STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS

ISSN 2052-5206

The index of dispersion as a metric of quanta – unravelling the Fano factor

Wilfred K. Fullagar,* Mahsa Paziresh, Shane J. Latham, Glenn R. Myers and Andrew M. Kingston

Applied Mathematics, RSPE, Oliphant Building 60, Mills Road, Australian National University, Canberra, ACT 2601, Australia. *Correspondence e-mail: wilfred.fullagar@anu.edu.au

Received 17 February 2017
Accepted 19 June 2017

Edited by Y.-S. Chen, Center for Advanced Radiation Sources, USA

Keywords: Monte Carlo methods; high-quality data refinement; spectrum modelling; polychromatic methods; quantum energy determination; detector development.

In statistics, the index of dispersion (or variance-to-mean ratio) is unity ($\sigma^2/\langle x\rangle = 1$) for a Poisson-distributed process with variance σ^2 for a variable x that manifests as unit increments. Where x is a measure of some phenomenon, the index takes on a value proportional to the quanta that constitute the phenomenon. That outcome might thus be anticipated to apply for an enormously wide variety of applied measurements of quantum phenomena. However, in a photon-energy proportional radiation detector, a set of M witnessed Poisson-distributed measurements $\{W_1, W_2, \ldots W_M\}$ scaled so that the ideal expectation value of the quantum is unity, is generally observed to give $\sigma^2/\langle W\rangle < 1$ because of detector losses as broadly indicated by Fano [*Phys. Rev.* (1947), **72**, 26]. In other cases where there is spectral dispersion, $\sigma^2/\langle W\rangle > 1$. Here these situations are examined analytically, in Monte Carlo simulations, and experimentally. The efforts reveal a powerful metric of quanta broadly associated with such measurements, where the extension has been made to polychromatic and lossy situations. In doing so, the index of dispersion's variously established yet curiously overlooked role as a metric of underlying quanta is indicated. The work's X-ray aspects have very diverse utility and have begun to find applications in radiography and tomography, where the ability to extract spectral information from conventional intensity detectors enables a superior level of material and source characterization.

1. Introduction

1.1. Foundations

Extraction of the maximum possible information from datasets has been critical to the practical establishment of many fields, among them X-ray charge-density analyses (Coppens, 1997; Bolotovsky *et al.*, 1995), and monochromatic and polychromatic photocrystallographies (Coppens *et al.*, 2009; Fullagar *et al.*, 2000; Helliwell *et al.*, 1989; Kim *et al.*, 2002). The capability for hyperspectral X-ray imaging *via* shot noise in energy-proportional X-ray imaging detectors does not seem widely used within X-ray radiography and tomography communities. The corresponding aspect of shot noise has been indicated elsewhere (Uhlig *et al.*, 2011), which motivated its further elaboration in the present study.

Quantum shot noise exists in many situations, where classical statistics may be applied. When doing so, the following two conditions may be indicated.

(i) A witnessed measurement is proportional to both the number of quanta reaching the detector, N, and the detected content of the quanta, E.

$$W = NET(E); \quad \langle W\rangle = \langle N\rangle ET(E). \tag{1}$$

(ii) The number of quanta shows shot noise (Poisson) variation about its expectation value $\langle N\rangle$, so that

$$\sigma = \sqrt{\langle N \rangle} ET(E). \quad (2)$$

The measurement W may be the calibrated output of an analog to digital converter, but it could equally be a mass, a concentration, a deflection, a current or any other witnessable magnitude. Depending on literary background and context, the product NE might be formally defined as an intensity; to avoid potential conflicts of understanding we avoid that notation. The transfer function $T(E)$ describes how E is transmitted through the detector system to yield the witnessed signal. Observe that a quantum efficiency $Q(E)$ governs the number of quanta N *captured* by the detector [and so effectively resides within N in equations (1) and (2)], but that $Q(E)$ is functionally distinct from $T(E)$. The product $ET(E)$ is the detector's *response* to what it captures.

Rearrangement of equation (2) gives $E = \left(\sigma/\sqrt{N}\right)[1/T(E)]$. This has been used to estimate the energy of hot electrons in matter (Uhlig *et al.*, 2011), which is usually a preliminary step of X-ray detection (Wilson, 1912; Fullagar *et al.*, 2008).

The role of $\sigma^2/\langle W \rangle$ as a metric of E follows from equations (1) and (2),

$$\frac{\sigma^2}{\langle W \rangle} = ET(E), \quad (3)$$

so is independent of $Q(E)$, but still a function of $T(E)$. Alternatively, $\langle W \rangle/\sigma = \sqrt{\langle N \rangle}$ (the signal-to-noise ratio for Poisson signals) follows by cancellation of $ET(E)$ from equations (1) and (2).

1.2. The polychromatic case

Equation (3) only really applies if E is single-valued (*i.e.* monodisperse). Otherwise, the spectrum may be conceptually divided into b evenly spaced energy bins of average energy E_1, $E_2 \ldots$, E_b. The filling of neighbouring spectral bins is assumed to be uncorrelated, so numerator and denominator contributions of the different energy bins are expanded separately to give

$$\frac{\sigma^2}{\langle W \rangle} = \frac{n_1 E_1^2 T^2(E_1) + n_2 E_2^2 T^2(E_2) + + n_b E_b^2 T^2(E_b)}{n_1 E_1 T(E_1) + n_2 E_2 T(E_2) + + n_b E_b T(E_b)} = \frac{\langle S(E) E T^2(E) \rangle}{\langle S(E) T(E) \rangle}. \quad (4)$$

Here the second equality gives the result as a ratio of expectation values. These involve the spectral density function $S(E) = E_i n_i(E)/N$, with $\sum_i n_i = N$ in the conceptual limit of infinite b and N. That suggests use of integral notation, although here the quest for properties of quanta motivates the summation notation of statistical mechanics and thermodynamics (Baker & Cousins, 1984; Schrödinger, 1989). Equation (4) reduces to equation (3) in the monochromatic case, but $S(E)$ and $T(E)$ otherwise become essential features within the expectation brackets. Spectral dispersion manifests through $S(E)$, which by itself always acts to increase the value of $\sigma^2/\langle W \rangle$ relative to the expectation energy $\langle S(E) \rangle$ (see *Appendix E*). Meanwhile partial registry manifests through $T(E)$ and suppresses $\sigma^2/\langle W \rangle$ *via* arguments introduced by Fano (1947) (see *Appendix H*). Their combined influence may be explored *via* equation (4) using models of $S(E)$ and $T(E)$. The ratio of expectation values recalls a Rayleigh quotient (Anton & Rorres, 1987), which would give an eigenvalue corresponding to $ET(E)$; while $\sigma^2/\langle W \rangle$ is the eigenfunction of the repeated self-convolution process corresponding to the central limit theorem (see *Appendix G*). The calibration of $T(E)$ for an energy-proportional detector is thus tantamount to making a crude spectrometer from it. It enables hyperspectral imaging from the corresponding detector, even in cases where individual quanta are not discernible.

1.3. Scope and organization

This document seeks to provide a useful collection of information concerning the index of dispersion as a metric of quanta. A mathematical review is avoided for accessibility reasons, recognizing that any claims of novelty would be impossible to defend in that arena. Readers may be able to apply the interpretation offered here, that includes polychromatic and lossy situations, to data of diverse natures. A cursory introduction to historical and other aspects is made next, before presenting X-ray observations from which applied developments have been growing from this work, and closing with some corollaries. Appendices are referred to throughout. Their structure is as linear as reasonably possible, offering a deeper look at the many necessary considerations under suitable subheadings. They provide essential insights to pertinent physics and mathematics, additional applied examples whose examination allows placement of general bounds on applicability, and entry points to further literature.

2. Applicability

2.1. Historical and mathematical

The ability of the equations above to access the magnitude of quanta applies in broad fields. For example, in the early 1800s, ideas of the quantization of matter were re-invigorated by the likes of Dalton; molecular weights m became known through the work of Avogadro *via* Dumas' vapour density measurements (soon also associated with the names Cannizzaro, Viktor Meyer and Hoffman) which then also led to atomic weight determinations (Glasstone, 1940; Pauling, 1956). These use the fact that $\rho T/P = m/R$ in accordance with the ideal gas law, for a gas of physical density ρ (which supplies the necessary mass dimension on the left side), at temperature T and pressure P with gas constant R. A connection to equation (3) is seen through the kinetic theory of gases, briefly motivated as follows. From the average velocity v of molecules, $P \propto mv\rho_n$ (molecular momentum times numerical density); while $T \propto mv^2$ (molecular kinetic energy) may be viewed as m-normalized dispersion of the molecular momentum. The ratio $T/P \propto \sigma_{mv}^2/(mv)$ analogous to equation (3). The recovery of m is by multiplying by ρ $(= m\rho_n)$, the measured physical density of the gas.

All the following have in common mathematics relating to $\sigma^2/\langle W \rangle$ for the purposes of extracting quantum information,

and are grounded in fluctuation analyses in statistical mechanics (Barnes & Silverman, 1934; Schrödinger, 1952): determination of the electron charge by shot noise measurements (Hull & Williams, 1925); diffusion and random walk models (Atkins, 1998); demonstration of the molecular constitution of matter and Avogadro's number determination *via* Brownian motion (Glasstone, 1940); optical galvanometer and cantilever noise in atomic force microscopy to measure Boltzmann's constant (Ising, 1926; von Smoluchowski, 1912; White, 1963). Raindrop measurements by corresponding arguments are presented in *Appendix I*. The closely related Fano factor has been examined in neuroscience contexts to evaluate whether neural spiking follows a Poisson distribution (Eden & Kramer, 2010); the underlying arguments are general. There are strong reasons to indicate a correspondence principle linking many related phenomena (Hughes & Ninham, 2016).

The Poisson distribution is broadly accepted in many contexts, but it can be difficult to pin specific *physical* causes for it. One derivation takes the binomial distribution in the limit of N (number of tries) going to infinity while the product Np (with p being the probability of 'success' per try) remains constant; while the binomial distribution can itself be seen as a limit of other distributions (Dudewicz & Mishra, 1988). Relevance in combinatorial problems is implied, as when adding random phases (Rayleigh, 1919) or when enumerating microstates in statistical mechanics. Boltzmann's combinatorial approach to statistical mechanics can be avoided (Darwin & Fowler, 1922; Schrödinger, 1952), whereupon the Poisson distribution may be derived in functional form by 'top–down' approaches (Morse, 1969; Kittel & Kroemer, 1980), or from quantum mechanics (Henri-Rousseau & Blaise, 2011). These allow the Poisson distribution's perception as a distribution in its own right and not just as an approximation to the binomial distribution (Noble, 1991); as an entropy-related function (Jaynes, 2003); and independently of any dimension chosen for its expression (*e.g.* the time domain is a popular choice; Lindgren, 1976; Davis, 1996). Mathematical consistency anticipates several approaches to the Poisson distribution's derivation, just as the Gaussian distribution also enjoys (Jaynes, 2003). That empowers its use in both bottom–up and top–down arguments, allowing robustness when facing questions like 'but what about quantum theory?' (Jaynes, 2003; Cabello *et al.*, 2016). Inevitably then, the Poisson distribution's deep roots lie beyond the scope of this manuscript. Rather, the present contribution draws attention to the connection it implies between $\sigma^2/\langle W\rangle$ and quanta, a relationship that is overlooked by many works that might have attached tremendous significance to it.

2.2. Requirements

For a property of quanta to be revealed by the index of dispersion, any patterns in data must be rigorously accounted for – dispersion otherwise arises from variance in expectation values too, just as it does from shot noise [equation (4) contains this, in $S(E)$]. Such patterns alter the expectation value for data points, which is the key to their removal by subtraction, if needs are pressing and experiments cannot eliminate the cause. On average, a pattern properly accounted for always brings expectation values closer to measured data (*Appendix I*). This reduces $\sigma^2/\langle w\rangle$, whose minimum reliably obtained value is therefore presumed best, provided that underlying patterns do not involve an excessive number of parameters (the optimum number of parameters given any prior information is provided by Bayesian analysis; Sivia & Carlile, 1992). Examining all the dimensions of a dataset often suggests patterns, while examination of known dimensions within existing data and improved experimental methodology stimulates the quest. (Typical experimental pattern searches may be seen in the determination of raindrop size *Appendix I*.) The shape of a Poisson distribution very rapidly approximates Gaussian even for quite modest expectation numbers of quanta, but with the key additional property $\langle x\rangle = \sigma^2$ (Prigogine, 1978). The two distributions are often experimentally indistinguishable, especially where differences are involved. Where individual quanta are observed, Poisson error-normalized residuals whose histogram matches a $\mu = 0$, $\sigma = 1$ Gaussian distribution indicate the exhaustion of a dataset's capacity to support further patterns (Uhlig *et al.*, 2013). (This is because the additional distribution caused by a pattern would convolve with the observed residuals to broaden their error-normalized histogram beyond the $\mu = 0$, $\sigma = 1$ Gaussian expected for the $\langle x\rangle$-subtracted Poisson distribution alone.) Where quanta are not directly observed, the underlying Poisson distribution may be open to question. However, using the ability of shot noise to numerically overwhelm other noise sources, together with spectral examinations, a picture of detector response can be built up. Poisson shot noise is then unambiguously quantifiable in proportional detectors that lack the sensitivity for single quantum measurements.

2.3. Key restrictions and the case for thermal detection

Practical measurement of a mean and standard deviation *via* equations (1) and (2) implies a set of M individual measurements, under experiment conditions that anticipate an expectation value. Such measurements might be done sequentially (in time domain) or in parallel (using identical detectors in spatial or angular domains). All three domains are common in movies and images. Why then, is $\sigma^2/\langle I\rangle$ not used to estimate colour information from black and white photographs? The X-ray experiments below show that at X-ray photon energies (far exceeding the detector's bandgap) the X-ray colour *can* be determined in this way. However, for energies not much above the detector bandgap, $ET(E) \simeq$ constant (*Appendix H*). This is because a quantum with only fractionally more energy than the bandgap cannot excite more than one electron-hole pair over that energy barrier. Other consequences of the photon capture (mostly heat) are usually not recorded. When the detector response $ET(E)$ loses dependence on E, equations (1)–(4) no longer offer the required metric of E. The result is the same where

thresholding (discrimination or windowing) is applied in the front-end electronics of X-ray detectors. A different form of ill-conditioning occurs when $T(E)$ for a detector is exponential or extremely large, for example in detectors that, like Geiger counters (Strong, 1938), use avalanche gain mechanisms (Fullagar *et al.*, 2011*b*; Tutt *et al.*, 2012), and methods of thermal neutron detection that depend on unleashing relatively enormous nuclear energies. There also, equations (1)–(4) can fail as a reliable metric of E.

