8-1-2022

# Light from the Darkness: Detecting Ultra-diffuse Galaxies in the Perseus Cluster through Over-densities of Globular Clusters with a Log-Gaussian Cox Process

Dayi David Li
*University of Toronto*

Gwendolyn M. Eadie
*University of Toronto*

Roberto Abraham
*University of Toronto*

Patrick E. Brown
*University of Toronto*

William E. Harris
*McMaster University, Faculty of Science*

*See next page for additional authors*

## Recommended Citation

Authors

Dayi David Li, Gwendolyn M. Eadie, Roberto Abraham, Patrick E. Brown, William E. Harris, Steven R. Janssens, Aaron J. Romanowsky, Pieter Van Dokkum, and Shany Danieli

# Light from the Darkness: Detecting Ultra-diffuse Galaxies in the Perseus Cluster through Over-densities of Globular Clusters with a Log-Gaussian Cox Process

Dayi David Li (李大一)[1] , Gwendolyn M. Eadie[1,2] , Roberto Abraham[2,3] , Patrick E. Brown[1,4] , William E. Harris[5] ,
Steven R. Janssens[2,6] , Aaron J. Romanowsky[7,8,9] , Pieter van Dokkum[10] , and Shany Danieli[11,12]

[1] Department of Statistical Sciences, University of Toronto, Ontario Power Building, 700 University Avenue, 9th Floor, Toronto, ON M5G 1Z5, Canada
dayi.li@mail.utoronto.ca
[2] David A. Dunlap Department of Astronomy and Astrophysics, University of Toronto, 50 St George Street, Toronto, ON M5S 3H4, Canada
[3] Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St George Street, Toronto, ON M5S 3H4, Canada
[4] Center for Global Health Research, St. Michael's Hospital, 30 Bond Street, Toronto, ON M5B 1W8 Canada
[5] Department of Physics and Astronomy, McMaster University, Hamilton, ON L8S 4M1, Canada
[6] Centre for Astrophysics and Supercomputing, Swinburne University, Hawthorn VIC 3122, Australia
[7] Department of Physics & Astronomy, San José State University, One Washington Square, San Jose CA 95192, USA
[8] University of California Observatories, 1156 High Street, Santa Cruz CA 95064, USA
[9] Department of Astronomy and Astrophysics, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[10] Department of Astronomy, Yale University, New Haven, CT 06511, USA
[11] Department of Astrophysical Sciences, 4 Ivy Lane, Princeton University, Princeton, NJ 08544, USA

## Abstract

We introduce a new method for detecting ultra-diffuse galaxies by searching for over-densities in intergalactic globular cluster populations. Our approach is based on an application of the log-Gaussian Cox process, which is a commonly used model in the spatial statistics literature but rarely used in astronomy. This method is applied to the globular cluster data obtained from the PIPER survey, a Hubble Space Telescope imaging program targeting the Perseus cluster. We successfully detect all confirmed ultra-diffuse galaxies with known globular cluster populations in the survey. We also identify a potential galaxy that has no detected diffuse stellar content. Preliminary analysis shows that it is unlikely to be merely an accidental clump of globular clusters or other objects. If confirmed, this system would be the first of its kind. Simulations are used to assess how the physical parameters of the globular cluster systems within ultra-diffuse galaxies affect their detectability using our method. We quantify the correlation of the detection probability with the total number of globular clusters in the galaxy and the anticorrelation with increasing half-number radius of the globular cluster system. The Sérsic index of the globular cluster distribution has little impact on detectability.

## 1. Introduction

Ultra-diffuse galaxies (UDGs) are a class of extended low-surface brightness systems first defined in van Dokkum et al. (2015) using data from the Dragonfly Telephoto Array (Abraham & van Dokkum 2014). These authors noted the existence of a large number of extended low-surface brightness galaxies in the Coma cluster, and they proposed a definition based on half-light radius ($R_e > 1.5$ kpc), and $g$-band central surface brightness ($\mu_{0,g} > 24.5$ mag arcsec$^{-2}$). Subsequent papers have identified thousands of UDGs in rich clusters (e.g., Yagi et al. 2016; Wittmann et al. 2017; Janssens et al. 2019; Lim et al. 2020) but also in smaller groups and sparser environments (e.g., Martínez-Delgado et al. 2016; Forbes et al. 2019, 2020a; Román et al. 2019; Danieli et al. 2020). This progress has led to a growing appreciation of the high abundance of UDGs, and also to the intriguing discovery that these objects exhibit a rich variety in their dark matter properties. For example, the iconic UDG Dragonfly 44 was found to be dominated by dark matter

(van Dokkum et al. 2015, 2019), while the qualitatively similar NGC 1052-DF2 shows little evidence for any dark matter (van Dokkum et al. 2018a; Danieli et al. 2019; Shen et al. 2021).

Investigations of their globular clusters (GCs) have added to the mystery. While many UDGs have normal GC populations consistent with dwarfs (Amorisco et al. 2018), several UDGs in Coma exhibit an excess in their GC populations (van Dokkum et al. 2016, 2017; Lim et al. 2018; Forbes et al. 2020b)—though this is under debate (e.g., Amorisco et al. 2018; Saifollahi et al. 2022). Several other UDGs host extremely peculiar GC systems. The dark matter free UDGs NGC 1052-DF2 and DF4 have a large population of overluminous GCs (van Dokkum et al. 2018b; Shen et al. 2021), which may be the result of these galaxies forming from a high-speed galactic collision (van Dokkum et al. 2022). Meanwhile, NGC 5846-UDG1 is the only known galaxy to date with a stellar mass fraction in GCs exceeding 10%, and this fraction could plausibly have been nearly 100% at formation—a pure GC galaxy—where the present day stellar body is composed of disrupted GCs (Danieli et al. 2022).

Some UDGs have been known for decades (e.g., Binggeli et al. 1984), but substantial recent progress has largely been driven by new approaches to detecting low-surface brightness galaxies, which have revealed the existence of huge numbers of these objects. It is by no means clear that the current detection process

[12] NASA Hubble Fellow.

has hit its limit. In fact, there are strong suggestions that selection biases continue to play an important role in defining samples of UDGs. For example, the majority of UDGs currently identified are in distant galaxy clusters, with the nearest examples being at distances around that of the Virgo cluster ($d = 16.5$ Mpc, Mei et al. 2007). Finding a very nearby UDG would be highly significant, since a dark matter-dominated UDG at $\lesssim 5$ Mpc would be an ideal testbed for probing dark matter models, most notably ultralight axionic (*fuzzy*) dark matter models (Hu et al. 2000; Walker & Peñarrubia 2011; Hui et al. 2017), which predict the existence of very compact soliton cores in massive UDGs (Wasserman et al. 2019; Burkert 2020).

UDGs have an extended, faint, and diffuse appearance. For this reason, a challenge to finding UDGs in the nearby universe is distinguishing them from imaging systematics and artifacts such as internal reflections and flat-fielding imperfections. Another issue is distinguishing them from small patches of Galactic cirrus, which are quite common in wide-field surveys (Wittmann et al. 2017; Greco et al. 2018). This challenge is most significant at distances from a few megaparsecs out to 30 Mpc or so. At larger distances, individual UDGs are small enough that they no longer resemble most types of cirrus, while at distances less than a few megaparsecs (just beyond the Local Group), the stellar content of UDGs becomes resolved in ground-based data. At such distances, resolved star counts are probably the most effective tool for finding UDGs.

There is also no physical theory requiring a UDG—or galaxy in general—to have a faint, diffuse appearance. While UDGs could represent the low-surface brightness end of the normal galaxy population, they might instead represent the *high* surface brightness end of an even more diffuse population—the most diffuse of which are composed almost entirely of dark matter except perhaps for a population of a few old stars including GCs. The total number or mass of the GCs appears to be correlated almost linearly with the dark matter content of galactic halos (e.g., Blakeslee 1997; Spitler & Forbes 2009; Harris et al. 2017; Forbes et al. 2020a; Gannon et al. 2021). Although the linear relation is not well suited for low-mass galaxies, it does not exclude the existence of such an ultralow-luminosity system. Such systems have not yet been discovered, though Peng & Lim (2016) explored the idea in the context of the Coma cluster UDG Dragonfly 17. Clearly, such *pure* dark matter halos would be nearly impossible to find using conventional approaches.

In this paper, we introduce a new method for finding GC-rich UDGs. The idea is simple—if a UDG has an associated population of GCs, these GCs should appear clustered compared to a random GC population from its surrounding intergalactic region. Therefore, if we detect a clustering signal of GCs, we can search for the presence of diffuse stellar light at the same location. A *dark* UDG or one with extremely low stellar light might be found in this way, assuming of course that such objects would have GCs present. Such a discovery would have important implications for the empirical relation between GC system mass and galaxy stellar mass (Eadie et al. 2022). If a UDG has no diffuse component and no associated population of globular clusters, it is undetectable via starlight (though other approaches, such as using gravitational lensing, or searching for diffuse line emission from cooling gas, may succeed).[13]

We describe our approach using concepts and methods from spatial statistics, using the formalism of *spatial point process*

*theory*. Several of our key ideas are widely used in other subject areas (most notably ecology, geology, and epidemiology). We will adopt the common formalism used in these subject areas. While point processes are ubiquitous in astronomy, the terminology used is specific to astronomy and unfamiliar to most working statisticians, which limits cross-talk between disciplines. In this paper we strive to map several of the key ideas of spatial statistics into an astronomical context. At the same time, we aim to remind statisticians that some key ideas in spatial statistics originated in astronomy, before evolving to become widely used in other disciplines. A good example of this is the Neyman–Scott process (Neyman & Scott 1958) originally developed to better understand the clustering of galaxies.

Our approach follows in the footsteps of other papers that apply spatial point process models to astronomy. For example, Tempel et al. (2016) developed the Bisous model for detecting galaxy filaments, and Li & Barmby (2021) employed a Gibbs point process model to study the formation of young star clusters in M33.

In the present paper, we focus on a class of point process models called the log-Gaussian Cox process (LGCP) and apply it to UDG detection within the Perseus cluster of galaxies. Interestingly, the LGCP also has its roots in astronomy, where the closely related log-Gaussian random field is used to model the distribution of matter in the universe by Coles & Jones (1991), who notably provided some examples on simulating galaxy distribution arising from an LGCP. The theoretical properties and applications of LGCP were developed and refined later in the lens of point process theory by Møller et al. (1998), and the method has since garnered a wide range of applications in ecology (Raynaud & Nunan 2014), epidemiology (Li et al. 2012), neuroscience (Samartsidis et al. 2019), environmental science (Serra et al. 2013), etc. But the use of the LGCP in astronomy seems to have remained restricted to investigations of the large-scale structure of the universe. Here, we attempt to broaden the application of the LGCP in astronomy, and describe developments in the LGCP made by statisticians working in other areas over the past two decades.

The outline of this paper is as follows. In Section 2, we describe the data used in our study. Section 3 contains a general introduction to the LGCP methodology and the associated techniques for cluster detection. Details on model construction for our specific problem are also included. Section 4 presents our results on UDG detection using the LGCP and Section 5 contains a preliminary analysis of a potential pure GC galaxy detected using our method. Section 6 describes a set of simulations used to test the performance and reliability of our method. Section 7 offers some concluding remarks and describes our plan for future research.

## 2. Data

The data used in our study come from the Program for Imaging of the PERseus cluster (PIPER; Harris et al. 2020). The images for this survey were obtained with the Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS)/WFC and WFC3/UVIS cameras, under program GO-15235. The target region is the rich Perseus cluster (A-426) at $d = 75$ Mpc (Gudehus 1995; Hudson et al. 1997). A series of fields, including many positioned through the intracluster medium (ICM) away from major galaxies, was imaged in the F475W and F814W filters for ACS, and F475X and F814W for WFC3. The data used in our study consist of 10 pairs of pointings (with the WFC3 camera as the Primary and the ACS as the Parallel camera) of these ICM regions; they

---

[13] Although one could argue if such an object can even be classified as a UDG since it is a pure dark matter halo with some gas content.

avoid any of the most giant Perseus members but include many less massive giant early-type galaxies (ETGs), numerous dwarfs, and previously known UDGs (see Figure 1 in Harris et al. 2020).

GCs in these pointings were extracted for analysis. As described in Harris et al. (2020), Perseus is distant enough that its GCs are virtually all unresolved (star-like) and so a relatively clean list of candidates can be constructed by rigorous elimination of non-stellar objects. For analysis in this paper, we select the GC candidates from this catalog of star-like objects based on the following additional criteria for magnitude and color range, which include the GC population:

$$22.0 < \mathrm{F814W}_0 < 25.5 \text{ and}$$
$$1.0 < (\mathrm{F475W} - \mathrm{F814W})_0 < 2.4,$$

where the subscript 0 denotes the de-reddened values of the magnitudes with foreground Galactic extinction removed (see Harris et al. 2020). For comparison, the turnover or peak point of the normal globular cluster luminosity function (GCLF) at the distance modulus of Perseus ($(m - M)_I = 34.62$) would be at an apparent magnitude $\mathrm{F814W}_{to} \simeq 26.2$, about 0.7 mag fainter than the imposed selection limit above (Harris et al. 2020). Our search for GC populations around UDGs is therefore biased toward using fairly luminous GCs.

Note also that the GC candidates selected will contain some fraction of contaminating objects that are either foreground stars or faint, compact background galaxies. We do not expect these to strongly influence our results as the number of contaminants should be fairly low. In the selected color and magnitude ranges used here, Harris et al. (2020) show that the contaminants will make up no more than about 20% of the total (see their Figures 7, 10, and 14 particularly and the accompanying discussion).

As a final note on the photometry, the WFC3 F475X band is slightly different from ACS F475W, so a transformation is applied to the WFC3 magnitudes to obtain the equivalent magnitude in ACS F475W (see Harris et al. 2020). All our final data are then based on the ACS (F475W, F814W) Vegamag system.

Aside from the PIPER Survey, for cross checking our detected UDGs we also utilize the low-surface brightness galaxy catalog from Wittmann et al. (2017) obtained from the Prime Focus Imaging Platform at the William Herschel Telescope through the Opticon program, as well as a more recent unpublished UDG catalog obtained by A. Romanowsky. The latter was based on archival Canada–France–Hawaii Telescope (CFHT)/MegaCam imaging of the Perseus cluster, and was constructed for the PIPER survey HST proposal. This has been superseded by much better data from Subaru Hyper Suprime-Cam (see initial results in Gannon et al. 2021). From here on, we refer to the UDGs from Wittmann et al. (2017) as WUDGs and UDGs from Romanowsky as RUDGs. The WUDG catalog contains a total of 89 UDG candidates, of which 31 fall within the PIPER survey fields. The RUDG catalog contains 148 identified UDG candidates but only 10 of them are within the PIPER fields.

## 3. Methodology: LGCP

Our goal is to introduce and apply the tools and formalism of spatial statistics, specifically LGCP, to extract information about the spatial distribution of GCs. Subsequently, we attempt to detect UDGs through the clustering signals of their associated GCs.

Before we introduce the basics of point processes, it is noteworthy that many concepts within spatial statistics are in fact already widely used in astronomy. However, due to the

isolated development between the two fields in the past decades, the terminologies of the same concepts have diverged. We could translate the terminologies in spatial statistics back to their corresponding ones in astronomy, but we believe it is best to keep the terminologies used in spatial statistics so that interested readers from astronomy can search the spatial statistics literature more easily for other novel methods. We will instead provide the corresponding astronomy concepts for terminologies in spatial statistics if we deem necessary.

We begin by specifying the basic terminology and notation for point process theory. A *random point process* is denoted by $X$, while $x$ denotes a configuration/realization of $X$. $S \subset \mathbb{R}^d$ is the region or observation window on which $X$ is defined and $s \in S$ denotes a location in the region. Here, we only focus on the process on the plane of the sky, so $d = 2$.

The most basic statistical model for point processes is the Poisson point process (PPP), which represents *complete spatial randomness*. It is solely determined by its intensity function $\lambda(s)(> 0)$, which represents the expected number of points per unit area at location $s$. In astronomy, the intensity function is often the mean surface number density or the mean point count per unit area. We include a formal definition of PPP and some of its defining properties in Appendix A.

Depending on the intensity function, a PPP can be classified as a *homogeneous* Poisson process (HPP) or an *inhomogeneous* Poisson process (IPP). An HPP has a constant intensity function everywhere, i.e., $\lambda(s) \equiv \lambda > 0$. In astronomy, an HPP is essentially a point pattern with points drawn from a uniform distribution. An IPP, on the other hand, has an intensity function that depends on the location $s$: the mean surface number density is different at different locations, but the occurrence of a point is independent of any other point.

Building on a PPP, a more flexible class of models is called log-Gaussian Cox process (LGCP; Møller et al. 1998). An LGCP is a type of doubly-stochastic (Cox) process that possesses a hierarchical structure. On the first (progenitor) level of the hierarchy, we have a log-Gaussian process (log-Gaussian random field), which generates an intensity function, while at the second level, the point process is assumed to be an IPP with intensity function being the one previously generated by the log-Gaussian process.

An example of the LGCP can be illustrated through the galaxy distribution discussed in Coles & Jones (1991), where at the first level of the hierarchy, the large-scale matter (baryonic and dark matter) distribution in the universe is modeled as a log-Gaussian random field. After a realization of such a random field is obtained, the second level of the hierarchy is then an IPP with intensity function being the realized log-Gaussian random field and locations of galaxies will then *spawn* according to the IPP. With a concrete depiction of the LGCP in mind, we now introduce the mathematical formalism of the LGCP.

### 3.1. Model Specification

We start from an IPP $X$ with intensity function $\lambda(s) \geqslant 0$, where $s \in S \subset \mathbb{R}^2$. Let $x = \{x_1, ..., x_N\} \subset S$ denote a realization of $X$, and $N$ denote the number of points in $x$. Then the probability density function of $X$ w.r.t. a unit rate HPP, i.e., $\lambda_{\mathrm{unit}}(s) \equiv 1$, is

$$\pi(x | \lambda(s)) = \exp\left(\int_S 1 - \lambda(s)\, ds\right) \prod_{i=1}^N \lambda(x_i). \tag{1}$$

If we further assume that the intensity function $\lambda(s)$ arises from a Gaussian process (GP) $\mathcal{U}(s)$ such that $\lambda(s) = \exp(\mathcal{U}(s))$, then $X$ is called an LGCP.

A GP or a Gaussian random field $\mathcal{U}(s)$ is a collection of random variables indexed by time or space such that any finite collection of these random variables is multivariate normally distributed. A GP is uniquely determined by its mean function $\mu(s)$ and its covariance function $\mathrm{Cov}(\mathcal{U}(s), \mathcal{U}(t))$. If we further assume the process is isotropic and stationary, then the covariance function can be written as $\mathrm{Cov}(\mathcal{U}(s), \mathcal{U}(t)) = \sigma^2 C(|s - t|)$, where $\sigma^2$ is the marginal variance and $C(|s - t|)$ is the correlation between $\mathcal{U}(s)$ and $\mathcal{U}(t)$ as a function of the Euclidean distance $|s - t|$.

For a more succinct representation, the model can be written as

$$X \sim \mathrm{IPP}(\lambda(s)),$$
$$\log(\lambda(s)) = X(s)'\beta + \mathcal{U}(s),$$
$$\mathcal{U}(s) \sim \mathcal{GP}(\mathbf{0}, \sigma^2 C(\cdot)),$$
$$\mathrm{Cov}(\mathcal{U}(s), \mathcal{U}(t)) = \sigma^2 C(|s - t|),\ s, t \in S. \quad (2)$$

Note that $X(s)$ is a row vector for some underlying covariates at location $s$, whose value is available at any location in the observation window. $\beta$ is the vector of covariates' coefficients. $X(s)'\beta$ is the fixed effect that captures the large-scale spatial variation while $\mathcal{U}(s)$ is set to a zero mean GP to capture the spatial random (residual) effect.

For the correlation function, the most widely used is the Matérn correlation function:

$$\mathcal{M}_\nu(|s - t|) = \frac{|s - t|^\nu}{\Gamma(\nu)2^{\nu-1}h^\nu} K_\nu\left(\frac{|s - t|}{h}\right), \quad (3)$$

where $h > 0$ is a range parameter that controls the scale of the correlation, $\nu > 0$ is the shape parameter that controls the smoothness of the GP, and $K_\nu$ is the modified Bessel function of the second kind. In most practical applications, $\nu$ is usually assumed to be known and excluded from inference. The parameters being inferred in the model are then $\beta$, $\sigma$, and $h$.

