

Euclid preparation

Controlling angular systematics in the *Euclid* spectroscopic galaxy sample

Euclid Collaboration: P. Monaco^{*1,2,3,4,5}, M. Y. Elkhachab^{2,3,1,4}, B. R. Granett⁶, J. Salvalaggio^{2,4,1,3}, E. Sefusatti^{2,4,3}, C. Scarlata⁷, B. Zabelle⁷, M. Bethermin⁸, S. Bruton⁹, C. Carbone¹⁰, S. de la Torre¹¹, S. Dusini¹², A. Eggemeier¹³, L. Guzzo^{14,6,15}, G. Lavaux¹⁶, S. Lee¹⁷, K. Markovic¹⁷, K. S. McCarthy^{17,18}, M. Moresco^{19,20}, F. Passalacqua^{21,12}, W. J. Percival^{22,23,24}, I. Risso^{6,25}, A. G. Sánchez²⁶, D. Scott²⁷, C. Sirignano^{21,12}, Y. Wang²⁸, B. Altieri²⁹, S. Andreon⁶, N. Auricchio²⁰, C. Baccigalupi^{4,2,3,30}, M. Baldi^{31,20,32}, S. Bardelli²⁰, A. Biviano^{2,4}, E. Branchini^{33,25,6}, M. Brescia^{34,35}, J. Brinchmann^{36,37,38}, S. Camera^{39,40,41}, G. Cañas-Herrera^{42,43}, V. Capobianco⁴¹, V. F. Cardone^{44,45}, J. Carretero^{46,47}, S. Casas^{48,49}, F. J. Castander^{50,51}, M. Castellano⁴⁴, G. Castignani²⁰, S. Caviuoti^{35,52}, A. Cimatti⁵³, C. Colodro-Conde⁵⁴, G. Congedo⁵⁵, C. J. Conselice⁵⁶, L. Conversi^{57,29}, Y. Copin⁵⁸, F. Courbin^{59,60,61}, H. M. Courtois⁶², H. Degaudenzi⁶³, G. De Lucia², H. Dole⁶⁴, F. Dubath⁶³, C. A. J. Duncan⁵⁵, X. Dupac²⁹, S. Escoffier⁶⁵, M. Farina⁶⁶, R. Farinelli²⁰, S. Ferriol⁵⁸, N. Fourmanoit⁶⁵, M. Frailis², E. Franceschi²⁰, M. Fumana¹⁰, S. Galeotta², K. George⁶⁷, W. Gillard⁶⁵, B. Gillis⁵⁵, C. Giocoli^{20,32}, J. Gracia-Carpio²⁶, A. Grazian⁶⁸, F. Grupp^{26,69}, S. V. H. Haugan⁷⁰, W. Holmes¹⁷, F. Hormuth⁷¹, A. Hornstrup^{72,73}, K. Jahnke⁷⁴, M. Jhabvala⁷⁵, B. Joachimi⁷⁶, E. Keihänen⁷⁷, S. Kermiche⁶⁵, A. Kiessling¹⁷, B. Kubik⁵⁸, M. Kümmel⁶⁹, M. Kunz⁷⁸, H. Kurki-Suonio^{79,80}, A. M. C. Le Brun⁸¹, S. Ligorì⁴¹, P. B. Lilje⁷⁰, V. Lindholm^{79,80}, I. Lloro⁸², G. Mainetti⁸³, D. Maino^{14,10,15}, E. Maiorano²⁰, O. Mansutti², S. Marcin⁸⁴, O. Marggraf¹³, M. Martinelli^{44,45}, N. Martinet¹¹, F. Marulli^{19,20,32}, R. J. Massey⁸⁵, E. Medinaceli²⁰, S. Mei^{86,87}, Y. Mellier^{88,16}, M. Meneghetti^{20,32}, E. Merlin⁴⁴, G. Meylan⁸⁹, A. Mora⁹⁰, L. Moscardini^{19,20,32}, C. Neissner^{91,47}, S.-M. Niemi⁴², C. Padilla⁹¹, S. Paltani⁶³, F. Pasian², K. Pedersen⁹², V. Pettorino⁴², S. Pires⁹³, G. Polenta⁹⁴, M. Poncet⁹⁵, L. A. Popa⁹⁶, L. Pozzetti²⁰, F. Raison²⁶, A. Renzi^{21,12}, J. Rhodes¹⁷, G. Riccio³⁵, E. Romelli², M. Roncarelli²⁰, E. Rossetti³¹, R. Saglia^{69,26}, Z. Sakr^{97,98,99}, D. Saponi¹⁰⁰, B. Sartoris^{69,2}, P. Schneider¹³, T. Schrabback¹⁰¹, M. Scodeggio¹⁰, A. Secroun⁶⁵, G. Seidel⁷⁴, S. Serrano^{51,102,50}, P. Simon¹³, G. Sirri³², A. Spurio Mancini¹⁰³, L. Stanco¹², J. Steinwagner²⁶, C. Surace¹¹, P. Tallada-Crespi^{46,47}, A. N. Taylor⁵⁵, H. I. Teplitz²⁸, I. Tereno^{104,105}, N. Tessore¹⁰⁶, S. Toft^{107,108}, R. Toledo-Moreo¹⁰⁹, F. Torradeflot^{47,46}, I. Tutusaus^{50,51,98}, L. Valenziano^{20,110}, J. Valiviita^{79,80}, T. Vassallo², A. Veropalumbo^{6,25,33}, D. Vibert¹¹, J. Weller^{69,26}, A. Zacchei^{2,4}, G. Zamorani²⁰, F. M. Zerbi⁶, E. Zucca²⁰, V. Allevalo³⁵, M. Ballardini^{111,112,20}, M. Bolzonella²⁰, A. Boucaud⁸⁶, E. Bozzo⁶³, C. Burigana^{113,110}, R. Cabanac⁹⁸, M. Calabrese^{114,10}, A. Cappi^{115,20}, J. A. Escartin Vigo²⁶, G. Fabbian⁶⁴, L. Gabarra¹¹⁶, W. G. Hartley⁶³, R. Maoli^{117,44}, J. Martín-Fleitas¹¹⁸, S. Matthew⁵⁵, N. Mauri^{53,32}, R. B. Metcalf^{19,20}, A. Pezzotta⁶, M. Pöntinen⁷⁹, V. Scottez^{88,119}, M. Sereno^{20,32}, M. Tenti³², M. Viel^{4,2,30,3,5}, M. Wiesmann⁷⁰, Y. Akrami^{120,121}, I. T. Andika^{122,123}, S. Anselmi^{12,21,124}, M. Archidiacono^{14,15}, F. Atrio-Barandela¹²⁵, S. Avila⁴⁶, D. Bertacca^{21,68,12}, A. Blanchard⁹⁸, L. Blot^{126,81}, M. Bonici^{22,10}, S. Borgani^{1,4,2,3,5}, M. L. Brown⁵⁶, A. Calabro⁴⁴, B. Camacho Quevedo^{4,30,2}, F. Caro⁴⁴, C. S. Carvalho¹⁰⁵, T. Castro^{2,3,4,5}, F. Cogato^{19,20}, S. Conseil⁵⁸, A. R. Cooray¹²⁷, O. Cucciati²⁰, S. Davini²⁵, G. Desprez¹²⁸, A. Díaz-Sánchez¹²⁹, J. J. Diaz⁵⁴, S. Di Domizio^{33,25}, J. M. Diego¹³⁰, A. Enia²⁰, Y. Fang⁶⁹, A. G. Ferrari³², A. Finoguenov⁷⁹, A. Fontana⁴⁴, A. Franco^{131,132,133}, J. García-Bellido¹²⁰, T. Gasparotto⁴⁴, V. Gautard¹³⁴, E. Gaztanaga^{50,51,135}, F. Giacomini³², F. Gianotti²⁰, G. Gozalias^{136,79}, M. Guidi^{31,20}, C. M. Gutierrez¹³⁷, A. Hall⁵⁵, C. Hernández-Monteagudo^{138,54}, H. Hildebrandt¹³⁹, J. Hjorth⁹², S. Joudaki⁴⁶, J. J. E. Kajava^{140,141}, Y. Kang⁶³, V. Kansal^{142,143}, D. Karagiannis^{111,144}, K. Kiiveri⁷⁷, J. Kim¹¹⁶, C. C. Kirkpatrick⁷⁷, S. Kruk²⁹, M. Lattanzi¹¹², V. Le Brun¹¹, L. Legrand^{145,146}, M. Lembo^{16,111,112}, F. Lepori¹⁴⁷, G. Leroy^{148,85}, G. F. Lesci^{19,20}, J. Lesgourgues⁴⁸, L. Leuzzi²⁰, T. I. Liaudat¹⁴⁹, S. J. Liu⁶⁶, A. Loureiro^{150,151}, J. Macias-Perez¹⁵², M. Magliocchetti⁶⁶, F. Mannucci¹⁵³, C. J. A. P. Martins^{154,36}, L. Maurin⁶⁴, M. Miluzio^{29,155}, C. Moretti^{2,4,3,30}, G. Morgante²⁰, S. Nadathur¹³⁵, K. Naidoo^{135,76}, A. Navarro-Alsina¹³, S. Nesseris¹²⁰, D. Paoletti^{20,110}, K. Paterson⁷⁴, L. Patrizii³², A. Pisani⁶⁵, D. Potter¹⁴⁷, S. Quai^{19,20}, M. Radovich⁶⁸, G. Rodighiero^{21,68}, S. Sacquogna¹⁵⁶, M. Sahlén¹⁵⁷, D. B. Sanders¹⁵⁸, E. Sarpa^{30,5,3}, A. Schneider¹⁴⁷, D. Sciotti^{44,45}, E. Sellentin^{159,43}, L. C. Smith¹⁶⁰, K. Tanidis¹¹⁶, C. Tao⁶⁵, G. Testera²⁵, R. Teyssier¹⁶¹, S. Tosi^{33,25,6}, A. Troja^{21,12}, M. Tucci⁶³, A. Venhola¹⁶², D. Vergani²⁰, F. Vernizzi¹⁶³, G. Verza¹⁶⁴, P. Vielzeuf⁶⁵, and N. A. Walton¹⁶⁰

(Affiliations can be found after the references)

ABSTRACT

We present the strategy to identify and mitigate potential sources of angular systematics in the *Euclid* spectroscopic galaxy survey, and we quantify their impact on galaxy clustering measurements and cosmological parameter estimation. We first survey the *Euclid* processing pipeline to identify all evident, potential sources of systematics, and classify them into two broad classes: angular systematics, which modulate the galaxy number density across the sky, and catastrophic redshift errors, which lead to interlopers in the galaxy sample. We then use simulated spectroscopic surveys to test our ability to mitigate angular systematics by constructing a random catalogue that represents the visibility mask of the survey; this is a dense set of intrinsically unclustered objects, subject to the same selection effects as the data catalogue. The construction of this random catalogue relies on a detection model, which gives the probability of reliably measuring the galaxy redshift as a function of the signal-to-noise ratio (S/N) of its emission lines. We demonstrate that, in the ideal case of a perfect knowledge of the visibility mask, the galaxy power spectrum in the presence of systematics is recovered, to within sub-percent accuracy, by convolving a theory power spectrum with a window function obtained from the random catalogue itself. In the case of only approximate knowledge of the visibility mask, we test the stability of power spectrum measurements and cosmological parameter posteriors by using perturbed versions of the random catalogue. We find that significant effects are limited to very large scales, and parameter estimation remains robust, with the most impacting effects being connected to the calibration of the detection model.

Key words. Surveys; Cosmology; observations; large-scale structure of Universe; cosmological parameters

1. Introduction

Euclid (Euclid Collaboration: Mellier et al. 2025) is a space telescope of the European Space Agency designed to survey the Universe out to redshift $z \sim 2$, with the main aim of measuring weak gravitational lensing and galaxy clustering. The goal is to use them to constrain the late-time evolution of the Universe, when dark energy becomes dominant and potential signatures of modified gravity become detectable in the growth of structure. With its visible imager VIS (Euclid Collaboration: Cropper et al. 2025) and near-infrared photometer and spectrometer (NISP, Euclid Collaboration: Jahnke et al. 2025), it will provide images for billions of galaxies and slitless spectra for tens of millions of emission-line galaxies (ELG).

The ongoing Euclid Wide Survey (EWS, Euclid Collaboration: Scaramella et al. 2022) will eventually cover a sky area of 14 000 deg², split into four connected regions separated by the Milky Way (MW) zone of avoidance and by the ecliptic, where zodiacal light hampers deep observations. The EWS is currently tiling the sky in a ‘shift and stare’ mode, with each tile characterised by a reference observing sequence (ROS) consisting of four dithers, each shifted following an S-shaped pattern to minimise the impact of detector gaps (Markovič et al. 2017). Each dither includes the simultaneous parallel observation of VIS photometric (I_E band) and NISP spectroscopic observations, followed by three shorter near-infrared photometric exposures in the J_E , H_E , and Y_E filters. Spectroscopy is performed in slitless mode with red grisms, which are sensitive in the wavelength range $\lambda \in [1.2, 1.9] \mu\text{m}$. To minimise the effect of spectral confusion, a different orientation for the dispersion of light is used in each dither of the ROS, corresponding to 0°, −4°, 184°, and 180° rotations with respect to the focal plane reference orientation. Such a scanning strategy delivers images with an AB magnitude depth of 24.5 (10 σ) and 24 (5 σ) respectively in the I_E and near-infrared bands, and spectra deep enough to measure emission lines with a line flux of $f_{\text{line}} \gtrsim 2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$. Because H α is the most prominent emission line in the spectrum of an ELG in the optical/near-infrared wavelength range, the main sample will consist of H α emitters in the redshift range from 0.85 to 1.88. This redshift range will be rounded to [0.9, 1.8] in the rest of the paper.

A smaller area of 53 deg², split in three separate regions of the sky, is also being observed, accumulating what eventually will be 15 independent wide-like passes. To these 25 further

passes will be added using a blue grism that covers the wavelength range $\lambda \in [0.9, 1.3] \mu\text{m}$. Thanks to its larger depth and increased spectral wavelength coverage, the EDS will allow us to extract a 99% pure and complete sample of the ‘target’ ELGs visible in the EWS. Moreover, the 15 repeated passes assess the probability that a galaxy is detected with a correct or catastrophically wrong redshift, thus quantifying the purity and completeness of the sample (defined in Sect. 2.3).

With the huge volume surveyed by the EWS, the error budget, on both the measured summary statistics and the inferred cosmological parameters, will be dominated by systematics.¹ For measured summary statistics, corresponding to a data vector $\mathbf{D}$, fitted with a model that depends on a set of cosmological and nuisance parameters θ , the posterior distribution on the parameters is evaluated in a Bayesian framework as

$$P(\theta|\mathbf{D}) = \frac{\mathcal{L}(\mathbf{D}|\theta) P(\theta)}{P(\mathbf{D})}, \quad (1)$$

where $\mathcal{L}(\mathbf{D}|\theta)$ is the likelihood of having the data given a set of parameter values, $P(\theta)$ are the priors on parameters, and the evidence $P(\mathbf{D})$ enters as a normalisation. The crucial ingredient of this equation is the likelihood.

We can classify systematics as: (1) ‘data systematics’, namely all those effects that perturb the data vector $\mathbf{D}$; (2) ‘theory systematics’, namely the inaccuracy in the model to be compared with the data; and (3) ‘likelihood systematics’, namely all the issues that affect the likelihood $\mathcal{L}$ and the process of parameter inference.

We concentrate on data systematics of spectroscopic galaxy clustering, and for the data vector $\mathbf{D}$ we use the first three even multipoles of the galaxy power spectrum. In Sect. 2 we present a formalism that describes the main observable, the galaxy density, and show that data systematics in a spectroscopic sample can be broadly divided into angular systematics, which originate from fluctuations of the survey depth on the sky, and those arising from errors in redshift measurements. The latter effect is addressed in Euclid Collaboration: Risso et al. (2025) and Euclid Collaboration: Lee et al. (in prep.), respectively in configuration and Fourier spaces. In Sect. 3 we identify all evident, potential sources of systematics along the *Euclid* data reduction pipeline (details given in Appendix B) and group them into a few classes,

¹ Following the conventional jargon, we use here the word ‘systematics’ as a shorthand for the more correct but longer ‘systematic effects’.

* e-mail: pierluigi.monaco@inaf.it

for which we discuss the adopted mitigation strategy, assesses the impact on the survey (how strongly the systematics affect the clustering measurement), and estimate its risk factor (how difficult it is to mitigate it). After describing in Sect. 4 the tools used to model the effect of angular systematics on the measured power spectrum of simulated (mock) galaxy catalogues, we quantify in Sect. 5 their effect and propagate the uncertainty of its mitigation to cosmological parameters. Finally, Sect. 6 gives the conclusions of the paper. To help the readers, we report in Table 1 some definitions that are used throughout the text.

2. Formalism

2.1. The observed galaxy density

The key observable for extracting cosmological information from a spectroscopic galaxy redshift survey is the galaxy density field $n_g(\mathbf{x})$, where $\mathbf{x}$ is the comoving position in redshift space (assuming here to be an exact redshift measurement). For the moment we do not specify how the galaxy sample is selected. Once the galaxy density has been defined, its density contrast reads

$$\delta_g(\mathbf{x}) := \frac{n_g(\mathbf{x}) - \bar{n}_g(z)}{\bar{n}_g(z)}, \quad (2)$$

where averages are performed over the survey volume and the average density $\bar{n}_g(z)$ is in principle redshift dependent, due to the observational selection function. The density contrast $\delta_g(\mathbf{x})$ can be used to obtain summary statistic such as the galaxy two-point correlation function in configuration space or the galaxy power spectrum in Fourier space. We formalise the effect of systematics by focusing on the galaxy number density, distinguishing between the target galaxy density n_t of an ideal reference sample from the observed one, n_o , which includes systematics. As a consequence, our mitigation strategy will hold for any summary statistics based on δ_g . The formalism presented here follows that of Colavincenzo et al. (2017) and Monaco et al. (2019).

To keep the formalism simple, we call f the flux of the H α line. The spectroscopic catalogue of the EWS will roughly be limited in f , and while the spectra of bright galaxies will be measured with high reliability, most galaxies will be detected in spectroscopy only through a single emission line with relatively low S/N. The spectroscopic sample will faithfully trace the underlying galaxy density field as long as: (i) the flux limit does not vary across the sky; (ii) the catalogue is free of interlopers with a catastrophically wrong redshift. Neither assumption is expected to be correct, so these two effects must be carefully characterised and mitigated.

Consider a volume around a comoving position $\mathbf{x}$ along our past light cone, large enough to contain many galaxies but small enough to neglect evolution effects inside it. Observed galaxy positions consist of a redshift z and a line-of-sight direction $\hat{\mathbf{n}}$, represented by two angles, such as the Galactic coordinates. Assuming a fiducial cosmology, they can be translated into three (redshift-space) comoving coordinates, measured in h^{-1} Mpc; we will use the two coordinate systems (redshift and $\hat{\mathbf{n}}$, or comoving position $\mathbf{x}$) interchangeably. We call $\phi_{\text{local}}(f|\mathbf{x})$ the local H α luminosity function² (LF) of that volume, expressed in terms

² Here we loosely call luminosity function a quantity that is defined in terms of the observed flux; at fixed z flux and luminosity are proportional so ϕ_{local} and Φ can be trivially reformulated in terms of proper luminosity functions.

of f and having units of $h^3 \text{ Mpc}^{-3} (\text{erg s}^{-1} \text{ cm}^{-2})^{-1}$. We define a fiducial limiting flux $f_0 := 2 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$ as the nominal flux limit of the galaxy sample we aim to measure. The *target* galaxy density is then defined as

$$n_t(\mathbf{x}) := \int_{f_0}^{\infty} \phi_{\text{local}}(f|\mathbf{x}) df. \quad (3)$$

We can define a global H α LF $\Phi(f|z)$ as the volume average of ϕ_{local} in a small redshift bin at a median redshift z ,

$$\Phi(f|z) := \langle \phi_{\text{local}}(f|\mathbf{x}) \rangle_z. \quad (4)$$

This global LF is related to the redshift-dependent average galaxy density of the survey as

$$\bar{n}_t(z) = \int_{f_0}^{\infty} \Phi(f|z) df. \quad (5)$$

Here both bars and $\langle \rangle_z$ brackets denote volume averages. To first approximation, the local LF is equal to the global one Φ rescaled by the galaxy density contrast δ_g ,

$$\phi_{\text{local}}(f|\mathbf{x}) \simeq [1 + \delta_g(\mathbf{x})] \Phi(f|z), \quad (6)$$

where δ_g is independent of f_0 . This relation is exact as long as the shape of the local LF is universal, independent of environment (meaning here the value of the local galaxy density on some scale larger than the size of a typical halo). We know that this is not true because the halo bias depends on the halo mass and the galaxy luminosity correlates with the latter, so that brighter galaxies are more clustered than fainter ones. For simplicity, we lay down our formalism by assuming that Eq. (6) is true, so that we can define a generic galaxy density contrast that is independent of the galaxy sample definition. We will see in Sect. 4.1 how to force this behaviour in the mock galaxy catalogue, and use it for testing our mitigation strategies in Sect. 5.1, while Appendix A will quickly discuss how the formalism laid down here changes when this simplification is removed.

Defining an observed sample that is exactly flux-limited would require a reliable measurement of line flux, which is possible only at high S/N, but this may not lead to an optimal sample for cosmology, where number density is crucial. We will then assume that the observed sample contains all galaxies that have a reliable redshift measurement. To formalise the *observed* galaxy density, consider the probability P_{det} that an ELG enters the observed sample with a correct redshift, meaning that its H α line is detected and correctly interpreted. This probability will depend on several parameters; on the one hand, it will depend on the pixel-level properties of the image in the stripe where the galaxy spectrum is dispersed, and more generally on several details of the observation and data reduction process. These details can be represented by a set of nuisance maps $\{N_i(\hat{\mathbf{n}}, z)\}$, which mostly depend on the sky position $\hat{\mathbf{n}}$ and may have a secondary dependence on redshift.

Detection will also depend on the intrinsic properties of the galaxy beyond its H α flux f and redshift z . In fact, at the NISP spectral resolution, H α is blended with the [N II] doublet, so the signal will be given by the whole H α + [N II] complex; then the list of galaxy properties that determine its detection will include the [N II] line luminosity, which is related to galaxy metallicity. Other factors that modulate detection probability are the line

fluxes of all other relevant emission lines, the galaxy magnitudes I_E , Y_E , J_E , and H_E , the galaxy size, morphology, ellipticity, and inclination (see the discussion in Appendix B.10). We define the completeness function C as the galaxy detection probability marginalised over all galaxy intrinsic properties; calling for simplicity $\mathbf{p}$ the m -dimensional vector of these properties, we have that

$$C(f, z, \{N_i\}) := \int d^m p P_{\text{det}}(f, z, \{N_i\} | \mathbf{p}). \quad (7)$$

This completeness function $C(f, z, \{N_i\})$ gives the fraction of galaxies with flux f at redshift z that enter the catalogue with a correct measurement of redshift. As such, it will automatically account for Malmquist bias induced by line flux or by any other observable property. Galaxies included in C that have a correct redshift measurement are called *correct* galaxies, to emphasise their difference from the target galaxies. Within the *Euclid* data reduction pipeline, the detection probability P_{det} is identified as the ‘visibility mask’ of the survey. We consistently identify C as its ‘marginalised visibility mask’ (MVM).

The MVM will drop to zero at $f \ll f_0$, but this cut will not be infinitely sharp, so the catalogue will always contain galaxies that are fainter than the nominal flux limit f_0 . The observed number density of correct galaxies, n_{oc} , will be

$$n_{\text{oc}}(\mathbf{x}) = \int_0^\infty C(f, z, \{N_i\}) \phi_{\text{local}}(f | \mathbf{x}) df \simeq [1 + \delta_g(\mathbf{x})] \int_0^\infty C(f, z, \{N_i\}) \Phi(f | z) df, \quad (8)$$

where the nuisance maps $\{N_i\}$ are evaluated at the sky position $\hat{\mathbf{n}}$ and the second equation is exact if Eq. (6) holds. The average of this density can be computed by assuming that δ_g is statistically independent of C , meaning that there are no systematics that depend on the cosmological signal. One can further assume that $\langle \delta_g \rangle_z = 0$, to obtain

$$\bar{n}_{\text{oc}} = \int_0^\infty \bar{C}(f, z) \Phi(f | z) df, \quad (9)$$

where $\bar{C}(f, z) := \langle C(f, z, \{N_i\}) \rangle_z$.³

2.2. Redshift errors

The density in Eq. (8) is a function of the true galaxy (redshift-space) comoving coordinate $\mathbf{x}$. While we can assume that the uncertainty on the galaxy sky position $\hat{\mathbf{n}}$ is negligible, the galaxy observed redshift z_0 will differ from the true z . For a correct galaxy the $H\alpha$ line is recognised and the difference $\Delta z = z_0 - z$ will be equal to the error in the redshift measurement, which is expected to be distributed with a standard deviation $\sigma_z \simeq 0.001(1 + z)$ – see *Euclid Collaboration: Le Brun et al. (2025)*.⁴ Calling the observed position of the galaxy $\mathbf{x}_0$, the observed density of correct galaxies is obtained by convolving $n_{\text{oc}}(\mathbf{x})$ along

³ As a matter of fact, forcing the density of a sample to be equal to the average density in each redshift bin amounts to forcing a strong radial integral constraint, removing radial modes. We can therefore assume that Eq. (9) holds when averaging over many realisations of a survey.

⁴ This is the original requirement on the redshift error, but the error is in fact independent of redshift; in *Euclid Collaboration: Risso et al. (2025)* we use a $\sigma_z = 0.001$ at all redshifts.

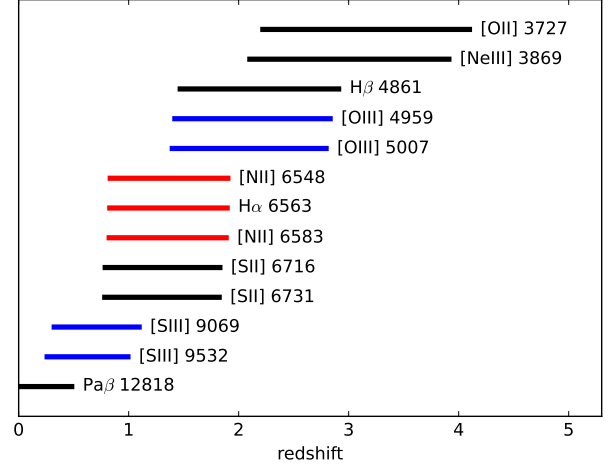


Fig. 1. Main emission lines visible in the galaxy spectra, together with their visibility range for a detector that is sensitive in the wavelength range from 1.2 to 1.9 μm . The red lines correspond to the $H\alpha$ complex and the blue lines to the main line interlopers we expect in the EWS. In *Euclid* spectra the $H\alpha$ + $[N\text{II}]$ complex is not resolved, so the measured ‘ $H\alpha$ flux’ is the integrated flux of the three lines.

the line of sight with the redshift probability distribution function (PDF) $P_z(z_0 | z)$ that represents the uncertainty, or the random error, of redshift measurements,

$$n_{\text{oc}}(\mathbf{x}_0) = n_{\text{oc}}(\mathbf{x}) * P_z(z_0 | z). \quad (10)$$

The process of redshift measurement can give rise to catastrophic errors, where in principle $\Delta z \gg \sigma_z$. The galaxies that enter the spectroscopic catalogue with a catastrophically wrong redshift are called ‘interlopers’. We divide them into two classes. ‘Line interlopers’ are emission-line galaxies for which we detect a line, different from $H\alpha$, that is, however, misinterpreted as $H\alpha$. ‘Noise interlopers’ are sources that do not have emission lines but acquire a false detection due to a noise fluctuation or an instrumental feature that is not properly masked and is mistaken for $H\alpha$. Noise interlopers may even be stars or spurious objects left over by the source extraction algorithm.

