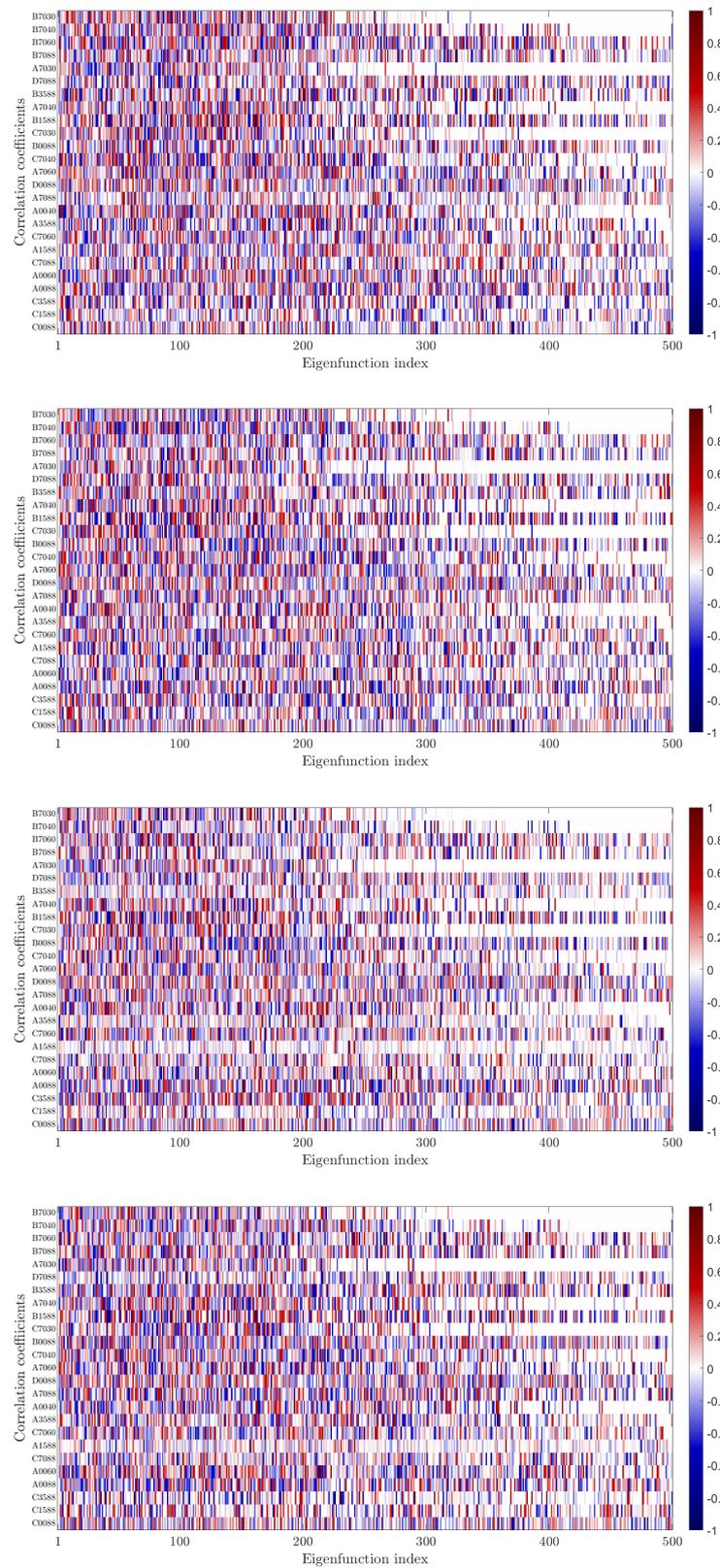


Figure 4.5. Correlation matrix between LASSO coefficients of all the different geometries and their (*first row*)  $k_s/k_{rms}$ ; (*second row*)  $k_s/k_p$ ; (*third row*)  $k_s/k_t$ ; (*fourth row*)  $k_s/k_z$  vectors. Colored scale indicates correlation coefficient values.

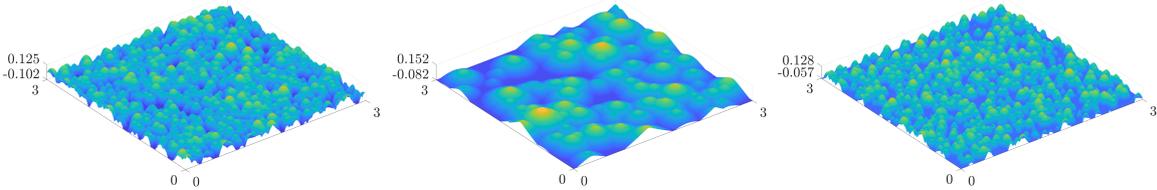


Figure 4.6. Rough surfaces characterized by (left)  $\Delta U^+ = 7.60$ ; (middle)  $\Delta U^+ = 7.67$ ; (right)  $\Delta U^+ = 7.99$ . All dimensions are normalized with effective channel half-height  $h$ . Colored scale indicates surface elevation.

# Conclusions

This paper proposes an innovative representation of roughness, which offers many useful information to efficiently cluster the different morphologies. Specifically, the surface elevation of each rough geometry is written as a linear combination of the Laplace-Beltrami (LB) eigenfunctions computed on a reference smooth wall. Coefficients of these linear systems are then computed with the method LASSO. Accordingly to the results obtained in this work, a relatively large number of these functions, with some of them at very high frequency, is needed to properly describe the spectrum of high-ES surfaces (dense roughness), while low-ES surfaces (long-wavelength or "wavy" roughness) require a smaller number of predictors, without the need to exploit those at the highest frequencies. A high number of surface peaks is indeed registered in the former case, while only few roughness elements, with high ratios of height to base diameter, characterize the latter surfaces. Furthermore, the equivalent sand-grain roughness height  $k_s$  related to a multiscale roughness is suitably normalized with a statistical scale, e.g. the root-mean-square roughness height  $k_{rms}$ , while a geometrical measure, e.g. the maximum peak-to-through elevation  $k_t$ , is a more appropriate choice for surfaces with uniform peak sizes. In these two scenarios, only the LASSO coefficients of few predictors have indeed a high correlation coefficient ( $\approx 0.9$ ) with the given normalization of  $k_s$ , while the others are characterized by a value which is small ( $\leq 0.3$ ) or null. This is due to the presence of surface peaks with either very different heights, on which a statistical scale can filter extreme asperities (e.g. sufficiently small scales to be considered hydrodynamically irrelevant), or identical sizes, on which a geometrical scale can directly describe these measures.