Recognition of this situation in photon-measuring direct-detection CCDs (Fullagar *et al.*, 2008) motivated many investments in ultracryogenic microcalorimetry (Enss, 2005; Mazin *et al.*, 2013; Fullagar *et al.*, 2007). By suitable measures, not only is each quantum *fully* measured, but it is split into a temperature-dependent number of low-energy quasiparticles, fluctuations of which bound the accuracy of measurements of the parent quantum according to thermalization temperature and the absorber's heat capacity (Moseley *et al.*, 1984; Mössbauer, 1991; Enss, 2005; Ullom & Bennett, 2015). A critical consequence of this low-temperature microcalorimetry approach for X-ray science is that the typically $\sim 10^{-5}$ Darwin–Bragg losses of conventional diffraction-based analyses (Als-Nielsen & McMorrow, 2001; Darwin, 1914) are avoided (Fullagar, Uhlig *et al.*, 2010; Fullagar *et al.*, 2017), opening scattering-law technique development (Uhlig, 2011; Uhlig *et al.*, 2013, 2015; Fullagar *et al.*, 2017) around temporally remarkable in-house polychromatic hard X-ray sources (Fullagar *et al.*, 2007; Ta Phuoc *et al.*, 2012).

2.4. Radiation application fields

Equations (1)–(4) apply in microcalorimeter systems and in many other real-world imaging systems where the original

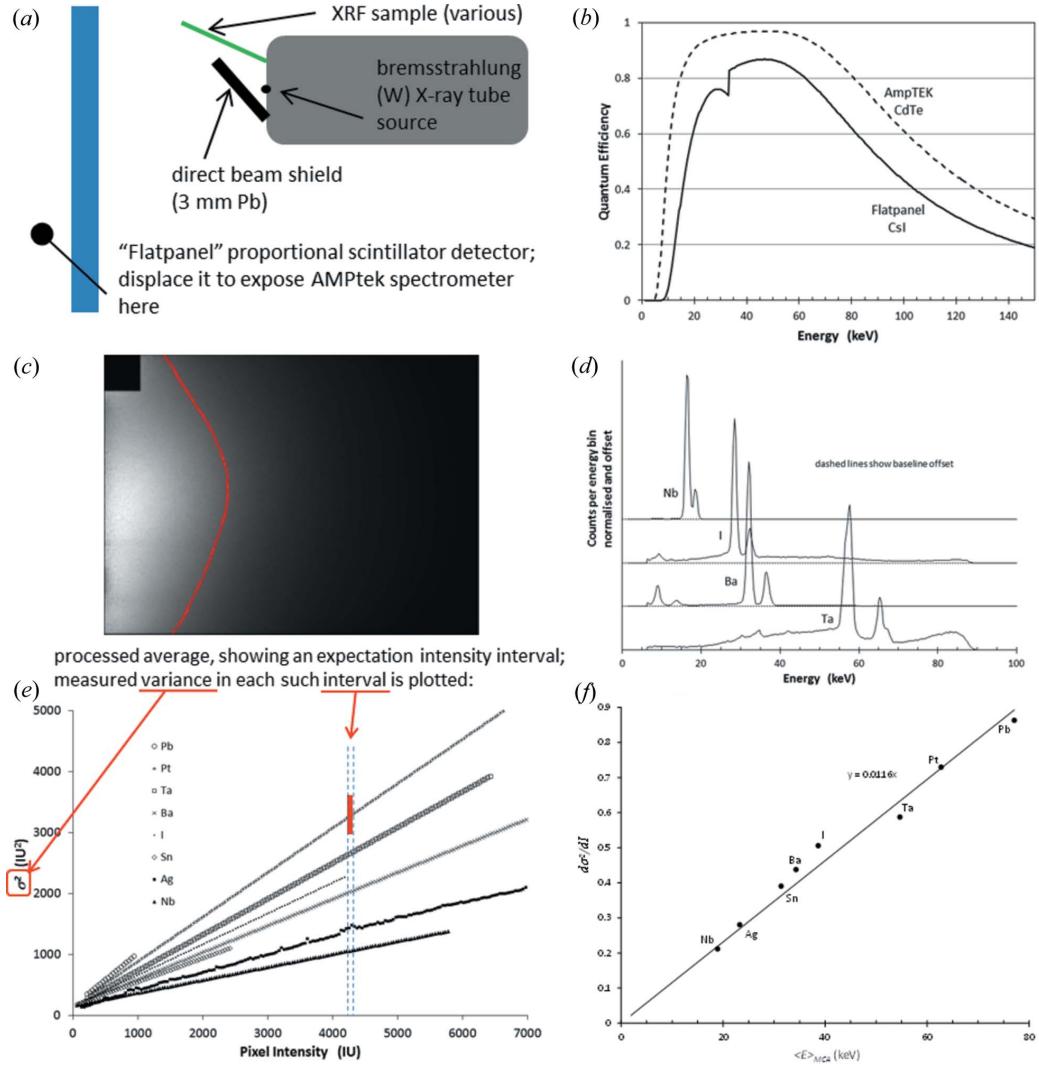


Figure 1
Pseudomonochromatic XRF data were generated in experiments using two different detectors (*a*) and (*b*). For the XRF images, expectation intensities (an example in *c*) were calculated from averages of 256 clearfield and 96 darkfield frames. (*d*) shows a selection of corresponding measured spectra. In (*c*) an expectation isointensity strip is highlighted, whose contributing pixels show measurable intensity variance. Plotting the variance against the expectation intensity in (*e*) yields an approximately straight line for each XRF sample, with a small *y*-axis offset due to inherent detector noise. In (*f*) the fitted line slopes are plotted against average photon energy observed in the spectrometer. A straight line fit through the points in panel *f* yields the zero-order estimate for the scintillator detector's response.

quantum energy greatly exceeds bandgaps. These include X-ray radiography/tomography (Paziresh *et al.*, 2016), Laue crystallography (Helliwell *et al.*, 1989), polychromatic X-ray phase contrast imaging (Wilkins *et al.*, 1996), cryogenic microcalorimetry approaches to astronomical (Mazin *et al.*, 2013) and THz (Becker *et al.*, 2014) imaging, and other disciplines (*e.g.* fast electron measurements; Uhlig *et al.*, 2011), and may be useful to estimate X-ray fluorescence (XRF) contributions to diffuse scatter in recent XFEL reports (Ayyer *et al.*, 2016). In such situations we indicate the potential utility to establish corresponding quantum energies and spectral parameters, when appropriate spectral models are applied to equation (4). Equation (4) is an entry point to the thermodynamics of colour, with Figs. 1 and 2 being demonstrations of it.

Laue crystallography leads to monochromatic (Bragg-diffracted) reflections on detector planes. This polychromatic technique is especially valuable for its ability to witness large volumes of reciprocal space in single radiation shots (Šrajer *et al.*, 2001), which brilliant monochromatic sources cannot do given ordered samples (Kim *et al.*, 2002; Moffat, 2003; Helliwell *et al.*, 1989; Fullagar 2017). The energy of quanta in the reflections is key to determining momentum transfer and thereby structure (van Hove, 1954; Higginbotham *et al.*, 2014; McDonald *et al.*, 2015), as well as molecular dynamics *via* spectroscopy (van Hove, 1954; Fullagar, 1999; Ament *et al.*, 2011). In combination with earlier work (Uhlig *et al.*, 2011), results presented here indicate how to deduce that energy using pixel-to-pixel intensity variance in single shot data.

X-ray radiography and tomography often uses a broadband source and scintillator-coupled optical detector. This breed of detector was used in experiments described below. Parametric models relate $S(E)$ to source spectra and (for example) density ρ and atomic number Z of intervening materials (Alvarez & Macovski, 1976). Owing to the great redundancy of projections traversing any given voxel, equation (4) offers clear scope for improved spectral understanding of statistical mechanics approaches to model-based refinement of

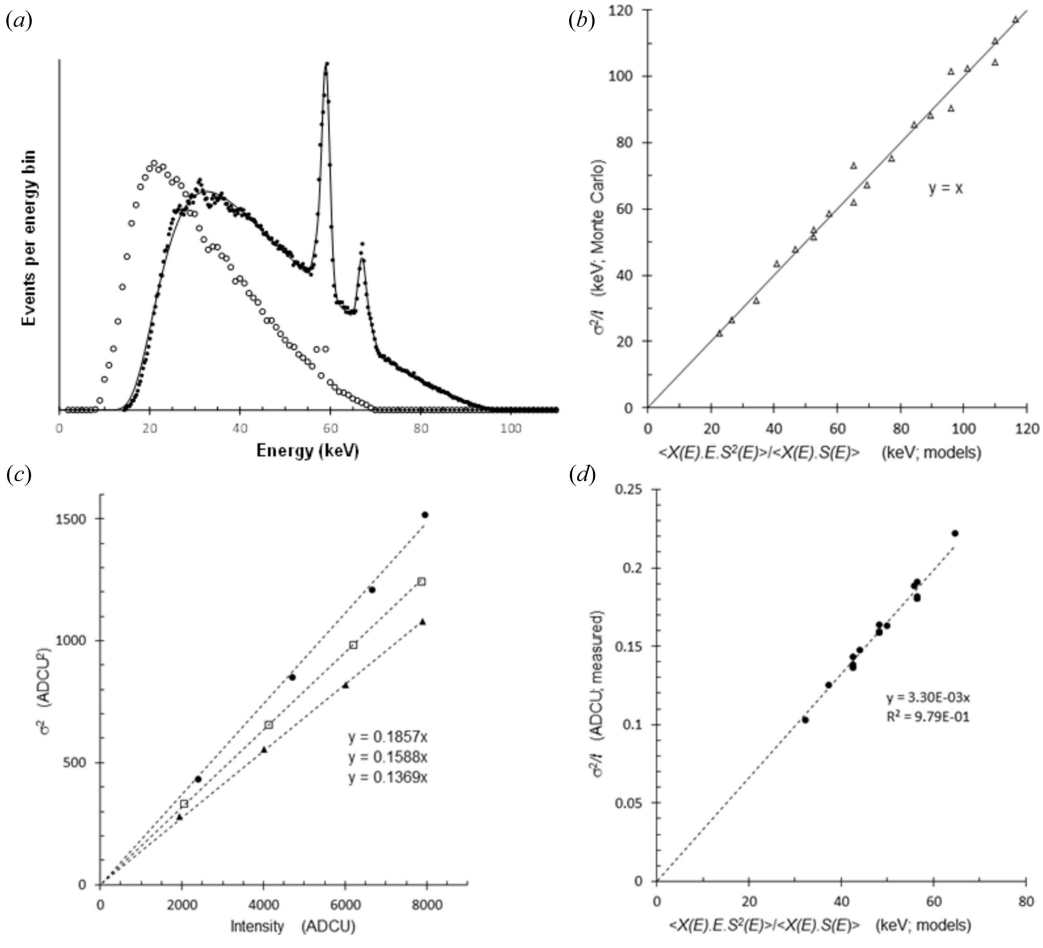


Figure 2
(*a*) An observed Amptek spectrum after spectral stripping (closed circles), the agreement obtainable in our models (line), and a typical Monte Carlo simulated spectrum (open circles; using different model parameters). (*b*) Monte Carlo tests of left- *versus* right-hand sides of equation (4) for several simulated spectra and applying $T(E) = 1$, anticipating direct correlation of the axes, as observed. (*c*) Physically measured data for three distinct spectra, observing three distinct slopes upon adjusting the intensity by tube current. In (*d*) (*cf.* *b*), physical measurements of $\sigma^2/\langle I\rangle$ are plotted against the correspondingly simulated right-hand side of equation (4). Nine measurements combining three filter arrangements and three tube voltage settings are evaluated, together with the data from (*c*). Random scatter about a straight line is seen, supporting the assumption $T(E)$ = constant. (Physically different detectors were used in Fig. 1 *versus* Fig. 2, explaining why their final panels show different slopes.)

polychromatic tomographic data (De Man *et al.*, 2001; Elbakri & Fessler, 2003; Nuyts *et al.*, 2013). Equation (4) thus opens spectral artefact correction and materials identification (Paziresh *et al.*, 2016) using statistical mechanics and a wide range of machine learning approaches to find variational minima.

2.5. Scaling laws

An easily overlooked aspect of Poisson (shot) noise is that the value of $\sigma^2/\langle W\rangle$ obtained from measurements is independent of the expectation number of photons $\langle N\rangle$ applied in the individual intensity measurements, even when $\langle N\rangle \in (0, 1]$. This is seen in the cancellation of N that led to equation (3) and in Monte Carlo simulations (*Appendix E*). Monte Carlo approaches shown there also confirm the analytical anticipation [equation (15)] that variance among sets of $\{\sigma^2/\langle W\rangle\}$ containing M determinations of it scales as $\Delta^2_{(\sigma^2/\langle W\rangle)} \propto M^{-1}$ (here using delta notation to avoid a confusion of sigmas). So in the absence of detector noise or instrumental digitization, the energy resolution of a monochromatic source obtainable by this averaging process is

$$\frac{\Delta E}{E} = \frac{\Delta_{(\sigma^2/I)}}{E} \propto \langle N\rangle^0 M^{-1/2}. \tag{5}$$

With that lack of constraints and repeated measurement, the resolution becomes arbitrarily good according to the standard error in the mean of $\sigma^2/\langle W\rangle$, essentially irrespective of the expectation number of events in individual measurements. In non-monochromatic situations the value of $\sigma^2/\langle I\rangle$ must not be too loosely identified with E (see *Appendix E*); equation (4) is instead appropriate.

2.6. A bound on utility

Instrument digitization must not pin measurement results to particular values, since it eliminates intermediate values that otherwise give meaningful contribution to the variance and average. In this sense, the use of digital signal channels is nothing other than a variant of quantum thresholding, but at a different level. In *Appendix I* it is motivated that if σ_p expresses the total measurable noise (from all sources including human, analog and digitization), and if the digitization increment is a, then the ability to quantitatively measure the index of dispersion is exhausted when

$$M > 144\left(\frac{\sigma_p}{a}\right)^4. \tag{6}$$

Equation (6) bounds the otherwise valuable situation that followed equation (5); digitization causes its own quantum 'bandgap' problem that constrains equation (1) and the Poisson-based uses that follow from it. For example, in the raindrop measurement of *Appendix I*, it curbs any hope of revealing the molecular weight of water! One cannot simply make infinitely measurements of rain [large M in equation (5)] in that experiment using the available digital scales, and compensate for dispersion as in *Appendix E* and equation (4). On the other hand, molecular weight determination from vapour density is essentially independent of the number of molecules due to $\propto N^0$ scaling; there the necessary average comes by *collectively* weighing the molecules in the respective volume, a process for which the digitization error can often be made insignificant in the macroscopic world.

