A model for multiproperty galaxy cluster statistics

August E. Evrard,[1,2,3]* Pablo Arnault,[1,4] Dragan Huterer[1] and Arya Farahi[1]

[1]*Department of Physics and Michigan Center for Theoretical Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[2]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[3]*Institut d'Astrophysique, 98bis Bd. Arago, F-75012 Paris, France*
[4]*École Normale Supérieure de Cachan, 61 Avenue du Président Wilson, F-94230 Cachan, France*

ABSTRACT

The massive dark matter haloes that host groups and clusters of galaxies have observable properties that appear to be lognormally distributed about power-law mean scaling relations in halo mass. Coupling this assumption with either quadratic or cubic approximations to the mass function in log space, we derive closed-form expressions for the space density of haloes as a function of multiple observables as well as forms for the low-order moments of properties of observable-selected samples. Using a Tinker mass function in a Λ cold dark matter cosmology, we show that the cubic analytic model reproduces results obtained from direct, numerical convolution at the 10 per cent level or better over nearly the full range of observables covered by current observations and for redshifts extending to $z = 1.5$. The model provides an efficient framework for estimating effects arising from selection and covariance among observable properties in survey samples.

Key words: galaxies: clusters: general – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Counts of galaxy clusters provide constraints on cosmological parameters (e.g. Voit 2005; Allen, Evrard & Mantz 2011, and references therein), and test fundamental theories of gravity and cosmic acceleration (e.g. Weinberg et al. 2013). Such studies typically use cluster samples identified via optical (Rozo et al. 2010), X-ray (Henry et al. 2009; Mantz et al. 2010), or thermal Sunyaev–Zel'dovich (SZ; Benson et al. 2013; Planck Collaboration 2013; Sievers et al. 2013) signatures of the baryons in the haloes that host cluster phenomena. These analyses are empowered by simulation studies that calibrate the space density as a function of halo mass, known as the *mass function*, within a given cosmology (Tinker et al. 2008; Bhattacharya et al. 2011; Murray, Power & Robotham 2013).

Modelling the expected counts of clusters in a wide-area observational survey requires combining the mass function with a statistical model that expresses the likelihood for a halo of mass M at redshift z to have some intrinsic observable signal, S, detectable in the survey. Evidence from observations (Arnaud, Pointecouteau & Pratt 2005; Maughan 2007; Pratt et al. 2009; Vikhlinin et al. 2009; Zhang et al. 2011; Ettori 2013; Ruel et al. 2013; Saliwanchik et al. 2013; Maughan 2014) and simulations (Evrard et al. 2008; Stanek et al. 2010; Fabjan et al. 2011; Munari et al. 2013; Jiang et al. 2014; Le Brun et al. 2013; Biffi et al. 2014) support a model in which the

scaling law behaviour is power law with mass in the mean, with approximately lognormal variance.

While scaling behaviour of cluster properties has been studied for decades (see Giodini et al. 2013, for a recent review), most works have focused on correlating pairs of observed signals, $\{S_2, S_1\}$, or on studying how a single observable scales with mass, $\{S_1, M\}$. Simulations provide a natural environment for the latter, since the true halo mass is known. For observations, mass estimates are made indirectly from measured signals, for example through assumption of virial or hydrostatic equilibrium, and this methodology introduces the potential for bias and additional variance that must be calibrated (e.g. Rasia et al. 2012; Battaglia et al. 2013; Nelson et al. 2014). Alternatively, masses can be inferred through inversion of a given observable–mass relation. In this way, an observable serves as proxy for halo mass.

Evidence of biases in mass proxies can arise from comparisons among different observable signals. *Planck* satellite measurements of the thermal SZ effect in the optically selected `maxBCG` sample (Planck Collaboration 2011) led to a detailed re-examination of X-ray, SZ, and optical scaling relations by Rozo et al. (2014a,b, 2014c). That study concluded that the *Planck* Y_{SZ} mass calibration was biased low by a few tens of per cent, a finding supported by independent weak gravitational lensing estimates of *Planck* clusters (von der Linden et al. 2014), although other studies are less supportive (Israel et al. 2014).

Rozo et al. (2014b) present a model for multivariate signal counts and other statistics under the assumption of a locally power-law

mass function. That model was employed to interpret a combined set of X-ray, SZ, and optical data, resulting in a set of preferred scaling relations presented in Rozo et al. (2014c, see their table 4).

In this paper, we present a non-local extension of that model that expands its scope to effectively cover the complete dynamic range of properties displayed by the population of galaxy clusters. Within a Λ cold dark matter (ΛCDM) cosmology, we show that the mass function of the massive haloes that host groups and clusters of galaxies can be represented by a low-order polynomial (in log space) to a typical accuracy of a few per cent, comparable to its calibrated level of precision from N-body simulations. Convolving this mass function representation with a multivariate Gaussian of logarithmic halo properties at fixed mass and redshift results in analytic expressions for the space density as a function of multiple observables and other derivative statistics. By offering a fast method for estimating survey sample and follow-up study outcomes, this formalism is intended to complement data analysis methods based on similar model assumptions (Maughan 2014).

We employ a halo mass convention of M_{500c}, the mass within a spherical region encompassing 500 times the critical density, $\rho_c(z)$, but the analytic expressions can be applied using any choice of halo mass convention. We use M_{500c} to be consistent with the scaling laws presented in Rozo et al. (2014c). In Section 2, we derive expressions for multi-observable cluster population statistics using low-order polynomial approximations of the mass function in log space. These are then applied to X-ray and SZ statistics in Section 3. In Section 4, we discuss some of the model's strengths and limitations, and we summarize our results in Section 5.

2 MODEL FOR LOW-ORDER MOMENTS

We first develop expressions for the space density of clusters as a function of observables such as temperature or luminosity, then compute first and second moments for properties of samples selected by a specific observable. Our model expands the results presented in Rozo et al. (2014b) and uses slightly different notation.

Consider a set of N bulk observable properties S_a where $a \in \{1, N\}$; these observables can be, for example, X-ray luminosity, L_X, temperature, T_X, gas thermal energy measured in the X-ray or SZ flux, Y_X or Y_{SZ}, number of galaxies, N_{gal} or λ, inferred lensing mass M_{lens}, etc., measured within some characteristic radius. Let $s_a \equiv \ln(S_a)$ represent the natural logarithms of these properties in some chosen basis of units (for example, 10^{44} erg s^{-1} for X-ray luminosity).