To facilitate inference, a traditional approach[14] is to divide the spatial domain into an $n \times m$ grid with equally sized cells, and then assume the intensity within each cell is constant. After the grid is constructed, the log-intensity within the $i$th cell $y(c_i) = y_i$ is constant and indexed by the centroid $c_i$, $i \in \{1,...,nm\}$. Denoting $\mathbf{y} = (y_1,...,y_{nm})'$, by the definition of a GP, $\mathbf{y}$ then follows a multivariate normal distribution $\mathbf{y} \sim \mathcal{N}(X(s)'\beta, \sigma^2 C_{nm})$, where $C_{nm}$ is the $nm \times nm$ correlation matrix with entry $C(|c_i - c_j|)$. Let $\theta = (\beta, \sigma^2, h)$, $n_i$ the number of points in the $i$th cell, and let $A$ denote the area of each individual cell. Under a Bayesian setting, the joint log-posterior density given the discretization rule is then

$$\log \pi(\theta, \mathbf{y}|\mathbf{x}) = \sum_i [y_i n_i - A \exp(y_i)]$$
$$- \frac{1}{2\sigma^2}(\mathbf{y} - X'\beta)' C_{nm}^{-1}(\mathbf{y} - X'\beta)$$
$$- \frac{nm}{2}\log(\sigma^2) - \frac{1}{2}\log(\det(C_{nm})) + \log \pi(\theta)$$
$$(4)$$

where $\pi(\theta)$ is the prior distribution of the parameters. If we are interested in inferring the parameters $\theta$, then we will focus on the posterior marginal distribution $\pi(\theta|\mathbf{x})$. If we are interested in the latent intensity function, then we will look at the posterior marginal distribution $\pi(\mathbf{y}|\mathbf{x})$ (i.e., the posterior predictive distribution).

For our purposes, GCs are divided into the following categories: (a) GCs within bright normal galaxies (including normal dwarfs), (b) GCs in the intergalactic medium (IGM), and (c) GCs within UDGs. In our modeling assumption, the spatial residual effect $\mathcal{U}(s)$ contains the latter two categories of GCs. Therefore, by studying the posterior predictive distribution of the intensity surface of the residuals $f(\mathcal{U}|\mathbf{x})$, we can identify over-dense regions of GCs with some user-defined level of confidence, which in turn provides us with positions of UDG candidates.

One could argue that the LGCP method is essentially an over-elaborated variant of the kernel intensity estimator (KIE; Diggle 1985). However, LGCP has distinct advantages over KIE for our problem such as theoretical guarantee of performance as well as a full probabilistic prediction of the latent intensity function instead of the point-wise prediction provided by KIE (Diggle et al. 2013).

Most importantly, KIE cannot incorporate covariate information. In our problem, a clear covariate effect is the presence of various bright, normal galaxies, where 11 out of the 20 pointings contain at least one of them. Since GCs are generally strongly clustered and abundant in normal galaxies, the clustering signals of GCs from these galaxies will dominate or contaminate the signals from smaller GC systems of UDGs, making UDGs undetectable. Using KIE cannot deal with the confounding signals from these normal galaxies. But for LGCP, we can easily construct covariate models, e.g., Sérsic models (Sérsic 1963), to eliminate the effect of GC systems from normal galaxies. Afterward, the signals from UDGs are then captured in the residual effect.

For inference, LGCP can be fitted using traditional methods such as Markov chain Monte-Carlo (MCMC). However, due to the high-dimensional nature of the underlying GP, inference through MCMC is computationally prohibitive. Here, we adopt the state-of-the-art method called the integrated nested Laplace approximation (INLA; Rue et al. 2009) to conduct Bayesian inference. INLA is an alternative method to MCMC that directly approximates the posterior marginal distribution using Laplace approximation, which is orders of magnitude faster. Note that in most cases, we are only interested in the marginal distribution of the posterior instead of the usual high-dimensional full posterior. For a detailed overview of INLA, readers can refer to Rue et al. (2009).

The main computational tool for model inference will be the R-INLA package as well as inlabru (Bachl et al. 2019) in R (R Core Team 2021), where inlabru is a much more integrated and streamlined version of R-INLA. In the spirit of reproducible research, we need to note that due to the automatic parallelization employed in the R-INLA package, different machines will produce slightly different results as the initialization values for inference are machine-dependent and out of user control. For the record, the results for the UDG detection in the Perseus cluster presented in Section 4 are obtained on a 64 bit Windows 10 desktop with an Intel i7-6800 K 3.40 GHz CPU. For the simulation analysis done in Section 6, they are carried out using a high performance

---

[14] Note that this is not the inference method we use, but due to its simplistic nature, we describe it here for illustrative purposes.

computing cluster with a Linux CentOS 7 operating system and an Intel Xeon E5-2683 v4 2.1 GHz CPU. Although the results from R-INLA can differ for different machines, in general the differences are insignificant unless the data are ill-conditioned, which is not of concern for our problem.

Aside from employing INLA, we also utilize another state-of-the-art method for fitting LGCP models, namely approximating the latent GP as the solution to a stochastic partial differential equation (SPDE; Lindgren et al. 2011; Simpson et al. 2015). Due to the highly technical nature of the method, we do not over-elaborate the details here. It suffices to know that the SPDE approach is used for obtaining fast and accurate approximation of the unknown integral in Equation (1) and it is computationally superior to the aforementioned method where the spatial domain is divided into equally spaced cells. The solution of the SPDE is instead a continuous representation of the latent GP and it requires sufficiently fewer *cells*[15] to achieve a similar accuracy compared to when equally spaced cells are used, thereby drastically reducing the computational time. The SPDE approach is also implemented in the R-INLA and inlabru.

Note that there exists other point process models aside from LGCP that could potentially be suitable for our problem, e.g., Poisson cluster process (Neyman & Scott 1958; Møller 2003). LGCP is chosen here among others due to computational reasons. In general, fitting point process models is difficult since the likelihood of any point process model requires integrating the unknown intensity function in Equation (1). The SPDE approach combined with LGCP can render the computation of this integral extremely fast. Moreover, within the context of point process models, LGCP is the only model framework that can enjoy the benefit of INLA as LGCP belongs to the class of latent Gaussian model (Rue et al. 2009) for which INLA is specially designed. For our data set, LGCP combined with INLA only requires a maximum of 10 minutes to run for one point pattern. For other point process models, however, the inference has to rely on MCMC and the computation can take days or hours. Since we also need to conduct a large-scale simulation in Section 6, LGCP is the only feasible option.

### 3.2. Detection through Excursion Sets and Excursion Function

To detect UDGs using LGCP, we require a further method proposed by Bolin & Lindgren (2015) to identify the so-called *excursion sets* of a stochastic process, i.e., regions in the posterior intensity surface that exceed some threshold values with high probability. These excursion sets will then provide us with potential locations of UDGs.

Formally, an excursion set with threshold $C$ and confidence level $1 - \alpha$ is given as follows:

$$E_{C,\alpha}(\mathcal{U}) = \arg\max_{D}\{|D|: P(D \subset A_C(\mathcal{U})|\boldsymbol{x}) \geqslant 1 - \alpha\}, \quad (5)$$

where

$$A_C(\mathcal{U}) = \{s: \mathcal{U}(s) > \log(C)\}. \quad (6)$$

The interpretation is that $E_{C,\alpha}(\mathcal{U})$ is the set where *all the points* within it exceed level $\log(C)$ with probability of at least $1 - \alpha$.

---

[15] *Cell* here is only a nomenclature, where in fact SPDE does not construct cells but a Delaunay triangulation of the observation window based on the point pattern. The latent GP and its associated integral are then approximated based on the vertices of the triangulation using basis-function expansion.

Note that a smaller value of $\alpha$ would indicate a higher significance level. As noted by Bolin & Lindgren (2015), however, $E_{C,\alpha}(\mathcal{U})$ only provides us with regions that surpass $C$ with probability $1 - \alpha$. There can be many more hidden structures within $E_{C,\alpha}(\mathcal{U})$ that are left out. For example, we can set $\alpha = 0.5$, but $E_{C,0.5}(\mathcal{U})$ can potentially contain regions such as $E_{C,0.05}(\mathcal{U})$. Only reporting $E_{C,0.5}(\mathcal{U})$ will not provide us information about these hidden regions. Therefore, the goal here is to obtain a description of the excursion set by something similar to that of *p*-value for each location. To address this issue, Bolin & Lindgren (2015) proposed the following excursion function:

$$F_C(s) = \sup\{1 - \alpha: s \in E_{C,\alpha}(\mathcal{U})\}. \quad (7)$$

The excursion function takes a value between zero and one. A value close to zero at a given location $s$ indicates that $\mathcal{U}(s)$ will only exceed $\log(C)$ for large values of $\alpha$, hence the process is unlikely to exceed $\log(C)$ at $s$. On the other hand, a value close to 1 indicates that the process is very likely to exceed $\log(C)$ at $s$. We will not elaborate on the computational details of excursion sets and functions here, and instead refer the reader to Bolin & Lindgren (2015). The actual computations of excursion sets and excursion functions are facilitated using the excursion package (Bolin & Lindgren 2018) in R.

The last part of our methodology that needs to be addressed is the excursion threshold $C$. Intuitively, we would expect that regions where $\exp(\mathcal{U})$ exceeds high values of $C$ with high probability are likely to indicate strong clustering signals and hence potential locations of UDGs. However, the notion of high excursion threshold is not well defined here. In contrast, in many disciplines such as environmental science and climate science, the threshold $C$ is often predetermined. For example, one might be interested in regions where an air pollutant exceeds a certain preset threshold, or regions where the $CO_2$ level exceeds, say, 300 PPM. These thresholds are generally determined before any study is carried out, and they possess clear, problem-specific motivation. But this is not the case for our problem since what is considered a strong clustering signal of GCs will vary for different GC point patterns. Thus, we utilize the quantile, $Q(p)$, of the posterior marginal distribution of $\mathcal{U}(s)$ as initially suggested by Diggle et al. (2005). Using the quantiles will provide us with a principled procedure for computing excursion functions that are comparable across different GC point patterns. However, the original suggestion by Diggle et al. (2005) to compute the quantiles is rather simplistic and not suited for a Bayesian setting. The mathematical details for obtaining the posterior marginal quantiles are given in Appendix B. For our problem, we consider $p = 0.5$, 0.75, 0.9, and 0.95 to compute the excursion functions, i.e., we consider $\log(C)$ to be the median, 75th, 90th, and 95th quantiles of the posterior marginal distribution of $\mathcal{U}(s)$.

### 3.3. LGCP Model Form for UDG Detection

Here, we describe our LGCP models for detecting UDGs in the Perseus cluster. We denote the point pattern of GCs in each pointing by $X$ and the field of view of each pointing by $S$, and we

specify the following model structure based on Equation (2):

$$X \sim \text{IPP}(\lambda(s)),$$

$$\log(\lambda(s)) = \beta_0 + \sum_{k=1}^{N_G} \beta_k g(s; R_k, \alpha_k) 1_{\{N_G \neq 0\}} + \mathcal{U}(s),$$

$$\mathcal{U} \sim \mathcal{GP}(\mathbf{0}, \sigma^2 \mathcal{M}_1(\cdot; h)),$$

$$\text{Cov}(\mathcal{U}(s), \mathcal{U}(t)) = \mathcal{M}_1(|s-t|; h), s, t \in S. \quad (8)$$

The model in Equation (8) consists of two main components. First, the fixed effects comprise the intercept term $\beta_0$ to model the baseline mean intensity, and the summation term to model the covariate effects from normal galaxies within the pointing. Specifically, $N_G$ represents the number of normal galaxies in the pointing. Note that these galaxies do not need to be completely within the pointing. If a pointing includes part of these galaxies, they will be included. The coefficients $\beta_k$ represent the log-relative central surface intensity of GCs within these galaxies. $1_{\{\}}$ denotes the indicator function and if there is no normal galaxy in the pointing, then the fixed effect only contains the intercept term.

The function $g$ in Equation (8) follows the functional form of the Sérsic model:

$$g(s; R_k, \alpha_k) = \exp\left(-\left(\frac{d(s, c_k)}{R_k}\right)^{\alpha_k}\right), \quad (9)$$

where $c_k$ is the center of the $k$th galaxy, which is known, and $d(s, c_k)$ is the distance from $s$ to $c_k$. $R_k$ is the scale parameter for the GC system of the $k$th galaxy and $\alpha_k$ is the inverse of the Sérsic index. For simplicity, we only follow the functional form of the Sérsic model instead of adhering to the traditional parameterization where the half-number radius and the Sérsic index are used. Moreover, we found that our parameterization and the traditional parameterization generally lead to similar results but the traditional parameterization of the Sérsic model led to higher computational cost (20% more computational time in some cases) and sometimes numerical instability.

On another note, the model in Equation (9) is a simplified version of what we actually employ—namely, we have dropped the information on position angle and ellipticity of the galaxies. However, since these two parameters can be easily measured with high accuracy, we treat them as known and do not explicitly include them in Equation (9). The ellipticities, position angles, as well as the half-light radii of normal galaxies are obtained using the same archival CFHT/MegaCam imaging of the entire Perseus cluster for the RUDG catalog. These quantities are obtained using SExtractor (Bertin & Arnouts 1996).

We need to point out that the position angle and the ellipticity can have a significant impact on the distribution of GC systems depending on the galaxy type. Although the exact impact varies from one galaxy to another, the general understanding is that the GC systems in ETGs tend to follow the galaxy isophotes, while GC systems are spherically distributed for disk type galaxies (Harris 1991; Wang et al. 2013). Thus, for simplicity, we adopt the previous observation by Harris (1991) and Wang et al. (2013) when constructing the GC profile structure of galaxies.

The second component of the model in Equation (8) is the residual effect modeled by the latent GP. Here, we adopt a Matérn correlation function for the latent GP with the shape parameter $\nu = 1$. The choice of $\nu$ here is rather arbitrary

(although under the SPDE approach, $\nu$ is required to be less than one for technical reasons), and we found that different values of $\nu$ do not seem to have any significant impact on the results.

It needs to be mentioned that there exists another model formulation instead of Equation (8): the GC point pattern can be regarded as the superposition of GCs from normal galaxies and GCs from the IGM and UDGs. This model formulation is probably more natural and used more often in an astrophysical context, e.g., when calculating background-subtracted GC counts in galaxies. Under such a formulation, the intensity function can be written as

$$\lambda(s) = \sum_{k=1}^{N_G} \beta_k g(s; R_k, \alpha_k) 1_{\{N_G \neq 0\}} + \beta_0 \exp(\mathcal{U}(s)), \quad (10)$$

i.e., the intensity function of all GCs is the sum of intensity functions of GCs in all normal galaxies and the intensity function of GCs in IGM and UDGs. However, inference is now impossible with INLA as the model is no longer an LGCP. The best available solution is to incorporate INLA within MCMC combined with data augmentation (Gómez-Rubio & Rue 2017), which is much more difficult to implement and extremely computationally expensive[16] without significant gains in UDG detection performance. The difference between the two model structures lies in the physical interpretation of the data generating process of the GC point patterns: while both interpretations are acceptable, they are irrelevant for the purpose of our problem. Thus, Equation (8) is the one we adopt for the sake of feasibility.

### 3.4. Prior Specification

After model construction, we specify the prior distributions for our models. We denote all parameters by $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathcal{R}, \sigma, h)$ where $\boldsymbol{\beta} = (\beta_0, ..., \beta_{N_G})$, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_{N_G})$, and $\mathcal{R} = (R_1, ..., R_{N_G})$. For the prior of $\boldsymbol{\beta}$, we assign a $(N_G + 1)$-dimensional standard normal distribution, i.e.,

$$\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, 1000 \times \boldsymbol{I}_{N_G+1}),$$

where $\boldsymbol{I}_{N_G+1}$ is the $(N_G + 1)$-dimensional identity matrix. For the prior distribution of the Sérsic indices $\boldsymbol{\alpha}$, since they are required to be positive, we assign an exponential distribution with rate parameter one to each individual component $\alpha_k$ of $\boldsymbol{\alpha}$ and assume that the individual components are independent. Note that the prior distributions chosen here are quite uninformative as we do not have a very good idea about the values of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$.

For the scale parameters $\mathcal{R}$, we construct the prior distribution for individual components $R_k$ loosely based on the results from Forbes & Forbes (2017). The author studied the scale ratio between the half-light radii and the half-number radii of GC systems for a sample of ETGs. It was found that on average, the half-number radii of the GC system of a galaxy is roughly four times its half-light radii, and there exists a positive linear relationship between the stellar mass and the scale ratio, but the amount of scatter is quite significant. It was also noted that for low-mass ETGs, the scale ratio is roughly three times while for more massive galaxies, the scale ratio increases to an average of five times.

---

[16] A mere 5000 iterations of INLA within MCMC will take roughly at least 17 hr to run for one point pattern in our data. It only takes ~10 minutes under Equation (8).

Due to the significant scatter in the scale ratio, we obtain a crude guess of $R_k$ by eye based on the galaxy's light profile and GC system with the observation of Forbes & Forbes (2017) as a rough guide. The contaminating normal galaxies in the PIPER survey can be divided into two distinctive subpopulations of giant ETGs and dwarf ellipticals (dEs). Thus, for giant ETGs, we construct a prior distribution for $R_k$ where it follows a log-normal distribution with a mean equal to five times the galaxy's half-light radius with a standard deviation equal to the half-light radius. The choice of a log-normal distribution is to ensure that $R_k$'s are positive. For dEs, we supply the prior distribution of $R_k$ with the same log-normal distribution but with the mean value being three times the half-light radius. For completeness, the prior distributions for $R_k$'s of all normal galaxies are provided in Table 4 in Appendix D.

We also need to mention that $R_k$ is not the half-number radius of the GC system, as the parameterization in our modeling is different from the standard Sérsic model. However, we found that the Sérsic indices of the GC systems for most of the galaxies in our data are relatively low ($0.5 \lesssim n \lesssim 1$). For a Sérsic index in the range of $0.5 \lesssim n \lesssim 1$, it can be shown that $R_k$ and the half-number radius differ less than a factor of 2. Hence, to reduce the complexity of our model, we assume here that $R_k$ and the half-number radius are the same.

Lastly, for the hyperparameters $\sigma$ and $h$, we assign the penalized complexity (PC) priors based on the results from Simpson et al. (2017) and Fuglstad et al. (2018). The PC priors are currently the standard prior choice for modeling the latent GP within an LGCP. As the name suggests, PC priors penalize the complexity of the fitted model so that the resulting model possesses more predictive power. It does so by forcing the range parameter $h$ to be as large as possible while pushing the marginal standard deviation $\sigma$ to be as small as possible. Such a scenario corresponds to a Gaussian random field that is inherently smooth with almost no local fluctuation. Therefore, the model will not be heavily influenced by local spurious noise structures. Furthermore, PC priors are scale invariant, meaning that changing the physical unit of distance measurement does not change the underlying model structures, which is highly desirable. Specifically, the hyperparameters are first reparameterized, where

$$h = \rho / \sqrt{8\nu},$$

Here, $\rho$ is the distance at which the correlation function for the GP is roughly 0.1 as per convention by Lindgren et al. (2011). Then the PC priors for $\sigma$ and $\rho$ are specified jointly as

$$f(\sigma, \rho) = \lambda_1 \lambda_2 \rho^{-2} \exp(-\lambda_1 \rho^{-1} - \lambda_2 \sigma), \quad \sigma, \rho > 0. \quad (11)$$

Furthermore, the parameters for the PC prior are specified implicitly by the constraints where $P(\rho < \rho_0) = a_1$ and $P(\sigma > \sigma_0) = a_2$ with $\lambda_1$ and $\lambda_2$ satisfying

$$\lambda_1 = -\log(a_1)\rho_0,$$

and

$$\lambda_2 = -\frac{\log(a_2)}{\sigma_0}.$$

The parameters $\rho_0$, $\sigma_0$, $a_1$, and $a_2$ are supplied by users.

For the prior of the range parameter $\rho$, we specify the prior parameters based on the sizes of identified UDGs within the Coma cluster obtained by van Dokkum et al. (2015). The median of the half-light radius of the 47 UDGs they discovered

within the Coma cluster is roughly 3 kpc. Forbes & Forbes (2017) also studied the half-number radii of the GC system of UDGs and found that the sizes of GC systems are roughly over twice the half-light radii, although their sample size is rather small (only three). Based on this information, since $\rho$ represents the distance at which the latent GP has a correlation of less than 0.1, we assign $\rho_0 \approx 7$ kpc and $a_1 = 0.5$. The motivation for the choice of these parameters is that the majority of the correlation of the latent GP in the model should be caused by the presence of UDGs; at distances greater than the half-number radius of the GC systems of UDGs ($\sim$7 kpc), the correlation of the latent GP governing the GC point patterns should be rather small. Hence, $\rho_0$ is set to $\sim$7 kpc while this value is the median of the distribution.

As for the variance parameter $\sigma$, there is not much prior information available to us. Hence, we set $\sigma_0 = 1.5$ and $a_2 = 0.5$ to represent a rather uninformative prior.

