

Probing bursty star formation by cross-correlating extragalactic background light and galaxy surveys

Guochao Sun ¹★, Adam Lidz,² Andreas L. Faisst^{2,3} and Claude-André Faucher-Giguère ¹

¹CIERA and Department of Physics and Astronomy, Northwestern University, 1800 Sherman Avenue, Evanston, IL 60201, USA

²Department of Physics & Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA

³Caltech/IPAC, MS314-6, 1200 East California Boulevard, Pasadena, CA 91125, USA

Accepted 2023 June 30. Received 2023 June 30; in original form 2023 May 15

ABSTRACT

Understanding the star formation rate (SFR) variability and how it depends on physical properties of galaxies is important for developing and testing the theory of galaxy formation. We investigate how statistical measurements of the extragalactic background light (EBL) can shed light on this topic and complement traditional methods based on observations of individual galaxies. Using semi-empirical models of galaxy evolution and SFR indicators sensitive to different star formation time-scales (e.g. $H\alpha$ and ultraviolet continuum luminosities), we show that the SFR variability, quantified by the joint probability distribution of the SFR indicators (i.e. the bivariate conditional luminosity function), can be characterized as a function of galaxy mass and redshift through the cross-correlation between deep, near-infrared maps of the EBL and galaxy distributions. As an example, we consider combining upcoming SPHEREx maps of the EBL with galaxy samples from Rubin Observatory Legacy Survey of Space and Time. We demonstrate that their cross-correlation over a sky fraction of $f_{\text{sky}} \sim 0.5$ can constrain the joint SFR indicator distribution at high significance up to $z \sim 2.5$ for mass-complete samples of galaxies down to $M_* \sim 10^9 M_\odot$. These constraints not only allow models of different SFR variability to be distinguished, but also provide unique opportunities to investigate physical mechanisms that require large number statistics such as environmental effects. The cross-correlations investigated illustrate the power of combining cosmological surveys to extract information inaccessible from each data set alone, while the large galaxy populations probed capture ensemble-averaged properties beyond the reach of targeted observations towards individual galaxies.

Key words: galaxies: star formation – cosmic background radiation – infrared: diffuse background.

1 INTRODUCTION

Both observations of star-forming galaxies at different cosmic epochs (Weisz et al. 2012; Emami et al. 2019; Faisst et al. 2019) and galaxy simulations resolving the gravitational collapse of star-forming gas and stellar feedback (Domínguez et al. 2015; Sparre et al. 2017; Gurvich et al. 2023; Hopkins et al. 2023) have led to an emerging picture where the star formation rate (SFR) of galaxies in certain regimes is highly time-variable – a situation often referred to as bursty star formation. Elucidating the physical origin of bursty star formation and the transition to time-steady star formation is a key task for galaxy formation theory (Faucher-Giguère 2018; Caplar & Tacchella 2019; Iyer et al. 2020; Furlanetto & Mirocha 2022; Orr et al. 2022; Hopkins et al. 2023). To this end, a crucial way to connect observations with theory is to investigate the variety of SFR indicators sensitive to different time-scales of star formation. Among the large number of SFR indicators proposed in the literature, the $H\alpha$ $\lambda 6563$ nebular line emission and the ultraviolet (UV) continuum emission are most commonly considered (e.g. Emami et al. 2019; Flores Velázquez et al. 2021). Because $H\alpha$ emission is

predominantly produced by recombinations in $H\text{II}$ regions ionized by young, massive stars, it is expected to be sensitive to recent SFR variations on time-scales as short as a few Myr. On the other hand, the UV continuum emission has substantial contributions from the non-ionizing radiation of older stellar populations and therefore is sensitive to significantly longer star formation time-scales (~ 10 Myr when the SFR is time-steady and ~ 100 Myr following extreme starbursts; see e.g. Flores Velázquez et al. 2021). The exact value depends on various factors, such as the wavelength of emission, the star formation history (SFH), and the stellar population synthesis model assumed.

Traditional methods relying on these SFR indicators usually require measuring the $H\alpha$ and UV luminosities of individual galaxies simultaneously from flux-limited surveys. Such measurements are expensive and likely susceptible to issues like selection bias that preferentially selects galaxies experiencing an ongoing burst of star formation (Domínguez et al. 2015; Faisst et al. 2019; Sun et al. 2023). Meanwhile, measuring the mean ratio $L_{H\alpha}/L_{UV}$ (where $L_{UV} = \nu L_\nu$ is the UV luminosity per logarithmic frequency) alone for a limited sample of galaxies is insufficient to probe the SFR variability because it can be very sensitive to complications such as dust attenuation, whereas characterizing the joint distribution of $L_{H\alpha}$ and L_{UV} , especially its width, with a large galaxy sample can

* E-mail: guochao.sun@northwestern.edu

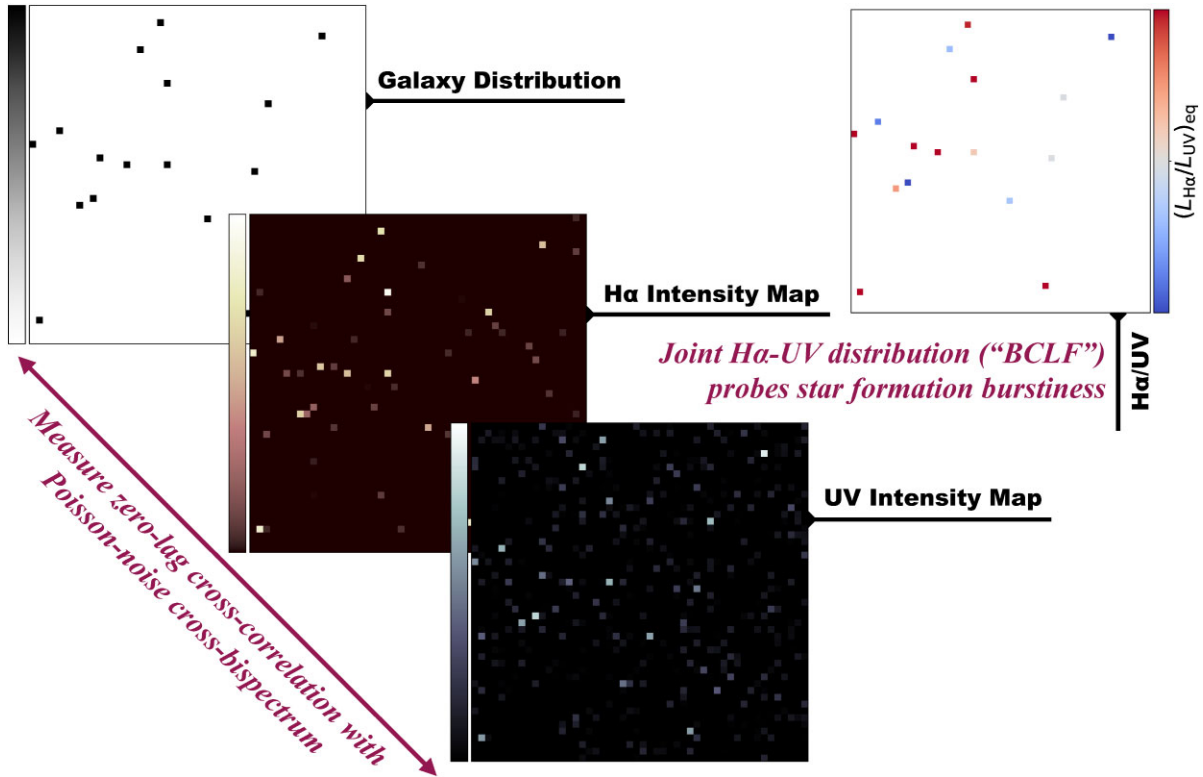


Figure 1. A graphical representation of the EBL–galaxy cross-correlation analysis investigated in this work for probing effects of bursty star formation on the joint distribution of SFR indicators $L_{H\alpha}$ and L_{UV} . Distributions of galaxies, $H\alpha$ line intensity, and UV continuum intensity are cross-correlated in Fourier space to measure the cross-bispectrum. This constrains the joint $H\alpha$ –UV luminosity distribution, especially its width, which reflects the scatter in $L_{H\alpha}/L_{UV}$ around the equilibrium value when star formation is time-steady. The Fourier-space cross-bispectrum analysis in the Poisson-noise-dominated limit is formally equivalent to a zero-lag cross-correlation (i.e. stacking) on galaxy positions in real space, as demonstrated in Section 2.2, but allows foregrounds and observational systematics to be more easily separated (Section 4).

be a lot more informative (Sparre et al. 2017; Emami et al. 2019). These limitations together make an extensive, mass-complete study of bursty star formation in galaxies of different properties at different cosmic times challenging.

Composed of the accumulated radiation from all the sources in the universe outside the Milky Way, the extragalactic background light (EBL) offers a wealth of information about the galaxy and star formation physics across cosmic time (Finke, Razzaque & Dermer 2010; Finke et al. 2022). At near-infrared (IR) wavelengths (corresponding to rest-frame optical/UV at high redshifts), its potential to constrain the star formation process in high-redshift galaxies has attracted increasing interest in recent years (see e.g. Sun et al. 2021; Mirocha, Liu & La Plante 2022; Scott, Upton Sanderbeck & Bird 2022; Sun 2022). Therefore, as an alternative approach to probe bursty star formation, we investigate in this work the possibility of statistically constraining the joint distribution of $L_{H\alpha}$ and L_{UV} by cross-correlating cosmological surveys of the near-IR EBL and galaxy distributions. Thanks to its unprecedented survey depth and sky coverage, the Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer mission (SPHEREx; Doré et al. 2014; Korngut et al. 2018; Crill et al. 2020) promises to accurately quantify sources of the EBL out to the epoch of reionization and thereby probe galaxy formation and evolution across a wide range of cosmic times. In synergy with wide-field galaxy surveys to be conducted by e.g. the Rubin Observatory Legacy Survey of Space and Time (Rubin/LSST; LSST Science Collaboration 2009) or the *Nancy Grace Roman Space Telescope* (Spergel et al. 2015), it

has been demonstrated that the EBL–galaxy cross-correlation can be detected at high significance in each spectral channel of SPHEREx, thereby allowing the mean rest-frame optical/UV emission spectrum of galaxies to be accurately measured (Cheng & Chang 2022). It is therefore interesting to explore whether the EBL–galaxy cross-correlation can help constrain bursty star formation in galaxies, including its mass and redshift dependence, and provide a test of galaxy formation theory.

In this paper, we conduct a proof-of-principle study of using the (near-IR) EBL–galaxy cross-correlation to probe bursty star formation. In particular, we focus on the cross-correlation between intensity maps of $H\alpha$ and UV continuum emission and the distribution of galaxies selected by their stellar mass. More specifically, we aim to constrain the joint distribution of $L_{H\alpha}$ and L_{UV} as a probe for the SFR variability by measuring the zero-lag cross-correlation of the distribution of mass-selected galaxy samples and intensity maps of $H\alpha$ and UV emission. As illustrated in Fig. 1, such a measurement can probe the decorrelation effect on the zero-lag cross-correlation caused by the scatter in the $L_{H\alpha}$ – L_{UV} joint distribution, which links to the SFR variability (though complications due to e.g. dust attenuation exist; see Section 4). To measure the zero-lag cross-correlation, we calculate the Poisson-noise cross-bispectrum in Fourier space, which is the optimal way to separate the signal of interest from other sources of confusion, including large-scale clustering, instrument noise, and observational systematics. We forecast the prospects for measuring this cross-correlation using SPHEREx and Rubin/LSST and demonstrate the utility for probing

bursty star formation in galaxies in different mass and redshift ranges.

