How the sensitivity of TOF-PET depends on the interplay between the temporal and spatial detector resolutions and the resolution required for the imaging task.

Johan Nuyts*, Michel Defrise, Christian Morel, Paul Lecoq

December 17, 2025

Abstract

Objective. Sensitivity is a key feature of positron emission tomography (PET). Here, sensitivity can be defined as the reciprocal of the amount of injected radioactivity needed to obtain a sufficient image quality for a particular scan duration. PET sensitivity can not only be increased by increasing the solid angle covered by the detectors, but also by improving their spatial and the temporal resolution, which in turn determines the time-of-flight (TOF) resolution. This paper analyzes how the interplay between the spatial detector resolution, the TOF resolution and the required imaging resolution affects the sensitivity of a TOF-PET system.

Approach. Two approaches are studied. The first computes the performance of the Hotelling observer for discriminating a small hot spot from a less small and less hot spot with the same total activity. This approach is flexible, and elegant closed form equations are obtained. In the second approach, analytical equations are derived for the variance of the reconstructed voxel values as a function of the TOF and spatial detector resolutions and of the reconstruction point spread function. To keep the mathematics tractable, the derivations are done for the center of a uniform sphere or cylinder, which is placed in the center of a cylindrical or spherical PET system. The results are verified with simulation experiments for 2D PET and 3D PET with septa.

Main results. Both approaches confirm that the sensitivity of the PET system increases, when the spatial detector resolution is improved, in agreement with simulation results published by Muehllehner, 1985. The same is true for improvements of the TOF resolution. Remarkably, when the TOF resolution (converted to a distance) approaches the spatial resolution, further improving it increases the sensitivity more than expected based on experience with current TOF-PET systems and on the analysis for moderate TOF resolution (Tomitani, 1981). This agrees with simulation results published recently by Toussaint et al., 2020. Finally, our new equations confirm that the value of a TOF-kernel is well characterized by the integral of its square, as was reported previously (Nuyts et al., 2022; Nuyts et al., 2023).

Significance. This analysis explains how spatial detector resolution and time-of-flight accuracy contribute to the sensitivity of PET systems, and predicts that pushing the

JN is with KU Leuven, University of Leuven, Department of Imaging and Pathology, Nuclear Medicine & Molecular imaging; Medical Imaging Research Center (MIRC), B-3000, Leuven, Belgium (e-mail: johan.nuyts@kuleuven.be). MD is with Department of Nuclear Medicine, Vrije Universiteit Brussel, B-1090, Brussels, Belgium (e-mail: michel.defrise@vub.be). CM is with Aix-Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France, (e-mail: morel@cppm.in2p3.fr). PL is with Polytechnic University of Valencia, Spain (e-mail: p.lecoq@cern.ch).

* Corresponding author.

time-of-flight resolution well below 100 ps will produce a larger benefit than expected based on current rules of thumb.

1 Introduction

In positron emission tomography (PET), sensitivity is very important, because the injected activities should be as low as reasonably achievable. The sensitivity of a PET system is often defined as the fraction of the emitted photon pairs that it can detect. Since the data are subject to Poisson noise, a higher sensitivity results in a higher signal-to-noise ratio (SNR). However, the SNR can also be increased by increasing the amount of information provided by each measured photon pair. In principle, this could be achieved in many ways, e.g. by increasing the energy resolution (and thereby reducing the disturbing Compton scatter contribution), by a better rejection of random coincidences, by improving the time-of-flight (TOF) resolution and/or the spatial resolution of the detectors, by measuring not only the position and the energy, but also the traveling direction of the photons, etc. To include such effects, a more general definition of sensitivity is used, often referred to as the "effective sensitivity". Budinger noted that in PET, the square of the SNR is proportional to the number of events (Budinger, 1983). Accordingly, he defined the *relative* or *effective sensitivity* as the square of the SNR. Consequently, a gain in effective sensitivity can be computed as the reduction of the total amount of activity required to obtain the same variance in the reconstructed image. In this paper, we adopt this definition and use it to derive equations quantifying its dependence on the spatial resolution and the temporal resolution, or more precisely the TOF resolution (also called the coincidence time resolution or CTR) of the PET system.

It is well-known that TOF contributes to the effective sensitivity (Conti et al., 2013; Schaefferkoetter et al., 2013). Tomitani quantified the benefit obtained from TOF by computing analytically the variance in the center of the image reconstructed from a 2D (TOF-)PET scan of a uniform cylinder (Tomitani, 1981). For the gain provided by TOF over non-TOF, he obtained the often cited equation

$$\frac{\mathrm{Var}_{\mathrm{nonTOF}}}{\mathrm{Var}_{\mathrm{TOF}}} = \sqrt{\frac{2\ln 2}{\pi}}\frac{D}{F_t} \simeq 0.664\frac{D}{F_t} \tag{1}$$

where D is the diameter of the uniform cylinder and F_t is the full width at half maximum of the Gaussian TOF point spread function (TOF-kernel), expressed as a length. A similar expression was given in (Budinger, 1983), but without the factor 0.664, making it overly optimistic[1]. An extension of the expression to account for random coincidences was given in (Conti, 2005), and a slightly different one in (Nuyts et al., 2022). Tomitani's result was extended in (Nuyts et al., 2022) to PET systems which are characterized by a distribution of possibly non-Gaussian TOF-kernels, and in (Nuyts et al., 2023) the benefit of TOF for joint reconstruction of the activity and the attenuation was computed. In these papers (Nuyts et al., 2022; Nuyts et al., 2023; Tomitani, 1981), it was assumed that the PET system had perfect (spatial) detector resolution, but the image was reconstructed with a finite spatial resolution to ensure a finite variance in the reconstructed image. In addition, approximations

[1]The factor $\sqrt{2\ln 2/\pi}$ results from converting the standard deviation to FWHM and from computing the integral of a squared Gaussian.

were introduced, based on the assumption that the TOF resolution was poor compared to the spatial resolution of the reconstructed image.
In (Nuyts et al., 2022; Nuyts et al., 2023), the benefit from TOF was also computed with another approach. A simple detection task in a signal-known exactly, background-known exactly (SKE-BKE) setting was designed. The detection was done by the optimal linear discriminant, often referred to as the Hotelling observer (Barrett & Myers, 2013; Barrett et al., 1993), and the effective sensitivity was quantified as the squared SNR of that observer. In both papers, we found that the observer SNR^2 agreed closely with the reciprocal of the variance in the reconstructed image.
The spatial resolution of PET detectors contributes to the effective sensitivity too. Gerd Muehllehner (Muehllehner, 1985) showed with computer simulations that when the PET system has a better spatial resolution, fewer counts have to be acquired to obtain the same visual image quality. In (Surti et al., 2013), Suleman Surti *et al.* examined the effects of improving the TOF resolution and the spatial resolution with simulations; it was found that improving the TOF resolution from 600 ps to 300 ps had a similar effect on lesion detection and localization as reducing the detector crystal size from 4 mm to 2.6 mm. In this paper, we will refer to these findings as the *Muehllehner effect*.
Maxime Toussaint *et al.* (Toussaint et al., 2020) observed in their simulation experiments, that for "ultrafast TOF", the TOF resolution improves the achievable spatial resolution in the reconstructed images. Remarkably, this improvement already sets in well before the TOF resolution becomes comparable to the detector resolution[2]. This surprised us at first, as we intuitively expected the resolution of the reconstructed image to be limited by either the detector resolution or the TOF resolution, depending on which was better. We will refer to this finding as the *Toussaint effect*.
The aim of this paper is to provide theoretical support for these findings and quantify the effects of the TOF resolution, the detector resolution and the spatial resolution required for the imaging task on the effective sensitivity of the PET system. Therefore, in these new derivations, approximations relying on a poor TOF resolution are avoided. Similar to (Nuyts et al., 2022; Nuyts et al., 2023), two complementary methods will be used to quantify the effective sensitivity: (1) the SNR of the Hotelling observer in a SKE-BKE discrimination experiment, and (2) the variance in the center of an image reconstructed from the PET scan of a uniform phantom.

2 Theory: the SNR^2 for discriminating different small blobs

Two approaches for quantifying the effective sensitivity of a PET system are proposed. The first approach computes the effective sensitivity as the performance of the Hotelling observer, for discriminating a small hot blob from a larger less hot blob that has the same total activity. This is done assuming that the signal and background are known exactly (SKE-BKE), the only source of uncertainty is the noise. Although this is a strong simplification, this task has still some similarity to the common clinical task of assessing the severity of the increased tracer uptake in a small lesion. This task requires a good resolution, because poor resolution would blur away the difference between the two blobs. As shown below, with well-designed blobs, elegant closed form equations are obtained for the observer SNR.

[2]In this paper, "detector resolution" refers to the spatial resolution of the detector, the resolution of the measured difference in photon arrival times (converted to a distance) is denoted as "TOF resolution".

The second approach computes the variance in the center of an image reconstructed from the PET scan of a uniform phantom. It is required that the reconstruction produces a predefined point spread function (PSF) in the final image. That ensures that the signal of the imaging process is always the same, such that the variance is inversely proportional to the square of the SNR. We obtain expressions for the reconstruction variance, as integrals which have to be evaluated numerically for evaluating a particular PET system.
For both approaches, we assume that all considered PET systems scan the same phantom with the same scan duration. A PET system is considered superior for an imaging task if it has a higher effective sensitivity, either quantified as the SNR^2 of the Hotelling observer or as the reciprocal of the variance in the reconstructed image. We anticipated that a PET system scoring better for the discrimination task, will also score better for the reconstruction task.

2.1 Discriminating a small hot blob from a larger less hot one

2.1.1 The Hotelling observer

Consider a measurement system, e.g. a pair of PET detector blocks in coincidence, that produces a measurement $m = \bar{m} + n$ when given an input λ, where m is an array with measurement values, n an array with corresponding measurement noise and $\bar{m}$ is the expectation of the measurement when the input equals λ.
In an SKE-BKE discrimination task, there are only two possible inputs λ_1 and λ_2. The observer knows the corresponding measurement expectations $\bar{m}_1$ and $\bar{m}_2$ and the noise covariance matrix V, which is assumed the same for both inputs. Based on a noisy measurement m, the observer must decide if the input λ was either λ_1 or λ_2. The Hotelling observer does that, by computing the probability of observing m for both possible inputs and choosing the one that produces the highest probability:

$$p_1 = p(m|\lambda_1), \quad p_2 = p(m|\lambda_2), \quad \text{if } t = \ln(p_2) - \ln(p_1) \geq 0 \text{ then } \lambda = \lambda_2 \text{ else } \lambda = \lambda_1, \tag{2}$$

where t is the logarithm of the likelihood ratio p_2/p_1. For data corrupted by multivariate Gaussian noise with covariance V, one finds that (Barrett & Myers, 2013; Barrett et al., 1993)

$$t = (\bar{m}_2 - \bar{m}_1)^T V^{-1} m \ + c \tag{3}$$

where c is a constant independent of the data m. If the difference $\bar{m}_2 - \bar{m}_1$ is very small compared to $\bar{m} = \bar{m}_1 \simeq \bar{m}_2$ then (3) also holds for Poisson noise, because then the data covariance $V = \text{diag}(\bar{m})$ and:

$$\begin{aligned} t &= \sum_i m_i \ln(\bar{m}_{2i}) - \bar{m}_{2i} - (m_i \ln(\bar{m}_{1i}) - \bar{m}_{1i}) = \sum_i m_i \ln\left(1 + \frac{\bar{m}_{2i} - \bar{m}_{1i}}{\bar{m}_{1i}}\right) - (\bar{m}_{2i} - \bar{m}_{1i}) \\ &\simeq \sum_i \frac{m_i(\bar{m}_{2i} - \bar{m}_{1i})}{\bar{m}_i} - (\bar{m}_{2i} - \bar{m}_{1i}) \quad = \quad (\bar{m}_2 - \bar{m}_1)^T V^{-1} m - \sum_i (\bar{m}_{2i} - \bar{m}_{1i}) \end{aligned} \tag{4}$$

This is equivalent to (3) because $\bar{m}_2 - \bar{m}_1$ is independent of the measurement m.
The expected change of t caused by changing the input from λ_1 to λ_2 equals

$$E(\Delta t) = (\bar{m}_2 - \bar{m}_1)^T V^{-1} (\bar{m}_2 - \bar{m}_1). \tag{5}$$