## 4. Results

### 4.1. Detection of Known UDGs

We now present the results for UDG detection in the Perseus cluster using our LGCP model and GC candidates from the PIPER survey data. To validate our results, we compare our detections to the WUDG (Wittmann et al. 2017) and RUDG catalogs. Recall that 31 of the WUDGs and 10 of the RUDGs are within the PIPER fields. We note that some of these UDGs contain bright nuclei and the nuclei are included in the GC data that we use for detection.

For illustrative purposes, we present the results obtained from our LGCP model for one of the pointings, V11-ACS, in Figure 1. The figures for all other detected UDGs are presented in Figures 11–17 in Appendix C. V11-ACS is the most contaminated pointing of all; a large lenticular (S0) galaxy (PGC 012437) takes up a significant portion of the pointing and another elliptical galaxy (UGC 02673) just outside the pointing has a GC system that clearly extends into the field of view. In Figure 1, previously detected UDGs from other catalogs are denoted by black crosses. Figure 1(a) shows the posterior mean intensity surface $\lambda(s)$ obtained from LGCP. Figure 1(b) presents the posterior spatial random effect $\mathcal{U}(s)$, shown on the exponential scale. Figure 1(c) presents the excursion function of $\mathcal{U}(s)$ with the excursion threshold being the median of the posterior marginal of $\mathcal{U}(s)$ or $C = 1$. For simplicity and to save space, we only show the results for the median here and omit the results for other quantile excursion thresholds.

We can see in Figure 1(a) the intensity is dominated by the elliptical galaxy at the left-center region of the pointing. The high intensity region near the top is the other elliptical galaxy that falls partially in the field. It is clear from Figure 1(a) that the signals from known UDGs are drowned out, and not picked up by $\lambda(s)$. However, if we strip away the fixed effects from the elliptical galaxies, the spatial random effect $\mathcal{U}(s)$ detects the UDG signals (Figure 1(b)). In fact, the locations of the two identified and previously known UDGs have the highest intensity in $\mathcal{U}(s)$. The excursion function in Figure 1(c) successfully picks out the two UDGs with almost no false-positive noise. The selection of UDG candidates is then based on the results as demonstrated in Figure 1(c), where regions with the highest excursion function values are chosen as potential locations for UDGs. Our results demonstrate the extraordinary power of LGCP for detecting UDGs.
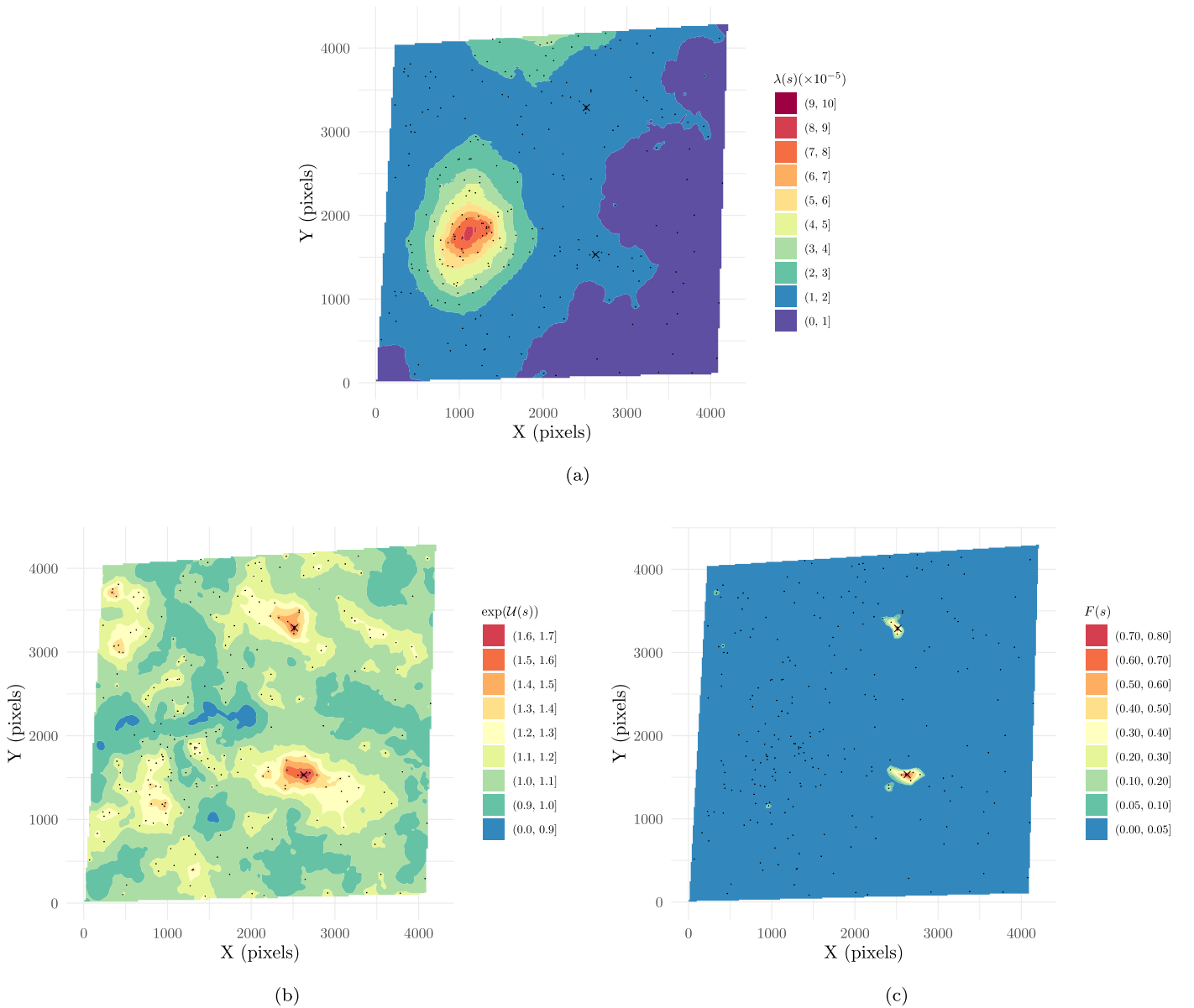
(a)



(b)



(c)

**Figure 1.** Posterior results for V11-ACS pointing. Black points are locations of GC candidates. Black crosses are locations of previously detected UDG candidates (WUDG-89 at top, WUDG-88 at bottom). North and east are approximately downward and rightward, respectively. (a) Posterior mean intensity $\lambda(s)$. The signals from UDGs are dominated by large GC systems associated with two giant ETGs (PGC 012437 at left center, and UGC 02673 just beyond the top of the field of view). (b) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$, shown as $\exp(\mathcal{U}(s))$. After eliminating the signals from the GC systems from the two ETGs, the signals of two UDGs clearly show up within the spatial random effect and they are the strongest signals in the field of view. (c) The excursion function, $F_C(s)$, of $\mathcal{U}(s)$, with $C = 1$, i.e., the median of the posterior marginal of $\mathcal{U}(s)$. The excursion function picks out the two UDGs in the field with almost no false-positive detection.

Based on the excursion functions obtained for all 20 PIPER survey pointings, LGCP picked up 11 UDG candidates within nine different pointings, which are shown in Figure 2. Among these, 10 UDGs were already reported in the WUDG and RUDG catalogs, with four in the WUDG catalog and six in the RUDG catalog. Five of these UDGs have cross-IDs from the PCC catalog (Wittmann et al. 2019): PCC 2251 (WUDG-33), PCC 3023 (WUDG-56), PCC 5374 (WUDG-88), PCC 5402 (WUDG-89), and PCC 4867 (RUDG-27). It turns out that these 10 UDGs are all the previously confirmed UDGs within the PIPER survey that have observed GC counts greater than two. For other confirmed UDGs, their observed GC counts are all fewer than two and thus undetectable.

Note that the observed GC count we consider here is the observed number of GCs within three times the half-number

radius of the UDG GC system where the half-number radii are from S. R. Janssens et al. (2022, in preparation). This GC count is our observable: it contains the majority of GCs in UDGs that allow for detection, but it also contains background contamination. GCs outside this region will not be clustered enough to have a significant effect on detection. Although one would be more interested in knowing how the background-subtracted GC count affects the UDG detection, its true value is never known in real data. Furthermore, it is traditionally obtained by fitting a Sérsic model to the GC system with background contamination subtracted, such an estimation can be subject to a high level of uncertainty especially for systems with small GC populations.

For an independent comparison based on GC counts, we also consider the traditional background-subtracted GC counts for
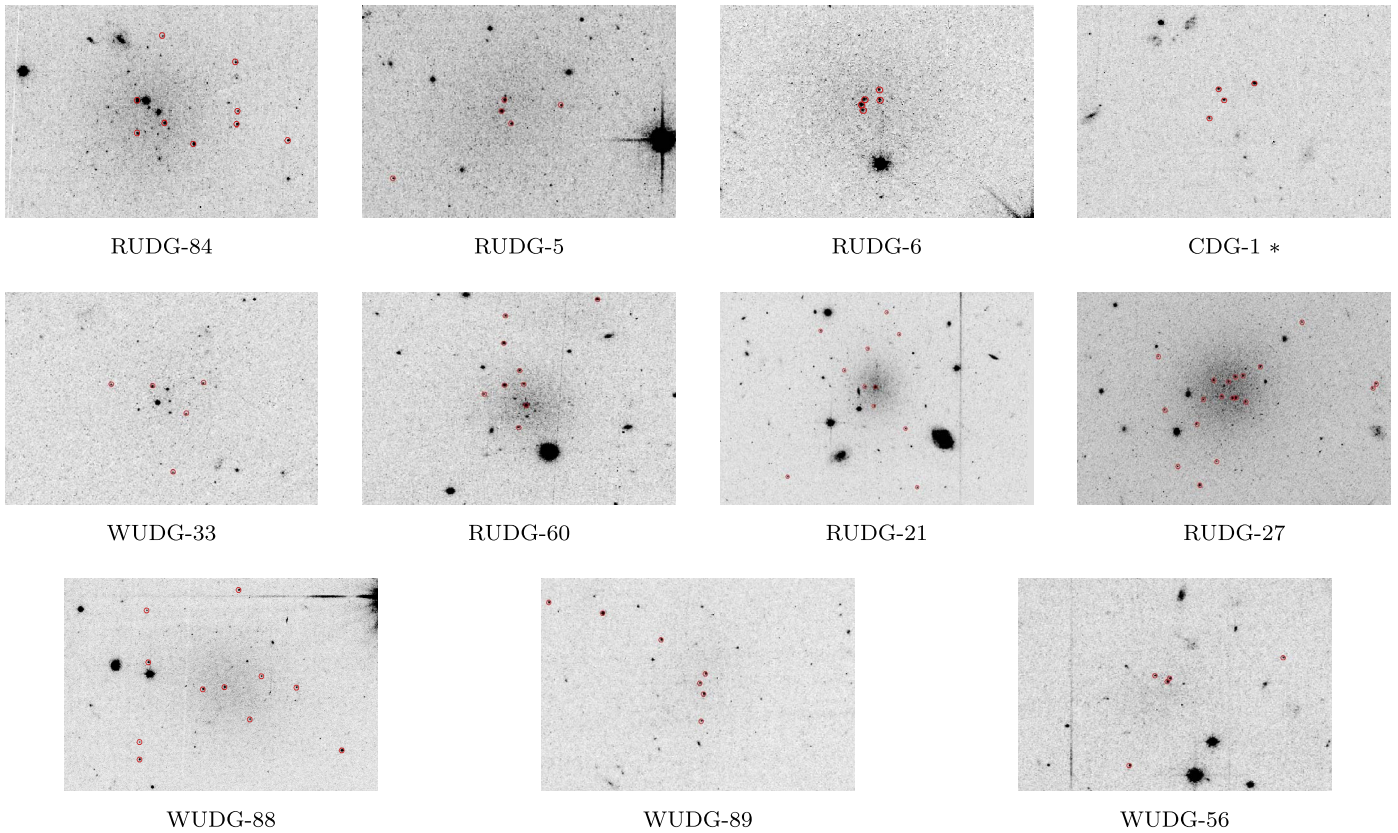
**Figure 2.** All UDG candidates detected within the PIPER survey data using LGCP. * represents the potential dark galaxy we found that has no detectable diffuse light. WUDG represents UDGs in the catalog of Wittmann et al. (2017) and RUDG represents UDGs in the catalog of Romanowsky et al. Red circles in the images are the selected GC candidates from the PIPER survey.

all 41 known UDGs. The background-subtracted GC counts are obtained by Janssens et al. (2022, in preparation). Out of the 41 known UDGs, 14 have background-subtracted GC counts greater than two. Of these, five were not detected using our method, namely, WUDG-5, WUDG-8, WUDG-13, WUDG-22, and WUDG-29. However, these UDGs all have background-subtracted GC counts just above two and the estimates are highly uncertain. Moreover, none of these UDGs has previously defined observed GC counts greater than two. On the other hand, there is one UDG we detected—WUDG-56—that has a background-subtracted GC count lower than two. But the PIPER GC data shows that there are in fact three GC candidates tightly clumped at the center of this UDG and our method is able to detect it.

### 4.2. Detection of CDG-1

In addition to detecting known UDGs, the most interesting object we have detected is CDG-1—*Candidate Dark Galaxy-1*. CDG-1 consists of four densely clumped observable GC candidates but has no detectable diffuse stellar content at that location. Since we are not yet certain of the nature of CDG-1, we will not classify it as a potential UDG.

CDG-1 is in the V8-ACS pointing, and it is the only overdensity that stands out. We include it here since, in theory, there is no condition that says a galaxy needs to have any diffuse stellar content. It is entirely possible to have galaxies that host GCs but are devoid of field stars. A more detailed detection analysis of this object will be covered in Section 5.

### 4.3. Effects of Neighboring Galaxies

Table 1 lists the detailed detection results for the 11 UDGs shown in Figure 2. The main results shown in Table 1 are the excursion function values for each detected UDG. Note that the computed excursion function values are presented for the four quantile excursion thresholds as mentioned in Section 3.2. Furthermore, the presented excursion function values are obtained by considering a region within the half-number radius from the center of each UDG,[17] and the maximum value of the excursion function in this region is chosen. We also present some of the characteristics of the UDGs that may have a significant impact on their detections, namely, the observed associated GC counts, estimated half-number radii of the GC systems, and whether or not there are bright normal galaxies within the pointing. The half-number radius for CDG-1 is obtained by fitting a separate Sérsic model to its observed GC system. We can see that many UDGs have extremely high excursion function values, which indicate that there is no issue in detecting them. However, there are some UDGs with excursion function values dropping to a very low level when the quantile thresholds are increased. Unsurprisingly, bright normal galaxies within the field of observations do seem to have a large impact on the excursion function values of the UDGs. It is worth noting that the excursion function values of CDG-1 are not the weakest in our list; in fact, it has a stronger detection signal than three out of the 10 known UDGs.

---

[17] Such a region is guaranteed to contain the maximum excursion function value associated with each UDG.

**Table 1**
Object ID, Coordinates, Visit Pointings, Observed Number of Associated GCs ($N_{GC}$), Half-number Radius of GC System ($R_{GC}$) from Janssens et al. (2022, in preparation), and the Maximum Excursion Function Values for the Detected UDGs

| ID | R.A. (J2000) | Decl. (J2000) | Pointing | $N_{GC}$ | $R_{GC}$ (kpc) | $F_{Q(0.5)}$ | $F_{Q(0.75)}$ | $F_{Q(0.9)}$ | $F_{Q(0.95)}$ | Galaxies? |
|---|---|---|---|---|---|---|---|---|---|---|
| RUDG-84 | 03 17 24.80 | +41 44 20.84 | V7-ACS | 9 | 1.24 | 0.999 | 0.999 | 0.996 | 0.987 | N |
| RUDG-5 | 03 17 34.50 | +41 45 21.97 | V7-ACS | 4 | 0.61 | 0.998 | 0.984 | 0.859 | 0.697 | N |
| RUDG-6 | 03 16 58.90 | +41 42 09.64 | V7-WFC3 | 5 | 0.53 | 0.999 | 0.999 | 0.997 | 0.984 | N |
| CDG-1 | 03 18 12.27 | +41 45 57.96 | V8-ACS | 4 | 0.67 | 0.870 | 0.700 | 0.518 | 0.385 | Y |
| WUDG-33 | 03 18 25.86 | +41 41 06.90 | V8-WFC3 | 5 | 0.86 | 0.732 | 0.520 | 0.350 | 0.262 | Y |
| RUDG-60 | 03 19 36.10 | +41 57 25.67 | V9-ACS | 8 | 1.45 | 0.999 | 0.999 | 0.999 | 0.997 | N |
| RUDG-21 | 03 20 29.40 | +41 44 50.70 | V10-ACS | 10 | 2.78 | 0.983 | 0.946 | 0.831 | 0.693 | Y |
| RUDG-27 | 03 19 43.40 | +41 42 47.72 | V10-WFC3 | 16 | 1.55 | 0.999 | 0.999 | 0.999 | 0.999 | N |
| WUDG-88 | 03 19 59.10 | +41 18 33.10 | V11-ACS | 7 | 1.41 | 0.800 | 0.607 | 0.421 | 0.322 | Y |
| WUDG-89 | 03 20 00.20 | +41 17 05.10 | V11-ACS | 6 | 1.19 | 0.639 | 0.363 | 0.171 | 0.100 | Y |
| WUDG-56 | 03 18 48.02 | +41 14 02.40 | V15-ACS | 3 | 0.60 | 0.943 | 0.882 | 0.603 | 0.469 | Y |

**Note.** The coordinates for WUDGs are from Wittmann et al. (2017), the coordinates for RUDGs are obtained from the RUDG catalog, the coordinate for CDG-1 is taken as the average of the coordinates of its constituent GCs. The excursion function values are presented for each of the quantile excursion thresholds (see details in the text). The excursion function here can be loosely interpreted as the probability that a UDG signal exceeds the excursion thresholds. The last column represents whether or not there are bright normal galaxies within or near the pointing.

To confirm the bright-galaxy effect, we conducted a preliminary analysis of the detection results using a simple linear regression with the excursion function values transformed into log-odds scale versus $N_{GC}$, $R_{GC}$, and whether or not there are bright normal galaxies within the pointing. We found that for all four quantiles, pointings with normal galaxies indeed have significantly lower excursion function values. This is somewhat expected since for pointings with no normal galaxies, as we did not include any covariate information in the LGCP models, i.e., there is only the intercept term. In other words, the presence of normal galaxies increases the uncertainty of the signal, which decreases the values of the excursion functions. This is especially clear with WUDG-89, which is extremely close (~40 kpc) to an elliptical galaxy (UGC 02673). This proximity causes the LGCP model to be uncertain about whether the GC over-density belongs to the elliptical or is a system of its own. Moreover, there is a larger UDG (WUDG-88) with a stronger signal in the same pointing, whose presence also negatively affects the detection probability of WUDG-89: its signal is the weakest among UDGs with similar $N_{GC}$ and $R_{GC}$. Nevertheless, we are still able to pick out the signal from WUDG-89.

The results from the linear regression model show that $N_{GC}$ and $R_{GC}$ have no significant impact on the detection probability. This is likely due to the small sample size, such that the actual effects of $N_{GC}$ and $R_{GC}$ are undetectable. Furthermore, the numerical ranges of $N_{GC}$ and $R_{GC}$ are too restricted to obtain any meaningful relationship.

### 4.4. Receiver Operating Characteristic Analysis of UDG Detection

The detection results presented previously only address LGCP's performance in providing true positive detection. We still require analysis of the overall performance of LGCP as a binary classifier which demands assessing LGCP's performance with respect to false positive, false negative, and true negative detection. To this end, we provide an overall detection analysis using the spatial receiver operating characteristic (ROC). Note that in what follows, the UDGs we are interested in are restricted to GC-rich UDGs.

The ROC (Fawcett 2006; Mas et al. 2013) is a widely used performance analysis tool in machine learning for assessing the overall performance of any binary classifier. In our problem, for any given excursion threshold $C$, the resulting excursion function obtained from LGCP can be viewed as a binary classifier. By choosing a probability threshold $q$, the excursion function will produce a corresponding Boolean map: pixels (regions) with excursion function values greater than $q$ will be assigned the true value, i.e., part of a UDG, while pixels with excursion function values less than $q$ will be assigned the false value, i.e., not part of a UDG. We can then compare the obtained Boolean map to the *truth* map: a reference map indicating which pixels are truly part of a UDG and which are not. The comparison will result in the predicted Boolean value of each pixel to fall into one of the four categories: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). To assess the performance of a classifier, the true positive rate (TPR) and the false positive rate (FPR) defined below are used jointly as performance measures:

$$TPR = \frac{TP}{TP + FN},$$
$$FPR = \frac{FP}{FP + TN}.$$

If we are able to find a value of $q$, so that the resulting TPR is one and FPR is zero, it then means the classifier has perfect performance.

