



Turbulent velocity field reconstruction using four-dimensional variational data assimilation

Naseer Sabah Abdullah

Submitted for the degree of PhD
School of Mathematics and Statistics

May, 2019

Supervisor: Dr Yi Li

University of Sheffield

I would like to thank my family. My great wife Sura, my little heroes my children Milan and Sam, my parents for giving birth to me in the first place Muna and Sabah, to my brothers Bashar and Ammar and sisters Rana and Reem for supporting me spiritually throughout writing this thesis and my life in general.

I would particularly like to thank my beloved country Iraq and the Iraqi authorities represented by The Higher Committee for Education Development in Iraq, for their trust by giving me a great opportunity to obtain a PhD degree from a respected university with a scholarship with their financial support.

I thank the mother university, the University of Mosul-Iraq, for the permission to leave my position as a lecturer for all previous years and provided me with the support necessary to study and obtain a PhD.

Besides those, I would like to thank my advisor Dr Julia Rees, Prof Elizabeth Winstanley, and Dr Prof Nick Monk, for their insightful comments and encouragement.

Last but not least, I would like to express my gratitude to the University of Sheffield for all the support they offered me to overcome the difficulties that arose along with my study.

Abstract

Important progress in computational fluid dynamics has been made recently by applying the data assimilation (DA) techniques. In this thesis, we apply the four-dimensional variational approach to rebuild the small scales of the velocity fields of fully developed three-dimensional turbulence, given a time sequence of measurement data on a coarse mesh of grid points. In this problem, we deal with new challenges since the flow is governed by the processes of nonlinear vortex stretching and forward energy cascade, which are absent in two-dimensional flows that have been investigated so far. Two different models are presented to examine their effects on the reconstruction quality: the Navier-Stokes equations as the model and when the large eddy simulations are applied as the model (Smagorinsky model). The investigations examine different statistics of the reconstructed fields. The results show that the agreement improves over time within the optimization horizon, where the rebuilt fields tend to the DNS target. Reasonable agreements are accomplished between the optimal initial fields and the target data. To assess the quality of the reconstruction of non-local structures, minimum volume enclosing ellipsoids are introduced, which enables us to perform quantitative comparisons for the geometry of non-local structures. The rebuilding of non-local structures with strong vorticity, strain rate and subgrid-scale energy dissipation gives satisfactory results. A small misalignment between the MVEE's axes can be obtained; structures in the rebuilt fields are reproduced with sizes smaller by a small percentage to what exists in the target field; the locations of the MVEE are different on average by around 20% of the axes lengths. Both Navier-Stokes equations and filtered Navier-Stokes equation (as models for the data) show satisfactory reconstruction. The imperfect model (Smagorinsky model) demonstrate the capability for recovering the target fields quicker in most of the presented examinations.

Acknowledgements

Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from a special person. Therefore, I would like to express my sincere gratitude to my supervisor Dr *Yi Li* for the continuous support of my PhD study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study.

Contents

1	Introduction	1
1.1	Homogeneous turbulence and large eddy simulation	1
1.2	The basic concepts and methods of data assimilation	5
1.2.1	Optimal Interpolation Methods	6
1.2.2	Variational optimization	8
1.2.3	Kalman filter	10
1.2.4	Comparison between 4DVAR and EnKF	13
1.3	The applications of DA in fluid mechanics	14
1.4	The objective and the outline of the dissertation	19
2	Turbulent fields reconstruction using 4D variational data assimilation with the Navier-Stokes equations as the model	22
2.1	Introduction	22
2.2	Formulation	23
2.3	Iterative Solution	24
2.4	Simulation and discussions	25
2.4.1	The energy spectrum $E(k)$	25
2.4.2	The energy spectrum $E_\delta(k)$	27
2.4.3	The probability density functions of the alignment between the vorticity and rate of the strain tensor	29
2.4.4	The probability density functions of the velocity increments	29
2.4.5	The subgrid-scale energy dissipations	32
2.4.6	The probability density functions of the shape factor S^*	33

2.4.7	DNS of the Gaussian random field with same energy spectrum from the optimal initial field	35
2.4.8	Contour plots of instantaneous distributions	41
2.4.9	Characterising structures with minimum volume enclosing ellipsoids	49
2.4.10	The minimum volume enclosing ellipsoid (MVEEs) for the vorticity field	53
2.4.11	The minimum volume enclosing ellipsoid (MVEEs) for the strain rate field	63
2.4.12	The minimum volume enclosing ellipsoid (MVEEs) for the Subgrid-scale energy dissipation	69
2.4.13	Comparisons between the three types of structures	70
3	Reconstructing turbulent fields using 4D variational data assimilation: large eddy simulations and the Smagorinsky model	77
3.1	Introduction	77
3.2	An introduction to large eddy simulations	78
3.3	Formulation for the 4DVAR problem	78
3.4	Simulation and discussions	79
3.4.1	Contour plots of instantaneous distributions	85
3.4.2	The minimum volume enclosing ellipsoid (MVEEs)	90
3.4.3	Summary	108
4	Conclusions	110
A	Derivation of the optimality system for the 4DVAR with the Navier-Stokes equations as the model	113
B	The adjoint equation for the filtered NS equations with the standard Smagorinsky model	119

List of Figures

2.1	The energy spectrum. Solid line with squares: DNS target. Dotted line with circles: OIC. Dashed line: Kolmogorov ($-5/3$).	26
2.2	Spectrum of $\mathbf{v} - \mathbf{u}$. Solid line with squares: $E(k)$ for DNS target. Dotted line with circles: $E_\delta(k)$	27
2.3	The ratio of the spectra: $E_\delta(k)/E(k)$. Solid line with: Square for $0T$. Downward-pointing triangle for $T/3$. Diamond for $T/2$. Upward-pointing triangle for $2/3T$. Circle for T	28
2.4	PDFs of the alignment between $\tilde{\omega}$ and the eigenvectors of \tilde{S}_{ij} . DNS target is solid line with squares, circles; OIC is dotted line with upward-pointing triangle, diamond; OIC is dashed line with asterisks, right-pointing triangle: intermediate eigen-direction and eigen-direction with negative eigenvalue, respectively.	30
2.5	PDFs of velocity increments. δu_r : Longitudinal. δv_r : Transverse. Solid line is DNS target: with squares (δu_r) and diamond (δv_r). Dotted line is OIC: with circles (δu_r) and upward-pointing triangle (δv_r). Dashed line is OIC: with asterisks (δu_r) and right-pointing triangle (δv_r). The thin line is normal distribution.	31
2.6	The mean subgrid-scale energy dissipations as a function of the filter scales Δ	32
2.7	PDFs of the subgrid-scale energy dissipation.	33
2.8	PDFs of the shape factor $S^*(\tilde{S}_{ij})$ for the filtered strain tensor \tilde{S}_{ij} . Solid line is DNS target: with squares for $0T$, diamond for $T/6$. Dotted line is OIC: with circles for $0T$ and upward-pointing triangle for $T/6$	34
2.9	PDFs of the shape factor $S^*(\tau_{ij})$ for the subgrid-scale stress tensor τ_{ij} . Solid line is DNS target: with squares for $0T$. Dotted line is OIC: with circles for $0T$	35

2.10	PDFs of the alignment between $\tilde{\omega}$ and the eigenvectors of \tilde{S}_{ij} . GR: solid line with squares, circles: intermediate eigen-direction and eigen-direction with negative eigenvalue, respectively. DNS target: dotted line with upward-pointing triangle, diamond, respectively.	36
2.11	PDFs of velocity increments. δu_r : Longitudinal. δv_r : Transverse. Solid line is from GR ICs: with squares (δu_r) and Diamond (δv_r). Dotted line is DNS target: with circles (δu_r) and Upward-pointing triangle (δv_r).	37
2.12	The mean subgrid-scale energy dissipations as a function of the filter scales Δ	38
2.13	PDFs of the subgrid-scale energy dissipation.	39
2.14	PDFs of the shape factor $S^*(\tilde{S}_{ij})$ for the filtered strain tensor \tilde{S}_{ij}	40
2.15	PDFs of the shape factor $S^*(\tau_{ij})$ for the subgrid-scale stress tensor τ_{ij}	40
2.16	Contour plots for $ \tilde{\omega} $ at $t = 0T$	41
2.17	Contour plots for $ \tilde{\omega} $ at $t = T/3$	42
2.18	Contour plots for $ \tilde{\omega} $ at $t = 2/3T$	42
2.19	Contour plots for $ \tilde{\omega} $ at $t = T$	43
2.20	Contour plots for $ \tilde{S} $ at $t = 0T$	43
2.21	Contour plots for $ \tilde{S} $ at $t = T/3$	44
2.22	Contour plots for $ \tilde{S} $ at $t = 2/3T$	44
2.23	Contour plots for $ \tilde{S} $ at $t = T$	45
2.24	Q criterion for $t = 0T$	46
2.25	Q criterion for $t = T/2$	46
2.26	Q criterion for $t = T$	47
2.27	Q Contour plots for $t = 0T$	47
2.28	Q Contour plots for $t = T/2$	48
2.29	Q Contour plots for $t = T$	48
2.30	Structures as clusters of points for $ \tilde{S} $ at $t = 0T$	51
2.31	Structures as clusters of points for $ \tilde{S} $ at $t = T$	52
2.32	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = 0T$	53
2.33	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = 0T$	54
2.34	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = 0T$	54
2.35	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = T/6$	55

2.36	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = T/6$.	55
2.37	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = T/6$.	55
2.38	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = T/3$.	56
2.39	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = T/3$.	56
2.40	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = T/3$.	56
2.41	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = T/2$.	57
2.42	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = T/2$.	57
2.43	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = T/2$.	57
2.44	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = 2/3T$.	58
2.45	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = 2/3T$.	58
2.46	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = 2/3T$.	58
2.47	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = 5/6T$.	59
2.48	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = 5/6T$.	59
2.49	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = 5/6T$.	59
2.50	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = T$.	60
2.51	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = T$.	60
2.52	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = T$.	60
2.53	Medians of $ \cos(e^T, e^D) $ for strong $ \tilde{\omega} $ structures. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.	61
2.54	Median ratio between the sizes of the clusters for strong $ \tilde{\omega} $ structures. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.	62
2.55	d_c/ℓ for strong $ \tilde{\omega} $ structures at $t = 0T$.	62
2.56	d_c/ℓ for strong $ \tilde{\omega} $ structures at $t = T/2$.	63
2.57	d_c/ℓ for strong $ \tilde{\omega} $ structures at $t = T$.	63
2.58	Alignment between the major axes for $ \tilde{S} $ at $t = 0T$.	64
2.59	Alignment between the middle axes for $ \tilde{S} $ at $t = 0T$.	64
2.60	Alignment between the minor axes for $ \tilde{S} $ at $t = 0T$.	64
2.61	Alignment between the major axes for $ \tilde{S} $ at $t = T/3$.	65
2.62	Alignment between the middle axes for $ \tilde{S} $ at $t = T/3$.	65
2.63	Alignment between the minor axes for $ \tilde{S} $ at $t = T/3$.	65

2.64	Alignment between the major axes for $ \tilde{S} $ at $t = 2/3T$	66
2.65	Alignment between the middle axes for $ \tilde{S} $ at $t = 2/3T$	66
2.66	Alignment between the minor axes for $ \tilde{S} $ at $t = 2/3T$	66
2.67	Alignment between the major axes for $ \tilde{S} $ at $t = T$	67
2.68	Alignment between the middle axes for $ \tilde{S} $ at $t = T$	67
2.69	Alignment between the minor axes for $ \tilde{S} $ at $t = T$	67
2.70	Medians of $ \cos(e^T, e^D) $ for $ \tilde{S} $. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.	68
2.71	Median ratio between the sizes of the clusters for $ \tilde{S} $. Black with circles: Median ratios for all clusters. Red with circles: median ratios for clusters of a given size.	69
2.72	d_c/ℓ for $ \tilde{S} $ at $t = 0T$	69
2.73	d_c/ℓ for $ \tilde{S} $ at $t = T/2$	70
2.74	d_c/ℓ for $ \tilde{S} $ at $t = T$	70
2.75	Alignment between the major axes for SGS energy dissipation at $t = 0T$. 71	
2.76	Alignment between the middle axes for SGS energy dissipation at $t = 0T$. 71	
2.77	Alignment between the minor axes for SGS energy dissipation at $t = 0T$. 71	
2.78	Alignment between the major axes for SGS energy dissipation at $t = T/3$. 72	
2.79	Alignment between the middle axes for SGS energy dissipation at $t = T/3$. 72	
2.80	Alignment between the minor axes for SGS energy dissipation at $t = T/3$. 72	
2.81	Alignment between the major axes for SGS energy dissipation at $t = 2/3T$. 73	
2.82	Alignment between the middle axes for SGS energy dissipation at $t = 2/3T$. 73	
2.83	Alignment between the minor axes for SGS energy dissipation at $t = 2/3T$. 73	
2.84	Alignment between the major axes for SGS energy dissipation at $t = T$. 74	
2.85	Alignment between the middle axes for SGS energy dissipation at $t = T$. 74	
2.86	Alignment between the minor axes for SGS energy dissipation at $t = T$. 74	
2.87	Medians of $ \cos(e^T, e^D) $ for SGS energy dissipation. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.	75
2.88	Median ratio between the sizes of the clusters for SGS energy dissipation. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.	75
2.89	d_c/ℓ for SGS energy dissipation at $t = 0T$	76

2.90	d_c/ℓ for SGS energy dissipation at $t = T/2$	76
2.91	d_c/ℓ for SGS energy dissipation at $t = T$	76
3.1	The energy spectrum. Solid line with squares: DNS target. Dotted line with circles: OICS. Dashed line: Kolmogorov $-5/3$ spectrum.	81
3.2	Spectrum of $\mathbf{v} - \bar{\mathbf{u}}$. Solid line with squares: $E(k)$ for DNS target. Dotted line with circles: $E_\delta(k)$	82
3.3	The ratio of the spectra: $E_\delta(k)/E(k)$	83
3.4	PDFs of the alignment between $\tilde{\tilde{\omega}}$ and the eigenvectors of $\tilde{\tilde{S}}_{ij}$. DNS target: solid line with squares, circles: intermediate eigen-direction and eigen-direction with negative eigenvalue, respectively. OICS: dotted line with upward-pointing triangle, diamond, respectively.	83
3.5	The mean subgrid-scale energy dissipations as a function of the filter scales Δ	84
3.6	PDFs of the subgrid-scale energy dissipation.	84
3.7	The probability density functions of the shape factor $S^*(\tilde{\tilde{S}}_{ij})$ for the filtered strain tensor $\tilde{\tilde{S}}_{ij}$. Solid line is DNS target: with squares for $0T$, circles for $T/8$. Dotted line is OICS: with upward-pointing triangle for $0T$ and diamond for $T/8$	85
3.8	PDFs of the shape factor $S^*(\bar{\tau}_{ij})$ for the subgrid-scale stress tensor $\bar{\tau}_{ij}$. Solid line is DNS target: with squares for $0T$. Dotted line is OICS: with circles for $0T$	86
3.9	Contour plots for $ \tilde{\tilde{\omega}} $ at $t = 0T$	86
3.10	Contour plots for $ \tilde{\tilde{\omega}} $ at $t = T/4$	87
3.11	Contour plots for $ \tilde{\tilde{\omega}} $ at $t = T/2$	87
3.12	Contour plots for $ \tilde{\tilde{\omega}} $ at $t = 3/4T$	88
3.13	Contour plots for $ \tilde{\tilde{\omega}} $ at $t = T$	88
3.14	Contour plots for $ \tilde{\tilde{S}} $ at $t = 0T$	89
3.15	Contour plots for $ \tilde{\tilde{S}} $ at $t = T/2$	89
3.16	Contour plots for $ \tilde{\tilde{S}} $ at $t = T$	90
3.17	Alignment between the major axes for strong $ \tilde{\tilde{\omega}} $ structures at $t = 0T$	90
3.18	Alignment between the middle axes for strong $ \tilde{\tilde{\omega}} $ structures at $t = 0T$	91
3.19	Alignment between the minor axes for strong $ \tilde{\tilde{\omega}} $ structures at $t = 0T$	91
3.20	Alignment between the major axes for strong $ \tilde{\tilde{\omega}} $ structures at $t = T/2$	92

3.21	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = T/2$.	92
3.22	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = T/2$.	92
3.23	Alignment between the major axes for strong $ \tilde{\omega} $ structures at $t = T$.	93
3.24	Alignment between the middle axes for strong $ \tilde{\omega} $ structures at $t = T$.	93
3.25	Alignment between the minor axes for strong $ \tilde{\omega} $ structures at $t = T$.	93
3.26	Medians of $ \cos(\mathbf{e}^T, \mathbf{e}^D) $ for strong $ \tilde{\omega} $ structures. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.	94
3.27	Median ratio between the sizes of the clusters for strong $ \tilde{\omega} $ structures. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.	95
3.28	d_c/ℓ for strong $ \tilde{\omega} $ structures at $t = 0T$.	95
3.29	d_c/ℓ for strong $ \tilde{\omega} $ structures at $t = T/2$.	96
3.30	d_c/ℓ for strong $ \tilde{\omega} $ structures at $t = T$.	96
3.31	Alignment between the major axes for $ \tilde{S} $ at $t = 0T$.	97
3.32	Alignment between the middle axes for $ \tilde{S} $ at $t = 0T$.	97
3.33	Alignment between the minor axes for $ \tilde{S} $ at $t = 0T$.	98
3.34	Alignment between the major axes for $ \tilde{S} $ at $t = T/2$.	98
3.35	Alignment between the middle axes for $ \tilde{S} $ at $t = T/2$.	98
3.36	Alignment between the minor axes for $ \tilde{S} $ at $t = T/2$.	99
3.37	Alignment between the major axes for $ \tilde{S} $ at $t = T$.	99
3.38	Alignment between the middle axes for $ \tilde{S} $ at $t = T$.	99
3.39	Alignment between the minor axes for $ \tilde{S} $ at $t = T$.	100
3.40	Medians of $ \cos(\mathbf{e}^T, \mathbf{e}^D) $ for $ \tilde{S} $. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.	101
3.41	Median ratio between the sizes of the clusters for $ \tilde{S} $. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.	102
3.42	d_c/ℓ for $ \tilde{S} $ at $t = 0T$.	102
3.43	d_c/ℓ for $ \tilde{S} $ at $t = T/2$.	103
3.44	d_c/ℓ for $ \tilde{S} $ at $t = T$.	103
3.45	Alignment between the major axes for SGS energy dissipation at $t = 0T$.	103

3.46	Alignment between the middle axes for SGS energy dissipation at $t = 0T$.	104
3.47	Alignment between the minor axes for SGS energy dissipation at $t = 0T$.	104
3.48	Alignment between the major axes for SGS energy dissipation at $t = T/2$.	104
3.49	Alignment between the middle axes for SGS energy dissipation at $t = T/2$.	105
3.50	Alignment between the minor axes for SGS energy dissipation at $t = T/2$.	105
3.51	Alignment between the major axes for SGS energy dissipation at $t = T$.	105
3.52	Alignment between the middle axes for SGS energy dissipation at $t = T$.	106
3.53	Alignment between the minor axes for SGS energy dissipation at $t = T$.	106
3.54	Medians of $ \cos(\mathbf{e}^T, \mathbf{e}^D) $ for SGS energy dissipation. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.	107
3.55	Median ratio between the sizes of the clusters for SGS energy dissipation. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.	108
3.56	d_c/ℓ for SGS energy dissipation at $t = 0T$	108
3.57	d_c/ℓ for SGS energy dissipation at $t = T/2$	109
3.58	d_c/ℓ for SGS energy dissipation at $t = T$	109

List of Symbols

K_{Kol}	Kolmogorov constant
ε	energy dissipation rate
$\langle \rangle$	a statistical average
l	characteristic length
L	Lagrangian function
λ	Lagrange multiplier
ρ	density
t	time
\mathbf{x}	vector coordinate system
∇	gradient operator
T	optimization horizon
$P_{jk}(\mathbf{k})$	projection tensor
τ_L	large eddy turn-over time scale
\mathbf{f}	force term in NS equations
$\bar{\mathbf{f}}$	force term in NS equations for imperfect model
\mathbf{k}	wave-number field
K_c	a certain threshold value in the target data
η_K	length scale of Kolmogorov
J	cost function of the variational analysis
ξ	velocity adjoint
\mathbf{N}	constraints
N	The number of grid points
μ	pressure adjoint/ the dynamic viscosity
$E_\delta(k)$	energy spectrum of a difference of two velocity fields
β_2	a constant number
P	pressure
B	periodic box
ν	kinematic velocity
\mathbf{u}	any velocity field/ a solution of the NS equations
\mathbf{v}	velocity field of the target data
$\bar{\mathbf{u}}$	filtered velocity field/ a solution of the filtered NS equations

\mathbf{u}'	residual (SGS) velocity field
\mathbf{u}^*	complex conjugate of \mathbf{u}
d_c	displacement of the centres of the MVEEs
ℓ	lengths of the axes of the MVEEs
Π	SGS kinetic energy dissipation
$\bar{\Pi}$	SGS kinetic energy dissipation for imperfect model
τ_{ij}	SGS stress tensor
$\bar{\tau}_{ij}$	SGS stress tensor for imperfect model
τ_{ij}^r	tensor of the anisotropic SGS stress
\bar{p}	modified filtered pressure
k_{max}	highest resolved wave-number
ν_r	residual subgrid-scale eddy viscosity
C_s	Smagorinsky coefficient
Δ	filter width
Ω	a specific domain
$\mathbf{\Omega}$	rate of rotation tensor
$\tilde{\mathbf{\Omega}}$	filtered rate of rotation tensor
Q	criterion for vortex identification
\mathbf{S}	strain rate tensor
$\tilde{\mathbf{S}}$	filtered strain rate tensor
$\bar{\mathbf{S}}$	filtered strain rate tensor for imperfect model
$ \mathbf{S} $	magnitude of strain rate field
$ \tilde{\mathbf{S}} $	magnitude of filtered strain rate field
$ \bar{\mathbf{S}} $	magnitude of filtered strain rate field for imperfect model
$\boldsymbol{\omega}$	vorticity field
$\tilde{\boldsymbol{\omega}}$	filtered vorticity
$\bar{\boldsymbol{\omega}}$	filtered vorticity for imperfect model
$ \boldsymbol{\omega} $	vorticity magnitude
$ \tilde{\boldsymbol{\omega}} $	filtered vorticity magnitude
$ \bar{\boldsymbol{\omega}} $	magnitude of filtered vorticity field for imperfect model
$\boldsymbol{\omega}^v$	vorticity field of the target data
X	state vector
X_t	true state
X_a	analysis
X_b	background state or first guess
P_f	background error covariance matrix

B	background error covariance matrix
R	observation error covariance matrix
Q_a	model error covariance matrix
η	model error
A	analysis error covariance matrix
P_a	analysis error covariance matrix
H	interpolation operator (observation operator)
y	observation
\mathbf{H}	tangent linear operator of the observation operator
H	
\mathbf{H}^T	adjoint operator
M	model of nonlinear of NWP
\mathbf{M}^T	adjoint model containing integration backward
K	linear operator
ε_b	vector errors for the background
ε_o	observation errors
ε_a	analysis errors
P	a point
n_p	neighbours of a point
Λ	a group of grid points
ε_p	a distance value
D	embedding space dimension
\mathbb{E}	a symmetric positive matrix
\mathbf{c}	a centre of an ellipsoid
\mathcal{P}	structure of points p_i
\mathcal{G}_T	DNS target field structure
\mathcal{G}_o	reconstructed field structure

List of Abbreviations

DA	Data assimilation
NWP	Numerical weather prediction
4DVAR	Four-dimensional variational assimilation
3DVAR	Three-dimensional variational assimilation
KF	Kalman filter
EKF	Extended Kalman filter
EnKF	Ensemble Kalman filter
NS	Navier-Stokes equations
LES	Large eddy simulation
RSM	Reynolds stress models
DNS	Direct numerical simulation (target data)
OIC	Optimal initial condition based on the NS equations model
OICS	Optimal initial condition based on the imperfect model
ICs	Initial conditions
BCs	Boundary conditions
MVEEs	Minimum volume enclosing ellipsoids
OI	optimal interpolation

Chapter 1

Introduction

Turbulence is the fluid motion state distinguished by chaotic and seemingly random three-dimensional vorticity. Usually, the effect of turbulence predominates all other phenomena of the flow, where it affects the increased energy dissipation, heat transfer, drag, and mixing. Real turbulence occurs only when having three-dimensional (3D) vorticity, and one of the fundamental reasons for this is that only in such flow, is the vorticity stretching through the flow field itself achievable.

An overview of the chapter

As an outline for this chapter, in section 1.1, a brief review is presented of homogeneous turbulence which is the main flow field investigated in the thesis. In section 1.2, the data assimilation (DA) concepts are presented as well as the related methods and the most important terms being used. Section 1.3 presents the literature review which shows the most significant applications of DA in fluid mechanics related to our approach. Lastly, the outline of the dissertation is presented in 1.4.

§ 1.1 Homogeneous turbulence and large eddy simulation

Homogeneous turbulence is a type of turbulence in which the average characteristics are position-independent, i.e. a type of turbulence that is invariant with the transition.

The turbulent motion that happens in engineering or in nature is always inhomogeneous. The statistics in such turbulence are different with the location, and the fluctuating velocity changes with the location as well. Interaction consequently happens between the fluctuating and mean components of the motion. However, it is helpful to study the homogeneous turbulent flow that avoids such interaction, which provides the understanding of a simple flow that helps to understand inhomogeneous turbulent flows.

According to [69], the equations that determine the turbulent velocity include the continuity equation. For an incompressible fluid, it is given as:

$$\nabla \cdot \mathbf{u} = 0, \quad (1.1)$$

where \mathbf{u} represents the velocity field in turbulent flows at a position in the fluid identified by \mathbf{x} , and ∇ is the gradient operator with respect to the coordinate system \mathbf{x} . The Navier Stokes (NS) equations describe the viscous fluid motion which are given as:

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\rho} \nabla P + \nu \nabla^2 \mathbf{u} + \mathbf{f}, \quad (1.2)$$

where the pressure is denoted by P , the density is denoted by ρ , \mathbf{f} is the force term, and ν is the kinematic viscosity.

The NS equations in Fourier space are often used when investigating the homogeneous turbulence. The Fourier mode of $\mathbf{u}(\mathbf{x}, t)$ is written as $\hat{\mathbf{u}}(\mathbf{k}, t)$, and the evolution equation for $\hat{\mathbf{u}}(\mathbf{k}, t)$ is:

$$\left(\frac{d}{dt} + \nu k^2 \right) \hat{u}_j(\mathbf{k}, t) = -ik_l P_{jk}(\mathbf{k}) \sum_{\mathbf{k}'} \hat{u}_k(\mathbf{k}', t) \hat{u}_l(\mathbf{k} - \mathbf{k}', t), \quad (1.3)$$

where $P_{jk}(\mathbf{k})$ represents the projection tensor and the left hand side includes $\hat{\mathbf{u}}$ only at \mathbf{k} , while the right hand side includes $\hat{\mathbf{u}}$ at \mathbf{k}' and \mathbf{k}'' , in which $\mathbf{k} = \mathbf{k}' + \mathbf{k}''$ ([73]).

The nonlinear term on the right-hand-side of Eq. 1.3 leads to energy exchange between different Fourier modes, a process known as the energy cascade. In turbulence, the energy cascade transfers energy from the larger to smaller scales of the motion. The famous poem presented by the meteorologist and mathematician L. F. Richardson has summarized this concept, as quoted in [5]:

Big whorls have little whorls,
Which feed on their velocity;
And little whorls have lesser whorls,
And so on to viscosity.

Energy cascade leads to spectral energy distribution. The energy spectrum tensor $E_{ij}(\mathbf{k})$ is defined as

$$E_{ij}(\mathbf{k}) = \langle \hat{u}_i(\mathbf{k}) \hat{u}_j^*(\mathbf{k}) + \hat{u}_j(\mathbf{k}) \hat{u}_i^*(\mathbf{k}) \rangle / 2,$$

where the angle brackets refer to a statistical average, and $\hat{\mathbf{u}}^*$ is the complex conjugate of $\hat{\mathbf{u}}$. However, it is more common to use the energy spectrum $E(k)$ which is related to the mean turbulent kinetic energy by the following equation

$$\int_0^\infty E(k)dk = \langle |\mathbf{u}|^2 \rangle / 2,$$

where $k \equiv |\mathbf{k}|$ is the magnitude of the wave-number \mathbf{k} . Therefore, $E(k)$ represents the contribution to turbulence kinetic energy from Fourier modes with wave-numbers from k to $k + dk$. According to the Kolmogorov phenomenology, the energy spectrum $E(k)$ is given as:

$$E(k) = K_{Kol} \langle \varepsilon \rangle^{2/3} k^{-5/3}, \quad (1.4)$$

where the Kolmogorov constant is denoted by K_{Kol} and the energy dissipation rate is referred to by ε .

The method of filtering and large eddy simulation (LES) is employed in this thesis. Therefore, a brief introduction is presented for the concepts used in this method.

In LES, the larger-scale turbulent motion is directly represented, while the impacts of the smaller-scale are modelled. In terms of the computational cost, LES is placed between the Direct Numerical Simulation (DNS) (DNS is a simulation method in which NS equations are solved by resolving all the scales with the absence of the turbulent models [73]), and Reynolds Stress Equation Model (RSM) (where the Reynolds stress tensor components are directly calculated and the hypothesis of eddy viscosity is avoided [50]). Since the representation of large scale unsteady motions is explicit, LES can be assumed to be more reliable and precise than RSM.

It is well known that DNS is not applicable for the high Reynolds number (Re) flows because of the computational expense (Re is a dimensionless number which is defined as the ratio of the inertial force to the viscous force, and therefore identifies the relative significance of these two forces. It is given in the formula, $Re = ul/\nu$, where u is a characteristic velocity scale and l is the characteristic length). The computation in DNS is mostly exhausted on the motion of the smallest scales, while the anisotropy and energy are included mostly in the motion of the larger scales. In LES, the large scale motions are calculated directly, and models are applied to express the impacts of the smaller scales. Therefore, unlike DNS, LES avoids the costly computation spent on the motions of small scales.

The author in [73] has explained the conceptual steps related to the LES method. Firstly, the velocity fields $\mathbf{u}(\mathbf{x}, t)$ can be broken down by filtering into two parts, a residual component which is denoted by $\mathbf{u}'(\mathbf{x}, t)$ and a resolved component which is referred to by $\bar{\mathbf{u}}(\mathbf{x}, t)$. Physically speaking, $\bar{\mathbf{u}}$ is the motion of larger eddies. We define the operation of the filtering as:

$$\bar{\mathbf{u}}(\mathbf{x}, t) = \int G_\Delta(\mathbf{r}) \mathbf{u}(\mathbf{x} - \mathbf{r}, t) d\mathbf{r},$$

where G_Δ is the filter with Δ being the width of the filter, and a filtered value is denoted by an overline. As an example, the Gaussian filter ([41] and [76]) is given by:

$$G_\Delta(\mathbf{r}) = \left(\frac{6}{\pi\Delta^2} \right)^{1/2} \exp \left(-\frac{6}{\Delta^2} |\mathbf{r}|^2 \right).$$

Secondly, the equations for filtered velocity field $\bar{\mathbf{u}}(\mathbf{x}, t)$ can be found by applying the filter to the governing equations. Therefore, the continuity equation becomes:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0. \quad (1.5)$$

We obtained the following when we filter the NS equations:

$$\frac{\partial \bar{u}_i}{\partial t} + \overline{u_j \frac{\partial u_i}{\partial x_j}} = 2\nu \frac{\partial \bar{S}_{ij}}{\partial x_j} - \frac{1}{\rho} \frac{\partial \bar{P}}{\partial x_i} + \bar{f}_i, \quad (1.6)$$

where \bar{S}_{ij} is the filtered strain rate tensor and \bar{f}_i is the filtered force term. The equations can be re-written to include the subgrid-scale stress tensor τ_{ij} that appears due to the motions of the residual component. Together with the anisotropic SGS stress tensor τ_{ij}^r and modified filtered pressure \bar{p} , they are defined as the following:

$$\tau_{ij} = (\overline{u_i u_j} - \bar{u}_i \bar{u}_j), \quad (1.7)$$

$$\tau_{ij}^r = \tau_{ij} - \frac{1}{3} \tau_{kk}^R \delta_{ij}, \quad (1.8)$$

$$\bar{p} = \frac{1}{\rho} \bar{P} + \frac{1}{3} \tau_{kk}^R. \quad (1.9)$$

Hence, we obtain the filtered momentum equation as:

$$\frac{\partial \bar{u}_i}{\partial t} + \bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = 2\nu \frac{\partial \bar{S}_{ij}}{\partial x_j} - \frac{\partial \tau_{ij}^r}{\partial x_j} - \frac{\partial \bar{p}}{\partial x_i} + \bar{f}_i \quad (1.10)$$

or

$$\partial_t \bar{u}_i + \bar{u}_j \partial_j \bar{u}_i = \partial_j (2\nu \bar{S}_{ij} - \tau_{ij}^r) - \partial_i \bar{p} + \bar{f}_i \quad (1.11)$$

Therefore, models for τ_{ij}^r are needed to close the equations and hence solve the modified filtered pressure $\bar{p}(\mathbf{x}, t)$ and the filtered velocity field $\bar{\mathbf{u}}(\mathbf{x}, t)$. The Smagorinsky model is one of those models, which is named after Joseph Smagorinsky, the American meteorologist who presented his model in 1963. The Smagorinsky model is the most simple SGS model, which has been confirmed to work reasonably well [73]. In this model, the anisotropic residual-stress τ_{ij}^r is given by

$$\tau_{ij}^r = -2\nu_r \bar{S}_{ij}. \quad (1.12)$$

The subgrid-scale eddy viscosity ν_r represents the effects of the residual motions. It is modeled as [73], [76]:

$$\nu_r = (C_s \Delta)^2 (2\bar{S}_{lk} \bar{S}_{lk})^{1/2}. \quad (1.13)$$

Here, we have the Smagorinsky coefficient C_s and the filter width Δ . Finally, we can write the filtered NS equations as

$$\partial_t \bar{u}_i + \bar{u}_j \partial_j \bar{u}_i = 2\partial_j \left((\nu + (C_s \Delta)^2 (2\bar{S}_{lk} \bar{S}_{lk})^{1/2}) \bar{S}_{ij} \right) - \partial_i \bar{p} + \bar{f}_i, \quad i = 1, 2, 3. \quad (1.14)$$

The model has some drawbacks, as summarized in [40]. $C_s = 0.16$ is used in the simulations in this thesis.

§ 1.2 The basic concepts and methods of data assimilation

The method to use data from observations or experiments to enhance the predictions of the computational model is called Data Assimilation (DA). It has been used, e.g., to define the optimal state estimation, and to define the initial conditions (ICs) of the prediction model. Errors in real world measurements could appear due to the lack of accuracy of the measurement tools and the inappropriate choice of the measurement location. DA is a systematic method to deal with these issues.

DA is particularly useful in complex chaotic systems. The basis of this complexity is that slight variation in the ICs may lead to substantial variations in the accuracy of the forecast. DA has the information from an adequate prior prediction and the current measurements and uses them to find improved predictions.

The fundamental approaches available to DA, according to [8], are:

- Intermittent (non-sequential) assimilation, in which the observations at a later time may be applied to improve prediction at a previous time. The observations are treated in small batches.
- Continuous (sequential) assimilation, which takes into account observation performed from the previous to the current time. This method also considers observation batches taken for longer times, and the improvement to the enhanced prediction happens continuously in time.

The combination of the measurement/target data and the “a priori” estimate/prediction led to the assimilation of the observations. This process gives an updated prediction,

which is known as the analysis or the a posteriori estimate. In order to reduce the error for the a posteriori estimate, suitable weights for the combination have to be chosen. The analysis can be expressed by X_a and it is what we are seeking. The problem of finding the best analysis is called the analysis problem.

Several concepts are important in the mathematical formalization of the analysis problem, which are defined as below:

- The true state is denoted by X_t and it is the true description of the reality at the analysis time.
- The background state, i.e. the a priori estimate, is denoted by X_b . It represents the prediction of the true state before applying DA.
- The analysis problem is to discover an analysis increment $\delta X \equiv X_a - X_b$ such that X_t and X_a are close to each other.
- Observations may be considered as a function of X_t . These observations are gathered into vector y , possibly perturbed by errors. The observation operator is denoted by H .
- Departures are the differences given by $y - H(X)$. This is also called the analysis residual $y - H(X_a)$ when the calculation is conducted with the analysis; and innovation, when it is calculated with the background, i.e., $y - H(X_b)$.

In the atmospheric sciences, the numerical weather prediction (NWP) models are used to characterize the atmospheric dynamical behaviour. NWP is an initial value problem [68]. Therefore, there is a need for the ICs to resemble the current state of the weather. As a result, there has been a long history of using DA in NWP, as presented, for instance, in [19] and [43]. Richardson is represented as the founder of this research, as mentioned in [75]. The early calculation in [12] relied on hand interpolations, as mentioned in [43] and [72]. [30] as well as [21] and [2] made big efforts to develop automatic objective analysis. [3] and [21] utilised the prior data to complement, to some extent, imperfect information. A series of papers in DA in oceanography and meteorology are presented in [38]. [29] sums up the state of the DA in the 1990's. Recent reviews can be found in [8] and [68]. We briefly introduce two of the most important methods.

1.2.1 OPTIMAL INTERPOLATION METHODS

The optimal interpolation method illustrates the basic ideas of the Kalman filters and the variational methods. A brief review can be seen in Zupanski and Kalnay [88] along with the respected publication of Kalnay [43]. Sasaki [77] and [78] made an earlier attempt connecting optimal interpolation (OI) with the variational technique.

We want to obtain the best state estimation by merging data including the background and the observation, which may contain errors. In OI, the estimation for the analysis X_a is determined by an interpolation equation as follows:

$$X_a = K(y - H[X_b]) + X_b. \quad (1.15)$$

in which the observation operator H is assumed to be linear, and K is known as the weight matrix (the Kalman gain) for the analysis which can be found by using the covariance matrix of the analysis error.

The vector errors for the background X_b are given by $\varepsilon_b = X_b - X_t$, which represents the difference between the background and the true state, and $\langle \varepsilon_b \rangle$ represents its average. Therefore, the background error covariance B is

$$B = \left\langle (\varepsilon_b - \langle \varepsilon_b \rangle)(\varepsilon_b - \langle \varepsilon_b \rangle)^T \right\rangle.$$

Also, R refers to the observation error covariance, i.e.:

$$R = \left\langle (\varepsilon_o - \langle \varepsilon_o \rangle)(\varepsilon_o - \langle \varepsilon_o \rangle)^T \right\rangle,$$

where the observation errors $\varepsilon_o = y - H(X_t)$ and $\langle \varepsilon_o \rangle$ refers to its average.

According to Eq. 1.15, the analysis error covariance matrix

$$A = \left\langle (\varepsilon_a - \langle \varepsilon_a \rangle)(\varepsilon_a - \langle \varepsilon_a \rangle)^T \right\rangle$$

is given by

$$A = KRK^{-1} + (I - KH)B(I - KH)^T, \quad (1.16)$$

where the analysis errors is ε_a and $\langle \varepsilon_a \rangle$ represents its average. By taking the trace of the Eq. 1.16, we obtain

$$Tr(A) = Tr(B) + Tr(KHBH^T K^T) - Tr(BH^T K^T) - Tr(KHB) + Tr(KRK^T). \quad (1.17)$$

K is chosen to minimize $Tr(A)$. The differentiation d_K of Eq. 1.17, with respect to K , can be expressed as follows:

$$\begin{aligned} d_K[Tr(A)] &= d_K \left[Tr(KHBH^T K^T) - Tr(BH^T K^T) - Tr(KHB) + Tr(KRK^T) \right], \\ &= \left(KHB^T H^T + KHBH^T - 2BH^T + KR + KR^T \right), \\ &= 2 \left(KHBH^T - BH^T \right) + 2KR, \\ &= 2K(HBH^T + R) - 2BH^T. \end{aligned} \quad (1.18)$$

Letting $d_K = 0$, we find

$$K = BH^T(HBH^T + R)^{-1}, \quad (1.19)$$

which is the optimal least-squares gain. With K given by Eq. 1.19, A turns into

$$A = B(I - KH), \quad (1.20)$$

and Eq. 1.15 is called the optimal interpolation. It can be shown that the optimal interpolation can be obtained as a solution of the variational optimization problem, with cost function $J(X)$ given by

$$J(X) = (y - H(X))^T R^{-1}(y - H(X)) + (X - X_b)^T B^{-1}(X - X_b). \quad (1.21)$$

The variational optimization method is briefly reviewed next.

1.2.2 VARIATIONAL OPTIMIZATION

The work of Euler and Lagrange is considered the beginning of variational methods, according to [68]. In the 1750s, Leonhard Euler and Joseph-Louis Lagrange developed the Euler-Lagrange equation. It gives a way to find the solution that extremizes a specified cost functional. It is widely applied to optimisation problems and trajectories calculation. The calculus of variation has had a broad attractiveness because of its capability to deduce the governing equation for an entire system without the details of the system ingredients. The general theory is interpreted by the works of [6], [18] and [49]. The Lagrange multipliers method is the basis of the theory. Generally, constrained optimization problems comprise minimizing or maximizing a multivariable function $f(x)$ whose input has any number of dimensions, subject to the constraint $g(x) = 0$.

By applying the first-order conditions of minimization, we calculate the first derivatives of the Lagrangian function L which is given by:

$$L(x, \lambda) = f(x) + \lambda g(x). \quad (1.22)$$

Hence, we have

$$\nabla_x L(x, \lambda) = \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x}, \quad (1.23)$$

and

$$\nabla_\lambda L(x, \lambda) = g(x). \quad (1.24)$$

Here the Lagrange multiplier is denoted by λ which measures the sensitivity of $f(x)$ at a fixed point to the differences in the constraint. The minimum is found at the critical points where $\nabla_x L = \nabla_\lambda L = 0$.

The three-dimensional variational assimilation

The first researchers on the adjoint DA were Lewis and Derber [54], Courtier [20], and Le Dimet and Tala-grand [51]. The fully nonlinear case was extended from adjoint sensitivity analysis by [11] and [10], while [67] presented in detail the procedure used to enforce integral invariants conservation.

The works of Sasaki are considered as examples of a standard implementation of variational techniques, see [81], [79] and [80], and could be seen as a three-dimensional variational assimilation (3DVAR). [66] applied it to impose the total energy conservation, total mass and total enstrophy in the 1D and 2D equations of shallow water. The idea of 3DVAR is to seek analysis as an approximate solution to a minimisation problem (Eq. 1.21) and to avoid computation of the gain K in Eq. 1.19. [58] showed that OI and 3DVAR are equivalent as long as the assumption of the cost function J is given as follows:

$$J = [y - H(X)]^T R^{-1} [y - H(X)] + [X - X_b]^T B^{-1} [X - X_b]. \quad (1.25)$$

The first term gives the distance of prediction X to observations y and the second term represents the distance from prediction to X_b . Therefore, the algorithm of 3DVAR considers the three space dimensions, where the observations are distributed at one time.

The four-dimensional variational assimilation

In [24], it has been shown that the basis for the four-dimensional variational assimilation (4DVAR) method is optimal control, as the problem is merely a constrained optimization. The assimilations of the data are done temporally and spatially within the model by the minimization of the variation between the data and the prediction of the model. 4DVAR has been given the name because the data assimilated are the time sequences of spatial data of three-dimensions. The aim of the 4DVAR is to find the estimate of the true state of the system (analysis), consistent with both observations distribution in time and the system dynamics.

The 4DVAR is a straightforward generalization of 3DVAR for data distributed in time. It uses similar equations, provided H is generalized to include a forecast model that can be used to compare X and y at the appropriate time. Its formula can be defined by the following cost function:

$$J(X) = (X - X_b)^T B^{-1} (X - X_b) + \sum_{i=0}^N (H_i(X_i) - y_i)^T R_i^{-1} (H_i(X_i) - y_i), \quad (1.26)$$

where H_i , X_i and y_i refer to the observation operator at the time i , model state at the time i and the observation at the time i , respectively.

The problem of the 4DVAR can be defined as minimising Eq. 1.26, subject to the constraint in which the model state sequence X_i has to be a solution of the numerical model:

$$X_i = M_{0 \rightarrow i}(X),$$

in which $M_{0 \rightarrow i}(X)$ represents the model prediction from the initial time to the i th time step. Therefore, 4DVAR is a non-linearly constrained optimising problem that is usually hard to solve.

The prediction X_i is the integration of a numerical forecast model starting with X as the IC. Letting $X_0 = X$ such that M_0 is the identity, then by denoting M_i the prediction step from $i - 1$ to i , we have $X_i = M_i X_{i-1}$ and by recurrence

$$X_i = M_i M_{i-1} \dots M_1 X.$$

The minimisation of J requires an optimisation algorithm which requires, at each iteration, calculating both J and ∇J . The gradient can be calculated by running the adjoint model backward, where the adjoint can be found as the transpose of the tangent linear model. Therefore, to calculate the cost function, it is needed to run the nonlinear model to calculate all X_i . While to calculate the gradient of J , we need one more run of the adjoint model (backward in time)

1.2.3 KALMAN FILTER

Kalman filter (KF) is an optimal estimation algorithm that is named after Rudolf Kalman (1930-2016), who led the development of its theory. It predicts the parameters of interest such as direction, speed, and location while accounting for noise in their measurement. KF is commonly applied to situations such as control system; navigation; guidance; signal processing; and computer vision system. An early application of the KF was in the 1960s, when it was applied in the Apollo programme.

KF is utilised to find optimal estimates of variables of interest when it is complicated to measure them directly, but indirect measurements are available. It can also be applied to obtain the best estimate of the state of a system by the combination of measurements from a variety of sensors exposed to noise.

The KF is an adaption of the optimal interpolation method in section 1.2.1 in the framework of sequential DA, where every background is supplied via a prediction that begins from the former analysis. The background is adjusted to assimilate the real-time observations to form a prediction. The KF combines the prediction and measurement to compute an unbiased estimate with minimum variance.

We assume a linear model (M) and a linear observation operator, allowing for model error (i.e. imperfect model). The model may be written as

$$X_t(i+1) = M(i)X_t(i) + \eta(i),$$

where η represents the error in the model, which is a random variable with zero mean and the covariance $Q_a(i) = \langle \eta(i)\eta(i)^T \rangle$. The notation i refers to the time step.

At each time step, the analysis of the previous time will be considered the background to base the forecast on. An estimation of the uncertainty of the analysis at this time is given by an uncertainty forecasting model. The same procedure moves forward in time until all observations are assimilated. Hence, for each time, it needs to calculate the analysis and analysis uncertainties and forecast them both in order to get the prior and the uncertainties for the next time. The gain matrix at time i is calculated as:

$$K(i) = P_f(i)H^T(i) \left(H(i)P_f(i)H^T(i) + R(i) \right)^{-1},$$

where P_f represents the error covariance matrix of the background. The analysis at the time we are at is given by

$$X_a(i) = X_b(i) + K(i) \left(y(i) - H(i)X_b(i) \right).$$

The error covariance of analysis is given by

$$P_a(i) = \left(I - K(i)H(i) \right) P_f(i).$$

Then, we obtain the next prior and the next prior uncertainty. They are given by the following equations:

State forecast:

$$X_b(i+1) = M(i)X_a(i). \tag{1.27}$$

Error covariance forecast:

$$P_f(i+1) = M(i)P_a(i)M^T(i) + Q_a(i), \tag{1.28}$$

where $Q_a(i)$ is the error covariance matrices of the computer model.

Extended Kalman Filter

It is possible to generalize the KF algorithm to nonlinear operators, leading to the extended Kalman filter (EKF). If H is nonlinear, let \mathbf{H} be its tangent linear operator in the neighbourhood of X_b . Likewise, if M is nonlinear, \mathbf{M} denotes the tangent linear prediction model in the neighbourhood of X_a . Using \mathbf{M} and \mathbf{H} in the formulas for KF, one obtains the extended KF.

Ensemble Kalman filter

The propagation of the error covariance matrix in KF and EKF schemes is very expensive to compute. Besides, if the nonlinearities of the model are significant, one may need to apply more general stochastic forecast schemes. Examples include an ensemble prediction scheme, identified as the Ensemble Kalman Filter (EnKF).

We consider here the perturbed observation ensemble Kalman filter. In this version of the EnKF, the error covariance forecast P_f is approximated by an ensemble of model runs. The solution of Eq. 1.28 is thus avoided, leading to potentially significant saving. In practice, the prediction step of EnKF evolve each analyses member using the nonlinear model

$$X_b^{(j)}(i) = M\left(X_a^{(j)}(i)\right) + \eta(i), \quad \text{for } j = 1, N, \quad (1.29)$$

where $\eta \sim \mathcal{N}(0, Q_a)$, and it generates ensemble of size N given X_b , and Q_a . We have an ensemble of states $X_a^{(j)}$, and (j) indicates the ensemble member. Therefore, we run the nonlinear model to obtain an ensemble of forecasts $X_b^{(j)}$, then form the ensemble mean

$$\langle X_b(i) \rangle = \frac{1}{N} \sum X_b^{(j)}(i). \quad (1.30)$$

Therefore, the sample covariance of the forecasts is:

$$P_f = \frac{1}{N-1} \sum \left(X_b^{(j)}(i) - \langle X_b \rangle \right) \left(X_b^{(j)}(i) - \langle X_b \rangle \right)^T, \quad (1.31)$$

where we subtract the mean from each of the ensemble members such that the covariance matrix is given in Eq. 1.31.

Next, the analysis step updates each ensemble member using the perturbed observations as follows:

$$X_a^{(j)}(i) = X_b^{(j)}(i) + K \left(y(i) + \epsilon_y - H X_b^{(j)}(i) \right), \quad (1.32)$$

where $K = P_f H^T (H P_f H^T + R)^{-1}$, and $\epsilon_y \sim \mathcal{N}(0, R)$. Perturbing the observation is necessary to ensure correct error statistics. Therefore, for each ensemble member, we have a different perturbation to the observation to ensure that the analysis error covariance depends on the forecast covariance according to

$$P_a(i) = \left(I - K(i)H(i) \right) P_f(i).$$

The analysis error covariance matrix can be calculated as the mean of this updated ensemble

$$X_a(i) = \frac{1}{N-1} \sum X_b^j(i)$$

$$P_a = \frac{1}{N-1} \sum (X_a^j(i) - X_a)(X_a^j(i) - X_a)^T. \quad (1.33)$$

More discussion can be found in [26].

1.2.4 COMPARISON BETWEEN 4DVAR AND ENKF

This thesis studies only one of the DA techniques, the 4DVAR. However, it is worth shedding light on the similarities and differences between, and advantages of, the 4DVAR techniques and Kalman filter. In terms of real-time assimilation, 4DVAR over a short period is a workable approach.

Many similarities as well as significant differences between two approaches have been previously mentioned in [27], [38], [39] and [44]:

- 4DVAR can assimilate asynchronous observations.
- The performance of 4DVAR converges to the full rank extended Kalman filter when applied over a long enough window.
- 4DVAR may runs for assimilation in a realistic NWP frame due to the fact that the computational procedure is cheaper than EnKF and KF.
- 4DVAR is superior than EnKF or KF inside the optimisation time interval as it utilises all the observations at the same time. This makes it a smoother, rather than a sequential estimator.
- 4DVAR is operated for a finite-time period, particularly if the dynamical model was nonlinear, while in principle EnKF may run continually.
- 4DVAR itself does not give a background error covariance estimation. A specific process for that purpose is needed, that requires a cost as much as operating the equivalent extended Kalman filter.

Some principal advantages of EnKF are shown below:

- It is relatively simple to implement and it is model independent.
- EnKF filters out fast processes such as convection in an automatic manner, through the nonlinear saturation.
- EnKF gives optimal initial ensemble disturbances as they represent the analysis error covariance matrix.

The 4DVAR technique needs the maintenance and development for the adjoint model, which can be considered as one of its disadvantages. The low-dimensionality of the ensemble of EnKF is considered an essential disadvantage due to the sampling errors in the estimations of the background error covariance matrices.

§ 1.3 The applications of DA in fluid mechanics

As mentioned in the previous section, research into numerical weather prediction is an example where there has been a long history of using DA, as presented, for instance, in [19] and [43]. Therefore, numerous sophisticated DA methods have been developed. 4DVAR methods and EnKF are considered, among others, the most common approaches; for more details see [43], [42] and [26].

Both 4DVAR and EnKF have been used in fluid dynamics recently, as reviewed in [34]. The paper presents a review of studies in which investigations are made into procedures for combining measurement and numerical simulation like the 4DVAR and Kalman filters. The 4DVAR might partially find a solution to the problem where the ICs and boundary conditions (BCs) are the control parameters. Creating a numerical representation of the characteristics of unstable flow is also resolved by Kalman filters by modifying mathematical models to stabilise the original dynamical system's unstable modes with feedback signals.

The variational method has been applied by [31] to construct the initial condition and the inflow condition. The procedure is concerned with the correct modelling and the reproduction of spatio-temporal dynamics for an experimentally examined flow. Two-dimensional (2D) mixing layers and wake flows behind a cylinder were investigated. The research concluded that the 4DVAR DA used was validated by the experiments within a controlled framework and demonstrated the reliability and feasibility of the procedure. The data used in the research was from a real world application and therefore the method has the potential to be used in computations of complex external flows thus reducing the computational high costs in the near-wall region.

2D wake flows have also been considered by [63], where the authors presented a comparison between various DA methods. Three different schemes are investigated within the NS equations framework: 4DVAR; ensemble Kalman smoother (EnKS) which may

be considered as the extension of EnKF; and four-dimensional ensemble-based variation (4DEnVAR) which merges the application of a Monte Carlo ensemble with the 4DVAR technique, without the need for adjoint operators. The outcomes of these techniques show that they can be successfully applied for the rebuilding of ICs and BCs for 2D unstable flows around a cylinder in the presence of coherent gusts at $Re = 100$. In these configurations, the 4DVAR is accomplished with the lowest error between observations and estimation for a fixed computational cost, based on the given observation and prior statistics. The 4DVAR produced a significant reduction in error. Whilst the EnKS and 4DEnVAR were not as successful, the reconstruction of the target trajectory could be satisfactory and give a comparable computational cost as 4DVAR. However, the EnKS and 4DEnVar schemes depend on ensembles the sizes of which are broadly smaller than the control vector dimension. Also, the adjoint procedure has an advantage, namely the adjoint variables give the first-order sensitivity with respect to the control variables. Both 4DEnVAR and EnKS procedures are also applied iteratively, and it is demonstrated that the outcomes are better than those obtained with only one-loop of the algorithm while the same total number of direct runs are required. The EnKS can be more sensitive to observation and past statistics than the 4DEnVAR and 4DVAR. However, in the situation when the appropriate previous statistic is given, an advantage of the EnKS is that it supplies posterior information in addition to the assimilated state. The 4DEnVAR procedure, by contrast, can be applied in a deterministic framework, showing itself to be less sensitive to both the ensemble statistics and its size than the EnKS.

The EnKF procedure has used in [45], [46], and [57] to enhance the prediction of Reynolds-Averaged Navier-Stokes (RANS) models.