The remainder of this paper is organized as follows. In Section 2, we first introduce a simple, semi-empirical model for the $L_{H\alpha}$ – L_{UV} joint distribution of galaxies conditioned on stellar mass. We then show, in the limit where Poisson fluctuations dominate over clustering, how the zero-lag cross-correlation in real space is equivalent to a measurement of the cross-bispectrum in Fourier space. Finally, we describe the full framework for constraining the $L_{H\alpha}$ – L_{UV} joint distribution with a set of correlation coefficients defined by cross-bispectra. In Section 3, we present main results of our analysis, including forecasts for the various cross-correlation signals and the implied constraints on the toy models considered in our case study for SPHEREx and Rubin/LSST. We discuss some limitations and caveats of the presented analysis in Section 4, before concluding in Section 5. A flat, Lambda cold dark matter cosmology consistent with the measurements from Planck Collaboration (2016) is adopted throughout this paper.

2 METHODS

2.1 Modelling the $L_{H\alpha}$ – L_{UV} joint distribution

2.1.1 Overview

While the modelling and analysis frameworks to be presented are generally applicable, for our proof-of-principle study in this paper, we investigate specifically the prospects for cross-correlating near-IR EBL maps measured by SPHEREx with distributions of galaxies from the Rubin/LSST photometric redshift survey, which is expected to measure the mean rest-UV/optical spectrum of galaxies at high significance up to $z \sim 4$ (Cheng & Chang 2022).

Given the wavelength coverage of SPHEREx (0.75–5 μm) and the redshift range over which high-quality photo- z measurements can be achieved by Rubin/LSST, we aim to optimize the chance of detecting the decorrelation between $H\alpha$ and UV luminosities due to bursty star formation, which is expected to be more pronounced in low-mass galaxies that are abundant but faint. For the longer time-scale SFR indicator, we choose the U -band (3500 \AA) luminosity¹ rather than the more commonly used far-UV (1500 \AA) luminosity because the former reaches lower redshifts ($z \simeq 1.2$) and maximizes the contrast in star formation time-scales compared to $H\alpha$ (Emami et al. 2021).

Performing the analysis at $z \sim 1$ rather than $z \sim 4$, is also motivated by the completeness limit of the Rubin/LSST photometric redshift survey, below which issues like selection bias due to incompleteness introduce significant systematics. Following Leauthaud et al. (2020), we can estimate the stellar mass range accessible by scaling from the 90 percent mass completeness limit of the COSMOS2015 catalogue (Laigle et al. 2016). For Rubin/LSST with i -band limiting magnitude of $i = 26.8$, the 90 percent mass completeness limits are $\log(M_*/M_\odot) = 8.55, 8.95, 9.25, 9.4$ at $z = 1, 1.5, 2, 2.5$, respectively, well below stellar masses at which simulations predict galaxies at these redshifts to exhibit a considerable level of scatter in $L_{H\alpha}/L_{UV}$ due to bursty star formation (Domínguez et al. 2015; Sparre et al. 2017).

¹Throughout this paper, we use UV and U band interchangeably when referring to the continuum emission to be studied together with $H\alpha$. For simplicity, we refer to it with the subscript U hereafter.

We analytically derive a conditional luminosity function (CLF)-based description of the different moments of $L_{H\alpha}$ and L_U necessary for the cross-correlation. Since galaxies in different stellar mass bins will be analysed separately, the luminosity distributions are conditioned on stellar mass M_* . The exact parametrization is based on semi-empirical models of galaxy evolution and $H\alpha$ and UV emission, which are verified against the matching between the observed U -band luminosity functions (e.g. Moutard et al. 2020) and stellar mass functions (e.g. Shuntov et al. 2022) at redshifts of interest.

2.1.2 $H\alpha$ – UV bivariate conditional luminosity function (BCLF)

Taking $\Phi(L)$ to be the probability distribution function (PDF) of the luminosity L such that $\int \Phi(L)dL = 1$, we can write the joint PDF of $L_{H\alpha}$ and L_U conditioned on M_* as

$$\Phi(L_{H\alpha}, L_U|M_*) = \Phi(L_{H\alpha}|L_U, M_*)\Phi(L_U|M_*), \quad (1)$$

where on the right-hand side the first term $\Phi(L_{H\alpha}|L_U, M_*)$ is given by a lognormal distribution around the mean $H\alpha$ luminosity $\bar{L}_{H\alpha} = L_{U,0}(L_U/L_{U,0})^\beta$, following the functional form from Mehta et al. (2015), with a logarithmic scatter of $\sigma_{\alpha U}(M_*)$. The second term, $\Phi(L_U|M_*)$, often referred to as the CLF (see e.g. Yang, Mo & van den Bosch 2003), is the distribution of L_U conditioned on M_* that can be determined by matching the observed stellar mass function and UV luminosity function. For $\Phi(L_U|M_*)$, we also consider a lognormal distribution specified by some mean relation $\bar{L}_U(M_*)$ and a logarithmic scatter σ_{LM} . Putting these ingredients together, we define a *bivariate conditional luminosity function* (BCLF) of $L_{H\alpha}$ and L_U , $\Phi(L_{H\alpha}, L_U|M_*)$, that is the product of

$$\Phi(L_{H\alpha}|L_U, M_*) = \frac{\exp\left\{-\frac{[\ln L_{H\alpha} - \ln \bar{L}_{H\alpha}(L_U)]^2}{2\sigma_{\alpha U}^2(M_*)}\right\}}{\sqrt{2\pi}\sigma_{\alpha U}(M_*)L_{H\alpha}} \quad (2)$$

and

$$\Phi(L_U|M_*) = \frac{1}{\sqrt{2\pi}\sigma_{LM}L_U} \exp\left\{-\frac{[\ln L_U - \ln \bar{L}_U(M_*)]^2}{2\sigma_{LM}^2}\right\}, \quad (3)$$

which satisfies

$$\bar{L}_U(M_*) = e^{-\sigma_{LM}^2/2} \int_0^\infty dL_U \Phi(L_U|M_*)L_U. \quad (4)$$

By the definition of the CLF, equation (3) can in principle be determined by finding the appropriate functional form of $\bar{L}_U(M_*)$ and the value of σ_{LM} that best matches the observed U -band luminosity function $\phi(L_U) = dn/dL_U$ and stellar mass function $\psi(M_*) = dn/dM_*$, where n is the number density of galaxies. In this work, however, we construct a simple, parametric model of $\bar{L}_U(M_*)$ and σ_{LM} based on the specific SFR–stellar mass relation from semi-empirical models of galaxy formation given by the UNIVERSEMACHINE code (Behroozi et al. 2019) and the observed U -band luminosities of galaxies from Zhou et al. (2017). As a sanity check, we have verified our simple model by comparing its predicted U -band luminosity function against the observed ones at redshifts where measurements are available (Moutard et al. 2020).

To describe $H\alpha$ and U -band continuum emission, we take $L_{H\alpha} = 2.1 \times 10^{41} \text{ erg s}^{-1} (\text{SFR}/M_\odot \text{ yr}^{-1})$, valid for the Chabrier initial mass function (Chabrier 2003) assumed in this work, and adopt the attenuation-corrected, empirical relation between U band and $H\alpha$ luminosities from Zhou et al. (2017), who provide a calibration of the U -band luminosity as an SFR indicator. Because both these luminosities and the stellar masses they are anchored to are dust-corrected, to properly model their observed strengths in our cross-correlation analysis, we must reapply dust attenuation. To do this

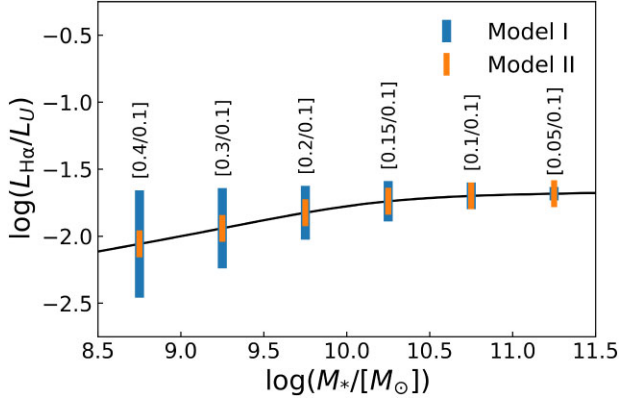


Figure 2. Illustration of the average $\log(L_{\text{H}\alpha}/L_U)$ and the scatter around it (specified in brackets in units of dex) as a function of stellar mass described by the baseline model (Model I) and its variant (Model II) considered in this work. The scatters are overplotted on the mean relation in the six stellar mass bins uniformly distributed over $8.5 < \log(M_*/M_\odot) < 11.5$. The growth of scatter with decreasing stellar mass as in Model I is often considered as an indication of an increasing level of bursty star formation.

self-consistently, we assume the $A_{\text{FUV}}(M_*)$ relation from McLure et al. (2018) that is derived for star-forming galaxies at $z \sim 2-3$,

$$A_{\text{FUV}} = 2.293 + 1.16\mathcal{M}_{10} + 0.256\mathcal{M}_{10}^2 + 0.209\mathcal{M}_{10}^3, \quad (5)$$

where $\mathcal{M}_{10} = \log(M_*/10^{10} M_\odot)$, and the Calzetti et al. (2000) dust attenuation curve, which implies $A_{\text{H}\alpha} = 0.44A_{\text{FUV}}$ and $A_U = 0.62A_{\text{FUV}}$, respectively.²

With the BCLF of $L_{\text{H}\alpha}$ and L_U , the ensemble averages that enter our cross-correlation analysis can then be written as

$$\langle L_U L_{\text{H}\alpha} \rangle \propto \int dM_* \Phi(L_{\text{H}\alpha}, L_U | M_*) \psi(M_*) \times \iint dL_U dL_{\text{H}\alpha} 10^{-0.4(A_{\text{H}\alpha} + A_U)} L_U L_{\text{H}\alpha}, \quad (6)$$

$$\langle L_{\text{H}\alpha}^2 \rangle \propto \int dM_* \Phi(L_{\text{H}\alpha}, L_U | M_*) \psi(M_*) \times \iint dL_U dL_{\text{H}\alpha} 10^{-0.8A_{\text{H}\alpha}} L_{\text{H}\alpha}^2, \quad (7)$$

and

$$\langle L_U^2 \rangle \propto \int dM_* \Phi(L_U | M_*) \psi(M_*) \int dL_U 10^{-0.8A_U} L_U^2, \quad (8)$$

where $\psi(M_*)$ is the stellar mass function that we self-consistently obtain from UNIVERSEMACHINE, and $\langle \dots \rangle$ implicitly assumes that the ensemble average is taken for the sample of stellar-mass-selected galaxies over the mass bin $[M_*, M_* + \Delta M_*]$. We have also confirmed that using the latest observed stellar mass functions (e.g. Shuntov et al. 2022) has little impact on our results.