The variance on t equals

$$\mathrm{Var}(t) = E\big((\bar{m}_2 - \bar{m}_1)^T V^{-1}(m - \bar{m})(m - \bar{m})^T V^{-1}(\bar{m}_2 - \bar{m}_1)\big) = (\bar{m}_2 - \bar{m}_1)^T V^{-1}(\bar{m}_2 - \bar{m}_1). \tag{6}$$

Consequently, the squared SNR of the observer decision equals

$$\mathrm{SNR}^2 = (E(\Delta t))^2/\mathrm{Var}(t) = (\bar{m}_2 - \bar{m}_1)^T V^{-1}(\bar{m}_2 - \bar{m}_1). \tag{7}$$

In (Nuyts et al., 2022) it was shown that the observer performance can be directly computed from the PET data, producing the same SNR as would be obtained from optimal discrimination in the image domain. With this approach, the SNR can be computed for a single detector pair. The SNR for multiple detector pairs is obtained as a simple weighted summation of the squared SNRs of all pairs. This works even when each detector pair is associated with a different (arbitrarily shaped) TOF-kernel (Nuyts et al., 2022).
In this paper, the derivations will be done assuming continuous images and measurements. In TOF-PET, the measurement m produced by two detector blocks in coincidence is a 3D image, and equation (7) becomes:

$$\mathrm{SNR}^2 = \int \frac{(\bar{m}_2(\vec{s}) - \bar{m}_1(\vec{s}))^2}{\bar{m}(\vec{s})} d\vec{s} = \int \frac{((\mathrm{PSF_A} \otimes (\lambda_2 - \lambda_1))(\vec{s}))^2}{(\mathrm{PSF_A} \otimes \lambda)(\vec{s})} d\vec{s} \tag{8}$$

with $\lambda = \lambda_1 \simeq \lambda_2$, and assuming that the measurement operator A can be well approximated as a convolution with a point spread function, which was denoted as $\mathrm{PSF_A}$.

2.1.2 Two blobs to be discriminated

In (Nuyts et al., 2022), the SNR gain obtained with TOF was analyzed with a SKE-BKE task, where λ_1 corresponded to a cylindrical phantom filled with a uniform activity concentration B and λ_2 differed from λ_1 only by a slight activity increase in the center of that phantom. However, that paper studied only poor to moderate TOF resolution and ignored the effects of the detector resolution. In contrast, here we wish to study what happens when the TOF resolution approaches the detector resolution. Most recent clinical PET systems have a detector resolution of 3 to 4 mm. To achieve a similar spatial resolution in the TOF direction, a TOF timing resolution of 20 to 30 ps would be needed [3]. This is well below the current state-of-the-art of commercial systems, which is around 200 ps, but in the lab, sub-50 ps coincidence timing resolutions have already been obtained (Kwon et al., 2021; Lecoq et al., 2020; Ota et al., 2019).
To analyze such future TOF-PET systems, a task requiring good spatial resolution is needed. In a first attempt, we studied the discrimination of a small and hot Gaussian blob from a slightly larger but less hot Gaussian blob with the same activity. Since both blobs have the same total activity, they cannot be discriminated by simply analyzing the total number of measured counts; resolution is needed for this discrimination, as desired. We obtained a closed form expression for the detection SNR as intended. Unfortunately, this expression provided little insight: it computed essentially the small difference between two large terms. We expected more insight from a single term describing directly that small difference. For that reason, we considered a limit case: the derivative of a Gaussian blob with respect to

[3] A TOF resolution of 10 ps FWHM corresponds to a spatial resolution of 1.5 mm FWHM.

its standard deviation. Defining $G_3(\vec{r}, \sigma_o)$ as a 3D Gaussian with standard deviation σ_o one finds:

$$\sigma_o^4 \frac{d}{d\sigma_o} G_3(\vec{r}, \sigma_o) = \sigma_o^4 \left(\frac{|\vec{r}|^2}{\sigma_o^3} - \frac{3}{\sigma_o} \right) G_3(\vec{r}, \sigma_o) \quad = \mathrm{SLoG}(\vec{r}, \sigma_o), \quad \vec{r} \in \mathbb{R}^3 \tag{9}$$

The scale factor σ_o^4 ensures that the maximum of this distribution is independent of its size σ_o; the reason for this choice will be explained later. Interestingly, this function is equal to the Laplacian of Gaussian function, except for a constant factor:

$$\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} \right) G_3(\vec{r}, \sigma_o) = \frac{1}{|\vec{r}|} \frac{d^2}{d|\vec{r}|^2} (|\vec{r}| G_3(\vec{r}, \sigma_o)) = \frac{1}{\sigma_o} \frac{d}{d\sigma_o} G_3(\vec{r}, \sigma_o) \tag{10}$$

which is why we named this function "SLoG", Scaled Laplacian of Gaussian, in (9). Equation (9) shows that the SLoG can be considered as the difference between two strictly positive blobs: a small hot Gaussian blob ($3\sigma_o^3 G_3(\vec{r}, \sigma_o)$) and a wider less hot blob with a cold interior ($|\vec{r}|^2 \sigma_o G_3(\vec{r}, \sigma_o)$). A central slice through both blobs is shown in figure 1. Both blobs have the same integral (equal to $3\sigma_o^3$), since the integral of the SLoG is zero.

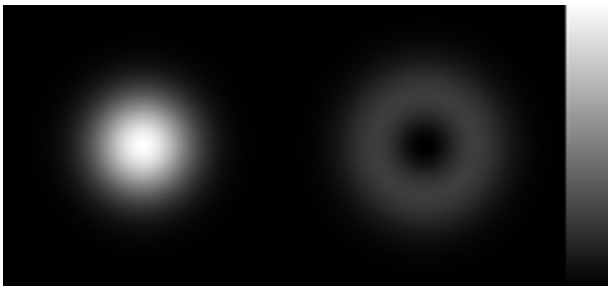

Figure 1: *The* SLoG *can be considered as the difference between a hot Gaussian blob, and a larger but less hot blob that has a cold interior. Both blobs have the same total activity. The figure shows a central slice through both blobs (at the left the hot Gaussian, at the right the less hot hollow blob).*

For the discrimination task, we assume that λ_1 and λ_2 have an identical uniform background activity concentration B. Superimposed on that background, in the center of the phantom, is the larger hollow blob in λ_1, and the hotter Gaussian blob in λ_2. Consequently, λ_1 and λ_2 have the same total activity. This task corresponds to the clinical task of discriminating a small hot lesion from a less hot and larger lesion. Alternatively, the very same task can be thought of as a the detection of the presence or absence of the SLoG (i.e. the difference of the two blobs). In this view, λ_1 and λ_2 have again the same uniform background activity. In the center of λ_2 (but not in λ_1), the SLoG is superimposed on that background. Because the SLoG activity is assumed very small compared to the background activity, the activity in λ_2 is still strictly positive. This task corresponds to identifying the presence or absence of a small heterogeneous structure.
The size of the SLoG in (9) is determined by σ_o; we will refer to σ_o as the size of the SLoG and use it as a measure of the spatial resolution required for the detection task. Note that σ_o is not the standard deviation of the SLoG, it is the standard deviation of the hot Gaussian blob of the SLoG (fig. 1).

2.2 Equations for the squared SNR when the TOF-kernel is Gaussian

Two opposing TOF-PET detector blocks form a so-called tube of response along which radioactivity is detected. We assume that both detector blocks are positioned with their center on the x-axis, and with the front faces of their crystal(s) aligned with the y- and z-axes. We consider only the lines of response (LOR) parallel to the x-axis that are seen by this detector pair. The TOF-PET measurement has a finite spatial resolution (Lecomte, 2009; Maas et al., 2009; Moses, 2011). For a set of parallel LORs, the TOF-measurement and its spatial uncertainty are three dimensional. We assume that this 3D spatial resolution can be well approximated with a 3D Gaussian distribution. In addition, we assume that it can be separated into a 1D Gaussian with standard deviation σ_t along the x-axis, which we will denote as the TOF resolution, and two 1D Gaussians along the y- and z-axes with standard deviations σ_y and σ_z, which we will denote as the detector resolutions. The TOF resolution σ_t incorporates any uncertainty along the LOR, including the x-component of the positron range and of the depth of interaction uncertainty. Likewise, the Gaussians representing the 2D detector resolutions σ_y and σ_z incorporate the blurring due to the acollinearity [4], the positron range perpendicular to the LOR and, for oblique incidence angles, a component of the depth of interaction uncertainty. Although the real 3D resolution is usually not exactly (multivariate) Gaussian and not separable into three 1D distributions, this model is a reasonable approximation for most systems, in particular near the center of a cylindrical PET system.

Consequently, the TOF-PET system outputs for each pair of detector blocks a 3D blurred and noisy image of the radioactivity in that tube. This image has a resolution of σ_t, σ_y and σ_z along the x-, y- and z-axes. This image is the measurement $m(x, y, z)$, acquired along LORs parallel to the x-axis by the detector pair.

In appendix 9.A this image has been inserted in (8) to compute the SNR^2 for detecting a SLoG blob of size σ_o with TOF-PET detectors having a Gaussian TOF resolution σ_t and detector resolutions σ_y and σ_z. We assumed that the activity in the SLoG was scaled with S, and that the background activity concentration was equal to B. In addition, we introduced a factor η, representing the detection sensitivity of the detector pair due to finite stopping power, attenuation, solid angle associated with the LORs and so on. This produces the following equation for the SNR^2:

$$\mathrm{SNR}^2 = \frac{\eta}{16\pi\sqrt{\pi}} \frac{(S\sigma_o^4)^2}{B} \frac{\sigma_o^2}{\sqrt{(\sigma_t^2+\sigma_o^2)(\sigma_y^2+\sigma_o^2)(\sigma_z^2+\sigma_o^2)}} \left(\frac{1}{2} \left(\frac{1}{\sigma_t^2+\sigma_o^2} + \frac{1}{\sigma_y^2+\sigma_o^2} + \frac{1}{\sigma_z^2+\sigma_o^2} \right)^2 + \frac{1}{(\sigma_t^2+\sigma_o^2)^2} + \frac{1}{(\sigma_y^2+\sigma_o^2)^2} + \frac{1}{(\sigma_z^2+\sigma_o^2)^2} \right) \tag{11}$$

If the TOF resolution is poor compared to the detector resolution, assumed here the same along y and z for convenience, i.e. $\sigma_t \gg \sigma_y = \sigma_z$, then (11) can be well approximated as

$$\mathrm{SNR}^2\big|_{\sigma_t \gg \sigma_y = \sigma_z} \simeq \frac{\eta}{4\pi\sqrt{\pi}} \frac{(S\sigma_o^4)^2}{B} \frac{1}{\sigma_t} \frac{\sigma_o^2}{(\sigma_y^2+\sigma_o^2)^3} \tag{12}$$

[4] The locus of all emission points producing the same TOF difference in two detectors is a hyperbola. Consequently, the acollinearity uncertainty produces a blurring along this hyperbola. At their intersection point, the hyperbola is perpendicular to the LOR between the two detectors. Consequently, since the acollinearity uncertainty is small (Lecomte, 2009), it can be treated as a blurring perpendicular to the LOR.

Extension of equation (11) to account for random coincidences is straightforward, if we can assume that the randoms contribution is uniform in the region of interest (as is almost always the case). Denoting the number of random coincidences per unit measurement volume by R, it suffices to simply replace B with $B + R$ in (11). More information about accounting for the presence of random coincidences and scattered photons is given in appendix 9.B.
Equation (11) summarizes the main results of this paper. It shows that the three resolutions σ_t, σ_y and σ_z play exactly the same role. The equation has two factors. Each resolution contributes multiplicatively to the first factor, and additively to the second. If the TOF resolution is poor compared to the detector resolutions ($\sigma_t \gg \sigma_y \simeq \sigma_z$), then improving TOF contributes an improvement of the SNR^2 through the first factor, but it contributes negligibly to the second factor, leading to approximation (12). This $1/\sigma_t$ dependence agrees with Tomitani's result (1). Equations (11) and (12) show that an improvement of the detector resolution produces a very strong improvement of the SNR, confirming the Muehllehner effect (Muehllehner, 1985). Only when the detector blurring (σ_y and σ_z) becomes very small compared to the SLoG-blob size σ_o, further improvements to the detector resolution start having a smaller effect on the SNR. When improvements to the TOF resolution decrease σ_t to the order of σ_y and σ_z, TOF starts to contribute significantly to the second factor of (11), confirming the Toussaint effect (Toussaint et al., 2020).
Figure 2 compares 1/SNR^2 computed with the accurate equation (11) to that computed with the approximation (12). The plots confirm the Toussaint effect: the gain from TOF becomes higher than predicted by Tomitani's equation or (12) when σ_t starts approaching the detector resolution. The effect is stronger when the detector blurring (σ_y and σ_z) is large compared to the SLoG blob size σ_o.