3. X-rays

The X-ray spectrum witnessed by a detector [$S(E)$ in equation (4)] can usually be accurately accounted for using parametric models, as is done below. Thereafter a combination of near-monochromatic and polychromatic X-ray examinations, together with physical understanding of a detector system, offer incremental confidence in the nature and causes of the detector response function $T(E)$, as well as changes of $S(E)$ arising from samples. Such measurements are particularly valuable given that specific details of radiation detectors and their inner workings may be withheld from customers for commercial reasons. The extraction of spectral functions from inner products is known as *unfolding*. Many robotic inference procedures (Jaynes, 2003) for spectral unfolding in radiation contexts have been implemented (Sanna & O'Brien, 1971; Weise & Matzke, 1989; Mukherjee, 2002; Ortiz-Rodriguez & Vega-Carrillo, 2012) and reviewed (Matzke, 2002, 2003), where this selection of references glimpses the history and diversity of applied algorithms. Helpful practical entry points for several of these algorithmic procedures are available to experimentalists (van Dam, 2009). While acknowledging the intimate relevance of those approaches, the present work views the situation from an essentially causative perspective, rather than the symptomatic post-collection perspective used when unfolding. Doing so enables a potentially separable investigation of the roles of spectral dispersion, partial registry, amplifier gains, energy flow, digitization and other aspects, at the same time permitting a collective overview of highly diverse detector systems. Moreover, as a result of this treatment we see that the ratio in equation (4) involves more than just one inner product. For these reasons our title uses the word 'unravelling'; an intention being to help reveal the causes.

3.1. A pseudo-monochromatic example

The sequence of panels in Fig. 1 shows pseudo-monochromatic XRF observations using a single pixel photon energy dispersive spectrometer (an Amptek XR-100CR X-ray detector with PX-4 digital pulse processor unit) in conjunction with a scintillator-based optical X-ray imaging detector (FlatPanel). A General Electric xs|180nf tungsten tube source excited XRF from a range of materials. In these analyses the detector's intrinsic noise, σ_d, manifests as an offset in panel *e* where fitted lines converge on the vertical axis: $\sigma_d^2 = 156$ IU^2 (intensity units). A provisional indication is that $T(E)$ is substantially independent of E over the measured X-ray energy range, leading to a constant slope in panel *f* and a zero-order estimate of $T(E) \simeq 0.0116$ IU keV^{-1} over the range ~20 to ~80 keV. There is insufficient reason to model higher-order energy dependence, given recognisable constraints in

the data thus far available. A pixel intensity signal level of (say) $3\sigma_d$ is achieved only when the total deposited X-ray energy corresponds to $E_{\rm tot} = 3\sigma_d/T(E) \simeq 3230$ keV (equivalent to ∼108 X-ray photons of 30 keV energy arriving in a single pixel). This is not a photon detecting detector by any stretch of the imagination; yet the basic spectroscopic capabilities of this X-ray proportional area detector are seen in panel *f*, *via* interpretation of the X-ray shot noise implicit in equation (3).

A more detailed account of Fig. 1 follows.

3.1.1. Image procedures. Panel *a* sketches how XRF was generated from samples of Ag (foil), Pb (sheet), $BaCO_3$ (powder), Ta (sheet), Nb (foil), SnO_2 (powder), NaI (2.0 mol L^{-1} aqueous solution) and Pt (spatula), using tube voltages of 70, 180, 80, 90, 70, 100, 100 and 120 keV, respectively, and without X-ray filtering. Each XRF spectrum was set up and recorded using an Amptek CdTe detector prior to clearfield collection using a FlatPanel proportional area detector. Darkfield frames (96) and clearfield frames (256) were collected from each sample, from which respective averages were calculated. Individual clearfield frames were then least-squares fitted as linear combinations of those two averages. Given few-percent flux variations from the tube source and smaller systematic variation among darkfields also, the fits prescribe X-ray expectation intensity, and background-subtracted expectation intensity of each pixel in every frame. Pixel locations corresponding to X-ray intensity expectation bins were then identified throughout the 256 frame dataset. Within each bin, the corresponding observed variance from expectation was calculated. In this way, plots of variance against expectation value are shown in panel *e*. They form a continuous range of intensities in each data set, that turn out to be essentially straight lines as anticipated by equation (3). The slopes are manifestly dependent on the XRF photon energy. A common ordinate intercept shows the detector system's inherent noise.

3.1.2. Spectrum aspects. Fig. 1 panel *d* shows selected typical and pathological Amptek spectra. The XRF lines (typ. K_α and K_β manifold) that were the basis for this pseudo-monochromatic experiment are clear, as are small and varying contamination by background spectral features (typ. coherent and Compton scatter of bremsstrahlung used to excite the XRF samples). Anisotropic behaviour of background spectra reaching the area detector and the effects of incident angle on its $T(E)$ were not quantified in these measurements, though such studies could be anticipated as usefully revealing of both sample and detector. Three points on the same curve in panel *e* with different intensities should not fall on the same straight line if their spectrum is different. Observable features in the background spectra therefore warn against over-interpreting the apparently straight lines. Knowing this, no attempt is made to model the possibility of slight nonlinearity of the data presented in Fig. 1 panel *e*; they are simply displayed 'as is'. Nevertheless, in practice the data readily support straight-line fits. Their slopes correlate well with expectation values calculated from Amptek spectra, as shown in panel *f*. There, ordinate values incorporate the observed background spectrum (but exclude the < 20 keV Cd, Te XRF escape peaks in the Ba spectrum; Redus *et al.*, 2009), also spectral weighting by the relative QE curves of the Flatpanel *versus* Amptek detectors (panel *b*). This incorporation of the background spectrum leads to correlation between both detectors of the otherwise unexpected energy inversion of XRF from Ba *versus* I.

3.2. A polychromatic example

Fig. 2 supports equation (4) *via* polychromatic X-ray measurements. Experimental circumstances meant that practical polychromatic measurements used a later model of the same type of FlatPanel brand area detector. By incorporating the bremsstrahlung, the polychromatic measurements extend the measured energy range to ∼20 to ∼160 keV. Here the observed intrinsic (darkfield) detector variance $\sigma_d^2 = 28.3$ IU^2 was subtracted from the measured σ^2, hence the fitted lines in panel *c* extrapolate to the origin. The same subtraction ensures that σ_d is not a potential cause of nonlinearity in panel *d*. In simulating the right hand side of equation (4), $T(E)$ is first assumed constant as motivated by Fig. 1. Those equation (4) simulations give the abscissa in panel *d* for physical measurements shown on the ordinate. Random and small scatter about a straight line further justifies the assumption $T(E) \simeq$ constant for this detector, and the results from panel *c* are incorporated in the fitted value $T(E) \simeq 0.00330$ IU keV^{-1}. The importance of the spectral models in determining the placement of datapoints along the lines in panels *b* and *d*, and the even scatter about those lines, show that shot noise in the proportional area detector again provides an intrinsic measure of spectroscopic information in this polychromatic situation, accessible *via* equation (4). (Additional Monte Carlo work in *Appendix E*3 shows the role of spectral dispersion in these polychromatic data.)

A more detailed account of Fig. 2 follows.

3.2.1. Spectral simulation. The polychromatic analyses of Fig. 2 incorporate spectral stripping algorithms (Redus *et al.*, 2009), and spectral simulations populated by Monte Carlo procedures (see *Appendix E*). To proceed with equation (4), the spectrum received by the detector is modelled

$$S(E) = \chi(E).\Pi_j\{\exp[-\mu_j(E)t_j]\}.\{1 - \exp[-\mu_{\rm abs}(E)t_{\rm abs}]\}. \quad (7)$$

Here $\chi(E)$ is the source spectrum (which potentially also includes temporal, spatial and angular dependencies); the middle product term describes spectral transmission projected through intervening windows, air and sample materials; and the final term is the fraction of incident photons that are stopped in the sensitive absorption region of the detector (quantum efficiency) such that detection may ensue. In typical applications, samples are factored out of the middle product term, but obviously their spectral contribution must remain within equation (4)'s expectation brackets. A few online databases for material $\mu(E)$ coefficients are available; this work used NIST's XCOM database (Berger *et al.*, 2010). Its minimally 1 keV grid includes photoelectric, Compton and small coherent scatter losses and suffices for the available

X-ray spectrometer resolution, while covering the X-ray output energies produced by the tungsten source with tube voltage variously in the range ~10–180 keV. Kramers' law simulates the bremsstrahlung spectrum, adding XRF contributions when the tube anode potential exceeds the respective inner shell binding energy. Fits to Amptek spectrometer observations provide the relative magnitudes of spectral components (K for bremsstrahlung and a_k for XRF), thus

$$\chi(E) = K\left(\frac{E_{\rm tube}}{E} - 1\right) + \sum\nolimits_k a_k \delta(E - E_{\rm XRF}) \qquad (8)$$

Substituting equation (8) into (7) completes the received spectrum $S(E)$, followed by Monte Carlo population (*Appendix E*). Part *a* of Fig. 2 shows typical examples of a measured spectrum, a spectral fit, and a spectral simulation, applied in the following context.

This parametrically modelled X-ray spectrum may be shaped in two ways: by filtering through varying thicknesses of different materials, and by adjusting the tube voltage, as evident in equations (7) and (8). Panel *b* shows the outcome from several simulations that did this. Since no actual detector exists in such simulations, an ideal detector $T(E) = 1$ may be modelled. That presupposes direct 1:1 correlation of the ordinate and abscissa, as this Monte Carlo test confirms.

3.2.2. Polychromatic experiments. In actual measurements of physical polychromatic radiation, the spectrum was received by a 700 μm CsI scintillator-based FlatPanel detector with 194 μm lateral pixel dimensions, its windows and stopping power modelled within equation (7). Fig. 1 *versus* Fig. 2 used two different FlatPanel detectors whose $T(E)$ differ, presumably due to different gain settings of an internal preamplifier [see *Appendix A* equation (13)]. Two types of physical polychromatic experiment were undertaken, as follows.

In both experiments, each measurement used 100 radiographs collected 302 mm from the source. The mean and variance of associated darkfield images was subtracted from the respective intensities and variance of clearfield images. Only the central 128 × 128 pixels were then considered, to avoid beam profile compensations away from the central axis. The mean intensity was calculated for each pixel, and the squared deviance $(I - \langle I\rangle)^2$ calculated from this. Finally all 16 384 mean intensities and squared deviances were averaged to give single estimates of intensity and variance.

The first type of experiment used 90 kV tube voltage with additional filtering by (*a*) nothing, (*b*) 2 mm Al and (*c*) 2 mm Al plus 0.25 mm Cu; yielding three distinct spectra. Intensity was then varied using the tube current. According to equation (4), plots of variance *versus* intensity should yield three linear trends whose slopes depend on the spectrum, as observed in Fig. 2(*c*).

The second type of experiment correlated the observed $\sigma^2/\langle I\rangle$ with the right-hand side of equation (4), applying equations (7) and (8) using the parameters of physical experiments. Fig. 2(*d*) shows the result, being a physically measured variant of panel *b*'s Monte Carlo simulations. Here, nine different spectra used the same three filter arrangements, with 60, 90 and 120 kV tube voltage. Intensity in the recorded flatfield images was kept approximately constant by adjusting the tube current. The zero-order approximation $T(E)$ = constant is an appropriate starting point and assumed in (*d*). A first-order energy term was also tried in $T(E)$ to reduce the scatter of points. However, as in Fig. 1(*f*), even a first-order correction is unwarranted by the available data. Models of $T(E)$ shall inevitably need adjustment as relevant details of the detector's internal workings come to light, as the nature of measured data diversifies, and as their quality and quantity increase.

4. Two corollaries

We close with two simple and foreseeably useful results that appear to follow from equation (4).

4.1. Propagation

When aspects of an experiment are changed, observations of $\sigma^2/\langle W\rangle$ will change, and a relationship between the values is anticipated. Adding a material filter in a polychromatic beam is an example; it alters measurable intensity (here $I = W$) and its variance, and may be associated with a calculable spectrum. That example allows developing the following argument. (Note that a change might alternatively come about by altering source or detector parameters; analyses follow a similar path.) The non-filtered intensites $\{I_{1n}, I_{2n}, \ldots\}$ and the preserved intensities after filtering $\{I_{1p}, I_{2p}, \ldots\}$ provide values for $(\sigma^2/\langle I\rangle)_n$ and $(\sigma^2/\langle I\rangle)_p$, while the effective intensity that 'disappears' due to the filter is

$$\langle I\rangle_d = \langle I\rangle_n - \langle I\rangle_p. \qquad (9)$$

These intensities are the corresponding cumulant sums in the denominator of the first equation (4), noting that the expectation intensity lost to the filter $\langle I\rangle_d$ is, like the other intensities, calculable from the incident spectrum and the filter's material and thickness parameters (of course incorporating the detector's spectral response). In the numerator of the first equation (4), the stochastic filling of energy bins according to their expectation distribution leads to noncorrelated quadrature addition. Now the key is that for the non-filtered variance, this allows its quadrature sum's separation into two parts: components that shall disappear in a filter, and components that shall not

$$\sigma_n^2 = \sum_i \sigma_{n,i}^2 = \sum_i \sigma_{d,i}^2 + \sum_i \sigma_{p,i}^2 = \sigma_d^2 + \sigma_p^2. \qquad (10)$$

Combining equations (9) and (10) shows how $\sigma^2/\langle I\rangle$ values propagate

$$\left(\frac{\sigma^2}{\langle I\rangle}\right)_n = a\left(\frac{\sigma^2}{\langle I\rangle}\right)_d + b\left(\frac{\sigma^2}{\langle I\rangle}\right)_p, \qquad (11)$$

here abbreviating $a = \langle I\rangle_d/\big(\langle I\rangle_d + \langle I\rangle_p\big)$ and $b = \langle I\rangle_p/\big(\langle I\rangle_d + \langle I\rangle_p\big)$ for the measurable intensity ratios. Observe that $(\sigma^2/\langle I\rangle)_d$ is not usually a measured quantity. [It could be, for example when measuring I_0 *versus* I in transmission XAS, or in other 'stacked' detectors (Hanley & Denton, 2005),

whereby equation (11) could be verified. Moreover, the expectation expression for $\langle I\rangle_d$ permits Monte Carlo estimates of its associated σ_d^2 if necessary.] The value of $(\sigma^2/\langle I\rangle)_d$ inferred from equation (11) allows parametric estimates of the absorbing material and the spectrum consumed by it, since it must be consistent with equation (4) in its own right.

4.2. Extension to photon energy-dispersing detectors

Equation (4) can be altered in an interesting way, in energy-dispersive photon detectors that can additionally reveal the full spectrum of the witnessed instances of W. Where that is possible its left-hand side (LHS) denominator may be written $\langle W(E)\rangle$ where the averaging is now over the measured energy bins as well as the multiple recorded spectra, while retaining direct equality to the first right-hand side (RHS) denominator. In that situation there is an opportunity to equate the denominator and numerator on the RHS by further incorporating energy bins E_i and the loss spectrum $T(E)$ in the inner product of the denominators, thus

$$\frac{\sigma^2}{\langle W(E)ET(E)\rangle} = \frac{n_1E_1^2T^2(E_1) + n_2E_2^2T^2(E_2) + \ldots + n_bE_b^2T^2(E_b)}{n_1E_1^2T^2(E_1) + n_2E_2^2T^2(E_2) + \ldots + n_bE_b^2T^2(E_b)} = 1 \quad (12)$$

The simple expression that results: $\sigma^2 = \langle W(E)ET(E)\rangle$ suggests its value where quantum energy dispersing imaging detectors are used in pileup mode.

