



Received 15 March 2021

Accepted 14 July 2021

Edited by I. A. Vartianians, Deutsches Elektronen-Synchrotron, Germany

Keywords: femtosecond studies; free-electron laser; correlated fluctuations; diffract-then-destroy; single particles; XFEL.

On incoherent diffractive imaging

Leon M. Lohse,* Malte Vassholz and Tim Salditt

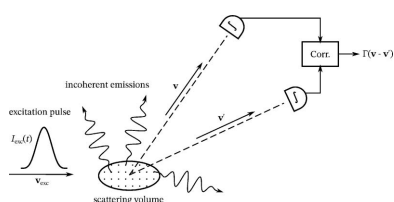
Institut für Röntgenphysik, Universität Göttingen, Germany. *Correspondence e-mail: llohse@uni-goettingen.de

Incoherent diffractive imaging (IDI) promises structural analysis with atomic resolution based on intensity interferometry of pulsed X-ray fluorescence emission. However, its experimental realization is still pending and a comprehensive theory of contrast formation has not been established to date. Explicit expressions are derived for the equal-pulse two-point intensity correlations, as the principal measured quantity of IDI, with full control of the prefactors, based on a simple model of stochastic fluorescence emission. The model considers the photon detection statistics, the finite temporal coherence of the individual emissions, as well as the geometry of the scattering volume. The implications are interpreted in view of the most relevant quantities, including the fluorescence lifetime, the excitation pulse, as well as the extent of the scattering volume and pixel size. Importantly, the spatiotemporal overlap between any two emissions in the sample can be identified as a crucial factor limiting the contrast and its dependency on the sample size can be derived. The paper gives rigorous estimates for the optimum sample size, the maximum photon yield and the expected signal-to-noise ratio under optimal conditions. Based on these estimates, the feasibility of IDI experiments for plausible experimental parameters is discussed. It is shown in particular that the mean number of photons per detector pixel which can be achieved with X-ray fluorescence is severely limited and as a consequence imposes restrictive constraints on possible applications.

1. Introduction

X-ray diffraction capitalizes on the fact that microscopic signals of scattered waves add up coherently and form a macroscopic interference pattern which can be captured by X-ray detectors in the far-field. Hence, within a volume defined by a coherence length of the radiation, the signal is enhanced by constructive interference proportional to the square of the scattering centers $\sim N^2$, most notably in the forward direction and, for diffraction from crystals, at the Bragg peak positions. Moreover, even for distances beyond the coherence length, the diffraction signal scales linearly with the interaction volume. Modulating the scattering intensity as a function of the outgoing wavevector $I(k\mathbf{v})$ with wavenumber k and unit direction vector \mathbf{v} , the measurable far-field interference pattern thus encodes the spatial Fourier transform of the ‘scattering length density’ $F(\mathbf{r})$ on atomic scales. Since the scattering process is coherent, the difference between the outgoing \mathbf{v} and incoming wavevector \mathbf{v}_0 determines the phase shifts between scattering centers, and in the kinematical scattering regime the intensity is proportional to the structure factor $S(\mathbf{q}) \propto |\mathcal{F}\{F\}(\mathbf{q})|^2$ as a function of the scattering vector $\mathbf{q} \equiv k(\mathbf{v}_0 - \mathbf{v})$.

Notwithstanding the abiding importance of coherent diffraction, the small cross section of Thomson scattering limits structural analysis at the level of single small crystallites or even single molecules. This has raised a desire to make



OPEN ACCESS

better use of the much higher cross section for photo-electric absorption, not only for spectroscopy, but also for a generalized diffraction method. Depending on photon energy $E = \hbar\omega$ and atomic number Z , the ratio of cross sections for photo-electric absorption and elastic (Thomson) scattering can easily result in hundreds of photons being absorbed per coherently scattered photon. Photo-ionization then leaves an excited atom with an inner-shell vacancy behind, from which a sizeable fraction decays via emission of an X-ray fluorescence photon. Fluorescence emission was long perceived not to convey any information about the microscopic sample structure, due to the random nature of the emission with independent phases.¹ Correspondingly, one would expect a random speckle pattern in the far-field, encoding not only the path length differences due to structure but also the unknown random phase differences. Yet, all structure in the far-field pattern ‘averages out’ because the *coherence time* of X-ray fluorescence is generally on the order of 1 fs, and all measurements are significantly longer than the coherence time.

However, as observed by Classen *et al.* (2017), the pulse length of modern XFELs (X-ray free-electron lasers) on the order of a few up to a hundred femtoseconds provides an intrinsic time gating for the fluorescence emission. Although not yet quite as short as the coherence time of X-ray fluorescence, the pulse lengths might just be short enough to leave some structural information in the fluorescence. In the work of Classen *et al.* (2017), a simple time-independent quantum-mechanical model was used to show that the two-point correlations of the fluorescence intensity are in fact proportional to the very same structure factor $\mathcal{S}(\mathbf{q})$ which emerges in coherent scattering plus a constant offset. They hence proposed a method to extract spatial information from incoherent diffraction patterns, which they termed incoherent diffractive imaging (IDI) in analogy to coherent diffractive imaging (CDI) (Miao *et al.*, 1999, 2015; Chapman *et al.*, 2006). The underlying principle, intensity interferometry, goes back to the work of Hanbury Brown & Twiss (1956), where it was used to measure the angular diameter and separation of stars. However, we are not aware of any successful realization of IDI with atomic resolution, to date.

A closely related experiment, exploiting two-point intensity correlations of X-ray fluorescence, was reported by Inoue *et al.* (2019). The authors used the correlations in the fluorescence from a thin copper foil, excited by XFEL pulses from SACLA (Spring-8 Angstrom Compact free electron Laser), to infer the duration of the exciting pulse and spatial extent of the focal spot. Unlike IDI, however, which aims for 3D *imaging* with atomic resolution, the latter experiment extracted only the spatial extent of the scattering volume with a resolution just below 1 μm , and thus served mainly for characterization of the pulse itself. Similarly, two-point intensity correlations have been studied analytically as a means to deduce the source size

as well as the detector point-spread-function (Gureyev *et al.*, 2017). While the experiment of Inoue *et al.* (2019) is hence still far from a realization of IDI with atomic resolution, it is conceptually similar, since in both cases structural information is extracted from two-point intensity correlations of hard X-ray fluorescence. Further, two-point and higher-intensity correlations of incoherent diffraction data have also been exploited to reconstruct 2D test structures imaged by FEL (free-electron laser) pulses (Schneider *et al.*, 2017). They are thus prototypical for an emerging class of experiments, where the information is not contained in the mean of the experimental data (which is homogeneous), but in its dependency structure.

The concept of IDI and the absence of its experimental demonstration have also sparked theoretical investigations. Ho *et al.* (2020) discussed the coherence time of hard X-ray fluorescence following the excitation by intense XFEL pulses. More recently, Trost *et al.* (2020) discussed the signal-to-noise ratio (SNR) of IDI, using a simple time-independent wave-optical model. Their model considers a set of ‘emitters that emit monochromatic spherical waves with random relative phases’ and includes the photon (detection) statistics of a pixel detector. Based on their model, they derive and simulate the scaling of the SNR of the two-point photon correlations as a function of several external parameters, such as the mean number of counts per pixel, the number of emitters and the number of modes. They also give an analytical expression for the number of modes as a function of the coherence time, the duration of the excitation pulse and the polarization state. Moreover, they mention that as the pixel size is increased, the number of modes increases. However, they do not explicitly treat this ‘speckle sampling’ effect, arguing that it can be considered by an appropriate adjustment of the parameters. Similarly, they discuss only qualitatively how ‘large crystals’ can lead to an increased number of modes, when the linear extent of the scattering volume is greater than the coherence length of the emissions, but do not quantify this effect. They conclude, in agreement with the seminal work of Classen *et al.* (2017), that, under optimized conditions, ‘IDI may offer utility in structure determination of single molecules at X-ray FELs’ and that ‘IDI could potentially provide element-specific structural information to complement weak coherent scattering’ (Trost *et al.*, 2020).

Here, we ask to what extent we can translate our intuition from the kinematic and dynamic theory of coherent diffraction to incoherent diffraction? In particular, can we develop a quantitative understanding of contrast formation in incoherent diffraction, including the effects that were mentioned qualitatively in the works of Classen *et al.* (2017) and Trost *et al.* (2020) in a self-contained way?

To answer these questions, we develop a *time-dependent* probabilistic model for incoherent emissions following a short excitation pulse that accounts for the geometry of scatterers and detector. We treat the fluorescence emissions with fully specified (self-) coherence functions and explicit emission and propagation times. Since two emission events may have lost temporal overlap when reaching the detector, in particular for

¹Here we leave aside effects due to secondary scattering of the emitted photons such as for Kossel lines (Kossel *et al.*, 1935; Laue, 1935) or fluorescence holography (Gog *et al.*, 1996), which do carry some information on the structure hosting the fluorescent emitters.

extended samples, the effective contrast can be severely degraded. We derive detailed estimates for the contrast in terms of a few parameters, including coherence time, sample extent, as well as the number and relative strength of emission lines. In particular, we can show that the contrast is inherently limited and derive a universal upper bound on the photon yield and correspondingly on the SNR.

The text is structured as follows. In Section 2, we describe the model, compute some statistical properties in terms of the radiated energy, and discuss contrast formation in the two-point correlations. We subsequently extend this model in Section 3 to account for quantized photodetection and derive an expression for the SNR. In contrast to the statistical treatment in the work of Trost *et al.* (2020), we treat the coincident measured counts as statistically dependent random variables and give rigorous lower bounds on the noise level and thereby rigorous upper bounds on the SNR. In the course of Section 4 we discuss different contributions to the contrast with special emphasis on the geometrical implications of the finite coherence time, propagation of the excitation pulse, and non-negligible size of the scattering volume. Section 5 is concerned with spatial sampling of the correlation signal and the implications of finite-sized detector pixels. In particular, we quantify, using rigorous upper bounds on the contrast, how the contrast decreases with increasing angular pixel size and size of the scattering volume, as was previously mentioned qualitatively in the works of Trost *et al.* (2020) and Classen *et al.* (2017). In Section 6 we estimate and show how the constraints imposed by the contrast relations inherently limit the photon yield. We use these estimates to discuss the feasibility of experiments in Section 7. Section 8 summarizes and concludes our findings.

2. A probabilistic model for incoherent emissions

We develop a simple probabilistic model to describe the random incoherent emissions from an ensemble of emitters, following excitation by a short pulse. A time-dependent description is necessary to capture the loss of contrast due to the finite coherence time in a closed form. We use this model to derive statistical properties of the energy that is radiated

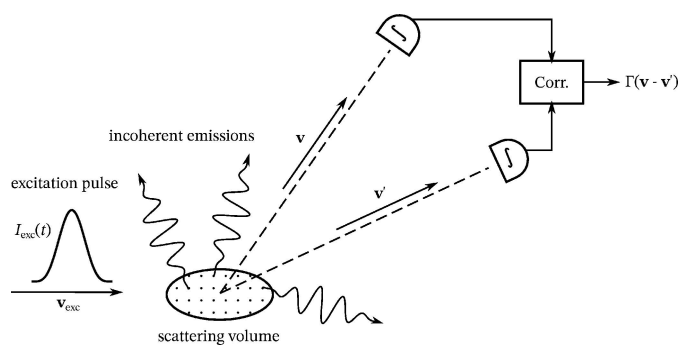


Figure 1 Sketch of the experimental principle based on coincident detection of incoherent emissions. The single-pulse two-point correlations $\Gamma(\mathbf{v} - \mathbf{v}')$ contain the structure factor of the emitter ensemble.

during each pulse and show how the structure factor of the emitter distribution enters the two-point intensity correlations. The main symbols used throughout the article are listed in Appendix C for reference.

2.1. Setting

Let us consider the following experimental setting, as sketched in Fig. 1. An ensemble of atoms is irradiated by an ultrashort excitation pulse that excites a sizeable number of them. Some fraction of the excited atoms subsequently undergoes a radiative transition and each emits a fluorescence photon. These photons are registered by (energy-sensitive) pixels of a 2D pixel array detector. The average radiated energy is isotropic if we neglect secondary interactions with the sample. However, the equal-pulse two-point intensity correlations Γ contain a structural signal of the emitter ensemble.

IDI, as originally proposed, uses pulses from an XFEL that produce inner-shell vacancies due to photo-absorption. There, the majority of the emitted photons stem from radiative transitions to the *K* shell, most frequently the $K\alpha_1$ and $K\alpha_2$ transitions. The model we describe, however, neither depends on the nature of the excitation pulse nor on the particular type of transitions, and should therefore be equally applicable to other spectral ranges, in particular the optical regime, and even other kinds of incoherent emissions. We formulate the theory as generally as possible, referring to the parameter range of $K\alpha$ radiation only for examples.

Significant correlations can only be observed when the pulse duration (implementing the time gating) is comparable to the coherence time of the emitted light. The coherence time of inner-shell fluorescence radiation is on the order of femtoseconds, which is significantly shorter than the time resolution of any conceivable detector. We thus assume that only the total energy, deposited over an entire pulse, is registered experimentally.

Although emissions of individual photons are typically discussed in a quantum-mechanical formulation, a large number of emissions can be conveniently described semi-classically. Interference in two-point intensity correlations from a pair of incoherently excited atoms is discussed for example in the book by Agarwal (2013), ch. 14. There it is shown that the emerging interference fringes have 100% visibility, which is a clear sign that the underlying process is non-classical (Mandel, 1999). However, the visibility gradually decreases with increasing number of emitters N_E and asymptotically approaches 50%. In fact, Classen *et al.* (2017) have calculated that the interference patterns produced, respectively, by single-photon emitters (SPE) and thermal light sources (TLS) converge with one another with $\mathcal{O}(1/N_E)$. It thus seems plausible that the interference from a large number of emitters can be approximated well with a semi-classical description, but special care has to be taken when discussing two-photon contributions.