We assume power-law forms for the observable–mass scaling relations. Choosing a halo pivot mass scale, M_p, at some fiducial redshift (values discussed below), and letting $\mu \equiv \ln(M/M_p)$, the vector of log-observables, $\boldsymbol{s}$, scales in the mean with mass as

$$\langle \boldsymbol{s}|\mu \rangle = \boldsymbol{\pi} + \boldsymbol{\alpha}\mu , \tag{1}$$

where the vectors $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$ are the normalizations and slopes of the relevant scaling laws. We consider redshift-dependent normalizations, $\boldsymbol{\pi}(z)$, which scale in a self-similar manner (Böhringer, Dolag & Chon 2012). While the slopes may also be redshift dependent, we take them as constant here.

Individual haloes drawn from the cosmic population deviate from this mean behaviour in a manner that we assume is lognormal. At all redshifts, the full probability density function (PDF), $P(\boldsymbol{s}|\mu)$, is described by a covariance matrix with elements $C_{ab} = \langle(s_a - \langle s_a|\mu\rangle)(s_b - \langle s_b|\mu\rangle)\rangle$, and where $C_{aa} = \sigma_a^2$ is the intrinsic log variance of the ath observable. We assume this covariance to be independent of mass and redshift.

2.1 First-order (local) mass function

The model of Rozo et al. (2014b) uses a first-order Taylor expansion of the mass function, $n(\ln M)$ (with dimension of number density per $\ln(M)$), around some pivot mass M_p,

$$n_1(\mu) = A e^{-\beta_1 \mu} , \tag{2}$$

where A and β_1 are the local amplitude and (negative) slope of the mass function evaluated at the pivot, $\mu = 0$. Note that A and β_1 are functions of redshift, cosmology, and pivot location, as explained in Section 3. The subscript on the space density indicates the first-order nature of the mass function expansion.

Using equations (1) and (2), Bayes theorem in the form $P(\mu|\boldsymbol{s}) = P(\boldsymbol{s}|\mu)P(\mu)/P(\boldsymbol{s})$ allows us to obtain the mean and variance of the log mass selected by a fixed combination of observables. In the first-order approximation to the mass function, this probability is Gaussian, with mean and variance

$$\langle \mu|\boldsymbol{s} \rangle_1 = \left[\boldsymbol{\alpha}^T \boldsymbol{C}^{-1}(\boldsymbol{s} - \boldsymbol{\pi}) - \beta_1\right]\sigma_{\mu|s,1}^2 , \tag{3}$$

$$\sigma_{\mu|s,1}^2 = (\boldsymbol{\alpha}^T \boldsymbol{C}^{-1}\boldsymbol{\alpha})^{-1} . \tag{4}$$

In the case of a single observable quantity s_a, these expressions reduce to $\langle \mu|s_a \rangle_1 = (s_a - \pi_a)/\alpha_a - \beta_1\sigma_{\mu|a,1}^2$ and $\sigma_{\mu|a,1}^2 = (\sigma_a/\alpha_a)^2$. The mean mass is biased low relative to the assumed scaling by an amount given by the product of the local slope of the mass function and the mass variance of the chosen observable.

Space density of multiple observables. Convolving equation (2) with the lognormal likelihood, $P(\boldsymbol{s}|\mu)$ yields the halo number density as a function of the full vector of observable properties as

$$n_1(\boldsymbol{s}) = A_1' \exp\left[-\frac{1}{2}\left((\boldsymbol{s} - \boldsymbol{\pi})^T \boldsymbol{C}^{-1}(\boldsymbol{s} - \boldsymbol{\pi}) - \frac{\langle \mu|\boldsymbol{s}\rangle_1^2}{\sigma_{\mu|s,1}^2}\right)\right] ,$$

$$A_1' = A\, \sigma_{\mu|s,1}\left((2\pi)^{N-1}|\boldsymbol{C}|\right)^{-1/2} . \tag{5}$$

Observable-selected samples. Now consider selecting a sample by a certain observable, s_a. With the full space density above, we can derive the PDF of any second observable, s_b, by $P(s_b|s_a) = n(s_a, s_b)/n(s_a)$. The result is also Gaussian with mean and variance as

$$\langle s_b|s_a \rangle_1 = \pi_b + \alpha_b[\langle \mu|s_a\rangle_1 + \beta_1 r_{ab}\sigma_{\mu|a,1}\sigma_{\mu|b,1}] , \tag{6}$$

$$\sigma_{b|a,1}^2 = \alpha_b^2\left[\sigma_{\mu|a,1}^2 + \sigma_{\mu|b,1}^2 - 2r_{ab}\sigma_{\mu|a,1}\sigma_{\mu|b,1}\right] , \tag{7}$$

where r_{ab} is the correlation coefficient between properties s_a and s_b at fixed mass. When the intrinsic correlation of these observables is non-zero, then a shift in the mean of s_b is induced that is similar in form to the bias of equation (3) but with opposite sign if r_{ab} is positive. This effect can be understood by the fact that the dominant lower mass haloes that scatter upwards into the chosen s_a bin will also have a positive deviation from the mean s_b if r_{ab} is positive. If r_{ab} is negative, the effect is reversed.

Along with these purely observable properties, one can compute the correlation coefficient between mass and s_b for samples chosen by s_a, finding

$$r_{(\mu b|a),1} = \frac{\sigma_{\mu|a,1}/\sigma_{\mu|b,1} - r_{ab}}{[1 - r_{ab}^2 + (\sigma_{\mu|a,1}/\sigma_{\mu|b,1} - r_{ab})^2]^{1/2}} . \tag{8}$$

In the case of uncorrelated observables ($r_{ab} = 0$), a positive correlation between mass and s_b is induced by the fact that haloes with lower mass that scattered up into the s_a bin will also have lower s_b, and vice-versa. In the limit that the selection property, s_a, is a much better mass proxy than s_b such that $\sigma_{\mu|a,1}/\sigma_{\mu|b,1} \simeq 0$, then the

correlation of s_b and mass takes on the opposite sign of the intrinsic correlation, r_{ab}. Relative to the mean behaviour in the selection bin, haloes of lower mass, $\Delta\mu < 0$, will have positively enhanced selection signal at that mass, $\Delta s_a > 0$, and then $\Delta s_b \simeq r_{ab}\Delta s_a \simeq -r_{ab}\Delta\mu$.

2.2 Higher order (non-local) mass functions

The first-order model functions well over a relatively narrow range in mass or observable near the chosen pivot point. We now wish to extend the range of the model by introducing quadratic and cubic terms into the mass function approximation. We derive here exact expressions for the quadratic case, and approximate expressions for the cubic case, and show below that the latter are accurate to better than 10 per cent for a wide range of halo mass scales and redshifts.

