Power spectrum estimation methods on intracluster medium surface brightness fluctuations

M. Bishop[1,*], Y. Perrott[1], T. Parashar[1], and S. Oughton[2]

[1]Victoria University of Wellington, New Zealand
[2]University of Waikato, New Zealand

Abstract.
Accurate estimation of galaxy cluster masses is a central problem in cosmology. Turbulence is believed to introduce significant deviations from the hydrostatic mass estimates. Estimation of turbulence properties is complicated by projection of the 3D cluster onto the 2D plane of the sky, and is commonly done in the form of indirect probes from fluctuations in the X-ray surface brightness and Sunyaev-Zeldovich effect maps. In this paper, we address this problem using simulations. We examine different methods for estimating the power spectrum on 2D projected fluctuation data, emulating data projected onto a 2D plane of the sky, and comparing them to the original, expected 3D power spectrum. Noise can contaminate the power spectrum of ICM observations, so we also briefly compare a few methods of reducing noise in the images for better spectral estimation.

1 Introduction

The total mass of a galaxy cluster is one of its most fundamental properties. Estimations of this mass often rely on the assumption of hydrostatic equilibrium (HSE), an approach with shortcomings since the intracluster medium (ICM) is continuously disturbed by mergers, feedback processes, and motions of galaxies. These processes generate gas motions that contribute nonthermal pressure, typically associated with turbulence, that leads to an underestimation of the mass by as much as 30% [2]. We can measure turbulence through indirect probes that come in the form of fluctuations in the X-ray surface brightness and Sunyaev-Zeldovich (SZ) effect maps. ICM observations via the X-ray and SZ effect give projected (along the line-of-sight ℓ, onto the 2D sky-plane vector coordinate $\boldsymbol{\theta}$) versions of the density and pressure, respectively:

$$I_X(\boldsymbol{\theta}) \propto \int n_e(\boldsymbol{\theta}, \ell)^2 \Lambda(T) \, d\ell \; ; \quad Y_{SZ}(\boldsymbol{\theta}) \propto \int P_e(\boldsymbol{\theta}, \ell) \, d\ell,$$

with some temperature dependent cooling function $\Lambda(T)$ for the X-ray emissivity [3]. The fluctuations can be obtained using Reynolds decomposition, *i.e.* $A = A_0 + \delta A$ where $A_0 = \langle A \rangle$ is a large scale (mean) estimate and δA is the turbulent fluctuations. Typically, $I_{X,0} = \langle I_X \rangle$ and $Y_{SZ,0} = \langle Y_{SZ} \rangle$ are represented by a circular or elliptical model [4–6]. Subtracting the mean profile from the observed image provides the surface brightness fluctuations ($\delta I_X, \delta Y_{SZ}$),

---

*e-mail: mark.bishop@vuw.ac.nz

which have zero mean. These fluctuations are interpreted as turbulence-induced departures from the large-scale HSE model.

It is common to analyze turbulence fluctuations using a power spectrum, i.e., by decomposing the energy into contributions at each scale. There are many different techniques for accomplishing this and here we implement and compare some of those employed in the intracluster/interstellar medium, solar wind, and general turbulence literature. The first section will briefly discuss the different spectrum estimation methods and compare their results on projected (simulated and synthetic) fluctuation data. As observations contain noise contributions that interfere with power spectrum estimates, we also look at different techniques for reducing the noise contribution so as to obtain improved spectral estimates. This proceeding serves as an initial exploration into these issues. A more thorough examination and description will be published at a later date.

