Astronomy & Astrophysics

Organised randoms: Learning and correcting for systematic galaxy clustering patterns in KiDS using self-organising maps

Harry Johnston[1,2], Angus H. Wright[3], Benjamin Joachimi[2], Maciej Bilicki[4], Nora Elisa Chisari[1], Andrej Dvornik[3], Thomas Erben[5], Benjamin Giblin[6], Catherine Heymans[3,6], Hendrik Hildebrandt[3], Henk Hoekstra[7], Shahab Joudaki[8,9], and Mohammadjavad Vakili[7]

[1] Institute for Theoretical Physics, Utrecht University, Princetonplein 5, 3584 CE Utrecht, The Netherlands
e-mail: h.s.johnston@uu.nl
[2] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[3] Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), German Centre for Cosmological Lensing, 44780 Bochum, Germany
[4] Center for Theoretical Physics, Polish Academy of Sciences, al. Lotników 32/46, 02-668 Warsaw, Poland
[5] Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany
[6] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[7] Leiden Observatory, Leiden University, PO Box 9513, Leiden 2300 RA, The Netherlands
[8] Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK
[9] Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

ABSTRACT

We present a new method for the mitigation of observational systematic effects in angular galaxy clustering through the use of corrective random galaxy catalogues. Real and synthetic galaxy data from the Kilo Degree Survey's (KiDS) 4th Data Release (KiDS-1000) and the Full-sky Lognormal Astro-fields Simulation Kit package, respectively, are used to train self-organising maps to learn the multivariate relationships between observed galaxy number density and up to six systematic-tracer variables, including seeing, Galactic dust extinction, and Galactic stellar density. We then create 'organised' randoms; random galaxy catalogues with spatially variable number densities, mimicking the learnt systematic density modes in the data. Using realistically biased mock data, we show that these organised randoms consistently subtract spurious density modes from the two-point angular correlation function $w(\vartheta)$, correcting biases of up to 12σ in the mean clustering amplitude to as low as 0.1σ, over an angular range of $7-100$ arcmin with high signal-to-noise ratio. Their performance is also validated for angular clustering cross-correlations in a bright, flux-limited subset of KiDS-1000, comparing against an analogous sample constructed from highly complete spectroscopic redshift data. Each organised random catalogue object is a clone carrying the properties of a real galaxy, and is distributed throughout the survey footprint according to the position of the parent galaxy in systematics space. Thus, sub-sample randoms are readily derived from a single master random catalogue through the same selection as applied to the real galaxies. Our method is expected to improve in performance with increased survey area, galaxy number density, and systematic contamination, making organised randoms extremely promising for current and future clustering analyses of faint samples.

Key words. cosmology: observations – large-scale structure of Universe – methods: data analysis

1. Introduction

Recent decades have seen the advent of precision cosmology, as inferred from the large-scale structures of the Universe (Ho et al. 2012; Font-Ribera et al. 2014; Anderson et al. 2014; Hildebrandt et al. 2017; Alam et al. 2017; Troxel et al. 2018; Hamana et al. 2019; eBOSS Collaboration 2020; Tröster et al. 2020; Asgari et al. 2021b). We can shed light on the mysterious dark matter scaffold by examining the distribution of the galaxies supported by it, and extend these measurements across broad redshift epochs in order to study dark energy through the evolution of structures and the universal expansion history. Tensions are beginning to emerge between near-universe (Freedman 2017; Riess et al. 2019; Hildebrandt et al. 2019; Joudaki et al. 2020; Heymans et al. 2021) and cosmic microwave background (CMB, e.g., Aghanim et al. 2020) measures of the expansion rate, and the amount and clustering of matter in the Universe. Resolv-

ing these tensions will require strict control of sources of systematic error in the new era of statistical precision soon to be explored (Dark Energy Spectroscopic Instrument; Aghamousa 2016, *Rubin* Observatory; LSST Science Collaboration 2009, *Euclid*; Laureijs et al. 2011).

Currently, the most competitive large-scale structure (LSS) constraints upon the matter energy density Ω_m versus normalisation of the matter power spectrum σ_8 plane come from combined analyses of galaxy clustering and weak gravitational lensing (van Uitert et al. 2018; Joudaki et al. 2017; Abbott et al. 2018; Asgari et al. 2021a; Heymans et al. 2021), making use of wide-area galaxy survey data (e.g., Dark Energy Survey; The Dark Energy Survey Collaboration 2005, Kilo Degree Survey; de Jong et al. 2013, Hyper Suprime-Cam; Miyazaki et al. 2012). However, these surveys are susceptible to complex source selection functions, which can destroy and/or mimic real, informative signals, such as cosmological density fluctuations.

These selection functions are generally imprinted on galaxy survey data in a way that correlates with observational and physical properties of the survey on-sky. Examples of such properties, which have the propensity to introduce spurious selection effects on galaxy data, include spatial variation in atmospheric seeing, the telescope point-spread function (PSF), stellar density, Galactic dust extinction, and more. Crucially, it is possible, or even likely, that some spurious selection functions may be introduced due to a complex confluence of many observational properties, some of which we may be unaware of.

Upcoming surveys are likely to build upon the many weak lensing analyses that currently aim to constrain cosmology with the 3x2pt analysis method: simultaneously fitting the two-point functions describing the galaxy shear-shear (cosmic shear), position-shear (galaxy-galaxy lensing), and position-position (galaxy clustering) correlations. Analysed simultaneously, these statistics help to constrain nuisance parameters (e.g., photometric redshift calibration uncertainties, or intrinsic alignments), thereby tightening cosmological constraints. However, this additional constraining power also brings sensitivity to systematic biases; any bias in (e.g.,) the galaxy clustering correlation, such as those that may be introduced by systematic variation of the galaxy density field, will propagate into the inferred (joint) cosmology in a pathological fashion.

This work focuses on sources of systematic error affecting the detection of galaxies, and propagating into statistics involving galaxy positional data, for instance correlation functions. Following recent studies on this topic, it is useful to divide mitigation strategies into three categories: (i) Monte Carlo simulation of synthetic objects, (ii) mode (de)projection, and (iii) regression.

The first approach seeks to inject artificial galaxies into realistic images, so as to gauge the detector-response to observational systematics (Bergé et al. 2013; Suchyta et al. 2016). As a result, a forward model for the survey selection function can be built, from which a Monte Carlo sampling can generate mock catalogues of unprecedented fidelity. This computationally demanding technique is extremely promising and under active development (Everett et al. 2020).

Techniques for systematics mode projection, for example assigning large variance to spurious modes so that they are ignored by power spectrum estimators, were developed by Leistedt et al. (2013) and Leistedt & Peiris (2014) for photometric quasar clustering, and similar methods have recently been applied to Hyper Suprime-Cam (HSC) data by Nicola et al. (2020). Interestingly, Weaverdyck & Huterer (2021) demonstrated that mode projection techniques can be consistently expressed under a regression framework, and showed that extensions to previous formulations can automate the selection of important systematic features and remain robust under scenarios of numerous correlated systematics.

Much work on this topic has focused upon regression-based approaches, where the aim is to model the functional relationships between systematic-tracing variables and galaxy number densities. Ross et al. (2011) and Ho et al. (2012) suppressed systematics in SDSS BOSS-like (Sloan Digital Sky Survey; York et al. 2000, Baryon Oscillation Spectroscopic Survey; Eisenstein et al. 2001) photometric luminous red galaxy (LRG) data by deriving per-galaxy inverse weights from number density-systematics ('one-point' or 'pixel') correlations or by computing signal corrections (assuming systematic-tracers relate linearly to galaxy number densities) from galaxy-systematic (two-point) cross-correlations. Vakili et al. (2020) also estimated weights from one-point functions to measure

the clustering of photometric LRGs in the Kilo Degree Survey (KiDS), decomposing systematic-tracers in an orthogonal basis and exploring second-order polynomials to characterise cross-talk between parameters. Elvin-Poole et al. (2018) iteratively derived weights for Dark Energy Survey (DES) LRGs from linear systematic-density fits, and Wagoner et al. (2021) recently improved upon this analysis by performing simultaneous likelihood fitting of linear coefficients to all systematics maps and the observed density contrast, and then calibrating for over-correction of clustering correlations with mock catalogues.

A possible sub-category of regression is comprised of studies that use machine learning to predict the relationship between observed galaxy number densities and multivariate systematics vectors, and then derive density field corrections without the least-squares methods that typically characterise regression-based approaches. For example, Rezaie et al. (2020) derived weights for emission line galaxies (ELGs) selected (following eBOSS; Raichoor et al. 2017) from the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019) using deep neural networks. Morrison & Hildebrandt (2015) tackled clustering biases in the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS; Erben et al. 2013), using k-means clustering in the high-dimensional density-systematics space to create weight-maps from which to draw random points. These random points then compensate for systematic density fluctuations through typical estimators for clustering correlations. Our methods belong to this machine-learning regression category, which has the advantage over most regression-based methods that the multivariate density-systematics models require no functional form and can be arbitrarily non-linear, allowing for the compensation of more complex modifications to the cosmic density field.

We seek to mitigate density field biases through the construction of tailored random galaxy catalogues (hereafter 'organised randoms'), which mirror systematically induced galaxy-density variations (similarly to Morrison & Hildebrandt 2015; Suchyta et al. 2016). Random galaxy catalogues (commonly referred to simply as 'randoms') are widely used when estimating galaxy clustering and galaxy-galaxy lensing (GGL), whereby correlations between galaxies and random points allow for reductions in methodological bias, improved covariance properties between the statistics, removal of systematics caused by edge or masking effects, and aid in the reduction of additional systematic correlations (Landy & Szalay 1993; Singh et al. 2017). For this task, studies typically employ high-density (relative to the survey galaxy number density) spatially uniform random points. However, while these will aid in the removal of systematic correlations with the observed galaxy distribution, they cannot account for spatial correlations stemming from systematically unobserved galaxies, or those systematically lost due to sample selection effects.

Our work aims to tackle this problem: we use a form of machine-learning assisted dimensionality reduction, the self-organising map (or 'SOM', Kohonen 1990), to infer from the observed galaxy distribution the high-dimensional mapping between survey systematics and galaxy number densities on-sky. We then create many clones of the real galaxies (i.e. copies retaining all photometric or other properties, see Farrow et al. 2015) and distribute them as random points throughout the survey footprint, in accordance with their systematically derived number density on-sky. This allows any selection effects in galaxy data to be trivially mirrored in the organised randoms, thereby preserving the systematic patterns and the systematically induced density variations for an arbitrarily defined galaxy sample.

Table 1. KiDS: systematic-tracer variables for training of self-organising maps, with units and descriptions.

Systematic-tracer variable	Unit	Description
MU_THRESHOLD (MU)	mag arcsec^{-2}	r-band detection threshold above background; the minimum surface brightness of objects detected after background-subtraction. Fainter objects will be lost from an area of observation where the detection threshold is brighter.
psf_fwhm	arcsec	Full-width at half-maximum of the r-band point-spread function, averaged over $n_{side} = 512$ (47 arcmin^2) HEALPIX[1] (Gorski et al. 2005) pixels and then interpolated to galaxy locations. The blurring of small, faint sources could cause objects to drop below detection thresholds.
psf_ell	dimensionless	Ellipticity ($1 - q$, where q is the 2D major/minor axis ratio) of the r-band point-spread function, also averaged over $n_{side} = 512$ pixels and interpolated to galaxy locations. A PSF ellipticity indicates non-isotropic blurring of object isophotes, creating challenges for shape estimation and inducing a directional dependence for detections.
MAG_LIM_x	mag	x-band limiting magnitude (5σ above background in a $2''$ aperture) at the object's location, for each of the 9 bands observed by KiDS-VIKING. KiDS object detection is performed in the r-band.
EXTINCTION_r	mag	Galactic dust extinction in the r-band, derived using the Schlafly & Finkbeiner (2011) coefficients for the Schlegel et al. (1998) dust maps. Dust preferentially scatters short-wavelength light from extragalactic objects; the loss of flux could prevent detections, and the modification of galaxy spectral energy distributions (SEDs) poses other problems, e.g., for photo-z estimation.
gaia_nstar	count arcmin^{-2}	Number density of *Gaia* DR2 (Arenou et al. 2018) $14 < G < 17$ stars in $n_{side} = 512$ pixels, interpolated onto galaxy locations. High stellar densities can hamper detections as the light from stars obscures background objects, and can also result in spurious galaxy-detections through the misidentification of PSF-blurred, or blended, stars as galaxies.

Notes. [1] http://healpix.sourceforge.net (Gorski et al. 2005). Systematic-tracer variables chosen from the KiDS DR4 data products, or from other public data e.g., *Gaia* n_{star}. These are variables thought to trace phenomena that may impact upon the observed number density of galaxies. See Kuijken et al. (2019) for details on threshold, PSF, and magnitude-limit parameters. Maps of each of these quantities are displayed in Figs. 1 and 2. Where specified, pixelated systematic-tracer values are interpolated from pixel centres onto galaxy locations via a two-dimensional (RA/Dec) nearest-neighbour interpolation.

The paper is organised as follows: In Sect. 2 we introduce our galaxy data from Kilo Degree Survey (KiDS, Kuijken et al. 2019) observations and from the Full-sky Lognormal Astrofields Simulation Kit (FLASK; Xavier et al. 2016) simulations. Section 3 describes self-organising maps and how they are used in this work. In Sect. 4 we assess the capability of the self-organising map to identify artificially created systematic trends in galaxy density. Section 5 then turns to KiDS data-driven systematic density fluctuations and demonstrates the utility of organised randoms in recovering unbiased clustering signals from realistically biased FLASK mocks. Final data applications are presented in Sect. 6, and we make concluding remarks in Sect. 7. Throughout this work, we quote AB magnitudes and assume a fiducial WMAP9+BAO+SN cosmology (Hinshaw et al. 2013): flat ΛCDM, with $\Omega_m = 0.2905$, $\Omega_b = 0.0473$, $\sigma_8 = 0.826$, $h = 0.6898$, and $n_s = 0.969$.

2. Data

We validated our random catalogues using both real and synthetic galaxy data, and invoking both realistic distributions of systematic parameters on-sky (drawn from the KiDS 4th Data Release, DR4, Kuijken et al. 2019) and artificially constructed systematics distributions. For our simulations, we used lognormal galaxy fields simulated with FLASK (Xavier et al. 2016).

2.1. KiDS

Both KiDS (de Jong et al. 2013) and its partner survey, the VISTA Kilo-Degree Infrared Galaxy (VIKING; Edge et al. 2013) survey, are now observationally complete, covering a combined area of 1350 deg^2 on-sky in a total of nine photometric bandpasses ($ugriZYJHK_s$). Over 1000 deg^2 of this combined dataset is publically available as part of KiDS DR4[1] (Kuijken et al. 2019), providing gravitational shear estimates (unused in this work), nine-band photometric redshift estimates (Hildebrandt et al. 2019, 2021; Wright et al. 2020), and observational information for over 100 million galaxies. This galaxy sample, typically referred to as KiDS-1000 (though the retained area after masking is ~900 deg^2), samples a wide range of observing conditions over a large area, and is consequently imprinted with an unknown combination of systematic galaxy depletion and enhancement patterns. Of all the various data products provided within the KiDS DR4, we identify a selection of systematic-tracer variables (detailed in Table 1 and mapped out in Figs. 1 and 2) that trace physical phenomena that have the greatest potential to imprint subtle galaxy selection functions on the dataset, based on our experience. These parameters, individually and in combination, form the dataset used to train our SOMs. It is not known a priori whether the chosen variables
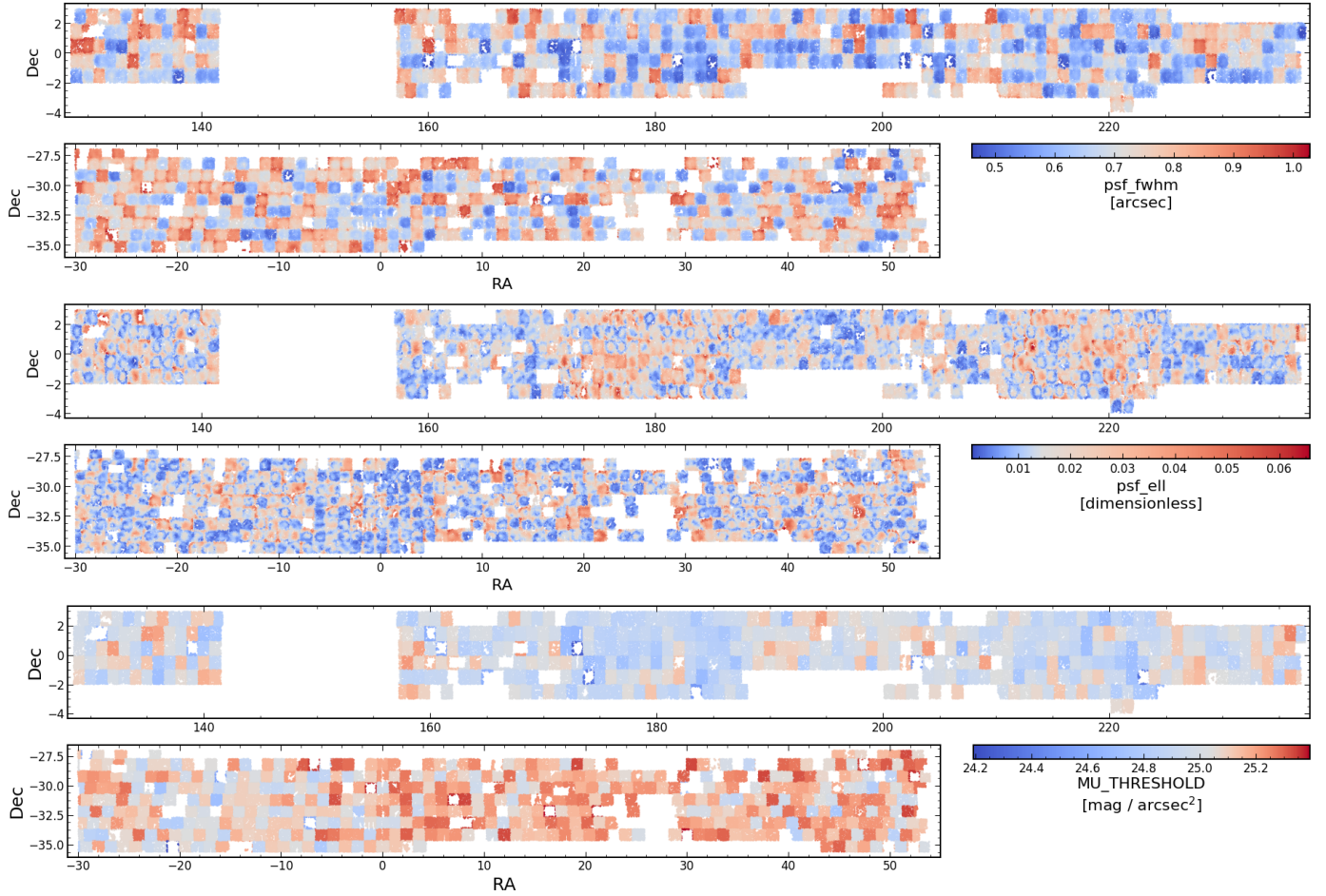
---

[1] http://kids.strw.leidenuniv.nl/DR4/

Fig. 1. Maps of systematic-tracer variables (from Table 1) from *r*-band (the detection band) imaging in the KiDS-North (*top panels*) and KiDS-South (*bottom panels*) areas. Grey denotes the 50th percentile of the systematics distribution in each case, and blue and red then denote good and bad observing conditions relative to the 50th percentile. As we show in Fig. 4, the majority of spatial variations in galaxy number density correlate with these parameters at ≲5%.

all trace real observational phenomena that cause the systematic loss of galaxies; we therefore explore the effect of unimportant, 'distracting' variables as we test our methodology (Sect. 4.1).