As illustrated in Fig. 2 and summarized in Table 1, two toy models of the BCLF of $L_{\text{H}\alpha}$ and L_U are considered for our subsequent analysis. The fiducial model, Model I, assumes that the scatter, $\sigma_{\alpha U}$, increases with decreasing stellar mass, whereas the contrasting model, Model II, assumes a constant $\sigma_{\alpha U} = 0.1$ dex

²Following McLure et al. (2018), we assume that $E(B - V)_{\text{star}} = 0.76E(B - V)_{\text{neb}}$ and derive $A_\lambda = k_\lambda E(B - V)$ from $k_\lambda = 2.659 \times (-2.156 + 1.509/\lambda - 0.198/\lambda^2 + 0.011/\lambda^3) + 4.05$ for $0.12 \mu\text{m} < \lambda < 0.63 \mu\text{m}$ (rest frame) or $k_\lambda = 2.659 \times (-1.857 + 1.040/\lambda) + 4.05$ for $0.63 \mu\text{m} < \lambda < 2.2 \mu\text{m}$, as in Calzetti et al. (2000).

Table 1. Specifications of the toy models considered in this work. The scatter $\sigma_{\alpha U}$ is allowed to vary across the six stellar mass bins uniformly distributed over $8.5 < \log(M_*/M_\odot) < 11.5$.

Model	$\sigma_{\alpha U}$ (dex)	σ_{LM} (dex)	β	$L_{U,0}^a$ (erg s ⁻¹)
I	0.4, 0.3, 0.2, 0.15, 0.1, 0.05	0.2	1.25	3.55×10^{51}
II	0.1, 0.1, 0.1, 0.1, 0.1, 0.1	0.2	1.25	3.55×10^{51}

^aThe exact value of $L_{U,0}$ does not impact the cross-correlation coefficients (Section 2.1.3) but affects the expected detectability of cross-correlation.

across all stellar mass bins. For both models, we further assume a constant $\sigma_{\text{LM}} = 0.2$ dex, consistent with the scatter in the light-to-mass ratio observed and commonly assumed in semi-empirical models of high- z galaxies (More et al. 2009; Sun & Furlanetto 2016), whereas $\beta = 1.25$ and $L_{U,0} = 3.55 \times 10^{51}$ erg s⁻¹ are suggested by the best-fitting relation to the observed correlation between H α and U-band luminosities of galaxies (Zhou et al. 2017). We note that even though more accurately modelling the H α -UV BCLF is beyond the scope of this study, our simple parametrization of the mean relations is grounded on empirical models that reliably describe galaxy evolution and the production of H α and U-band emission at the redshifts of interest. The two contrasting cases for $\sigma_{\alpha U}$ are chosen to roughly bracket the range of possible mass dependence of the width of $L_{\text{H}\alpha}$ - L_U distribution as a proxy for star formation burstiness, motivated by observations and numerical simulations (Weisz et al. 2012; Domínguez et al. 2015; Sparre et al. 2017; Emami et al. 2019; Faisst et al. 2019).

2.1.3 Connection to the EBL-galaxy cross-correlation

From the ensemble averages defined above and their dependence on our BCLF model parameters, we can obtain a few simple and useful expressions that connect cross-correlation observables to these model parameters. The observable most directly related to the cross-correlation analysis is the cross-correlation coefficient, $r_{\times}^g(\ell)$, which characterizes how correlated the two SFR tracer fields are for the galaxy population g of interest. As will be shown in Section 2.3, when measured in the Poisson-noise limit in Fourier space, the cross-correlation coefficient $r_{\times,P}^g \equiv r_{\times}^g(\ell \gg \ell_c)$ takes the simple form

$$r_{\times,P}^g = \frac{B_{\ell,P}^{U,\text{H}\alpha,g}}{\sqrt{B_{\ell,P}^{U,U,g} B_{\ell,P}^{H\alpha,H\alpha,g}}} \propto \frac{\langle L_U L_{\text{H}\alpha} \rangle}{\sqrt{\langle L_U^2 \rangle \langle L_{\text{H}\alpha}^2 \rangle}}. \quad (9)$$

Here, the multipole moment ℓ_c denotes some characteristic scale (to be estimated from a power spectrum analysis) at which non-linear clustering is comparable to the Poisson noise, and $B_{\ell,P}^{i,j,k}$ denotes the Poisson-noise-limit cross-bispectrum of fields i , j , and k . In Section 2.2, we will first motivate the understanding of the cross-correlation of interest in both real and Fourier spaces. We will then detail how to arrive at the proportionality, and derive the components of $r_{\times,P}^g$ and their uncertainties, in Section 2.3.

Combining equations (1) through (9), we can show that $r_{\times,P}^g$ is in fact insensitive to the $\tilde{L}_U(M_*)$ parametrization or the value of $L_{U,0}$, and obtain

$$\ln[r_{\times,P}^g] = - \left[\frac{\sigma_{\alpha U}^2}{2} + \frac{\sigma_{\text{LM}}^2(\beta - 1)^2}{2} \right]. \quad (10)$$

It is easy to see that $r_{\times,P}^g$ drops below unity if either $\sigma_{\alpha U}$ or σ_{LM} (as long as β is not strictly 1) is non-zero. While the latter characterizes the intrinsic scatter in the mass-to-light ratio of galaxies due to stochasticity in e.g. mass accretion rates (McBride, Fakhouri & Ma 2009; Fakhouri, Ma & Boylan-Kolchin 2010; van den Bosch et al. 2014), the former may be largely attributed to the time variability of the SFR. Because constraints on bursty star formation mainly come from the comparison of $r_{\times,P}^g$ in different stellar mass bins instead of its exact values, factors that are generally mass independent will not significantly complicate the interpretation. For reference, assuming $\sigma_{LM} = 0$, we have $r_{\times,P}^g = 0.97, 0.79$, and 0.52 for $\sigma = 0.1, 0.3$, and 0.5 dex, respectively.

By analogy to the cross-correlation coefficient, $r_{\times,P}^g$, we can also define and derive the following autocorrelation coefficients for H α and UV emission

$$\ln(r_{H\alpha,P}^g) = \ln\left(\frac{C_{\ell,P}^{H\alpha,g}}{\sqrt{B_{\ell,P}^{H\alpha,H\alpha,g}}}\right) = -\frac{\sigma_{\alpha U}^2 + \sigma_{LM}^2 \beta^2}{2} \quad (11)$$

and

$$\ln(r_{UV,P}^g) = \ln\left(\frac{C_{\ell,P}^{UV,g}}{\sqrt{B_{\ell,P}^{UV,UV,g}}}\right) = -\frac{\sigma_{LM}^2}{2}, \quad (12)$$

where $C_{\ell,P}^{H\alpha,g}$ and $C_{\ell,P}^{UV,g}$ are the Poisson-noise terms of the angular cross-power spectra of H α and UV emission, respectively, with galaxies to be defined in Section 2.3. Equations (10)–(12) therefore connect correlation coefficients directly measurable from the EBL–galaxy cross-correlation to parameters of our BCLF model, which can be individually constrained by solving these equations. Although we will focus on the analysis of the BCLF hereafter, for completeness, in Appendix A we also derive the mean and variance of the luminosity ratio, $L_{H\alpha}/L_U$, as two examples of other potentially useful measures of the BCLF and thus the star formation burstiness.

2.2 Relationship between the real-space zero-lag cross-correlation and the Fourier-space cross-bispectrum

Here, before presenting the full cross-correlation analysis framework in Fourier space, we start with a demonstration of how the Poisson-noise cross-bispectrum to be analysed relates to the zero-lag cross-correlation (i.e. stacking) in real space, which might be more intuitive to understand as a well-established method to probe astrophysics beyond the reach of individually targeted observations (see e.g. Viero et al. 2022, for a recent stacking analysis of the dust-obscured star formation in high- z galaxies). By showing that they are essentially equivalent, we aim to build up the physical intuition to comprehend details of the full, Fourier-space treatment to be described in Section 2.3.

To demonstrate the equivalence of cross-correlation analyses performed in Fourier and real spaces, it is sufficient to compare the signal-to-noise ratio (S/N) derived in both cases as a measure of the information available. For a zero-lag cross-correlation of intensity maps j and k with galaxies in real space, in the Poisson-noise-dominated limit, the S/N scales as

$$\left(\frac{S}{N}\right)_{rs} \sim (N_{gal})^{1/2} \left(\frac{\langle \nu I_v^j \rangle}{\sigma_{pix,N}^j}\right) \left(\frac{\langle \nu I_v^k \rangle}{\sigma_{pix,N}^k}\right) \frac{\langle L_j L_k \rangle}{\langle L_j \rangle \langle L_k \rangle}, \quad (13)$$

which is a product of the cross-correlation coefficient, the S/N per pixel of the intensity maps, and a scaling factor for the noise

reduction when ‘stacking’ on N_{gal} galaxies. Using definitions of cross-bispectrum and its uncertainty to be introduced in Section 2.3, we can show that the S/N of cross-bispectrum B_{ℓ}^{ijk} defined in Fourier space resembles equation (13) in the Poisson-noise limit. Specifically, we have (see Section 2.3.2 for details)

$$\begin{aligned} \left(\frac{S}{N}\right)_{\times} &= \left(\frac{B_{\ell,P}^{ijk}}{\delta B_{\ell,P}^{ijk}}\right)^2 \approx \sum_{\ell_1, \ell_2, \ell_3} \frac{[B_{\ell}^{ijk}(\ell_1, \ell_2, \ell_3)]^2}{C_{\ell}^i(\ell_1) C_{\ell}^j(\ell_2) C_{\ell}^k(\ell_3)} \Omega_s \ell_{max} \Delta \ell_1 \Delta \ell_2 \Delta \ell_3 \\ &\approx \ell_{max}^4 \Omega_s \frac{(B_{\ell,P}^{ijk})^2}{C_{\ell,P}^i C_{\ell,P}^j C_{\ell,P}^k}, \end{aligned} \quad (14)$$

where the approximation $N_{trip} \approx \ell_{max}^2 \Omega_s^2 \Delta \ell_1 \Delta \ell_2 \Delta \ell_3 \approx \ell_{max}^4 \Omega_s^2$ is applied. Note that here $\ell_{max} \approx \theta_{pix}^{-1}$, where θ_{pix} is the pixel size in steradian, and Ω_s is the survey size. As will be shown in Section 2.3.2, we can write the angular power spectra as $C_{\ell,P}^i = \Omega_s N_{gal}^{-1}$, $C_{\ell,P}^j = (\sigma_{pix,N}^j)^2 \theta_{pix}^2$, and $C_{\ell,P}^k = (\sigma_{pix,N}^k)^2 \theta_{pix}^2$, whereas the cross-bispectrum scales as

$$B_{\ell,P}^{ijk} \propto \langle \nu I_v^j \rangle \langle \nu I_v^k \rangle \frac{\langle L_j L_k \rangle}{\langle L_j \rangle \langle L_k \rangle}. \quad (15)$$

Putting together, we can recover the form of equation (13) from equation (14).