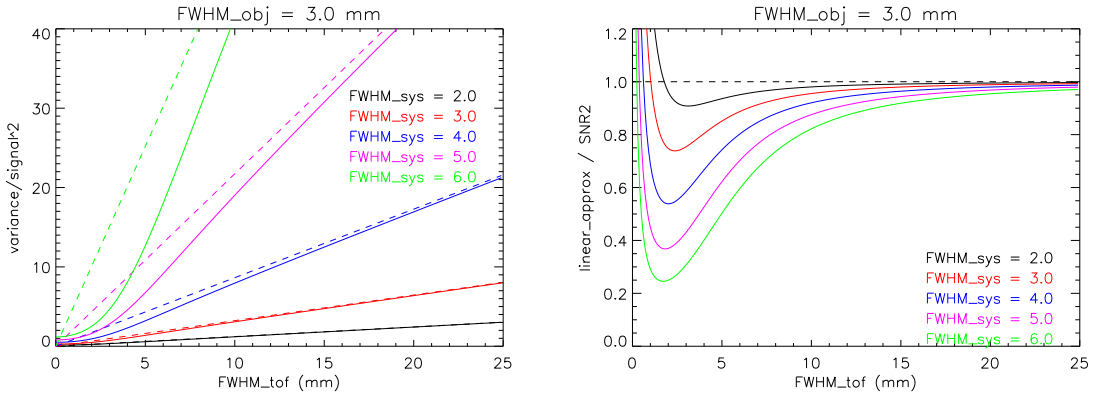

Figure 2: *Left: 1/SNR^2 as a function of the TOF resolution F_t ranging from 0 to 25 mm, for F_o = 3 mm and FWHM_sys = $F_y = F_z \in \{2, 3, 4, 5, 6\}$ mm. Note that F_t is the TOF resolution expressed as FWHM in mm, so $F_t = \sqrt{8 \ln 2}\, \sigma_t$, and similar for F_o, F_y, F_z. The solid lines are the accurate values computed with (11), the dashed lines show the approximation with (12). Right: the ratio between the accurate value (11) and its approximation (12). The finding that the accurate SNR^2 values are (much) better than those based on the linear approximation confirms the Toussaint effect.*

Figure 3 illustrates the Muehllehner effect revealed by equation (12): improving the detector resolution produces an improvement of the effective sensitivity or SNR^2. This improvement is very strong when the detector resolution becomes better than the SLoG blob size.

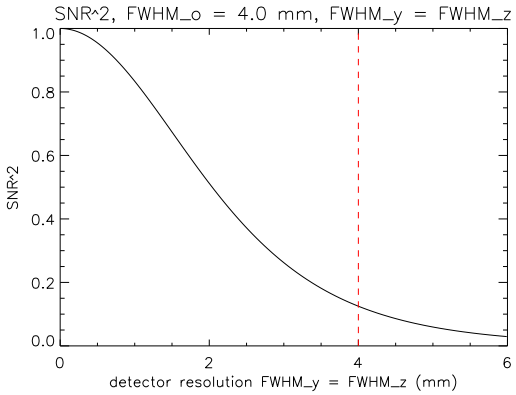

Figure 3: *Plot of the approximate equation for moderate TOF (12) as a function of detector resolution $F_y = F_z$ in mm, normalized to its maximum value. The red dashed line indicates the value of the SLoG blob size $F_o = 4$ mm. When $F_y = F_z \leq F_o$, the SNR^2 improves strongly with improving detector resolution, confirming the Muehllehner effect.*

2.3 Using FWHM instead of standard deviations.

For the derivations it is convenient to work with standard deviations $\sigma_o, \sigma_t, \sigma_y$ and σ_z. However, in the PET community it is common practice to use the full width at half maximum (FWHM) instead. Therefore, we use standard deviations in the derivations, but represent the resolutions with their FWHM in mm when plotting results.[5]
For a Gaussian, the FWHM F equals $F = \sqrt{8 \ln 2}\,\sigma$. Consequently, the equations can be easily rewritten using FWHM instead of σ. E.g., (12) is equivalent with

$$\mathrm{SNR}^2\big|_{F_t \gg F_y = F_z} \simeq \frac{1}{(\sqrt{8 \ln 2})^3} \frac{\eta}{4\pi\sqrt{\pi}} \frac{(S\, F_o^4)^2}{B} \frac{1}{F_t} \frac{F_o^2}{(F_y^2 + F_o^2)^3} \tag{13}$$

where F_o, F_t, F_y, F_z denote the FWHM of the SLoG's hot blob, the TOF resolution and the detector resolutions. Thus, to compute the SNR for resolutions in FWHM, one can use the expressions for standard deviations and simply divide the result by $(\sqrt{8 \ln 2})^3$, a constant independent of the resolutions. The equations contain other constants which can be chosen freely: the activity S, the background concentration B, and the value of η, accounting for the solid angle and stopping power of the detectors, the scan duration etc. When comparing PET systems with different TOF and detector resolutions, those arbitrary constants don't matter: as long as they are kept the same for all systems, they don't affect the ranking.

2.4 Extension of (11) to the case where the TOF-kernel is not small compared to the object size.

In the derivation of (11) it was assumed that the diameter of the background cylinder was large compared to the width of the TOF-kernel. Appendix 9.B derives an equation avoiding this assumption. Interestingly, this equation (53) is also valid for a non-uniform background activity.

[5] We express the TOF resolution everywhere in mm; divide by 0.15 mm/ps to convert it to ps.

This equation is more complicated than (11) and requires the numerical integration of three integrals. Fortunately, all three only involve the integration of 1D smooth functions, which can be done in a few ms. Figure 4 shows a result for $F_o = 4$ mm, $F_y = F_z = 4$ mm, an object diameter of 100 mm and for TOF resolutions ranging from 0 to 100 mm. When the FWHM of the Gaussian TOF-kernel is about 2.5 times smaller than the background diameter, the two models agree very well. The model that considers the diameter converges to the non-TOF SNR^2, which is indicated by the dashed blue line.

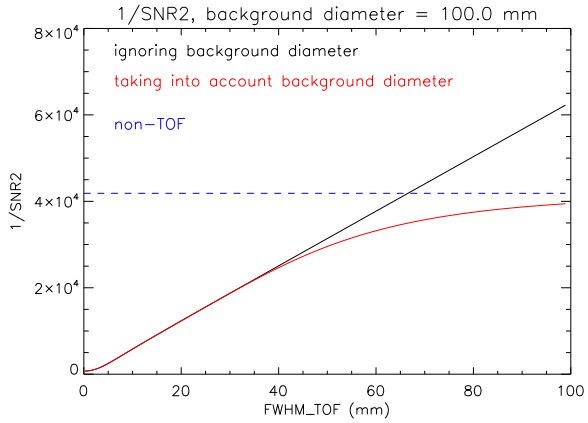

Figure 4: *Comparison of 1/SNR^2 computed while ignoring (black) or accounting for (red) the finite diameter of the uniform background object (100 mm). The dashed blue line is the 1/SNR^2 for non-TOF.*

2.5 Approximation for a wide TOF-kernel of arbitrary shape

Using (8), the performance of the Hotelling observer for detecting the SLoG is computed as

$$\text{SNR}^2 = \frac{\eta S^2}{B} \iiint d\hat{x}\, d\hat{y}\, d\hat{z} \left(\iiint k(x) G_1(y, \sigma_y) G_1(z, \sigma_z)\, \text{SLoG}(\hat{x} - x, \hat{y} - y, \hat{z} - z) dx\, dy\, dz \right)^2 \tag{14}$$

where $k(x)$ is a TOF-kernel of arbitrary shape, producing resolution loss along the x-axis, and $G_1(y, \sigma_y)$ and $G_1(z, \sigma_z)$ represent the Gaussian detector resolutions along the y- and z-axes. If the TOF-kernel $k(x)$ is wide compared to the size of the SLoG blob, then we can ignore its variation over the range where SLoG is different from zero and move it out of the integral, producing

$$\text{SNR}^2 = \frac{\eta S^2}{B} \iiint d\hat{x}\, d\hat{y}\, d\hat{z}\; k^2(\hat{x}) \left(\iiint G_1(y, \sigma_y) G_1(z, \sigma_z)\, \text{SLoG}(x, \hat{y} - y, \hat{z} - z) dx\, dy\, dz \right)^2 \tag{15}$$

where the function between brackets is independent of the TOF-kernel k and of $\hat{x}$; the first argument of SLoG could be replaced by x, because the integral is over $x = -\infty$ to ∞. This shows that these new equations still support the notion that the value of a (wide) TOF-kernel is well characterized by the integral of its square, as previously reported in (Nuyts et al., 2022; Nuyts et al., 2023) and validated in (Merlin & Visvikis, 2022). In these papers, the equivalent Gaussian TOF-kernel was computed, by finding the Gaussian kernel that has the same integral of its square as the non-Gaussian one.

2.6 Extension of (11) for a TOF-kernel consisting of a sum of Gaussians

Recent detector designs push the TOF resolution by exploiting the Cherenkov photons in BGO or LYSO (Efthimiou et al., 2020; Gundacker et al., 2019; Maebe & Vandenberghe, 2023) or by combining fast and dense scintillators into a single detector (Konstantinou et al., 2021; Shultzman et al., 2024; Turtos et al., 2019). These detectors often produce detection events with an event-dependent TOF resolution, resulting in a non-Gaussian TOF-kernel. TOF-kernels are typically smooth and always non-negative, and as a result can be well approximated as a sum of a few Gaussians, possibly shifted relative to each other. Such a TOF-kernel can be written as

$$k(x) = \frac{1}{\sum_{i=1}^{n} A_i} \sum_{i=1}^{n} A_i \, G_1(x, \sigma_i, x_i) \tag{16}$$

$$G_1(x, \sigma, x_1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_1)^2}{2\sigma^2}} \tag{17}$$

where n is the number of Gaussians, and A_i, σ_i and x_i denote the weight, the standard deviation and the shift of the i-th Gaussian. The SNR^2 for this TOF-kernel is derived in appendix 9.C, producing the following result:

$$\Phi(v_o, v_t, v_y, v_z) = v_o \frac{\frac{1}{2}\left(\frac{1}{v_t+v_o} + \frac{1}{v_y+v_o} + \frac{1}{v_z+v_o}\right)^2 + \frac{1}{(v_t+v_o)^2} + \frac{1}{(v_y+v_o)^2} + \frac{1}{(v_z+v_o)^2}}{\sqrt{(v_t+v_o)(v_y+v_o)(v_z+v_o)}} \tag{18}$$

$$\Psi(v_o, v_t, v_y, v_z, \delta_x) = \frac{\delta_x^2 \, v_o}{\sqrt{v_t+v_o}\sqrt{v_y+v_o}\sqrt{v_z+v_o}} \left(\frac{\delta_x^2}{16(v_t+v_o)^4} - \frac{1}{2(v_t+v_o)^3} - \frac{1}{4(v_t+v_o)^2}\left(\frac{1}{v_t+v_o} + \frac{1}{v_y+v_o} + \frac{1}{v_z+v_o}\right)\right) \tag{19}$$

$$\mathrm{SNR}^2 = \frac{\eta \, S^2}{16\pi\sqrt{\pi}B(\sum_i A_i)^2} \left\{ \sum_{i=1}^{n} A_i^2 \Phi(\sigma_o^2, \sigma_i^2, \sigma_y^2, \sigma_z^2) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2A_i A_j e^{-\frac{(x_i-x_j)^2}{2(\sigma_i^2+\sigma_j^2+2\sigma_o^2)}} \left(2\Psi\left(\sigma_o^2, \frac{\sigma_i^2+\sigma_j^2}{2}, \sigma_y^2, \sigma_z^2, x_i - x_j\right) + \Phi\left(\sigma_o^2, \frac{\sigma_i^2+\sigma_j^2}{2}, \sigma_y^2, \sigma_z^2\right)\right) \right\} \tag{20}$$

If desired, the equivalent Gaussian TOF-kernel width can be obtained by finding (iteratively) the σ_t that produces the same SNR^2 with equation (11). Since this approach would account for the effect of the detector resolution, it would be more accurate for ultrafast TOF, than simply computing the equivalent Gaussian from the integral of the squared TOF-kernel. In a test with a sum of two Gaussians, the two approaches differed less than 10% as long as the standard deviation of the smallest of the two Gaussians was not smaller than σ_o.