Angular clustering correlations are typically measured in bins of galaxy redshift, so as to constrain the galaxy bias of redshift samples, thus combining powerfully with lensing probes such as GGL (e.g., Yoon et al. 2019), and to assess the growth of large-scale structure over cosmic time. Accurate redshifts (typically from spectroscopy) are required for the optimal binning of galaxies and modelling of correlations, but these are expensive to obtain for large samples of galaxies over a wide area. Consequently, such wide-field surveys rely upon photometric redshifts (photo-*z*), estimated from broadband photometry (such as the nine filters used in KiDS-1000) that sample the spectral energy distributions (SEDs) of the galaxies. In this work, we focus on a subsample of KiDS-1000 with high-quality photo-*z* estimates: the ~1M bright galaxy subsample (with $r \lesssim 20$), whose photo-*z* are computed using ANNz2 neural networks (Sadeh et al. 2016), trained on spectroscopically observed galaxies from the Galaxy And Mass Assembly (GAMA) survey (Driver et al. 2009). This approach to photo-*z* estimation for bright KiDS galaxies was originally presented by Bilicki et al. (2018) using only optical (*ugri*) photometry from KiDS DR3. In our work we use the updated DR4 bright-sample described in Bilicki et al. (2021), which leverages the expanded nine-band photometric dataset to

achieve photo-*z*s with a typical accuracy of $\sigma_{z_{\rm phot.}} \sim 0.02(1 + z)$ in terms of the normalised median absolute deviation (nMAD).

The photometric redshift distribution $n(z_{\rm phot.})$ for this GAMA-like photometric sample is shown in the top panel of Fig. 3. Following van Uitert et al. (2018), we define two redshift bins for our GAMA-like sample, with edges at $z_{\rm phot.} = \{0.02, 0.2,$ and $0.5\}$, within (and between) which we measured angular clustering correlations. For the purpose of additional testing, we defined two additional bins within the range of $z_{\rm phot.} = \{0.02$ and $0.5\}$, but with inner boundaries well separated in photo-*z*, at $z_{\rm phot.} = \{0.22$ and $0.28\}$. This separation minimises the overlap between the true redshift distributions in the bins, which is induced by photo-*z* scatter. We display the approximate 95% scatter $2\sigma_{z_{\rm phot.}} \sim 0.04(1+z)$ as a red line, with values on the right-hand axis. We used these GAMA-like KiDS data, 'KiDS-Bright' henceforth, for intermediate tests and to assess the ultimate performance of our organised randoms (Sects. 5 and 6).

In our companion letter, Wright et al. (in prep.), we also explore an application of our organised randoms to measurements of galaxy clustering in the KiDS-1000 shear sample. This 'gold' sample (see Wright et al. 2020; Giblin et al. 2021; Hildebrandt et al. 2021, for details) is ~5 magnitudes deeper than KiDS-Bright, and a factor ~20 more dense on-sky. This increased statistical power should allow for a more faithful sampling of the multivariate systematics-density relation that is
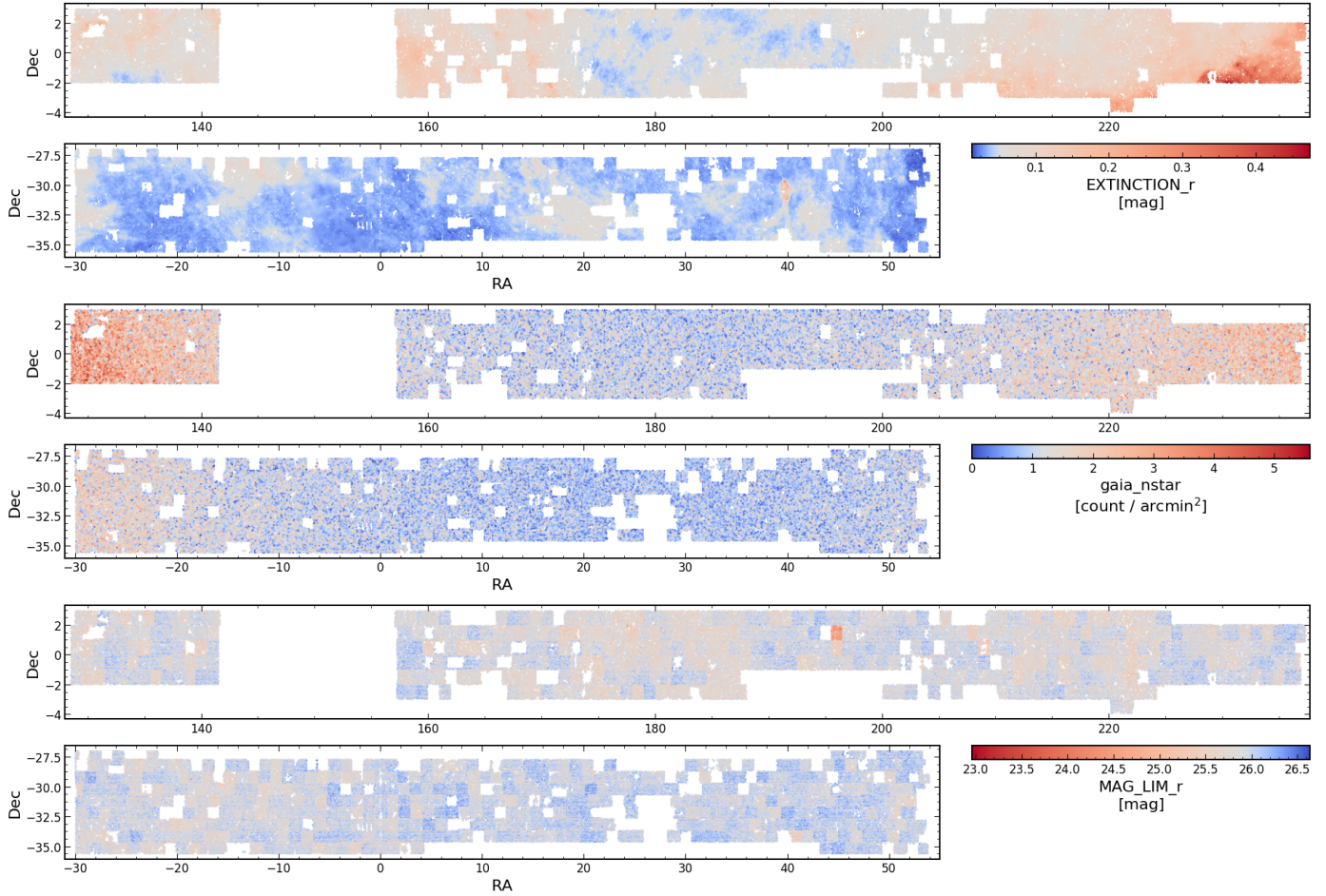
Fig. 2. Same as Fig. 1, but for the remaining systematic-tracer variables from Table 1.

hidden in the data, which should also be easier to disentangle from cosmic structure as the faint data are more heavily biased. We expect the performance of organised randoms to improve on application to faint datasets, posing intriguing possibilities for the future of deep galaxy clustering analyses.

2.2. FLASK

FLASK (Xavier et al. 2016) is a public code designed to simulate lognormal (or Gaussian) random fields on the celestial sphere, with configurable tomography and preservation of all relevant correlations between galaxy density and weak lensing convergence fields, to the sub-percent level. We estimate the error on $w(\vartheta \geq 3\,\mathrm{arcmin})$ to be $\gtrsim 2.5\%$ for KiDS-Bright-like statistics from sample variance and Poisson noise considerations alone, hence <1% accuracy from FLASK is sufficient for our purposes.

For the cosmology specified at the end of Sect. 1 and the $n(z_{\mathrm{phot.}})$ displayed in the top panel of Fig. 3, we computed a 'truth' angular power spectrum C_ℓ with which we used FLASK to generate many mock galaxy catalogues from lognormal random fields; these form the basis of initial testing for our organised randoms, as is described in Sects. 4 and 5. The bottom panel of Fig. 3 demonstrates that measurements of $w(\vartheta)$ in our mocks reliably recover the analytical input clustering and sample variance (plus shot-noise) over 30 realisations (see Eqs. (5.1) and (5.2)), modulo noise. These statistics are for FLASK realisations with average galaxy densities $0.36\,\mathrm{arcmin}^{-2}$ (i.e. the

same as KiDS-Bright) simulated within the true KiDS-Bright masked survey footprint, and with a galaxy bias equal to unity. We used these KiDS-Bright-like FLASK catalogues, as well as simpler $10\,\mathrm{deg} \times 100\,\mathrm{deg}$ windows (straddling the celestial equator, to loosely mimic the KiDS survey geometry) initialised with $1\,\mathrm{galaxy\,arcmin}^{-2}$ to test our self-organising maps and random galaxy catalogues.

3. Self-organising maps

Self-organising maps (Kohonen 1990) are a class of unsupervised neural network methods, designed to project high-dimensional data onto (typically) two-dimensional maps that preserve the topological features of the input space. Proximity on the map therefore tends to denote proximity within the high-dimensional space. SOMs are fast, simple, and useful for problems benefiting from dimensionality reduction, unsupervised classification, and ease of data visualisation. Their use within cosmology has included object selection and classification for large datasets (Geach 2012), template photo-zs (Speagle & Eisenstein 2015, 2017), characterisation of galaxy properties from observables (Davidzon et al. 2019), and calibration of the colour-redshift relation (Masters et al. 2017, 2019), enabling direct photo-z calibration (Buchs et al. 2019; Wright et al. 2020; Hildebrandt et al. 2021).

This work makes use of SOMs for dimensionality reduction and unsupervised classification to identify on-sky areas of observation (or simulation) with correlated observing conditions, as
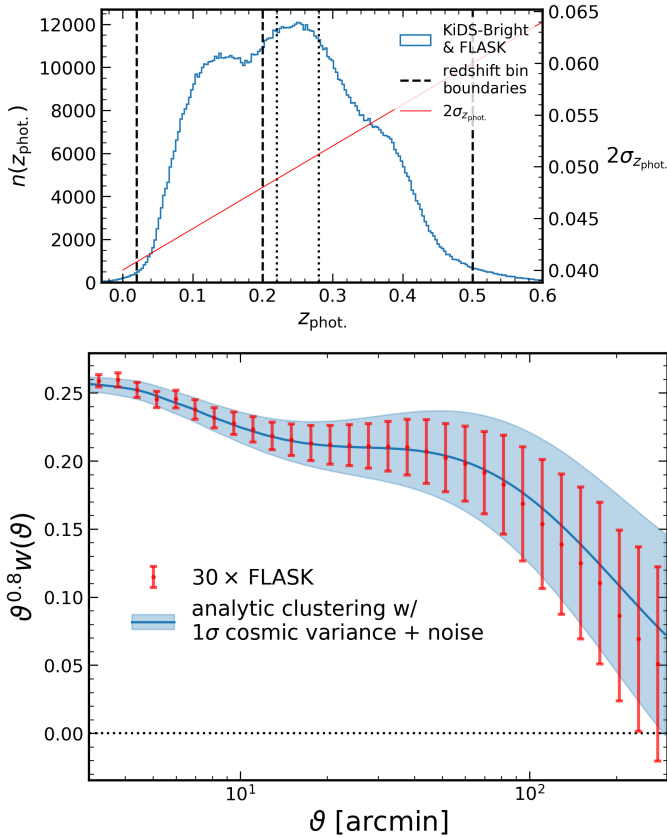
Fig. 3. *Top:* photometric redshift distribution of the KiDS-Bright GAMA-like KiDS DR4 bright sample (Bilicki et al. 2021), with the $z_{phot.} = \{0.02, 0.2, 0.5\}$ redshift bins (dashed lines) employed in our clustering analysis (Sect. 6) and the additional bins $z_{phot.,1a} = \{0.02, 0.22\}$ and $z_{phot.,2a} = \{0.28, 0.5\}$ (dotted lines) defined to have minimal photo-z overlap, as reckoned by the 95% scatter (red line); $2\sigma_{z_{phot.}} \sim 0.05$, at that redshift. We use this $n(z_{phot.})$ distribution to generate angular power spectra and FLASK mock galaxy catalogues. *Bottom:* full redshift-range angular clustering (red) averaged over 30 independent FLASK lognormal random field realisations of the input power spectrum, which is displayed in blue (see Eq. (5.2)) with the theoretical 1σ error for a KiDS-Bright-like survey, with 0.36 galaxies per square arcminute over 900 deg². The galaxy bias is set to unity. Error bars are the root-diagonal of the covariance across the 30 realisations. We note that for the 30 data points shown, the covariance over just 30 realisations of the field is quite noisy. We are clearly able to recover the analytical input cosmology with these 30 realisations from FLASK.

indicated by various systematic survey variables. These variables, such as those describing atmospheric effects and Galactic foreground properties, ought to correlate with phenomena causing systematic alterations to the observed galaxy number density in a wide-field survey. The variables define a systematics-vector[2] V per galaxy, which jointly describe an $\mathcal{R}^n$ [where $n = \text{length}(V)$] dimensional space to be mapped by the SOM. While this space can hypothetically consist of all derived data within the galaxy catalogue, in practice it is beneficial to select variables from the data products available that are likely to trace the true galaxy density variations, so as to allow the maximum density-variation information to be encoded in the SOM.

We therefore explore different choices of tracers that we detail in Table 1.

The SOM algorithm starts by instantiating a grid with user-specified dimensions, for example 100×100 for a two-dimensional SOM containing 10^4 cells. Each cell is then assigned a randomised weights-vector W of the same length as the galaxy systematics-vectors, that is, the number of systematic-tracer variables n. To train the SOM, galaxy systematics-vectors V are then chosen at random and presented to the SOM lattice. At each step of the training, the SOM cell with weights W most closely matching the training galaxy systematics-vector V is termed the best-matching unit (BMU). The match is typically quantified through the Euclidean distance d between the SOM cell weights-vector W and the galaxy systematics-vector V as

$$d = \sqrt{\sum_{i=1}^{n} (V_i - W_i)^2}, \qquad (3.1)$$

where the minimum d over the grid belongs to the BMU. Next we identify SOM cells within some radius $\sigma(t)$ (the neighbourhood) of the BMU, and modify their weights-vectors W to be closer to V, with more significant modifications for cells nearer the BMU. The resulting weights-vectors are given by

$$W(t + 1) = W(t) + L(t)\,\Theta(t, \sigma)\,[V(t) - W(t)], \qquad (3.2)$$

where t denotes a time-step (i.e. the presentation of a new training galaxy to the SOM), the learning rate L sets the strength of modifications, and Θ implements the distance-to-BMU dependence thereof. The final feature is that all of (i) the radius σ within which cell weights-vectors are to be modified, (ii) the learning rate L, and (iii) the BMU-distance dependence Θ, are exponentially decaying with each time-step, hence their dependence upon t. In this way, the SOM converges to a final representation as the training data are exhausted. Once all galaxies have been presented to the SOM, each cell on the SOM grid carries a weights-vector describing some unique position in the n-dimensional systematics parameter space, and the full collection of 10^4 cells spans the entire space sampled by the galaxies.

In our implementation, the resulting 2D map then represents the landscape of possible systematics-vectors realised by the data in question. By computing the distances between points in the space, we can then divide the landscape into N_{HC} maximally separated hierarchical clusters (see Wright et al. 2020, for a description of this process). Briefly, hierarchical clusters are defined by assigning each SOM cell to its own cluster, and then iteratively combining the two least-separated clusters (in our case by Euclidean distance between the cell weight-vectors W) into one, until only a single cluster remains. At each iteration, the cluster centres are recomputed using the Lance-Williams dissimilarity formula (see Defays 1977; Iezzi 2014), invoking complete-linkage clustering to generate the most similar clusters. This iterative process constructs a cell-merger dendrogram that can then trivially be sliced at the desired number of clusters N_{HC}. Each cluster of cells then contains a unique subset of the total galaxy sample, which is described by similar systematics-vectors V. In this way, the combination of the SOM and hierarchical clustering is able to construct N_{HC} distinct groups of sources with similar systematics properties, but which occupy non-contiguous areas of the sky. The average total area on-sky that is spanned by each of the N_{HC} groups is therefore determined by the area of the dataset in question (i.e. ~900 deg² for KiDS-Bright) and N_{HC}.