Potential line interlopers are all those galaxies that have an emission line. In Fig. 1 we report the main emission lines that might be visible in the NISP spectra, where the $H\alpha$ complex is denoted in red; each segment denotes the redshift range in which the line is visible. While the existence of an emission line opens the possibility of having a line interloper, the likelihood of this to happen depends on many crucial details, and a one-to-one mapping of lines to line interlopers yields only an upper limit to their number.

Assuming an unspecified number of line interlopers, we call $n_{oi}(\mathbf{x}_0)$, with i an integer number, the number density of the i -th line interloper, and n_{on} the number density of noise interlopers. The observed galaxy number density can thus be written as

$$n_o(\mathbf{x}_0) = n_{\text{oc}}(\mathbf{x}) * P_z(z_0 | z) + \sum_i n_{oi}(\mathbf{x}_0) + n_{on}(\mathbf{x}_0). \quad (11)$$

The role of interlopers is analysed in depth in *Euclid Collaboration: Risso et al. (2025)* and *Euclid Collaboration: Lee et al. (in prep.)*; here the number densities of interlopers are directly reported in the observed coordinates $\mathbf{x}_0$, thus leaving the convolution with the redshift PDF implicit.

2.3. Purity and completeness of a galaxy sample

The need for denser spectroscopic samples pushes us to go down in the S/N of the detected spectra. As such, a sample that is optimised for cosmology may still include a significant fraction of interlopers. The quality of the sample can be quantified through the concepts of ‘purity’ and ‘completeness’, following the definitions of Euclid Collaboration: Granett et al. (in prep.).

The spectroscopic sample will consist of all sources for which: (i) a reliable redshift is provided by the data reduction pipeline; and (ii) it lies in the interval $[0.9, 1.8]$. We call $\bar{n}_o(z)$ the average density of such sample at redshift z . We define ‘sample purity’ as the fraction of galaxies in the spectroscopic sample that: (i) are correct galaxies (their $H\alpha$ line is detected and their redshift is correctly measured); and (ii) are target galaxies (their $H\alpha$ line has a true flux $f \geq f_0$). The numerator of this quantity, which we call $\bar{n}_{ot}$, can be formalised as

$$\bar{n}_{ot}(z) = \int_{f_0}^{\infty} \bar{C}(f, z) \Phi(f|z) df. \quad (12)$$

The difference with $\bar{n}_{oc}(z)$ of Eq. (9) lies in the lower limit of the integral: correct galaxies with $f < f_0$ are not target galaxies. Sample purity is then

$$\text{sample purity} := \frac{\bar{n}_{ot}(z)}{\bar{n}_o(z)}. \quad (13)$$

This quantity is a function of redshift, but it is easy to define a sample purity for the whole sample by computing the numerator and the denominator over the whole redshift range of the survey.

The MVM $C(f, z)$ is not expected to be a Heaviside step function in true flux f , since any measure of redshift reliability would correlate with flux with some unavoidable scatter. Due to this fact and to the steepness of the LF $\Phi(f|z)$ at f_0 , a fair fraction of correct galaxies will not be target galaxies and will not contribute to sample purity. Nonetheless, these galaxies trace the same matter density field and are thus useful for cosmology. We thus define ‘redshift purity’ as the fraction of correct galaxies in the spectroscopic sample,

$$\text{redshift purity} := \frac{\bar{n}_{oc}(z)}{\bar{n}_o(z)}. \quad (14)$$

While sample purity is of interest for assessing the performance of the instrument, redshift purity is the quantity that is mostly relevant for cosmology.

We define ‘sample completeness’ as the fraction of target galaxies that are present in the sample as correct galaxies,

$$\text{sample completeness} := \frac{\bar{n}_{ot}(z)}{\bar{n}_t(z)}. \quad (15)$$

As such, this definition ignores galaxies below the line flux threshold f_0 , but has the merit of quantifying the ability of the telescope to detect target galaxies. From Eqs. (5) and (12), it is clear that sample completeness is the average of the completeness function C , weighted by the LF Φ .

Finally, it is convenient to define the ‘redshift efficacy’ of the spectroscopic sample as its ability to include galaxies with correct redshifts compared to the target sample,

$$\text{redshift efficacy} := \frac{\bar{n}_{oc}(z)}{\bar{n}_t(z)}. \quad (16)$$

Contrary to what happens to the other quantities, redshift efficacy is not constrained to be ≤ 1 .

If the sample is selected at lower S/N of galaxy spectra, sample completeness and redshift efficacy increase while redshift purity decreases, and the optimal point in a redshift purity versus sample completeness plane should be decided by maximising the figure of merit of the final results. In Euclid Collaboration: Granett et al. (in prep.) it is shown that, based on simplified end-to-end simulations of the measurement process, an optimal combination is achieved at a redshift purity of 80% and sample completeness of 50%. This implies that 20% of the spectroscopic sample will be interlopers. These simulations also show that the main line interlopers are expected to be detected through their [O III] and [S III] lines (denoted in blue in Fig. 1).

2.4. The observed density contrast

In the analyses presented in this paper we will not consider redshift errors, meaning that our simulations will have no interlopers and will consistently ignore the small difference between the true position of the galaxy $\mathbf{x}$ and the observed one $\mathbf{x}_o$. Then for the rest of the paper

$$n_o(\mathbf{x}_o) = n_o(\mathbf{x}) = n_{oc}(\mathbf{x}). \quad (17)$$

The estimation of the galaxy average density will rely on the generation of a ‘random catalogue’, which will represent the distribution of a dense set of unclustered galaxies subject to known observational systematics, and will be the concrete representation of the visibility mask (see Sect. 3.3). However, the construction of the random catalogue has its uncertainty and the true visibility will be unknown; the random catalogue will be based on an approximated visibility mask, whose marginalisation will give the approximate MVM $\hat{C}(f, z, \{N_i\})$, which represents our best knowledge of the true $C(f, z, \{N_i\})$.

The number density of random galaxies will be

$$n_r(\mathbf{x}) = N_r \int_0^{\infty} \hat{C}(f, z, \{N_i\}) \Phi(f|z) df, \quad (18)$$

where N_r is the replication factor, which is planned to be $N_r = 50$ (see Euclid Collaboration: de la Torre et al. 2025; Euclid Collaboration: Salvalaggio et al., in prep.). This density acquires a dependence on position due to the angular variation of the nuisance maps. We call $\alpha := 1/N_r$ the inverse of the replication factor. The random density n_r , having both an angular and a redshift dependence, is in fact a function of the comoving position $\mathbf{x}$. It is important to notice that the random catalogue will be constructed such that $\alpha \bar{n}_r(z)$, averaged over angular coordinates, reproduces $\bar{n}_o(z)$, not $\bar{n}_t(z)$; the two average densities of observed and target galaxies will in general be different. The observed galaxy density contrast will then be defined as

$$\delta_o(\mathbf{x}) := \frac{n_o(\mathbf{x}) - \alpha n_r(\mathbf{x})}{\alpha n_r(\mathbf{x})}. \quad (19)$$

In Eq. (18) we have assumed an approximate knowledge of the true MVM C , but the integral also contains the true galaxy luminosity function $\Phi(f|z)$. This quantity will be obtained from the EDS, under the assumption that the latter provides a highly pure and complete representation of the galaxies visible in the EWS. We will discuss below how inaccurate knowledge of $\Phi(f|z)$, due to sample variance of the EDS, can affect the construction of the

random catalogue; here for simplicity we absorb this uncertainty into that of C , and assume for the moment that the equality of n_r and $\bar{n}_o$ can be forced by suitably calibrating the detection probability.

When interlopers are present in the catalogue, the random catalogue will also represent interlopers as separate classes of unclustered galaxies subject to the relevant systematics, which are not the same as those of correct galaxies. We will get back to this point in Sect. 3.3, where the construction of the random catalogue is presented in detail.

Under the approximation that Eq. (6) is valid, the observed galaxy density contrast (Eq. 19) is equal to the galaxy density contrast only if we have perfect knowledge of the MVM, $\hat{C} = C$. Following Colavincenzo et al. (2017) we define the angular mask $A(\hat{n}, z)$ as

$$A(\hat{n}, z) := \frac{\int_0^\infty C(f, z, \{N_i\}) \Phi(f|z) df}{\int_0^\infty \hat{C}(f, z, \{N_i\}) \Phi(f|z) df} - 1. \quad (20)$$

Here the notation emphasises that A will mostly depend on angular coordinates, with a possible additional redshift dependence. This angular mask is a quantification of the relative error on the MVM, and it vanishes if this is perfectly known. Being a perturbation of the galaxy density, it propagates to the galaxy density contrast as

$$1 + \delta_o(\mathbf{x}) \simeq [1 + \delta_g(\mathbf{x})] [1 + A(\hat{n}, z)], \quad (21)$$

and this shows that modulations of the flux limit imprint a systematic effect to δ_g which is neither additive nor multiplicative. The angular dependence of A will be determined by how the MVM, and our lack of knowledge of it, is sensitive to the set of nuisance fields $\{N_i\}$. It is important to remark here that this apparently simple relation between δ_o and δ_g is obtained by assuming that galaxy bias is independent of luminosity.

In Appendix A we discuss, both under the assumption of luminosity-independent galaxy bias (Eq. 6) and more generally, how to compute the response of n_o to nuisance maps. This is related in a non-trivial way to the shape of the galaxy luminosity function, strengthening the argument of the convenience of using mock catalogues to characterise the visibility mask.

3. A classification of systematics

3.1. Expected number of sources in a ROS

As a first step, it is useful to quantify the number of sources that are expected to be found in a ROS; the results of this computation are reported in Table 2. Our main interest here is to understand what happens at the level of NISP spectroscopic detection, so we will quantify the sources that fall on the NISP focal plane. Since the dithering pattern covers the detector gaps, we will refer here to the whole focal plane including gaps, amounting to 0.55 deg^2 ; each pixel of $18 \mu\text{m}$ corresponds to an angle of $0''.3$, so the field area translates to 8×10^7 pixels. The number of sources with $H_E < 24$ is obtained from the Flagship galaxy mock catalogue, which is calibrated on a set of observations (see Euclid Collaboration: Castander et al. 2025). An assessment of the number of ELGs with line flux $f \geq f_0$ is provided by the WISP survey (Bagley et al. 2020), which however has a smaller wavelength coverage of $1.25 \mu\text{m} < \lambda < 1.7 \mu\text{m}$. To extend WISP number densities to Euclid's redshift range we use the Flagship catalogue, specifically computing the ratio between galaxy counts

Table 1. Frequently used definitions

Target galaxy sample	all galaxies with $f > f_0$
Observed galaxy sample	all galaxies with reliably measured redshift $z \in [0.9, 1.8]$
Correct galaxy sample	the subset of observed galaxies that have a correct redshift measurement
(True) MVM	marginalised visibility mask, $C(f, z, \{N_i\})$
Approximate MVM	our best knowledge of the MVM, $\hat{C}(f, z, \{N_i\})$
Tabulated detection probability	numerical representation of the MVM
Master data catalogue	50 EuclidLargeMocks representing galaxies with $f > 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$
Master random catalogue	a set of unclustered galaxies with number density $N_r = 50$ times that of 1000 EuclidLargeMocks
Ad-hoc random catalogue	a uniform random catalogue with no angular systematics
Ideal mitigation	obtained using a random catalogue selected with the true MVM
Realistic mitigation	obtained using a random catalogue selected with the perturbed MVM
No mitigation	obtained with ad-hoc random catalogue
Baseline systematics	a combination of image noise, MW extinction and exposure time map

Table 2. Estimates of the number of expected sources in a typical ROS.

sources in NISP	Number
$H_E < 24$ galaxies	100 000
$H_E < 20$ galaxies with continuum	1900
$H\alpha + [\text{N II}]$	2100
$[\text{O III}]$	200
$[\text{S III}]$	80
tot. ELGs	2380
stars	5000–15 000

Notes. All numbers are valid at the order-of-magnitude level. The number of ELGs refers to the main line visible in the grism, so that $[\text{O III}]$ and $[\text{S III}]$ counts refer to galaxies where the $H\alpha$ complex is not observed in the spectrum, while $H\alpha + [\text{N II}]$ galaxies may have other lines visible. All emission lines are counted when $f_{\text{line}} \geq f_0$.

in the WISP and Euclid wavelength ranges and using this ratio to correct WISP counts.

We approximate the photometric catalogue to be limited to the magnitude $H_E < 24$, where we expect to find 1.8×10^5 galaxies deg^{-1} ; this implies 10^5 galaxies per NISP field. In the dispersed images, each galaxy projects a 530×5 pixel stripe on the focal plane, so galaxies will nominally cover 2.7×10^8 pixels, more than three times the number of actual pixels in the focal plane; this means that the focal plane is in principle completely covered by potential source spectra. However, in a typical image the continuum falls below the background for $H_E > 20$ (Euclid Collaboration: Passalacqua et al., in prep.), giving ~ 1900 galaxies per field, whose spectral tracks cover 6% of the pixels. Taking the nominal flux limit f_0 as a threshold for line detection, we expect to have 2100 $H\alpha + [\text{N II}]$ complexes in the field, plus 280 other lines, mainly $[\text{O III}]$ and $[\text{S III}]$. The bulk of these ELGs will have effective radii of about $0''.25$, with a very small fraction going beyond $1''$ (Euclid Collaboration: Passalacqua et al., in prep.). A pixel in the detector plane subtends an angle of

0''.3, while light is dispersed on the detector at $13.54 \text{ \AA pixel}^{-1}$, which translates to a velocity from 210 to 340 $\text{km s}^{-1} \text{ pixel}^{-1}$. In these conditions the flux of the $H\alpha$ complex of a typical ELG is dispersed on an area of 5×5 pixels (Euclid Collaboration: Copin et al. 2025). The lines will then cover a negligible fraction, less than 0.1%, of the NISP pixels. To these numbers we should add stars, which are quantified in Euclid Collaboration: Scaramella et al. (2022) to range from 10 000 to 30 000 deg^{-2} , which translates to 5000 to 15 000 per ROS. Stars will add to the contaminating sources, with tracks covering from 16% to 50% of the pixels, although, as for the galaxies, their continuum will fall below the noise level when their magnitude is $H_E > 20$.

3.2. The SGS pipeline

The data reduction and analysis is performed using a pipeline developed by Euclid's Science Ground segment (SGS). This structure is organised in operational units (OUs), which develop the data processing codes where each step is performed by a processing function (PF).

The analysis and reconstruction of dispersed images is based on detailed input source catalogues provided by the VIS (Euclid Collaboration: McCracken et al. 2025), NIR (Euclid Collaboration: Polenta et al. 2025), and MER PFs (Euclid Collaboration: Romelli et al. 2025). VIS and NIR perform the extraction of photometric sources on single and stacked images, while the merging of their products is computed by the MER PF, which detects objects in all the Euclid imaging data, measures their properties, and prepares a single multi-wavelength catalogue that includes sources that are identified in either visible or NIR images. One crucial point in MER processing is image deblending, which is accomplished by two density-based clustering algorithms tuned on pixel-level simulations. The MER catalogue of photometric sources is expected to have a depth in the H_E -band of approximately 24.

The SIR PF (Euclid Collaboration: Copin et al. 2025) processes the NISP dispersed images and produces fully calibrated one-dimensional spectra for (potentially all) sources in the MER catalogue. SIR performs calibration in wavelength and flux, corrects instrumental effects, subtracts cross-contaminations of different spectra, and finally produces one-dimensional spectra for scientific use. A crucial step of SIR is the decontamination of overlapping spectra. As shown above (Sect. 3.1), galaxies with $H_E < 24$ completely fill the detector plane in the dispersed images; on average, one can expect that the 1st-order spectral trace of a source of interest overlaps with 10 to 30 other spectral traces. Contamination can also be induced by the zeroth-order image of relatively bright sources and, to a lower extent, by other orders. Decontamination is performed starting from NIR photometry, which gives a robust estimate of the flux to be subtracted out. Modelling of the dispersed spectrum makes use of other photometric information from the MER catalogue, like source brightness profile, ellipticity, and inclination. NISP detectors are affected by the so-called 'persistence' of the signal (see Appendix B.5). This can be masked out or subtracted, making use of the information from previous images in the observing sequence, building a persistence model that estimates the level of contamination of each pixel.

Starting from one-dimensional spectra produced by SIR, the SPE PF (Euclid Collaboration: Le Brun et al. 2025) performs redshift measurement, reliability estimation, and spectroscopic classification of all the extracted sources. SPE applies a modified version of the algorithm for massive automated z evaluation and determination (AMAZED), which classifies the spectra into three

categories: galaxies, stars, and quasars, with an associated probability of belonging to a specific class. Then, for each spectral class, the five best redshift solutions are computed, each complemented with an estimate of the measurement reliability. To maximise sample completeness, SPE applies a prior that favours solutions containing the $H\alpha$ line. The final SPE output is a catalogue that, for each source, contains a list of redshift solutions, and for each redshift solution it provides line flux, S/N, central wavelengths, and full width at half maximum for each line.

The resulting catalogue is analysed by OU-LE3, responsible for the galaxy clustering measurements. The first step of this processing is the selection of a sample and the construction of the random catalogue, performed respectively by SEL and VMSP. In particular, VMSP creates a master random catalogue, where each galaxy is assigned full galaxy properties (from the EDS) and a detection probability as described below in Sect. 3.3. The selection of the master data and random catalogues is performed by the SEL PF, following criteria that will in general be aimed at maximising the figure of merit of cosmological parameters. Besides selecting the data and random catalogues, SEL estimates the (sample) purity and completeness of the data catalogue in the areas covered by the EDS.

The selected data and random catalogues are then fed to the other LE3 PFs, named 2PCF (Euclid Collaboration: de la Torre et al. 2025) and 3PCF (Euclid Collaboration: Veropalumbo et al., in prep.) for the two-point and three-point correlation functions in configuration space, PK (Euclid Collaboration: Salvalaggio et al., in prep.) and BK (Euclid Collaboration: Rizzo et al., in prep.) for the corresponding power spectrum and bispectrum in Fourier space. The PK code is described below in Sect. 4.4. At the end of LE3 pipeline, the covariance matrix PFs, CM-2PCF and CM-PK, take care of computing the covariance of two-point estimators in configuration and Fourier space (see Sect. 3.4).

A detailed discussion of all the potential *entry points* of systematics along the pipeline, from photometry to redshift measurement, is given in Appendix B. We focus in the main text on the last part of the pipeline, including a discussion on the random catalogue and on the computation of covariance matrices.

3.3. The random catalogue as a tool to mitigate angular systematics

A large part of the observational systematics will be mitigated by suitably constructing a random catalogue that represents the visibility mask P_{det} (introduced in Sect. 2.4). This will be done by the VMSP PF, relying on a bona fide sample of galaxies drawn from the EDS that represents the galaxies for which a reliable redshift measurement is possible. It is required that the sub-sample of target galaxies (Eq. 3) is recovered with a very high purity greater than 99%.

VMSP picks up galaxies from this sample and places them in random positions on the sky, then evaluates the probability P_{det} that the sources are detected at the EWS depth. In principle this step could be done by injecting the source spectra in the four NISP dispersed dithers, then processing the ROS with SIR and SPE; this complete process will be performed for limited areas, but it would be too costly to populate the whole EWS this way. Using VMSP, we resort to a bypass simulation algorithm, called *pypelid* and described in Euclid Collaboration: Granett et al. (in prep.). For each random galaxy, VMSP reads the noise at the location of the main emission lines in the NISP dithers. It then simulates the emission lines to obtain their S/N. The mapping from S/N (called S in the equations) to the detection probability,

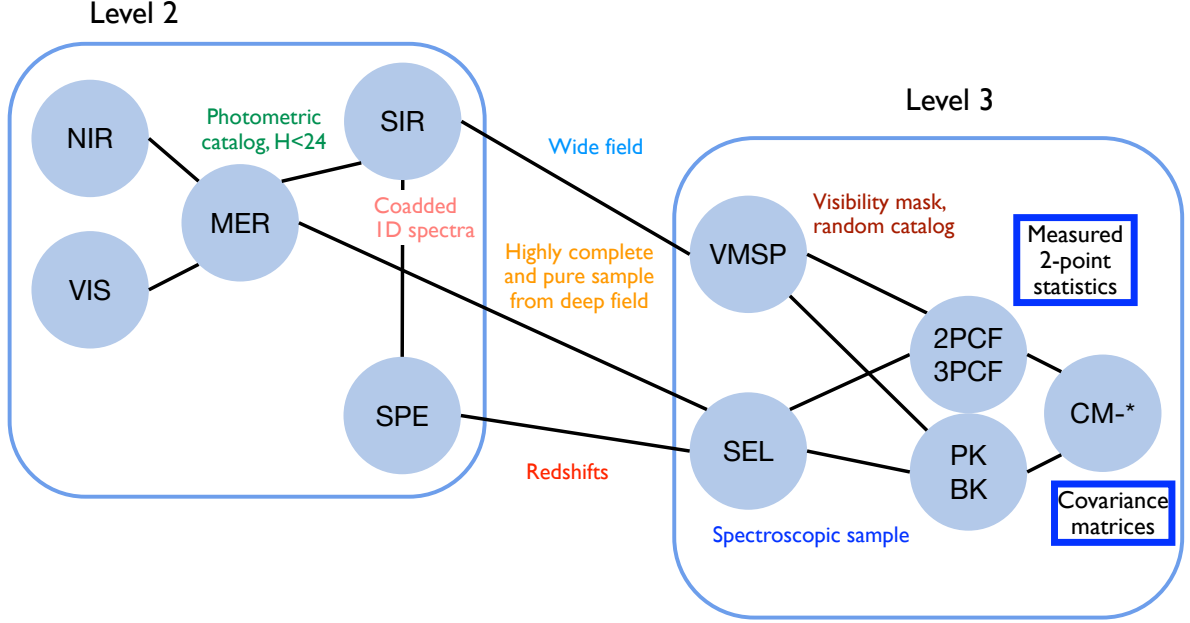


Fig. 2. Schematic view of the SGS pipeline for galaxy clustering. The two boxes mark level-2 and level-3 processing, with the circles corresponding to OUs in the left box and to OU-LE3 PFs in the right box. Black segments outline the dependences of the various blocks, some of which are commented by coloured text. The text highlighted by blue boxes gives the final products of the galaxy clustering pipeline.

$P_{\text{det}}(S)$, is then given by a ‘detection model’, that is assumed to be a sigmoid function,

$$P_{\text{det}}(S) = \frac{c}{1 + \left(\frac{S}{S_0}\right)^{-\beta}}, \quad (22)$$

where c , S_0 , and β are free parameters to be calibrated; in particular, the function crosses $c/2$ at S_0 with a slope determined by β , and then levels off to c .

Using the 15 wide-like passes of the EDS, it will be possible to estimate for these galaxies both the S/N of their emission lines and their detection probability, thus creating a calibration set for the detection model $P_{\text{det}}(S)$. In Euclid Collaboration: Granett et al. (in prep.) we use a calibration set obtained from spectroscopic simulations processed by SPE; the detection probability thus relies on a choice of the threshold for spectrum reliability used in sample selection. The P_{det} of this set is shown in the left panel of Fig. 3, and is fit by a sigmoid (Eq. 22) with parameter values $S_0 = 3.69$, $\beta = 4.98$ (the red line), while c is found consistent with unity, thus reducing the parameters to two. For later use, we adopt a simplified approach to assign a 1σ error to the parameters: we count the calibration points above and below the curve, then vary one parameter at a time and mark when the smallest of these numbers gets as low as 16% of the whole set.⁵ This gives $S_0 = 3.69 \pm 0.42$ and $\beta = 4.98 \pm 1.25$. Given that the calibration set is small and the parameters may be correlated, we use this as a pessimistic estimation of the error on the calibration of the detection model. Figure 3 shows four sigmoid curves obtained by varying one of the parameters by 1σ . We also compute the scatter around this relation; the right panel of Fig. 3 reports the standard deviation of P_{det} from the calibration set divided by its best-fit value; this scatter is reproduced by a function $\sigma_P = 2.4 S^{-2.1}$.

⁵ To find the error in β , we count what is above the model for $S < S_0$ plus what is below it at higher S

This is capped at the value 0.4 to avoid divergence at very low S , which have little impact since detection probabilities are very low in that range.

It is clear that such a detection model will be valid as long as the relation between S and P_{det} is found to be tight in the calibration set, and the residuals are not correlated to any other galaxy property; in case any secondary dependence of detection probability is found, the algorithm will be revised to take them into account. It is also worth noticing that the algorithm deals with the uncertainties in the dispersed images in a realistic way, but assumes a perfect photometric measurement.

Adding line interlopers will be performed using a very similar approach: starting from all EDS ELGs that have at least one emission line, the analysis of the 15 wide-like passes will determine the probability that these galaxies are misclassified as $H\alpha$ ELGs, then `pypelid` will compute the S/N of these spectra and a suitable (mis-)detection model will be calibrated on these results and used to create a set of random galaxies that represent line interlopers. For noise interlopers a different strategy will be adopted, with two possible options: one is to add a completely unclustered set of points to the random catalogue, another is to add a fraction of the photometric sample, equal to N_r times the small probability that a photometric source is mistaken as a noise interloper. As a result, random galaxies will be labelled according to the kind of galaxies they are supposed to represent, be it target, i -th line, or noise interloper.