There is the need to enlarge the dataset on which this method is applied. Despite allowing systematic variations of the main surface properties in a relatively wide range, these geometries, which are reproduced following the procedure presented in [9], are not obtained by realistic surface measurements. The geometric scales of some of the rough surfaces generated in this work, reported in table 3.1, reach a quarter of channel height. As such, this paper actually deals with "roughness corrugations", with expected turbulent motions different from those over superficial roughness. Moreover, although some important features of naturally formed roughness, e.g. randomness in the size and positioning of roughness peaks as well as many surface parameters, are included, these geometries still contain many simplifications compared to real rough surfaces. Due to high computational costs, a moderate friction-Reynolds number ( $Re_\tau \approx 500$ ), where the separation of scales is not perfectly reached, was also chosen.

The authors of the original paper suggested to use higher values in the next works, to ensure that the fully-rough regime hypothesis, necessary to tie  $k_s$  to  $\Delta U^+$ , is verified. Lastly, this database, composed of eleven statistical copies of twenty-five different rough surfaces, may be too narrow to be considered statistically significant. These issues might have brought a non-negligible numerical error on the results presented in this work, thus the need to consider a larger database, possibly taken from realistic measurements. As such, a recently developed roughness database, which contains surface profiles and statistics, as well as experimental measurements and simulation results for flows over a wide range of roughness topographies [6], can be downloaded free on the web site <http://roughnessdatabase.org/>.

There are many future developments which naturally arise from this work. Firstly, the surface representation here presented could be further improved. The weak point of this formulation is indeed that each shape collection, composed of a specific geometry with its ten statistical copies, is completely independent of the others. There are many techniques along different computer graphics applications which exploit the advantages in representing surfaces belonging to different collections. For example, ref. [15] proposes to extract a limit shape in a collection. Although it is possible to endow this "average shape" with a natural geometric structure, the limit shape is directly computed in the frequency domain constructing a Canonical Consistent Latent Basis (CCLB), as described in Algorithm 1 of [15]. Therefore, the surface elevation vectors could be transferred to this limit shape rather than the reference plane, and write  $k$  as a linear combination of the CCLB. This approach has many advantages in terms of informativeness and computational efficiency and might become a powerful tool to study and compare rough surfaces. On the other hand, this method is way less intuitive and more complex than that proposed in this work, since it requires to familiarize with the concept of functional map network [28]. Lastly, the method proposed in this paper is well-suited to many machine learning techniques. In particular, the limits of the state-of-art approach to address the roughness problem are well known. As a matter of fact, a universal roughness correlation tying either  $\Delta U^+$  or  $k_s$  related to a rough surface to its topographical properties, which is the main goal of that method, has not been developed yet, and it is probably not feasible. Existing roughness correlations are computed only on a limited number of data points (surfaces) covering a relatively narrow region of the parameter space [6], in which the flows over these rough surfaces are supposed to have a similar physics. Specifically, there are many flows characteristics, e.g. flow patterns around roughness protuberances, flow separation locations, shear layers associated with the separation bubbles, that cannot be catch neither by the existing correlations nor by this method. The results of this experiment show indeed how very different roughness topographies, which surely cause different outcomes on the flows over them, might have similar values of  $\Delta U^+$ . However, there are many very recent papers, e.g. [18, 21], claiming that a ML network trained to correlate these flow characteristics (as outputs) to the roughness geometry (as inputs) might be an efficient tool for determining the

sets of roughness geometrical features which are important for characterizing these effects, thus reaching an unprecedented level of comprehension of roughness [38]. The roughness representation here presented might be an efficient input of this network. As a matter of fact, a physically reliable description of roughness is outlined through a model which dimension can be set to adequately meet requirements of model accuracy and model complexity. It is indeed possible to set a correlation coefficient value below which the effect of the corresponding predictor, and thus its presence, can be neglected. The lower is this value, the smaller is the considered number of predictors and the higher is the lost information. This is a huge advantage when building inputs of a ML network. Furthermore, the method described in this work allows to analyse and compare rough surfaces independently from their topographical features, since it only requires to know the surface elevation  $k$  of each rough surface and to compute the LB eigenfunctions on the reference smooth wall. As such, there is no need to address the critical point in the state-of-art framework, i.e. identify and measure the infamous bare-minimum set of invaluable topographical properties (e.g.  $\lambda_f, \lambda_p, S_k, \dots$ ) required to estimate the drag on a surface.



# Appendix A

## MATLAB Code to compute LB eigenfunctions

The MATLAB algorithm to discretize the Laplace-Beltrami operator on a shape by means of the Cotangent-Weight Scheme is here reported. The code follows the procedure described in ref. [29].

Inputs of the `compute_LB.m` function are:

- $S$ : MATLAB structure containing two fields:  $S.VERT$ , a  $[n_{grid-points}, 3]$  matrix which columns contain the streamwise, wall-normal and spanwise (from left to right) coordinates of each mesh grid-point;  $S.TRIV$ , a  $[n_{triangles}, 3]$  matrix describing the connectivity list of the 3D triangulation describing the surface, i.e. indices of the grid-points defining each triangle.
- $numEigs$ : number of LB eigenfunctions to compute.

This algorithm outputs two variables, saved as a specific field in the MATLAB structure  $S$ :  $S.evecs$ , a  $[n_{grid-points}, n]$  matrix containing the first  $n$  LB eigenfunctions computed on the shape  $S$ ;  $S.evals$ , a  $[1, n_{grid-points}]$  containing their corresponding eigenvalues.