5. Conclusion

A measure of the contribution of individual quanta may be had from macroscopic measurements involving the index of dispersion. Where linear response conditions are met *and* the measured quanta are monodisperse, the index of dispersion $\sigma_W^2/\langle W\rangle$ is proportional to the mean contribution of each quantum [equation (3)]. The Fano factor (Fano, 1947) expresses that proportionality inasmuch as it relates to partial registry of quanta by the measurement apparatus. Otherwise, energy dependence of losses $T(E)$ and dispersion of the spectrum $S(E)$ are involved according to equation (4). The situations where it lies behind the provision of quantum information are diverse. For electromagnetic radiation, requirements are that the measurement apparatus' response is not compromised by any measurement nonlinearity at the level of the quantum (thresholding and exponential gain being particularly offensive but common situations), and that the quanta obey a Poisson distribution in their contribution to witnessed signals W_i. Under these conditions $\sigma^2/\langle W\rangle$ can act as a metric of the quanta's individual content when M measurements are made, although only up to the point where equation (6) applies. Equation (6) invokes variance due to digitization/pixellation intervals. When that happens it is time to review and upgrade aspects of the experiment. Parametric models of $S(E)$ and $T(E)$ can be substantially examined *via* equation (4). This allows insight into an often 'black box' chain of events within apparatus, and the determination of samples by hyperspectral approaches employing statistics. A close relationship exists to unfolding procedures, which are thereby usefully extended.

APPENDIX *A*
Expansions of equation (1) and the part played by *T*(*E*)

Equation (1) contains terms expandable according to the context. Thus, for X-rays, the photon energy may be written $E = h\nu$. The number of quanta captured by a detector may be considered in its spectral form, as is done, for example, in equations (7) and (8).

The transmission function $T(E)$ in equation (1) covers a great multitude of evils. The mere stopping of a quantum within the sensitive region of the detector by no means guarantees that its energy content will be fully measured, and therein lies much of the nature of this work. In X-ray contexts, the transfer of energy within the system to produce the eventual witnessed measurement may involve many fundamental changes, involving ion and photoelectron generation, partial (or complete) thermalization, surmounting of bandgaps, exciton formation, fluorescence generation, analog current and voltage manipulations *etc.* with partial losses and introduction of noise at each step. Our formalism intentionally retains flexibility for a huge variety of possible physical causes, corresponding to a great variety of detector signal chains that are known to exist. The fraction of a quantum's energy that is actually deposited in the detector may vary depending on the precise location within the absorber where its wavefunction opens a new energy channel, and for many reasons the subsequent transformations and losses leading to an eventually witnessed signal can depend on the quantum energy. The factor *T(E)* contains all these matters. The variety and nature of conversions and losses obviously depends on the particular detector, though in broad terms one might be tempted to write something like:

$$T(E) = A(E)B(E)C(E). \quad (13)$$

Here $A(E)$ could permit description of non-detected heat formation in non-calorimetric detectors, recombination losses, finite optical and charge collection efficiencies; *B(E)* could cover opamp gains, and potentially nonlinear processes such as avalanche gains, while $C(E)$ can describe matters such as finite dynamic range, offset errors, artificial or intrinsic thresholds, activation barriers and sources of spectral redistribution. Given a particular detector, one typically has reason to quantitatively examine and research all such aspects to the extent permitted by warranties and industrial product disclosure agreements. Clearly, for present purposes we must be satisfied with general and broad awareness of the roles played by the many factors that always do contribute to $T(E)$. Bandgaps and any artificial discrimination or thresholding, digitization intervals, nonlinear gains, and spectral redistribution are matters of particular awareness throughout this work.

APPENDIX *B*
The Poisson distribution

$$P(\langle x\rangle, k) = \frac{\langle x\rangle^k \exp(-\langle x\rangle)}{k!} \tag{14}$$

can be derived in various ways, allowing substantial decoupling from matters of temporal resolution (see *Appendix F*).

Given the imposition of $T(E) = 1$ permissible in theory, the Poisson distribution has the distinguishing feature that $\sigma^2 = \langle x\rangle$ (Prigogine, 1978). Its index of dispersion is thus $\sigma^2/\langle x\rangle = 1$. In the context of this work this may be interpreted as saying that for the theoretical ideal system with $T(E) = 1$, its index of dispersion is nothing other than the quantized interval between successive integers in the denominator of equation (14); in other words, $\Delta k = 1$.

For sufficient k the Poisson distribution may be considered almost continuous in k. The situation motivates the Stirling approximation $\ln k! \simeq k \ln k - k$ for the high k values typical in statistical mechanics; factorials may also be rendered continuous using the Γ-function, being valuable for calculus operations.

A feature of the Poisson distribution is its very rapid convergence to Gaussian shape (Kittel & Kroemer, 1980) even for quite small $\langle x\rangle$ (*Appendix G*). Measurements of residuals in photon counting experiments are typically quite unable to distinguish the two. By appropriate normalization of known-Poisson residuals, observation of Gaussian distributed residuals becomes a strong indicator that no further parameterization of a given dataset is possible (Uhlig *et al.*, 2013).

APPENDIX *C*
Standard deviation essentials

We use the usual notation for variance σ^2 and its positive square root σ, the standard deviation. For observations $\{x_1, x_2, \ldots, x_M\}$ of an observed random variable x, the mean $\langle x\rangle = \left(\sum_i x_i\right)/M$. By definition

$$\begin{aligned}\sigma^2 &= \langle x^2\rangle - \langle x\rangle^2 \\ &= \sum\nolimits_i (x_i - \langle x\rangle)^2/M.\end{aligned} \tag{15}$$

The first equality is valuable when given only the histogram of outcomes from the process, and to parameterize continuum situations. As a centrally important example of its use, a top-hat function of width a associated with pixellation or digitization (Thompson *et al.*, 2002) contributes the following variance by evaluating the integral over the binwidth

$$\sigma_a^2 = \frac{1}{a}\left\{\int_{-a/2}^{a/2} x^2 \mathrm{d}x - \left(\int_{-a/2}^{a/2} x \mathrm{d}x\right)^2\right\} = \frac{a^2}{12}. \tag{16}$$

It often appears in detector design in connection with the least significant digital bit of analog to digital converters, in which context it may be referred to as 'quantization' noise (Janesick, 2001; Holst & Lomheim, 2007). The second equality in (15) comes into play when faced with a corresponding dataset. It is possible to make a 'running' estimate of the standard deviation that grows as more of a dataset is absorbed into the calculation, just as a 'running' average is possible. Quadrature addition of uncorrelated variances is equivalent to convolution of the corresponding Gaussian distributions.

Using equation (15) the index of dispersion becomes

$$\frac{\sigma^2}{\langle x\rangle} = \frac{\langle x^2\rangle}{\langle x\rangle} - \langle x\rangle = \frac{\sum_i x_i^2}{\sum_i x_i} - \langle x\rangle \tag{17}$$

which also applies to any distribution. The index of dispersion has the same units as x, being a clue to its significance.

The witnessed observables x_i may have arrived *via* the central limit theorem (*Appendix G*), as happens in our X-ray measurements. In the case of non-calorimetric X-ray detection, intensity observations are the post-calibration consequence of measurements that access only a fraction of the original X-ray photons' energy (see *Appendix F* and *Appendix H*). Each x_i in the dataset is burdened in this way; it carries a part of the energy E that is its root cause, but it also carries the scars of transfer through all aspects of the detection process and measurement apparatus, $T(E)$ [equation (13)].

Progress often starts by supposing that $T(E) = 1$, anticipating $\sigma^2/\langle x\rangle = 1$ for a Poisson distribution, as described earlier. Further progress then comes from deeper analysis of $T(E)$ and recognition of the role played by spectral dispersion in equation (4), which is an extension of the monochromatic equation (3). In the remaining appendices, evidence is accumulated that equation (24) is a closer expression of physical reality than equation (3), although both are valid.

In its original context (Fano, 1947), the Fano factor expressed essentially the fraction of monochromatic parent quantas' energy that can be accounted for in statistically observed measurements. As such it is intimately connected to the partial registry described in *Appendix H*. In other words, the Fano factor may be said to equal the measured quantity $ET(E)$ (this being the instrument-related factor that normalizes $\sigma^2/\langle x\rangle$ upon practical measurement of a Poisson process), but only for a strictly monodisperse situation as described by equation (3). Because spectral breadth acts to increase the index of dispersion *via* equation (4), the interpretation of a Fano factor may be rather complicated in practice, and perhaps not always appropriate.

APPENDIX *D*
A random walk/Brownian motion analogy

Consider a flat field radiation image on an ideal proportional X-ray area detector (no intrinsic noise), and imagine each measured pixel intensity is represented by the length of a piece of string. The analogy facilitates logical connection to one-dimensional random walk diffusion, with its mature contexts in many disciplines. (The string length might instead be governed by the start and end positions of an ant, which for every step forward randomly decides to take a step either forward or backwards, the total number of steps chosen according to a Poisson distribution.) Unlike radiation detectors, there are normally no surreptitious losses when

measuring string lengths. Here the string lengths will be distributed with mean $\langle x\rangle$ and standard deviation σ. Also importantly, for any individual pixel within the field, its repeated measurement will give those same values of mean $\langle x\rangle$ and standard deviation σ. For a Gaussian distribution the probability of deviance by amount x from the mean $\langle x\rangle$ is

$$P \propto \exp\left(-\frac{x^2}{2\sigma^2}\right). \tag{18}$$

Statistical mechanics arguments show that the probability P of ending such a walk an amount x different from the mean $\langle x\rangle$ is approximately (Atkins, 1998)

$$P = \left(\frac{2}{\pi N}\right)^{1/2} \exp\left(-\frac{x^2}{2N\Lambda^2}\right), \tag{19}$$

where N is the total number of steps and Λ is the average step size. By relating the coefficients of the exponents of (18) and (19) (as is done to arrive at the Einstein–Smoluchowski equation; Atkins, 1998; Mott & Gurney, 1948)

$$\sigma = \sqrt{N}\Lambda. \tag{20}$$

A key property of the Poisson distribution is $\sigma^2 = x$, so that as implied by equations (1)–(3)

$$N = (\langle x\rangle/\sigma)^2. \tag{21}$$

Substituting (21) into (20) and rearranging finally reveals the anticipated result for this lossless analogy: equality between the average step size and the index of dispersion

$$\Lambda = \sigma^2/\langle x\rangle. \tag{22}$$

Brownian motion mathematics has been widely studied for its strong relevance to statistical sciences. It may be useful to mention a potential pitfall involving the limit of a ratio (Klafter *et al.*, 1996) as opposed to the ratio of a limit, described further by Jaynes (1991).

APPENDIX *E*
Monte Carlo simulations

Spreadsheets easily demonstrate several key points of this work. This work used Microsoft Excel and Python according to scripting needs.

E1. Procedure

To populate an evenly binned $S(E)$ spectrum [*e.g.* equations (7) and (8)], the normalized cumulant function of $S(E)$ is formed. In its range, each energy bin thus acquires the width $S(E_i)/\left(\sum_i S(E_i)\right)$, collectively spanning the interval [0,1). Next, a field of maximally entropic random numbers is established on the same interval [0,1), with their total count N elected according to a Poisson distribution about an expectation value $\langle N\rangle$. By raining the random numbers into the cumulant intervals, each bin of $S(E)$ randomly receives events in proportion to its local value of $S(E)$. The corresponding Monte Carlo histogram of $S(E)$ is built in this way. In other words, its total population is Poisson distributed, as is each bin, with each bin having the required expectation value of $NS(E_i)/[\sum_i S(E_i)]$ in the limit of many simulations. This is the Poisson-populated maximum-entropy Bayesian approach indicated by Jaynes (1983); it can equivalently be seen as a Monte Carlo approach to Lebesgue's integration (Poole Jr, 1998). [As a practical approach it is worth noticing that because of this bin-based approach, the outcome is subject to the form of variance described by equation (16). While not a problem in the immediate context, it can be avoided by forming a list of maximally entropic random numbers that, when arranged in order, may be effectively mapped to the cumulant function in a linear fit (Jaynes, 1991).]

As a simple but key demonstration, using the Poisson distribution of equation (14) with some parametric value of x, each value in the field of random numbers can be mapped to its corresponding k value. Given a large field of random numbers, this Monte Carlo approach confirms that $\sigma^2/\langle x\rangle \simeq 1$ by either histogram or set-based approaches of equations (15) and (17). However, if each looked-up k value is now multiplied by a factor g, then the same histogram or set-based approaches yield $\sigma^2/\langle x\rangle \simeq g$, as expected from the role of $T(E)$ in equations (3) and (4).

E2. From monochromatic to bichromatic

Fig. 3 shows an informative example populated by this Monte Carlo approach. It is an artificial bichromatic situation, applied as an analytical test on the way to incorporating more advanced spectral models, to ensure the expected quantitative behaviour. Two photon energy bins (E_1 and E_2) with nominally equal population $n_1 = n_2$ diverge in energy from their fixed average value $S(E)$, such that $E_1 = S(E) - \delta$ and $E_2 = S(E) + \delta$. The average total intensity obeys $i_1 + i_2 = I$ and the intensities' standard deviations add in quadrature: $\sigma_1^2 + \sigma_2^2 = \sigma^2$. Equations (1), (2) and (4) are obeyed, and for now assume a mathematically ideal detector [$T(E) = 1$]. With these conditions, equation (4) simplifies to

$$\begin{aligned}\frac{\sigma^2}{I} &= \frac{n_1E_1^2T^2(E_1) + n_2E_2^2T^2(E)}{n_1E_1T(E_1) + n_2E_2T(E_2)} = \frac{E_1^2 + E_2^2}{E_1 + E_2}\\ &= \langle S(E)\rangle + \frac{\delta^2}{\langle S(E)\rangle}\\ &\geq \langle S(E)\rangle.\end{aligned} \tag{23}$$

The result is quadratic with respect to the difference δ and is plotted as the solid curve in the second panel of Fig. 3. In particular it is not a flat line, as it would be if equation (4) merely gave the expectation energy, $\langle S(E)\rangle$. Populating the spectra using the procedure described above agrees well with equation (23), supporting the viability of both the Monte Carlo code and equation (4).

The observation in equation (23) and Fig. 3 warns that an interpretation $\sigma^2/\langle x\rangle \propto \langle S(E)\rangle$ is fundamentally wrong (it is *not* the mean value), *except* in monodisperse situations. Rather, it shows to expect an estimate that will be slightly high [by a fraction of $\langle S(E)\rangle$], the overestimate increasing according to how disperse the spectrum is. The Cauchy–Schwarz inequality offers generalization in this context (Polad, 2008).