From the on-sky distribution and galaxy-count corresponding to each hierarchical cluster, we can compute an estimate of

---

[2] These are simply rows in a galaxy catalogue, where each element is a number describing the amplitude of some potential source of systematic error at the location of the galaxy.

Table 2. Analysis choices for various self-organising map and randoms-creation configurations.

Setup ID	N_{HC}	res (smth) [arcmin]		Systematics
T1mock	100	5	(0)	A1,A2,A3,B1,B2,B3,C1,C2
100A	100	2.8	(0.1)	MU,psf_ell,psf_fwhm
800A	800	2.8	(0.1)	MU,psf_ell,psf_fwhm
800Ares2	800	2	(0)	MU,psf_ell,psf_fwhm
100B	100	2.8	(0.1)	MU,psf_ell,psf_fwhm,MAG_LIM_r
800B	800	2.8	(0.1)	MU,psf_ell,psf_fwhm,MAG_LIM_r
100C	100	2.8	(0.1)	MU,psf_ell,psf_fwhm,MAG_LIM_r,gaia_nstar,EXTINCTION_r
800C	800	2.8	(0.1)	MU,psf_ell,psf_fwhm,MAG_LIM_r,gaia_nstar,EXTINCTION_r

Notes. Columns are: setup identifier, number of hierarchical clusters, resolution on-sky, and systematic-tracers for training. Various configurations for our SOM and organised randoms-creation pipeline. Throughout the text, we refer to these by their setup ID. This shorthand typically just indicates the number of hierarchical clusters defined on the SOM, e.g., 800 or 100, and the number of systematic-tracer variables used in training: A is just the PSF and threshold parameters, B then includes the r-band magnitude limit, and C further includes Galactic dust extinction and stellar density (all described in Table 1). The exception T1mock refers to the SOM trained against artificial systematic-tracers A1-3, B1-3, and C1-2, which are described in Sect. 4.1. N_{HC} is the number of hierarchical clusters to be defined on a given SOM, 'res (smth)' are the resolution and Gaussian smoothing width in arcminutes of the sky-grid used to group galaxies and populate random fields, and 'systematics' refers to the set of systematic-tracer variables presented to each SOM in training (see Table 1). All SOMs employed in our analysis have dimensions 100×100.

the galaxy number density associated with each region of the high-dimensional systematics space (i.e. the space of systematic-tracer variables from Table 1 and Figs. 1 and 2). Our randoms creation algorithm uses this information to populate the survey volume with variably dense but locally unclustered random points. Moreover, each random point is constructed as a copy of one of the training galaxies from the hierarchical cluster to which it belongs, carrying all of the parent galaxy physical and photometric properties: a clone. Clones are therefore scattered only within the (again, non-contiguous) areas of sky represented by the parent hierarchical cluster. As a result, galaxy sample selection effects and associated selection of specific systematic clustering patterns are easily reproduced in the randoms catalogue by a simple selection of clones satisfying the chosen galaxy selection function. Acting as the reference points in a galaxy clustering measurement, the randoms ought to then compensate for density variations correlating with any combination of the tracked systematic variables, for any selection of galaxies used to train the SOM.

The number of clusters N_{HC} is an important tuning parameter for this analysis: in addition to determining the area-per-cluster, N_{HC} must also trade off against the amount of discretisation of the systematics space to be mapped. If we invoke too few clusters in our high-dimensional space, each will span a wide range of observational (i.e. systematic) and cosmological regimes, and could thus lose discerning power as a result (indeed, $N_{HC} = 1$ returns completely uniform randoms by definition). Conversely, invoking too many clusters in the same space will result in too much freedom, and begin to over-fit the number density versus systematics relation. This over-fitting can itself be pathological, as we show below, if the individual clusters begin to trace the cosmic (i.e. not systematic) structure.

Our various choices for SOMs are detailed in Table 2, and Fig. 4 displays SOMs of the configuration 100A after training on KiDS-Bright. The bottom left, top left and top right panels are coloured by the MU,psf_fwhm,and psf_ell (Table 1) values for these cells, and the bottom right panel connects them to the galaxy density contrast, $(n_{gal} - \langle n_{gal} \rangle)/\langle n_{gal} \rangle$, in 100 hierarchical clusters (denoted by discrete patches of colour with black borders). Inspecting any region of the density contrast map, we can easily identify correlations between density contrast and various systematic parameters. For example, we can note that the upper

left quadrant of the density contrast SOM is populated by clusters of generally below-average (i.e. blue) density contrast. This is clearly correlated with above-average values of the r-band imaging surface brightness limit MU_THRESHOLD. The opposite effect (i.e. above-average density contrast) is seen for below-average values of MU_THRESHOLD, except for clusters where the value of the PSF full width at half maximum (FWHM) is above average; in this case, the density contrast is once again below average. These conclusions are easily drawn from the SOM and are a significant advantage of our methods presented here.

4. Validating SOMs with artificial systematics

First, we elucidate the capabilities and limitations of SOMs in recognising galaxy density-systematic correlations through the use of artificial systematic fields with complex correlations to the depletion of the galaxy number density.

4.1. Creating artificial systematic density fluctuations

Figure 5 shows on-sky distributions of 1 arcmin^{-2} FLASK galaxies, colour-coded by our set of eight artificial systematics variables, which were designed to mimic realistic spatial patterns in KiDS-like wide-field observations. A-type variables vary smoothly over large angles, in a manner similar to Galactic foregrounds. B-type variables have a two-dimensional Gaussian form, which varies independently and discretely in $1 \times 1 \deg^2$ tiles, thereby mimicking telescope and camera effects such as PSF variations over the focal plane. Finally, C-type variables vary discretely between tiles but are constant within them, thereby mimicking per-exposure effects such as limiting depth variations that arise from the use of a step-and-stare observing strategy (meaning that each tile is only observed once, per-band, over the course of the entire survey, thereby increasing sensitivity to the variable sky brightness on any given night). Each effect varies in the range $[0, 1]$ for simplicity. By construction, these analytic systematics have no serious outliers.

To create spurious density modes as multivariate functions of our artificial systematics, we invented an independent depletion function for each parameter, shown in Fig. 6 as red dashed lines. These curves each describe the probability that an object will be discarded from the catalogue as a function of the
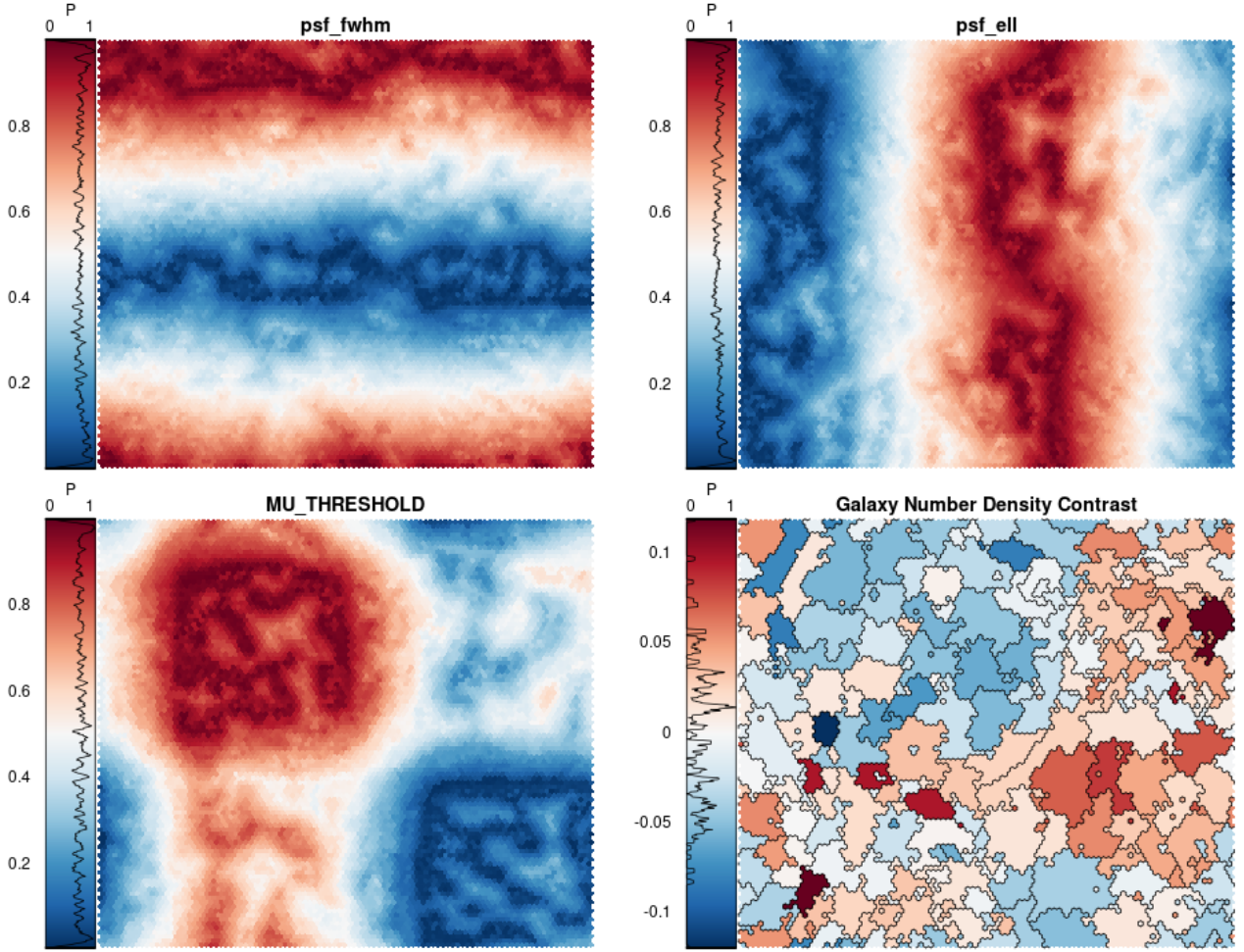
Fig. 4. Self-organising maps, with dimension 100×100, coloured (*top left, top right, bottom left*) by the systematics values taken on by each cell during the training procedure described in Sect. 3. In the density contrast panel (*bottom right*), colours and black borders mark the 100 hierarchical clusters (see Sect. 3) defined according to groupings of cells with similar systematics-vectors for KiDS parameters `psf_fwhm,psf_ell,MU_THRESHOLD` (Table 1). Systematic-tracer variables are linearly mapped onto the interval $[0, 1]$ before being passed to the SOM, hence the colour-bar ranges. The density contrast panel (*bottom right*) maps clusters of SOM pixels from their vectors of systematic-tracer values back onto a relative number density on-sky and reveals almost all systematic density fluctuations to be at $\lesssim 10\%$, as reckoned by this SOM configuration (`100A`; Table 2).
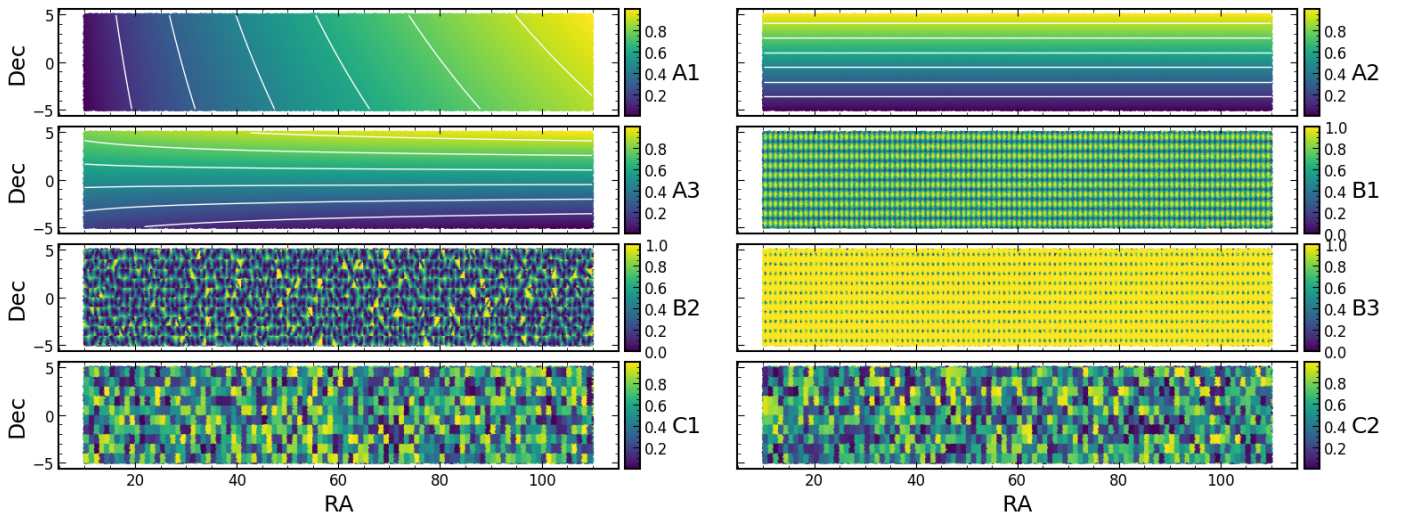
Fig. 5. FLASK-generated uniform random field, binned hexagonally (cell scale ~10 arcmin) in RA/Dec and coloured according to our various artificial systematic-tracer variables. A-types vary smoothly over large angles, B-types vary over 1 deg^2 tiles with a 2D Gaussian form, and C-types are single-valued for each tile. Iso-contours (white lines) are lain over A-type variables for clarity, and to illustrate the subtle difference between A2 and A3.
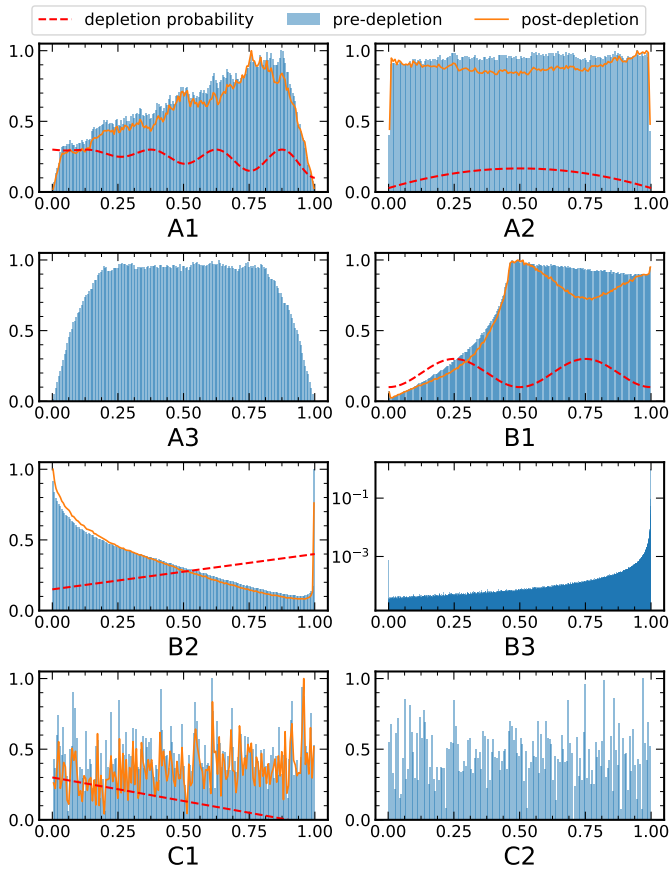
Fig. 6. Histograms, each normalised to a maximum of unity, of the artificial model systematics (see Fig. 5) we imposed upon our FLASK lognormal random field galaxies, both before (blue) and after (orange) applying our probabilistic depletion functions, shown as red dashed lines. Dummy variables (A3,B3,and C2) have no depletion applied. They are intended to distract the SOMs.

position of the object in systematics space, thereby depleting the galaxy number density at that point in the space. We note that some systematic variables cause no depletion (A3,B3,and C2) and act instead as dummy parameters for the SOM to navigate. In Fig. A.1 we show the spatially variable excess probability of depletion resulting from the depletion functions of Fig. 6 as applied to the systematics-maps of Fig. 5.

We applied each depletion function individually[3] to a set of 30 FLASK mocks. We then trained a SOM using each depleted mock galaxy sample with the T1mock configuration from Table 2 and created a hierarchy of cell clusters on each trained SOM. The number densities of the clusters should then reflect the input depletion functions per individual systematic from Fig. 6. This test is thus a simple verification of the ability of the SOM to characterise galaxy-density variations as a function of some systematic effect, with a particular pattern on-sky. Furthermore, we can use this test to explore possible systematic patterns that may cause the SOM to fail in its corrective pursuit.

---

[3] The case of multiple potentially correlated systematic biases working in confluence is more realistic, but a SOM could then only infer the combined multivariate depletion function, making an assessment of success more complicated. We explore the consequences of multiple biases when we consider data-driven systematic density fluctuations in the next section.

4.2. Characterisation of artificial systematic fluctuations

We find that the SOMs are able to recover individual depletion functions with good accuracy. Figure 7 displays number densities versus systematic values for hierarchical clusters defined on the T1mock SOM (Table 2), along with our input depletion functions in red. Error bars are the root-diagonal of the covariance over 30 realisations, and we display the result for both lognormal (green) and uniform (orange) random fields (GRF and URF respectively). The additional noise seen in n_{gal} for the GRF case is due to the simulated cosmic structure, and can be seen to correlate with systematics that localise finite regions on-sky (i.e. A and C types). In contrast, B-type parameters show little difference between uniform and lognormal random fields, as individual systematics clusters cover a much less localised on-sky area. They are consequently much less sensitive to (local) variations in cosmic number density. We also verified that the SOM does not create depletion functions where none exist. Figure A.2 shows that the SOM recovers the mean galaxy density in all cases when training the SOM on the same parameters but without any depletion of galaxy number densities.