Listing A.1. Cotangent-Weight Scheme MATLAB algorithm

---

```
1 %% compute_LB.m
2 function S = compute_LB(S, numEigs)
3 % Function to compute the first numEigs Laplace-Beltrami eigenfunctions on
4 % the shape S.
5
6 fprintf('Computing %d Eigenfunctions...\n', numEigs); tic;
7
8 % COTANGENT_WEIGHT SCHEME
9 S.W = cotWeights(S.VERT, S.TRIV);
10 S.A = diag(vertexAreas(S.VERT, S.TRIV));
11
12 % GENERALIZED EIGENVALUE-PROBLEM.
13 [S.evecs, tmp] = eigs(S.W, S.A, numEigs, 'SM');
```

```

14
15 % CHECK THAT ALL EIGENVALUES ARE REAL
16 tmp_full = diag(tmp);
17 tmp = abs(tmp_full);
18 if imag(tmp_full) == 0
19     fprintf('All eigenvalues are real\n')
20 else
21     fprintf('Complex eigenvalues\n')
22 end
23
24 % Eigenfunctions are sorted into increasing eigenvalue order.
25 [S.evals,indices] = sort(tmp,'ascend'); clear tmp; clear tmp_full;
26 S.evecs = S.evecs(:,indices); clear indices;
27
28 fprintf('done. ');
29
30 end
31
32
33 %% cotWeights.m
34 function [W, A] = cotWeights(X, T)
35 % Function to compute the cotangent-weights matrix W.
36
37 % FIND ORIGINAL EDGE LENGTHS AND ANGLES
38 nv = size(X,1);
39 L1 = normv(X(T(:,2),:)-X(T(:,3),:));
40 L2 = normv(X(T(:,1),:)-X(T(:,3),:));
41 L3 = normv(X(T(:,1),:)-X(T(:,2),:));
42 EL = [L1,L2,L3];
43 A1 = (L2.^2 + L3.^2 - L1.^2) ./ (2.*L2.*L3);
44 A2 = (L1.^2 + L3.^2 - L2.^2) ./ (2.*L1.*L3);
45 A3 = (L1.^2 + L2.^2 - L3.^2) ./ (2.*L1.*L2);
46 A = [A1,A2,A3];
47 A = acos(A);
48
49 % COTANGENT LAPLACIAN
50 I = [T(:,1);T(:,2);T(:,3)];
51 J = [T(:,2);T(:,3);T(:,1)];
52 S = 0.5*cot([A(:,3);A(:,1);A(:,2)]);
53 In = [I;J;I;J];
54 Jn = [J;I;I;J];
55 Sn = [-S;-S;S;S];
56
57 W = sparse(double(In),double(Jn),double(Sn),nv,nv);
58

```

```

59 end
60
61
62 %% normv.m
63 function nn = normv(v)
64 % Function to compute vector norm.
65 nn = sqrt(sum(v.^2,2));
66
67 end
68
69
70 %% vertexAreas.m
71 function [A,At] = vertexAreas(X, T)
72 % Function to compute lumped area-weights matrix A.
73
74 % TRIANGLE AREAS
75 N = cross(X(T(:,1),:)-X(T(:,2),:), X(T(:,1),:) - X(T(:,3),:));
76 At = normv(N)/2;
77
78 % VERTEX AREAS
79 I = [T(:,1);T(:,2);T(:,3)];
80 J = ones(size(I));
81 S = double([At(:,1);At(:,1);At(:,1)]);
82 nv = size(X,1);
83
84 A = sparse(I,J,S,nv,1)/3;
85
86 end

```

---



# Appendix B

## Details of the experiment

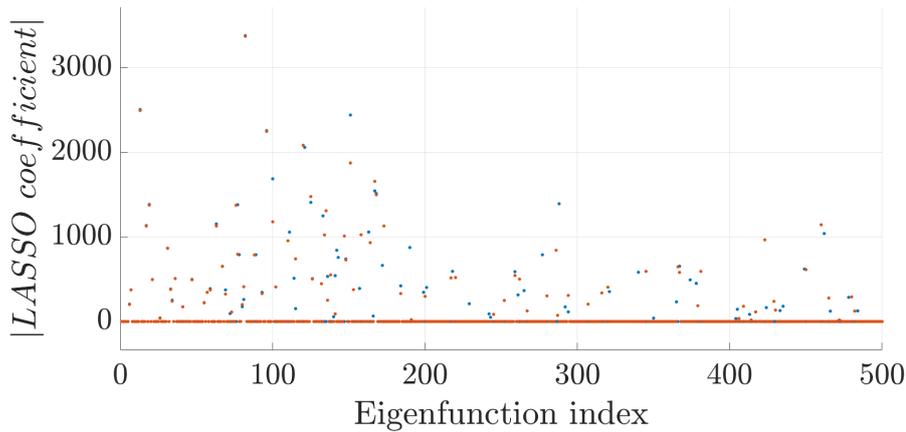
### B.1 Numerical error

This section proposes a qualitative analysis of the numerical error computed whenever the surface elevation  $k$  of a rough geometry is written as a linear combination of the first  $n$  LB eigenfunctions discretized on a smooth wall with different domain dimensions (in viscous units). About that, LASSO coefficients are computed between vector  $k$  of all the rough surfaces and the first 500 LB eigenfunctions (without considering the first one) discretized on the reference plane of dimensions either equal to that of the corresponding geometries (*case 1*) or equal to  $(3h \cdot 500, 3h \cdot 500)$  (as done in Chapter 4) (*case 2*). Figure B.1 reports the results obtained, in these two cases, for the eleven A7088 geometries. As initially guessed, these results slightly differ in both the highlighted eigenfunction indices and their LASSO coefficient. However, this numerical error, which has to be introduced to fairly compare different surfaces, is negligible for almost every predictor.