2.2. Basic assumptions

First, we discuss the emissions as randomly parameterized classical electromagnetic waves (see Fig. 2). Later, in Section 3, we additionally include quantized photodetection, following the book of Goodman (1985). We initially suppose that the fields are perfectly polarized to simplify the notation. The unpolarized nature of fluorescence is taken into account in Section 4.3. Let N_E be the number of individual emitters and let \mathbf{r}_m for $m \in \{1, \dots, N_E\}$ denote their positions. We assume the system to be stationary in the sense that each pulse has the same initial conditions and can be represented as a realization of the same statistical ensemble.

We make the following assumptions:

(i) The event in which a specific atom with index m emits a photon follows a Bernoulli distribution with probability s_m . We associate the random variables $b_m \in \{0, 1\}$ with the emission of a photon.

(ii) The photon energy $\hbar\omega_m$ takes values from a finite set of discrete transition energies (emission lines) ℓ_m with probabilities corresponding to the relative line intensities.

(iii) The emission times t_m are randomly distributed with probability density proportional to the cycle-averaged local intensity $I_{\text{exc}}(\mathbf{r}_m, t)$ of the excitation pulse (the lifetime of the excited state is described by the time evolution of the field amplitudes) at position \mathbf{r}_m .

(iv) The emissions are independent, such that the random variables characterizing the emissions, b_m , t_m and $\hbar\omega_m$, are mutually independent.

(v) The emissions can be described by outgoing spherical waves.

Let $u_m(\mathbf{v}r, t)$ be the complex-valued analytic signal associated with the electromagnetic disturbance due to the emission from the m th atom. The normalized vector \mathbf{v} denotes the

observation direction relative to the center of the scattering volume. Let

$$w_m \, d\Omega = \int u_m(\mathbf{v}r, t) u_m(\mathbf{v}r, t)^* \, dt \, d\Omega \quad (1)$$

denote the energy flow of u_m into the infinitesimal solid angle $d\Omega$ at some point $\mathbf{v}r$ in the far-field. The total electromagnetic disturbance at position $\mathbf{v}r$ due to all the emissions can be written as

$$U(\mathbf{v}r, t) = \sum_m b_m u_m(\mathbf{v}r, t). \quad (2)$$

The total energy flow through $d\Omega$ becomes

$$W \, d\Omega = \int U(\mathbf{v}r, t) U(\mathbf{v}r, t)^* \, dt \, d\Omega. \quad (3)$$

2.3. Correlations

These basic assumptions suffice to significantly simplify the average intensity and two-point intensity correlations. The expectation value of W can be expressed as

$$E W \, d\Omega = \int E[U(\mathbf{v}r, t) U(\mathbf{v}r, t)^*] \, dt \, d\Omega. \quad (4)$$

Inserting (2) and exploiting the independence of the individual emissions show

$$\begin{aligned} E[U(t)U(t)^*] &= \sum_m E[b_m b_m u_m(t) u_m(t)^*] \\ &+ \sum_m \sum_{n \neq m} E[b_m b_n] E[u_m(t)] E[u_n(t)^*] \\ &= \sum_m E b_m E[|u_m(t)|^2]. \end{aligned} \quad (5)$$

Here, we have used $E[u_m(t)] = 0$ and $b_m^2 = b_m$. Performing the time integral and inserting $E b_m = s_m$ yields

$$E W \, d\Omega = \sum_m s_m E w_m \, d\Omega. \quad (6)$$

Next, we calculate the two-point correlations of the intensity. Consider the energy flow into the direction \mathbf{v}' through the solid angle $d\Omega'$ as a second observable. We are interested in the correlation of $W \, d\Omega$ and $W' \, d\Omega'$, expressed as

$$\Gamma \, d\Omega \, d\Omega' = E[WW'] \, d\Omega \, d\Omega'. \quad (7)$$

Importantly, the random variables that appear are identical for both observation directions. We obtain

$$\begin{aligned} E[WW'] \, d\Omega \, d\Omega' \\ = \int \int E[U(\mathbf{r}, t) U(\mathbf{r}, t)^* U(\mathbf{r}', t') U(\mathbf{r}', t')^*] \, dt \, dt' \, d\Omega \, d\Omega' \end{aligned} \quad (8)$$

and

$$\begin{aligned} E[U(\mathbf{r}, t) U(\mathbf{r}, t)^* U(\mathbf{r}', t') U(\mathbf{r}', t')^*] \\ = \sum_m \sum_n \sum_{m'} \sum_{n'} E[b_m b_n b_{m'} b_{n'}] \\ \times E[u_m(\mathbf{r}, t) u_n(\mathbf{r}, t)^* u_{m'}(\mathbf{r}', t') u_{n'}(\mathbf{r}', t')^*]. \end{aligned} \quad (9)$$

Here, we use the notation $\mathbf{r} \equiv \mathbf{v}r$ and $\mathbf{r}' \equiv \mathbf{v}'r'$. Since the emission times are mutually independent and $E u_m = 0$ as well as $E u_m^2 = 0$, only certain combinations of the indices survive the expectation value: $m = n = m' = n'$, $m = n \neq m' = n'$ and $m = n' \neq m' = n$. Consequently, (9) becomes

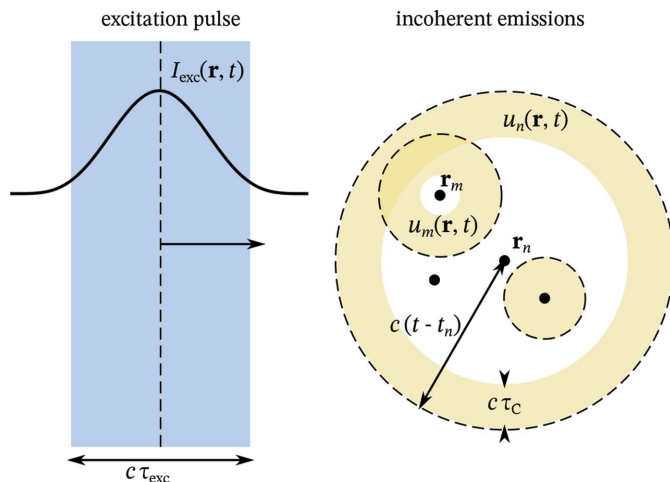


Figure 2 Sketch of the probabilistic model. The emitters at fixed positions \mathbf{r}_m randomly emit spherical waves with finite coherence time τ_c . The moment of emission is random and proportional to the instantaneous intensity of the excitation pulse, which is assumed to be a traveling wave with pulse duration τ_{exc} .

$$\begin{aligned}
 & E[U(\mathbf{r}, t)U(\mathbf{r}, t)^*U(\mathbf{r}', t')U(\mathbf{r}', t')^*] \\
 &= \sum_m E b_m E \left[|u_m(\mathbf{r}, t)|^2 |u_m(\mathbf{r}', t')|^2 \right] \\
 &+ \sum_m \sum_{n \neq m} E b_m E b_n E [u_m(\mathbf{r}, t)u_m(\mathbf{r}, t)^*] E [u_n(\mathbf{r}', t')u_n(\mathbf{r}', t')^*] \\
 &+ \sum_m \sum_{n \neq m} E b_m E b_n E [u_m(\mathbf{r}, t)u_m(\mathbf{r}', t')^*u_n(\mathbf{r}', t')u_n(\mathbf{r}, t)^*].
 \end{aligned} \tag{10}$$

Performing the time integration yields

$$\begin{aligned}
 E[WW'] d\Omega d\Omega' &= \sum_m s_m E [w_m w_m'] d\Omega d\Omega' \\
 &+ \sum_m \sum_{n \neq m} s_m s_n E w_m E w_n' d\Omega d\Omega' \\
 &+ \sum_m \sum_{n \neq m} s_m s_n E \int \int u_m(\mathbf{r}, t)u_m(\mathbf{r}', t')^* \\
 &\times u_n(\mathbf{r}', t')u_n(\mathbf{r}, t)^* dt dt' d\Omega d\Omega'.
 \end{aligned} \tag{11}$$

To evaluate the remaining integrals, we express the u_m as

$$u_m(\mathbf{r}, t) \equiv u_m(t - t_m - \phi_m) \tag{12}$$

with geometric phases $\phi_m = |\mathbf{r} - \mathbf{r}_m|/c$. The integral in (11) can then be expressed as

$$\begin{aligned}
 & E \int \int u_m(\mathbf{r}, t)u_m(\mathbf{r}', t')^* u_n(\mathbf{r}', t')u_n(\mathbf{r}, t)^* dt dt' d\Omega d\Omega' \\
 &= E [J_{mn}(-\tau_{mn} - \phi_m + \phi_n)J_{mn}(-\tau_{mn} - \phi_m' + \phi_n')^*] d\Omega d\Omega'
 \end{aligned} \tag{13}$$

with $\tau_{mn} = t_m - t_n$ and interference terms

$$J_{mn}(\tau) = \int u_m(t + \tau)u_n^*(t) dt. \tag{14}$$

In the far-field we have $|\mathbf{r}_m| \ll |\mathbf{r}|$ such that the usual far-field approximation gives $\phi_m - \phi_n \simeq -\mathbf{v} \cdot \mathbf{T}_{mn}$ with

$$\mathbf{T}_{mn} = (\mathbf{r}_m - \mathbf{r}_n)/c. \tag{15}$$

In the following we suppress the solid-angle differentials.

To further evaluate these interference terms (14), we require additional assumptions on the temporal structure of the individual emissions. We assume that the emissions produce electromagnetic signals that are characteristic for the involved atomic energy levels (emission line ℓ_m) up to the spatial origin \mathbf{r}_m and emission time t_m . More precisely, we parameterize the waves in terms of their respective temporal and spatial origin and their emission line ℓ_m such that

$$u_m(t - t_m - \phi_m) \equiv u_0(\ell_m; t - t_m - \phi_m), \tag{16}$$

where $u_0(\ell; t)$, for each emission line ℓ , describes an outgoing spherical wave. Inserting this into (14), we obtain

$$J_{mn}(\tau) = \delta_{\ell_m \ell_n} \int u_0(\ell_m; t + \tau)u_0^*(\ell_m; t) dt. \tag{17}$$

We have assumed that signals with $\ell_m \neq \ell_n$ do not interfere because the emission lines do not overlap.

2.4. Spectrum and self-coherence

It turns out that (17) is fully determined by the complex degree of coherence (CDC) of U , which in turn is fully determined by the (phase-less) spectrum of U .

For simplicity, we ignore the spatial dimensions and also consider the simplest situation of only a single contributing emission line ℓ . Then, we may express the self-coherence at time delay τ of the total signal U as the ensemble average (Mandel & Wolf, 2015):

$$\begin{aligned}
 E[U(t)U^*(t + \tau)] &= \sum_m E [u_0(t + \tau - t_m)u_0^*(t - t_m)] \\
 &+ \sum_{m \neq n} E [u_0(t + \tau - t_m)u_0^*(t - t_n)].
 \end{aligned} \tag{18}$$

Here the expectation value is taken over the time offsets t_m which are distributed over some time interval longer than the coherence time. Since $u_0(t)$ oscillates rapidly with some central frequency ω_0 and the time offsets are mutually independent, each non-diagonal term must vanish. The diagonal terms become

$$E [u_0(t + \tau - t_m)u_0^*(t - t_m)] \propto \int u_0(t' + \tau)u_0^*(t') dt', \tag{19}$$

for $|\tau| \ll \tau_{\text{exc}}$. Here the duration of the excitation pulse τ_{exc} quantifies the width of the distribution of t_m . Then, also the self-coherence (18) is proportional to

$$E[U(t)U^*(t + \tau)] \propto \int u_0(t' + \tau)u_0^*(t') dt'. \tag{20}$$

Normalizing the left-hand side gives by definition the CDC γ_ℓ associated with the emission line ℓ , such that

$$\gamma_\ell(\tau) = \frac{1}{w_\ell} \int u_0(\ell; t' + \tau)u_0^*(\ell; t') dt'. \tag{21}$$

The temporal autocorrelation of u_0 is hence directly related to the macroscopically accessible CDC of an isolated emission line. Importantly, equation (21) implies that the temporal autocorrelation of u_0 is fully determined by γ_ℓ , which is determined by the coherence time $\tau_{C,\ell}$ and the line shape.

2.5. Interference terms

Substituting (21) in (17) yields

$$J_{mn}(\tau) = \delta_{\ell_m \ell_n} w_m \gamma_{\ell_m}(\tau). \tag{22}$$

The CDC factorizes into a rapidly oscillating part and an envelope, such that

$$\gamma_\ell(\tau) = \exp(i\omega_\ell \tau) \tilde{\gamma}_\ell(\tau). \tag{23}$$

The envelope is real for symmetric line shapes but may take negative values in general. However, for Gaussian and Lorentzian line shapes, it is restricted to non-negative real values. It can hence be expressed as $\tilde{\gamma}_\ell = |\tilde{\gamma}_\ell|$. In particular for Lorentzian line shapes, we have $\tilde{\gamma}_\ell(\tau) = \exp(-|\tau|/\tau_{C,\ell})$ where $\tau_{C,\ell}$ is the coherence time. Inserting the factorized $\gamma_\ell(\tau)$ yields

$$\begin{aligned}
 & J_{mn}(\mathbf{v} \cdot \mathbf{T}_{mn} - \tau_{mn}) \times J_{mn}(\mathbf{v}' \cdot \mathbf{T}_{mn} - \tau_{mn})^* \\
 &= w_m w_m' \exp[ik_m(\mathbf{v} - \mathbf{v}') \cdot (\mathbf{r}_m - \mathbf{r}_n)] \\
 &\times \left| \gamma_{\ell_m}(\mathbf{v} \cdot \mathbf{T}_{mn} - \tau_{mn}) \right| \times \left| \gamma_{\ell_m}(\mathbf{v}' \cdot \mathbf{T}_{mn} - \tau_{mn}) \right| \delta_{\ell_m \ell_n}
 \end{aligned} \tag{24}$$

where $k_m = \omega_m/c$.