In [45], the EnKF process is used. A pseudo experiment is conducted and the pseudo experimental values are created from the computed flow of the flat plate boundary layer. It considers the Spalart-Allmaras model and the EnKF is used to estimate the values of the model parameters, in which the parameters are combined with the state variables to accomplish such estimation. These parameters are considered the uncertainty factors in the simulation. The Spalart-Allmaras (SA) model is a one equation model with an equation for the kinematic eddy turbulent viscosity [82]. Six parameters are used to represent the turbulent effects, which are traditionally given as follows: $C_{b1} = 0.1355$; $\sigma = 2/3$; $C_{b2} = 0.622$; $C_{\omega2} = 0.3$; $C_{\omega3} = 2$; $C_{v1} = 7.1$. The fast aerodynamic routine (FaSTAR) developed by the Japan Aerospace Exploration Agency is applied, and the EnKF is performed in FaSTAR. The results for the model show that the original values given by the SA model are precise in the statistical perspective, thus demonstrating the EnKF effectiveness to define the parameters' values in a model statistically or objectively.

The researchers in [46] consider complex aeronautics. The ensemble transform Kalman Filter (ETKF) is one of the extended Kalman Filter methods. The ETKF is used in this study as a method which overall reduces cost in ensemble based DA techniques. The ETKF is applied to 2D transonic flows around the RAE 2822 airfoil, and 3D transonic

flows around the ONERA M6 wing. To show how effective this approach would be, three factors with uncertainty, namely, the angle of attack; Mach number; and the turbulent viscosity, were estimated in the two test cases. The ETKF showed itself to be a more effective method than the SA and Shear Stress Transport (SST) models. (The SST is a two-equation eddy viscosity turbulence model using the $k - \epsilon$ and $k - \omega$ turbulence models [61]). The ETKF approach with the three factors of uncertainty, provided more satisfactory corrected values than either the SA or the SST.

In [57] the authors defined a new data-driven adaptive computational model, for imperfect measurement data. By using the compressible RANS $k - \omega$ model and a data driven algorithm, they proposed the D-DARK model. The D-DARK model is tested in three validation cases: flow around an airfoil; the backward-facing step; and pipe flow. In the simulation experiment, it was observed that the D-DARK model adapted the $k - \omega$ model coefficients to improve the accuracy of CFD.

The application of the EnKF method in the detached eddy simulation (DES) and LES is introduced in [60], where a given sequential estimator is integrated into the separated solvers for the incompressible flow. An augmented solver is supplied by this method that integrates observations in the CFD model. Two procedures of model reduction are considered, to decrease the model computational expenses that show 10%-15% improvement with respect to the traditional numerical simulation. Simplified formulas representing the accuracy of the numerical procedure are improved. The estimator was used to compute a number of flow fields. The first one is a 2D flow around a cylinder when $Re = 100$, while the second case considers a 3D simulation flow around a circular cylinder. The outcomes demonstrated that the augmented solver affects positively the forecast of the tested physical quantities. Finally, it shows that these DA methods have the possibility in the numerical simulation to use the augmented forecast for the optimisation as a robust device for the free variables.

The authors of [28] introduce the mathematical framework of the 4DVAR DA method for the recapture of the mean velocity fields that are solved from the RANS equations. Procedures that seek a reconstruction of the state from limited data are needed because measurements in fluid mechanics may cover partial information of the total state of the flow. The DA framework is discussed with the 2D flow around an infinite cylinder at $Re = 150$ as a test case. After examining a number of measures on the flow, the DA process is used to build the whole mean flow field. Spatial interpolation, extrapolation, the reconstruction of the mean velocity ingredients from measurements of the magnitude only, and noise filtering were tested independently. The mean velocity field exhibits satisfying results for all cases, to the extent that it is more accurate than simpler approaches which rely on high order interpolation which ignores physical limitations on the mean flow.

In [22], [1] and [74], reduced order models based on Galerkin truncation are combined with the variational techniques.

The work of [22] is about the capability of a variational DA method to extract a proper orthogonal decomposition (POD) based Galerkin low-order dynamical systems (LODS)

from the particle image velocimetry (PIV) measurements. A variational framework is used to estimate the variables that characterise the flow, such as velocity components or density or pressure, to improve the first initial estimation gained by polynomial identification. It uses a least squares estimation of the control variable with respect to the full measurements sequence usable in a set time range. The control variable can be set to the IC or to any parameter for the evolution law. The investigation chooses a wake flow as an experiment case. The outcome shows that the POD assimilation method considerably outperforms the least-squares fitting techniques. The POD assimilation procedure does not give the coefficients that characterise the LODS. However, those coefficients may be successfully recaptured by polynomial identification on the assimilated outcomes. Therefore, the method seems to be a valuable tool to rectify noisy experimental data.

In order to reduce the deficiency of polynomial identification of the dynamics from experimental data, [1] developed a scheme of variational DA with the initial state of the POD modes and the model coefficients as the controls. The assimilation is conducted on the basis of POD modes computed from a sequence of snapshots of the velocity fields. The authors evaluate two variational DA methods, which involve different control variables. One of them presents a weak dynamical model expressed only up to an additional time-dependent noise function which represents a time-dependent control variable on the dynamics. This uncertainty function is considered as part of the cost function which is designed to improve the IC to a priori value, up to Gaussian noise. The other model deals with a strong dynamical constraint in which the control variables are the coefficients of the dynamical system. It aims to look for the control that minimises a cost function encoding the errors between the unknown coefficients subject to the constraint set by the dynamics and the measurement errors. The rebuilt dynamical model demonstrated it can largely recapture the dynamics of a large number of modes for short experimental and numerical sequences.

In [74], the authors suggest a variational technique to optimize non-linear eddy viscosity (EV) for POD models. The EV is presented as an unknown function of the resolved energy. The optimization is achieved by using the 4DVAR method, POD models with the optimal EV are built for three shear flows. Firstly, an investigation is used to assess the two dimensional POD model for the transient behaviour in the 2D cylinder wake. Secondly, the 20 modes POD model of the 2D mixing layer at Reynolds number 500 with velocity ratio three is studied. Finally, the 100 modes POD model of the wake of a 3D Ahmed body at Reynolds number 300000 is built. The formulation leads to some important physical understanding regarding the EV ansatz. The 20 modes mixing layer model shows a negative EV when the fluctuations are low. Also, an accurate energy distribution can be obtained by the 100 modes wake models.

The 4DVAR methods have been applied in state estimation in the flow control context by the research in [7], which handles the exact state reconstruction problems and the estimation of the approximate state from wall measurement in a wall-bounded incompressible unstable flow. It presents a strategy for defining a turbulent channel flow at

time t , based only on the wall skin friction and pressure measurements obtainable in a neighbourhood of time t , regardless of the ICs at some time $t_0 < t$.

The extended Kalman filter procedure is studied in [14], based on a linearized model. It discusses estimating the turbulent channel flow at $Re_\tau = 100$ from noisy measurements on the wall. It shows that, through defining a suitable second-order statistic in a fully nonlinear DNS of the system, then combining the information of the statistic in the linear feedback gains computation, an effective estimator may be constructed depending on the availability of measurements at the wall.

The EnKF method is used in [15]. The authors discuss the estimation of turbulent channel flow based on wall information only, which is considered the key issue in model-based flow control. Only low-Reynolds number flows are considered. Its outcome reveals the law of the wall, in which statistics are given as a function of the distance from the wall.

The variational methods are applied to extrapolate the data from experiments in [35]. The research aims to direct attention to the capability of such procedures in integrating with correlation based PIV techniques, to enhance the consistency of the simulation of flow fields.

The EnKF technique is investigated by [16] with the joint use of a Kinect depth sensor and a stochastic DA method to recapture free surface flow fields from a sequence of images. The Kinect sensor captures observations of surfaces with wavelengths and amplitudes small enough for applications such as monitoring flow or characterising medium to large scale flows.

Variational DA is applied to obtain the optimal placement of sensors in [64]. The technique is used to the reconstruction of unstable 2D flows past a rotationally oscillating cylinder at $Re = 100$, with a view to maximise the DA method's efficiency. It is aimed at the a priori design of a sensor mesh and depends on a first order adjoint technique. The goal is to identify the flow areas in which the sensitivities are the biggest regarding the change in control variables such as the ICs, BCs, or model parameters. Two steps are taken in the investigation. First, preliminary experiments for the DA method are made to identify the possible solutions of inverse problems in the unstable wake flow fields. The outcomes show that despite the fact that the DA method is able to fit the obtainable measurements, there is a difficulty for recovering the IC and the cylinder rotational speed in the non-lock-on system. Second, the observation optimization process is applied for the optimal placement of velocity sensors downstream of the cylinder, which makes it possible to improve the DA's functionalities. In addition, it is used to find the most essential sensors at a selected mesh. Optimal sensor placement leads to a fall in computational cost compared to a regular observation mesh and improved the quality of the reconstruction.

The 4DVAR has also been used to examine the decaying isotropic turbulent flow, by using the eddy-damped quasi-normal Morkovian (EDQNM) model in [62].

§ 1.4 The objective and the outline of the dissertation

The mentioned research has shown the usefulness of DA in the simulations of turbulence and in improving its physical understanding. As illustrated in NWP research, DA has an unparalleled capability in the enhancement of the instantaneous prediction or modelling of turbulent flows, as well as improving the statistics. However, few investigations have explored this feature in the three-dimensional turbulent flow simulation. [87] has studied the capability of DA in recovering small scales in the field of the three-dimensional isotropic turbulence instantaneously, as has [48]. The data presented as the Fourier modes in this research, which replace the predictions of the NS equations at each time step, directly. It has shown that the velocity field for the instantaneously small scale can be asymptotically produced precisely when the Fourier modes with the wave-number up to the threshold at $0.2\eta_K^{-1}$ are given, where η_K is the length scale of Kolmogorov.

However, the methods of variation or EnKF have not been considered; many questions remain to be answered. The thesis attempts to answer some of these questions. The fully developed Kolmogorov flow in a three-dimensional periodic box is considered in this thesis. The assumption that a time sequence for the velocity data is provided on a set of grid points has been made. The 4DVAR is used for reconstructing the initial velocity field such that this velocity at later times agrees with the velocity of the given data. The velocity field time sequences obtained from the initial velocity field are compared with the known velocity field. The goal is to understand how well the small scale velocity fields are rebuilt instantaneously with 4DVAR. The three-dimensional turbulent Kolmogorov flow in a periodic box is one of the simplest flows where the nonlinear inter-scale interactions have the controlling role in the dynamical process. Consequently, despite the similarity between the mathematical problems of two-dimensions and three-dimensions, considerable new challenges are presented by the latter.

The application of the 4DVAR technique is one of the central contributions of this thesis. To assess the reconstruction procedures, it is essential to quantify the variation instantaneously between the target and the rebuilt fields. Important information can be discovered from the statistics of point-wise correlation. Nevertheless, such statistics are not the most suitable for capturing the geometry for the non-local structures that fill the small scales of turbulence (the filaments of the vortex as an example). The lack of understanding becomes an obstacle when the distribution for the non-local structures becomes one of our most meaningful subjects. The morphology of the non-local structures has long been described qualitatively by visualizations. Many efforts have been made recently to improve techniques for the description quantitatively.

The curvatures of the structure surface as its signature are calculated in [4]. [86] illustrates the geometry in a channel flow in terms of curvelets and angular spectra. In [52], indices named shapefinders, determined in terms of Minkowski functionals, are applied for the classification of the structures. These techniques provide a huge

amount of detail about the descriptions of the structures, quantitatively. Nevertheless, they mainly focus on the intrinsic geometry. The data of the locations, orientations and the structure sizes are absent, but they are important if we discuss the geometrical comparisons between two different fields. Hence, for the geometrical description of a nonlocal structure, we utilise minimum volume enclosing ellipsoids (MVEEs). The estimation of statistics, analysis of the cluster and image processing are some examples for the wide uses of MVEEs ([84] and [9]). The application of MVEEs in research into turbulent flows has not been investigated. The outcomes show that MVEEs are a helpful tool for non-local geometrical analysis in turbulent flows.

The non-local structure studies allow us to capture the geometry of these structures in the small scales of turbulence. In what follows, the analyses of the non-local structures of these quantities are presented: vorticity magnitude; strain rate tensor magnitude; subgrid-scale dissipation, which gives a full characterization of the small scales in the flow fields.

The structures of \mathbf{u} can be represented by the velocity gradient $A_{ij} \equiv \partial_j u_i$. The strain rate tensor describes the rate of shearing and stretching as [23], [71] and [73]. It is given by the following formula:

$$S_{ij} = \frac{1}{2} (A_{ij} + A_{ji}).$$

Vorticity plays an essential function in self-sustaining the turbulent energy cascade process. Vorticity is visualized intuitively as a three-dimensional filamentary body with spatially concentrated combinations of strong vorticity called vortex filaments ([13] and [47]). The vorticity field is denoted as $\boldsymbol{\omega}$, and defined as follows:

$$\nabla \times \mathbf{u} \equiv \boldsymbol{\omega} \Leftrightarrow \omega_i \equiv \epsilon_{ijk} A_{kj}.$$

The subgrid-scale (SGS) dissipation of turbulent kinetic energy is described in [36], [37] and [83], which is proportional to the product between the SGS stress tensor τ_{ij} and strain rate S_{ij} . Specifically, the SGS kinetic energy dissipation is given as

$$\Pi = -\tau_{ij} S_{ij}.$$

The SGS energy dissipation describes the kinetic energy rate which is transferred from resolved to the SGS motions. Therefore, it is considered important for the characterizations of τ_{ij} .

The reconstruction of the velocity fields using 4DVAR depends on the assumption we make about the dynamics of the system. In Chapter 2 of the thesis, we investigate the reconstruction using the NS equations as the model. The main focus is to reconstruct the instantaneous distributions for the small scales. The target data fields are used to compare the quality of the optimal initial fields. The vorticity field, the strain rate tensor field, and the subgrid-scale energy dissipation fields are investigated. Several

geometrical statistics are studied, including the PDFs of the alignment between the filtered vorticity and the eigenvectors of the filtered strain rate tensor; the PDFs of velocity increments; in addition to other well-known statistics. These statistics illustrate how different scales interact with each other, which is the key feature of turbulent flows. They enable us to examine the quality of the reconstruction and how far we can recapture the statistics at the target data fields. Also, the investigation of the DNS of the Gaussian random field with the same energy spectrum from the optimal initial field is presented. The investigation tests the ability of the reconstructed fields as an initial condition to recover the statistics in the DNS target fields. Finally, MVEEs is used to assess the non-local structures reconstruction quantitatively.

Obviously, it is interesting to investigate the ability to perform the reconstruction using simplified models, because, for example, it might be much more efficient to do so. Therefore, in Chapter 3, we present the results of such a study with the filtered NS equations as the model for the data. The optimality system is derived. Similar analyses are performed and compared with those in Chapter 2.

The main conclusion and the outlook are summarized in Chapter 4. The Appendix includes the derivations of the optimality system for the 4DVAR with the NS equations as a model as well as the adjoint equations for the filtered NS equations with standard Smagorinsky model.

Chapter 2

Turbulent fields reconstruction using 4D variational data assimilation with the Navier-Stokes equations as the model

§ 2.1 Introduction

We look into a turbulent system where incomplete velocity measurement data over a time period are available. Assuming that the dynamical evolution of the system is governed by the Navier-Stokes equation, we use the 4D variational DA, to reconstruct the initial state of the system, that the evolution of the system matches the measurement data. We formulate the problem as an optimization problem, where the initial field is taken as the control variable. The goal is to find the optimal control variable to minimize the differences between the measurement and the velocity field evolved from the reconstructed initial field, subject to the constraint imposed by the NS equation. The reconstructed field is compared with DNS data. We look into the difference between the instantaneous variations of the reconstructed and the DNS field. The instantaneous high vorticity structures, the filtered strain rate tensor, and the subgrid-scale energy dissipation for the optimal initial condition fields are compared with the DNS field.

The motivation for the study is two-fold. By building turbulence, we can learn its internal mechanisms, which is not fully understood. Also, the re-constructed turbulent field can be used as initial and inlet boundary conditions for numerical simulations,

which have crucial effects on the quality of the simulations.

In this chapter, the Navier-Stokes equations as the model is presented to examine the effects on the reconstruction quality. The formulation of the optimization problem is given in section 2.2 and the iteration procedure is presented in section 2.3. To assess the quality of the reconstruction, section 2.4 presents different investigations that examine different statistics of the reconstructed fields. Also, the assessment of the reconstruction quality of the non-local structures is presented in different investigations. Furthermore, we consider in subsection 2.4.9 a quantitative description of the structures using the minimum volume enclosing ellipsoid which enables us to define the location, dimensions, and orientation of the structure bounded by the MVEE.

§ 2.2 Formulation

We assume the following is known in the problem:

- A set of data $\mathbf{v}(\mathbf{x}, t)$ is known, for $t \in [0, T]$ and $\mathbf{x} \in \Omega$, where T is the optimization horizon.
- \mathbf{v} is a part of the solution of the NS equations for unknown IC $\mathbf{u}(\mathbf{x}, 0) = \boldsymbol{\varphi}(\mathbf{x})$.
 \mathbf{u} is a solution of the NS equations.

The method we used here is 4D variational DA (i. e. flow optimization), in which we find $\boldsymbol{\varphi}(\mathbf{x})$ to minimize the cost function given by

$$J = \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} |\mathbf{u} - \mathbf{v}|^2, \quad (2.1)$$

subject to constraints:

$$\mathbf{N}(\mathbf{u}) = \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nu \nabla^2 \mathbf{u} - \mathbf{f} = 0, \nabla \cdot \mathbf{u} = 0, \mathbf{u}(\mathbf{x}, 0) = \boldsymbol{\varphi}(\mathbf{x}). \quad (2.2)$$

The assimilation procedure is formulated as a constrained optimization problem. A control variable then has to be chosen which will be optimized in order to identify the minimum of a cost function which measures how far the control variable deviates from measurements. The control parameter $\boldsymbol{\varphi}(\mathbf{x})$ here represents the initial condition. We monitor the difference between the simulated and the measured velocity data by the cost function.

The optimization problem is solved with the adjoint method. The Lagrangian function is given as

$$L = J - \int_0^T \int_B \boldsymbol{\xi} \cdot \mathbf{N}(\mathbf{u}) dt d\mathbf{x} - \int_0^T \int_B \mu \nabla \cdot \mathbf{u} dt d\mathbf{x} - \int_B \boldsymbol{\lambda} \cdot [\mathbf{u}(\mathbf{x}, 0) - \boldsymbol{\varphi}(\mathbf{x})] d\mathbf{x}, \quad (2.3)$$

where $\boldsymbol{\xi}(\mathbf{x}, t)$, $\mu(\mathbf{x})$, and $\boldsymbol{\lambda}(\mathbf{x})$ are the adjoint variables corresponding to the constraints. B represents the $[0, 2\pi]^3$ periodic box.

The state equations are given by

$$\frac{\partial L}{\partial \boldsymbol{\xi}} = 0, \frac{\partial L}{\partial \mu} = 0, \frac{\partial L}{\partial \boldsymbol{\lambda}} = 0. \quad (2.4)$$

The adjoint equations are given by

$$\frac{\partial L}{\partial \mathbf{u}} = 0, \frac{\partial L}{\partial p} = 0. \quad (2.5)$$

Therefore, the state equations are

$$\begin{aligned} \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nu \nabla^2 \mathbf{u} - \mathbf{f} &= 0, \\ \nabla \cdot \mathbf{u} &= 0, \\ \mathbf{u}(\mathbf{x}, t = 0) &= \boldsymbol{\varphi}(\mathbf{x}). \end{aligned} \quad (2.6)$$

The adjoint equations are

$$\begin{aligned} \partial_t \boldsymbol{\xi} + \mathbf{u} \cdot \nabla \boldsymbol{\xi} - \nabla \mathbf{u} \cdot \boldsymbol{\xi} + \nabla \mu + \nu \nabla^2 \boldsymbol{\xi} &= \mathbf{F}, \\ \nabla \cdot \boldsymbol{\xi} &= 0, \\ \boldsymbol{\xi}(\mathbf{x}, t = T) &= 0, \end{aligned} \quad (2.7)$$

where $\mathbf{F}(\mathbf{x}, t) = -[\mathbf{u}(\mathbf{x}, t) - \mathbf{v}(\mathbf{x}, t)]$ when $\mathbf{x} \in \Omega$ and $\mathbf{F} = 0$ otherwise. Taking into account that $\boldsymbol{\xi}$ represents the velocity adjoint and μ is the pressure adjoint.

The gradient of the cost function is

$$\frac{DJ}{D\boldsymbol{\varphi}} = \boldsymbol{\xi}(\mathbf{x}, 0), \quad (2.8)$$

which is given by the adjoint. It is needed to update the initial condition $\boldsymbol{\varphi}$.

§ 2.3 Iterative Solution

The iteration procedure is started from an estimate $\boldsymbol{\varphi} = \boldsymbol{\varphi}^0$, where $\boldsymbol{\varphi}^0$ is the initial guess for $\boldsymbol{\varphi}$, and a tolerance e . We then repeat the following for $i = 0, 1, 2, \dots$:

- Integrate the state equation from $t = 0$ to $t = T$.

- Calculate the cost function $J^i \equiv J(\mathbf{u}^i)$. Exit if $J^i < e$, otherwise continue to the next step.
- Integrate the adjoint equation from $t = T$ to $t = 0$, hence find $\frac{DJ}{D\varphi^i}$.
- Use $\frac{DJ}{D\varphi^i}$ to update φ^i , and then continue from step 1.

In the last step above, a nonlinear Conjugate-Gradient method with the Brents line search algorithm is used to update φ .

§ 2.4 Simulation and discussions

The optimal φ is the reconstructed/rebuilt initial condition (OIC). An interesting question is how well φ mimics real turbulence. Therefore, the statistics of φ are compared with those in turbulent fields. In the comparisons, we refer to φ as the OIC field, while we refer to the DNS field as the DNS target field.

Before starting the analysis of the results, we need a description of the simulations that we use in the following sections in this chapter except where indicated.

Kinetic energy is injected into the flow field by using the force in NS equations which is given as $\mathbf{f} = (0, f_a \cos k_f x_1, 0)$ where $f_a = 0.15$ and $k_f = 1$, see Eq. 2.6. A set of velocity data $\mathbf{v}(\mathbf{x}, t)$ is known, for $t \in [0, T]$. Where $\hat{\mathbf{v}}(\mathbf{k}, t) = 0$ for $|\mathbf{k}| \geq K_c$, $K_c = 4$. This means that only the low wave-number Fourier modes are utilised in the optimisation problem. Here, K_c is an index of the spatial resolution of the target data. Therefore, the cost function J can be presented in term of the Fourier mode as

$$J = \frac{1}{2} \int_0^T dt \int_{k \leq K_c} d\mathbf{k} (\hat{\mathbf{u}}(\mathbf{k}, t) - \hat{\mathbf{v}}(\mathbf{k}, t)) \cdot (\hat{\mathbf{u}}^*(\mathbf{k}, t) - \hat{\mathbf{v}}^*(\mathbf{k}, t)), \quad (2.9)$$

where $k = |\mathbf{k}|$, and $\hat{\mathbf{u}}(\mathbf{k}, t)$ and $\hat{\mathbf{v}}(\mathbf{k}, t)$ denote the Fourier modes of $\mathbf{u}(\mathbf{x}, t)$ and $\mathbf{v}(\mathbf{x}, t)$, respectively. Allowing $\hat{\mathbf{F}}(\mathbf{k}, t)$ be the Fourier transform of $\mathbf{F}(\mathbf{x}, t)$, we obtain $\hat{\mathbf{F}}(\mathbf{k}, t) = -[\hat{\mathbf{u}}(\mathbf{k}, t) - \hat{\mathbf{v}}(\mathbf{k}, t)]$ when $k \leq K_c$ and $\hat{\mathbf{F}}(\mathbf{k}, t) = 0$ otherwise.

We use a three-dimensional pseudo-spectral code for isotropic turbulence where the boundary conditions are periodic in a domain $[0, 2\pi]^3$. Optimum is found when $J/J(\mathbf{u} = 0) \leq e$, where tolerance $e = 0.0002$. The number of grid points N^3 is 128^3 . We use the kinematic viscosity $\nu = 0.006$. The optimization horizon is up to $T = 0.4\tau_L$, where $\tau_L \approx 4.35$ is the large eddy turn-over time scale. The highest resolved wave-number is $k_{max} = N/2 = 64$, and the grid size is $\delta_x = (2\pi)/N = \pi/64$. The filter scale is $8\delta_x$.

2.4.1 THE ENERGY SPECTRUM $E(k)$

We look into the energy spectrum in this subsection, the energy spectrum for the DNS field and the OIC fields are shown as follows

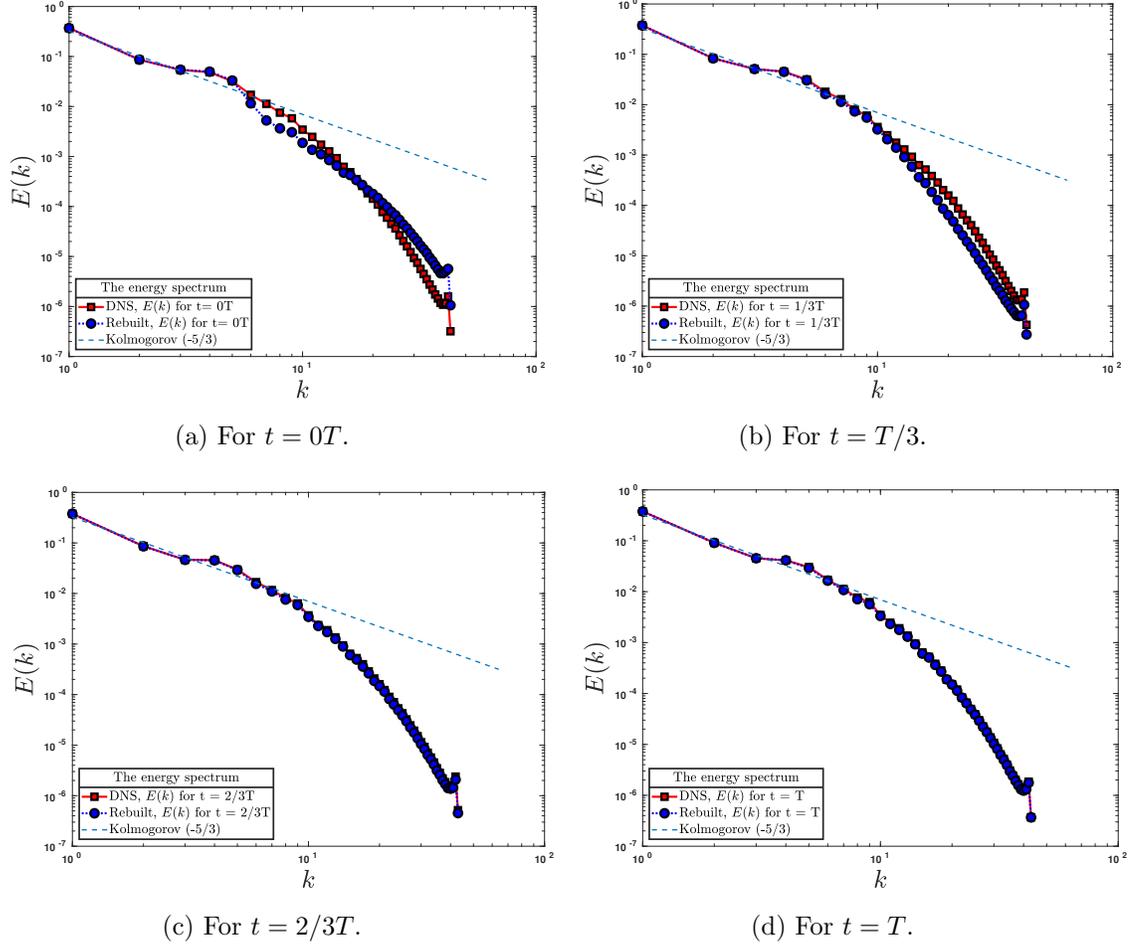


Figure 2.1: The energy spectrum. Solid line with squares: DNS target. Dotted line with circles: OIC. Dashed line: Kolmogorov $(-5/3)$.

Figs 2.1 plots the energy spectrum of the DNS target and the OIC fields. As we can see, the error between the fields, especially at the small scales and the high wavenumbers where the dissipative region are located, is big at $t = 0T$. Then the agreement between the fields improves from the initial time at $T/3$. The error between the target data energy and the energy of the OIC field disappears when the time equals $2/3T$. We notice the same agreement of the fields, at the final time when $t = T$. Another observation is the perfect agreement of the energy spectrum for both fields for $k \leq 4$ during the whole time due to the fact that the large scales are matched with the target data.

2.4.2 THE ENERGY SPECTRUM $E_\delta(k)$

In this subsection, we focus on the instantaneous difference between the OIC and the DNS target fields. $E_\delta(k)$ is used to denote the energy spectrum for the difference between the two velocity fields.

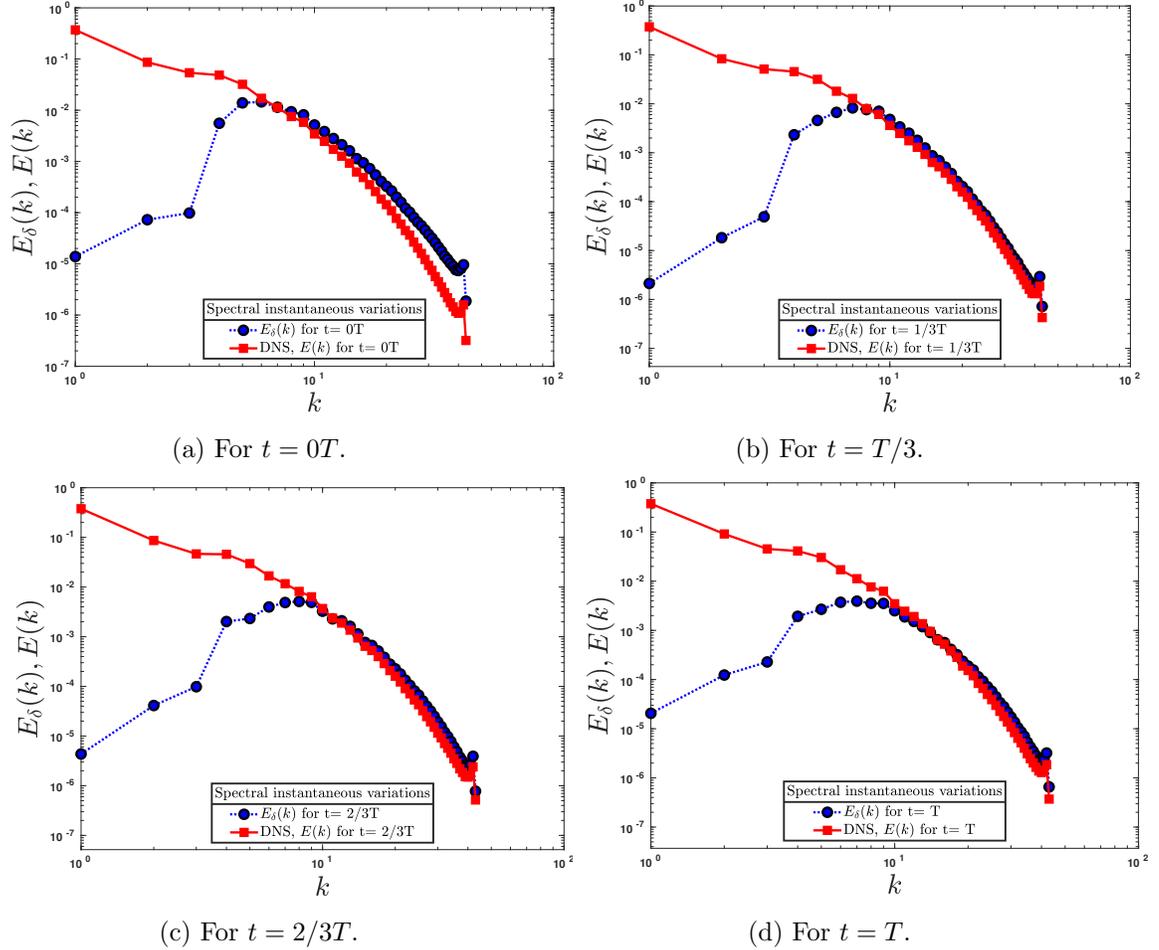


Figure 2.2: Spectrum of $\mathbf{v} - \mathbf{u}$. Solid line with squares: $E(k)$ for DNS target. Dotted line with circles: $E_\delta(k)$.

The blue curves in Fig. 2.2 represent the energy spectra of $\mathbf{v} - \mathbf{u}$, while the red ones are the energy spectra from the target field \mathbf{v} . As we can notice, the energy spectra of the instantaneous differences for $k \leq K_c$ (i.e. large scales) are small. This is a consequence of minimizing the cost function such that at these wave-numbers \mathbf{u} mimic \mathbf{v} quite well. The value of $E_\delta(k)$ when the time increases to $T/3$ is lower than the value at the initial time for the large scales, which means that differences between the velocities are even less than before. Such a perfect match between the reconstruction fields and the DNS target field continues at $2/3T$, where the spectra of $E_\delta(k)$ remains extremely

small with compare to the $E(k)$ of the target field in the low wave-number region until the end when $t = T$. Interestingly, we observe that the spectra of the instantaneous differences for $k > K_c$ are affected as well, where the agreement between both fields enhances with time. This means that \mathbf{u} tends to close to \mathbf{v} for the high wave-numbers and the reconstruction of the small scales is progresses over time. Therefore, the results of this experiment consequently demonstrate the effects of the DA method because they show that the reproduction of the scales at different sizes in the \mathbf{u} field is feasible.

Fig. 2.3 calculates the ratio between the energy spectrum for the difference between the OIC fields and the DNS target field, $E_\delta(k)$, and the energy spectrum of the DNS target field $E(k)$.

The ratio calculations cover several time steps in the optimization horizon to show the development with time.

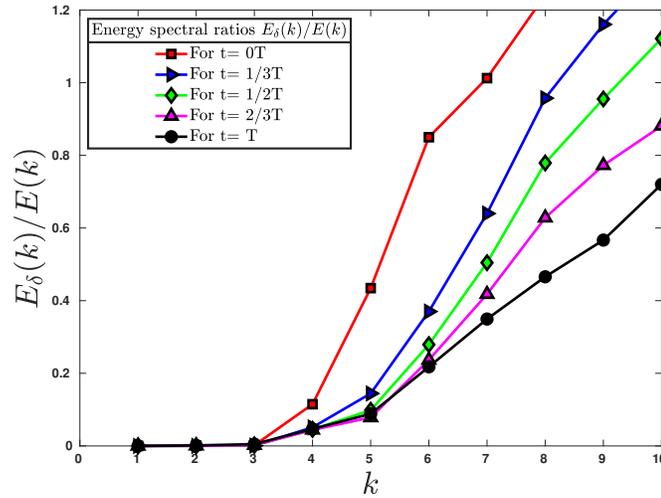


Figure 2.3: The ratio of the spectra: $E_\delta(k)/E(k)$. Solid line with: Square for $0T$. Downward-pointing triangle for $T/3$. Diamond for $T/2$. Upward-pointing triangle for $2/3T$. Circle for T .

We can see that the OIC field tends to the corresponding target field as time increases. These positive observations are shown in Fig. 2.3 which covers the whole time within the optimization horizon. The order of the curves shows that as time increases, the ratio reduces and reaches the minimum at the end. However, the ratios when $k = 8$ actually show that the agreements between Fourier modes are not very good because Fourier modes are affected by the whole flow field. The ratio at $k = 8$ reaches about 45%. Later analysis will show improved agreement can be obtained for local flow field parameters.

2.4.3 THE PROBABILITY DENSITY FUNCTIONS OF THE ALIGNMENT BETWEEN THE VORTICITY AND RATE OF THE STRAIN TENSOR

In this subsection, we present a number of results that involved with the local deformation of the flow. Starting with the most significant geometric statistics, alignment between the filtered strain rate field \tilde{S}_{ij} and the filtered vorticity field $\tilde{\omega}$.

The results for two of the three eigenvectors are presented. The missing one is the extensional eigenvector; we focus on the intermediate and the compressible eigenvectors.

We have $\alpha_{\tilde{S}}, \beta_{\tilde{S}}$ and $\gamma_{\tilde{S}}$ as the strain-rate tensor eigenvalues. Each one of the three eigenvectors is corresponding to one of the eigenvalues. The eigenvalues are defined such that $\alpha_{\tilde{S}} \geq \beta_{\tilde{S}} \geq \gamma_{\tilde{S}}$, with the sum of all of them equal zero ($\alpha_{\tilde{S}} + \beta_{\tilde{S}} + \gamma_{\tilde{S}} = 0$), which thus requires $\alpha_{\tilde{S}} \geq 0$, and $\gamma_{\tilde{S}} \leq 0$, and $\beta_{\tilde{S}}$ is either negative or positive. The names of the eigenvectors are given by their corresponding eigenvalues: the eigenvector $\gamma_{\tilde{S}}$ is the contractive, the eigenvector $\beta_{\tilde{S}}$ is the intermediate and the eigenvector $\alpha_{\tilde{S}}$ is the extensive.

The majority of the studies about the alignment showed that the extensional eigenvector has no preferred alignment with the vorticity vector although it corresponds to the biggest eigenvalue.

The vorticity vector tend to be aligned with the intermediate eigendirection, as mention in [17], [32], [33] and [70]. For that reason, we are interested in the intermediate and compressible eigenvectors which correspond to the intermediate and the smallest eigenvalue, respectively.

We present the following results to examine the possible differences between the optimal fields and the DNS target field.

The results show that the development of the alignment between the target data field and the optimal fields at different times within the optimization horizon happens gradually at an early time, as we can see in Fig. 2.4. The agreement at the initial time in Fig. 2.4a is not quite well, where the error between the curves for the intermediate eigenvectors is big. The variations between the curves for the compressible eigenvectors at the initial time also are quite big, as it starts with a lower value for the OIC fields than DNS target field at 1.4, and ends with a higher probability at 0.5 while the value of DNS target is much smaller. However, the green curves show that the differences decrease when $t = T/6$, although we still see some significant variations between fields. Fig. 2.4c shows that in the following time steps, the agreements between the fields increase, as we notice at time $T/3$ and later until the end when $t = T$.

2.4.4 THE PROBABILITY DENSITY FUNCTIONS OF THE VELOCITY INCREMENTS

The intermittency in turbulence refers to irregular, large fluctuations in the motion of the small-scales, which is reflected in the statistics of the gradients of the velocity, velocity increments, and the rate of the dissipation. Because of the intermittency of the

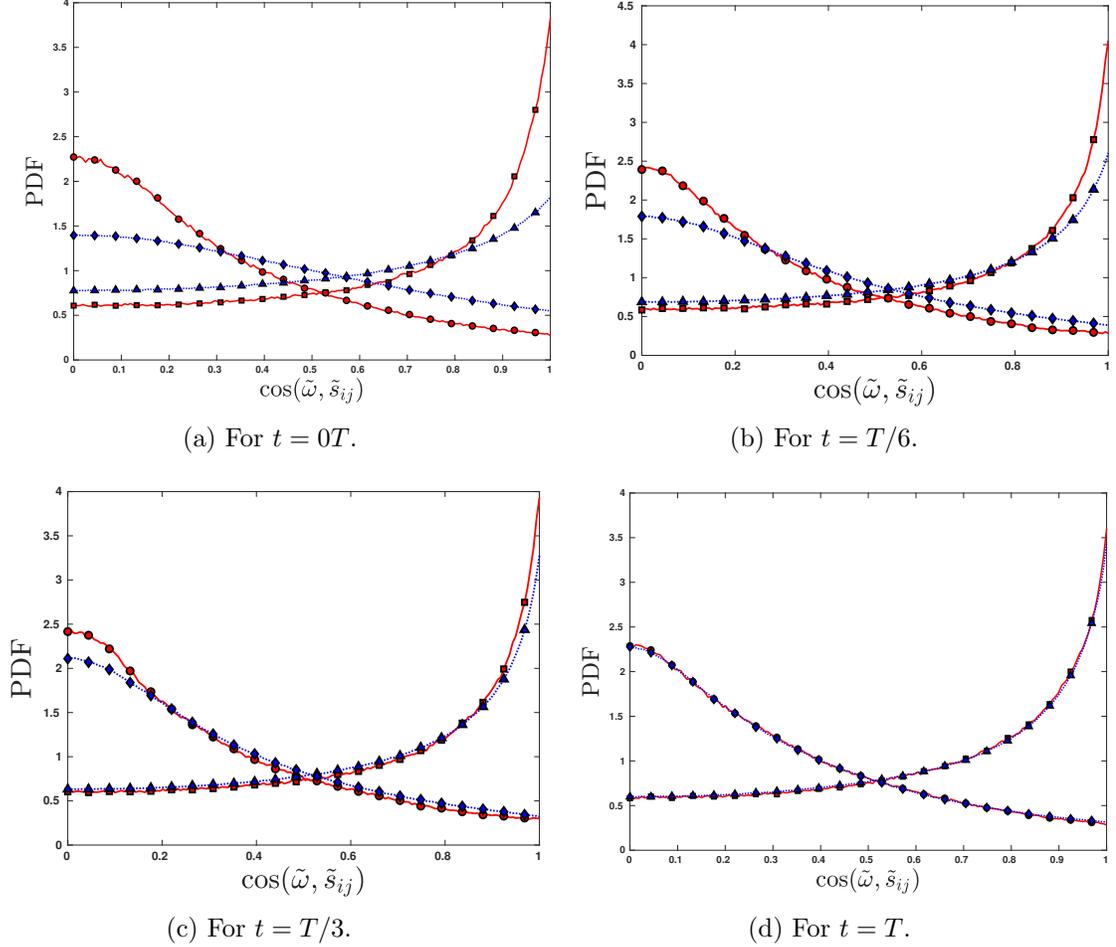


Figure 2.4: PDFs of the alignment between $\tilde{\omega}$ and the eigenvectors of \tilde{S}_{ij} . DNS target is solid line with squares, circles; OIC is dotted line with upward-pointing triangle, diamond; OIC is dashed line with asterisks, right-pointing triangle: intermediate eigen-direction and eigen-direction with negative eigenvalue, respectively.

small-scale, observations from simulations and experiments show that the probability density functions for the quantity of the small scale have exponential or stretched exponential tails. Also, the exponents of the scaling laws of the velocity increment moments significantly diverge from the prediction of the classical theory of Kolmogorov, as mentioned in [55].

We follow the original work of Kolmogorov in 1941, where the turbulent flows were investigated in the framework of Eulerian velocity increments $\delta\mathbf{u}(\mathbf{x}|\mathbf{r}, t) = \mathbf{u}(\mathbf{x} + \mathbf{r}, t) - \mathbf{u}(\mathbf{x}, t)$, as mentioned in [53] and [73].

In the isotropic and homogeneous turbulent flow, the statistics of $\delta\mathbf{u}(\mathbf{x}|\mathbf{r}, t)$ depend on the scale of the separation only $r = \|\mathbf{r}\|$, and hence, $\delta\mathbf{u}(\mathbf{x}|\mathbf{r}, t)$ can be projected onto

the directions perpendicular or parallel to $\hat{\mathbf{r}} = \mathbf{r}/r$.

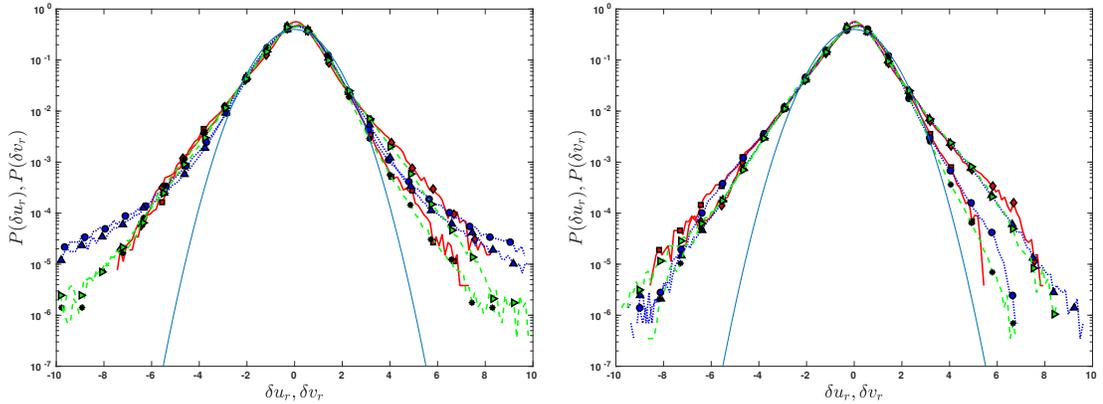
This leads to longitudinal and transverse increments, as written in equations 2.10 and 2.11, respectively:

$$\delta u_r(\mathbf{x}|r, t) = \delta \mathbf{u}(\mathbf{x}|\mathbf{r}, t) \cdot \hat{\mathbf{r}}. \quad (2.10)$$

$$\delta v_r(\mathbf{x}|r, t) = \delta \mathbf{u}(\mathbf{x}|\mathbf{r}, t) \cdot \hat{\mathbf{t}}, \quad (2.11)$$

where $\hat{\mathbf{t}}$ is a direction perpendicular to $\hat{\mathbf{r}}$.

We present in this section the PDFs of the longitudinal velocity increments and the transverse velocity increments. We look into the difference between the OIC fields and the DNS target field during the optimization horizon at different stages. Also, we add the normal distribution to show how the PDFs deviates from the Gaussian field.



(a) For $t = 0T$ (DNS/OIC), $T/3$ (OIC).

(b) For $t = T/2$ (DNS/OIC), $5T/6$ (OIC).

Figure 2.5: PDFs of velocity increments. δu_r : Longitudinal. δv_r : Transverse. Solid line is DNS target: with squares (δu_r) and diamond (δv_r). Dotted line is OIC: with circles (δu_r) and upward-pointing triangle (δv_r). Dashed line is OIC: with asterisks (δu_r) and right-pointing triangle (δv_r). The thin line is normal distribution.

From the PDFs results of the DNS field, we can see some essential observations regarding the behaviours of this kind of statistics in the real turbulence. Our aim is to see similar features in the PDFs of the OIC fields, such as the negative skewness that occurs in the longitudinal PDFs and also the transverse PDF being wider than the longitudinal PDF.

The negative skewness seems to be absent at the beginning, However, when the time equals $T/3$, the negative skewness seems more clear for the OIC fields.

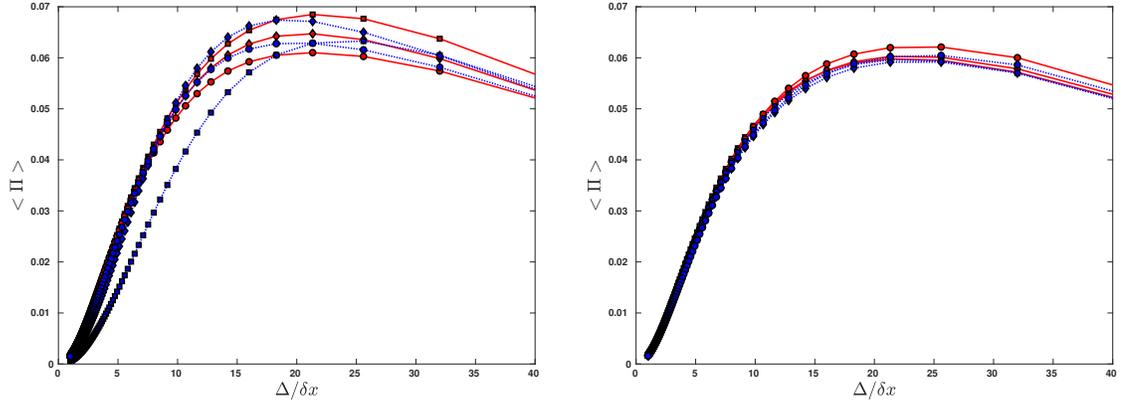
Fig. 2.5b shows that when we arrive at the mid-time step, each PDF agrees with the corresponding one of the other field. Also, some gaps still appear with PDFs behaviour.

At $t = 5T/6$ of the whole duration, both PDFs of the OIC fields tend to be quite close to the target field with minimum errors between them.

2.4.5 THE SUBGRID-SCALE ENERGY DISSIPATIONS

In this section, we conduct an analysis of the statistics of the SGS energy dissipations generated by the OIC field. Using a filtering approach, the energy flux from the large scales to the small scales is represented by the SGS energy dissipation, which provides useful insights into the local structures of the velocity field. The mean SGS energy dissipations as a function of the filter scales represent a vital statistic to understand the characteristics of the small scales of isotropic homogenous turbulence.

As we did in the previous section, we examine the SGS energy dissipation at different times during the optimization horizon between the OIC fields and the target data fields.



(a) Solid line is DNS target: with squares for $0T$, diamond for $T/6$ and circles for $T/3$. Dotted line is OIC: with squares for $0T$, diamond for $T/6$ and circles for $T/3$.

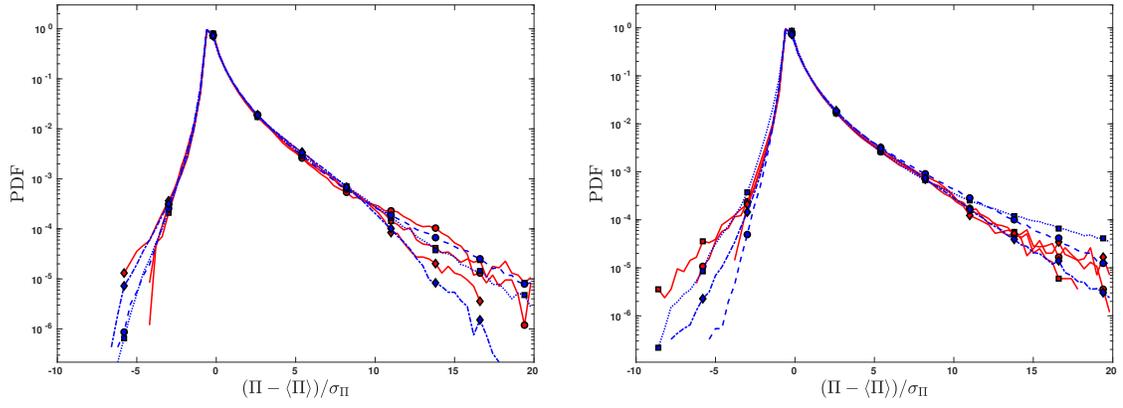
(b) Solid line is DNS target: with squares for $T/2$, diamond for $2T/3$ and circles for T . Dotted line is OIC: with squares for $T/2$, diamond for $2T/3$ and circles for T .

Figure 2.6: The mean subgrid-scale energy dissipations as a function of the filter scales Δ .

Regarding the mean SGS energy dissipation, we can see the blue curve at the initial time in Fig. 2.6a is underestimated from its corresponding red curve. However, at $T/6$ in the same figure, the energy dissipation shows a big jump compared with other times to see that the peak value reaches at 0.07, where the OIC fields values are higher than the target data field. When the time equals $T/3$, the energy dissipation slightly reduces from the peak value, but it remains close to the DNS target field, and the energy dissipation values for optimal fields still exceed the target field. At Fig. 2.6b, we display reasonable agreement between the different fields as for the times of $T/2$, $2/3T$ and T , respectively, although the average between the two fields give variations in their behaviours as the reconstructed initial fields are under-estimation when $t = T$.

Therefore, we are able to reconstruct a reasonable mean of the SGS energy dissipation which matches with the target field in a short time.

In the following probability density functions of the normalised SGS energy dissipation at a given scale are given. Also, we examine the OIC fields by studying their PDFs at different stages and find out whether they simulate the real turbulence by comparing the DNS and the OIC fields. Note that σ_{Π} stands for the root mean square value of Π .



(a) Solid line is DNS target: with squares for $0T$ and diamond for $T/6$. Dotted line is OIC: with circles for $0T$ and upward-pointing triangle for $T/6$.

(b) Solid line is DNS target: with squares for $T/3$ and diamond for $T/2$. Dotted line is OIC: with circles for $T/3$ and upward-pointing triangle for $T/2$.

Figure 2.7: PDFs of the subgrid-scale energy dissipation.

Fig. 2.7a shows that, at the initial time, the PDFs have some differences in their features where the left tails of the OIC PDFs are relatively weak. Although the agreement enhances when the time is $T/6$ as the differences are smaller, the distribution of the OIC field is still weaker, especially with the positive tail. At later times (shown in Fig. 2.7b), the PDFs of the OIC fields seem quite close to their corresponding ones at the DNS target field. Therefore, we are able to produce some features by the OIC fields for $\Delta = 8\delta_x$, such as the positive skewness in the PDFs, which is indicating the forward energy cascade, although, there are some quantitative differences.

2.4.6 THE PROBABILITY DENSITY FUNCTIONS OF THE SHAPE FACTOR S^*

To characterize the joint distributions for the three eigenvalues for the rate of the filtered strain tensor, we present a non-dimensional parameter of the eigenvalues. We have $\alpha_{\tilde{s}}, \beta_{\tilde{s}}$ and $\gamma_{\tilde{s}}$ as the strain-rate tensor eigenvalues, where $\alpha_{\tilde{s}} \geq \beta_{\tilde{s}} \geq \gamma_{\tilde{s}}$, with the condition $\alpha_{\tilde{s}} + \beta_{\tilde{s}} + \gamma_{\tilde{s}} = 0$. The non-dimensional shape factor for the filtered strain tensor S_{ij} is given by

$$S^*(\tilde{S}_{ij}) = \frac{-3\sqrt{6}\alpha_{\tilde{S}}\beta_{\tilde{S}}\gamma_{\tilde{S}}}{(\alpha_{\tilde{S}}^2 + \beta_{\tilde{S}}^2 + \gamma_{\tilde{S}}^2)^{3/2}}. \quad (2.12)$$

The usefulness of $S^*(\tilde{S}_{ij})$ comes from its ability to indicate the local motion types. For instance, $S^*(\tilde{S}_{ij}) = 1$ corresponds to axisymmetric extension (i. e., $\alpha_{\tilde{S}} = \beta_{\tilde{S}} > 0, \gamma_{\tilde{S}} < 0$), $S^*(\tilde{S}_{ij}) = 0$ corresponds planer straining (i. e., $\beta_{\tilde{S}} = 0$), and $S^*(\tilde{S}_{ij}) = -1$ corresponds to axisymmetric contraction (i. e., $\alpha_{\tilde{S}} > 0, \beta_{\tilde{S}} = \gamma_{\tilde{S}} < 0$). For incompressible flow, state parameter of the strain $S^*(\tilde{S}_{ij})$ is restricted between -1 and 1.

In the following, we study the probability density functions of $S^*(\tilde{S}_{ij})$ calculated from the reconstructed turbulent velocity fields and compare them with the probability density functions evaluated from the target velocity field. We compare these PDFs at different steps during the optimization horizon, in order to understand when the results give reasonable agreement among both fields.