The second-order model uses a mass function

$$n_2(\mu) = Ae^{-\beta_1\mu - \frac{1}{2}\beta_2\mu^2}, \tag{9}$$

where β_2 is the magnitude of the second derivative of the mass function at the pivot mass scale, which is negative for massive haloes in ΛCDM cosmologies.

The convolution remains analytic, and the halo number density as a function of multiple observables retains the form of equation (5), but adding local curvature reduces the weight of lower mass haloes scattered upwards into the signal bin. Consequently, the Gaussian distribution of halo mass at fixed observable properties has a compressed mean and variance relative to the first-order treatment as

$$\langle\mu|s\rangle_2 = x_s\,\langle\mu|s\rangle_1, \tag{10}$$

$$\sigma^2_{\mu|s,2} = x_s\,\sigma^2_{\mu|s,1}, \tag{11}$$

where the compression factor, $x_s \equiv (1 + \beta_2\,\sigma^2_{\mu|s,1})^{-1}$, is less than unity and is well approximated by $1 - \beta_2\sigma^2_{\mu|s,1}$ for most of the applications discussed below.

Consider again the case of two observables s_a and s_b. The PDF of having observable s_b in a population selected by observable s_a also remains Gaussian, and the expressions written in mass equivalents are somewhat simpler. Letting $\delta_i = (s_i - \pi_i)/\alpha_i$, then the mean and variance for the second-order mass function approximation are

$$\langle\delta_b|s_a\rangle_2 = x_a\left[\langle\mu|s_a\rangle_1 + (\beta_1 + \beta_2\delta_a)\,r_{ab}\,\sigma_{\mu|a,1}\,\sigma_{\mu|b,1}\right], \tag{12}$$

$$\frac{\sigma^2_{b|a,2}}{\alpha_b^2} = x_a\left[\frac{\sigma^2_{b|a,1}}{\alpha_b^2} + \beta_2\,\sigma^2_{\mu|a,1}\,\sigma^2_{\mu|b,1}\,(1 - r^2_{ab})\right]. \tag{13}$$

The first expression indicates that the mean observable now senses the curvature in the mass function through the $\beta_2\delta_a$ term, where recall that δ_a is measuring the equivalent log-mass distance from the pivot location.

In the limit of uncorrelated observables ($r_{ab} = 0$), the second expression reduces to $x_a\sigma^2_{\mu|a,1} + \sigma^2_{\mu|b,1}$, as it should since the mass function curvature affects the mass variance in the selection variable but not that of the non-selection variable.

Finally, the correlation coefficient between mass and property s_b at fixed s_a is now given by

$$r_{(\mu b|a),2} = \frac{\sigma_{\mu|a,1}/\sigma_{\mu|b,1} - r_{ab}}{[(1 - r^2_{ab})/x_a + (\sigma_{\mu|a,1}/\sigma_{\mu|b,1} - r_{ab})^2]^{1/2}}. \tag{14}$$

For uncorrelated observables, this expression again reduces to equation (8) with $\sigma^2_{\mu|a,2}$ replacing $\sigma^2_{\mu|a,1}$.

In the appendix, we show a further extension to the third order, with $n_3(\mu) = Ae^{-\beta_1\mu - \frac{1}{2}\beta_2\mu^2 - \frac{1}{6}\beta_3\mu^3}$.

3 APPLICATIONS TO OBSERVABLE CLUSTER PROPERTIES

We now evaluate the utility of the above expressions by comparing their predictions to expectations calculated via explicit local convolution of the Tinker mass function (Tinker et al. 2008). For hot gas observables, we examine the X-ray luminosity at soft-photon energies, L_X, and the total gas thermal energy as determined by the thermal SZ, Y_{SZ}. These cases represent examples of relatively high- and low-scatter mass proxies, respectively. We also examine the case optical richness, N_{gal}, a relatively high-scatter proxy for which the correlation with hot gas properties at fixed mass is currently poorly understood.

We perform analysis at $z = 0.23$, the redshift where local mass–observable relations used here are calibrated, and also at $z = 1.5$. The higher redshift is chosen to be representative of the outer reaches of near-term cluster surveys and is also an epoch at which the mass function is both steeper and more strongly curved compared to low redshift, aspects that make the higher order corrections more important.

3.1 Mass function and log-space polynomial fits

The Tinker mass function employs an updated version of the normalized mass fraction functional, $f(\sigma)$, in the form originally calibrated for CDM cosmologies by Sheth, Mo & Tormen (2001) and Jenkins et al. (2001) as

$$n_{\text{Tinker}}(\mu, z) = \frac{dn}{d\mu} = \frac{\overline{\rho}_m(z)}{M}\frac{d\ln\sigma^{-1}(M, z)}{d\mu}f(\sigma), \tag{15}$$

where $\overline{\rho}_m(z)$ is the mean cosmic matter density and $\sigma^2(M, z)$ is the linearly evolved variance of matter density fluctuations filtered on a mass scale $M \equiv M_p e^\mu$, both evaluated at redshift z. In this work, we use the mass function calculator published by Murray et al. (2013) and employ the Tinker fit for $f(\sigma)$ (Tinker et al. 2008) using the CAMB transfer function of a *Wilkinson Microwave Anisotropy Probe* 7 (*WMAP*7) cosmology. The cosmological parameters for the ΛCDM model are: scaled Hubble constant, $h = 0.704$, baryon, CDM, and dark energy parameters of $\Omega_b = 0.0455$, $\Omega_c = 0.226$, $\Omega_{DE} = 0.728$, respectively, spectral index, $n_s = 0.967$, and present amplitude of matter density fluctuations, $\sigma_8 = 0.81$.

At our two fiducial redshifts, we compute the coefficients A, β_1, β_2, and β_3 by taking numerical derivatives of $n_{\text{Tinker}}(\mu, z)$ at the chosen pivot mass. Values of the pivot mass, M_p, in units of 10^{14} M$_\odot$ along with the fit parameters are shown in Table 1. At higher redshifts, the mass function steepens and becomes more strongly curved. For observables with large mass variance, the second-order compression factor, $1 + \beta_2\sigma^2_{\mu|s}$, can be important since $\beta_2 = 0.70$ and 1.22 at redshifts $z = 0.23$ and 1.5, respectively.