Suppose that the emissions are spectrally filtered so that only photons from a relatively narrow energy band are registered in the detectors. More precisely, suppose that

$(k_m - \bar{k})/\bar{k}$ is very small for all m , where $\bar{k} = E[\omega_m]/c$ is the mean wavenumber of the involved frequencies. In the same sense suppose that the pulse energies are approximately equal, i.e. $w_m, w'_m \simeq \bar{w}$. The expectation value of (24) can then be written as

$$E[J_{mn}(\mathbf{v} \cdot \mathbf{T}_{mn} - \tau_{mn}) \times J_{mn}(\mathbf{v}' \cdot \mathbf{T}_{mn} - \tau_{mn})^*] \simeq \bar{w}^2 C_{mn} \exp[i\bar{k}(\mathbf{v} - \mathbf{v}') \cdot (\mathbf{r}_m - \mathbf{r}_n)] \quad (25)$$

with

$$C_{mn} = E\left[\left| \gamma_{\ell_m}(\mathbf{v} \cdot \mathbf{T}_{mn} - \tau_{mn}) \right| \times \left| \gamma_{\ell_m}(\mathbf{v}' \cdot \mathbf{T}_{mn} - \tau_{mn}) \right| \times \delta_{\ell_m \ell_n} \right]. \quad (26)$$

The coupling coefficients C_{mn} take values in the interval $[0, 1]$. Note that their diagonal elements are $C_{mm} = 1$ and that $C_{mn} = C_{nm}$ as can be readily seen. From here on set $\bar{w} \equiv 1$ and $\mathbf{q} \equiv \bar{k}(\mathbf{v} - \mathbf{v}')$ to simplify the notation.

After inserting (25) into (13), (11) finally yields

$$\begin{aligned} \Gamma &= \sum_m s_m + \sum_m \sum_{n \neq m} s_m s_n + \sum_m \sum_{n \neq m} s_m s_n C_{mn} \\ &\quad \times \exp[i\mathbf{q} \cdot (\mathbf{r}_m - \mathbf{r}_n)] \\ &= \sum_m \sum_n s_m s_n + \sum_m \sum_n s_m s_n C_{mn} \\ &\quad \times \exp[i\mathbf{q} \cdot (\mathbf{r}_m - \mathbf{r}_n)] + \sum_m (s_m - 2s_m^2). \end{aligned} \quad (27)$$

Comparing this expression with the result of Classen *et al.* (2017), which was derived with a time-independent quantum-mechanical description of SPE, shows two differences. First, their model is time independent and as such implicitly assumes $C_{mn} \equiv 1$. Second, our classical formulation fails to reproduce that the diagonal terms, corresponding to coincident detection of the same atomic emission, must vanish. However, since the diagonal terms scale with $\mathcal{O}(1/N_E)$ relative to the other terms, they can be neglected for large N_E and the two expressions converge.

Using $EW = EW' = \sum_m s_m$ and neglecting the single sum simplifies the correlation to

$$\Gamma \simeq EWEW' \left\{ 1 + \frac{\sum_m \sum_n s_m s_n C_{mn} \exp[i\mathbf{q} \cdot (\mathbf{r}_m - \mathbf{r}_n)]}{(\sum_m s_m)^2} \right\}. \quad (28)$$

The right-hand side of (28) consists of two parts: a constant background $EWEW'$ and a structural signal $\Sigma = \Gamma - EWEW'$. It follows that

$$\Sigma = \text{Cov}[W, W'], \quad (29)$$

using $\text{Cov}[W, W'] = E[WW'] - EWEW'$.

2.6. Effective contrast and structure factor

The two-emitter contributions in (28) are individually attenuated by the coupling coefficients C_{mn} . Although these coefficients could be computed individually for a specific model, it is instructive to consider their global average,

$$C_{\text{eff}} = \langle C_{mn} \rangle_{mn}. \quad (30)$$

Using the approximation $C_{mn} \simeq C_{\text{eff}}$ decouples the contributions of the individual emitters in (28) and the right-hand side simplifies to

$$\Gamma \simeq EWEW'[1 + C_{\text{eff}}\mathcal{S}(\mathbf{q})], \quad (31)$$

with

$$\mathcal{S}(\mathbf{q}) = \frac{|\sum_m s_m \exp(i\mathbf{q} \cdot \mathbf{r}_m)|^2}{|\sum_m s_m|^2}. \quad (32)$$

Equation (32) resembles the definition of the *structure factor* defined in the context of elastic scattering, although the coefficients have a different meaning.

The sums can be expressed as integrals. Defining the (unit-less) emitter distribution

$$s(\mathbf{X}) = \sum_m s_m \delta(\mathbf{X} - \bar{k}\mathbf{r}_m), \quad (33)$$

we can write the structure factor (32) as

$$\mathcal{S}(\mathbf{q}) = \frac{|\mathcal{F}\{s\}(\mathbf{q}/\bar{k})|^2}{|\mathcal{F}\{s\}(0)|^2}, \quad (34)$$

where $\mathcal{F}\{s\}$ denotes the 3D Fourier transform of $s(\mathbf{X})$. We scaled the emitter distribution with the wavenumber \bar{k} , because it is a property of the emission line and hence determined by the emitters. We will mostly use the unit-less scattering vector, defined as $\mathbf{x} \equiv \mathbf{v} - \mathbf{v}'$, in place of \mathbf{q} . The two are related by $\mathbf{q} = \bar{k}\mathbf{x}$ so that $|\mathbf{x}|$ probes a length scale of $\lambda/|\mathbf{x}|$. We will use the two notations $\mathcal{S}(\mathbf{q}) = \mathcal{S}(\mathbf{x})$ synonymously.

Equation (31) is equally applicable for $W' = W$, which provides the variance $\text{Var}W = E[WW] - (EW)^2$. Using $E[WW] = \Gamma(0)$ and $\mathcal{S}(0) = 1$, we read off that

$$\text{Var}W \simeq C_{\text{eff}}(EW)^2. \quad (35)$$

For clarity and later reference, we rewrite (31) more explicitly to include the two observation directions and their solid-angle differentials as

$$\Gamma(\mathbf{x}) d\Omega d\Omega' = EWEW'[1 + C_{\text{eff}}\mathcal{S}(\mathbf{x})] d\Omega d\Omega'. \quad (36)$$

2.7. Comparison with elastic scattering

Compare (36) with the elastically scattered intensity, which can be written as

$$I_{\text{el}}(\mathbf{v}) d\Omega \propto \mathcal{S}_{\text{el}}[k(\mathbf{v} - \mathbf{v}_{\text{in}})] d\Omega, \quad (37)$$

where here \mathbf{v}_{in} is the propagation direction of the *incoming* beam and k the wavenumber of the incoming and scattered beam. Recall that the elastic structure factor \mathcal{S}_{el} is defined as in (32); however, instead of the emission probabilities s_m , the coefficients are interpreted as atomic form factors.

There are two profound formal differences between (37) and (36) besides the dependence on one and two directions, respectively. First, (36) contains constant (unit) offset. Second, the signal is attenuated by the constant C_{eff} quantifying the

contrast of the structure factor to this constant background. The correlation function thus contains a constant *intrinsic background* and the *intrinsic contrast* of the structural signal to this background is always less than unity. The coherently scattered intensity on the other hand has no intrinsic background and therefore no intrinsic contrast limitations. Only additional processes that contribute to the intensity form a background and limit the contrast. This extrinsic contrast, however, is not limited to unity.

3. Photon statistics and noise

The goal of this section is to estimate the noise level of measurements of Σ . Apart from the fluctuations of the classical energies W , which are caused by the randomness of the emissions (source noise) [Trost *et al.* (2020) call this contribution phase noise, because it originates from the random phases in their model], additional fluctuations emerge due to the quantization of the electromagnetic field or the detection process (shot noise). In fact, in the low-photon regime, this second source of noise is the bigger contribution to the noise level. Importantly, it is independent of the signal level of the correlation.

3.1. Statistics at a single observation point

We follow the semi-classical formalism described in the textbook of Goodman (1985), starting by repeating some fundamentals for reference purposes. Let K be the number of counts registered by the detector at observation direction \mathbf{v} and covering the solid-angle differential $d\Omega$. The conditional distribution $K | W$, conditioned with the energy W entering the detector, is Poissonian with parameter ηW , where η is the quantum efficiency of the detector. (Here we include the photon energy $\hbar\omega$ in the definition of the quantum efficiency.) The expectation value and variance of the (unconditional) K can be expressed in terms of moments of W as

$$EK = \eta EW, \quad \text{Var}K = \eta^2 \text{Var}W + \eta EW. \quad (38)$$

Inserting (35), which gives $\text{Var}W$ in terms of EW for the model discussed in Section 2, we obtain

$$\text{Var}K = C_{\text{eff}}(EK)^2 + EK. \quad (39)$$

Equation (39) gives the fluctuations of the number of counts measured with a single detector as a function of the mean counts. The fluctuations have two contributions: the source noise, $C_{\text{eff}}(EK)^2$, and the shot noise, EK . As we are going to see in Section 6, the mean count numbers that can be expected are usually well below $\bar{K} \lesssim 1$. In this regime, the shot noise strongly dominates the source noise.

3.2. Count-correlations

The situation is more complicated for coincident measurements at two or more points, because the emissions into the two respective directions are correlated. We can, however, assume that the two detectors do not influence each other, such that the conditional random variables $K | W$ and $K' | W'$

are in fact independent. Under this assumption we have that $E[KK'] = \eta^2 E[WW']$ and similarly $\text{Cov}[K, K'] = \eta^2 \text{Cov}[W, W']$ (Goodman, 1985). As a result, the correlations and covariances of the two classical energies and the two photon counts are related simply by a constant factor. In particular, we can define the signal as

$$\Sigma = E[KK'] - EKEK' = \text{Cov}[K, K'], \quad (40)$$

and use the expressions for $\text{Cov}[W, W']$. Note that (40) assumes $K \neq K'$, because otherwise the two conditional distributions are no longer independent. For $K = K'$ we get $\text{Cov}[K, K] = \text{Var}K$ which includes a shot-noise term that is absent in (40) (Singer & Vartanyants, 2013).

3.3. Fluctuations in the count-correlations

The covariance on the right-hand side of (40) can be expressed as $\text{Cov}[K, K'] = E[\Delta K \Delta K']$, where $\Delta K = K - EK$ and $\Delta K'$ alike. We would like to calculate the second central moment, $\text{Var}[\Delta K \Delta K']$. Unfortunately, there is no simple identity relating this expression to moments of W and W' . A simple but rather lengthy calculation (see Appendix B) shows that

$$\begin{aligned} \text{Var}[\Delta K \Delta K'] &= \eta^2 EWEW' + \eta^2 \text{Cov}[W, W'] \\ &+ \eta^3 E[(\Delta W)^2 W'] + \eta^3 E[(\Delta W')^2 W] \\ &+ \eta^4 \text{Var}[\Delta W \Delta W']. \end{aligned} \quad (41)$$

In particular, since all terms on the right-hand side are non-negative, we conclude that

$$\text{Var}[\Delta K \Delta K'] \geq EKEK' + \underbrace{\eta^2 \text{Cov}[W, W']}_{\geq 0}. \quad (42)$$

This proves that the variance is greater than or equal to the product of the shot noise of two individual (independent) intensity measurements, irrespective of the joint statistics of W and W' . It is obvious that the other terms in (41) will contribute significantly for higher average counts EK , so that the variance scales differently. However, this high-photon regime is of little relevance here due to the low photon numbers that can be expected (see Section 6). For further discussion of this high-photon-count regime, see the work of Trost *et al.* (2020).

3.4. Measurements and SNR

We now turn our attention to the SNR for estimates of the covariance Σ . We first specify, for reference, how to compute Σ from measured coincident realizations of K and K' . Suppose that (κ_j, κ'_j) for $j \in \{1, \dots, R\}$ are R independent (coincident) realizations of (K, K') . Define the sample mean as

$$\bar{\kappa} = \frac{1}{R} \sum_{j=1}^R \kappa_j, \quad (43)$$

and $\bar{\kappa}'$ alike, and set

$$\hat{\Sigma} = \frac{1}{R} \sum_{j=1}^R (\kappa_j - \bar{\kappa})(\kappa'_j - \bar{\kappa}'). \quad (44)$$

Then $\bar{\kappa}$ converges to $E\bar{K}$ and $\widehat{\Sigma}$ converges to $\text{Cov}[K, K']$ for increasing R . The SNR of $\widehat{\Sigma}$ can be defined as

$$\text{SNR} = \frac{E\widehat{\Sigma}}{(\text{Var}\widehat{\Sigma})^{1/2}}. \quad (45)$$

Using

$$\text{Var}\widehat{\Sigma} \simeq \frac{1}{R} \text{Var}[\Delta K \Delta K'] \quad (46)$$

and inserting $E\widehat{\Sigma} \simeq \Sigma$ we find that

$$\text{SNR} \simeq \frac{R^{1/2} \times \Sigma}{(\text{Var}[\Delta K \Delta K'])^{1/2}}. \quad (47)$$

Inserting (42) and $\Sigma \simeq EKEK' \times C_{\text{eff}}\mathcal{S}(\mathbf{x})$, and assuming isotropic radiation with $\bar{K} = EK = EK'$, we conclude that the SNR is bounded by

$$\text{SNR} \leq R^{1/2} \times \Sigma / \bar{K} = R^{1/2} \times \bar{K} \times C_{\text{eff}}\mathcal{S}(\mathbf{x}). \quad (48)$$

Recall that (42) bounds the variance by the shot noise of the intrinsic background, which results in a strict upper bound on the SNR. For higher photon counts, however, the source noise may exceed the shot noise so that (48) overestimates the SNR. In fact, for increasing \bar{K} , the SNR saturates at $R^{1/2}\mathcal{S}(\mathbf{x})$ (Trost *et al.*, 2020).