2 Power Spectra: Comparison of methods
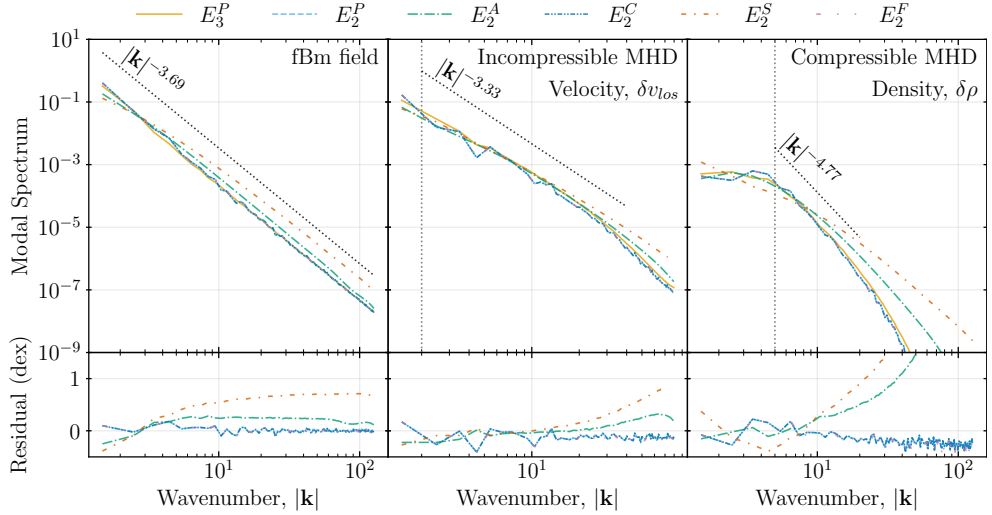
Figure 1. Comparison of the 3D (periodogram method) power spectrum, E_3^P, to 2D projected power spectra, E_2^X, obtained using the periodogram, Arévalo, correlogram, structure function, and flatsky methods, for three different datasets. Left: Using a fractional Brownian motion field. Middle: The *z*-axis (line-of-sight) component of the velocity fluctuations of an incompressible MHD simulation. Right: The density fluctuations of a compressible MHD simulation. Note that the periodogram, correlogram, and flatsky methods are equivalent.

We use synthetic data including fields generated using fractional Brownian motion (fBm) and data from 3D (incompressible and compressible) MHD simulations. Fractional Brownian motion fields are *not* turbulent but have a similar statistical structure to turbulent fields. The astrophysical observations can provide information about density [4], pressure [5, 6], and velocity [7–9]. The synthetic data that we use represent these variables. The fBm field could represent any of these. We use the density fluctuations from the compressible MHD simulation. For the incompressible MHD simulation, we use the line-of-sight component of the velocity (v_{los}).

	Method	MSE	(dex)	MAPE	(dex)
fBm	Periodogram	$5.32 \cdot 10^{-5}$	$(1.35 \cdot 10^{-3})$	$5.92 \cdot 10^{-2}$	$(6.95 \cdot 10^{-3})$
	Arevalo	$1.59 \cdot 10^{-4}$	$(3.96 \cdot 10^{-2})$	$5.67 \cdot 10^{-1}$	$(3.84 \cdot 10^{-2})$
	SF	$2.90 \cdot 10^{-4}$	$(4.52 \cdot 10^{-1})$	$3.71 \cdot 10^{0}$	$(1.16 \cdot 10^{-1})$
IMHD	Periodogram	$3.93 \cdot 10^{-5}$	$(1.97 \cdot 10^{-2})$	$2.49 \cdot 10^{-1}$	$(2.91 \cdot 10^{-2})$
	Arevalo	$2.75 \cdot 10^{-5}$	$(4.11 \cdot 10^{-2})$	$5.17 \cdot 10^{-1}$	$(3.33 \cdot 10^{-2})$
	SF	$3.92 \cdot 10^{-5}$	$(2.51 \cdot 10^{-1})$	$2.17 \cdot 10^{0}$	$(7.66 \cdot 10^{-2})$
CMHD	Periodogram	$1.35 \cdot 10^{-9}$	$(5.56 \cdot 10^{-2})$	$3.90 \cdot 10^{-1}$	$(2.27 \cdot 10^{-2})$
	Arevalo	$2.66 \cdot 10^{-10}$	$(5.39 \cdot 10^{0})$	$1.38 \cdot 10^{3}$	$(1.73 \cdot 10^{-1})$
	SF	$4.62 \cdot 10^{-9}$	$(1.22 \cdot 10^{1})$	$6.21 \cdot 10^{4}$	$(2.70 \cdot 10^{-1})$

Table 1. The MSE and MAPE statistical measures of the accuracy of the 2D power spectra versus the true 3D power spectrum for the different spectrum estimation methods shown in figure 1: a fBm field, an incompressible MHD simulation (IMHD), and a compressible MHD simulation (CMHD).