3 Theory: the variance in the reconstructed PET image

This section starts with clarifying the relation between the image reconstruction variance and the effective sensitivity. Then, a general equation for the variance in the center of a reconstructed image is provided. Based on this equation, specific integral equations are derived for a spherical PET system, a multi-ring PET system with septa, a circular 2D PET system and a

fully 3D cylindrical PET system. The sections give the integral equations, details about their derivation and modifications to improve numerical stability are provided in the appendices. In the final subsection, predictions obtained from numerical evaluation of the integrals are shown.

3.1 Effective sensitivity and variance

To compare different PET systems, this approach assumes that all systems acquire a measurement of the same radioactive object with the same scan duration, and reconstruct an image from it with a predefined reconstruction point spread function (PSF). The PET system that achieves the lowest variance in its reconstructed image is best system. Because all systems have the same reconstruction PSF (and therefore the same signal), the SNR^2, and therefore the effective sensitivity, is proportional to the reciprocal of the variance in the reconstructed image:

$$\text{effective sensitivity} \sim \mathrm{SNR}^2 \sim \frac{1}{\mathrm{V}AR}. \tag{21}$$

The same approach was used in (Nuyts et al., 2022). Equation (21) only gives a proportionality; below also quantitative expressions for the variance will be derived, allowing us to compare the predictions to the results of simulations. Here, the standard deviation σ_{recon} of the Gaussian reconstruction PSF represents the required imaging resolution, playing a similar role as σ_o in the detection method explained above.

To simplify the problem, we only consider cylindrical or spherical PET systems, which acquire a scan of a uniform radioactive sphere (or cylinder) that was positioned exactly in the center of the field of view. In addition, we only consider the variance in the center of the reconstructed image. The assumption is that if a cylindrical PET system has a higher effective sensitivity for the center of a uniform sphere, it will most likely also have a higher effective sensitivity for other, non-uniform objects, such as the tracer distribution in a patient's body.

In the next section, a general equation for the variance in the center of the reconstructed image is derived. In the subsequent sections, this general equation will be applied to different PET systems, quantifying the effective sensitivity as the reciprocal of the variance in the center of the reconstructed image of the radioactive sphere.

Although the derived equations are based on analytical reconstruction, the numerical verification is done with maximum likelihood expectation maximization (MLEM). This makes sense, because MLEM can be well approximated with weighted least squares reconstruction. For the center of a uniform sphere, this reduces to unweighted least squares reconstruction, because all weights are equal. And the minimum variance analytical reconstruction derived below should also produce an unweighted least squares reconstruction.

3.2 Optimal variance of 3D TOF-PET activity reconstruction

In this section, an analytical expression is derived for the minimum variance that can be achieved with a reconstruction algorithm based on filtering and weighted backprojection. The derivation is based on two assumptions: the pixel variance is relatively smooth and it does not depend on the direction of the LOR. These assumptions make sense for the center of a uniform radioactive sphere.

We consider a 3D radioactive distribution $f(\vec{r})$, which is positioned in a PET scanner. We assume that this PET system has an (approximately) shift-invariant solid angle aperture

$\Omega \subset \mathbb{S}^2$ ($\mathbb{S}^2$ is the set of all unit vectors in $\mathbb{R}^3$), so that the measured PET data can be written as

$$p(\hat{u}, \vec{s}) = \int_{\mathbb{R}^3} f(\vec{r}) h(\hat{u}, \vec{s} - \vec{r}) d\vec{r}, \quad \hat{u} \in \Omega, \vec{s} \in \mathbb{R}^3 \tag{22}$$

where $h(\hat{u}, \vec{r})$ is the system response for a line of response (LOR) parallel to unit vector $\hat{u}$. Thus, $\vec{r} \cdot \hat{u}$ is the TOF variable. We define $\vec{r} \perp \hat{u}$ as the projection of $\vec{r}$ on the plane perpendicular to $\hat{u}$. The measurement at fixed $\hat{u}$ is modeled as a convolution of the image with system response $h(\hat{u}, \vec{r})$, which accounts for the TOF resolution (along $\hat{u}$) and the finite spatial resolution of the detector (perpendicular to $\hat{u}$). We treat these effects as separable, locally shift invariant and independent of the LOR direction $\hat{u}$:

$$h(\hat{u}, \vec{r}) = k(\vec{r} \cdot \hat{u})\ s(\vec{r} \perp \hat{u}) \tag{23}$$

where $k(x)$ is the 1D TOF-kernel along the LORs parallel to $\hat{u}$, and $s(\vec{r} \perp \hat{u})$ is the 2D blurring kernel in the plane perpendicular to the LOR. Note that for fixed $\hat{u}$, the measurement image $p(\hat{u}, \vec{s})$ has the same dimension as the reconstruction image $f(\vec{r})$; it corresponds to the TOF-PET histo-image of (Matej et al., 2009).

In appendix 10.A, the data noise is analytically propagated through an FBP reconstruction which imposes the reconstruction PSF $r(\vec{r})$, and which selects the reconstruction filter and backprojection weight to minimize the variance. The following equation for the variance $\mathrm{Var}(f_{\mathrm{recon}}(\vec{r}))$ in the reconstructed image was obtained:

$$\boxed{\frac{\mathrm{Var}(f_{\mathrm{recon}}(\vec{r}))}{\mathrm{Var}(p(\vec{r}))} \simeq \int_{\mathbb{R}^3} \frac{R^2(\vec{\nu})}{\int_\Omega |K(\vec{\nu} \cdot \hat{u})|^2 |S(\vec{\nu} \perp \hat{u})|^2 d\hat{u}}\, d\vec{\nu}} \tag{24}$$

where $K(\nu)$ is the 1D Fourier transform of $k(x)$ and $S(\vec{\nu} \perp \hat{u})$ is the 2D Fourier transform of $s(\vec{r} \perp \hat{u})$, Ω is the solid angle covered by the PET system, $\hat{u}$ is a unit vector (in the direction of the LOR) and $\vec{\nu} \perp \hat{u}$ is the projection of $\vec{\nu}$ on the plane orthogonal to $\hat{u}$. $R(\vec{\nu})$ is the 3D Fourier transform of $r(\vec{r})$, so it is the reconstruction MTF, and $\mathrm{Var}(p(\vec{r}))$ is the noise variance of measured sinogram values affected by activity near $\vec{r}$.

Since the derivation accounted for the finite spatial resolution of the PET system, equation (24) predicts the variance for image reconstruction with resolution recovery.

In the following sections, this equation will be used to compute the effective sensitivity for a spherical PET system with complete angular coverage, for a cylindrical PET system equipped with septa, for a circular 2D PET system operating in a 2D world, and finally for a cylindrical PET system with fully 3D sampling (no septa) and a finite axial length. For convenience, we will assume that the TOF-kernel, the detector resolution and the reconstruction PSF are all Gaussian, although equation (24) also holds (within the assumptions and approximations) for kernels with non-Gaussian shapes.

By requiring a Gaussian shape for all kernels, it becomes impossible to obtain a reconstruction PSF that is narrower than the measurement PSF. This is because deblurring a Gaussian to a narrower Gaussian in image space, requires filtering with an amplification that increases exponentially with the spatial frequency. This produces an infinite variance in the continuous analysis, and instability (extremely high variances) in a corresponding discretized implementation.

3.3 Spherical PET with full angular coverage

As before, we use a 1D Gaussian with standard deviation σ_t to represent the TOF-kernel along the LOR. A 3D isotropic Gaussian with σ_{recon} is used for the 3D reconstruction resolution. Thus, σ_{recon} represents the required imaging resolution, playing a similar role as σ_o in the detection task. The detector resolution is represented with a 2D isotropic Gaussian with standard deviation σ_{sys}, because allowing for an anisotropic detector resolution as in the previous section (σ_y and σ_z) would complicate the derivations considerably. Appendix 10.B shows that equation (24) can then be rewritten as

$$\frac{\text{Var}(f_{\text{3D}}(\vec{r}))}{\text{Var}(p_{\text{3D}}(\vec{r}))} \simeq \int_0^\infty \nu^3 d\nu \frac{e^{-4\pi^2\nu^2\sigma_{\text{recon}}^2}}{\int_0^\nu d\xi\ e^{-4\pi^2(\sigma_t^2\xi^2+\sigma_{\text{sys}}^2(\nu^2-\xi^2))}} \tag{25}$$

The subscript "3D" is used to emphasize that this equation holds for a spherical PET with full 3D sampling. For accurate results, equation is (25) is rewritten for better numerical stability (see appendix 10.B) and integrated numerically. However, in some cases or with some approximations, closed form expressions are obtained.
Because we consider 3D TOF-PET with complete angular sampling, a finite variance is obtained when $\sigma_{\text{sys}} < \sigma_{\text{recon}}$ and/or $\sigma_t < \sigma_{\text{recon}}$, but not otherwise:

$$\left.\frac{\text{Var}(f_{\text{3D}}(\vec{r}))}{\text{Var}(p_{\text{3D}}(\vec{r}))}\right|_{\sigma_{\text{recon}}\leq\sigma_{\text{sys}},\sigma_t} = \infty \tag{26}$$

In the special case where $\sigma_t = \sigma_{\text{sys}} < \sigma_{\text{recon}}$, equation (25) reduces to

$$\left.\frac{\text{Var}(f_{\text{3D}}(\vec{r}))}{\text{Var}(p_{\text{3D}}(\vec{r}))}\right|_{\sigma_t=\sigma_{\text{sys}}} \simeq \frac{\sqrt{\pi}}{32\pi^3}\frac{1}{(\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2)^{3/2}} \tag{27}$$

If $\sigma_t \gg \sigma_{\text{recon}}$ and $\sigma_t \gg \sigma_{\text{sys}}$, then a good approximation is given by

$$\left.\frac{\text{Var}(f_{\text{3D}}(\vec{r}))}{\text{Var}(p_{\text{3D}}(\vec{r}))}\right|_{\sigma_t\gg\sigma_{\text{sys}},\sigma_{\text{recon}}} \simeq \frac{\sqrt{\pi}}{8\pi^4}\frac{\sqrt{\sigma_t^2-\sigma_{\text{sys}}^2}}{(\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2)^2} \simeq \frac{\sqrt{\pi}}{8\pi^4}\frac{\sigma_t}{(\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2)^2} \tag{28}$$

3.4 Multi-ring PET with septa

We consider a cylindrical (multi-ring) TOF-PET system with septa that eliminate all LORs except those perpendicular to axis z of the cylinder. Ω is the unit circle in the plane $z = 0$. The resolution in the z-direction is determined by σ_{sys}, the resolution perpendicular to z is determined by the combination of the TOF resolution σ_t and the detector resolution σ_{sys}. Inserting the corresponding Gaussians in (24) produces (see appendix 10.C):

$$\frac{\text{Var}(f_{\text{septa}}(\vec{r}))}{\text{Var}(p_{\text{septa}}(\vec{r}))} \simeq \frac{\sqrt{\pi}}{\sqrt{\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2}}\int_0^\infty \nu_r d\nu_r \frac{e^{-4\pi^2\sigma_{\text{recon}}^2\nu_r^2}}{\int_0^{2\pi} d\theta e^{-4\pi^2\nu_r^2(\sigma_t^2\cos^2\theta+\sigma_{\text{sys}}^2\sin^2\theta)}} \tag{29}$$

Because TOF does not contribute to resolution in the z-direction, a finite variance can only be obtained if $\sigma_{\text{sys}} < \sigma_{\text{recon}}$.

$$\left.\frac{\text{Var}(f_{\text{septa}}(\vec{r}))}{\text{Var}(p_{\text{septa}}(\vec{r}))}\right|_{\sigma_{\text{recon}}\leq\sigma_{\text{sys}}} = \infty \tag{30}$$