Therefore, we stress that, while measuring a zero-lag cross-correlation in real space is mathematically equivalent to measuring a Poisson-noise cross-bispectrum in Fourier space, we choose to work in Fourier space below given practical considerations in observational data analysis that favour it as a more robust and unbiased method. For example, the finite angular and spectral resolution of SPHEREx imply that the pure zero-lag cross-correlation is not strictly observable. The separation between the clustering contributions and Poisson fluctuations is then more transparent in Fourier space, as are the treatment of the beam, spectral resolution, foreground contamination, and pixel noise, while the analysis may also be more easily generalized to incorporate clustering terms.

2.3 The EBL–galaxy cross-correlation: signals and errors

2.3.1 Cross-power spectra and cross-bispectra

Following Cheng & Chang (2022), we can write the cross-power spectra between H α /UV emission and galaxies in the Poisson-noise limit as

$$C_{\ell,P}^{H\alpha,g} = \frac{1}{\sigma_g} \Delta z_g \frac{d\nu I_v}{dz} \Big|_{H\alpha} \quad (16)$$

and

$$C_{\ell,P}^{UV,g} = \frac{1}{\sigma_g} \Delta z_g \frac{d\nu I_v}{dz} \Big|_U. \quad (17)$$

By analogy to the definition of cross-power spectra, three fields (two factors of intensity map and one factor of galaxy distribution) are required to calculate ensemble averages involving the second moment of luminosity, $\langle \mathcal{O}(L^2) \rangle$. We therefore define the cross-bispectrum as an integral of the differential flux densities $d(\nu I_v)/dz$ of H α and UV emission [which themselves are mass integrals over the galaxy population described by the stellar mass function $\psi(M_*)$] over redshift, conditioned on the subgroup of galaxies selected by stellar mass. When a narrow redshift range $\Delta z_g \ll 1$ is considered, the redshift integral $\int_{\Delta z_g} F(z) dz$ can be approximated as $F(z_g) \Delta z_g$, which simplifies the calculations.

For H α (line) and UV (continuum) emission, we can write their differential flux densities as³

$$\left. \frac{d\nu I_\nu}{dz} \right|_{H\alpha} = \frac{1}{\Delta z_{H\alpha}} \int_{M_{*,\min}}^{M_{*,\max}} dM_* \psi(M_*) \frac{\nu L_{H\alpha}}{4\pi D_L^2} \frac{d\chi}{d\nu} D_{A,\text{com}}^2$$

$$\simeq \frac{1}{\Delta z_g} \frac{c/H(z_g)}{4\pi(1+z_g)} \int dM_* \psi(M_*) L_{H\alpha}, \quad (18)$$

$$\left. \frac{d\nu I_\nu}{dz} \right|_U = \int_{M_{*,\min}}^{M_{*,\max}} dM_* \psi(M_*) \frac{L_U}{4\pi D_L^2} \frac{d\chi}{dz} D_{A,\text{com}}^2$$

$$= \frac{c/H(z_g)}{4\pi(1+z_g)^2} \int dM_* \psi(M_*) L_U, \quad (19)$$

and the density of galaxies (per unit solid angle) is

$$\sigma_g \simeq \Delta z_g \frac{dN_g}{dz d\Omega} = \frac{\chi^2(z_g) c \Delta z_g}{H(z_g)} \int dM_* \psi(M_*), \quad (20)$$

where $H(z)$, χ , D_L , and $D_{A,\text{com}} = \chi$ are the Hubble parameter, the comoving radial distance, the luminosity distance, and the comoving angular diameter distance, respectively. The χ gradients are given by $d\chi/d\nu = c(1+z)/[\nu H(z)]$ for the observed frequency ν , and $d\chi/dz = c/H(z)$. We assume $\Delta z_g \approx \Delta z_{H\alpha} = (1+z)/R$ with R being the spectral resolving power. Note that both L_U and $L_{H\alpha}$ are defined to be non-specific luminosities in units of erg s^{-1} that, to the first order, scale with the SFR and thus M_* . Unless otherwise specified when the mass integral spans the full range of stellar mass from $M_{*,\min} = 10^{7.5} M_\odot$ to $M_{*,\max} = 10^{11.5} M_\odot$ (as in equations 18 and 19), the stellar mass integral is by default over ΔM_* , which selects the subgroup of galaxies in the stellar mass bin of interest.

With equations (18) and (19), the Poisson-noise-limit cross-bispectrum of the H α line, U -band continuum, and galaxy fields can be written as

$$B_{\ell,P}^{\times} \equiv B_{\ell,P}^{H\alpha,U,g}$$

$$= \frac{1}{\sigma_g} \int dz \frac{H(z)}{c\chi^2(z)\Delta z_{H\alpha|g}} \int dM_* \psi(M_*) \times$$

$$\left[\frac{\nu L_{H\alpha}}{4\pi D_L^2} \frac{d\chi}{d\nu} D_{A,\text{com}}^2 \right] \left[\frac{L_U}{4\pi D_L^2} \frac{d\chi}{dz} D_{A,\text{com}}^2 \right] \Phi(L_{H\alpha}, L_U | M_*)$$

$$= \frac{c \int dM_* \Phi(L_{H\alpha}, L_U | M_*) \psi(M_*) L_{H\alpha} L_U}{16\pi^2 \sigma_g H(z_g) (1+z_g)^3 \chi^2(z_g)}, \quad (21)$$

where $\Delta z_{H\alpha|g} \approx \Delta z_{H\alpha} \approx \Delta z_g$ denotes the redshift range over which galaxy and emission intensity fields overlap. Similarly, for the $\langle L_U^2 \rangle$ and $\langle L_{H\alpha}^2 \rangle$ (autocorrelation) terms in the denominator of equation (9), we have

$$B_{\ell,P}^{U,U,g} = \frac{1}{\sigma_g} \int dz \frac{H(z)}{c\chi^2(z)} \int dM_* \psi(M_*, z)$$

$$\times \left[\frac{L_U}{4\pi D_L^2} \frac{d\chi}{dz} D_{A,\text{com}}^2 \right]^2 \Phi(L_U | M_*)$$

$$= \frac{c \Delta z_g \int dM_* \Phi(L_U | M_*) \psi(M_*, z_g) L_U^2}{H(z_g) 16\pi^2 \sigma_g (1+z_g)^4 \chi^2(z_g)}, \quad (22)$$

³Note that we omit the convolution with the conditional PDFs of the luminosities in the two expressions below for brevity, but include them in the full expressions for $B_{\ell,P}$ below.

and

$$B_{\ell,P}^{H\alpha,H\alpha,g} = \frac{1}{\sigma_g} \int dz \frac{H(z)}{c\chi^2(z)} \frac{1}{\Delta z_{H\alpha}} \int dM_* \psi(M_*, z)$$

$$\times \left[\frac{\nu L_{H\alpha}}{4\pi D_L^2} \frac{d\chi}{d\nu} D_{A,\text{com}}^2 \right]^2 \Phi(L_{H\alpha}, L_U | M_*)$$

$$= \frac{1}{\sigma_g} \frac{c}{H(z_g) \Delta z_g} \frac{\int dM_* \psi(M_*, z_g) L_{H\alpha}^2}{16\pi^2 (1+z_g)^2 \chi^2(z_g)}. \quad (23)$$

2.3.2 Uncertainties on cross-power spectra and cross-bispectra

For the cross-power spectrum between an intensity map i (H α or U -band intensity map here) and galaxies, the uncertainty for a given multipole moment ℓ binned in a width of $\Delta\ell$ can be expressed as

$$\left(\delta C_{\ell,P}^{i,g} \right)^2 = \frac{1}{f_{\text{sky}}(2\ell+1)\Delta\ell} \left[\left(C_{\ell,P}^{i,g} \right)^2 + C_{\ell,N}^i C_{\ell,P}^g \right], \quad (24)$$

where f_{sky} is the sky covering fraction and we assume here that autocorrelations of the intensity map and galaxies are dominated by the instrument noise and the Poisson noise, respectively, on the small scales considered in our analysis. In practice, to obtain the net effective uncertainty of the cross-power spectrum, we further scale down equation (24) by a factor of 300 to approximate the gain in sensitivity from binning together modes over $10^4 < \ell < 10^5$. This renders S/N of $C_{\ell,P}^{H\alpha,g}$ (or $C_{\ell,P}^{U,g}$) substantially higher than that of $B_{\ell,P}^{H\alpha,H\alpha,g}$ (or $B_{\ell,P}^{U,U,g}$), as will be detailed below, and therefore the S/N of autocorrelation coefficients $r_{H\alpha,P}^g$ (or $r_{U,P}^g$) can be simply approximated as twice of that of $B_{\ell,P}^{H\alpha,H\alpha,g}$ (or $B_{\ell,P}^{U,U,g}$).

Following Kayo, Takada & Jain (2013), we can write the bispectrum variance in the Gaussian approximation as

$$\text{Var} \left[B_{\ell}^{ijk}(\ell_1, \ell_2, \ell_3) \right] = \frac{\Omega_s C_{\ell_1}^i(\ell_1) C_{\ell_2}^j(\ell_2) C_{\ell_3}^k(\ell_3)}{N_{\text{trip}}(\ell_1, \ell_2, \ell_3)}, \quad (25)$$

where Ω_s is the total survey area over which EBL and galaxy surveys overlap ($\Omega_s \approx 5.5$ sr for SPHEREx and Rubin/LSST), and the number of triplets that form closed triangles in Fourier space $N_{\text{trip}}(\ell_1, \ell_2, \ell_3) \equiv \sum_{\mathbf{q}_i: \mathbf{q}_i \in \ell_i} \Delta \mathbf{q}_{123}$, which can be approximated in the limit of large multipole bins as

$$N_{\text{trip}} \simeq \frac{\Omega_s^2 \ell_1 \ell_2 \ell_3 \Delta \ell_1 \Delta \ell_2 \Delta \ell_3 / 2\pi^3}{\sqrt{2\ell_1^2 \ell_2^2 + 2\ell_1^2 \ell_3^2 + 2\ell_2^2 \ell_3^2 - \ell_1 - \ell_2 - \ell_3}}. \quad (26)$$

Each of the three angular autopower spectra in the numerator of equation (25) has contributions from clustering,⁴ Poisson noise, and instrument noise (for intensity maps of H α and UV emission) whose relative importance varies across ℓ . Specifically, assuming Limber approximation and narrow redshift range $\Delta z_g \ll 1$, we have (Cheng & Chang 2022)

$$C_{\ell,\text{cl}}^g(\ell) = \frac{H(z_g) \langle b \rangle_g^2(z_g)}{\Delta z_g c \chi^2(z_g)} P_{\delta\delta} \left[k = \frac{\ell}{\chi(z_g)}, z_g \right] \quad (27)$$

and

$$C_{\ell,P}^g = \left(\frac{dN_g}{d\Omega} \right)^{-1} = \sigma_g^{-1} \quad (28)$$

for the autopower spectrum of galaxies, where $\langle b \rangle_g$ is the galaxy bias averaged over the ensemble of galaxies in the stellar mass bin

⁴For simplicity, we ignore the non-linear clustering whose impact on scales smaller than $\ell \sim 10^4$ is expected to be subdominant to that of the Poisson noise (see e.g. Cheng & Bock 2022).