Table 1. Mass function expansion parameters at two redshifts for a *WMAP*7 cosmology. The pivot mass, M_p, is in units of 10^{14} M$_\odot$ and the amplitude, A, is in units of 10^{-6} Mpc^{-3}.

z	M_p	A	β_1	β_2	β_3
0.23	2.0	1.944	1.97	0.70	0.40
1.5	1.0	0.293	3.07	1.20	0.73

Table 2. Observable–mass scaling parameters at $z = 0.23$ and pivot mass, $M_p = 2 \times 10^{14}\,M_\odot$, from Rozo et al. (2014c). See text for unit definitions. Along with the published scatter values for N_{gal}, we also consider a smaller value in parentheses based on the λ richness estimator (Rykoff et al. 2012).

S	$S_p = e^{\pi_s}$	α_s	$\sigma_{\ln S}$	$\sigma_{\mu|s}$
L_X	0.61	1.55	0.39	0.252
$D_A^2 Y_{SZ}$	0.62	1.71	0.15	0.088
N_{gal}	37	0.94	0.42 (0.21)	0.45 (0.23)

3.2 Scalings for L_X, Y_{SZ}, and N_{gal}

As specific examples of scaling laws we use results derived in the panchromatic study of Rozo et al. (2014c). In that work, L_X is the X-ray luminosity in the rest-frame [0.1, 2.4] keV band, expressed in units of $10^{44}\,\mathrm{erg\,s^{-1}}$ and measured within a cylindrical aperture of radius R_{500c}. The thermal SZ signal, $D_A^2 Y_{SZ}$, where D_A is the angular distance of the source, is given in units of $10^{-5}\,\mathrm{Mpc^2}$ and is measured within the spherical aperture of radius R_{500c}. The optical richness, N_{gal}, is the count of red sequence galaxies determined by the `maxBCG` cluster-finding algorithm within an estimated sphere of radius R_{200c} (Koester et al. 2007).

We use the observable–mass parameters, π_s, α_s, and $\sigma_{\ln S}$, at $z = 0.23$ given in table 4 of Rozo et al. (2014c). The parameters, with normalizations rescaled to our choice of pivot mass, $2 \times 10^{14}\,M_\odot$, are summarized in Table 2. At $z = 1.5$, we employ self-similar scalings for normalizations of L_X and Y_{SZ} discussed below.

For the optical richness, we use a simple inversion of the scaling of lensing mass at fixed N_{gal} and consider both the published value of the scatter as well as a more optimistic value that is appropriate for the multicolour richness estimator, λ (Rykoff et al. 2012; Rozo & Rykoff 2014). The improvement of a factor of 4 in variance has significant implications that we illustrate below.

3.3 Application to cluster number counts

We first compare the approximate analytical formulae, $n_i(\ln S)$, to $n_{Tinker}(\ln S)$, where the latter quantity is obtained by explicit convolution of the Tinker mass function, $\int n_{Tinker}(\mu) P(\ln S|\mu)\mathrm{d}\mu$.

In Fig. 1, we show the X-ray luminosity function at redshifts $z = 0.23$ and 1.5, derived using the parameters in Tables 1 and 2. The normalization of the high-redshift L_X–M relation is scaled using the assumption that the soft-band luminosity at fixed mass scale in a self-similar fashion, $L(z) \sim H(z)^2$ (Böhringer et al. 2012). While the first-order model is locally accurate, the second- and third-order models extend the accuracy over increasingly wider ranges of X-ray luminosity. The third-order model traces well the local Tinker convolution results across more than two decades in luminosity.

In Fig. 2, we evaluate the accuracy of the different orders for all observables listed in Table 2. Values of the properties range from 0.03 to 100 for $L_X/10^{44}\,\mathrm{erg\,s^{-1}}$ and $D_A^2 Y_{SZ}/10^{-5}\,\mathrm{Mpc^2}$ and from 3 to 1000 for N_{gal}.

At first order, the first-order counts always overestimate the Tinker-convolved estimates. The normalization error is approximately $1 - \sqrt{x_s} \simeq \beta_2 \sigma_{\mu|a,1}^2/2$. The first-order approximation is thus most accurate for Y_{SZ} at low redshift and worst for N_{gal} at high redshift. With $x_s = 0.80$, the first-order N_{gal} counts at $z = 1.5$ lie more than 10 per cent above the Tinker convolution, even near the pivot point. When the optical richness scatter is reduced by a factor of 2 (dotted grey lines), the first order improve dramatically. In all
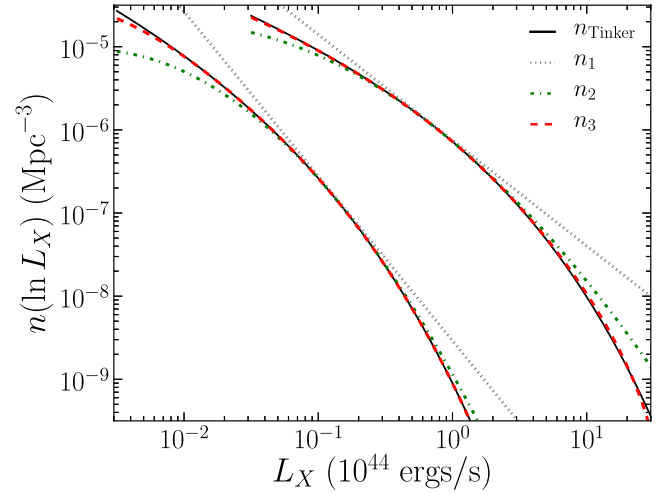

Figure 1. Comoving space density of haloes as a function of L_X evaluated at redshift $z = 0.23$ (upper curves) and 1.5 (lower), assuming self-similar redshift evolution of the X-ray normalization for the latter. Black lines show expectations from local convolution of the Tinker mass function, while the dotted (grey), dot–dashed (green), and dashed (red) curves show our analytic expressions based on first-, second-, and third-order mass function expansions, respectively. Pivot masses of $M_p = 2$ and $1 \times 10^{14}\,M_\odot$ are used at $z = 0.23$ and 1.5, respectively, and the latter are offset by a factor of 0.1 in L_X for clarity. Deviations of the approximations from the direct Tinker convolution are plotted in Fig. 2.

cases, the range within which the first-order estimates are accurate is limited to a factor of a few close to the pivot location.

At second order, the dynamic range over which the counts lie within 5 per cent of the Tinker expectations widens, reaching nearly a decade for Y_{SZ} and L_X at low redshift. The approximation is always better at low redshifts where the second and third derivatives, β_2 and β_3, of the mass function are lower. Note that the zero crossing is shifted upwards from the pivot location by an amount that scales with the signal variance.