When only a single pair of detectors is used then the number of independent realizations R is given by the number of pulses N_{pulses} . In practice, however, it makes sense to use a pixel array detector and acquire a large number N_{pixel} of measurements simultaneously. Since not all possible pairs of pixels correspond to distinct scattering vectors, every frame provides multiple realizations for the same \mathbf{x} simultaneously. The multiplicity $\nu(\mathbf{x})$ corresponding to the scattering vector \mathbf{x} decays from N_{pixel} for $\mathbf{x} = 0$ to 1 for large \mathbf{x} , where its exact shape depends on the chosen discretization and the detection geometry. Note that these simultaneous realizations are in general not independent. Even if the individual intensity measurements were independent (if the coincident intensity measurements were strictly independent, their correlation, and thereby the signal, would vanish), then the individual terms in (44) would still be correlated, because some κ_j appear more than once (Trost *et al.*, 2020). Nevertheless, for the purpose of an optimistic upper bound on the SNR, we can use $R \leq N_{\text{pulses}} \nu(\mathbf{x})$ and implicitly count the simultaneous realizations as if they were independent.

Solving (48) for N_{pulses} gives the minimum number of pulses required for a certain target SNR and a given signal level:

$$N_{\text{pulses}} \gtrsim \frac{\text{SNR}^2}{\nu(\mathbf{x})\bar{K}^2 C_{\text{eff}}^2 \mathcal{S}(\mathbf{x})^2}. \quad (49)$$

The main results of this section are (48) and (49), giving an optimistic bound on the relation between the SNR and the number of pulses. Although we derived the SNR only for a specific method of estimating the two-point covariances, which we do not claim to be optimal, the result illustrates a general characteristic: for low photon counts, the noise in the estimated covariances is dominated by the shot noise of the

individual measurements, which is independent of the signal level. As a result, the signal level does not cancel out and the SNR is linearly proportional to the signal level. These exponents, together with the scaling of contrast and photon counts, sensitively govern the achievable SNR for different experimental settings, as we will discuss in Section 6.

4. Contrast estimates

Since the contrast C_{eff} is a crucial parameter in (36) and (48), it seems appropriate to discuss it in more detail. We make some simplifying assumptions to factorize C_{eff} and discuss the main constituents individually. Each can be estimated from a few parameters, such as the duration of the excitation pulse, the spatial extent of the scattering volume, and the spectrum of the emitted radiation.

Suppose for simplicity that all involved emission lines have the same coherence time τ_c and therefore that $|\gamma_\ell|$ is identical for all emission lines. Then, the coefficients (26) factorize into a spatiotemporal (ST) and spectral (L) factor,

$$C_{mn} = E[\delta_{\ell_m \ell_n}] \times E[|\gamma(\mathbf{v} \cdot \mathbf{T}_{mn} - \tau_{mn})| \times |\gamma(\mathbf{v}' \cdot \mathbf{T}_{mn} - \tau_{mn})|], \\ = C_{mn}^L \times C_{mn}^{\text{ST}}. \quad (50)$$

Here we dropped the index ℓ since all $|\gamma_\ell|$ are alike.

Equation (50) gives an approximate expression for the coupling coefficient between two specific emitters. The macroscopic contrast C_{eff} , however, is determined by the average coupling constant between all pairs of emitters. Assuming that the spectral factor is constant for all pairs of emitters allows one to take the averages separately, so that

$$C_{\text{eff}} = \langle C_{mn} \rangle_{m,n} = C_{\text{eff}}^L \times C_{\text{eff}}^{\text{ST}}, \quad (51)$$

with $C_{\text{eff}}^L = \langle C_{mn}^L \rangle_{m,n}$ and $C_{\text{eff}}^{\text{ST}} = \langle C_{mn}^{\text{ST}} \rangle_{m,n}$.

4.1. Spectral overlap

We discuss the spectral contribution first. Suppose, for simplicity, that all emitters have identical emission spectra, consisting of discrete emission lines with relative intensities p_ℓ . In that case, the first contribution to the coupling coefficients becomes

$$E[\delta_{\ell_m \ell_n}] = \begin{cases} \sum_\ell p_\ell^2, & \text{for } m \neq n \\ 1, & \text{for } m = n. \end{cases} \quad (52)$$

The factor therefore does not depend on m and n , but differs for diagonal (identical emitter) and non-diagonal (distinct emitters) contributions. Since there are only N_E diagonal terms, but $N_E^2 - N_E$ non-diagonal terms, only the latter contribute significantly to C_{eff} . In particular, the $K\alpha$ line splits into the $K\alpha_1$ and $K\alpha_2$ lines with relative intensities of 2/3 and 1/3, so that $C_{\text{eff}}^L(K\alpha) \simeq 5/9$.

4.2. Spatiotemporal overlap

A pair of emissions contributes to the structural signal only when the two emissions arrive within the coherence time at each pixel. This temporal overlap is purely governed by the duration of the excitation pulse when the scattering volume is

smaller than the coherence length. For larger scattering volumes, however, the finite-speed propagation of the excitation pulse and of the emissions also enter the picture, and we have to consider the combined temporal and spatial overlap.

The overlap C_{mn}^{ST} could be approximated analytically for certain symmetric configurations; however, covering all possibilities would be lengthy and cumbersome. Instead, our strategy is to write down the full expression and to derive some complementary upper bounds for it. In this way, we can include all configurations with little effort and give rigorous estimates on the maximum achievable contrast.

The factor $C_{\text{eff}}^{\text{ST}}$ is governed by the emitter locations \mathbf{r}_m and the probability distribution of the emission times t_m . We write the probability density function of the temporal separation $\tau_{mn} = t_m - t_n$ as $\Psi(\mathbf{r}_m, \mathbf{r}_n, \tau)$. The probability density of t_m is proportional to the intensity of the excitation pulse $I_{\text{exc}}(\mathbf{r}_m, t)$, by assumption, so that

$$\Psi(\mathbf{r}_m, \mathbf{r}_n, \tau) = \frac{\int I_{\text{exc}}(\mathbf{r}_m, t + \tau) I_{\text{exc}}(\mathbf{r}_n, t) dt}{\int I_{\text{exc}}(\mathbf{r}_m, t) dt \times \int I_{\text{exc}}(\mathbf{r}_n, t) dt}. \quad (53)$$

The spatial diagonal $\Psi(\mathbf{r}, \mathbf{r}, \tau)$ is the temporal autocorrelation of the excitation pulse, which, neglecting dispersion, does not depend on \mathbf{r} . For brevity we write $\Psi_0(\tau) = \Psi(\mathbf{r}, \mathbf{r}, \tau)$.

We model the excitation pulse as a dispersion-less plane wave with propagation direction \mathbf{v}_{exc} , group velocity c and duration τ_{exc} . Since a plane wave can be written as a function of a single argument, $I_{\text{exc}}(\mathbf{r}_m, t) \equiv I_{\text{exc}}(\mathbf{r}_m \cdot \mathbf{v}_{\text{exc}} - ct)$, a simple shift in the integration variable of (53) shows that the cross-correlation can be expressed as

$$\Psi(\mathbf{r}_m, \mathbf{r}_n, \tau) = \Psi_0(\tau - \mathbf{v}_{\text{exc}} \cdot \mathbf{T}_{mn}). \quad (54)$$

Note that $\Psi_0(0) \sim 1/\tau_{\text{exc}}$ and $\int \Psi_0(\tau) d\tau = 1$.

The expectation value can be expressed as

$$\begin{aligned} C_{mn}^{\text{ST}} &= \mathbb{E}[|\gamma(\mathbf{v} \cdot \mathbf{T}_{mn} - \tau_{mn})| \times |\gamma(\mathbf{v}' \cdot \mathbf{T}_{mn} - \tau_{mn})|] \\ &= \int |\gamma(\tau - \mathbf{v} \cdot \mathbf{T}_{mn})| |\gamma(\tau - \mathbf{v}' \cdot \mathbf{T}_{mn})| \Psi(\mathbf{r}_m, \mathbf{r}_n, \tau) d\tau \\ &= \int |\gamma(\tau - \mathbf{x} \cdot \mathbf{T}_{mn}/2)| |\gamma(\tau + \mathbf{x} \cdot \mathbf{T}_{mn}/2)| \\ &\quad \times \Psi_0[\tau + (\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}}) \cdot \mathbf{T}_{mn}] d\tau, \end{aligned} \quad (55)$$

where $\mathbf{v}_{\text{cen}} = (\mathbf{v} + \mathbf{v}')/2$. For the last identity, we have substituted (54) and shifted the integration variable. First, we focus on the purely temporal components of the overlap, corresponding to a small scattering volume. Suppose that all

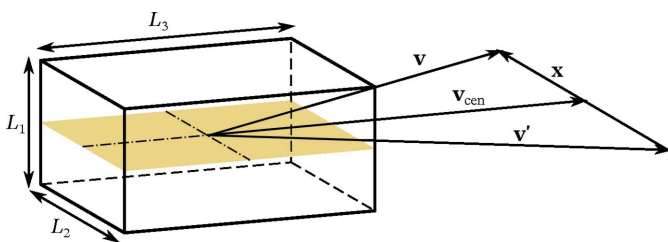


Figure 3 Consider a collection of emitters that is homogeneously distributed in a cuboid scattering volume with edge lengths L_j . The edges of the cuboid are aligned with respect to the scattering vector \mathbf{x} and $\mathbf{v}_{\text{cen}} = (\mathbf{v} + \mathbf{v}')/2$.

emitters lie close together such that $|\mathbf{T}_{mn}| \ll \tau_c$ and $|\mathbf{T}_{mn}| \ll \tau_{\text{exc}}$. Under that assumption (55) simplifies to

$$C_{mn}^{\text{ST}} \simeq \int |\gamma(\tau)|^2 \Psi_0(\tau) d\tau. \quad (56)$$

Importantly, this quantity is position-independent. We can give upper bounds to (56) using the inequality

$$\int f(x)g(x) dx \leq \max_x \{f(x)\} \int g(x) dx, \quad (57)$$

which holds for non-negative functions f and g . Noting that $\int |\gamma(\tau)|^2 d\tau = \tau_c$, we obtain

$$C_{mn}^{\text{ST}} \lesssim \min\{1, \tau_c \Psi_0(0)\}. \quad (58)$$

This shows that the contrast is independent of the emitter geometry and hence ‘pulse limited’, for small scattering volumes.

Larger scattering volumes, on the other hand, do affect the contrast. Consider a cuboid scattering volume, aligned with the observation directions as shown in Fig. 3. Assuming a homogeneous emitter distribution, we can calculate the average over all emitter pairs. The average can be expressed as

$$\begin{aligned} C_{\text{eff}}^{\text{ST}} &= \frac{1}{T_1 T_2 T_3} \int_{-T_1}^{T_1} \int_{-T_2}^{T_2} \int_{-T_3}^{T_3} \Lambda(t_1/T_1) \Lambda(t_2/T_2) \Lambda(t_3/T_3) \\ &\quad \times \int |\gamma(\tau - |\mathbf{x}|t_2/2)| |\gamma(\tau + |\mathbf{x}|t_2/2)| \\ &\quad \times \Psi_0(\tau + |\alpha_1|t_1 + |\alpha_2|t_2 + |\alpha_3|t_3) d\tau dt_1 dt_2 dt_3, \end{aligned} \quad (59)$$

where $T_j = L_j/c$ are the propagation durations, $\alpha_j = (\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_j$ are the respective projections to the coordinate axes, and $\Lambda(t) = \max\{1 - |t|, 0\}$ is the unit triangle function. Note that the dependence on α_j can be neglected when $\tau_{\text{exc}} \gg T_j$ and $\tau_{\text{exc}} \gg \tau_c$ because then Ψ_0 is essentially independent of its argument. Exchanging the order of integration, we obtain

$$\begin{aligned} C_{\text{eff}}^{\text{ST}} &= \frac{1}{T_2} \int_{-T_2}^{T_2} \int \tilde{\Psi}_0(\tau + |\alpha_2|t_2) \Lambda(t_2/T_2) \\ &\quad \times |\gamma(\tau - |\mathbf{x}|t_2/2)| |\gamma(\tau + |\mathbf{x}|t_2/2)| dt_2 d\tau, \end{aligned} \quad (60)$$

with

$$\begin{aligned} \tilde{\Psi}_0(\tau) &= \frac{1}{T_1 T_3} \int_{-T_1}^{T_1} \int_{-T_3}^{T_3} \Lambda(t_1/T_1) \Lambda(t_3/T_3) \\ &\quad \times \Psi_0(\tau + |\alpha_1|t_1 + |\alpha_3|t_3) d\tau dt_1 dt_3. \end{aligned} \quad (61)$$

