On SNR estimation over N-frame movie images

Technical notes and mathematical details to accompany CARYON
(**C**orrection **A**nd **R**ecover**Y** **O**f **N**-dimensional movies)

Christopher H S Aylett
c.aylett@imperial.ac.uk

Section for Structural Biology, Department of Infectious Disease,
Imperial College Road, South Kensington, London, SW7 2BB, UK

Abstract

Cryogenic electron microscopes are now typically fitted with direct electron detectors that are capable of identifying individual electron impacts with high fidelity and at high frame rate. However, the dose that can be applied to biological samples before they undergo radiation damage remains exceptionally small, on the order of tens of electrons per Ångström squared. The correction of such “movie” images requires the estimation of motion in X and Y, the sample position in Z and accompanying lens parameters, and the signal-to-noise ratio over the spectrum. While algorithms providing high fidelity in the estimation of position and lens parameters are available, current methods for signal-to-noise ratio estimation make incomplete use of the information available within the frames of modern movie images. Here I extend upon current signal-to-noise ratio estimation methods to optimise their application to high-frame-count movie images. This manuscript details the underlying approaches that will be used within the updated version of CARYON, a cryo-EM movie image filter.

1. Introduction to the problems underlying image SNR estimation, and explicit assumptions underlying an extension to N-dimensional movies

Biological samples' structural integrity depends on weak interactions, making them exceptionally radiation-soft, however it has been established that imaging their structures with high-energy electrons at cryogenic temperatures provides optimal information recovery in relation to the radiation damage engendered (Glaeser, 1971; Knapek and Dubochet, 1980; Henderson, 1995; Grant and Grigorieff, 2015). The electron dose that can be applied before radiation damage becomes intolerable remains exceptionally small, however, and therefore the signal-to-noise ratio (SNR) of such images is poor.

Cryogenic electron microscopes are now typically fitted with direct electron detectors that are capable of identifying individual electron impacts with high fidelity and at high frame rates (Li *et al.*, 2013; McMullan *et al.*, 2014). The correction of such "movie" images requires the estimation of motion in X and Y, the sample position in Z and accompanying lens parameters defining the contrast transfer function (CTF), and the SNR over the spectral range in which the signal of interest in the image is found (Zemlin, 1987; Brilot *et al.*, 2012; Li *et al.*, 2013). Algorithms providing high fidelity in the estimation of motion and the CTF are available (Mindell and Grigorieff, 2003; Brilot *et al.*, 2012; Campbell *et al.*, 2012; Li *et al.*, 2013; Rohou and Grigorieff, 2015; Zhang, 2016). However, current methods for SNR estimation either rely on comparison to known signals or are of lower fidelity, being based either on correlation or power-ratio sums, both requiring limiting assumptions on the nature of the signal distribution (Bershad and Rockmore, 1974; Frank and Ali, 1975; Harauz and van Heel, 1986; Unser *et al.*, 1987). Here I extend upon power-ratio SNR estimation methods to optimise their application to the high-frame-count movie images that are now available.

Considering the statement of the problem of SNR estimation for a general cryo-EM "movie", we have a number (N) of "frames", which we consider to have been previously aligned to one another in X and Y, i.e. they have already been previously "motion-corrected". Each "frame" consists of a large number (D) of pixel values, which originally would have been electron counts, but after alignment

represent a value derived from these. We assume that there is a “true”, invariant, value of the signal from the uncorrupted ground-truth image at each of the D pixels, thereby accounting for the effects of radiation damage only as part of the random noise per-pixel. We assume that there is a noise-distribution “falling-off” around each of these ground-truth pixel values, corrupting the true signal. Therefore, each of the D pixels in each of the N frames can be considered to represent a “draw” from a random variable, and our final movie represents N draws each (one per frame) from D random variables (one random variable per pixel). The SNR estimator I will consider is a power ratio closely related to that explored by Unser and colleagues (1987): the ratio of the variance of the signal to that of the noise. The original method was intended for application in Fourier space but can equally be applied in real space (or filtered bands Fourier transformed to real space) providing that the underlying assumptions remain fulfilled. These assumptions are quite stringent placing constraints on the statistical distributions of both signal and the noise. For estimation of the SNR of an image two half-images are constructed, by selection of even versus odd frames for example, each representing an independent sum of half of the movie frames. Current power ratio estimators assume that both the signal and the noise are distributed as jointly bivariate Gaussian random variables, the signal corresponding to their covariance and the noise corresponding to the variance of their differences. For the jointly bivariate Gaussian random variable the estimator of the signal-to-noise ratio is then:

$$\mathrm{SNR} = \frac{\sum_{[D]}\left(X_{[even-N,d]} + X_{[odd-N,d]}\right)^2}{\sum_{[D]}\left(X_{[even-N,d]} - X_{[odd-N,d]}\right)^2} - 1$$

... the numerator being considered the joint power of the “signal and noise” together, the denominator the power of the “noise” alone, and unity being subtracted to account for the fact that the power of the “noise” is included in both numerator and denominator, yielding the ratio between “signal” and “noise” power alone.