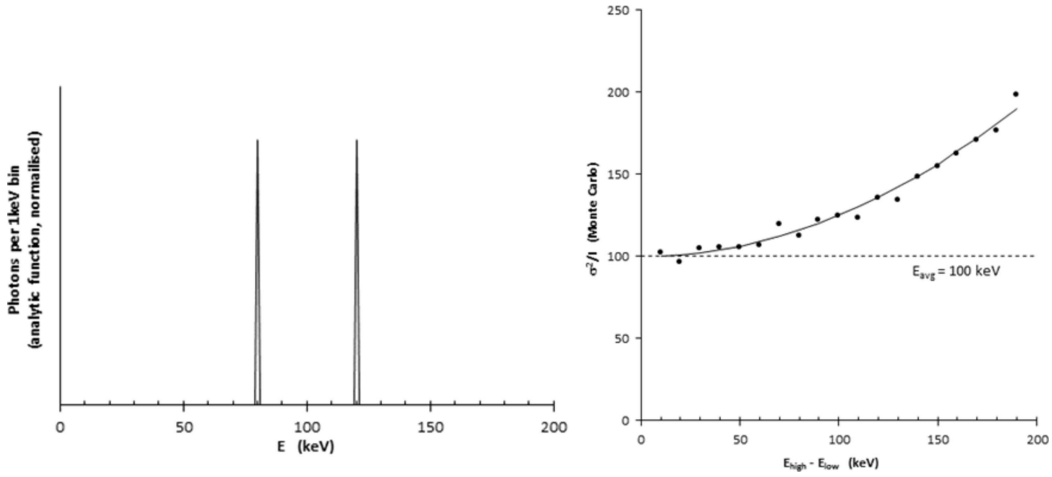


Figure 3
A spectrum consisting of two nominally equally populated peaks is separated by an interval. This may be Monte Carlo populated with events, using the scheme described in the text. Each point in the plot at right shows the outcome from a large number of such simulations, assuming ideal detection [$T(E) = 1$]. As the gap between the peaks grows, the simulated values of σ^2/I diverge from the mean, in accordance with equations (23) and (4).

The name 'index of dispersion' is clearly appropriate for the ratio $\sigma^2/\langle x\rangle$, with the example above illustrating that it can be at most 2 in the extreme case of divergent dispersion in lossless measurements. The character of the test in Fig. 3 suggests its roots in the calculation of central moments in statistics, which is more clearly evident in equations (15)–(17).

A further aspect of this demonstration is that $\sigma^2/\langle x\rangle$ has the potential to indicate an average value for the quanta where in fact none of them actually have that precise value. (For example, vapour density measurement of a mixed gas such as air will not give the molecular weight of any actual existing molecule; to be effective, the gas must be pure.)

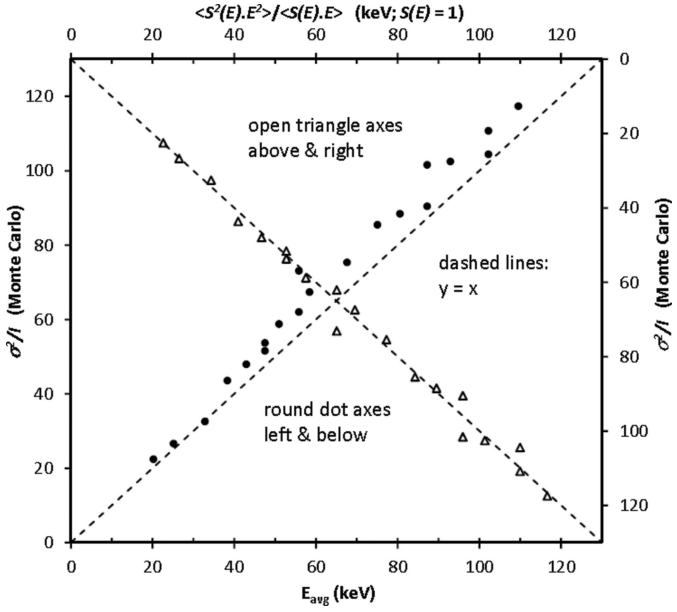


Figure 4
This extends Fig. 2(*b*), showing how polychromatic spectral dispersion alters the interpretation of the index of dispersion away from the average energy $S(E)$. Here Monte Carlo tests of equation (4) used bremsstrahlung and material filtering spectral simulation as per equations (7) and (8). Each of the closed circles (read from left and lower axes) shows σ^2/I from a set of simulations; scatter in the results are seen to slightly overestimate the average photon energy $S(E)$. Open triangles (read from upper and right axes) show the same results plotted against the argument of equation (4).

E3. Dispersion in the polychromatic case

Fig. 4 extends the test of Fig. 3 to the full spectral simulation of material filtered bremsstrahlung, applying equations (7) and (8) with $T(E) = 1$ using the Monte Carlo procedure inspired by Jaynes (1983). Again in this broadband case, it is seen that the polychromatic spectral dispersion leads to good correlation of σ^2/I with equation (4), while fractionally overestimating the simple average $E_{\rm avg} = \langle S(E)\rangle$.

E4. Tests of scaling laws

Other powerful results can be had from Monte Carlo simulations. The following are obtained by this approach, independently of prior expectations and without noise sources that often frustrate real-world experiments. Consequently, they are strong indicators of the correctness of (or mistakes in!) corresponding analytical arguments. Simulations in Fig. 5 demonstrate that:

(i) The error in σ^2/I is independent of the expectation number of photons in individual measurements.

(ii) Variance among measurements of σ^2/I scale as M^{-1}, where M is the number of measurements made. Equivalently: $\sigma_{(\sigma^2/I)} \propto 1/\sqrt{M}$ (like the standard error in the mean).

Unsurprisingly, further analogous simulations show that the latter two outcomes apply equally well in polychromatic cases.

APPENDIX *F*
A pause on temporal aspects

Electromagnetic intensity is regularly encountered expressed as a flux; the energy through a specified increment of [position$_{x,y}$, divergence$_{x,y}$] phase space, per unit of *time*. The

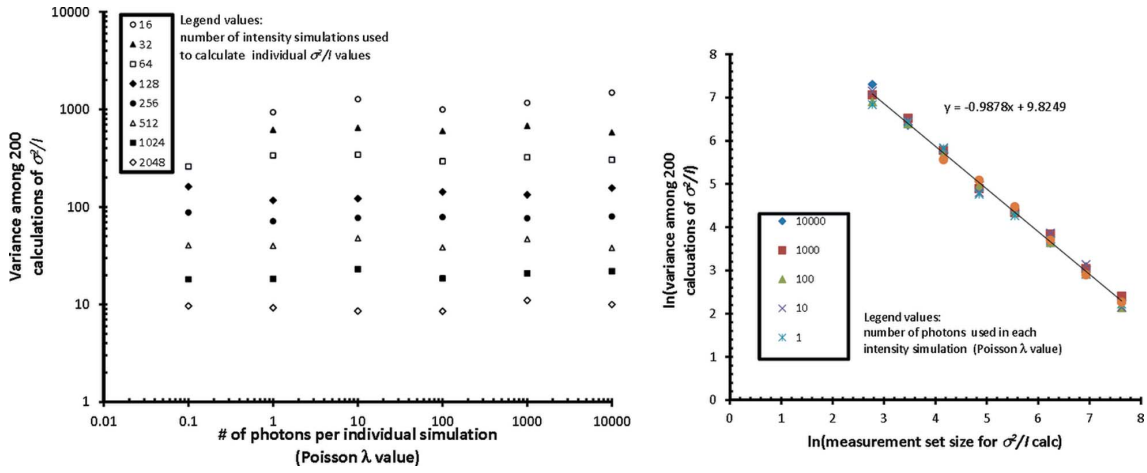


Figure 5
Each point in the left panel shows the variance among 200 Monte Carlo calculations of σ^2/I for a monochromatic 100 keV spectrum. Each of these 200 σ^2/I calculations simulated a set of intensities, where the number of individual simulations (M in the text) is denoted by legend symbols. Within each individual simulation, the Poisson expectation value for photon observation (N in the text) is indicated on the horizontal axis. The right panel is a log–log plot using $\ln(M)$ as the ordinate axis; the fitted slope confirms the scaling $\sigma^2_{(\sigma^2/I)} \propto M^{\sim -1}$, without recourse to algebraic analysis.

present work has deliberately avoided expressions involving time. To see why, consider first a continuously emitting source. Our concern is with situations where stochastic quantization of the energy applies. In that very traditional situation, it would seem only natural to attend to exposure durations. However, a modified form of mathematical support is then needed in ultrafast situations. There, photons travel as a packet whose physical length is far shorter than the length of the detector absorber in which their effects shall statistically manifest. In the latter situation, what temporal basis is appropriate for the observed flux? A temporally based formalism might be arranged that encompasses a continuum of such situations – however, its inclusion of the temporal domain would find subsequent needs to encompass the timescales of various primary processes (Hart & Anbar, 1970), gain-bandwidth products at the levels of primary physics (Rose, 1963), spectral distribution of power associated with detector noise sources (Robinson, 1962), secondary amplification (Robinson, 1962; Horowitz & Hill, 1989) and their inevitable Fourier transformation into spectral resolution. Meanwhile, accounting for those aspects does nothing to clarify or dispel the actual objective of this work, namely, to witness the role of the index of dispersion as a metric of quanta. To be sure, temporal expressions are absolutely mandatory in their respective contexts, and can only be collectively consistent with our present quarry. As a developmental analogy, the Maxwell Boltzmann distribution for molecular velocities is necessarily consistent with the ideal gas law – yet the ideal gas law is mostly preferred when entering a description of gases. Desirable and inevitable consistency across broad timescales (Coppens, 2015) must here be moderated by a kind of artificial horizon (Jaynes, 1991). So we have sought to relieve analyses of temporal aspects where possible, at least for now.

The present quest appears serviceable by a time-independent view of statistical mechanics, in which temporal aspects can be at least allayed until they are required, as when dealing with temporal fluctuations (Dugdale, 1966). In taking this approach, we observe the equivalence of time averages for stationary processes, and ensemble averages, typically denoted $\bar{x}$ and $\langle x \rangle$ respectively (Robinson, 1962).

F1. Typical situations

In radiation detection contexts, the $T(E)$ function invariably *is* a function of time. This relates to the fact that different physical phenomena are measurable as detectable signals, on different timescales, depending on the detector. All too often, X-ray detection has evolved from what is the easiest or most empirically convenient trap-state symptom that can be measured. All trap states (energetic photoelectrons, electrons in conduction bands, colour centres, excitonic states, fluorescence *etc.*) must eventually thermalize and dissipate their energy as heat (in the end motivating calorimetry). A complex flow of temporally overlapping intermediate states can be anticipated (Schmidt *et al.*, 2003), and is inevitable in any detector system. Fig. 6 implies some typical intermediate processes *via* time-resolved photography of a cloud chamber (Wilson, 1912). Insights from such images have been fundamental to charged particle and X-ray detector design. In the photographed situation, either the scattered light or the electrically measurable ionization are typically considered as measurable output signals (*cf.* ionization chamber). The optical signal is made ill-conditioned by a diffusion-moderated avalanche-gain process, further involving the fundamentally independent energy of the scattered light (Fullagar, Paganin *et al.*, 2010; Fullagar *et al.*, 2011*a*). For the electrical signal, spontaneous charge recombination has consequences in the measured temporal domain, at the same time implicating the independent energy supply of its electronic amplifier's power supply (Washtell, 1958). Depending on bias voltage and geometric configurations one may end up with a gas-proportional detector in the sense used by Fano (1947) or a Geiger tube (Strong, 1938). In truth then, either signal evolves

with time in nonlinear ways. Using the first of these two approaches, the images in Fig. 6 next involved a chain of basic physical events (Mott & Gurney, 1948) leading to the photographs. The final outcome is stable on human and indeed *trans*-generational timeframes, today using digital storage techniques' binary information metastably trapped in flip-flop circuits and bistable memory elements (Horowitz & Hill, 1989). Potentially usefully, the light scattered (Fullagar *et al.*, 2011*b*) from the many photoelectron trails should, under suitable conditions, satisfy equation (1), from which the rest follows. The total scattered light will obviously be subject to many parameters including delay and exposure times, refractive indices, temperatures, wavelengths and so forth; but to dwell on them cannot serve immediate interests. Different physical, chemical, optical, thermal, electronic and other detection schemes are each associated with temporal dependency of $T(E)$. Each has corresponding complexity that is typically swept into equation (13).

F2. What non-thermal detectors do

One sees that the purpose of the X-ray detector is to integrate an observable signal arising from one or other ephemeral intermediate state, with most approaches offering a proxy for just a *fraction* of the quantum's original energy on its pathway to thermalization. The detector system must then present this integral to the user in a permanent form amenable to calibration, if the apparatus is to be useful as a proportional detector. The observable phenomenon must manifest on a timescale that is short compared with the integration, while the integration must occur on a timescale that is short compared with the user's needs. The final result represents a calibrated accumulation of whatever quanta are associated with the observable fraction of the photons' original energy. The permanent form can be delivered as a photograph, a logbook recording of a needle's deflection, or binary numbers for digital storage *etc.* The temporal and potentially lossy processes that go on 'under the bonnet' of all radiation detectors are assumed to have been addressed to the best reasonable extent by detector manufacturers when one makes a purchase; and they usually are.

F3. Time and statistical inference in the bigger picture

Being just another part of the electromagnetic spectrum, the frequency (energy) dependence of X-ray noise assessments is necessarily analogous to that observed elsewhere in the spectrum. Noise *versus* bandwidth assessments in electronic circuits (Horowitz & Hill, 1989; Robinson, 1962) translate quite generally into optical (Davis, 1996; Goodman, 1985) and X-ray spectral (Als-Nielsen & McMorrow, 2001) considerations, duly motivating studies of mathematical behaviour in the complex frequency plane in each case. In absorption and scattering experiments matter is viewed as an assembly of coupled oscillators (Thompson, 1919) and much effort goes into calculating dispersion properties of the assembly (Brillouin, 1946; Zangwill, 2015). The assembly is *collectively* interrogated by *individual* quanta (Schrödinger, 1952), according to their energies and wave properties. Despite this nonlocality the quantum can only open alternative energy channels in *one* place according to probability (Rayleigh, 1880, 1919; Norton *et al.*, 1955; Feynman, 1985; Susskind & Friedman, 2014) and the conservation of energy. This occurs with particle-like behaviour, according to the quantum's finite and range-independent energy. Appreciation of this peculiar situation has been an everyday experience for some generations of short-λ diffraction scientists (Bragg, 1962), while the wave-particle delocalization of the quantum and its seemingly magical collapse to open a highly localized detection channel has never been an altogether comfortable fact (Ives, 1951; Jaynes, 2003). Taken together with well founded statistical interpretations (Landau & Lifshitz, 1969; Glansdorff & Prigogine, 1971; Prigogine, 1978), it has been the source of deep speculation (Bohm & Vigier, 1954; Hazelett & Turner, 1979; LaViolette, 2010; Ananthaswamy, 2017). Contemporary explorations of temporal symmetry (Hadad *et al.*, 2016) and other temporal developments (Yao *et al.*, 2017) suggest ongoing relevance to crystallography.