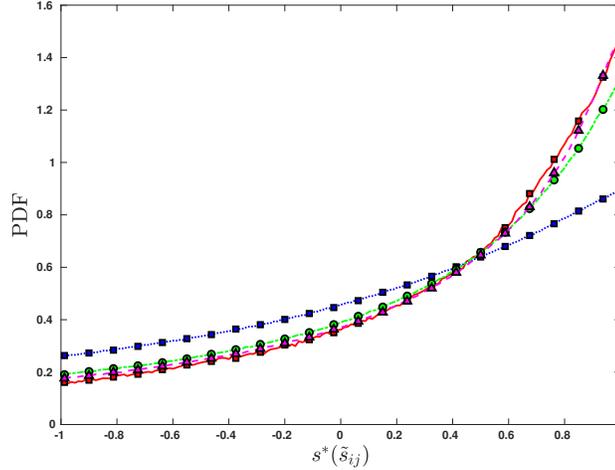


Figure 2.8: PDFs of the shape factor $S^*(\tilde{S}_{ij})$ for the filtered strain tensor \tilde{S}_{ij} . Solid line is DNS target: with squares for $0T$, diamond for $T/6$. Dotted line is OIC: with circles for $0T$ and upward-pointing triangle for $T/6$.

From the PDF shown with the blue curve in the initial time at Fig. 2.8, we see the difference with the DNS target as its probability is much lower at the peak. However, its behaviour enhances quickly as it reaches nearly the same peak with small gaps with the target data, at the time $T/6$. At the later time steps, all the PDFs behaviours of the OIC fields give quite similar behaviours of the DNS field PDFs.

Therefore, show that we are able to reconstruct PDFs of S^* in a short time, such that these PDFs are giving satisfactory agreement between both fields.

We have also $\alpha_{\tau}, \beta_{\tau}$ and γ_{τ} are the SGS stress tensor eigenvalues, where $\alpha_{\tau} \geq \beta_{\tau} \geq \gamma_{\tau}$.

The names of the eigenvectors depend on their corresponding eigenvalues: the eigenvector γ_τ is the contractive, the eigenvector β_τ is the intermediate and the eigenvector α_τ is the extensive. These statistics are mentioned in [36] [37], [59] and [83]. Note that the isotropic part of τ_{ij} , i. e. $\frac{1}{3}\delta_{ij}\tau_{kk}$ has been subtracted from τ_{ij} .

The non-dimensional shape parameter for the stress state $S^*(\tau_{ij})$ is given in the formula

$$S^*(\tau_{ij}) = \frac{-3\sqrt{6}\alpha_\tau\beta_\tau\gamma_\tau}{(\alpha_\tau^2 + \beta_\tau^2 + \gamma_\tau^2)^{3/2}}. \quad (2.13)$$

The probability density functions of the stress shape parameters are presented at Fig. 2.9.

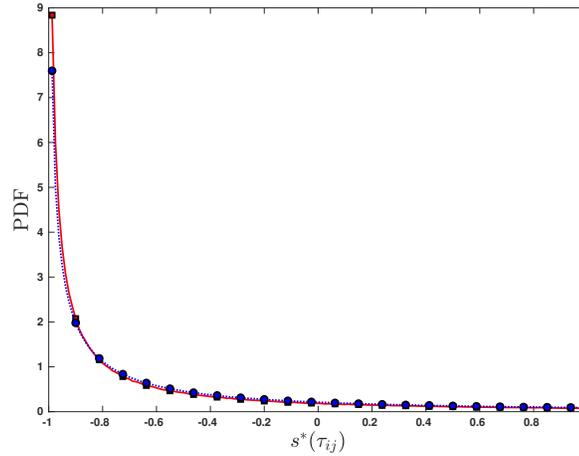


Figure 2.9: PDFs of the shape factor $S^*(\tau_{ij})$ for the subgrid-scale stress tensor τ_{ij} . Solid line is DNS target: with squares for $0T$. Dotted line is OIC: with circles for $0T$.

The PDFs of the OIC fields show very good agreements with the PDFs of the target field and at the early time as it is shown in Fig. 2.9, as their probabilities peak at -1 for all results. The PDF of the OIC field is slightly different, but the difference is very small.

2.4.7 DNS OF THE GAUSSIAN RANDOM FIELD WITH SAME ENERGY SPECTRUM FROM THE OPTIMAL INITIAL FIELD

In this section, we use the energy spectrum from the rebuilt field to generate Gaussian random (GR) fields and use them as initial conditions to generate time sequences of velocity fields. (These fields are called GR fields later). We study the new GR field and follow their changes in time within the optimization horizon and compare them with the DNS target field as well as the fields obtained from OIC. Our goal is to show using OIC may speed up the recovery of the statistics in the DNS target fields. All the

calculations have been made in a cubic periodic box with a length of 2π , 128^3 resolution and a filter of grid size equal $8\delta x$ unless stated otherwise.

The probability density functions of the alignment of vorticity with the strain rate eigenvectors, the PDFs of $\cos(\tilde{\omega}, \tilde{\lambda}_i)$, where $i = \alpha, \beta$ and γ , are displayed first.

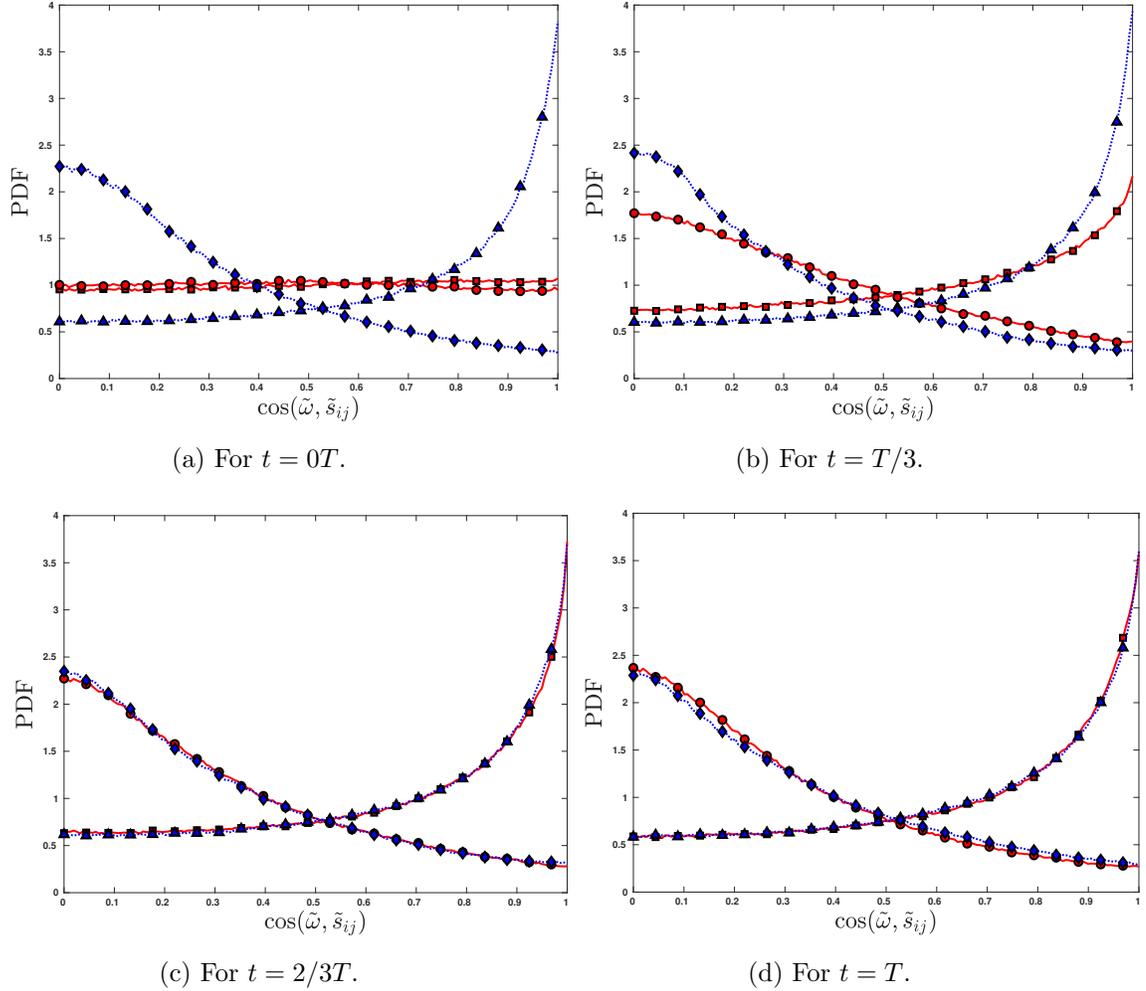


Figure 2.10: PDFs of the alignment between $\tilde{\omega}$ and the eigenvectors of \tilde{S}_{ij} . GR: solid line with squares, circles: intermediate eigen-direction and eigen-direction with negative eigenvalue, respectively. DNS target: dotted line with upward-pointing triangle, diamond, respectively.

For the GR field, at the initial time, we notice that both PDFs for the red curves tend to be flat, as is expected. The correct behaviour of the curves builds up over time. When the time reaches two-thirds of the whole duration, the results in fully developed DNS target fields have been recovered.

Comparing the GR with OIC fields, it takes longer to reproduce the DNS result for

the former. Fig. 2.4 shows that it takes only about $T/3$ to recover the fully developed DNS behaviour whereas here it takes about $2/3T$.

The probability density functions of the velocity increments are now discussed in Fig. 2.11. We observe that the PDFs of the current results reflect the differences between the GR fields and the non-Gaussian field. The initial time step clearly shows the nature of the GR field, where there is no positive nor negative skewness with the PDFs. The negative tails in the PDFs for δu_r and the flat tails in the PDFs for δv_r develop over time, as expected. Good agreement with target data appears at $t = 5/6T$.

On the other hand, same or better agreement is already obtained at $t = T/2$ for results from OIC field (see Fig. 2.5).

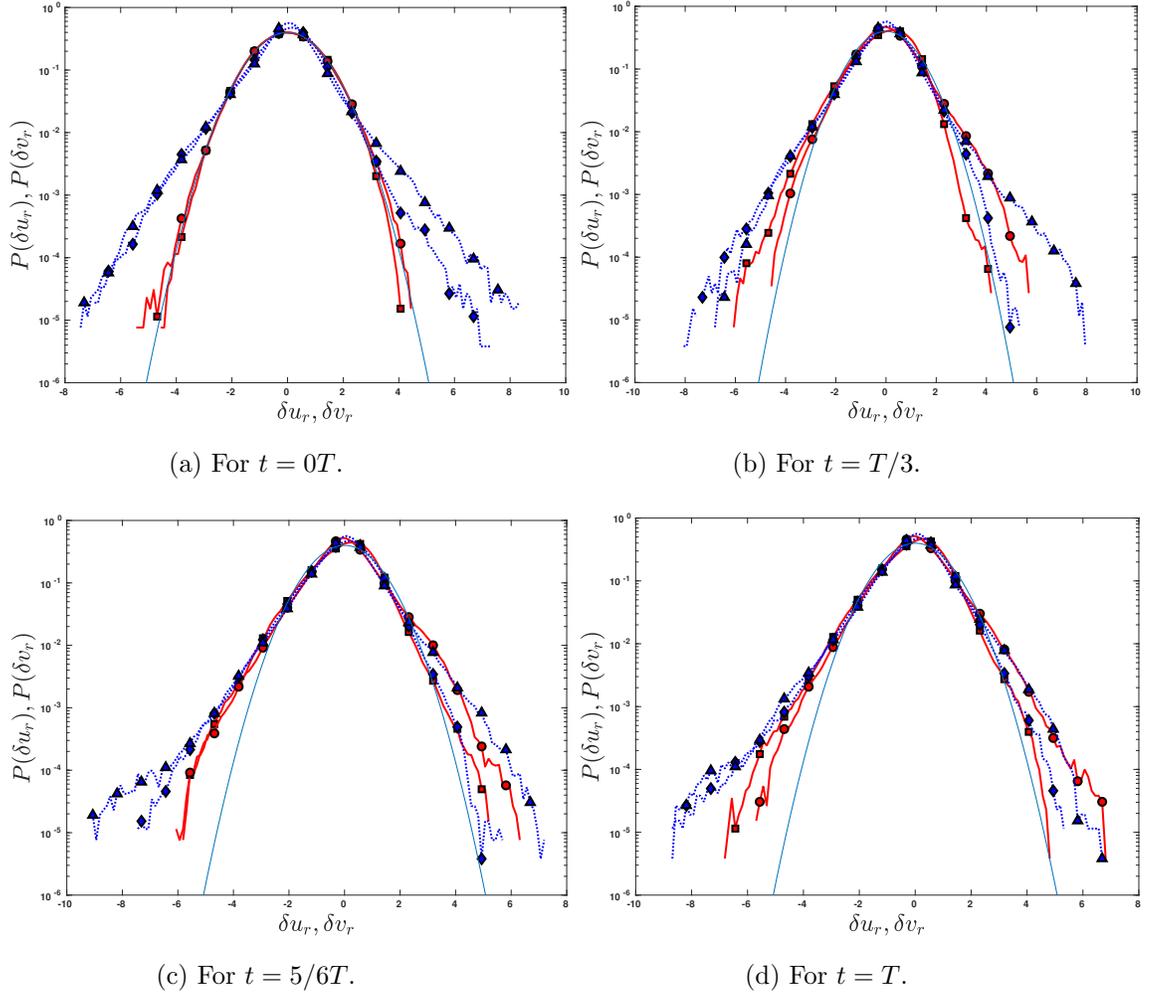
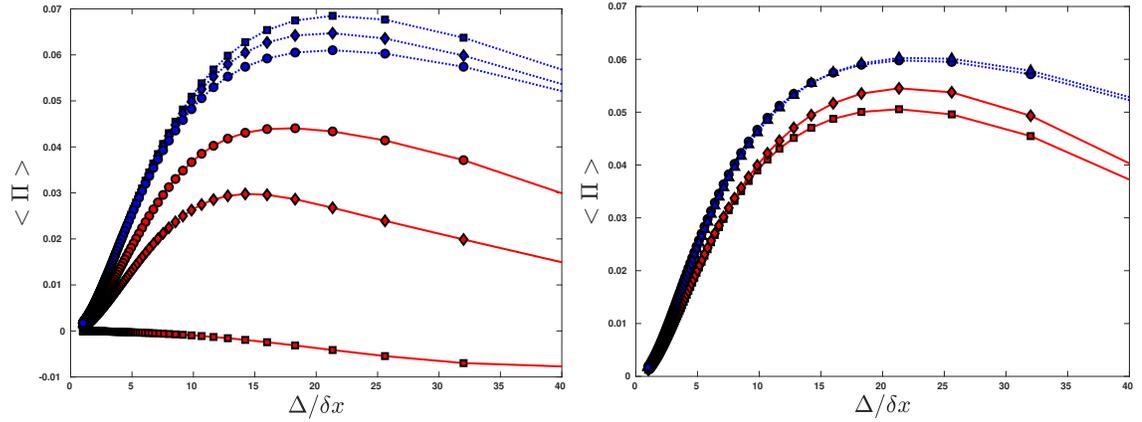


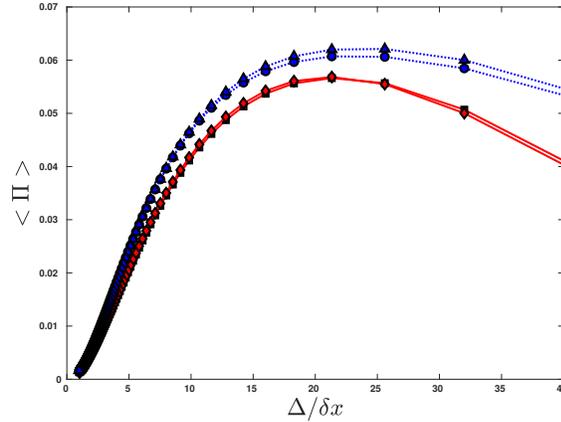
Figure 2.11: PDFs of velocity increments. δu_r : Longitudinal. δv_r : Transverse. Solid line is from GR ICs: with squares (δu_r) and Diamond (δv_r). Dotted line is DNS target: with circles (δu_r) and Upward-pointing triangle (δv_r).

The mean SGS energy dissipations as a function of the filter scales study are plotted in Fig. 2.12. From the red curve of the GR field, we can see clearly that there is no mean SGS energy dissipation at the initial time due to no energy cascade in the Gaussian field. When the time increases, energy cascade quickly builds up. However, it is interesting to observe that there is still significant underestimate even at $t = T$. This is in clear contrast with the results from the OIC field shown in Fig. 2.6.



(a) Solid line is GR: with squares for $0T$, diamond for $T/6$ and circles for $T/3$. Dotted line is DNS target: with squares for $0T$, with diamond for $T/6$ and circles for $T/3$.

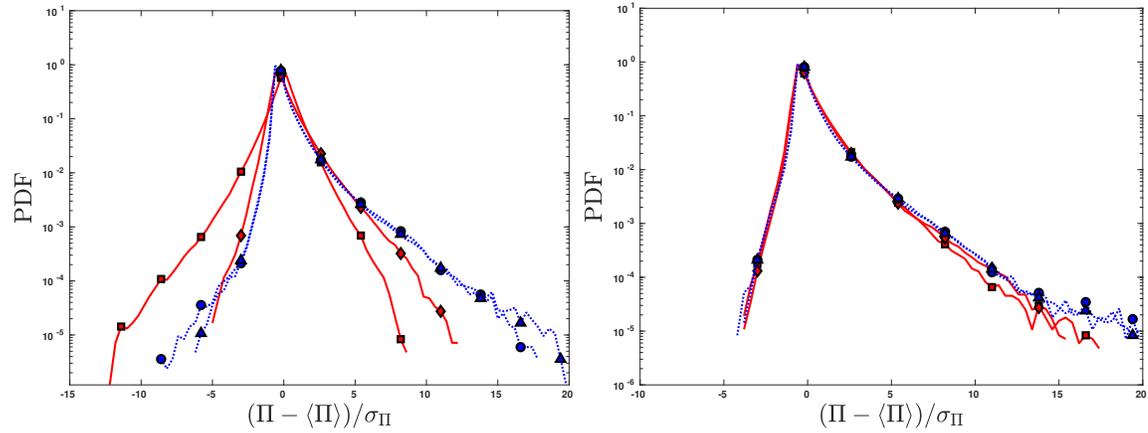
(b) Solid line is GR: with squares for $T/2$ and diamond for $2/3T$. Dotted line is DNS target: with circles for $T/2$ and upward-pointing triangle for $2/3T$.



(c) Solid line is GR: with squares for $5/6T$ and diamond for T . Dotted line is DNS target: with circles for $5/6T$ and upward-pointing triangle for T .

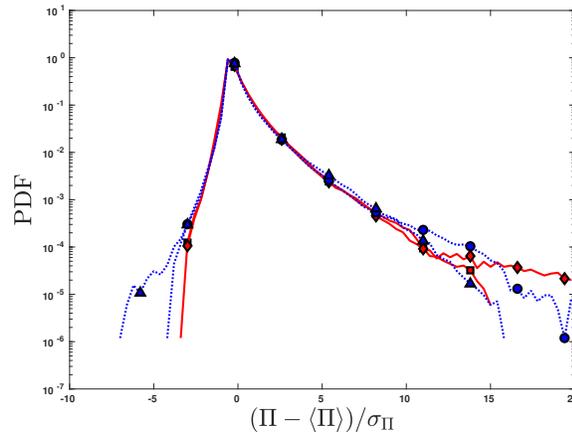
Figure 2.12: The mean subgrid-scale energy dissipations as a function of the filter scales Δ .

The PDFs of the normalised SGS energy dissipation at a given scale are given in Fig.



(a) Solid line is GR: with squares for $0T$ and diamond for $T/6$. Dotted line is DNS target: with circles for $0T$ and upward-pointing triangle for $T/6$.

(b) Solid line is GR: with squares for $T/3$ and diamond for $T/2$. Dotted line is DNS target: with circles for $T/3$ and upward-pointing triangle for $T/2$.

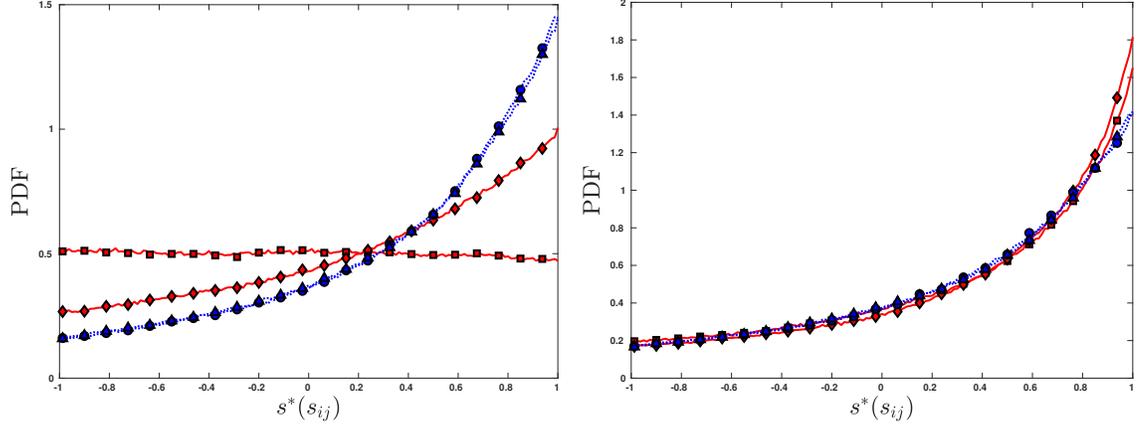


(c) Solid line is GR: with squares for $2/3T$ and diamond for $5/6T$. Dotted line is DNS target: with circles for $2/3T$ and upward-pointing triangle for $5/6T$.

Figure 2.13: PDFs of the subgrid-scale energy dissipation.

2.13. At the initial time, the PDF of the GR field is a symmetrical distribution, which is not surprising as it is the natural behaviour of the GR field. When the time reaches one-third and half of T at Fig. 2.13b, the PDFs of the GR field features have been clearly enhanced and the agreements are already satisfactory. Therefore, the PDF recovers quicker than the mean.

We calculate next the probability density functions of the shape parameter for the

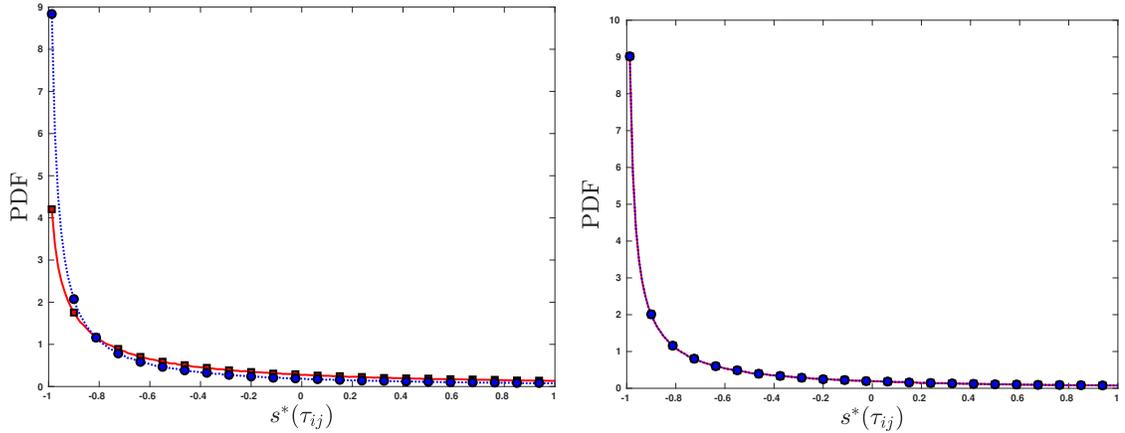


(a) Solid line is Gaussian random: with squares for $0T$, diamond for $T/6$. Dotted line is DNS target: with circles for $0T$ and upward-pointing triangle for $T/6$.

(b) Solid line is GR: with squares for $T/3$, diamond for $T/2$. Dotted line is DNS target: with circles for $T/3$ and upward-pointing triangle for $T/2$.

Figure 2.14: PDFs of the shape factor $S^*(\tilde{S}_{ij})$ for the filtered strain tensor \tilde{S}_{ij} .

strain rate tensor in the GR field. We observe from Fig. 2.14a that PDFs of the GR field tends to be flat. With $t = T/3$ (Fig. 2.14b), we obtain satisfactory PDFs, as the result are quite close to the DNS target field. Therefore, the PDFs of the GR field tend to the DNS target field quite quickly. However, compared with Fig. 2.8, we see that it is still slower than OIC results.



(a) Solid line is GR: with squares for $0T$. Dotted line is DNS target: with circles for $0T$.

(b) Solid line is GR: with squares for $T/2$. Dotted line is DNS target: with circles for $T/2$.

Figure 2.15: PDFs of the shape factor $S^*(\tau_{ij})$ for the subgrid-scale stress tensor τ_{ij}

Fig. 2.15 shows the behaviour of the probability density functions of the shape parameter for the SGS stress evaluated from the GR field. In this case, we are able to obtain a good agreement between the two fields.

2.4.8 CONTOUR PLOTS OF INSTANTANEOUS DISTRIBUTIONS

In this section, we calculate the magnitude of the vorticity $|\tilde{\omega}| \equiv \sqrt{\tilde{\omega}\tilde{\omega}}$ and the magnitude of the strain rate tensor $|\tilde{S}| \equiv \sqrt{2\tilde{S}_{ij}\tilde{S}_{ij}}$. This investigation qualitatively shows the improvement in agreement between target fields and the OIC fields over time. Quantitative analysis is left for the next section.

We first plot the contours for $|\tilde{\omega}|$. The contour results are calculated when the resolution equals 128, and the filter scale equals $8\delta_x$. Figs. 2.16 to 2.19 show the agreement improvement for the $|\tilde{\omega}|$ between two fields. $|\tilde{\omega}|$ in the rebuilt field is, at the initial time, weaker than the later times.

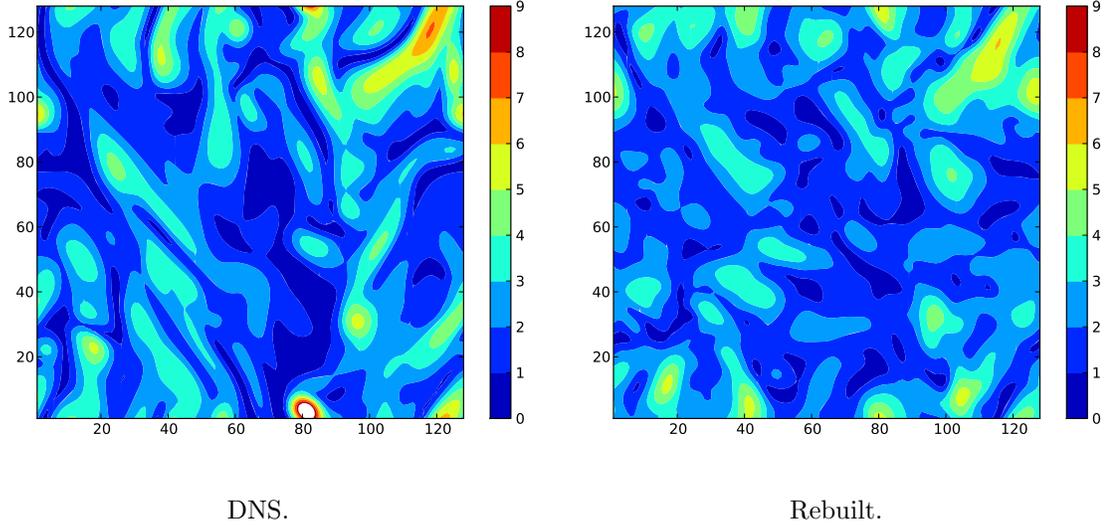
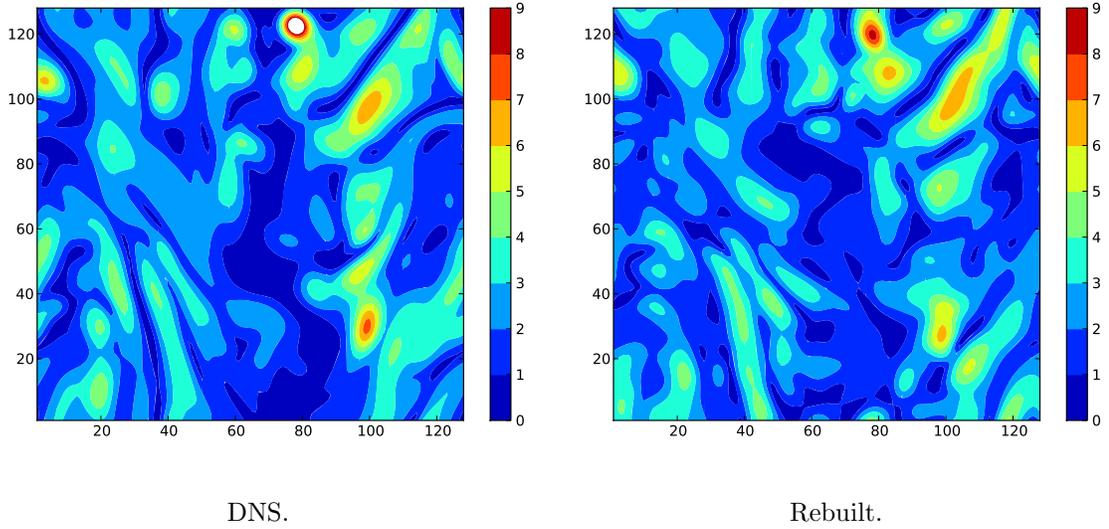
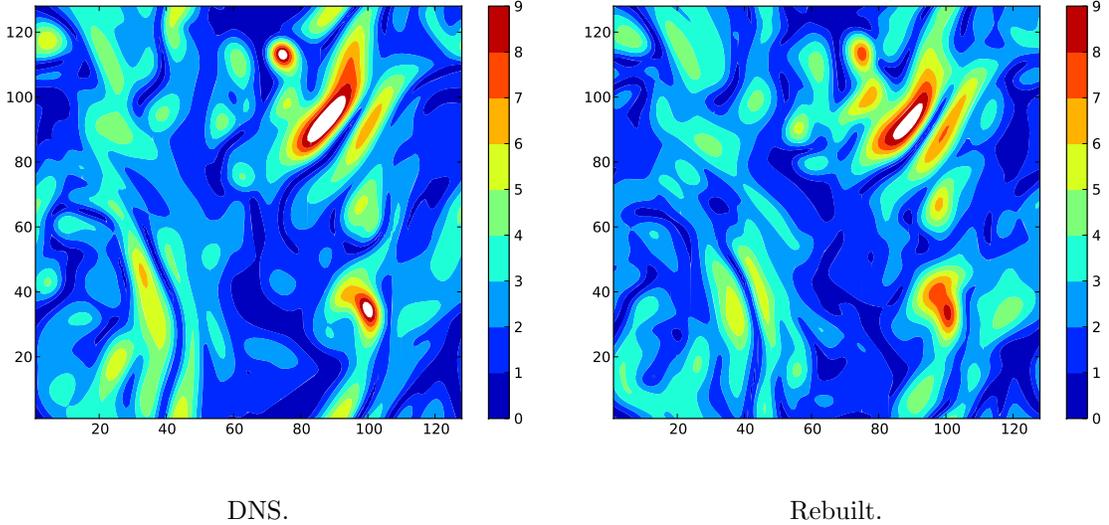


Figure 2.16: Contour plots for $|\tilde{\omega}|$ at $t = 0T$.

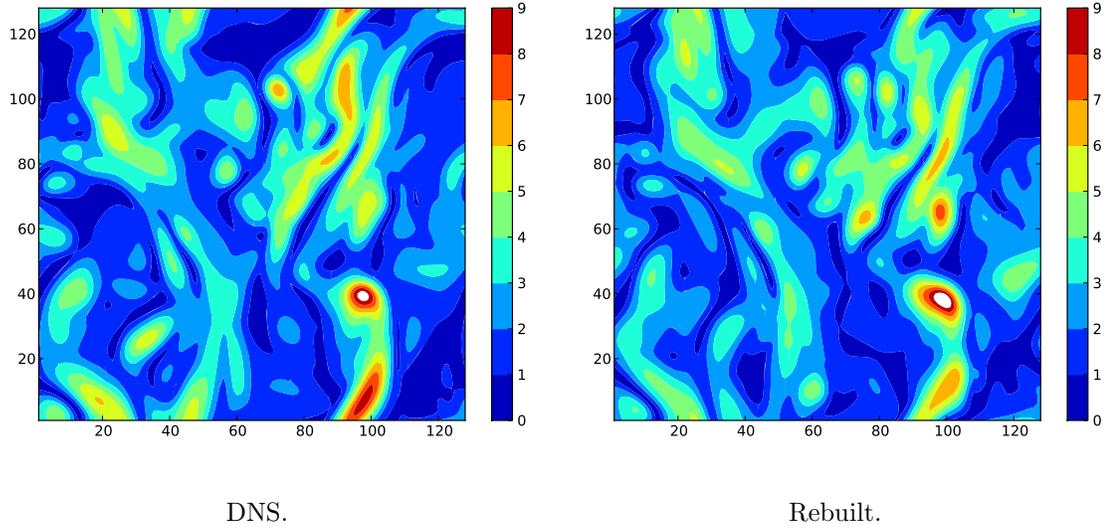
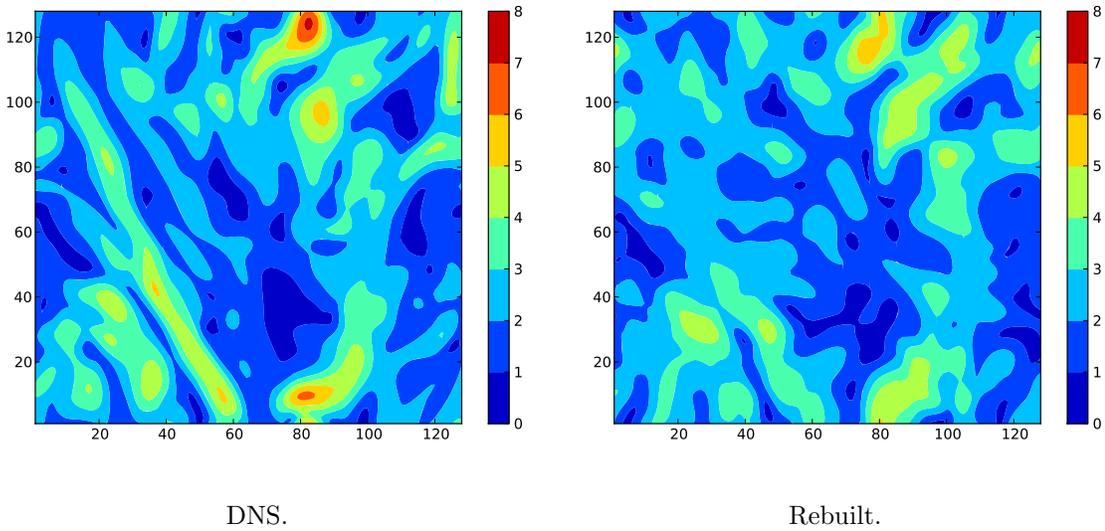
The vorticity contours in the rebuilt field become clearer and stronger with time. Also, the shapes or geometries for the rebuilt field tend closer to the DNS target field. In other words, the differences reduce with time. At the final time, we can see a reasonable qualitative agreement between two fields by looking into the geometry of the contours. For example, Figs. 2.18 and 2.19 shows that high $|\tilde{\omega}|$ spots are observed on right in both the DNS target field and the rebuilt field. Although the regions are not exactly the same, they do show similar orientations and sizes.

Figs. 2.20 to 2.23 show the contour plots for the strain rate field $|\tilde{S}|$ at different time steps. Like the vorticity field, the initial time step shows that some similarity for the

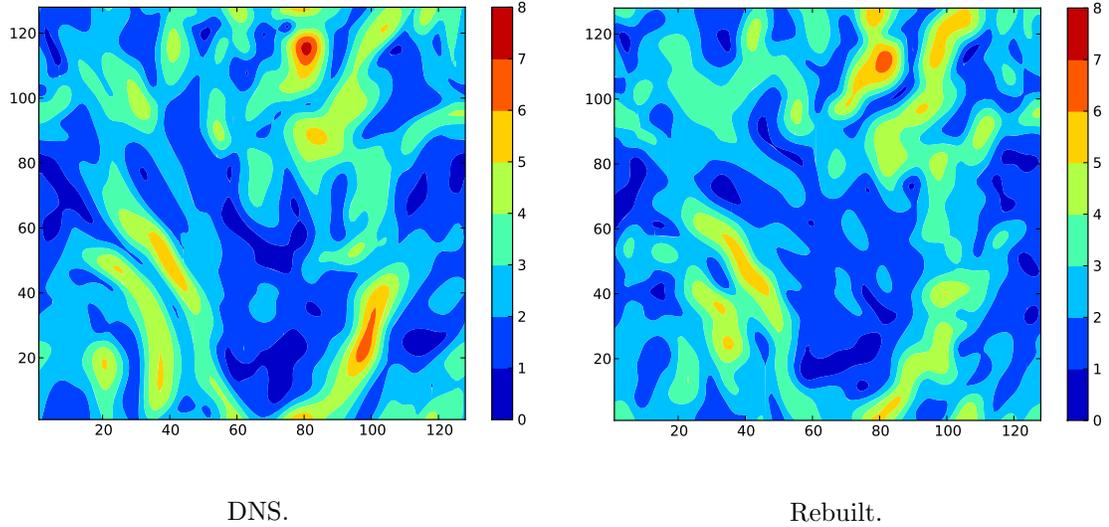
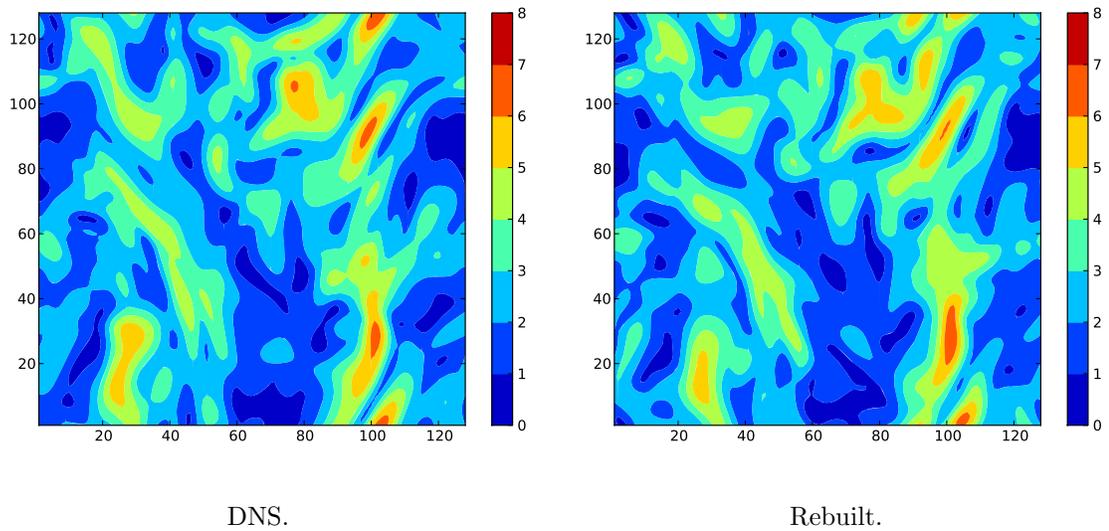
Figure 2.17: Contour plots for $|\tilde{\omega}|$ at $t = T/3$.Figure 2.18: Contour plots for $|\tilde{\omega}|$ at $t = 2/3T$.

contours plots are visible at this early time, although the strong strain spots in the top middle and at the bottom middle are absent from the OIC field. The similarity between the two fields increases over time. At the time T , we observe the best agreement between the target field and the OIC fields, with almost identical structures in their contour plots.

Figs. 2.20 to 2.23 show the contour plots for the strain rate field $|\tilde{S}|$ at different time

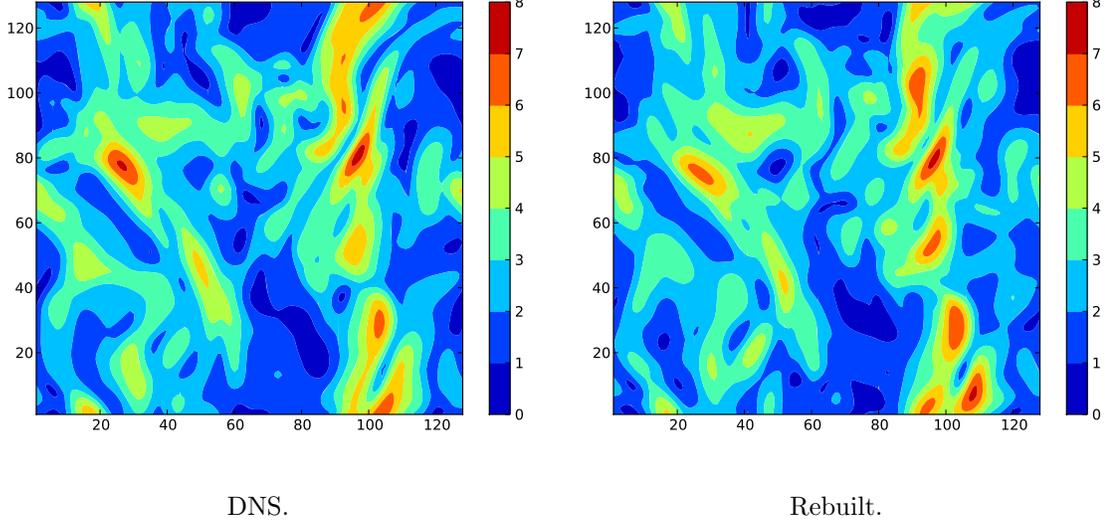
Figure 2.19: Contour plots for $|\tilde{\omega}|$ at $t = T$.Figure 2.20: Contour plots for $|\tilde{S}|$ at $t = 0T$.

steps. Like the vorticity field, the initial time step shows that some similarity for the contours plots are visible at this early time, although the strong strain spots in the top middle and at the bottom middle are absent from the OIC field. The similarity between the two fields increases over time. At the time T , we observe the best agreement between the target field and the OIC fields, with almost identical structures in their contour plots.

Figure 2.21: Contour plots for $|\tilde{S}|$ at $t = T/3$.Figure 2.22: Contour plots for $|\tilde{S}|$ at $t = 2/3T$.

We noticed, as a general conclusion, both the rate of the strain field and the vorticity at the rebuilt field become closer to the DNS field over time.

In the literature, there are many methods to identify the vortices in the velocity fields. In the following, we present another method for vortex identification. It is known as the Q criterion for vortex identification. The quantity Q is defined as:

Figure 2.23: Contour plots for $|\tilde{\mathcal{S}}|$ at $t = T$.

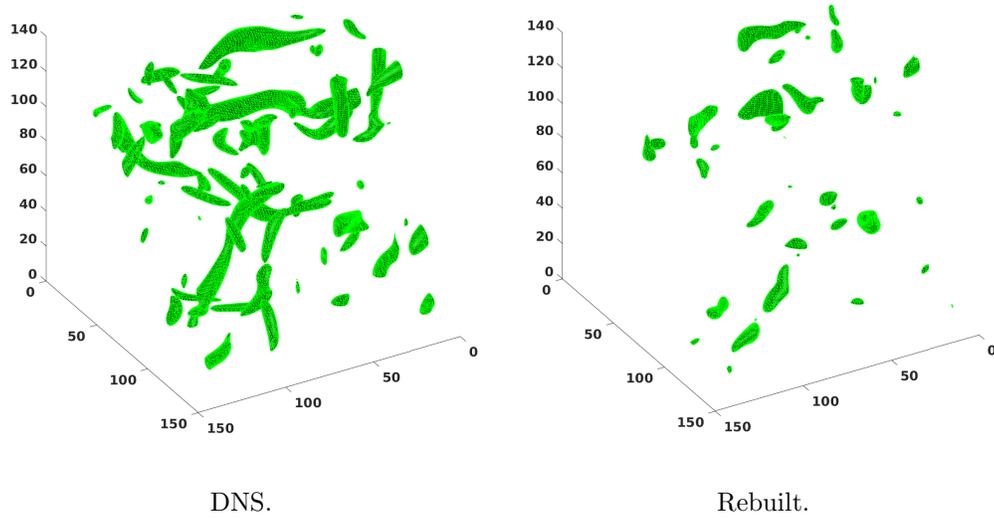
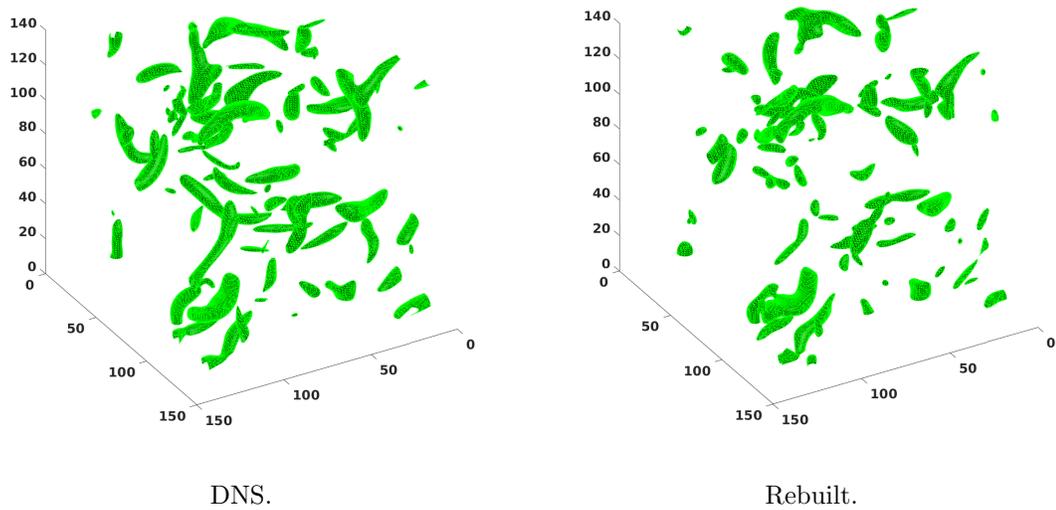
$$Q = \frac{1}{2}(\|\tilde{\mathbf{\Omega}}\|^2 - \|\tilde{\mathbf{S}}\|^2), \quad (2.14)$$

where $\tilde{\Omega}_{ij} = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i})$, is the filtered rate of rotation tensor, corresponding to the motion of pure rotation, and \tilde{S}_{ij} is the filtered strain rate tensor, which corresponds to the pure irrotational motion. Therefore, Q is a measure of the excess rotational rate relative to the rate of strain. One may define connected areas with non-negative Q as vortices. Note that the norm $\|\mathbf{G}\|$ of tensor \mathbf{G} is defined by $\|\mathbf{G}\| = [\text{tr}(\mathbf{G}\mathbf{G}^T)]^{1/2}$, i. e. $\|\tilde{\mathbf{S}}\| = [\text{tr}(\tilde{\mathbf{S}}\tilde{\mathbf{S}}^T)]^{1/2}$ and $\|\tilde{\mathbf{\Omega}}\| = [\text{tr}(\tilde{\mathbf{\Omega}}\tilde{\mathbf{\Omega}}^T)]^{1/2}$.

The vortex structure is shown in terms of the isosurfaces, in Figs 2.24 to 2.26, where on the left are the isosurfaces from the target field while on the right are the isosurfaces from the rebuilt fields. The surfaces are shown in green. The threshold value for Q in our calculation here equals 3.5.

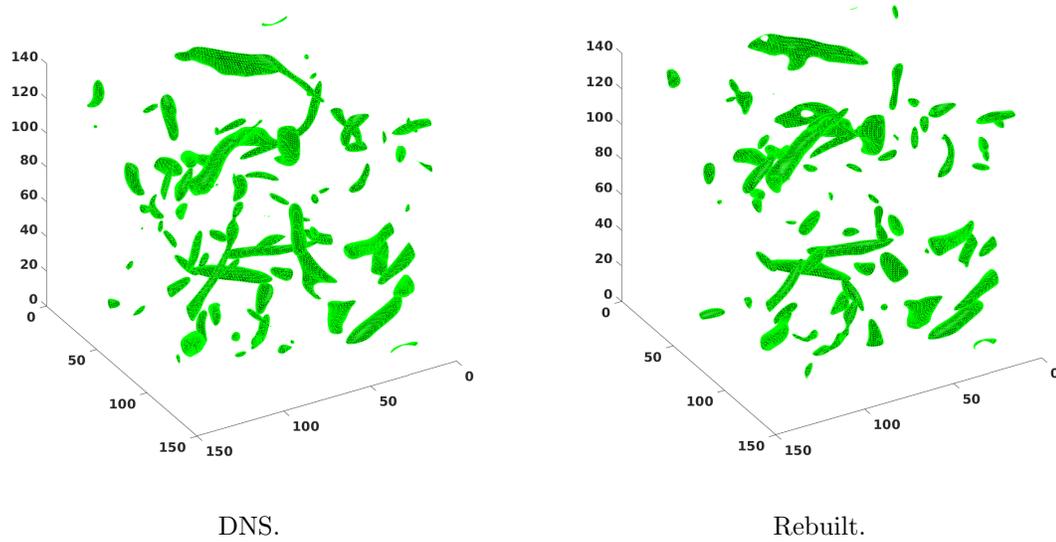
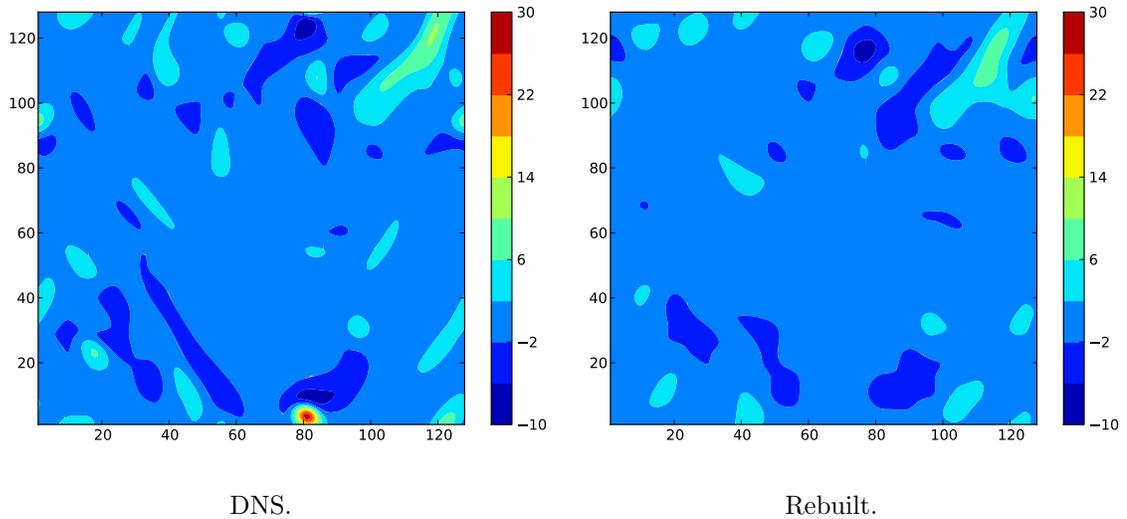
In a general analysis of this method, at the initial time in Fig. 2.24, the optimal fields are able to build some structures. However, the number of these structures is smaller than the target data and the agreement between them is not big enough.

Fig.2.25 shows the isosurfaces for Q at $t = T/2$. We can see the number of vortical structures has rapidly risen for the optimal field. Besides, we can also observe, their locations in the space are closer to those of the DNS target field. The best agreement between the fields of the optimal and DNS target appears at the final time in Fig. 2.26, where the structures of the optimal field have an almost identical number to the target data and their appearance in the three-dimensional space demonstrates the perfect agreement between the fields.

Figure 2.24: Q criterion for $t = 0T$.Figure 2.25: Q criterion for $t = T/2$.

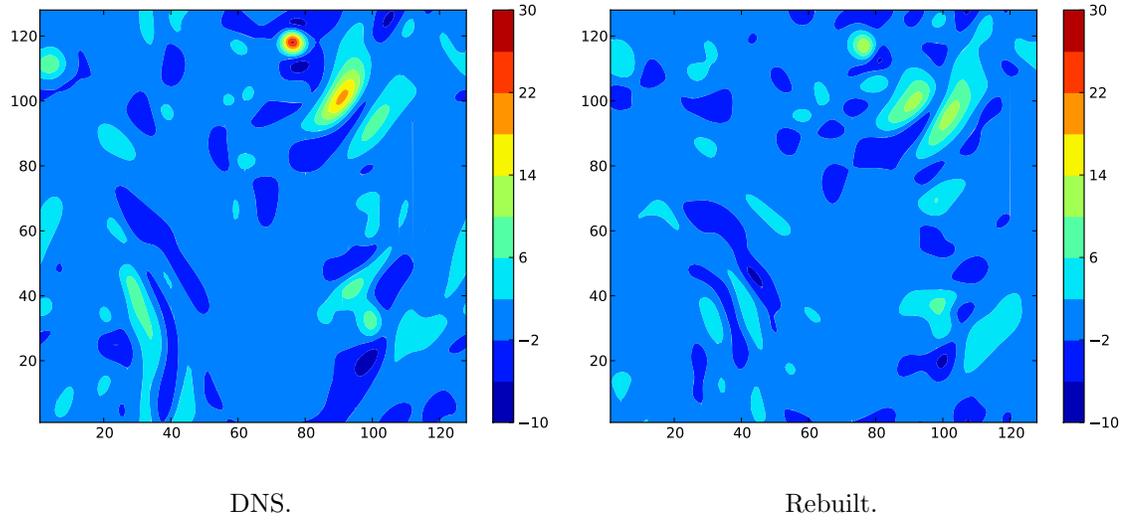
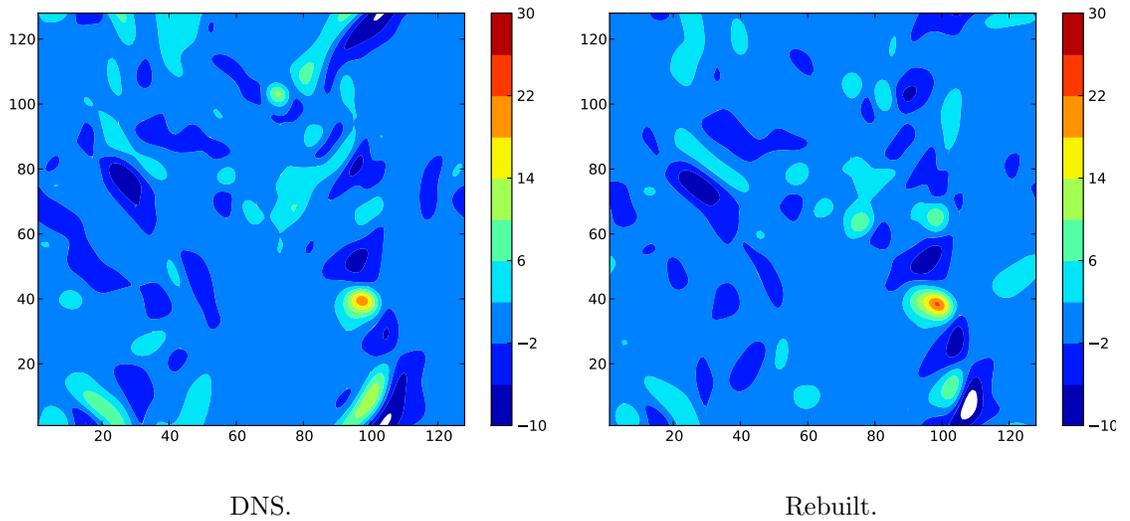
Figs. 2.27, ..., 2.29 present the contour plots of Q on a 2D slice.