2. Extension of power-ratio SNR estimator calculations to N dimensions can improve the precision of the SNR estimate by up to √2-fold

Whilst retaining the assumptions of Gaussian signal and noise, we can improve the precision of power-ratio SNR estimation through the use of all frames within the movie. From our now generally available cryo-EM movie images, N “frames”, rather than 2 subsets, are present, and with modern detectors fine slicing can be used to increase N substantially if it is beneficial to do so. Half sums, as are generally used for SNR estimation or correlation calculation within the field, can be used. However, they discard information that can be used for the calculation, and also exclude movies with odd numbers of frames. The estimate of the “signal and noise” power in such an extended calculation remains the variance of the sum; i.e. on projection onto the column vector:

$$\begin{matrix} 1 \\ 1 \\ 1 \\ . \\ . \\ . \end{matrix}$$

... and therefore, the precision of the estimate of the joint “signal and noise” power cannot be improved upon from current power-ratio SNR calculations. N-1 orthogonal “noise” dimensions are available from the movie images allowing (N-1)-fold improvement of the estimation of the “noise” power parameters, however. The power-ratio SNR methodology corresponds to rotation and uniform scaling by a factor of √2 of the two-dimensional half-image space with the matrix:

$$\begin{matrix} 1 & 1 \\ -1 & 1 \end{matrix}$$

We can extend this to a higher N-dimensional space by rotation to place one vector along the √N diagonal of the (hyper)cube (Fig. 1).

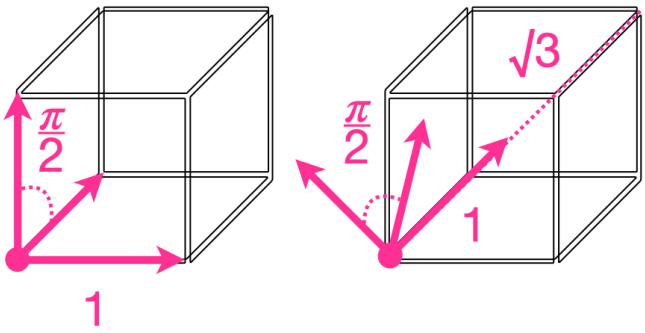

Figure 1: Rotation of the unit basis vectors in three dimensions to place the new unit basis vector along the cube diagonal. The same rotation is extended to N dimensions for N-vector hypercube geometry.

Such a rotation matrix for N dimensions can be generated trivially by taking an N by N square from the top left of the matrix shown below:

```
 1    1    1    1    1    .    .    .
-1    1    1    1    1    .    .    .
 0   -2    1    1    1    .    .    .
 0    0   -3    1    1    .    .    .
 0    0    0   -4    1    .    .    .
 .    .    .    .    .    .    .    .
 .    .    .    .    .    .    .    .
 .    .    .    .    .    .    .    .
```

... with each column vector being divided by its L_2 norm to yield a unit determinant as required for a rotation matrix. Essentially an upper triangular unit matrix is constructed and the sums are then subtracted below the matrix diagonal which guarantees orthogonality of each successive vector from the last vector as their projection into it will sum to zero by definition. A methodology essentially identical to that in current use, under the same simplifying assumption that all distributions are jointly Gaussian, can then be used to determine an SNR parameter for the rotated space, but with the "noise power" being estimated over the N-1 marginal noise distributions instead of only a single dimension, yielding improved accuracy. For 20 “frames” of unit variance the precision is increased by slightly below one third over that from half-images, becoming

better correlated with the value of the calculation for the sum given the use of the ground truth value of the noise variance (Fig. 2).