The third-order approximation is accurate to within 5 per cent in number over more than two decades in both L_X and Y_{SZ}, with the larger errors only occurring at high signal values that correspond to M_{500} masses above 1 and $0.5 \times 10^{15}\,M_\odot$ at low and high redshift, respectively. Such massive haloes are quite rare, with space densities of roughly 20 and $0.2\,\mathrm{Gpc^{-3}}$ at $z = 0.23$ and 1.5. For statistical cluster samples that typically reach space densities many times higher, the third-order model is capable of yielding highly accurate estimates of counts as a function of observable properties.

One can always vary the pivot mass scale in order to improve the quality fit over a certain range of signal. However, even with a tuned pivot, the first- and second-order models cannot reach 10 per cent accuracy over the wide dynamic range of observables covered by current surveys. In particular, unless the pivot mass scale is chose to be very high, it is very difficult for these models to obtain good accuracy at large masses and high redshifts.

3.4 Observable-selected sample expectations

Dedicated follow-up observations or joint studies of overlapping surveys at different wavelengths can allow multiple properties to be measured for clusters selected by a particular observable. Here, we explore expectations for such secondary properties based on the different orders of the multivariate model of Section 2.
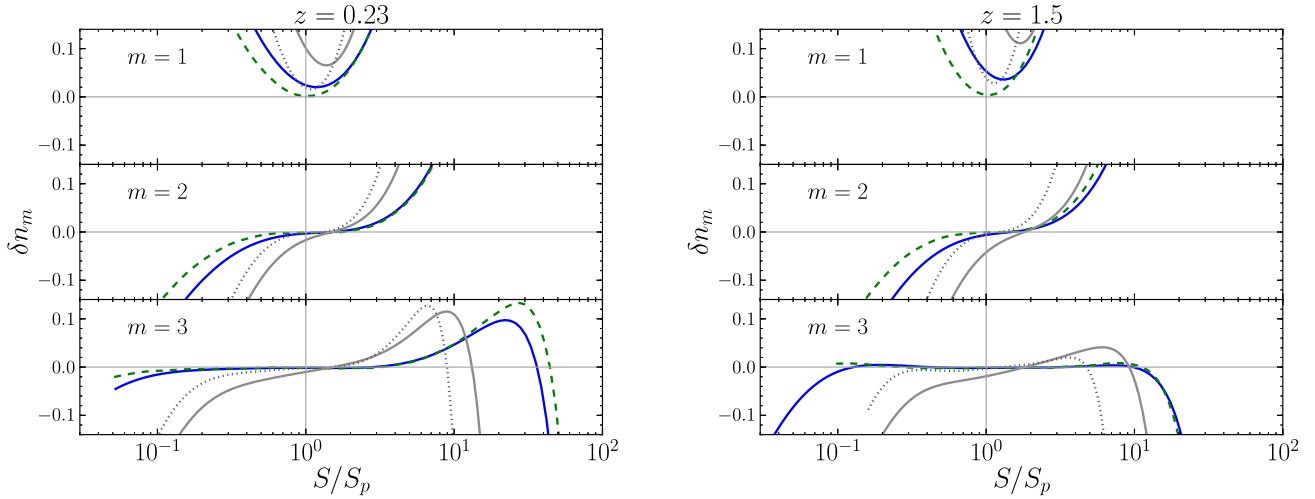
Figure 2. Fractional error in number counts, $\delta n_m(\ln S) \equiv n_m(\ln S)/n_{\text{Tinker}}(\ln S) - 1$, are shown for first-, second-, and third-order expansions of the log-space mass function (top to bottom) at $z = 0.23$ (left) and 1.5 (right). The observable signals S represented are X-ray luminosity, L_X (blue solid lines), the SZ flux, $D_A^2 Y_{SZ}$ (green dashed), and the optical richness, N_{gal}, using either the high-scatter value, $\sigma_\mu = 0.45$ (grey solid) or low value, $\sigma_\mu = 0.23$ (grey dotted). Signals are plotted relative to the log-mean values, S_p, expected at the pivot mass scales of 2 and 1×10^{14} M$_\odot$ (left and right, respectively) using parameters in Table 2.
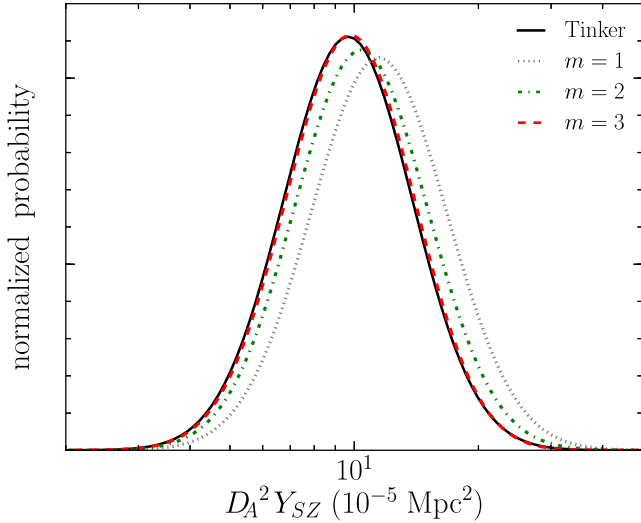
Figure 3. PDF, $P(\ln S_b | S_a)$, for the SZ signal, $S_b = D_A^2 Y_{SZ}$, at a chosen X-ray luminosity, $S_a = L_X = 10^{45}$ erg s^{-1} at $z = 0.23$ as inferred from direct convolution of the Tinker mass function (black line) or from the approximate model using first- (grey dotted), second- (green dot–dashed), and third-order (red dashed) forms.

Fig. 3 shows an example for low redshift ($z = 0.23$) in the form of the SZ thermal decrement expected for clusters selected at fixed X-ray luminosity. We chose a relatively bright luminosity of $L_X = 10^{45}$ erg s^{-1} appropriate for a mass scale of 1.2×10^{15} M$_\odot$, a factor of 6 above the low-redshift pivot point. The correlation coefficient, r, between $\ln(D_A^2 Y_{SZ})$ and $\ln(L_X)$ is expected to be positive since both signals scale as positive powers of the intracluster gas mass and temperature (Stanek et al. 2010; Angulo et al. 2012). We take $r = 0.5$ as an example value.

In Fig. 3, the first-order likelihood overestimates the mean by ~20 per cent. This relatively small error arises from near-cancellation of the much larger errors made in both $n_1(\ln L_X, \ln D_A^2 Y_{SZ})$ and $n_1(\ln L_X)$ relative to their Tinker values at this high mass. The first-order model is surprisingly accurate for conditional likelihoods.