At the initial time, weak agreements are shown, where the structures of the contours in the OIC fields show some similarity with the contours in the DNS target field. Fig. 2.28 shows that the geometries of the vortices present a bigger agreement between the

Figure 2.26: Q criterion for $t = T$.Figure 2.27: Q Contour plots for $t = 0T$.

two fields at $t = T/3$ and $t = T/2$. However, the red spots are still absent. In the final time $t = T$, the similarities between the contours in the two fields are quite big. In particular, the biggest magnitudes are presented, indicating the good agreement are achieved between the two corresponding fields.

Therefore, the agreement between the vortical structures identified with the Q criterion

Figure 2.28: Q Contour plots for $t = T/2$.Figure 2.29: Q Contour plots for $t = T$.

is also very good. There is no difference from using $|\tilde{\omega}|$. Therefore, we will use $|\tilde{\omega}|$ only in what follows.

2.4.9 CHARACTERISING STRUCTURES WITH MINIMUM VOLUME ENCLOSING ELLIPSOIDS

The previous subsection demonstrates qualitatively the similarity between the structures in the target field and the OIC field. We consider in this subsection quantitative description of the structures using the Minimum Volume Enclosing Ellipsoid (MVEE).

As we presented in the previous section, the contours show some simple descriptions about the structures. However, it calls for investigations about more accurate descriptions for quantities such as shapes, sizes and locations. The use of the MVEE method presents more information about these parameters. The MVEE is the ellipsoid which has the smallest volume between all the ellipsoids that contain a certain structure. We use the centre and length for the three axes of the MVEE to determine the location and dimensions of the structure and figure out the orientation of the structure from the orientation of the three axes. Therefore, we use the information obtained from the MVEE to define the location, dimensions and orientation of the structure which is bounded by the MVEE. Such a comparison involves four steps to complete the calculations, which are discussed as follows:

In the first step, a group Λ of the grid points are extracted, where the condition of a certain threshold is met. Usually, Λ consists of a number of separate regions. The structures we are going to consider refer to these kinds of regions, such that every single structure corresponds to a separate region.

In the second step, we extract and distinguish a structure in Λ from others. This is done by formulating it as a clustering problem, by using a density-based clustering algorithm called DBSCAN. In the DBSCAN algorithm, we identify firstly the neighbours of a point P , where a neighbour is determined as a point whose distance to P is smaller than a given value ε_p . We consider the point which has less than n_p neighbours as noise. The identification of the cluster is accomplished according to the rule in what follows: the neighbours for a non-noise point P are considered to be in the same cluster where point P belongs. The algorithm accomplishes the identification of the points in a cluster by recursively using such a rule until all the points to be identified are noise. The procedure is repeated for other non-noise points that have not been associated with any structures. The details about the algorithm of DBSCAN are presented in [25]. The result of applying the algorithm is that we find all the points in a structure and store the points separately from the points of different structures. The distance normalisation is done in such a way that the distance between any two sequential grid points is 1. As a result, $\varepsilon_p = D^{1/2}$, and we use $n_p = D + 1$, where D is the embedding space dimension.

In the third step, we calculate the MVEE for each structure. An ellipsoid is defined by a symmetric positive definite matrix \mathbb{E} , which identifies the orientation of the ellipsoid and its axes lengths, and a vector \mathbf{c} which is specifying the centre of the ellipsoid.

The MVEE volume is proportional to the determinant of $\sqrt{\mathbb{E}}$, $\det \sqrt{\mathbb{E}}$. We assume in the structure \mathcal{P} we have points $p_i (i = 1, \dots, N)$. The MVEE is defined by the optimal \mathbb{E} and the centre \mathbf{c} that minimize $\det \sqrt{\mathbb{E}}$, subject to the constraints.

$$(p_i^T - \mathbf{c}^T)\mathbf{E}^{-1}(p_i - \mathbf{c}) \leq 1, \quad \forall p_i \in \mathcal{P}. \quad (2.15)$$

We apply the constraints to ensure that p_i are in the ellipsoid. The Khachiyan algorithm can be used to solve this problem [85]. The MATLAB implementation by Moshtagh is used [65].

We find the correspondence between a structure in the DNS target field and its reconstruction in the rebuilt field in the 4th step (and last). We apply the above three steps to the DNS target field and the rebuilt fields. When the rebuilt field perfectly mimics the DNS target field, we will gain two groups of MVEEs, with 1 – 1 correspondence between the members. However, the differences between the DNS target and rebuilt fields are large enough sometimes to stop the 1 – 1 correspondence; as a result, such a simple technique fails in reality. In order to prevent that happening, we apply the following procedure, where the identification of the rebuilt field structures is by the location instead of their sizes.

According to the number of grid points of a certain structure, the size of the structure is determined. The DNS target field structures are first arranged into a list in descending order according to cluster sizes. Assuming that, we have a structure \mathcal{G}_T in the list, the reconstructed structure whose distance to \mathcal{G}_T is the minimum specified. This is where we define the distance between the two structures as the shortest distance from one to the other. Suppose that \mathcal{G}_o is that structure. If \mathcal{G}_o is not unique, the chosen structure will be the one that has the largest size. This \mathcal{G}_o is taken as the reconstruction of \mathcal{G}_T . \mathcal{G}_o and \mathcal{G}_T are named the matching structures. We repeat the process above for each structure in the DNS target field, starting from the largest one until we identify all the matching structures. We compare the matching MVEEs in terms of their statistics, including the alignment, relative displacement, and relative sizes.

Firstly, we give a quick demonstration of the agreement between the structures by plotting the structures as clusters of the points inside the structures. These results are shown in Fig. 2.30 and Fig. 2.31 below, where the blue squares represent the clusters of the points in the OIC fields and the red circles represent the clusters of the points in the DNS target fields.

The main purpose of this comparison is to provide some intuition about the structures. Although only selected samples are shown, one may still see that the agreement between the structures improves over time. Some structures are highly irregular. Such as the 1st structure in Fig 2.30. It is interesting to see that the reconstructed structure mimics the irregular shape quite well.

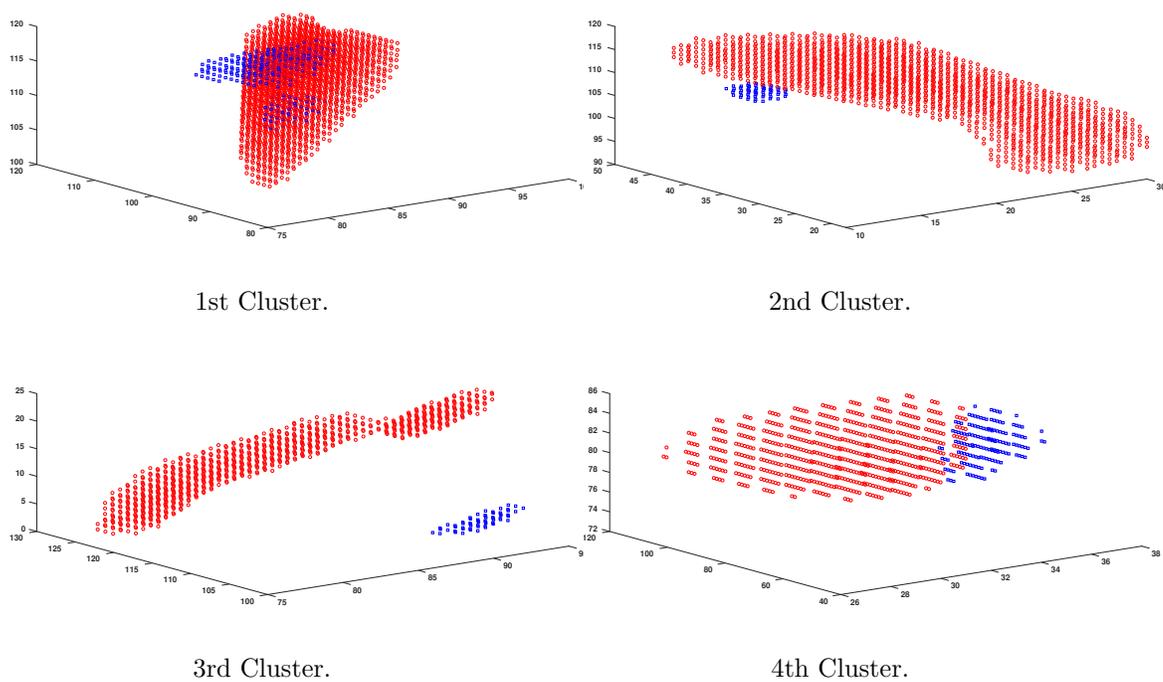


Figure 2.30: Structures as clusters of points for $|\tilde{S}|$ at $t = 0T$.

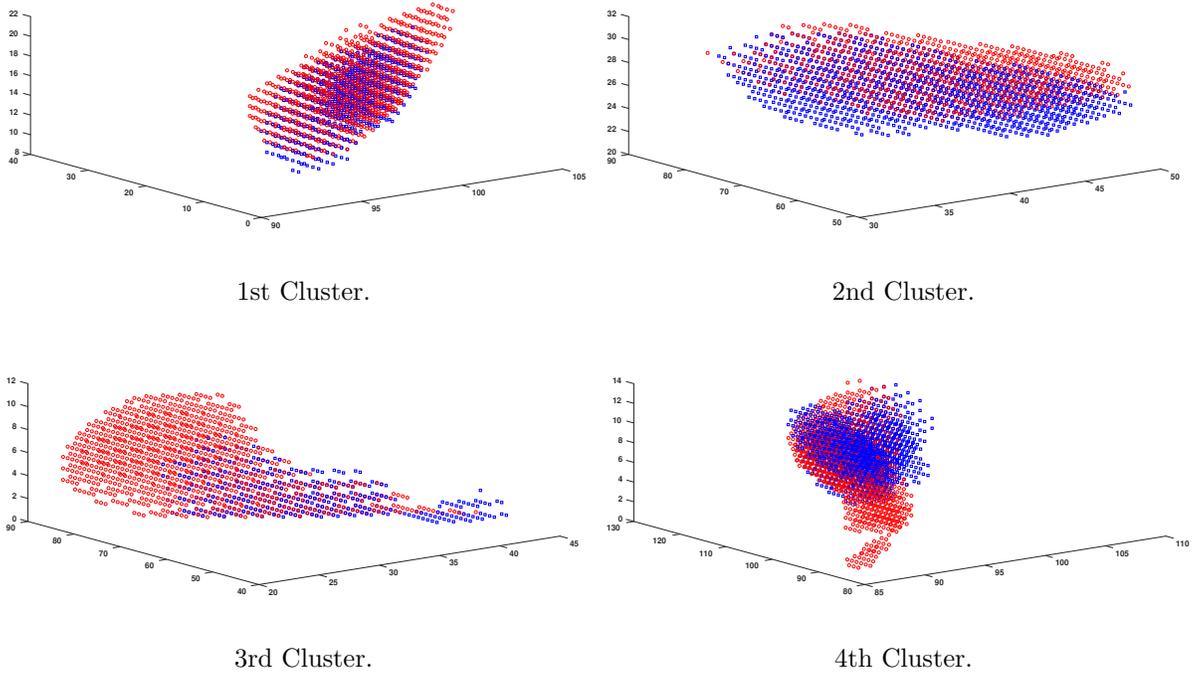


Figure 2.31: Structures as clusters of points for $|\tilde{S}|$ at $t = T$.

2.4.10 THE MINIMUM VOLUME ENCLOSING ELLIPSOID (MVEEs) FOR THE VORTICITY FIELD

We present now results related to the MVEE for the vorticity. All the calculations have been done with the filter scale equal, $8\delta_x$, and the structures are found with vorticity magnitude given by $\sqrt{\tilde{\omega}\tilde{\omega}} \geq 2.5\langle\tilde{\omega}\tilde{\omega}\rangle^{1/2}$.

The alignment data are represented by the histograms in Figs 2.32, ..., 2.52. We investigate these statistics at different times within the optimization horizon, starting from the initial time and ending with the optimization horizon, including several time-steps in-between. At each single time step, we calculate the alignment between the three axes of the corresponding MVEEs in the rebuilt and the DNS target fields.

Three cases are considered. The first case is for alignment between all corresponding MVEEs. The second is for the alignment between MVEEs for which the ratio of the major axis to the middle axes is bigger than an allocated sphericity value. In other words, the statistics of the alignment will be conditioned on rod-like MVEEs. In the third case, we find the alignment when it is conditioned on pancake-like MVEEs. An MVEE is considered pancake-like when the ratio of the middle axis to the minor axis is bigger than the same sphericity value which has been used in the previous case. The value is 2.

At each time step $t = 0T, \dots, t = T$, we have three sets of histograms including the unconditional and conditional histograms which they represent: the alignments for the major axes; the alignment results of the middle axes; and the results of the alignments for the minor axes.

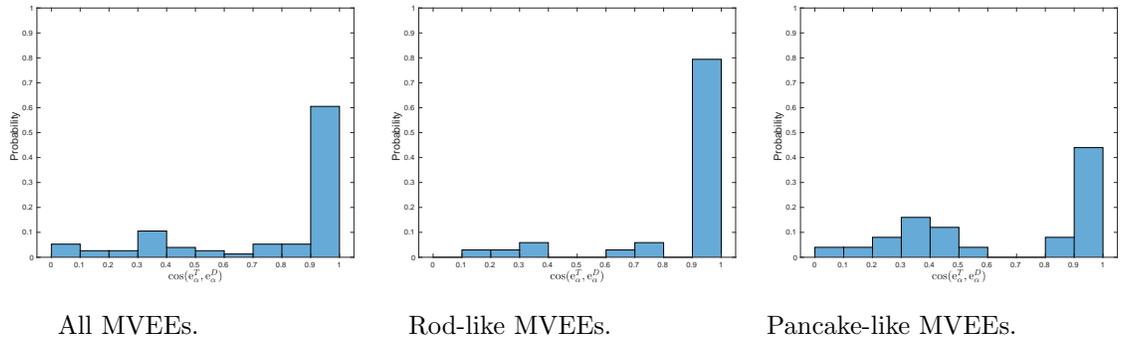


Figure 2.32: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = 0T$.

The alignment is given by the histograms for the cosine of the angles between the axes. Perfect alignment is obtained when the cosine equals 1.

At the initial time (Fig. 2.32), the probability of 0.6 shows acceptable agreement in the alignment of the MVEEs for all clusters. In the same figure, the conditional cases show different observations, as we notice that the rod-like MVEEs give the best result for the major axes, as the probability value is 0.8. However, this is not the case with

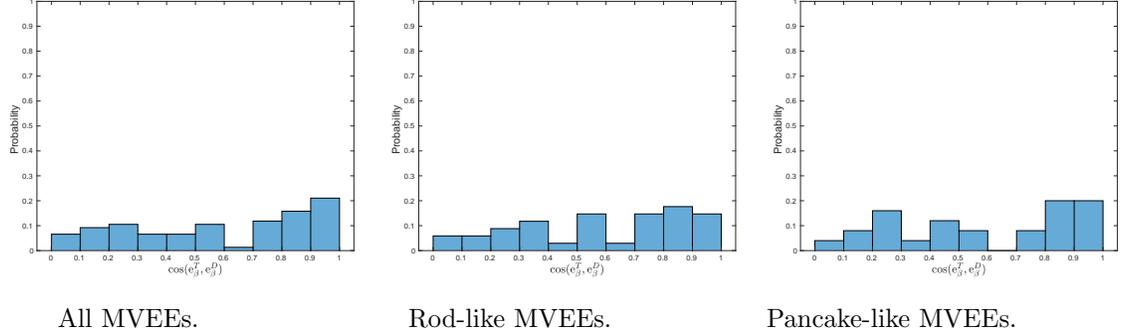


Figure 2.33: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = 0T$.

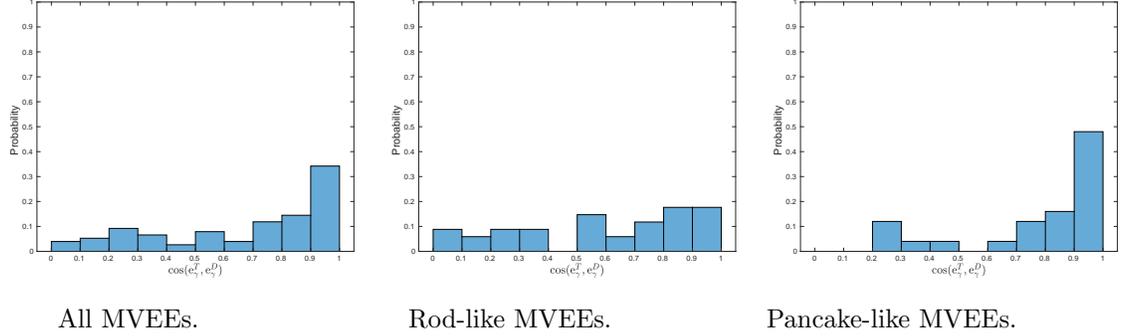


Figure 2.34: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = 0T$.

the alignments of pancake-like MVEEs where the histogram gives a lower value of 0.45 at 1.

In Fig. 2.33, we see that the histogram for the middle axes MVEEs shows a probability value at 0.22 for the perfect alignment. The other conditional probabilities are lower as they give for rod-like MVEEs and pancake-like MVEEs at 0.15 and 0.2, respectively.

According to the results at Fig. 2.34, the minor axes present slightly better agreement than the middle axes, but it is still much lower than the major axes. The histogram on the left gives alignment a percentage at 0.34. The conditional alignment of pancake-like MVEEs shows agreement with a probability value at 0.48 and rod-like MVEEs is the much lower at 0.19.

Therefore, we started at the initial time with some agreements between the corresponding clusters between the OIC and DNS target fields, especially with the major axes, but we do not see a big agreement at such an early stage. However, the agreement improves as the time increases, as we will show with Figs 2.35, . . . , 2.52.

Considering the unconditional alignment for the major axes, the value of the probability of the major axes at $\cos(e_\gamma^T, e_\gamma^D) = 1$ improves over time and reach the peak at 0.88 at

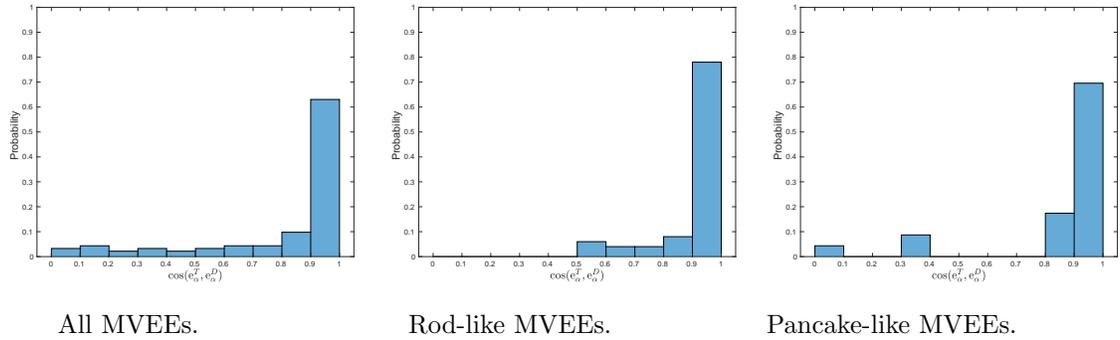


Figure 2.35: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = T/6$.

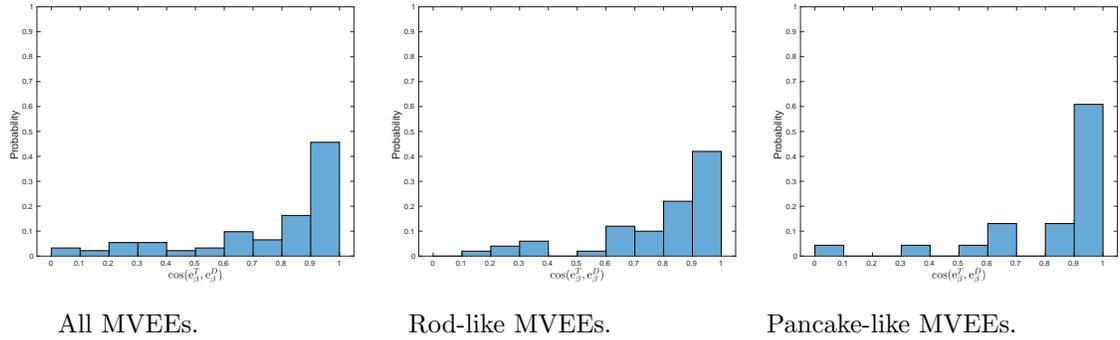


Figure 2.36: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = T/6$.

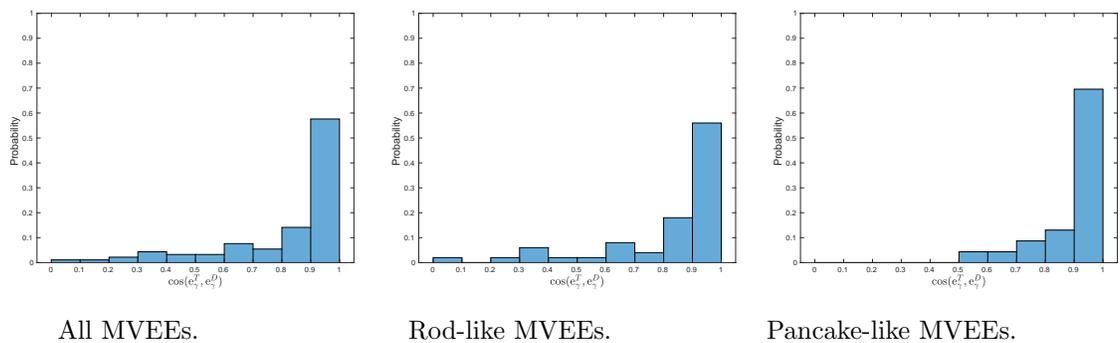


Figure 2.37: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = T/6$.

Fig. 2.50. It indicates that we are able to reproduce MVEEs in the OIC fields quite closely to what is already in the DNS target data. The conditional histograms have different indications, as rod-like MVEEs is also considered as the highest peak at 0.93. However, pancake-like MVEEs gives a value at 0.55, which is not quite satisfactory, where it drops from its peak at just the previous time step from the value at 0.97 when

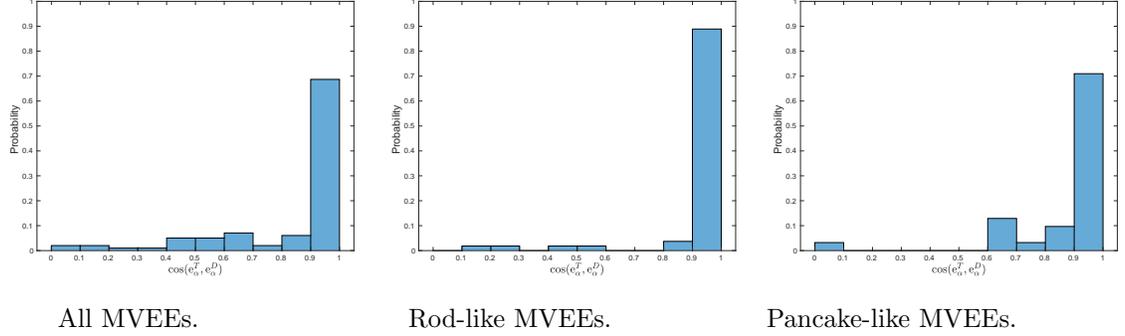


Figure 2.38: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = T/3$.

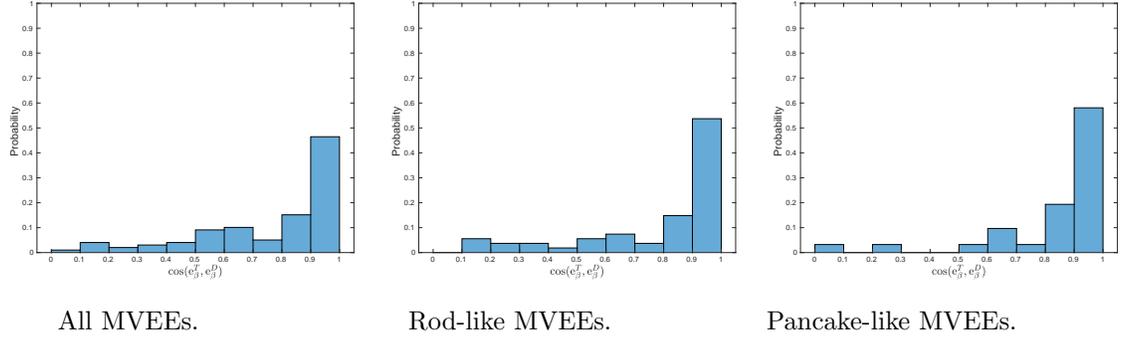


Figure 2.39: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = T/3$.

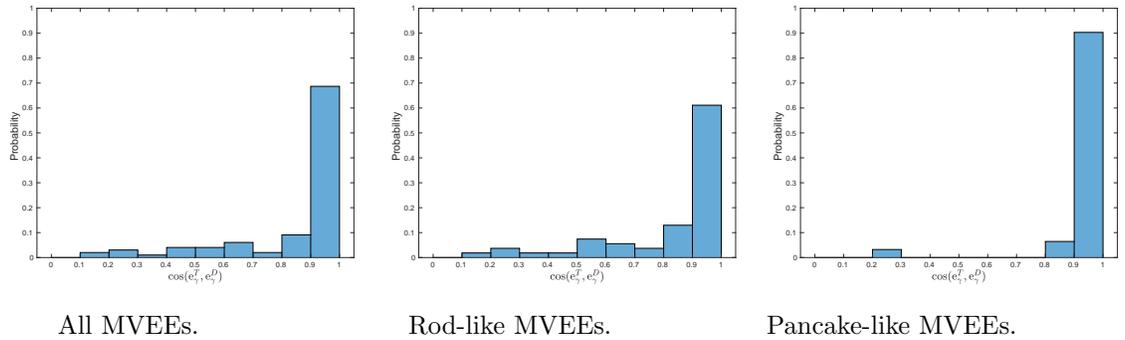


Figure 2.40: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = T/3$.

the time equals $5/6T$.

The probabilities in Fig. 2.51 shows that there are acceptable alignments for the middle axes at time T . For instance, the highest peak for the middle axes for all MVEEs is approximately 0.54. Also, the conditional alignment for rod-like MVEEs is the highest at 0.6. However, pancake-like MVEEs gives a somewhat smaller value.

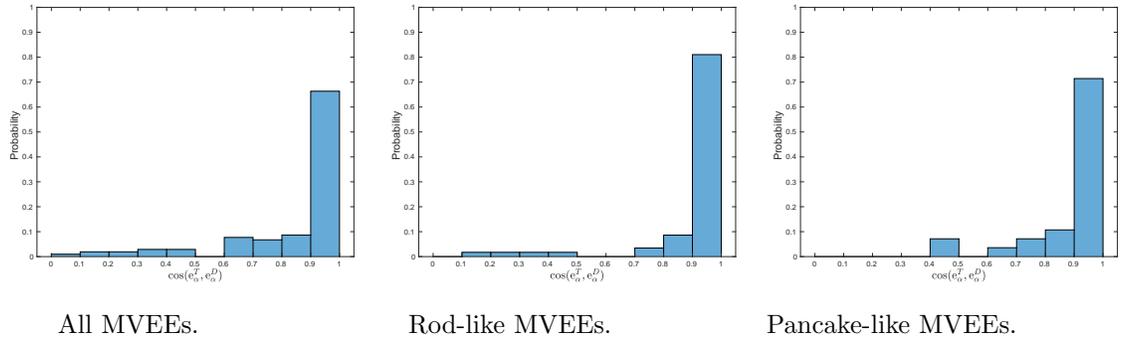


Figure 2.41: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = T/2$.

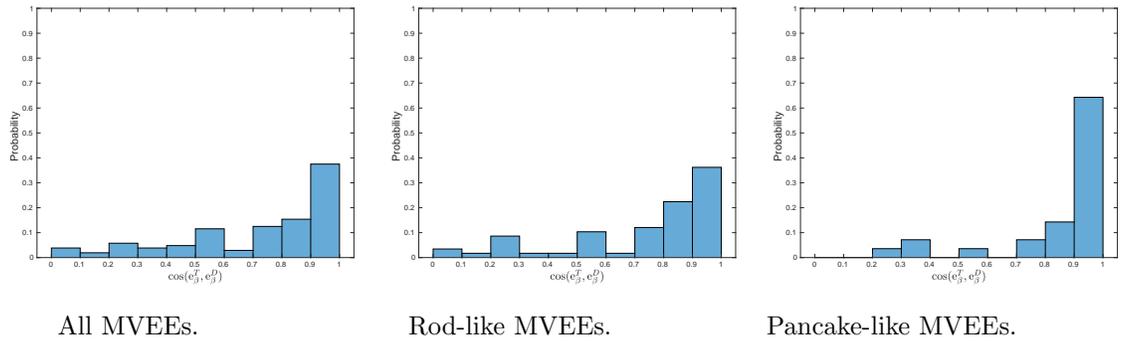


Figure 2.42: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = T/2$.

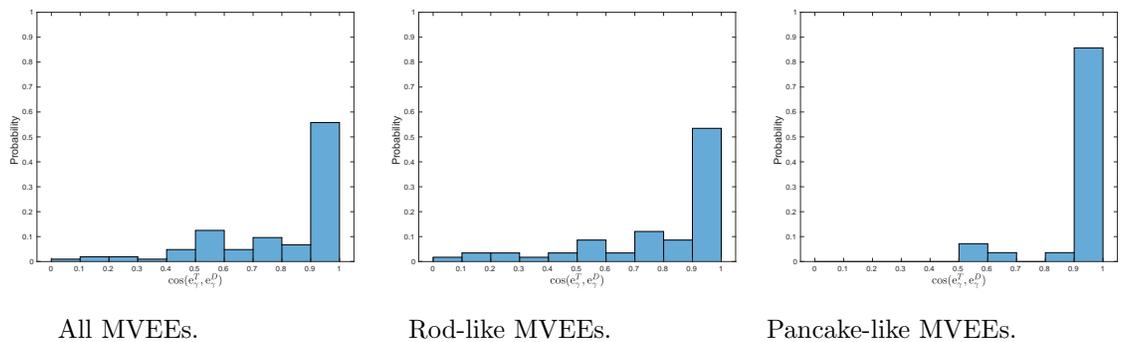


Figure 2.43: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = T/2$.

Fig. 2.52 shows that the probability for the minor axes reaches to 0.6, indicating a good alignment. It is bigger than the middle axes value and reflects better agreement. Rod-like MVEEs gives an acceptable value of 0.65 while pancake-like MVEEs reaches a lower value of 0.48.

To summarize, the agreements between the axes are shown clearly by the histograms,

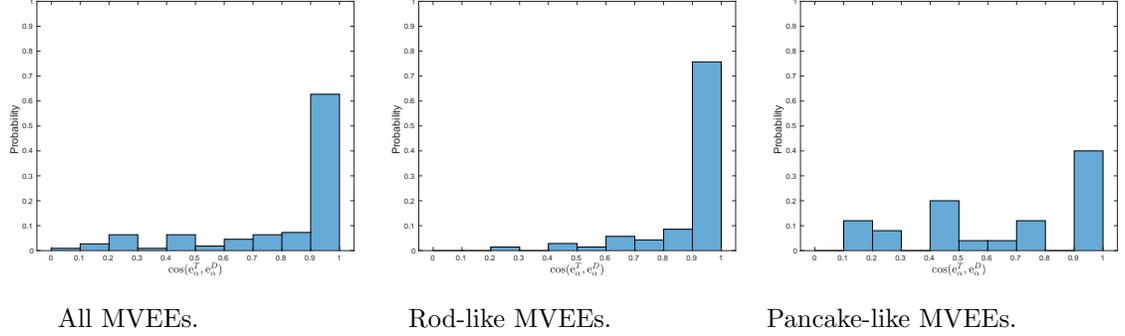


Figure 2.44: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = 2/3T$.

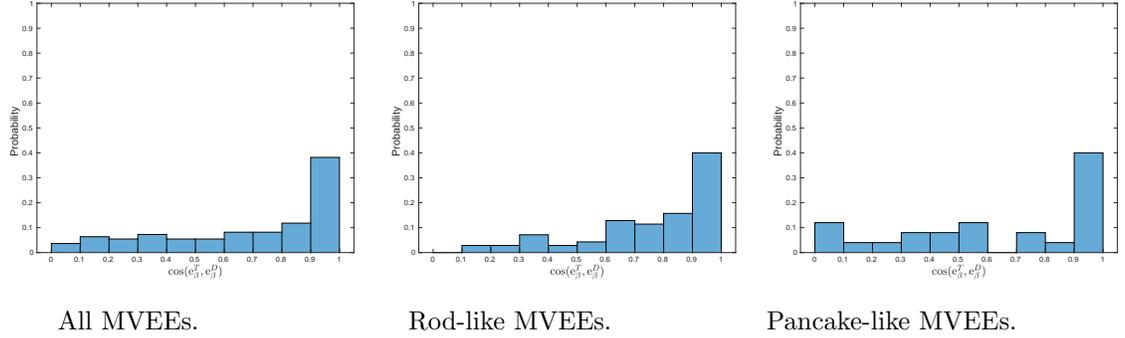


Figure 2.45: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = 2/3T$.

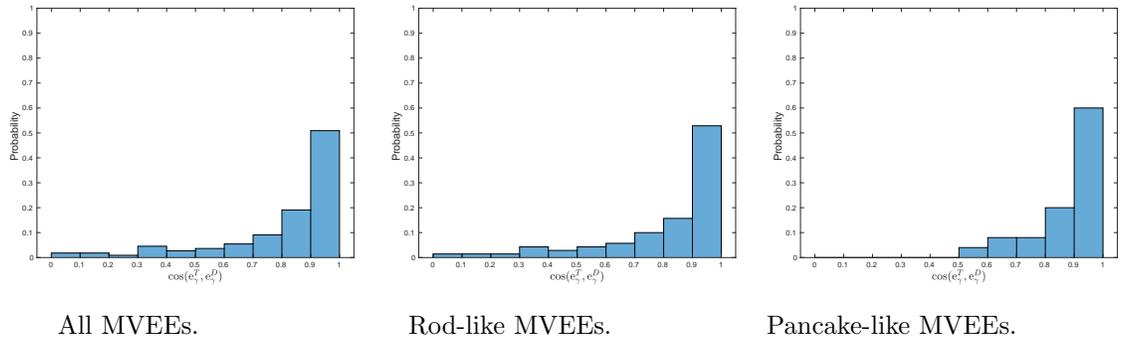


Figure 2.46: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = 2/3T$.

as there exist high probability values at 1 for all MVEEs, which means that the angles between the corresponding axes are small. The major axes give the highest peaks with the best agreement, and then the minor axes, and finally the middle axes come as the third.

The conditional alignments for rod-like MVEEs and pancake-like MVEEs also indicate

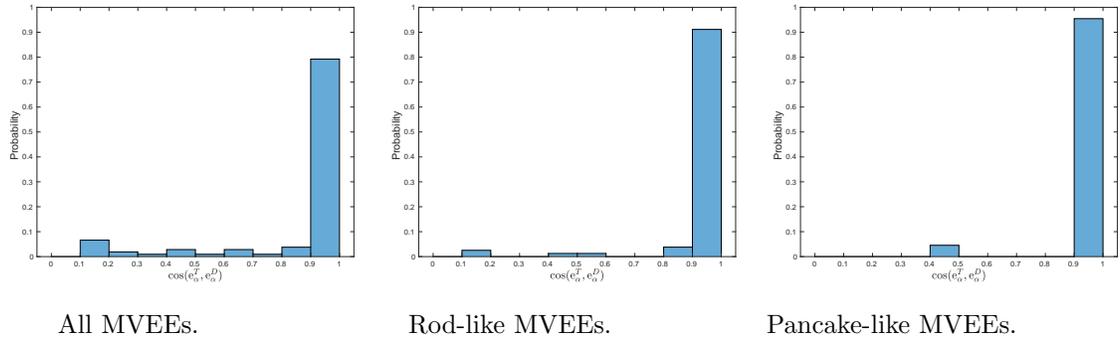


Figure 2.47: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = 5/6T$.

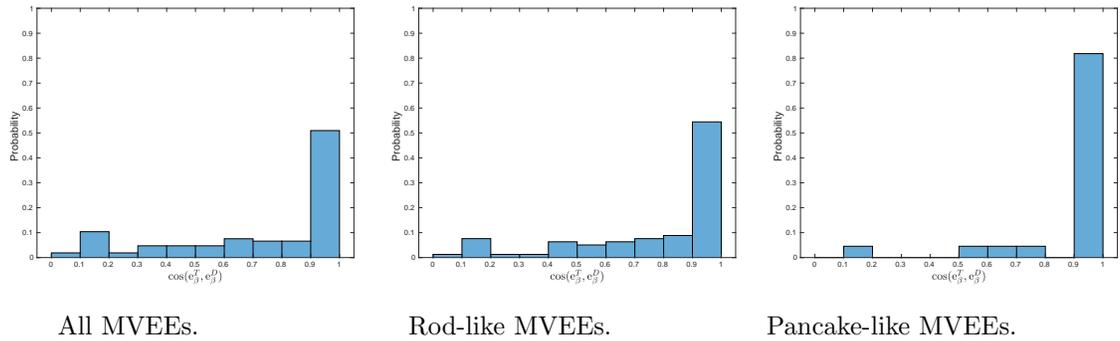


Figure 2.48: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = 5/6T$.

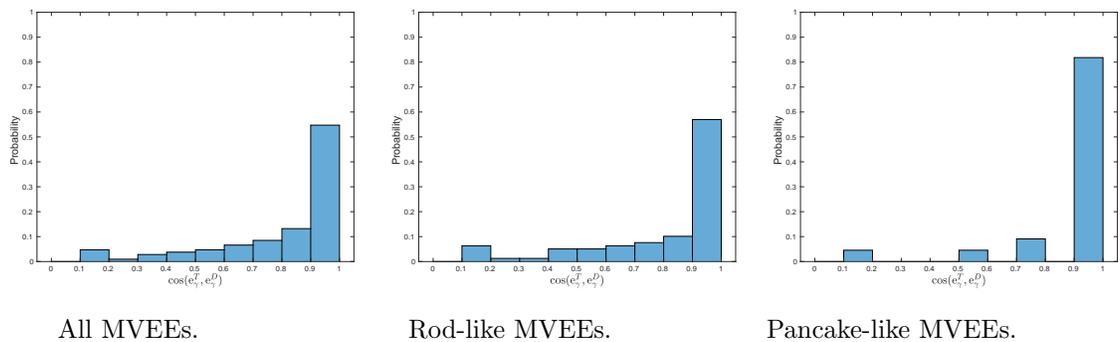


Figure 2.49: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = 5/6T$.

strong alignments between the corresponding axes. Rod-like MVEEs tend to show better alignment than others.

Fig. 2.53 presents the median of the alignment distributions as functions of time. The red curve represents the median for the major axes. We can see that the median value starts at 0.95, at the initial time, and stay stable at the peak at 0.99 until the end. The

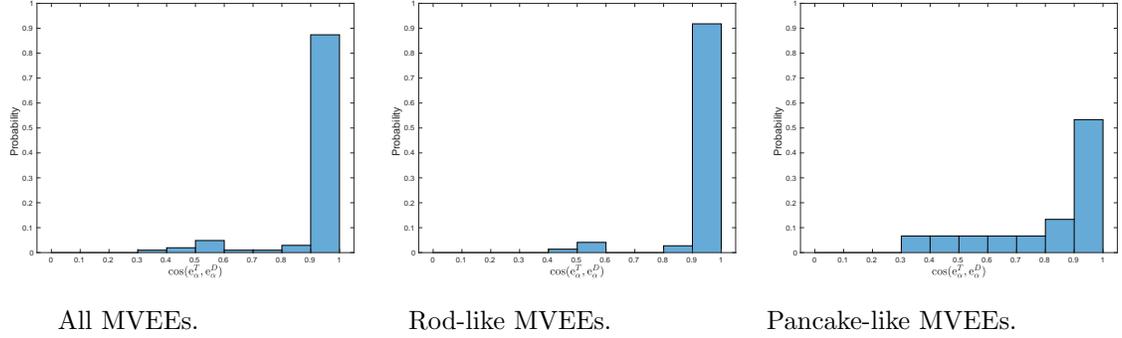


Figure 2.50: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = T$.

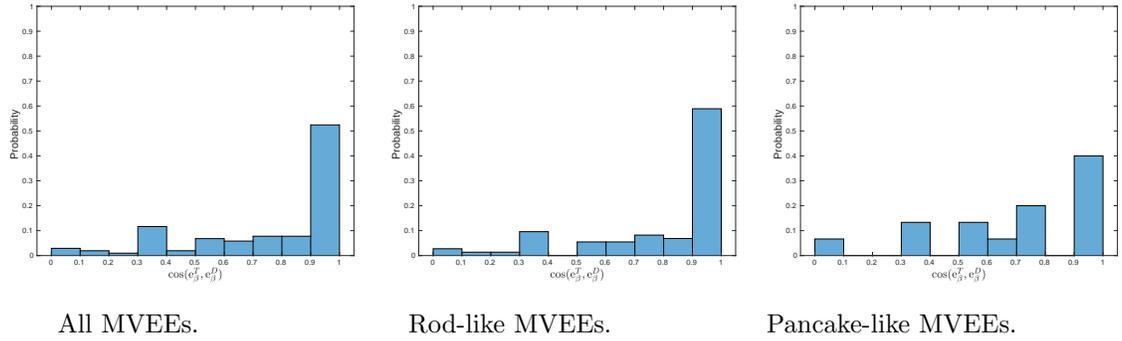


Figure 2.51: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = T$.

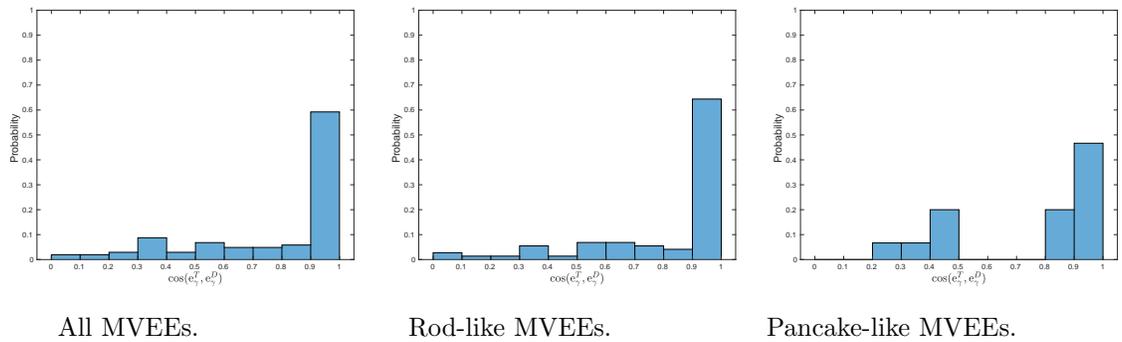


Figure 2.52: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = T$.

big value of the median reflects that the angle between the axes in the OIC fields and the DNS target field is very small, (i. e. the axes of the OIC fields have the similar orientation as the axes of the DNS target data). The major axes have the best results among the three, with a higher median at most times.

The median of the middle axes at the same figure is represented by the blue curve. It

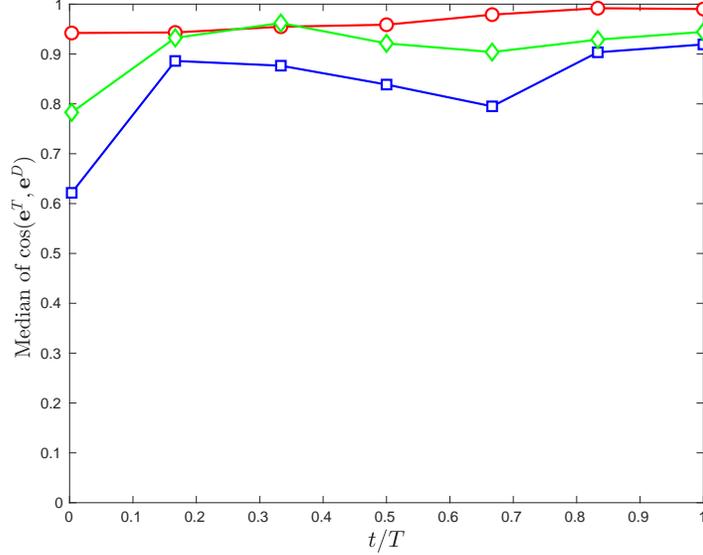


Figure 2.53: Medians of $|\cos(\mathbf{e}^T, \mathbf{e}^D)|$ for strong $|\tilde{\omega}|$ structures. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.

starts with a value of 0.63. However, the value increases at the end to 0.92. It is the lowest value between all axes. The green curve illustrates the median of the alignment data of the minor axes. The median starts at 0.79 and stops at 0.95 at the time T .

Therefore, the medians also reflect the good agreement between the target and the OIC fields.

Fig. 2.54 shows the median ratios of the sizes between the clusters in the optimal fields to the corresponding clusters of the target data field. As we can see from the results the curve starts with a small value at 0.15, increases to a value 0.9 which shows good agreement in the size of the clusters. In the same figure, the green lines refer to the median ratios for clusters of a given size. As we can see the green lines are either over the median ratios or below it. Moreover, there are some significant values which exceed 1. However, the behaviours of the green curves are similar to what we have seen in the black curve; both curves increase as the time passes.

In what follows, we compare the locations of the MVEEs in the target fields and the OIC fields using the probability of the displacement of the centres of the MVEEs d_c for both OIC and DNS target fields relative to the lengths of the axes ℓ at the DNS target field, which is represented by d_c/ℓ .

The results are given for three different time steps, which are $0T$, $T/2$ and T . Each includes results for Major axes, Middle axes, and Minor axes, which represent the displacement of the centres of the MVEEs in the major, middle and minor axis directions

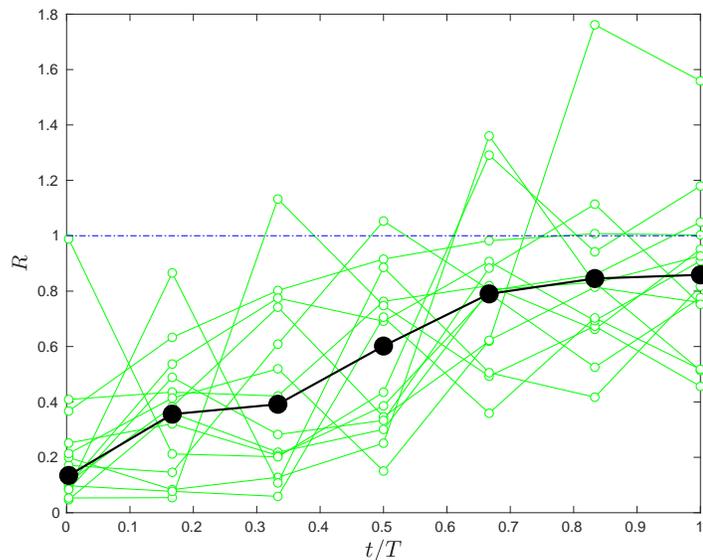
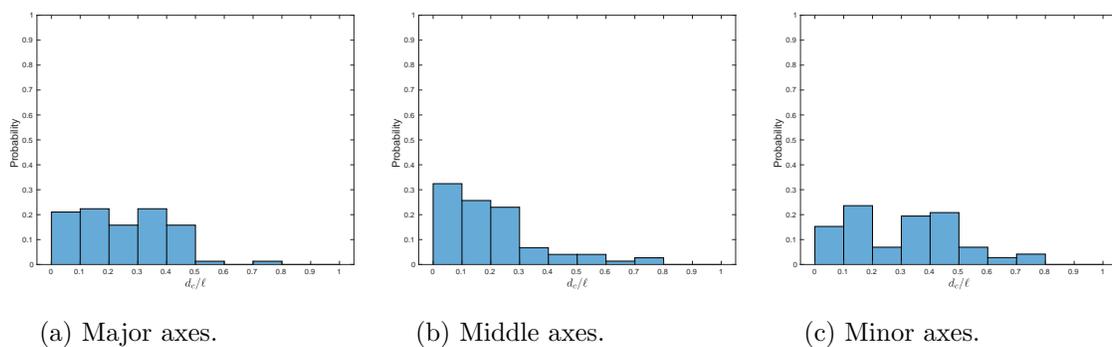


Figure 2.54: Median ratio between the sizes of the clusters for strong $|\tilde{\omega}|$ structures. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.

relative to the axis lengths, respectively.



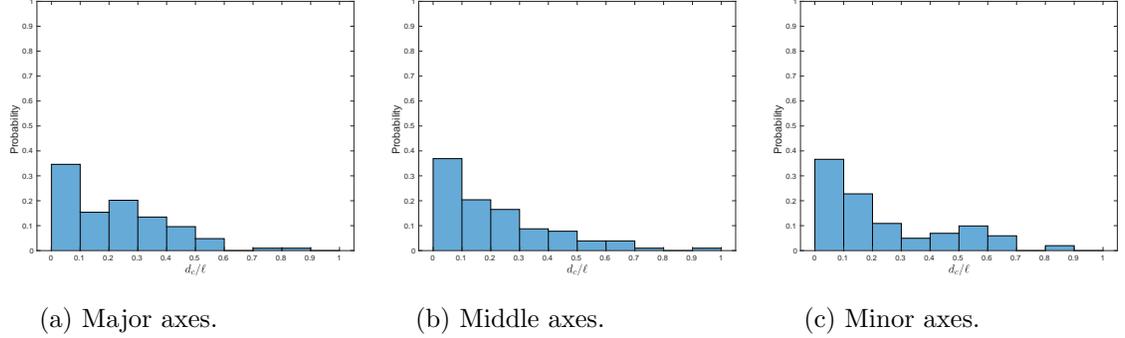
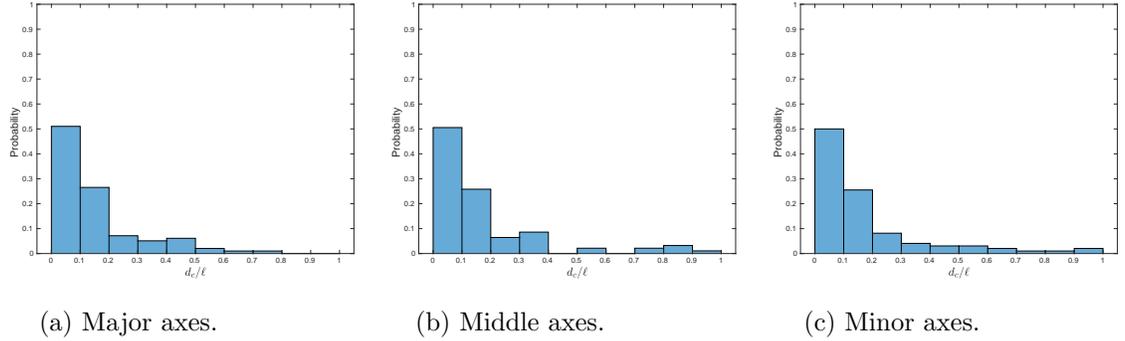
(a) Major axes.

(b) Middle axes.

(c) Minor axes.

Figure 2.55: d_c/ℓ for strong $|\tilde{\omega}|$ structures at $t = 0T$.

Fig. 2.55 shows that at the initial time there are wide spreads in d_c/ℓ , which implies that the MVEEs can be significantly displaced. However, the peaks around $d_c/\ell=0$ increase over time, showing that the agreement in the locations of the MVEEs enhances over time. There is no obvious difference between the axes. At $t = T$, the probability for $d_c/\ell \leq 0.2$ is approximately 80%.

Figure 2.56: d_c/ℓ for strong $|\tilde{\omega}|$ structures at $t = T/2$.Figure 2.57: d_c/ℓ for strong $|\tilde{\omega}|$ structures at $t = T$.

2.4.11 THE MINIMUM VOLUME ENCLOSED ELLIPSOID (MVEES) FOR THE STRAIN RATE FIELD

In this section, we present clusters analysis for the magnitude of the strain rate field. All the comparisons happen between the OIC and DNS target fields. The calculations are made with the filter scale equal $8\delta_x$. We use the value 2.5 times of the mean as the threshold value, i. e. structures with $|\tilde{S}| \geq 2.5\langle|\tilde{S}|\rangle$ are extracted.

The histograms for alignment are shown in Fig. 2.58 to 2.69. Fig. 2.58 shows that, at the initial time, the agreement is already visible in the major axes, as the probability value equals 0.6. The other conditional cases also show acceptable agreements where pancake-like MVEEs gives a high peak at 0.9. However, the conditional alignment for rod-like MVEEs shows a lower peak at 0.5.

The results for the other two axes, however, do not show clear alignment between the MVEEs from the two fields.

Fig. 2.61 to 2.63 Show that good alignment for the axes has already developed at $t = T/3$. The alignment between the minor axes is the strongest, which means the alignment in this direction is developed the earliest. The pancake-like MVEEs show

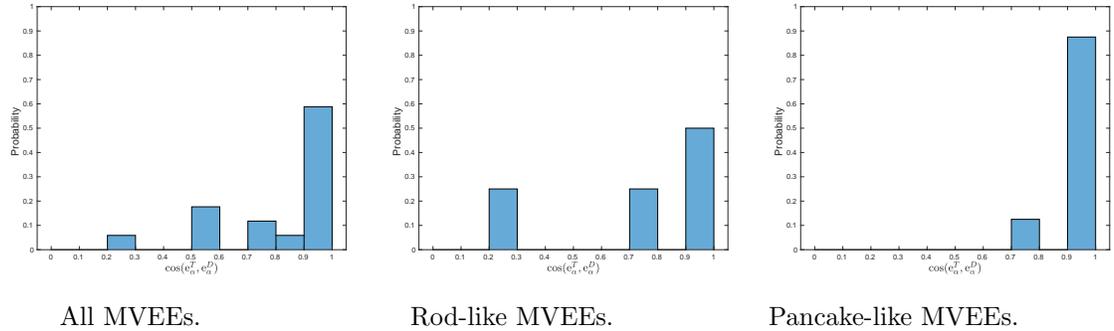


Figure 2.58: Alignment between the major axes for $|\tilde{S}|$ at $t = 0T$.

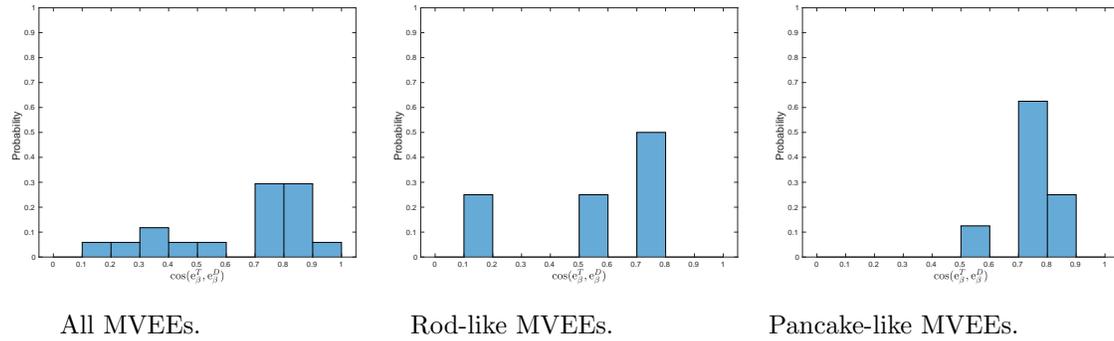


Figure 2.59: Alignment between the middle axes for $|\tilde{S}|$ at $t = 0T$.