We note and address the presence of some irregularities in the distributions of Fig. 7, bearing in mind that the test here is merely to recover the form of depletions, and that these are superseded by two-point statistics in the next section. In extrema, A-type parameters appear to have outlier clusters with very high or low number density. These are artefacts from binning galaxies on the 5 arcmin Cartesian grid. At the footprint edges, some grid cells are only partially filled, resulting in seemingly under- or overdense clusters. The B1 clusters are also affected by footprint edges, where B1 is predominantly low (Fig. 5), and the amplitude of recovered fluctuations is suppressed with respect to the expectation for decreasing values of B1. The distribution of B1 (Fig. 6; blue histogram) drops sharply below values of ~0.5, whilst the probability of depletion rises (Fig. 6; dashed red line), resulting in a thin, sparsely populated, grid-like distribution of objects across the footprint that begins to noisily sample the mean density and causes further suppression of inferred fluctuations when the periodicities of the B1 fluctuation and the on-sky grid are misaligned. We verified that such effects can be mitigated by modifying the on-sky grid used to discretise the galaxy sample, but we caution that a reduction in the grid size can cause artificial masking of empty patches of sky, given sparse galaxy data. In practice, these effects can be mitigated by ensuring mutual masking of the randoms and galaxy catalogues, and by tuning the on-sky grid size and grid-smoothing parameters to account for the range of number densities in the dataset under investigation. We further note that these artefacts are also seen in Fig. A.2, demonstrating that they are not related to the applied systematic density fluctuations.

Given our experience with this test, we settled on a fiducial setup with an on-sky resolution of 2.8 arcmin, and invoked a Gaussian smoothing kernel with a standard deviation of 0.1 arcmin. This allows empty grid-cells on the sky to be affected by their populated neighbours, thus ensuring that the randoms respect the survey mask even for sparsely distributed galaxy data (such as the KiDS-Bright sample).

5. Validating organised randoms with data-driven systematics

The ultimate goal for our organised randoms is to provide a simple and unsupervised method with which to debias galaxy clustering signals upon measurement. In the previous section,
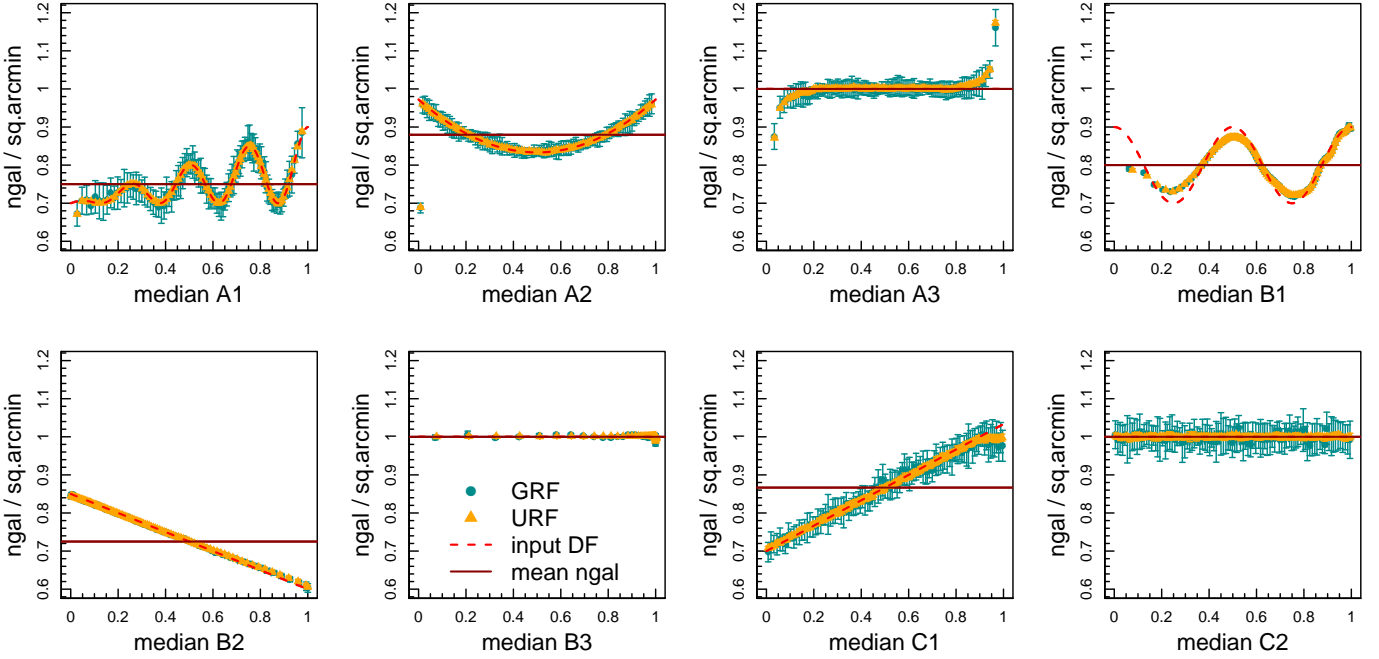
Fig. 7. Recovery by the T1mock SOM (Table 2) of the artificial biases applied to 30 FLASK realisations (see Figs. 5 and 6) of random fields. Panels show the galaxy number density (arcmin^{-2}) of 100 hierarchical clusters, defined on the SOM, against the median systematic value in each of these clusters. Green points give the results for SOMs trained against depleted lognormal random fields (GRF), and orange points give the same for uniform random fields (URF), each generated using FLASK. The functions describing systematic galaxy depletion relative to our artificial systematic parameters (from Fig. 6) are converted into expected number densities and shown here as dashed red lines ('input DF'; depletion function). The solid horizontal lines indicate the global average number densities expected per systematic after depletion. Points where both our URF and GRF data sit systematically away from the expected values are artefacts due to footprint edge-effects in our FLASK simulations. They do not affect the conclusions of this test (see Sect. 4.2).

we demonstrated that self-organising maps are capable of identifying systematic density fluctuations in data. We now use the SOMs to infer realistic systematic fluctuations from KiDS-Bright data and apply them to simulated FLASK catalogues. Training organised randoms against these biased data, we can then assess their ability to recover unbiased two-point statistics.

We first note that many typical approaches to systematics-handling for galaxy clustering feature the correction of pixel density-systematics one-point correlations (e.g., Suchyta et al. 2016; Elvin-Poole et al. 2018; Rezaie et al. 2020; Kitanidis et al. 2020; Vakili et al. 2020). We find these metrics to be unsuitable, at least in isolation, for validation of our own methods, where a strict correction of individual pixel density-systematic correlations is not necessary for the recovery of unbiased clustering two-point functions, and may be a misleading measure of performance. A perfect performance with regard to one-point density-systematic correlations could mask a total failure of our methods at the two-point clustering level, featuring a dramatic suppression of real clustering signals. For further details, see Appendix B.

5.1. Galaxy clustering

For all measurements of galaxy clustering two-point correlations, we considered the angular correlation function $w(\vartheta)$, and made use of the public software TREECORR[4] (Jarvis et al. 2004). The standard Landy-Szalay (Landy & Szalay 1993) estimator

for angular galaxy clustering is given as

$$\hat{w}^{ij}(\vartheta) = \frac{\Theta(\vartheta)\left\{D_i D_j - D_i R_j - R_i D_j + R_i R_j\right\}}{\Theta(\vartheta)\left\{R_i R_j\right\}}, \quad (5.1)$$

where i, j denote the galaxy samples being correlated; photo-z bins, in this work. DD, DR, and RR denote normalised[5] pair-counts between galaxies D and randoms R, and the bin-filter operator $\Theta(\vartheta)$ rejects pairs with separations falling outside of the angular bin centred at ϑ. We considered both auto- ($i = j$) and cross-correlations ($i \neq j$) when evaluating our organised randoms.

Assuming a flat universe and a linear galaxy bias, and under the Limber and flat-sky approximations, we obtain a theoretical expectation for $w(\vartheta)$ between galaxy samples i and j as follows (Limber 1953; Loverde & Afshordi 2008):

$$w^{ij}(\vartheta) = 2\pi \, b_g^i b_g^j$$
$$\times \int_0^\infty \frac{\mathrm{d}\ell}{\ell} J_0(\ell\vartheta) \int_0^{\chi_h} \mathrm{d}\chi \, \frac{p^i(\chi)p^j(\chi)}{\chi^2} P_\delta\left(\frac{\ell + 1/2}{\chi}, \chi\right), \quad (5.2)$$

where b_g denotes the linear galaxy bias, ℓ the angular frequency, J_0 the zeroth-order Bessel function of the first kind, χ the comoving distance to a given redshift, $p(\chi)$ the normalised comoving distance distribution of the sample, and P_δ the matter power spectrum with non-linear corrections (for which we employed the halofit model of Smith et al. 2003; Takahashi et al. 2012).

---

[4] https://github.com/rmjarvis/TreeCorr

[5] Normalisation is by the product of total counts $N_i N_j$ for samples i and j, that is, the count of all possible pairs. For auto-correlations, N is subtracted from the product, as a galaxy cannot be paired with itself.

The right-most integral is equal to the (Limber-approximate) angular power spectrum C_ℓ.

We note here that an alternative approach to the construction of organised randoms, which correct the clustering through the R terms in the above estimator (Eq. (5.1)), would be to correct the galaxy densities directly, for example using galaxy weights, and then employ standard, uniform randoms in the estimator (as in Ross et al. 2011; Elvin-Poole et al. 2018; Rezaie et al. 2020; Kitanidis et al. 2020; Vakili et al. 2020; Wagoner et al. 2021). Our SOM methods lend themselves naturally to this approach, and we intend to explore this with angular power spectra in an upcoming KiDS analysis.

5.2. Realistic biasing of FLASK fields

An application of SOMs to real data lacks a clear notion of truth: we do not know exactly how our tracer variables relate to the deprecation of observed galaxy densities. Thus we cannot know whether the systematic density relations inferred by the SOMs are contaminated by cosmological density variations; for example, it could be that a real, local north-to-south galaxy density gradient happens to correlate with some systematic-tracer. The SOMs could then falsely identify such a variable to be tracing the source of the density gradient. Any randoms organised accordingly would then act to erase real cosmological clustering signals. Moreover, we do not know the true underlying clustering signal: if our corrections were already sufficient to recover the unbiased clustering, but one-point pixel density-systematics correlations (e.g., Elvin-Poole et al. 2018; Rezaie et al. 2020, see Appendix B) still revealed correlations with systematics, how would we know to stop?

Here we assess the performance of organised randoms in recovering unbiased clustering signals, and also address the question of over-fitting to spatial, possibly cosmological trends in galaxy density. For the reasons detailed above, this is difficult to do for real galaxy data such as KiDS-Bright, where any biases are already present. We sidestepped these concerns with synthetic galaxy distributions from FLASK. The test method is as follows: we fed real KiDS-Bright galaxies and systematics data to a SOM and inferred the spatial pattern of depletion. It is unimportant whether the pattern is contaminated by LSS at this stage. We performed a nearest-neighbour interpolation in RA/Dec to port the real spatial distributions of systematics from KiDS-Bright onto FLASK galaxies simulated within the same mask. We then applied the SOM-inferred systematic density fluctuations to the mocks, probabilistically, as

$$P_{\text{depl.}}(\boldsymbol{x}) = 1 - \frac{n_{\text{gal}}}{n'_{\text{gal}}}[1 + m\,\delta_{\text{OR}}(\boldsymbol{x})], \qquad (5.3)$$

where $P_{\text{depl.}}(\boldsymbol{x})$ is the probability that a galaxy at position $\boldsymbol{x}$ will be lost; a uniform random draw in the range $[0, 1]$ must exceed $P_{\text{depl.}}(\boldsymbol{x})$ for the galaxy to be retained. n_{gal} and n'_{gal} are the target and initial FLASK number density. We can only remove galaxies from the mocks, so we initialised the FLASK realisations with $n'_{\text{gal}} = 0.72$ galaxies arcmin^{-2} and then generated systematic under- or overdensities with respect to the mean (target) KiDS-Bright density of $n_{\text{gal}} = 0.36$ arcmin^{-2}. $\delta_{\text{OR}}(\boldsymbol{x})$ is the density contrast at position $\boldsymbol{x}$ sourced by systematics[6] according to the

---
[6] This inferred density contrast from systematics forms the basis for our organised randoms generation algorithm, although for this test we created organised randoms only after training separate SOMs against the depleted FLASK mocks.

SOM, and m is a scalar variable that we can use to manually modify the amplitude of the applied depletion whilst retaining its functional relation to the KiDS-Bright systematics distribution. Taking $m > 1$ would intensify the depletion relative to that present in KiDS-Bright, whilst $m < 1$ would yield a comparatively soft depletion. We display an example map of δ_{OR}, inferred from KiDS-Bright galaxies, in Fig. 8.

Repeating this procedure for many realisations of the underlying cosmology, we created the truth case of a single, global pattern of galaxy depletion. Running the SOMs again, now against the depleted FLASK mocks, we assessed how consistently they are able to recover this truth for many different realisations of the constant cosmological background. If the SOMs are able to retrieve the fixed systematic depletion pattern from many independent realisations of the cosmic structure, then we can assert that the inferred $\delta_{\text{OR}}(\boldsymbol{x})$ is uncontaminated by cosmology.

Having thus created many realisations of KiDS-Bright-like data for which we can turn systematic biases on or off, we ran our SOMs and assessed the corrective performance of our organised randoms with various measurements of $w(\vartheta)$. A caveat is that our depletion of the FLASK galaxies was derived from runs of the SOM against KiDS-Bright, thus the bias can be said not to be entirely realistic (although far more so than the artificial systematics of the previous section, for which the accuracy of capture was excellent) but dependent upon the configuration of the initial SOM. We therefore considered many different configurations along with modifications (through the m parameter from Eq. (5.3)) to the intensity of depletion that we applied to the FLASK mocks.

Over the course of testing, we recognised that systematic-tracer set B (Table 1) was a relatively uninformative yardstick between sets A and C; therefore we limit our discussion from here to parameter sets A and C, which are instructive for our work with the relatively unbiased KiDS-1000 bright sample. We note also that the soft systematic biases present in KiDS-Bright may well be treatable with a simpler linear regression formalism (e.g., Ross et al. 2011; Elvin-Poole et al. 2018), although we have not tested this claim. Our more complex SOM-derived model is, however, more likely to be necessary for our work with the faint KiDS shear sample (Wright et al., in prep.), which features stronger systematic biases.

5.3. Correction of data-driven systematic fluctuations

We devised a battery of $w(\vartheta)$ tests to assess the performance of our organised randoms at the two-point level. The key variables that we changed were (i) the parameters used to train a SOM against KiDS-Bright and (ii) the number of hierarchical clusters N_{HC} defined on that KiDS-Bright SOM. These determine the spatial depletion to be ported onto FLASK realisations, as described in Sect. 5. We also changed (iii) the parameters used to train a second SOM against biased FLASK realisations and (iv) the number of hierarchical clusters N_{HC} defined on that FLASK SOM. This allowed us to break any circularity by comparing independent SOMs. Finally, we also experimented with (v) intensifying the depletion through the m parameter (Eq. (5.3)).

We find that the performance and flexibility of our organised randoms are excellent. The randoms were able to consistently mitigate biases in FLASK fields, even when they had limited sensitivity to smaller scales (through reduced N_{HC} with respect to the SOM used to infer systematic modes), or when trained on incomplete systematics information. Figure 9 displays some performance examples of organised randoms, which we continue to explore in Appendix C. The panel titles in these figures give
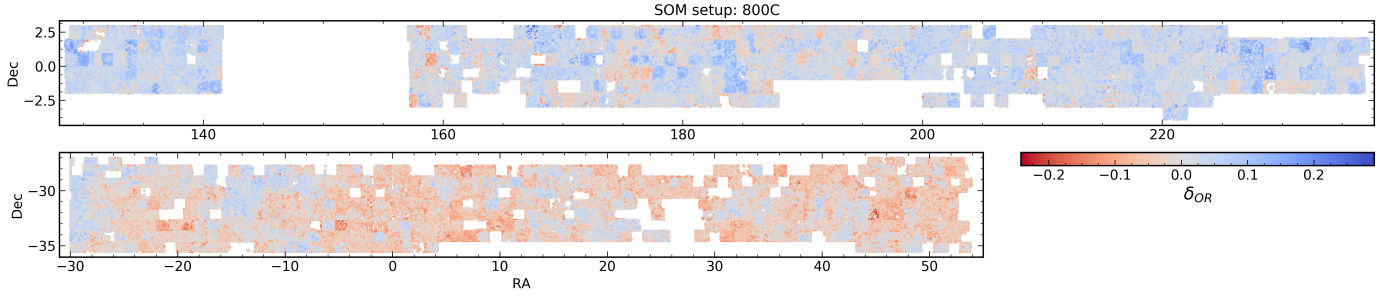
Fig. 8. Systematic density contrast δ_{OR} (Eq. (5.3)) inferred for the KiDS-North (*top panel*) and KiDS-South (*bottom panel*) areas of the KiDS-1000 bright sample (KiDS-Bright) by the 800C SOM setup. We use these maps to construct our organised randoms, populating the footprint to mirror the systematic density modes.
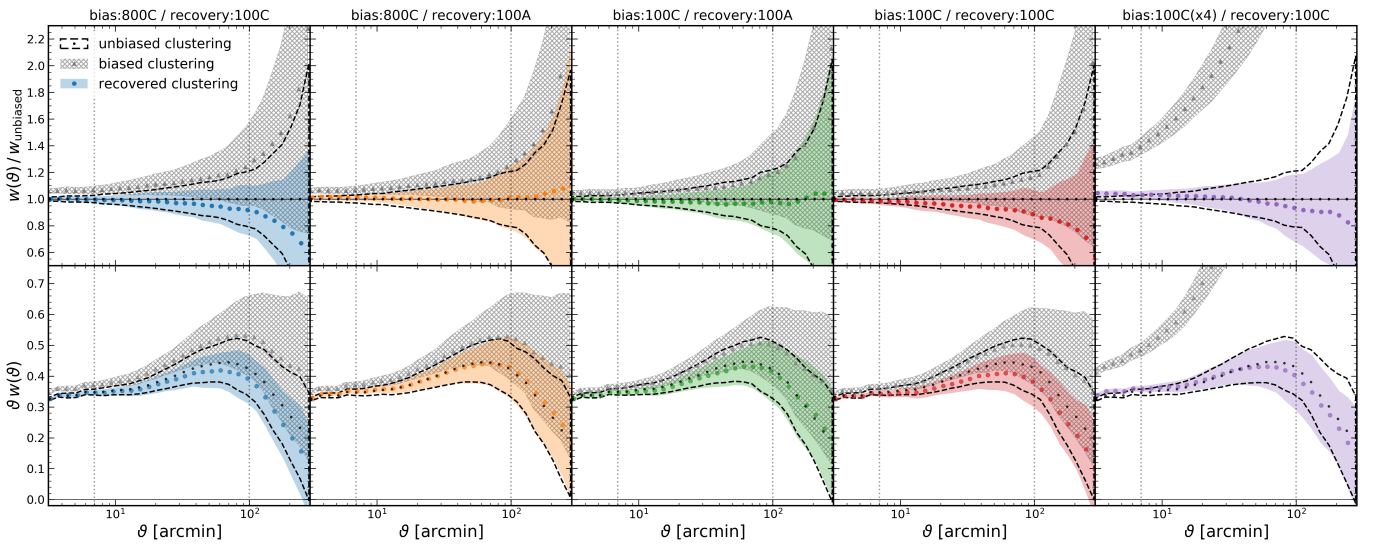