By sensing the local mass function curvature, the second-order estimate, equation (12), corrects the first-order logarithmic mean by roughly $\alpha_b(x_a - 1) \simeq -0.07$, which reduces, but does not fully eliminate, the discrepancy with the local Tinker convolution. The third-order succeeds in matching the local Tinker probability with high accuracy.

The behaviour of the low-order analytical likelihoods depends primarily on the mass variance of the selection variable and the signal value relative to the pivot location. At the pivot, the first-order estimate, $\langle s_b | s_a \rangle_1$, overestimates the second-order value, $\langle s_b | s_a \rangle_2$, by approximately $\beta_2 \sigma_{\mu|a,1}^2$. At $z = 0.23$, this error is 0.04 for L_X selection, 0.005 for Y_{SZ} selection, and 0.12 for N_{gal} selection. These corrections are smaller than or comparable to the quoted uncertainties in the scaling relation amplitudes derived by Rozo et al. (2014c) using a first-order approach.

Selecting at an X-ray luminosity *below*, rather than above, the pivot luminosity, L_p, results in different behaviour because the corrections at odd and even orders have different sign. Selecting at $L_X = L_p/A$, where $A > 1$, the first-order mean $\ln D_A^2 Y_{SZ}$ lies lower than the Tinker expectation, but the overshoot is smaller in magnitude than for the case of selecting at $L_X = AL_p$. The second-order correction then slightly overcorrects, with the mean *above* the Tinker estimate. The third-order applies a small, negative correction to closely align with the Tinker value.

3.5 Covariance between high-scatter mass proxies

As a final demonstration of the model, we explore the relatively poor mass proxies, N_{gal} and L_X, with $\ln M$ scatter of 0.45 and 0.25, respectively (Table 2). There are currently no theoretical or empirical constraints on the covariance between cold and hot baryons at fixed halo mass and redshift, so there is complete freedom in choosing the correlation coefficient, r, linking their deviations about the mean scaling behaviours.

From equation (12), with the aforementioned mass scatter for the two proxies, the shift in $\langle \ln N_{\text{gal}} | L_X \rangle$ as r is varied from -0.5 to 0.5 will be of the order of 0.11 times the magnitude of the local logarithmic slope of the mass function. At a mass scale of 5×10^{14} M$_\odot$, the latter is of order three, implying a shift of roughly
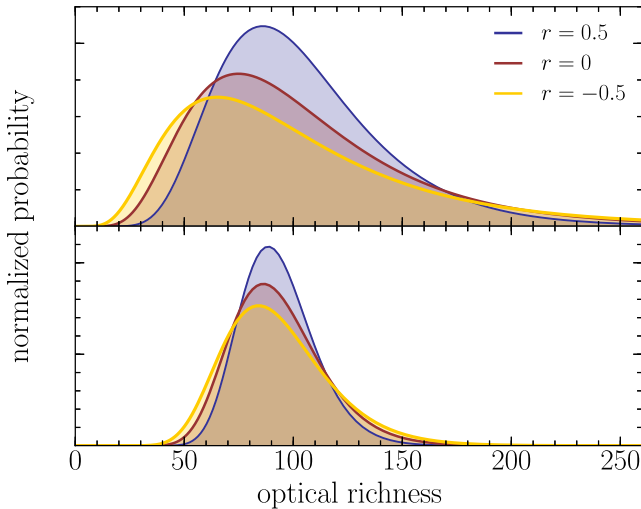
Figure 4. Expectations for optical richness at $z = 0.23$ are demonstrated for two extreme cases. The top panel selects haloes with $L_X = 2.5 \times 10^{44}$ erg s^{-1} and assumes a log-mass scatter at fixed richness of 0.45, appropriate for the original `maxBCG` sample N_{gal} richness estimator. The lower panel selects haloes with $D_A^2 Y_{SZ} = 3 \times 10^{-5}$ Mpc^2 and assumes a log-mass scatter at fixed richness of 0.23, appropriate for the improved λ richness estimator. In both panels, the different PDFs arise from different choices of correlation coefficient between richness and the selection variable, as indicated by the legend. The third-order, log-space approximation to the mass function is used in all calculations.

30 per cent. In addition, the variance, equations (7) and (13), is maximized when the properties strongly anticorrelate while it is minimized as $r \to 1$.

In the top panel of Fig. 4, we show the likelihood of optical richness at $z = 0.23$ for haloes selected to have an X-ray luminosity of 2.5×10^{44} erg s^{-1}, which highlights a mass $M_{500} \simeq 5 \times 10^{14}$ M$_\odot$. The third-order estimator is used to calculate the conditional likelihood, and we show expectations for three discrete values of the correlation coefficient, $r = -0.5, 0, 0.5$.

As anticipated, the large mass scatter in these proxies means that predictions for the mean and variance in N_{gal} shift considerably as r is varied. The modal value of N_{gal} is 62 for $r = -0.5$, increasing to 86 at $r = 0.5$, while the scatter in $\ln N_{gal}$ drops from 0.60 to 0.38 for these two cases. These effects conspire to dramatically change the 2.5σ lower limits for N_{gal}, with values of 14 and 33 for $r = -0.5$ and 0.5, respectively.

Observational data demonstrate large variance between N_{gal} and L_X (e.g. Rykoff et al. 2008; Andreon & Hurn 2010), but a new approach to estimating optical richness from multicolour photometry offers significant improvement. The λ richness estimator (Rykoff et al. 2012) uses a probabilistic membership trained on spectroscopic calibration of the red sequence in multicolour space. This richness estimator shows considerably smaller scatter in L_X compared to N_{gal} and the technique has been extended for use as a cluster finder in photometric surveys (Rykoff et al. 2014). Based on matching clusters found in the SDSS DR-8 sample to known X-ray clusters, Rozo & Rykoff (2014) demonstrate that the implied mass scatter of the λ richness measure is ∼25 per cent, considerably reduced relative to the original N_{gal} richness.

The lower panel of Fig. 4 demonstrates how the use of low-scatter mass proxies can significantly tighten the conditional likelihood of optical richness. Shown is the likelihood, $p(\lambda|Y_{SZ})$ at $z = 0.23$ for $D_A^2 Y_{SZ} = 3 \times 10^{-5}$ Mpc^2, a value that selects roughly the same

mass as the L_X choice used in the top panel. We assume 23 per cent mass scatter in λ and 8.8 per cent mass scatter in Y_{SZ}. Relative to the top panel, the scatter in $P(\lambda|Y_{SZ})$ is reduced by more than a factor of 2, and the sensitivity of $\langle\lambda|Y_{SZ}\rangle$ to the correlation coefficient is also weakened.