Systematics coming from the EDS are addressed in Euclid Collaboration: Bruton et al. (in prep.); we discuss here some relevant points. For a $\sim 50 \text{ deg}^2$ collective area of the EDS and a replication factor $N_r = 50$, it is easy to see that each EDS galaxy will be replicated about 14 000 times in the random catalogue. In particular, the $H\alpha$ LF $\Phi(f|z)$ sampled by EDS galaxies will be subject to significant fluctuations in redshift, due to sample variance in the limited area surveyed. If propagated to the number density of the random catalogue for the much larger EWS, these

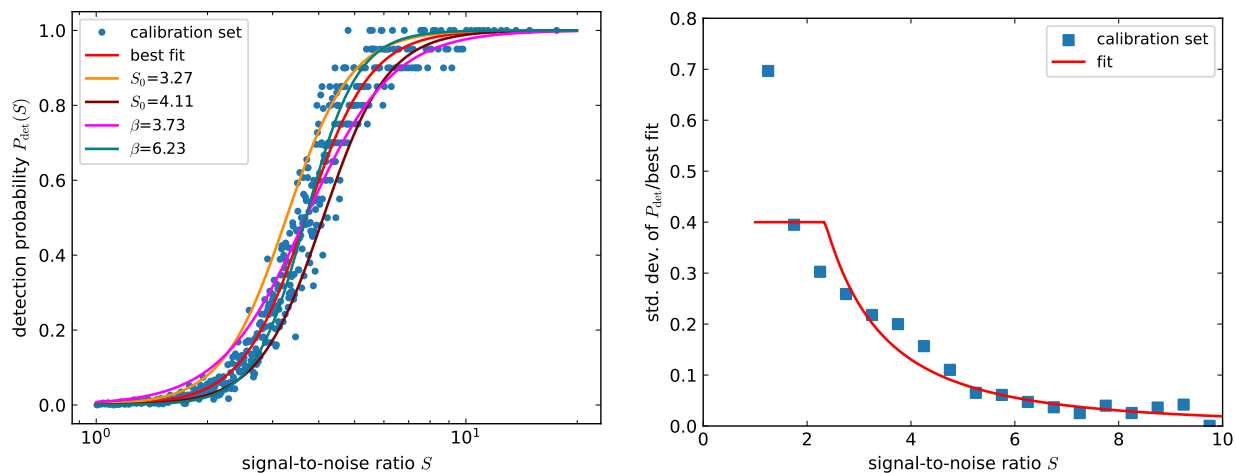


Fig. 3. *Left:* detection probability $P_{\text{det}}(S)$, that is the probability of detecting a galaxy in the spectroscopic sample as a function of the S/N of its $H\alpha$ line. The points give the measurements from the calibration set used to fix the parameters of the detection model. The five curves correspond to the best-fit detection model, together with those obtained by varying the two fitting parameters by 1σ (see the legend for the assumed parameter values). *Right:* the standard deviation of the ratio of the measurements from the calibration set and the best fit model.

fluctuations would create spurious radial modulations of the density contrast that, due to their anisotropy, would bias the higher multipoles of two-point correlations. To mitigate this effect, we will fit the number density of EDS galaxies with a smooth function, thus reducing the impact of sample variance on $\bar{n}_r(z)$ to the 1–2% level.

The addition of interlopers will imply a further uncertainty, driven again by sample variance; because the interlopers come from a different region of the comoving space, the sample variances of target galaxies and the various classes of line interlopers do not cancel out in the interloper fractions.

If photometry is used to select the spectroscopic sample (for instance to limit interlopers), the uncertainty in the photometry must be propagated to the random catalogue, which in its standard version assumes exact magnitudes. While some effects of over- or under-deblending, if significant, may require specific strategies, propagating the error in photometry may simply require us to perturb the magnitudes of deep field galaxies when injecting them into the master random catalogue; this alone would reproduce any Malmquist-like bias due to selection based on photometry.

As a matter of fact, no information from the EWS is used to construct the random catalogue, so there is no guarantee that $\bar{n}_r(z) = N_r \bar{n}_o(z)$ at all redshifts. A failure to achieve this equality would be a clear sign of some problems in the pipeline and would prompt a reassessment of the systematics budget and of the construction of the random catalogue (see Sect. 5.1.2). However, this equality can be reached only up to the uncertainty due to EDS sample variance; once this level of agreement is reached, it is legitimate to rescale the number of random galaxies so that the random number density $n_r(z)$ is exactly equal to $N_r \bar{n}_o(z)$. Any effect that perturbs the average number density by less than the sample variance of the EDS cannot be recognised with this test.

3.4. Covariances

The calculation of covariance matrices is based on a brute-force computation of 2PCF and PK of a set of thousands of mock catalogues (Euclid Collaboration: Monaco et al. 2025). The CM-2PCF and CM-PK PFs then compute the sample average and covariance of the corresponding measurements; we will call

this the ‘numerical covariance’ of the corresponding estimator. While trivial from the technical point of view, these PFs will perform a crucial step in the analysis that involves codes that are outside the *Euclid* SGS pipeline.

Given a large set of mock galaxy catalogues, their numerical covariance allows us to take into account sample variance due to the finite volume of the survey and to account for mode coupling induced by non-linearities and by the window function. At the same time, the random catalogue will mitigate the systematics induced by fluctuations of survey depth, and in part by redshift errors. Due to the complicated (both additive and multiplicative) relation that connects the galaxy density contrast with the observed one (Eq. 21, which is a simplification anyway), the contributions of sample variance (coming through δ_g) and uncertainty on the visibility mask (coming through the mask A) are intertwined in a complicated way, as shown for instance by Colavincenzo et al. (2017). This makes a numerical approach to the problem unavoidable.

The uncertainty in this process may be obtained by perturbing each step in the construction of the random catalogue and producing a set of thousands of random catalogues, one for each mock, which represent the expected uncertainty in the mitigation process. However, this is an inefficient way to tackle the problem; a much more convenient option is to perturb the visibility mask applied to the galaxy mocks, while using a single random catalogue for all clustering measurements, which represents the average visibility mask (in principle the same random catalogue used for the observed data catalogue). It is important to point out that imposing angular systematics to a galaxy mock catalogue is a very similar operation to constructing the random catalogue, the main difference being that galaxy positions are not randomly chosen. But each galaxy in the data or random catalogue is described by several properties beyond sky coordinates and redshift, indicated by the $\mathbf{p}$ vector of Eq. (7). We will describe in Sect. 4.3 how it is possible to exploit the VMSP/pypeid algorithm to build an MVM, which depends only on $H\alpha$ flux and redshift, thus enabling us to impose angular systematics to thousands of catalogues that contain minimal information for each galaxy.

Table 3. Table of expected systematics along the pipeline.

Systematic effect	Entry point	Class	Consequences	Mitigation strategy	Impact	Risk
Fluctuation of photometric depth + photometric selection	MER +SEL	angular	loss of sources for SIR modulation of catalogue	random +photo errors	low	low mild
Under-deblending	MER	angular	loss of sources for SIR mis-centred sources	none	v. low	v. low
Over-deblending	MER	noise int.	fake sources	random	low	low
Persistence: spectro to photo	MER	noise int.	fake sources	random	mild	low
Zodiacal light	astro	angular	increased background	random	v. high	v. low
Nearby galaxies	astro	angular	increased background and confusion	random	low	low
MW extinction	astro	angular	decreased incoming flux	random (incomplete)	high	v. high
MW emission	astro	angular	contribution to background	random	v. low	v. low
Flagged pixels (bad pixels, saturated stars, cosmic rays, persistence)	SIR	angular	degraded image quality	random	v. high	low
Stray light	SIR	angular	increased noise	random	high	low
Persistence: photo to spectro	SIR	noise int.	fake emission lines	random	mild	high
Persistence: spectro to photo, photo to photo spectro to spectro	SIR	angular + noise int.	biased decontamination	none	mild	high
Calibration of SIR noise	SIR	angular	modulation of random	none	low	mild
Wavelength calibration	SIR	AP-like	redshift bias	none	v. low	v. low
Confusion from foregrounds	SIR	angular + noise int.	increased noise in spectra potential interlopers	random	high	low
Confusion from $z > 0.9$ sources	SIR	angular	correlated to signal	no	low	v. high
Coaddition of dithers	SIR	angular noise int.	error propagation	random	mild	low
Templates for spectral fit	SPE	AP-like	bias in redshift measure	no	low	low
Star-galaxy separation	SPE	noise int.	spurious sources	random	mild	low
Redshift uncertainty	SPE	theory	anisotropic smoothing	model	mild	low
Line misidentification	SPE	line int.	wrong redshift	model	high	high
Spectral reliability selection	SPE +SEL	noise int. + line int.	sub-optimal cleaning of interlopers	random	mild	low
Photometric selection	SPE +SEL	noise int. + line int.	introducing photometric systematics	random +photo errors	high	high
Galaxy properties	astro	—	marginalised over by model	no	none	none
Intrinsic alignments	astro	theory	detection correlated with tidal field	no	v. low	mild
Sample variance in EDS	VMSP	radial	fluctuations in $\bar{n}(z)$	random +mocks	mild	low
Sample variance in galaxy properties from EDS	VMSP	angular	uncertainty in detection probability	random +mocks	low	low
Mapping of S/N to detection	VMSP	angular	uncertainty in bypass	random +mocks	low	low

Notes. The columns report a very brief description of the systematic, its entry point, its classification (see the text), its consequences, the mitigation strategy, its impact on the overall survey, and the risk connected to its mitigation.

3.5. Classification and ranking

The results of this discussion are reported in Table 3, where each entry corresponds to each of the systematics mentioned in this section and in Appendix B. Each row of this table gives a very brief description of the effect and its entry point, followed by its classification, a description of the consequences of the effect, the implemented mitigation strategy, the impact of this effect on the survey (i.e., how much it affects its quality), and the risk associated to its mitigation (i.e., how difficult it to achieve an accurate mitigation).

This long list of systematics results in a classification, given in the third column of the table: (i) ‘angular’, an effect that mod-

ulates the effective flux limit of the spectroscopic catalogue on the sky; (ii) ‘noise interlopers’, an effect that adds spurious objects to the catalogue that have a measured redshift unrelated to true redshift; (iii) ‘line interlopers’, where a line is detected but misidentified; (iv) ‘AP-like’, leading to a distortion (similar to the Alcock–Paczynski effect) in the mapping between measured redshift and line-of-sight distance; (v) ‘radial’, leading to a contamination of the galaxy density limited to the radial coordinate; and (vi) ‘theory’ systematics, which should be addressed by adding sophistication to the theory model.

The list of mitigation strategies is even more limited. Mitigation will mostly rely on the random catalogue or on the theory

model, with the support of a set of mock catalogues that are accurately calibrated on the data.

3.6. Comparison with other surveys

The treatment of systematics in ground-based spectroscopic galaxy surveys such as the Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al. 2013), the extended Baryon Oscillation Spectroscopic Survey (eBOSS, Dawson et al. 2016), and the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration: Aghamousa et al. 2016) differs from the approach outlined here for *Euclid*. These surveys define their samples by applying selection criteria to a pre-existing photometric parent catalogue, from which targets are selected for spectroscopic follow up. This allows the assessment of sample completeness relative to a well-defined parent catalogue. Moreover, the quality of such ground-based spectra makes catastrophic redshift errors much less likely to happen. Conversely, potential angular systematics affecting clustering statistics can be introduced in the target catalogue (e.g., due to seeing variations in the photometric data or in the calibration of different photometric systems) or during the subsequent spectroscopic observations (e.g., the inability to place spectroscopic fibres on nearby targets, commonly referred to as fibre collisions).

BOSS identified foreground stellar density as the primary source of systematics, with additional contributions from variations in observing conditions such as seeing, sky background, airmass, and MW extinction. These effects were mitigated using regression-based corrections in the form of a systematics weighting scheme (Ross et al. 2012) applied to the galaxy sample. Additionally, specific weights were applied to account for fibre collisions and redshift failures, ensuring a uniform sampling of the underlying large-scale structure (Reid et al. 2016). eBOSS extended these methodologies while introducing refinements to handle new galaxy tracers, including luminous red galaxies (LRGs), ELGs, and quasars (Ross et al. 2020). Additionally, improvements were made in fibre assignment corrections using the nearest-neighbour approach to mitigate fibre collision effects (Bautista et al. 2021; De Mattia et al. 2021).

The approach followed by DESI builds upon the methodologies developed for BOSS and eBOSS, while incorporating more advanced techniques for systematics mitigation (Adame et al. 2025; Krolewski et al. 2025; Rosado-Marín et al. 2025). One of the key innovations in DESI is the use of machine-learning techniques to model and correct for imaging systematics (Rezaie et al. 2021). Additionally, DESI employs improved fibre assignment algorithms to minimise selection biases and maximise the completeness of spectroscopic observations (Bianchi et al. 2025). In all cases, mock catalogues have played a crucial role in assessing and validating these corrections (e.g. Kitaura et al. 2016).

The different observational approaches lead to a difference in the mitigation strategies, with ground-based surveys mostly relying on a-posteriori regression of the density maps with known nuisance maps, as opposed to the forward modelling of the visibility mask performed on the dispersed images. More specifically, here we decide to weigh the random galaxies, which is another way of expressing the construction of the visibility mask, rather than weighting the data catalogue. Nonetheless, the methodologies developed in ground-based surveys provide valuable insights for *Euclid* systematic mitigation strategies, particularly in the treatment of angular systematics affecting the observed galaxy number densities.

4. Modelling angular systematics in power spectrum measurements

In this section we present the methods that we use to model the impact of angular systematics on the measured power spectrum of mock galaxy catalogues, and on the inferred cosmological parameters. We introduce the mock galaxy catalogues, the nuisance maps, the tabulated detection probability, and the codes to measure the power spectrum and perform parameter inference from it.

4.1. Mock galaxy catalogues

Several sets of thousands of mock galaxy catalogues were produced to support the study of systematics in *Euclid*'s spectroscopic galaxy clustering; these are described in detail in Euclid Collaboration: Monaco et al. (2025). We will limit our description here to the EuclidLargeMocks set, which is presented in that paper. This has been produced using the *pinocchio* fast code (Monaco et al. 2002, 2013; Munari et al. 2017); this is based on excursion set theory, ellipsoidal collapse, and Lagrangian perturbation theory (LPT), and can be seen as a halo finder in Lagrangian space, where haloes are moved to their final position using a single 3LPT displacement. The configuration is that of a cubic box of $3380 h^{-1} \text{Mpc}$ sampled by 6144^3 particles, for a total volume of $38.61 h^{-3} \text{Gpc}^3$ per simulation. This allows us to reach a particle mass of $1.48 \times 10^{10} M_{\odot}$; being *pinocchio* a semi-analytic code, we can push the mass threshold down to 10 particles, thus resolving haloes with $M_h \geq 1.48 \times 10^{11} M_{\odot}$. The total volume sampled by these 1000 realisation is thus $38614 h^{-3} \text{Gpc}^3$, 11 times the volume of the visible Universe.

Haloes were populated with galaxies using an halo occupation distribution (HOD) model extracted from the Flagship galaxy mock catalogue (Euclid Collaboration: Castander et al. 2025), that is calibrated on the $H\alpha$ luminosity function given by model 3 of Pozzetti et al. (2016). As a result of this process, the average number density and average galaxy power spectrum of the 1000 mocks were found to be consistent with the number density and galaxy power spectrum of the Flagship galaxies, and this is true for different redshift bins and flux limits.

The mocks were created using an idealised footprint of a circle of radius 30° , covering 2763 deg^2 on the sky; this area is just larger than the 2500 deg^2 originally planned for *Euclid* Data Release 1 (DR1), and can (almost) contain its north and south islands, as described in Euclid Collaboration: Scaramella et al. (2022). The volume subtended by a 30° circle and lying at redshift from 0.9 to 1.8 can be fully immersed in the parent simulation box of $3380 h^{-1} \text{Mpc}$, so our analysis is free from problems connected with repetitions of the simulated box (see details in Euclid Collaboration: Monaco et al. 2025). This footprint was used as a reference in the paper describing and testing the PK code (Euclid Collaboration: Salvalaggio et al., in prep.).

It is useful, for specific tests, to work with catalogues where the effect of luminosity-dependent bias is removed. To achieve this, we binned galaxies in redshift, with a small bin width of $\Delta z = 0.01$, and for each bin we shuffled all the line fluxes f among galaxies belonging to that bin. As a consequence, a selection in flux leads to a sparse sampling of galaxies, preserving their clustering amplitude, while the redshift-dependent luminosity function is preserved (within the relatively fine redshift binning). Each mock catalogues contain both standard and shuffled line fluxes.

The numerical covariance matrices used in this paper are obtained by measuring the power spectra of all the 1000 galaxy

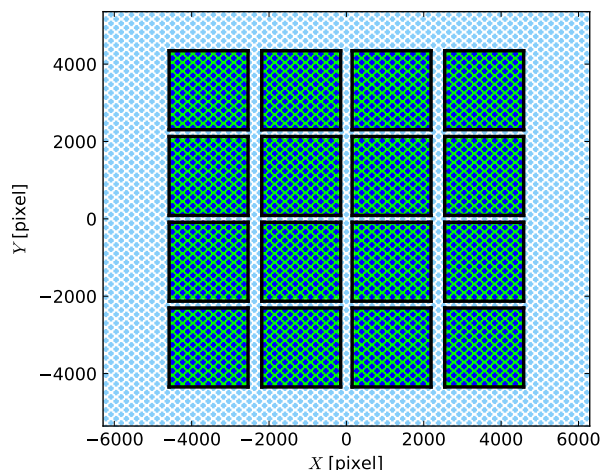


Fig. 4. Adopted model for NISP detector geometry, the smaller squares denote the 16 detectors, crosses are the centres of healpix sky pixels (with $N_{\text{side}} = 4096$) that are inside (blue) or outside (cyan) the detectors.

catalogues, in a few relevant configurations. However, in this paper we are more concerned with the average value of the power spectrum, so we will routinely analyse a smaller number of realisations. To set the number, we define the effective volume of a survey (Tegmark 1997) that covers a sky area with a solid angle Ω from z_1 to z_2 as

$$V_{\text{eff}} = \frac{\Omega}{3} \left[d_c^3(z_2) - d_c^3(z_1) \right] \left[\frac{\bar{n} P(k_{\text{BAO}})}{\bar{n} P(k_{\text{BAO}}) + 1} \right]^2, \quad (23)$$

where $\bar{n}$ is the galaxy density in that redshift bin and the power spectrum value is computed at the baryonic acoustic oscillations (BAO) scale, $k_{\text{BAO}} \approx 0.07 h \text{ Mpc}^{-1}$. We divide the survey into four redshift bins, bounded by the values 0.9, 1.1, 1.3, 1.5, and 1.8. For a survey area of 2763 deg^2 , the effective volume of the four redshift bins is, respectively, 0.99, 1.24, 1.34, and $2.10 h^{-3} \text{ Gpc}^3$. We compute the effective volume of the survey as the sum of these four values, amounting to $5.67 h^{-3} \text{ Gpc}^3$. The total effective volume of the EWS is obtained by rescaling these figures to 14000 deg^2 , and amounts to $28.7 h^{-3} \text{ Gpc}^3$. We will mostly work with the second redshift bin, $z \in [1.1, 1.3]$, and we decide to use 50 realisations, amounting to ~ 2.3 times the EWS effective volume. Cosmological inference will be obtained by combining the four redshift bins. In this case, five mocks cover $2763 \times 5 = 13815 \text{ deg}^2$, just 1.3% lower than the 14000 deg^2 expected final area of EWS, so the 50 mocks will be split into ten groups of five mocks each, thus covering the EWS survey volume ten times.

4.2. Nuisance maps: exposure time, image noise, MW extinction

In our path towards complexity, we decided to immerse the 30° circle footprint in the largest island of the planned final survey, thus creating a footprint with realistic holes and dithering pattern but a regular overall shape similar to the one used in the PK paper.

To produce a realistic survey footprint, we created a model for the NISP detector (Fig. 4) by tiling 16 squares of 2048^2 pixels, exactly aligned on a grid, where each pixel corresponds to $0''.299$ (see Euclid Collaboration: Scaramella et al. 2022, for a

very similar model). Gaps between detectors are identical in the y direction (amounting to 168.63 pixels), while the central gap in the x direction is larger than the other two (346.47 pixels versus 288.45 pixels). Moreover, the first eight external pixels of the detectors were assumed to be not usable. This model was placed on the sky at $(0, 0)$ galactic angular coordinates, using a flat-sky approximation, and sampled with a healpix grid⁶ of size N_{side} ranging from 512 to 8192. We stored the list of sky pixels whose centre falls inside the detectors; Fig. 4 reports the healpix grid for $N_{\text{side}} = 4096$. We checked that with $N_{\text{side}} < 2048$ the detector gaps are not properly represented.

The mapping from sky coordinates to position on the detector is well defined for a photometric pointing, but the grism dispersion distorts the photon paths, spreading a point source into a stripe of ~ 530 pixels. If a galaxy is detected through its $H\alpha$ line, the position of this line is shifted along the dispersion direction by a number of pixels equal to $(\lambda - \lambda_0)/\lambda_{\text{pix}}$, where λ_0 is the wavelength whose path is not bent by the grism; we assume $\lambda_0 = 1.52 \mu\text{m}$ (so that for a galaxy at $z = 1.31$ the $H\alpha$ line falls at the location of the photometric source). Moreover, $\lambda_{\text{pix}} = 13.54 \text{ \AA}$ is the wavelength dispersion on a pixel (as in Sect. 3.1). We create the exposure time maps by computing how many times the $H\alpha$ line of a source is observed, so these maps are redshift-dependent.

The exposure time map for the EWS was created by going through the survey timeline, as planned before launch and presented in Euclid Collaboration: Scaramella et al. (2022), which lists for each planned pointing the position of its centre and its orientation on the sky. For each pointing, we rotated the sky pixels that sample the NISP detector to the pointing position, shifted them to compute the position of the $H\alpha$ line, and sampled them with another healpix tessellation at half N_{side} , flagging each sky pixel that contains at least one NISP pixel; this oversampling was done to avoid that the rotation creates spurious holes in the footprint. The survey footprint was then created as a boolean healpix map that is true in any sky pixel that is visited at least once, false otherwise. The exposure time map was created by counting in how many pointings a sky pixel is observed; this produces an integer map, but rotation or resampling can make it a floating point map, giving the effective number of exposures in a sky pixel. The left panel of Fig. 5 gives the resulting exposure time map in the 30° circle, at a map resolution of $N_{\text{side}} = 2048$ (obtained using a NISP resolution of $N_{\text{side}} = 4096$). For this map, we simply used the photometric position of the galaxy. Although immersed in a simple overall geometry, this map shows a complicated pattern with a wheel-like geometry around the north ecliptic pole, where overlaps among pointings are stronger. Moreover, several holes are present to avoid bright objects.

The middle panel of Fig. 5 gives MW reddening $E(B - V)$ in the same region of the sky. Reddening is discussed in depth in Appendix B.3. Here we use the reddening map produced by the Planck Collaboration in their 2013 results (Planck Collaboration: Abergel et al. 2014, hereafter P13). They provide the map at $N_{\text{side}} = 2048$, slightly oversampling the observational beam.

The right panel of Fig. 5 gives a noise map, again at $N_{\text{side}} = 2048$. The noise map was created using a model for the zodiacal light and a model for the stray light, produced within the Eu-

⁶ Hierarchical, Equal Area, and iso-Latitude Pixelation (healpix) (Górski et al. 2005) is a code, widely used in cosmology, that performs a tessellation of the sphere into $N_{\text{pix}} = 12 N_{\text{side}}^2$ sky pixels. Here we reserve the word ‘pixel’ for the detector pixels, and call ‘sky pixels’ those related to an healpix tessellation.

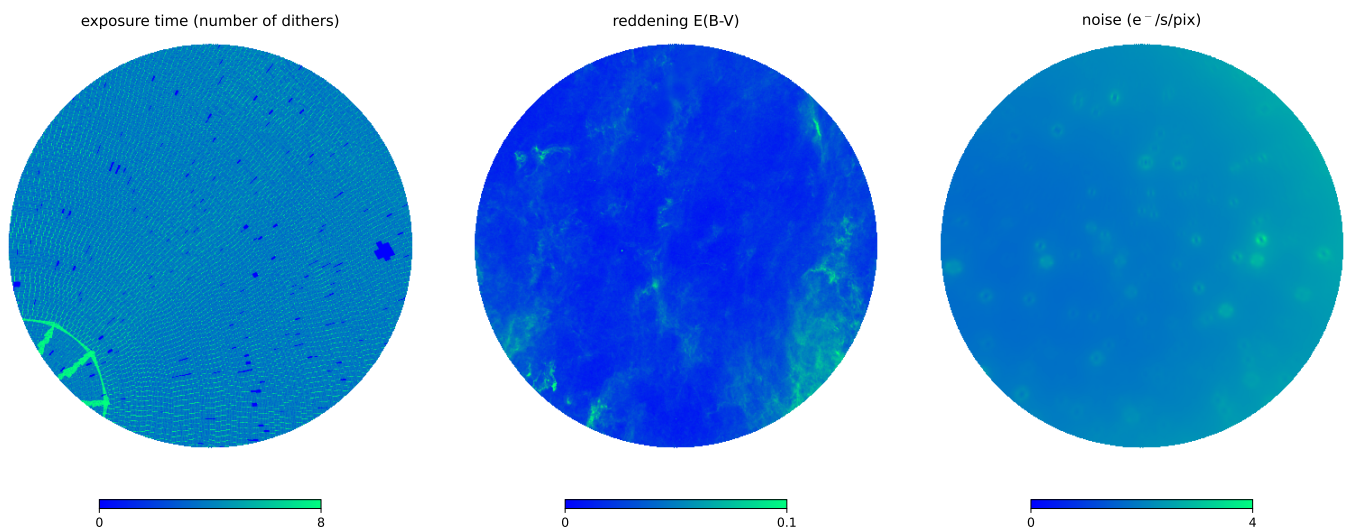


Fig. 5. Gnomonic projection of maps used for the average MVM, the marginalised visibility mask. Left panel: the exposure time map at the photometric position of the galaxy, in units of the number of times (dithers) the sky pixel is visited; a full dithering sequence can provide up to four visits, higher values happen at pointing overlaps. Mid panel: the $E(B - V)$ reddening map. Right panel: the noise map in units of $e \text{ pix}^{-1} \text{ s}^{-1}$.

clid Collaboration. These models were interpolated on `healpix` tessellations of varying N_{side} , so as to have a map at the same resolution as the exposure time map. While the exposure time and reddening maps are realistic, the noise map is clearly optimistic and simplified, since it lacks small-scale power coming from all the effects discussed in Sect. 3 and Appendix B. Since these features are very hard to predict before real images are taken, we decided to represent them by adding a detector noise term that has a constant value of $1 e \text{ pix}^{-1} \text{ s}^{-1}$.⁷ As a consequence, in these tests small-scale power in the random catalogue is mostly given by exposure time variations.

4.3. The tabulated detection probability

The mock catalogues described in Sect. 4.1 provide a minimal set of properties for each galaxy, namely flux f , redshift z , sky coordinates, and few other non-observable properties (like host halo mass, redshift without peculiar velocity contribution, central or satellite galaxy). To decide if a galaxy is part of the observed catalogue (to compute P_{det}) we could use the `pypelid` bypass of VMSP, but this requires knowledge of many other galaxy properties, called $\mathbf{p}$ in Eq. (7), so we should find a way to assign these properties to our galaxies. Since the mocks were constructed to reproduce the Flagship mock catalogue, one could assign $\mathbf{p}$ by randomly extracting that of a Flagship galaxy with very similar flux and redshift. An equivalent and more convenient procedure is to directly use the MVM $C(f, z, \{N_i\})$, obtained by marginalising $P_{\text{det}}(f, z, \{N_i\} | \mathbf{p})$ over galaxy properties (Eq. 7). This can be computed by applying `pypelid` to a large set of galaxies, namely those of the Flagship mock; marginalisation will simply consist of computing an overall detection probability for galaxies in bins of z , f , and nuisance map values (exposure time, reddening, and noise).