In the special case where $\sigma_t = \sigma_{\text{sys}} < \sigma_{\text{recon}}$, one finds

$$\left.\frac{\text{Var}(f_{\text{septa}}(\vec{r}))}{\text{Var}(p_{\text{septa}}(\vec{r}))}\right|_{\sigma_t=\sigma_{\text{sys}}} \simeq \frac{\sqrt{\pi}}{16\pi^3}\frac{1}{(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})^{3/2}} \tag{31}$$

If $\sigma_t \gg \sigma_{\text{recon}} > \sigma_{\text{sys}}$, then a good approximation is given by

$$\left.\frac{\text{Var}(f_{\text{septa}}(\vec{r}))}{\text{Var}(p_{\text{septa}}(\vec{r}))}\right|_{\sigma_t\gg\sigma_{\text{sys}},\sigma_{\text{recon}}} \simeq \frac{\sqrt{\pi}}{32\pi^2}\frac{\sqrt{\sigma_t^2-\sigma^2_{\text{sys}}}}{(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})^2} \simeq \frac{\sqrt{\pi}}{32\pi^2}\frac{\sigma_t}{(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})^2} \tag{32}$$

3.5 Circular 2D PET

A PET system in a 2D world is obviously unrealistic, but it is interesting because often much can be learned from studying the simple 2D tomographic PET problem. The situation is similar to the previous case, except that there is no z-axis. Consequently, the detector resolution and TOF resolution play the same role, and expressions symmetrical in σ_t and σ_{sys} are obtained.

$$\frac{\text{Var}(f_{\text{2D}}(\vec{r}))}{\text{Var}(p_{\text{2D}}(\vec{r}))} \simeq 2\pi\int_0^\infty \nu d\nu \frac{e^{-4\pi^2\nu^2(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})}}{\int_0^{2\pi} d\theta e^{-4\pi^2\nu^2(\sigma_t^2-\sigma^2_{\text{sys}})\sin^2\theta}} \tag{33}$$

For perfect detector resolution $\sigma_{\text{sys}} = 0$, this expression is equivalent to Tomitani's equation (17) in (Tomitani, 1981) (where it was written differently, using modified Bessel function I_0). A finite variance is obtained provided that $\sigma_t < \sigma_{\text{recon}}$ or $\sigma_{\text{sys}} < \sigma_{\text{recon}}$:

$$\left.\frac{\text{Var}(f_{\text{2D}}(\vec{r}))}{\text{Var}(p_{\text{2D}}(\vec{r}))}\right|_{\sigma_{\text{recon}}\le\sigma_{\text{sys}},\sigma_t} = \infty \tag{34}$$

When $\sigma_{\text{sys}} = \sigma_t < \sigma_{\text{recon}}$, equation (33) reduces to

$$\left.\frac{\text{Var}(f_{\text{2D}}(\vec{r}))}{\text{Var}(p_{\text{2D}}(\vec{r}))}\right|_{\sigma_t=\sigma_{\text{sys}}} = \frac{1}{8\pi^2(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})} \tag{35}$$

If $\sigma_t \gg \sigma_{\text{recon}} > \sigma_{\text{sys}}$, then a good approximation is given by

$$\left.\frac{\text{Var}(f_{\text{2D}}(\vec{r}))}{\text{Var}(p_{\text{2D}}(\vec{r}))}\right|_{\sigma_t\gg\sigma_{\text{sys}},\sigma_{\text{recon}}} \simeq \frac{1}{16\pi}\frac{\sqrt{\sigma_t^2-\sigma^2_{\text{sys}}}}{(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})^{3/2}} \simeq \frac{1}{16\pi}\frac{\sigma_t}{(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})^{3/2}} \tag{36}$$

See appendix 10.D for a few intermediate equations.

3.6 Cylindrical fully 3D PET

Finally, the equation for a realistic cylindrical PET with axial acceptance angle of 2β is given below (β is the maximum co-polar angle at the center of the scanner). Even if we assume that R, K and S are all Gaussians, the expression is more complicated than those obtained above. We start again from (24), which is repeated as the first equation below. The covered solid angle Ω is defined by limiting the co-polar acceptance angle to $[-\beta, \beta]$ with $\beta \le \pi/2$,

and equals $\Omega = 4\pi \sin\beta$. We rewrite $\vec{\nu}$ as $\nu\hat{\nu}_{\varphi,\theta}$, where $\hat{\nu}$ is the unit vector defined by copolar angle φ and azimuthal angle θ.

$$\begin{aligned}
\frac{\mathrm{Var}(f_{\mathrm{cyl}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{cyl}}(\vec{r}))} &\simeq \int_{\mathbb{R}^3} \frac{|R(\vec{\nu})|^2}{\int_\Omega |K(\hat{u}\cdot\vec{\nu})|^2 |S(\vec{\nu}\perp\hat{u})|^2 d\hat{u}}\, d\vec{\nu} \\
&= \int_0^\infty d\nu \int_{-\pi/2}^{\pi/2} d\varphi \int_0^{2\pi} d\theta \frac{\nu^2\cos\varphi\ \ |R(\nu\ \hat{\nu}_{\varphi,\theta})|^2}{\int_{-\beta}^{\beta}\cos\varphi' d\varphi' \int_0^{2\pi} d\theta' |K(\nu\hat{\nu}_{\varphi,\theta}\cdot\hat{u}_{\varphi',\theta'})|^2 |S(\nu\hat{\nu}_{\varphi,\theta}\perp\hat{u}_{\varphi',\theta'})|^2} \\
&= 2\pi \int_0^\infty d\nu \int_0^{\pi/2} d\varphi \frac{\nu^2\cos\varphi\ \ |R(\nu)|^2}{\int_{-\beta}^{\beta}\cos\varphi' d\varphi' \int_0^{\pi} d\theta' |K(\nu\ \hat{u}_{\varphi',\theta'}\cdot\hat{\nu}_{\varphi,0})|^2 |S(\nu\hat{\nu}_{\varphi,0}\perp\hat{u}_{\varphi',\theta'})|^2}
\end{aligned} \tag{37}$$

with

$$\begin{aligned}
\hat{\nu}_{\varphi,\theta} &= (\cos\varphi\cos\theta, \cos\varphi\sin\theta, \sin\varphi) \\
\hat{u}_{\varphi',\theta'} &= (\cos\varphi'\cos\theta', \cos\varphi'\sin\theta', \sin\varphi') \\
|\hat{\nu}_{\varphi,0}\cdot\hat{u}_{\varphi',\theta'}|^2 &= (\cos\varphi'\cos\varphi\cos\theta' + \sin\varphi'\sin\varphi)^2 \\
|\hat{\nu}_{\varphi,0}\perp\hat{u}_{\varphi',\theta'}|^2 &= 1 - |\hat{\nu}_{\varphi,0}\cdot\hat{u}_{\varphi',\theta'}|^2 \\
|R(\nu\hat{\nu})|^2 &= |R(\nu)|^2 = e^{-4\pi^2\sigma_{\mathrm{recon}}^2\nu^2} \\
|K(\nu)|^2 &= e^{-4\pi^2\sigma_t^2\nu^2} \\
|S(\nu\hat{\nu})|^2 &= e^{-4\pi^2\sigma_{\mathrm{sys}}^2\nu^2}
\end{aligned} \tag{38}$$

The last equality of (37) holds because we assume that $R(\vec{\nu})$, $K(\nu)$ and $S(\vec{\nu})$ are Gaussians, because of the circular symmetry in θ and θ', because we integrate over all possible angles θ and θ', and because of the mirror-symmetry in φ.
It may be possible to further simplify equation (37) analytically, but most likely not all the way to a closed form solution. For now, we have simply computed the integral numerically. Using only about a hundred samples along each dimension, the variance for a single value of σ_{recon}, σ_{sys} and σ_t could be computed in 1.25 s with a straightforward implementation and running on a single core.
When $\sigma_{\mathrm{sys}} = \sigma_t$, the equation reduces to

$$\begin{aligned}
\left.\frac{\mathrm{Var}(f_{\mathrm{cyl}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{cyl}}(\vec{r}))}\right|_{\sigma_t=\sigma_{\mathrm{sys}}} &\simeq 2\pi \int_0^\infty d\nu \int_0^{\pi/2} d\varphi\ \frac{\nu^2\cos\varphi\ e^{-4\pi^2\sigma_{\mathrm{recon}}^2\nu^2}}{\int_{-\beta}^{\beta}\cos\varphi'\ d\varphi' \int_0^{\pi} d\theta'\ e^{-4\pi^2\sigma_{\mathrm{sys}}^2\nu^2}} \\
&= \frac{1}{\sin\beta}\ \frac{\sqrt{\pi}}{32\pi^3}\ \frac{1}{(\sigma_{\mathrm{recon}}^2 - \sigma_{\mathrm{sys}}^2)^{3/2}}.
\end{aligned} \tag{39}$$

Appendix 10.E shows how a good approximation can be obtained for wide TOF-kernels and PET systems with short axial length, i.e. $\sigma_t \gg \sigma_{\mathrm{recon}} > \sigma_{\mathrm{sys}}$ and $|\beta| \ll \pi/2$. The result is

$$\left.\frac{\mathrm{Var}(f_{\mathrm{cyl}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{cyl}}(\vec{r}))}\right|_{\sigma_t\gg\sigma_{\mathrm{recon}}>\sigma_{\mathrm{sys}}} \simeq \frac{1}{2\beta}\frac{\sqrt{\pi}}{32\pi^2}\frac{\sqrt{\sigma_t^2 - \sigma_{\mathrm{sys}}^2}}{(\sigma_{\mathrm{recon}}^2 - \sigma_{\mathrm{sys}}^2)^2} \simeq \frac{1}{2\beta}\frac{\sqrt{\pi}}{32\pi^2}\frac{\sigma_t}{(\sigma_{\mathrm{recon}}^2 - \sigma_{\mathrm{sys}}^2)^2} \tag{40}$$

3.7 Application to discrete data

To apply the continuous equations above, the equations have to be converted to their corresponding discrete versions. This process is complicated by the fact that continuous TOF-PET

projections represent count densities (counts per length^n, where n is the number of dimensions of the image), whereas in practice, a discrete TOF-PET data set contains counts. This conversion is explained in appendix 10.F. The analytical results can be verified quantitatively with PET simulation experiments by setting

$$\begin{aligned}
\mathrm{Var}(p_{\mathrm{3D}}(\vec{r}_k)) &= p^b(\vec{r}_k)\frac{4\pi}{T\,N_{\mathrm{proj}}} \quad \text{in equation (25)} \\
\mathrm{Var}(p_{\mathrm{septa}}(\vec{r}_k)) &= p^b(\vec{r}_k)\frac{2\pi}{T\,N_{\mathrm{proj}}} \quad \text{in equation (29)} \\
\mathrm{Var}(p_{\mathrm{2D}}(\vec{r}_k)) &= p^b(\vec{r}_k)\frac{2\pi}{T\,N_{\mathrm{proj}}} \quad \text{in equation (33)} \\
\mathrm{Var}(p_{\mathrm{cyl}}(\vec{r}_k)) &= p^b(\vec{r}_k)\frac{4\pi\ \sin\beta}{T\,N_{\mathrm{proj}}} \quad \text{in equation (37)}
\end{aligned} \tag{41}$$

where $p^b(\vec{r}_k)$ is the expected count in the central data bin of the simulated TOF-PET scan, T is the ratio of the TOF bin size to the detector size d, and d also equals the reconstruction voxel size (assuming cubic voxels). To justify the approximation of sums by integrals, d should be small compared to σ_{recon} and σ_{sys}. N_{proj} is the number of acquired parallel beam TOF-PET projections, and β is the co-polar acceptance angle.