Note that (60) reduces to (58) in the limit of small T_j . First, consider the effect of T_1 and T_3 on $\tilde{\Psi}_0$. Using (57) on both integrals and extending the bounds to infinity, we obtain

$$\tilde{\Psi}_0(\tau) \leq \min \left\{ \Psi_0(\tau), \frac{1}{|\alpha_1|T_1}, \frac{1}{|\alpha_3|T_3} \right\}. \quad (62)$$

Next, applying (57) on the integral over t_2 in (60) shows that

$$\begin{aligned}
 C_{\text{eff}}^{\text{ST}} &\leq \frac{1}{T_2} \int \int \tilde{\Psi}_0(\tau + |\alpha_2|t) |\gamma(\tau - t|\mathbf{x}|/2)| |\gamma(\tau + t|\mathbf{x}|/2)| dt d\tau \\
 &\leq \frac{2}{|\mathbf{x}|T_2} \int \tilde{\Psi}_0(\tau) \int |\gamma(\tau - t)| |\gamma(\tau + t)| dt d\tau \\
 &= \frac{2}{|\mathbf{x}|T_2} \int \tilde{\Psi}_0(\tau) \text{AC}_{|\gamma|}(2\tau) d\tau. \tag{63}
 \end{aligned}$$

For the second inequality we have set $|\alpha_2| = 0$, which does not decrease the value of $C_{\text{eff}}^{\text{ST}}$. Here $\text{AC}_{|\gamma|}$ denotes the temporal autocorrelation function of $|\gamma(\tau)|$,

$$\begin{aligned}
 \text{AC}_{|\gamma|}(2\tau) &= \int |\gamma(t + 2\tau)| |\gamma(t)| dt \\
 &= \tau_C (1 + 2|\tau|/\tau_C) \exp(-2|\tau|/\tau_C). \tag{64}
 \end{aligned}$$

For the second identity we have used $|\gamma(\tau)| = \exp(-|\tau|/\tau_C)$, corresponding to a Lorentzian line shape. The time integral reads

$$\int \text{AC}_{|\gamma|}(2\tau) d\tau = 2\tau_C^2. \tag{65}$$

Inserting (62) into (63), using (57) to bound the integral over $\Psi_0(\tau)$, and then inserting (65) yields

$$\begin{aligned}
 C_{\text{eff}}^{\text{ST}} &\lesssim \frac{2c\tau_C}{|\mathbf{x}|L_2} \\
 &\times \min \left\{ 1, 2\tau_C \Psi_0(0), \frac{2c\tau_C}{|(\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_1|L_1}, \frac{2c\tau_C}{|(\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_3|L_3} \right\}, \tag{66}
 \end{aligned}$$

which is the equivalent of (58) for larger volumes and large scattering vectors \mathbf{x} .

Equation (66) is only useful for $|\mathbf{x}| > 0$. Going back to (63), we find another inequality for $|\alpha_2| > 0$. More precisely, (63) implies

$$\begin{aligned}
 C_{\text{eff}}^{\text{ST}} &\leq \frac{1}{|\alpha_2|T_2} \int \int \tilde{\Psi}_0(\tau + t) |\gamma(\tau)|^2 dt d\tau \\
 &= \frac{\tau_C}{|\alpha_2|T_2} \int \tilde{\Psi}_0(\tau) d\tau. \tag{67}
 \end{aligned}$$

Here we have set $|\mathbf{x}| = 0$, which only increases the value of the integral. Inserting $\tilde{\Psi}_0(\tau) \leq \Psi_0(\tau)$ yields

$$C_{\text{eff}}^{\text{ST}} \leq \frac{c\tau_C}{|(\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_2|L_2}. \tag{68}$$

This shows that even for small scattering vectors, the contrast can be significantly below the pulse-limited contrast, when the detector is not placed in the forward direction with respect to the excitation pulse. This effect is most pronounced with $|(\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_2| \simeq 2$ in a ‘back-scattering’ geometry but also enters with $|(\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_2| = |\mathbf{v}_{\text{cen}}|$ in the 90° geometry that has been proposed by Classen *et al.* (2017) and is depicted in the work of Trost *et al.* (2020, Fig. 1).

In summary, we have discussed that for small scattering volumes (*i.e.* negligible propagation time), the spatiotemporal contrast is given by (58), which we refer to as the pulse-limited case. For larger scattering volumes, this estimate is complemented by (66) and (68). Since all estimates are rigorous upper bounds, the smallest one takes preference. In particular, the contrast can only be pulse limited when all of $|\mathbf{x}|L_2 \lesssim 2c\tau_C$,

$|(\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_2|L_2 \lesssim c\tau_C$ and $|(\mathbf{v}_{\text{cen}} - \mathbf{v}_{\text{exc}})_j|L_j \lesssim 2c\tau_C$ for $j \in \{1, 3\}$ are satisfied. This is the case, irrespective of the scattering volume, when measuring at small scattering vectors $|\mathbf{x}| \ll 1$ in the forward direction with respect to the excitation pulse, so that $\mathbf{v}_{\text{exc}} \parallel \mathbf{v}_{\text{cen}}$. For all other cases, however, optimal (pulse-limited) contrast requires effectively that $L_j \lesssim c\tau_C$ in all three dimensions. Importantly, we have that $C_{\text{eff}}^{\text{ST}} \lesssim 2c\tau_C/(|\mathbf{x}|L_2)$ even in the best case.

Note that the scattering vector corresponds to a feature size a by $|\mathbf{x}| \sim \lambda/a$. The constraint on L_2 can thus be expressed as

$$L_2/a \lesssim 2c\tau_C/\lambda. \tag{69}$$

The right-hand side is on the order of 10^4 for inner-shell fluorescence.

4.3. Polarization effects

In the preceding discussion we have assumed the emissions to be perfectly polarized such that they can be expressed as scalar functions $u(\mathbf{r}, t)$. In an ensemble of emitters without a preferred spatial orientation however, the individual emitters produce (on average) *unpolarized* emissions. This further reduces the contrast.

Consider two fixed observation directions \mathbf{v} and \mathbf{v}' . For any $\tilde{\mathbf{v}}$ that lies in the plane spanned by \mathbf{v} and \mathbf{v}' the emissions can be decomposed into two orthogonal polarization components, one parallel component (π) and one orthogonal component (σ). A detector that does not distinguish the polarization registers both components, $W^{(\pi)}$ and $W^{(\sigma)}$, independently, such that the total deposited energy can be written as $W = W^{(\pi)} + W^{(\sigma)}$. It follows that

$$\begin{aligned}
 \Gamma &= E[W^{(\sigma)}W^{(\sigma)} + W^{(\pi)}W^{(\pi)} + W^{(\sigma)}W^{(\pi)} + W^{(\pi)}W^{(\sigma)}] \\
 &= \Gamma^{(\sigma)} + \Gamma^{(\pi)} + E[W^{(\sigma)}W^{(\pi)}] + E[W^{(\pi)}W^{(\sigma)}]. \tag{70}
 \end{aligned}$$

Assuming that the two polarization components are uncorrelated and equally intense, we obtain

$$\begin{aligned}
 \Gamma &= \Gamma^{(\sigma)} + \Gamma^{(\pi)} + 2EW^{(\sigma)} \times EW^{(\sigma)} \\
 &\simeq 4EW^{(\sigma)} \times EW^{(\sigma)} \left[1 + \frac{C_{\text{eff}}^{(\sigma)} + C_{\text{eff}}^{(\pi)}}{4} \mathcal{S}(\mathbf{q}) \right]. \tag{71}
 \end{aligned}$$

Comparison with (31) and using $EW = 2EW^{(\sigma)}$ shows that effectively

$$C_{\text{eff}} = \frac{C_{\text{eff}}^{(\sigma)} + C_{\text{eff}}^{(\pi)}}{4}. \tag{72}$$

When both polarizations produce the same contrast, the contrast for unpolarized light is reduced by 2. On the other hand, when one component produces no contrast at all, the contrast is reduced by 4. We can thus incorporate unpolarized emissions simply by introducing $C_{\text{eff}}^{\text{p}}$, with $1/4 \leq C_{\text{eff}}^{\text{p}} \leq 1/2$, as an additional factor into (51).

5. Spatial sampling and finite detectors

We have seen that the intensity correlation $\Gamma(\mathbf{v} - \mathbf{v}')$ between two observation directions \mathbf{v} and \mathbf{v}' gives access to the struc-

ture factor $\mathcal{S}(\mathbf{x})$ at the scattering vector $\mathbf{x} = \mathbf{v} - \mathbf{v}'$. Next we shall discuss requirements on the sampling of observation directions in order to map the structure factor as a *function* of the scattering vector \mathbf{x} . Note that we consider $\mathcal{S}(\mathbf{x})$ as the quantity of interest and ignore the topic of reconstructing the real-space emitter distribution. In particular, *resolution* exclusively refers to the resolution of $\mathcal{S}(\mathbf{x})$. Moreover, we quantify and discuss how finite detectors influence the contrast.

5.1. Spatial spectrum of the correlation signal

Set $EW = EW' = 1$ for this section. Consider $\Sigma(\mathbf{x})$ as a function of the unit-less scattering vector \mathbf{x} . Its spatial spectrum is easily computed from (28) and can be expressed as

$$\mathcal{F}\{\Sigma\}(\mathbf{X}) = \frac{\sum_m \sum_n s_m s_n C_{mn} \delta[\mathbf{X} + \bar{\mathbf{k}}(\mathbf{r}_m - \mathbf{r}_n)]}{|\sum_m s_m|^2}. \quad (73)$$

Note that the spectral amplitudes are real and positive since $\Gamma(\mathbf{x})$ is real and symmetric by definition. The spectrum of the homogeneous background is a single δ function. The spectrum of Σ , on the other hand, is similar to the spatial self-correlation or generalized Patterson function (Cowley, 1995) of the emitter distribution, but attenuated by the C_{mn} coefficients. The coefficients can be pulled out of the sums in the same way as in (31) by approximating them with C_{eff} , so that

$$\mathcal{F}\{\Sigma\}(\mathbf{X}) = C_{\text{eff}} \frac{\sum_m \sum_n s_m s_n \delta[\mathbf{X} + \bar{\mathbf{k}}(\mathbf{r}_m - \mathbf{r}_n)]}{|\sum_m s_m|^2}. \quad (74)$$

Integrating the spectrum over \mathbb{R}^3 gives approximately C_{eff} .

Consider the previously introduced geometry, a homogeneous emitter distribution contained in a box with edge

lengths L_1, L_2 and L_3 as shown in Fig. 3. Approximating the homogeneous emitter distribution by a continuous density, the spatial spectrum (73) can be expressed as

$$\mathcal{F}\{\Sigma\}(\mathbf{X}) = \Sigma(0) \prod_{j=1}^3 \frac{\Lambda(X_j / \Delta X_{s,j})}{\Delta X_{s,j}} \quad (75)$$

with $\Delta X_{s,j} = \bar{k}L_j$, $\Sigma(0) = C_{\text{eff}}$ and the unit triangle function $\Lambda(t) = \max\{1 - |t|, 0\}$.

More generally, *any* compact emitter distribution can be enclosed in such a box and its spatial frequencies are therefore limited by $\Delta X_{s,j}$ for all dimensions j . These frequencies only depend on the respective linear extent but not on the exact form of the scattering volume.

In other words, the structure factor and $\Sigma(\mathbf{x})$ are *band limited* with bandwidth ΔX_s . They can thus be sampled aliasing-free with the Nyquist rate $2\Delta X_s$, that is on a 3D grid with spacing less than or equal to $\pi/\Delta X_{s,j}$.

5.2. Finite pixels

The argument of $\Gamma(\mathbf{x})$ in (36) is determined by the difference of two unit vectors \mathbf{v} and \mathbf{v}' , representing two observation directions. In a real system these observation directions are not points but finite solid angles, given by the pixel size and emitter-to-detector distance. We describe these solid angles as finite patches on the (normalized) Ewald sphere, *i.e.* $\Delta\Omega, \Delta\Omega' \subset \mathbb{S}^2$. The measured correlation function can then be expressed as

$$\begin{aligned} \Gamma(\Delta\Omega, \Delta\Omega') &= \int \int \frac{\mathbf{1}_{\Delta\Omega}(\mathbf{v}) \mathbf{1}_{\Delta\Omega'}(\mathbf{v}')}{|\Delta\Omega| |\Delta\Omega'|} \Gamma(\mathbf{v} - \mathbf{v}') d\Omega d\Omega' \\ &= \int \int \frac{\delta_{\Delta\Omega}(\mathbf{x}) \delta_{\Delta\Omega'}(\mathbf{x}')}{|\Delta\Omega| |\Delta\Omega'|} \Gamma(\mathbf{x} - \mathbf{x}') d\mathbf{x} d\mathbf{x}' \\ &= \{H * \Gamma\}(\mathbf{v}_0 - \mathbf{v}'_0), \end{aligned} \quad (76)$$

with

$$H(\mathbf{x}; \Delta\Omega, \Delta\Omega') = \frac{\int \delta_{\Delta\Omega}(\mathbf{y} + \mathbf{v}_0 - \mathbf{x}) \delta_{\Delta\Omega'}(\mathbf{y} + \mathbf{v}'_0) d\mathbf{y}}{|\Delta\Omega| |\Delta\Omega'|}, \quad (77)$$

where $*$ denotes a 3D convolution, $\mathbf{1}_{\Delta\Omega}(\mathbf{v})$ the indicator function of $\Delta\Omega$ on the sphere \mathbb{S}^2 and $\delta_{\Delta\Omega}(\mathbf{x})$ the Dirac surface delta function (Laplacian of the indicator) of $\Delta\Omega$ on \mathbb{R}^3 . The vectors $\mathbf{v}_0 \in \Delta\Omega$ and $\mathbf{v}'_0 \in \Delta\Omega'$ are the central directions of the two observation regions.