We thus run `pypelid` on the full Flagship catalogue, placing the octant on the sky so as to minimise the area heavily obscured

by the MW. We assumed constant noise and exposure time, then computed the average galaxy detection probability in 10 bins in f , 90 bins in z , and 3 bins in $E(B - V)$, and then constructed a table that gives the detection probability already marginalised over galaxy properties, using the detection model of Eq. (22) shown in Fig. 3. The tabulated detection probability is thus a numerical representation of the MVM. Here it is convenient to use for f the flux of the $H\alpha$ line *after* correction for extinction; in this way the effect of reddening is only to effectively shift the multivariate distribution of galaxy properties we are marginalising over. Since the dependence of S/N on exposure time t_{exp} and noise level N is predictable, the detection probability can be rescaled to any value of these variables.

The fine binning in redshift was applied to be able to follow sharp features connected to the entry of a line in the red grism; the only significant feature we find is a jump in detection probability at $z = 1.42$, when the $[\text{O III}]\text{b}$ line enters the red grism. To avoid sample variance being propagated to the MVM, for each combination of values of f , t_{exp} , N , and $E(B - V)$, the detection probability was smoothed in redshift by applying a Gaussian smoothing over five $\Delta z = 0.01$ bins,⁸ and this was done separately below and above $z = 1.42$, so as not to smooth the jump. Figure 6 shows the tabulated detection probability as a function of z , f , t_{exp} , and N , for relevant combinations of the other parameters; in this figure the exposure time is given as a multiple of the exposure time of a single dispersed dither, 547 s.

The procedure described here relies on a mock galaxy catalogue, like the Flagship mock, which provides knowledge of the *true universe*. When dealing with real data, this role will be played by the EDS. Constructing the tabulated detection probability will not, in principle, require more work than constructing the random catalogue: VMSP will use `pypelid` for assigning a detection probability to galaxies, so it will be possible to collect the results of this task into a table, while constructing the relevant noise map. Moreover, the very large number of galaxies in

⁷ This is a very prudent value, higher by one or two orders of magnitude than the noise found in detector tests (Euclid Collaboration: Cogato et al., in prep.). It is supposed to represent the overall effect of pixel-scale noise terms like bad pixels, cosmic rays and so on.

⁸ To avoid border effects, the data vector to be smoothed was linearly extrapolated beyond the nominal redshift range by 15 bins before being smoothed.

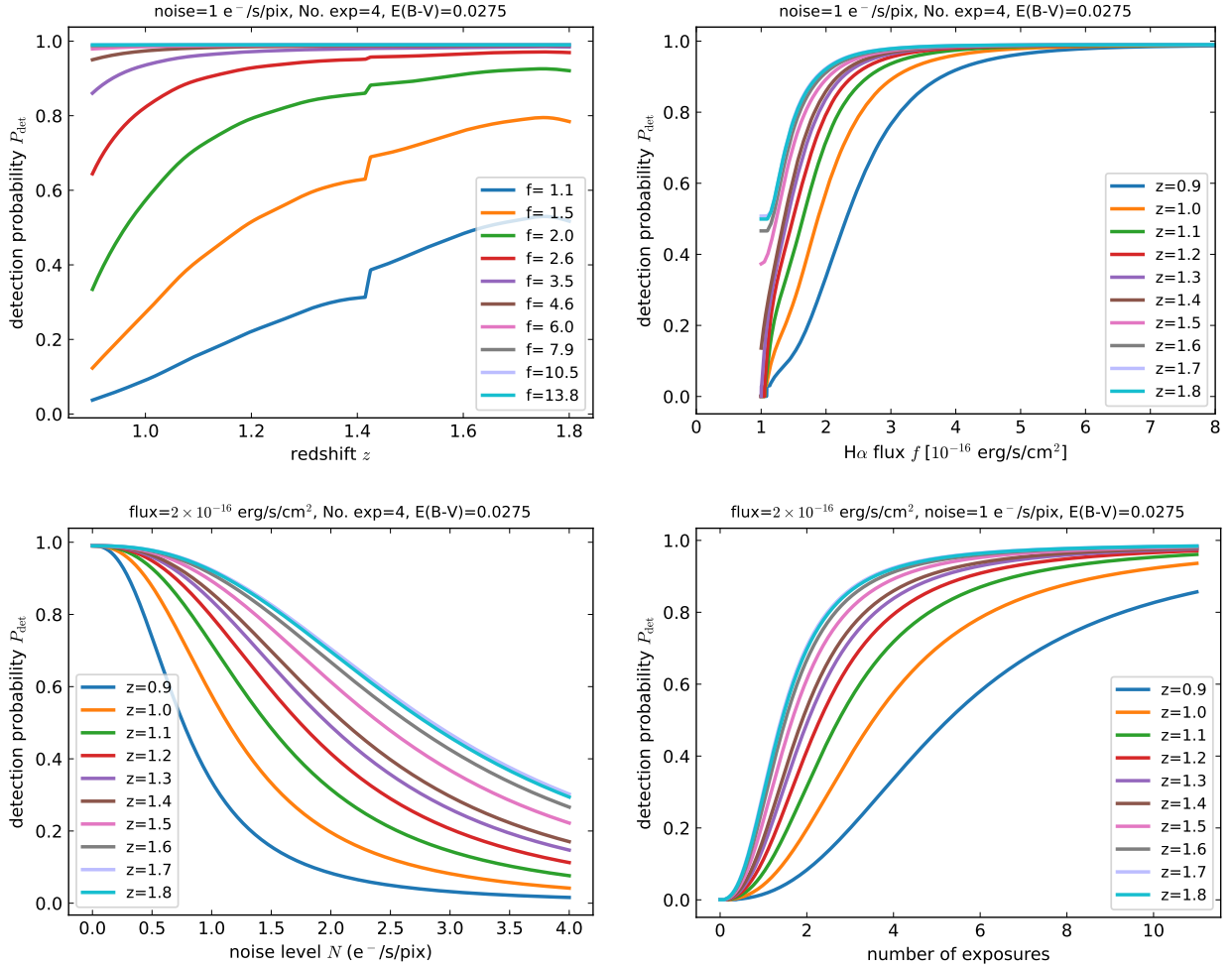


Fig. 6. Tabulated detection probability P_{det} , marginalised over galaxy properties. *Upper left:* as a function of redshift z for several flux limits (in units of $10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$). *Upper right:* as a function of H α flux f for several redshifts. *Lower left:* as a function of noise level N for several redshift. *Lower right:* as a function of number of exposures for several redshifts. The values of the parameters that are kept fixed are shown above each panel. In all panels $E(B - V)$ is set to the value 0.0275; the dependence on reddening is very mild and not very informative, so we do not show it here.

the master random catalogue will ensure that the table is properly populated.

4.4. Measurements and cosmological inference

The power spectra of galaxy catalogues were computed using *Euclid*'s PK PF (Euclid Collaboration: Salvalaggio et al., in prep.). The code computes the galaxy density from a catalogue using a piecewise cubic spline interpolation on a grid, then computes its power spectrum using fast Fourier transforms (FFTs) and the technique of interlacing to suppress aliasing near the Nyquist frequency (Sefusatti et al. 2016). It implements a Yamamoto–Bianchi estimator (Yamamoto et al. 2006; Bianchi et al. 2015; Scoccimarro 2015) to measure the first five multipoles of the power spectrum.

The Fourier-space estimator by construction provides a measurement that represents the convolution of the true power spectrum with a window function, defined as

$$W(\mathbf{x}) = W_{\text{fp}}(\hat{\mathbf{n}}) w_{\text{fkp}}(z) \alpha n_r(\mathbf{x}). \quad (24)$$

Here $W_{\text{fp}}(\hat{\mathbf{n}})$ gives the angular footprint of the survey and is 1 where there are observations and 0 otherwise, w_{fkp} is a weight

defined in Feldman et al. (1994), and $n_r(\mathbf{x})$ is the density of the random catalogue. A theoretical prediction of the measured power spectrum can then be obtained by convolving the theory power spectrum, valid for an infinite volume, with the Fourier transform of the window. This convolution was performed using a mixing matrix computed as follows. We used the same PK code to compute the power spectrum of the window function, $P_\ell^W(k)$. This was then Hankel-transformed to obtain the two-point correlation function of the window and from it the mixing matrix $\mathcal{W}_{\ell\ell'}$, such that the expectation value of the power spectrum measurement can be related to the theoretical power spectrum P_ℓ^{Th} as

$$\langle \hat{P}_\ell(k) \rangle = \sum_{\ell'=0}^4 \sum_{k'=k_1}^{k_2} [\mathbb{W}(k, k')]_{\ell, \ell'} P_{\ell'}^{\text{Th}}(k'), \quad (25)$$

where

$$[\mathbb{W}(k, k')]_{\ell, \ell'} = \Delta \log_{10} k' (k')^3 [\mathcal{W}_{\ell\ell}(k, k') - \mathcal{W}_{\ell\ell}(0, k')]. \quad (26)$$

Here $\Delta \log_{10} k'$ (meaning k' over $h \text{ Mpc}^{-1}$) denotes the logarithmic k' bin size and the limits of the summation operator, (k_1, k_2) ,

are determined by the sampling of the power spectrum window function $P_\ell^W(k)$ over the Fourier-space grid.

For all the measurements shown in this work, we embedded the catalogues in a box with side length $L_{\text{fit}} = 7000 h^{-1} \text{Mpc}$, with $N^3 = 1200^3$ cells. This allowed us to measure the power spectrum multipoles from the fundamental frequency $k_f = 2\pi/7000 = 8.98 \times 10^{-4} h \text{Mpc}^{-1}$ to the Nyquist frequency $k_{\text{ny}} = Nk_f = 1.08 h \text{Mpc}^{-1}$. To limit the data vector size with minimal information loss, measurements were binned in bins of $6k_f$. For the window function power spectrum measurements, we computed ten spectra for each mixing matrix, varying only the measurement box size and, consequently, the fundamental frequency of each measurement. Using this multiscale approach, we could sample the window function power spectrum over the range $7.8 \times 10^{-6} h \text{Mpc}^{-1} < k < 0.91 h \text{Mpc}^{-1}$. The resulting spectra were then used to compute the mixing matrix.

For the theoretical model of the power spectrum, we considered two perturbative models that differ in their treatment of redshift-space distortions; the performance of these two models is analysed in the context of the *Euclid* survey in Euclid Collaboration: Camacho-Quevedo et al. (in prep.). The first model implements the effective field theory (EFT) of large-scale structure to describe galaxy clustering observables through perturbative expansions (at one loop) of the nonlinear density and velocity fields (for a recent review, see Ivanov 2024). The second model, called velocity-difference generator (VDG), also relies on a one-loop perturbative expansion of the nonlinear density field, but replaces the perturbative treatment of real-to-redshift-space mapping with a non-perturbative approach that effectively captures the impact of small-scale velocity divergences on large-scale structures (Scoccimarro 2004; Eggemeier et al. 2025). Both models decompose the theoretical power spectrum into three main contributions. The first originates from the dynamical, bias, and redshift-space distortion terms at leading and next-to-leading order within standard perturbation theory (SPT). The second contribution arises from the stochastic components of the density and velocity fields. Finally, the third component accounts for small-scale corrections missing in the SPT terms, modelled via a set of parameters known as counter-terms.

We can categorise the parameters of both models into four distinct subsets (see the cited papers for detail). The first subset comprises the cosmological parameters, $\Theta_{\text{cos}} = \{h, \omega_c, \omega_b, A_s, n_s\}$. The second subset includes the bias parameters, which relate the galaxy overdensity field to the matter density field: $\Theta_{\text{bias}} = \{b_1, b_2, \gamma_2, \gamma_{21}\}$. The third subset consists of the stochastic shot-noise parameters: $\Theta_{\text{sn}} = \{N_0^p, N_{20}^p, N_{22}^p\}$. Finally, the fourth subset contains the counter-term parameters: $\Theta_{\text{ctr}} = \{c_0, c_2, c_4\}$. Each model introduces an additional model-specific degree of freedom. In the EFT model, this is the parameter c_{nlo} , an extra counter-term that captures next-to-leading-order corrections to the perturbative expansion sourced by the velocity-difference generating function. In the VDG model, this term is unnecessary as the velocity-difference distribution is modelled analytically with an additional parameter, a_{vir} , for its kurtosis.

For both models, we use the Gaussian-process emulator comet (Eggemeier et al. 2023) to compute model predictions. Following that paper, all parameters listed above are varied in the analysis, except for γ_{21} and N_{22}^p . The former is determined using the coevolution relation,

$$\gamma_{21} = \frac{2}{21}(b_1 - 1) + \frac{6}{7}\gamma_2, \quad (27)$$

Table 4. Prior distributions used for parameter inference in the EFT and VDG models.

Parameter	EFT	VDG
h	$\mathcal{U}(0.5, 1.0)$	★
ω_c	$\mathcal{U}(0.085, 0.155)$	★
A_s	$\mathcal{U}(1.4, 2.6)$	★
n_s	$\mathcal{N}(0.96, 0.0041)$	★
ω_b	$\mathcal{N}(0.02218, 0.00055)$	★
b_1	$\mathcal{U}(0.25, 4.0)$	★
b_2	$\mathcal{U}(-5, 5)$	★
γ_2	$\mathcal{U}(-5, 5)$	★
γ_{21}	Coevolution	★
$c_0 [\text{Mpc}^2]$	$\mathcal{N}(0, 200)$	$\mathcal{U}(-400, 400)$
$c_2 [\text{Mpc}^2]$	$\mathcal{N}(0, 200)$	$\mathcal{U}(-400, 400)$
$c_4 [\text{Mpc}^2]$	$\mathcal{N}(0, 200)$	$\mathcal{U}(-400, 400)$
$c_{\text{nlo}} [\text{Mpc}^4]$	$\mathcal{N}(0, 200)$	–
$a_{\text{vir}} [\text{Mpc}]$	–	$\mathcal{U}(0, 20)$
N_0^p	$\mathcal{N}(1, 3)$	$\mathcal{U}(0, 3)$
$N_{20}^p [\text{Mpc}^2]$	$\mathcal{N}(0, 3)$	$\mathcal{U}(-3, 3)$
$N_{22}^p [\text{Mpc}^2]$	–	–

Notes. Cosmological parameters are separated by a horizontal line from the bias and nuisance parameters. In fits performed on multiple redshift bins, the nuisance parameters are repeated for each redshift bin. The symbols $\mathcal{N}$ and $\mathcal{U}$ represent Gaussian and uniform distributions, respectively. Finally, a star indicates a shared prior between the two models, while a dash denotes an unused parameter.

and the latter, that is highly degenerate with c_{nlo} , is set to zero. Finally, for parameter inference, we employ the importance nested-sampling algorithm for Bayesian posterior reconstruction, as implemented in the Nautilus package (Lange 2023) and use the python package `getdist` (Lewis 2025) for the plotting and analysis of the parameter chains.

We want to stress here that the analyses described in this paper are in continuation with those we presented in Euclid Collaboration: Salvalaggio et al. (in prep.), where the consistency of power spectrum measurements and window function modelling was tested in three steps. Firstly, a ‘faux-cone’ wedge-like survey was extracted from a set of 10 000 periodic boxes at $z = 1$, and its mixing matrix measured from its random catalogue; the convolution of the theory model, consisting in the average power spectrum measured on the periodic boxes, and the window function was consistent with the average measurement of the faux-cone power spectrum to better than 0.1%. Secondly, we measured the power spectra of the EuclidLargeMocks, without systematics, and fitted them with a model, checking that the original cosmological parameters used to run the simulation were recovered. Thirdly, we imposed simple angular systematics, caused by the varying exposure time due to the tiling of pointings, and checked that the fitting model, convolved with the new window function (which has small-scale power), fits the measurements at the same level as the previous case without angular systematics. In this paper we use the same approach, with the same mock catalogues and theory modelling, but with angular systematics represented by more complicated MVMs. Results presented in Sect. 5 are, as a matter of fact, an extended testing of the power spectrum estimator in the case of a realistic random catalogue.

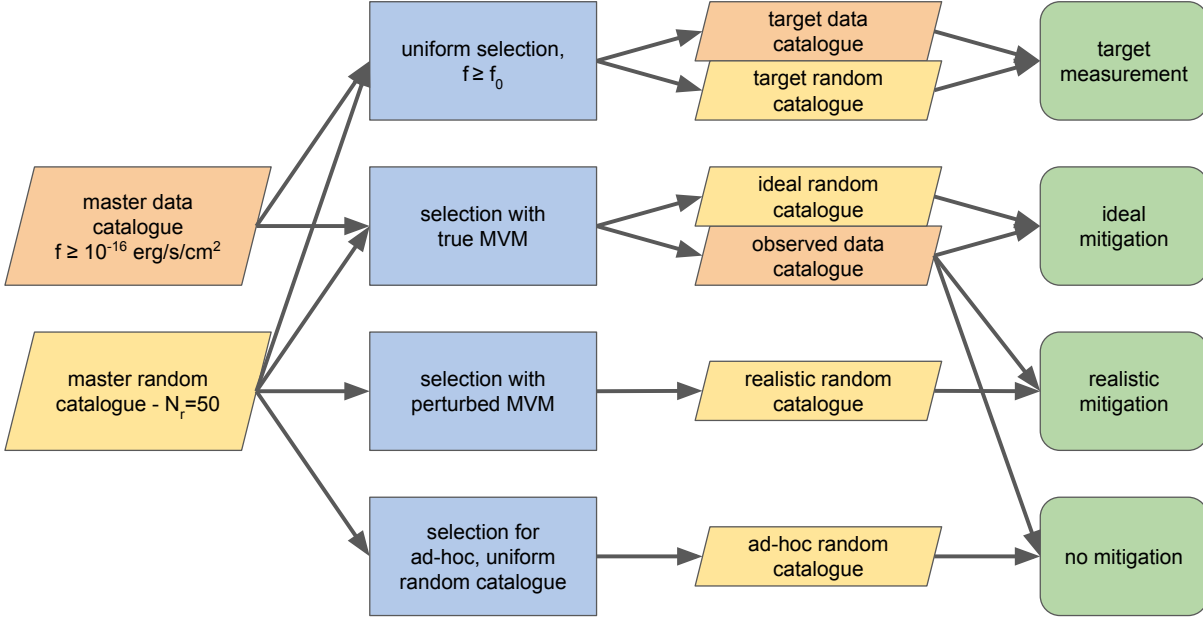


Fig. 7. Scheme of the strategy adopted for the processing of the mock catalogues; see Sect. 4.5 for details.

4.5. Strategy for the tests

We start from a set of 50 master data catalogues, taken from the EuclidLargeMocks, which are limited at $f \geq 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$. From these we construct a master random catalogue by picking up a number of galaxies equal to $N_r = 50$ times the galaxy density $n_g(z)$ averaged over the full set of 1000 mocks, which we regard as our theoretical number density. Of these galaxies we keep the observed redshifts and H α fluxes, so as to preserve the redshift-dependent galaxy LF, but we create random angular coordinates that sample the 30° circle footprint. This single master random catalogue is used for all the mock catalogues.

Then, as detailed in Fig. 7, both master data and random catalogues are subject to a selection, which may be as simple as selecting (target) galaxies with $f \geq f_0$ ('uniform selection'), or may be based on an MVM. The data and random catalogues may be selected using the same MVM, or the random catalogue may be selected with a perturbed MVM that accounts for the uncertainty in its construction. The 'ad-hoc selection' of the random catalogue is based on sparse-sampling it to reproduce (N_r times) the number density of the data catalogue subject to the MVM (averaged over 50 realisations). The resulting random catalogue will have a uniform density on the sky, so we use this selection to show what happens if a given angular systematics is not mitigated.

The study of the effect of systematics will make use of these combinations. (i) Target measurement: this is obtained by selecting the master data and random catalogues with a uniform selection $f \geq f_0$ (specific of target galaxies), with no angular dependence. (ii) Ideal mitigation: the master data and random catalogues are selected by applying the same MVM, so that the random catalogue corrects the angular systematics in the best possible way. (iii) Realistic mitigation: while the master data catalogues are selected with an MVM, the master random catalogue is selected with an approximated one. (iv) No mitigation: the master data catalogues are selected with an MVM, but the master random catalogue is subject to the ad-hoc selection.

5. Results

We test here how systematics impact the measurement of the galaxy power spectrum and the inferred cosmological parameters. The analysis is carried out in two steps. In the first step we use shuffled fluxes to remove luminosity-dependent bias; in this case the galaxy power spectrum is not expected to change (apart from the different window and level of shot noise) when a different selection is applied. This allows us on the one hand to quantify the effect of specific systematics, and on the other hand to stress-test our ability to correct the effects of an MVM with a random catalogue. Limiting ourselves to the analysis of the redshift bin $[1.1, 1.3]$, we first fit the measurement of the power spectrum of target galaxies (that have $f \geq f_0$) with a model, using the mixing matrix computed for the 30° cone without angular systematics. We then consider a generic MVM, apply it to the master catalogues and to the random catalogue, and measure the power spectra for this ideal mitigation. With this random catalogue we compute the window function and its mixing matrix, then use it to obtain a convolved model, where the theory model is the one we obtained with the fit of the target measurement. We then compare the measurements that have ideal mitigation with the convolved model, and test that their agreement is as good as that obtained when fitting the target measurements.

The second step addresses the stability of inferred cosmological parameters in the presence of systematics, both with ideal and with realistic (or even ad-hoc) mitigation. In this case we use the standard fluxes in place of the shuffled ones, and process the mock catalogues dividing them into four redshift bins, bounded by 0.9, 1.1, 1.3, 1.5, and 1.8 (see Sect. 4.1), and fit their measured power spectra at the same time, using a numerical covariance computed on the 1000 mocks with baseline systematics defined below. We present results for cosmological parameter posteriors and check their stability with respect to using an approximate MVM.

Table 5. List of cases of combinations of systematics used in the analysis.

Case reference name	Zodiacal light	Stray light	Detector noise	MW extinction	Exposure time	Detection model	Calibration error
<i>zodiacal</i>	yes	no	no	no	4	standard	no
<i>stray_light</i>	no	yes	no	no	4	standard	no
<i>img_noise</i>	yes	yes	yes	no	4	standard	no
<i>MW_P13</i>	no	no	no	P13	4	standard	no
<i>tiling</i>	no	no	no	no	map	standard	no
<i>baseline</i>	yes	yes	yes	P13	map	standard	no
<i>calib_2%</i>	yes	yes	yes	P13	map	standard	2%
<i>calib_10%</i>	yes	yes	yes	P13	map	standard	10%
<i>MW_P15</i>	no	no	no	P15	4	standard	no
<i>MW_SFD</i>	no	no	no	SFD	4	standard	no
<i>detmodel1</i>	yes	yes	yes	P13	map	$S_0 = 3.27$	no
<i>detmodel2</i>	yes	yes	yes	P13	map	$S_0 = 4.11$	no
<i>detmodel3</i>	yes	yes	yes	P13	map	$\beta = 3.73$	no
<i>detmodel4</i>	yes	yes	yes	P13	map	$\beta = 6.23$	no

Notes. Each case is defined by a set of choices, declared in each column. The first six cases are discussed in Sect. 5.1.1, the other eight in Sect. 5.1.2. The contributions to image noise (zodiacal light, stray light, detector noise) can be present (yes) or absent (no); MW extinction can be applied using the P13 (see Sect. 4.2), P15 or SFD reddening maps (see Sect. 5.1.2); the exposure time can be set to 4 dithers, equal for all sky pixels, or to the (redshift-dependent) exposure time map discussed in Sect. 4.2 assuming a redshift of $z = 1.2$. A realistic mitigation is obtained by perturbing the detection model or by adding a calibration error; the relative columns report the value of the perturbed parameter of the detection model or the value of the largest calibration error. In all columns, values in bold face highlight the non-trivial choices.

5.1. Testing measurements with shuffled fluxes

We fit the master catalogues without angular systematics using the VDG model (see Sect. 4.4), which has the advantage, with respect to the EFT model, of being more stable beyond the fitting range. For the numerical covariance, we compute it from 1000 measurements of all the EuclidLargeMocks, already presented in Euclid Collaboration: Monaco et al. (2025). We fit the monopole and quadrupole power spectra in the range $z \in [1.1, 1.3]$ up to $k_{\max} = 0.3 \, h \, \text{Mpc}^{-1}$, in bins of size $\Delta k = 6k_f$, using the parameter priors shown in Table 4. With the exception of the ω_b and n_s parameters, for which we adopt informative Gaussian priors, all the varied parameters have wide uniform priors.

We group the 50 mocks in ten groups of five mocks each (Sect. 4.1), so that the error on the fit is relative to the area of the final survey. Figure 8 shows the results of this fit; the upper panel shows the first three even multipoles of the power spectrum (here and in all similar plots shot noise has been subtracted), with the blue line giving the measurement and the red line the (convolved) fitting model. The lighter blue shaded area gives the variance of the measurements, computed for the 50 mocks and rescaled to represent the uncertainty on five measurements. The lower panels give, for the three multipoles, the residuals of the measurements with respect to the model, in percent, the vertical pink area highlights the first BAO region, and the vertical dotted lines show the range of scales used for the fit, from $k = 0.009 \, h \, \text{Mpc}^{-1}$ to $0.3 \, h \, \text{Mpc}^{-1}$. The model is in excellent agreement with the measurements in the fitting range, then it slowly departs from the data at higher wavenumbers, showing a 2% difference at $k \sim 0.5 \, h \, \text{Mpc}^{-1}$ for the monopole, while the agreement with the quadrupole worsens in a more marked way.

5.1.1. Ideal mitigation

We now analyse six cases of systematics, obtained by gradually switching on the various contributions to the noise map, the MW extinction, and the exposure time map; this last map is either set to a constant exposure value of 4 dithers, or to the exposure time map computed, as explained in Sect. 4.2, at the median redshift

$z = 1.2$. These six cases are reported in the upper part of Table 5, where the value of each nuisance term is reported. For these six cases, the detection model is perfectly known, and no calibration error is present. The cases tagged as *zodiacal*, *stray_light*, *MW_P13* and *tiling* have just one nuisance active, *img_noise* is a combination of the three contributions to image noise (zodiacal light, stray light and detector noise), while for *baseline* all the nuisances are present.

The upper panel of Fig. 9 shows the galaxy number density for the sample of target galaxies and for the six cases of systematics. This is quantified in observational units of galaxies per redshift interval per deg^2 . The number density of target galaxies decreases in redshift, following the evolution of the $H\alpha$ luminosity function of Pozzetti et al. (2016), while the other number densities have a peak at a redshift around 1.2. This different behaviour is the result of the increase in the detection probability with redshift shown in Fig. 6. Also, number densities with systematics can be higher than those of target galaxies, since at low noise levels the detection of galaxies with $f < f_0$ is more probable. From this figure we see that, when taken alone, stray light and zodiacal light have respectively the weakest and the highest impact on number density. The *baseline* case has the lowest number density, and this is expected since it is the combination of all other cases. The lower panel of Fig. 9 shows, for the *baseline* case, the purity and completeness metrics defined in Sect. 2.3; here redshift purity is unity because there are no interlopers, while redshift efficacy gets values larger than unity thanks to the high probability of detecting a galaxy at high redshift.