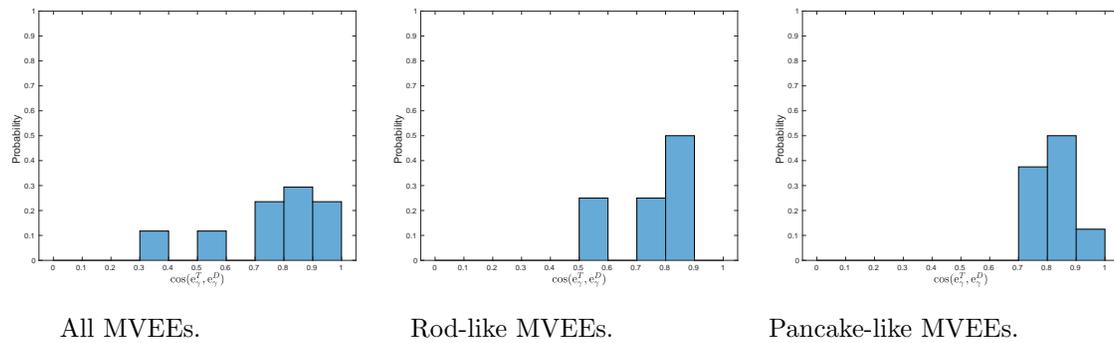


Figure 2.60: Alignment between the minor axes for $|\tilde{S}|$ at $t = 0T$.

better alignment in all three axis directions.

The alignment continues to improve over time, as shown by the results at $t = 2/3T$ (Fig. 2.64 to 2.66).

At $t = T$, the good agreement appears in Fig. 2.67, where the high probability for the

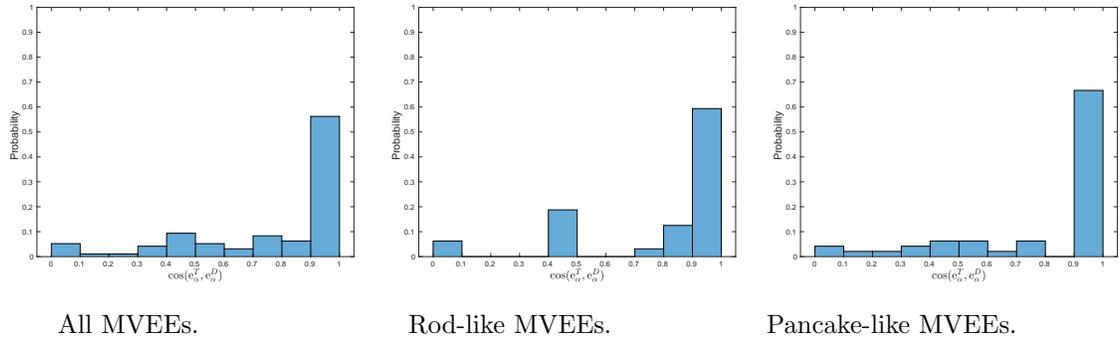


Figure 2.61: Alignment between the major axes for $|\tilde{S}|$ at $t = T/3$.

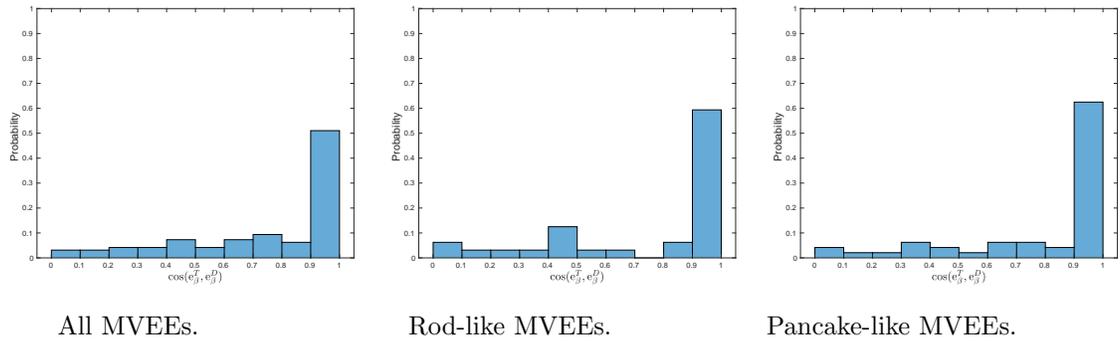


Figure 2.62: Alignment between the middle axes for $|\tilde{S}|$ at $t = T/3$.

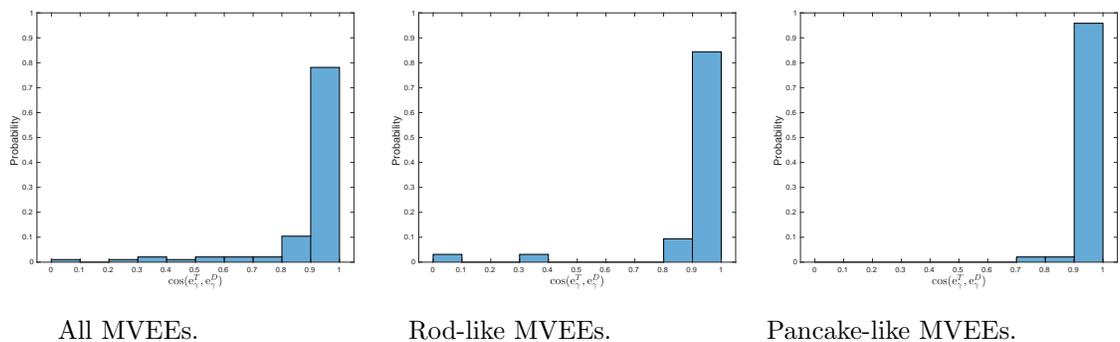


Figure 2.63: Alignment between the minor axes for $|\tilde{S}|$ at $t = T/3$.

major axes with a value 0.88 is found. Also, rod-like MVEEs and pancake-like MVEEs alignment reflect reasonable agreements with values of 0.98, and 0.79, respectively.

The probabilities for the middle axes also give acceptable alignments, since it gives a value of 0.72. The conditional alignment show close values with a peak for rod-like MVEEs valued at 0.74 and for pancake-like MVEEs at 0.78.

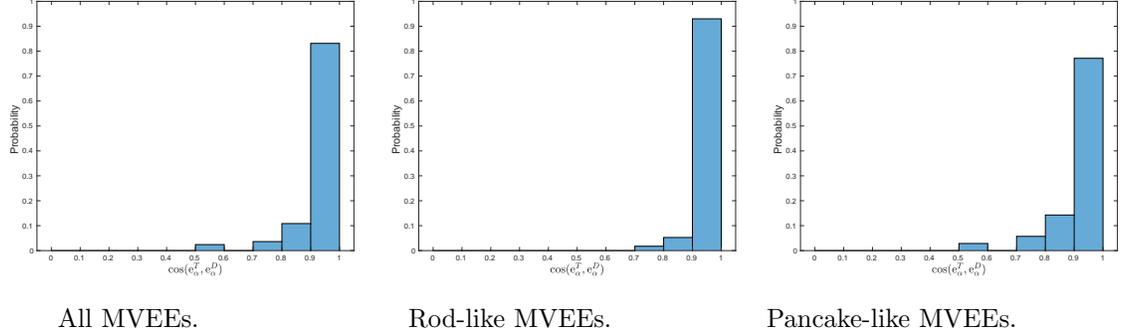


Figure 2.64: Alignment between the major axes for $|\tilde{S}|$ at $t = 2/3T$.

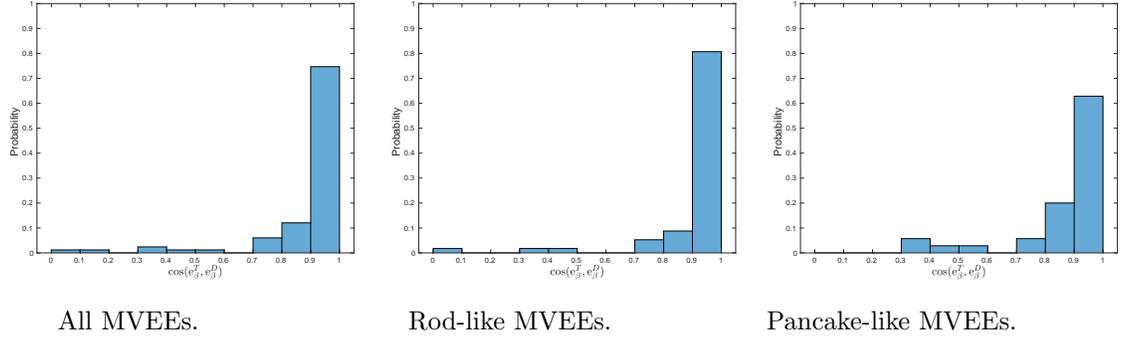


Figure 2.65: Alignment between the middle axes for $|\tilde{S}|$ at $t = 2/3T$.

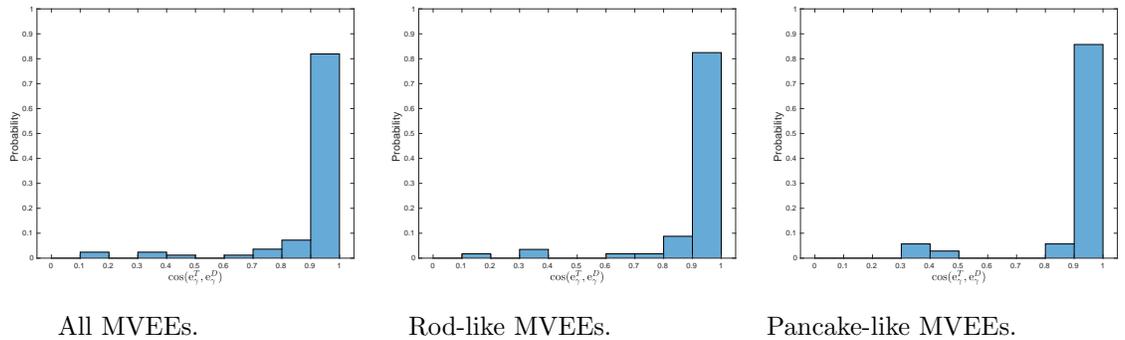


Figure 2.66: Alignment between the minor axes for $|\tilde{S}|$ at $t = 2/3T$.

The peak probability for the minor axes in Fig. 2.69 reaches 0.8. The conditional MVEEs alignments show a similar value for rod-like MVEEs and bigger for pancake-like MVEEs, found at 0.94.

Fig. 2.70 shows the median of the distributions. The red curve represents the median for the major axes. We can see that the median starts at 0.95, in the initial time, and

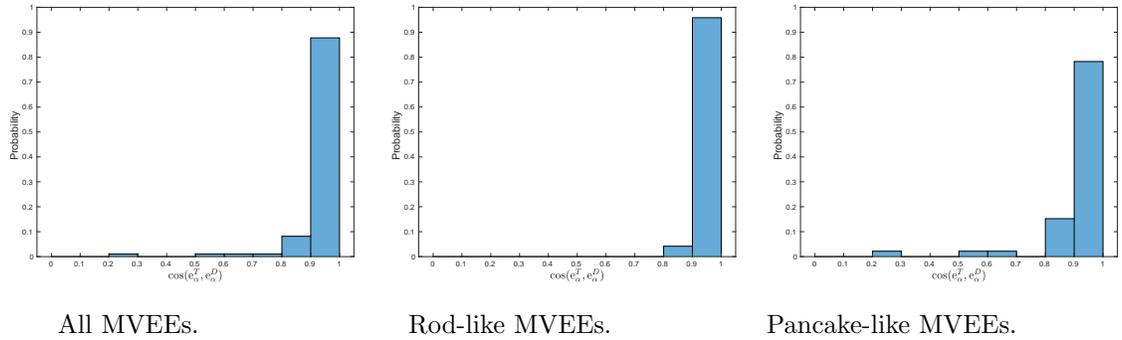


Figure 2.67: Alignment between the major axes for $|\tilde{S}|$ at $t = T$.

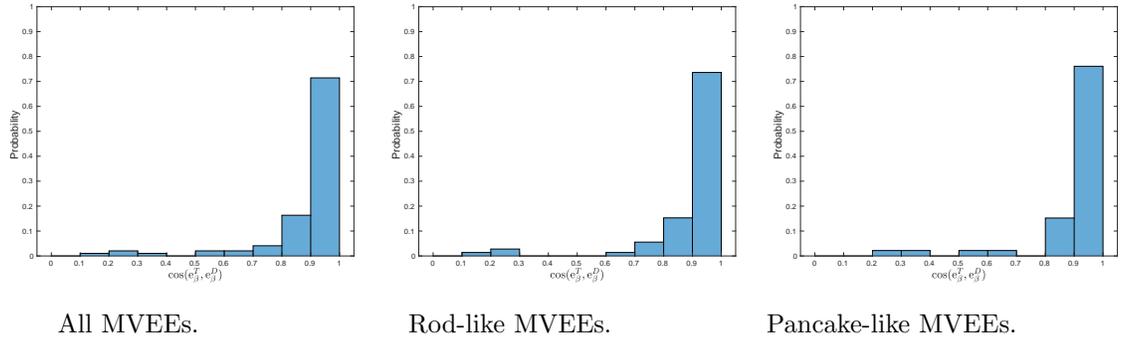


Figure 2.68: Alignment between the middle axes for $|\tilde{S}|$ at $t = T$.

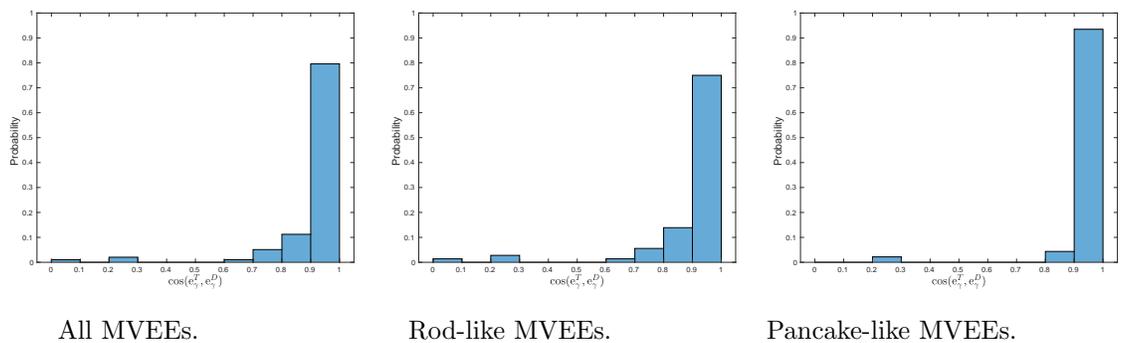


Figure 2.69: Alignment between the minor axes for $|\tilde{S}|$ at $t = T$.

reaches almost 1 at $t/T = 1$.

Fig. 2.71 shows median ratios for the sizes between the clusters of the OIC fields and the corresponding ones at the DNS target field.

The black and red curves and their symbols have the same meaning as we have presented

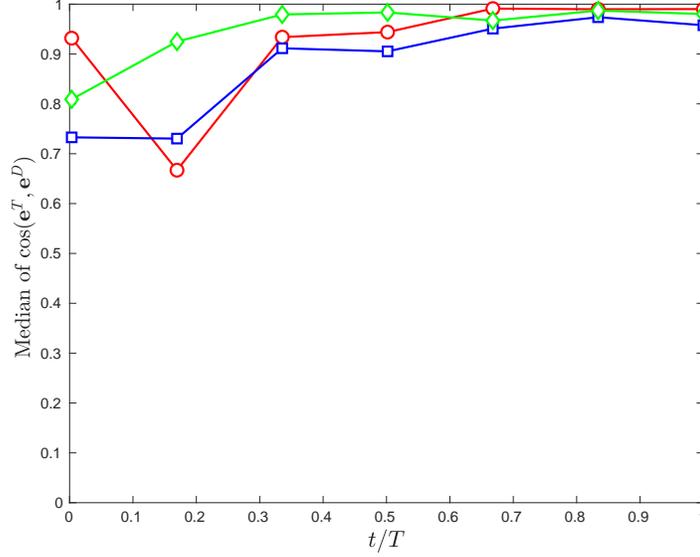


Figure 2.70: Medians of $|\cos(\mathbf{e}^T, \mathbf{e}^D)|$ for $|\tilde{S}|$. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.

before for $|\tilde{\omega}|$. The curves show how the median ratios develop with time, as we see from the black curve, starting at the time $T/6$, the median ratio is 0.35 and it grows to the value 0.7 in the end time T . In the same figure, the red lines refer to the median ratios for the clusters of a given size, which shows the variation of the ratios.

In what follows, we present the probability of the displacement of the centres of the MVEEs d_c/ℓ for the structures in the strain rate field $|\tilde{S}|$. We follow a similar procedure to the discussion of $|\tilde{\omega}|$.

Fig. 2.72 shows the displacement probability at the initial time. As we see that the major axes present a high peak in the smallest bin with a peak value of 0.5 which means the agreement for the major axes starts at an early time. However, it is the case for the other axes. We have weak agreements in middle and minor axes between the corresponding clusters.

For the middle time at $T/2$, Fig. 2.73 shows that the agreement has enhanced, as the peak probabilities for the middle and minor axes increase, although the major axes are slightly reduced from before.

Fig. 2.74 presents the displacement between the axes when the time equals T . The major axes increase again as they reach at 0.5 as a probability value. Also, the middle axes reach the highest peak at 0.44, while the minor axes slightly drop from $t = T/2$, but they improve from the initial time.

The histogram shows that there are about 60-70% chances for the relation displacement

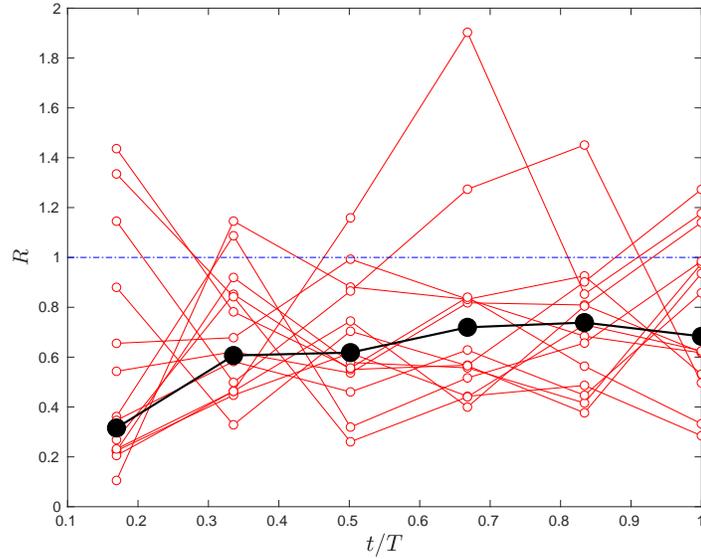


Figure 2.71: Median ratio between the sizes of the clusters for $|\tilde{S}|$. Black with circles: Median ratios for all clusters. Red with circles: median ratios for clusters of a given size.

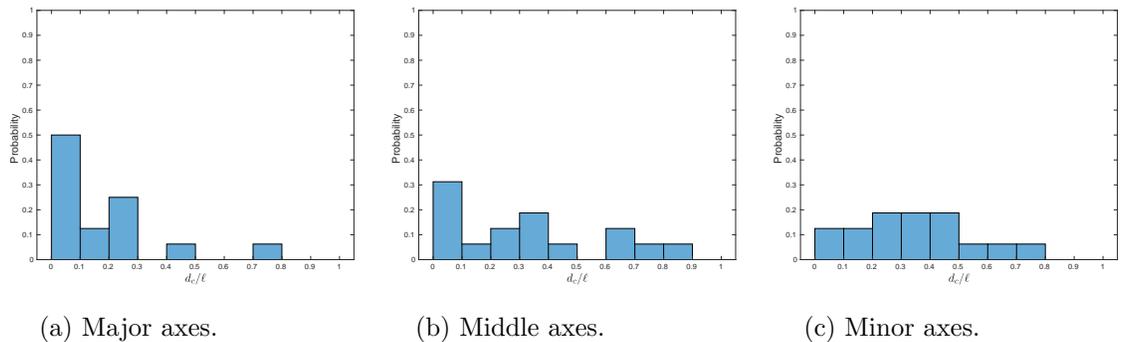
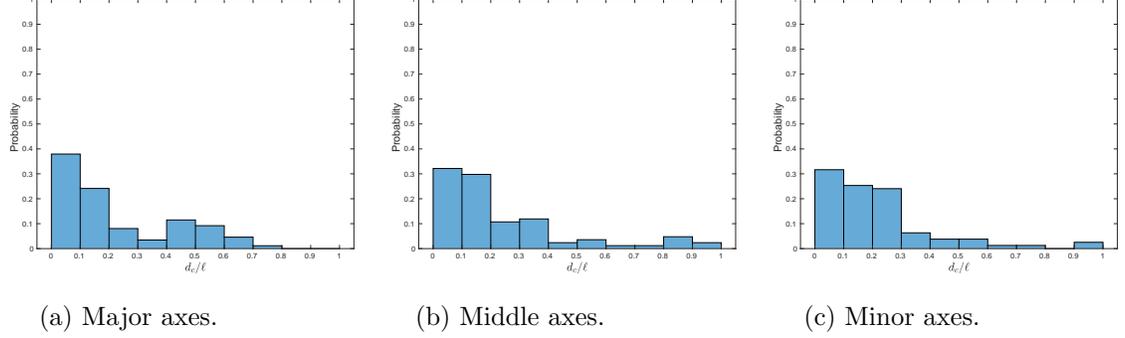
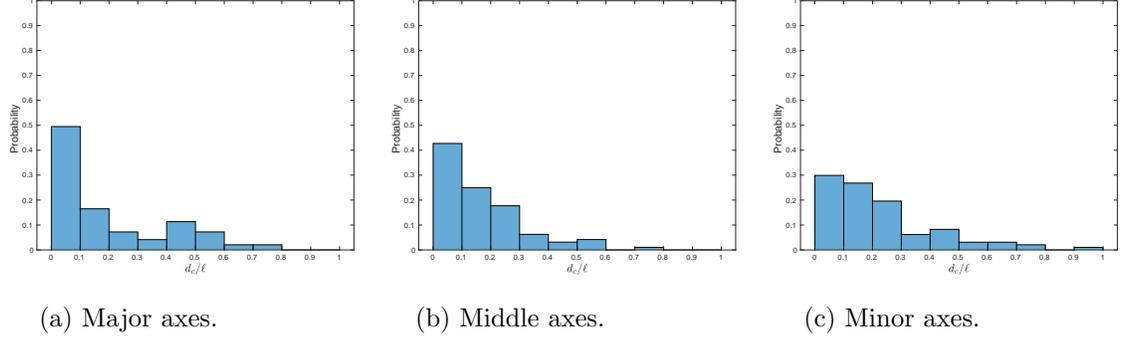


Figure 2.72: d_c/ℓ for $|\tilde{S}|$ at $t = 0T$.

to be less than 20%.

2.4.12 THE MINIMUM VOLUME ENCLOSED ELLIPSOID (MVEEs) FOR THE SUBGRID-SCALE ENERGY DISSIPATION

In the following section, we present some results related to the clusters analysis for the Subgrid-scale energy dissipation (SGS) quantities. For the calculation of the SGS energy dissipation MVEEs, we use 10 times of the mean as the threshold value, to extract the structures. We follow a similar procedure in presenting these results to

Figure 2.73: d_c/ℓ for $|\tilde{S}|$ at $t = T/2$.Figure 2.74: d_c/ℓ for $|\tilde{S}|$ at $t = T$.

what we presented for $|\tilde{\omega}|$ and $|\tilde{S}|$.

Fig. 2.75 to 2.86 plot the histograms of the alignment cosine. Fig. 2.87 plots the median of the corresponding histograms. The median ratios are given in Fig. 2.88 and finally the histograms of d_c/ℓ are shown in Fig. 2.89 to 2.91. Fig. 2.87 plots the median of the corresponding histograms.

These results show similar trends as these for $|\tilde{\omega}|$ and $|\tilde{S}|$. Reasonable agreement again is observed at $t/T = 1$.

2.4.13 COMPARISONS BETWEEN THE THREE TYPES OF STRUCTURES

We present in this section comparisons between the clusters for the three quantities (i.e. vorticity magnitude, strain rate, and subgrid-scale energy dissipation).

As we have mentioned before, the corresponding clusters have good agreements for all three quantities. However, we also see that the developments of the clusters for the three quantities have shown different behaviours although all have tended to be closer to the ones in the DNS target over time. The main point is that the agreements between

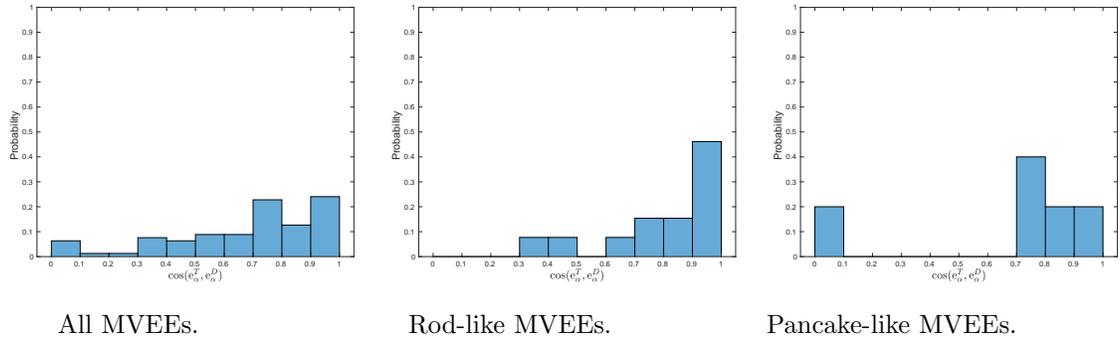


Figure 2.75: Alignment between the major axes for SGS energy dissipation at $t = 0T$.

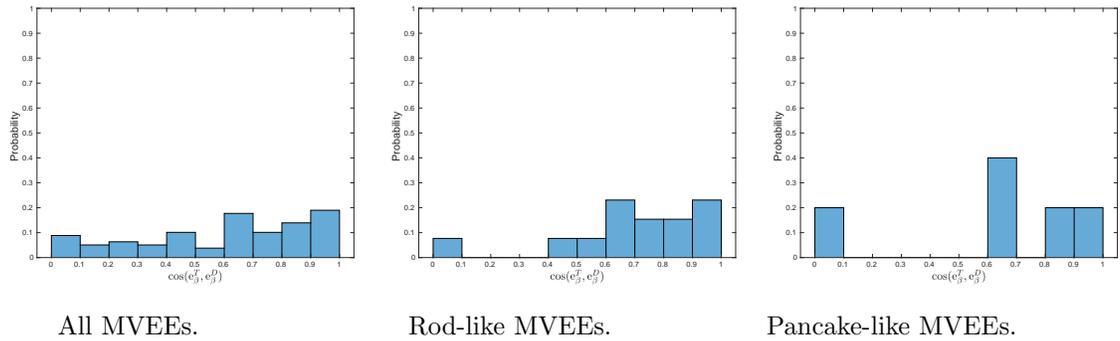


Figure 2.76: Alignment between the middle axes for SGS energy dissipation at $t = 0T$.

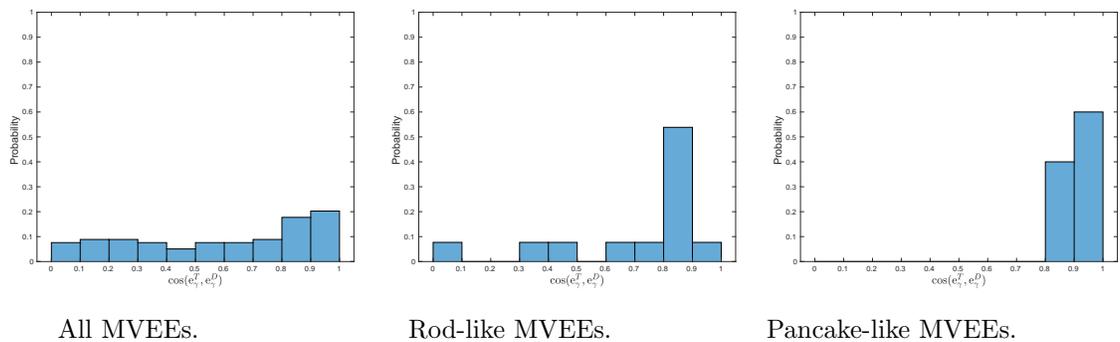


Figure 2.77: Alignment between the minor axes for SGS energy dissipation at $t = 0T$.

the clusters are improved after a short time from the initial time for the major, middle and minor axes, reaching the highest probability values close to T . Also, the major axes present the best agreement, then the minor axes, while the middle axes show the lowest agreement.

Variations certainly exist between the three different quantities, as the vorticity magni-

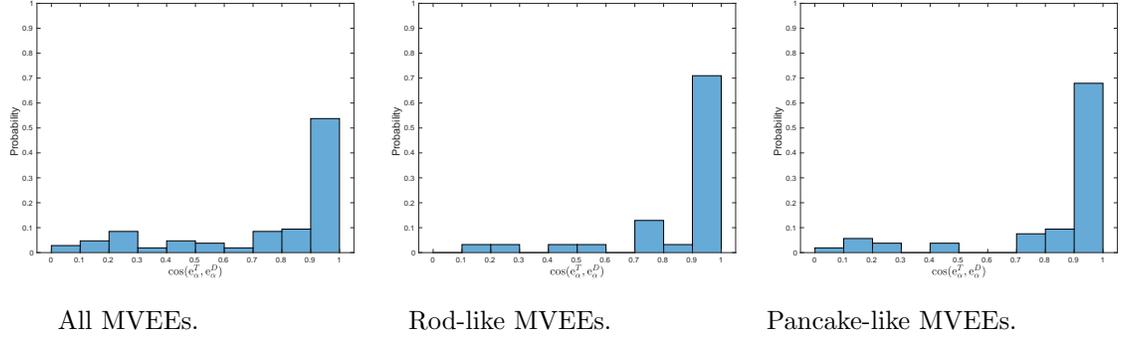


Figure 2.78: Alignment between the major axes for SGS energy dissipation at $t = T/3$.

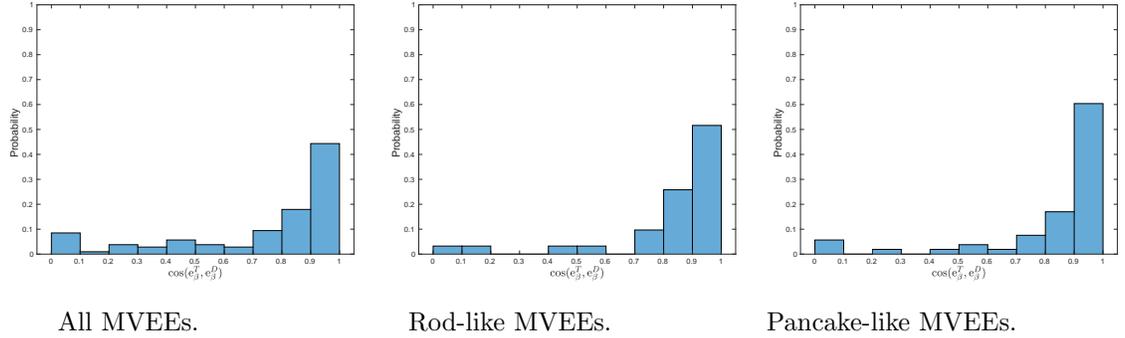


Figure 2.79: Alignment between the middle axes for SGS energy dissipation at $t = T/3$.

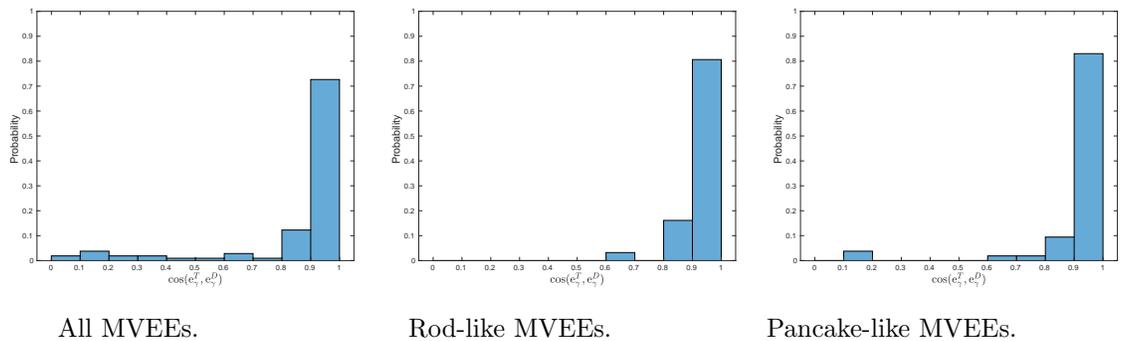


Figure 2.80: Alignment between the minor axes for SGS energy dissipation at $t = T/3$.

tude and the strain rate magnitude are similar than the SGS energy dissipation. One of the observations is that both the vorticity and strain rate starts at the initial time with a reasonable agreement for the major axes, while the initial agreement for SGS energy dissipation is much lower. However, this argument is different for the middle axes, for which the strain rate field presents the lowest agreement compared with the vorticity

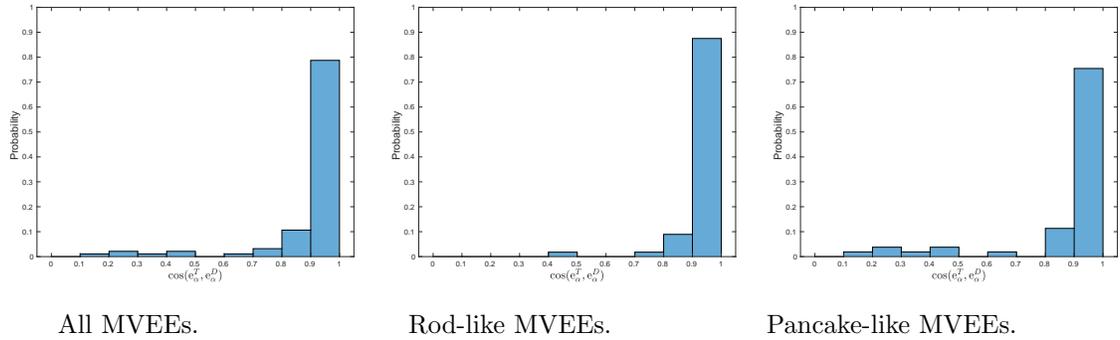


Figure 2.81: Alignment between the major axes for SGS energy dissipation at $t = 2/3T$.

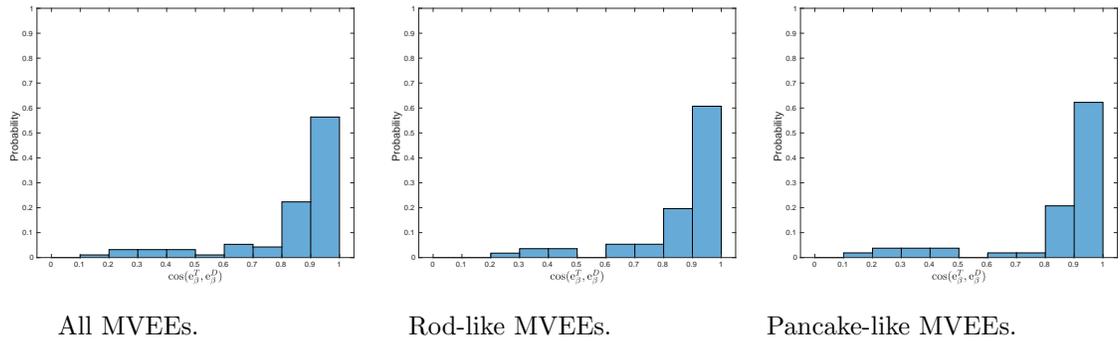


Figure 2.82: Alignment between the middle axes for SGS energy dissipation at $t = 2/3T$.

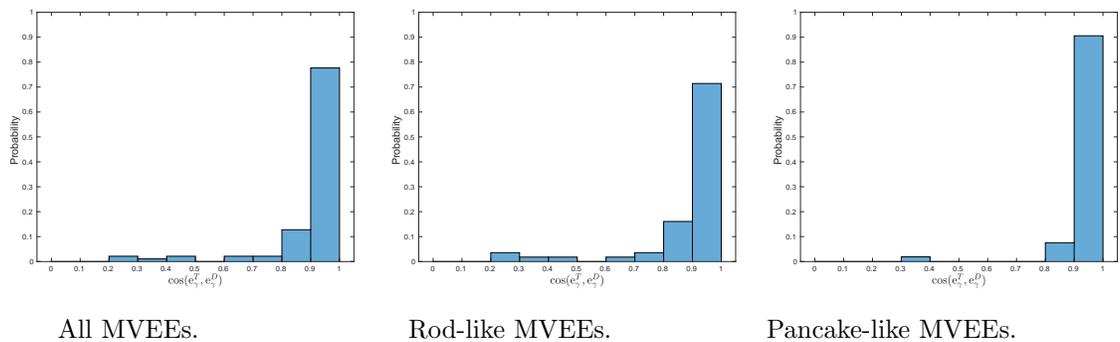


Figure 2.83: Alignment between the minor axes for SGS energy dissipation at $t = 2/3T$.

and SGS energy dissipation. Adding to that, the agreement between the minor axes when the time reaches T is lower for the vorticity than the other two quantities. In general, the vorticity shows the best agreement, then the strain rate field, and finally the SGS energy dissipation.

The results for the size ratios show that the vorticity has the highest ratio at around

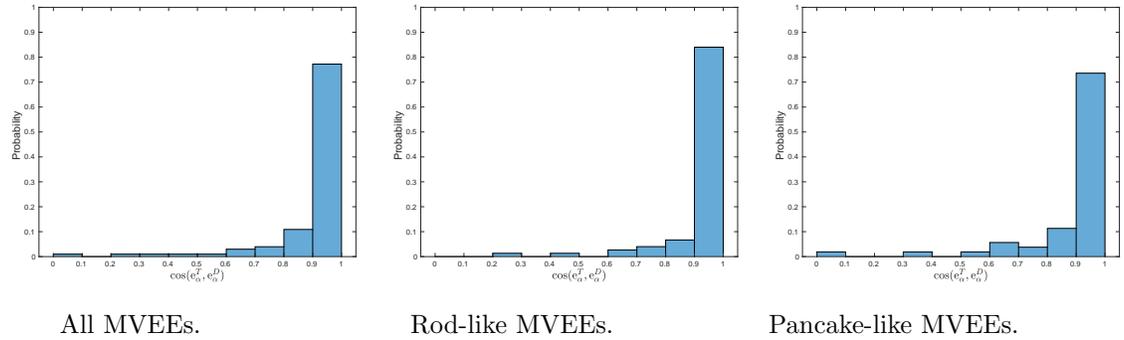


Figure 2.84: Alignment between the major axes for SGS energy dissipation at $t = T$.

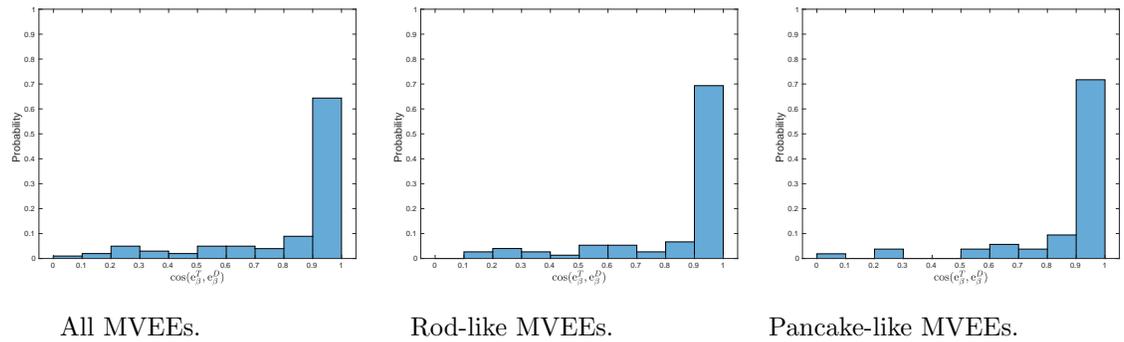


Figure 2.85: Alignment between the middle axes for SGS energy dissipation at $t = T$.

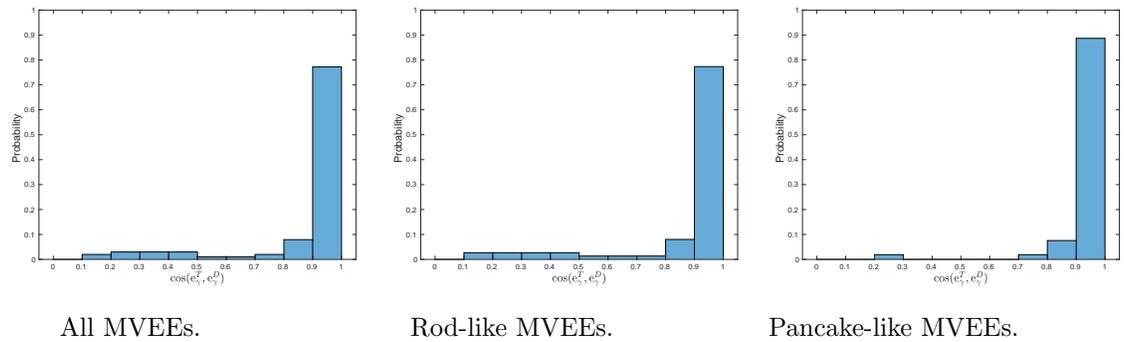


Figure 2.86: Alignment between the minor axes for SGS energy dissipation at $t = T$.

0.9 while the strain rate field was the lowest one which is just under 0.8.

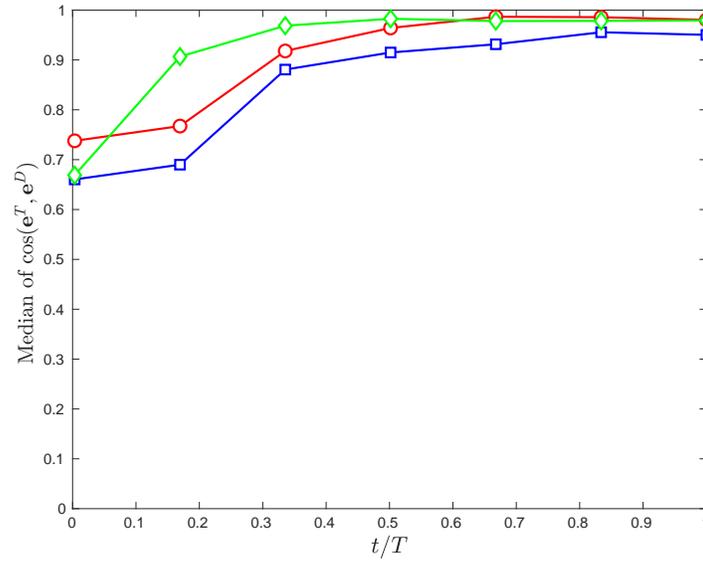


Figure 2.87: Medians of $|\cos(e^T, e^D)|$ for SGS energy dissipation. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.

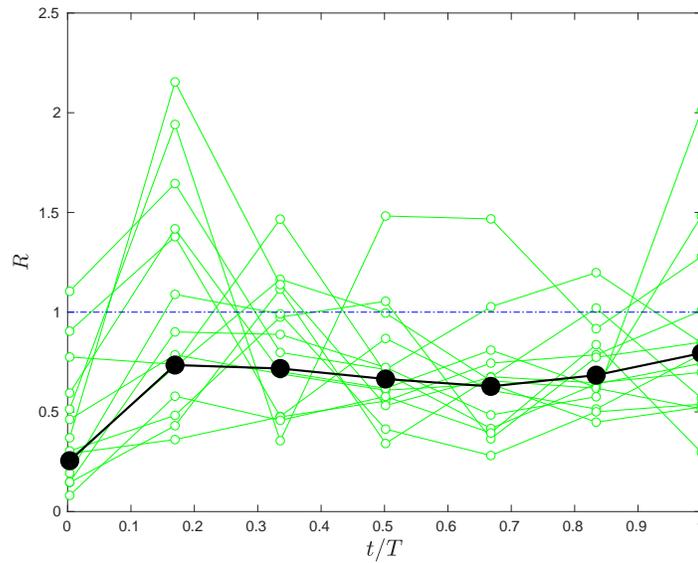
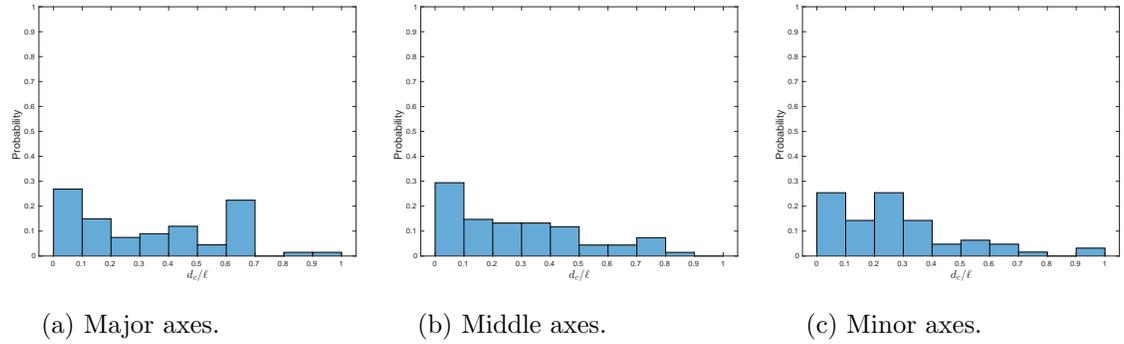
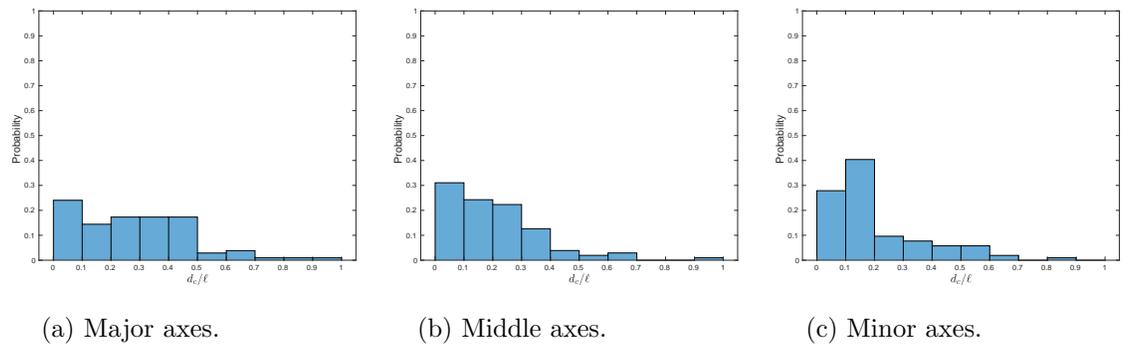
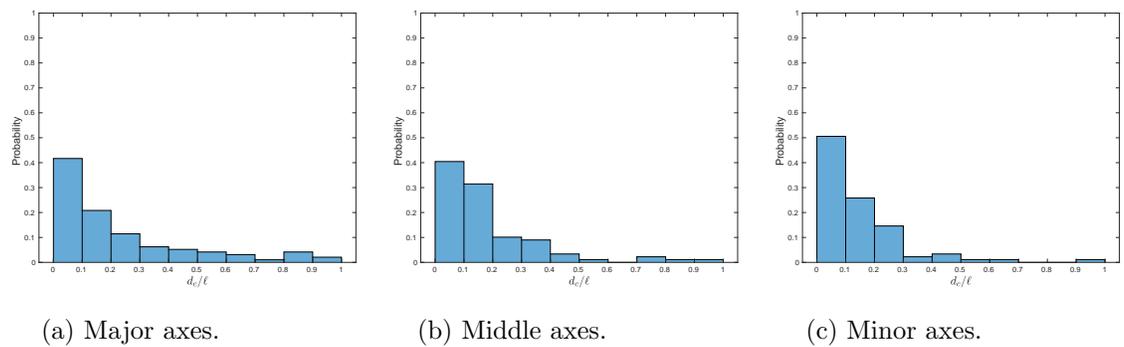


Figure 2.88: Median ratio between the sizes of the clusters for SGS energy dissipation. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.

Figure 2.89: d_c/ℓ for SGS energy dissipation at $t = 0T$.Figure 2.90: d_c/ℓ for SGS energy dissipation at $t = T/2$.Figure 2.91: d_c/ℓ for SGS energy dissipation at $t = T$.

Chapter 3

Reconstructing turbulent fields using 4D variational data assimilation: large eddy simulations and the Smagorinsky model

§ 3.1 Introduction

The model is presented in this chapter is to examine the effects on the reconstruction quality when the large eddy simulations are applied as the model using the Smagorinsky model. The presented assumption is that the dynamical evolution of the system is governed by the filtered Navier-Stokes equation. The 4D variational DA is applied to reconstruct the initial state of the system. The formulation of the system is given as an optimization problem. The main purpose of this investigation is to reconstruct the instantaneous distributions for the small scales using such a model.

The structure of this chapter can be displayed as follows. Section 3.2 presents a brief introduction of the large eddy simulations. In section 3.3, the formulation of the optimization problem is presented including the description of the simulations based on the filtered NS equations. The simulation and discussions part are given in section 3.4, where the target data fields and the data of the reconstructed initial condition based on the imperfect model are compared to test the quality of the data generated by such model. The presented experiments are similar to what has been used in the previous chapter when the NS equations as the model have used. Also, the discussions include a comparison between the outcomes of both optimal fields.

§ 3.2 An introduction to large eddy simulations

The numerical solutions for the NS equations of turbulent flows are called Direct Numerical Simulation, as we mentioned before. However, the cost of the computation is a rapidly increasing function of the Reynolds number of the flow [73]. When the Reynolds number is very high, DNS calculations in the near future are not feasible. In practical applications, a recently developed method is large eddy simulations. While the large-scale motions are represented directly in the large eddy simulations, a model is used to represent the smaller scale motions in the same flow; see [73].

§ 3.3 Formulation for the 4DVAR problem

The problem setup is similar to the one in Chapter 2. We assume:

- A set of data $\mathbf{v}(\mathbf{x}, t)$ is known, for $t \in [0, T]$ and $\mathbf{x} \in \Omega$, where T is the optimization horizon.
- \mathbf{v} is a part of the solution of the NS equations for unknown IC. $\bar{\mathbf{u}}$ is a solution of the filtered NS equations with the Smagorinsky model for the SGS stress, and $\bar{\mathbf{u}}(\mathbf{x}, 0) = \boldsymbol{\varphi}(\mathbf{x})$.

The goal, is to find $\boldsymbol{\varphi}(\mathbf{x})$, such that the solution $\bar{\mathbf{u}}(\mathbf{x}, t)$ matches $\mathbf{v}(\mathbf{x}, t)$ for $t \in [0, T]$ and $\mathbf{x} \in \Omega$. The solution method is similar to the one used in Chapter 2. The cost function is defined in the same way:

$$J = \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} |\bar{\mathbf{u}} - \mathbf{v}|^2, \quad (3.1)$$

J is minimized subject to constraints:

$$\begin{aligned} \mathbf{N}(\bar{\mathbf{u}}) &= \partial_t \bar{u}_i + \bar{u}_j \partial_j \bar{u}_i = 2\partial_j \left((\nu + (C_S \Delta)^2 (2\bar{S}_{lk} \bar{S}_{lk})^{1/2}) \bar{S}_{ij} \right) - \partial_i \bar{p} + \bar{f}_i, \quad (i = 1, 2, 3), \\ \nabla \cdot \bar{\mathbf{u}} &= 0, \\ \bar{\mathbf{u}}(\mathbf{x}, 0) &= \boldsymbol{\varphi}(\mathbf{x}). \end{aligned} \quad (3.2)$$

The Lagrangian function is given as

$$\begin{aligned} L &= J - \int_0^T \int_B \boldsymbol{\xi} \cdot \mathbf{N}(\bar{\mathbf{u}}) dt d\mathbf{x} - \int_0^T \int_B \mu \nabla \cdot \bar{\mathbf{u}} dt d\mathbf{x} \\ &\quad - \int_B \boldsymbol{\lambda} \cdot [\bar{\mathbf{u}}(\mathbf{x}, 0) - \boldsymbol{\varphi}(\mathbf{x})] d\mathbf{x}, \end{aligned} \quad (3.3)$$

where $\boldsymbol{\xi}(\mathbf{x}, t)$, $\mu(\mathbf{x})$, and $\boldsymbol{\lambda}(\mathbf{x})$ are adjoint variables corresponding to the constraints. The state equations are given by

$$\frac{\partial L}{\partial \boldsymbol{\xi}} = 0, \frac{\partial L}{\partial \mu} = 0, \frac{\partial L}{\partial \boldsymbol{\lambda}} = 0. \quad (3.4)$$

The adjoint equations are given by

$$\frac{\partial L}{\partial \bar{\mathbf{u}}} = 0, \frac{\partial L}{\partial \bar{p}} = 0. \quad (3.5)$$

Therefore, the state equations are

$$\begin{aligned} \partial_t \bar{u}_i + \bar{u}_j \partial_j \bar{u}_i - 2\partial_j ((\nu + (C_S \Delta)^2 (2\bar{S}_{lk} \bar{S}_{lk})^{1/2}) \bar{S}_{ij}) + \partial_i \bar{p} - \bar{f}_i &= 0, \quad (i = 1, 2, 3.) \\ \nabla \cdot \bar{\mathbf{u}} &= 0, \\ \bar{\mathbf{u}}(\mathbf{x}, 0) &= \boldsymbol{\varphi}(\mathbf{x}). \end{aligned} \quad (3.6)$$

The adjoint equations are

$$\begin{aligned} \partial_t \boldsymbol{\xi} + \bar{\mathbf{u}} \cdot \nabla \boldsymbol{\xi} - \nabla \bar{\mathbf{u}} \cdot \boldsymbol{\xi} + \nabla \mu + \nu \nabla^2 \boldsymbol{\xi} + (C_S \Delta)^2 \nabla \cdot \left[\frac{2X(\bar{S} : \bar{S}) + 2\bar{S}^T(\bar{S} : X)}{|\bar{S}|} \right] &= \mathbf{F}, \\ \nabla \cdot \boldsymbol{\xi} &= 0, \\ \boldsymbol{\xi}(\mathbf{x}, t = T) &= 0. \end{aligned} \quad (3.7)$$

where $\mathbf{F}(\mathbf{x}, t) = -[\bar{\mathbf{u}}(\mathbf{x}, t) - \mathbf{v}(\mathbf{x}, t)]$ when $\mathbf{x} \in \Omega$ and $\mathbf{F} = 0$ otherwise. We note that $\boldsymbol{\xi}$ represents the velocity adjoint, μ is the pressure adjoint. Also, $X_{ki} = \frac{1}{2}(\partial_k \xi_i + \partial_i \xi_k)$. The gradient of the cost function

$$\frac{DJ}{D\boldsymbol{\varphi}} = \boldsymbol{\xi}(\mathbf{x}, 0), \quad (3.8)$$

is given by the adjoint. It is needed to update the initial condition $\boldsymbol{\varphi}$.