For almost all binary classifiers, perfect performance is rare. Furthermore, the performance of the classifier depends on the choice of $q$: a low value of $q$ will produce high TPR while a high value of $q$ will produce low FPR. Therefore, an overall performance assessment of the classifier is obtained by plotting the TPR against the FPR with the corresponding value of $q$ ranging from $q = 0$ to $q = 1$. This procedure will result in a curve called the ROC curve. For illustration purposes, we present in Figure 3 the ROC curves of the four quantile excursion thresholds for the pointing V11-ACS. For the truth map of the V11-ACS pointing, true values are given to regions within twice the GC system half-number radius (shown in Table 1) from the center of the UDGs in the pointing, while the rest of the regions are assigned false values. If we can find $q$ such that the classifier has perfect performance, then its ROC curve will correspond to the black solid curve in Figure 3 since
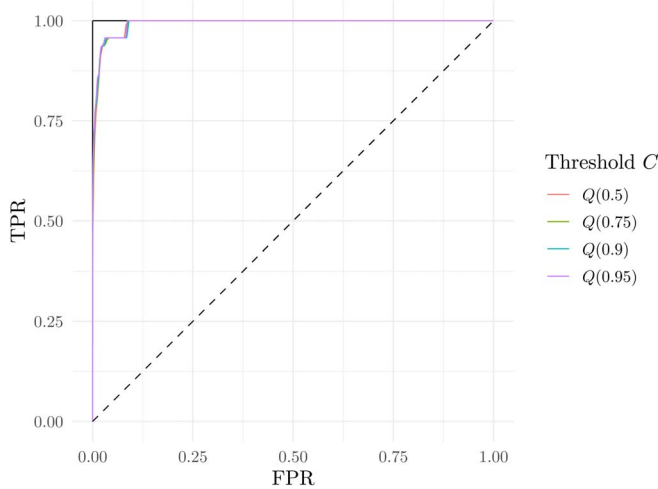
**Figure 3.** ROC curves of the four quantile excursion thresholds obtained for the pointing V11-ACS. Dashed black curve represents a random classifier; black solid curve represents a perfect classifier.

**Table 2**
AUC Computed from the ROC Curve of Each of the Quantile Excursion Thresholds for All Pointings with Confirmed UDGs

| Pointing | $\mathrm{AUC}_{Q(0.5)}$ | $\mathrm{AUC}_{Q(0.75)}$ | $\mathrm{AUC}_{Q(0.9)}$ | $\mathrm{AUC}_{Q(0.95)}$ |
|---|---|---|---|---|
| V7-ACS | 0.99721 | 0.99777 | 0.99795 | 0.99806 |
| V7-WFC3 | 0.99974 | 0.99984 | 0.99981 | 0.99981 |
| V8-WFC3 | 0.99917 | 0.99928 | 0.99952 | 0.99950 |
| V9-ACS | 0.99592 | 0.99709 | 0.99756 | 0.99763 |
| V10-ACS | 0.99517 | 0.99812 | 0.99833 | 0.99808 |
| V10-WFC3 | 0.99667 | 0.99812 | 0.99871 | 0.99915 |
| V11-ACS | 0.99222 | 0.99254 | 0.99270 | 0.99311 |
| V15-ACS | 0.99942 | 0.99961 | 0.99942 | 0.99937 |

it will have to hit the point (0, 1) (TPR = 1, FPR = 0). If a classifier is a random classifier, i.e., the classifier's performance is the same as random guessing, its ROC curve then corresponds to the dashed line in Figure 3. Hence, when comparing two classifiers, the better classifier will effectively have a ROC curve that approaches closer to the point (0, 1). We can see from Figure 3 that all the ROC curves for the four thresholds are essentially the same and they are all extremely close to the point (0, 1), indicating that the performance of our method is very good.

Intuitively, we might expect the performance of our classifier to be exactly perfect since the LGCP method can pick out the two UDGs in the V11-ACS pointing perfectly without any noise (see Figure 1). However, we do not have a perfect classifier as shown in Figure 3 because the truth map used in our ROC analysis has a mismatch with the intuitive notion of truth in our problem. The intuitive truth for our problem is the true locations of the UDGs instead of the projected regions occupied by the UDGs. However, to conduct the ROC analysis, we require the projected regions occupied by UDGs. It is impossible to obtain a perfect recovery of said regions since a UDG occupies a region with a poorly defined boundary. Furthermore, even if the region occupied by a UDG is the same as the ones we have set to, it is still not possible to obtain a perfect performance since the detected regions returned by LGCP will unlikely coincide exactly with the ones in the truth map. Nevertheless, we only need the ROC analysis for a bulk part assessment of the overall performance of LGCP, and it is completely adequate for our purpose here.

As shown in Figure 3, the overall performance of LGCP is quite spectacular for V11-ACS. However, it is highly inconvenient to present the performance results in the format of Figure 3 for all the pointings. A simple method to extract a one-number summary from a ROC curve is by the area under the curve (AUC), since the total area under the ROC curve will generally be greater for a better classifier. A perfect classifier will then have an AUC of one while a random classifier has AUC being a half.

Table 2 provides the computed AUCs obtained for the four quantile excursion thresholds for all the pointings that contain detected UDGs. Note that the pointing V8-ACS, which

contains CDG-1, is excluded as the nature of CDG-1 is not yet confirmed. Furthermore, we only consider the performance on detecting UDGs with sufficient GC populations. The truth map for each of the pointings is obtained in the same way as the one we have set for obtaining the ROC curves in Figure 3. As shown in Table 2, the AUC values for all excursion thresholds and pointings are almost one, indicating that the LGCP's performance as a binary classifier for identifying UDGs with significant GC populations is essentially perfect.

The detection results of the known UDGs and CDG-1 show a promising future for our proposed method, and the potential confirmation of CDG-1 as a *dark* galaxy would have profound implications for the understanding of galaxy evolution as well as fundamental physics. Thus, in the next section, we present some preliminary analysis to determine the nature of CDG-1 using observational data at hand.

For the sake of rigor and a better understanding of the performance and applicability in a more general setting, we require a calibrated assessment of our method using simulation. For interested readers, the simulation results are presented in Section 6. Otherwise, one can skip to Section 7 for a brief summary and the future plan of our work.

## 5. CDG-1: A Pure GC Galaxy?

### 5.1. Detection Results

Here, we present the detection results for CDG-1. Figures 4 and 5 show the results for the V8-ACS pointing. Figure 4 shows the posterior mean spatial random effect $\exp(\mathcal{U}(s))$ and Figure 5 presents the excursion functions for $\mathcal{U}(s)$ computed using the four quantile excursion thresholds. The highest intensity region in the lower-left corner of Figure 4 is the location of CDG-1. Figure 5 confirms with high probability across all thresholds that CDG-1 is indeed something worth our attention. Although there are some other signals (lower-right corner) in the excursion function when the excursion threshold is low, these signals are rather weak in comparison with CDG-1. They also completely disappear ($<0.1$) for higher thresholds and only the signal from CDG-1 persists.

The weaker signals in Figure 5 seem to be arising from the remote halo region of the massive ETG NGC 1265. This giant galaxy is $\simeq 120$ kpc away, well outside the field of view of the ACS, but the globular cluster systems of giant ellipticals often extend more than 100 kpc outward.[18] This weak signal can be

---

[18] A giant galaxy with the luminosity of NGC 1265 would normally have a population of ~5000 GCs for a GC specific frequency $S_N \simeq 4$ (Harris et al. 2013). Though there are individual differences from galaxy to galaxy, some dozens of GCs could easily extend beyond 100 kpc Galactocentric radius, a segment of which seems visible in the V8-ACS pointing.
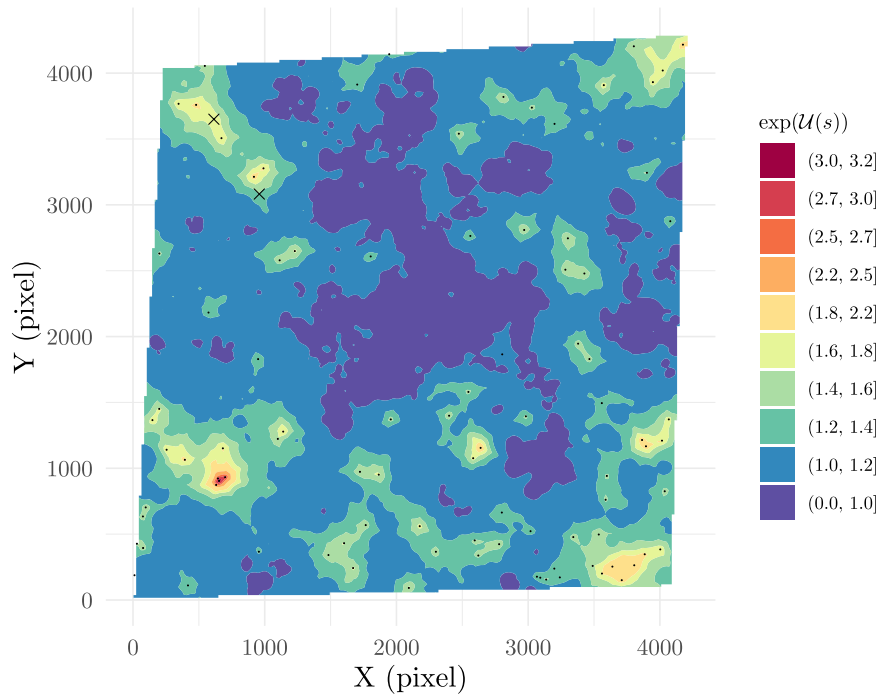
**Figure 4.** Posterior spatial random effect $\mathcal{U}(s)$ for V8-ACS, shown in exponential scale. Black points are GC candidates from the PIPER survey. Two known UDGs (WUDG-28 and WUDG-29) are in the upper left corner shown as black crosses. The object CDG-1 is in the lower-left corner, with the strongest signal in the entire pointing.

safely regarded as contaminating noise. We did not include NGC 1265 as a covariate because there is also a bright foreground star at the lower-right corner in the same direction. Thus, modeling the GC system of NGC 1265 would be extremely difficult and add additional unwanted uncertainty. In this sense, the actual signal strength of CDG-1 is quite strong as we did not remove the contaminating noise from NGC 1265.

A similar but lower-amplitude issue is that a smaller ETG, the lenticular IC 312, lies just off the lower-left corner of V8-ACS and generates a contaminating signal of its own. Due to its closeness to CDG-1, IC 312 is included as a covariate effect in our model. Further comments about IC 312 and NGC 1265 are given below.

### 5.2. Is CDG-1 Real?

Despite the strong signal from CDG-1, it is possible the four GCCs in CDG-1 are a random clustering signal arising from a PPP, i.e., they are a random association of intergalactic GCs belonging to the Perseus ICM. In V8-ACS and the adjacent V8-WFC3, the mean intensity of GCs in the magnitude and color range we are using (22.0 < F814W$_0$ < 25.5, 1.0 < (F475W−F814W)$_0$ < 2.4 as defined above) is ~13.7 arcmin$^{-2}$. If we make the simplest assumption that the GC distribution in V8-ACS follows an HPP, then the probability that four random GCs would fall within a similar area as in CDG-1 is ~3 × 10$^{-6}$, just a few parts per million.[19] However, this probability is an overestimate since V8-ACS and V8-WFC3 are contaminated by multiple ETGs; the actual GC mean intensity in the ICM should be much lower than 13.7 per arcmin$^2$ and so should the previously computed

probability. After a crude removal of regions with contaminating ETGs, the previous probability drops to ~1 × 10$^{-6}$.

Another alternative is that the CDG-1 clump may be some substructure that belongs to the halo of either IC 312 or NGC 1265, the bigger galaxies that are in the neighborhood. We need the complete GC catalogs of both galaxies to test the association of CDG-1 further, but these are not yet available.

Nevertheless, based on the observed features and location of CDG-1, it is unlikely that CDG-1 is a substructure in the halo of either galaxy; If it is, then it is either an infalling satellite or a clump of GCs within the larger galaxy. An infalling satellite, however, should still exhibit diffuse light.

CDG-1 is also unlikely to be a clump of GCs belonging to either galaxy. For NGC 1265, CDG-1 is too far (120 kpc) from the center of the giant ETG for its GC system to have such an over-dense clump of GCs. For IC 312, CDG-1 has a projected distance of ~26 kpc to its center (assuming the distance to IC 312 is 75 Mpc) while IC 312 has a half-light radius of ~2 kpc. If we assume the *five times* scale ratio between the half-number radius of the GC system and the half-light radius of IC 312, the half-number radius of its GC system is then ~10 kpc. From the Hyperleda catalog, the total absolute magnitude of IC 312 is $M_B = -20.32$ or $M_V \simeq -21.22$. For a specific frequency, $S_N = 2$ typical for galaxies in that range (Harris et al. 2013), the estimated total GC population of IC 312 would be $N_{GC} \sim 600$. Of these, about 160 will be brighter than our adopted limit of F814W at 25.5 by assuming a normal Gaussian-like luminosity distribution for the GCs in IC 312. Moreover, CDG-1 lies in the direction of the minor axis of IC 312 while IC 312 has an ellipticity of ~0.6. If the result in Wang et al. (2013) is assumed where the GC distribution follows the galaxy isophotes, then a very generous estimate of the number of GCs lying further than CDG-1 would be at most six. This suggests it is statistically improbable for IC 312 to have four GCs lying as far and clumping as close as in CDG-1.

---

[19] Here, we assume the area is the smallest circle encompassing all GCCs in CDG-1, which leads to an area of 0.0074 arcmin$^2$. Following condition 1 in Appendix A with $\lambda = 13.7$ arcmin$^{-2}$ and $|B| = 0.0074$ arcmin$^2$ would give the estimated probability by computing $\mathbb{P}(N(B) = 4)$.

(a) $F_{Q(0.5)}(s)$

(b) $F_{Q(0.75)}(s)$

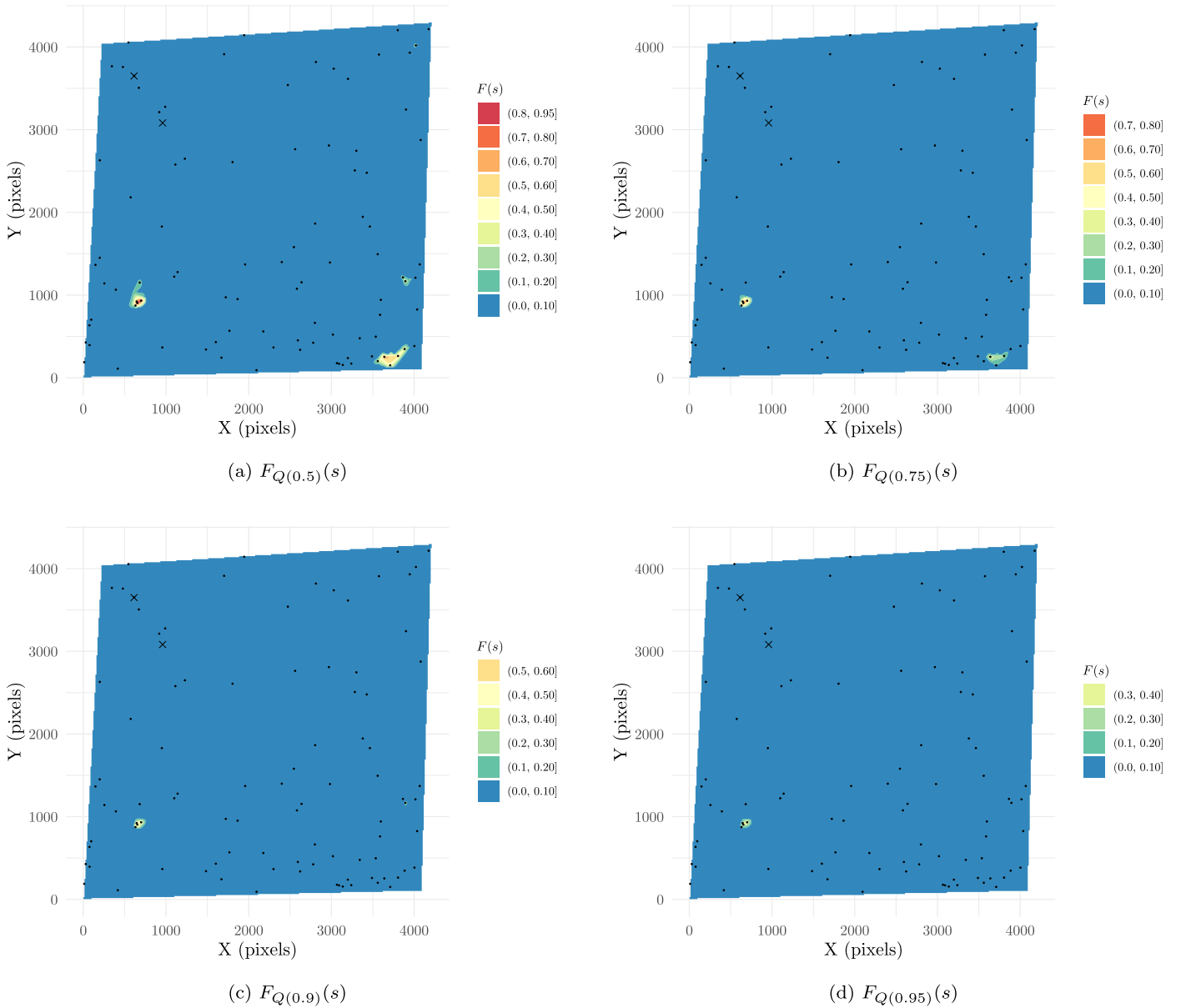(c) $F_{Q(0.9)}(s)$

(d) $F_{Q(0.95)}(s)$

**Figure 5.** Excursion functions for V8-ACS pointing. Black points are the locations of GC candidates from the PIPER survey. (a)–(d) are excursion functions for $\mathcal{U}(s)$ with four quantile excursion thresholds. Two known UDGs (WUDG-28 and WUDG-29) are in the upper left corner but have not detected significantly through GC counts. The weaker signals in the lower-right corner in (a) and (b) are most likely contaminating GCs from the neighboring giant radio galaxy NGC 1265 (see text for details). The *dark galaxy* candidate CDG-1 is clearly shown in the lower-left corner of all the panels.

One last possibility can also be considered: the four GCs defining CDG-1 are a group of faint background galaxies whose near-stellar morphologies and colors happen to match those of GCs seen at the Perseus distance. Close inspection of the objects in the region adjacent to CDG-1 shows that other similarly faint and obviously non-stellar objects are present. However, similarly faint background galaxies are scattered everywhere across the field and the region close to CDG-1 does not particularly stand out. The majority of these faint background galaxies are also much redder or much bluer than the color range of true GCs (Harris et al. 2020). Any Milky Way foreground stars are almost all distinctly redder than the GC range, so the same argument applies to them (see Harris et al. 2020, for a more extensive discussion of the background population). Though this option cannot be definitively ruled out at present, it must be viewed as unlikely. Nevertheless, we carefully inspect the properties of the GCs in the next section.

### 5.3. Properties of the GCs in CDG-1

The detailed aspects of the four globular cluster candidates (GCCs) in CDG-1 are provided in Table 3. Specifically, we provide their coordinates, magnitudes, and colors, luminosities, half-light radii, and metallicities.

The four GCCs in CDG-1 have magnitudes and color indices appropriate for genuine GCs. Very close inspection of their morphologies on the images with *iraf/imexamine* shows that all of them have profiles with FWHM $\simeq 2.2$ pixels, whereas an ideal star-like object would have FWHM = 2.0 px. The GCCs may then be marginally resolved. To quantify their intrinsic scale sizes, we applied ISHAPE (Larsen 1999), which convolves a King (1962) profile with the point spread function to produce a best-fit profile for each object and hence a half-light radius. The results for the four targets are presented in Table 3 under the column $r_h$. These scale radii are about twice as large as the most common GCs in the

<div align="center">

**Table 3**
Data for the GCCs in CDG-1

</div>

| ID | R.A.<br>(J2000) | Decl.<br>(J2000) | $F814W_0$<br>(Vegamag) | $(F475W-F814W)_0$ | $r_h$<br>(pc) | $\log(L/L_\odot)$ | (Fe/H)<br>(dex) |
|---|---|---|---|---|---|---|---|
| GCC-1 | 3 18 12.23 | +41 45 58.03 | $24.709 \pm 0.035$ | $1.445 \pm 0.088$ | 4.3 | 5.49 | $-1.19 \pm 0.28$ |
| GCC-2 | 3 18 12.43 | +41 45 59.55 | $23.535 \pm 0.046$ | $1.606 \pm 0.066$ | 6.7 | 5.97 | $-0.68 \pm 0.21$ |
| GCC-3 | 3 18 12.06 | +41 45 57.60 | $24.926 \pm 0.050$ | $1.446 \pm 0.098$ | 6.7 | 5.41 | $-1.19 \pm 0.31$ |
| GCC-4 | 3 18 12.29 | +41 45 57.23 | $24.773 \pm 0.039$ | $1.684 \pm 0.076$ | 6.5 | 5.48 | $-0.44 \pm 0.24$ |

Milky Way, for example, but fall well within the range of observed GCs considered more widely, particularly in dwarf galaxies (e.g., Georgiev et al. 2010). A more extended discussion of the $r_h$ distribution for the intergalactic GCs in Perseus is given by Harris et al. (2020), for which $r_h \sim 5$–7 pc is also a very typical value.