This dramatic improvement results from the combination of improved mass selection and tighter mass scaling of λ. Results for λ richness using L_X as the selection variable would lie intermediate to the two panels shown in Fig. 4.

4 DISCUSSION

We have emphasized an application of the model to the massive haloes that host galaxy clusters at late cosmic times, but the mathematical framework is general, so the model could be applied at earlier epochs to describe phenomenology associated with the high-mass end of the mass function. Galaxies and quasars at redshifts of a few or early star formation at redshifts of tens are potential applications. The key requirements are observables or properties that of haloes that scale as power laws with halo mass in the mean, with variability described by a lognormal covariance.

Applied to groups and clusters, the third-order model is essentially global in scope. Compared to local Tinker convolution estimates, the cubic approximation achieves better than 10 per cent accuracy over nearly the whole signal ranges covered by current observations, and for redshifts $z < 1.5$. This level of accuracy is comparable to the current level of systematic uncertainty in the mass function derived from simulations, particularly when the effects of baryon physics are included (Stanek, Rudd & Evrard 2009; Cusworth et al. 2014; Martizzi et al. 2014; Cui, Borgani & Murante 2014).

The pivot location sets the range of accuracy for the lower order approximations. Since the third-order approximation is based on a Taylor expansion of $e^{-\beta_3\mu^3/6} \simeq 1 - \beta_3\mu^3/6$, where $\mu = \ln(M/M_p)$, the model breaks down as $\mu^3 \to 6/\beta_3$. For the pivots chosen here, this occurs only for very rare, massive systems with $M_{500} \gtrsim 10^{15}$ M$_\odot$. To achieve the widest possible dynamic range, the second-order expressions could be interpolated using multiple pivot points, $M_{p,k}$, requiring values of $\beta_1(M_{p,k},z)$ and $\beta_2(M_{p,k},z)$ to be provided. For light-cone applications, these derivatives could be modelled as low-order polynomials in redshift.

The pivot masses we employ at the two demonstration redshifts correspond closely to those satisfying a fixed sky surface density condition, $dN(>M)/dz = $ const., in a ΛCDM cosmology (Evrard et al. 2002; Mortonson, Hu & Huterer 2011). For cluster survey applications, this would seem a natural choice.

While power-law scaling with lognormal covariance is supported for intrinsic properties of clusters, the available evidence is often limited. In particular, covariance among different signals is poorly understood (Maughan 2014) and there are few constraints as to whether the slope and variance of a particular signal's scaling with mass is indeed constant with mass and redshift, as is assumed here (e.g. Balaguera-Antolínez et al. 2012, provided a hint of evidence for curvature in the L_X–M relation). Redshift dependences are easily incorporated into the existing framework by writing the slopes, $\alpha(z)$, intercepts, $\pi(z)$, and covariance, $C(z)$, as explicit functions of z. Extensions to the model that incorporate weak mass dependence in the scaling parameters are also possible.

We interpret the model covariance as reflecting that of intrinsic halo properties, but observed properties of clusters inevitably include some contribution from projection effects and other sources

of noise. If these additional sources combine with intrinsic scatter to produce a lognormal form of property covariance, then model presented here remains directly applicable. The measured covariance matrix in this case will be a mixture of intrinsic and extrinsic sources. Characterizing these effects, particularly SZ and optical projection, deserves further investigation (White, Cohn & Smit 2010; Angulo et al. 2012; Noh & Cohn 2012).

5 SUMMARY

Using polynomial log-space approximations to the high-mass end of the cosmic mass function, we present analytic forms for statistics of multi-observable properties of the haloes that host groups and clusters of galaxies. The model employs scaling laws between observables and mass that are power law in the mean with lognormal covariance.

The model provides closed-form expressions for cluster counts as a function of multiple observables and for conditional likelihoods of observable-selected samples. Comparing to a locally convolved Tinker mass function, we show that the first-order model is generally accurate within a narrow range near the pivot mass, except for very high mass-scatter proxies. The second- and third-order extensions provide increasingly wider coverage in observables irrespective of the mass scatter. The third-order model is nearly global in scope.

The mass variance in a particular observable determines many expected features, as does the covariance between pairs of observables at fixed mass. As multiwavelength surveys and dedicated follow-up campaigns provide increasingly rich, uniform samples of clusters, opportunities to apply this model to better constrain the statistical properties of massive haloes will become apparent. Such knowledge will provide useful constraints on the physical processes that govern baryon evolution in massive haloes.

REFERENCES

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Andreon S., Hurn M. A., 2010, MNRAS, 404, 1922
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Arnaud M., Pointecouteau E., Pratt G. W., 2005, A&A, 441, 893
Balaguera-Antolínez A., Sánchez A. G., Böhringer H., Collins C., 2012, MNRAS, 425, 2244
Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2013, ApJ, 777, 123