Fig. 9. Angular clustering correlation functions $w(\vartheta)$ measured in FLASK fields after they have been depleted (as described in Sect. 5.2; Eq. (5.3)) according to the output of various SOMs trained against the KiDS-1000 bright sample (KiDS-Bright). For each configuration, panel titles give the bias:SOM trained against KiDS-Bright, and the recovery:SOM trained against biased FLASK data to create organised randoms. *Top:* ratios of measured clustering signals to the true, unbiased clustering. *Bottom:* unbiased angular clustering signature (black points and dashed curves; measured with uniform randoms on unbiased FLASK fields), compared with biased (grey triangles and hatching; measured with uniform randoms) and recovered (coloured points and shading; measured with organised randoms) clustering signals measured in the depleted fields. Errors are the root-diagonal of the covariance over 30 FLASK realisations, and all are given to $\pm 1\sigma$. *From left to right*, the FLASK fields are biased according to δ_{OR} from the 800C, 800C, 100C, 100C, and 100C SOMs, trained against KiDS-Bright. The m-parameter (Eq. (5.3)) is set to unity in all but the purple panel (where $m = 4$). Each correlation measured in the depleted fields using organised randoms displays improved consistency with the unbiased signal, compared to the signals measured using uniform randoms.

the SOM setups trained (i: bias) against KiDS-Bright to infer the clustering bias, and (ii: recovery) against the biased FLASK mocks to recover the true clustering with organised randoms. Multiplicative factors in the panel titles indicate where $m \neq 1$. Black points and dashed curves give the unbiased clustering signal $\pm 1\sigma$ errors from 30 FLASK realisations (equivalent to red points in Fig. 3). Grey triangles and hatched curves are the measured (biased) clustering after depletion of the FLASK fields, and solid filled curves show the corrected clustering, measured with organised randoms. In all meaningful (see Appendix B, where we present a deliberate failure mode: 800Ares2) bias and recovery cases we considered, our organised randoms yield clustering correlations that are more consistent with the truth than the biased signals (i.e. those measured with uniform randoms).

Figure 9 shows that the recovery:100C organised randoms yield an effective correction of clustering biases from the bias:800C SOM (blue panel), with the recovered signal much closer to the unbiased measurement (mean absolute deviation over $7 - 100$ arcmin: $1.48\sigma \rightarrow 0.31\sigma$, where σ is the uncertainty

in the unbiased signal). This indicates that a relatively insensitive SOM setup (100C) is able to characterise the more complex biases inferred from a SOM with greater sensitivity to small-scale systematic structure (800C). We see an even better correction ($1.43\sigma \rightarrow 0.09\sigma$) when passing incomplete systematics information to the SOM, as demonstrated by bias:800C versus recovery:100A (orange panel). This demonstrates that our SOM methods are able to correctly infer systematic density fluctuation patterns even when the patterns are sourced by systematic-tracers that are unknown to the SOM. Inter-parameter correlations therefore serve to make organised randoms robust against missing training variables, and the additional freedom afforded to N_{HC} clusters in a systematics space of reduced dimensionality can improve the accuracy of the correction.

We note that these statements are specific to this particular setup. It is not generally true that SOMs will be able compensate for unknown systematics, especially if those unknowns levy significant density fluctuations and are uncorrelated with the variables presented in training of the SOM. Thus the

selection of important systematic-tracers remains an intrinsic challenge for this type of study (one that can and should be tested with simulation-based methods similar to those presented here; also see Weaverdyck & Huterer 2021, who introduced methods for selecting the most important features from a large initial set). However, this setup does demonstrate that the elimination of redundant systematic-tracer features can improve the performance of clustering corrections by organised randoms (much as Rezaie et al. 2020 demonstrated with their artificial neural network methods).

Homogenising the scale sensitivity (through N_{HC}) between bias and recovery SOMs, but keeping the incomplete systematics set for recovery (100C versus 100A; green panel), we still see an excellent recovery of the unbiased clustering signature ($0.94\sigma \rightarrow 0.28\sigma$), as we might expect for a less complex bias. For identical bias and recovery setups 100C (red panel), we begin to see a slight over-correction by the organised randoms. This comes about as clusters on the SOM begin to over-fit to the cosmic structure around the density-systematics relation.

Whilst the recovery is still preferable to the biased statistic here ($0.96\sigma \rightarrow 0.54\sigma$), we acknowledge that over-corrections could be problematic for cosmological inference (see Wagoner et al. 2021; Weaverdyck & Huterer 2021). However, the KiDS-Bright sample is a bright subset of the KiDS-1000 photometric sample, specifically chosen to be ~5 magnitudes shallower than the survey flux limit, and thus is less sensitive to systematic detection failures; these data are already relatively unbiased. For more pathological biases, resulting in a higher amplitude of systematic density contrast, such over-fitting to cosmic structure is less likely to occur. Thus we expect our randoms to perform even better for samples with stronger systematics imprints, for instance the KiDS-1000 shear sample, dominated by faint galaxies. We tested this assertion using the same setup, but setting $m = 4$ in Eq. (5.3). In this case [100C($\times$4) versus 100C; purple panel], a massively inflated bias of 12.17σ is once again reliably corrected to 0.34σ.

This result is particularly important, as measuring accurate clustering statistics for faint, systematics-dominated samples has historically been extremely challenging; we therefore often choose bright subsamples for clustering analyses (also for typically more reliable estimation of the galaxy photo-z; see Porredon et al. 2021). Our companion letter, Wright et al. (in prep.), thus presents an application of our organised randoms pipeline to the KiDS-1000 shear sample, finding excellent and robust performance in correcting for systematic clustering bias. Moreover, the shear sample is a factor of ~20 denser on-sky than the bright sample, enabling us to increase the resolution of the cartesian RA/Dec grid whilst still respecting the survey mask (see Sect. 4.2). In this way, we become more sensitive to small-scale systematic density fluctuations at fixed N_{HC}, which Wright et al., (in prep.) show is important for these faint data.

6. Clustering in the KiDS-1000 bright sample

Having validated the performance of our SOMs and organised randoms in recognising systematic trends in galaxy density and removing their traces from synthetic galaxy two-point correlations, we applied our methods to measurements of galaxy clustering in the KiDS-Bright sample and compared tomographic cross-correlations with analogous signals measured in the highly complete GAMA (Driver et al. 2009; Liske et al. 2015) sample.

We used randoms organised by the SOM configuration 100A (detailed in Table 2), which we show (in Figs. 9 and C.1) to robustly improve the fidelity of signal recovery in all bias:SOM

scenarios from our FLASK tests (which vary in scale sensitivity through N_{HC}, and in the systematic-tracer parameters), thus we expect a reasonable correction even if the additional systematic-tracers (MAG_LIM_r, EXTINCTION_r, gaia_nstar) are in reality uncorrelated with the KiDS-Bright systematic density contrast. Moreover, the bias:800C/recovery:100A test case (Fig. 9; orange panel) demonstrates that the bias inferred from all parameters (set C: threshold, PSF FWHM or ellipticity, limiting magnitude, Galactic density of stars and extinction), with high small-scale sensitivity ($N_{HC} = 800$), is corrected with great precision by 100A organised randoms, having a residual bias of $<0.1\sigma$. Table C.1 summarises the performance of differently configured organised randoms and reveals 100A to be the net-best performer across all bias:SOMs.

We note that these particular SOM setups and parameter sets may not be ideal for other surveys with different areas, geometries, systematics imprints, etc., and that any organised randoms should be thoroughly tested with simulations, as we have done (in Sect. 5).

Figure 10 indicates the advantages of creating cloned galaxy randoms in which each random point is a clone of a real object used in training of the SOM, and clones are spatially restricted to on-sky areas occupying similar positions in the systematics space. The figure illustrates how we reproduced systematic density variations in a strip of KiDS-Bright ($172 < RA < 205$, $2 < Dec < 3$ degrees) for a coarse red and blue (bottom and top) sample selection[7] merely by restricting our organised randoms to the relevant clones. The figure shows the density contrast of 1 degree columns in RA, and we expand the radial dimension ($z_{phot.}$) to draw attention to non-physical density modes, that is, underdensities that are localised to single pointings, seen as fainter vertical strips of colour in the right-most column (e.g., at RA $\in [175.5, 197.5]$), and to the redshift evolution of galaxy number density, which varies for red and blue galaxies; these trends are mirrored in the 800Ares2[8] organised randoms (right column) in contrast with the uniform randoms (left column). The cloning utility will be fully realised in future work. Systematics that differentially affect arbitrary galaxy selections can be automatically compensated for, allowing for easy splitting of analyses into red and blue galaxies, for instance, or into bins of galaxy luminosity. For now, we focus only on redshift tomography.

In addition to the tomographic bins 1 and 2, with edges at $z_{phot.} = \{0.02, 0.2,$ and $0.5\}$, we cross-correlated two additional, non-overlapping bins, 1a and 2a, with edges at $z_{phot.} = \{0.02$ and $0.22\}$ and $z_{phot.} = \{0.28$ and $0.5\}$, respectively. The gap between the inner edges of the two bins is more than the typical 95% photo-z scatter at this redshift (Fig. 3, red line in the top panel), therefore we tested for a null cross-correlation between them. Barring the auto-correlations of the full (total) sample, we refer to correlations between bins i and j as i–j, that is, 1-1 is the auto-correlation of tomographic bin 1 and 1-2 is the cross-correlation between bins 1 and 2, etc.

We generated organised randoms with 20 times the number density of KiDS-Bright, so as to combat Poisson-noise in the relevant pair-counts (see Eq. (5.1)). Our measurements of angular clustering auto-correlations within the two primary redshift bins and across the full KiDS-Bright sample are displayed in Fig. 11.

---

[7] We defined the red and blue boundary at observer-frame $u - i = 2.66$, which sits in the trough of the bimodal colour distribution of KiDS-Bright.

[8] Chosen for illustrative purposes, as the inferred δ_{OR} is wider for this setup (see Appendix B) and the increased resolution results in more clearly visible systematic density modes in the organised randoms (discussed in Appendices B and C).
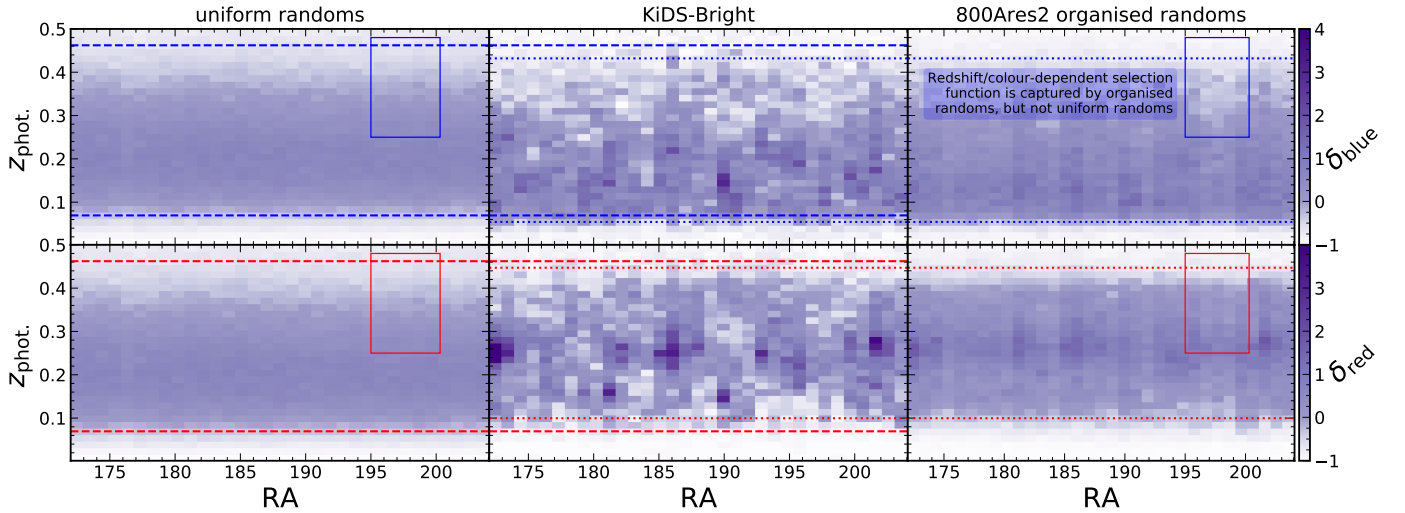
Fig. 10. Galaxy number density contrast δ in 2D bins of photometric redshift vs. RA for uniform randoms (*left*), the KiDS-1000 bright sample (KiDS-Bright; *middle*), and relevant clones from `800Ares2` organised randoms (*right*). Rows display blue (*top*) and red (*bottom*) galaxies, according to a boundary at observer-frame $u - i = 2.66$. We only show galaxies in the range $172 < $ RA $ < 205$, $2 < $ Dec $ < 3$ in order to reveal systematic density variations in the data and randoms. Unlike for uniform randoms, the organised randoms can be seen to reproduce the systematic trends in the KiDS-Bright data, namely under-densities over different pointings (e.g., RA $= 175.5, 197.5$), visible as vertical bands of fainter pixels, and the differential evolution of densities in the samples with redshift, as evidenced by the boxes (highlighting a specific volume with differential density as a function of sample colour) and horizontal lines (dashed lines give the $z_{phot.}$ corresponding to the 3rd and 97th percentile number counts for uniform randoms, and dotted lines give these for KiDS-Bright data). We elect to display `800Ares2` randoms here for a clearer illustration of the cloning mechanism, which is more subtle in our other favoured randoms (see Sect. 5.3).

Figure 12 displays the redshift bin cross-correlations in comparison with equivalently binned signals measured in the GAMA sample. We used our ANNz2 photo-z estimates to define these GAMA bins and employed the spatially uniform, windowed, cloned-galaxy randoms presented by Farrow et al. (2015). As we discuss in Appendix D, auto-correlations in GAMA are not suitable for validating KiDS-Bright correlations here, hence they are not displayed in Fig. 11.

For all $w(\vartheta)$ correlations in Figs. 11 and 12, we estimated errors with the delete-one jackknife technique. We divided the KiDS-Bright/GAMA footprints into N_{patch} roughly equal-area patches, and computed $w(\vartheta)$ upon the successive removal of individual patches. For these N_{patch} signals w_α, the covariance is then

$$\hat{C}_{jack.} = \frac{N_{patch} - 1}{N_{patch}} \sum_{\alpha=1}^{N_{patch}} (w_\alpha - \bar{w})(w_\alpha - \bar{w})^{\mathrm{T}}, \qquad (6.1)$$

where T denotes the conjugate transpose of a vector and $\bar{w}$ is the average of the N_{patch} measurements w_α. For KiDS-Bright, $N_{patch} = 31$, and $N_{patch} = 20$ for GAMA[9]. We note that for our chosen binning of 30 log-spaced bins in the range $3 \leq \vartheta \leq 300$ arcmin, so few jackknife samples will yield singular covariance matrices (see Hartlap et al. 2007). We are only interested in measurement errors (the square root of the matrix diagonal) here, and because we did not perform any fitting, we proceeded accordingly.

Figure 11 shows that the `100A` organised randoms make more sizeable corrections (blue points) to the measured total and 2-2 correlations (red points) than to the 1-1 correlation. This is perhaps to be expected because at higher redshifts the objects with faint apparent magnitudes, and predominantly small

angular extents, will be more prone to dipping below the detection threshold (or out of sample selection criteria) as a result of observational effects; the smaller correction to 1-1 could then be the opposite manifestation of this effect, wherein apparently brighter or larger galaxies are more robustly detected. The large-scale corrections to cross-correlations (Fig. 12) are also small. This is the desired behaviour, particularly in the case of disjoint bins 1a-2a, as the cross-correlation is expected to be zero. The rising 1a-2a signals at low values of ϑ (red points) are largely corrected for by the organised randoms. Given the vast gap ($\delta z = 0.06$) between the disjoint bins, these small-ϑ signals are highly unlikely to correspond to real structures. The organised randoms thus compensate for small-scale systematic density modes that are shared across the redshift range.

We note the (still relatively small) amplitude of the $w(\vartheta)$ correction at $\vartheta \gtrsim 10$ arcmin in the total sample correlation (in Fig. 11) and its similarity to the signal recoveries displayed in Fig. 9. This is encouraging as we can intuit that our extraction and replication of depletion patterns, from KiDS-Bright and in FLASK realisations (Sect. 5.2), is realistic enough to result in consistent inferences of the required correction to $w(\vartheta)$. Thus we can expect that the corrected total signal is closer to the true clustering, as is the case for the correlations from Fig. 9. In the intermediate range $7 < \vartheta < 100$ arcmin, the mean corrections to $w(\vartheta)$ for each correlation are {1-1: -5.6%, 1-2: -30.0%, 2-2: -10.9%, and total: -9.4%}.