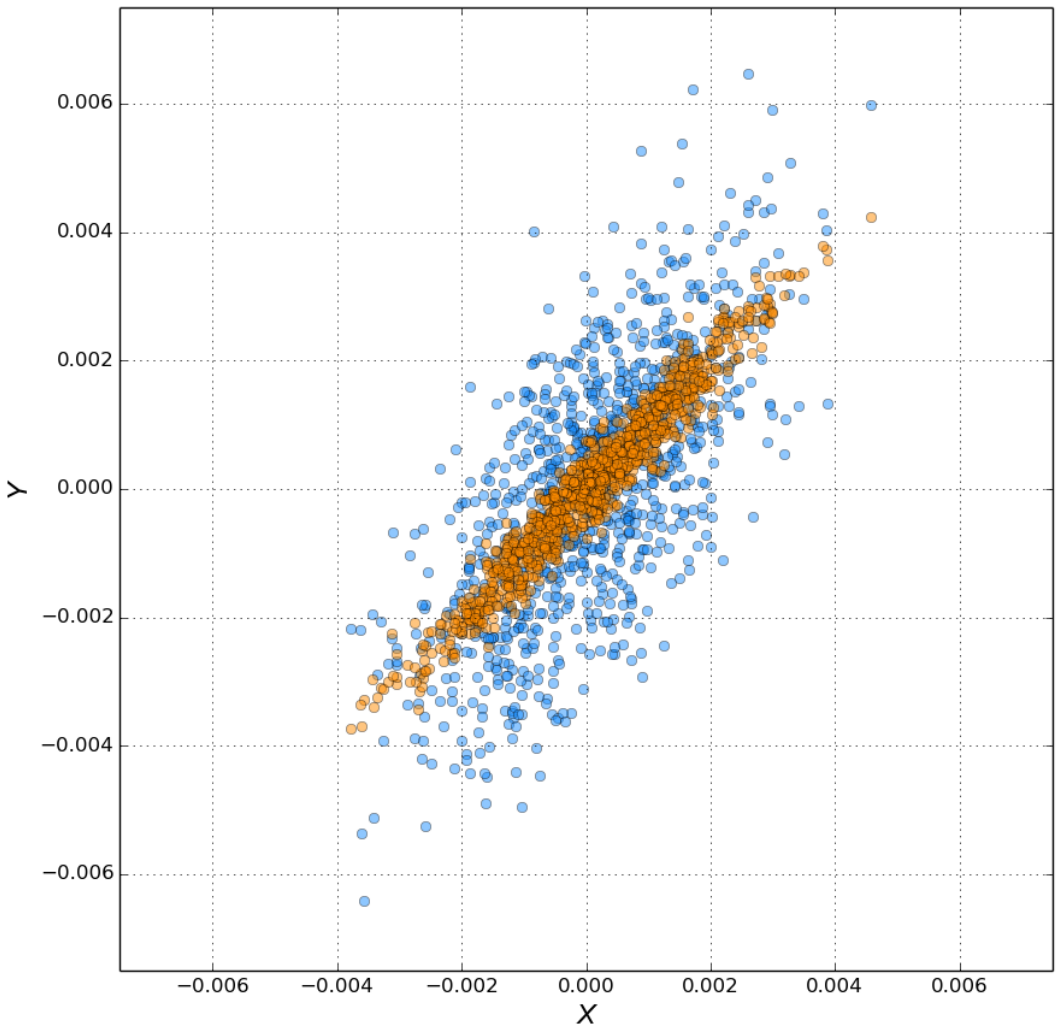

Figure 2: Scatter plot showing power-ratio estimates for the SNR (Y-dimension) plotted against the error in the sum of squares in the "signal and noise" observation alone (X-dimension). Even versus odd half-image estimate in blue, 20 frame estimate in orange. 20 frames of 1024^2 pixels were drawn from scipy.stats.norm.rvs(loc=0.0, scale=1.0) to form 1000 trial observations with a known SNR of zero.

Perfect linear correlation with the variance in the "signal and noise" sum, improving precision by $\sqrt{2}$, is therefore the optimum that can be achieved by incorporating N frames into the power-ratio calculation.

3. Current power-ratio SNR estimates are linearly related to the cross-entropy from the noise distributions to the sum distribution

The cross entropy is defined as:

$$\mathrm{Hx} = \frac{1}{\mathrm{J}} \cdot \sum_{[\mathrm{J}]} \left(-\log\left(P_{exp}[\mathrm{X_j}] \right) \right)$$

... where X represents the vector of J observations and P_{exp} represents the expected probability of observing each value, X_j, according to a known probability distribution function. The cross-entropy yields the expected increase or decrease in information required to represent the distribution actually observed over the J points in X in terms of the expected distribution P_{exp}. The cross entropy of the "signal and noise", or sum, distribution in terms of the "noise", or difference, distribution for a bivariate Gaussian corresponds to:

$$\mathrm{Hx} = \frac{1}{\mathrm{D}} \cdot \sum_{[\mathrm{D}]} \left(-\log\left(\frac{1}{\sqrt{2\pi \cdot \sigma^2}} \cdot \mathrm{e}^{\left(\frac{(\mu - \mathrm{X_d})^2}{-2\sigma^2} \right)} \right) \right)$$

... from the definitions of the Gaussian and the cross-entropy, and where μ represents noise mean and σ^2 represents noise variance ...

$$\mathrm{Hx} = \log 1 - \log\left(\sqrt{2\pi \cdot \sigma^2} \right) + \frac{1}{\mathrm{D}} \cdot \sum_{[\mathrm{D}]} \left(\frac{(\mu - \mathrm{X_d})^2}{2\sigma^2} \right)$$

... by separating the log terms and cancelling the exponential ...

$$\mathrm{Hx} = -\frac{1}{2} \cdot \log(2\pi \cdot \sigma^2) + \frac{1}{2\mathrm{D}} \cdot \sum_{[\mathrm{D}]} \left(\frac{\mathrm{X_d}^2}{\sigma^2} \right)$$