For each synthetic dataset we calculate the 3D modal spectrum [16] averaged over spherical surfaces at radius $|\mathbf{k}|$, $E_3^P(|\mathbf{k}|)$. This is commonly called the power spectrum in astronomical literature [1]. These are used as the 'ground truth' to compare against various 2D spectra of the fluctuations projected along the line-of-sight. The 2D modal spectra are averaged over circular annuli of radius $|\mathbf{k_\theta}|$. The five spectral estimation methods we test are:

Periodogram: The periodogram spectral estimate is simply the squared magnitude of the Fourier transform of the data [10], denoted with the superscript P, E^P.

Correlogram: The correlogram spectral estimator (E^C) takes the Fourier transform of the auto-correlation function (called Blackman-Tukey if a window function is applied) [10, 11]. The Weiner-Khinchin theorem states that this and the periodogram give the same result.

Arévalo: A scale-space decomposition via Gaussian convolutions, with the standard deviation of the Gaussian approximating the scale, can be used to estimate the power spectrum (E^A) [12]. This is the most common method used in the ICM turbulence literature [4, 6, 13].

Structure function: A structure function [14] describes the moments of a lagged difference and loosely represents the variance as a function of a length scale (the lag ℓ). A crude approximation can turn the second order structure function into an equivalent power spectrum (E^S) as a function of equivalent wavenumber ($k \equiv 2\pi/\ell$) [14, 15].

Spherical harmonics: Given that the ICM is observed on the sky-sphere, a spectral decomposition using spherical harmonics is natural. The small size of the cluster allows for a flat-sky approximation [5] to estimate the power spectrum (E^F).

Figure 1 shows the true, averaged modal spectrum for the fluctuations, E_3^P, and the above five 2D proxies for E_3^P. These are calculated for: (left) a 3D fBm field with a spectral slope of $|\mathbf{k_3}|^{-11/3}$, where $|\mathbf{k_3}|$ is the magnitude of the 3D wavevector, (middle) an incompressible MHD simulation, and (right) a compressible MHD simulation. The plots at the bottom of the figure 1 show the residual (difference) in units of order of magnitude (dex) of the 2D projected power spectra (scaled by a factor of $\Delta k/2\pi$ to account for the difference in dimensions) compared to the true 3D spectrum.

The projection-slice theorem states that the Fourier transform of the 2D projected data should be a slice in the Fourier transform of the original 3D data (*i.e.* the line-of-sight wavenumber $k_\ell = 0$). Under the assumption of isotropy, this implies same modal spectra for 3D and 2D projected. The Fourier methods—periodogram, correlogram, and flatsky— yield the same results: their 2D and 3D spectra are equivalent. Evidently, this is not the

case for the structure function and Arévalo spectral estimates. These two methods obtain an appropriate powerlaw for the fBm field, and for the inertial range in the incompressible and compressible MHD simulations (indicated by the black dotted lines in Fig. 1). The Arévalo spectral estimate has a known bias for a pure powerlaw spectrum that can be corrected for, if the powerlaw of the 3D spectrum is already known [6, 12]. This bias does not necessarily account for the order of magnitude difference seen in the compressible MHD case beyond the inertial range. This has interesting implications for the capabilities of the Arévalo method to actually capture the dissipation scale [13]. Similarly, the structure function method can obtain reasonably correct powerlaws, but has significant bias in the fBm field, and compressible MHD beyond the inertial range.

To further quantify the accuracy of these methods, we compute the mean squared error (MSE) and mean absolute percentage error (MAPE) for each 2D power spectrum ($E_2^X(|\mathbf{k_\theta}|)$). These are listed in table 1.

$$\text{MSE} = \left\langle \left(E_2^X(|\mathbf{k_\theta}|) - E_3^P(|\mathbf{k_3}|) \right)^2 \right\rangle; \quad \text{MAPE} = \left\langle \left| \frac{E_2^X(|\mathbf{k_\theta}|) - E_3^P(|\mathbf{k_3}|)}{E_3^P(|\mathbf{k_3}|)} \right| \right\rangle$$

where the angle brackets $\langle \cdot \rangle$, represent the average over available wavenumbers. We also calculate the MSE and MAPE in units of *dex*, which has the same formula except we take the base-10 log of $E_2^X(|\mathbf{k_\theta}|)$ and $E_3^P(|\mathbf{k_3}|)$. As indicated by figure 1, periodogram works the best, followed by Arévalo. The structure function equivalent spectrum has the largest error, especially for wavenumbers in the dissipation range of the MHD simulations.