In practice, the two solid angles are small enough to neglect the curvature of the Ewald sphere. They can hence be approximated as planar patches with respective normals \mathbf{v}_0 and \mathbf{v}'_0 [see Fig. 4(a)]. We discuss the symmetric case of two partially aligned planar squares with relative angle φ and side lengths (in practice, the side lengths $\Delta\theta$ are given by the angular size of the detector pixels) $\Delta\theta$ as sketched in Fig. 4(a). Here, the convolution kernel H can be computed analytically. However, the exact expression for H is less important than its support, which is sketched in Fig. 4(b). We approximate H by

$$H(\mathbf{x}) \simeq \prod_{j=1}^3 \Lambda(x_j / \Delta x_{H,j}) / \Delta x_{H,j} \quad (78)$$

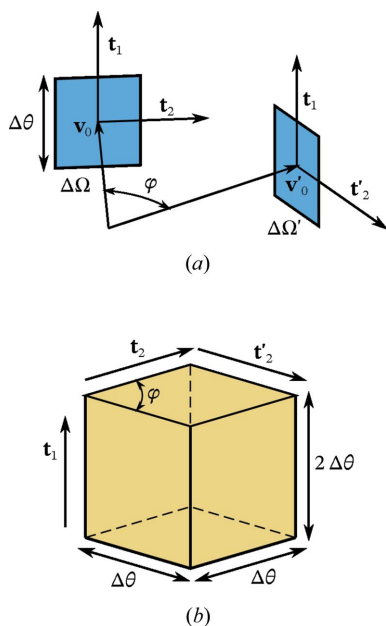


Figure 4
The effect of finite-sized detector pixels (a) on the correlations $\Gamma(\mathbf{x})$ can be described by a 3D convolution $H * \Gamma$. (b) Support of the convolution kernel H .

where

$$\Delta x_{H,1} = \Delta\theta, \Delta x_{H,2} = (1 - |\mathbf{x}/2|^2)^{1/2} \Delta\theta, \Delta x_{H,3} = |\mathbf{x}/2| \Delta\theta. \quad (79)$$

The Fourier transform of this approximate kernel reads

$$\mathcal{F}\{H\}(\mathbf{X}) = \prod_{j=1}^3 \left[\frac{1}{2\pi} \text{sinc}^2(X_j \Delta x_{H,j}/2) \right]. \quad (80)$$

As expected, the finite pixels cause a low-pass filtering of $\Gamma(\mathbf{x})$ with angular frequency scale $1/\Delta x_H$ and thereby limit the resolution with which $\Gamma(\mathbf{x})$ can be measured, irrespective of the sampling rate. The more general case of unaligned detectors corresponds to a less regular convolution kernel but qualitatively has the same effect.

5.3. Contrast

Finite pixels limit the resolution with which the structure factor $\mathcal{S}(\mathbf{x})$ can be measured since higher spatial frequencies are attenuated. However, due to the presence of the intrinsic background, they also limit the achievable intrinsic contrast. We will illustrate this with a simple example that can be computed analytically: the homogeneous emitter distribution in a box as shown in Fig. 3.

For that purpose we decompose $\Gamma = 1 + \Sigma$ into a sum of constant unit background and signal $\Sigma(\mathbf{x}) = C_{\text{eff}} \mathcal{S}(\mathbf{x})$. The measured correlation becomes $H * \Gamma = 1 + H * \Sigma$, using that the background is unaffected by the convolution. We may factorize the approximate convolution kernel $H(\mathbf{x}) = \prod_j H_j(x_j)$ as well as the signal $\Sigma(\mathbf{x}) = \prod_j \Sigma_j(x_j)$ into 1D functions, such that

$$|H * \Sigma| = \prod_j |H_j * \Sigma_j|. \quad (81)$$

Each factor satisfies the inequality

$$|H_j * \Sigma_j| \leq \max_x \{ \Sigma_j(x) \} \int |H_j(x)| dx = \Sigma_j(0). \quad (82)$$

Exploiting the Fourier convolution theorem we find a second triple of inequalities,²

$$\begin{aligned} |H_j * \Sigma_j| &\leq 2\pi \int |\mathcal{F}\{H_j\}(X) \mathcal{F}\{\Sigma_j\}(X)| dX \\ &\leq 2\pi \max_X \{ \mathcal{F}\{\Sigma_j\}(X) \} \int |\mathcal{F}\{H_j\}(X)| dX \\ &= \Sigma_j(0) \frac{2\pi}{\Delta x_{H,j} \Delta X_{s,j}}. \end{aligned} \quad (83)$$

Combining the two inequalities yields

$$|H * \Sigma| \leq \Sigma(0) \prod_{j=1}^3 \min \left\{ 1, \frac{2\pi}{\Delta x_{s,j} \Delta X_{H,j}} \right\}. \quad (84)$$

Substituting $\Delta X_{H,j}$ and $\Delta x_{s,j}$ shows

² Here we implicitly approximate the discrete emitter distribution by a smooth density, to keep the derivation short. A more general derivation is discussed in Appendix A.

$$\begin{aligned} |H * \Sigma| &\leq \Sigma(0) \min \left\{ 1, \frac{L_{\text{crit}}}{L_1} \right\} \min \left\{ 1, \frac{L_{\text{crit}}}{(1 - |\mathbf{x}/2|^2)^{1/2} L_2} \right\} \\ &\quad \times \min \left\{ 1, \frac{L_{\text{crit}}}{|\mathbf{x}/2| L_3} \right\}, \end{aligned} \quad (85)$$

with

$$L_{\text{crit}} = \lambda / \Delta\theta. \quad (86)$$

A generalization of this result, not relying on specific assumptions on the shape of Σ and H , is presented in Appendix A. Equation (85) implies that the contrast falls off with $1/L_j$ for each linear extent L_j of the scattering volume that exceeds a critical length in the order of L_{crit} . Note that (85) is based on a not necessarily sharp upper bound on the signal strength, so that the contrast may suffer even for smaller L_j .

To put this into perspective, consider $K\alpha$ radiation of iron ($\lambda \simeq 0.19$ nm) and a typical pixel detector with 50 μm pixel size. Placing the detector 10 cm, 1 m or 10 m from the emitters corresponds to L_{crit} of 0.38 μm , 3.8 μm or 38 μm , respectively.

6. Photon count estimates

We have seen in (48) that the SNR is bounded by an expression proportional to $C_{\text{eff}} \times \bar{K}$. Here, we derive universal upper bounds on this product, based on the fact that the emitter density $\bar{\rho}_E$ is finite and that the contrast decreases with increasing extent of the scattering volume.

The mean number of photons emitted into 4π from a cuboid scattering volume with edge lengths L_j (see Fig. 3) is $nL_1L_2L_3\bar{\rho}_E$, where n is the emission efficiency in the considered energy band. A detector pixel of angular size $\Delta\theta$ and quantum efficiency η registers on average

$$\bar{K} = \eta n \bar{\rho}_E L_1 L_2 L_3 \frac{\Delta\theta^2}{4\pi} \quad (87)$$

counts if we neglect self-absorption within the scattering volume. We have derived two independent constraints between scattering volume and contrast. First, (66) implies, in particular, that the contrast falls off with $1/L_2$ for $L_2 \gtrsim 2c\tau_c/|\mathbf{x}|$ (coherence time constraint). We set

$$\beta_C = \frac{2c\tau_c}{|\mathbf{x}|L_2} \quad (88)$$

to quantify the margin by which the coherence time constraint is satisfied. Second, (85) implies that the contrast falls off with $1/L_j$ for each dimension j in which the linear extent of the scattering volume L_j is larger than a certain critical size (sampling constraint). We analogously set

$$\begin{aligned} \beta_{s,1} &= \frac{\lambda}{L_1 \Delta\theta}, \beta_{s,2} = \frac{\lambda}{(1 - |\mathbf{x}/2|^2)^{1/2} L_2 \Delta\theta}, \\ \beta_{s,3} &= \frac{\lambda}{|\mathbf{x}/2| L_3 \Delta\theta} \end{aligned} \quad (89)$$

to quantify the margin by which the sampling constraints are satisfied. Inserting (88) and (89) into (87) yields

$$\bar{K} \leq \frac{\eta n}{\pi \beta_{S,1} \beta_{S,3} \beta_C} \times \frac{\bar{\rho}_E \lambda^2 c \tau_C}{|\mathbf{x}|^2}. \quad (90)$$

The unit-less scattering vector corresponds to a feature size a via $|\mathbf{x}| \sim \lambda/a$, so that (90) can be expressed as

$$\bar{K} \leq \frac{\eta n}{\pi \beta_{S,1} \beta_{S,3} \beta_C} \times \bar{\rho}_E a^2 c \tau_C. \quad (91)$$

Having expressed \bar{K} in terms of the parameters β , we can optimize the product $C_{\text{eff}} \times \bar{K}$. Equations (66) and (85) imply that

$$C_{\text{eff}} = \tilde{C} \times \min\{1, \beta_{S,1}\} \times \min\{1, \beta_{S,2}\} \times \min\{1, \beta_{S,3}\} \times \min\{1, \beta_C\}, \quad (92)$$

where $\tilde{C} \leq 1$. Since (92) becomes proportional to every β that is less than one, it follows that

$$C_{\text{eff}} \times \bar{K} \leq \tilde{C} \times \bar{K}|_{\beta=1}, \quad (93)$$

where $\beta = 1$ refers to all the coefficients. In particular, increasing the mean photon count \bar{K} beyond its value for $\beta \simeq 1$ does not increase the SNR, because the contrast is reduced just as much. In other words, there is an optimal size of the scattering volume, beyond which the SNR does not increase any further.

Although some β -dependent factors are included in \tilde{C} and are not explicitly written in (92), they only reduce the contrast further. Because of these factors and since the individual inequalities are not sharp (there can be significant margins between the left- and right-hand side), the maximum SNR may actually be attained with some $\beta \gtrsim 1$ but, nevertheless, will not exceed the given bound.

For $|\mathbf{x}| \ll 1$, the sampling constraint on L_3 becomes arbitrarily large and thereby exceeds the self-absorption length (the $1/e$ length of the emitted light in the material), so that (90) strongly overestimates the achievable \bar{K} . We can use $\beta_{S,2}$ instead of β_C to quantify L_2 to obtain

$$\bar{K} \leq \frac{\eta n}{4\pi \beta_{S,1} \beta_{S,2}} \times \bar{\rho}_E \lambda^2 L_{\text{SA}}, \quad (94)$$

where L_{SA} is the self-absorption length. This estimate is independent of the coherence time.

To put (91) into perspective, consider a dense mono-elemental iron crystal with $\tau_C \simeq 2.6$ fs, $a \simeq 0.29$ nm and $\bar{\rho}_E \simeq 85$ nm⁻³. For these values, (69) yields $L_2 \lesssim 2.3$ μm . Optimistically assuming $\eta = 1$, $\beta_C = 1$ and $\beta_S = 4$, (91) shows that $n \simeq 0.1\%$ of all atoms are required to emit a photon for an average photon count of $\bar{K} = 0.1$ photons per detector pixel.

The right-hand side of (91) is proportional to the squared feature size a^2 and therefore grows quadratically with increasing a . However, when measuring atomic distances, it effectively scales with $1/a$, because the mean emitter density $\bar{\rho}_E$ scales with $1/a^3$. Consider a crystal structure with lattice constant a , such that $\bar{\rho}_E \sim 1/a^3$. The number $\bar{\rho}_E a^2 c \tau_C$, which bounds the mean photon count \bar{K} in (91), is then essentially governed by $c \tau_C/a$. In general, increasing the feature size a increases the SNR only as long as the mean emitter density decays slower than $1/a^2$.

As an example consider a crystal with ten times the lattice constant of iron, $a \simeq 2.9$ nm, and one iron atom per unit cell. The emission efficiency needs to be ten times as high for the same photon yield as for the pure iron crystal, so that, based on our previous estimate, $n \simeq 1\%$ for $\bar{K} = 0.1$.

In summary, we have shown that an optimal size of the scattering volume exists. Larger volumes do not result in higher SNR, even though the photon count is increased. The mean photon count corresponding to optimal SNR can be estimated by (90), (91) or (94), with all coefficients $\beta \gtrsim 1$.

7. Discussion

Next, we use our findings to discuss the feasibility of different experiments and to highlight the challenges. We chose two particular experiments to illustrate the fundamentally different experimental regimes.

7.1. Determination of illumination spot size

First, consider an experiment to determine the illumination spot size, following Inoue *et al.* (2019). A spot size of $L_2 = 0.5$ μm with copper $K\alpha$ radiation ($\lambda = 0.154$ nm) corresponds to a scattering vector $|\mathbf{x}| \sim \lambda/L_2 = 3 \times 10^{-4}$. Correspondingly, the width of the zeroth-order correlation peak has to be resolved, similar to the primary beam profile in classical small-angle X-ray scattering geometry. To this end, the detector is placed in the forward direction a few metres downstream of the sample. The sample can be a foil so that the focal spot size of the excitation pulse determines the transverse extent of the scattering volume. Here, since the scattering vectors are small, the sample thickness L_3 can be chosen largely independently of the detection geometry, in practice up to the self-absorption depth. In particular, the contrast is pulse limited, because $|\mathbf{x}|L_2 \sim \lambda \ll c \tau_C$ and $\mathbf{v}_{\text{exc}} - \mathbf{v}_{\text{cen}}$ vanishes.