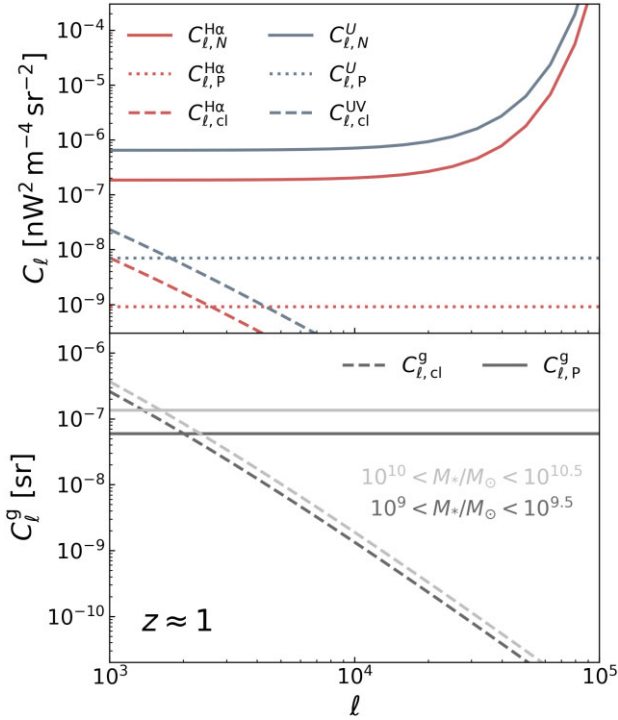


Figure 3. A comparison of the error budget for C_ℓ of $H\alpha$ and UV emission (top panel), as well as the galaxy distribution (bottom panel) at $z \approx 1.5$. At high multipoles $\ell \geq 10^4$, uncertainties of the intensity (galaxy) power spectra are strongly dominated by the instrument noise $C_{\ell,N}$ (Poisson noise $C_{\ell,P}^g$). Note that, unlike in the bottom panel, the sample variances in the top panel are evaluated by integrating over the full range of stellar mass $[M_{*,\min}, M_{*,\max}]$.

of width ΔM_* (see Appendix B for a more detailed description of the various bias factors involved) and $P_{\delta\delta}$ is the dark matter power spectrum. Similarly, for $H\alpha$ and UV emission, the autopower spectra are

$$C_{\ell,cl}^{H\alpha}(\ell) = \int dz \frac{H(z)}{c\chi^2(z)} b_{H\alpha}^2(z) \left[\frac{dv I_v}{dz} \Big|_{H\alpha}(z) \right]^2 \times P_{\delta\delta} \left[k = \frac{\ell}{\chi(z)}, z \right], \quad (29)$$

$$C_{\ell,P}^{H\alpha} = \int dz \frac{H(z)}{c\chi^2(z)} \frac{1}{\Delta z_{H\alpha}^2} \int_{M_{*,\min}}^{M_{*,\max}} dM_* \psi(M_*, z) \times \left[\frac{v L_{H\alpha}}{4\pi D_L^2} \frac{d\chi}{dv} D_{A,com}^2 \right]^2 \Phi(L_{H\alpha}, L_U | M_*), \quad (30)$$

$$C_{\ell,cl}^U(\ell) = \int dz \frac{H(z)}{c\chi^2(z)} b_U^2(z) \left[\frac{dv I_v}{dz} \Big|_U(z) \right]^2 \times P_{\delta\delta} \left[k = \frac{\ell}{\chi(z)}, z \right], \quad (31)$$

and

$$C_{\ell,P}^U = \int dz \frac{H(z)}{c\chi^2(z)} \int_{M_{*,\min}}^{M_{*,\max}} dM_* \psi(M_*, z) \times \left[\frac{L_U}{4\pi D_L^2} \frac{d\chi}{dz} D_{A,com}^2 \right]^2 \Phi(L_U | M_*). \quad (32)$$

As shown by Fig. 3, on small scales the Poisson noise and instrument noise dominate the angular power spectra of galaxies and emission fields, respectively. Therefore, we take $C_\ell^i = C_{\ell,P}^i = \sigma_g^{-1}$, $C_\ell^j = C_{\ell,N}^{H\alpha} = \sigma_{\text{pix},N}^2 |_{\lambda_{H\alpha}(1+z)} \Omega_{\text{pix}} e^{\Omega_{\text{pix}} \ell^2}$ and, $C_\ell^k =$

$C_{\ell,N}^U = \sigma_{\text{pix},N}^2 |_{\lambda_U(1+z)} \Omega_{\text{pix}} e^{\Omega_{\text{pix}} \ell^2}$, where $\sigma_{\text{pix},N}$ is the projected surface brightness sensitivity of the SPHEREx all-sky survey.⁵ To estimate the detectability of the bispectrum in terms of its total S/N, we adopt a universal bin size of $\Delta\ell = 1000$ and sum the S/N of individual ℓ bins over $\ell_{\min} = 10^4$ to $\ell_{\max} = 10^5$, where the angular power spectra are well within the Poisson-noise-dominated regime, namely

$$\left(\frac{S}{N} \right)_\times^2 = \sum_{\{\ell_1, \ell_2, \ell_3\}=\ell_{\min}-\frac{\Delta\ell}{2}}^{\ell_{\max}+\frac{\Delta\ell}{2}} \frac{(B_{\ell,P}^{ijk})^2}{\text{Var}[B_{\ell}^{ijk}(\ell_1, \ell_2, \ell_3)]}. \quad (33)$$

Finally, from the definition of $r_{\times,P}^g$, we have

$$\left(\frac{S}{N} \right)_{r_\times}^{-2} = \left(\frac{S}{N} \right)_\times^{-2} + \frac{1}{4} \left[\left(\frac{S}{N} \right)_{H\alpha}^{-2} + \left(\frac{S}{N} \right)_U^{-2} \right]. \quad (34)$$

3 RESULTS

In this section, we first present the detectability of the various cross-bispectra related to our case study, where we cross-correlate EBL maps of rest-frame $H\alpha$ and UV (U -band) emission and photometric galaxies to be observed with SPHEREx and Rubin/LSST, respectively (Section 3.1). Then, we show the constraints on BCLF model parameters derived from the predicted sensitivity to the correlation coefficients, $r_{\times,P}^g$, $r_{H\alpha,P}^g$, and $r_{U,P}^g$ (Section 3.2). The toy models considered here suffice to forecast the potential for EBL–galaxy cross-correlations to distinguish these limiting cases and thereby shed light on bursty star formation.

3.1 Detectability of cross-correlation signals

In the left-hand panel of Fig. 4, we show the predicted detectability of the cross-bispectrum, $B_{\ell,P}^{H\alpha,U,g}$, of $H\alpha$ and U -band intensity maps measured by the all-sky survey with SPHEREx and photo- z galaxies surveyed by Rubin/LSST in each of the six stellar mass bins. The S/N numbers quoted here are evaluated for a single pair of spectral channels corresponding to a narrow redshift range of $\Delta z = (1 + z)/R$ around $z = 1.5$, where $R = 41$ is the spectral resolving power of SPHEREx in bands relevant to this study. We note that at the redshifts of interest for this study ($z \sim 1.5$ – 2.5), the adopted Δz happens to be comparable to the level of photometric redshift uncertainty expected for the nominal 10-yr Rubin/LSST survey, which may be further improved over the course of the survey by the addition of near-IR and UV photometry from other existing/concurrent surveys, such as *Roman*, *Euclid*, and SPHEREx (Graham et al. 2018, 2020). Due to the trade-off between the brightness of sources and the number density of galaxies contributing to the intensity fields and available for cross-correlation, the expected S/N of $B_{\ell,P}^{H\alpha,U,g}$ peaks at intermediate mass scales $M_* \sim 10^{10.5} M_\odot$, although a high-significance detection can be achieved in all but the lowest mass bins. Meanwhile, from the comparison between cases with and without dust attenuation, it is clear that the expected detectability of the EBL–galaxy cross-correlation is highly sensitive to the treatment of dust attenuation (especially for massive galaxies that are more dust-rich), which has sometimes been neglected for simplicity in previous work, although dust attenuation will likely reduce the SNR of EBL observations with SPHEREx (e.g. Gong et al. 2017).

⁵See the public data product of surface brightness sensitivity available at https://github.com/SPHEREx/Public-products/blob/master/Surface_Brightness_v28_base_cbe.txt.

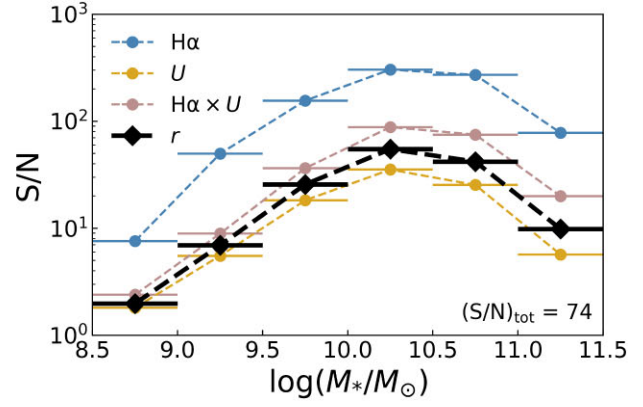
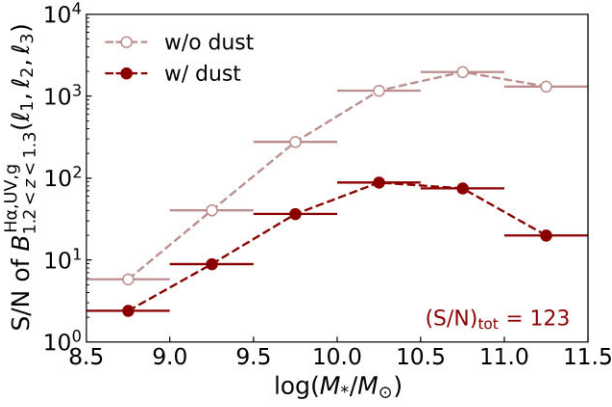


Figure 4. Left: S/N of the Poisson-noise cross-bispectra of H α , UV, and galaxies at $z \approx 1.25$ in different stellar mass bins, before and after including dust attenuation. The fiducial model (Model I) is assumed and the total S/N is quoted for the sum over all stellar mass bins. Right: a comparison of the detectability of the cross-correlation coefficient, $r^g_{x,p}$ (black), as well as its three components, namely $B^{H\alpha,U,g}_{l,p}$ (red), $B^{H\alpha,H\alpha,g}_{l,p}$ (blue), and $B^{U,U,g}_{l,p}$ (yellow). The fiducial model (Model I) is assumed, after including dust attenuation. All the data displayed here are evaluated for a single redshift interval, without redshift binning (see Section 3.1).

In the right-hand panel of Fig. 4, we show how the S/N of each bispectrum involved in the definition of $r^g_{x,p}$ can be propagated to obtain the S/N of $r^g_{x,p}$ (see equations 9 and 34). As shown by the comparison, the detectability of $r^g_{x,p}$, from which constraints on $\sigma_{\alpha U}$ (and other BCLF model parameters) are drawn, evolves across the mass bins in a similar way to the bispectra and is mainly set by how well $B^{H\alpha,U,g}_{l,p}$ can be measured.