... cancelling and assuming that the distributions share a mean ...

$$\mathrm{Hx} = -\frac{1}{2} \cdot \log(2\pi \cdot \sigma^2) + \frac{1}{2} + \frac{1}{2} \cdot \sum_{[\mathrm{D}]} \left(\frac{\mathrm{X_d}^2}{\mathrm{D} \cdot \sigma^2} - 1 \right)$$

... rearranging to yield the noise entropy and power-ratio SNR ...

$$\mathrm{Hx} = \frac{1}{2} \cdot \mathrm{SNR} - \mathrm{H}_{\mathrm{noise}}$$

... where $\mathrm{H}_{\mathrm{Noise}}$ represents the entropy of the Gaussian noise distribution itself. The cross entropy may be evaluated in cases in which multivariate Gaussian assumptions do not hold and for cases in which SNR cannot be calculated as defined for power-ratio estimates. The cross entropy will not be linear to the SNR calculated on non-multivariate Gaussian distributed data in many cases, but may in fact be more suitable as a measure in most, as the strict Gaussian assumptions for current SNR estimators hold only rarely and the estimators provide an incorrect relative scaling factor wherever they do not. The cross entropy as shown above represents the average power required to convert the “noise” distribution into the “signal and noise” distribution, yielding a widely applicable value.

4. Given a large number of observed pixel values a copula of N dimensions is constructed by ranking the N•D observations, preserving information on their relationship but excluding confounding factors

The copula is a probability distribution consisting of an N-dimensional hypercube constructed from N unit uniform marginal distributions, each of which represents the probability integral transform of an original marginal distribution. The copula structure therefore isolates all the dependence information between variables of interest from any confounding factors, and the negative entropy of the copula has been shown to be identical to the mutual information between the marginal variables (Ma and Sun, 2011).

Under the assumptions that both the D ground truth location values and the N•D pixel values are separable and can be ordered, replacing each sample with its sorted and weighted rank yields a well-approximated unit uniform marginal distribution. Projecting these into the unit hypercube approximates the copula. Rotation of the unit hypercube copula in exactly the same manner as described above in section 2 is trivial and preserves the useful properties of the construction. In the case of zero mutual information between the variables, and therefore zero SNR, the transformed marginal distributions will correspond to Bates' distribution of order N after rotation:

$$\mathrm{Bates}(X) = \frac{N}{2(N-1)!} \cdot \sum_{[k]} (-1)^k \binom{N}{k} (NX-k)^{N-1} \cdot \mathrm{sgn}(NX-k)$$

... which represents the orthogonal volume along each diagonal of the hypercube as a series of simplices of weights (N-choose-k, $^{N}_{k}$) from Pascals triangle. Here x! represents the factorial operator, x(x-1)(x-2)...(x-(x-1)), and sgn(x) the sign operator $x/\sqrt{x^2}$. Calculation of Bates' distribution can be simplified using Heaviside's step function U(x), returning the value of x only for positive x, and 0 otherwise, which halves the number of terms to be considered, and by reflection around its mean at 0.5, also reducing in the number of terms by either half or one less than half:

$$\text{Bates}(X) = \frac{N}{(N-1)!} \cdot \sum_{[k]} (-1)^k \binom{N}{k} U(NX-k)^{n-1}$$

... while for large values of N (ideally greater than 10) Bates' distribution can be approximated with a normal distribution of variance $(12N)^{-1}$, simplifying calculations substantially.

5. Cross-entropy calculated from the N-dimensional copula can yield accurate estimates of the SNR for Gaussian data and accurate estimates of an "SNR-equivalent" for non-Gaussian distributions that do not yield reproducible SNR estimates by standard means

As the SNR approaches zero, projections of the rotated copula approach Bates' distribution (and its Gaussian approximation for large N) which may be used to calculate the cross-entropy directly. The N-1 orthogonal dimensions drop in power directly proportionally to signal power accumulation on the diagonal sum. An "SNR-equivalent" parameter may therefore be calculated from the copula alone without parameterising the distributions of the data at any point (Fig. 3).