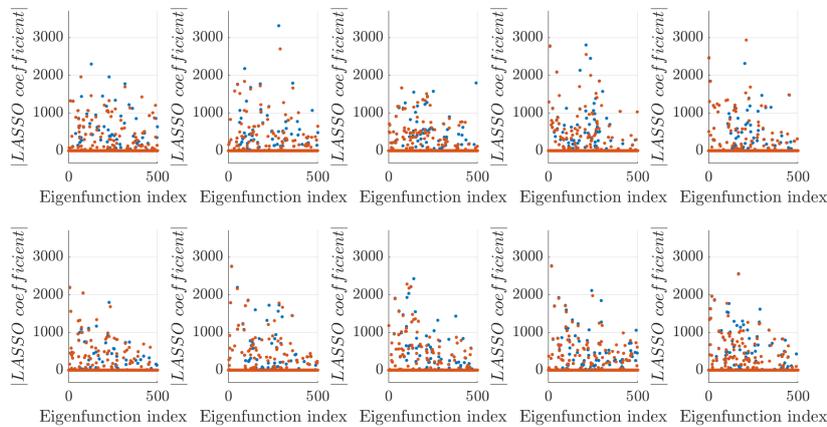
To quantify this error, the linear system 4.1 must be solved in the *case 1*. Then, coefficients  $\{\alpha_i\}_1$  are transferred to the reference plane represented in *case 2* through a *functional map*  $C$ :

$$\{\alpha_i\}_2 \approx C\{\alpha_i\}_1 \quad (\text{B.1})$$

The functional map  $C$  can be computed as  $C = \Phi_2^+ \Pi_{21}^T \Phi_1$  [28], where  $\Phi_1$  and  $\Phi_2$  are the matrices of the first 500 LB eigenfunctions computed on the flat plate in the *case 1* and *case 2*, respectively, while  $\Pi$  is a  $(n_{grid-points_2}, n_{grid-points_1})$  binary matrix encoding the pointwise map  $T$  between the flat plate in *case 2* and *case 1* using:  $\Pi_{21}(p, q) = 1$  if  $T(p) = q$ , 0 elsewhere. In the case of interest,  $\Pi = I$  (same grid resolution) and  $\Phi_1 \approx \Phi_2$ , thus making the numerical error in approximating  $\{\alpha_i\}_1$  with  $\{\alpha_i\}_2$  very small. However, this passage needs to be executed whenever the rough surface and the flat plate have very different domain sizes (in viscous units) and if the two do not share the same grid resolution.



(a)



(b)

Figure B.1. Absolute value of LASSO coefficients between surface elevation vectors of A7088 geometries and the first 500 LB eigenfunctions on the reference plane, computed with  $\lambda = 1$ . (a) Original geometry; (b) *first row*: geometries obtained by shuffling the reference A7088 center locations; *second row*: geometries obtained by shuffling the reference A7088 peak sizes. Blue points are computed when *case 1* (explained in section B.1) is considered; orange points are computed when *case 2* is considered.

## B.2 Parameters tuning

The same experiment conducted in Chapter 4 is here reproduced with different values of the two characteristic parameters ( $n, \lambda$ ). The purpose is to understand both advantages and disadvantages in increasing or decreasing these values.

### B.2.1 Effect of $n$

The first parameter to be set is the number  $n$  of LB eigenfunctions beyond which the series 4.1 is truncated. This parameter should be chosen as the lowest value

guaranteeing that all the important predictors for each shape collection are included. The main work proposed to use  $n = 500$ ; here, the same results are presented when considering  $n = 300$  and  $n = 1000$ .

Figure B.2 reports the matrices obtained with  $n = 300$ . Each correlation outlines that, in all the twenty-five cases, the entire frequency spectrum is excited. Moreover, most of the geometries need very relevant predictors at the highest frequencies, suggesting that  $n = 300$  is too low to describe the present surfaces. Many important features at higher frequencies may have to be taken into account.

Figure B.3 presents the results computed with  $n = 1000$ . As assumed in Chapter 4, these correlations highlight some features with high absolute value of their correlation coefficient at very high index, especially for roughness topographies with  $ES = 0.88$ . However, all the different geometries retain almost every important features within the first 500 eigenfunctions, with a average "relevance trend" decreasing with the increase of the eigenfunction frequency. As such, it is a good choice to truncate the series at lower values of  $n$  (e.g. 500) and accept to loose information coming from few relevant predictors, rather than double-size the model dimension, i.e. its complexity.

## B.2.2 Effect of $\lambda$

It is here discussed the setting of the regularization penalty  $\lambda$ . This parameter has a crucial importance since it establishes the weight of the penalization term, relative to that of the OLS cost function, in equation 3.6. This section aims to point out that the choice of  $\lambda$  is really problem-dependent, and one should take care on which is the dominant constraint between model prediction accuracy, useful to develop model predictions, and model complexity reduction.

A modern, machine learning approach to set  $\lambda$  solves a cross validation code. In this case, one should choose a set of  $P$  values of  $\lambda$  to test, split the dataset into  $K$  folds, and follow this algorithm:

- for  $p$  in  $1 : P$ 
  - for  $k$  in  $1 : K$ 
    - \* keep fold  $k$  as hold-out data
    - \* use the remaining folds and  $\lambda = \lambda_p$  to estimate  $\hat{\alpha}_{LASSO}$
    - \* predict hold-out data  $f_{test,k} = \Phi_{test,k} \hat{\alpha}_{LASSO}$
    - \* computed a sum of squared residuals:  $SSR_k = \|f - f_{test,k}\|^2$
  - end for  $k$
  - average  $SSR$  over the folds:  $SSR_p = (1/K) \sum_{k=1}^K SSR_k$
- end for  $p$
- choose optimal value  $\lambda_{opt} = \operatorname{argmin}_p SSR_p$

In MATLAB, it is possible to enable the input option 'CV' in function *lasso*, and choose the value of  $k$  to perform the  $k$ -fold cross-validation. A typical value of  $k$  is  $k = 10$  [10]. Following this procedure, a possibly different value of  $\lambda$  is obtained for each of the twenty-five geometries and their copies, but all of these values are approximately equal to 0.1. Therefore, to fairly compare all the surfaces, the experiment is conducted with  $\lambda = 0.1$  and its results are presented in figure B.4. As initially guessed, this value, which is an "average optimum" in terms of model accuracy, does not allow to distinguish the relevant predictors for the different rough surfaces, similarly to OLS. Consequently, it is not a good choice for the case of interest

The regularization penalty must be set to a higher value to reduce the model dimension and highlight relevant features. However, an exaggerated  $\lambda$  might cancel at all some important contributors, thus making the results completely unreliable. Figure B.5 shows the results obtained with  $\lambda = 2$ . It is clear how the LASSO algorithm retains way less predictors than the case of  $\lambda = 1$ , especially for high-ES surfaces. Specifically, an increase in the value of  $\lambda$  translates in a sort of "downward shift" of the LASSO coefficients presented in figure 4.3, explaining why low-ES surfaces in figure B.5 retain much more predictors than the high-ES ones. While the results obtained for the former case might be useful, those obtained for the latter are certainly unreliable.