3 Noise reduction methods
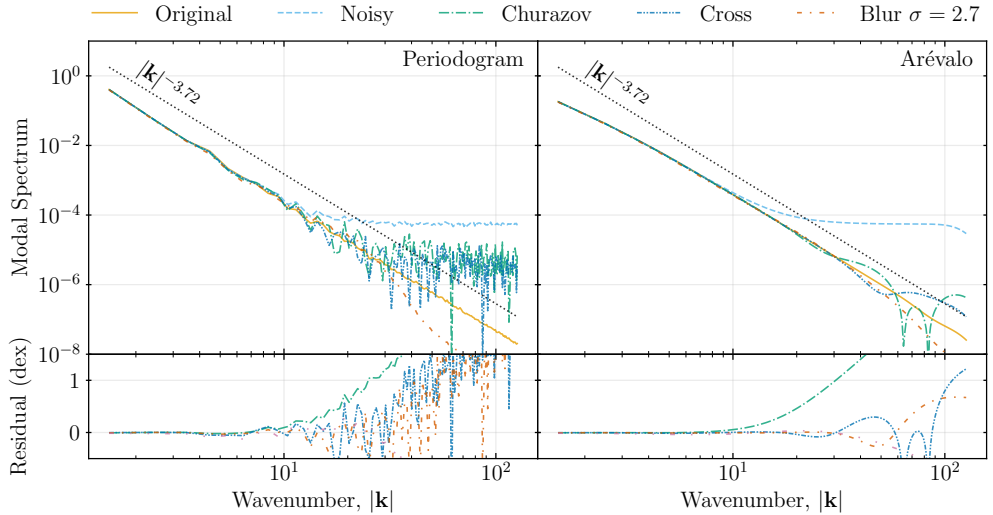
Figure 2. Comparison of the original (noiseless) fBm spectrum to the version with added Poisson noise, and the versions obtained using three different noise removal methods. Spectra are calculated using the periodogram (left) and Arévalo (right) methods.

We now address the issue of Poisson noise, which is typically introduced by the counting of photons by the telescope detectors. We *apply* Poisson noise to a (noiseless) fluctuation

image by treating each pixel as the mean (and therefore, the variance) of the Poisson distribution to sample from. The noise reduction methods described below can be applied to any method described in section 2; here we only show results for the periodogram and Arévalo methods. We test the following noise reduction methods:

Churazov et al: From a noisy observation (with a single realization of the Poisson noise), we can generate additional realizations by sampling each (noisy) pixel as the mean and variance of the Poisson distribution. This method requires generating many (we use 100) additional realizations of the noisy image and averaging their power spectrum. The difference between the original noisy spectrum and the realization-averaged power spectrum gives an estimate of the (constant) noise floor. Subtracting the noise floor from the original noisy spectrum obtains an estimate on the noiseless spectrum [4].

Cross spectrum: Data are split into two observations A and B; either by independent observations, or splitting one observation cube into two parts, and the cross-power spectrum is computed. It retains the spectral properties of features common to the two observations whereas the *independent* noise realizations cancel out. This is equivalent to the 'jack-knife' method: $(A + B)/2$ contains the data of interest and $(A - B)/2$ contains the noise component [17]. Cross spectra can be computed for all the methods in section 2 (see [6] for an Arévalo cross-spectrum method).

Gaussian blur: Lastly, a simple Gaussian blur with some selected scale, ideally one at which the signal to noise ratio (SNR) drops to a predetermined value, can be used to remove noise. This blurs the fluctuations below the selected scale, and retains structures at larger scales. However, the choice of an appropriate blurring scale (Gaussian σ) is difficult without knowledge of the true spectrum. For the case shown here, we iterate over different scales and select the value that best reconstructs the original spectrum.

Amongst these methods, the cross-spectrum and blurring methods are the simplest to implement and understand. We apply all methods to the fractional Brownian motion fields with the goal of recreating the original 2D (projected) spectrum. We generate two realizations of the noise onto the original image for use with the cross-spectrum. The results are shown in figure 2 and table 2, where we show the periodogram and Arévalo methods of power spectrum estimation (left and right respectively).

For both the periodogram and Arévalo methods, the noisy spectra are essentially equal to the original spectrum, until the spectral level drops to that of the constant noise floor, as is expected. All noise reduction methods tested are able to recover more of the original spectral range compared than is available in the noisy spectrum. The cross spectrum and Gaussian smoothing techniques seem to work best in spite of their simplicity. For the Arévalo-produced spectra both these methods yield comparable results, while for periodogram-produced spectra Gaussian smoothing produces a smoother end result.