Different from the predicted constraints on the bispectra presented in Fig. 4, which are evaluated for a single redshift interval using one pair of spectral channels of SPHEREx, we consider broader redshift bins for measuring the BCLF of $L_{H\alpha}$ and L_U from the correlation coefficients to optimize the parameter constraints. Specifically, we define three redshift bins with bin centres $z_c = 1.5$, $z_c = 2.0$, and $z_c = 2.5$, and bin edges [1.25, 1.75], [1.75, 2.25], and [2.25, 2.75], respectively. We further divide each redshift bin into $\mathcal{N} = 0.5R/(1 + z_c)$ redshift intervals with $R = 41$, which yields $\mathcal{N} = 8, 7, 6$, respectively. The uncertainties in the correlation coefficients evaluated for z_c and $\Delta z = (1 + z_c)/R$ are consequently scaled by a factor of $1/\sqrt{\mathcal{N}}$ to approximate the effect of binning together \mathcal{N} redshift intervals.

Fig. 5 shows the constraints on the cross-correlation coefficient, $r^g_{x,p}$, in each stellar mass bin predicted by Models I and II in three broad redshift bins as labelled on the vertical axis. With the help of the additional statistical power from redshift binning, we expect cross-correlating EBL maps from SPHEREx with photo- z galaxies from Rubin/LSST to distinguish Model II from Model I by detecting the decrease of $r^g_{x,p}$ towards lower stellar masses at high significance up to $z \sim 3$. It is noteworthy that even though the difference between the two toy models is modest in intermediate-mass bins, strong evidence for decorrelation may still be obtained thanks to the expected high sensitivity to the bispectra at these mass scales. Detecting such a decorrelation between H α and UV luminosities in low-mass galaxies and characterizing the mass dependence via the EBL–galaxy cross-correlation described can be a smoking gun for an elevated level of bursty star formation, although alternative explanations may exist (see discussion in Section 4). We note that, for simplicity, instead of estimating the actual galaxy counts taking into account of the mass incompleteness, we show the 90 per cent mass completeness limit in Fig. 5 and note that the constraining power in lower mass bins should therefore be taken as an upper limit due to incompleteness.

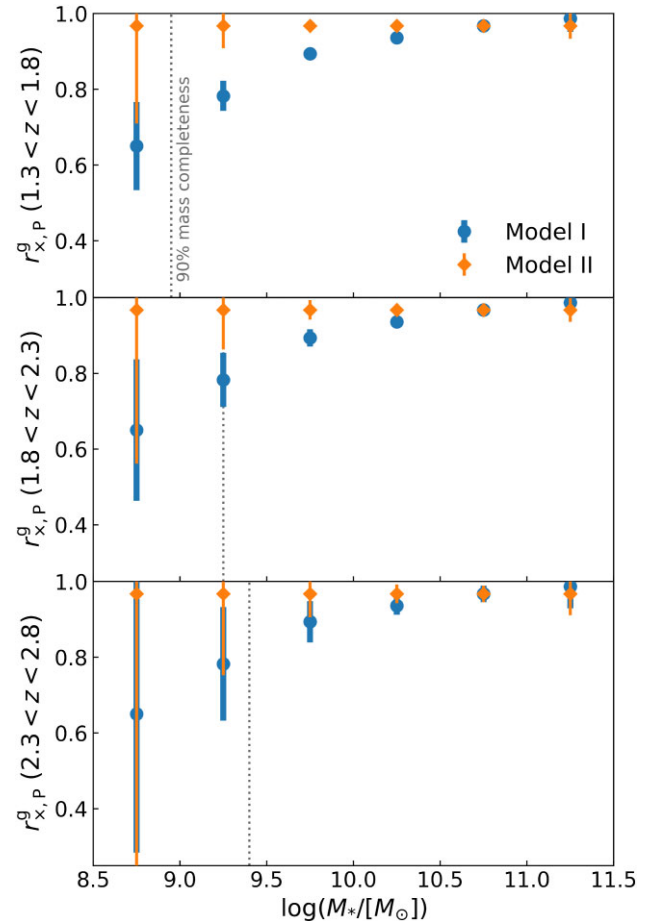


Figure 5. Capability of distinguishing Model I (blue) and Model II (orange) implied by the constraints on $r^g_{x,p}$ in individual stellar mass bins, after binning in redshift. From the top to the bottom, the three panels show the expected constraints evaluated in the three broad redshift bins, respectively. The vertical dotted lines indicate the 90 per cent mass completeness limits of the Rubin/LSST photometric galaxy redshift survey expected at these redshifts (Section 2.1.1).

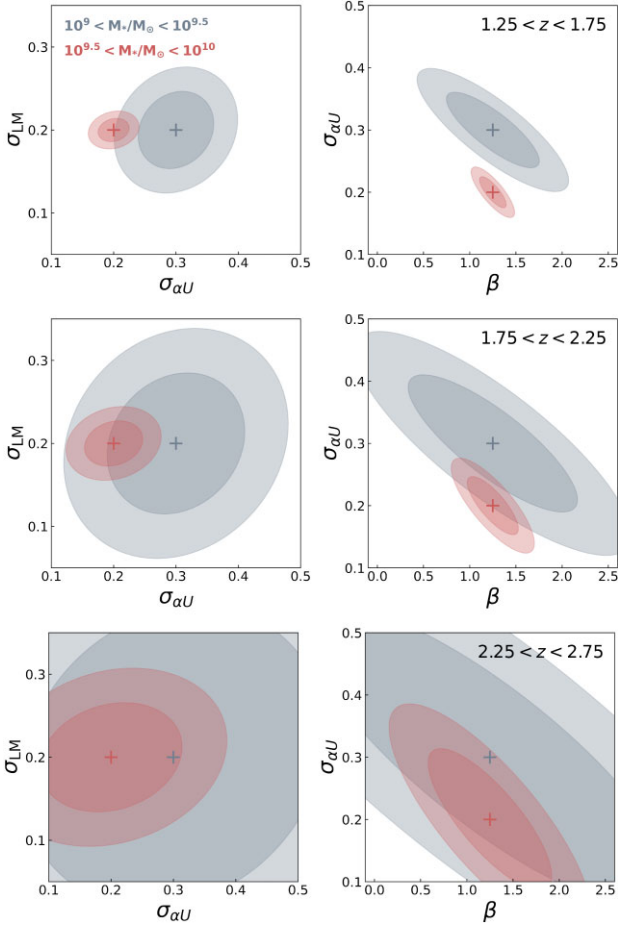


Figure 6. Constraints on the BCLF model parameters $\sigma_{\alpha U}$, σ_{LM} , and β in Model I drawn from the cross-correlation analysis for two example stellar mass bins in different colours and in the three redshift bins from top to bottom. The dark and light shaded ellipses represent the 1σ and 2σ confidence intervals, respectively.

3.2 Constraints on BCLF model parameters

From the expected constraints on $r_{\times, P}^g$ shown in Fig. 5, together with the similarly derived constraints on the autocorrelation coefficients $r_{H\alpha, P}^g$ and $r_{U, P}^g$ (see equations 11 and 12), we can directly constrain the H α –UV BCLF model assumed. To estimate the parameter constraints, we employ a Fisher matrix formalism, which performs a quadratic expansion around the loglikelihood of the data vector $\hat{\mathbf{r}}$, namely

$$F_{ij} = \sum_k \frac{1}{\text{var}(r_k)} \frac{\partial r_k}{\partial \theta_i} \frac{\partial r_k}{\partial \theta_j}, \quad (35)$$

with $\mathbf{r}(\theta) = (r_{\times, P}^g(\theta), r_{H\alpha, P}^g(\theta), r_{U, P}^g(\theta))$ being the model vector for $\theta = (\sigma_{\alpha U}, \sigma_{LM}, \beta)$. We neglect the covariance between the correlation coefficients, which is likely a reasonable approximation in the instrument-noise-dominated regime relevant to this work (Section 2.3.2), and no priors are assumed on the parameters.

The resulting constraints on the BCLF model parameters in Model I are shown in Fig. 6 for two example stellar mass bins where strong evidence for a decorrelation between H α and U-band luminosities may exist. As shown by the ellipses, the cross-correlation between SPHEREx and Rubin/LSST surveys can place useful constraints on our main proxy for bursty star formation, $\sigma_{\alpha U}$, up to $z \sim 2.5$, despite the degeneracy between $\sigma_{\alpha U}$ and β . At redshifts where the

Table 2. Fractional uncertainties in the BCLF model parameters in different redshift and mass bins derived from the diagonal of the inverse of the Fisher matrix.

Redshift	$\log(M_*/M_\odot)$	$\sigma_{\alpha U}/\sigma_{\alpha U}^{\text{fid}}$	β/β^{fid}	$\sigma_{LM}/\sigma_{LM}^{\text{fid}}$
$1.25 < z < 1.75$	9–9.5	0.133	0.265	0.153
$1.25 < z < 1.75$	9.5–10	0.082	0.076	0.046
$1.75 < z < 2.25$	9–9.5	0.242	0.482	0.280
$1.75 < z < 2.25$	9.5–10	0.154	0.145	0.090
$2.25 < z < 2.75$	9–9.5	0.532	1.046	0.595
$2.25 < z < 2.75$	9.5–10	0.375	0.352	0.216

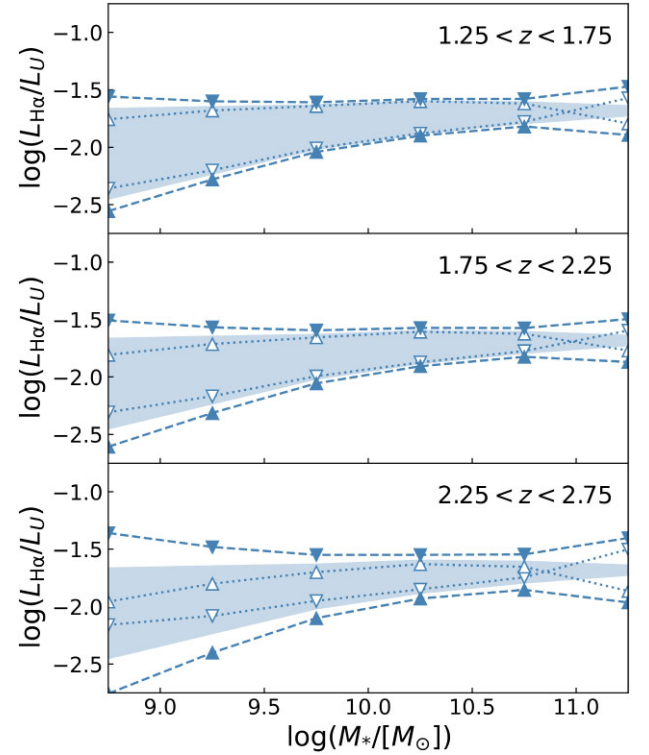


Figure 7. Constraints on $\log(L_{H\alpha}/L_U)$ as a function of M_* in the three redshift bins expected from Model I (shaded band) and the Fisher matrix analysis. Marginalized $\pm 1\sigma$ bounds on the width of the $\log(L_{H\alpha}/L_U)$ distribution are indicated by the outer, dashed curves with filled triangles (1σ upper bound) and the inner, dotted curves with empty triangles (1σ lower bound), respectively. Note that the dotted (lower-bound) curves cross at the high-mass end as a result of increased uncertainties.

constraints are tight enough (e.g. $z \sim 1.5$), it is also possible to quantify by how much $\sigma_{\alpha U}$ differs between different mass bins, which serves as another way to probe the strength of star formation burstiness (see Section 4 for further discussion). Table 2 summarizes the constraints on the three BCLF model parameters in terms of the fractional uncertainties derived from the diagonal of the inverse of the Fisher matrix in each redshift and mass bin.