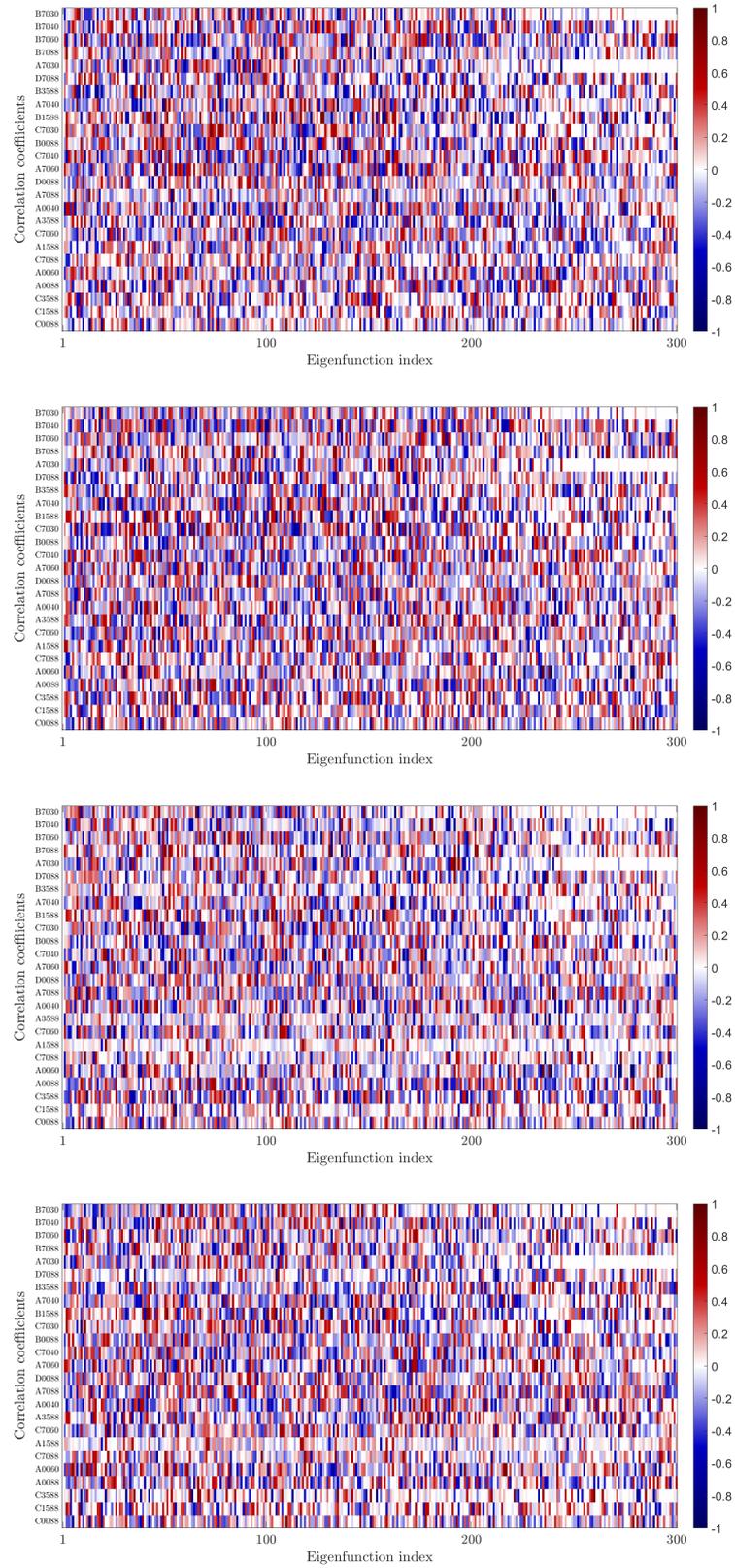


Figure B.2. Correlation matrix between LASSO coefficients of all the different geometries and their (*first row*)  $k_s/k_{rms}$ ; (*second row*)  $k_s/k_p$ ; (*third row*)  $k_s/k_t$ ; (*fourth row*)  $k_s/k_z$  vectors. Coefficients are computed with  $\lambda = 1$  and  $n = 300$ . Colored scale indicates correlation coefficient values.