The four candidates, as GCs, are moderately bright: converting their F814W magnitude to luminosity gives the $\log(L/L_\odot)$ values in Table 3. These are similar to the bright Messier-type GCs in the Milky Way. If these four candidates are the brighter members of a GC population that follows anything like a normal luminosity distribution, then deeper exposures could be expected to reveal many less luminous ones and to add to the known population. Such data would enable a strong test of our interpretation. Conversely, they may turn out to resemble the very top-heavy GCLF seen in NGC1052-DF2 and DF4 (Shen et al. 2021), where clusters fainter than the classic turnover point are present but at surprisingly low numbers.

Since GC integrated color is primarily a function of their heavy-element abundance, the de-reddened color indices $(F475W-F814W)_0$ can be converted to give rough estimates of metallicity [Fe/H] (Harris et al. 2016). These are listed in the last column of Table 3 and again fall in the typical range for luminous GCs, though on average they are somewhat more metal-rich (redder) than most of the ones found in previously known dwarfs in the Local Group, Virgo, and nearby groups (e.g., Peng et al. 2006; Georgiev et al. 2010). However, the photometric data from Peng et al. (2006) show that a significant range of colors (metallicities) can be found for the GCs in any one dwarf.

In brief, the available data and inferred properties are consistent with the interpretation that the four objects within CDG-1 are genuine GCs. A clinching argument for their identities, though observationally quite challenging, would be direct measurement of their radial velocities.

### 5.4. Comparison to GCs in Other Systems

Aside from the spatial features, the colors of GCCs in CDG-1 should also satisfy certain properties if CDG-1 is indeed a dark UDG or dark galaxy. Based on the observation in Peng et al. (2006), GC color dispersion can exhibit quite a range of variation within dwarfs. Hence, we should expect the GC color dispersion within CDG-1 to be less than the GC color dispersion of the whole UDG population. However, since there are only four GCs in CDG-1, any statistical tests conducted to determine color dispersion differences have little meaning. Especially since it is much more likely to produce a false negative result than a false positive result for a small sample size. Nevertheless, a simple numerical comparison shows that the GC color dispersion of CDG-1 is not out of the ordinary with a mean color of 1.546 and a dispersion of 0.119. The average GC color for the 10 known UDGs detected in this

paper is 1.503 and the average color dispersion is 0.143. Compared to the 10 known UDGs, the GC color dispersion in CDG-1 is lower than eight of them. Hence, one should not be surprised if CDG-1 is indeed some sort of UDG or dwarf.

We also compare the color distribution of CDG-1 GCs and that of GCs within IC 312. This requires the complete GC data of IC 312, but there are only five GCs within V8-ACS that should belong to IC 312. The colors of these five GCs are 1.813, 1.85, 1.845, 1.502, and 1.384. Again, statistical tests for differences between the color distribution are not meaningful due to small sample sizes. Nevertheless, the average color of the five GCs in IC 312 is 1.678 with a standard deviation of 0.219. Comparing this with the four GCs in CDG-1, we see that, at least with the available data, the CDG-1 GCs are bluer with a much smaller color dispersion than IC 312 GCs.

In brief, observational evidence in hand is entirely consistent with the interpretation that CDG-1 is a dark—or, at least, extremely faint—galaxy consisting of at least four GCs and (as yet) no other stellar population to the current limit of detection. Deeper imaging should be capable of reinforcing (or disproving) this hypothesis. Moreover, for a normal GC mass-to-light ratio $(M/L)_V \simeq 2$, the total mass of the four GCs in CDG-1 adds up to $M_{GCS} = 3.3 \times 10^6 M_\odot$. The normal relation between GC system mass and host galaxy stellar mass, for dwarf galaxies that contain GCs, is

$$\log M_{GCS} = -0.725 + 0.788 \log M_* \qquad (12)$$

(Eadie et al. 2022). This relation would predict that CDG-1 should show up as a dwarf with $M_* \simeq 2.7 \times 10^9 M_\odot$, which would be easily visible on the HST images. Similarly, if the standard near-uniform ratio between GCS mass and host galaxy halo mass of $M_{GCS}/M_h = 2.9 \times 10^{-5}$ (Harris et al. 2017) applies in the dwarf regime (Forbes et al. 2018; Burkert & Forbes 2020), then CDG-1 should have had $M_h \sim 10^{11} M_\odot$ at time of formation. The fact that nothing is visible of any dwarf now may then be a signal that considerable stripping of its stellar mass has happened since.

### 6. Simulation

In this section, we analyze the performance of LGCP for UDG detection in more general scenarios by simulating the GC systems of UDGs, assuming the GC system is dictated by a Sérsic model. The question of interest here is: if there is a UDG in a certain region (in our case this would be the observational pointing made by HST), how well can LGCP detect the UDG given its physical structure? In other words, given the Sérsic model, we seek to understand how UDG detection is affected by (i) the Sérsic index ($n$), (ii) the half-number radius ($R_e$), and (iii) the number of GCs within the GC systems ($N_{GC}$).

Another variable of interest is environment. By environment, we mean the spatial distribution of GCs that do not belong to

the UDGs we want to detect, e.g., intergalactic GCs and GCs within normal galaxies. We shall call these GCs noise GCs. Intuitively, one would expect the more noise GCs there are in a region, the harder it is to detect the UDGs. However, generating realistic simulations that encompass the broad range of environments where UDGs exist is quite challenging. Therefore, we simply place simulated UDGs inside the pointings obtained by the PIPER survey. This ensures that the environment variable we consider is realistic while also being computationally manageable.

There are a total of 20 pointings surrounding the center of the Perseus cluster in the PIPER survey, but for the simulation conducted here, we only consider placing simulated UDGs in the eight pointings listed in Table 2 (i.e., those containing identified UDGs with sufficient GC populations).

Below we introduce some mathematical notations for our simulation procedures:

1. Denote the eight point patterns of GCs obtained by the PIPER survey as $X_k = \{x_{k,1},\ldots,x_{k,N_k}\} \subset S_k$, $k = 1,\ldots,8$. $N_k$ is the number of GCs in $X_k$ and $S_k$ is the observational window (field of view) for the $k$th pointing.
2. For $S_k$, we select 10 locations $p_{k,j} \in S_k$, $j = 1,\ldots,10$ to place the simulated UDG. These 10 locations are chosen to cover most of $S_k$. Note that regions where there are normal galaxies and confirmed UDGs are avoided.
3. For location $p_{k,j}$, generate a simulated UDG based on the Sérsic model with parameters $\theta_i = (n, R_e, N_{GC})_i$. Denote the point pattern of GCs within such a simulated UDG as $U_{kji}$. We then obtain a simulated point pattern $X_{kji} = X_k \cup U_{kji}$. The parameter values are set to: $n = \{0.5, 1, 1.5\}$; $R_e$ (pixel) = $\{80, 110, 140, 170, 200\}$; $N_{GC} = \{8, 12, 16, 20, 24\}$. Hence, there is a total of 75 configurations for $\theta_i$ with $i = 1,\ldots,75$.
4. Fit the LGCP model to the point pattern $X_{kji}$ and determine the detection measure, $D_{kji}$, of the simulated UDG $U_{kji}$.
5. Carry out 15 iterations of steps 3 and 4 to obtain a sample of detection measures of the simulated UDG for pointing $k$, location $j$, and parameter configuration $i$. Denote this sample as $\boldsymbol{D}_{kji} = \{D_{kji}^l\}_{l=1}^{15}$. We then integrate out the variations in the locations $j$ and the iterations $l$ by computing the average detection measure of the simulated UDG for the $k$th pointing and the $i$th parameter configuration:

$$\bar{D}_{ki} = \frac{\sum_{l=1}^{15}\sum_{j=1}^{10} D_{kji}^l}{150}.$$

We then investigate how the environment and the structural parameters of UDGs affect the performance of our method.

Note that the unit of $R_e$ specified in step 3 is set to pixels instead of the physical unit of kiloparsec. The reason for doing this is that the pixel widths in the ACS and WFC3 pointings have different physical scalings, but the dimensions of the observational windows for the two pointings are the same in pixels. The difference in physical scaling causes the same UDG to have different apparent sizes in the ACS and WFC3 pointings. Apparent size here means the size of the UDG with respect to the size of the observation window. The difference in the apparent size will then artificially introduce differences on the effect of $R_e$ in the two camera pointings if the unit of $R_e$ is

kpc. Therefore, the solution is to let the simulated UDGs have different physical sizes but the same apparent sizes in different pointings. The values of $R_e$ given in step 3 roughly corresponds to a physical sizes in kpc of $\{1.45, 2, 2.5, 3.1, 3.6\}$ in the ACS pointing and $\{1.2, 1.6, 2, 2.5, 3\}$ in the WFC3 pointing.

Indeed, the variable that ultimately affects the UDG detection is not the actual physical sizes of the UDGs but their apparent sizes. To see this, we can consider placing the same UDG at two different distances (assuming the same numbers of GCs are observed) but keeping the size of the observation window unchanged. It is easier to detect the UDG strictly through its GC when it is placed at a farther distance than at a closer distance. If the UDG is too close, it can take up the entire observation window, rendering itself undetectable. Although the root of the issue in our problem here is not difference in distance but difference in the pixel scaling, the underlying logic is the same. Note that we do not consider distance as a physical parameter that affects the UDG detection since in this work all UDGs have the same distances in the Perseus cluster.

To ensure the UDG simulation is as realistic as possible, we also consider the HST limit for detecting GCs. Specifically, for each of the simulated GCs, we also simulate a luminosity in the $I$ band based on the GC luminosity function derived in Harris et al. (2020). The resulting simulated apparent magnitude in F814W is then obtained by assuming $M_I = $ F814W. The apparent magnitude in F475W is obtained by assuming $M_B = $ F475W and $M_B - M_I = 1.5$ with a Gaussian scatter of 0.25 (Harris et al. 2020). We then discard the simulated GCs whose brightness is below the HST detection limit. The detection limits are given in Harris et al. (2020) where F475W > 27.67 and F814W > 27.07. The remaining simulated GCs will be the GC system that we utilize for analysis.

When we remove the GCs that are below the detection limit, the number of GCs, $N_{GC}$, provided in step 3 above is reduced, and the actual *observable* number of GCs in the simulated GC system is now random. Since the observable number of GCs potentially affects the UDG detection and it is the actual variable we want to investigate, we consider the average number of observable GCs of all simulated UDGs by averaging out all other variables. The observable number of GCs here does not include the background contamination since we know the exact value of observable number of GCs due to simulation.

Before presenting the simulation results, we also require the quantification of the environment (noise GCs) for each $S_k$. Since there can be numerous features characterizing the environment, e.g., number of noise GCs and the presence of normal galaxies, it is difficult to consider all such features. For simplicity, we consider the noise GC intensity level, $\lambda_{noise}$, as a general characterization of the environment. $\lambda_{noise}$ here is computed for each $S_k$ and defined as the total number of noise GCs in $S_k$ divided by the area of $S_k$.

We present the simulation results and analysis next. Similar to the detection measures employed in Tables 1 and 2, we will examine how the $\lambda_{noise}$ and the structural parameters affect both the detection probability and the overall performance measured by AUC.

### 6.1. Detection Probability

The detection probability considered here is only restricted to the simulated UDGs instead of real UDGs already in the pointings. Hence, when we compute $\lambda_{noise}$, the GCs from the
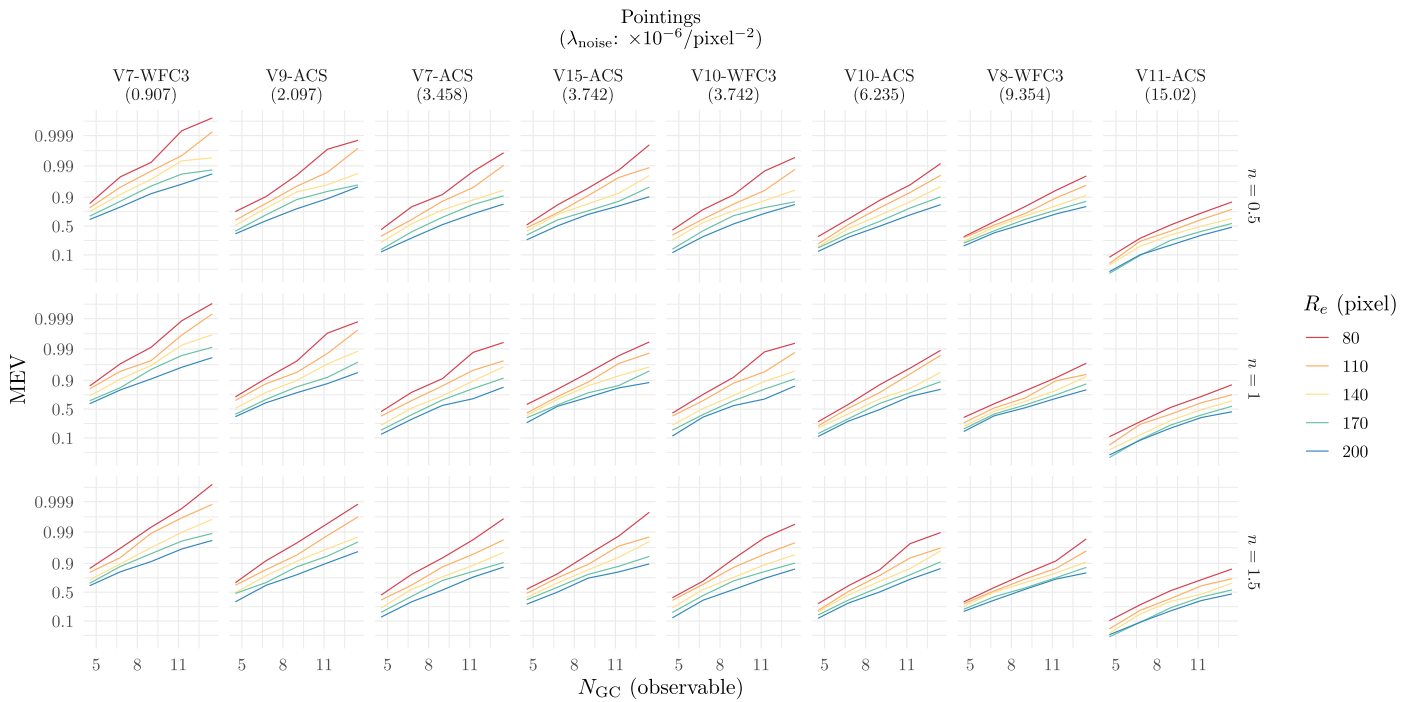
**Figure 6.** MEVs for the 95th quantile threshold against the *observable* number of GCs ($N_{GC}$), the Sérsic index ($n$), the GC system half-number radius ($R_e$), and the noise GCs intensity level ($\lambda_{noise}$). Every row has the same Sérsic index, every column shows the observational pointing and the corresponding $\lambda_{noise}$. The MEVs are shown in log-odds scale.

real UDGs within a pointing is also considered as noise GCs since these GCs are confounding noise signal with respect to the simulated UDG we try to detect.

The measure of the detection probability is similar to the one used in Table 1, i.e., for each of the simulated UDG, we consider a circle centered at the location of the simulated UDG with radius equal to the half-number radius of the simulated GC system, and we find the maximum excursion function value within the circle for a given excursion threshold. For simplicity, we shall call this the maximum excursion value (MEV).

Figure 6 shows the results obtained from simulation. To save space, we only present the results for the MEVs computed with the excursion threshold being $Q(0.95)$. The results are more or less similar for other excursion thresholds. Moreover, we found that the MEVs for lower levels of excursion thresholds are subject to more noise. This is caused by the relatively small number of iterations (15 times). Higher levels of excursion thresholds are more resistant to such a noise since the range of MEVs for higher threshold is smaller, hence the variance of MEVs is smaller.

As shown in Figure 6, the observable number of GCs heavily influences the detection probability. As expected, the more GCs there are within a UDG, the easier it is to detect the UDG. Moreover, Figure 6 shows that the MEV is above 0.5 if the observable number of GCs is $\gtrsim 13$. We focus on the value of 0.5 as a reference point since a region with MEV greater than 0.5 means that it exceeds the excursion threshold more often than not. If we exclude the simulation results for V11-ACS, we only need the observable number of GCs to be $\gtrsim 9$ for the MEV to be above 0.5. Hence, the performance of LGCP for detecting UDGs with larger GC populations is of no concern. Moreover, since the results we have shown here are for the highest excursion threshold—$Q(0.95)$, the resulting MEVs are inevitably low. The results obtained for the excursion threshold $Q(0.5)$ indicate that the MEVs are above 0.5 for all observed number of GCs except

for V11-ACS. This means that in using the excursion threshold $Q(0.5)$, we are very likely to detect UDGs in most scenarios, as long as the "observable" number of GCs is $\gtrsim 4$. This is an extremely promising result; the majority of the pointings within which we want to detect UDGs, they are unlikely to be as noisy as V11-ACS (at least within an environment similar to the Perseus cluster). Since our ultimate goal is to discover nearby UDGs with median to high number of GC population, the results from the simulation indicate we should have no issue detecting such a UDG using LGCP, if one does exist. One caveat of using a lower threshold such as $Q(0.5)$ is its tendency to produce false positive signals. Hence, our advice is to look at the excursion function for all thresholds and focus on signals that persist through all thresholds.

Figure 6 shows that the half-number radius also has a strong negative effect on UDG detection. This coincides with our intuition since the larger the half-number radius, the more dispersed the GCs within UDGs would appear, which reduces the strength of clustering of the GCs. To better detect UDGs with larger sizes, we can then look at the detection probability for lower excursion thresholds. Again, the MEVs obtained for the median ($Q(0.5)$) threshold are above 0.5 for all half-number radii except for V11-ACS. Another method to mitigate the negative effect of larger sizes on UDG detection is to consider a larger observation window, since the relationship we have discovered here is regarding the apparent sizes of the UDGs in a given observation window. A UDG with a smaller apparent size in an observation window would be easier to detect. However, we certainly need to balance the noise increase introduced by increasing the size of the observation window.

Unlike the number of GCs and the half-number radius, the Sérsic index does not appear to have any significant impact on the detection probability at all. This result is somewhat surprising since higher Sérsic indices would lead to a more concentrated distribution of GCs within a UDG, hence increasing the

clustering strength of GCs. We tested whether this is due to the range of Sérsic index being too restrictive. Hence, we performed another smaller set of simulations (results not shown here) with the Sérsic index ranging from {0.5, 1, 1.5, 2, 2.5, 3}. The result is similar to the one shown in Figure 6. We believe that the Sérsic index does have an impact on UDG detection, but the effect would only be apparent when we compare, say, $n = 0.5$ to $n = 8$. For the vast majority of currently detected UDGs, few have Sérsic indices exceeding 3. Therefore, we conclude that the Sérsic index is not a strong factor in UDG detection using LGCP.

Lastly, in Figure 6, we have also ordered the results of each pointing in ascending order of $\lambda_{\mathrm{noise}}$. We can see that as $\lambda_{\mathrm{noise}}$ increases, the MEVs decrease, which is expected. UDG detection will be harder in a more contaminated environment. For a better visualization of the numerical relationship, Figure 7 shows the average MEVs computed for each pointing (averaging out all other variables) against the corresponding $\lambda_{\mathrm{noise}}$. We can see an almost perfect negative linear relationship exists between the MEVs and $\lambda_{\mathrm{noise}}$. Certainly, there are fluctuations in the relationship, e.g., V15-ACS has a higher average MEV than V7-ACS, while V7-ACS is less *noisy* than V15-ACS. This is an expected random fluctuation since $\lambda_{\mathrm{noise}}$ is an extremely simple one-number summary of the environment. Upon detailed inspection, V15-ACS consists mostly of empty space where a relatively small region contains the majority of the noise GCs that belong to a dE within the pointing. In contrast, V7-ACS has GCs evenly distributed everywhere. Hence, $\lambda_{\mathrm{noise}}$ is not a summary statistics that can sufficiently highlight the aforementioned difference between V15-ACS and V7-ACS but it is simple enough to provide us with useful insights.