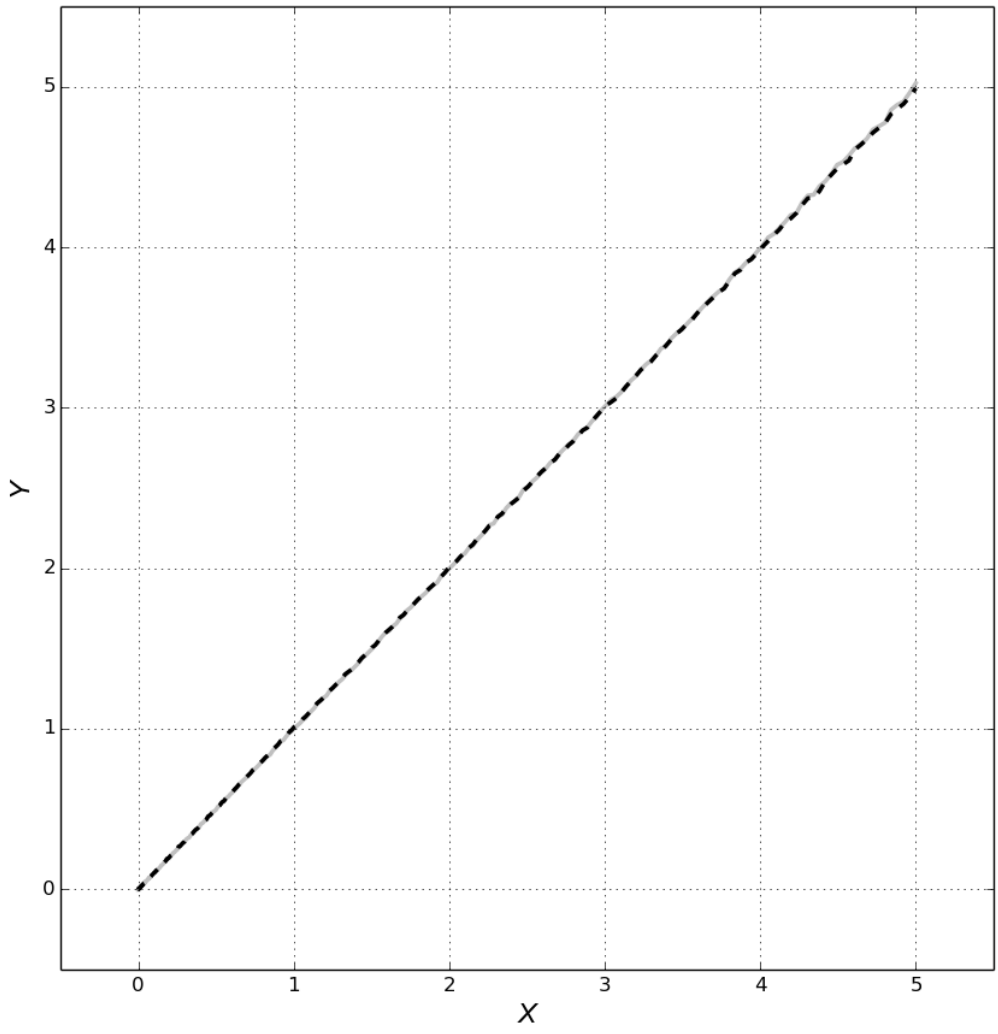

Figure 3: Scatter plot of power-ratio (silver-solid) and copula-based (black-dotted) SNR estimates (Y-dimension) plotted against the ground-truth SNR (X-dimension) from 0 to 5. 20 frames of 1024^2 pixels were drawn from scipy.stats.norm.rvs(loc=scipy.stats.norm.rvs(loc=0.0, scale=X), scale=1.0) to produce 250 evenly spaced observations.

This approach can be considered analogous to the extension of Spearman's rank correlation coefficient upon Pearson's correlation coefficient. Bate's distribution is dependent only on N and therefore the calculation is quite accurate, having marginally higher variance than the extended power ratio calculation for perfectly Gaussian data (estimator accuracy increasing with increasing N), although it becomes non-linear (giving slight underestimation) at high SNR (Fig. 4).