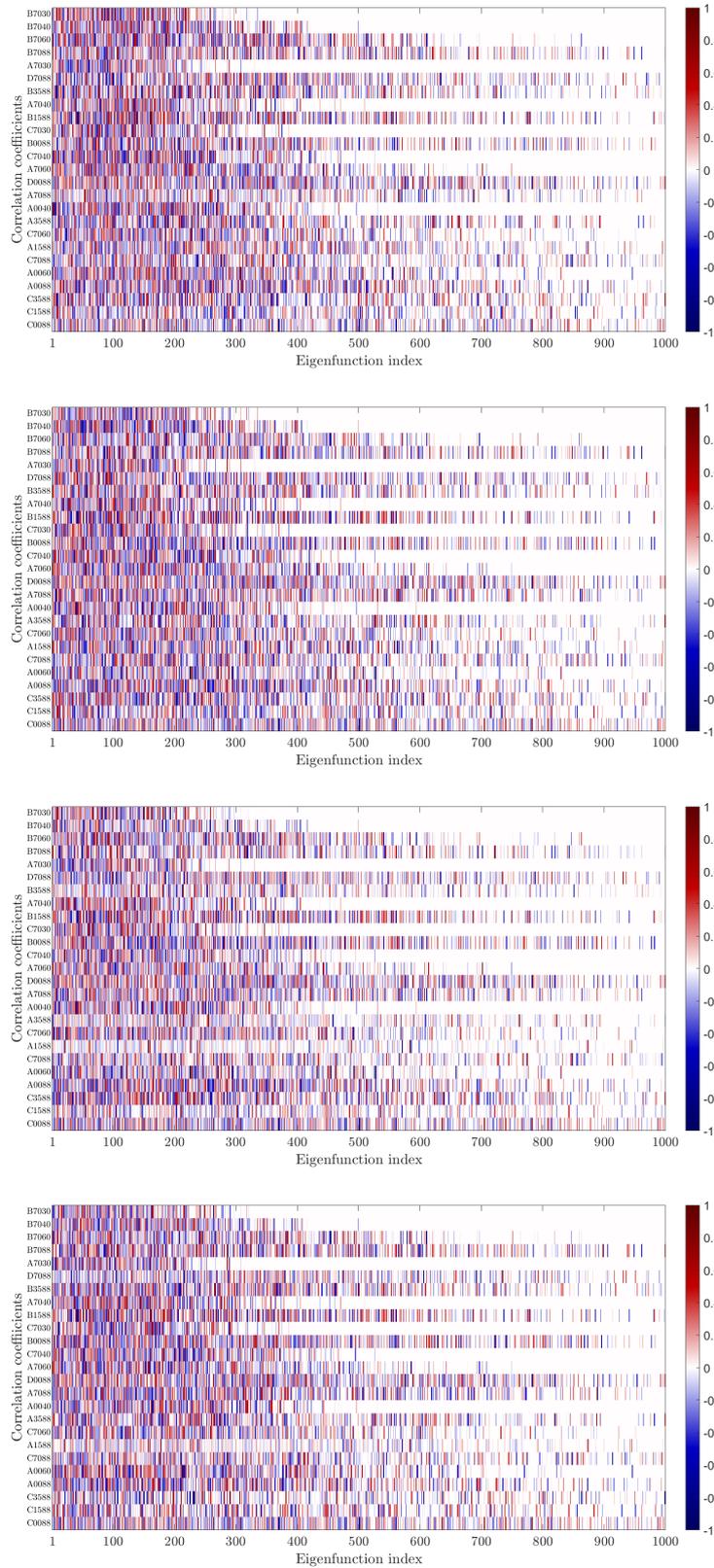


Figure B.3. Correlation matrix between LASSO coefficients of all the different geometries and their (*first row*)  $k_s/k_{rms}$ ; (*second row*)  $k_s/k_p$ ; (*third row*)  $k_s/k_t$ ; (*fourth row*)  $k_s/k_z$  vectors. Coefficients are computed with  $\lambda = 1$  and  $n = 1000$ . Colored scale indicates correlation coefficient values.

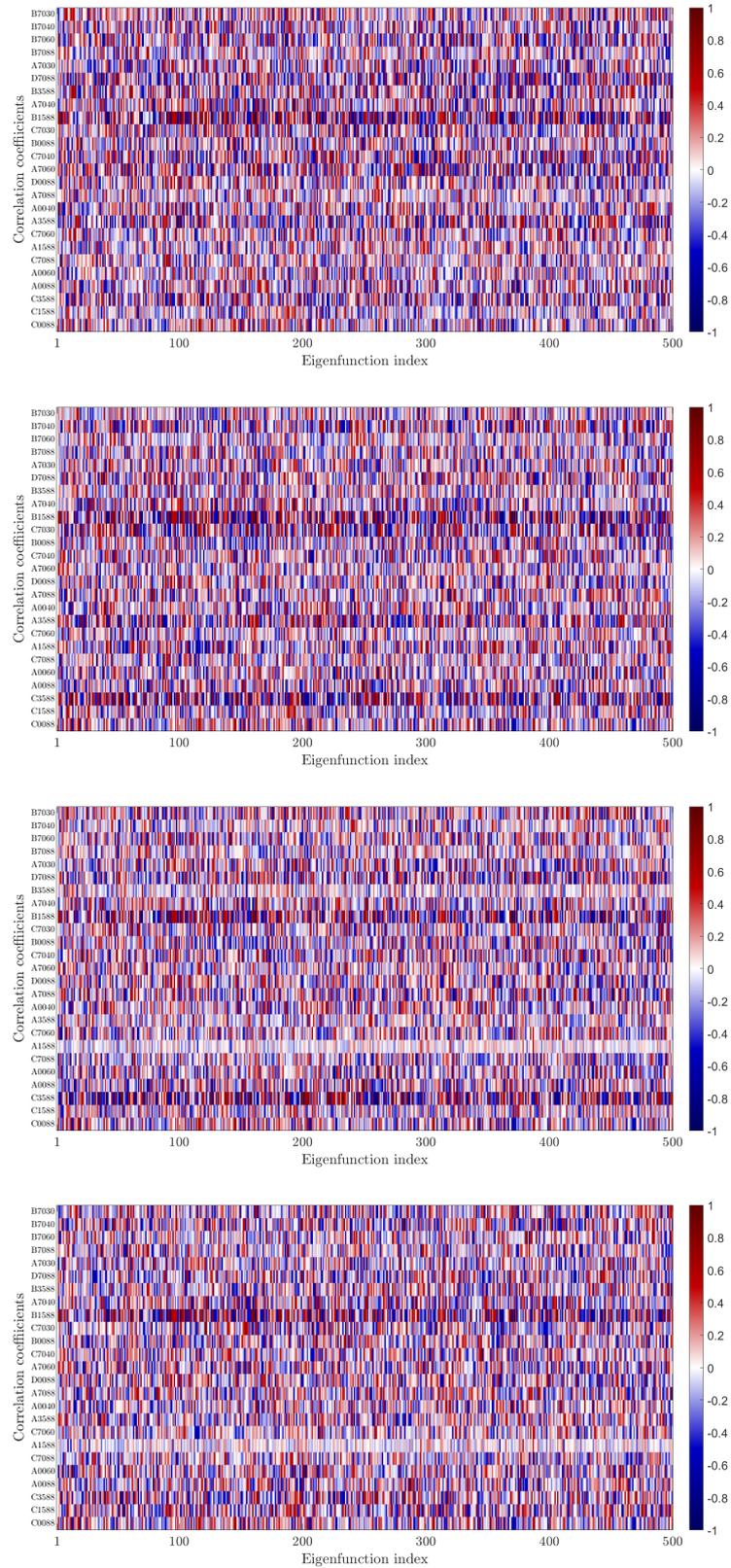


Figure B.4. Correlation matrix between LASSO coefficients of all the different geometries and their (*first row*)  $k_s/k_{rms}$ ; (*second row*)  $k_s/k_p$ ; (*third row*)  $k_s/k_t$ ; (*fourth row*)  $k_s/k_z$  vectors. Coefficients are computed with  $\lambda = 0.1$  and  $n = 500$ . Colored scale indicates correlation coefficient values.