§ 3.4 Simulation and discussions

The optimal solution for $\boldsymbol{\varphi}$ is the reconstructed/rebuilt initial condition (OICS) based on the imperfect model (the filtered NS equations). An interesting question is how well $\boldsymbol{\varphi}$ mimics real turbulence. Therefore, the statistics of $\boldsymbol{\varphi}$ are compared with those in turbulent fields. In the comparisons, we refer to $\boldsymbol{\varphi}$ as the OICS field, while we refer to the DNS field as the DNS target field. The results will be compared with those in

Chapter 2 as well. This comparison will reveal the deficit of using an imperfect model to assimilate the data, if any.

Before starting the analysis of the results, we need a description of the simulations based on the filtered NS equations that we use in this chapter.

- We use a three-dimensional pseudo-spectral code for isotropic turbulence where the boundary conditions are periodic in a domain $[0, 2\pi]^3$. The number of grid points N^3 is 64^3 . The highest resolved wave-number is $k_{max} = N/2 = 32$, and the grid size is $\delta_x = (2\pi)/N = \pi/32$. The filter scale used in the analysis of the data is $4\delta_x$.
- kinetic energy is injected into the flow field by using the term of the force in filtered NS equations which is given as $\bar{\mathbf{f}} = (0, f_a \cos k_f x_1, 0)$ where $f_a = 0.15$ and $k_f = 1$.
- A set of velocity data $\mathbf{v}(\mathbf{x}, t)$ is known, for $t \in [0, T]$ and $\mathbf{x} \in \Omega$. The data assimilation is accomplished in this model in the spectral space in a similar procedure as mentioned in 2.4 when we have used the reconstruction using 4DVAR DA with the NS equations as the model.
- Optimum is found when $J/J(\bar{\mathbf{u}} = 0) \leq e$, where tolerance $e = 0.0002$.
- We use the kinematic viscosity $\nu = 0.006$.
- The optimization horizon is up to $T = 0.5\tau_L$, where $\tau_L \approx 4.35$ is the large eddy turn-over time scale.

The energy spectra of the DNS target and the OICS Fields are presented at Fig. 3.1. The observation at Fig. 3.1a shows that the error between the two fields is bigger than the late times and specifically, at high wave-numbers, i. e. small scales. Figs. 3.1b to 3.1d show that the agreement improves over time as expected. A perfect agreement is observed for the energy spectrum for $k \leq 4$ the whole time because, by our procedure, the large scales are matched with the DNS target data. Comparisons with Fig. 2.1 shows that the behaviours are similar.

The energy spectra of the instantaneous differences $\mathbf{v} - \bar{\mathbf{u}}$ is presented in the blue curves with the $E(k)$ of the target field in the red curves at Fig. 3.2. The behaviour of the blue curve shows that the spectra start very small in the low wave-numbers region for $k \leq K_c$ which indicates a good agreement between the imperfect model with the target data at an early time. Furthermore, Figs. 3.2b to 3.2d reflect the agreement's improvement over time. Also, such an improvement of the agreement between both fields can be observed for $k > K_c$, in which the value of the blue curve beyond $K_c = 4$ decreases overtime where the high wave-numbers are located. This indicates that the reconstruction of the small scales has happened broadly, where the differences between $\bar{\mathbf{u}}$ and \mathbf{v} become smaller as the optimization horizon time increases. Thus, the successful reconstruction of the different size scales has been achieved in this experiment.

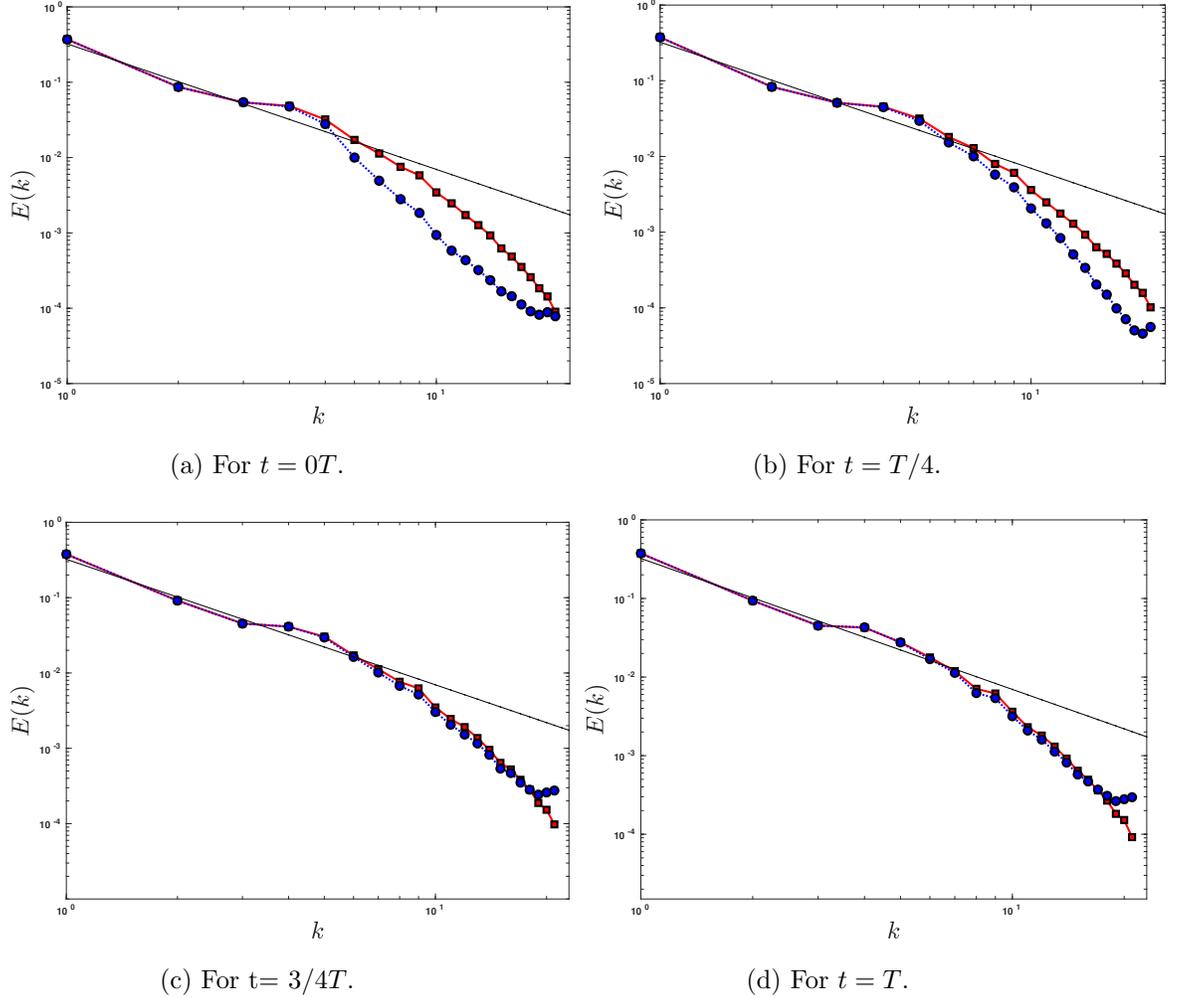


Figure 3.1: The energy spectrum. Solid line with squares: DNS target. Dotted line with circles: OICS. Dashed line: Kolmogorov $-5/3$ spectrum.

In term of the comparisons between the two models, the OICS field seems to recover the DNS target field quicker than the OIC field, since the spectra at $t = 3/4T$ gives a smaller difference than the ones in Fig. 2.2 when $t = 2/3T$.

We look also into the ratio $E_\delta(k)/E(k)$. We notice that the OICS fields tend to the DNS target field as the time increases. These positive observations are shown in Figs. 3.3a and 3.3b. Fig. 3.3b shows that, at $k = 8$ and $t = T$, the ratio can be reduced to about 20%, significantly smaller than that one shown in Fig. 2.3.

We now consider the alignment between the strain rate field \tilde{S}_{ij} and the vorticity field $\tilde{\omega}$. Fig. 3.4 plots the PDFs of the alignment between the vorticity and strain rate fields for the OICS and DNS target fields. The PDFs at the initial time in Fig. 3.4a shows

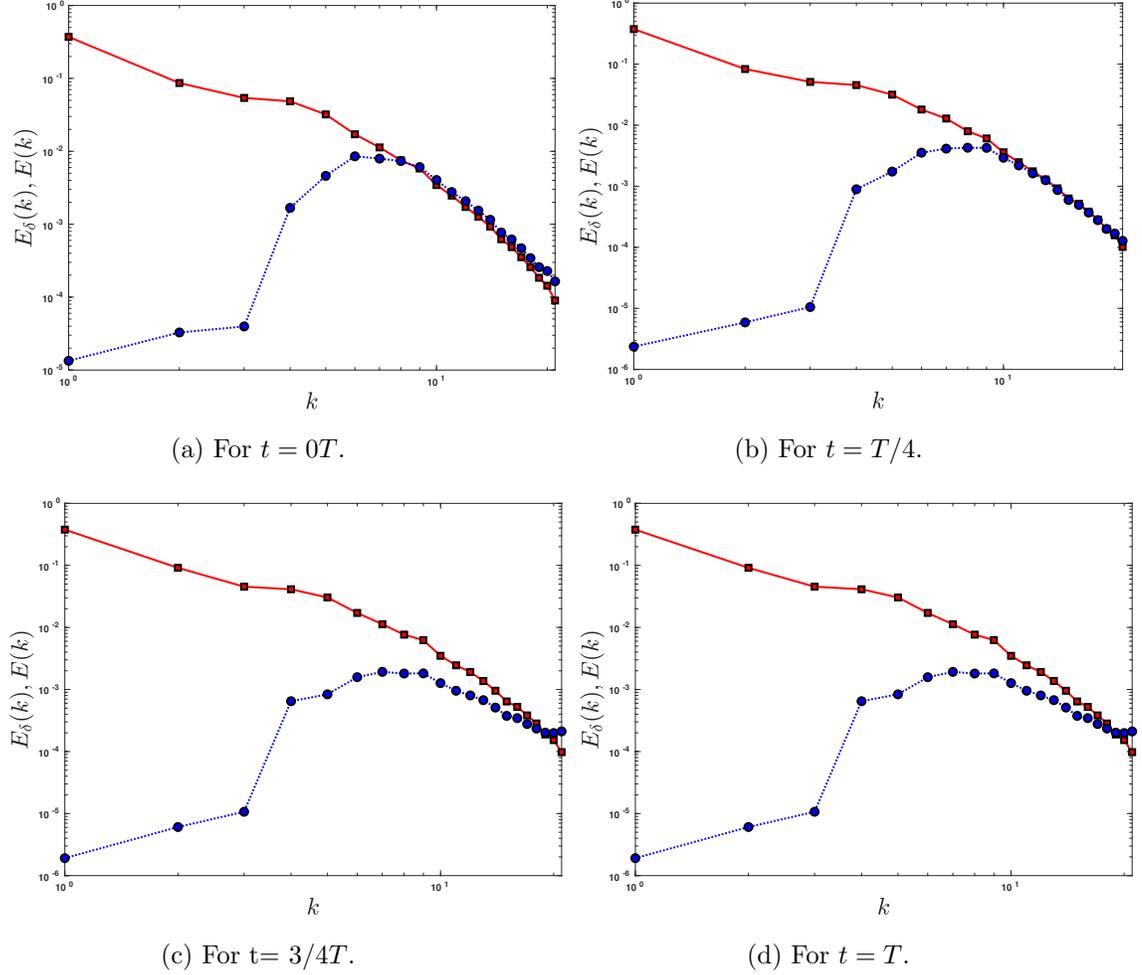
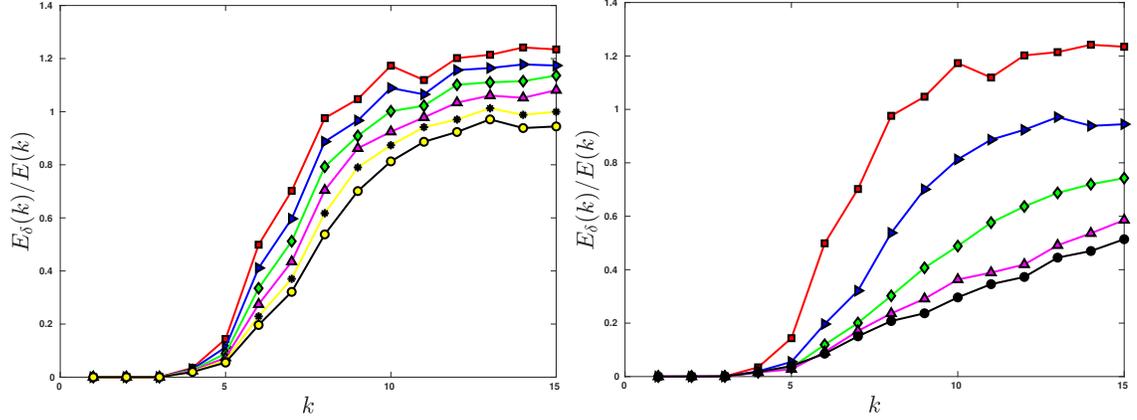


Figure 3.2: Spectrum of $\mathbf{v} - \bar{\mathbf{u}}$. Solid line with squares: $E(k)$ for DNS target. Dotted line with circles: $E_\delta(k)$.

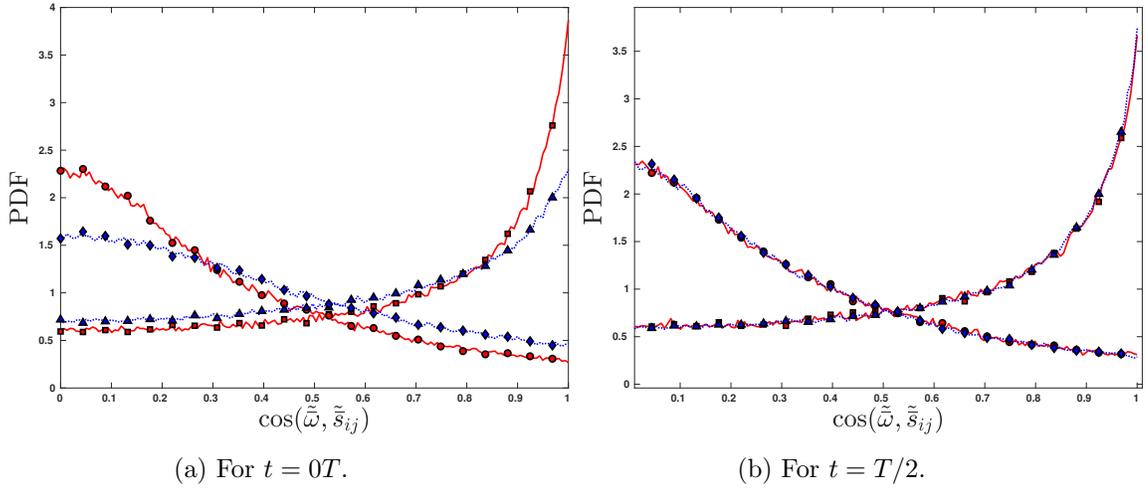
a significant result of OICS fields, as both PDFs deviate from the corresponding PDFs of the DNS target fields by only a small amount. The agreement improvement occurs quickly after a short time, to reach the perfect agreement at $T/2$, shown in Fig. 3.4b. The comparison between the OIC fields in Fig. 2.4 and OICS fields in Fig. 3.4 show that the current model speeds up the agreement with DNS target field, and obtains better reconstruction at $t = 0$.

The observation in Fig. 3.5a for the mean SGS energy dissipation as a function of filter scale Δ shows that the curve of the OICS field at the initial time underestimates its corresponding DNS target curve. This difference, however, decreases when the time equals $T/4$ in Fig. 3.5b. Good agreements are continued until time T . The probability density functions of the normalised SGS energy dissipation at scale $\Delta = 4\delta_x$



(a) Solid line with: Square for $0T$. Downward-pointing triangle for $T/20$. Diamond for $T/10$. Upward-pointing triangle for $3/20T$. Asterisk for $T/5$. Circle for $T/4$.
 (b) Solid line with: Square for $0T$. Downward-pointing triangle for $T/4$. Diamond for $T/2$. Upward-pointing triangle for $3/4T$. Circle for T .

Figure 3.3: The ratio of the spectra: $E_\delta(k)/E(k)$.

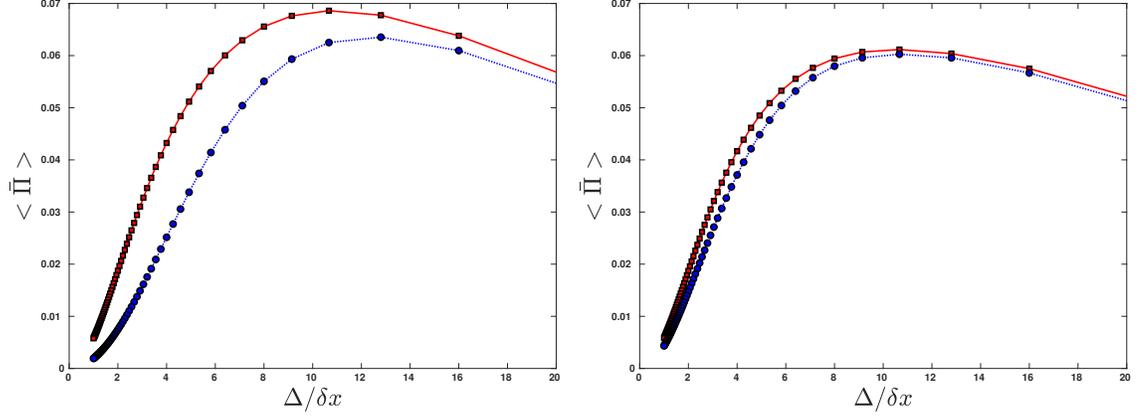


(a) For $t = 0T$.

(b) For $t = T/2$.

Figure 3.4: PDFs of the alignment between $\tilde{\omega}$ and the eigenvectors of \tilde{S}_{ij} . DNS target: solid line with squares, circles: intermediate eigen-direction and eigen-direction with negative eigenvalue, respectively. OICS: dotted line with upward-pointing triangle, diamond, respectively.

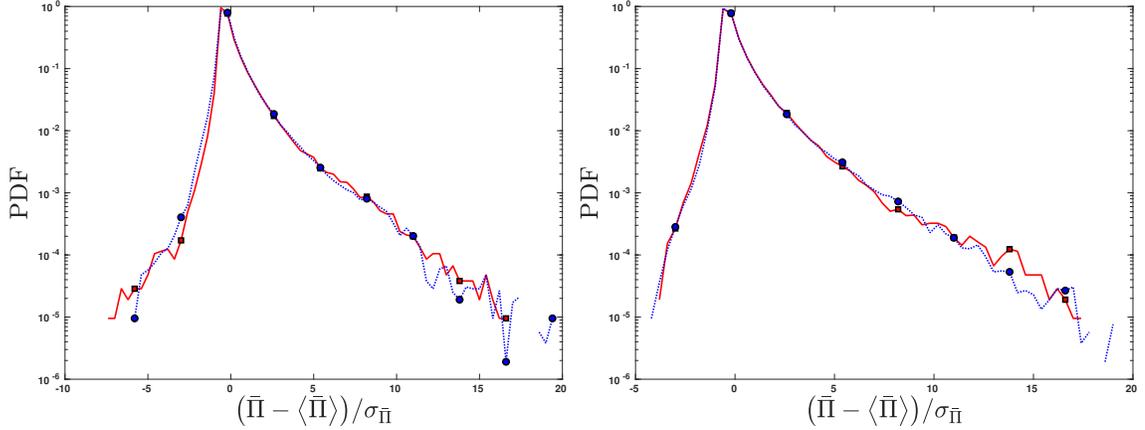
are presented in Fig. 3.6. Fig.3.6a shows the agreement between the OICS and DNS target fields at the initial time. The agreement improves when the time increases. Thus, the feature of the PDFs of the SGS energy dissipation is reproduced by the OICS fields, with positive skewness that indicates the forward energy cascade.



(a) Solid line is DNS target: with squares for $0T$. Dotted line is OICS: with circles for $0T$.

(b) Solid line is DNS target: squares for $T/4$. Dotted line is OICS: with circles for $T/4$.

Figure 3.5: The mean subgrid-scale energy dissipations as a function of the filter scales Δ .



(a) Solid line is DNS target: with squares for $0T$. Dotted line is OICS: with circles for $0T$.

(b) Solid line is DNS target: squares for $T/2$. Dotted line is OICS: with circles for $T/2$.

Figure 3.6: PDFs of the subgrid-scale energy dissipation.

The probability density functions of the shape parameter for the strain rate tensor are calculated in what follows. Fig. 3.7 shows that the PDF of S^* at the initial time for the OICS underestimates than the corresponding PDF of the target data. However, the PDF recovers quickly when the time equals $T/8$, such that we have quite a good agreement already. At later time steps, all the PDFs of the OICS fields give similar features as the PDFs of the DNS field. Fig. 3.7, compared with Fig. 2.8, shows a quicker recovering for the OICS field and a better reconstruction at $t = 0$.

We calculate the probability density functions of the shape factor $S^*(\bar{\tau}_{ij})$ for the

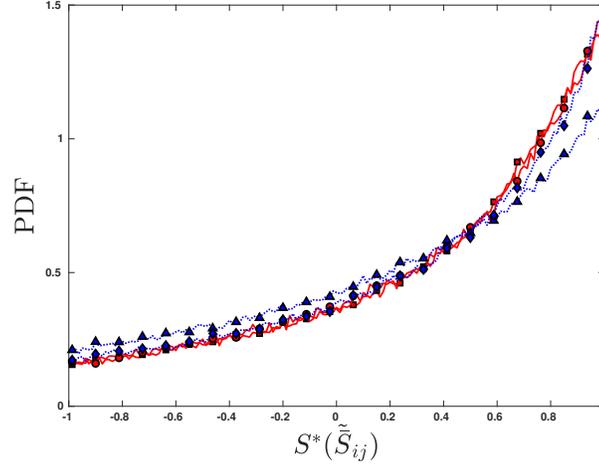


Figure 3.7: The probability density functions of the shape factor $S^*(\tilde{S}_{ij})$ for the filtered strain tensor \tilde{S}_{ij} . Solid line is DNS target: with squares for $0T$, circles for $T/8$. Dotted line is OICS: with upward-pointing triangle for $0T$ and diamond for $T/8$.

subgrid-scale stress tensor $\bar{\tau}_{ij}$ in Fig. 3.8. The PDFs reflect a perfect agreement between the OICS and DNS target data fields at the initial time, where both fields give a similar probability peak at the value -1 . The perfect agreement stays the same during the whole time horizon T . Comparing with OIC fields, the PDF in Fig. 2.9 needs slightly more time to give such a perfect agreement.

3.4.1 CONTOUR PLOTS OF INSTANTANEOUS DISTRIBUTIONS

The contour plots are presented for the magnitude of the vorticity $|\tilde{\omega}|$ and the magnitude of the strain rate tensor $|\tilde{S}|$ in this section, which illustrates the qualitative agreements between the OICS fields and the DNS target field.

For vorticity magnitude, the contour plots are calculated when the resolution equals 64, and the filter scale equals $4\delta_x$. As is represented in the two-dimensional contours, the red areas represent the vortical structures with strong vorticity. The result is presented for three different time steps: the initial time step, the middle time and the final time step.

Fig. 3.9 shows that initially, the agreement between the OICS field and the DNS target field is weak, where the structures of strong vorticity magnitudes are absent from OICS fields. Fig. 3.10 shows the contour levels in the OICS fields are increased compared to what we have at the initial time. These observations are continued when we do the comparisons at the middle time in Fig. 3.11, as the contour structures for both fields are quite similar to each other, such that it indicates the existence of a good agreement. We can see similar observations in Fig. 3.12 and 3.13, where $t = 3/4T$ and

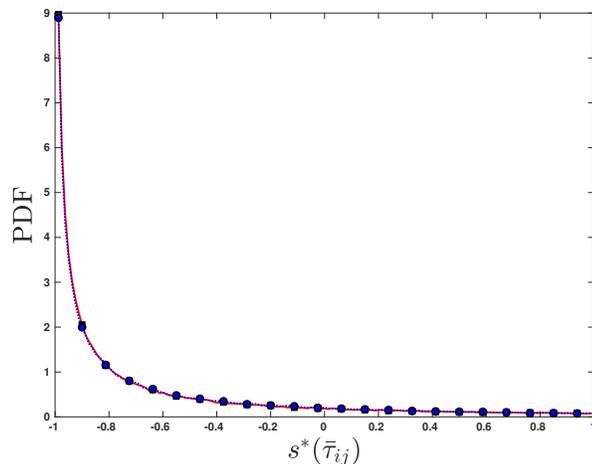


Figure 3.8: PDFs of the shape factor $S^*(\bar{\tau}_{ij})$ for the subgrid-scale stress tensor $\bar{\tau}_{ij}$. Solid line is DNS target: with squares for $0T$. Dotted line is OICS: with circles for $0T$.

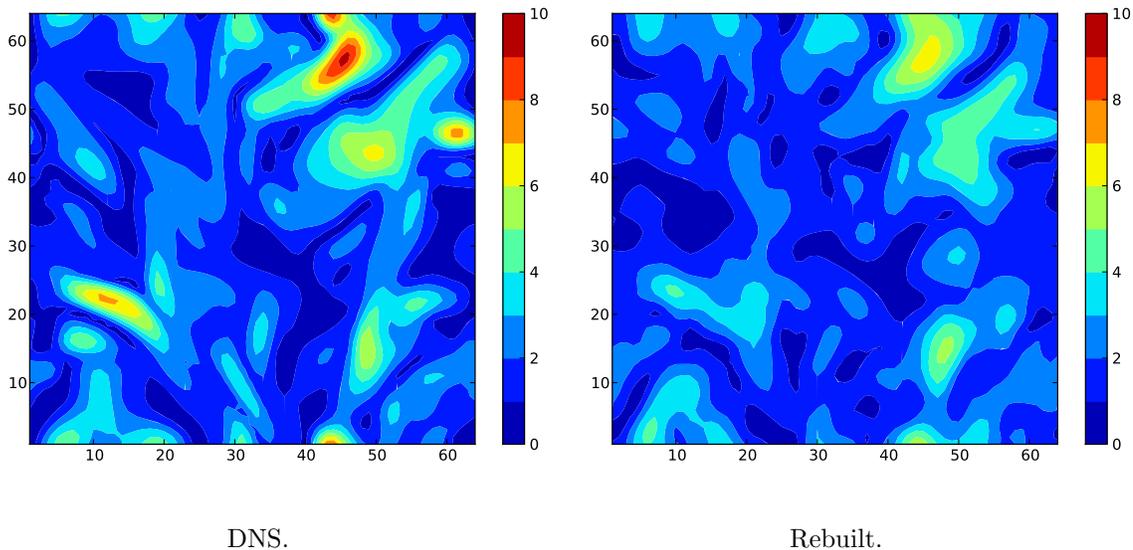
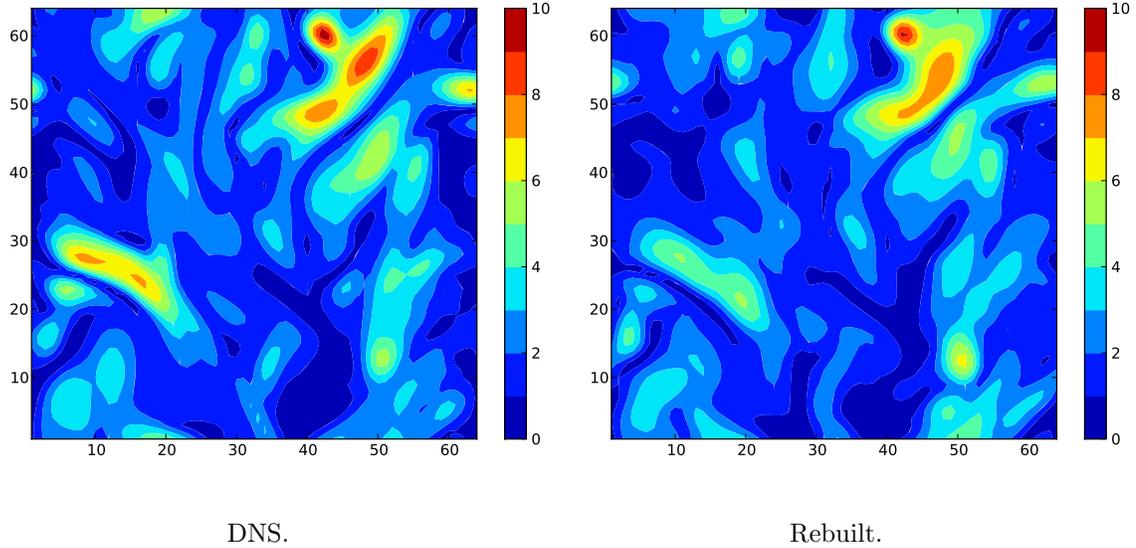
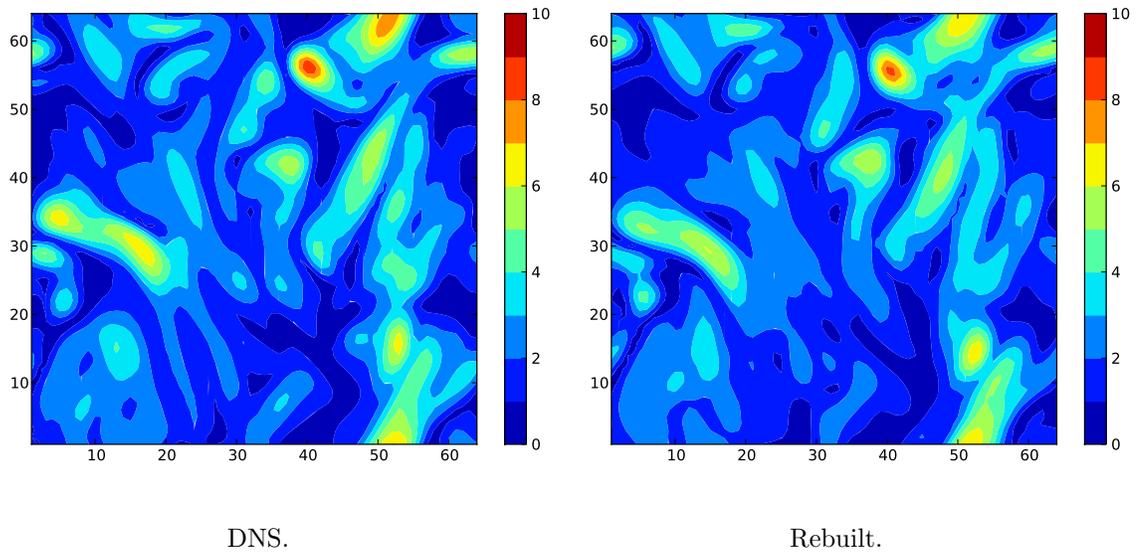


Figure 3.9: Contour plots for $|\tilde{\omega}|$ at $t = 0T$.

$t = T$ respectively.

For the OIC fields in Fig. 2.16 to Fig. 2.19, the vorticity contours in OIC and DNS target fields are similar as here. Therefore, there is no obvious difference with the agreement between the OICS and DNS target field.

The contour plots for the magnitude of the strain rate field are shown in Figs 3.14, . . . , 3.16. Starting from the initial time step shown in Fig. 3.14, the similarity is not big

Figure 3.10: Contour plots for $|\tilde{\omega}|$ at $t = T/4$.Figure 3.11: Contour plots for $|\tilde{\omega}|$ at $t = T/2$.

enough between the contours for the two fields, where the strong spots are absent at the OICS fields. The improvement of the contour plots of the OICS fields is quite clear in Fig. 3.15, as we can see that the agreement between the OICS and the DNS target data fields is nearly perfect. The perfect agreement remains stable until the end time T , as we see in Fig. 3.16 when the contour plots in the OICS fields reconstruct similar

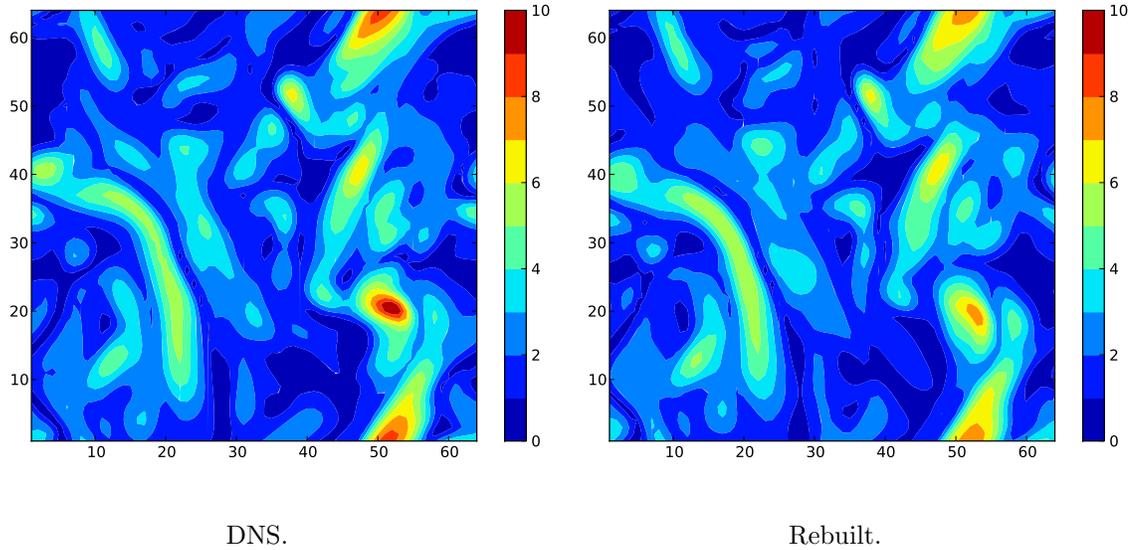


Figure 3.12: Contour plots for $|\tilde{\omega}|$ at $t = 3/4T$.

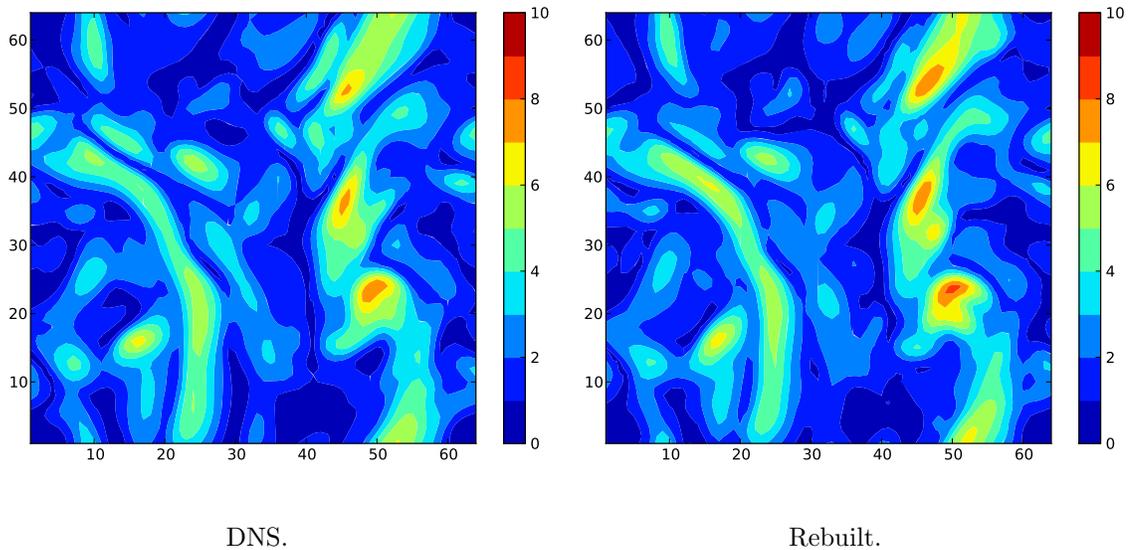
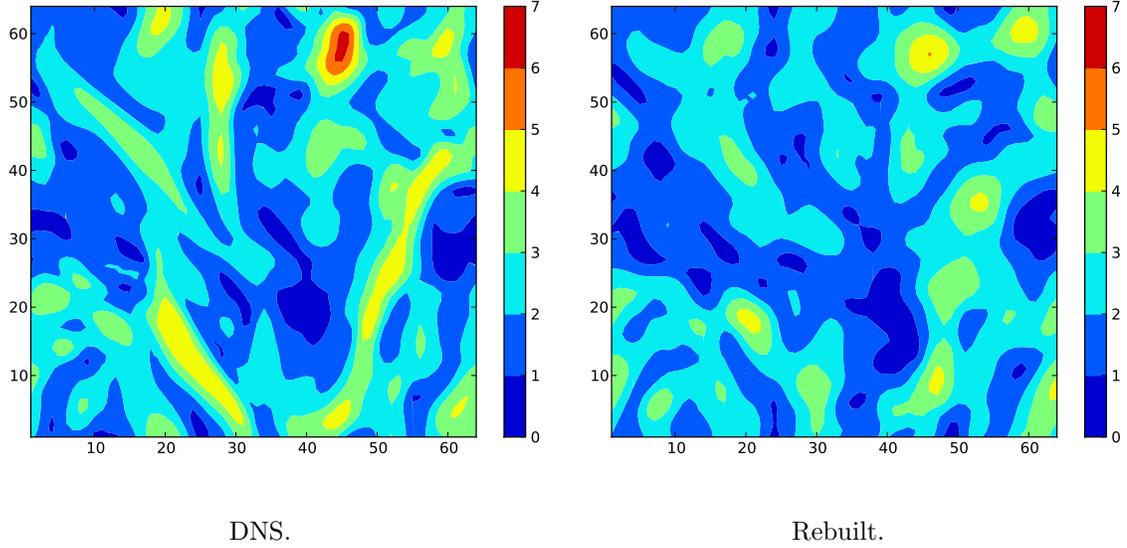
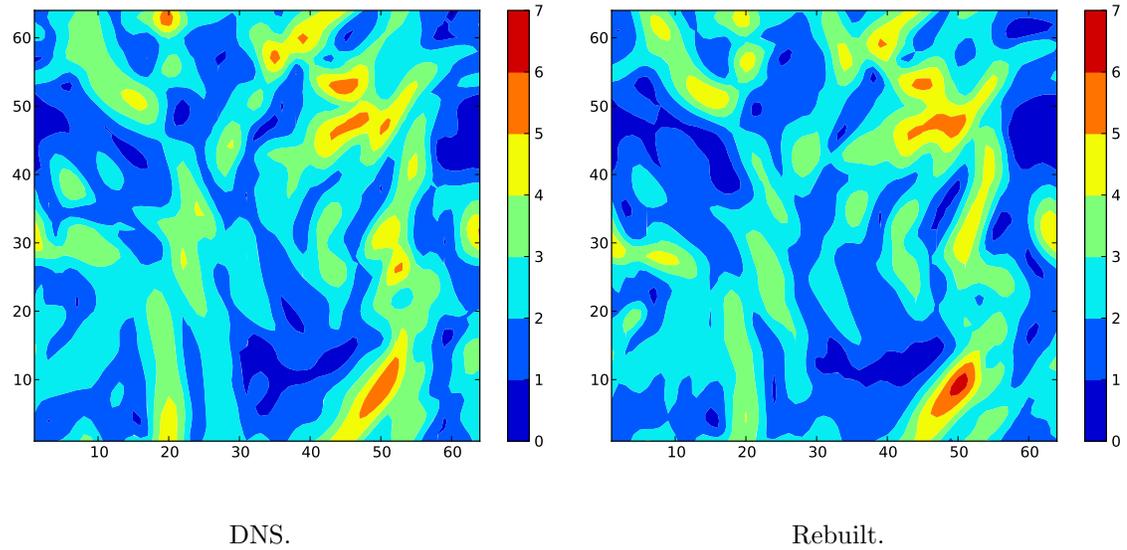


Figure 3.13: Contour plots for $|\tilde{\omega}|$ at $t = T$.

patterns to what we have in the DNS target field.

As a general conclusion for this part, the strain rate field in the OICS fields recovers in an early time at the $t = T/2$ to be close to the DNS target field. Compared with the result in Chapter 2, it seems that good or even better reconstructions can be obtained

Figure 3.14: Contour plots for $|\tilde{S}|$ at $t = 0T$.Figure 3.15: Contour plots for $|\tilde{S}|$ at $t = T/2$.

using the imperfect model.

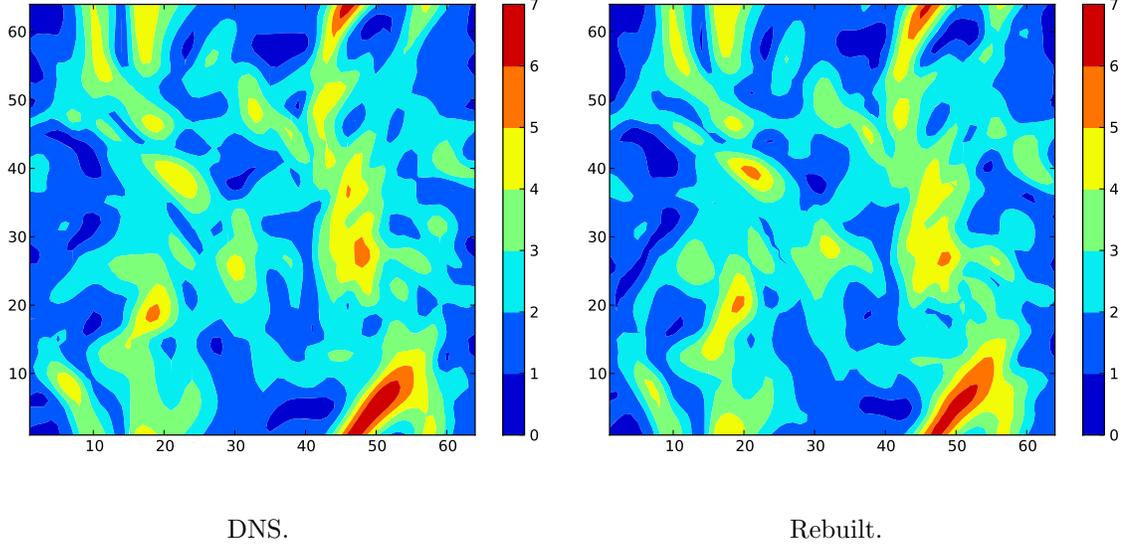


Figure 3.16: Contour plots for $|\tilde{S}|$ at $t = T$.

3.4.2 THE MINIMUM VOLUME ENCLOSING ELLIPSOID (MVEES)

We present now results related to the MVEE for the enstrophy in the OICS fields. The structures are defined as the regions with vorticity magnitude given by $\sqrt{\tilde{\omega}\tilde{\omega}} \geq 1.5\langle\tilde{\omega}\tilde{\omega}\rangle^{1/2}$.

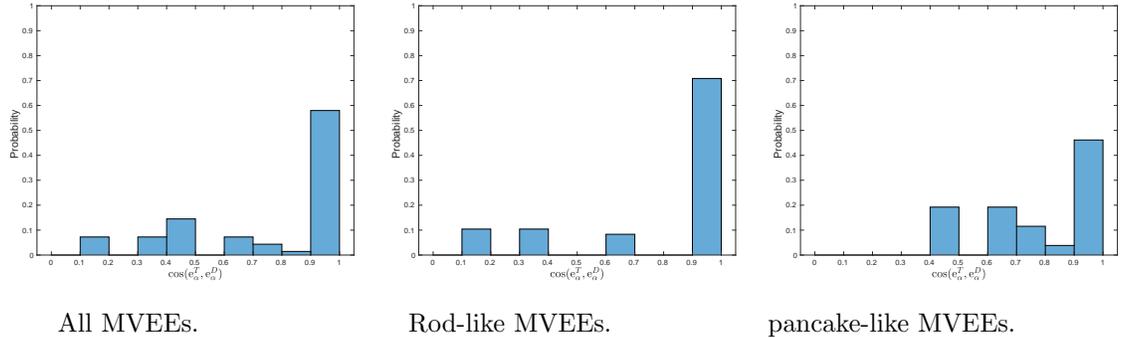


Figure 3.17: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = 0T$.

Fig. 3.17 shows that the agreement at the initial time between the major axes is already acceptable with a value of 0.58 for the peak probability, for all MVEEs. In the conditional cases, good agreement is also found, where the values of peak probabilities are approximately 0.7, for rod-like MVEEs, while pancake-like MVEEs is lower than that at 0.47.

For the middle axes in Fig. 3.18, the probability at $|\cos(e_\beta^D, e_\beta^T)| = 1$ is much lower than the major axes as it reaches 0.12, indicating weak agreement between the axes. Also, we see the conditional histograms give small values in rod-like MVEEs and pancake-like MVEEs.

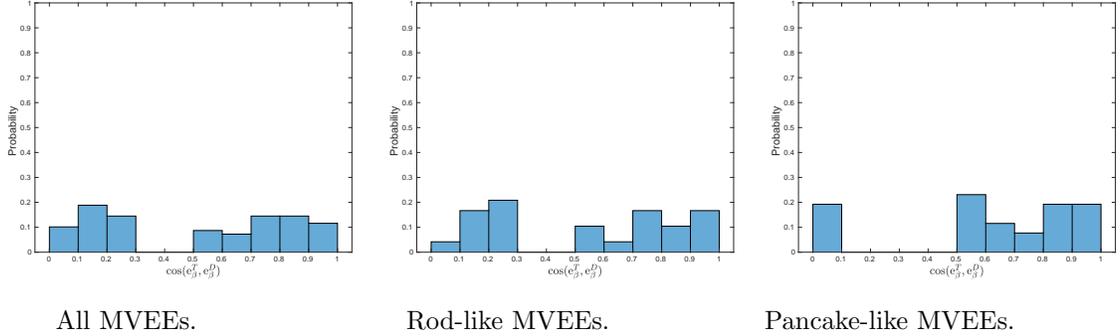


Figure 3.18: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = 0T$.

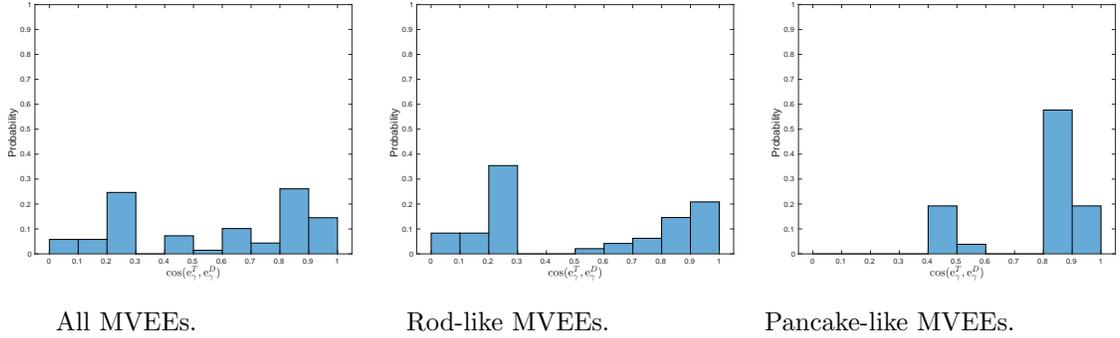


Figure 3.19: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = 0T$.

For the minor axes in Fig. 3.19, we notice that the probability values $|\cos(e_\gamma^D, e_\gamma^T)| = 1$ are slightly bigger than the middle axes. Rod-like MVEEs and pancake-like MVEEs show also small peak probability values.

For $t = T/2$ shown in Figs. 3.20 to 3.22, the peak probability values show significant improvements over the values that we have at the initial time for the three axes. Strongest alignment is observed for the major axes.

Figs 3.23 to 3.25 show the results for the major, middle and minor axes at the end of the optimization horizon, when $t = T$. For the major axes in Fig. 3.23, the peak probability value slightly decreases to 0.7. However, it still higher than the initial time. The peak probability values of rod-like MVEEs and pancake-like MVEEs also decline from their maximum values at $T/2$. For the middle axes, despite the peak probability value (Fig. 3.24) drops from its previous value to approximately 0.4, it still is bigger than the one

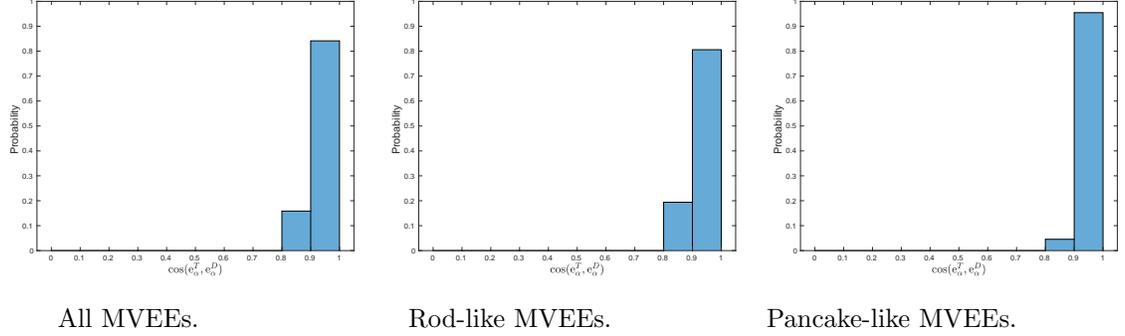


Figure 3.20: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = T/2$.

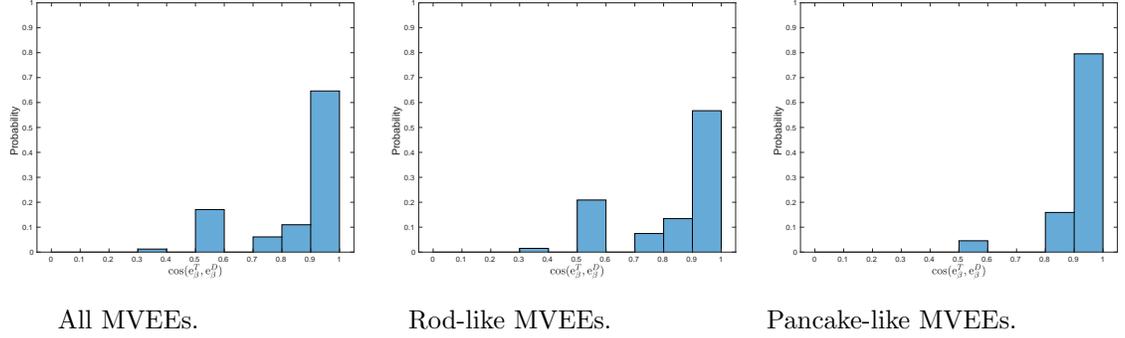


Figure 3.21: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = T/2$.

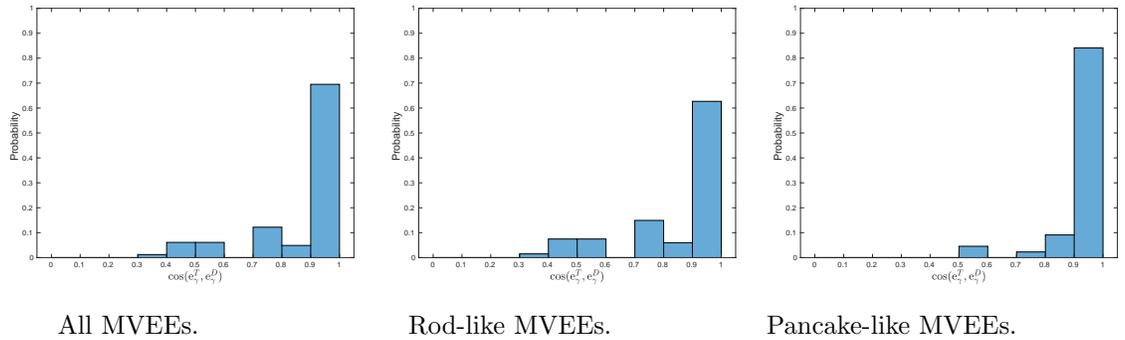


Figure 3.22: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = T/2$.

at $0T$. Both rod-like MVEEs and pancake-like MVEEs improve on their corresponding histograms at the initial time, although the peak probabilities also decrease from the maximum values in the mid-time. Fig. 3.25 shows similar observations for the minor axes to what we have already observed in the middle and the major axes at the end at T .

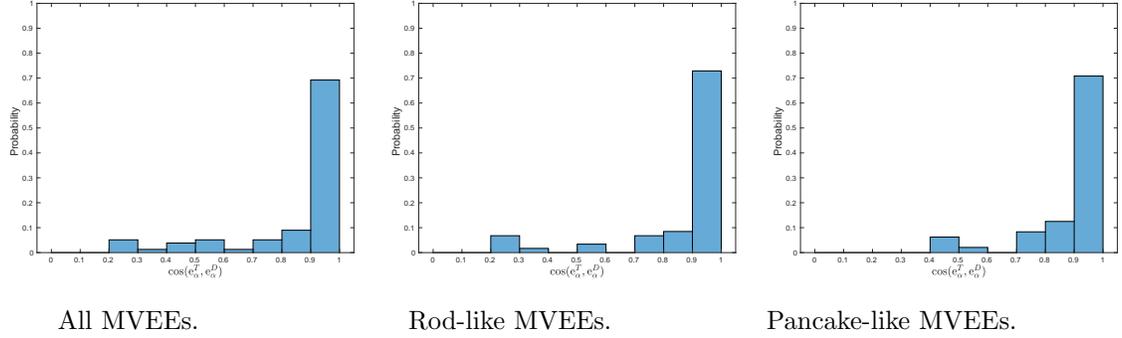


Figure 3.23: Alignment between the major axes for strong $|\tilde{\omega}|$ structures at $t = T$.

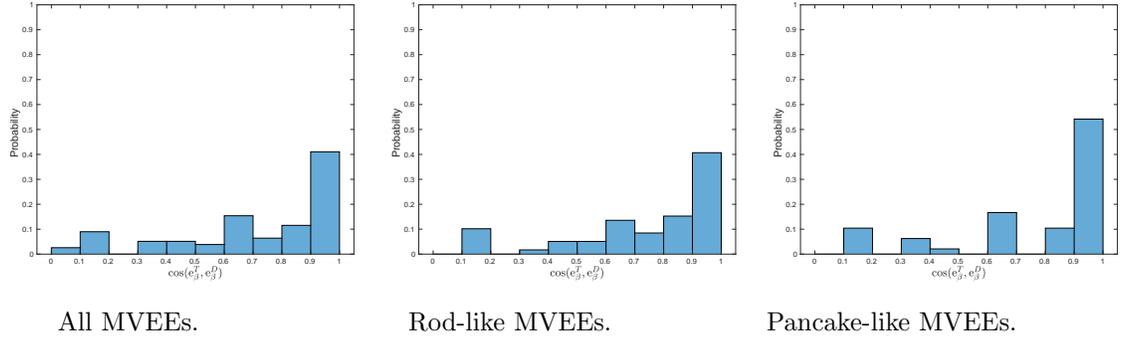


Figure 3.24: Alignment between the middle axes for strong $|\tilde{\omega}|$ structures at $t = T$.