3.8 Reconstruction variance predictions based on the integral equations

In this section, the integral equations derived above are numerically evaluated to analyze the effect of the detector resolution, the TOF resolution and the reconstructed image resolution on the variance in the center of the radioactive sphere. We compare these variances to those obtained from the simple linear predictions for moderate TOF resolution, which corresponds to the current rule of thumb.
For relatively wide TOF-kernels, equations (28), (32), (36) and (40) assert that this variance is approximately proportional to

$$\mathrm{Var}(f_{\mathrm{recon}}(\vec{r})) \sim \frac{\sigma_t}{(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})^{\frac{n+1}{2}}} \tag{42}$$

where $n = 2$ for 2D PET and $n = 3$ for the other PET systems. This shows that improving the detector resolution improves the effective sensitivity. This can be attributed to the amount of smoothing that can be applied during reconstruction: a smoothing with a kernel width proportional to $\sqrt{\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}}}$ can be applied in each dimension. If the reconstruction noise were uncorrelated, then such a smoothing would decrease the noise by a factor proportional to $\left(\sqrt{\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}}}\right)^n$. However, it is correlated high frequency noise [6], and as a result, smoothing is more effective and suppresses this noise by a factor $\sim \left(\sqrt{\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}}}\right)^{n+1}$. This effect can be strong: e.g., when aiming for a reconstruction resolution of 6 mm, decreasing the detector resolution from 5 to 4 mm improves the effective sensitivity with a factor of 3.3. This result agrees with the findings published by Muehllehner (Muehllehner, 1985), Surti *et al.* (Surti et al., 2013) and others.

[6] All image reconstruction algorithms must amplify high frequencies, as illustrated by the ramp filter in non-TOF tomography. As a result, the uncorrelated data noise propagates into correlated image noise.

Equation (42) also shows that the variance is proportional to the width of the TOF-kernel, implying that improving the TOF resolution with a particular factor, improves the effective sensitivity with the same factor, a well known result (Tomitani, 1981). However, Toussaint *et al.* observed in their "ultrafast TOF-PET" simulations an unexpectedly strong improvement in the resolution of the reconstructed image, as the TOF resolution approaches the detector resolution. Translated to our methodology, this effect should improve the variance in the reconstruction even more than predicted by (42). Therefore, we used our equations to look for a better than linear relationship between variance and TOF, for narrow TOF-kernels.
For this comparison, it is not necessary to have absolute quantification; it suffices to quantify the variance up to a constant scale factor, ensuring that this scale factor is the same everywhere. As mentioned above, the figures show plots with the resolution expressed as FWHM in mm, as is common practice in the PET community.

3.8.1 Spherical PET with full angular coverage

The left plot of figure 5 shows the variance as a function of the TOF resolution for a fixed reconstruction PSF of 6 mm, and for detector resolutions between 1 and 5.8 mm. The curves are compared to the approximate linear prediction of equation (28). For wide TOF-kernels ($\sigma_t \gg \sigma_{\text{recon}}$), the variances from the accurate model (25) agree well with the linear prediction, but for narrow TOF-kernels, the variances are lower than these predictions. The right panel of figure 5 plots the ratio of the accurate variances and their linear predictions.

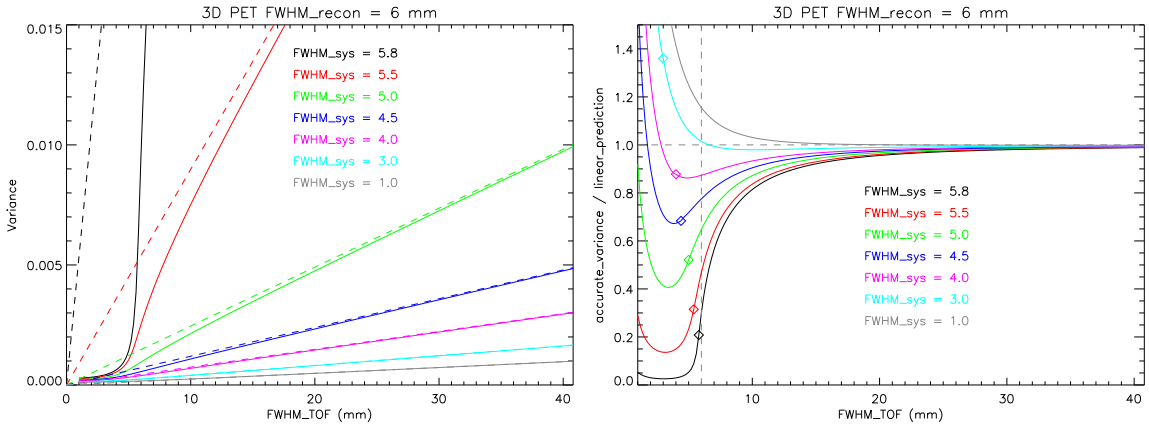

Figure 5: *Left: plots of the variance in the reconstructed image from a spherical PET, for a reconstruction PSF of 6 mm FWHM, as a function of the TOF resolution (in mm). Curves are shown for detector resolutions varying between 1 and 5.8 mm (solid lines). The dashed lines show the variance predicted with the approximate moderate TOF equation (28). Right: the ratio of the accurate variance calculations and their linear predictions. The vertical line is where $F_t = F_{\text{recon}}$; the diamonds indicate on each curve where $F_t = F_{\text{sys}}$.*

The variance improvement due to ultrafast TOF is larger for poorer detector resolutions σ_{sys}. The better than linear benefit from TOF kicks in well before the TOF resolution becomes similar to the reconstruction and detector resolutions. However, there is no such benefit for very high detector resolutions, presumably because the resolution provided by the detector is already so good, that further resolution improvements with TOF are essentially smoothed away to achieve the target reconstruction resolution.

Because spherical PET provides complete angular sampling, good image resolution σ_{recon} can be obtained with good TOF resolution, even when the detector resolution is poor. Of course, good image resolution can also be achieved with good detector resolution, even when the TOF resolution is poor. Figure 6 compares the variances in the reconstructed image for these two situations. The left panel plots the variance for a fixed "poor" TOF resolution of 10 mm, when the detector resolution is varied between 1 and 5.8 mm (black curve), and for a fixed poor detector resolution of 10 mm and TOF varying between 1 and 5.8 mm (red curve). In all cases, the resolution of the final image was 6 mm. In the first case, that resolution was achieved thanks to the detector resolution, because the TOF resolution of 10 mm exceeded the reconstruction resolution. In the second case, it was the other way around. The right panel shows the ratio between the variance obtained with TOF and the variance obtained with the detector. Achieving the image resolution with the detector produces a lower variance. That is expected, because the detector provides good resolution in two dimensions, whereas TOF provides it only in one dimension.

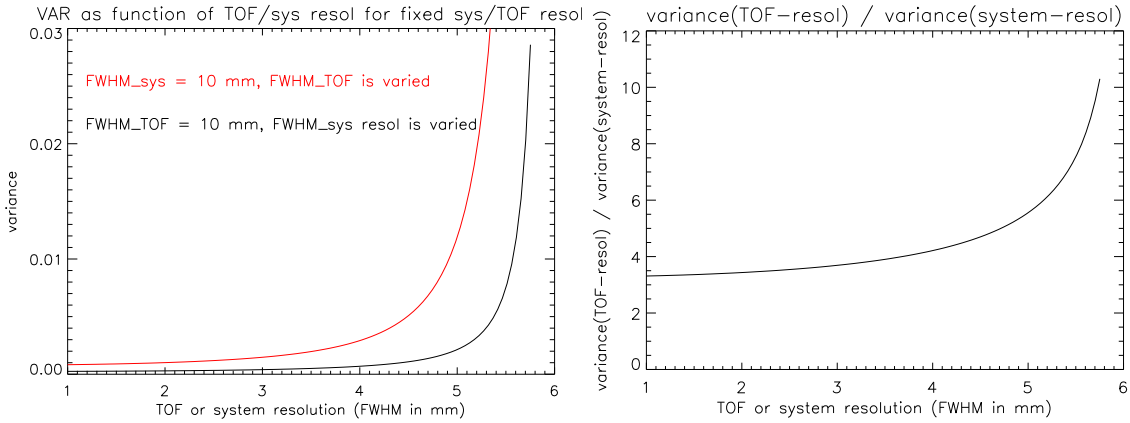

Figure 6: *Left: the variance as a function of TOF (red) or detector (black) resolution (in mm), for a detector or TOF resolution fixed at 10 mm and a spherical PET. The reconstruction resolution was 6 mm FWHM. Right: the ratio of the red to the black curves in the left panel.*

3.8.2 Multi-ring PET with septa

The left panel of figure 7 shows the variances of a TOF-PET system with septa as a function of the TOF resolution, for different detector resolutions. The results are very similar to those for a spherical TOF-PET system in figure 5. The right panel shows the ratio of the variance of the system with septa to that of the spherical PET system, for a fixed detector resolution of 4 mm. These results hold if both systems acquired the same total number of counts, which is achieved automatically by using the definitions of equations (41) for $\text{Var}(p_{\text{3D}}(\vec{r}))$ and $\text{Var}(p_{\text{septa}}(\vec{r}))$: the counts in each projection equal the activity in the image, and both systems have the same number of projections N_{proj}, distributed over their solid angle (4π for spherical PET and 2π for PET with septa). Of course, in practice, a system with septa would have a much smaller N_{proj} and therefore need to scan much longer for acquiring the same total number of counts. The curves show that spherical PET has always a lower variance, except when the TOF resolution equals the detector resolution. In that case, the projections

of both systems are identical to the image, isotropically blurred with $\sigma_{\text{sys}} = \sigma_t$, see equations (27) and (31), in combination with (41).
For moderate TOF resolutions $\sigma_t \gg \sigma_{\text{recon}}$, the variance of the system with septa is a factor $\pi^2/8 \simeq 1.23$ higher than that of the spherical system, see equations (32) and (28), using (41). This ratio $8/\pi^2 \simeq 0.81$ between the variances of these PET systems was also obtained in (Defrise et al., 1990).
Finally, making the TOF resolution superior to the detector resolution suppresses the variance more in spherical PET than in PET with septa, because in the latter, the TOF resolution does not contribute to the axial resolution of the reconstructed image.

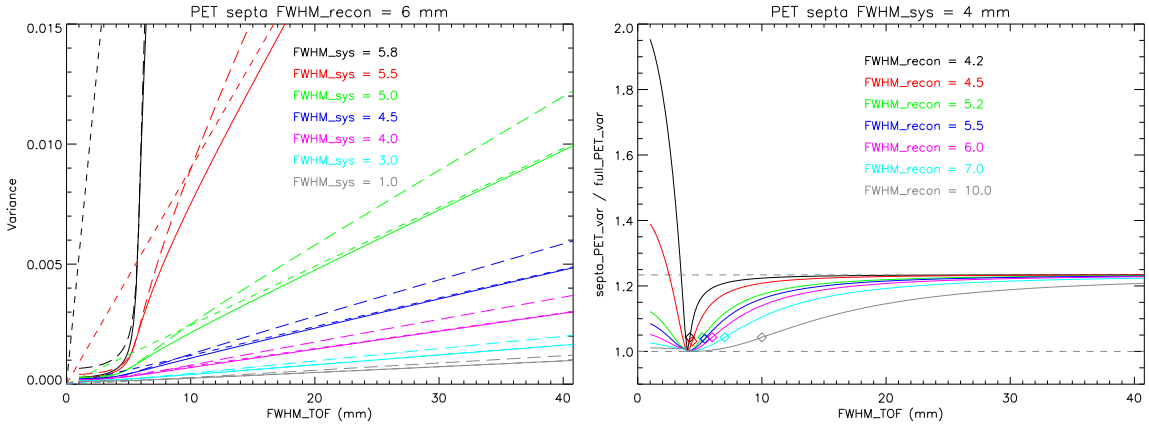

Figure 7: *Left: reconstruction variance of PET with septa, for a reconstruction PSF of 6 mm FWHM, as a function of the TOF resolution (mm). Curves for detector resolutions between 1 and 5.8 mm (solid lines) are shown. The dashed lines show the variance predicted with the approximate moderate TOF equation (32). Right: the ratio of the variance from PET with septa to that from spherical PET, assuming both systems had seen the same total number of counts. In this plot, the detector resolution is fixed at 4 mm, while curves are shown for different reconstruction resolutions ranging from 4.2 to 10 mm. The diamonds indicate on each curve where* $F_t = F_{\text{recon}}$.

3.8.3 Circular 2D PET

Figure 8 shows the results for a circular TOF-PET system operated in a 2D world. The curves are similar to those of the two 3D TOF-PET systems shown above, except that the benefit from TOF is stronger: when the TOF-kernel was wider than the reconstruction resolution, improving TOF always helped more than predicted by the linear approximation, even for extremely high detector resolutions. A stronger benefit from TOF was expected, because in 2D, the role of the TOF resolution is identical to that of the detector resolution, which is not the case in a 3D world.