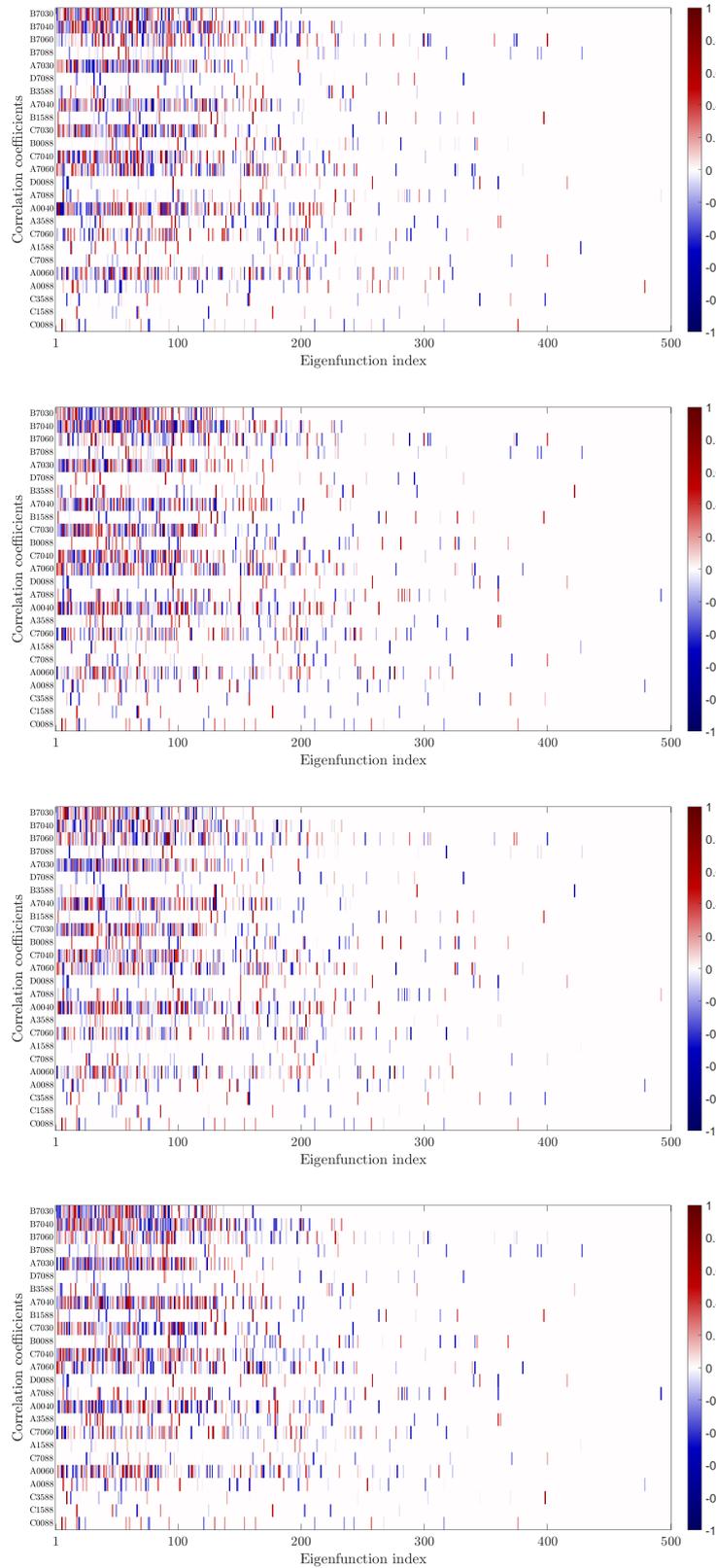


Figure B.5. Correlation matrix between LASSO coefficients of all the different geometries and their (*first row*)  $k_s/k_{rms}$ ; (*second row*)  $k_s/k_p$ ; (*third row*)  $k_s/k_t$ ; (*fourth row*)  $k_s/k_z$  vectors. Coefficients are computed with  $\lambda = 2$  and  $n = 500$ . Colored scale indicates correlation coefficient values.

# Bibliography

- [1] ANGUELOV, D., SRINIVASAN, P., KOLLER, D., THRUN, S., RODGERS, J., AND DAVIS, J. Scape: shape completion and animation of people. In *ACM SIGGRAPH 2005 Papers*. 2005, pp. 408–416.
- [2] BONS, J. A critical assessment of Reynolds analogy for turbine flows. *Journal of Heat Transfer* 127, 5 (2005), 472–485.
- [3] BUSSE, A., AND JELLY, T. O. Influence of surface anisotropy on turbulent flow over irregular roughness. *Flow, Turbulence and Combustion* 104, 2 (2020), 331–354. Publisher: Springer.
- [4] BUSSE, A., LÜTZNER, MARK, AND SANDHAM, NEIL D. Direct numerical simulation of turbulent flow over a rough surface based on a surface scan. *Computers & Fluids* 116 (2015), 129–147. Publisher: Elsevier.
- [5] CHAN, L., MACDONALD, M., CHUNG, D., HUTCHINS, N., AND OOI, A. A systematic investigation of roughness height and wavelength in turbulent pipe flow in the transitionally rough regime. *Journal of Fluid Mechanics* 771 (2015), 743–777. Publisher: Cambridge University Press.
- [6] CHUNG, D., HUTCHINS, N., SCHULTZ, M. P., AND FLACK, K. A. Predicting the Drag of Rough Surfaces. *Annual Review of Fluid Mechanics* 53 (2021), 439–471. Publisher: Annual Reviews.
- [7] FLACK, K., SCHULTZ, M., AND BARROS, J. Skin friction measurements of systematically-varied roughness: Probing the role of roughness amplitude and skewness. *Flow, Turbulence and Combustion* 104, 2 (2020), 317–329. Publisher: Springer.
- [8] FLACK, K. A., AND SCHULTZ, M. P. Review of hydraulic roughness scales in the fully rough regime. *Journal of Fluids Engineering* 132, 4 (2010). Publisher: American Society of Mechanical Engineers Digital Collection.
- [9] FOROOGHI, P., STROH, A., MAGAGNATO, F., JAKIRLIĆ, S., AND FROHNAPFEL, B. Toward a universal roughness correlation. *Journal of Fluids Engineering* 139, 12 (2017). Publisher: American Society of Mechanical Engineers Digital Collection.