Aside from the effects of the related parameters on the UDG detection probability, we are also interested in whether the data from simulation can be used to provide us with any prediction capability. Essentially, we are aiming to answer the following interesting and important question: for a UDG with given physical parameters regarding its GC systems, what would be its detection probability under LGCP? Being able to answer this question would ensure our method is applicable in more general settings. We hence fit the following polynomial regression model using the simulated data:

$$\log\left(\frac{\mathrm{MEV}}{1 - \mathrm{MEV}}\right) = P_2(N_{\mathrm{GC}}^o, R_e, \lambda_{\mathrm{noise}}, p) + \epsilon. \qquad (13)$$

$P_2(\cdot)$ here is a second-order polynomial. The observed number of GCs ($N_{\mathrm{GC}}^o$), the half-number radius ($R_e$), $\lambda_{\mathrm{noise}}$, and the value $p$ for the quantile excursion thresholds $Q(p)$ are used to construct the basis for the polynomial. We did not include the Sérsic index since it is demonstrated in Figure 6 that the Sérsic index does not affect the MEVs. $\epsilon$ here is a mean-zero Gaussian error term. Note that $N_{\mathrm{GC}}^o$ here is obtained the same way as the $N_{\mathrm{GC}}$ values given in Table 1 since in real data we do not know the true number of observable GCs associated with a UDG. We have also fitted the following linear regression model:

$$\log\left(\frac{\mathrm{MEV}}{1 - \mathrm{MEV}}\right) = \beta_N N_{\mathrm{GC}}^o + \beta_R R_e + \beta_\lambda \lambda_{\mathrm{noise}} + \beta_p p + \epsilon,$$
$$(14)$$

but we found that Equation (13) fits the data much better compared to Equation (14). We then compute the predicted MEVs and the 95% prediction intervals of MEVs for all the
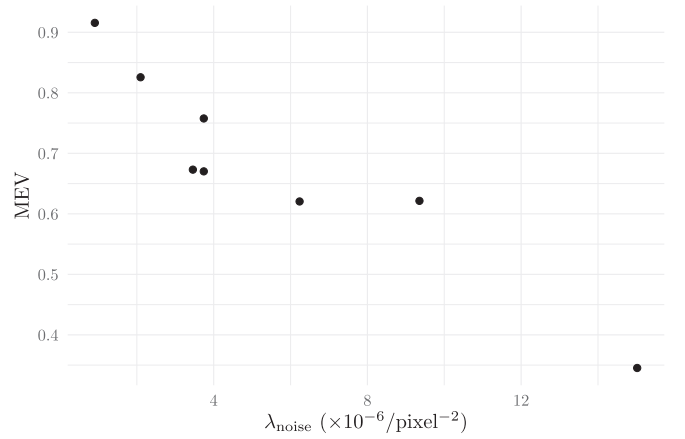


**Figure 7.** Average MEV for the 95th quantile threshold vs. the noise GC intensity level of each pointing.

detected UDGs listed in Table 1 and compare the prediction to the true MEVs listed in Table 1. The results are shown in Figure 8.

Figure 8 shows that almost all the true MEVs for the detected UDGs are within the 95% prediction intervals produced by Equation (13). The UDGs whose MEVs are outside of the prediction intervals are those with extremely high MEVs ($\gtrsim 0.98$), and their predicted MEVs are extremely high as well. Hence, the prediction error is in fact minuscule and negligible. For UDGs with lower true MEVs, the majority of the predictions from the model match with the true MEVs extremely well, except for WUDG-89 in the V11-ACS pointing. The larger prediction error for this UDG is somewhat expected since as mentioned, V11-ACS is a severely contaminated pointing and WUDG-89 is extremely close to the boundary of the giant ETG UGC 02673. Thus, we do not expect the prediction to match the true MEVs well in this case.

The results in Figure 8 suggest that the simulated data can be quite useful in determining the expected detection probability for any UDGs, at least within an environment similar to the Perseus cluster. The prediction model thus can be used as a screening mechanism for UDG detection: we will obtain the MEVs for UDG candidates in any pointing and then serve the corresponding parameters of the detected candidate into the prediction model and compute the predicted MEVs. If the true MEVs are close or much higher than the predicted MEVs, then it suggests the detection may be real and subsequent confirmation and verification will follow suit.

We also want to highlight the prediction for CDG-1. We can see that the predicted MEVs for CDG-1 match quite well with the true MEVs for all thresholds, although there does seem to be a small degree of overestimation from Equation (13). We suspect the overestimation is likely due to the contaminating noise from NGC 1265, which may have driven down the MEVs of CDG-1. Since the pointing (V8-ACS) where CDG-1 resides is not used in the simulation, in machine learning language, CDG-1 is essentially a test data point for model 13. The matching between the predicted MEVs and the true MEVs for CDG-1 suggests the prediction produced by model 13 is quite generalizable.

### 6.2. ROC Analysis

In this section, we analyze how the physical parameters affect the overall detection performance of LGCP using ROC analysis. Unlike the analysis for detection probability where we
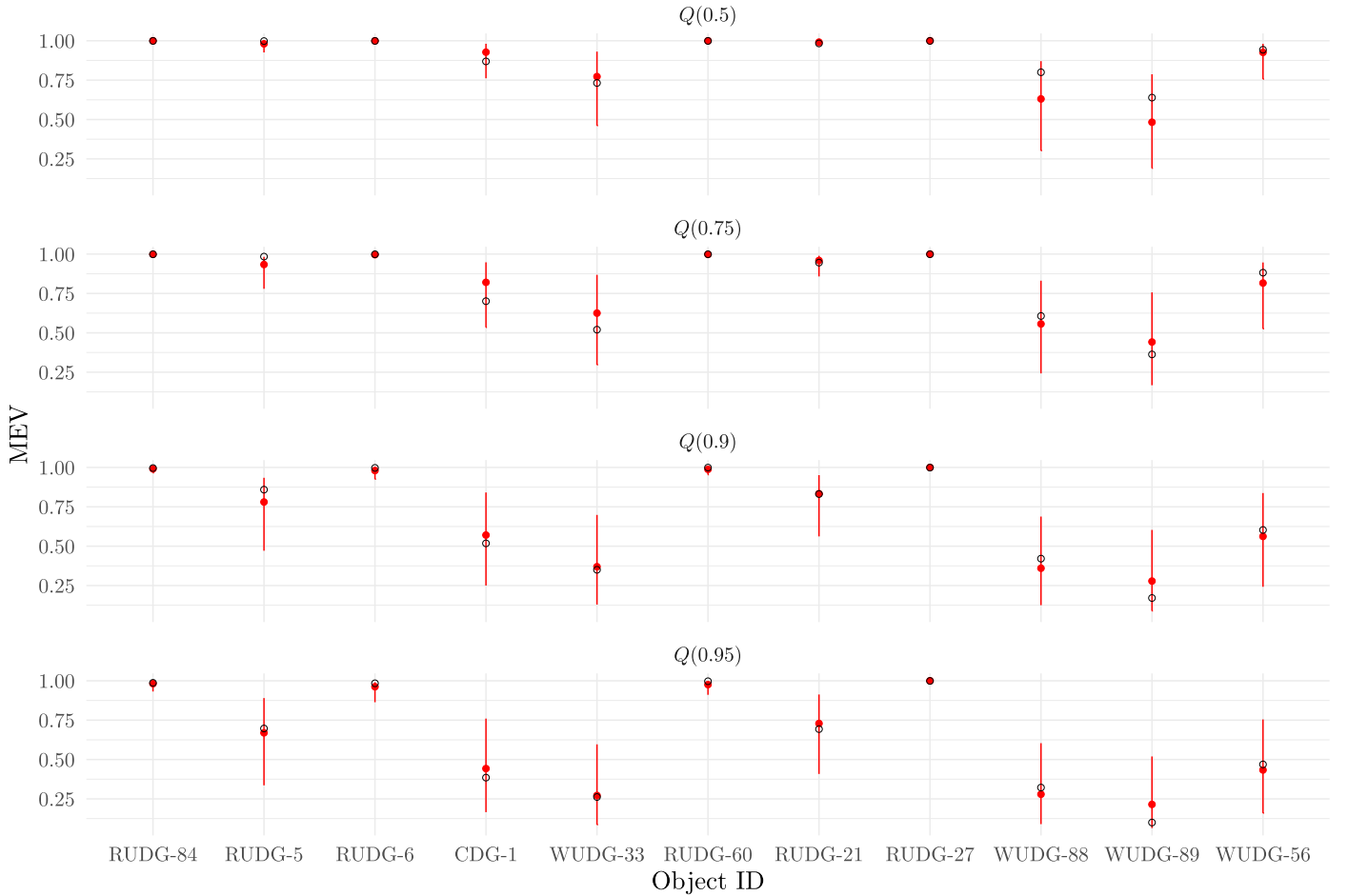
**Figure 8.** Predicted MEVs for all detected UDGs given in Table 1 using Model 13 fitted with the simulated data. Red dots represent the predicted values while red vertical lines indicate the 95% prediction intervals. Black circles show the true MEVs obtained in Table 1.

only focused on detecting the simulated UDGs, we consider the detection of both the already identified UDGs and simulated UDGs in each pointing for the overall performance. This is done to ensure we understand how the overall performance of LGCP behaves when there are multiple UDGs in a pointing. Moreover, if we only consider simulated UDGs, the real UDGs in each pointing will be considered as *false positive* when we conduct ROC analysis, which is not desirable. Therefore, for the ROC analysis, the truth map of each pointing consists of regions with detected UDGs, simulated UDGs, and no UDGs. For regions where there are UDGs, we determine them the same way as in Section 4.4 where regions within twice the GC system half-number radius from the center of a UDG are assigned the true values while other regions are assigned false values.

Since it is not possible to present the ROC curves for all the parameter configurations, we present in Figure 9 example average ROC curves obtained for the 95th quantile excursion threshold from four different sets of simulations. These roughly encompass the best to worst performance of LGCP in terms of AUC in our simulation. Figure 9 shows that even when LGCP has the worst performance, the corresponding AUC has a value of 0.76. Thus, LGCP classification is still much better than random guessing.

To fully investigate how the overall performance of LGCP is affected by the physical parameters, we consider the AUC values. The AUC values for each parameter configuration are

computed based on the simulation procedure described in Section 6 where AUC is now the detection measure *D*. The results are presented in Figure 10. Note that similar to the detection probability, we also only present the ROC analysis results for the 95% quantile excursion threshold. The ROC analysis results are almost exactly the same across all excursion thresholds.

As shown in Figure 10, similar to the detection probability, the observable number of GCs positively affects the AUC while the half-number radius of the GC system negatively impacts the AUC. The Sérsic index still does not seem to affect the AUC. The explanation of these observed effects is essentially the same as the ones previously illustrated for detection probability. Furthermore, for the majority of parameter configurations, the AUC values are above 0.9, indicating that the overall performance of LGCP for UDG detection is quite remarkable. However, despite the similarity to the results for detection probability, the effects of the physical parameters on the AUCs also have a few notable differences.

First, the effect of the observable number of GCs and the half-number radius on the AUCs are highly non-linear. We can see in Figure 10 that the AUCs will reach a plateau as the number of observed GCs increases. We believe this is unlikely due to a physical reason. As mentioned in Section 4.4, the Boolean map obtained by LGCP is highly unlikely to match exactly with the truth map we have employed. Therefore, once the AUC reaches a certain level, it cannot increase any further
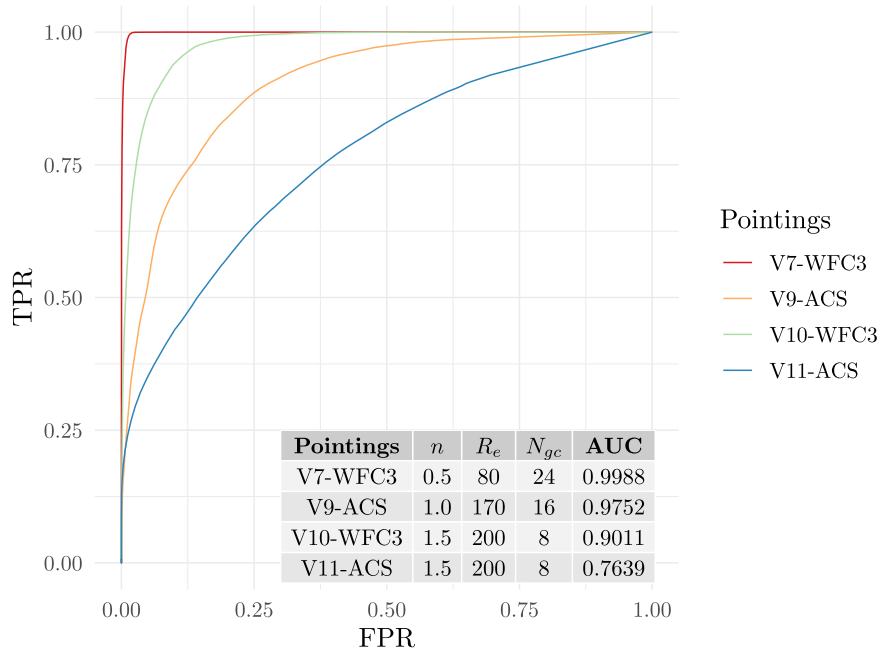
| Pointings | $n$ | $R_e$ | $N_{gc}$ | AUC |
|-----------|-----|-------|----------|-----|
| V7-WFC3 | 0.5 | 80 | 24 | 0.9988 |
| V9-ACS | 1.0 | 170 | 16 | 0.9752 |
| V10-WFC3 | 1.5 | 200 | 8 | 0.9011 |
| V11-ACS | 1.5 | 200 | 8 | 0.7639 |

**Figure 9.** Example ROC curves computed for the 95th quantile threshold from four different sets of simulations. The sets of simulations chosen are to represent the complete range of performance of LGCP from best to worst in terms of AUC.

as there are always a very small number of pixels in the truth map that LGCP simply cannot correctly classify.

Second, the noise GCs intensity level ($\lambda_{\mathrm{noise}}$) of each pointing does not seem to have any clear relationship with the AUCs. This is in contrast to the observed negative relationship between $\lambda_{\mathrm{noise}}$ and the detection probability as demonstrated in Figures 6 and 7. We suspect the environmental variables affecting the AUC cannot be properly characterized by $\lambda_{\mathrm{noise}}$. We can easily see that one of the environmental variables that would affect the AUC is whether there are over-densities of noise GCs within a pointing since these can become false positive detection and severely affect the AUC. However, $\lambda_{\mathrm{noise}}$ cannot characterize whether a pointing has over-densities of noise GCs or not, and there are no good summary statistics for characterizing such an environmental variable.

Note that we are not suggesting the $\lambda_{\mathrm{noise}}$ has absolutely no effect on the AUC since we can see from Figure 10, V11-ACS pointing still has the lowest AUC values out of all pointings. Since V11-ACS is extremely contaminated, the high level of noise does ultimately cause the AUC to be low. Another thing to note is that V7-WFC3 and V15-ACS have two of the highest AUC values. These two pointings have large areas where there are no GCs, and they also have no significant clumps of noise GCs. Thus, the resulting AUCs for these two fields are high.

In contrast, the detection probability is a rather locally defined quantity: having over-densities of noise GCs far away from a UDG is unlikely to affect the detection probability of the UDG. However, it will be strongly affected by $\lambda_{\mathrm{noise}}$ of the pointing since, intuitively, raising the base level of noise inevitably reduces the strength of the signal.

Another conclusion we can draw from Figure 10 is that we should have no trouble of detecting multiple UDGs within the same pointing. Looking at the highest AUC values in Figure 10 for each pointing, they are very similar to the ones obtained in Table 2. We focus on the highest AUC values since it ensures the simulated UDGs in the pointings are correctly identified. Furthermore, these cases correspond to the scenarios where the

signals from simulated UDGs are the strongest, which makes the signals from real detected UDGs most likely to be drowned out. Nevertheless, the AUC values in Figure 10 and Table 2 are nearly identical, suggesting that aside from the simulated UDGs, the real detected UDGs in each pointing are also detected.

## 7. Summary and Future Works

In this paper, we have introduced the LGCP—a novel method for detecting ultra-diffuse galaxies by identifying the over-density of their associated GCs instead of first finding their associated diffuse light. We have demonstrated the usage and effectiveness of the method by applying it to data from the PIPER Survey in the Perseus cluster. The LGCP was able to successfully detect all the confirmed UDGs with sufficient GC populations within the PIPER Survey field of view. ROC analysis has shown that the method also produced no false positive detection at least within the data used.

Using our method, we have detected a potential *dark* galaxy in the Perseus cluster—CDG-1 in our catalog—which is not visible with the traditional detection method through diffuse light. Examination of the spatial features of its constituent GCs showed it is unlikely a random alignment of intracluster GCs in Perseus. However, it is near a normal lenticular galaxy IC 312 and currently the GC catalog of IC 312 is still proprietary, hence we cannot yet definitively confirm whether CDG-1 is some substructure of IC 312. The giant ETG NGC 1265 is also near enough in projection (120 kpc) that its large GC system may overlap the region of CDG-1. Nevertheless, preliminary analysis of its spatial and color distribution reveals that the clump of objects defining CDG-1 is also highly unlikely to be associated with these galaxies. In future work, we intend to obtain deeper imaging of this object and the neighboring galaxies using HST. Combined with data from the Subaru Suprime camera once it is available, we aim to test and confirm the nature of CDG-1.
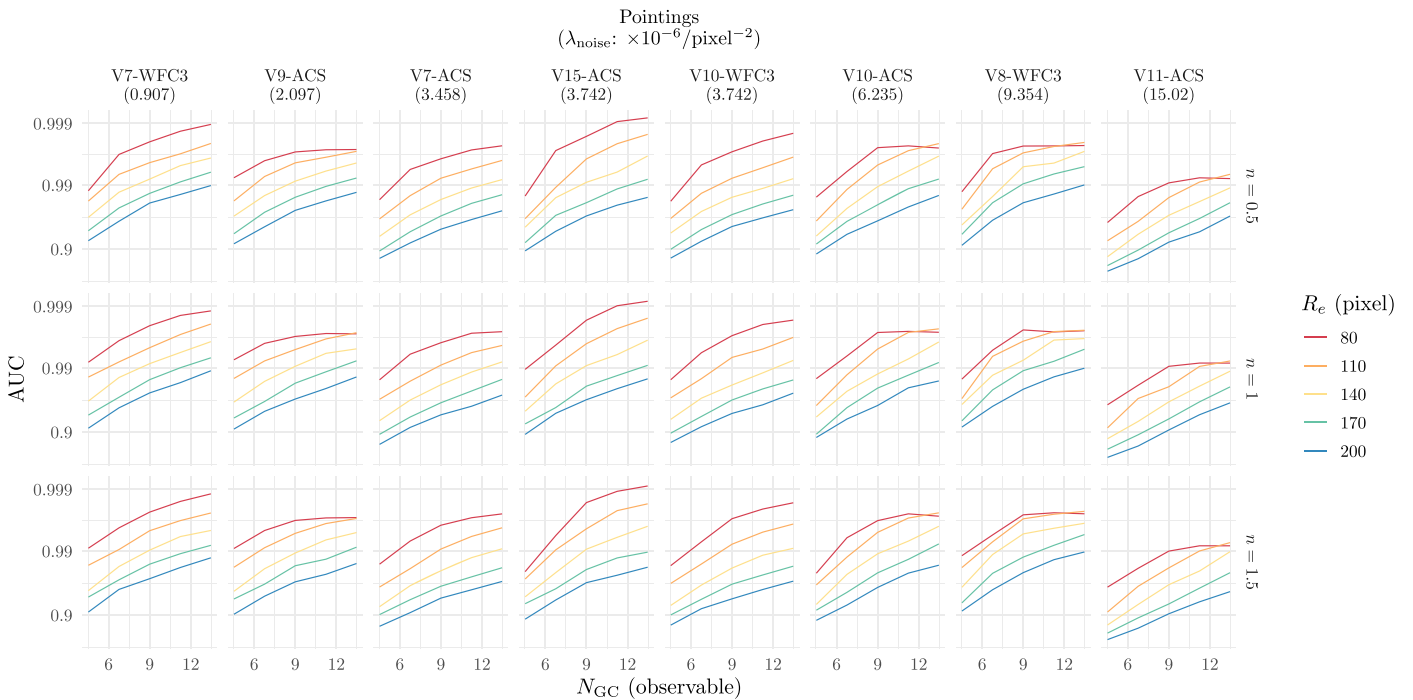
**Figure 10.** AUCs for the 95th quantile threshold against the *observable* number of GCs ($N_{GC}$), the Sérsic index ($n$), the GC system half-number radius ($R_e$), and the noise GCs intensity level ($\lambda_{noise}$). Every row has the same Sérsic index, every column shows the observational pointing and the corresponding $\lambda_{noise}$. The AUCs are shown in log-odds scale.