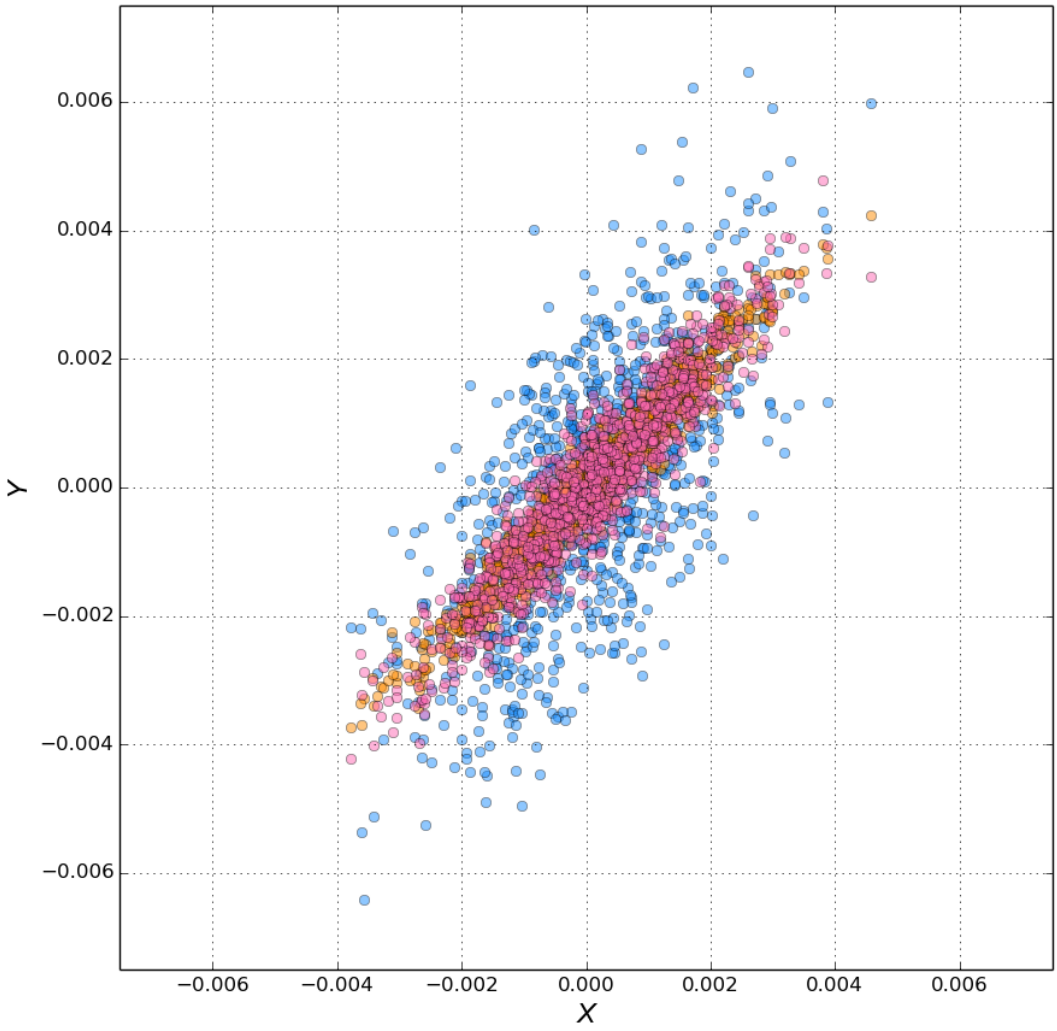

Figure 4: Scatter plot showing SNR estimates (Y-dimension) plotted against the error in the sum of squares in the "signal and noise" alone (X-dimension). Even versus odd half-image estimate in blue, 20 frame estimate in orange, copula-based estimate in pink. 20 frames of 1024^2 pixels were drawn from scipy.stats.norm.rvs(loc=0.0, scale=1.0) to form each of 1000 trial observations with a known SNR of zero.

This is beneficial as the exact signal distribution in cryo-EM images, and any image of scientific interest, is unknown and therefore difficult to parameterise. Consistent "SNR-equivalent" values can also be obtained for distributions that cannot be readily measured, for instance Cauchy-distributed random variation, which has undefined variance because of heavy tails leading to large "outliers" (Fig. 5).