In order that analyses be conducted free of any confirmation bias, it has become common in large-scale structure analyses to practice blinding (e.g., Kuijken et al. 2015), wherein several modified versions of key statistics are produced[10] from the data alongside the true version. The truth is revealed to the team by an independent entity only after the entire analysis is complete, such that no critical decisions can be taken in favour of

---

[9] Since the total equatorial GAMA area $\sim 180 \deg^2$, each of the 20 patches is $\sim 9 \deg^2 \equiv [180 \text{ arcmin}]^2$ in size, hence errors on $w(\vartheta \gtrsim 180$ arcmin) are likely to be under-estimated.

[10] The best methods for blinding are active areas of research; see for example Muir et al. (2020), Sellentin (2020), and Brieden et al. (2020).
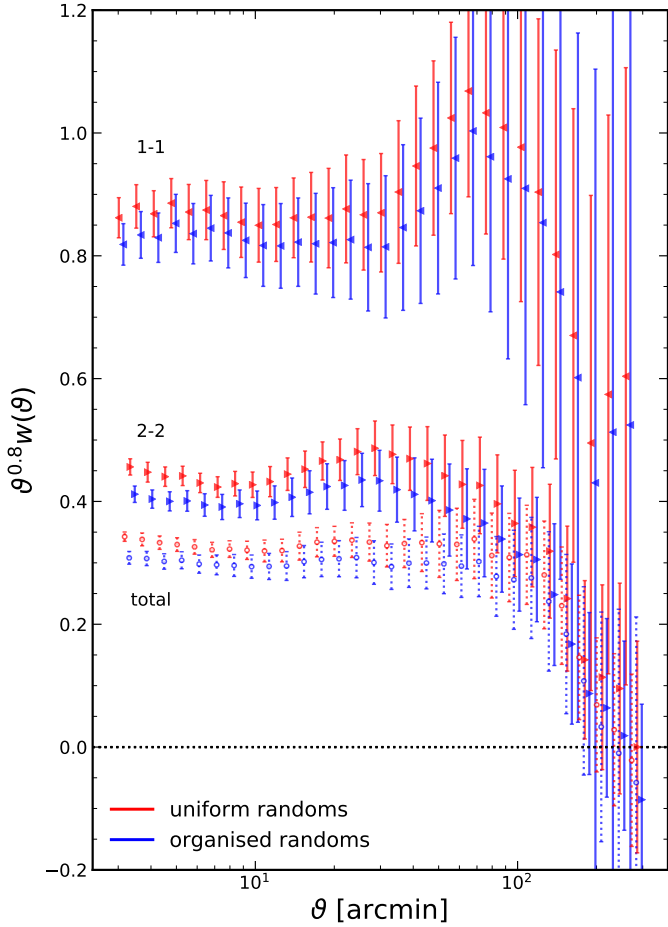
Fig. 11. Angular auto-correlation functions $w(\vartheta)$ measured in the total (i.e. unselected) KiDS-1000 bright sample (bottom), and within redshift bins 1 (1-1; top) and 2 (2-2; middle), with edges $\in$ [0.02, 0.2, 0.5], as shown in Fig. 3. Measurements using uniform randoms are shown in red, and those made with `100A` organised randoms are shown in blue. Errors are estimated with a 2D delete-one jackknife with 31 pseudo-independent patches of the footprint. Points are horizontally offset to aid clarity. The total sample correlation is more clearly visible in Fig. D.1.

Fig. 12. Angular cross-correlation functions $w(\vartheta)$ measured between KiDS-1000 bright sample (KiDS-Bright) redshift bins 1 and 2 (*top*), with edges $\in$ [0.02, 0.2, and 0.5], and 1a and 2a (*bottom*), with edges $\in$ [0.02 and 0.22] and [0.28 and 0.5], as shown in Fig. 3. Measurements using uniform randoms are shown in red, and those made with `100A` organised randoms are shown in blue. Errors are estimated with a 2D delete-one jackknife with 31 pseudo-independent patches of the footprint. Grey points and shading give the equivalently binned (by ANNz2 photo-z) correlations measured for the GAMA sample, with errors estimated again with a jackknife, but from 20 sub-regions of the GAMA window. Errors on scales $\vartheta \gtrsim 180$ arcmin are thus likely to be underestimated for GAMA correlations.

GAMA data are highly complete (>98%; Liske et al. 2015) and can be considered as the unbiased truth here. They are what we expect for photo-z of this quality in the absence of systematic galaxy density patterns. We thus argue that our corrections to the redshift bin cross-correlations successfully remove systematic correlations from KiDS-Bright data.

7. Summary

We have developed and tested a method for the construction of organised random galaxy catalogues, which mirror systematic trends in galaxy density, using self-organising maps. We made extensive use of lognormal random field simulations from FLASK to test the abilities of SOMs to recognise both artificial and real systematic loss of galaxies, and demonstrated that organised randoms constructed using this information are able to reliably correct the measurable angular clustering of the synthetic data.

With the present data volume, constructing effective organised randoms relies upon a balance between the area of sky probed by each hierarchical cluster (essentially an n-dimensional bin) defined on the SOM, the variables and dimensionality of the systematics space given to the SOM, and the width of the distribution of systematic density contrast. As we have demonstrated, this balance is readily assessed with simulations. If systematic modes are very much smaller than cosmological modes, organised randoms become more prone to over-correction of clustering biases, although our SOM methods are able to test for the necessity of any correction, as they estimate the distribution of systematic density contrast $\delta_{\rm OR}$. Conversely, for strong

some expected result. At the time of writing, we were blind to the truth of our KiDS data, and thus chose not to compare with theoretical models for galaxy clustering. Our companion work (Wright et al., in prep.) considers the best-fit cosmological model from the latest KiDS 3x2-point analysis (Heymans et al. 2021) for comparison with corrected clustering statistics measured in the KiDS-1000 shear sample.

A direct comparison of GAMA clustering auto-correlations with our own is likewise unsuitable here for reasons we discuss in Appendix D. We therefore only considered the redshift bin cross-correlations from GAMA (in Fig. 12) for validation. With the alignment of sample properties less important here, these correlations instead probe (i) the photo-z scatter and (ii) any systematic correlations shared across the KiDS-Bright redshift bins, where the former should be negligible in the 1a-2a correlation owing to the large gap (Fig. 3) between the bins. The rising signals at small-ϑ (Fig. 12; red points) are absent from GAMA (grey), and our organised randoms' corrections (blue) result in greater consistency between GAMA and KiDS-Bright. This is also true of the slightly negative (but not significant) 1a-2a blue data-points in the $\vartheta \sim 10$ arcmin range, where the negative GAMA points indicate possible LSS fluctuations. The

pathological density modes, our randoms are highly effective. Moreover, regardless of our analysis choices when testing with FLASK simulations, our recovered clustering results are always consistent with the underlying truth, having an average bias correction across all meaningful runs of $2.31\sigma \rightarrow 0.34\sigma$. Our recovery of the truth is particularly striking in the amplified bias ($m = 4$) case, where we shift from catastrophic bias with uniform randoms (12.2σ) to full consistency with the truth (0.34σ).

We found that the importance of certain systematics-tracing variables at the level of the two-point correlation function is not necessarily determined by the strength of the one-point (pixel) correlation of the tracer with galaxy density (see Appendix B). Whilst this finding may only hold for the KiDS-1000 bright sample (KiDS-Bright), we note that simply correcting for these one-point correlations may yet be problematic for the general goal of recovering unbiased galaxy clustering two-point functions. The amount of one-point correction required to achieve this goal is not necessarily clear and is further complicated by correlations between systematic-tracer parameters. We recommend the use of principle component analysis (PCA) to alleviate the latter concern, but two-point functions should also be considered for validation in analyses of this type.

We worked with FLASK simulations, modelling the footprint and number density of the KiDS-1000 bright sample, to create realistically biased synthetic galaxy fields within which to test the performance of our organised randoms. We inferred the field of systematic galaxy density contrast directly from KiDS-1000 bright sample data under various assumptions modifying sensitivity to angular scales and to different systematic-tracers, and modifying the amplitude of systematic fluctuations. Applying these data-driven systematic clustering imprints to many independent realisations from FLASK, we then generated organised randoms by training against the biased FLASK fields, again varying the scales and tracers we used. Under several scenarios of biasing due to the spatially variable PSF, detection threshold, survey depth, and Galactic stellar density and extinction, we found that training upon the `psf_fwhm,psf_ell,MU_THRESHOLD` parameters and defining 100 hierarchical clusters on the trained SOM (setup `100A`) was sufficient to yield organised randoms that consistently remove the various realistic systematic density modes from the FLASK galaxy fields. These `100A` organised randoms recovered clustering signals deviating on average from the unbiased signal at $\sim 0.3\sigma$ over the relevant FLASK-testing setups with average bias $\sim 1.1\sigma$. For the most pessimistic clustering bias scenarios, where uncorrected signals deviate from the truth at up to $\sim 12\sigma$, the performance of organised randoms remains robust, with the bias of recovery at $\sim 0.3\sigma$.

We presented the first measurement of photometric galaxy clustering from KiDS for bright GAMA-like galaxies from the $1000\,\mathrm{deg}^2$ 4th Data Release. Defining two tomographic bins with edges $z_{\mathrm{phot.}} = \{0.02, 0.2, \text{ and } 0.5\}$, we measured the angular auto- and cross-correlation functions over $3 < \vartheta < 300$ arcmin with uniform and organised randoms. We saw that our organised randoms make variable corrections to tomographic auto- and cross-correlations, editing amplitudes at intermediate angular scales ($7 \lesssim \vartheta \lesssim 100$ arcmin) by up to $\sim 10\%$ ($\sim 1\sigma$) in the auto-correlations, and $\sim 30\%$ in the cross-correlations.

We implemented our randoms such that each random point was a clone of a real galaxy, scattered within regions of the survey footprint that are similar to the location of the parent galaxy in terms of the position it occupies in systematics space. Thus by mimicking galaxy sample selections in the randoms, we compensated for distinct sample-specific systematic

correlations such as those induced by selections in galaxy photo-z. For tomographic cross-correlations, our randoms were found to correct significant systematic density modes at small-ϑ, which are shared between disparate redshift populations, whilst making nearly negligible corrections throughout the remaining angular range. This indicates similar small-scale, but distinct larger-scale systematic clustering imprints for the different redshift populations. This utility is easily generalised to any galaxy sample selections in luminosity, colour, etc.

An extension of this work to increased areas and galaxy number densities is extremely promising. Larger areas will result in better handling of large-ϑ systematic density modes due to more redundant sampling and in a smoother distribution of density contrast, which will minimise contamination of the randoms by cosmic structure. Higher galaxy densities offer better sampling of small-ϑ modes, a smoother description of the systematics space, and the possibility of increasing the resolution of the randoms without fear of greater contamination by structure. Thus the performance of organised randoms should improve on all scales with next-generation datasets; our companion letter (Wright et al., in prep.) moves to verify our assertions here, applying our testing pipeline to measurements of clustering in the faint KiDS-1000 shear sample, thus exploring a deep-survey, high number density scenario.

Upon publication of our companion letter, we will make our code and methods public, such that independent teams can experiment with the handling of systematic density variations using organised random clones. Future surveys that will make powerful use of galaxy clustering (e.g., the *Rubin* Observatory and *Euclid*) can then include the construction of organised randoms as a pipeline module, adding to the panoply of complementary means for the accurate measurement of galaxy positional statistics.

contributed to the development and writing of this paper. The authorship list is given in two groups: the lead authors (H. J., A. H. W., B. J.), followed in alphabetical order by authors who made a significant contribution to either the data products or to the scientific analysis.

References

Abbott, T. M., Abdalla, F. B., Alarcon, A., et al. 2018, Phys. Rev. D, 98, 043526
Aghanim, N., Akrami, Y., Ashdown, M., et al. 2020, A&A, 641, A6
Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617
Anderson, L., Aubourg, É., Bailey, S., et al. 2014, MNRAS, 441, 24
Arenou, F., Luri, X., Babusiaux, C., et al. 2018, A&A, 616, A17
Asgari, M., Lin, C. A., Joachimi, B., et al. 2021a, A&A, 645, A104
Asgari, M., Friswell, I., Yoon, M., et al. 2021b, MNRAS, 501, 3003
Bergé, J., Gamper, L., Réfrégier, A., & Amara, A. 2013, Astron. Comput., 1, 23
Bilicki, M., Hoekstra, H., Brown, M. J., et al. 2018, A&A, 616, A69
Bilicki, M., Dvornik, A., Hoekstra, H., et al. 2021, ArXiv e-prints [arXiv:2101.06010]
Brieden, S., Gil-Marín, H., Verde, L., & Bernal, J. L. 2020, JCAP, 2020, 052
Buchs, R., Davis, C., Gruen, D., et al. 2019, MNRAS, 489, 820
Davidzon, I., Laigle, C., Capak, P. L., et al. 2019, MNRAS, 489, 4817
Defays, D. 1977, Comput. J., 20, 364
de Jong, J. T., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, Exp. Astron., 35, 25
DESI Collaboration (Aghamousa, A., et al.) 2016, ArXiv e-prints [arXiv:1611.00036]
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, Astron. Geophys., 50, 5.12
eBOSS Collaboration (Alam, S., et al.) 2020, ArXiv e-prints [arXiv:2007.08991]
Edge, A., Sutherland, W., Kuijken, K., et al. 2013, Messenger, 154, 32
Eisenstein, D. J., Annis, J., Gunn, J. E., et al. 2001, AJ, 122, 2267
Elvin-Poole, J., Crocce, M., Ross, A. J., et al. 2018, Phys. Rev. D, 98, 042006
Erben, T., Hildebrandt, H., Miller, L., et al. 2013, MNRAS, 433, 2545
Everett, S., Yanny, B., Kuropatkin, N., et al. 2020, ApJS, submitted [arXiv:2012.12825]
Farrow, D. J., Cole, S., Norberg, P., et al. 2015, MNRAS, 454, 2120
Font-Ribera, A., Kirkby, D., Busca, N., et al. 2014, JCAP, 2014, 027
Freedman, W. L. 2017, Nat. Astron., 1, 0121
Geach, J. E. 2012, MNRAS, 419, 2633
Giblin, B., Heymans, C., Asgari, M., et al. 2021, A&A, 645, A105
Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hamana, T., Shirasaki, M., Miyazaki, S., et al. 2019, J. Biochem., 166, 16
Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399
Heymans, C., Tröster, T., Asgari, M., et al. 2021, A&A, 646, A140
Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454
Hildebrandt, H., Köhlinger, F., Van Den Busch, J. L., et al. 2019, A&A, 633, 1
Hildebrandt, H., van den Busch, J. L., Wright, A. H., et al. 2021, A&A, 647, A124
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Ho, S., Cuesta, A., Seo, H. J., et al. 2012, ApJ, 761, 14
Iezzi, D. F. 2014, Cluster Analysis (Dordrecht: Springer Netherlands), 966
Jarvis, M., Bernstein, G., & Jain, B. 2004, MNRAS, 352, 338
Jolliffe, I. T. 2002, Principal Component Analysis, Springer Series in Statistics (Springer)

Joudaki, S., Blake, C., Johnson, A., et al. 2017, MNRAS, 474, 4894
Joudaki, S., Hildebrandt, H., Traykova, D., et al. 2020, A&A, 638, L1
Kitanidis, E., White, M., Feng, Y., et al. 2020, MNRAS, 496, 2262
Kohonen, T. 1990, Proc. IEEE, 78, 1464
Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, MNRAS, 454, 3500
Kuijken, K., Heymans, C., Dvornik, A., et al. 2019, A&A, 625, A2
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints [arXiv:1110.3193]
Leistedt, B., & Peiris, H. V. 2014, MNRAS, 444, 2
Leistedt, B., Peiris, H. V., Mortlock, D. J., Benoit-Lévy, A., & Pontzen, A. 2013, MNRAS, 435, 1857
Limber, D. N. 1953, ApJ, 117, 134
Liske, J., Baldry, I. K., Driver, S. P., et al. 2015, MNRAS, 452, 2087
Loverde, M., & Afshordi, N. 2008, Phys. Rev. D, 78, 123506
LSST Science Collaboration (Abell, P. A., et al.) 2009, ArXiv e-prints [arXiv:0912.0201]
Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2017, ApJ, 841, 111
Masters, D. C., Stern, D. K., Cohen, J. G., et al. 2019, ApJ, 877, 81
Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, Ground-based and Airborne Instrumentation for Astronomy IV, 8446, 84460Z
Morrison, C. B., & Hildebrandt, H. 2015, MNRAS, 454, 3121
Muir, J., Bernstein, G. M., Huterer, D., et al. 2020, MNRAS, 494, 4454
Nicola, A., Alonso, D., Sánchez, J., et al. 2020, JCAP, 2020, 044
Porredon, A., Crocce, M., Fosalba, P., et al. 2021, Phys. Rev. D, 103, 043503
Raichoor, A., Comparat, J., Delubac, T., et al. 2017, MNRAS, 471, 3955
Rezaie, M., Seo, H.-J., Ross, A. J., & Bunescu, R. C. 2020, MNRAS, 495, 1613
Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, ApJ, 876, 85
Ross, A. J., Ho, S., Cuesta, A. J., et al. 2011, MNRAS, 417, 1350
Sadeh, I., Abdalla, F. B., & Lahav, O. 2016, PASA, 128, 104502
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Schlegel, D. J. J., Finkbeiner, D. P. P., & Davis, M. 1998, ApJ, 500, 525
Sellentin, E. 2020, MNRAS, 492, 3396
Singh, S., Mandelbaum, R., Seljak, U., Slosar, A., & Gonzalez, J. V. 2017, MNRAS, 471, 3827
Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311
Speagle, J. S., & Eisenstein, D. J. 2015, ArXiv e-prints [arXiv:1510.08080]
Speagle, J. S., & Eisenstein, D. J. 2017, MNRAS, 469, 1186
Suchyta, E., Huff, E. M., Aleksić, J., et al. 2016, MNRAS, 457, 786
Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, ApJ, 761, 152
The Dark Energy Survey Collaboration 2005, ArXiv e-prints [arXiv:astro-ph/0510346]
Tröster, T., Sánchez, A. G., Asgari, M., et al. 2020, A&A, 633, L10
Troxel, M. A., Maccrann, N., Zuntz, J., et al. 2018, Phys. Rev. D, 98, 043528
Vakili, M., Hoekstra, H., Bilicki, M., et al. 2020, A&A, submitted [arXiv:2008.13154]
van Uitert, E., Joachimi, B., Joudaki, S., et al. 2018, MNRAS, 476, 4662
Wagoner, E. L., Rozo, E., Fang, X., et al. 2021, MNRAS, 503, 4349
Weaverdyck, N., & Huterer, D. 2021, MNRAS, 503, 5061
Wright, A. H., Hildebrandt, H., van den Busch, J. L., & Heymans, C. 2020, A&A, 637, A100
Xavier, H. S., Abdalla, F. B., & Joachimi, B. 2016, MNRAS, 459, 3693
Yoon, M., James Jee, M., Anthony Tyson, J., et al. 2019, ApJ, 870, 111
York, D. G., Adelman, J., Anderson, J. E., Jr, et al. 2000, AJ, 120, 1579