We also analyzed the performance of the LGCP method in terms of true detection and overall performance for UDG detection within the Perseus cluster using simulation. We found that under a Sérsic model for the simulated GC systems of UDGs, the Sérsic index does not have any significant impact on the detection probability of UDGs, while the number of GCs has a strong positive effect on the detection probability, and the half-number radius of the GC systems negatively affects the detection probability. We also concluded that the environment where a UDG resides strongly influences its detection probability: as the environment becomes more noisy, the detection probability decreases. Nevertheless, we found that for the majority of the environments considered, LGCP performs remarkably at detecting a UDG.

Regarding the overall performance of LGCP measured by AUC, we have found that the physical parameters of the GC systems of UDGs have similar influences on the AUCs as they do on the detection probability. However, the effect of the environment on the overall performance of LGCP is rather unclear as it depends on a set of descriptive statistics that are hard to quantify. Moreover, the simulated results showed that the LGCP has no issue of detecting multiple UDGs within a pointing which is a characteristic that is highly desired.

Lastly, other applications of LGCP can also lead to fruitful results. Specifically, LGCP could be used to find ultra-faint satellite galaxies in the Local Group, as well as faint star clusters in the halos of nearby galaxies. In both of these cases, we will be looking for clumps of stars embedded in a background population. The significant advantage of LGCP at tackling these problems is that satellites and star clusters in the halo will be close to their host galaxies, thus the ability of LGCP to incorporate covariate information to exclude the signals from host galaxies can be extremely efficient at detecting the weaker signals from satellites and star clusters.

In conclusion, LGCP is quite a reliable method for detecting UDGs as demonstrated by its efficacy at detecting confirmed UDGs, as well as simulated UDGs. However, the results shown in this paper only serve as a preliminary first step, since the method was only applied to UDG detection in the Perseus cluster. In the future, work is needed to calibrate the method in a much more general setting. In short, we intend to apply the method to other data sets to better understand the capability of LGCP for UDG detection in different environments other than the Perseus cluster.

Furthermore, more simulation studies would also assist in understanding and calibrating the detection procedure and criteria for UDG detection using LGCP. The ultimate goal here is to detect potential dark UDGs in the nearby universe which only consist of observable GC populations and are devoid of individual stars. The discovery of such an object in the nearby universe would be groundbreaking and significantly advance the fields of astronomy and fundamental physics.

discussion and recommendation of the method to compute the marginal quantile of the posterior spatial random effect of LGCP. The high performance computing cluster used in this work is SHARCNET (Shared Hierarchical Academic Research Computing Network).

*Facilities:* HST (ACS, WFC3).

*Software:* R Statistical Software Environment (R Core Team 2021); Main R packages include: R-INLA, inlabru (Bachl et al. 2019), excursions (Bolin & Lindgren 2018); Secondary R packages: tidyverse (Wickham et al. 2019), spatial(Venables & Ripley 2002).

# Appendix A
# PPP

Consider any Borel-measurable set $B \subset S \subset \mathbb{R}^2$, let $X$ denote a point process defined on $S$ and $N(B) \in \mathbb{N}$ denote the random variable counting how many points of $X$ there are in $B$. Then $X$ is a PPP with intensity function $\lambda(s)$, denoted by $X \sim \mathrm{PPP}(\lambda(s))$, if it satisfies the following conditions:

1. $N(B) \sim \mathrm{Poisson}(\mu(B))$ where $\mu(B) = \int_B \lambda(s)ds$. That is the number of points within $B$ is a Poisson random variable with mean $\mu(B)$.
2. For any Borel-measurable sets $A, B \subset S \subset \mathbb{R}^2$ such that $A \cap B = \varnothing$, $N(A)$ and $N(B)$ are independent.

The above definition implies some important properties of PPP. First, if we assume that the intensity function $\lambda(s)$ is constant, say $\lambda(s) \equiv \lambda > 0$, then condition 1 from above implies $\mathbb{E}(N(B)) = \mu(B) = \lambda|B|$, where $|\cdot|$ is the area of $B$. It is then easy to see that $\lambda$ here is the familiar mean surface number density in astronomy.

Second, assume again that $\lambda(s) \equiv \lambda$, if we want to simulate a PPP in $S$ with intensity $\lambda$, we only need to generate a number from a Poisson random variable with parameter $\lambda|S|$, say $n$, then generate $n$ points uniformly across $S$. On another note, for the simulation of the GC system of UDG carried out in Section 6, given the intensity function $\lambda(s)$ (a Sérsic profile) and the number of GCs ($N_{\mathrm{GC}}$) we want to generate, the locations of GCs can be generated by randomly sampling $N_{\mathrm{GC}}$ number of points from the probability density function $\pi(s) = \lambda(s)/\int_S \lambda(s)ds$.

The last and the most important property of a PPP is that the occurrence of a point has absolutely no effect on the occurrence of any other points. It is due to this last property that a PPP with constant intensity function is used as a reference point process to represent the notion of *complete spatial randomness*.

# Appendix B
# Computation of the Quantiles of Posterior Marginal Distribution of $\mathcal{U}(s)$

The marginal distribution of a random field $\mathcal{U}(s)$ is effectively its one-point distribution function in astronomy. Hence, if $\mathcal{U}(s)$ is a zero mean GP with marginal variance $\sigma^2$, its marginal distribution is then $\mathcal{N}(0, \sigma^2)$. Thus, Diggle et al. (2005) first suggested to compute the quantiles of the posterior marginal distribution of $\mathcal{U}(s)$ by assuming $\mathcal{U}(s)|\boldsymbol{x}$ approximately follows $\mathcal{N}(0, \sigma^2)$ where $\sigma$ is a point estimate from the posterior distribution of $\sigma$. We then choose our quantiles from $\mathcal{N}(0, \sigma^2)$. However, this procedure is only suitable for visualization purposes, and is not suited for accurate numerical analysis since it is a frequentist approach imposed on a fully Bayesian framework. Moreover, even though the prior distribution of $\mathcal{U}(s)$ is a Gaussian process, the posterior distribution $\mathcal{U}(s)|\boldsymbol{x}$ is not, and it is in fact analytically intractable. This means that, theoretically, the true posterior marginal distribution $\mathcal{U}(s)|\boldsymbol{x}$ is unknown. Hence, the accuracy of the approximation of $\mathcal{U}(s)|\boldsymbol{x}$ by a normal distribution is not guaranteed. However, a workaround is by the following identity:

$$\pi(\mathcal{U}(s)|\boldsymbol{x}) = \int \pi(\mathcal{U}(s)|\sigma, \boldsymbol{x})\pi(\sigma|\boldsymbol{x})d\sigma. \tag{B1}$$

The workaround method is using the fact that $\mathcal{U}(s)|\sigma, \boldsymbol{x}$ can be very well approximated by a GP through INLA. This means that the marginal distribution for $\mathcal{U}(s)|\sigma, \boldsymbol{x}$ is now well approximated by $\mathcal{N}(0, \sigma^2)$. Therefore, Equation (B1) means that the posterior marginal of $\mathcal{U}(s)|\boldsymbol{x}$ is well approximated by a mixture distribution of $\mathcal{N}(0, \sigma^2)$ where $\sigma$ is generated from the posterior distribution of $\sigma$.

Based on the previous argument, a principled framework we adopt is the following: we draw a sample $(\sigma_1,...,\sigma_n)$ from the posterior distribution $\pi(\sigma|\boldsymbol{x})$. For each of the individual $\sigma_i$, $i \in \{1,...,n\}$, we generate a random value from $\mathcal{N}(0, \sigma_i^2)$. From this sample of a mixture of normal distributions, we compute the empirical $p$th quantile, $Q(p)$. This is then the actual $p$th quantile of the posterior marginal distribution $\mathcal{U}(s)|\boldsymbol{x}$.
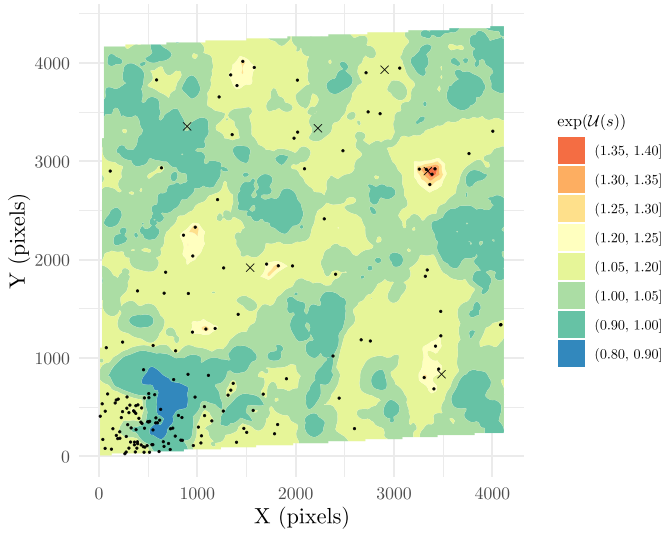
# Appendix C
# Detection Results for Pointings with Detected UDGs

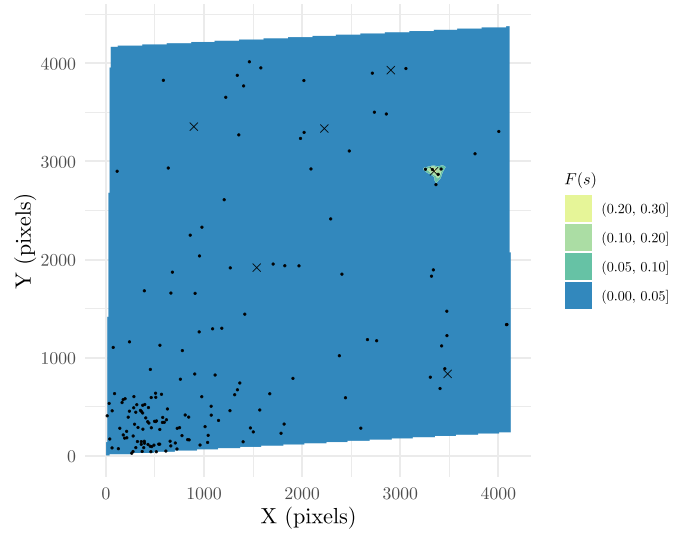## C.1. Pointings with Contaminating Bright Galaxies

Figures 11–13 show the detection results for pointings with contaminating bright galaxies.
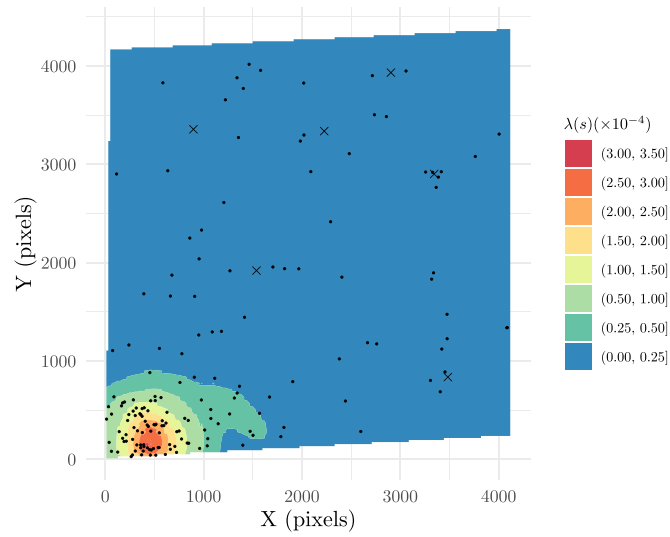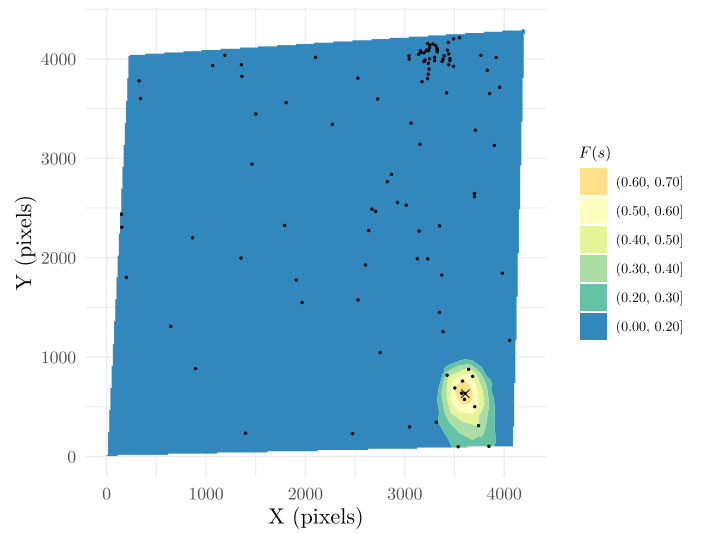
(a)



(b)



(c)

**Figure 11.** Posterior results for V8-WFC3 pointing. (a) Posterior mean intensity $\lambda(s)$; (b) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$; (c) Excursion function $F_C(s)$ with $C = Q(0.95)$. Black points are locations of GC candidates. Black crosses are the locations of previously known UDG candidates.
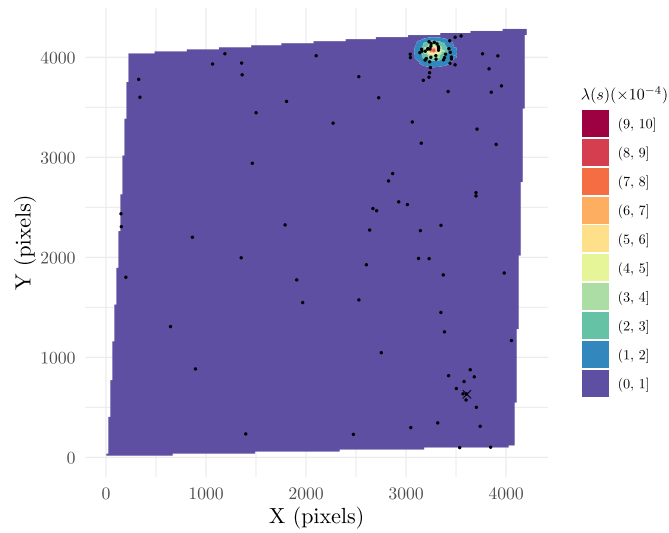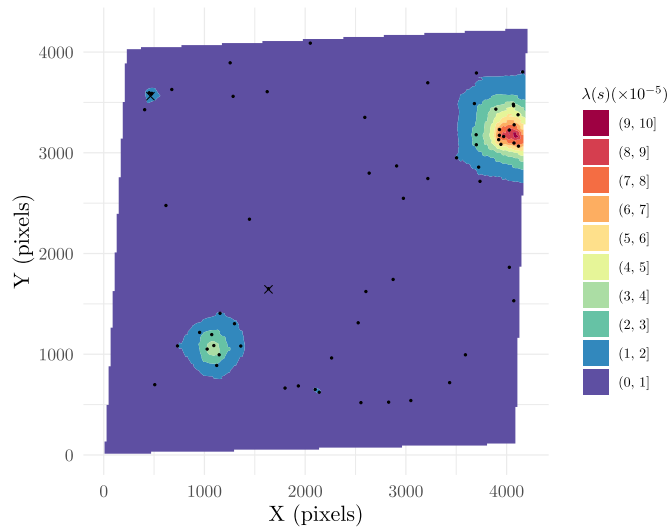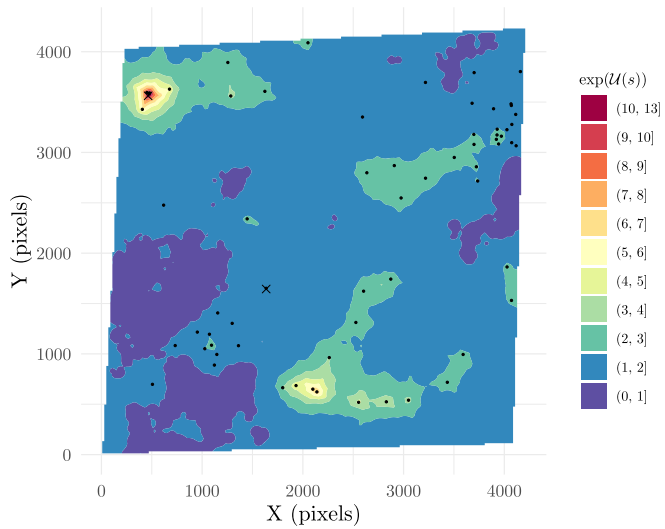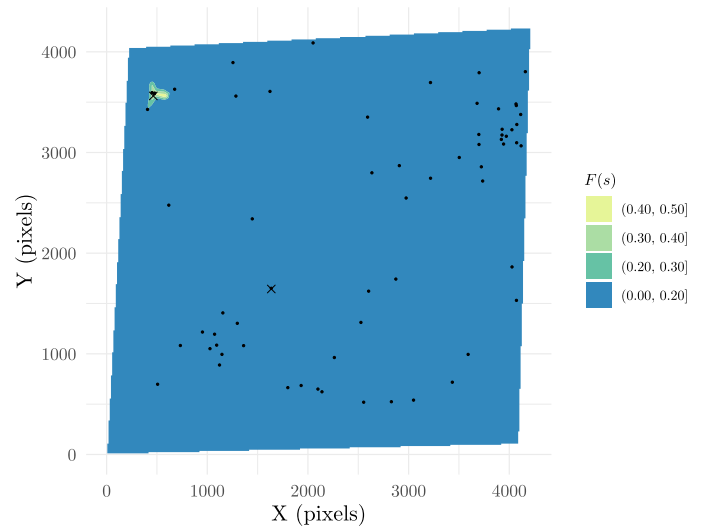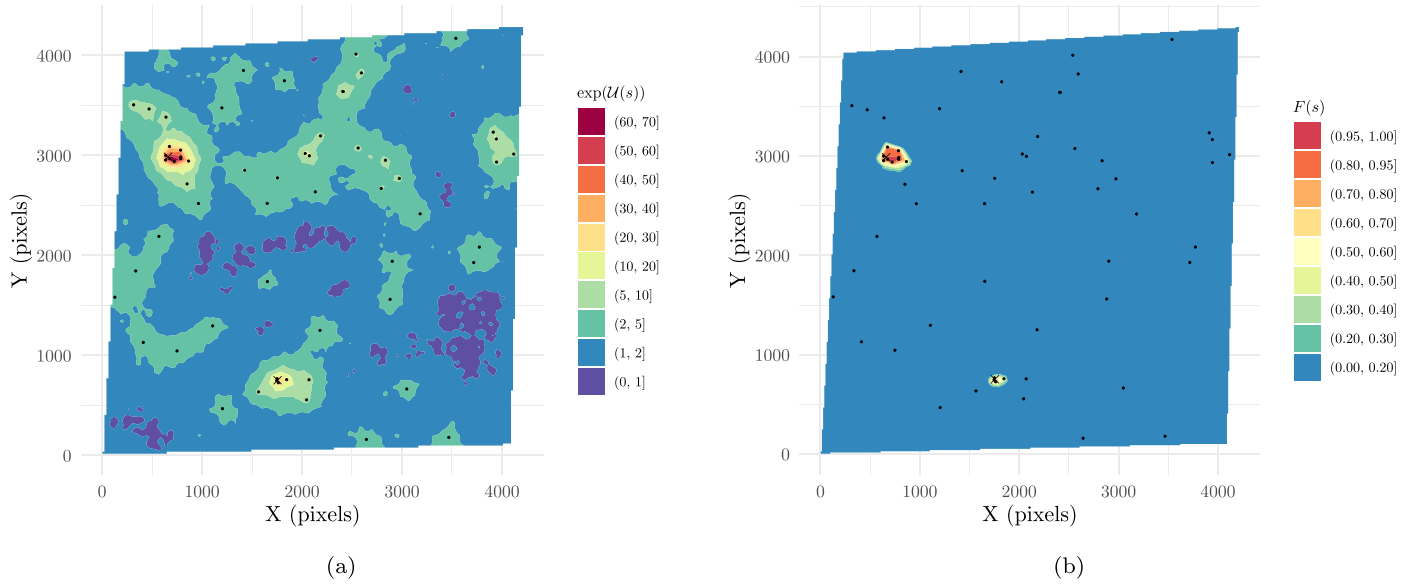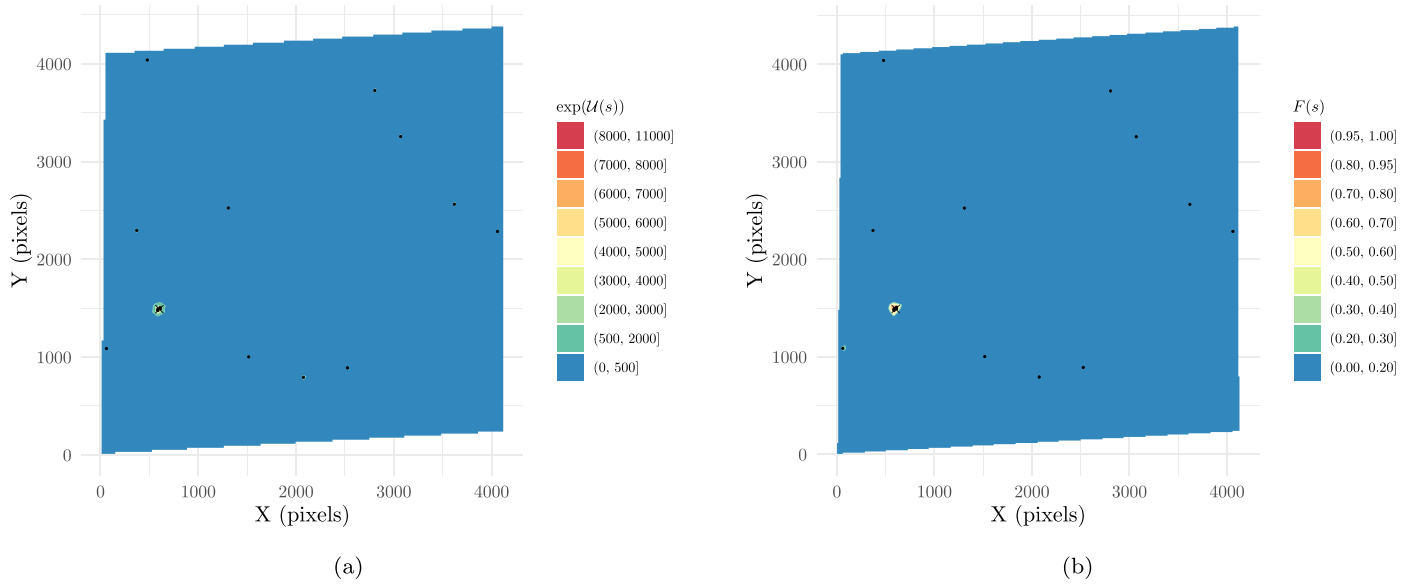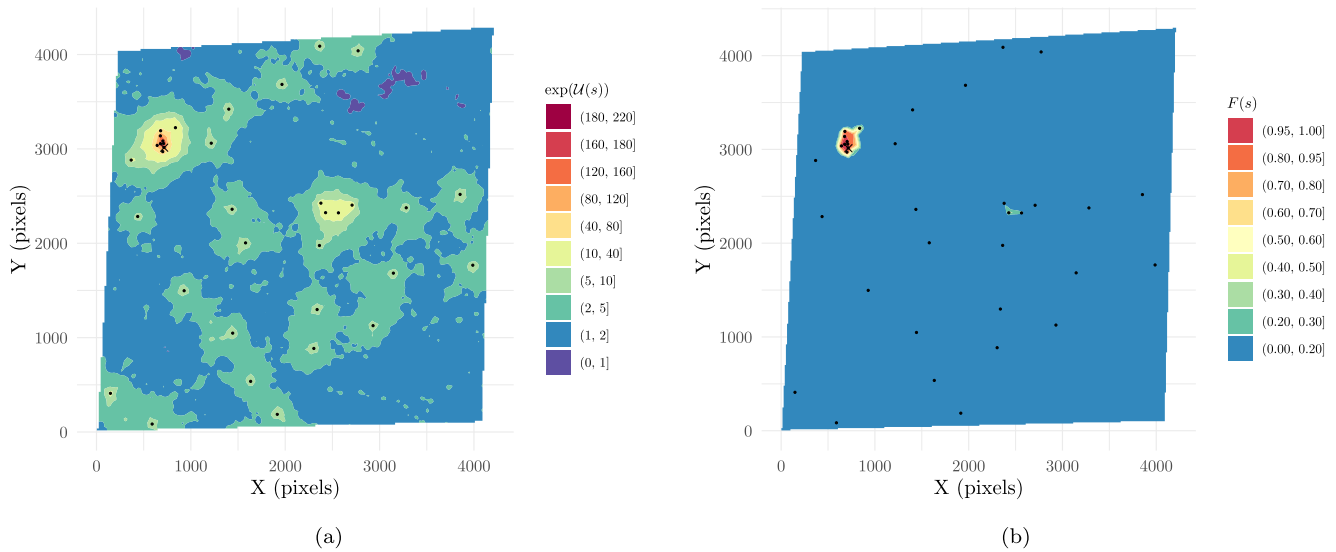
(a)
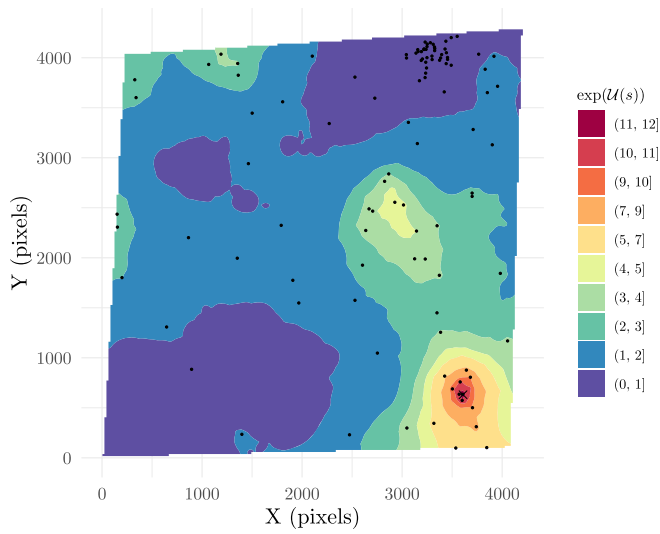