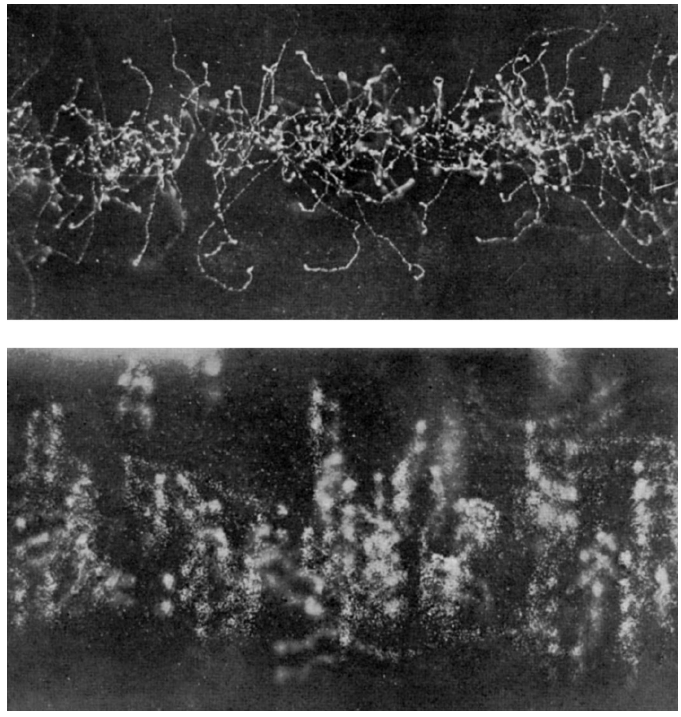


Figure 6
Time-resolved X-ray pump, optical probe photographs of photoelectron trails from an X-ray beam propagating in the left–right direction in a cloud chamber (Wilson, 1912). In the lower image an electric field acting in the vertical direction separated positive and negative droplets prior to the flash of optical illumination that exposed the photograph. A doubling of the trails shows the presence of droplets with both charge, while their diffusion during the temporal interval is also evident. Reproduced with permission.

APPENDIX *G*
The central limit theorem

The central limit theorem is the saving grace of the complexity that otherwise besets the $T(E)$ function. Fig. 7 outlines its

modus operandii. Part (*a*) shows, on average, how a quantum from a monochromatic source might contribute to a witnessable response, as a fraction of its quantum energy, in some different detectors. Normalizing the area under each curve corresponds to assuming that the quantum contributes to a witnessed signal. The quantum's energy is measured with $E/h\nu = 1$ only if it is witnessed entirely as heat, since in any other detection scenario, a portion is invariably lost as (non-detected) heat. Semiconductor detectors register just a fraction of an original X-ray photon's energy (*e.g.* ~1.1 eV/3.65 eV $\simeq$ 0.3 in silicon; the bandgap divided by the average X-ray energy per electron-hole pair; the fraction may be associated with the Fano factor as originally formulated; Fano, 1947). Scintillation or scatter based detectors are worse still, with many optical losses, trap states and recombination phenomena to compound and smear variations in optical fluorescence yields and scatter directionality. Obviously many other detector types exist too. Such a single quantum may or may not be individually detectable, depending on the magnitude of those losses, the energy of the quantum, internal gain in the detector, and the envelope of the detector's noise, σ_d. In the present work we rely on high photon energies far above the scintillator material's bandgap such that equation (1) is valid. Their number N is sufficiently large to cause both the signal ($\propto N$) and shot noise ($\propto \sqrt{N}$) to outgrow the σ_d envelope of detector noise. Whatever the detector, its average response curve to monochromatic quanta could be depicted as in Fig. 7(*a*). Any such curve has a mean value μ_1 and a standard deviation σ_1.

When registering many quanta in a single measurement, their pileup is observed. Just one quantum would have given a detector response somewhere on the horizontal axis in panel *a* of Fig. 7, with likelihood weighted by the vertical axis value. A second quantum adds to this by doing the same; statistically the result is a self-convolution. Subsequent quanta repeat this convolution again and again, P times for P quanta. (The process can also be modelled as a Markov chain.) In the limit of many convolutions, the central limit theorem applies, and whatever function appears in (*a*) smears to a Gaussian, shown in (*b*). This Gaussian has mean $\mu = P\mu_1$ and standard deviation $\sigma = \sqrt{P}\sigma_1$ (quanta's mutual effects are non-correlated in linear regimes of intensity measurement). In the limit $P \rightarrow \infty$, the ratio $\sigma^2/\mu = \sigma_1^2/\mu_1$ is uniquely preserved, being the eigenfunction of the process (a property that is perhaps insufficiently recognized). It should be clear that μ is identifiable with witnessed measurements W and σ with their variation, indicating the connection to equations (3) and (4).

APPENDIX *H*
Partial registry

Multiple forms of partial registry (Moseley *et al.*, 1984; Prigozhin *et al.*, 2003; Fullagar *et al.*, 2008; Redus *et al.*, 2009)

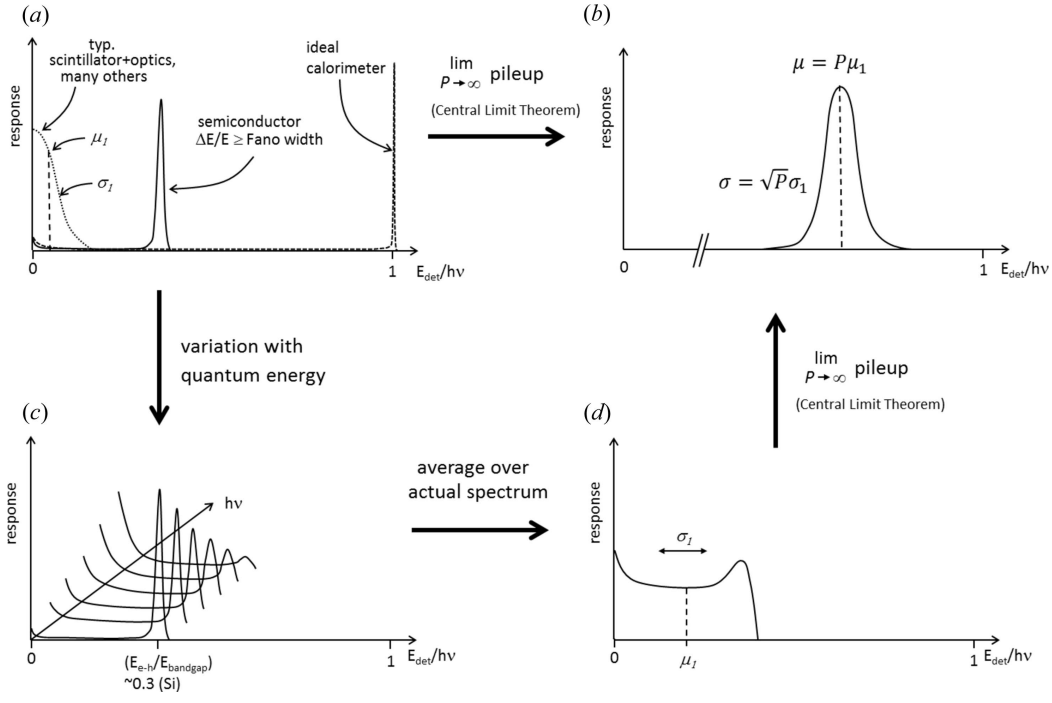


Figure 7
The fractional energy of a monochromatic photon registered by a detector is sketched in (*a*) for a few detector types, and in any case may be less than the detector's intrinsic noise envelope (not shown). Pileup corresponds to repeated self-convolution, which in the limit of many photons leads to a Gaussian distribution *via* the central limit theorem (*b*). In polychromatic cases the detector response can be different for different photon energies (*c*). After modulating the detector response by the corresponding spectral weight, the average photon response may be as shown in (*d*). Their pileup again leads to a Gaussian distribution *via* the central limit theorem.

cause spectral redistribution of observed spectra (Prigozhin *et al.*, 2000). From *Appendix F* and *Appendix G* we see that in non-calorimetric cases a detection procedure witnesses only a fraction of the photon's original energy $\mu_1(E)E$, noting the general energy dependence of μ_1 (Fig. 7*c*). The quantity $\mu_1(E)$ relates directly to the 'internal quantum efficiency' (Hartmann *et al.*, 2006). This fraction is transmitted through a detection chain that hopefully(!) gives an output response proportional to its input signal, but which can contain changes of form of the signal *en route* that introduce more or less deliberate nonlinearities, as outlined in *Appendix A*. Denoting the detector chain's actual response as $T'(E)$, it is now seen that

$$\sigma^2/\langle W\rangle = \mu_1(E)ET'(E), \tag{24}$$

whereby energy dependence can arise either for 'fundamental' reasons [the energy dependence of $\mu_1(E)$], or for reasons relating to the construction of the signal transfer apparatus [$T'(E)$]. In applications, proportionality to the actual photon energy is what makes equation (24) useful. In eagerness to exploit that, one typically lumps $\mu_1(E)T'(E) = T(E)$, as is done in equation (3). The crime can be made with an awareness that permits progress (Uhlig *et al.*, 2011), but sometimes it may also be at the expense of satisfactory knowledge that the resulting observation has potential to be messed up in *two* ways: at a 'fundamental' signal generation level [$\mu_1(E)$] *and/or* by nonlinearity in the signal transfer apparatus [$T'(E)$].

Awareness of the deeper situation is required for the polychromatic extension to equation (4). Such an understanding in due course also motivates cryogenic microcalorimetry (Enss, 2005; Fullagar *et al.*, 2007; Fullagar, Uhlig *et al.*, 2010; Mazin *et al.*, 2013), in which partial registry due to trap state, photoelectron (Fullagar *et al.*, 2008) and X-ray fluorescence (Redus *et al.*, 2009) losses are minimized at the design stage (Moseley *et al.*, 1984). There, $\mu_1(E) \simeq 1$, and in a related way, $T'(E)$ does not suffer the consequences of physical issues (*e.g.* bandgaps), or thresholding/discrimination, or exponential gain (Robbins & Hadwen, 2003; Tutt *et al.*, 2012), that often are imposed between users and the physical phenomena of radiation detection.

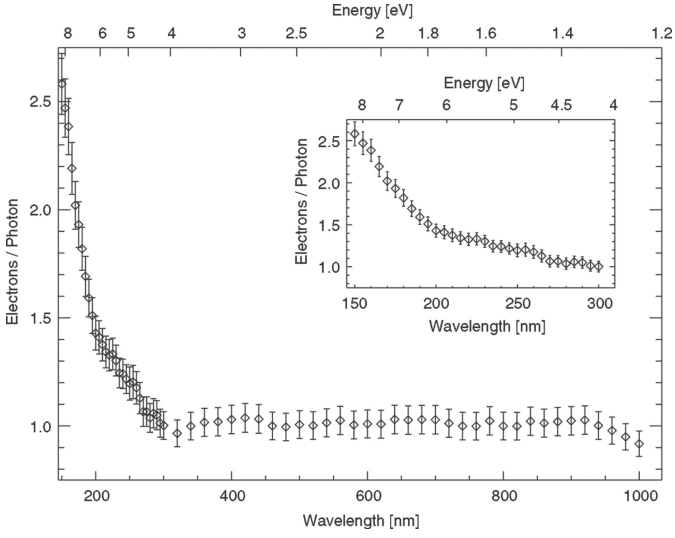


Figure 8
The spectral production of a detectable electronic signal in silicon, reproduced with permission (Hartmann *et al.*, 2006). Like any other non-calorimetric technique, the excitation of an electron over a bandgap cannot produce a corresponding fractional number of conduction electrons if the photon energy is only fractionally greater than the bandgap energy. The correspondingly flat response caused by the bandgap (here extending over the full visible spectrum) removes the essential energy dependence in equation (1) for photons in this energy range.

H1. The role of energy barriers at sufficiently low photon energies

In particular, at optical and lower photon energies, this work's approach would surely have been very broadly applied long ago (*e.g.* for estimation of colours from black and white footage), if not for an important catch! The catch is simply that a threshold photon energy (energetic barrier) exists in non-calorimetric detectors, below which no detectable response is produced other than by thermal or defect-assisted leakage. This governs the nature of $T(E)$ at the corresponding energies and is easily illustrated by example. In silicon the bandgap is ∼1.1 eV, but the electron yield does not increase beyond one until photon energies exceed ∼4 eV, as shown in Fig. 8. These considerations underpin the Shockley–Queisser efficiency limit when presented with blackbody radiation (Shockley & Queisser, 1961) and are an intrinsic aspect within the $T(E)$ term in equation (6). Such barriers are central to our relationship with energy and lie at the heart of debates between photovoltaic and photosynthetic *versus* thermal approaches to energy capture. The equivalent 'chemical' Arrhenius barriers allow chemical energy to be barricaded (Gonick & Criddle, 2005) for use on demand between the hot source of the sun and the cold sink of space (Raman *et al.*, 2014). The resulting capacity for nonequilibrium in turn drives the organization that is characteristic of life (Turing, 1952; Prigogine, 1978). However, in this work the consequence is that no radiation receiver based on a suprathermal energy barrier can satisfactorily obey equation (1) until the energy of the detected quanta is very substantially higher than that barrier. Equation (1) cannot be valid when the quantum to be detected faces barriers comparable to or greater than its own energy. When the proportionality of equation (1) is lost, the utility of $\sigma^2/\langle W\rangle$ as a metric of the quanta breaks down.

The virtues of low-temperature thermal measurement are thus quickly seen: metallic absorbers lack a bandgap except at the quantum pseudo-continuum level. Sensitive calorimeter detectors should thus enable spectroscopic information at optical and sub-optical photon energies *via* equation (4), even when the quanta themselves may not be individually discernible above the detector noise (Mazin *et al.*, 2013; Becker *et al.*, 2014). The latter is the situation that applies in the X-ray experiments of Fig. 1 and Fig. 2, where the X-ray photon energies are vastly greater than the absorbing scintillator's bandgap.

APPENDIX *I*
Information from macroscopic examples

I1. Statistical inference of raindrop size

Signal losses are avoidable in a simple wet-weather experiment, inspired by a 'bucket' analogy for CCD operation attributed to Kristian and Blouke (Janesick, 2001). This experiment has the objective of measuring an average raindrop weight $\langle w \rangle$ by weighing the rainwater collected in an array of cups. As indicated in *Appendix F*, the arrival time of each drop is irrelevant, provided that all drops occur within the designated exposure/integration interval. The rain's overall nature and size distribution will of course vary from one shower to the next, and we have no model for its spectral size dispersion (*Appendix E*3), so will assume it is monochromatic. If one raindrop does not significantly influence others, we can suppose a Poisson distribution for their arrival, both in time, and from one cup to the next. The cups were individually pre-weighed, and water was carefully wiped off their outside surface prior to weighing, so that the mass of water fallen inside each cup could be established. See Fig. 9.