4 Conclusion

We compare different spectrum calculation techniques, constructed in very different ways, with different pros and cons. The differences in the outputs of these methods could potentially introduce biases in physical interpretations. There is a clear advantage to the Fourier methods (periodogram, correlogram, and flatsky), but these methods can behave poorly when data is masked, which is common for astrophysical observations. Arévalo and structure function methods are more robust to missing data but struggle to capture the dissipation/kinetic range of the spectrum. Simple minded noise reduction techniques appear to work reasonably well in extending the spectral range below the noise floor. Future work will expand on the information presented in this proceeding.

	Method	MSE	(dex)	MAPE	(dex)
P.gram	Noisy	$6.27 \cdot 10^{-8}$	$(5.47 \cdot 10^{0})$	$6.08 \cdot 10^{2}$	$(3.07 \cdot 10^{-1})$
	Churazov	$9.63 \cdot 10^{-9}$	$(2.20 \cdot 10^{0})$	$6.44 \cdot 10^{1}$	$(1.80 \cdot 10^{-1})$
	Cross	$6.07 \cdot 10^{-8}$	$(1.74 \cdot 10^{0})$	$3.84 \cdot 10^{1}$	$(1.57 \cdot 10^{-1})$
	Smooth	$9.87 \cdot 10^{-7}$	$(9.62 \cdot 10^{-1})$	$7.36 \cdot 10^{-1}$	$(1.26 \cdot 10^{-1})$
Arévalo	Noisy	$4.03 \cdot 10^{-8}$	$(4.62 \cdot 10^{0})$	$3.50 \cdot 10^{2}$	$(2.85 \cdot 10^{-1})$
	Churazov	$2.10 \cdot 10^{-8}$	$(3.00 \cdot 10^{-1})$	$2.23 \cdot 10^{0}$	$(5.33 \cdot 10^{-2})$
	Cross	$3.90 \cdot 10^{-8}$	$(1.83 \cdot 10^{-1})$	$1.54 \cdot 10^{0}$	$(4.75 \cdot 10^{-2})$
	Smooth	$3.01 \cdot 10^{-7}$	$(2.62 \cdot 10^{-1})$	$4.85 \cdot 10^{-1}$	$(5.76 \cdot 10^{-2})$

Table 2. The MSE and MAPE statistical measures of the 2D projected noiseless power spectra versus noise reduction methods for the fBm field shown in figure 2.

References

[1] C.M. Brunt *et al.*, The Astrophysical Journal **595**, 824 (2003)

[2] A. Simionescu *et al.*, Space Science Reviews **215**, 24 (2019)

[3] R.S. Sutherland, M.A. Dopita, The Astrophysical Journal Supplement Series **88**, 253 (1993)

[4] E. Churazov *et al.*, Monthly Notices of the Royal Astronomical Society **421**, 1123 (2012)

[5] R. Khatri, M. Gaspari, Monthly Notices of the Royal Astronomical Society **463**, 655 (2016)

[6] C.E. Romero *et al.*, The Astrophysical Journal **951**, 41 (2023)

[7] Hitomi Collaboration, Nature **535**, 117 (2016)

[8] Hitomi Collaboration, Publications of the Astronomical Society of Japan **70**, 9 (2018)

[9] Y. Li *et al.*, The Astrophysical Journal **889**, L1 (2020)

[10] P. Stoica, R. Moses, *Spectral Analysis of Signals* (Pearson Prentice Hall Upper Saddle River, NJ, 2005)

[11] R.B. Blackman *et al.*, Physics Today **13**, 52 (1960)

[12] P. Arévalo *et al.*, Monthly Notices of the Royal Astronomical Society **426**, 1793 (2012)

[13] I. Zhuravleva *et al.*, Nature Astronomy **3**, 832 (2019)

[14] P.A. Davidson, M. Nelkin, Physics Today **58**, 80 (2005)

[15] R. Chhiber *et al.*, Journal of Geophysical Research: Space Physics **123**, 9941 (2018)

[16] S. Oughton *et al.*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **373**, 20140152 (2015)

[17] G. Cerini *et al.*, The Astrophysical Journal **945**, 152 (2023)