**Raw cumulants**

The sample is composed of M stochastically independently fluctuating emitters. The intensity distribution over time per pixel can be written as

where , and denote the brightness, position and normalized temporal fluctuations of the kth-emitter. is the point-spread function (PSF) and *b* represents a temporally constant background. Supposing an infinite resolution and zero noise contribution, we can write

which represents the object. Ultimately, we would like to design an imaging system such that in an ideal case, we would image perfectly the object

To approach this perfect imaging, we will take N images for achieving a “super-resolved image”. In fact, we exploit the temporal fluctuations registered as a sequence of individual images. can be split in a spatial component and a temporal fluctuating component , the time traces of the kth-emitter. Ideal imaging fully exploits to generate a super-resolved image.

Using time – or more specifically, the temporal fluctuations of emitters – as a way to achieve super-resolution can be performed by employing ***higher-order statistics***: the photon traces per pixel are analysed based on moments or cumulants . The 1st and 2nd order moments are widely used in engineering and are known as the mean and variance of a stochastic process.

The 4th and all higher order moments of two statistically independent stochastic processes does not equal the sum of the moments of the individual processes meaning moments are not additive:

for

In other words, the nth order moment of would result mainly from crosstalks between the two fluorophores. This principle can be further extended resulting in many more cross-talks as the number of fluorophore in is increasing.

An elegant way to resolve this issue is to use ***cumulants*** rather than *moments.* A rigorous analytical description of cumulants is beyond the scope of this tutorial and we will thus only list a few of its key properties and explain how they enable to provide background-free super-resolution.

Defintion: The nth-order cumulant of random variables is defined as the coefficient of in the Taylor series expansion of the cumulant generating function:

To compare, the moment generating function is the following:

Therefore, the nth-order cumulants can be inferred from its joint moments of orders up to n as follows:

Where denotes summation over all partitions of set *I*. See *Table A1 from Mendel et al (1991).*

Property 1: If , are constants, and , are random variables, then

Property 2: If is a constant, then

Property 3: If the random variable are indepedent of the random variables then

As explained previsouly, cross-terms (cross-talks between many fluorophores) are present in moments and absent in cumulants. This difference is formulated by property 3 which is exclusive to cumulants (can be easily demonstrated using « Definition »).

Returning to the equation of , the intensity distribution measured over time on a detector pixel of a sample composed of M independetly fluctuating emitters can be written as:

The nth order cumulants is described by:

contains information about the brightness , location and induced point-spread function of the fluorophore . It is straightforward to notice that the position of the fluorophore, as described in , is unchanged and thus does not affect the true location of fluorophores. We can also notice that the point-spread function has been replaced by . If we approximate the point-spread function by a Gaussian function, we obtain:

In that case, the point-spread function of becomes thinner by a factor and as consequence, there is a resolution improvement of a factor in as compared to . In addition, due to Property 2, the background was in theory removed. Finally, since is neither necessary for resolution improvement or background suppression, we arbitrarily compute for an impressively fast computation (3-4 seconds for ). Finally, since cumulants are by nature blind to random Gaussian processes, a wide range of time-varying noise are significantly reduced in the super-resolved image as compared to the original diffraction-limited image. In other words, if {x(t)} is Gaussian then the cumulants of {x(t)} are all zero.

The cumulants are presently computed for each pixel detector individually to provide a background-free and noise-reduced super-resolved image. Nevertheless, it is also possible to combine the information of many pixels to even further reduce any time-varying noise and determine inter-pixels values by computing ***cross-cumulants*** . In this case, the inter-pixel positions are defined by .

In the case of 2nd order statistics, the cross-cumulants can be written as

To compare:

assuming a Gaussian point-spread function. As shown in the equation above, can be easily computed since it equals the temporal cross-correlation of with . Since time-varying noise, including shot noise, is spatially and temporally uncorrelated (noise varies randomly with time and position), the temporal cross-correlation function of noise tends to zero. In addition, the thinner point spread function is now centered at instead of , and thus a virtual pixel between and is computed.

We can further push the reasoning as such: the pixel at location can also be computed with two other neighbouring pixels (i.e. in vertical direction rather than horizontal) . We could therefore calculate the pixel value of in two different ways to then average them leading to a reduced amount of noise in the new pixel value. We can extend the reasoning to nth order cross-cumulants where we would need *n* pixels for the computation and thus have a much higher number of ways of computing the cumulant pixel andaveraging some together as described in Dertinger et al. 2009 to even further reduce noise. .

The resulting nth order cross-cumulant can thus be described by the following equation:

Figure 1 shows the comparison of with a spline interpolation of . Fluorophores are seen “sharper” (due to a thinner point spread function) on than on . In addition, in , noise has been completely removed from . Figure 2 and 3 helps understanding why only considering and not or is appropriate. Indeed, since there are no practical gain in reducing noise and improving the resolution, computing only is sufficient. Only the first few coefficients provide sufficient information: rapidly decays toward zero as increases.