Applying the Fisher matrix formalism to all mass and redshift bins and extracting the variance on $\sigma_{\alpha U}$, we derive ultimately the constraints on the $\log(L_{H\alpha}/L_U)$ – M_* relation available from the EBL–galaxy cross-correlation using SPHEREx and Rubin/LSST data, which can be readily compared with observations of individual galaxies. The resulting constraints are illustrated in Fig. 7 in terms of the upper (dashed curves and filled triangles) and lower bounds (dotted curves and empty triangles) on the width of the $\log(L_{H\alpha}/L_U)$ –

M_* distribution. From these constraints, it can be seen that any stellar mass dependence of $\sigma_{\alpha U}$ resulting from changes in the SFR variability may be tested by the cross-correlation analysis up to $z \sim 2$, beyond which data from SPHEREx and Rubin/LSST cannot provide sufficient constraining power.

4 DISCUSSION

By cross-correlating EBL and galaxy surveys to be conducted by SPHEREx and Rubin/LSST as an example, we have so far demonstrated that statistical constraints on the BCLF of $L_{H\alpha}$ and L_{UV} may be obtained at high significance and used to probe bursty star formation across a wide range of galaxy mass and redshift. Next, we supplement the presented analysis with a semiquantitative discussion of the caveats, limitations, and implications of the method explored in this work. In particular, we focus on ways to identify and reduce the potential ambiguity from dust attenuation, and compare the statistical approach presented in this paper with the characterization of SFR indicators like $L_{H\alpha}$ and L_{UV} for samples of individual galaxies.

4.1 Ambiguity associated with dust attenuation

For individual galaxies, both $L_{H\alpha}$ and L_U are subject to non-negligible dust attenuation, but the amount of attenuation can vary substantially and with different time dependence for $H\alpha$ and UV continuum emission from galaxy to galaxy, as a result of the different sites and mechanisms these photons are created in galaxies (the UV continuum can be much more extended than $H\alpha$ emission produced in star-forming regions). Consequently, part of the observed scatter $\sigma_{\alpha U}$ may actually be associated with variations of the level of dust attenuation rather than star formation burstiness for a given galaxy sample (Reddy et al. 2015). For the analysis presented, we do not consider the effect of dust on the BCLF. We do, however, take into account dust attenuation in estimating the detectability of various cross-correlation signals. While methods have been proposed to apply appropriate dust corrections for accurate comparison of $H\alpha$ and UV SFR indicators (see e.g. Weisz et al. 2012, for an example method based on energy balance), they do not directly apply to the statistical approach considered in this paper. Here, through a similar cross-correlation analysis to estimate the Balmer decrement ($L_{H\alpha}/L_{H\beta}$) variations, we discuss a possible way to reduce the ambiguity associated with unknown dust attenuation variations in the interpretation of results like those shown in Section 3.

The attenuation $A_\lambda = k_\lambda E(B - V)$ and the Balmer decrement are related by (Domínguez et al. 2013)

$$A_{FUV} = C \log \left(\frac{L_{H\alpha}/L_{H\beta}}{2.86} \right), \quad (36)$$

where the coefficient $C = 2.5k_{FUV}/(k_{H\beta} - k_{H\alpha}) = 19.6$ and $L_{H\alpha}/L_{H\beta} = 2.86$ is the intrinsic Balmer decrement that remains roughly constant for typical star-forming galaxies. Assuming perfectly correlated scatters in dust-attenuated $L_{H\alpha}$ and L_U induced by a scatter in A_{FUV} as defined in equation (5), we find that a scatter of about 4 in A_{FUV} , corresponding to a 0.2 dex scatter in the Balmer decrement $\log(L_{H\alpha}/L_{H\beta})$, results in a 0.3 dex scatter in $\log(L_{H\alpha}/L_U)$ comparable to what one might expect from a strongly time-variable SFH. Therefore, to see whether or not an observed scatter in $\log(L_{H\alpha}/L_U)$ can be explained entirely by variations in the dust attenuation, we can use the cross-correlation between $H\alpha$ and $H\beta$ to constrain the scatter σ_{BD} in $\log(L_{H\alpha}/L_{H\beta})$. Since $L_{H\alpha}$ and $L_{H\beta}$ are almost strictly proportional to each other, their cross-correlation coefficient is simply related to σ_{BD} as $\ln(r_{H\alpha \times H\beta, P}^g) = -\sigma_{BD}^2/2$,

which implies a more than 10 percent decorrelation for a Balmer decrement scatter of $\sigma_{BD} = 0.2$ dex.

At $z \sim 1.5$, for example, by performing a detectability analysis for the $H\alpha$ – $H\beta$ cross-correlation similar to that shown in Fig. 4 for the case of $H\alpha$ and U -band luminosities, we expect $r_{H\alpha \times H\beta, P}^g$ to be detected at $S/N \gtrsim 40$ (after redshift binning, see Section 3.1) by cross-correlating SPHEREx and Rubin/LSST surveys in all stellar mass bins except the least massive one, which is somewhat below the expected mass completeness limit of the Rubin/LSST galaxy survey anyway. Such a high S/N should allow us to reliably test whether or not a notable decorrelation, e.g. $r_{H\alpha \times H\beta, P}^g < 0.9$, between $L_{H\alpha}$ and $L_{H\beta}$ exists as a sign of large variations in the Balmer decrement. This can be compared in turn with level of dust attenuation variations required to fully account for the measured scatter $\sigma_{\alpha U}$ in the $H\alpha$ – UV BCLF.

Finally, it is also noteworthy that the scatter from dust attenuation variations will likely increase with stellar mass, since the massive galaxies tend to be more dust-rich. Therefore, the expected trend with stellar mass is opposite to that of the burstiness, which may also help clarify the ambiguity associated with dust attenuation.

4.2 Limitations and implications of the presented method

Despite its great potential for constraining bursty star formation using forthcoming cosmological survey data sets, the presented framework based on the EBL–galaxy cross-correlation has a few noteworthy limitations due to either simplified assumptions or the methodology itself. First, while being motivated by observations, a rather simplistic description of the BCLF is adopted in this proof-of-concept study. Potentially more self-consistent and physically grounded models can be constructed from the combination of analytical arguments and results from detailed galaxy simulations, in order to better connect burstiness observables such as $\log(L_{H\alpha}/L_U)$ to realistic representations of the time-variable SFHs. Meanwhile, in the presented analysis we have focused almost entirely on constraining the scatter in the $H\alpha$ – UV BCLF, whereas any trend between the mean value and M_* may be an additional way to probe bursty star formation.

Perhaps more importantly, we emphasize the pros and cons of the presented method when compared to observations of individual galaxies with instruments such as *JWST*, which will remain the mainstream approach for studying bursty star formation in the foreseeable future. Because of the requirement of a reference galaxy catalogue with not only sufficient depth but also a large sky coverage (nearly 2π in the presented case study), it is impossible for the cross-correlation analysis to reach a comparable mass limit to the galaxy observations with *JWST*, which can obtain mass-complete samples of galaxies with more than 100 times lower stellar masses (see e.g. Bagley et al. 2023). The most important advantage of the cross-correlation method is the access to a huge sample size of galaxies that probe a much larger cosmic volume than typical pencil-beam-like surveys with *JWST*. At $z = 2$, for example, the cumulative number densities of galaxies with stellar mass $M_* > 10^7 M_\odot$ and $M_* > 10^9 M_\odot$ only differ by a factor of 10 given the shape of the stellar mass function, whereas photometric galaxy surveys with e.g. Rubin/LSST and *Roman* typical probe cosmic volumes 1000–10 000 times larger than *JWST* programmes. The huge statistical sample of galaxies available for cross-correlation analysis makes it suitable for investigating impacts of e.g. selection bias, cosmic variance, and environmental dependence on the interpretation of bursty star formation, in addition to simply providing constraints based on large number statistics

that can be cross-checked with direct observations of individual galaxies.

Notably, the EBL–galaxy cross-correlation as presented already has some interesting implications for better understanding bursty star formation and its role in the process of galaxy formation. With constraints on the $H\alpha$ –UV BCLF in different stellar mass and redshift bins, it is possible to identify and distinguish trends predicted by rival galaxy formation and evolution theories involving different assumptions/treatments of the physics of e.g. star formation and stellar feedback. This is particularly of interest during $1 \lesssim z \lesssim 2$ when simulations predict the transition from bursty to steady star formation to happen in many Milky Way progenitor galaxies, which also correlates with the vertical disc settling process (e.g. Stern et al. 2021; Gurvich et al. 2023; Hopkins et al. 2023; Yu et al. 2023) as a key milestone in the galaxy formation history. Meanwhile, the expected constraints on the scatter of $\log(L_{H\alpha}/L_U)$, and thereby the corresponding level of SFR variability, may also shed light on the possible connection between bursty star formation and the cuspcore transformation of the dark matter halo profiles (e.g. Pontzen & Governato 2012; Teyssier et al. 2013; Chan et al. 2015; Oñorbe et al. 2015). Thanks to the huge sample size available for analyses of multiple subsamples, the EBL–galaxy cross-correlation can test whether the level of stellar feedback and SFR variability required for modulating the dark matter profiles is satisfied in different mass and redshift regimes and cosmological environments.

5 CONCLUSIONS

We have presented a simple, semi-empirical framework to study the possibility of probing bursty star formation in galaxies at $1 \lesssim z \lesssim 3$ using the cross-correlation between data sets from EBL and galaxy redshift surveys to be available in the next decade. By constructing an observationally motivated toy model for the BCLF of $H\alpha$ and U -band continuum luminosities, two commonly used SFR indicators probing the recent SFH on different time-scales, we demonstrate how useful constraints on the BCLF can be obtained from Fourier-space analysis of the EBL–galaxy cross-correlation signals in the Poisson-noise limit. Taking the synergy between SPHEREx and Rubin/LSST surveys as an example, we forecast the detectability of key observables derived from the summary statistics, in particular the correlation coefficients, and showcase the expected constraints on the parameter space from these forthcoming data sets.

Our analysis suggests that useful constraints on the mass and redshift evolution of the BCLF as a key measure of the time variability of the SFH can be placed by the EBL–galaxy cross-correlation, in a complementary manner to traditional methods based on observations of individual galaxies. A similar approach may also be applied to the same data set to investigate the potential ambiguity that can be caused by dust attenuation. In summary, the presented framework provides a novel way to probe bursty star formation and the related physics in high-redshift galaxies using cosmological data sets. Constraints from the EBL–galaxy cross-correlation will be useful additions to deep observations of individual galaxies to be conducted by e.g. *JWST*, thanks to the much greater sample size accessible.