Appendix A: Artificial systematics

Here we extend our discussion of the characterisation of artificial systematic density fluctuations with self-organising maps (Sect. 4). We created artificial systematic-tracer variables (Fig. 5) with arbitrary depletion functions (Fig. 6) to apply to FLASK mock data (the excess probability of depletion for each artificial variable is given in Fig. A.1), finding the T1mock SOM (Table 2) to be capable of characterising the depletion functions directly from the mock catalogue (Fig. 7). As a sanity check, we also ran the same SOM against the 30 FLASK catalogues before applying any depletion. The SOM should not find any density-systematic correlations where none exist. The results are shown in Fig. A.2, where the depletion functions from Fig. 6 are included for reference, and we see the desired behaviour. The SOM recognises no significant, systematic trends in galaxy density (barring some of the irregularities related to grid resolution, discussed in Sect. 4.2).
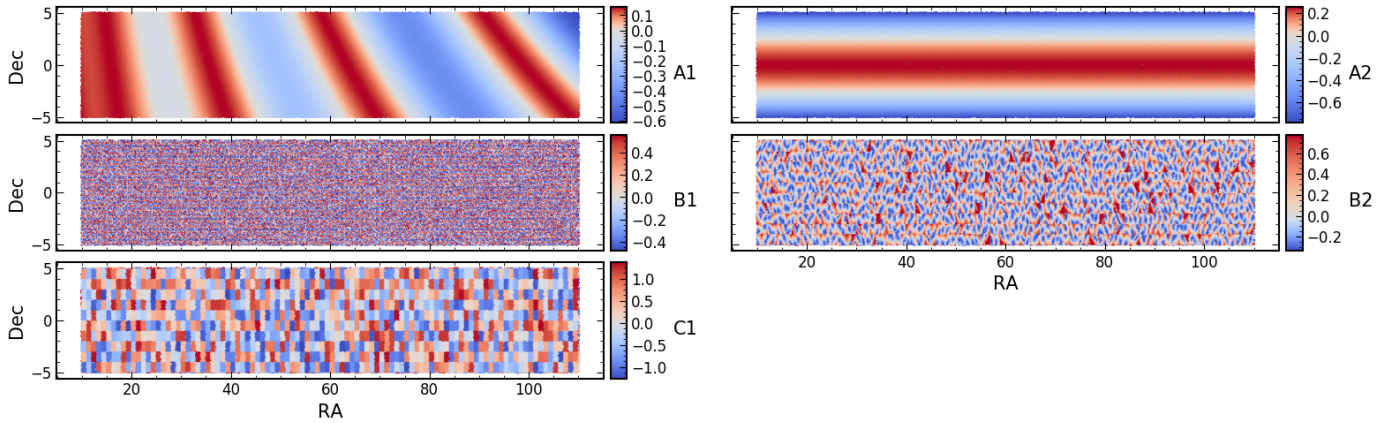

Fig. A.1. Same as Fig. 5, but only for parameters with non-zero depletion functions. The colouring now denotes the excess probability of depletion as a function of each artificial systematic. A galaxy with an excess probability of 1 is 100% more (or twice as) likely to be depleted when compared with an excess probability of zero.
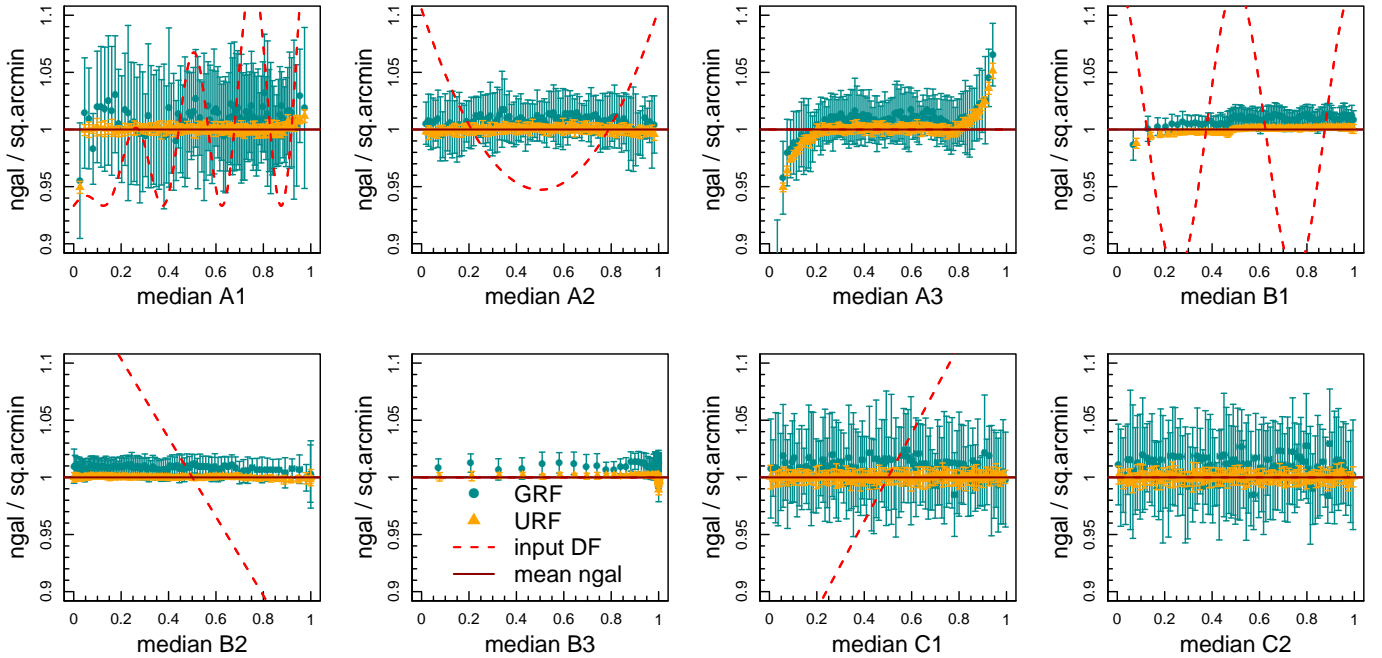

Fig. A.2. Same as Fig. 7, but for a SOM trained on an unbiased FLASK field. The depletion functions are again shown as dashed red lines, although no depletions were applied. Irregularities in the A3 panel, for example, are discussed in Sect. 4.2. Data-points derived from lognormal random fields (green) are on average higher in n_{gal} than those from uniform fields (orange) because of the clustering of galaxies. The Cartesian grid we used to estimate the area of the sample has more empty cells for the GRF, hence the area is under-estimated and n_{gal} is over-reported.

Appendix B: Pixel density systematic one-point correlations

Here we consider correlations between the densities of galaxies in pixels on-sky and the values of systematic-tracer variables in those pixels. Many versions of these statistics appear in the literature, often as metrics to demonstrate successful decorrelation of galaxy densities and systematics by other methods (e.g., Suchyta et al. 2016; Rezaie et al. 2020; Kitanidis et al. 2020, where some also use two-point metrics, which we argue is necessary) and occasionally in order to directly derive the corrections themselves (Elvin-Poole et al. 2018; Vakili et al. 2020). We note that the drawbacks of one-point metrics that we explore in detail here may be specific to our methods and the parameters of our survey data. For these tests, we constructed galaxy density and systematics maps using `healpy` with $n_{side} = 256$, corresponding to on-sky pixels of roughly 14×14 arcmin in size. We computed the mean systematics in each pixel, applying requisite scaling to pixels that were only partially sampled by the survey window. Taking the means-in-pixels will constitute some loss of information, but this should primarily affect sub-pixel scales. We leave any exploration of higher resolution maps and/or improved pixel statistics to future work with denser samples.

This approach is somewhat analogous to that of Elvin-Poole et al. (2018), who used individual density systematic one-point functions directly to derive per-galaxy weights for the mitigation of clustering systematics. In our case, however, the one-point correlations serve only as a diagnostic. Our SOM approach improves upon direct pixel-correlation methods and drops some commonly made approximations. In particular, we mapped the galaxy depletion function through a higher-dimensional systematics space, correcting galaxies directly in this space, rather than iteratively suppressing only the worst individual correlations until crossing some threshold. We note that Weaverdyck & Huterer (2021) showed with DES year-one-like simulations that the latter point does not damage the success of corrections. Additionally, whilst usually approximated as linear in the density contrast, the depletion of galaxies as a multivariate function of systematics in our method can be arbitrarily non-linear, allowing greater freedom to correct for more complex systematic modes

We note that, similarly to this analysis, Rezaie et al. (2020) tested their mitigation scheme with contaminated mock catalogues, finding that linear models for density-systematic correlations were sufficiently descriptive to yield successful corrections. However, their contamination of mocks was itself based upon the linear model, thus the necessity of non-linear corrections was not fully explored for those data. We introduced methods for contamination of mocks based upon the non-linear density-systematics modelling from SOMs (Sect. 5.2). Future work can apply linearly derived corrections to these mocks and test for the necessity of non-linearity in the modelling for any dataset.

We made some changes to the typical formulation of the galaxy density versus systematic pixel correlations. First, we normalised per-pixel number counts (or pixel counts) of galaxies, N_{gal}, not by the global average galaxy pixel counts, $\langle N_{gal} \rangle$, but instead by pixel counts for a uniform random field, N_{rand} (given the same on-sky pixelisation). These considerations are particularly relevant for our method, which constructs a high-resolution mask of the sky using data and random points. The two estimators are generally equivalent in the limit of large area, modulo normalisation between the average random and galaxy number densities. Second, for the corrected case, we normalised the galaxy pixel counts by the organised randoms' pixel counts.

Any systematic correlations in the data should thus be cancelled by the mirrored trends in the organised randoms.

We note here, however, that the erasure of such trends can be a misleading measure of performance. Inter-tracer correlations can induce trends in the N_{gal}/N_{rand} distribution as a function of systematics that make individual systematics appear more or less significant than they truly are (sometimes referred to as 'chance correlations' in the literature). Furthermore, the strength of correlation per systematic that is needed to introduce a significant clustering bias is also an open question, but is likely dependent upon the sample being analysed, the distribution and dynamic range of the systematic parameter, and the ultimate angular scales of interest. Finally, these distributions can behave particularly pathologically when correcting systematic density variations using variable randoms, as we did here. As a simple demonstration, should the organised randoms produce a catalogue that exactly reproduced the input galaxy catalogue (i.e. the randoms encoded all galaxy density fluctuations, both systematic and cosmological), then these correlations would appear perfectly flat. This would constitute the most abject failure of our corrective method, and yet this diagnostic would be labelled a success.

These concerns led us to derive and consider orthogonal principle component maps (i.e. principle component analysis, or PCA; see Jolliffe 2002, for a comprehensive text on PCA) alongside our systematics maps (see also Wagoner et al. 2021, who applied such a decomposition to their systematics maps). Principle components are defined as new orthogonal basis vectors for n-dimensional data, constructed from linear combinations of the original dimensions. The new basis then describes in order, from component 1 to component n, the directions of greatest variance in the data. As the basis vectors are orthogonal, principle components have no covariance and are thus more instructive for assessing the significance of density-systematic correlations, although they are more difficult to interpret physically.

Lastly, we assessed the expected spread of pixel density-PCA component (or density-systematic) correlations for individual unbiased realisations of the large-scale structure within our survey parameters. These are given in Fig. B.1, where solid lines give the 1σ and 2σ widths of the N_{gal}/N_{rand} versus systematic-tracer distributions from 100 individual FLASK realisations. As Fig. B.1 shows, the intrinsic spread of galaxy number density versus systematic-tracers over 100 independent realisations of KiDS-Bright-like synthetic data is often around a few percent and rises for more long-varying tracers such as Galactic extinction and stellar density. Thus, for our data, galaxy density-component or systematic correlations within a few percent of unity cannot be distinguished from those arising stochastically. These are the limits of useful bias-correction dictated by sample variance, and we mark them as purple shading in the following pixel-correlation figures.

Correction of one-point correlations

After creating maps of each of the systematic-tracers listed in Table 1, we masked pixels in the 1st and 99th percentile tails of each distribution and made additional cuts by hand, excluding the few sparse pixels with extreme outlier values for systematic-tracers. These would otherwise contaminate correlations with poorly sampled unrepresentative galaxy counts. We then whitened the parameter distributions (transforming each to zero-mean and unit variance) before retrieving 14 orthogonal principle component maps. The matrix of linear coefficients
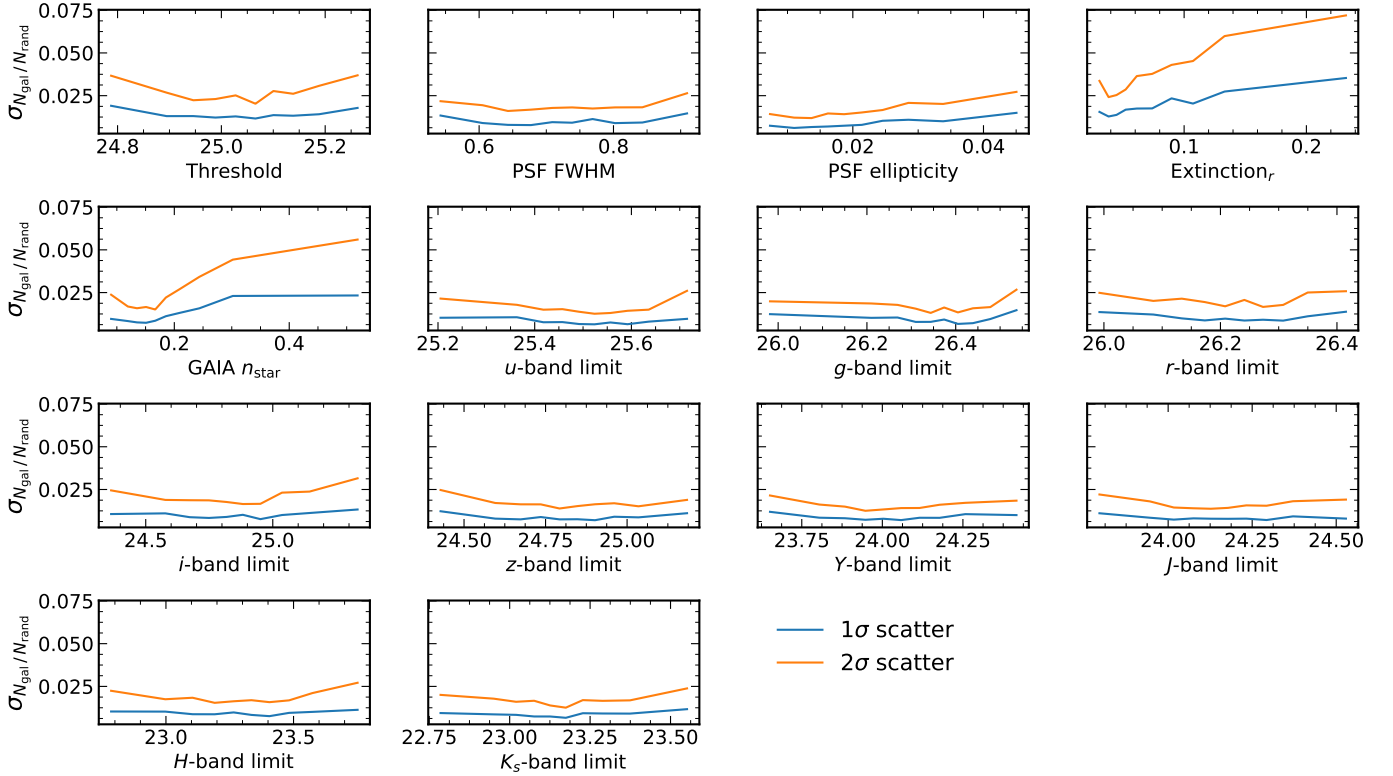
Fig. B.1. Intrinsic scatter of N_{gal}/N_{rand} vs. various systematics tracer variables (Table 1) as measured over 100 FLASK realisations with the KiDS-Bright survey parameters and footprint. Without any biases applied to the FLASK realisations, the blue (1σ) curves give half the difference between the 16th and 84th percentiles of the distribution of 100 individual N_{gal}/N_{rand} vs. systematics correlations. 2σ (orange) curves give half the difference between the 2.5th and 97.5th percentiles. Galaxy/random counts and mean systematics are taken in $n_{side} = 256$ pixels, about 14 arcmin in size.
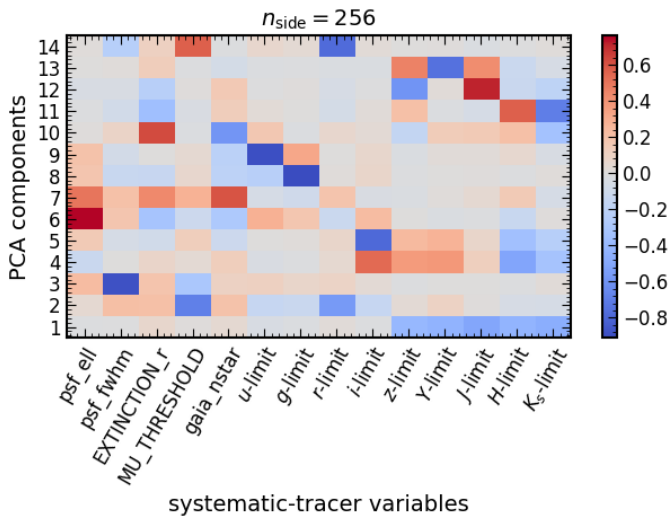

Fig. B.2. Linear coefficients as indicated by the colour bar, describing the transformation of systematic-tracer variable maps into principle component maps for $n_{side} = 256$.

connecting PCA components to systematics tracers is shown in Fig. B.2.