To better visualise these MVMs, from the selected random catalogues we count the random galaxies from $z = 0.9$ to 1.8 that fall onto a sky pixels for a healpix tessellation with $N_{\text{side}} = 512$, and from them we compute density contrasts simply as $(N_{\text{pix}} - \bar{N})/\bar{N}$. Figure 10 shows these maps; all of them show large-scale features, which are especially strong in the very smooth zodiacal light, while exposure time is the nuisance map that gives most small-scale power to the *baseline* selection.

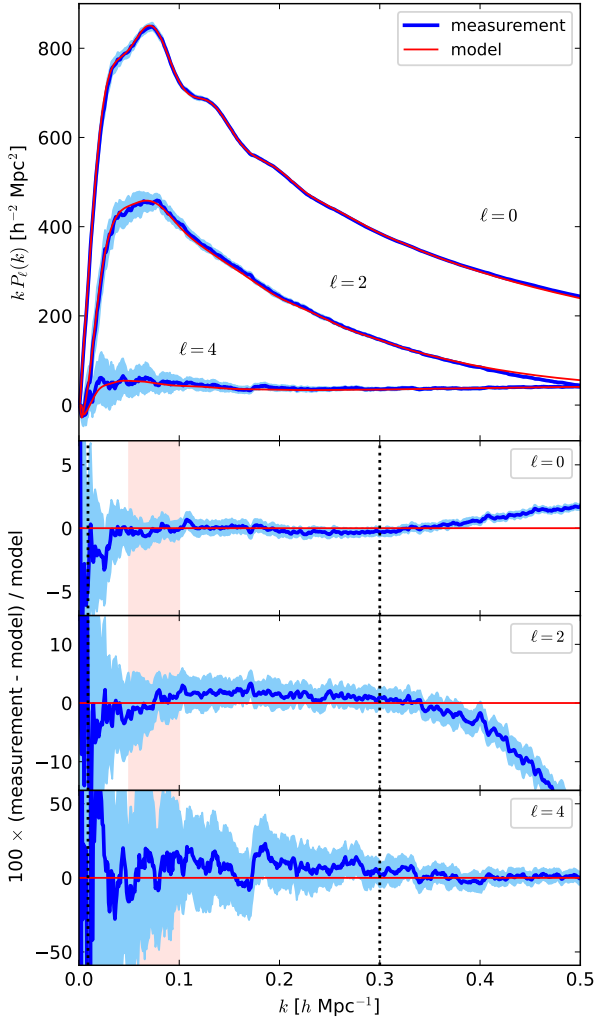


Fig. 8. Upper panel: power spectra even multipoles ($\ell = 0, 2, 4$) of target galaxies in the $z \in [1.1, 1.3]$ redshift bin, averaged over 50 realisations, with shuffled fluxes to remove luminosity-dependent bias. Blue lines give the target measurement, while the lighter blue shaded areas give the error on an average of five measurements, covering an area similar to the final EWS. The red line gives the best fit model, convolved with the 30° cone window function. Lower panels give, for the three multipoles, the relative difference with respect to the model fit to the master mocks, convolved with the relative window. The pink shaded area highlights the region of the first BAO, while the dotted vertical lines mark the range of scales used for the fit.

We first show, in Fig. 11, the power spectrum of target galaxies, from Fig. 8, compared with the power spectra obtained for the six cases of systematics when using the ad-hoc random catalogue that has a uniform density on the sky, so that no mitigation of angular systematics is performed. As expected, all the power spectra show very significant deviations from the one of target galaxies, with the largest effect present on large scales for the *zodiacal* and *MW_P13* cases, while *tiling* and the *baseline* selections show peaks that beat with the projected dimension of the NISP field.

Ideal mitigation can remove these features. This is shown in detail, separately for each case of systematics, in Appendix C. Here we investigate, in Fig. 12 and only for the power spectrum monopole, the agreement between the measurements with ideal mitigation and the model convolved with the corresponding win-

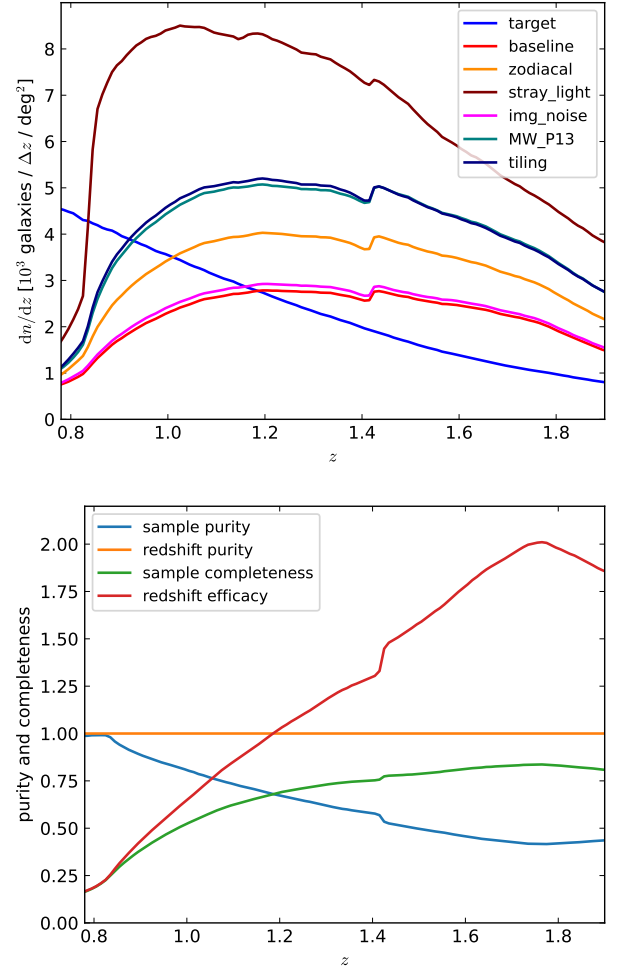


Fig. 9. Upper panel: number density dn/dz , in units of galaxies per redshift interval per deg^2 , of the target galaxies and of the galaxies subject to the six cases of systematics listed in Table 5. The lower panel reports the metrics of sample purity and completeness defined in Sect. 2.3 for the *baseline* case.

dow function. In the top panel we report the ratio between the ideal and target measurements, $P_i/P_t - 1$, quantified in percent. The measurements differ by up to 6%, and this difference should be accounted for by the convolution with the window. In the middle panel we show the residuals of measurement minus model, $P_i - P_{th}$, divided again by the target measurement P_t . Here the theory prediction is convolved each time with a different window; for the target measurement this line is identical to what was shown in Fig. 8 so its consistency with zero is a result of the fit; however, for the other lines the theory prediction is obtained by convolving the same model with the relative window function. All the lines fall on top of each other, showing that the level of agreement of model and (ideal) measurement is the same for a large range of MVMs, to sub-percent accuracy. The bottom panel shows another way of quantifying this excellent agreement. We use the difference of the measurement-minus-theory residuals between the given systematics and the target result, which we denote as

$$\Delta\Delta P := \frac{(P_i - P_{th,i}) - (P_t - P_{th,t})}{P_t}, \quad (28)$$

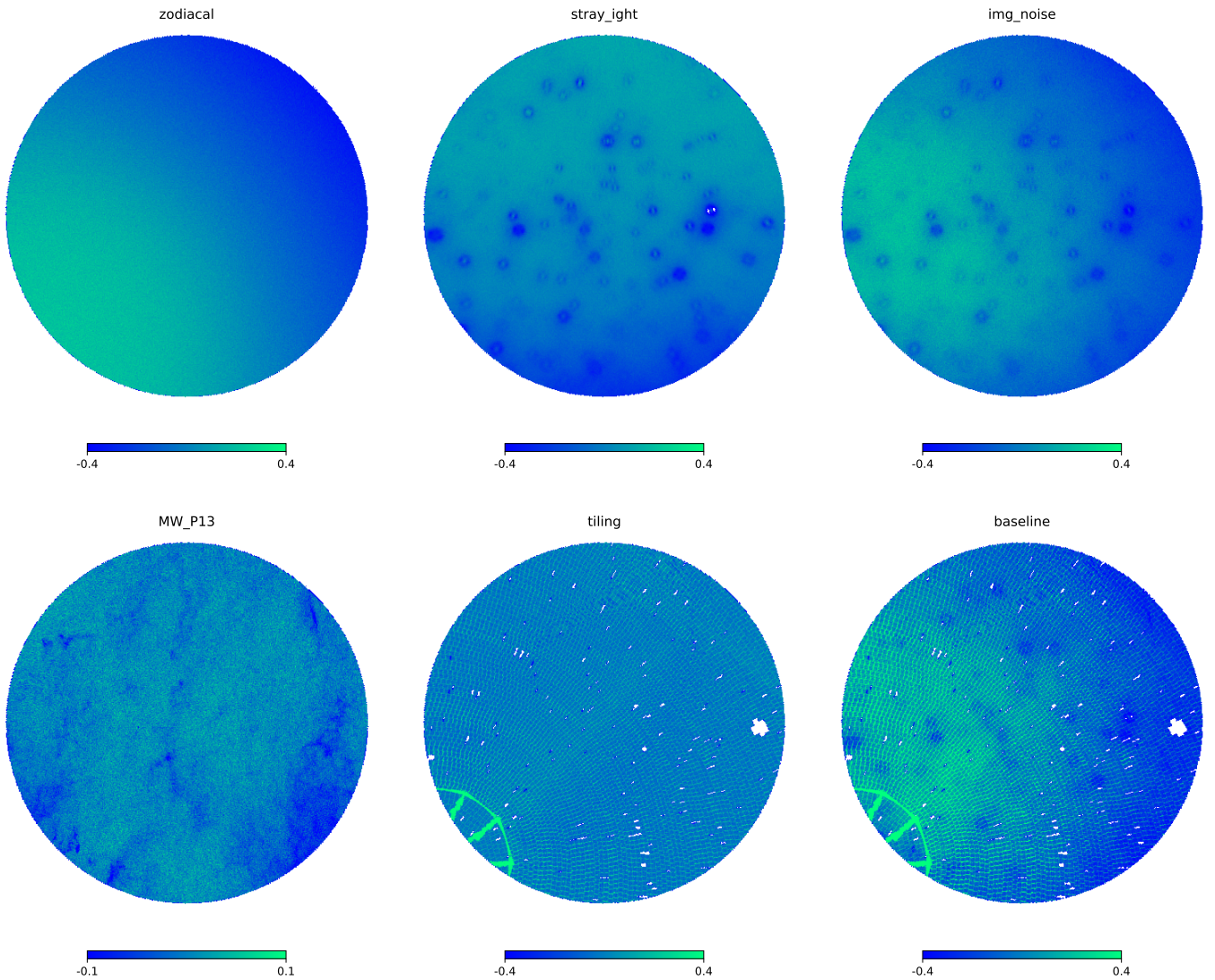


Fig. 10. Projected density contrasts of the random catalogues that are subject to the MVMs of the systematic cases declared above each panel and listed in Table 5. White spots in the last two maps denote unobserved regions. Notice that the limits of the colour bar for the *MW_P13* case are not the same as the other the maps.

where the theory model is convolved with the window of the specific systematic, $P_{th,i}$, or with the 30° cone window of the target measurement, $P_{th,t}$. This statistics quantifies the bias induced by the convolution process, and it is consistent with zero to sub-percent level. This result provides a very strong validation of the whole analysis pipeline.

We draw the following conclusions. (i) In all cases, the convolved model fits the power spectrum of the ideal mitigation cases to within sub-percent accuracy in the monopole and correspondingly good accuracy for the quadrupole and hexadecapole (see Fig. C.1). (ii) In the no-mitigation cases, the window convolution fails to achieve adequate accuracy. (iii) Fluctuations induced by zodiacal light and stray light have most power on large scales, and their effect on the measurements is a more or less constant relative shift on the monopole and quadrupole. (iv) MW extinction has effects only on the largest scales, and produces no shift at small scales. (v) The exposure time map induces power on small scales, and produces the most complicated features. (vi) The baseline MVM shows a combination of all the above fea-

tures, again reproduced by the convolved model with excellent accuracy.

5.1.2. Realistic mitigation

We now investigate to what level a realistic mitigation can recover the target power spectrum. We can perturb the three nuisance maps (exposure time, noise, and MW extinction), or we can perturb the detection model. We can assume that the exposure time is known exactly.

As discussed in Appendix B.3, assessing the uncertainty on MW extinction is not easy. We represent it here using the difference between the three maps: besides P13, we use another map presented by the Planck Collaboration (hereafter P15, Planck Collaboration: Ade et al. 2016) and the older Schlegel et al. (1998) reddening map (hereafter SFD). These are reported in Table 5. To show the effect of the uncertainty in this systematic, we compute (similarly to what we did in Fig. 10) a projected density contrast based on the true MVM, that is the true completeness function C , and its approximation $\hat{C}$ (see Sect. 2.4),

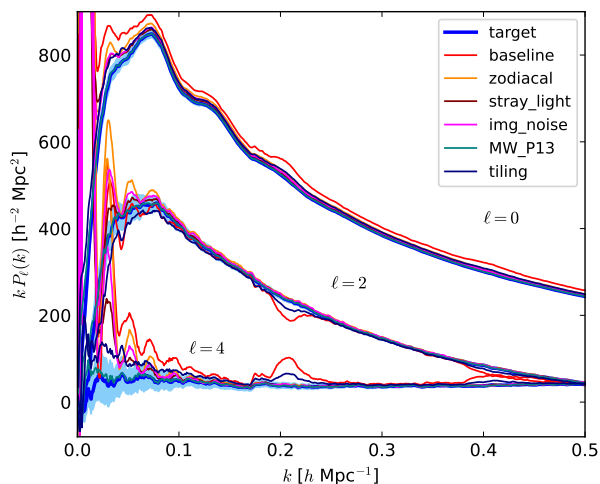


Fig. 11. Power spectrum multipoles for the sample of target galaxies and for the galaxy catalogues subject to systematics for the six cases reported in Table 5, when no mitigation of angular systematics is performed (we use the random catalogue with the ad-hoc, uniform selection). Line colours are given in the legend, the lighter blue shaded area gives the variance of the target measurement, rescaled to the average of five catalogues.

as $\Delta_C = (N_{\hat{C}} - N_C)/N_C$, where N is again the number of random galaxies in a sky pixel. The upper left panel of Fig. 13 shows this quantity computed using *MW_P13* and *MW_P15*. The strongest difference is due to the modelling of dust spurs, but this difference induces modest projected density contrasts. In Fig. 14, left panel, we show the power spectrum multipoles, with a logarithmic scale on the x -axis to better highlight the behaviour on large scales. Here the data catalogues have been selected using P13, while the random catalogue has been processed with the other two reddening maps; the extinction curve is the same in all cases. While the BAO scale is again accurate to the sub-percent level, large scales are somehow sensitive to the specific map, and while SFD and P13 give fairly consistent results, P15 deviates significantly.

Another source of uncertainty that should be addressed comes from an extinction curve that varies on the sky (e.g. Schlafly et al. 2016), due for instance to a variation of dust composition in the Solar neighbourhood. This case is discussed in Appendix B, we expect this effect to be important on the largest scales and do not test it in this paper.

The noise map will be read from the dispersed NISP images, so its measurement will be subject to a statistical uncertainty that will be localised in the pixel or, at most, in the surrounding pixels. Because the quantity that VMSP uses is the error on pixel electron counts, to perturb this we should know the error on the error. While it is possible to compute this quantity and perturb the pixel-level noise, we can anticipate that the impact on the results would be negligible: this perturbation is degenerate with the effect of the scatter in the detection model, which will be shown below to be negligible. We will thus neglect this perturbation term.

A more important and potentially relevant effect is the accuracy in the calibration of the photometric zero-point in the NISP images, discussed in Appendix B.6. A drift of this zero-point would bias the noise map input to the tabulated detection probability, and since it couples with the survey timeline, it would create large-scale features in the MVM. The calibration of NISP

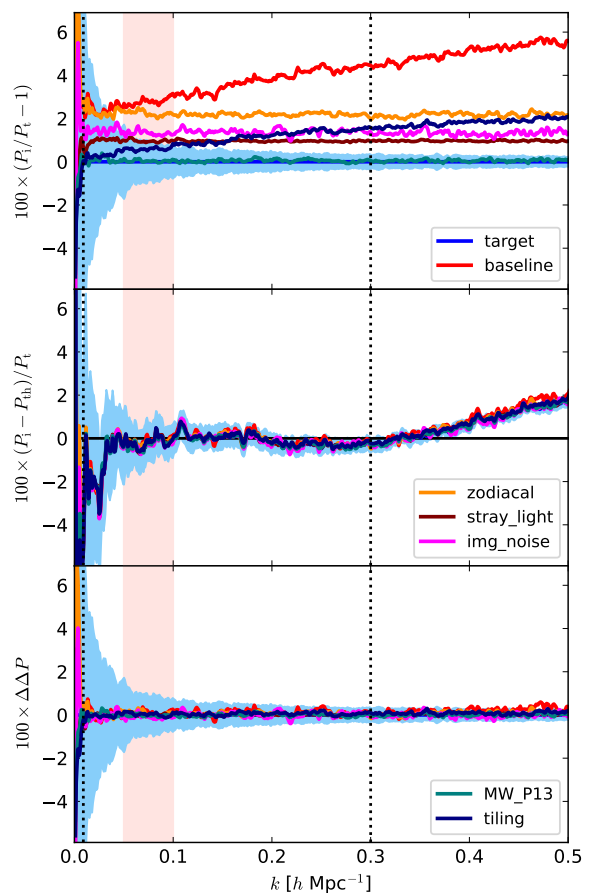


Fig. 12. Comparison of the measurements of the power spectrum monopole, for the sample of target galaxies (P_t) and for the galaxy catalogues subject to systematics for the six cases reported in Table 5 (denoted as P_i), when ideal mitigation of angular systematics is performed (the data and random catalogues are selected with the same MVM). The legend is distributed over the various panels for better readability, the lighter blue shaded areas give the variance of the target measurement, rescaled to the average of five catalogues. In the top panel we show the relative difference, in percent, of P_i with respect to P_t ; the measurements differ, despite their ideal mitigation, because the different random catalogues define different window functions. The middle panel gives, in percent, the quantity $(P_i - P_{th})/P_t$, where P_{th} is the best-fit model of the sample of target galaxies, convolved with the corresponding window. For the target sample, this is just the residuals of the fit, but the other lines give the difference between the measurement and the same theory model convolved with the relative window with systematics. Their agreement confirms that the differences observed in the upper panel are due to the different window functions. The bottom panel shows, again in percent, the quantity $\Delta\Delta P$ defined in Eq. (28), which gives difference between the residuals for a given systematic and the residuals of the fit of the target measurement.

images, described in Appendix B.6, relies on repeated, monthly visits of a field named SelfCal, aided by repeated visits of the EDS fields. We present here an experiment where the calibration has a systematic variation that is reset at any visit of the SelfCal field or of the EDS. Because the time span between two successive visits of a calibration field can be a month, a drift of 0.7% (as required for the stability of calibration of nearby fields) between successive ROSs would give catastrophically large offsets, so we do not use this figure but assume that the offset in zero point of flux measurement can have a maximum value. We construct a toy model for the calibration error by assuming that:

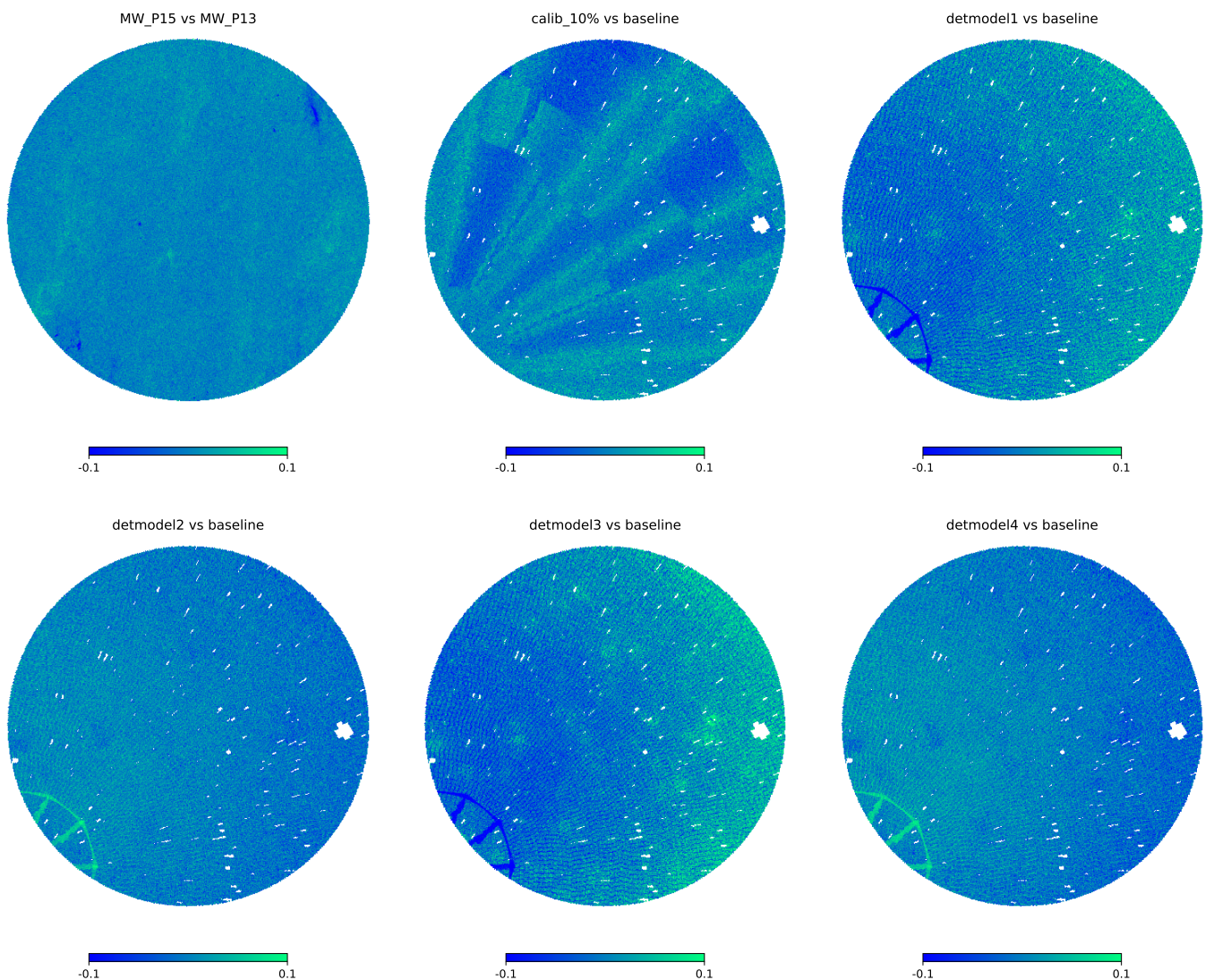


Fig. 13. These maps show how an error in the calibration of the MVM emerges as a pattern on the sky. The maps report projected density contrasts obtained using two random catalogues, one representing the true MVM, C , and the other the perturbed one, $\hat{C}$. For each random catalogue we compute its projected number density N on a healpix map, and show $\Delta_C = (N_{\hat{C}} - N_C)/N_C$. From the upper left to the lower right panels we show cases selected from Table 5: *MW_P15* vs. *MW_P13*; *calib_10%* against *baseline*; *detmodel1* to *detmodel4* against *baseline*.

(i) the zero point drifts as a parabola (as a function of time) between two successive visits of either the SelfCal field or one of the EDS fields; (ii) the slope of this parabola at the recalibration time is always the same; and (iii) the largest possible value of the zero-point (computed for the largest time span between visits of the calibration fields) is either 2% or 10% (Table 5), the first value being pessimistic and the second unrealistic, used just to stress-test the scheme. The value of noise used in the tabulated detection probability is then increased by this relative amount.

The upper-middle panel of Fig. 13 shows the map obtained using the *baseline* case for C , and *calib_10%* for $\hat{C}$ to make the effect more visible. This calibration error, coupled with the survey timeline, creates features that are easy to recognise, yet relatively weak. In Fig. 14, middle panel, we show the power spectra of the target and ideal measurements, compared with realistic mitigation for 2% and 10% calibration errors; for simplicity we do not again report the ad-hoc mitigation, and give in the residual panels only the ratio of the measurements; the scale of the x -axis is logarithmic to better show the large-scale behaviour. It is ev-

ident that a pessimistic 2% calibration error hardly changes the measurement, while a 10% error creates some significant features on large scales, but gives no effect at or below the BAO scale. We conclude that the calibration error is not expected to lead to significant effects on the clustering measurements.

To assess the impact of an error on the detection model, we follow the quantification of its uncertainty presented in Sect. 3.2 and perturb the sigmoid function adding $\pm 1\sigma$ to the value of each of the fitting parameters. We thus obtain four perturbed MVMs, which are also listed in Table 5. To construct these MVMs we take the tabulated detection probability described in Sect. 4.3, for each bin in f , z and reddening we work out an effective S/N using the true detection model, then feed it to the perturbed detection model to recompute the detection probability. This algorithm should in principle be applied to the galaxies processed by *pypelid* and not to the probability averaged over many galaxies, but the fine binning in redshift and flux used to compute the tabulated detection probability should guarantee that all galaxies

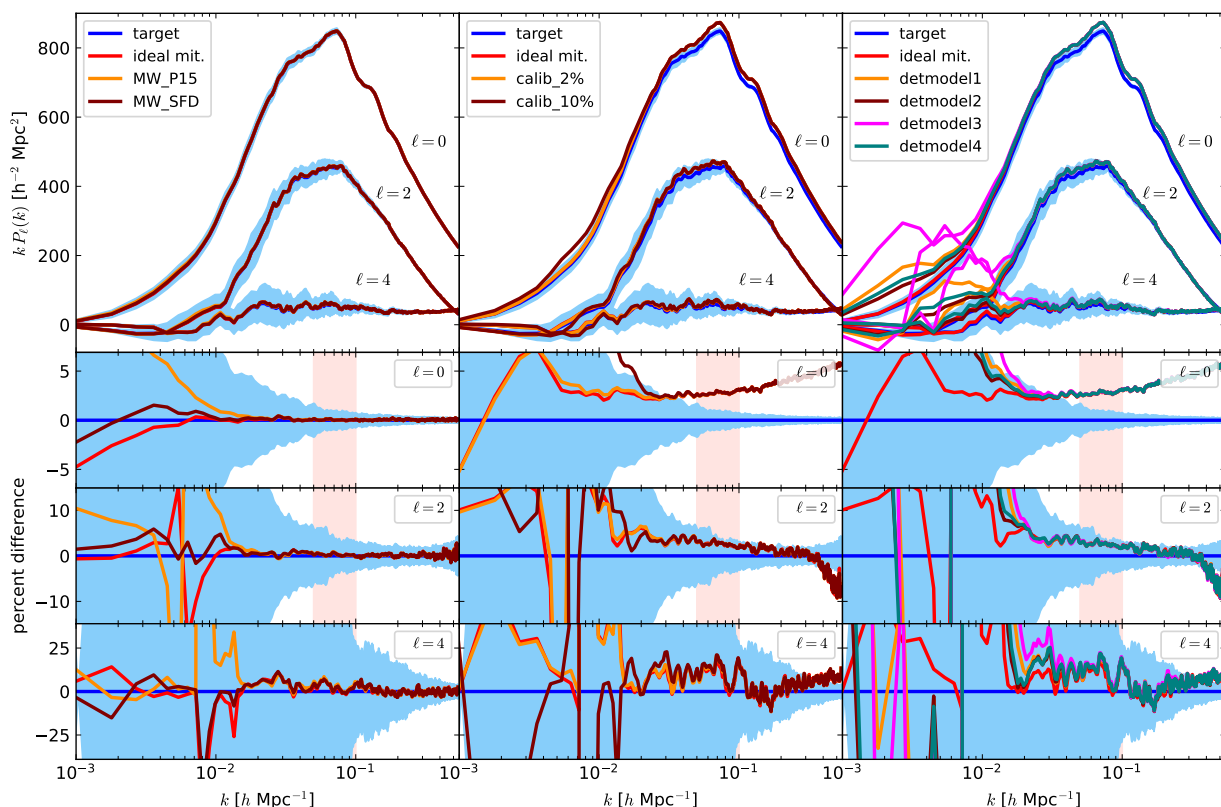


Fig. 14. Monopole, quadrupole and hexadecapole of the galaxy power spectrum in the case of realistic mitigation, for MW extinction on the left, calibration error in the middle, and error on the detection model on the right. The lighter blue shaded areas give the variance of the target measurement, rescaled to the average of five catalogues. Here the x-axis scale is logarithmic. The residual panels report the ratio of the measurements with respect to the target one, all in per-cent.

that end up in a bin have similar S/N, so this approximation is considered adequate for this test.