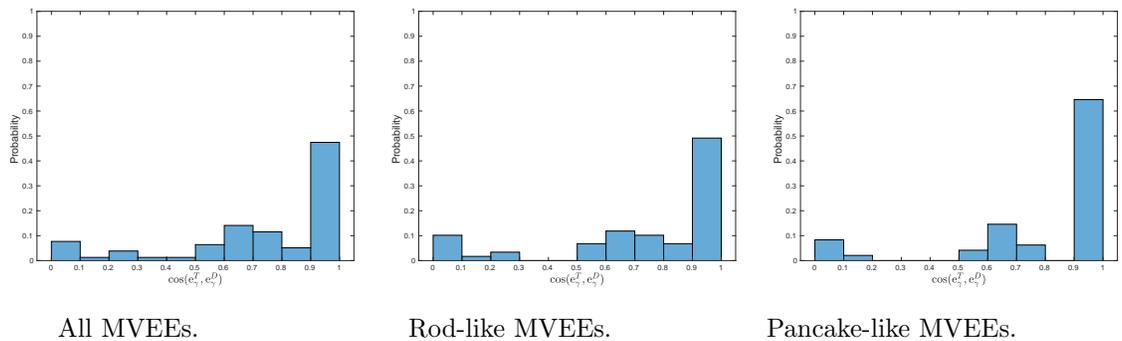


Figure 3.25: Alignment between the minor axes for strong $|\tilde{\omega}|$ structures at $t = T$.

Therefore, the histograms show reasonable agreement between the corresponding axes, especially when the time reaches the mid-time. The agreement decreases for some reasons at the end time at T . It might have been due to statistical fluctuations. The major axes show the best agreement between the corresponding clusters in the OICS and DNS target fields, then the minor axes and the middle axes. This observation is

similar to that in the OIC fields.

We present the medians of the histograms for $|\tilde{\omega}|$, in Fig. 3.26. From the three curves, we notice that the major axes reflect good alignment, as its curve starts at a high peak at 0.96 and it remains at ≈ 1 . However, it is not the case with the middle and the minor axes, as they start at 0.58 and 0.64, respectively. They tend to increase to show some enhancements and reach the maximum value of 0.97 at $t = 3/4T$. However, unlike the major axes, they decrease at the final time at T to arrive at the value 0.9. As a summary, reasonable agreement is observed which means that the axes of the MVEEs in the OICS fields have close orientations to the DNS target ones. The comparison with

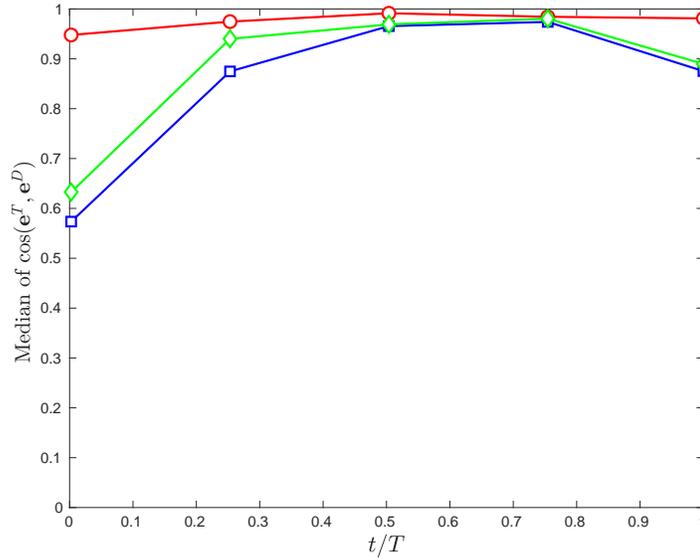


Figure 3.26: Medians of $|\cos(e^T, e^D)|$ for strong $|\tilde{\omega}|$ structures. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.

Fig. 2.53 shows slightly different behaviours, where the medians in the OICS fields give more stable trends for the middle and minor axes and slightly higher peaks.

In Fig. 3.27, we present the median ratios of the sizes between the clusters in the OICS fields and the corresponding clusters in the DNS target data field.

We notice that the black curve reflects a good agreement between the corresponding clusters. The agreement at the initial time is the smallest as it gives a value of 0.6. It increases to reach a maximum of 0.9 when $t = 3/4T$. Nevertheless, it slightly drops to stop at 0.85. The green symbols in the same figure present the median ratios of the clusters of a given size. These symbols show different behaviours, as we can see that some of them exceed the value 1. The trends of the symbols agree with what we have for the black curve.

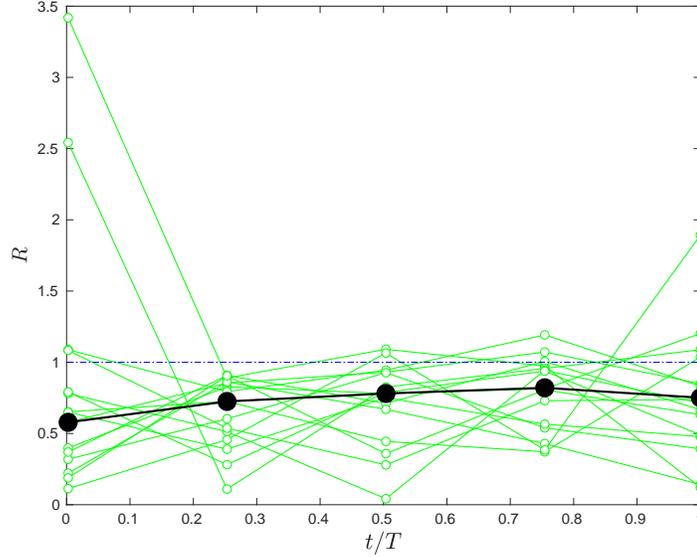


Figure 3.27: Median ratio between the sizes of the clusters for strong $|\tilde{\omega}|$ structures. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.

Therefore, the clusters in the OICS fields are produced with close similarity to the ones in DNS target field. The agreement improves over time quicker and the ratios show higher peak values than the ones in OIC fields, (Fig. 2.54).

In the following histograms, we show the probability of the displacement of the centres of the MVEEs for both OICS and DNS target fields relative to the lengths of the axes in the DNS target field, which is represented by d_c/ℓ for the vorticity magnitude.

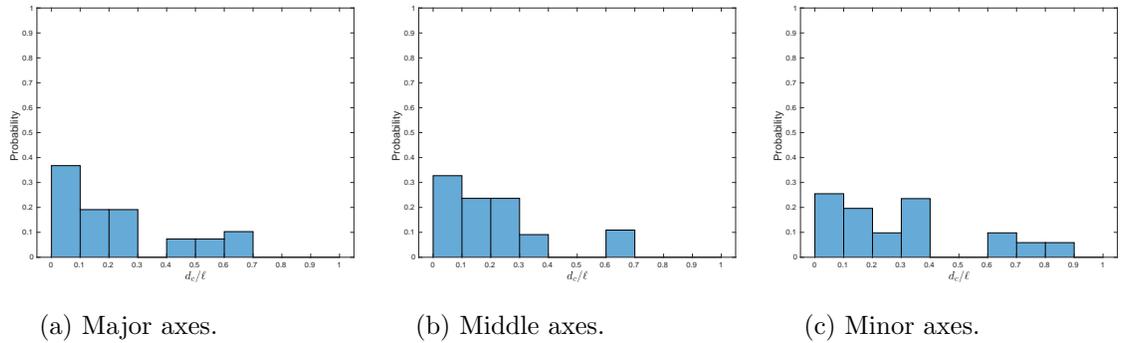


Figure 3.28: d_c/ℓ for strong $|\tilde{\omega}|$ structures at $t = 0T$.

As we can see in the initial time from Fig. 3.28, the peak probabilities of the three histograms are small. These lower values are the indication of the weak agreement

between the corresponding axes, where the major, middle and minor axes show values at 0.38, 0.32 and 0.26, respectively. In Fig. 3.29, we document the improvement of the histograms, as we notice the smaller bins show the highest values so far for the three axes. Figs. 3.30 displays the probabilities at the end time when $t = T$. We notice that the major and middle axes give the lower values than the middle time with a small amount, at 0.41 and 0.45, respectively. However, the minor axes still improve at this time to reach a peak probability value of 0.96.

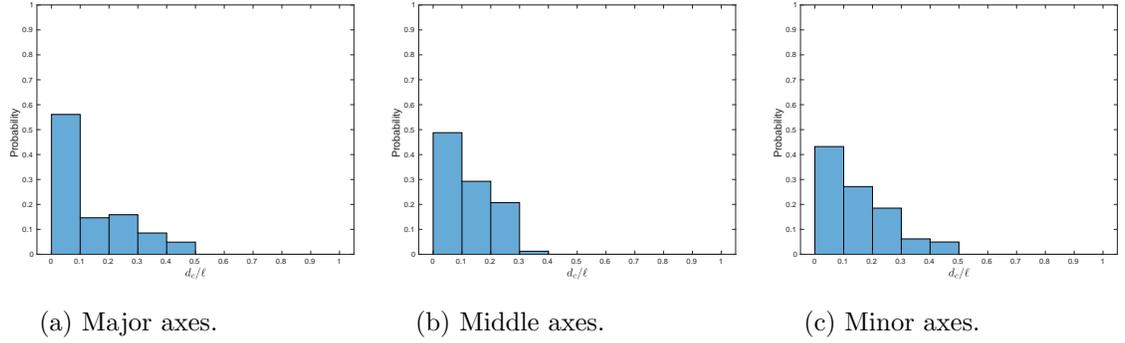


Figure 3.29: d_c/ℓ for strong $|\tilde{\omega}|$ structures at $t = T/2$.

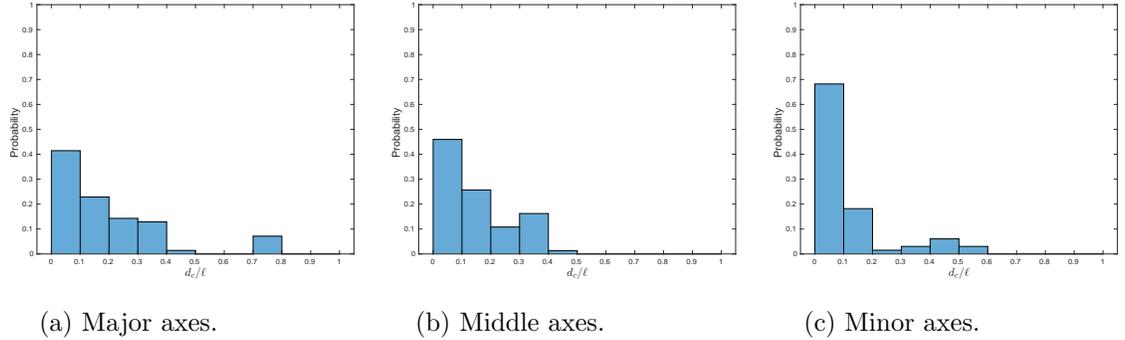


Figure 3.30: d_c/ℓ for strong $|\tilde{\omega}|$ structures at $t = T$.

The results of the MVEEs are now presented for the magnitude of the strain rate field. The rate of the filtered strain magnitude is written as $|\tilde{S}| = \sqrt{2\tilde{S}_{ij}\tilde{S}_{ij}}$; we extract the structures with $|\tilde{S}| \geq 1.7\langle|\tilde{S}|\rangle$.

Fig. 3.31 shows that the agreement between the major axes is moderate initially when the peak probability is found to be 0.59, for all MVEEs. In the conditional cases, better agreements are shown, with peak probabilities equal 0.71 for rod-like MVEEs, and 0.7 for pancake-like MVEEs.

For the middle axes in Fig. 3.32, the histogram present a lower peak value than the major axes as it reaches 0.34, which reflects the weak agreement between these

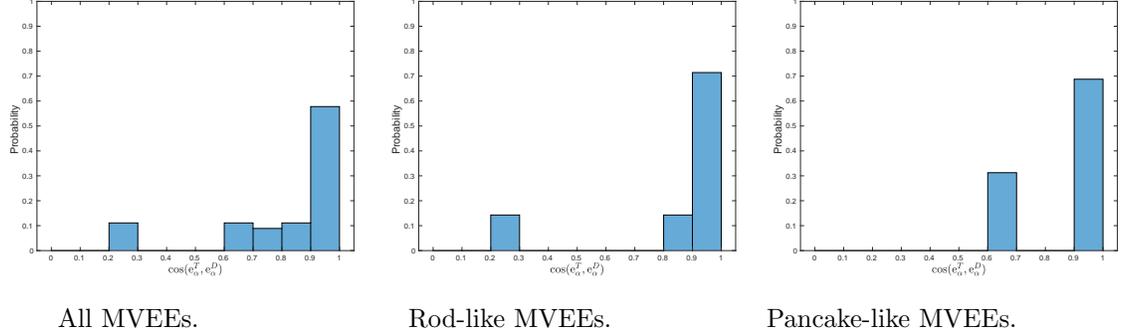


Figure 3.31: Alignment between the major axes for $|\tilde{S}|$ at $t = 0T$.

axes. Also, the conditional MVEEs alignments show smaller probability values in rod-like MVEEs and pancake-like MVEEs than the conditional ones in the major axes. Although the minor axes in Fig. 3.33 present a bigger peak probability value than the major and middle, it still reflects a moderate agreement at 0.68. We notice that the peak probability value with rod-like MVEEs is bigger than pancake-like MVEEs probability value.

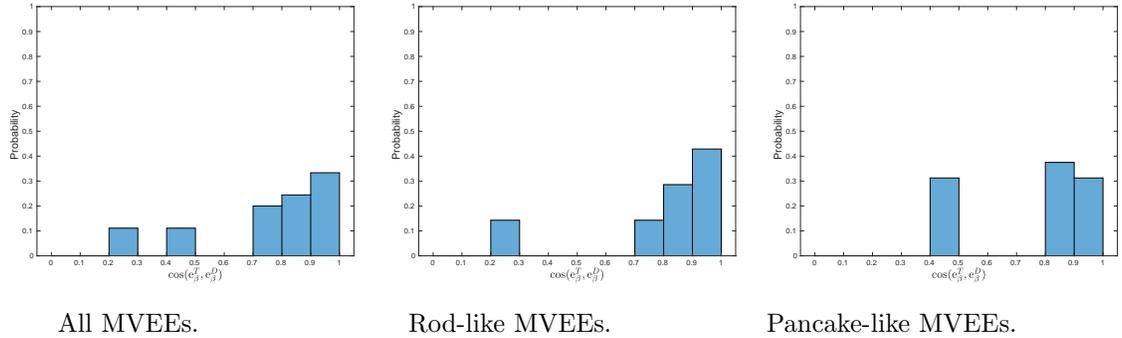


Figure 3.32: Alignment between the middle axes for $|\tilde{S}|$ at $t = 0T$.

At the time $T/2$, shown with Figs. 3.34 to 3.36, the agreement is considered the best through the time T , as the peak probability values increase to their maxima. The same observations are presented in the conditional MVEEs alignments. Figs 3.37 to 3.39 describe the observations for the major, middle and minor axes at $t = T$, respectively. The major axes in Fig. 3.37, show that the peak probability value slightly reduces to be at 0.74. Nevertheless, it is still bigger than the one at the initial time. The peak probability value for rod-like MVEEs increases to a peak at 0.89, unlike pancake-like MVEEs which drops from the perfect agreement at the mid-time.

The peak probability value for the middle axes in Fig. 3.38 decreases from its previous value as it reaches 0.66. However, it remains bigger than the one at $0T$. The peak value

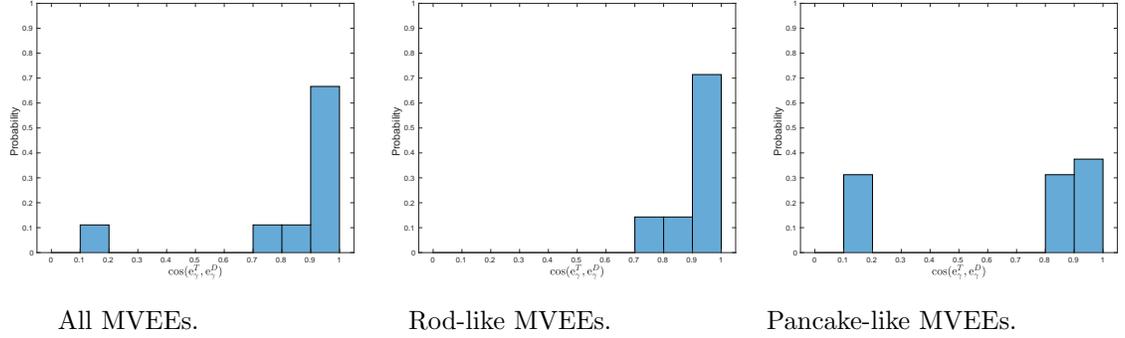


Figure 3.33: Alignment between the minor axes for $|\tilde{S}|$ at $t = 0T$.

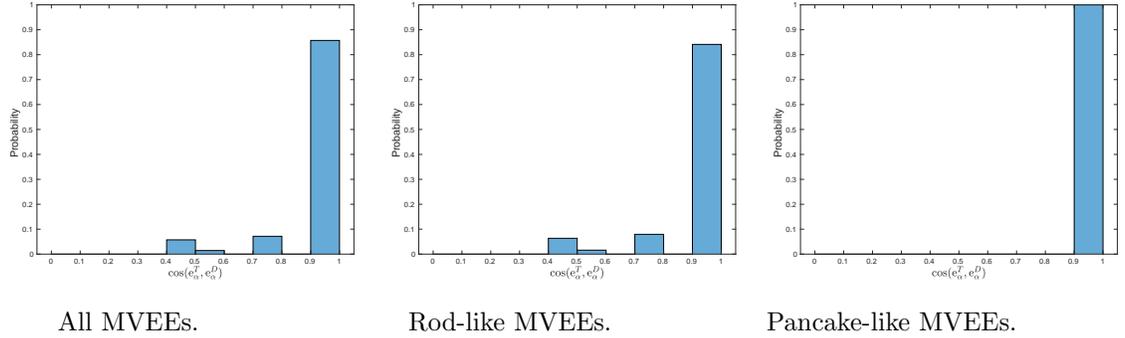


Figure 3.34: Alignment between the major axes for $|\tilde{S}|$ at $t = T/2$.

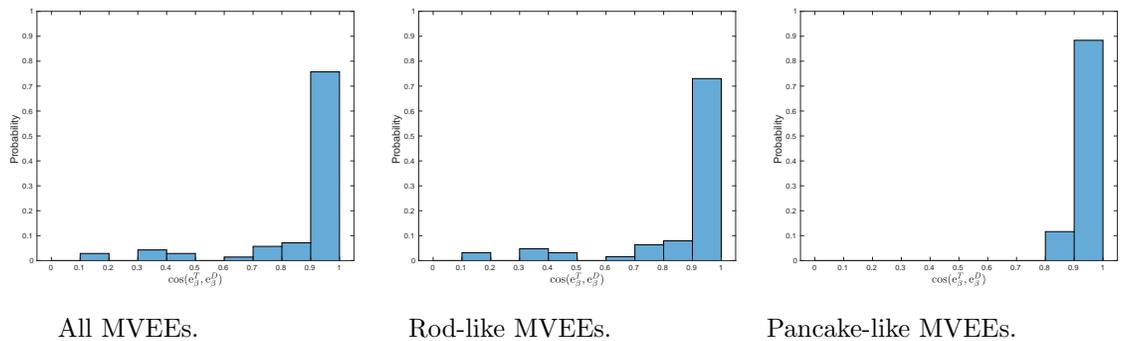


Figure 3.35: Alignment between the middle axes for $|\tilde{S}|$ at $t = T/2$.

on rod-like MVEEs improves to reach 0.79, unlike pancake-like MVEEs which decreases from the one at $T/2$. The minor axes in Fig. 3.39 present the highest peak value of 0.91 among all the results even for the major and middle axes. Same observations happen for the conditional alignments with the perfect agreement in pancake-like MVEEs while it gives a value of 0.89 for rod-loke MVEEs. Therefore, in general, the results show

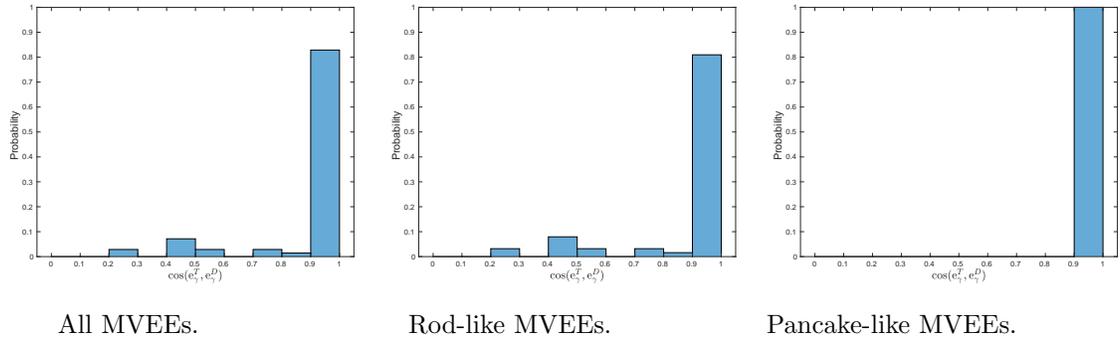


Figure 3.36: Alignment between the minor axes for $|\tilde{S}|$ at $t = T/2$.

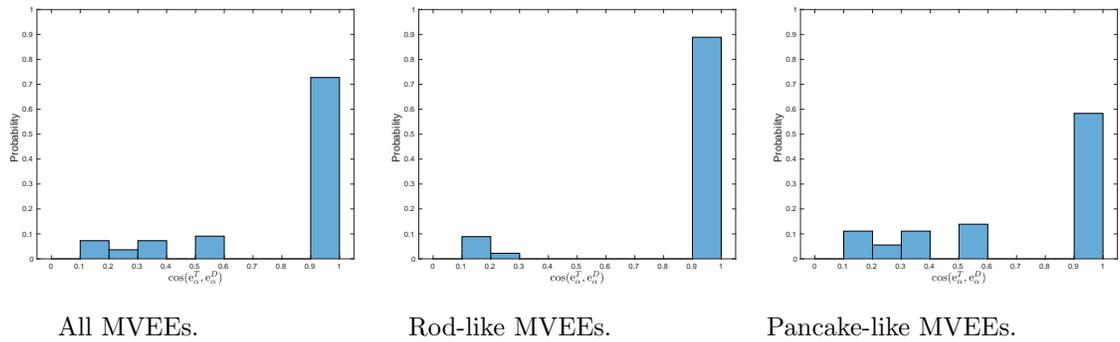


Figure 3.37: Alignment between the major axes for $|\tilde{S}|$ at $t = T$.

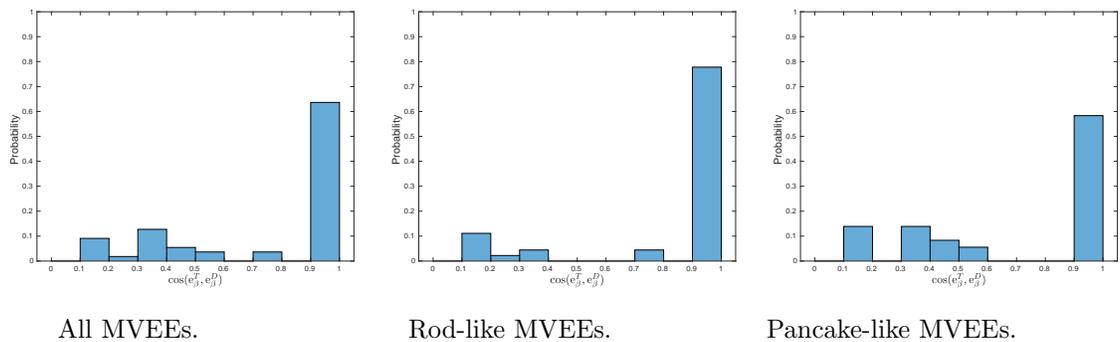


Figure 3.38: Alignment between the middle axes for $|\tilde{S}|$ at $t = T$.

stable improvement for the agreement between the corresponding axes over time for $|\tilde{S}|$. Also, the minor axes give the best agreement, the major axes come in the second while the middle axes show the weakest agreement.

The median of the MVEEs alignments between the axes of the major, middle and

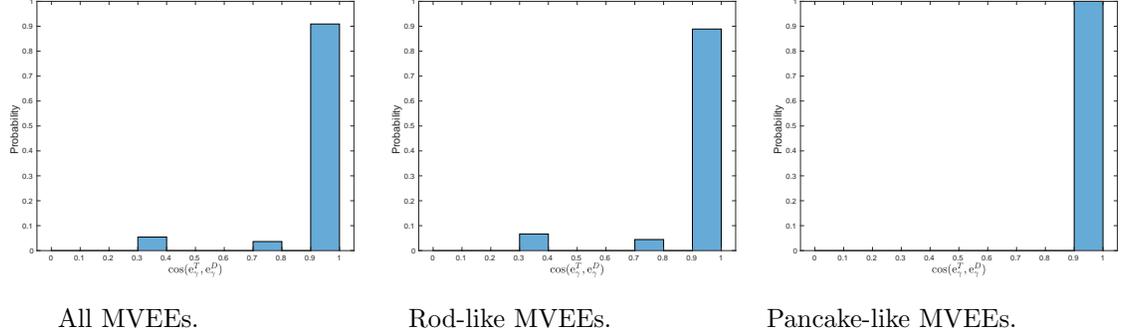


Figure 3.39: Alignment between the minor axes for $|\tilde{S}|$ at $t = T$.

minor is presented in Fig. 3.40. We notice the growth of the agreement between the corresponding axes over time. The investigations occur at the times $0T$, $T/4$, $T/2$, $3/4T$ and T . We notice the very good agreement represented by the medians, where they roughly stay at ≈ 1 , after a short time. The main variations between the axes appear at the initial time, where median for the major axes starts big of 0.9, and it increases to 0.97. Furthermore, it reaches the maximum value at the time $3/4T$. The median for the middle axes starts at the lowest of 0.84 to reach 0.97.

Therefore these results reflect a quite good agreement between the two fields. The major axes and the minor axes are the best while the middle axes are the weakest. By looking into Fig. 2.70, we see that the trends and the stabilities are better for the axes produced here by the OICS fields for the three medians in the \tilde{S} .

We introduce also the median ratios of the sizes of the clusters in the OICS fields and the DNS target field in Fig. 3.41. The development of the black curve, which shows the ratios between the sizes of the clusters, is clear. Although in the initial time it gives a small value of 0.3, it reaches a maximum value at $t = 3/4T$. It drops at the end when $t = T$ to 0.82. The median ratios of the clusters of a given size are presented by the green symbols. We notice that there are some values which are bigger than 1.

Therefore we can produce clusters in the strain rate field in the OICS fields having close sizes to the corresponding clusters in the DNS target field. Again, Fig. 2.71 shows that the improvement of the size ratios is weaker in the OIC field than the OICS fields.

In what follows, we present the probability of the displacement of the centres of the MVEEs relative to the lengths of the axes from the DNS target field d_c/ℓ for $|\tilde{S}|$. In Fig. 3.42, we show the probability of relative displacement at $0T$, as we see that the major axes present the lowest probability value in the smallest bin, with a value of 0.25, indicating that the agreement for the major axes starts weak. The middle axes are slightly better than the major axes, with a peak value of 0.29. However, this is not the case with minor axes as they give a high probability value of 0.7. Overall, at the initial time, we document a weak agreement between the OICS and the DNS target

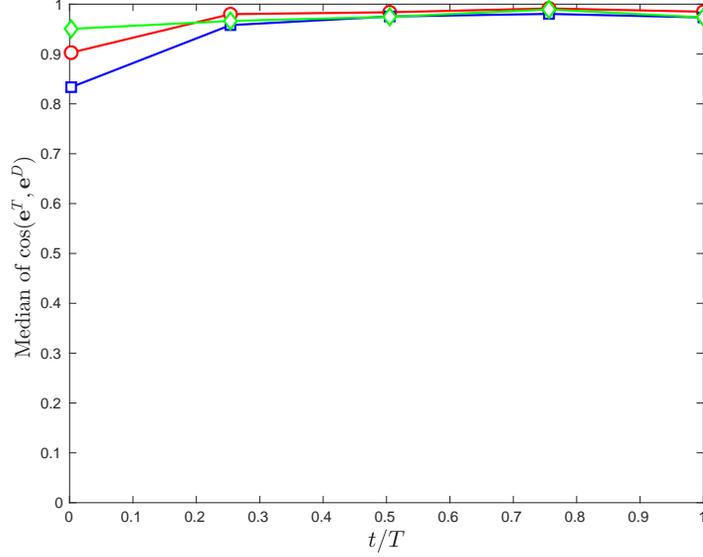


Figure 3.40: Medians of $|\cos(\mathbf{e}^T, \mathbf{e}^D)|$ for $|\tilde{S}|$. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.

fields.

At time $T/2$, Fig. 3.43 shows that the agreement improves for the axes, as the peak probability values for the axes rose from their probability values at the initial time. Fig. 3.44 shows that all the axes give lower values for the relative displacement when the time equals T compared to what we present at the middle time. However, for the major and middle axes, probabilities give higher values than those at the initial time, while the minor axes drop from the peak to reach 0.64.

The comparison with the histograms in OIC fields in Fig. 2.72 to 2.74 show that in some cases the agreement is better, but the differences are very small between the two models.

We calculate now the SGS energy dissipation MVEEs. We use 3 times of the mean as the threshold value to extract the structures.

The SGS energy dissipation displays the most significant results compared with the previous two quantities, where we notice the gradual growth of the agreement between the corresponding axes over the time to give the highest peaks at $t = T$. We can summarize how these values are developed over time as follows.

For the major axes in Figs. 3.45, 3.48 and 3.51, we see the peak probability values are 0.57, 0.7 and 0.93, respectively. The conditional MVEEs results also show the same trend, and give peak values, at $t = T$, 0.99 and 0.87 for rod-like MVEEs and pancake-like MVEEs, respectively.

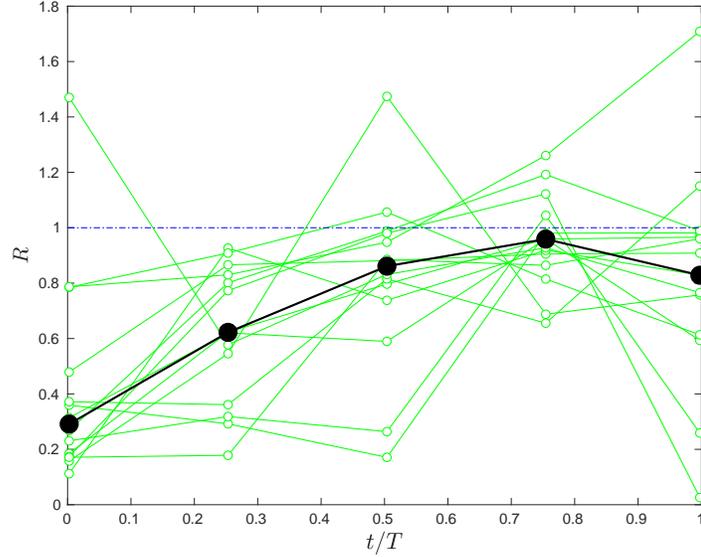
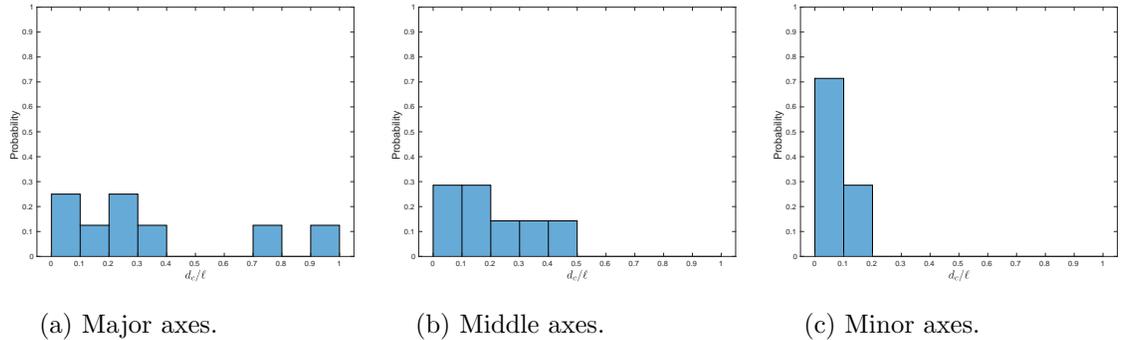


Figure 3.41: Median ratio between the sizes of the clusters for $|\tilde{S}|$. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.



(a) Major axes.

(b) Middle axes.

(c) Minor axes.

Figure 3.42: d_c/ℓ for $|\tilde{S}|$ at $t = 0T$.

For the middle axes in Figs. 3.46, 3.49 and 3.52, we see the peak probability values are 0.22, 0.27 and 0.86, respectively. The same observations exist for the conditional MVEE alignments. For the minor axes, shown in Figs. 3.47, 3.50 and 3.53, we see the peak probability values show of 0.4, 0.49 and 0.83, respectively. Also, we notice similar trends in the conditional MVEEs alignments.

Finally, the major axes show the best agreement between the two fields, and the middle axes are considered as the second best, while the minor axes reflect weaker agreement. These observations are slightly better than the ones in the OIC fields, shown in Figs. 2.75 to 2.86.

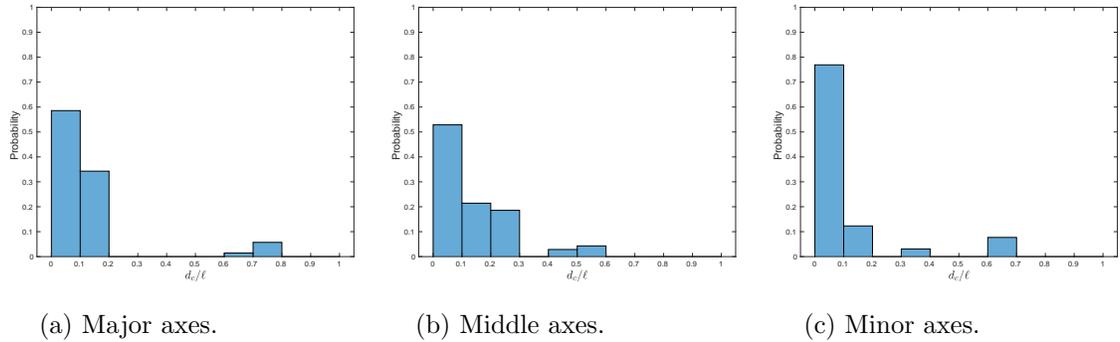


Figure 3.43: d_c/ℓ for $|\tilde{S}|$ at $t = T/2$.

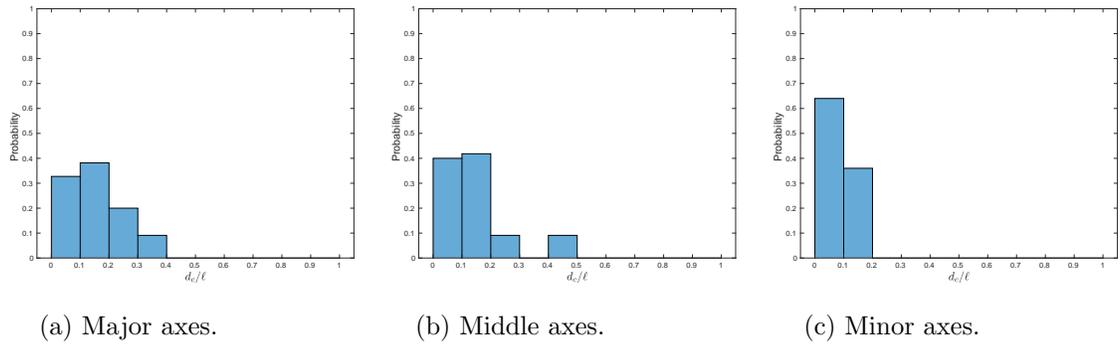


Figure 3.44: d_c/ℓ for $|\tilde{S}|$ at $t = T$.

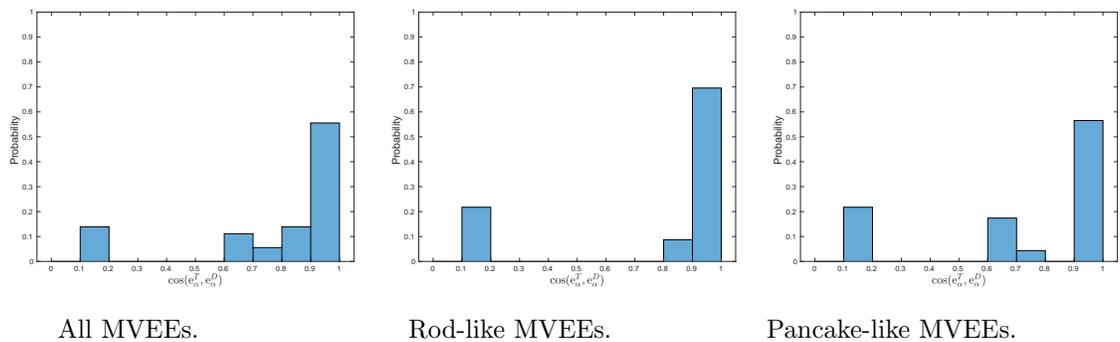


Figure 3.45: Alignment between the major axes for SGS energy dissipation at $t = 0T$.

Fig. 3.54 shows the medians of the MVEE alignments. The red curve, which represents the trend of the major axes shows the highest starting point of 0.95 at the initial time to reach a value of 0.96 at $t = T$. However, some decline occurs at $t = T/4$. The middle axes show the lowest values for the median between all the axes, in general, as the blue curve starts at 0.52 to register a significant increase to reach the maximum value at

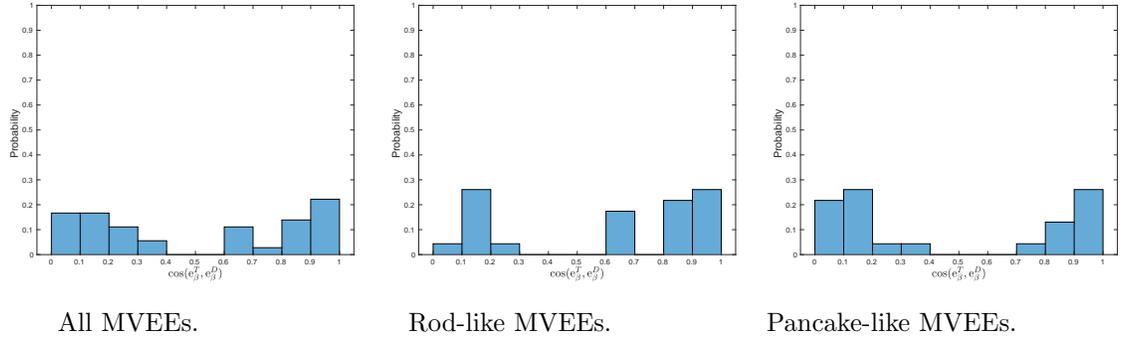


Figure 3.46: Alignment between the middle axes for SGS energy dissipation at $t = 0T$.

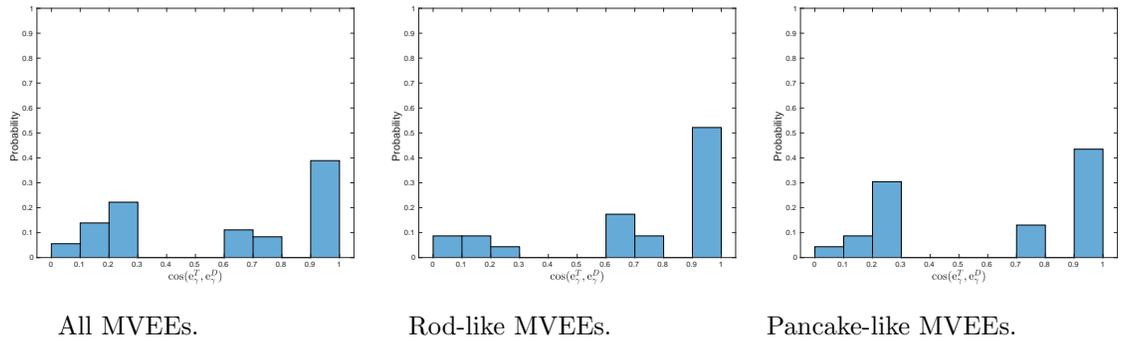


Figure 3.47: Alignment between the minor axes for SGS energy dissipation at $t = 0T$.

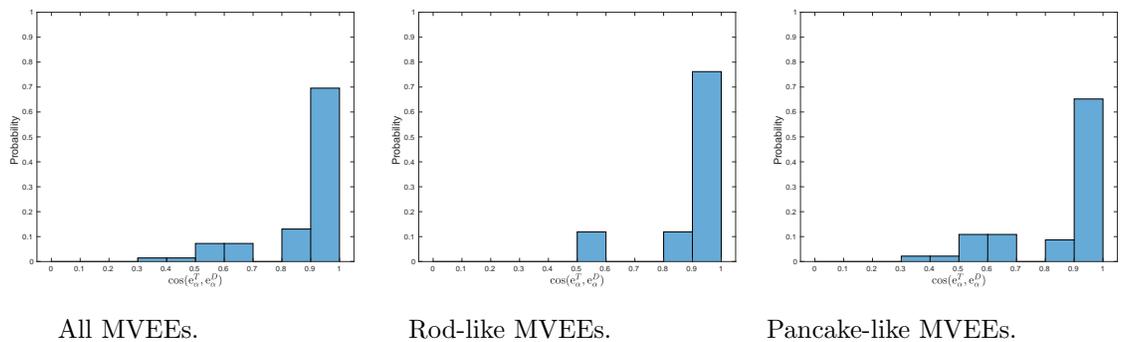


Figure 3.48: Alignment between the major axes for SGS energy dissipation at $t = T/2$.

0.94. The curve of the minor axes presents some different behaviour comparing to the other axes, as it gives value at 0.69 in the starting point, to reach 0.99. at the end time $t = T$. Therefore, the major and minor medians represent the strongest results as they show reasonable agreements, while the middle axes are the third.

The comparison between Fig. 2.87 and Fig. 3.54 show there is no clear winner between

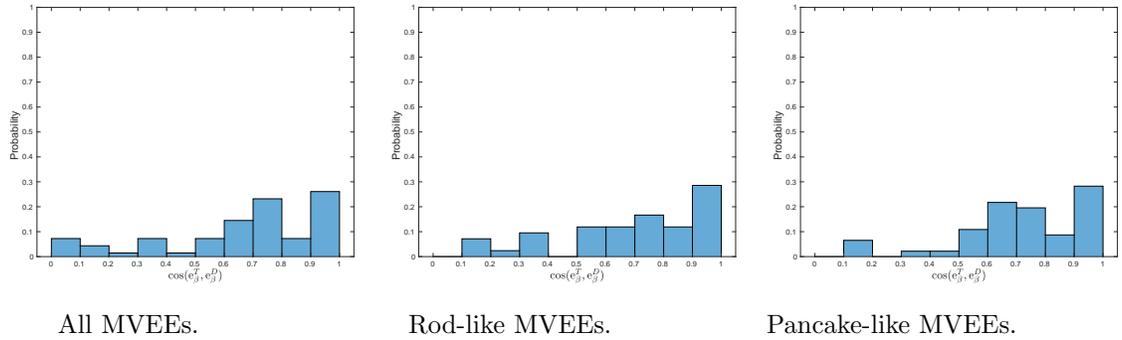


Figure 3.49: Alignment between the middle axes for SGS energy dissipation at $t = T/2$.

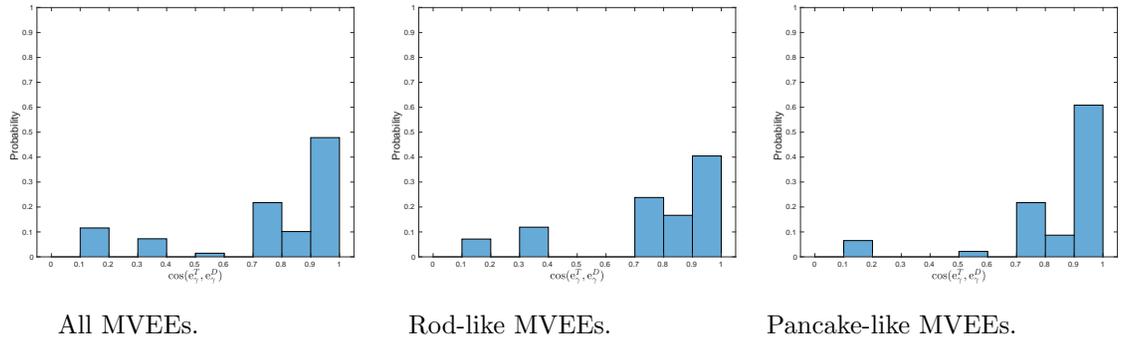


Figure 3.50: Alignment between the minor axes for SGS energy dissipation at $t = T/2$.

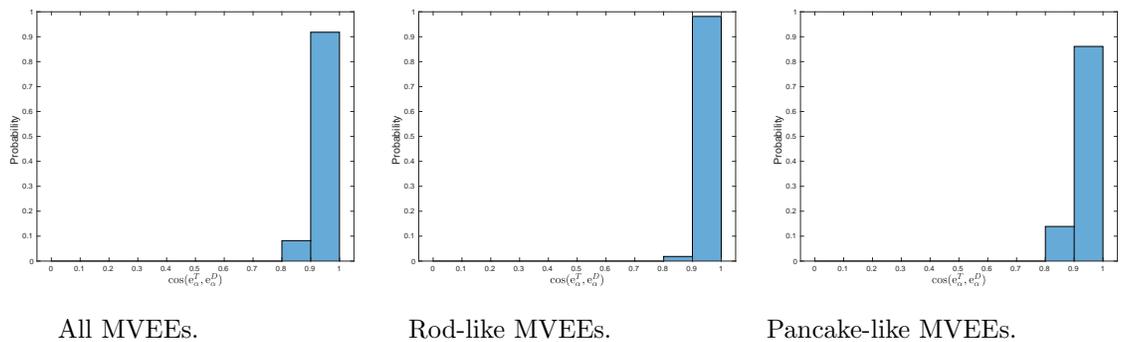


Figure 3.51: Alignment between the major axes for SGS energy dissipation at $t = T$.

them, where both the OIC and OICS fields are showing high values at the end for this specific statistic for the SGS energy dissipation.

As we presented before about the investigation of the median ratios of the sizes, we do the same with the quantity of the SGS energy dissipation in Fig. 3.55.

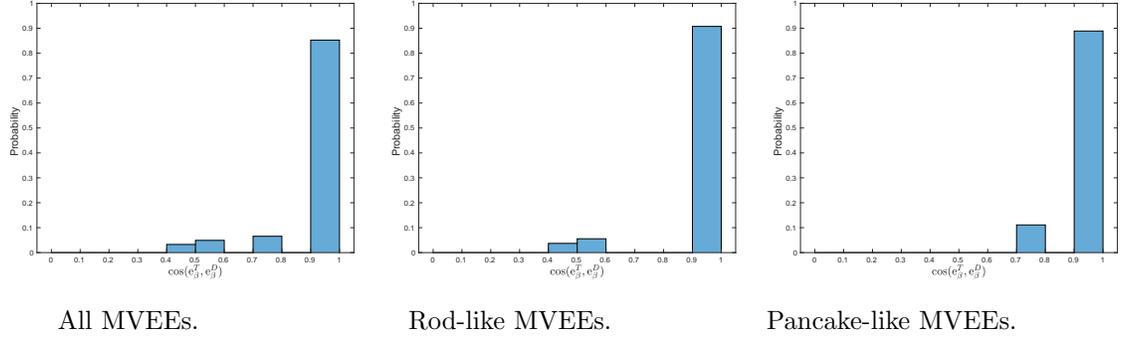


Figure 3.52: Alignment between the middle axes for SGS energy dissipation at $t = T$.

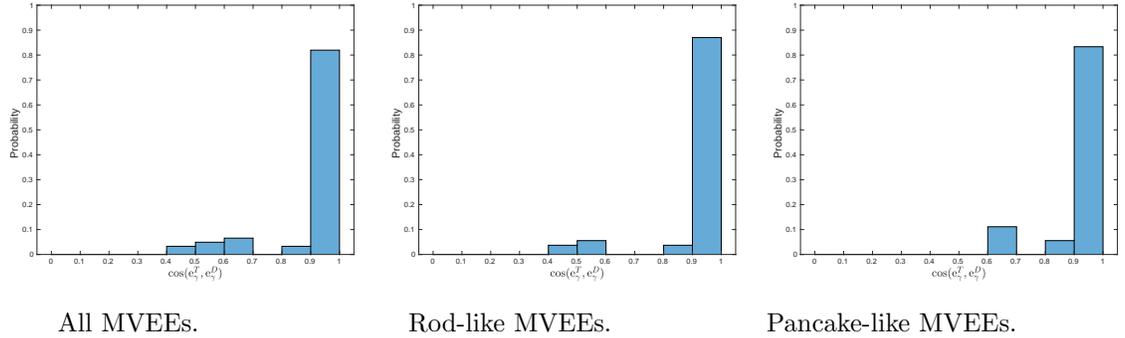


Figure 3.53: Alignment between the minor axes for SGS energy dissipation at $t = T$.

Despite the lower starting value at around 0.25 for the black curve, a rapid increase happens to reach 0.94. Therefore, the agreement improves as the time increase between the corresponding clusters for the OICS fields and the DNS target field. Besides, we observe the same trends for the median ratios of the clusters of a given size which are presented by the thinner curves. In general, we are able to reproduce close sizes of the structures of the SGS energy dissipation in the fields of OICS.

The comparison between Fig. 2.88 produced by the OIC fields and Fig 3.55 for the imperfect model shows that the latter field is the winner due to giving higher ratios over time.

The next three histograms in Figs 3.56, . . . , 3.58 express the relative displacement of the MVEEs for the subgrid-scale energy dissipation. The idea here is similar to all the previous sections that deal with these statistics.

At the initial time, we notice that the peak values of the relative displacement probabilities are small, at 0.24, 0.37 and 0.22 for the major, middle and minor axes, respectively. Therefore, the agreement for the OICS fields is not very good with the DNS target field, as shown in Fig. 3.56.

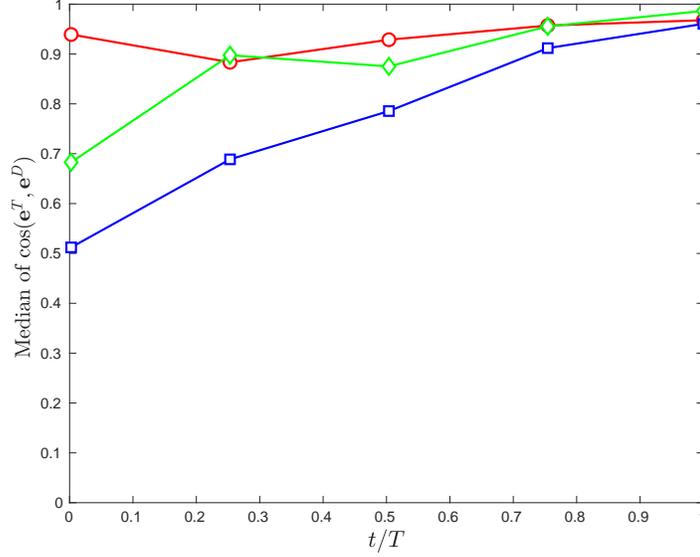


Figure 3.54: Medians of $|\cos(\mathbf{e}^T, \mathbf{e}^D)|$ for SGS energy dissipation. Red with circles: major axes. Blue with squares: middle axes. Green with diamond: minor axes.

Clearly, the peak values for all three axes have increased in Fig. 3.57, with an acceptable agreement. At $t = T$, Fig. 3.58 shows that the major axes reach a value of 0.64 and the peak probability for the middle and minor axes equal 0.69. Thus, the histogram results at the final time indicate a good agreement between the corresponding axes.

The comparisons between the displacement histograms for the SGS energy dissipation in the OIC fields in Fig. 2.89 to 2.91 and those in the imperfect model show that the latter gives better agreement between the corresponding axes for this specific result.

As a general observation for the three quantities, the agreement between the clusters for the OICS and the DNS target fields improves over time.

The histograms, in general, show that the MVEEs alignments for the SGS energy dissipation give highest peaks. The MVEEs for $|\tilde{S}|$ is considered the second in terms of the quantity of alignment. The agreement for $|\tilde{\omega}|$ is the weakest. The medians of the histograms are consistent with the previous arguments. Also, most of them show very good alignment, approximately 1, (Figs. 3.26, 3.40 and 3.54).

Generally, all of the quantities give the best agreement at the middle time $T/2$, and the weaker agreement is observed in the middle axes. Comparing the results with those in Chapter 2, little difference is observed. Therefore, for these statistics, the quality of DA is not changed.

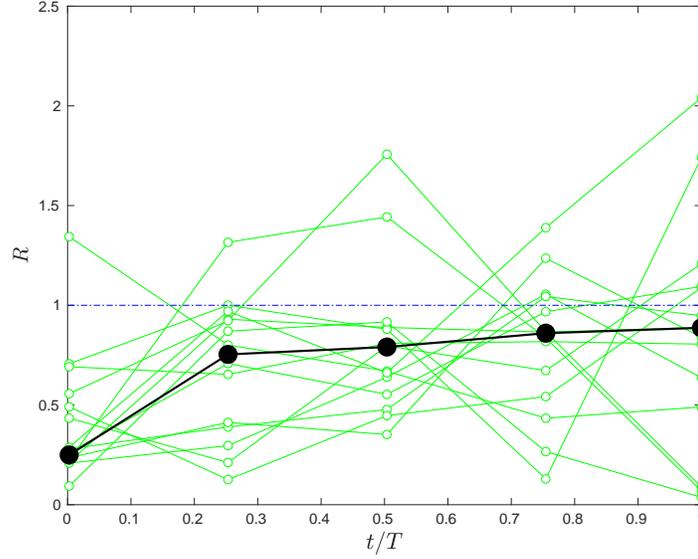


Figure 3.55: Median ratio between the sizes of the clusters for SGS energy dissipation. Black with circles: Median ratios for all clusters. Green with circles: median ratios for clusters of a given size.