Figures B.3 and B.4 illustrate why these one-point correlation metrics are not suitable for assessing the performance of

our corrective randoms. In Fig. B.3 we select the subset of four component-density pixel correlations in bias:800C FLASK realisations that are best corrected to unity after normalisation by the 100A organised randoms' density in these pixels (Fig. B.2 shows that these components are closely related to the PSF or threshold training variables from parameter set A). The implication is that residual correlations will cause two-point clustering correlations to be biased, as the 100A SOM is insufficiently tracing systematic density fluctuations at the pixel level. However, the right-hand side in Fig. B.3 also shows that these organised randoms are capable of recovering unbiased clustering two-point functions.

Conversely, when we raise the grid resolution of our organised randoms (setup 800Ares2 from Table 2), we see that the systematic density one-point correlations in Fig. B.4 are almost perfectly corrected (we show systematic-tracers in this figure as they are easier to interpret and because the component corrections are similarly nearly perfect), whilst the two-point functions reveal a pathological over-correction of the clustering signature. This is because the 800Ares2 randoms reproduce the real cosmic structure in the data too closely (demonstrated by the unbiased data-organised randoms cross-correlation, which is at ~50% of the unbiased signal), and thus act to destroy real clustering signals. We therefore restricted our validation of organised randoms to two-point correlation statistics, which offer insight into potential over-corrections as organised randoms become contaminated with cosmic structure.
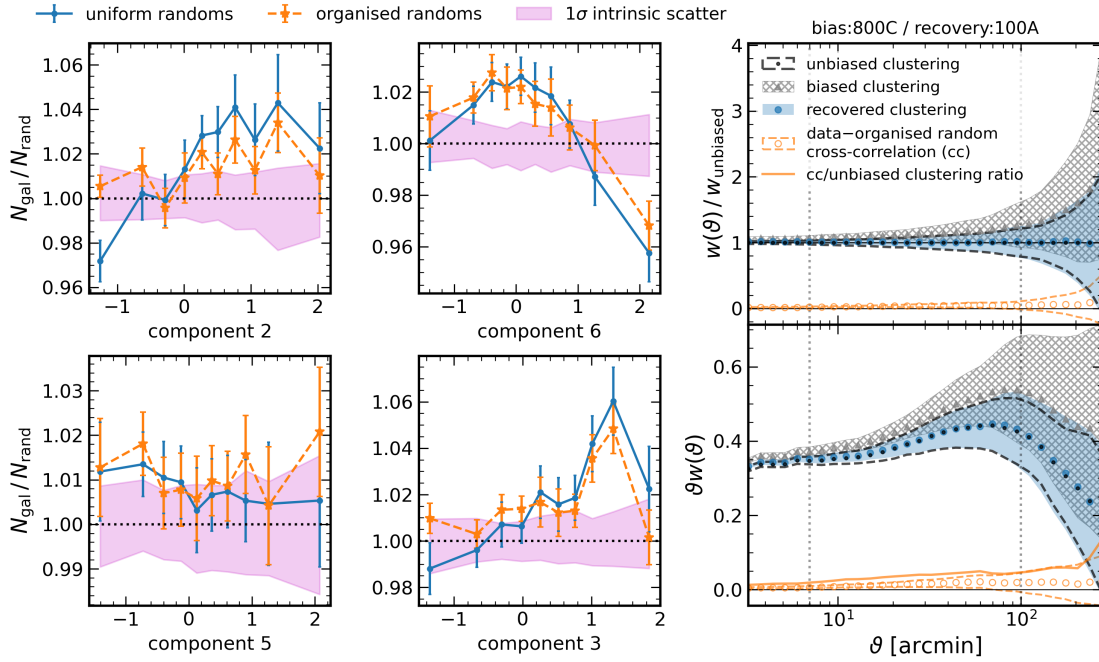
Fig. B.3. Illustration of divergent one- and two-point correlation corrections from organised randoms. *Left:* pixel (one-point) correlations between the galaxy number density $N_{\rm gal}/N_{\rm rand}$ in biased (according to the `800C` SOM) Flask mocks and the mean PCA components in the same pixels (see Appendix B). Correlations normalised by uniform randoms' counts-in-pixels are given in blue, and those normalised (corrected) by the pixel counts of `100A` organised randoms are given in orange. Correlations and errors are the mean and the root-diagonal of the covariance over 30 Flask realisations, respectively. Purple shading indicates the expected intrinsic 1σ spread in these correlations, as calculated over 100 unbiased Flask realisations (Fig. B.1). Shown here are the 4 (out of 14) correlations for which the organised randoms correction most improves consistency with unity. *Right:* corresponding biased and recovered two-point angular clustering correlations $w(\vartheta)$, shown in ratio to (*top*) and overlaid on (*bottom*) the unbiased clustering signature, as in Fig. 9. We also show the cross-correlation between organised randoms and the unbiased data and its ratio to the unbiased clustering signal in orange. Correlations and errors are again the mean and the root-diagonal of the covariance over 30 Flask realisations.
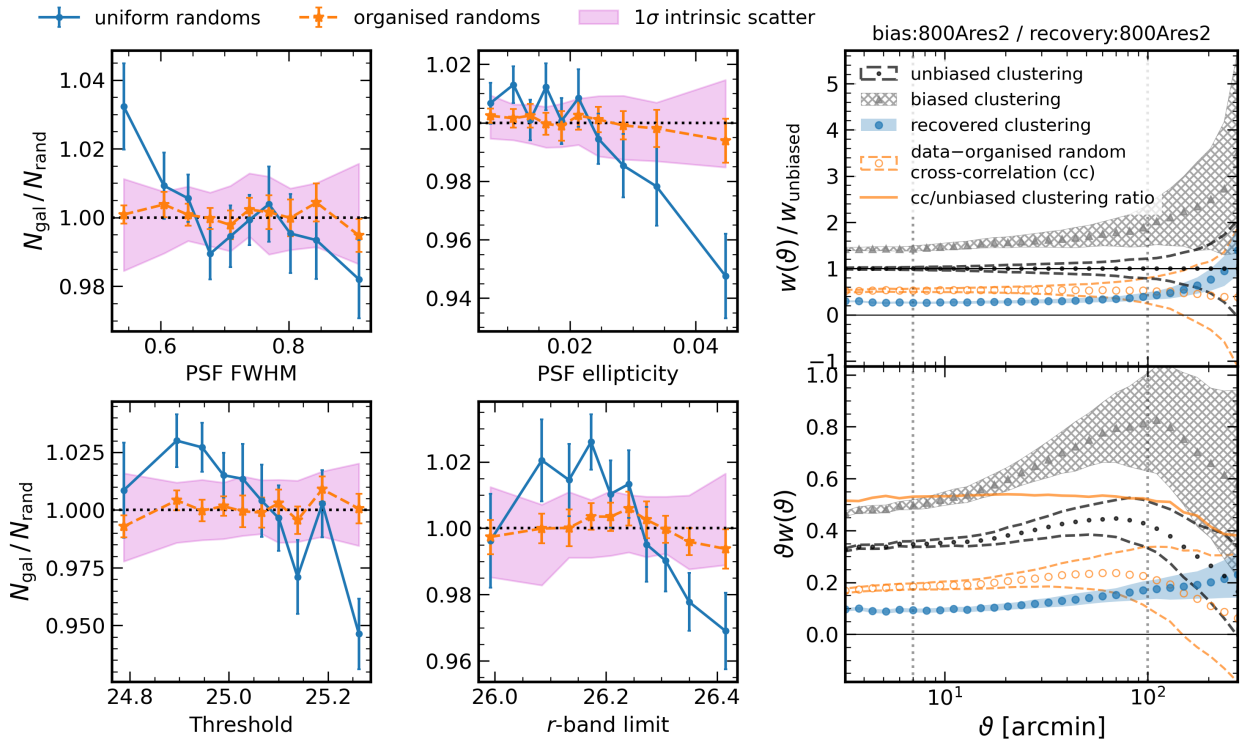

Fig. B.4. Same as Fig. B.3, now for bias:`800Ares2`/recovery:`800Ares2`, i.e. the high-resolution setup, and considering systematic-tracers rather than PCA components. Nearly perfect one-point correlation corrections (*left*) are accompanied by the cross-correlation between unbiased Flask data and organised randoms (*right;* orange) rising to ~50% of the unbiased clustering signature, compared with the negligible cross-correlation in Fig. B.3.

Appendix C: Data-driven systematics

Figure C.1 presents additional configurations of the bias:[800A,100A]/recovery:[100A,100C] two-point correlations. As stated in the main text, each meaningful case yields a recovered clustering signal that is more consistent with the unbiased clustering. For bias:800A cases, where the bias is slightly weaker than for bias:800C (Fig. 9) due to the omission of MAG_LIM_r,EXTINCTION_r,gaia_nstar parameters from training, the performance of organised randoms is similarly strong, and a slight preference for recovery:100A over 100C is negligible with respect to the noise level ($\sim 0.1\sigma$). When clustering biases are particularly weak, as for the bias:100A case, our methods become more prone to over-correction. Still,

the mean absolute deviation from the unbiased case (in units of the standard deviation of the unbiased clustering; summarised for all configurations in Table C.1) is improved in the recovery with respect to the biased case here, primarily due to the success of corrections at lower values of ϑ. We note that the bias:100A case constitutes our most optimistic bias scenario, having low sensitivity to small-ϑ (via small N_{HC}) and the minimum set of systematic-tracer training variables (psf_fwhm,psf_ell,MU_THRESHOLD). For more pessimistic scenarios, the organised randoms' performance is consistently improved. As also stated in the text, this bodes extremely well for the potential salvage of unbiased clustering signals from heavily systematics-contaminated data, such as the faint KiDS-1000 shear sample (see Wright et al., in prep.).
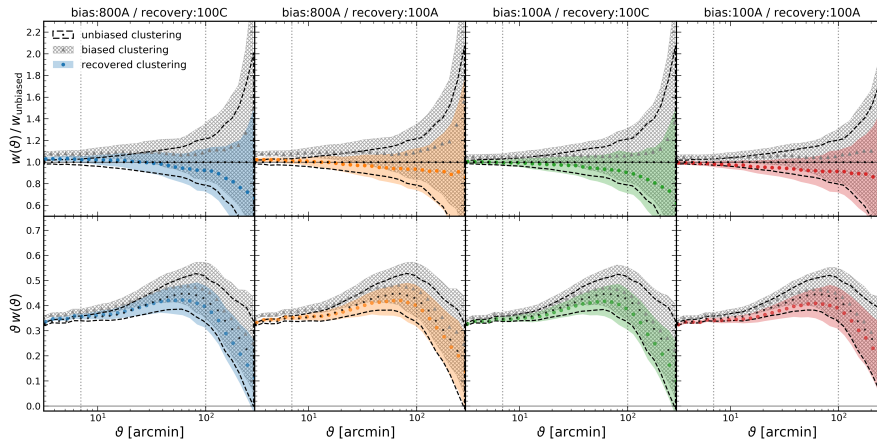

Fig. C.1. Same as Fig. 9, now for some additional bias:SOM and recovery:SOM configurations, as indicated by panel titles (with reference to Table 2).

Table C.1. Relative performance of corrections to angular clustering correlations $w(\vartheta)$ on the intermediate scales $7 < \vartheta < 100$ arcmin for differently configured clustering biases and organised randoms.

bias:SOM	recovery:SOM	Biased	Recovered
100A	100A	0.69σ	0.61σ
100A	100C	0.79σ	0.35σ
100C	100A	0.94σ	0.28σ
100C	100C	0.96σ	0.54σ
100C(x4)	100C	12.17σ	0.34σ
800A	100A	1.13σ	0.27σ
800A	100C	1.19σ	0.31σ
800Ares2	800Ares2	7.62σ	9.58σ
800C	100A	1.43σ	0.09σ
800C	100C	1.48σ	0.31σ

Notes. With reference to Table 2, the bias:SOM column gives the SOM trained against KiDS-Bright to infer δ_{OR} (Eq. (5.3)) and inform the density-field bias applied to 30 Flask realisations of KiDS-Bright-like data. The recovery:SOM column gives the SOM trained against those biased Flask mocks to produce organised randoms (see Sect. 5.2 for details). On average over the 30 Flask realisations, the Biased and Recovered columns then give the absolute deviation from the unbiased clustering signal $|w_{\mathrm{biased/recovered}} - w_{\mathrm{unbiased}}|$ in units of the uncertainty in the unbiased signals σ, averaged over the intermediate scales 7–100 arcmin. Non-uniformity of biased column values with self-similar bias:SOM setups arises stochastically due to our probabilistic application of biases to Flask realisations, and all variations are at $<0.1\sigma$. The 800Ares2 configuration illustrates a failure mode of our organised randoms methods, where the resolution and scale sensitivity of the setup are too high, resulting in over-fitting to cosmic structure and a consequent over-correction of the clustering signal.

Appendix D: GAMA comparisons

As discussed in Sect. 2.1, our KiDS-Bright selection is deliberately GAMA-like for the training of photo-z estimation (Bilicki et al. 2021). Thus we might expect a similar clustering signature from KiDS-Bright, but for a reduction in amplitude due to photometric redshift scatter. Scattering disconnected and connected galaxies in and out of each bin will dilute the measurable clustering signals. To replicate this effect in the spectroscopic GAMA sample, we cross-matched GAMA with KiDS-Bright and defined redshift bins using the KiDS ANNz2 photo-z. The resulting total correlations, shown as grey points and shading in the top-panel of Fig. D.1, exceeded even the uncorrected KiDS-Bright correlations in amplitude. We determined that KiDS-Bright is not sufficiently GAMA-like to expect agreement here by remeasuring correlations for all KiDS-Bright galaxies in the GAMA window (Fig. D.1; middle panel), and finding a reduced clustering amplitude which is more consistent with the corrected (by organised randoms) correlation. We measured correlations once more for KiDS-Bright in the GAMA window, now with a flux limit $r < 19.8$ applied to KiDS magnitudes[11] (Fig. D.1; bottom panel), and again see the clustering amplitude rise to exceed all KiDS-Bright signals. We conclude that a comparison with the GAMA auto-clustering signals is not entirely valid for KiDS-Bright, noting also that with only ~1/5 of the KiDS-Bright area, the impacts of increased sample variance in GAMA present further complications.
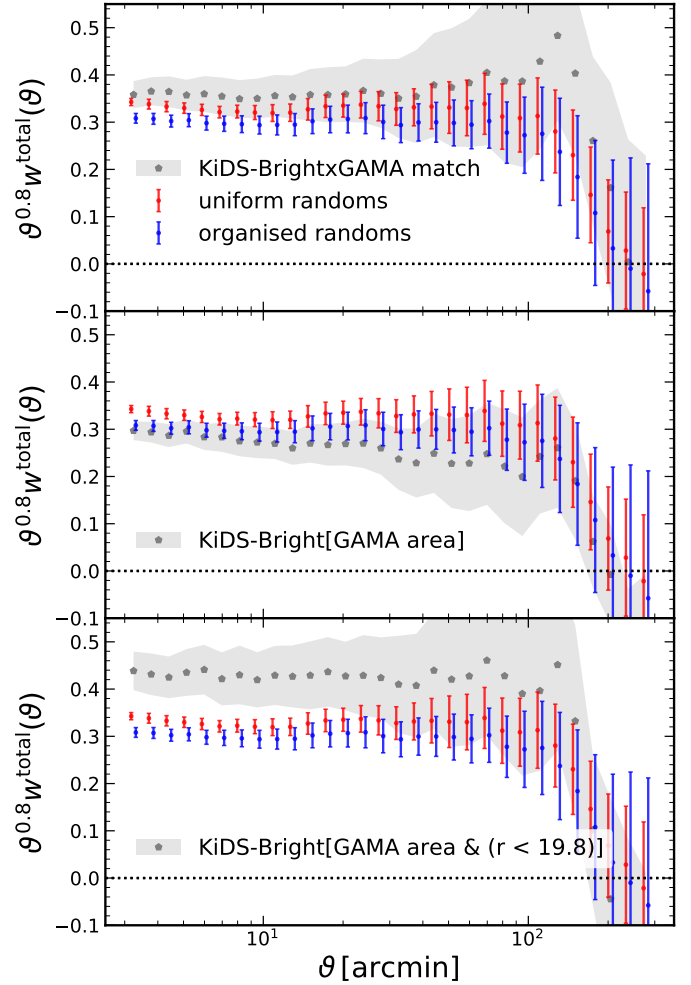

Fig. D.1. Comparison of different clustering measurements in the GAMA area and how they compare with the KiDS-Bright correlations measured with (blue) and without (red) `100A` organised randoms. Red and blue data-points are the same in all rows, whilst grey points and shading are the angular clustering measured with uniform randoms for (*top*) the KiDS-Bright-GAMA cross-matched catalogue, (*middle*) KiDS-Bright galaxies within the GAMA window, and (*bottom*) KiDS-Bright galaxies within the GAMA window, and with $r < 19.8$. Errors are estimated with delete-one jackknifes from 31 sub-regions of KiDS-Bright or 20 sub-regions of the GAMA window. Errors on scales $\vartheta \gtrsim 180$ arcmin (the approximate scale of each sub-region) are thus likely to be underestimated for the GAMA-area correlations.

---

[11] KiDS magnitudes are not equivalent to the Petrosian magnitudes against which GAMA targets were selected, therefore this flux limit is only approximately GAMA-like (see Bilicki et al. 2018).