Figure 13 shows the projected density contrast maps of the four perturbed random catalogues with respect to the *baseline* one, and Fig. 14 (right panel) shows the impact of the perturbations of the detection model on the power spectrum multipoles. Errors in the detection model mostly impact large scales, while all scales smaller than the BAO scale are largely unaffected. However, from Fig. 13 it is clear that these systematics leave in the realistic random catalogue a very predictable residual map. This will be recognisable in the density contrast field traced by $n_o - \alpha n_r$, opening the possibility of calibrating the detection model by requiring no correlation of the mitigated density contrast field of the spectroscopic catalogue with this potential residual map. In addition, these perturbations of the detection model induce a relative change in the number density of the random catalogue of the order of 20–30% with respect to the *baseline* case, while the accuracy of the LF $\Phi(f|z)$ from the EDS will be 1 to 2% (Euclid Collaboration: Bruton et al., in prep.). This choice would lead to a mismatch between random catalogue and catalogue number densities, already discussed in Sect. 3.3, which could be used to fine-tune the calibration of the detection model. This confirms that the variations of the detection model we are implementing here are pessimistic.

A final source of systematics lies in the intrinsic scatter of the detection model. We assume here that this scatter has no correlation with galaxy properties, an assumption that will be directly verified in the calibrator set. We implement this scatter at the catalogue level by modulating the probability to detect a

galaxy (in the *baseline* case) with a Gaussian-distributed multiplicative term with variance provided by the fitting function shown in Fig. 3. The random catalogue is consistently selected with the *baseline* case, since the scatter is intrinsic to the data but cannot be modelled in the MVM. The resulting catalogues show a power spectrum that is nearly indistinguishable from that obtained using the *baseline* selection. Clearly, such a pixel-level noise does not propagate to clustering on cosmological scales. We finally notice that, as anticipated above, adding a scatter to the detection probability is degenerate with adding an error to the noise map, and more generally to any pixel-level source of uncertainty that adds scatter but whose average is well represented in the random catalogue. These uncertainties do not propagate to clustering measurements, at least on the scales that are used to infer cosmology.

5.2. The effect of systematics on cosmological inference

When the MVM is applied to non-shuffled fluxes, the intrinsic clustering signal changes as a function of selection, due to the luminosity dependence of galaxy bias. We then test the impact of realistic mitigation directly on the inferred cosmological parameters; the metric for the success of a mitigation strategy will be how well the inference procedure recovers the cosmological parameters as compared to the analysis of target galaxies.

We fit the four redshift bins simultaneously, using the model described in Sect. 4.4 and a numerical covariance obtained by applying the *baseline* combination of systematics to all the 1000 EuclidLargeMocks and measuring its power spectra in the four

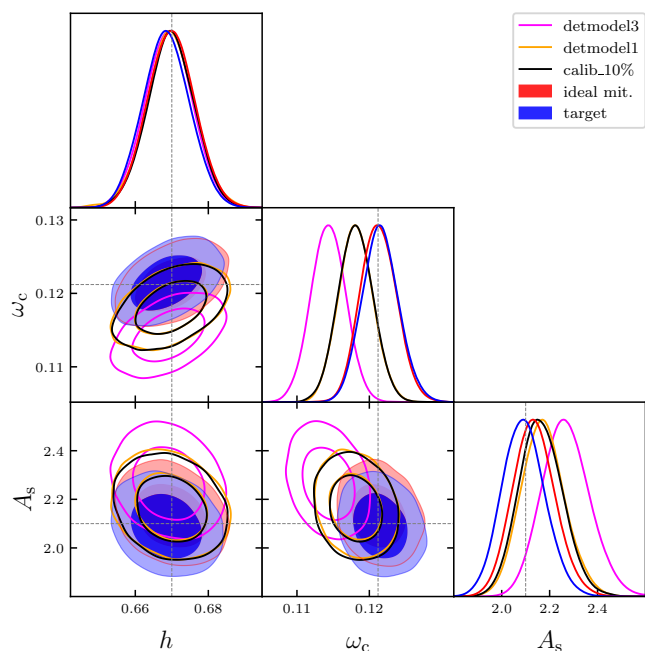


Fig. 15. Shifts in the derived cosmological constraints due to angular systematics under various mitigation strategies. The colour scheme matches that of Fig. 14.

redshift bins. We process five different cases: the target catalogues, the ideal mitigation in the *baseline* case, and three of the pessimistic scenarios listed in Table 5, namely *detmodel1*, *detmodel3*, and *calib_10%*. For each data vector in the fit, we average over five mocks and similarly rescale the covariance (see Sect. 5.1). We use the EFT model, with the priors listed in the second column of Table 4. To speed up parameter inference, we apply analytical marginalisation as in Carrilho et al. (2023), a technique that has been shown to be effective without biasing the un-marginalised parameters (see figure 8 in Pezzotta et al. 2025) and can be applied to any parameter that enters the model linearly and has analytically integrable priors. With this technique and adopting wide Gaussian priors we analytically marginalise over $\theta_{\text{AM}} = \{c_0, c_2, c_4, c_{\text{nl0}}, N_0^P, N_{20}^P\}$.

In Fig. 15, we present the posteriors of cosmological parameters obtained using a single set of five simulations. The colour scheme matches that of Fig. 14. The ideal mitigation achieved using the *baseline* random catalogue almost perfectly matches the constraints of the target catalogue. Additionally, we observe that realistic mitigations primarily affect the ω_c parameter, while their impact on h is minimal and on A_s is limited.

To show results across different realisations, in Fig. 16, left panel, we report the 1 and 2σ confidence intervals on the cosmological parameters, represented by thick and thin bars, respectively, for all ten sets of realisations. The figure shows that *detmodel3* consistently biases the estimates of A_s and ω_c , while the ideal mitigation consistently recovers parameter values similar to those of the target samples, with small systematic differences which are well below 1σ . The right panel of the same figure reports the inferred linear bias parameter b_1 for the four redshift bins. Due to the different selections and to luminosity-dependent galaxy bias, the b_1 value of target galaxies differs from the other selections, with this discrepancy becoming more pronounced at higher redshifts. It is evident though that the model is effective in marginalising over the differences in galaxy bias.

These results demonstrate that the strategy described in this paper to mitigate angular systematics through a random catalogue allows unbiased inference of cosmological parameters when the systematics are perfectly modelled. This work lays down a roadmap for assessing the error budget for the *Euclid* spectroscopic sample, which is dominated by the uncertainty in the calibration of the detection model. Appendix D gives a more detailed assessment of the metrics of the goodness of fit in the five cases considered here.

6. Conclusions

We presented the strategy devised by the Euclid Collaboration for dealing with data systematics in the cosmological exploitation of the *Euclid* spectroscopic sample, starting from the definition of the observable, the galaxy density n_g and its density contrast δ_g . We surveyed the processing pipeline to have a wide and organised view of all the possible contributions to the cosmological error budget. Differently from previous surveys, the spectroscopic Euclid Wide Survey will have minimal biases coming from the photometric pre-selection of targets, at the cost of having a higher number of catastrophic redshift errors, leading to a sizeable fraction of line and noise interlopers in the observed sample. We classified systematics into two main classes, angular systematics that imprint spurious clustering at all redshifts, and redshift errors that mix clustering coming from different ranges along the line of sight. To first approximation, angular systematics are multiplicative in the galaxy density n_g while redshift errors give additive terms, but both are more complicated than additive or multiplicative for the density contrast δ_g .

We then focused on angular systematics, where mitigation will rely on the construction of a visibility mask in the form of a random catalogue. This can be seen as a forward model of the selection function, based on a detection model that gives the probability of detecting yet another galaxy in the observed images as a function of the S/N of its $H\alpha$ line. The random catalogue is constructed with a (bypassed) source injection algorithm, while weights (alias detection probabilities) are assigned to random galaxies and not to observed ones.

We used 50 EuclidLargeMocks (Euclid Collaboration: Monaco et al. 2025), split into ten groups of five mocks that cover the expected final EWS area, to test our ability to recover unbiased power spectrum measurements and cosmological parameters in the presence of angular systematics. This was done in the case of ideal mitigation, by applying to the mocks the same visibility mask represented by the random catalogue, and in the case of more realistic mitigation, where we took into account the uncertainty of the detection model. Weighting random galaxies implies an average density that varies on the sky, which can be taken into account in Fourier space, at the sub-percent level in the power spectrum monopole, by convolving the theory model with a window function obtained by Fourier-transforming the random catalogue itself. The resulting cosmological parameters are very consistent for ideal mitigation, as revealed by our FoB, FoM, credibility interval width, and p -value analyses (see Appendix D). The model can effectively marginalise over galaxy bias parameters, even though their values depend on galaxy selection. While photometric calibration does not seem to have a strong impact, when perturbing the detection model we stress-test parameter estimation, obtaining in one case a worse FoB and a bad p -value, together with a smaller FoM. However, such miscalibration of the detection model would be easily recognisable as a residual correlation of galaxy density contrast with nuisance maps.

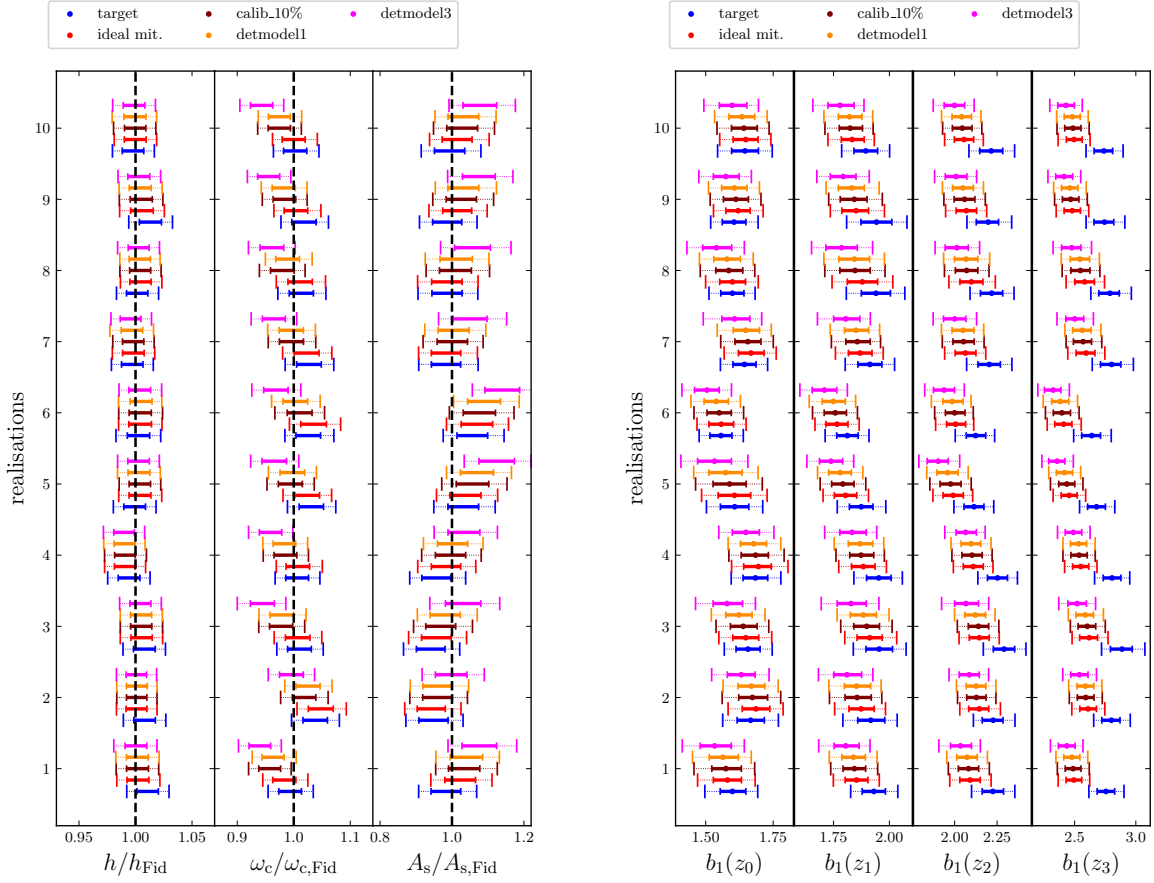


Fig. 16. Averages of the posteriors obtained for the three cosmological parameters (left) and of linear bias parameter (right) included in our analysis, for all realisations across the five cases considered. The 68% and 95% confidence intervals are overplotted as thick and thin bars, respectively.

Most of the work presented in this paper was done before the launch of *Euclid*, although it was finalised after Quick Data Release 1, and while DR1 data analysis is in preparation. Early data have fully confirmed the extent of the challenge of extracting spectra from thousands of square degrees of dispersed images, and the pipelines have quickly evolved to meet the challenge and are now ready for the first wave of cosmological results. The EDS is progressing, but, as planned in [Euclid Collaboration: Scaramella et al. \(2022\)](#), none of the surveyed sky patches is presently available at full depth. This means that the present strategy will be fully usable starting from the next data release (DR2), when at least half of the EDS will be surveyed at full depth.

Acknowledgements. The Euclid Consortium acknowledges the European Space Agency and a number of agencies and institutes that have supported the development of *Euclid*, in particular the Agenzia Spaziale Italiana, the Austrian Forschungsförderungsgesellschaft funded through BMIMI, the Belgian Science Policy, the Canadian Euclid Consortium, the Deutsches Zentrum für Luft- und Raumfahrt, the DTU Space and the Niels Bohr Institute in Denmark, the French Centre National d’Etudes Spatiales, the Fundação para a Ciência e a Tecnologia, the Hungarian Academy of Sciences, the Ministerio de Ciencia, Innovación y Universidades, the National Aeronautics and Space Administration, the National Astronomical Observatory of Japan, the Nederlandse Onderzoekschool Voor Astronomie, the Norwegian Space Agency, the Research Council of Finland, the Romanian Space Agency, the State Secretariat for Education, Research, and Innovation (SERI) at the Swiss Space Office (SSO), and the United Kingdom Space Agency. A complete and detailed list is available on the *Euclid* web site (www.euclid-ec.org/consortium/community/). This work has made use of CosmoHub, developed by PIC (maintained by IFAE and CIEMAT) in collaboration with ICE-CSIC. CosmoHub received funding from the Spanish government (MCIN/AEI/10.13039/501100011033), the EU NextGeneration/PRTR (PRTR-C17.11), and the Generalitat de Catalunya. This paper is supported by the

Italian Research Centre on High Performance Computing Big Data and Quantum Computing (ICSC) and by the Fondazione ICSC National Recovery and Resilience Plan (PNRR) Project ID CN-00000013 “Italian Research Centre on High-Performance Computing, Big Data and Quantum Computing” funded by MUR Missione 4 Componente 2 Investimento 1.4: “Potenziamento strutture di ricerca e creazione di campioni nazionali di R&S (M4C2-19)” - Next Generation EU (NGEU); by the PRIN 2022 PNRR project “Space-based cosmology with *Euclid*: the role of High-Performance Computing” (code no. P202259YAF), funded by “European Union – Next Generation EU”; and by MUR PRIN 2022 (grant 2022NY2ZRS 001). We acknowledge usage of Pleiadi computing system of INAF ([Taffoni et al. 2020](#); [Bertocco et al. 2020](#)). PM thanks Yi-Kuan Chiang for discussions on MW extinction.

References

- Adame, A. G., Aguilar, J., Ahlen, S., et al. 2025, JCAP, 07, 017
- Avila, S., Gonzalez-Perez, V., Mohammad, F. G., et al. 2020, MNRAS, 499, 5486
- Bagley, M. B., Scarlata, C., Mehta, V., et al. 2020, ApJ, 897, 98
- Bautista, J. E., Pavot, R., Vargas Magaña, M., et al. 2021, MNRAS, 500, 736
- Bertocco, S., Goz, D., Tornatore, L., et al. 2020, in Astronomical Society of the Pacific Conference Series, Vol. 527, Astronomical Data Analysis Software and Systems XXIX, ed. R. Pizzo, E. R. Deul, J. D. Mol, J. de Plaa, & H. Verkerke, 303
- Bianchi, D., Gil-Marín, H., Ruggeri, R., & Percival, W. J. 2015, MNRAS: Letters, 453, L11
- Bianchi, D., Hanif, M. M. S., Carnero Rosell, A., et al. 2025, JCAP, 04, 074
- Cagliari, M. S., Granett, B. R., Guzzo, L., et al. 2024, A&A, 689, A166
- Carrilho, P., Moretti, C., & Pourtsidou, A. 2023, JCAP, 01, 028
- Chiang, Y.-K. 2023, ApJ, 958, 118
- Chiang, Y.-K. & Ménard, B. 2019, ApJ, 870, 120
- Colavincenzo, M., Monaco, P., Sefusatti, E., & Borgani, S. 2017, JCAP, 03, 052
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
- De Mattia, A., Ruhlmann-Kleider, V., Raichoor, A., et al. 2021, MNRAS, 501, 5616

- DESI Collaboration: Aghamousa, A., Aguilar, J., Ahlen, S., et al. 2016, arXiv e-prints, arXiv:1611.00036
- Desjacques, V., Jeong, D., & Schmidt, F. 2018, *Phys. Rep.*, 733, 1
- Draine, B. T. & Li, A. 2007, *ApJ*, 657, 810
- Eggemeier, A., Camacho-Quevedo, B., Pezzotta, A., et al. 2023, *MNRAS*, 519, 2962
- Eggemeier, A., Lee, N., Scoccimarro, R., et al. 2025, *Phys. Rev. D*, 112, 063532
- Euclid Collaboration: Castander, F., Fosalba, P., Stadel, J., et al. 2025, *A&A*, 697, A5
- Euclid Collaboration: Copin, Y., Fumana, M., Mancini, C., et al. 2025, *A&A*, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554627>, arXiv:2503.15307
- Euclid Collaboration: Cropper, M., Al-Bahlawan, A., Amiaux, J., et al. 2025, *A&A*, 697, A2
- Euclid Collaboration: de la Torre, S., Marulli, F., Keihänen, E., et al. 2025, *A&A*, in press, arXiv:2501.16555
- Euclid Collaboration: Jahnke, K., Gillard, W., Schirmer, M., et al. 2025, *A&A*, 697, A3
- Euclid Collaboration: Kubik, B., Barbier, R., Clemens, J., et al. 2025, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202555777>, arXiv:2507.11326
- Euclid Collaboration: Le Brun, V., Bethermin, M., Moresco, M., et al. 2025, *A&A*, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554607>, arXiv:2503.15308
- Euclid Collaboration: McCracken, H. J., Benson, K., Dolding, C., et al. 2025, *A&A*, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554594>, arXiv:2503.15303
- Euclid Collaboration: Mellier, Y., Abdurro'uf, Acevedo Barroso, J., et al. 2025, *A&A*, 697, A1
- Euclid Collaboration: Monaco, P., Paribelli, G., Elkhatab, M. Y., et al. 2025, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202556459>, arXiv:2507.12116
- Euclid Collaboration: Polenta, G., Frailis, M., Alavi, A., et al. 2025, *A&A*, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554657>, arXiv:2503.15304
- Euclid Collaboration: Risso, I., Veropalumbo, A., Branchini, E., et al. 2025, *A&A*, submitted, arXiv:2505.04688
- Euclid Collaboration: Romelli, E., Kümmel, M., Dole, H., et al. 2025, *A&A*, in press (Euclid Q1 SI), <https://doi.org/10.1051/0004-6361/202554586>, arXiv:2503.15305
- Euclid Collaboration: Scaramella, R., Amiaux, J., Mellier, Y., et al. 2022, *A&A*, 662, A112
- Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, *ApJ*, 426, 23
- Ferreras, I., Tress, M., Bruzual, G., et al. 2021, *MNRAS*, 505, 283
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
- Green, G. M., Zhang, X., & Zhang, R. 2025, *ApJ*, 988, 5
- Hirata, C. M. 2009, *MNRAS*, 399, 1074
- Ivanov, M. M. 2024, in *Handbook of Quantum Gravity* (Springer), 213–260
- Karim, T., Singh, S., Rezaie, M., et al. 2025, *JCAP*, 02, 045
- Kitaura, F.-S., Rodríguez-Torres, S., Chuang, C.-H., et al. 2016, *MNRAS*, 456, 4156
- Krolewski, A., Yu, J., Ross, A. J., et al. 2025, *JCAP*, 01, 147
- Lamman, C., Eisenstein, D., Aguilar, J. N., et al. 2023, *MNRAS*, 522, 117
- Lange, J. U. 2023, *MNRAS*, 525, 3181
- Lewis, A. 2025, *JCAP*, 08, 025
- Lim, S., Hill, R., Scott, D., et al. 2023, *MNRAS*, 525, 1443
- Marković, K., Percival, W. J., Scodreggio, M., et al. 2017, *MNRAS*, 467, 3677
- Monaco, P., Di Dio, E., & Sefusatti, E. 2019, *JCAP*, 04, 023
- Monaco, P., Sefusatti, E., Borgani, S., et al. 2013, *MNRAS*, 433, 2389
- Monaco, P., Theuns, T., & Taffoni, G. 2002, *MNRAS*, 331, 587
- Munari, E., Monaco, P., Sefusatti, E., et al. 2017, *MNRAS*, 465, 4658
- Obulj, A., Percival, W. J., & Dalal, N. 2020, *JCAP*, 10, 058
- Pezzotta, A., Eggemeier, A., Gambardella, G., et al. 2025, *Phys. Rev. D*, 112, 023520
- Planck Collaboration: Abergel, A., Ade, P. A. R., Aghanim, N., et al. 2014, *A&A*, 571, A11
- Planck Collaboration: Ade, P. A. R., Aghanim, N., Alves, M. I. R., et al. 2016, *A&A*, 586, A132
- Pozzetti, L., Hirata, C. M., Geach, J. E., et al. 2016, *A&A*, 590, A3
- Reid, B., Ho, S., Padmanabhan, N., et al. 2016, *MNRAS*, 455, 1553
- Rezaie, M., Ross, A. J., Seo, H.-J., et al. 2021, *MNRAS*, 506, 3439
- Rosado-Marín, A. J., Ross, A. J., Seo, H., et al. 2025, *JCAP*, 07, 018
- Ross, A. J., Bautista, J., Tojeiro, R., et al. 2020, *MNRAS*, 498, 2354
- Ross, A. J., Percival, W. J., Sánchez, A. G., et al. 2012, *MNRAS*, 424, 564
- Schlafly, E. F., Meisner, A. M., Stutz, A. M., et al. 2016, *ApJ*, 821, 78
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
- Scoccimarro, R. 2004, *Phys. Rev. D*, 70, 083007
- Scoccimarro, R. 2015, *Phys. Rev. D*, 92, 083532
- Sefusatti, E., Crocce, M., Scoccimarro, R., & Couchman, H. M. P. 2016, *MNRAS*, 460, 3624
- Taffoni, G., Becciani, U., Garilli, B., et al. 2020, in *Astronomical Society of the Pacific Conference Series*, Vol. 527, *Astronomical Data Analysis Software and Systems XXIX*, ed. R. Pizzo, E. R. Deul, J. D. Mol, J. de Plaa, & H. Verkouter, 307
- Tegmark, M. 1997, *Phys. Rev. Lett.*, 79, 3806
- Yamamoto, K., Nakamichi, M., Kamino, A., et al. 2006, *PASJ*, 58, 93
- ¹ Dipartimento di Fisica - Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34131 Trieste, Italy
 - ² INAF-Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
 - ³ INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste TS, Italy
 - ⁴ IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34151 Trieste, Italy
 - ⁵ ICSC - Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing, Via Magnanelli 2, Bologna, Italy
 - ⁶ INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy
 - ⁷ Minnesota Institute for Astrophysics, University of Minnesota, 116 Church St SE, Minneapolis, MN 55455, USA
 - ⁸ Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
 - ⁹ California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, USA
 - ¹⁰ INAF-IASF Milano, Via Alfonso Corti 12, 20133 Milano, Italy
 - ¹¹ Aix-Marseille Université, CNRS, CNES, LAM, Marseille, France
 - ¹² INFN-Padova, Via Marzolo 8, 35131 Padova, Italy
 - ¹³ Universität Bonn, Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
 - ¹⁴ Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
 - ¹⁵ INFN-Sezione di Milano, Via Celoria 16, 20133 Milano, Italy
 - ¹⁶ Institut d'Astrophysique de Paris, UMR 7095, CNRS, and Sorbonne Université, 98 bis boulevard Arago, 75014 Paris, France
 - ¹⁷ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA
 - ¹⁸ Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, Chiba 277-8583, Japan
 - ¹⁹ Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, 40129 Bologna, Italy
 - ²⁰ INAF-Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
 - ²¹ Dipartimento di Fisica e Astronomia "G. Galilei", Università di Padova, Via Marzolo 8, 35131 Padova, Italy
 - ²² Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
 - ²³ Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
 - ²⁴ Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada
 - ²⁵ INFN-Sezione di Genova, Via Dodecaneso 33, 16146, Genova, Italy
 - ²⁶ Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, 85748 Garching, Germany
 - ²⁷ Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada
 - ²⁸ Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA
 - ²⁹ ESAC/ESA, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, Spain
 - ³⁰ SISSA, International School for Advanced Studies, Via Bonomea 265, 34136 Trieste TS, Italy
 - ³¹ Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93/2, 40129 Bologna, Italy
 - ³² INFN-Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
 - ³³ Dipartimento di Fisica, Università di Genova, Via Dodecaneso 33, 16146, Genova, Italy