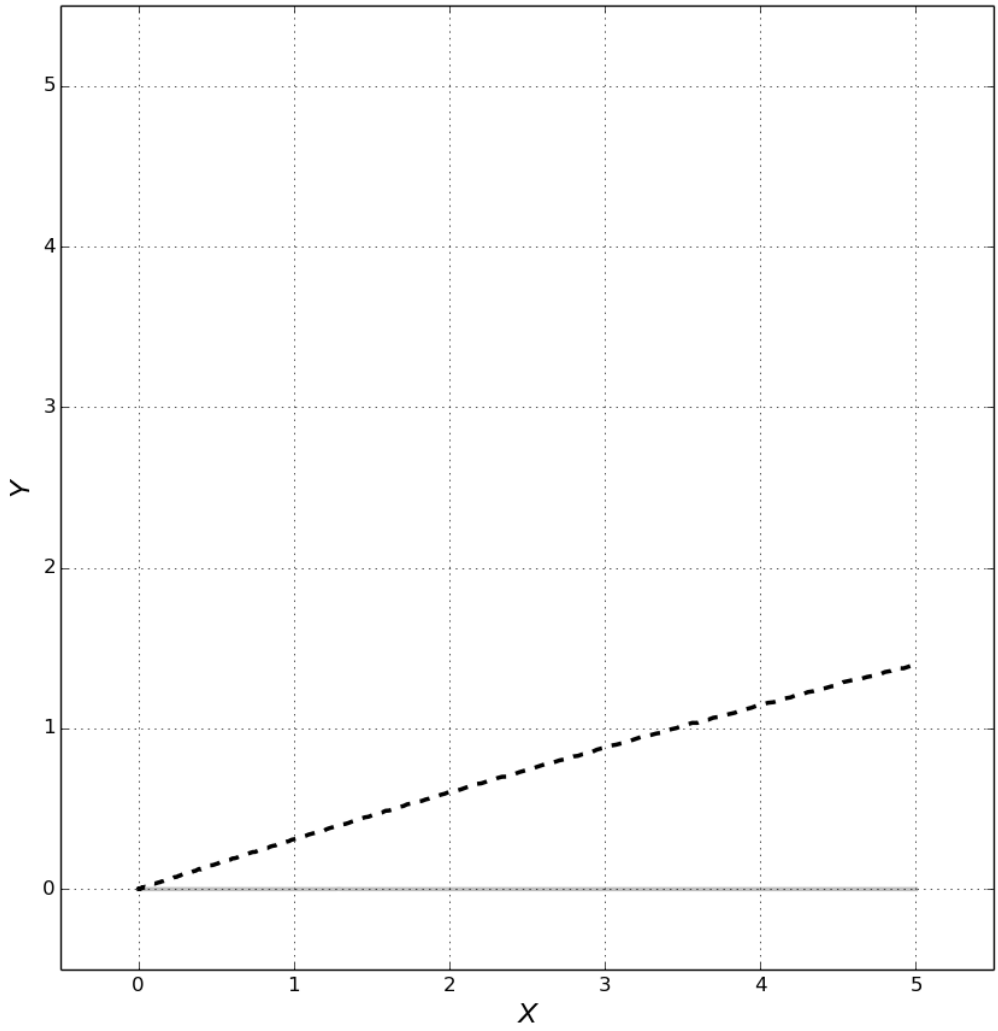

Figure 5: Scatter plot of power-ratio (silver-solid) and copula-based (black-dotted) SNR estimates (Y-dimension) plotted against the ratio of the standard deviation to the Cauchy scale parameter (X-dimension). 20 frames of 1024^2 pixels were drawn from scipy.stats.cauchy.rvs(loc=scipy.stats.norm.rvs(loc=0.0, scale=X), scale=1.0) for each of 250 evenly distributed observations with ratios from 0 to 5.

Consistent "SNR-equivalent" values can also be obtained for difficult distributions with correlated error, such as gamma-distributed noise with defined signal manifested through changes in rate parameter. This results in skew and bias in sum-of-squares estimates of SNR because correlation between the variance and the mean of the probability distribution leads larger values to dominate the estimate (Fig. 6).