(b)



(c)

**Figure 12.** Posterior results for V10-ACS pointing. (a) Posterior mean intensity $\lambda(s)$; (b) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$; (c) Excursion function $F_C(s)$ with $C = Q(0.95)$. Black points are locations of GC candidates. Black cross is the location of previously known UDG candidate.

(a)



(b)



(c)

**Figure 13.** Posterior results for V15-ACS pointing. (a) Posterior mean intensity $\lambda(s)$; (b) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$; (c) Excursion function $F_C(s)$ with $C = Q(0.95)$. Black points are locations of GC candidates. Black crosses are the locations of previously known UDG candidates.

### C.2. Pointings without Contaminating Bright Galaxies

Figures 14–17 show the detection results for pointings without contaminating bright galaxies.



**Figure 14.** Posterior results for V7-ACS pointing. (a) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$; (b) Excursion function $F_C(s)$ with $C = Q(0.95)$. Black points are locations of GC candidates. Black cross is the location of previously known UDG candidate.



**Figure 15.** Posterior results for V7-WFC3 pointing. (a) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$; (b) Excursion function $F_C(s)$ with $C = Q(0.95)$. Black points are locations of GC candidates. Black cross is the location of previously known UDG candidate.

**Figure 16.** Posterior results for V9-ACS pointing. (a) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$; (b) Excursion function $F_C(s)$ with $C = Q(0.95)$. Black points are locations of GC candidates. Black cross is the location of previously known UDG candidate.
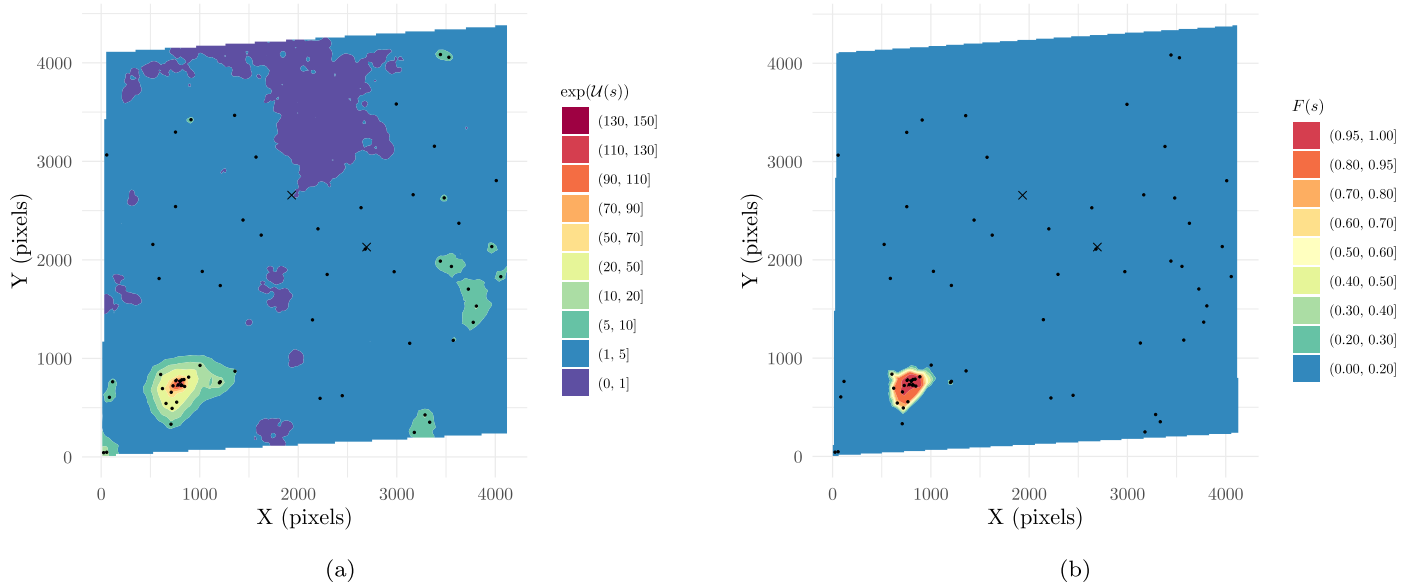


**Figure 17.** Posterior results for V10-WFC3 pointing. (a) Posterior mean spatial random effect $\exp(\mathcal{U}(s))$; (b) Excursion function $F_C(s)$ with $C = Q(0.95)$. Black points are locations of GC candidates. Black crosses are the locations of previously known UDG candidates.

# Appendix D
## Prior Distribution Parameter for $R_L$

Table 4 shows the prior distribution specifications for the effective radii of the GC systems of contaminating bright galaxies.

**Table 4**
Galaxy ID, Coordinates, Visit Pointings, Half-light Radius $R_e$, and Mean of Prior Distribution for $R_k$

| ID | R.A.<br>(J2000) | Decl.<br>(J2000) | Pointing | $R_e$<br>(kpc) | Mean $R_k (\times R_e)$<br>(kpc) |
|---|---|---|---|---|---|
| IC 312 | 03 18 08.41 | +41 45 15.75 | V8-ACS | 2.0 | 10.0 (5×) |
| SDSS J031813.09+414809.0 | 03 18 13.20 | +41 48 09.17 | V8-ACS | 1.3 | 3.9 (3×) |
| LEDA 12319[a] | 03 18 38.91 | +41 40 08.19 | V8-WFC3 | 2.2 | 11.0 (5×) |
| LEDA 12315 | 03 18 35.39 | +41 40 01.03 | V8-WFC3 | 1.0 | 3.0 (3×) |
| [BM99]−317[b] | 03 20 19.11 | +41 42 42.90 | V10-ACS | 3.4 | 10.2 (3×) |
| PGC 012437 | 03 19 52.97 | +41 18 08.52 | V11-ACS | 2.6 | 13.0 (5×) |
| UGC 02673 | 03 20 01.64 | +41 15 04.15 | V11-ACS | 3.3 | 16.5 (5×) |
| PCC 3164 | 03 18 52.25 | +41 16 02.94 | V15-ACS | 3.5 | 10.5 (3×) |
| LEDA 12356 | 03 19 04.30 | +41 13 59.00 | V15-ACS | 2.0 | 10.0 (5×) |

**Notes.** Galaxies listed here are all contaminating galaxies in pointings with detected UDGs.
[a] This galaxy is a dE but a simple inspection by eye shows its GC system is much more extended than its light profile. Hence, a five times ratio is adopted.
[b] [BM99]-NNN catalog was obtained by Brunzendorf & Meusinger (1999).

## ORCID iDs

Dayi David Li (李大一) ⓘ https://orcid.org/0000-0002-5478-3966

Gwendolyn M. Eadie ⓘ https://orcid.org/0000-0003-3734-8177

Roberto Abraham ⓘ https://orcid.org/0000-0002-4542-921X

William E. Harris ⓘ https://orcid.org/0000-0001-8762-5772

Steven R. Janssens ⓘ https://orcid.org/0000-0003-0327-3322

Aaron J. Romanowsky ⓘ https://orcid.org/0000-0003-2473-0369

Pieter van Dokkum ⓘ https://orcid.org/0000-0002-8282-9888

Shany Danieli ⓘ https://orcid.org/0000-0002-1841-2252

## References

Abraham, R. G., & van Dokkum, P. G. 2014, PASP, 126, 55

Amorisco, N. C., Monachesi, A., Agnello, A., & White, S. D. M. 2018, MNRAS, 475, 4235

Bachl, F. E., Lindgren, F., Borchers, D. L., & Illian, J. B. 2019, Methods Ecol. Evol., 10, 760

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Binggeli, B., Sandage, A., & Tarenghi, M. 1984, AJ, 89, 64

Blakeslee, J. P. 1997, ApJL, 481, L59

Bolin, D., & Lindgren, F. 2015, J.R. Statist. Soc. B, 77, 85

Bolin, D., & Lindgren, F. 2018, J. Stat. Softw., 86, 1

Brunzendorf, J., & Meusinger, H. 1999, A&AS, 139, 141

Burkert, A. 2020, ApJ, 904, 161

Burkert, A., & Forbes, D. A. 2020, AJ, 159, 56

Coles, P., & Jones, B. 1991, MNRAS, 248, 1

Danieli, S., van Dokkum, P., Abraham, R., et al. 2020, ApJL, 895, L4

Danieli, S., van Dokkum, P., Conroy, C., Abraham, R., & Romanowsky, A. J. 2019, ApJL, 874, L12

Danieli, S., van Dokkum, P., Trujillo-Gomez, S., et al. 2022, ApJL, 927, L28

Diggle, P. 1985, J.R. Statist. Soc. C, 34, 138

Diggle, P., Rowlinson, B., & Su, T.-l. 2005, Environmetrics, 16, 423

Diggle, P. J., Moraga, P., Rowlinson, B., & Taylor, B. M. 2013, StaSc, 28, 542

Eadie, G. M., Harris, W. E., & Springford, A. 2022, ApJ, 926, 162

Fawcett, T. 2006, PaReL, 27, 861

Forbes, D. A., Alabi, A., Romanowsky, A. J., Brodie, J. P., & Arimoto, N. 2020a, MNRAS, 492, 4874

Forbes, D. A., Dullo, B. T., Gannon, J., et al. 2020b, MNRAS, 494, 5293

Forbes, D. A., & Forbes, A. D. 2017, MNRAS, 472, L104

Forbes, D. A., Gannon, J., Couch, W. J., et al. 2019, A&A, 626, A66

Forbes, D. A., Read, J. I., Gieles, M., & Collins, M. L. M. 2018, MNRAS, 481, 5592

Fuglstad, G.-A., Simpson, D., Lindgren, F., & Rue, H. 2018, J. Am. Stat. Assoc., 114, 445

Gannon, J. S., Forbes, D. A., Romanowsky, A. J., et al. 2021, MNRAS, 510, 946

Georgiev, I. Y., Puzia, T. H., Goudfrooij, P., & Hilker, M. 2010, MNRAS, 406, 1967

Gómez-Rubio, V., & Rue, H. 2017, Stat. Comput., 28, 1033

Greco, J. P., Greene, J. E., Price-Whelan, A. M., et al. 2018, PASJ, 70, 19

Gudehus, D. 1995, A&A, 302, 21

Harris, W. E. 1991, ARA&A, 29, 543

Harris, W. E., Blakeslee, J. P., & Harris, G. L. H. 2017, ApJ, 836, 67

Harris, W. E., Blakeslee, J. P., Whitmore, B. C., et al. 2016, ApJ, 817, 58

Harris, W. E., Brown, R. A., Durrell, P. R., et al. 2020, ApJ, 890, 105

Harris, W. E., Harris, G. L. H., & Alessi, M. 2013, ApJ, 772, 82

Hu, W., Barkana, R., & Gruzinov, A. 2000, PhRvL, 85, 1158

Hudson, M. J., Lucey, J. R., Smith, R. J., et al. 1997, MNRAS, 291, 488

Hui, L., Ostriker, J. P., Tremaine, S., & Witten, E. 2017, PhRvD, 95, 043541

Janssens, S. R., Abraham, R., Brodie, J., Forbes, D. A., & Romanowsky, A. J. 2019, ApJ, 887, 92

King, I. 1962, AJ, 67, 471

Larsen, S. S. 1999, A&AS, 139, 393

Li, D., & Barmby, P. 2021, MNRAS, 501, 3472

Li, Y., Brown, P., Gesink, D. C., & Rue, H. 2012, Stat. Methods Med. Res., 21, 479

Lim, S., Côté, P., Peng, E. W., et al. 2020, ApJ, 899, 69

Lim, S., Peng, E. W., Côté, P., et al. 2018, ApJ, 862, 82

Lindgren, F., Rue, H., & Lindström, J. 2011, J.R. Statist. Soc. B, 73, 423

Martínez-Delgado, D., Läsker, R., Sharina, M., et al. 2016, AJ, 151, 96

Mas, J. F., Filho, B. S., Pontius, R. G., Gutiérrez, M. F., & Rodrigues, H. 2013, IJGI, 2, 869

Mei, S., Blakeslee, J. P., Cote, P., et al. 2007, ApJ, 655, 144

Møller, J. 2003, Adv. Appl. Probab., 35, 614

Møller, J., Syversveen, A. R., & Waagepetersen, R. P. 1998, Scand. J. Stat., 25, 451

Neyman, J., & Scott, E. L. 1958, J.R. Statist. Soc. B, 20, 1

Peng, E. W., & Lim, S. 2016, ApJL, 822, L31

Peng, E. W., Jordán, A., Côté, P., et al. 2006, ApJ, 639, 95

R Core Team 2021, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/

Raynaud, X., & Nunan, N. 2014, PLoSO, 9, e87217

Román, J., Beasley, M. A., Ruiz-Lara, T., & Valls-Gabaud, D. 2019, MNRAS, 486, 823

Rue, H., Martino, S., & Chopin, N. 2009, J.R. Statist. Soc. B, 71, 319

Saifollahi, T., Zaritsky, D., Trujillo, I., et al. 2022, MNRAS, 511, 4633

Samartsidis, P., Eickhoff, C. R., Eickhoff, S. B., et al. 2019, J.R. Statist. Soc. C, 68, 217

Serra, L., Saez, M., Mateu, J., et al. 2013, Environ. Ecol. Stat., 21, 531

Sérsic, J. L. 1963, BAAA, 6, 41

Shen, Z., Danieli, S., van Dokkum, P., et al. 2021, ApJL, 914, L12

Simpson, D., Illian, J. B., Lindgren, F., Sørbye, S. H., & Rue, H. 2015, Biometrika, 103, 49

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. 2017, StaSc, 32, 1

Spitler, L. R., & Forbes, D. A. 2009, MNRAS, 392, L1

Tempel, E., Stoica, R. S., Kipper, R., & Saar, E. 2016, A&C, 16, 17

van Dokkum, P., Abraham, R., Brodie, J., et al. 2016, ApJL, 828, L6

van Dokkum, P., Abraham, R., Romanowsky, A. J., et al. 2017, ApJL, 844, L11

van Dokkum, P., Cohen, Y., Danieli, S., et al. 2018b, ApJL, 856, L30

van Dokkum, P., Danieli, S., Cohen, Y., et al. 2018a, Natur, 555, 629

van Dokkum, P., Shen, Z., Keim, M. A., et al. 2022, Natur, 605, 435

van Dokkum, P., Wasserman, A., Danieli, S., et al. 2019, ApJ, 880, 91

van Dokkum, P. G., Abraham, R., Merritt, A., et al. 2015, ApJL, 798, 45

Venables, W. N., & Ripley, B. D. 2002, Modern Applied Statistics with S (4th ed.; New York: Springer)

Walker, M. G., & Peñarrubia, J. 2011, ApJ, 742, 20

Wang, Q., Peng, E. W., Blakeslee, J. P., et al. 2013, ApJ, 769, 145

Wasserman, A., van Dokkum, P., Romanowsky, A. J., et al. 2019, ApJ, 885, 155

Wickham, H., Averick, M., Bryan, J., et al. 2019, JOSS, 4, 1686

Wittmann, C., Kotulla, R., Lisker, T., et al. 2019, ApJS, 245, 10

Wittmann, C., Lisker, T., Tilahun, L. A., et al. 2017, MNRAS, 470, 1512

Yagi, M., Koda, J., Komiyama, Y., & Yamanoi, H. 2016, ApJS, 225, 11