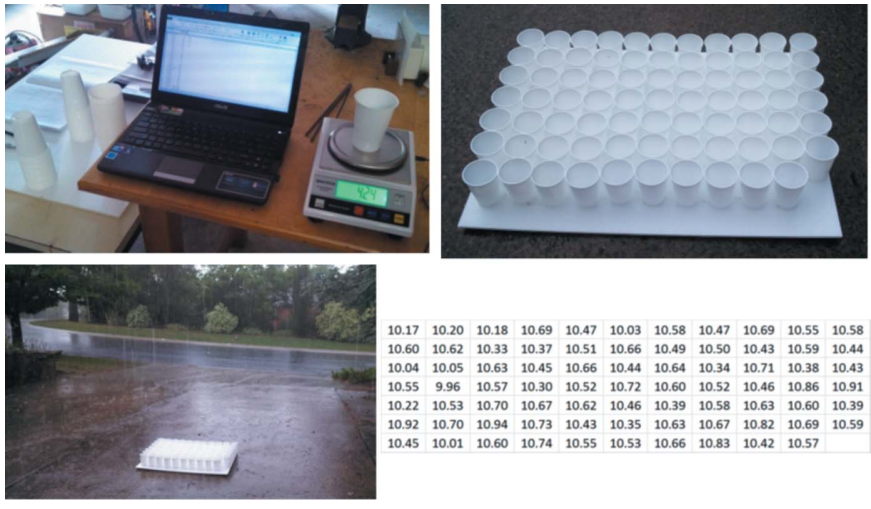
We do not (at first) anticipate any spatial variation. However, the first measurement (photographed) did show some, giving high values for the edge cups due to splatter from the surrounding pavement. This was addressed on subsequent occasions using bricks under the tray to raise the cup lips 28 cm array away from the pavement; but now a second measurement received little rain during a brief shower on a dry day, and noticeable evaporation occurred during the hour it took to measure the weights. A third measurement received an intermediate amount of rain, and obviated evaporation using a cover; the weights are shown in the table at bottom right of Fig. 9. Digital scales had a display error of ± 0.01 g, giving a nominal error of ± 0.02 g after subtracting the dry cup weight. Here the complications appear to have been addressed by the experiment and no patterns could be inferred when plotting weights along the known dimensions of x, y or measurement order. On this occasion we find $\sigma^2/\langle w \rangle = 0.004161$ g (equivalent spherical diameter 2.00 mm). It is presumed slightly higher than the mean collected raindrop diameter due to dispersion (*Appendix E*3).

I2. A non-Poisson example, where futility meets utility

It is useful to consider if any meaning might be attached to a measured value of $\sigma^2/\langle x \rangle$ for a distribution that we know not to associate with a Poisson distribution. The normal (Gaussian) distribution is a special and valuable case to consider. This is because: (i) a Poisson distribution rapidly converges to Gaussian shape with increasing $\langle x \rangle$; (ii) a difference of Poisson distributions is very commonly encountered [*e.g.* from the full or partial nulling of currents in bridge circuits and servo control mechanisms of scientific instrumentation (Moore *et al.*, 2003)]; and (iii) a vast number of everyday measurements lead to Gaussian behaviour.

Given two nulled Poissonian distributions it can be seen how the index of dispersion will fail, as follows. When two equivalent Poissonians each involving the same kind of quanta and each characterized by the same variance σ_P^2 are mutually subtracted, the resulting difference distribution will be symmetric (and if not analytically Gaussian, then ordinarily so close to Gaussian that there is rarely any reason to make a distinction). For an uncorrelated subtraction the variances will add in quadrature, so that $\sigma^2 = 2\sigma_P^2$. The average $\mu = 0$ makes it clear that the index of dispersion σ^2/μ says nothing about the quanta that we know are involved in this example.

That rationale shows more generally to avoid situations where noisy subtractions or additions are involved (such as when estimating expectation values to subtract). The special property $\sigma^2 = \langle x \rangle$ of the Poisson distribution is essential to the denominator if the index of dispersion is to be used as a metric of quanta, and must not be interfered with by such



10.17	10.20	10.18	10.69	10.47	10.03	10.58	10.47	10.69	10.55	10.58
10.60	10.62	10.33	10.37	10.51	10.66	10.49	10.50	10.43	10.59	10.44
10.04	10.05	10.63	10.45	10.66	10.44	10.64	10.34	10.71	10.38	10.43
10.55	9.96	10.57	10.30	10.52	10.72	10.60	10.52	10.46	10.86	10.91
10.22	10.53	10.70	10.67	10.62	10.46	10.39	10.58	10.63	10.60	10.39
10.92	10.70	10.94	10.73	10.43	10.35	10.63	10.67	10.82	10.69	10.59
10.45	10.01	10.60	10.74	10.55	10.53	10.66	10.83	10.42	10.57	

Figure 9
Pre-weighed cups (top left) were arranged as an array (top right) and placed in the rain (lower left). At the end of the collection, the mass of water within each cup was recorded (lower right; masses in grams). One cup in the array was missing (budget constraints).

manipulations. Any uncorrelated distribution added or subtracted will convolve into an observed σ^2. However, as long as that distribution has zero mean it will not alter the index of dispersion's denominator, only its numerator. Quadrature addition in the numerator (arising from convolutions) can only increase it, giving a falsely high index of dispersion where such noise is present. One must be a savvy and careful experimentalist if a measured index of dispersion is to be trusted as a metric of quanta.

I2.1. Establishing a bound on a useful number of measurements – analog meets digital. Situations arise where one has some digital instrument and would like to know how many measurements might be usefully made, before its digitization variance confounds the best efforts of the best experimentalist. To proceed, consider the standard error in the mean $\Delta_\mu = \mu/\sqrt{M}$. Here we use Δ notation (instead of the customary σ), because we shall apply it to the variance σ^2 in equation (15), in which σ^2 is itself expressed as a mean value. Thus

$$\Delta_{\sigma^2} = \frac{\sigma^2}{\sqrt{M}}. \tag{25}$$

In the absence of any digitization intervals, it may be presumed that the accuracy of our knowledge of σ^2 would become arbitrarily good as $M \to \infty$. However, when using digital apparatus the results get pinned according to the digitization increment, and increasing $M \to \infty$ can do very little to reduce the uncertainty. The consequences are contained in equation (16). So we extend equation (25) as follows, using σ_g^2 to denote the contribution of variances due to purely analog sources

$$\Delta_{\sigma^2} = \frac{\sigma^2}{\sqrt{M}} = \frac{\sigma_g^2}{\sqrt{M}} + \frac{a^2}{12}. \tag{26}$$

Viewing the second equality, it is evident that the right-hand side will come to be entirely dominated by the digitization variance as M increases. In due course analog errors become negligible while digital errors remain. Then the remaining inequality may be rearranged to

$$M > 144\left(\frac{\sigma}{a}\right)^4, \tag{27}$$

corresponding to the situation where further measurements will not reduce uncertainties. This is equation (6).

The indications above can be seen in the measurement below, where 160 pieces of ∼5 mm diameter plastic pipe were manually snipped to ∼5 mm lengths for a construction project. The first length was used as an approximate visual template for the remaining ones, but no particular effort was otherwise made. Experience tells us to expect a roughly Gaussian distribution of weights. The snipping procedure could be interpreted as a kind of human servo mechanism that seeks to give a null in the difference of pipe lengths, so from the arguments above, the index of dispersion is not expected to reveal any putative quantized phenomenon. The pieces were individually weighed on the same scales used for the raindrop experiment (with digital interval a = 0.01 g), taking care to re-zero between each item. In this way the digitization interval of the scales rounds down, so the average weight of each digitization bin is 0.005 g higher than observed values. Table 1 shows the result, plotted in Fig. 10 using a Gaussian fit and also a Poisson fit to show the similarity of the two. From the sum of variances $\sigma^2 = \sigma_g^2 + a^2/12$ the analog component is σ_g = 0.0106 g and dominates the error as expected. The bound established above shows that for this measurement $M >$ ∼182 would only be wasted effort on the part of the experimenter, because of the scales' digitization increment. Unless the digitization interval a is reduced, the experimenter can do nothing to better determine real variance (of the kind sought in the index of dispersion's numerator) and thereby deconvolve it from problems worthy of interest; while as soon as that is done, the scaling $M \simeq a^{-4}$ will very quickly test patience unless non-digital aspects of experiment error are also addressed.

Table 1
Weights of ∼5 mm diameter plastic pipe cut to ∼5 mm lengths using snips (human error).

Bin average	0.115	0.125	0.135	0.145	0.155	0.165
Count	7	33	57	45	16	2

I2.2. The role of experimental iteration. Accurate knowledge of the causes of error bars, and thereafter their magnitude, is mandatory in the iterative experimental quest to reduce the variance in the numerator of the index of dispersion. The situation is challenging when unsure of the causes. They might come from splatter on pavements or evaporation (examples from the raindrop experiment in *Appendix I*1), a faint breeze, temperature variability of device parameters, small electrostatic fields in the surrounding, power supply fluctuations, gradual sublimation of the sample, or of course any number of other things, depending on circumstances. As in

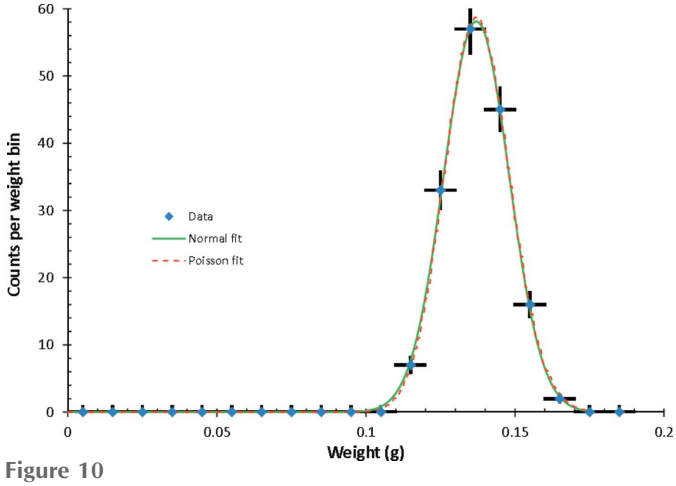


Figure 10
A histogram of manually snipped pipe weights, after compensating for the rounding-down of the scales on the horizontal axis (see text). The illustrated error bars show a total of one σ in both horizontal and vertical directions. The vertical error is assumed to be $\sigma = \sqrt{\text{counts}}$, or 1 if zero counts (no Bessel correction). A normal (Gaussian) fit is shown in green. By fitting a suitable scaling factor for the horizontal axis it can be modelled comparably well by a Poisson distribution.

the raindrop experiment, identification is the key to elimination.

Error bars and confidence in their causes are the decisive thing in that quest. Consequently we have refrained from suggesting error bars in instances where to do so may be misguiding, or where it would anticipate patterns that are known to exist, but whose quantitative models would be at first a wild guess then quickly a major and unnecessary distraction. Fig. 1 offers an example here, where estimates of anisotropic energy dependent Compton scatter of background radiation and the role of detector plane obliquity mean we have decided to turn away in the interest of progress. The mentioned errors surely are there. It might be possible with considerable difficulty (or might not) to crudely estimate their contributions, or better, somehow eliminate them by experiment refinement (as was done when unwelcome observations were seen in provisional raindrop measurements, *Appendix I*1). The linearity of the σ^2 *versus I* plots in Fig. 1 is compelling enough that scrutinizing error bars with regard to those matters would only retard the object of this work.

Where confidence exists in the mathematical robustness of Monte Carlo simulations, they can be a powerful way to avoid unknown causes of errors whose practical examination becomes an unnecessary distraction. We have developed and used them in that way, to demonstrate and guide this work.

Acknowledgements

WKF thanks the Department of Applied Mathematics (ANU) for supporting this work which began independently, and has appreciated the hospitality of Professor P. Coppens and kind interest shown by Professor J. W. White and many other colleagues throughout its development. AMK, GRM and SJL acknowledge funding through both ARC and FEI, from ARC Linkage Project LP150101040.

Funding information

The following funding is acknowledged: Australian Research Council (award No. Linkage Project LP150101040).