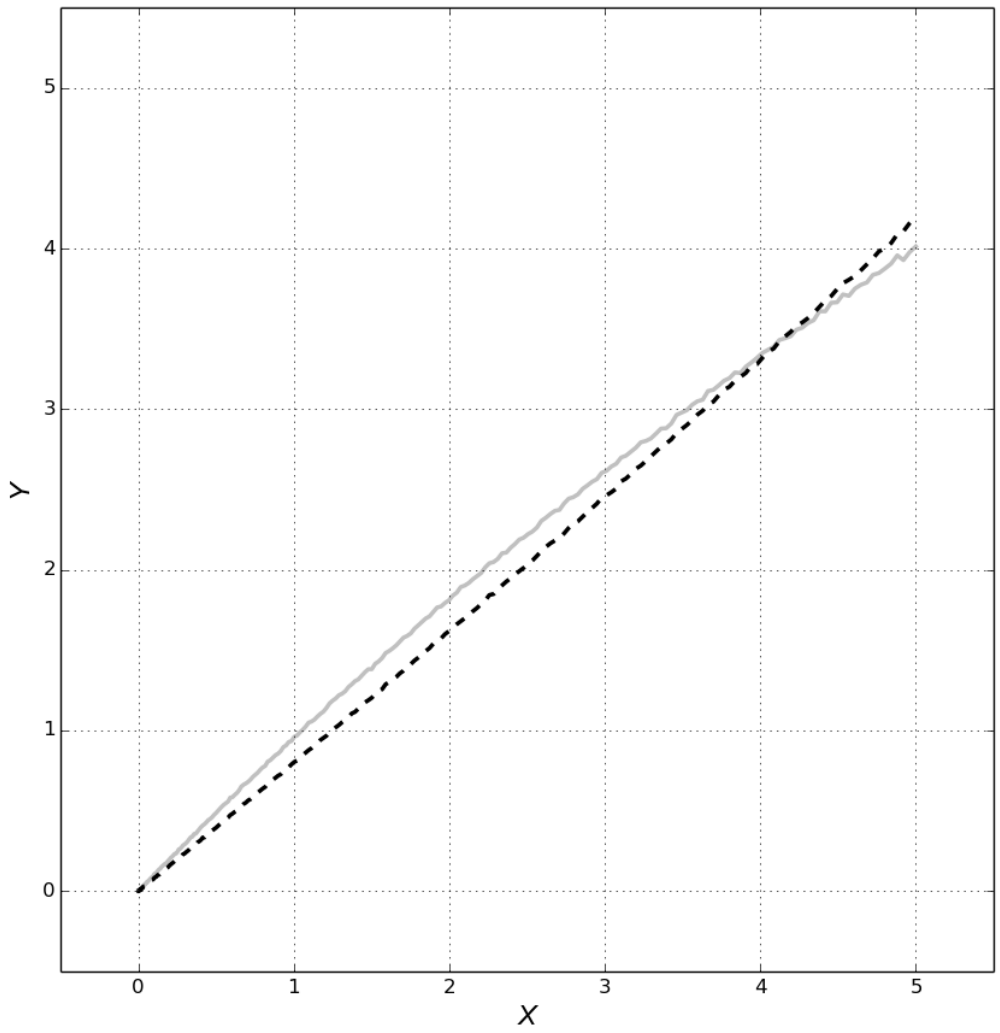

Figure 6: Scatter plot of power-ratio (silver-solid) and copula-based (black-dotted) SNR estimates (Y-dimension) plotted against the reciprocal squared shape parameter of their scale parameter for the gamma distribution (X-dimension) from values of 0 to 5. 20 frames of 1024^2 pixels were drawn from scipy.stats.gamma.rvs(1.0, loc=0.0, scale= scipy.stats.gamma.rvs(1.0 / X^2, loc=0.0, scale=1.0), scale=1.0) for each of 250 evenly spaced observations.

It should be noted that data derived from electron, or other similar statistically detected low-dose particle, impacts are distributed according to Poisson or analogous continuous distributions that are truncated at zero, concomitantly skewed, and have such strong confounding relationships between their mean and variance. These improvements therefore have immediate utility for our subject matter.

6. Use of information from movie frames allows calculation of SNR or equivalent with greater accuracy and from pathological distributions

I have extended current power-ratio methods into N-dimensional geometry to allow their use with improved precision of up to √2 for the estimation of SNR from movie images through use of the information within frames. I have also devised and demonstrated a copula and cross-entropy based approach that has similar accuracy to the N-dimensional extension of the power-ratio estimator for idealised Gaussian random variables, but which is generally applicable to pathological or poorly behaved marginal distributions such as Cauchy, gamma, or uniformly distributed data. The “SNR-equivalent” values obtained in this way are close to identical to the improved N-dimensional method for idealised Gaussian data, but are not affected by outliers, heavy tails, or correlated noise in the way that power-ratio methods necessarily are, allowing their use to estimate SNR for arbitrary signal distributions. These methods are to be implemented for MRC movie images, alongside CTF deconvolution, as part of CARYON.

References

Bershad NJ & Rockmore AJ (1974) On estimating signal-to-noise ratio using the sample correlation coefficient. *IEEE Trans. Inf. Theory* **20**: 112–113.

Brilot AF, Chen JZ, Cheng A, Pan J, Harrison SC, Potter CS, Carragher B, Henderson R & Grigorieff N (2012) Beam-induced motion of vitrified specimen on holey carbon film. *J. Struct. Biol.* **177**: 630–637

Campbell MG, Cheng A, Brilot AF, Moeller A, Lyumkis D, Veesler D, Pan J, Harrison SC, Potter CS, Carragher B & Grigorieff N (2012) Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure* **20**: 1823–1828

Frank J & Al-Ali L (1975) Signal-to-noise ratio of electron micrographs obtained by cross-correlation. *Nature* **256**: 376-379.

Grant T & Grigorieff N (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *Elife* **4**:

Glaeser RM (1971) Limitations to significant information in biological electron microscopy as a result of radiation damage. *J. Ultrasructure Res.* **36**: 466–482

Harauz G & van Heel M (1986) Exact filters for general geometry three-dimensional reconstruction. *Optik* **73**: 146-156

Henderson R (1995) The Potential and Limitations of Neutrons, Electrons and X-Rays for Atomic Resolution Microscopy of Unstained Biological Molecules. *Q. Rev. Biophys.* **28**: 171–193

Knapek E & Dubochet J (1980) Beam damage to organic material is considerably reduced in cryo-electron microscopy. *J. Mol. Biol.* **141**: 147–161

Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA & Cheng Y (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**: 584–90

Ma J & Sun Z (2011) Mutual information is copula entropy. *Tsinghua Sci. Technol.* **16**: 51-54

McMullan G, Faruqi AR, Clare D & Henderson R (2014) Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy* **147**: 156–163

Mindell JA & Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142(3)**: 334-347.

Rohou A & Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192(2)**: 216-21.

Unser M, Trus BL & Steven AC (1987) A new resolution criterion based on spectral signal-to-noise ratios. *Ultramicroscopy.* **23(1)**: 39-51.

Zemlin F (1978) Image synthesis from electron micrographs taken at different defocus. *Ultramicroscopy* **3**: 261–263

Zhang K (2016) Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**: 1–12

Acknowledgments

I would like to thank JS Aylett, CM Palmer, JT MacDonald, and K Ramlaul, for taking the time to read and review various aspects of this manuscript.

Funding

This work was supported by the Wellcome Trust and the Royal Society through a Sir Henry Dale Fellowship (206212/Z/17/Z) to CHSA.

Conflict of interest

I know of no conflicts of interest with respect to this work.