- [10] FUSHIKI, T. Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing* 21, 2 (2011), 137–146. Publisher: Springer.
- [11] GOLDSTEIN, D., HANDLER, R., AND SIROVICH, L. Modeling a no-slip flow boundary with an external force field. *Journal of Computational Physics* 105, 2 (1993), 354–366. Publisher: Elsevier.
- [12] HAMA, FRANCIS R. Boundary layer characteristics for smooth and rough surfaces. *Transactions - The Society of Naval Architects and Marine Engineers*. 62 (1954), 333–358.
- [13] HASLER, N., STOLL, C., SUNKEL, M., ROSENHAHN, B., AND SEIDEL, H.-P. A statistical model of human pose and body shape. In *Computer graphics forum* (2009), vol. 28, Wiley Online Library, pp. 337–346. Issue: 2.
- [14] HUANG, Q., WANG, F., AND GUIBAS, L. Functional map networks for analyzing and exploring large shape collections. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 1–11. Publisher: ACM New York, NY, USA.
- [15] HUANG, R., ACHLIOPTAS, P., GUIBAS, L., AND OVSJANIKOV, M. Limit shapes—a tool for understanding shape differences and variability in 3d model collections. In *Computer Graphics Forum* (2019), vol. 38, Wiley Online Library, pp. 187–202. Issue: 5.
- [16] JELLY, T. O., AND BUSSE, A. Reynolds and dispersive shear stress contributions above highly skewed roughness. *Journal of Fluid Mechanics* 852 (2018), 710–724. Publisher: Cambridge University Press.
- [17] JIMÉNEZ, J. Turbulent flows over rough walls. *Annual Review of Fluid Mechanics* 36 (2004), 173–196. Publisher: Annual Reviews.
- [18] JOUYBARI, M. A., YUAN, J., BRERETON, G. J., AND MURILLO, M. S. Data-driven prediction of the equivalent sand-grain height in rough-wall turbulent flows. *Journal of Fluid Mechanics* 912 (2021). Publisher: Cambridge University Press.
- [19] KATO, T. *Perturbation theory for linear operators*, vol. 132. Springer Science & Business Media, 2013.
- [20] KUWATA, Y., AND KAWAGUCHI, Y. Direct numerical simulation of turbulence over systematically varied irregular rough surfaces. *Journal of Fluid Mechanics* 862 (2019), 781–815. Publisher: Cambridge University Press.
- [21] LEE, S., YANG, J., FOROOGHI, P., STROH, A., AND BAGHERI, S. Predicting drag on rough surfaces by transfer learning of empirical correlations. *arXiv preprint arXiv:2106.05995* (2021).

- 
- [22] LUCHINI, P., MANZO, FERNANDO, AND POZZI, AMILCARE. Resistance of a grooved surface to parallel flow and cross-flow. *Journal of Fluid Mechanics* 228 (1991), 8–109. Publisher: Cambridge University Press.
- [23] MACDONALD, M., OOI, A., GARCÍA-MAYORAL, R., HUTCHINS, N., AND CHUNG, D. Direct numerical simulation of high aspect ratio spanwise-aligned bars. *Journal of Fluid Mechanics* 843 (2018), 126–155. Publisher: Cambridge University Press.
- [24] MACDONALD, R., GRIFFITHS, R., AND HALL, D. An improved method for the estimation of surface roughness of obstacle arrays. *Atmospheric environment* 32, 11 (1998), 1857–1864. Publisher: Elsevier.
- [25] MCDONALD, G. C. Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics* 1, 1 (2009), 93–100. Publisher: Wiley Online Library.
- [26] MEYER, M., DESBRUN, M., SCHRÖDER, P., AND BARR, A. H. Discrete differential-geometry operators for triangulated 2-manifolds. In *Visualization and mathematics III*. Springer, 2003, pp. 35–57.
- [27] NIKURADSE, J. Stromungsgesetze in rauhen Rohren. *VDI-Forschungsheft* 361 (1933), 1.
- [28] OVSJANIKOV, M., BEN-CHEN, M., SOLOMON, J., BUTSCHER, A., AND GUIBAS, L. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG)* 31, 4 (2012), 1–11. Publisher: ACM New York, NY, USA.
- [29] PINKALL, U., AND POLTHIER, K. Computing discrete minimal surfaces and their conjugates. *Experimental mathematics* 2, 1 (1993), 15–36. Publisher: Taylor & Francis.
- [30] RANSTAM, J., AND COOK, J. LASSO regression. *Journal of British Surgery* 105, 10 (2018), 1348–1348. Publisher: Oxford University Press.
- [31] RAUPACH, M. R., ANTONIA, R. A., AND RAJAGOPALAN, S. Rough-wall turbulent boundary layers.
- [32] REUTER, M., BIASOTTI, S., GIORGI, D., PATANÈ, G., AND SPAGNUOLO, M. Discrete Laplace–Beltrami operators for shape analysis and segmentation. *Computers & Graphics* 33, 3 (2009), 381–390. Publisher: Elsevier.
- [33] ROSENBERG, S., AND STEVEN, R. *The Laplacian on a Riemannian manifold: an introduction to analysis on manifolds*. 31. 1997.
- [34] SARAkinos, S., AND BUSSE, A. Influence of spatial distribution of roughness elements on turbulent flow past a biofouled surface.

- [35] THAKKAR, M., BUSSE, A., AND SANDHAM, N. Surface correlations of hydrodynamic drag for transitionally rough engineering surfaces. *Journal of Turbulence* 18, 2 (2017), 138–169. Publisher: Taylor & Francis.
- [36] VRIEZE, S. I. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods* 17, 2 (2012), 228. Publisher: American Psychological Association.
- [37] WANG, Y., ASAFI, S., VAN KAICK, O., ZHANG, H., COHEN-OR, D., AND CHEN, B. Active co-analysis of a set of shapes. *ACM Transactions on Graphics (TOG)* 31, 6 (2012), 1–10. Publisher: ACM New York, NY, USA.
- [38] YANG, J., STROH, A., CHUNG, D., AND FOROOGHI, P. DNS-based characterization of pseudo-random roughness in minimal channels. *arXiv preprint arXiv:2106.08160* (2021).
- [39] ZOU, H., AND HASTIE, T. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)* 67, 2 (2005), 301–320. Publisher: Wiley Online Library.