ACKNOWLEDGEMENTS

We thank the anonymous reviewer for their comments, as well as Tzu-Ching Chang and Jordan Mirocha for helpful conversations. GS was supported by a CIERA Postdoctoral Fellowship. CAFG was supported by NSF through grant AST-2108230 and CAREER award AST-1652522, by NASA through grants 17-ATP17-0067 and

21-ATP21-0036, by STScI through grant HST-GO-16730.016-A, by CXO through grant TM2-23005X, and by the Research Corporation for Science Advancement through a Cottrell Scholar Award.

DATA AVAILABILITY

The data associated with results presented in this article can be shared upon reasonable request to the corresponding author.

REFERENCES

- Bagley M. B. et al., 2023, preprint (arXiv:2302.05466)
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, 488, 3143
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
- Caplar N., Tacchella S., 2019, *MNRAS*, 487, 3845
- Chabrier G., 2003, *PASP*, 115, 763
- Chan T. K., Kereš D., Oñorbe J., Hopkins P. F., Muratov A. L., Faucher-Giguère C. A., Quataert E., 2015, *MNRAS*, 454, 2981
- Cheng Y.-T., Bock J. J., 2022, *ApJ*, 940, 115
- Cheng Y.-T., Chang T.-C., 2022, *ApJ*, 925, 136
- Crill B. P. et al., 2020, in Lystrup M., Perrin M. D., eds, Proc. SPIE Conf. Ser. Vol. 11443, Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 114430I
- Domínguez A. et al., 2013, *ApJ*, 763, 145
- Domínguez A., Siana B., Brooks A. M., Christensen C. R., Bruzual G., Stark D. P., Alavi A., 2015, *MNRAS*, 451, 839
- Doré O. et al., 2014, preprint (arXiv:1412.4872)
- Emami N., Siana B., Weisz D. R., Johnson B. D., Ma X., El-Badry K., 2019, *ApJ*, 881, 71
- Emami N. et al., 2021, *ApJ*, 922, 217
- Faist A. L., Capak P. L., Emami N., Tacchella S., Larson K. L., 2019, *ApJ*, 884, 133
- Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, *MNRAS*, 406, 2267
- Faucher-Giguère C.-A., 2018, *MNRAS*, 473, 3717
- Finke J. D., Razzaque S., Dermer C. D., 2010, *ApJ*, 712, 238
- Finke J. D., Ajello M., Domínguez A., Desai A., Hartmann D. H., Paliya V. S., Saldana-Lopez A., 2022, *ApJ*, 941, 33
- Flores Velázquez J. A. et al., 2021, *MNRAS*, 501, 4812
- Furlanetto S. R., Mirocha J., 2022, *MNRAS*, 511, 3895
- Gong Y., Cooray A., Silva M. B., Zemcov M., Feng C., Santos M. G., Dore O., Chen X., 2017, *ApJ*, 835, 273
- Graham M. L., Connolly A. J., Ivezić Ž., Schmidt S. J., Jones R. L., Jurić M., Daniel S. F., Yoachim P., 2018, *AJ*, 155, 1
- Graham M. L. et al., 2020, *AJ*, 159, 258
- Gurvich A. B. et al., 2023, *MNRAS*, 519, 2598
- Hopkins P. F. et al., 2023, *MNRAS*
- Iyer K. G. et al., 2020, *MNRAS*, 498, 430
- Kayo I., Takada M., Jain B., 2013, *MNRAS*, 429, 344
- Korngut P. M. et al., 2018, in Lystrup M., MacEwen H. A., Fazio G. G., Batalha N., Siegler N., Tong E. C., eds, Proc. SPIE Conf. Ser. Vol. 10698, Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 106981U
- Laigle C. et al., 2016, *ApJS*, 224, 24
- Leauthaud A., Singh S., Luo Y., Ardila F., Greco J. P., Capak P., Greene J. E., Mayer L., 2020, *Phys. Dark Univ.*, 30, 100719
- LSST Science Collaboration, 2009, preprint (arXiv:0912.0201)
- McBride J., Fakhouri O., Ma C.-P., 2009, *MNRAS*, 398, 1858
- McLure R. J. et al., 2018, *MNRAS*, 476, 3991
- Mehta V. et al., 2015, *ApJ*, 811, 141
- Mirocha J., Liu A., La Plante P., 2022, *MNRAS*, 516, 4123
- More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, *MNRAS*, 392, 801
- Moutard T., Sawicki M., Arnouts S., Golob A., Coupon J., Ilbert O., Yang X., Gwyn S., 2020, *MNRAS*, 494, 1894

- Oñorbe J., Boylan-Kolchin M., Bullock J. S., Hopkins P. F., Kereš D., Faucher-Giguère C.-A., Quataert E., Murray N., 2015, *MNRAS*, 454, 2092
- Orr M. E., Fielding D. B., Hayward C. C., Burkhardt B., 2022, *ApJ*, 932, 88
- Planck Collaboration, 2016, *A&A*, 594, A13
- Pontzen A., Governato F., 2012, *MNRAS*, 421, 3464
- Reddy N. A. et al., 2015, *ApJ*, 806, 259
- Scott B. R., Upton Sanderbeck P., Bird S., 2022, *MNRAS*, 511, 5158
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Shuntov M. et al., 2022, *A&A*, 664, A61
- Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2017, *MNRAS*, 466, 88
- Speigel D. et al., 2015, preprint (arXiv:1503.03757)
- Stern J. et al., 2021, *ApJ*, 911, 88
- Sun G., 2022, *ApJ*, 931, L29
- Sun G., Furlanetto S. R., 2016, *MNRAS*, 460, 417
- Sun G., Mirocha J., Mebane R. H., Furlanetto S. R., 2021, *MNRAS*, 508, 1954
- Sun G., Faucher-Giguère C.-A., Hayward C. C., Shen X., 2023, preprint (arXiv:2305.02713)
- Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, *MNRAS*, 429, 3068
- van den Bosch F. C., Jiang F., Hearin A., Campbell D., Watson D., Padmanabhan N., 2014, *MNRAS*, 445, 1713
- Viero M. P., Sun G., Chung D. T., Monceli L., Condon S. S., 2022, *MNRAS*, 516, L30
- Weisz D. R. et al., 2012, *ApJ*, 744, 44
- Yang X., Mo H. J., van den Bosch F. C., 2003, *MNRAS*, 339, 1057
- Yu S. et al., 2023, *MNRAS*, 523, 6220
- Zheng Z. et al., 2005, *ApJ*, 633, 791
- Zhou Z. et al., 2017, *ApJ*, 835, 70

APPENDIX A: CONNECTIONS TO OTHER STATISTICS

While $r_{\times, \text{P}}^g$ itself already encodes valuable information about the burstiness of star formation, there are other observables that are also informative about the variability of the SFR. Previous studies have suggested that the full PDF, $P(L_{\text{H}\alpha}/L_{\text{U}})$, may be the most reliable probe of bursty star formation. Thus, it might also be useful to inspect basic measures of $P(L_{\text{H}\alpha}/L_{\text{U}})$ like the mean, $\langle L_{\text{H}\alpha}/L_{\text{U}} \rangle$, and the variance, $\text{Var}(L_{\text{H}\alpha}/L_{\text{U}})$. Using Taylor expansion to the second order, we can approximate them as

$$\left\langle \frac{L_{\text{H}\alpha}}{L_{\text{U}}} \right\rangle = \frac{\langle L_{\text{H}\alpha} \rangle}{\langle L_{\text{U}} \rangle} - \frac{\text{Cov}(L_{\text{H}\alpha}, L_{\text{U}})}{\langle L_{\text{U}} \rangle^2} + \frac{\text{Var}(L_{\text{U}}) \langle L_{\text{H}\alpha} \rangle}{\langle L_{\text{U}} \rangle^3} \quad (\text{A1})$$

and

$$\text{Var} \left(\frac{L_{\text{H}\alpha}}{L_{\text{U}}} \right) = \left[\frac{\text{Var}(L_{\text{H}\alpha})}{\langle L_{\text{H}\alpha} \rangle^2} - \frac{2\text{Cov}(L_{\text{H}\alpha}, L_{\text{U}})}{\langle L_{\text{H}\alpha} \rangle \langle L_{\text{U}} \rangle} + \frac{\text{Var}(L_{\text{U}})}{\langle L_{\text{U}} \rangle^2} \right] \times \frac{\langle L_{\text{H}\alpha} \rangle^2}{\langle L_{\text{U}} \rangle^2}. \quad (\text{A2})$$

By definition, the covariance and variances can be written as $\text{Cov}(L_{\text{H}\alpha}, L_{\text{U}}) = \langle L_{\text{H}\alpha} L_{\text{U}} \rangle - \langle L_{\text{H}\alpha} \rangle \langle L_{\text{U}} \rangle$, $\text{Var}(L_{\text{H}\alpha}) = \langle L_{\text{H}\alpha}^2 \rangle - \langle L_{\text{H}\alpha} \rangle^2$, and $\text{Var}(L_{\text{U}}) = \langle L_{\text{U}}^2 \rangle - \langle L_{\text{U}} \rangle^2$. As shown in Section 2.3.1, while $\langle \mathcal{O}(L) \rangle$ terms can be readily derived from cross-power spectra of intensity maps and galaxies, we must resort to the UV–UV–galaxy, H α –H α –galaxy, and H α –UV–galaxy cross-bispectra to estimate the $\langle \mathcal{O}(L^2) \rangle$ terms, as in the case of $r_{\times, \text{P}}^g$.

APPENDIX B: BIASES OF GALAXIES AND INTENSITIES OF H α AND UV EMISSION

In what follows, we specify how we estimate biases of the tracers of interest (galaxies, H α and UV emission). While these bias factors do not directly enter our main analysis, which is limited to the Poisson-noise-dominated regime, they are essential for verifying the range of valid ℓ 's where our analysis should be performed, without substantial contamination from the clustering signals. Generally, we can express the average bias factor as

$$b_i(z) = \frac{\int dM_* \psi(M_*, z) b(M_* f_{\text{SHMR}}^{-1}, z) \mathcal{W}_i(M_*, z)}{\int dM_* \psi(M_*, z) \mathcal{W}_i(M_*, z)}, \quad (\text{B1})$$

where $b(M, z)$ is the halo bias (Sheth, Mo & Tormen 2001), f_{SHMR}^{-1} is the inverse of the stellar-to-halo mass ratio (Behroozi et al. 2019), and \mathcal{W}_i is some weighting function that varies for different signals. For the bias factor of galaxies, $\langle b \rangle_g(z)$, we assume $\mathcal{W}_g(M_*, z) = \langle N_g \rangle (M_* f_{\text{SHMR}}^{-1}, z)$, where $N_g = N_{\text{cen}} + N_{\text{sat}}$ follows the halo occupation distribution parametrization in Zheng et al. (2005). To estimate the bias factors of H α and UV emission, $b_{\text{H}\alpha}(z)$ and $b_{\text{U}}(z)$, we assume $\mathcal{W}_{\text{H}\alpha}(M_*, z) = \bar{L}_{\text{H}\alpha}[\bar{L}_{\text{U}}(M_*, z)]$ and $\mathcal{W}_{\text{U}}(M_*, z) = \bar{L}_{\text{U}}(M_*, z)$, respectively.

This paper has been typeset from a \LaTeX file prepared by the author.