- ³⁴ Department of Physics "E. Pancini", University Federico II, Via Cinthia 6, 80126, Napoli, Italy
- ³⁵ INAF-Osservatorio Astronomico di Capodimonte, Via Moiriello 16, 80131 Napoli, Italy
- ³⁶ Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal
- ³⁷ Faculdade de Ciências da Universidade do Porto, Rua do Campo de Alegre, 4150-007 Porto, Portugal
- ³⁸ European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany
- ³⁹ Dipartimento di Fisica, Università degli Studi di Torino, Via P. Giuria 1, 10125 Torino, Italy
- ⁴⁰ INFN-Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy
- ⁴¹ INAF-Osservatorio Astrofisico di Torino, Via Osservatorio 20, 10025 Pino Torinese (TO), Italy
- ⁴² European Space Agency/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
- ⁴³ Leiden Observatory, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands
- ⁴⁴ INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00078 Monteporzio Catone, Italy
- ⁴⁵ INFN-Sezione di Roma, Piazzale Aldo Moro, 2 - c/o Dipartimento di Fisica, Edificio G. Marconi, 00185 Roma, Italy
- ⁴⁶ Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Avenida Complutense 40, 28040 Madrid, Spain
- ⁴⁷ Port d'Informació Científica, Campus UAB, C. Albareda s/n, 08193 Bellaterra (Barcelona), Spain
- ⁴⁸ Institute for Theoretical Particle Physics and Cosmology (TTK), RWTH Aachen University, 52056 Aachen, Germany
- ⁴⁹ Deutsches Zentrum für Luft- und Raumfahrt e. V. (DLR), Linder Höhe, 51147 Köln, Germany
- ⁵⁰ Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain
- ⁵¹ Institut d'Estudis Espacials de Catalunya (IEEC), Edifici RDIT, Campus UPC, 08860 Castelldefels, Barcelona, Spain
- ⁵² INFN section of Naples, Via Cinthia 6, 80126, Napoli, Italy
- ⁵³ Dipartimento di Fisica e Astronomia "Augusto Righi" - Alma Mater Studiorum Università di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
- ⁵⁴ Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
- ⁵⁵ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
- ⁵⁶ Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
- ⁵⁷ European Space Agency/ESRIN, Largo Galileo Galilei 1, 00044 Frascati, Roma, Italy
- ⁵⁸ Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne, F-69100, France
- ⁵⁹ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, 08028 Barcelona, Spain
- ⁶⁰ Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig de Luíís Companys 23, 08010 Barcelona, Spain
- ⁶¹ Institut de Ciències de l'Espai (IEEC-CSIC), Campus UAB, Carrer de Can Magrans, s/n Cerdanyola del Vallès, 08193 Barcelona, Spain
- ⁶² UCB Lyon 1, CNRS/IN2P3, IUF, IP2I Lyon, 4 rue Enrico Fermi, 69622 Villeurbanne, France
- ⁶³ Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
- ⁶⁴ Université Paris-Saclay, CNRS, Institut d'astrophysique spatiale, 91405, Orsay, France
- ⁶⁵ Aix-Marseille Université, CNRS/IN2P3, CPPM, Marseille, France
- ⁶⁶ INAF-Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere, 100, 00100 Roma, Italy
- ⁶⁷ University Observatory, LMU Faculty of Physics, Scheinerstr. 1, 81679 Munich, Germany
- ⁶⁸ INAF-Osservatorio Astronomico di Padova, Via dell'Osservatorio 5, 35122 Padova, Italy
- ⁶⁹ Universitäts-Sternwarte München, Fakultät für Physik, Ludwig-Maximilians-Universität München, Scheinerstr. 1, 81679 München, Germany
- ⁷⁰ Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, 0315 Oslo, Norway
- ⁷¹ Felix Hormuth Engineering, Goethestr. 17, 69181 Leimen, Germany
- ⁷² Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
- ⁷³ Cosmic Dawn Center (DAWN), Denmark
- ⁷⁴ Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
- ⁷⁵ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
- ⁷⁶ Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- ⁷⁷ Department of Physics and Helsinki Institute of Physics, Gustaf Hållströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- ⁷⁸ Université de Genève, Département de Physique Théorique and Centre for Astroparticle Physics, 24 quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland
- ⁷⁹ Department of Physics, P.O. Box 64, University of Helsinki, 00014 Helsinki, Finland
- ⁸⁰ Helsinki Institute of Physics, Gustaf Hållströmin katu 2, University of Helsinki, 00014 Helsinki, Finland
- ⁸¹ Laboratoire d'étude de l'Univers et des phénomènes eXtremes, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 92190 Meudon, France
- ⁸² SKAO, Jodrell Bank, Lower Withington, Macclesfield SK11 9FT, UK
- ⁸³ Centre de Calcul de l'IN2P3/CNRS, 21 avenue Pierre de Coubertin 69627 Villeurbanne Cedex, France
- ⁸⁴ University of Applied Sciences and Arts of Northwestern Switzerland, School of Computer Science, 5210 Windisch, Switzerland
- ⁸⁵ Department of Physics, Institute for Computational Cosmology, Durham University, South Road, Durham, DH1 3LE, UK
- ⁸⁶ Université Paris Cité, CNRS, Astroparticule et Cosmologie, 75013 Paris, France
- ⁸⁷ CNRS-UCB International Research Laboratory, Centre Pierre Binétruy, IRL2007, CPB-IN2P3, Berkeley, USA
- ⁸⁸ Institut d'Astrophysique de Paris, 98bis Boulevard Arago, 75014, Paris, France
- ⁸⁹ Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
- ⁹⁰ Telespazio UK S.L. for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ⁹¹ Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
- ⁹² DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 155, 2200 Copenhagen, Denmark
- ⁹³ Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
- ⁹⁴ Space Science Data Center, Italian Space Agency, via del Politecnico snc, 00133 Roma, Italy
- ⁹⁵ Centre National d'Etudes Spatiales – Centre spatial de Toulouse, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9, France
- ⁹⁶ Institute of Space Science, Str. Atomistilor, nr. 409 Măgurele, Ilfov, 077125, Romania
- ⁹⁷ Institut für Theoretische Physik, University of Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
- ⁹⁸ Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
- ⁹⁹ Université St Joseph; Faculty of Sciences, Beirut, Lebanon
- ¹⁰⁰ Departamento de Física, FCFM, Universidad de Chile, Blanco Encalada 2008, Santiago, Chile

- ¹⁰¹ Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
- ¹⁰² Satlantís, University Science Park, Sede Bld 48940, Leioa-Bilbao, Spain
- ¹⁰³ Department of Physics, Royal Holloway, University of London, Surrey TW20 0EX, UK
- ¹⁰⁴ Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, PT1749-016 Lisboa, Portugal
- ¹⁰⁵ Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências, Universidade de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
- ¹⁰⁶ Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK
- ¹⁰⁷ Cosmic Dawn Center (DAWN)
- ¹⁰⁸ Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen, Denmark
- ¹⁰⁹ Universidad Politécnica de Cartagena, Departamento de Electrónica y Tecnología de Computadoras, Plaza del Hospital 1, 30202 Cartagena, Spain
- ¹¹⁰ INFN-Bologna, Via Irnerio 46, 40126 Bologna, Italy
- ¹¹¹ Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
- ¹¹² Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Giuseppe Saragat 1, 44122 Ferrara, Italy
- ¹¹³ INAF, Istituto di Radioastronomia, Via Piero Gobetti 101, 40129 Bologna, Italy
- ¹¹⁴ Astronomical Observatory of the Autonomous Region of the Aosta Valley (OAVdA), Loc. Lignan 39, I-11020, Nus (Aosta Valley), Italy
- ¹¹⁵ Université Côte d’Azur, Observatoire de la Côte d’Azur, CNRS, Laboratoire Lagrange, Bd de l’Observatoire, CS 34229, 06304 Nice cedex 4, France
- ¹¹⁶ Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH, UK
- ¹¹⁷ Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, 00185 Roma, Italy
- ¹¹⁸ Aurora Technology for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹¹⁹ ICL, Junia, Université Catholique de Lille, LITL, 59000 Lille, France
- ¹²⁰ Instituto de Física Teórica UAM-CSIC, Campus de Cantoblanco, 28049 Madrid, Spain
- ¹²¹ CERCA/ISO, Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
- ¹²² Technical University of Munich, TUM School of Natural Sciences, Physics Department, James-Frank-Str. 1, 85748 Garching, Germany
- ¹²³ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
- ¹²⁴ Laboratoire Univers et Théorie, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, 92190 Meudon, France
- ¹²⁵ Departamento de Física Fundamental, Universidad de Salamanca, Plaza de la Merced s/n. 37008 Salamanca, Spain
- ¹²⁶ Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
- ¹²⁷ Department of Physics & Astronomy, University of California Irvine, Irvine CA 92697, USA
- ¹²⁸ Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
- ¹²⁹ Departamento Física Aplicada, Universidad Politécnica de Cartagena, Campus Muralla del Mar, 30202 Cartagena, Murcia, Spain
- ¹³⁰ Instituto de Física de Cantabria, Edificio Juan Jordá, Avenida de los Castros, 39005 Santander, Spain
- ¹³¹ INFN, Sezione di Lecce, Via per Arnesano, CP-193, 73100, Lecce, Italy
- ¹³² Department of Mathematics and Physics E. De Giorgi, University of Salento, Via per Arnesano, CP-I93, 73100, Lecce, Italy
- ¹³³ INAF-Sezione di Lecce, c/o Dipartimento Matematica e Fisica, Via per Arnesano, 73100, Lecce, Italy
- ¹³⁴ CEA Saclay, DFR/IRFU, Service d’Astrophysique, Bat. 709, 91191 Gif-sur-Yvette, France
- ¹³⁵ Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
- ¹³⁶ Department of Computer Science, Aalto University, PO Box 15400, Espoo, FI-00 076, Finland
- ¹³⁷ Instituto de Astrofísica de Canarias, E-38205 La Laguna; Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
- ¹³⁸ Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
- ¹³⁹ Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing (GCCL), 44780 Bochum, Germany
- ¹⁴⁰ Department of Physics and Astronomy, Vesilinnantie 5, University of Turku, 20014 Turku, Finland
- ¹⁴¹ Serco for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹⁴² ARC Centre of Excellence for Dark Matter Particle Physics, Melbourne, Australia
- ¹⁴³ Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia
- ¹⁴⁴ Department of Physics and Astronomy, University of the Western Cape, Bellville, Cape Town, 7535, South Africa
- ¹⁴⁵ DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
- ¹⁴⁶ Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA, UK
- ¹⁴⁷ Department of Astrophysics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
- ¹⁴⁸ Department of Physics, Centre for Extragalactic Astronomy, Durham University, South Road, Durham, DH1 3LE, UK
- ¹⁴⁹ IRFU, CEA, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex, France
- ¹⁵⁰ Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, Stockholm, SE-106 91, Sweden
- ¹⁵¹ Astrophysics Group, Blackett Laboratory, Imperial College London, London SW7 2AZ, UK
- ¹⁵² Univ. Grenoble Alpes, CNRS, Grenoble INP, LPSC-IN2P3, 53, Avenue des Martyrs, 38000, Grenoble, France
- ¹⁵³ INAF-Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125, Firenze, Italy
- ¹⁵⁴ Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
- ¹⁵⁵ HE Space for European Space Agency (ESA), Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, 28692 Madrid, Spain
- ¹⁵⁶ INAF - Osservatorio Astronomico d’Abruzzo, Via Maggini, 64100, Teramo, Italy
- ¹⁵⁷ Theoretical astrophysics, Department of Physics and Astronomy, Uppsala University, Box 516, 751 37 Uppsala, Sweden
- ¹⁵⁸ Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
- ¹⁵⁹ Mathematical Institute, University of Leiden, Einsteinweg 55, 2333 CA Leiden, The Netherlands
- ¹⁶⁰ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- ¹⁶¹ Department of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ 08544, USA
- ¹⁶² Space physics and astronomy research unit, University of Oulu, Pentti Kaitekan katu 1, FI-90014 Oulu, Finland
- ¹⁶³ Institut de Physique Théorique, CEA, CNRS, Université Paris-Saclay 91191 Gif-sur-Yvette Cedex, France
- ¹⁶⁴ Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010, New York, NY, USA

Appendix A: Response of galaxy density to nuisance maps

If the galaxy bias is independent of luminosity, and the galaxy (H α line) luminosity function $\Phi(f|z)$ is universal (meaning that its shape is the same in all space location, independent of galaxy density), it is useful to define the effective flux limit f_{lim} of the survey such that the number density of galaxies brighter than this flux is equal to the number density of observed galaxies,

$$f_{\text{lim}}(\hat{n}, z, \{N_i\}) : \int_{f_{\text{lim}}}^{\infty} \Phi(f|z) df = \int_0^{\infty} C(f, z, \{N_i\}) \Phi(f|z) df, \quad (\text{A.1})$$

where, as in Sect. 2, $\hat{n}$ gives the angular position on the sky, $\{N_i\}$ is the set of nuisance maps and C is the completeness function, alias MVM. This flux limit acquires a dependence on sky coordinates $\hat{n}$ through the nuisance maps $\{N_i\}$. The observed number density of correct galaxies is then

$$n_{\text{oc}} = (1 + \delta_g) \int_{f_{\text{lim}}}^{\infty} \Phi(f|z) df, \quad (\text{A.2})$$

while the number density of target galaxies n_t , given in Eq. (3), is expressed by the same integral starting from f_0 . The relative residual $\mathcal{R}$ of n_{oc} with respect to n_t , can be written as

$$\mathcal{R} := \frac{n_{\text{oc}} - n_t}{n_t} = -\frac{1}{\bar{n}_t} \int_{f_0}^{f_{\text{lim}}} \Phi(f|z) df. \quad (\text{A.3})$$

This shows that a variation of flux limit gives rise to a multiplicative systematic effect on the density.

As discussed in Monaco et al. (2019), it is convenient to Taylor-expand the integral in flux limit from f_0 to f_{lim} ; because the modulations of f_{lim} are not necessarily expected to be small, the expansion cannot be limited to the first order. We call $\epsilon := (f_{\text{lim}} - f_0)/f_0$ the relative variation of the flux limit with respect to the fiducial one, and define two redshift-dependent coefficients that are determined by the shape of number counts: $S_\Phi := f_0 \Phi(f_0)/\bar{n}_t$ and the logarithmic slope $\alpha_\Phi := \partial \ln \Phi / \partial \ln f(f_0)$. By expanding the integral to second order, the relative residual becomes

$$\mathcal{R} = -S_\Phi \left(\epsilon + \frac{1}{2} \alpha_\Phi \epsilon^2 + O(\epsilon^3) \right). \quad (\text{A.4})$$

To compute the response of density residuals to systematics we need to relate f_{lim} to the nuisance maps. We call the exposure time $\mathcal{T}$ and assume that the noise map represents the variance of the pixel-level counts, and is the sum of a series of noise terms $\mathcal{N}_i$ (with calligraphic font to distinguish this from the list of nuisance maps), so that the total noise is $\mathcal{N} := \Sigma \mathcal{N}_i$. For the reddening $E(B-V)$, we define the relative decrease of the observed flux as

$$\mathcal{E} := 10^{-0.4 R(z) E(B-V)}, \quad (\text{A.5})$$

where $R(z) = R_\lambda [\lambda_{\text{H}\alpha}(1+z)]$. Then the S/N S of the H α line can be related to its flux as

$$S = \kappa_1 f \mathcal{E} \sqrt{\frac{\mathcal{T}}{\mathcal{N}}}, \quad (\text{A.6})$$

where κ_1 is a suitable constant that translates the right-hand side of that equation into a dimensionless number. The completeness function C in Eq. (A.1) can be expressed as the marginalised detection probability P_{det} (Eq. 22), which is a function of S alone. This function goes from 0 to 1 around the value S_0 , so it is clear that

$$f_{\text{lim}} = \kappa_2(z) \frac{S_0}{\kappa_1 \mathcal{E}} \sqrt{\frac{\mathcal{N}}{\mathcal{T}}}, \quad (\text{A.7})$$

where $\kappa_2(z)$ will be determined by the solution of the integral in Eq. (A.1) and will depend on the shape of the number counts $\Phi(f|z)$, which is again redshift dependent.

The response of galaxy density to a specific nuisance can be computed as the partial derivative of $\mathcal{R}$ with respect to that nuisance. For a noise term $\mathcal{N}_i$,

$$\frac{\partial \mathcal{R}}{\partial \mathcal{N}_i} = \frac{\partial \mathcal{R}}{\partial \epsilon} \frac{\partial \epsilon}{\partial \mathcal{N}_i} = -\frac{1}{2} S_\Phi (1 + \alpha_\Phi \epsilon + O(\epsilon^2)) \frac{1 - \epsilon}{\mathcal{N}}, \quad (\text{A.8})$$

where the (total) noise $\mathcal{N}$ is computed at the unperturbed value. The response to a variation in exposure time can be computed similarly,

$$\frac{\partial \mathcal{R}}{\partial \mathcal{T}} = \frac{\partial \mathcal{R}}{\partial \epsilon} \frac{\partial \epsilon}{\partial \mathcal{T}} = \frac{1}{2} S_\Phi (1 + \alpha_\Phi \epsilon + O(\epsilon^2)) \frac{1 - \epsilon}{\mathcal{T}}, \quad (\text{A.9})$$

while the response to a variation in reddening (in $\mathcal{E}$) would be

$$\frac{\partial \mathcal{R}}{\partial \mathcal{E}} = \frac{\partial \mathcal{R}}{\partial \epsilon} \frac{\partial \epsilon}{\partial \mathcal{E}} = S_\Phi (1 + \alpha_\Phi \epsilon + O(\epsilon^2)) \frac{1 - \epsilon}{\mathcal{E}}. \quad (\text{A.10})$$

Clearly the response to a nuisance is generally non-linear and depends on the shape of the number counts and through them on redshift. This is why an approach based on mock catalogues, paired with a random catalogue created from a (bypassed) source injection procedure, is to be preferred, since it is able to provide unbiased and robust measurements. A mitigation based on response to a set of potential nuisance fields would be equivalent only if the response is properly modelled.

An important assumption that we make here is that the physical process including a systematic effect on the galaxy density is statistically independent of the galaxy density field itself. This is not always true: (i) the map (based on FIR observations from *Planck*) used to correct extinction from MW dust may have contamination from the cosmic infrared background (CIB, see Chian & Ménard 2019), so the correction for extinction may contain a (spurious) term that couples with the density field (this is discussed below in Appendix B.3); moreover, (ii) the effect of confusion in the spectral images is independent of the density field only as long as the main contaminating sources, which are bright enough to have a significant continuum in the dispersed image, are at $z < 0.9$. Such cases, if not deemed to be negligible, deserve a dedicated study.

Luminosity-dependent bias makes the scheme even more complicated. We assume again that we can treat the completeness function as a Heaviside θ -function stepping from 0 to 1 at f_{lim} . We also call $\beta_1 := \partial b_1 / \partial f$ the linear bias parameter of galaxies in a small range of flux (b_1 is reserved for the bias of the total galaxy sample), and we assume that it depends on flux, $\beta_1 = \beta_1(f|z)$. Then

$$n_{\text{oc}} = \int_{f_{\text{lim}}}^{\infty} [1 + \beta_1(f) \delta] \Phi(f|z) df. \quad (\text{A.11})$$

The Taylor-expanded expression of $n_{\text{oc}} - n_t$ will then be

$$n_{\text{oc}} - n_t = f_0 \Phi(f_0) \left\{ [1 + \beta_1(f_0) \delta] \epsilon + \frac{1}{2} \alpha_\Phi [1 + \beta_1(f_0) \delta] \epsilon^2 + \frac{1}{2} \beta_1(f_0) \alpha_\beta \delta \epsilon^2 + O(\epsilon^3) \right\}, \quad (\text{A.12})$$

where we have defined $\alpha_\beta := \partial \ln \beta_1 / \partial \ln f$ as the logarithmic slope of the $\beta_1(f)$ function.

Luminosity-dependent bias has two main effects: to add a further term of order ϵ^2 that is proportional to α_β , and to break the simple dependence on matter density that allows us to divide this quantity by n_t keeping the expression treatable. This confirms that the response of relative residuals to nuisances given in Eqs. (A.8) to (A.10) is only an approximation.

Appendix B: Entry points of systematics

We present here a more detailed discussion of the SGS pipeline, described in Sect. 3.2 (see also Fig. 2). We highlight the entry points, which are the steps where systematics may contaminate the galaxy catalogue. This discussion is at the basis of several of the entries of Table 3.

Appendix B.1: The photometric sample

Slitless spectroscopy allows us, in principle, to define a sample purely selected in line flux. This is not exactly true in the *Euclid* pipeline analysis, since the spectra will be extracted starting from a list of photometric sources. The risk is then to contaminate the spectroscopic sample with fluctuations in survey depth of the photometric sample; although space has a definite advantage over Earth in the cleanness of photometric selection, the characterisation of the depth of a photometric sample obtained from the merging of VIS and NISP photometric catalogues is tricky, since images from the two instruments are affected at a varying degree by many of the systematics that will be described below. Moreover, the deblending algorithm of NISP sources adds some complexity to the process, due to the poorer PSF in the NIR and to persistence effects. However, photometry goes far deeper than spectroscopy, and this is already evident from Table 2, where the ratio between $H_E < 24$ sources and ELGs is of order 50:1. If the number of ELGs with photometric fluxes near or below the threshold is negligible, fluctuations of this threshold will not influence the resulting catalogue. The number of ELGs with $H\alpha$ line flux above f_0 and with $H_E > 24$ was estimated by Bagley et al. (2020) to be around 2%, so a modulation of this flux limit would have an influence on a few percent of the sample, inducing negligible effects on the galaxy density. This claim is a bit optimistic, however, since the threshold for line detection is not a Heaviside theta function of the true line flux, so a significant part of the survey will be at lower fluxes, where the probability of being hosted in an $H_E \geq 24$ galaxy is higher.

Photometric information, both from *Euclid* and ground-based images, can be used by SEL to significantly improve our ability to reduce both line and (especially) noise interlopers, thanks to the broader spectral coverage (Cagliari et al. 2024).

However, this is done at the cost of introducing subtle systematics, especially when ground-based data are used; in this case a pure forward model of systematics would be impossible. Systematics in the ground-based photometry may modulate on the sky the effectiveness of the removal of interlopers; this modulation would impact on a ten-percent-level contribution, so its effect would possibly be small. However, it would also remove a significant fraction of faint correct galaxies, giving a potentially larger effect. Careful tests will be needed when assessing the effects of any photometric-based selection. A similar concern applies to usage of photometric redshifts to inform the choice among different redshift solutions, aimed at limiting interlopers.

Photometry is used by SIR to decontaminate the spectra of ELGs from those of brighter, continuum-dominated sources like stars and nearby galaxies. Here a degradation of photometry could bias spectral decontamination and affect galaxy spectra in a subtle way, but such bright sources will hardly be affected by modulations of a flux limit that is at least four magnitudes fainter.

Another issue of photometry that can affect the spectroscopic sample is deblending of sources, which can fail in crowded fields. Deblending will mostly rely on VIS images, which have a better PSF, but very red sources may be missing in VIS, and occasional problems of some VIS pointings may worsen the ability of deblending. Under-deblending will result in a missing source in the photometric catalogue, and in a mis-centring of the main blended galaxy (which will likely be at low redshift), impacting on sample completeness in a way that is not represented by the random catalogue. Conversely, over-deblending will result in fake sources, contributing to noise interlopers; in some cases it may induce a double-counting of ELGs, but this would only happen to very near and extended galaxies. These effects will strongly depend on source surface density, but will likely be negligible with respect to confusion in spectroscopic images.

Appendix B.2: Astrophysical foregrounds

Background noise in the dispersed images will have an important contribution from astrophysical foregrounds, mostly zodiacal light and, to a lesser extent, nearby galaxies. We do not consider here stars in this class because they mostly impact through stray light and spectral confusion, which will be addressed below. Zodiacal light is, along with MW extinction, the main astrophysical killer of detection probability, and one of the main drivers of the EWS design. However, its behaviour is very smooth, both in angular coordinates and in SED, and it is moreover very predictable, so we expect to be able to remove its impact through the random catalogue. As a consequence, in Table 3 zodiacal light has a very high impact but a very low risk.

Nearby galaxies have a much smaller angular footprint, but their collective area can hide galaxies in the background. In principle VMSP will take account of their obscuration, unless they are due to a failure of the MER deblending algorithm: as an example, VMSP may place a random galaxy in the outskirts of a nearby galaxy, and the $H\alpha$ line may fall in a region where it is detectable, but the deblending algorithm may fail to recognise it as a target galaxy and not as a star-forming region of the foreground object; as a consequence VMSP would inject a random galaxy where a true one would go undetected, thus creating a fake underdensity. This may be recognised by correlating the galaxy density field with a map of nearby galaxies. However, this effect is expected to be small.

Appendix B.3: MW extinction

Light coming from extragalactic sources is absorbed or scattered by dust in the MW in a way that is very hard to control. Galactic dust is heated to a temperature of around 20 K by the local radiation field of stars, and it re-emits this energy in the FIR as a modified blackbody, peaking at $150\ \mu\text{m}$. This emission is an important foreground for CMB observations, and it can be effectively separated from the CMB signal due to its different spectrum. The most accurate measurement of dust emissivity has been presented by Planck Collaboration in 2013 (P13), which thanks to its high-frequency instrument HFI, augmented with IRAS data at $100\ \mu\text{m}$, has been able to measure all-sky dust radiance (the integral in frequency ν of the monochromatic surface brightness I_ν). In that paper an SED was constructed for each $5'$ sky pixel of a healpix tessellation, and dust emissivity in four bands was fit by a single modified blackbody, characterised by a temperature T , an index β , and an optical depth τ_{353} at 353 GHz. Dust radiance was found to be robust against the known degeneracy of T and β . The Planck Collaboration provided best fit parameters and dust radiance as healpix maps, along with their uncertainty and with an estimation of induced reddening, $E(B - V)$.

While the measurement of dust radiance can be considered robust, its relation with reddening and extinction is much more uncertain. In P13 it was assumed that radiance is correlated with reddening, and the correlation coefficient was found using a set of 54 492 quasars from the Sloan Digital Sky Survey, for which reddening was estimated by comparing their spectra with a template. This procedure is valid as long as dust physical properties and radiation field are constant in the volume of the dust disc that intersects the survey line of sight. We can estimate the size of this volume by assuming a height of 100 pc for the dust disc and a zone of avoidance at Galactic declination $b < 20^\circ$; the intersection is thus a cylinder of diameter $\lesssim 500$ pc. A gradient in dust composition, temperature or in stellar radiation field in this volume will influence the coefficient that translates dust radiance to reddening.

In P15, a more extended analysis of dust emission was presented, adding more bands at shorter wavelengths from WISE data and fitting the pixel-level SEDs with the model of [Draine & Li \(2007\)](#), to produce more physically-motivated estimates of dust masses and V -band extinction, A_V . However, extinction values turned out to be off by a factor of two, and were recalibrated by using a more extended sample of SDSS quasars. Considering the high quality of the dust model used, such a recalibration casts legitimate doubt on the level of control that we can achieve on the physics behind dust extinction.