As a rough estimate, consider a 10 fs excitation pulse, 2 fs coherence time (corresponding to copper $K\alpha$), $C_{\text{eff}}^{\text{p}} = 1/2$ and $C_{\text{eff}}^{\text{l}} = 5/9$ (the two $K\alpha$ lines) so that $C_{\text{eff}} \sim 0.05$. The self-absorption length depends on the material and the emission line. For the $K\alpha$ lines of copper with 8.96 g cm⁻³ mass density the $1/e$ length is 22 μm . Moreover, by using a highly focused XFEL beam it should be possible to ionize a large fraction of the sample atoms, so that a mean photon number of 0.1 per pixel can be easily achieved according to (94). Using a pixel detector with 1000×1000 pixels provides $\nu \sim 10^6$ parallel realizations. Inserting these numbers into (49) estimates that at least 14 pulses would be required to sample the half maximum $\mathcal{S} = 0.5$ with a signal-to-noise level of 10. The experiment is relatively robust, because translations and rotations of the scattering volume do not affect the signal in first order. Inoue *et al.* (2019) used such a measurement to determine the pulse duration and the focal size at SACLA. In contrast, if one wanted to perform an analogous experiment at a laser-driven plasma source (Schoenlein *et al.*, 2019), where on the order of 10^8 copper $K\alpha$ emissions occur per excitation pulse of about 100 fs, the emission efficiency can be estimated to 10^{-5} based on the atom density of copper, a spot size of

about 2 μm and a target thickness of 10 μm . The low emission efficiency results in at most 10^{-3} photons per pixel according to (94). The squared dependence on the photon count and contrast in (49) then would require more than 10^7 pulses. Ironically, initial attempts of such an endeavor had motivated the present work.

7.2. Atomic resolution from Bragg scattering

Second, consider an experiment to resolve crystal planes. As an example, consider the $K\alpha_1$ radiation from iron with $\tau_C = 2.6$ fs, $\lambda = 0.19$ nm and a total K -shell fluorescence yield of 35% (Schoonjans *et al.*, 2011). Pure iron at room temperature has a b.c.c. (body-centered cubic) lattice with lattice constant 0.29 nm and number density $\bar{\rho}_E = 85$ nm $^{-3}$. Since the wavelength is fixed by the emission line, the lowest (110) reflections have a (unit-less) scattering vector of $\mathbf{q}/k = 0.927$. To reach this peak, the pixel array detector needs a field of view of about 55° in one dimension. The accessible scattering vectors are bounded by $|\mathbf{v} - \mathbf{v}'| \leq 2$, so that the highest accessible reflection is the (220). As previously discussed, the coherence time constrains the sample size to about $\simeq 2$ μm through (69).

Consider a cubic perfect crystal with diameter 0.5 μm ($\beta_C \simeq 4$) so that the contrast is excitation-pulse limited. The sampling constraint with $\beta_S = 4$ requires an angular pixel size of $\Delta\theta \lesssim 6$ mdeg, so that a detector covering 60° requires 10^4 pixel-columns in that direction. This could be realized by arranging ten 1M detectors in an arc. Fewer detector modules with larger gaps could also be used in principle, at the cost of covering only specific Bragg peaks.

A very optimistic estimate on the contrast for a 10 fs-long XFEL pulse gives $C_{\text{eff}} = 0.07$, for perfect conditions. This estimate assumes a perfect single crystal and that the system is perfectly stationary and stable. In particular, the orientation of the sample has to be stable on the order of magnitude of $\Delta\theta/|\mathbf{x}|$. An uncertainty σ in the sample orientation that exceeds this limit smears out the Bragg peaks over several resolution elements and therefore decreases the contrast by $(\Delta\theta/|\mathbf{x}|/\sigma)^2$. An uncertainty of $\sigma \sim 60$ mdeg, for example, would decrease the contrast by 10^{-2} down to 7×10^{-4} . Whereas a stable sample orientation of 6 mdeg is easily obtained for static and extended samples, it is non-trivial to reach this accuracy in single-particle experiments with random orientations. The coherent diffraction signal of the particles could be used to calculate the particle orientation in each pulse. However, solely from a sampling perspective, detectors with a huge pixel number would be necessary to reach the desired angular accuracy σ , neglecting further experimental inaccuracies of 3D orientation determination. In this specific example, the coherent diffraction signal needs to be sampled on a detector with a minimum of about $10^4 \times 10^4$ pixels to obtain $\sigma \sim 6$ mdeg. Note that additional factors such as lattice vibrations (Debye–Waller factor), lattice strain and defects have not been considered.

Next, we discuss the multiplicity of correlation measurements from a single pulse (detector frame). The scattering

vectors given by the sets of all pixel pairs are distributed in a volume (Classen *et al.*, 2017). Most of the realizations correspond to small scattering vectors, while the larger scattering vectors of the Bragg peaks have a strongly reduced multiplicity ν . Assuming 10^3 pixel rows gives a conservative estimate of $\nu \sim 10^3$ parallel realizations for a Bragg peak signal.

Assuming a mean photon count of $\bar{K} = 0.1$ and using (49) we see that on the order of 2×10^6 realizations are required for an SNR of 10, which could be acquired within 2×10^3 pulses, assuming optimal conditions. Since the number of pulses depends quadratically on the contrast, an uncertainty in the sample orientation of 0.06° would increase the required number of pulses to 2×10^7 , which is a sizeable number for a dense mono-elemental iron sample.

It will be challenging to realize a mean photon count of even $\bar{K} = 0.1$ for dilute samples. Using (91) with $\beta_C = \beta_S = 4$ and assuming 100% detection efficiency shows that about 0.4% of the atoms in the pure iron sample need to emit a $K\alpha$ photon for a per-pixel photon yield of 0.1. Correspondingly, to achieve the same resolution in a sample with 1% iron content, 40% of the iron atoms would have to emit a $K\alpha$ photon, which is already above the K -shell fluorescence yield. Therefore, such dilute samples can only produce a photon count of less than 0.1 photons per pixel, even when fully ionized. Optimizing the geometry for smaller scattering vectors, *i.e.* coarser resolution, enables one to use larger scattering volumes and can improve the photon yield to some degree. However, the given estimate already corresponds to a sample with diameter 0.5 μm , which does not leave much room for increase in the case of nanocrystallography or single-molecule diffraction.

8. Summary and conclusions

We have derived comprehensive equations relating the two-point intensity correlations to the structure factor \mathcal{S} of the emitter configuration. We have reproduced the expression given by Classen *et al.* (2017) up to an additive term of order $\mathcal{O}(1/N_E)$, which is hence irrelevant for a large number of atomic emitters. This agreement with the results of Classen *et al.* (2017) and the classical description presented by Trost *et al.* (2020) underlines that IDI does neither rely on any non-classical states of light nor beat any classical limits.

By including time dependence, we have obtained an explicit expression for the contrast between the structural signal and the inherent background in the correlation functions. Equation (28) shows that the total signal can be decomposed into a sum of terms from individual pairs of emitters and that each term is attenuated by a coupling constant C_{mn} that takes values from 0 to 1. Averaging the coupling constant over all pairs of emitters gives an effective contrast C_{eff} of less than one.

We have given an estimate for the SNR, equation (48). In the low-photon regime, it scales linearly with the mean photon count \bar{K} and the signal strength $C_{\text{eff}}\mathcal{S}$. Importantly, we have obtained a rigorous upper bound on the SNR for two statistically *dependent* coincident measurements.

Based on our model, we have identified several factors influencing the contrast and have quantified them in terms of experimentally accessible parameters. First, the fact that multiple emission lines contribute to the total signal decreases the contrast by a factor C_{eff}^L that is inversely proportional to the number of emission lines and depends on their relative strengths. This emphasizes the need for an effective energy discrimination of the emitted radiation. Second, the lack of polarization of the emitted radiation decreases the contrast by a factor of C_{eff}^P , where $1/4 \leq C_{\text{eff}}^P \leq 1/2$ depending on the angular separation of the two observation directions. Third, the finite coherence time τ_C , the finite propagation time through the sample, and the finite duration of the excitation pulse τ_{exc} strongly affect the contrast by a spatiotemporal factor $C_{\text{eff}}^{\text{ST}}$. The scaling of the contrast depends on the relative magnitude of the three involved time scales. In particular, the pulse-limited scaling, (58), does only apply for small scattering volumes, or when both observation directions are in a small cone around the propagation direction of the excitation pulse \mathbf{v}_{exc} . In general, the contrast is additionally affected by the extent of the scattering volume, as detailed in (66). It implies that $C_{\text{eff}}^{\text{ST}} \lesssim 2c\tau_C/|\mathbf{x}|/L_2$ which effectively restricts the linear extent L_2 of the scattering volume for larger scattering vectors $|\mathbf{x}|$. Although for two fixed directions (or pixel coordinates), the size restriction applies to only one dimension, for effective use of a 2D pixel array, involving simultaneous measurements at many different scattering vectors $\mathbf{v} - \mathbf{v}'$, it effectively applies to all dimensions.

The contrast is also affected by the integration associated with finite detector pixels. If the detector pixels are too large to properly resolve the speckle patterns, the signal is decreased while the inherent background remains unaffected, resulting in a loss of contrast. The scaling of the contrast can be expressed in terms of a critical length $L_{\text{crit}} = \lambda/\Delta\theta$ depending on the wavelength λ and the angular pixel size $\Delta\theta$. According to (85), the contrast is decreased proportional to L_{crit}/L_j for each dimension j in which the diameter of the scattering volume exceeds a critical length. For small scattering vectors, only the transverse dimensions L_1 and L_2 are relevant while the thickness L_3 is effectively unconstrained, whereas for large scattering vectors, all dimensions contribute approximately equally.

Importantly, the constraints of the critical linear extent of the scattering volume, as discussed above, also directly limit the total photon yield per pixel for SPE due to their finite number density $\bar{\rho}_E$. We have shown in particular that the SNR cannot be improved beyond an optimal value, which is significantly lower than anticipated. Increasing the photon count beyond its corresponding optimum decreases the contrast and does not improve the SNR. The optimal photon count and SNR depend on the magnitude of the scattering vectors $|\mathbf{x}|$ and thus on the probed length scales. For small scattering vectors, the sample thickness can be increased independently of the transverse size to optimize the photon yield, so that the photon yield is bounded by (94). In contrast, for larger scattering vectors, all dimensions enter with the same scaling and the photon yield is bounded by (91).

Although both constraints had been mentioned by Classen *et al.* (2017) and Trost *et al.* (2020), their implications for the photon yield had not yet been quantified or discussed.

Based on these bounds, we have discussed examples of two fundamentally different experimental regimes. First, an experiment aiming to resolve the geometry of the scattering volume as demonstrated by Inoue *et al.* (2019). Second, an experiment to resolve the crystal structure within the scattering volume as was proposed by Classen *et al.* (2017). For the latter case, we have shown that the best possible SNR is inversely proportional to the lattice constant when aiming for atomic resolution. We have also given estimates for the best possible photon count and the corresponding fractions of ionized atoms. Moreover, we have discussed how the pulse-to-pulse orientational stability influences the SNR.

We would like to stress that the simulation presented by Classen *et al.* (2017) strongly overestimates the achievable photon count and SNR for the discussed geometry. More than 5 photons per pixel are only possible with a large and dense sample with optimal geometry, which is inconsistent with the stated assumptions. Similarly, Trost *et al.* (2020) used mean photon counts in the range from 10^{-2} to 10^3 in their simulations. We have shown that for atomic resolution IDI even $\bar{K} \sim 10^{-1}$ can be achieved without sacrificing the SNR only for samples with a high density of emitters, whereas for dilute samples with a lower emitter density, such as macromolecules, $\bar{K} \lesssim 10^{-1}$ is more realistic.

In light of the low SNR, even under idealized conditions, and the required pulse-to-pulse stability, we come to the conclusion that utilizing IDI for serial crystallography will be extremely challenging in general – even more so for diffractive imaging of single molecules, as was envisioned by Classen *et al.* (2017) and deemed achievable by Trost *et al.* (2020).

We hope that our quantitative estimates may serve as a solid basis for discussing the use of structure determination based on incoherent emissions. In particular, we hope that our results may be useful to assess the limit of length scales that can be reasonably probed. Finally, we would like to mention that the derived limits, which are quite fundamental, scale very favorably with the coherence time of the emissions. It should therefore not escape our attention that emissions with longer coherence time such as visible light fluorescence could result in quite realistic IDI experiments on larger length scales.

APPENDIX A Contrast for arbitrary emitter distributions

We utilized concrete assumptions on the shape of the functions Σ and H in the derivation of (85). In particular we assumed the multiplicative separability of the two functions to estimate each dimension separately. Clearly, this strategy fails for more general classes of Σ . Here we outline a similar derivation without the need the factorize Σ and H .