3.8.4 Cylindrical fully 3D PET

Figure 9 compares the results obtained above for PET with septa and spherical PET, to the two corresponding extreme cases of the cylindrical fully 3D PET (37). As before, all variances were scaled with equations (41) such that they correspond to acquisitions with the same total number of counts. The first case considered a very small axial acceptance angle of $\beta = \pi/50$,

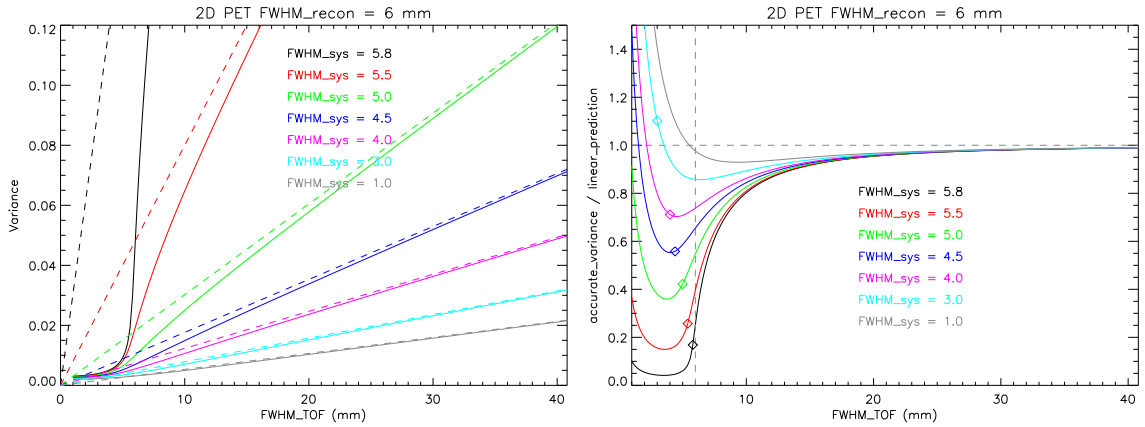

Figure 8: *Results for a 2D PET system. Left: the variance in the reconstructed image of a uniform disk, for a reconstruction PSF of 6 mm, as a function of the TOF resolution (mm), with detector resolutions varying between 1 and 5.8 mm (solid lines). For each curve, also the variance predicted with the approximate moderate TOF equation (36) are given (dashed lines). Right: the ratio of the accurate variance calculations and their linear predictions. The vertical line is where $F_t = F_{\text{recon}}$; the diamonds indicate on each curve where $F_t = F_{\text{sys}}$.*

which corresponds to PET with septa. The second case used $\beta = \pi/2$, which corresponds to spherical PET. The figure confirms that the numerical integration of (37) agrees well with the results obtained previously, despite the modest sampling applied during integral evaluation.

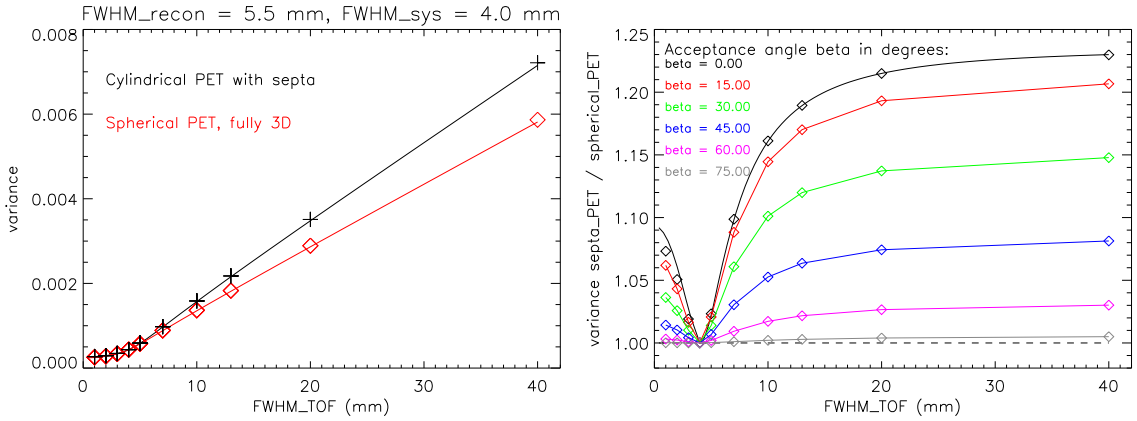

Figure 9: *Left: The solid lines show the variances for PET with septa (black) and spherical PET (red) as a function of F_t in mm, computed with numerical integration of (29) and (88) or (89), for the case when $F_{\text{recon}} = 5.5$ mm and $F_{\text{sys}} = 4$ mm. The points ($\diamond$, $+$) are computed with numerical integration of (37) for $\beta = \pi/50$ and $\beta = \pi/2$. Right: the ratio of the variances of the cylindrical PET (at different axial acceptance angles) to the variances of spherical PET.*

.

The right panel of figure 9 plots the ratio of the variances obtained for different axial acceptance angles β to the corresponding variances obtained with spherical PET. The values of β range from 0 (PET with septa) to $\pi/2$ (spherical PET).

The figure shows that the variance depends non-linearly on β: when $\beta = 45°$, the variance is closer to that of the spherical PET than that of PET with septa. For a TOF resolution of 40 mm, the variance of a cylindrical PET equals the mean of the two extreme cases when $\beta \simeq 37°$.

4 Numerical verification with simulations

4.1 Experiments

To verify the results obtained for 2D PET and for cylindrical 3D PET with septa, simulations with multiple noise realizations were performed. Variances have been computed as a function of different TOF resolutions, for a fixed detector resolution and a few reconstruction resolutions.
Since comparison with simulations requires properly scaled variance values, it is essential to convert the predicted variances to those for binned PET data, using equations (41).
The simulated phantom was a uniform radioactive cylinder with a radius of 80 mm. Its attenuation was zero. It was discretized into a 3D image of 200 × 200 × 19 voxels for PET with septa and a 2D image of 200 × 200 pixels for 2D PET. The voxel (or pixel) size was 1 mm. The TOF sinograms had 200 angles uniformly distributed over 180°. The TOF bin width was 0.8 mm for the 2 mm FWHM TOF resolution, and increased with increasing TOF resolution to 8 mm for 20 mm TOF. The mean total count in all sinograms was the same, the count in the central TOF bin increased from 8 at 2 mm TOF to 80 at 20 mm TOF. The finite system resolution was simulated by 2D or 1D Gaussian smoothing of the projections, for PET with septa or 2D PET, respectively. For each case, 50 Poisson noise realizations were computed.
Images were reconstructed from the noisy sinograms with MLEM using resolution modeling. For each case, the number of iterations was determined by monitoring the reconstruction PSF. This was done with reconstructions from two noise-free simulated sinograms, one with and one without a small hot spot in the center of the image. The hot spot consisted of a 10% increased activity in a central region of $2 \times 2\,(\times 2)$ voxels. After each iteration, we determined the FWHM of the Gaussian post-smoothing, which was required to minimize the squared difference between the reconstruction PSF and the target Gaussian PSF. This was done for two target PSFs: 5.0 mm and 5.5 mm FWHM. The iterations were terminated if the postsmoothing FWHM was at least 4.7 mm for the target resolution of 5.0 mm, and 5.0 mm for the target resolution of 5.5 mm. The fact that the required postsmoothing kernel is almost as wide as the target PSF indicates that the reconstruction is close to convergence.
However, because of the resolution modeling, the reconstruction PSF obtained after the MLEM iterations is not Gaussian at all, since it suffers from Gibbs artifacts. Consequently, although the procedure above gives a reasonable metric to monitor convergence, it produces a different PSF for each case, because the Gibbs artifacts are different for different TOF resolutions. Initial tests suggested that this mismatch in reconstruction PSF created also a mismatch in the observed variances. For that reason, the Gaussian postsmoothing was only done to determine the required number of iterations. To obtain a well matched PSF for all cases, the final postsmoothing was done with a filter defined in the Fourier domain, by dividing the FFT of the target reconstruction PSF by the FFT of the MLEM reconstruction PSF obtained after the last iteration. Denoting the target MTF as R and the MLEM

reconstruction MTF as P, the filter was defined as

$$\text{Filter} = |R| \, \frac{|P|}{|P|^2 + (\max(|P|)/100)^2} \tag{43}$$

The pixel variance was computed in a central region (radius of 6.7 mm), by using the noise-free reconstructions as the estimate of the mean. This produced 50 estimates of the pixel variance, from which the final estimate of the mean variance and its standard deviation were computed.
By using the equations (41) with $p^b(\vec{r}_k)$ equal to the count of the central bin in the simulations, the predictions are quantitative and can be compared directly to the variances obtained from the simulations.

4.2 Results

Figure 10 compares the variances obtained in the simulations with their predictions. The agreement is better for better TOF resolutions, for the wider reconstruction PSF and it is better for 2D PET than for 3D PET with septa. For 3D PET with septa, the largest error (at 20 mm TOF) was 26%, for 2D PET it was 17%. On the other hand, for ultrafast TOF resolutions of 2 and 4 mm, the errors were less than 10% and within the error bars.
The simulations confirm the Toussaint effect: the benefit from TOF is better than linear when the TOF resolution approaches the detector resolution.
Table 1 lists the number of iterations that was needed to obtain sufficient convergence, as estimated by requiring a minimum amount of Gaussian postsmoothing to obtain an optimal match with the target reconstruction PSF. As expected, more iterations were needed for poorer TOF resolutions, for a reconstruction at higher resolution, and for the 3D PET system with septa.

Table 1: *The number of iterations required to ensure that a postsmoothing of 4.7 mm FWHM was required for imposing the target reconstruction PSF of 5.0 mm FWHM, and a postsmoothing of 5.0 mm FWHM for imposing a reconstruction PSF of 5.5 mm.*

		TOF FWHM [mm]				
		2	4	7	10	20
		number of iterations				
2D PET	FWHM_{recon}=5.5 mm	3	5	13	21	48
PET septa	FWHM_{recon}=5.5 mm	5	6	15	25	56
2D PET	FWHM_{recon}=5.0 mm	5	9	25	42	92
PET septa	FWHM_{recon}=5.0 mm	10	11	30	51	114

5 Comparison of the two analysis methods

Equation (12) gives the SNR^2 of the Hotelling observer for detecting the SLoG when the TOF resolution is moderate, and when the detector resolution is the same in both directions