References

Als-Nielsen, J. & McMorrow, D. (2001). *Elements of Modern X-ray Physics.* New York: John Wiley and Sons, Ltd.
Alvarez, R. E. & Macovski, A. (1976). *Phys. Med. Biol.* **21**, 733–744.
Ament, L. J. P., van Veenendaal, M., Devereaux, T. P., Hill, J. P. & van den Brink, J. (2011). *Rev. Mod. Phys.* **83**, 705–767.
Ananthaswamy, A. (2017). *New Sci.* 4 February, pp. 28–32.
Anton, H. & Rorres, C. (1987). *Elementary Linear Algebra with Applications, 1987.* New York: John Wiley and Sons, Inc.
Atkins, P. (1998). *Physical Chemistry*, 6th ed. Oxford University Press.
Ayyer, K. *et al.* (2016). *Nature*, **530**, 202–206.
Baker, S. & Cousins, R. D. (1984). *Nucl. Instrum. Methods Phys. Res.* **221**, 437–442.
Barnes, R. B. & Silverman, S. (1934). *Rev. Mod. Phys.* **6**, 162–192.
Becker, D., Gentry, C., Smirnov, I., Ade, P., Beall, J., Cho, H.-M., Dicker, S., Duncan, W., Halpern, M., Hilton, G., Irwin, K., Li, D., Paulter, N., Reintsema, C., Schwall, R. & Tucker, C. (2014). *Proc. SPIE*, **9078**, 9078041–9078048.
Berger, M. J., Hubbell, J. H., Seltzer, S. M., Chang, J., Coursey, J. S., Sukumar, R., Zucker, D. S. & Olsen, K. (2010). *XCOM: Photon Cross Section Database*, Version 1.5. National Institute of Standards and Technology, Gaithersburg, MD.
Bohm, D. & Vigier, J. P. (1954). *Phys. Rev.* **96**, 208–216.
Bolotovsky, R., White, M. A., Darovsky, A. & Coppens, P. (1995). *J. Appl. Cryst.* **28**, 86–95.
Bragg, W. L. (1962). *The Universe of Light.* London: G. Bell and Sons.
Brillouin, L. (1946). *Wave Propagation in Periodic Structures: Electric Filters and Crystal Lattices.* New York: McGraw-Hill.
Cabello, A., Gu, M., Gühne, O., Larsson, J.-Å. & Wiesner, K. (2016). *Phys. Rev. A*, **94**, 0521271–0521275.
Coppens, P. (1997). *X-ray Charge Densities and Chemical Bonding.* Oxford University Press.
Coppens, P. (2015). *Struct. Dyn.* **2**, 020901–020908.
Coppens, P., Pitak, M., Gembicky, M., Messerschmidt, M., Scheins, S., Benedict, J., Adachi, S., Sato, T., Nozawa, S., Ichiyanagi, K., Chollet, M. & Koshihara, S. (2009). *J. Synchrotron Rad.* **16**, 226–230.
Dam, B. van (2009). *Artificial Intelligence – 23 Projects to bring your Microcontroller to Life!* Elektor International Media BV.
Darwin, C. G. (1914). *Philos. Mag.* **27**, 675–690.
Darwin, C. G. & Fowler, R. H. (1922). *Philos. Mag.* **44**, 823–842.
Davis, C. C. (1996). *Lasers and Electro-Optics: Fundamentals and Engineering.* Cambridge University Press.
De Man, B., Nuyts, J., Dupont, P., Marchal, G. & Suetens, P. (2001). *IEEE Trans. Med. Imaging*, **20**, 999–1008.
Dudewicz, E. J. & Mishra, S. N. (1988). *Modern Mathematical Statistics.* John Wiley and Sons.
Dugdale, J. S. (1966). *Entropy and Low Temperature Physics.* London: Hutchinson and Co. Ltd.
Eden, U. T. & Kramer, M. A. (2010). *J. Neurosci. Methods*, **190**, 149–152.
Elbakri, I. A. & Fessler, J. A. (2003). US Patent number 6, 507, 633.
Enss, C. (2005). *Cryogenic Particle Detection.* Berlin: Springer-Verlag.
Fano, U. (1947). *Phys. Rev.* **72**, 26–29.
Feynman, R. P. (1985). *QED – the Strange Theory of Light and Matter.* London: Penguin Books.
Fullagar, W. K. (1999). *Fullerene Sci. Technol.* **7**, 1175–1179.
Fullagar, W., Harbst, M., Canton, S., Uhlig, J., Walczak, M., Wahlström, C.-G. & Sundström, V. (2007). *Rev. Sci. Instrum.* **78**, 115105.
Fullagar, W., Paganin, D. & Hall, C. (2010). *FEL2010: Proceeding of the 32nd International Free Electron Laser Conference*, p. THPA11. Malmö, Sweden.
Fullagar, W. K., Paganin, D. M. & Hall, C. J. (2011*a*). *arXiv*: 1102.4680.
Fullagar, W. K., Paganin, D. M. & Hall, C. J. (2011*b*). *Ultramicroscopy*, **111**, 768–776.
Fullagar, W., Uhlig, J., Gador, N., Kinnuen, K., Maasilta, I., Wahlström, C. G., Sundström, V., Garrett, R., Gentle, I., Nugent, K. & Wilkins, S. (2010). *AIP Conf. Proc.* pp. 919–922.
Fullagar, W. K., Uhlig, J., Mandal, U., Kurunthu, D., El Nahhas, A., Tatsuno, H., Honarfar, A., Parnefjord Gustafsson, F., Sundström, V., Palosaari, M. R. J., Kinnunen, K. M., Maasilta, I. J., Miaja-Avila, L., O'Neil, G. C., Joe, Y. I., Swetz, D. S. & Ullom, J. N. (2017). *Struct. Dyn.* **4**, 044011.
Fullagar, W., Uhlig, J., Walczak, M., Canton, S. & Sundström, V. (2008). *Rev. Sci. Instrum.* **79**, 103302.
Fullagar, W. K., Wu, G., Kim, C., Ribaud, L., Sagerman, G. & Coppens, P. (2000). *J. Synchrotron Rad.* **7**, 229–235.
Glansdorff, P. & Prigogine, I. (1971). *Thermodynamic Theory of Structure, Stability and Fluctuations.* New York: Wiley-Interscience.
Glasstone, S. (1940). *Textbook of Physical Chemistry.* New York: MacMillan and Co., Inc.
Gonick, L. & Criddle, C. (2005). *The Cartoon Guide to Chemistry*, p. 225. New York: HarperCollins Publishers.
Goodman, J. W. (1985). *Statistical Optics.* New York: John Wiley and Sons, Inc.

Hadad, Y., Soric, J. C. & Alu, A. (2016). *Proc. Natl Acad. Sci.* **113**, 3471–3475.
Hanley, Q. S. & Denton, M. B. (2005). *J. Synchrotron Rad.* **12**, 618–625.
Hart, E. J. & Anbar, M. (1970). *The Hydrated Electron.* New York: Wiley-Interscience.
Hartmann, R., Buttler, W., Gorke, H., Herrmann, S., Holl, P., Meidinger, N., Soltau, H. & Strüder, L. (2006). *Nucl. Instrum. Methods Phys. Res. A*, **568**, 118–123.
Hazelett, R. & Turner, D. (1979). *The Einstein Myth and the Ives Papers: A Counter-Revolution in Physics.* Illinois: Hope Publishing House.
Helliwell, J. R., Habash, J., Cruickshank, D. W. J., Harding, M. M., Greenhough, T. J., Campbell, J. W., Clifton, I. J., Elder, M., Machin, P. A., Papiz, M. Z. & Zurek, S. (1989). *J. Appl. Cryst.* **22**, 483–497.
Henri-Rousseau, O. & Blaise, P. (2011). *Quantum Oscillators.* New York: John Wiley and Sons.
Higginbotham, A., Patel, S., Hawreliak, J. A., Ciricosta, O., Collins, G. W., Coppari, F., Eggert, J. H., Suggit, M. J., Tang, H. & Wark, J. S. (2014). *Rev. Sci. Instrum.* **85**, 033906.
Holst, G. C. & Lomheim, T. S. (2007). *CMOS/CCD Sensors and Camera Systems.* SPIE Press/JCD Publishing.
Horowitz, P. & Hill, W. (1989). *The Art of Electronics*, 2nd ed. Cambridge University Press.
Hove, L. van (1954). *Phys. Rev.* **95**, 249–262.
Hughes, B. D. & Ninham, B. W. (2016). *Physica A (Amsterdam)*, **443**, 495–517.
Hull, A. W. & Williams, N. H. (1925). *Phys. Rev.* **25**, 147–173.
Ising, G. (1926). *London, Edinb. Dubl. Philos. Mag. J. Sci.* **1**, 827–834.
Ives, H. E. (1951). *Proc. Am. Acad. Arts Sci.* **81**, 6–31.
Janesick, J. R. (2001). *Scientific Charge-Coupled Devices.* Washington: SPIE Press.
Jaynes, E. T. (1983). *Papers on Probability, Statistics and Statistical Physics*, edited by R. D. Rosenkrantz, pp. 149–209. Dordrecht: D. Reidel Publishing Co.
Jaynes, E. T. (1991). *Straight Line Fitting – A Bayesian Solution.* Unpublished.
Jaynes, E. T. (2003). *Probability Theory – the Logic of Science.* Cambridge University Press.
Kim, C. D., Pillet, S., Wu, G., Fullagar, W. K. & Coppens, P. (2002). *Acta Cryst.* A**58**, 133–137.
Kittel, C. & Kroemer, H. (1980). *Thermal Physics*, 2nd ed. New York: W. H. Freeman and Co.
Klafter, J., Shlesinger, M. F. & Zumofen, G. (1996). *Phys. Today*, **49**, 33–39.
Landau, L. D. & Lifshitz, E. M. (1969). *Statistical Physics.* Oxford: Pergamon Press.
LaViolette, P. A. (2010). *Subquantum Kinetics – a Systems Approach to Physics and Cosmology.* Virginia: Starlane Publications.
Lindgren, B. W. (1976). *Statistical Theory*, 3rd ed. New York: Macmillan Publishing Co.
Matzke, M. (2002). *Nucl. Instrum. Methods Phys. Res. A*, **476**, 230–241.
Matzke, M. (2003). *Radiat. Prot. Dosimetry*, **107**, 155–174.
Mazin, B. A., Meeker, S. R., Strader, M. J., Szypryt, P., Marsden, D., van Eyken, J. C., Duggan, G. E., Walter, A. B., Ulbricht, G., Johnson, M., Bumble, B., O'Brien, K. & Stoughton, C. (2013). *Publ. Astron. Soc. Pac.* **125**, 1348–1361.
McDonald, S. A., Reischig, P., Holzner, C., Lauridsen, E. M., Withers, P. J., Merkle, A. P. & Feser, M. (2015). *Sci. Rep.* **5**, 14665.
Moffat, K. (2003). *Faraday Discuss.* **122**, 65–77.
Moore, J. H., Davis, C. C. & Coplan, M. A. (2003). *Building Scientific Apparatus: A Practical Guide to Design and Construction*, 3rd ed. Colarado: Westview Press Perseus Books.
Morse, P. M. (1969). *Thermal Physics.* New York: W. A. Benjamin.
Moseley, S. H., Mather, J. C. & McCammon, D. (1984). *J. Appl. Phys.* **56**, 1257–1262.
Mössbauer, R. L. (1991). *J. Phys. G Nucl. Part. Phys.* **17**, S1–S13.
Mott, N. F. & Gurney, R. W. (1948). *Electronic Processes in Ionic Crystals.* New York: Clarendon Press.
Mukherjee, B. (2002). *Nucl. Instrum. Methods Phys. Res. A*, **476**, 247–251.
Noble, A. (1991). *New Applications of Mathematics*, edited by C. Bondi: London: Penguin Books.
Norton, K. A., Vogler, L. E., Mansfield, W. V. & Short, P. J. (1955). *Proceedings of the IRE* **43**, 1354–1361.
Nuyts, J., De Man, B., Fessler, J. A., Zbijewski, W. & Beekman, F. J. (2013). *Phys. Med. Biol.* **58**, R63–R96.
Ortiz-Rodriguez, J. M. & Vega-Carrillo, H. R. (2012). *International Symposium on Solid State Dosimetry 2012.*
Pauling, L. (1956). *General Chemistry – an Introduction to Descriptive Chemistry and Modern Chemical Theory.* New York: W. H. Freeman and Co.
Paziresh, M., Kingston, A. M., Latham, S. J., Fullagar, W. K. & Myers, G. M. (2016). *J. Appl. Phys.* **119**, 214901.
Polad, M. S. (2008). *Phys. Med. Biol.* **53**, 1475.
Poole, C. P. Jr (1998). *The Physics Handbook: Fundamentals and Key Equations*, p. 436. New York: John Wiley and Sons.
Prigogine, I. (1978). *Science*, **201**, 777–785.
Prigozhin, G., Butler, N. R., Kissel, S. E. & Ricker, G. R. (2003). *IEEE Trans. Electron Devices*, **50**, 246–253.
Prigozhin, G., Jones, S., Bautz, M., Ricker, G. & Kraft, S. (2000). *Nucl. Instrum. Methods Phys. Res. A*, **439**, 582–591.
Raman, A. P., Anoma, M. A., Zhu, L., Rephaeli, E. & Fan, S. (2014). *Nature*, **515**, 540–544.
Rayleigh, L. (1880). *Philos. Mag. Ser. 5*, **10**, 73–78.
Rayleigh, L. (1919). *Philos. Mag. Ser. 6*, **37**, 321–347.
Redus, R. H., Pantazis, J. A., Pantazis, T. J., Huber, A. C. & Cross, B. J. (2009). *IEEE Trans. Nucl. Sci.* **56**, 2524–2532.
Robbins, M. S. & Hadwen, B. J. (2003). *IEEE Trans. Electron Devices*, **50**, 1227–1232.
Robinson, F. N. H. (1962). *Noise in Electrical Circuits.* Oxford University Press.
Rose, A. (1963). *Concepts in Photoconductivity and Allied Problems.* New York: Wiley-Interscience.
Sanna, R. & O'Brien, K. (1971). *Nucl. Instrum. Methods*, **91**, 573–576.
Schmidt, M., Rajagopal, S., Ren, Z. & Moffat, K. (2003). *Biophys. J.* **84**, 2112–2129.
Schrödinger, E. (1952, 1989). *Statistical Thermodynamics.* New York: Dover Publications.
Shockley, W. & Queisser, H. J. (1961). *J. Appl. Phys.* **32**, 510–519.
Sivia, D. S. & Carlile, C. J. (1992). *J. Chem. Phys.* **96**, 170–178.
Smoluchowski, M. von (1912). *Physik Z.*, **XIII**, 1069–1080.
Šrajer, V., Ren, Z., Teng, T.-Y., Schmidt, M., Ursby, T., Bourgeois, D., Pradervand, C., Schildkamp, W., Wulff, M. & Moffat, K. (2001). *Biochemistry*, **40**, 13802–13815.
Strong, J. (1938). *Procedures in Experimental Physics.* New York: Lindsay Publications.
Susskind, L. & Friedman, A. (2014). *Quantum Mechanics – the Theoretical Minimum.* London: Penguin Books.
Ta Phuoc, K., Corde, S., Thaury, C., Malka, V., Tafzi, A., Goddet, J. P., Shah, R. C., Sebban, S. & Rousse, A. (2012). *Nat Photon* **6**, 308–311.
Thompson, S. P. (1919). *Light Visible and Invisible*, 2nd ed., pp. 100–104. New York: MacMillan and Company.
Thompson, R. E., Larson, D. R. & Webb, W. W. (2002). *Biophys. J.* **82**, 2775–2783.
Turing, A. M. (1952). *Philos. Trans. R. Soc. B Biol. Sci.* **237**, 37–72.
Tutt, J. H., Holland, A. D., Hall, D. J., Harriss, R. D. & Murray, N. J. (2012). *IEEE Trans. Electron Devices*, **59**, 167–175.
Uhlig, J. (2011). Life of a Photon in X-ray Spectroscopy, PhD thesis, Lund University.
Uhlig, J. *et al.* (2015). *J. Synchrotron Rad.* **22**, 766–775.
Uhlig, J. *et al.* (2013). *Phys. Rev. Lett.* **110**, 138302.
Uhlig, J., Wahlström, C.-G., Walczak, M., Sundström, V. & Fullagar, W. (2011). *Laser Particle Beams* **29**, 415–424.

Ullom, J. N. & Bennett, D. A. (2015). *Supercond. Sci. Technol.* **28**, 084003.
Washtell, C. C. H. (1958). *An Introduction to Radiation Counters and Detectors.* London: George Newnes Ltd.
Weise, K. & Matzke, M. (1989). *Nucl. Instrum. Methods Phys. Res. A*, **280**, 103–112.
White, K. S. (1963). *Am. J. Phys.* **31**, 922–925.
Wilkins, S., Gureyev, T., Gao, D., Pogany, A. & Stevenson, A. (1996). *Nature*, **384**, 335–338.
Wilson, C. T. R. (1912). *Proc. R. Soc. London A*, **87**, 277–292.
Yao, N. Y., Potter, A. C., Potirniche, I. D. & Vishwanath, A. (2017). *Phys. Rev. Lett.* **118**, 0304011–0304016.
Zangwill, A. (2015). *Phys. Today*, **68**, 34–39.