Calling $A_V = R_V E(B - V)$ the extinction in the Johnson V band, where R_V is a scaling parameter to which we assign the standard value of $R_V = 3.086$, dust extinction at wavelength λ is computed as $A_\lambda = k(\lambda) E(B - V)$, where $k(\lambda)$ is an extinction curve whose integral over the V -band wavelength range gives R_V . There is ample evidence that both R_V and the extinction curve are not universal but change with the line of sight (e.g., [Schlafly et al. 2016](#); [Ferreras et al. 2021](#); [Green et al. 2025](#)). Then, even a perfect knowledge of reddening $E(B - V)$ does not guarantee a precise knowledge of extinction at the wavelengths of the red grism. Moreover, a straightforward propagation of the nominal error in the extinction would not represent the true uncertainty, since for a two-point clustering measurement, and even more for its covariance, the important quantity is not the 1-point PDF of the error but its correlations on the sky. This point was addressed in [Monaco et al. \(2019\)](#), where a more extended discussion can be found.

A known obstacle to the measurement of dust properties is given by the already mentioned cosmic infrared background, or CIB, due to FIR emission of background galaxies. If not properly subtracted, this background can modulate the measurement of dust temperature T , especially through its degeneracy with β . While the CIB monopole is easy to be characterised and subtracted, its anisotropies due to galaxy clustering can affect dust extinction estimates. CIB contamination of dust maps was carefully assessed by [Chiang & Ménard \(2019\)](#) and [Chiang \(2023\)](#), who proposed a data-driven, map-level reconstruction method to subtract CIB contribution from dust radiance maps. If not properly subtracted, CIB provides a very worrying systematic effect, as it correlates with the cosmological signal δ_g we are measuring, breaking the basic assumption that C and δ_g are statistically independent.

A second problem is related to the fact that dust emission has structure on small scales that are not sampled at a resolution of about $5'$; such intra-field variations are not mitigated by applying extinction from the *Planck* maps, or from maps at similar resolution. To assess the impact of this effect, it is useful to analyse the angular power spectrum of dust reddening, or equivalently of a random catalogue subject to extinction. This spectrum is shown in figures 3 and 6 of [Monaco et al. \(2019\)](#), and is well represented by a decreasing power law with slope -1 . At the same time, the galaxy angular power spectrum is raising and dominates over the dust power spectrum at $\ell \gtrsim 100$. An extrapolation of this power law to smaller scales would give a sub-dominant contribution to galaxy clustering from any intra-field dust. This is confirmed by the analysis presented in Sect. 5: MW extinction adds spurious power on very large scales, so a small-scale contribution due to an extrapolation of the dust power spectrum would give a negligible effect.

This argument can be extended to estimate the impact of CIB anisotropies: a strong contamination from galaxies would give rise to a growing power spectrum at high ℓ that is not observed. At the same time, a contamination, as quantified in [Chiang & Ménard \(2019\)](#), which does not perturb much the angular power spectrum of dust adds power on scales that have little effect on the measurement of galaxy clustering. However, on the long term mitigation of MW extinction should be done using maps where CIB is properly subtracted.

MW extinction is expected to give rise to fake power on the largest scales, so it is a killer systematic effect for any science case based on scales at or below (in k -space) the peak of the power spectrum, like scale-dependent bias from primordial non-Gaussianity or relativistic effects near the horizon. As an example, [Karim et al. \(2025\)](#) found that parameter estimation in the cross correlation of DESI and CMB lensing data depends on the assumed extinction map. Nonetheless, there is no reason at the moment not to use a reddening map to correct for extinction, although a straightforward propagation of the quoted errors cannot represent its true uncertainty. Following [Monaco et al. \(2019\)](#), the most effective mitigation of this uncertainty is to consider the residuals after correction with a map as an unknown residual systematic that must be controlled a posteriori, with methods based on the cross correlation of different redshift bins.

As a final remark, dust has emission down to NIR wavelengths ([Lim et al. 2023](#)), so it can contribute to the general background at low Galactic latitudes. This contribution is expected to be subdominant with respect to zodiacal light and stray light, and will sum to the measured background that VMSP uses to construct the random catalogue. Its intra-field variations will be taken into account by the usage of a local measurement of noise in the detection model.

Appendix B.4: Systematics in NISP dispersed images

The first crucial step in the detection of spectroscopic galaxies is performed by SIR, which extracts, decontaminates, and coadds the spectra from the four NISP dispersed dithers. Pixel-level noise of the signal is later read by VMSP to compute P_{det} using the detection model (Sect. 3.3). NISP dispersed images are affected by systematics that are in common with all exposures: (i) detector noise; (ii) bad pixels; (iii) masked regions around saturated stars; (iv) cosmic rays; (v) ghosts from saturated objects; (vi) zeroth- and second-order spectra from bright objects (in dispersed images); (vii) stray light; (viii) persistence from previous images; and (ix) zero-point calibration errors. Spectra are then separated using a decontamination algorithm, taking advantage of the four dithers dispersing in four different directions. The final S/N of the spectrum will further depend on modulations of noise due to astrophysical foregrounds (e.g. zodiacal light and nearby galaxies) and MW extinction.

For detector noise, bad pixels, saturated stars, cosmic rays, ghosts, and zeroth-order spectra, all these effects influence the image quality, and their impact is encoded into pixel-level masks, which allow us to exclude contaminated pixels; moreover, they modulate the pixel-level variance of the signal, in case the feature is modelled and subtracted. We can envisage that the handling of these effects will have two phases. In the first phase of data analysis, the algorithms to remove these signals will be sub-optimal, leaving angular imprints in the galaxy density. These can be identified by creating maps of potential contaminants and seek for a positive cross-correlation of the galaxy density with these maps. As an example, an incomplete treatment of saturated stars will create an anti-correlation of the galaxy density with the density field of bright stars; problems in the removal of cosmic rays will create a correlation with a measure of (time-dependent) cosmic ray hits, and so on. The absence of such cross correlations will be a very useful validation test for the algorithms. In a second phase, when images are reprocessed with optimised algorithms, all these effects will be encoded in a pixel-level map, which will be used to compute the variance of the signal. The VMSP algorithm will then be in a position to propagate this uncertainty to the MVM.

One of the most important contribution to noise in SIR images will be the stray light from bright stars scattered inside the instrument. This contribution will be contained in the measurement of the pixel-level signal, used by VMSP to construct the random catalogue, so the impact of stray light is high in the determination of the S/N, but it can be fully mitigated by the random catalogue so the associated risk is considered low.

Appendix B.5: Persistence

When the photons of a bright source hit a pixel of the NISP detector, some of the electrons produced are not processed immediately by the readout hardware but are trapped between the semiconductor layers and later released to create a spurious signal that slowly dies off. As a consequence, a bright source imaged in a pointing remains visible in the successive exposures, in a way that is hard to characterise (Euclid Collaboration: Kubik et al. 2025). This gives rise to four main effects. (i) Spectro-to-photo: dispersed trails of bright sources are visible in the next photometric pointings as long stripes. This has a strong impact because the MER deblending code (discussed above, Appendix B.1), if not perfectly calibrated, can fragment these stripes into a stream of fake sources. (ii) Photo-to-spectro: a bright source in a photometric pointing can contaminate a dispersed image and mimic an

emission line in the spectrum of a random source in one of the dithers. (iii) Photo-to-photo: the three consecutive photometric pointings J_E , H_E , and Y_E , obtained without moving the telescope (with the exceptions of small movements due to the rotation of filter wheels, where an imperfect compensation can lead to pixel-level shifts), will be subject to self-persistence, where each flux in the H_E and Y_E exposures will contain a 1–2% systematic contribution from the previous pointing. (iv) Spectro-to-spectro: signal in very saturated pixels may persist from a dispersed spectroscopic image to the next, creating a feature that is not corrected for by the decontamination algorithm. In addition to these four effects, since the detector is not closed during the rotation of the grism wheel, persistence can produce streaks corresponding to the apparent movement of bright sources when the pointing is adjusted; this is a minor contribution which, however, cannot be removed using the previous pointings.

A detailed characterisation of persistence was performed before launch (Euclid Collaboration: Kubik et al. 2025), but the details of the detectors behaviour in space showed that laboratory tests do not represent the full complexity of actual images. As such, modelling of this effect is an important ongoing activity, and a further improved characterization and subtraction of the persistent signal is expected for the data releases beyond DR1.

Appendix B.6: Calibration error

Another potential killer systematic effect on large scales is the stability of the zero point of the photometric system. This is especially true for ground-based surveys, where target selection is based on available photometry. The spectroscopic sample is, to first-order, free from this effect (see Appendix B.1); moreover, once an emission line is detected, its selection is mostly based on the reliability flag of the redshift measurement (Appendix B.9 below) and not on the absolute measurement of its flux. However, there is a point where the absolute measurement of the signal is used in the pipeline: VMSP will base its estimation of the detection probability of a random galaxy on a measurement of the pixel-level noise at the location of the galaxy spectrum, and this is where the absolute flux calibration in dispersed images is relevant.

Even in this case, the calibration error is expected to be low. Observations will be done by the same instrument throughout the survey, in absence of varying atmospheric conditions, and the survey will continually visit the SelfCal calibration field (Euclid Collaboration: Scaramella et al. 2022), making it possible to have very stable calibration throughout the six years of the survey. The baseline requirement for fluctuations in calibration zero-point is 0.7% between different NISP fields, and a set of algorithms are being devised to guarantee this calibration at the level of a single detector, of the whole field, and at the level of ‘UberCalibration’ (Markovič et al. 2017), where the overlap of adjacent pointings is used to further check the stability of calibration on larger scales. At the present stage, we can say that it is unlikely that the UberCalibration will be worse than 1–2%.

Appendix B.7: Extraction of 2D spectra and wavelength calibration

The extraction of 2D spectra performed by SIR (Euclid Collaboration: Copin et al. 2025) requires the mapping of pixels to wavelengths; this requires wavelength calibration and correction for non-linearities in the grism response. The formal requirement on the precision of wavelength calibration is 0.8 pixels, correspond-

ing to 10.8 Å, below the spectral point-spread function. Calibration is performed by observing planetary nebulae, which have a very well-known spectrum dominated by emission lines, achieving sub-pixel precision. Sometimes, a dispersed image of 530×5 pixels can fall across a detector gap and thus have signal in two different the detectors; in this case the reconstruction of the spectrum relies on knowledge of the exact geometry of detectors. The accuracy of wavelength calibration can also be tested a posteriori by comparing the measured spectra with ground-based ones for archive sources, where ground-based spectra have much higher resolution and robust wavelength calibration.

Since the uncertainty in wavelength is below the spectral point-spread function, the impact of this uncertainty is expected to be very low. Its effect would be a distortion in the anisotropic power spectrum, which may be considered equivalent to a (negligible) bias in the estimation of Alcock–Paczynski parameters.

Appendix B.8: Confusion and spectral decontamination

Spectral confusion is one of the most important systematics in the extraction of *Euclid*’s spectroscopic catalogues. Galaxies are dispersed in 530×5 pixel stripes, and photometric sources will oversample the focal plane, so each galaxy will overlap with 10 to 30 other sources. As discussed above (Table 2 and Euclid Collaboration: Passalacqua et al., in prep.), only galaxies with $H_E \lesssim 20$ can have a detectable continuum, so confusion will be driven by these relatively bright objects. However each galaxy will overlap with different objects in the four dithers, making decontamination from brighter sources possible. The decontamination algorithm, described in Euclid Collaboration: Copin et al. (2025), will separate the sources, but the bright spectra will act as a background noise term for the fainter ones. As such, the main effect of confusion will be to increase the spectral noise so that a galaxy that would have barely been detected in isolation may be missed when contaminated. A less effective decontamination method would just make it more probable that a galaxy is missed, or could create noise interlopers, with a probability that acquires a dependence on source surface density. Because J_E and Y_E magnitudes are used to aid decontamination, its effectiveness will be affected by spectro-to-photo and photo-to-photo persistence.

We can separate the contaminating sources in two categories: (1) stars and foreground galaxies at redshift $z < 0.9$ (or beyond $z > 1.8$), and (2) galaxies at $z > 0.9$. The first class of sources is statistically independent of the spectroscopic galaxies, so their effect can be absorbed into the class of astrophysical foregrounds that create angular modulations of the density. As a matter of fact, VMSP accounts for of these sources, as the noise used to estimate the variance is read directly from the NISP dispersed image. If, however, there is a significant contamination from sources at similar redshift as the spectroscopic sample, the assumption of statistical independence of foreground and cosmological signal is not valid. This may be important, say, in the cores of galaxy clusters, creating a dangerous density-dependent bias. This point is addressed in the companion paper Euclid Collaboration: Passalacqua et al. (in prep.), where it is shown that the relative impact of contamination from galaxies at $z > 0.9$ on the observed galaxy density amounts to around 4%.

Appendix B.9: Fit of the 1D spectrum, line detection and redshift measurement

Once SIR has coadded the four 2D spectra into a single calibrated 1D spectrum, SPE (Euclid Collaboration: Le Brun et al. 2025) has the task of fitting it to classify the source, measure its redshift and line fluxes, and assess the reliability of the redshift measurement. The main condition for a source to get into the spectroscopic sample will be to have a reliable redshift measurement in the nominal redshift range of $H\alpha$ ELGs. To maximise sample density, emission lines down to a relatively low S/N of 3.5 are selected. However, since the photometric sources are much more numerous than the detectable ELGs, the probability that a noise fluctuation is misinterpreted as a spectral line is not negligible. Moreover, since the 2D spectra of photometric sources oversample the NISP field (see the discussion at the beginning of Sect. 3.1), any spurious feature in the dispersed image that is not perfectly subtracted or masked will fall over the stripe of some other source and may thus create a fake spectral feature. We then expect the spectroscopic catalogue to be contaminated by a relatively high fraction of noise interlopers.

Many sources will have a single (correctly) detected emission line. To maximise the density of the spectroscopic sample, fits will use a prior that enhances the interpretation that a single line is $H\alpha$, thus boosting completeness at the expenses of purity.

Spectral templates play an important role in the redshift measurement. For instance, the centroid of the $H\alpha$ –[N II] blended complex can be systematically offset from a standard reference position, due to the asymmetry of the [N II] doublet. Since the spectral resolution is not sufficient to cleanly separate these lines, the blend can lead to a systematic offset in the centroid, leading to a biased redshift (Euclid Collaboration: Passalacqua et al., in prep.; Euclid Collaboration: McCarthy et al., in prep.). The overall bias in the sample can be minimised by using spectral templates that span a realistic range of [N II]/ $H\alpha$ line ratios, and thus of shapes of the unresolved triplet combination. The maximum centroid offset over this range occurs when [N II] contributes more than half of the total flux; at the extreme we expect a redshift error of at most 0.5 pixels, of the same order as the wavelength calibration error.

Since the fitting procedure provides a redshift for any extracted spectrum, a sample with permissive cuts will have very low purity. Cleaning the catalogue of interlopers will be firstly based on a spectral quality flag; this will be defined as the integral of the normalised PDF of the redshift posterior around the main peak, a high value marking the most reliable measurements. Furthermore, a machine-learning algorithm will be trained on EDS galaxies to identify the region in the space of *Euclid* magnitudes that corresponds to the locus of potential $H\alpha$ ELGs, thus cleaning out a significant fraction of interlopers. As mentioned in Appendix B.1, usage of photometric information can be the entry point for systematics in photometry, so this option will be carefully tested before being used; already, Cagliari et al. (2024) have shown that a cut at $H_E < 22.5$ is effective in increasing purity.

Correctly identified emission lines lead to a valid redshift measurement, which is expected to have a statistical error with standard deviation of $\sigma_z \simeq 0.001(1+z)$ – see Euclid Collaboration: Le Brun et al. (2025); in fact, due to the grism response, this error does not depend on redshift (see Sect. 2.2). The effect of redshift errors is to imprint an anisotropic smoothing to galaxy clustering, which must then be added to the model; the exact width of the redshift error can be left as a nuisance parameter with a tight prior, to be measured and marginalised over at

the likelihood level. A straightforward way to validate the estimation of the redshift error is to compare *Euclid* redshift measurements with ground-based archival data; this will also allow us to check the shape of the PDF, which is not expected to be Gaussian (see [Euclid Collaboration: Le Brun et al. 2025](#)).

Appendix B.10: Dependence of detection on galaxy properties

A whole sector of systematics comes in principle from the dependence of detection probability on the intrinsic properties of the galaxy other than its line luminosity. These properties determine the precise mapping from dark matter haloes to observed galaxies, and thus determine the galaxy bias parameters. The state of a galaxy is mostly determined by the host halo mass and its merger history, which means that the bias is likely to have a component from assembly bias. However, our cosmological models are not supposed to be predictive of galaxy bias parameters; these are treated as nuisance parameters, and cosmological inference is obtained by marginalising over them. So the question shifts to whether the parametrisation of galaxy bias used in the fitting model, obtained from a perturbative expansion ([Desjacques et al. 2018](#)) or from an HOD based on an halo model (e.g. [Avila et al. 2020](#)), is general enough to account for such an effect. From this point of view, the most worrying potential systematics are those that are not included in a standard bias model. One of these comes from the dependence of line detection on the inclination of galaxy images with respect to the dispersion direction in the grism. Galaxy inclination is known to be correlated with the large-scale tidal fields ([Hirata 2009](#); [Obuljen et al. 2020](#); [Lamman et al. 2023](#)), implying that there is a link between selection, galaxy orientation, the larger scale density field, and galaxy velocity, which is anisotropic with respect to the line of sight. This can then cause a non-local bias term that is not present in standard models. Being this contingent on the selection, accounting for this term requires careful modelling of the specific samples – for *Euclid* this would involve coupling models of $H\alpha$ emission as a function of galaxy orientation with simulations that include the coupling between orientation and tidal fields.

Appendix C: Effect of each case of systematics on the galaxy power spectrum

In this Appendix, we present the galaxy power spectra for the six cases of systematics identified in Sect. 5.1.1 and listed in Table 5. They are shown in Fig. C.1. We use here catalogues selected using shuffled fluxes to remove the luminosity-dependent bias. As a consequence, the convolution of the best fit model (Fig. 8) with the window function measured from the random catalogue should fit the measurement obtained with ideal mitigation (as demonstrated in Fig. 12). Each panel shows measurements of the first three even multipoles of $P(k)$, averaged over 50 mocks, with shot noise subtracted. All panels report the target measurement in blue, with the lighter blue shaded area giving the relative error of the average of groups of five mocks. The red lines give measurements based on ideal mitigation, grey line those relative to no mitigation (also shown in Fig. 11). The three sub-panels below the main panel report, for the three multipoles, the relative variance of the target measurements as the shaded area, the ratio of the measurements $P_i/P_t - 1$ (like in the top panel of Fig. 12) as thin dashed lines, and the quantity $\Delta\Delta P$ discussed above as thicker lines, this time shown for all multipoles.

Appendix D: Metrics for the goodness of fit in the presence of systematics

To quantitatively estimate the impact of different mitigation strategies on cosmological parameters, we employ the following metrics.

Figure of Bias (FoB): the first metric we use to evaluate the performance of each mitigation procedure is the consistency of the inferred parameters with their true values. To quantify this consistency, we define the figure of bias as the covariance-weighted difference between the posterior means and their fiducial values,

$$\text{FoB} = \left[(\bar{\theta} - \theta_{\text{fid}}) S^{-1} (\bar{\theta} - \theta_{\text{fid}})^T \right]^{1/2}, \quad (\text{D.1})$$

where, $\bar{\theta}$ and θ_{fid} are vectors containing the posterior means of the three cosmological parameters and their fiducial values, and S is the parameter covariance matrix. We compute the parameter means and covariance for the cosmological parameters using the `getdist` package. The covariance matrix converts the parameter differences into levels of significance, such that a value of FoB=1.88 (2.83) can be interpreted as corresponding to the 68% (95%) confidence level.

Figure of Merit (FoM): this quantifies the constraining power of the measurement over the cosmological parameters and is computed via

$$\text{FoM} = \frac{1}{\det S}. \quad (\text{D.2})$$

Credibility interval width: this is the fractional change in the average width of the parameter 68% credibility intervals with respect to the target catalogues,

$$\delta\sigma = \frac{(\text{FoM})^{1/3}}{(\text{FoM}_{\text{target}})^{1/3}} - 1. \quad (\text{D.3})$$

Goodness of fit: we quantify it by estimating the probability that the power spectrum realisations are generated by the best-fit model. This probability, usually dubbed p -value, is computed using the χ^2 statistic obtained for each realisation. Specifically, for a given mitigation strategy that we denote with an index a and a realisation i yielding a value $\chi^2_{i,a}$, the p -value is defined as

$$p_a = 1 - \frac{1}{N_{\text{real}}} \sum_{i=1}^{N_{\text{real}}} \int_0^{\chi^2_{i,a}} \mathcal{P}(\chi^2, n_d) d\chi^2, \quad (\text{D.4})$$

where $\mathcal{P}(\chi^2, n_d)$ denotes the χ^2 distribution with n_d degrees of freedom. We set 0.05 as the threshold probability, below which we reject the hypothesis that the model is the underlying truth of the measurement in a particular case.

Figure D.1 presents the performance metrics for all five cases. In the top panel, which displays the FoB values, we observe that the estimated parameters do not exhibit significant bias for most of the mitigation strategies. However, in *detmodel3*, the bias exceeds the 68% confidence threshold, indicating a systematic deviation, which is evident in Fig. 15. The second panel from

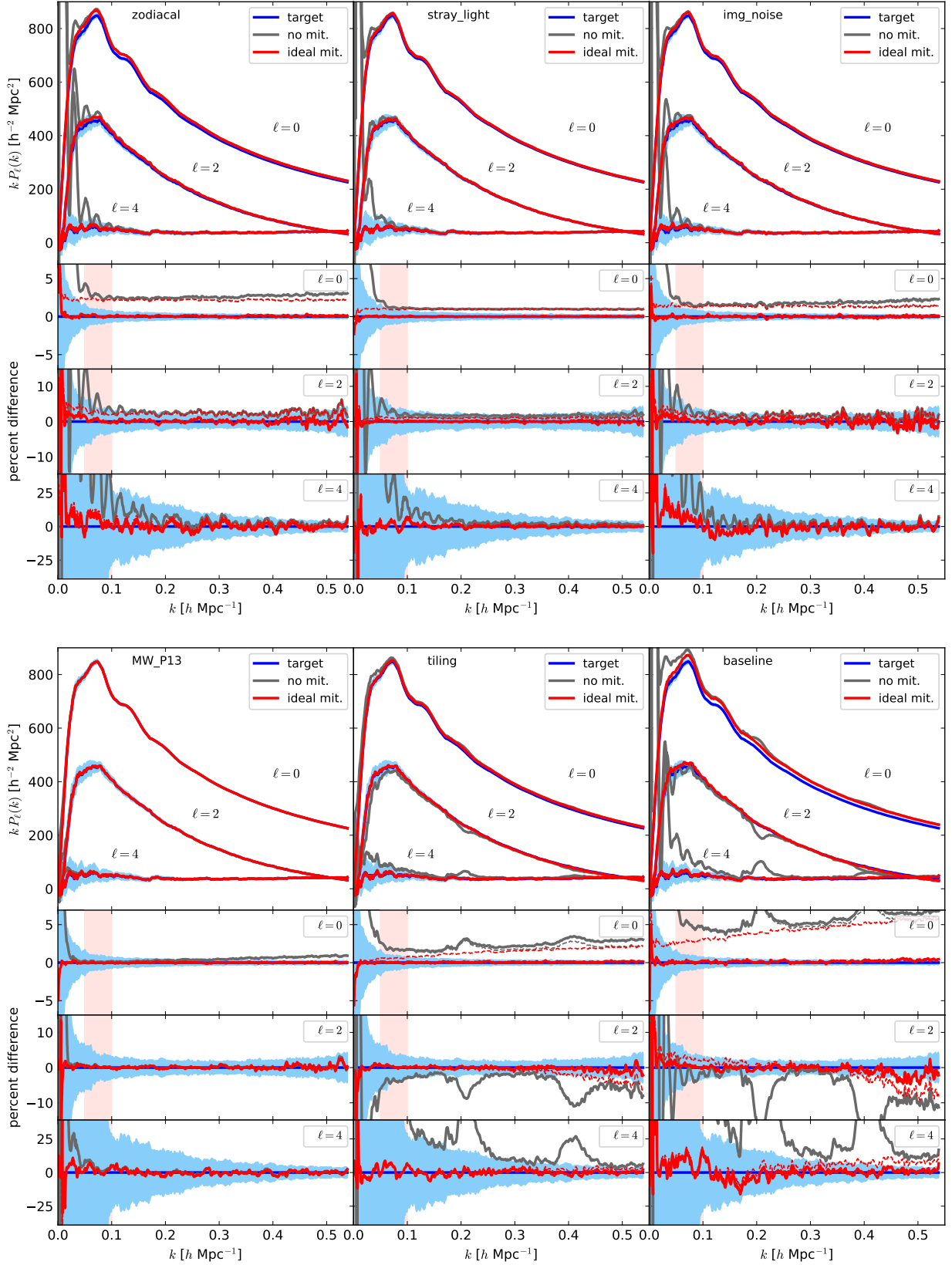


Fig. C.1. Monopole, quadrupole and hexadecapole of the galaxy power spectrum, for the six cases of systematics of Table 5. For each panel the layout is as in Fig. 8. Blue lines give the target measurement, the lighter blue shaded area gives the error on an average of five measurements. Red lines give the measurement with the systematic declared above each panel. Grey lines give the measurement relative to the ad-hoc random catalogue with no mitigation of angular systematics. In the sub-panels with residuals, thin dashed lines give the ratio of measurements, thick lines the quantity $\Delta\Delta P$ (Eq. 28) as in the lower panel of Fig. 12, all in percent.

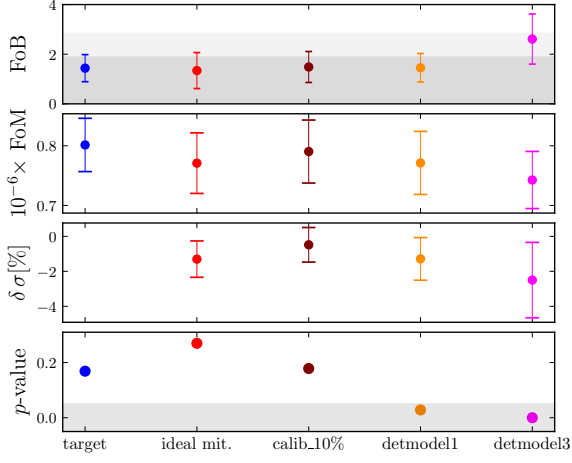


Fig. D.1. From the top, first panel: the average and scatter of the FoB values for the cosmological parameters $\{h, A_s, \omega_c\}$, respectively, across the different models. We shade the regions corresponding to 68% (95%) confidence levels. Second panel: the average and scatter of the FoM values. Third panel: the average and scatter of the $\delta\sigma$ values. Fourth panel: the p -value metric for the various cases; the shaded area shows the $p = 0.05$ rejection threshold.

the top shows, through the FoM, that the mitigation strategy has little impact on the model's constraining power. The third panel shows that parameter error bars are not strongly influenced by the mitigation strategy, the largest deviations are below 5%. Finally, in the bottom panel the p -value metric indicates that the model successfully fits most cases, except for *detmodel1* and *detmodel3*, which are below the rejection threshold. We note that despite its low p -value, *detmodel1* does induce significant biases, as seen in Fig. 16.