Following the same approach with the full 3D functions shows that $|H * \Sigma| \leq \Sigma(0)$ and

$$|H * \Sigma| \leq (2\pi)^3 \int |\mathcal{F}\{H\}(\mathbf{X})| \mathcal{F}\{\Sigma\}(\mathbf{X}) d^3\mathbf{X}. \quad (95)$$

Unfortunately the next inequality in (83) cannot be applied here, because $\mathcal{F}\{\Sigma\}$ is a distribution. Nevertheless, $\mathcal{F}\{H\}$ is smooth and slowly oscillating because H is compactly supported. We may therefore approximate the emitter distribution $s(\mathbf{X})$ by a continuous emitter density $\rho_E(\mathbf{X})$ such that

$$\Sigma(\mathbf{x}) \simeq \Sigma(0) \frac{|\mathcal{F}\{\rho_E\}(\mathbf{x})|^2}{|\mathcal{F}\{\rho_E\}(0)|^2} = \Sigma(0)(2\pi)^6 \frac{|\mathcal{F}\{\rho_E\}(\mathbf{x})|^2}{|\int \rho_E(\mathbf{X}) d^3\mathbf{X}|^2}. \quad (96)$$

We can thus express the Fourier transform of Σ by the normalized cross-correlation of ρ_E ,

$$\begin{aligned} \mathcal{F}\{\Sigma\}(\mathbf{X}) &\simeq \Sigma(0) \frac{\{\rho_E * \rho_E\}(\mathbf{X})}{|\int \rho_E(\mathbf{X}') d^3\mathbf{X}'|^2} \\ &\equiv \Sigma(0) \frac{\int \rho_E(\mathbf{X}') \rho_E(\mathbf{X}' + \mathbf{X}) d^3\mathbf{X}'}{|\int \rho_E(\mathbf{X}') d^3\mathbf{X}'|^2}. \end{aligned} \quad (97)$$

Substituting $\mathcal{F}\{\Sigma\}$ in (95) gives

$$\begin{aligned} &\int |\mathcal{F}\{H\}(\mathbf{X})| \mathcal{F}\{\Sigma\}(\mathbf{X}) d^3\mathbf{X} \\ &\leq \Sigma(0) \frac{\max_{\mathbf{X}} \{\{\rho_E * \rho_E\}(\mathbf{X})\}}{|\int \rho_E(\mathbf{X}') d^3\mathbf{X}'|^2} \int |\mathcal{F}\{H\}(\mathbf{X})| d^3\mathbf{X} \\ &= \Sigma(0) \frac{\int \rho_E(\mathbf{X}')^2 d^3\mathbf{X}'}{|\int \rho_E(\mathbf{X}') d^3\mathbf{X}'|^2} \int |\mathcal{F}\{H\}(\mathbf{X})| d^3\mathbf{X}. \end{aligned} \quad (98)$$

We obtain

$$|H * \Sigma| \leq \Sigma(0) \frac{(2\pi)^3}{V_\rho V_H} \left[\frac{\max_{\mathbf{X}} \{\rho_E(\mathbf{X})\}}{\min_{\mathbf{X}} \{\rho_E(\mathbf{X})\}} \right]^2 \quad (99)$$

where V_ρ denotes the volume of the emitter distribution [in units of $(2\pi/\lambda)^3$] and

$$V_H := \left[\int |\mathcal{F}\{H\}(\mathbf{X})| d^3\mathbf{X} \right]^{-1}. \quad (100)$$

Our approximate expression (78) yields $V_H = (\Delta\theta)^3 |\mathbf{x}/2| (1 - |\mathbf{x}/2|^2)^{1/2}$, roughly corresponding to the volume of the support of H . Equation (99) is a generalization of (83) for arbitrary detector shapes and scattering volumes.

Approximating $\mathcal{F}\{\Sigma\}$ by a continuous emitter density requires $\mathcal{F}\{H\}$ to be approximately constant on the (dimension-less) length scale of the emitter distance. On the other hand, in the limit of very wide H and thus very sharply peaked $\mathcal{F}\{H\}$ the individual δ -distributions are isolated in (95). We then obtain that

$$|H * \Sigma| \leq \Sigma(0) \frac{\sum_m s_m^2}{|\sum_m s_m|^2} \leq \Sigma(0) \frac{\max_m \{s_m\}}{\min_m \{s_m\}} \frac{1}{N_E} \quad (101)$$

independent of the scattering volume and detector size.

APPENDIX B

Variance of dependent random variables

Let $(X, X') \sim \mathcal{P}_{X, X'}$ be a pair of possibly dependent random variables on \mathbb{R}^+ . Define Y, Y' conditionally via

$\mathcal{P}_{Y, Y'|X=x, X'=x'} = \text{Poi}(x) \otimes \text{Poi}(x')$ as a tensor product of Poisson distributions. We seek an expression for $\text{Var}[(Y - \text{E}Y)(Y' - \text{E}Y')]$ in terms of X and X' .

We have $\text{E}Y = \text{E}[\text{E}Y | X] = \text{E}X$ and $\text{E}Y' = \text{E}X'$, as well as

$$\begin{aligned} \text{Cov}[Y, Y'] &= \text{E}[YY'] - \text{E}Y\text{E}Y' \\ &= \text{E}[\text{E}[YY' | X, X'] - \text{E}X\text{E}X'] \\ &= \text{E}[XX'] - \text{E}X\text{E}X' = \text{Cov}[X, X']. \end{aligned} \quad (102)$$

The second moment can be written as

$$\begin{aligned} &\text{E}[(Y - \text{E}Y)^2(Y' - \text{E}Y')^2] \\ &= \text{E}[(Y - \text{E}X)^2(Y' - \text{E}X')^2] \\ &= \text{E}[\text{E}[(Y - \text{E}X)^2(Y' - \text{E}X')^2 | X, X']] \\ &= \text{E}[(X + (X - \text{E}X)^2)(X' + (X' - \text{E}X')^2)], \end{aligned} \quad (103)$$

where the last step can be shown as follows. Let $x, x' \geq 0$ be two non-negative numbers and define $(Z, Z') \sim \text{Poi}(x) \otimes \text{Poi}(x')$. Then

$$\begin{aligned} &\text{E}[(Y - \text{E}X)^2(Y' - \text{E}X')^2 | X = x, X' = x'] \\ &= \text{E}[(Z - \text{E}X)^2(Z' - \text{E}X')^2] \\ &= \text{E}[(Z - \text{E}X)^2] \text{E}[(Z' - \text{E}X')^2] \\ &= (x + (x - \text{E}X)^2)(x' + (x' - \text{E}X')^2). \end{aligned} \quad (104)$$

The second *central* moment then becomes

$$\begin{aligned} &\text{Var}[(Y - \text{E}Y)(Y' - \text{E}Y')] \\ &= \text{E}[(Y - \text{E}Y)^2(Y' - \text{E}Y')^2] - (\text{Cov}[Y, Y'])^2 \\ &= \text{E}[(X + (X - \text{E}X)^2)(X' + (X' - \text{E}X')^2)] - (\text{Cov}[X, X'])^2 \end{aligned} \quad (105)$$

which can be expressed as

$$\begin{aligned} &\text{Var}[(Y - \text{E}Y)(Y' - \text{E}Y')] \\ &= \text{E}[XX'] + \text{E}[(X' - \text{E}X')^2 X] \\ &\quad + \text{E}[(X - \text{E}X)^2 X'] + \text{E}[(X' - \text{E}X')^2 (X - \text{E}X)^2] \\ &\quad - (\text{Cov}[X, X'])^2. \end{aligned} \quad (106)$$

APPENDIX C

List of the main symbols used in the paper

- N_E , number of emitters.
- \mathbf{r}_m , position of emitter m .
- b_m, s_m , random variable and probability that emitter m emits a photon.
- $\hbar\omega_m$, photon energy of the emission from emitter m .
- t_m , time when emitter m emits a photon.
- $u_m(\mathbf{r}, t)$, scalar field amplitude of the emission from emitter m .
- $w_m d\Omega$, energy flow of u_m into the solid angle $d\Omega$.
- $U(\mathbf{r}, t)$, total scalar field amplitude from all emissions.
- $W d\Omega$, total energy flow into the solid angle $d\Omega$.
- Γ , two-point correlation of energy flows (or photon counts).

Σ , two-point covariance of energy flows (or photon counts).
 $\gamma(\tau)$, complex degree of coherence (CDC).
 τ_C , coherence time.
 C_{mn} , coupling coefficients of the two-photon contributions.
 C_{eff} , average coupling coefficient/contrast.
 \mathbf{v}, \mathbf{v}' , observation directions.
 $\mathbf{x} = \mathbf{v} - \mathbf{v}'$, unit-less scattering vector.
 $S(\mathbf{x})$, structure factor.
 K , photon count (random variable).
 \bar{K} , average photon count (number).
 R , number of independent realizations.
 $\nu(\mathbf{x})$, multiplicity/number of realizations for scattering vector \mathbf{x} .
 $\mathbf{T}_{mn} = (\mathbf{r}_m - \mathbf{r}_n)/c$, distance from emitter m to n .
 $\tau_{mn} = t_m - t_n$, difference of emission times.
 $\Psi(\mathbf{r}_m, \mathbf{r}_n, \tau)$, probability density of τ_{mn} .
 $I_{\text{exc}}(\mathbf{r}, t) \equiv I_{\text{exc}}(\mathbf{v}_{\text{exc}} \cdot \mathbf{r} - ct)$, cycle-averaged intensity of the excitation pulse.
 $\Psi_0(\tau)$, temporal autocorrelation of the excitation pulse $I_{\text{exc}}(\tau)$.
 \mathbf{v}_{exc} , propagation direction of the excitation pulse.
 τ_{exc} , duration of the excitation pulse.
 $\mathbf{v}_{\text{cen}} = (\mathbf{v} + \mathbf{v}')/2$, arithmetic mean of observation directions.
 $L_j, j = 1, \dots, 3$, side lengths of scattering volume.
 a , feature size or lattice constant.
 $\Delta\theta$, angular pixel size.
 c , speed of light.
 $\bar{\rho}_E$, mean emitter density in the scattering volume.
 λ , (mean) wavelength of the emitted light.
 η , detector quantum efficiency.
 n , effective emission efficiency.
 β_C , number that quantifies the coherence time constraint.
 $\beta_{S,j}$, numbers that quantify the sampling constraints.

Acknowledgements

We thank Thomas Staudt for his help with the derivation of the variances of dependent random variables. Open access funding enabled and organized by Projekt DEAL.

Funding information

LML, MV and TS are part of the Max Planck School of Photonics supported by BMBF, Max Planck Society and Fraunhofer Society. We also acknowledge support by the

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through grant SFB 1456-C03.

References

- Agarwal, G. S. (2013). *Quantum Optics*. Cambridge: Cambridge University Press.
- Chapman, H. N., Barty, A., Bogan, M. J., Boutet, S., Frank, M., Hau-Riege, S. P., Marchesini, S., Woods, B. W., Bajt, S., Benner, W. H., London, R. A., Plönjes, E., Kuhlmann, M., Treusch, R., Düsterer, S., Tschentscher, T., Schneider, J. R., Spiller, E., Möller, T., Bostedt, C., Hoener, M., Shapiro, D. A., Hodgson, K. O., van der Spoel, D., Burmeister, F., Bergh, M., Caleman, C., Hultdt, G., Seibert, M. M., Maia, F. R. N. C., Lee, R. W., Szöke, A., Timneanu, N. & Hajdu, J. (2006). *Nat. Phys.* **2**, 839–843.
- Classen, A., Ayyer, K., Chapman, H. N., Röhlberger, R. & von Zanthier, J. (2017). *Phys. Rev. Lett.* **119**, 053401.
- Cowley, J. (1995). *Diffraction Physics*. North-Holland Personal Library. Amsterdam: Elsevier Science.
- Gog, T., Len, P. M., Materlik, G., Bahr, D., Fadley, C. S. & Sanchez-Hanke, C. (1996). *Phys. Rev. Lett.* **76**, 3132–3135.
- Goodman, J. (1985). *Statistical Optics*. New York: Wiley.
- Gureyev, T. E., Kozlov, A., Paganin, D. M., Nesterets, Y. I., De Hoog, F. & Quiney, H. M. (2017). *J. Opt. Soc. Am. A*, **34**, 1577.
- Hanbury Brown, R. & Twiss, R. Q. (1956). *Nature*, **178**, 1046–1048.
- Ho, P. J., Knight, C. & Young, L. (2020). *Phys. Rev. A*, **101**, 043413.
- Inoue, I., Tamasaku, K., Osaka, T., Inubushi, Y. & Yabashi, M. (2019). *J. Synchrotron Rad.* **26**, 2050–2054.
- Kossel, W., Loeck, V. & Voges, H. (1935). *Z. Phys.* **94** (1–2), 139–144.
- Laue, M. (1935). *Ann. Phys.* **415**(8), 705–746.
- Mandel, L. (1999). *More Things in Heaven and Earth*, pp. 460–473. New York: Springer.
- Mandel, L. & Wolf, E. (2015). *Optical Coherence and Quantum Optics*. Cambridge University Press.
- Miao, J., Charalambous, P., Kirz, J. & Sayre, D. (1999). *Nature*, **400**, 342–344.
- Miao, J., Ishikawa, T., Robinson, I. K. & Murnane, M. M. (2015). *Science*, **348**, 530–535.
- Schneider, R., Mehringer, T., Mercurio, G., Wenthaus, L., Classen, A., Brenner, G., Gorobtsov, O., Benz, A., Bhatti, D., Bocklage, L., Fischer, B., Lazarev, S., Obukhov, Y., Schlage, K., Skopintsev, P., Wagner, J., Waldmann, F., Willing, S., Zaluzhnyy, I., Wurth, W., Vartanyants, I. A., Röhlberger, R. & von Zanthier, J. (2017). *Nat. Phys.* **14**, 126–129.
- Schoenlein, R., Elsaesser, T., Holldack, K., Huang, Z., Kapteyn, H., Murnane, M. & Woerner, M. (2019). *Philos. Trans. R. Soc. Am* **377**, 20180384.
- Schoonjans, T., Brunetti, A., Golosio, B., Sanchez del Rio, M., Solé, V. A., Ferrero, C. & Vincze, L. (2011). *At. Spectrosc.* **66**, 776–784.
- Singer, A. & Vartanyants, I. A. (2014). *J. Synchrotron Rad.* **21**, 5–15.
- Trost, F., Ayyer, K. & Chapman, H. N. (2020). *New J. Phys.* **22**, 083070.