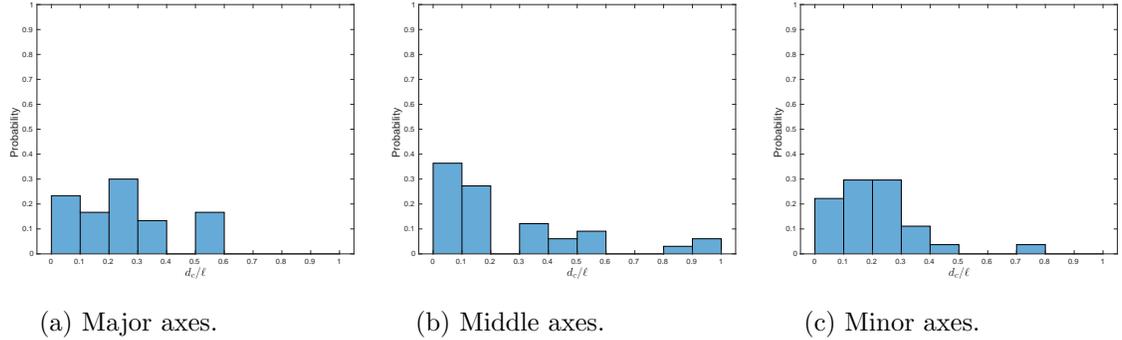
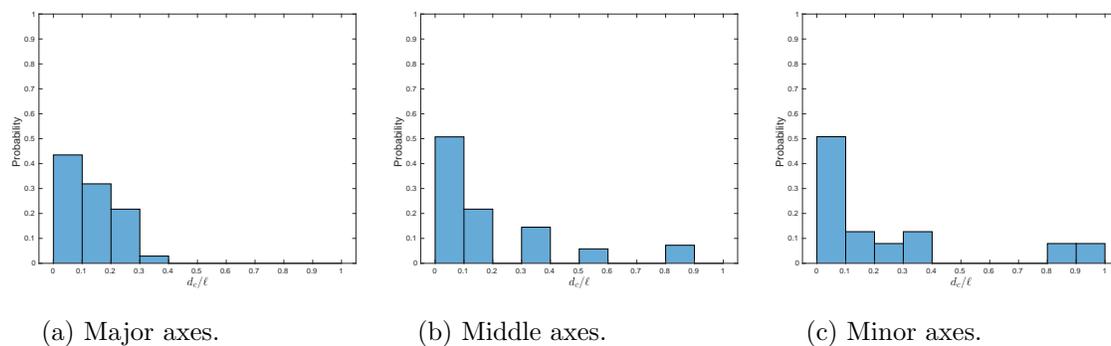
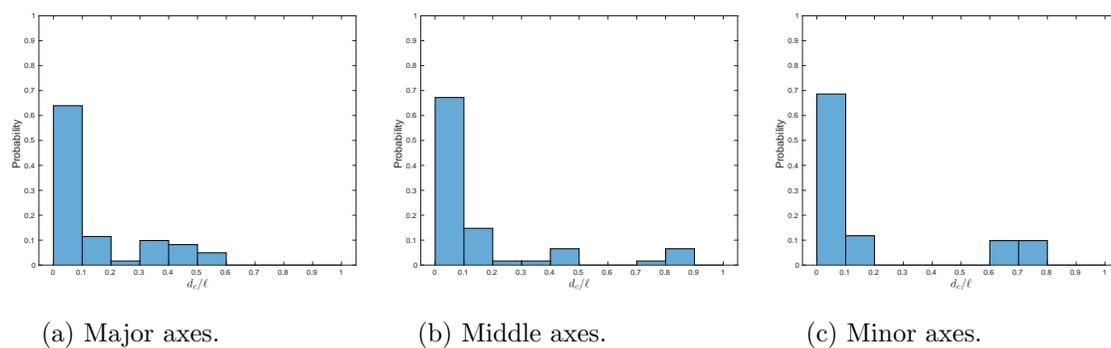


Figure 3.56: d_c/ℓ for SGS energy dissipation at $t = 0T$.

3.4.3 SUMMARY

To summarize, in this chapter we examine the reconstruction of 3D turbulent velocity fields using filtered NS equations with the standard Smagorinsky model as the closure for the SGS stress. The results show that, interestingly, for filter scale $\Delta = 4\delta_x$, the quality of reconstruction is usually better than using the full NS equations. The observation suggests that the smallest scales might not be important for the reconstruction, at least for T at the order of half of the large eddy turn-over time scale and filter scales around $4\delta_x$.

Figure 3.57: d_c/ℓ for SGS energy dissipation at $t = T/2$.Figure 3.58: d_c/ℓ for SGS energy dissipation at $t = T$.

Chapter 4

Conclusions

The reconstruction of the small scales of a sequence of 3D homogeneous turbulent velocity fields has been achieved by applying the four-dimensional variational method (4DVAR). The velocity on a coarse grid from a time sequence of DNS velocity fields is applied as the target. The length of the time sequence is of the order of $1/2$ of the scale of large eddy turn-over time. The main focus is to rebuild the instantaneous distributions for the small scales, and show the accuracy assessment of the data at a grid size corresponding to half of the grid of measurement. In Fourier modes, the number of modes contained in this grid is eight times that of the grid of measurement. In the optimal initial fields, the field of filtered vorticity, the filtered strain rate tensor, and the subgrid-scale energy dissipation Π are studied. Furthermore, some of the geometrical statistics are examined, including the PDFs of the alignment between the filtered vorticity and the eigenvectors of the filtered strain rate tensor; the PDFs of velocity increments, both longitudinal and transverse; in addition to other well-known statistics. These statistics describe how eddies interact with each other which is the key feature of turbulent flows.

The DNS of the Gaussian random field with the same energy spectrum from the optimal initial field is investigated as well. This aims to examine the efficiency of the reconstructed fields as an initial condition to recapture the statistics in the DNS target fields. The statistics of this field are compared with those in the DNS target field, and its improvement with time is compared to the DNS initialized with the optimal field.

Finally, the investigations of the non-local structures that allow focusing on the extraction, description, and classification of turbulent structures are essential because they enable us to capture the geometry of these structures in the small scales of turbulent fields. Therefore, the assessments for the non-local structures reconstruction are presented quantitatively, using MVEE to compare the reconstructed fields and the DNS target fields.

The investigation for the quality of reconstruction has been made, firstly by applying the technique with the NS equations as the model. Then, we filter the NS equations and use

large eddy simulations as a model for the data with the standard Smagorinsky model as the SGS model. Large eddy simulations are used to avoid the vast computational cost in DNS which are essential in order to consider flows with higher Reynolds numbers. Both models have quite similar optimality systems. Nevertheless, the filtered NS equations have an extra term due to the SGS stress tensor which consequently gives a different adjoint equation.

We summarize the results as follows, where rough numbers are given to illustrate the efficacy of the method.

- 1 The results (both non-local structures and geometrical statistics) have shown that the agreement is improved over time, where the rebuilt field tends towards the DNS target. The best agreement is achieved in the cases when the time $t = 2/3T$ for the NS model and $t = 3/4T$ for the imperfect model.
- 2 The DNS of the Gaussian random field with the same energy spectrum from the optimal initial field needs a longer time to recover the DNS target field. The time needed is longer by approximately 10%-15%.
- 3 The reconstruction of the non-local structures with strong $|\omega|$, $|\mathcal{S}|$ and Π is satisfactory. The angle between the MVEEs axes is small, around 20-25 degrees. The reconstructed fields are able to produce close structures with smaller sizes by about 20%-25% to that which exists in the DNS target fields. The locations of the MVEEs are missed by 18%-22% of the axes lengths.
- 4 The results of both models are encouraging because they confirm the efficiency of the 4DVAR in the reconstruction procedures. The imperfect model applying the traditional Smagorinsky model shows the ability to recover the statistics of the DNS target fields quicker than the application of the DA methods based on the NS model.

This study shows that MVEEs have the ability to present useful information about non-local structures for computational fluid dynamics with DA. Also, it demonstrates that only a small portion of the data of the total modes is needed to reconstruct the data of high wave-number modes in the DNS field. The method is expected to have more applications in turbulence research. Several questions to be answered are as follows:

- 1 It is well-known that Smagorinsky models inherit plenty of uncertainties associated with turbulence closure. However, the dynamical Smagorinsky model is used more widely and stable than the classical one, which makes the LES model most applicable. Therefore, using the dynamical Smagorinsky model can be one of the plans for the future as a LES model.
- 2 Increasing the optimisation horizon can improve reconstruction qualities. Our recent work in [56] has shown that effect, where it has been observed that the

agreement between different fields improves with a larger optimisation horizon over time (when T increased to $0.5\tau_L$, and τ_L). However, it is necessary to further extend T beyond one eddy turn over time scale in order to find the limit of the improvement.

- 3 The vorticity field represents the rotation of the velocity field (i.e. both fields characterise the whole velocity flow fields in turbulence). Because of the sensitive nature of the turbulent flow, vorticity involvement in the problem formulation may enhance the quality of the reconstruction. Therefore, investigating the effect of the vorticity in the cost function to reduce the difference to the minimum between the measurement and the velocity field evolved from the reconstructed initial field can also be one of our future plans.

Appendix A

Derivation of the optimality system for the 4DVAR with the Navier-Stokes equations as the model

The method we used here is 4D variational data assimilation where flow optimization is used to find $\boldsymbol{\varphi}(\mathbf{x})$ to minimize the cost function which consists of the velocity and vorticity and given by the formula:

$$J = \frac{1}{2} \int_0^T dt \int_{\Omega} d\mathbf{x} (|\mathbf{u} - \mathbf{v}|_2^2 + \beta_2^2 |\boldsymbol{\omega} - \boldsymbol{\omega}^v|_2^2), \quad (\text{A.1})$$

where β_2 is a constant, $\boldsymbol{\omega} \equiv \nabla \times \mathbf{u}$, and $\boldsymbol{\omega}^v \equiv \nabla \times \mathbf{v}$, subject to constraints:

$$\mathbf{N}(\mathbf{u}) = \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nu \nabla^2 \mathbf{u} - \mathbf{f} = 0, \nabla \cdot \mathbf{u} = 0, \mathbf{u}(\mathbf{x}, 0) = \boldsymbol{\varphi}(\mathbf{x}). \quad (\text{A.2})$$

The optimization problem is solved with the adjoint method. The Lagrangian function is given as

$$L = J - \int_0^T \int_B \boldsymbol{\xi} \cdot \mathbf{N}(\mathbf{u}) dt d\mathbf{x} - \int_0^T \int_B \mu \nabla \cdot \mathbf{u} dt d\mathbf{x} - \int_B \boldsymbol{\lambda} \cdot [\mathbf{u}(\mathbf{x}, 0) - \boldsymbol{\varphi}(\mathbf{x})] d\mathbf{x}, \quad (\text{A.3})$$

where $\boldsymbol{\xi}(\mathbf{x}, t)$, $\boldsymbol{\mu}(\mathbf{x})$, and $\boldsymbol{\lambda}(\mathbf{x})$ are the adjoint variables corresponding to the constraints. B represents the $[0, 2\pi]^3$ periodic box.

The state equations are given by

$$\frac{\partial L}{\partial \boldsymbol{\xi}} = 0, \frac{\partial L}{\partial \boldsymbol{\mu}} = 0, \frac{\partial L}{\partial \boldsymbol{\lambda}} = 0. \quad (\text{A.4})$$

The adjoint equations are given by

$$\frac{\partial L}{\partial \mathbf{u}} = 0, \quad \frac{\partial L}{\partial p} = 0. \quad (\text{A.5})$$

Therefore, the state equations are

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nu \nabla^2 \mathbf{u} - \mathbf{f} = 0, \quad \nabla \cdot \mathbf{u} = 0, \quad \mathbf{u}(\mathbf{x}, t = 0) = \boldsymbol{\varphi}(\mathbf{x}). \quad (\text{A.6})$$

Here, we will tackle the algebra to find $\frac{\partial L}{\partial \mathbf{u}}$, where $\frac{\partial L}{\partial \mathbf{u}}$ is the Gateaux derivative, which is defined by

$$\left\langle \frac{\partial L}{\partial \mathbf{u}}, \tilde{\mathbf{u}} \right\rangle = \lim_{\epsilon \rightarrow 0} \frac{L(\mathbf{u} + \epsilon \tilde{\mathbf{u}}) - L(\mathbf{u})}{\epsilon}, \quad (\text{A.7})$$

where $\tilde{\mathbf{u}}$ is an arbitrary variation. $\frac{\partial L}{\partial \mathbf{u}}$ is the function such that the above equation is true for arbitrary $\tilde{\mathbf{u}}$. To calculate the limit in Eq.A.7, we consider different parts of Eq. A.3 separately. Starting with the integral of $\beta_2^2 |\boldsymbol{\omega} - \boldsymbol{\omega}^v|^2$, its contribution can be obtained by the following steps:

$$\begin{aligned} & \beta_2^2 \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \left[|\nabla \times (\mathbf{u} + \tilde{\mathbf{u}}\epsilon) - \boldsymbol{\omega}^v|^2 - |\nabla \times \mathbf{u} - \boldsymbol{\omega}^v|^2 \right], \\ &= \beta_2^2 \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \left[|(\nabla \times \mathbf{u} - \boldsymbol{\omega}^v) + \epsilon \nabla \times \tilde{\mathbf{u}}|^2 - |\nabla \times \mathbf{u} - \boldsymbol{\omega}^v|^2 \right], \\ &= \beta_2^2 2 \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \left[(\nabla \times \mathbf{u} - \boldsymbol{\omega}^v) \cdot (\nabla \times \tilde{\mathbf{u}}) \right]. \end{aligned} \quad (\text{A.8})$$

Let $\mathbf{A} = 2\beta_2^2 (\nabla \times \mathbf{u} - \boldsymbol{\omega}^v)$. Using $\partial_j (A_i \tilde{u}_k) = A_i \partial_j \tilde{u}_k + \tilde{u}_k \partial_j A_i$, we have

$$\begin{aligned} A_i \epsilon_{ijk} \partial_j \tilde{u}_k &= \epsilon_{ijk} \partial_j (A_i \tilde{u}_k) - \epsilon_{ijk} \tilde{u}_k \partial_j A_i, \\ &= \epsilon_{ijk} \partial_j (A_i \tilde{u}_k) + \tilde{u}_k \epsilon_{kji} \partial_j A_i, \\ &= \epsilon_{ijk} \partial_j (A_i \tilde{u}_k) + \tilde{u}_k (\nabla \times \mathbf{A})_k. \end{aligned} \quad (\text{A.9})$$

Therefore, Eq. A.8 becomes

$$\begin{aligned} & \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \tilde{\mathbf{u}} \cdot (\nabla \times \mathbf{A}) = 2\beta_2^2 \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \tilde{\mathbf{u}} \cdot [\nabla \times (\nabla \times \mathbf{u} - \boldsymbol{\omega}^v)] \\ &= -2\beta_2^2 \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \tilde{\mathbf{u}} \cdot [\nabla^2 \mathbf{u} + \nabla \times \boldsymbol{\omega}^v]. \end{aligned} \quad (\text{A.10})$$

Note that the contribution from the first term on the right-hand side of Eq. A.9 is

$$\begin{aligned}
 & - \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \nabla \cdot [\mathbf{A} \times \tilde{\mathbf{u}}], \\
 & = - \int_0^T dt \int_{\mathbf{x} \in \Gamma} d\Gamma [\mathbf{A} \times \tilde{\mathbf{u}}] \cdot \mathbf{n}, \\
 & = - 2 \int_0^T dt \int_{\mathbf{x} \in \Gamma} d\Gamma [(\nabla \times \mathbf{u} - \boldsymbol{\omega}^v) \times \tilde{\mathbf{u}}] \cdot \mathbf{n},
 \end{aligned} \tag{A.11}$$

where Γ is the boundary. This term is zero due to the periodic BC.

The contribution of $|\mathbf{u} - \mathbf{v}|_2^2$ in the cost function is formed as follows:

$$\begin{aligned}
 & \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left[\frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} |(\mathbf{u} + \tilde{\mathbf{u}}\epsilon) - \mathbf{v}|^2 - \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} |\mathbf{u} - \mathbf{v}|^2 \right], \\
 & = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left[\frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} (\mathbf{u})^2 + \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} 2\epsilon (\mathbf{u} \cdot \tilde{\mathbf{u}}) \right. \\
 & \quad - \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} 2(\mathbf{u} \cdot \mathbf{v}) - \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} 2\epsilon (\mathbf{u} \cdot \mathbf{v}) \\
 & \quad + \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} (\mathbf{v})^2 - \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} (\mathbf{u})^2 \\
 & \quad \left. + \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} 2(\mathbf{u} \cdot \mathbf{v}) - \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} (\mathbf{v})^2 \right], \\
 & = \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \tilde{\mathbf{u}} \cdot (\mathbf{u}) - \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \tilde{\mathbf{u}} \cdot (\mathbf{v}) \\
 & = \frac{1}{2} \int_0^T dt \int_{\mathbf{x} \in \Omega} d\mathbf{x} \tilde{\mathbf{u}} \cdot (\mathbf{u} - \mathbf{v}).
 \end{aligned} \tag{A.12}$$

From the third term of Lagrangian function in Eq. A.3, we have

$$\begin{aligned}
 & \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left[- \int_0^T dt \int_B d\mathbf{x} \left[\mu \nabla \cdot (\mathbf{u} + \tilde{\mathbf{u}}\epsilon) - \mu \nabla \cdot \mathbf{u} \right] \right] \\
 & = - \int_0^T dt \int_B d\mathbf{x} \mu \nabla \cdot \tilde{\mathbf{u}}.
 \end{aligned} \tag{A.13}$$

By using the product rule, Eq. A.13 becomes

$$\begin{aligned}
& \int_0^T dt \int_B d\mathbf{x} \tilde{\mathbf{u}} \cdot \nabla \mu - \int_0^T dt \int_B d\mathbf{x} \nabla \cdot (\mu \tilde{\mathbf{u}}), \\
& = \int_0^T dt \int_B d\mathbf{x} \tilde{\mathbf{u}} \cdot \nabla \mu - \int_0^T dt \int_{\Gamma} d\Gamma \mu \tilde{\mathbf{u}} \cdot \mathbf{n}
\end{aligned} \tag{A.14}$$

For the 2nd term of Lagrangian function in Eq. A.3, we obtain

$$\begin{aligned}
& - \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T dt \int_B d\mathbf{x} \left[\boldsymbol{\xi} \cdot \left[\frac{\partial(\mathbf{u} + \tilde{\mathbf{u}}\epsilon)}{\partial t} + (\mathbf{u} + \tilde{\mathbf{u}}\epsilon) \cdot \nabla(\mathbf{u} + \tilde{\mathbf{u}}\epsilon) \right. \right. \\
& \left. \left. + \nabla p - \nu \frac{\partial^2(\mathbf{u} + \tilde{\mathbf{u}}\epsilon)}{\partial x_j \partial x_j} - f \right] - \left[\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nu \frac{\partial^2 \mathbf{u}}{\partial x_j \partial x_j} - f \right] \right], \\
& = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T dt \int_B d\mathbf{x} \left[-\boldsymbol{\xi} \cdot \epsilon \frac{\partial \tilde{\mathbf{u}}}{\partial t} - \boldsymbol{\xi} \cdot [\tilde{\mathbf{u}}\epsilon \cdot (\nabla \mathbf{u})] + \boldsymbol{\xi} \cdot \epsilon \nu \frac{\partial^2 \tilde{\mathbf{u}}}{\partial x_j \partial x_j} - \boldsymbol{\xi} \cdot [\epsilon \mathbf{u} \cdot (\nabla \tilde{\mathbf{u}})] \right], \\
& = \int_0^T dt \int_B d\mathbf{x} \left[-\boldsymbol{\xi} \cdot \frac{\partial \tilde{\mathbf{u}}}{\partial t} - \boldsymbol{\xi} \cdot [\tilde{\mathbf{u}} \cdot (\nabla \mathbf{u})] + \boldsymbol{\xi} \cdot \nu \frac{\partial^2 \tilde{\mathbf{u}}}{\partial x_j \partial x_j} - \boldsymbol{\xi} \cdot [\mathbf{u} \cdot (\nabla \tilde{\mathbf{u}})] \right].
\end{aligned} \tag{A.15}$$

Also for the 2nd term of Eq. A.3, we find

$$\begin{aligned}
& - \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T dt \int_B d\mathbf{x} \left[\boldsymbol{\xi} \cdot [\nabla(p + \tilde{p}\epsilon) - \nabla p] \right] \\
& = - \int_0^T dt \int_B d\mathbf{x} \boldsymbol{\xi} \cdot \nabla \tilde{p}, \\
& = \int_0^T dt \int_B d\mathbf{x} \nabla \cdot \boldsymbol{\xi} \tilde{p} - \int_0^T dt \int_B d\mathbf{x} \nabla \cdot (\boldsymbol{\xi} \tilde{p}), \\
& = \int_0^T dt \int_B d\mathbf{x} \nabla \cdot \boldsymbol{\xi} \tilde{p} - \int_0^T dt \int_{\Gamma} d\Gamma \boldsymbol{\xi} \tilde{p} \cdot \mathbf{n}
\end{aligned} \tag{A.16}$$

For the simplification of Eq. A.15, we have

$$\text{First term} = \int_0^T dt \int_B d\mathbf{x} \tilde{\mathbf{u}} \cdot \partial_t \boldsymbol{\xi} - \int_0^T dt \int_B d\mathbf{x} \partial_t (\tilde{\mathbf{u}} \cdot \boldsymbol{\xi}), \tag{A.17}$$

$$\text{Second term} = - \int_0^T dt \int_B d\mathbf{x} \tilde{\mathbf{u}} \cdot [(\nabla \mathbf{u}) \cdot \boldsymbol{\xi}], \tag{A.18}$$

$$\begin{aligned}
\text{Third term} & = \int_0^T dt \int_B d\mathbf{x} \nu \tilde{\mathbf{u}} \cdot \nabla^2 \boldsymbol{\xi} + \int_0^T dt \int_B d\mathbf{x} \nu \left[\partial_j (\partial_j \tilde{u}_k \xi_k) - \partial_j (\tilde{u}_k \partial_j \xi_k) \right], \\
& \tag{A.19}
\end{aligned}$$

$$\text{Fourth term} = \int_0^T dt \int_B d\mathbf{x} \tilde{\mathbf{u}} \cdot (\mathbf{u} \cdot \nabla \boldsymbol{\xi}) - \int_0^T dt \int_B d\mathbf{x} \nabla \cdot [\mathbf{u} (\tilde{\mathbf{u}} \cdot \boldsymbol{\xi})]. \tag{A.20}$$

Eqs. A.17, A.19 and A.20 can be simplified as

$$\text{First term} = \int_0^T dt \int_B dx \tilde{\mathbf{u}} \cdot \partial_t \boldsymbol{\xi} - \int_B dx \tilde{\mathbf{u}} \cdot \boldsymbol{\xi} \Big|_{t=T} + \int_B dx \tilde{\mathbf{u}} \cdot \boldsymbol{\xi} \Big|_{t=0}, \quad (\text{A.21})$$

$$\text{Third term} = \int_0^T dt \int_B dx \nu \tilde{\mathbf{u}} \cdot \nabla^2 \boldsymbol{\xi} + \int_0^T dt \int_\Gamma d\Gamma \nu \frac{\partial \tilde{\mathbf{u}}}{\partial n} \cdot \boldsymbol{\xi} - \int_0^T dt \int_\Gamma d\Gamma \nu \frac{\partial \boldsymbol{\xi}}{\partial n} \cdot \tilde{\mathbf{u}}, \quad (\text{A.22})$$

$$\text{Fourth term} = \int_0^T dt \int_B dx \tilde{\mathbf{u}} \cdot (\mathbf{u} \cdot \nabla \boldsymbol{\xi}) - \int_0^T dt \int_\Gamma d\Gamma \mathbf{u} \cdot \mathbf{n} (\tilde{\mathbf{u}} \cdot \boldsymbol{\xi}). \quad (\text{A.23})$$

For the last term in Eq. A.3, we have

$$-\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_B dx \left[\boldsymbol{\lambda} \cdot [(\mathbf{u} + \tilde{\mathbf{u}}\epsilon) - \boldsymbol{\varphi}] - \boldsymbol{\lambda} \cdot (\mathbf{u} - \boldsymbol{\varphi}) \right] = - \int_B \boldsymbol{\lambda} \cdot \tilde{\mathbf{u}}(x, 0) dx. \quad (\text{A.24})$$

By invoking Eqs. A.10, A.11, A.12, A.14, A.18, A.21, A.22, A.23 and A.24, and $\frac{\partial L}{\partial \mathbf{u}} = 0$, we gain the following:

$$\begin{aligned} 0 &= \int_0^T dt \int_B dx \left[\tilde{\mathbf{u}} \cdot \nabla \mu - 2\beta_2^2 \tilde{\mathbf{u}} \cdot [\nabla^2 \mathbf{u} + \nabla \times \boldsymbol{\omega}^v] + \tilde{\mathbf{u}} \cdot \partial_t \boldsymbol{\xi} - \tilde{\mathbf{u}} \cdot [(\nabla \mathbf{u}) \cdot \boldsymbol{\xi}] \right. \\ &\quad \left. + \nu \tilde{\mathbf{u}} \cdot \nabla^2 \boldsymbol{\xi} + \tilde{\mathbf{u}} \cdot (\mathbf{u} \cdot \nabla \boldsymbol{\xi}) \right] + \int_0^T dt \int_{x \in \Gamma} d\Gamma \left[-\mu \tilde{\mathbf{u}} \cdot \mathbf{n} - 2\beta_2^2 [(\nabla \times \mathbf{u} - \boldsymbol{\omega}^v) \times \tilde{\mathbf{u}}] \cdot \mathbf{n} \right. \\ &\quad \left. - \mathbf{u} \cdot \mathbf{n} (\tilde{\mathbf{u}} \cdot \boldsymbol{\xi}) + \nu \frac{\partial \tilde{\mathbf{u}}}{\partial n} \cdot \boldsymbol{\xi} - \nu \frac{\partial \boldsymbol{\xi}}{\partial n} \cdot \tilde{\mathbf{u}} \right] + \left[- \int_B dx \tilde{\mathbf{u}} \cdot \boldsymbol{\xi} \Big|_{t=T} + \int_B dx \tilde{\mathbf{u}} \cdot \boldsymbol{\xi} \Big|_{t=0} \right] \\ &\quad - \int_B \boldsymbol{\lambda} \cdot \tilde{\mathbf{u}}(x, 0) dx, \end{aligned} \quad (\text{A.25})$$

i. e.,

$$\begin{aligned} 0 &= \int_0^T dt \int_B dx \left[\tilde{\mathbf{u}} \cdot [\nabla \mu - 2\beta_2^2 [\nabla^2 \mathbf{u} + \nabla \times \boldsymbol{\omega}^v] + \partial_t \boldsymbol{\xi} - \nabla \mathbf{u} \cdot \boldsymbol{\xi} + \nu \nabla^2 \boldsymbol{\xi} + \mathbf{u} \cdot \nabla \boldsymbol{\xi}] \right. \\ &\quad \left. + \int_0^T dt \int_{x \in \Gamma} d\Gamma \left[-\mu \tilde{\mathbf{u}} \cdot \mathbf{n} - \mathbf{u} \cdot \mathbf{n} (\tilde{\mathbf{u}} \cdot \boldsymbol{\xi}) + \nu \frac{\partial \tilde{\mathbf{u}}}{\partial n} \cdot \boldsymbol{\xi} - 2\beta_2^2 [(\nabla \times \mathbf{u} - \boldsymbol{\omega}^v) \times \tilde{\mathbf{u}}] \cdot \mathbf{n} \right. \right. \\ &\quad \left. \left. - \nu \frac{\partial \boldsymbol{\xi}}{\partial n} \cdot \tilde{\mathbf{u}} \right] + \left[- \int_B dx \tilde{\mathbf{u}} \cdot \boldsymbol{\xi} \Big|_{t=T} + \int_B dx \tilde{\mathbf{u}} \cdot \boldsymbol{\xi} \Big|_{t=0} \right] - \int_B \boldsymbol{\lambda} \cdot \tilde{\mathbf{u}}(x, 0) dx. \end{aligned} \quad (\text{A.26})$$

Since the variation $\tilde{\mathbf{u}}$ is arbitrary we obtain from A.26

$$\partial_t \boldsymbol{\xi} + \mathbf{u} \cdot \nabla \boldsymbol{\xi} - \nabla \mathbf{u} \cdot \boldsymbol{\xi} + \nabla \mu + \nu \nabla^2 \boldsymbol{\xi} - 2\beta_2^2 (\nabla^2 \mathbf{u} + \nabla \times \boldsymbol{\omega}^v) = \mathbf{F}, \boldsymbol{\xi}(\mathbf{x}, t = T) = 0. \quad (\text{A.27})$$

where $\mathbf{F}(\mathbf{x}, t) = -[\mathbf{u}(\mathbf{x}, t) - \mathbf{v}(\mathbf{x}, t)]$ when $\mathbf{x} \in \Omega$ and $\mathbf{F} = 0$ otherwise. And simultaneously $\boldsymbol{\xi}(\mathbf{x}, 0) = \boldsymbol{\lambda}$.

As we can notice, Eqs. A.27 has an extra term which related to the vorticity field at the cost function. Numerical tests show that ignoring this term does not change the results significantly. Therefore, in Chapter 2, we have set $\beta_2 = 0$.

Since variation \tilde{p} in the state is arbitrary, using $\frac{\partial L}{\partial p} = 0$ and Eq. A.16, we have

$$0 = \nabla \cdot \boldsymbol{\xi}. \quad (\text{A.28})$$

Thus, the adjoint equations for Chapter 2 is given as 2.7. The boundary integrals are identically zero because of the periodic BC.

By the same Eq. A.3, we also find

$$-\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_B d\mathbf{x} \left[L(\boldsymbol{\varphi} + \tilde{\boldsymbol{\varphi}}\epsilon) - L(\boldsymbol{\varphi}) \right] = - \int_B \tilde{\boldsymbol{\varphi}} \cdot \boldsymbol{\lambda} \, d\mathbf{x}. \quad (\text{A.29})$$

Therefore, by the definition of $\frac{\partial L}{\partial \boldsymbol{\varphi}}$, we have

$$\left\langle \frac{\partial L}{\partial \boldsymbol{\varphi}}, \tilde{\boldsymbol{\varphi}} \right\rangle = \int_B \tilde{\boldsymbol{\varphi}} \cdot \boldsymbol{\lambda} \, d\mathbf{x}.$$

Therefore,

$$\boldsymbol{\lambda} = \frac{\partial L}{\partial \boldsymbol{\varphi}}. \quad (\text{A.30})$$

By the substitution of $\boldsymbol{\xi}(\mathbf{x}, 0) = \boldsymbol{\lambda}$, in Eq. A.30, we obtain the gradient of the cost function.

$$\frac{\partial L}{\partial \boldsymbol{\varphi}} = \boldsymbol{\xi}(\mathbf{x}, 0),$$

i. e.

$$\frac{DJ}{D\boldsymbol{\varphi}} = \boldsymbol{\xi}(\mathbf{x}, 0). \quad (\text{A.31})$$

Appendix B

The adjoint equation for the filtered NS equations with the standard Smagorinsky model

The optimality system is quite similar to what we presented in Appendix A. However, the filtered NS equations have an extra term due to the SGS stress tensor. Therefore, the adjoint equation for the filtered NSE with standard Smagorinsky model is different, and given below.

The filtered Navier-Stocks equations are given in the form:

$$(FNS)_i = \partial_t \bar{u}_i + \bar{u}_m \partial_m \bar{u}_i + \partial_i \bar{p} - \partial_m \tau_{im}^r - \nu \nabla^2 \bar{u}_i = 0, \quad (\text{B.1})$$

where \bar{u}_i is the filtered velocity. The SGS stress contribution to the adjoint equation is given by an additional term

$$T_k = \frac{\partial}{\partial u_k} \int_0^T \int_B \xi_i (-\partial_m \tau_{im}^r) d^3 \mathbf{x} dt. \quad (\text{B.2})$$

Here $\frac{\partial}{\partial u_k}$ has been used to denote the functional derivative $\frac{\partial}{\partial u_k(\mathbf{x}', t')}$. In Appendix A, the derivative is calculated using its relation with the Gateaux derivative. However, formally, $\frac{\partial}{\partial u_k}$ can also be calculated following the same rules for standard differentiation, with the additional relation:

$$\frac{\partial u_m(\mathbf{x}, t)}{\partial u_k(\mathbf{x}', t')} = \delta_{mk} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t'). \quad (\text{B.3})$$

In this section, we use Eq. B.3 to calculate the derivative in Eq. B.2. This contribution, with the non-dynamical Smagorinsky model for the SGS stresses, will be our focus, where

$$\tau_{im}^r = -\nu_t \bar{S}_{im} = -(C_s \Delta)^2 |\bar{S}| \bar{S}_{im} = -(C_s \Delta)^2 (2\bar{S}_{pq} \bar{S}_{qp})^{1/2} \bar{S}_{im}. \quad (\text{B.4})$$

Therefore,

$$T_k = \frac{\partial}{\partial u_k} \int_0^T \int_B \xi_i \partial_m [-(C_s \Delta)^2 |\bar{S}| \bar{S}_{im}] d^3 \mathbf{x} dt. \quad (\text{B.5})$$

$$= \frac{\partial}{\partial u_k} \int_0^T \int_B (C_s \Delta)^2 |\bar{S}| \bar{S}_{im} (\partial_m \xi_i) d^3 \mathbf{x} dt, \quad (\text{B.6})$$

using the periodic BCs. The derivative of the integral, therefore, leads us to the Smagorinsky model contribution in the adjoint of the FNS equation. We have

$$T_k = \int_0^T \int_B (C_s \Delta)^2 \frac{\partial}{\partial u_k} [|\bar{S}| \bar{S}_{im}] (\partial_m \xi_i) d^3 \mathbf{x} dt + \int_0^T \int_B \left[\frac{\partial (C_s \Delta)^2}{\partial u_k} \right] |\bar{S}| \bar{S}_{im} (\partial_m \xi_i) d^3 \mathbf{x} dt. \quad (\text{B.7})$$

We split the integration above into separate parts to make the derivations easy to understand. Thus, for the first term of Eq. B.7 we have

$$\begin{aligned} \frac{\partial}{\partial u_k} [|\bar{S}| \bar{S}_{im}] &= |\bar{S}| \frac{\partial}{\partial u_k} \bar{S}_{im} + \bar{S}_{im} \frac{\partial}{\partial u_k} (2\bar{S}_{pq} \bar{S}_{qp})^{1/2} \\ &= |\bar{S}| \frac{\partial \bar{S}_{im}}{\partial u_k} + \frac{1}{2} \bar{S}_{im} |\bar{S}|^{-1} (2 \times 2) \bar{S}_{pq} \frac{\partial \bar{S}_{pq}}{\partial u_k} \\ &= |\bar{S}| \frac{\partial \bar{S}_{im}}{\partial u_k} + 2 |\bar{S}|^{-1} \bar{S}_{im} \bar{S}_{pq} \frac{\partial \bar{S}_{pq}}{\partial u_k}. \end{aligned} \quad (\text{B.8})$$

By substituting Eq. B.8 into the first integral (called A_k) of Eq. B.7,

$$A_k = \int_0^T \int_B (C_s \Delta)^2 \left[|\bar{S}| \frac{\partial \bar{S}_{im}}{\partial u_k} + 2 |\bar{S}|^{-1} \bar{S}_{im} \bar{S}_{pq} \frac{\partial \bar{S}_{pq}}{\partial u_k} \right] (\partial_m \xi_i) d^3 \mathbf{x} dt. \quad (\text{B.9})$$

To simplify Eq. B.9, we have

$$\begin{aligned} A_k &= \int_0^T \int_B (C_s \Delta)^2 |\bar{S}| \frac{\partial \bar{S}_{im}}{\partial u_k} \partial_m \xi_i d^3 \mathbf{x} dt + \\ &\quad \int_0^T \int_B 2 (C_s \Delta)^2 |\bar{S}|^{-1} \bar{S}_{im} \bar{S}_{pq} \frac{\partial \bar{S}_{pq}}{\partial u_k} \partial_m \xi_i d^3 \mathbf{x} dt. \end{aligned} \quad (\text{B.10})$$

Therefore, the first integration in Eq. B.10 can be simplified as

$$\begin{aligned}
a_k &= \frac{1}{2} \int_0^T \int_B (C_s \Delta)^2 |\bar{S}| (\partial_m \xi_i) \left[\frac{\partial}{\partial x_i} \left(\frac{\partial u_m}{\partial u_k} \right) + \frac{\partial}{\partial x_m} \left(\frac{\partial u_i}{\partial u_k} \right) \right] d^3 \mathbf{x} dt, \\
&= -\frac{1}{2} \int_0^T \int_B \partial_i \left[(C_s \Delta)^2 |\bar{S}| (\partial_m \xi_i) \right] \delta_{mk} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') d^3 \mathbf{x} dt - \\
&\quad \frac{1}{2} \int_0^T \int_B \partial_m \left[(C_s \Delta)^2 |\bar{S}| (\partial_m \xi_i) \right] \delta_{ik} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') d^3 \mathbf{x} dt, \tag{B.11}
\end{aligned}$$

where integration by parts has been used. Therefore, Eq. B.11 becomes

$$\begin{aligned}
a_k &= -\frac{1}{2} \partial_i \left[(C_s \Delta)^2 |\bar{S}| (\partial_k \xi_i) \right] - \frac{1}{2} \partial_m \left[(C_s \Delta)^2 |\bar{S}| (\partial_m \xi_k) \right] \\
&= -\partial_i \left[(C_s \Delta)^2 |\bar{S}| X_{ki} \right], \tag{B.12}
\end{aligned}$$

where $X_{ki} = \frac{1}{2} (\partial_k \xi_i + \partial_i \xi_k)$.

Also, the second integration in Eq. B.10, denoted by b_k , is simplified as

$$\begin{aligned}
b_k &= \int_0^T \int_B (C_s \Delta)^2 |\bar{S}|^{-1} \bar{S}_{im} \bar{S}_{pq} (\partial_m \xi_i) \left[\frac{\partial}{\partial x_p} \left(\frac{\partial u_q}{\partial u_k} \right) + \frac{\partial}{\partial x_q} \left(\frac{\partial u_p}{\partial u_k} \right) \right] d^3 \mathbf{x} dt, \\
&= -\int_0^T \int_B \partial_p \left[(C_s \Delta)^2 |\bar{S}|^{-1} \bar{S}_{im} \bar{S}_{pq} (\partial_m \xi_i) \right] \delta_{qk} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') d^3 \mathbf{x} dt - \\
&\quad \int_0^T \int_B \partial_q \left[(C_s \Delta)^2 |\bar{S}|^{-1} \bar{S}_{im} \bar{S}_{pq} (\partial_m \xi_i) \right] \delta_{pk} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') d^3 \mathbf{x} dt. \tag{B.13}
\end{aligned}$$

Thus, Eq. B.13 becomes

$$\begin{aligned}
b_k &= -\partial_p \left[(C_s \Delta)^2 |\bar{S}|^{-1} \bar{S}_{im} X_{im} \bar{S}_{pk} \right] - \partial_q \left[(C_s \Delta)^2 |\bar{S}|^{-1} \bar{S}_{im} X_{im} \bar{S}_{qk} \right], \\
&= -2(C_s \Delta)^2 \partial_p \left[\bar{S}_{kp} |\bar{S}|^{-1} \bar{S}_{im} X_{im} \right]. \tag{B.14}
\end{aligned}$$

Thus, by invoking Eqs. B.12 and B.14 and knowing that the last part of the integration of Eq. B.7 equals to zero, we have:

$$\begin{aligned}
T_k &= -(C_s \Delta)^2 \partial_i \left[|\bar{S}| X_{ki} \right] - (C_s \Delta)^2 \partial_i \left[\frac{2\bar{S}_{ki} \bar{S}_{pm}}{|\bar{S}|} X_{pm} \right], \\
&= -(C_s \Delta)^2 \partial_i \left[\frac{2\bar{S}_{pm} \bar{S}_{pm} X_{ik} + 2\bar{S}_{ki} \bar{S}_{pm} X_{pm}}{|\bar{S}|} \right]. \tag{B.15}
\end{aligned}$$

Therefore, the adjoint equation is :

$$\begin{aligned}\partial_t \boldsymbol{\xi} + \bar{\mathbf{u}} \cdot \nabla \boldsymbol{\xi} - \nabla \bar{\mathbf{u}} \cdot \boldsymbol{\xi} + \nabla \mu + \nu \nabla^2 \boldsymbol{\xi} + (C_s \Delta)^2 \nabla \cdot \left[\frac{2X(\bar{S} : \bar{S}) + 2\bar{S}^T(\bar{S} : X)}{|\bar{S}|} \right] &= \mathbf{F}, \\ \nabla \cdot \boldsymbol{\xi} &= 0, \\ \boldsymbol{\xi}(\mathbf{x}, t = T) &= 0.\end{aligned}\tag{B.16}$$

The rest of the derivations can be seen in the previous section.

References

- [1] Guillermo Artana, Ada Cammilleri, Johan Carlier, and Etienne Memin. Strong and weak constraint variational assimilations for reduced order fluid flow modeling. *Journal of Computational Physics*, 231(8):3264–3288, 2012.
- [2] Stanley L Barnes. A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology*, 3(4):396–409, 1964.
- [3] Páll Bergthörsson and Bo R Döös. Numerical weather map analysis. *Tellus*, 7(3):329–340, 1955.
- [4] Ivan Bermejo-Moreno and DI Pullin. On the non-local geometry of turbulence. *Journal of Fluid Mechanics*, 603:101–135, 2008.
- [5] Luigi Carlo Berselli, Traian Iliescu, and William J Layton. *Mathematics of large eddy simulation of turbulent flows*. Springer Science & Business Media, 2005.
- [6] Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.
- [7] Thomas R Bewley and Bartosz Protas. Skin friction and pressure: the footprints of turbulence. *Physica D: Nonlinear Phenomena*, 196(1-2):28–44, 2004.
- [8] F Bouttier and P Courtier. Data assimilation concepts and methods march 1999. *Meteorological training course lecture series. ECMWF*, 718:59, 2002.
- [9] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [10] Dan G Cacuci. Sensitivity theory for nonlinear systems. i. nonlinear functional analysis approach. *Journal of Mathematical Physics*, 22(12):2794–2802, 1981.
- [11] Dan G Cacuci, CF Weber, EM Oblow, and JH Marable. Sensitivity theory for general systems of nonlinear equations. *Nuclear Science and Engineering*, 75(1):88–110, 1980.
- [12] Jules G Charney, Ragnar Fjørtoft, and John Von Neumann. Numerical integration of the barotropic vorticity equation. In *The Atmosphere A Challenge*, pages 267–284. Springer, 1990.

- [13] Qigang Chen, Qiang Zhong, Meilan Qi, and Xingkui Wang. Comparison of vortex identification criteria for planar velocity fields in wall turbulence. *Physics of Fluids*, 27(8):085101, 2015.
- [14] Mattias Chevalier, Jérôme Hœpffner, Thomas R Bewley, and Dan S Henningson. State estimation in wall-bounded flow systems. part 2. turbulent flows. *Journal of Fluid Mechanics*, 552:167–187, 2006.
- [15] CH Colburn, JB Cessna, and TR Bewley. State estimation in wall-bounded flow systems. part 3. the ensemble kalman filter. *Journal of Fluid Mechanics*, 682:289–303, 2011.
- [16] Benoit Combés, Dominique Heitz, Anthony Guibert, and Etienne Mémin. A particle filter to reconstruct a free-surface flow from a depth camera. *Fluid Dynamics Research*, 47(5):051404, sep 2015.
- [17] Peter Constantin, Itamar Procaccia, and Daniel Segel. Creation and dynamics of vortex tubes in three-dimensional turbulence. *Physical Review E*, 51(4):3207, 1995.
- [18] Richard Courant and David Hilbert. *Methods of Mathematical Physics: Partial Differential Equations*. John Wiley & Sons, 2008.
- [19] Philippe Courtier, John Derber, Ron Errico, Jean-Francois Louis, and Tomislava VukiEvi. Important literature on the use of adjoint, variational methods and the kalman filter in meteorology. *Tellus A: Dynamic Meteorology and Oceanography*, 45(5):342–357, 1993.
- [20] Philippe Courtier and Olivier Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation. ii: Numerical results. *Quarterly Journal of the Royal Meteorological Society*, 113(478):1329–1347, 1987.
- [21] George P Cressman. An operational objective analysis system. *Mon. Wea. Rev*, 87(10):367–374, 1959.
- [22] Juan D’adamo, Nicolas Papadakis, Etienne Mmin, and Guillermo Artana. Variational assimilation of pod low-order dynamical systems. *Journal of Turbulence*, 8:N9, 2007.
- [23] Peter A. Davidson. *Turbulence: An Introduction for Scientists and Engineers*. Oxford University Press, Oxford, 2004.
- [24] DE Edmunds. *Optimal control of systems governed by partial differential equations*, 1972.
- [25] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

-
- [26] G. Evensen. The ensemble kalman filter for combined state and parameter estimation. *IEEE Control Systems Magazine*, 29(3):83–104, June 2009.
- [27] Michael Fisher, Mand Leutbecher, and GA Kelly. On the equivalence between kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 131(613):3235–3246, 2005.
- [28] Dimitry P. G. Foures, Nicolas Dovetta, Denis Sipp, and Peter J. Schmid. A data-assimilation method for reynolds-averaged navierstokes-driven mean flow reconstruction. *Journal of Fluid Mechanics*, 759:404431, 2014.
- [29] Michael Ghil and Paola Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. In *Advances in geophysics*, volume 33, pages 141–266. Elsevier, 1991.
- [30] Bruce Gilchrist and George P Cressman. An experiment in objective analysis. *Tellus*, 6(4):309–318, 1954.
- [31] A. Gronskis, D. Heitz, and E. Mmin. Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. *Journal of Computational Physics*, 242:480 – 497, 2013.
- [32] Peter E Hamlington, Jörg Schumacher, and Werner JA Dahm. Direct assessment of vorticity alignment with local and nonlocal strain rates in turbulent flows. *Physics of Fluids*, 20(11):111703, 2008.
- [33] Peter E Hamlington, Jörg Schumacher, and Werner JA Dahm. Local and nonlocal strain rate fields and vorticity alignment in turbulent flows. *Physical Review E*, 77(2):026303, 2008.
- [34] Toshiyuki Hayase. Numerical simulation of real-world flows. *Fluid Dynamics Research*, 47(5):051201, sep 2015.
- [35] Dominique Heitz, Etienne Mémin, and Christoph Schnörr. Variational fluid flow measurements from image sequences: synopsis and perspectives. *Experiments in Fluids*, 48(3):369–393, Mar 2010.
- [36] Chad Higgins, C Meneveau, and MB Parlange. Energy dissipation in les: dependence on flow structure and effects of eigenvector alignments. Technical report, 2004.
- [37] Kiyosi Horiuti. Roles of non-aligned eigenvectors of strain-rate and subgrid-scale stress tensors in turbulence generation. *Journal of Fluid Mechanics*, 491:65–100, 2003.

- [38] Kayo Ide, Philippe Courtier, Michael Ghil, and Andrew C Lorenc. Unified notation for data assimilation: Operational, sequential and variational (gtspecial issue\data assimilation in meteorology and oceanography: Theory and practice). *Journal of the Meteorological Society of Japan. Ser. II*, 75(1B):181–189, 1997.
- [39] Heikki Järvinen, Erik Andersson, and François Bouttier. Variational assimilation of time sequences of surface observations with serially correlated errors. *Tellus A*, 51(4):469–488, 1999.
- [40] Volker John. *Large eddy simulation of turbulent incompressible flows: analytical and numerical results for a class of LES models*, volume 34. Springer Science & Business Media, 2012.
- [41] Volker John and William J Layton. Analysis of numerical errors in large eddy simulation. *SIAM Journal on Numerical Analysis*, 40(3):995–1020, 2002.
- [42] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.
- [43] Eugenia Kalnay. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press, 2003.
- [44] Eugenia Kalnay, Hong Li, Takemasa Miyoshi, Shu-Chih Yang, and Joaquim Ballabrera-Poy. 4-d-var or ensemble kalman filter? *Tellus A: Dynamic Meteorology and Oceanography*, 59(5):758–773, 2007.
- [45] Hiroshi Kato and Shigeru Obayashi. Approach for uncertainty of turbulence modeling based on data assimilation technique. *Computers & Fluids*, 85:2 – 7, 2013. International Workshop on Future of CFD and Aerospace Sciences.
- [46] Hiroshi Kato, Akira Yoshizawa, Genta Ueno, and Shigeru Obayashi. A data assimilation methodology for reconstructing turbulent flows around aircraft. *Journal of Computational Physics*, 283:559 – 581, 2015.
- [47] Václav Kolář. Vortex identification: New requirements and limitations. *International journal of heat and fluid flow*, 28(4):638–652, 2007.
- [48] Cristian C. Lalescu, Charles Meneveau, and Gregory L. Eyink. Synchronization of chaos in fully developed turbulence. *Phys. Rev. Lett.*, 110:084102, Feb 2013.
- [49] Cornelius Lanczos. *The variational principles of mechanics*. Courier Corporation, 2012.
- [50] B. E. Launder, G. J. Reece, and W. Rodi. Progress in the development of a reynolds-stress turbulence closure. *Journal of Fluid Mechanics*, 68(3):537–566, 1975.
- [51] François-Xavier Le Dimet and Olivier Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A: Dynamic Meteorology and Oceanography*, 38(2):97–110, 1986.

- [52] T. Leung, N. Swaminathan, and P. A. Davidson. Geometry and interaction of structures in homogeneous isotropic turbulence. *Journal of Fluid Mechanics*, 710:453481, 2012.
- [53] Emmanuel Lévêque and Aurore Naso. Introduction of longitudinal and transverse lagrangian velocity increments in homogeneous and isotropic turbulence. *EPL (Europhysics Letters)*, 108(5):54004, 2014.
- [54] John M Lewis and John C Derber. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus A*, 37(4):309–322, 1985.
- [55] Yi Li, Eric Perlman, Minping Wan, Yunke Yang, Charles Meneveau, Randal Burns, Shiyi Chen, Alexander Szalay, and Gregory Eyink. A public turbulence database cluster and applications to study lagrangian evolution of velocity increments in turbulence. *Journal of Turbulence*, (9):N31, 2008.
- [56] Yi Li, Jianlei Zhang, Gang Dong, and Naseer S. Abdullah. Small scale reconstruction in three-dimensional kolmogorov flows using four-dimensional variational data assimilation, 2019.
- [57] Zhiyong Li, Huaibao Zhang, Sean C.C. Bailey, Jesse B. Hoagg, and Alexandre Martin. A data-driven adaptive reynolds-averaged navier-stokes $k\omega$ model for turbulent flow. *Journal of Computational Physics*, 345:111 – 131, 2017.
- [58] Andrew C Lorenc. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112(474):1177–1194, 1986.
- [59] Thomas S Lund and Michael M Rogers. An improved measure of strain state probability in turbulent flows. *Physics of Fluids*, 6(5):1838–1847, 1994.
- [60] M. Meldi and A. Poux. A reduced order model based on kalman filtering for sequential data assimilation of turbulent flows. *Journal of Computational Physics*, 347:207 – 234, 2017.
- [61] F. R. Menter. Two-equation eddy-viscosity turbulence models for engineering applications, August 1994.
- [62] V. Mons, J.-C. Chassaing, T. Gomez, and P. Sagaut. Is isotropic turbulence decay governed by asymptotic behavior of large scales? an eddy-damped quasi-normal markovian-based data assimilation study. *Physics of Fluids*, 26(11):115105, 2014.
- [63] V. Mons, J.-C. Chassaing, T. Gomez, and P. Sagaut. Reconstruction of unsteady viscous flows using data assimilation schemes. *Journal of Computational Physics*, 316:255 – 280, 2016.
- [64] Vincent Mons, Jean-Camille Chassaing, and Pierre Sagaut. Optimal sensor placement for variational data assimilation of unsteady flows past a rotationally oscillating cylinder. *Journal of Fluid Mechanics*, 823:230277, 2017.

- [65] Nima Moshtagh et al. Minimum volume enclosing ellipsoid. *Convex optimization*, 111:112, 2005.
- [66] IM Navon. Implementation of a posteriori methods for enforcing conservation of potential enstrophy and mass in discretized shallow-water equations models. *Monthly Weather Review*, 109(5):946–958, 1981.
- [67] IM Navon and R De Villiers. Combined penalty multiplier optimization methods to enforce integral invariants conservation. *Monthly Weather Review*, 111(6):1228–1243, 1983.
- [68] Ionel M Navon. Data assimilation for numerical weather prediction: a review. In *Data assimilation for atmospheric, oceanic and hydrologic applications*, pages 21–65. Springer, 2009.
- [69] O.G.S. The theory of homogeneous turbulence. by g. k. batchelor. cambridge monographs on mechanics and applied mathematics. cambridge university press, 1953. pp. xi, 197; 28 figs. 25s. *Quarterly Journal of the Royal Meteorological Society*, 79(341):457–458, 1953.
- [70] Julio M Ottino and JM Ottino. *The kinematics of mixing: stretching, chaos, and transport*, volume 3. Cambridge university press, 1989.
- [71] Oktay Özcan, Knud Erik Meyer, and Poul Scheel Larsen. Measurement of mean rotation and strain-rate tensors by using stereoscopic piv. *Experiments in fluids*, 39(4):771–783, 2005.
- [72] RA Panofsky. Objective weather-map analysis. *Journal of Meteorology*, 6(6):386–392, 1949.
- [73] Stephen B. Pope. *Turbulent flows*. Cambridge University Press, Cambridge, 2000.
- [74] Bartosz Protas, Bernd R. Noack, and Jan Osth. Optimal nonlinear eddy viscosity in galerkin models of turbulent flows. *Journal of Fluid Mechanics*, 766:337367, 2015.
- [75] Lewis F. Richardson. *Weather prediction by numerical process*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2007. With a foreword by Peter Lynch.
- [76] Pierre Sagaut. *Large eddy simulation for incompressible flows: an introduction*. Springer Science & Business Media, 2006.
- [77] Yoshikazu Sasaki. A fundamental study of the numerical prediction based on the variational principle. *Journal of the Meteorological Society of Japan. Ser. II*, 33(6):262–275, 1955.
- [78] Yoshikazu Sasaki. An objective analysis based on the variational method. *Journal of the Meteorological Society of Japan. Ser. II*, 36(3):77–88, 1958.

-
- [79] Yoshikazu Sasaki. Numerical variational analysis formulated under the constraints as determined by longwave equations and a low-pass filter. *Mon. Wea. Rev.*, 98(12):884–898, 1970.
- [80] Yoshikazu Sasaki. Numerical variational analysis with weak constraint and application to surface analysis of severe storm gust. *Mon. Wea. Rev.*, 98(12):899–910, 1970.
- [81] Yoshikazu Sasaki. *Some basic formalisms in numerical variational analysis*. Cite-seer, 1970.
- [82] Philippe Spalart and Steven Allmaras. A one-equation turbulence model for aerodynamic flows. In *30th aerospace sciences meeting and exhibit*, page 439, 1992.
- [83] Bo Tao, Joseph Katz, and Charles Meneveau. Statistical geometry of subgrid-scale stresses determined from holographic particle image velocimetry measurements. *Journal of Fluid Mechanics*, 457:35–78, 2002.
- [84] Michael J. Todd. *Minimum-volume ellipsoids*, volume 23 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2016. Theory and algorithms.
- [85] Michael J Todd and E Alper Yıldırım. On khachiyan’s algorithm for the computation of minimum-volume enclosing ellipsoids. *Discrete Applied Mathematics*, 155(13):1731–1744, 2007.
- [86] Yue Yang and D. I. Pullin. Geometric study of lagrangian and eulerian structures in turbulent channel flow. *Journal of Fluid Mechanics*, 674:6792, 2011.
- [87] Kyo Yoshida, Junzo Yamaguchi, and Yukio Kaneda. Regeneration of small eddies by data assimilation in turbulence. *Phys. Rev. Lett.*, 94:014501, Jan 2005.
- [88] Milija Zupanski and Eugenia Kalnay. Principles of data assimilation. *Cambridge Univ Press, Cambridge*, 18(4):48–54, 1999.