Benson B. A. et al., 2013, ApJ, 763, 147
Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122
Biffi V., Sembolini F., De Petris M., Valdarnini R., Yepes G., Gottlöber S., 2014, MNRAS, 439, 588
Böhringer H., Dolag K., Chon G., 2012, A&A, 539, A120
Cui W., Borgani S., Murante G., 2014, preprint (arXiv:1402.1493C)
Cusworth S. J., Kay S. T., Battye R. A., Thomas P. A., 2014, MNRAS, 439, 2485
Ettori S., 2013, MNRAS, 435, 1265
Evrard A. E. et al., 2002, ApJ, 573, 7
Evrard A. E. et al., 2008, ApJ, 672, 122
Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801
Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, Space Sci. Rev., 177, 247
Henry J. P., Evrard A. E., Hoekstra H., Babul A., Mahdavi A., 2009, ApJ, 691, 1307
Israel H., Reiprich T. H., Erben T., Massey R. J., Sarazin C. L., Schneider P., Vikhlinin A., 2014, A&A, 564, A129
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115
Koester B. P. et al., 2007, ApJ, 660, 239
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2013, preprint (arXiv:e-prints)
Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, MNRAS, 406, 1759
Martizzi D., Mohammed I., Teyssier R., Moore B., 2014, MNRAS, 440, 2290
Maughan B. J., 2007, ApJ, 668, 772
Maughan B. J., 2014, MNRAS, 437, 1171
Mortonson M. J., Hu W., Huterer D., 2011, Phys. Rev. D, 83, 023015
Munari E., Biviano A., Borgani S., Murante G., Fabjan D., 2013, MNRAS, 430, 2638
Murray S. G., Power C., Robotham A. S. G., 2013, Astron. Comput., 3, 23
Nelson K., Lau E. T., Nagai D., Rudd D. H., Yu L., 2014, ApJ, 782, 107
Noh Y., Cohn J. D., 2012, MNRAS, 426, 1829
Planck Collaboration, 2011, A&A, 536, A12
Planck Collaboration, 2013, preprint (arXiv:1303.5080P)
Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361
Rasia E. et al., 2012, New J. Phys., 14, 055018
Rozo E., Rykoff E. S., 2014, ApJ, 783, 80R
Rozo E. et al., 2010, ApJ, 708, 645
Rozo E., Rykoff E. S., Bartlett J. G., Evrard A., 2014a, MNRAS, 438, 49
Rozo E., Evrard A. E., Rykoff E. S., Bartlett J. G., 2014b, MNRAS, 438, 62
Rozo E., Bartlett J. G., Evrard A. E., Rykoff E. S., 2014c, MNRAS, 438, 78
Ruel J. et al., 2013, preprint (arXiv:e-prints)
Rykoff E. S. et al., 2008, ApJ, 675, 1106
Rykoff E. S. et al., 2012, ApJ, 746, 178
Rykoff E. S. et al., 2014, ApJ, 785, 104R
Saliwanchik B. R. et al., 2013, preprint (arXiv:e-prints)
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Sievers J. L. et al., 2013, J. Cosmol. Astropart. Phys., 10, 60
Stanek R., Rudd D., Evrard A. E., 2009, MNRAS, 394, L11
Stanek R., Rasia E., Evrard A. E., Pearce F., Gazzola L., 2010, ApJ, 715, 1508
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Vikhlinin A. et al., 2009, ApJ, 692, 1033
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
von der Linden A. et al., 2014, preprint (arXiv:e-prints)
Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87
White M., Cohn J. D., Smit R., 2010, MNRAS, 408, 1818
Zhang Y.-Y., Andernach H., Caretta C. A., Reiprich T. H., Böhringer H., Puchwein E., Sijacki D., Girardi M., 2011, A&A, 526, A105

APPENDIX A: THIRD-ORDER MODEL

A1 Multiproperty halo number density

The accuracy gained in going from $n_1(s)$ to $n_2(s)$ motivates a third-order approach. We thus now consider

$$n_3(\mu) = A e^{-\beta_1 \mu - \frac{1}{2}\beta_2 \mu^2 - \frac{1}{6}\beta_3 \mu^3}, \tag{A1}$$

where $\beta_3 = -[(\mathrm{d}^3/\mathrm{d}\mu^3)\ln n_{\mathrm{exact}}](\mu = 0) > 0$. We could not find a closed-form solution at this order, so we instead consider the approximation that limits the mass range to be near enough to the pivot point so that

$$n_3(\mu) \simeq n_2(\mu)\left(1 - \frac{1}{6}\beta_3 \mu^3\right). \tag{A2}$$

Convolving this approximated form of $n_3(\mu)$ with $P(s|\mu)$ yields

$$n_3(s) = n_2(s)\left\{1 - \frac{\beta_3}{2}\left(\sigma_{\mu|s,2}^2\langle\mu|s\rangle_2 + \frac{1}{3}\langle\mu|s\rangle_2^3\right)\right\}. \tag{A3}$$

One can see clearly that the signal range will be limited from above by the requirement that the space density be non-negative.

A2 Selecting on an observable property

We have $P_3(s_b|s_a) = n_3(s_a, s_b)/n_3(s_a)$, which, using the previous formula, gives

$$P_3(s_b|s_a) = P_2(s_b|s_a)$$
$$\times \frac{1 - \frac{\beta_3}{2}\left(\sigma_{(\mu|a,b),2}^2\langle\mu|s_a, s_b\rangle_2 + \frac{1}{3}\langle\mu|s_a, s_b\rangle_2^3\right)}{1 - \frac{\beta_3}{2}\left(\sigma_{\mu|a,2}^2\langle\mu|s_a\rangle_2 + \frac{1}{3}\langle\mu|s_a\rangle_2^3\right)}. \tag{A4}$$

We can calculate analytically the mean and variance of $P_3(s_b|s_a)$, which gives

$$\langle s_b|s_a\rangle_3 = K_1 J_1 - K_2 C(J_2 + D J_1)$$
$$- K_3 C^3(J_4 + 3D J_3 + 3D^2 J_2 + D^3 J_1) \tag{A5}$$

$$\sigma_{b|a,3}^2 = K_1 J_2 - K_2 C(J_3 + D J_2)$$
$$- K_3 C^3(J_5 + 3D J_4 + 3D^2 J_3 + D^3 J_2) - \langle s_b|s_a\rangle_3^2, \tag{A6}$$

where

$$\begin{cases} J_n = \int_{-\infty}^{+\infty} s_b^n P_2(s_b|s_a)\mathrm{d}s_b \tag{A7} \\[2mm] \quad = \frac{1}{\sqrt{\pi}}\sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k}M^{n-2k}(2\sigma^2)^k\Gamma(k+1/2) \tag{A8} \\[2mm] M = \langle s_b|s_a\rangle_2 \tag{A9} \\[2mm] \sigma = \sigma_{b|a,2} \tag{A10} \end{cases}$$

(in particular : $J_0 = 1, J_1 = M, J_2 = \sigma^2 + M^2$)
and

$$\begin{cases} C = \frac{X\,\sigma_{(\mu|a,b),1}}{1 - r_{ab}^2} \tag{A11} \\[2mm] D = \frac{1}{X}\left[Y - X\boldsymbol{\pi}_b - \beta_1\left(1 - r_{ab}^2\right)\right] \tag{A12} \\[2mm] X = \frac{\alpha_b}{\sigma_b^2} - \frac{r_{ab}\alpha_a}{\sigma_a\sigma_b} \tag{A13} \\[2mm] Y = (s_a - \boldsymbol{\pi}_a)\left(\frac{\alpha_a}{\sigma_a^2} - \frac{r_{ab}\alpha_b}{\sigma_a\sigma_b}\right) \tag{A14} \end{cases}$$

$$\begin{cases} K_1 = \left[1 - \frac{\beta_3}{2}\left(\sigma_{\mu|s,2}^2\langle\mu|s\rangle_2 + \frac{1}{3}\langle\mu|s\rangle_2^3\right)\right]^{-1} \tag{A15} \\[2mm] K_2 = K_1\frac{\beta_3}{2}\sigma_{(\mu|a,b),1}\,x_{a,b}^2 \tag{A16} \\[2mm] K_3 = K_1\frac{\beta_3}{6}\,x_{a,b}^3. \tag{A17} \end{cases}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.