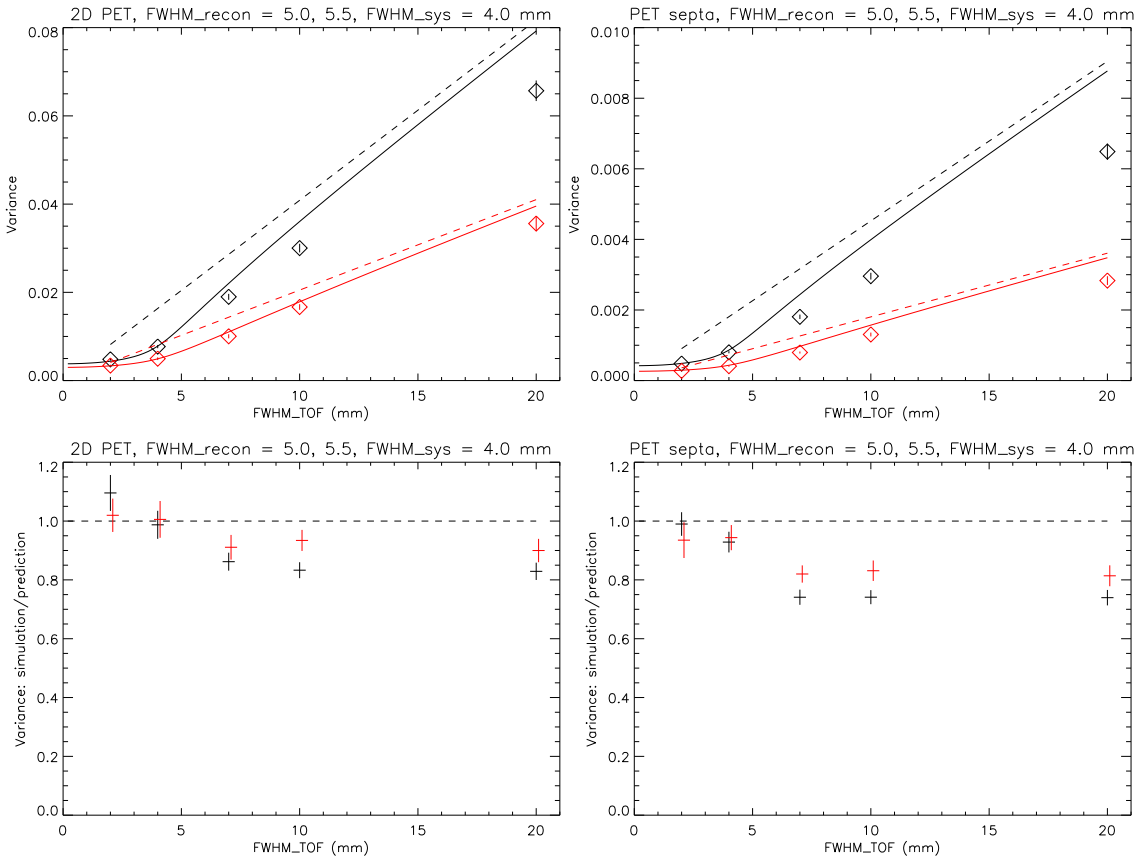

Figure 10: *Plots of the variance as a function of the TOF resolution in mm, for two different reconstruction PSF widths: 5.0 mm (black curves) and 5.5 mm FWHM (red curves). The detector resolution was 4.0 mm FWHM. The solid lines are the accurate predictions, the dashed lines the approximate predictions for moderate TOF resolution and the diamonds are the variances from the simulations. The error bars are ± 2 standard deviations. The left panel is for 2D PET, the right one for 3D PET with septa.*
The second row shows the same simulation variances, but now relative to the accurate predictions (the height of the + symbol is the error bar).

($\sigma_t \gg \sigma_y = \sigma_z$). It can be written as

$$\mathrm{SNR}^2_{detect} = \frac{C_1}{\sigma_t} \frac{\sigma_o^{10}}{(\sigma_y^2 + \sigma_o^2)^3} \tag{44}$$

Equation (28) gives the reconstruction variance for moderate TOF resolution. The corresponding SNR^2 is proportional to the reciprocal of that variance and can be written as

$$\mathrm{SNR}^2_{recon} = \frac{C_2}{\sigma_t}(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})^2 \quad \text{if } \sigma_{\mathrm{recon}} > \sigma_{\mathrm{sys}} \text{ and } \mathrm{SNR}^2_{recon} = 0 \text{ otherwise.} \tag{45}$$

We compare both functions setting $C_1/\sigma_t = C_2/\sigma_t = 1$. Figure 11 plots both squared SNRs as a function of the detector resolution $\sigma_{\mathrm{sys}} = \sigma_y = \sigma_z$, for a required imaging resolution $\sigma_{\mathrm{recon}} = \sigma_o = 4$. Both SNRs are very similar and confirm the Muehllehner effect: the effective sensitivity is a very strong function of the detector resolution.

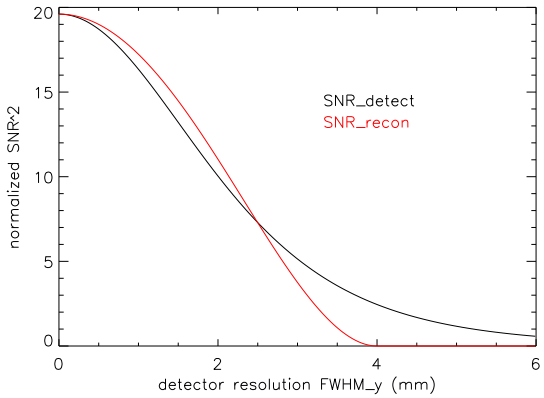

Figure 11: *Plot of squared detection SNR and squared reconstruction SNR as a function of detector resolution* $F_{\text{sys}} = F_y = F_z$ *in mm, for* $F_o = F_{recon} = 4$ *mm.*

Figure 12 compares both squared SNRs as a function of the required imaging resolution $\sigma_{\text{recon}} = \sigma_o$, for a detector resolution $\sigma_y = 5$. For imaging resolutions larger than the detector resolution, the squared SNRs are similar. This is why the SLoG was defined with a scale factor σ_o^4 in (9). Without this scaling, increasing the SLoG size would reduce its maximum activity, which would make it harder to detect and could therefore decrease the SNR. In contrast, the reciprocal of the reconstruction variance always increases for a wider reconstruction PSF.

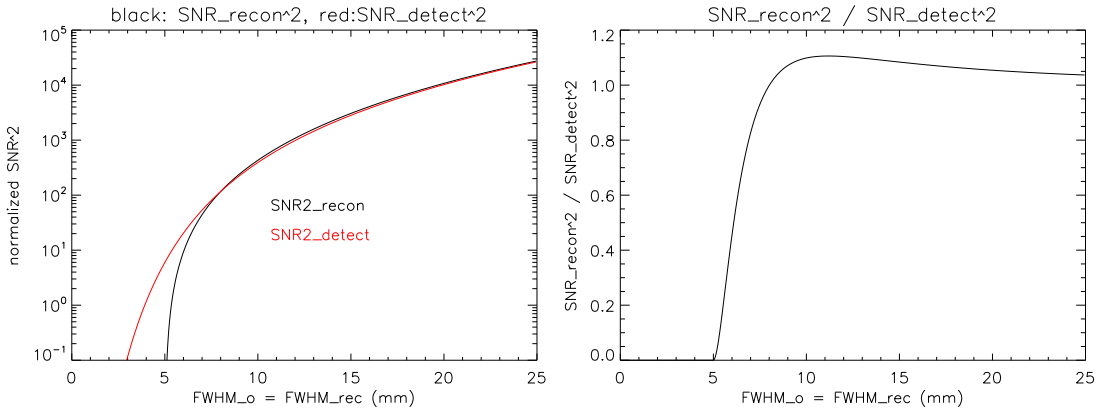

Figure 12: *Left: plot of squared detection SNR and squared reconstruction SNR as a function of the required imaging resolution* $F_o = F_{recon}$ *in mm, for detector resolution* $F_{\text{sys}} = F_y = F_z = 5$ *mm. Right: the ratio of both SNRs.*

Because both squared SNRs have a very similar behaviour, we anticipate that they will produce a similar ranking when comparing different PET systems. In most cases, it will be more convenient to rank PET systems with the detection SNR. It is more general than the reconstruction variance, because it can be computed on a LOR by LOR basis, enabling the evaluation of PET systems that have different resolutions along different LORs. It also avoids the numerical integration that is required for the variance-based approach, even if the considered LOR has an event-based TOF resolution, which may produce a distribution of possibly non-Gaussian TOF-kernels (Nuyts et al., 2022).

6 Discussion

Our equations based on the detection SNR, and our analysis of the reconstruction variance both confirm the Toussaint effect: if the TOF resolution approaches the detector resolution, the effective sensitivity of the PET system increases (much) faster than linearly with the TOF resolution. This effect is already suggested by the approximate equations for moderate TOF resolution, where the reconstruction variance is proportional to $\sqrt{\sigma_t^2 - \sigma_{\text{sys}}^2}$, and therefore depends stronger than linearly on the TOF-kernel width. However, these equations do not hold for small σ_t, which is why we have presented a more accurate analysis that also holds for ultrafast TOF.

The equations show that for 2D PET and spherical PET, ultrafast TOF can push the resolution of the reconstructed image well beyond the detector resolution. That is because in these systems, the TOF resolution can propagate perfectly into the reconstructed image thanks to the complete angular sampling (the activity is reconstructed from all possible LORs). This implies that ultrafast TOF can suppress the acollinearity blurring, because that blurring is perpendicular to the LOR (and can be incorporated in the 2D detector blurring). A related result based on simulations has been reported by Toussaint et al., 2025. In contrast, the positron range reduces the spatial resolution in all directions, and therefore slightly degrades the 1D TOF resolution as well as the 2D detector resolution. However, it can be reduced by modeling it during iterative reconstruction (or deblurring for it after reconstruction), which will be more effective if other sources of resolution loss are better suppressed, as is the case in ultrafast TOF-PET.

Our equations for the variance in the reconstructed image have been derived for the center of a uniform cylindrical or spherical phantom, and relied on the symmetry of that setup. In contrast, the equations based on the SNR of the SLoG detection are valid for individual LORs. As a result, they can be applied also to off-center regions in irregular objects, such as the body of a patient. The squared SNR for detecting the SLoG at a particular position is obtained by simply adding the squared SNRs for all LORs intersecting that position. We have shown that in the center of uniform cylinders or spheres, there is excellent agreement between the SNR^2 of the SLoG detection and the reciprocal of the variance in the reconstructed image. However, that agreement may be poorer for off-center regions in asymmetrical objects.

A good but imperfect agreement is obtained between the analytical predictions of the reconstruction variances and the variances observed in the simulations. As mentioned above, the agreement is worse (with errors of up to 26%) for poorer TOF resolutions and for reconstructions at higher resolution, and it is worse for 3D PET with septa than for 2D PET. On the other hand, the agreement is excellent for high TOF resolutions, which is good news, because high TOF resolution is what PET designers are aiming for (Kwon et al., 2021; Lecoq et al., 2020; Ota et al., 2019).

We do not know the cause of the observed discrepancy between predictions and simulations. Our calculations ignore the finite size of the cylinder and rely on "local shift invariance", i.e. the assumption that the PSF and variance change slowly with position, and can be treated as constants within the range of the TOF-kernel. These approximations should be better for larger objects. However, changing the cylinder radius from 40 mm (not shown) to 80 mm (shown) in the simulations had only a very small effect. Local shift invariance was also invoked to justify a uniform filtering of the entire image to obtain the target PSF. In reality, the reconstruction PSF is position dependent, and therefore a complicated position

dependent filtering would be needed to impose the target PSF everywhere in the image. When the TOF resolution is improved, MLEM has more uniform convergence, making the PSF less position dependent. This may (partly) explain why the agreement is better for better TOF resolutions. A more uniform convergence should also be achieved by increasing the number of iterations, and we found indeed that increasing that number (to the values reported) increased the agreement, but again the effect was small.

MLEM is known to be biased, and more so when the images are noisier. This bias often comes with a change in variance. Consequently, the discrepancy might have been caused by the higher noise on the reconstructions from data with poorer TOF resolution and sharper reconstruction PSF. This was verified by increasing the number of counts with a factor of 25, which should reduce noise induced bias. However, the agreement did not improve, the variances (and the pixel values) were simply 25 times higher.

Finally, the derivation is continuous, the discretization into measurement bins and image pixels is not part of it, it is only applied at the end to make the predictions quantitative. Therefore, it is possible that also a fundamental difference between the continuous and discrete computations has contributed to the discrepancy. Recovering resolution from blurred measurements is an ill-posed problem, and more so if one aims for a higher resolution. As a result, at higher iterations, a very small gain in resolution comes with a very strong increase in variance. This implies that a slight mismatch in resolution can create a strong mismatch in variance. This is illustrated by figure 10, where improving the reconstruction PSF from 5.5 mm to 5.0 mm increases the variance with a factor of two. For that reason, it is not surprising that the results are worse when aiming for a higher resolution in the final image, and making use of a poorer TOF resolution. Probably, our straightforward way of matching the discrete PSFs to their continuous counterparts is not optimal. If we repeat the analysis, assuming that the discrete system resolution of 4 mm FWHM corresponds to an analytical system resolution of 3.85 mm (a change of only 3.75%), then the discrepancies reduce to less than 10%. This assumption is entirely heuristic, but the result shows that for the numerical verification, very accurate resolution matching is critical. The variance plots in figure 13 clearly show that the shape of the variance curves from the simulations agrees very well with that of the slightly adjusted model, confirming the Toussaint effect.

An intuitive explanation of this effect is provided by the method based on the detection performance, equation (11). When a wide TOF-kernel is improved to a less wide one, the effective sensitivity is improved with the same factor, in agreement with the linear prediction, by suppressing the noise contributions from surrounding activity. However, when the TOF-kernel width approaches the detector resolution, TOF not only contributes by suppressing the influence from surrounding activity, but also directly benefits resolution recovery, producing an even larger gain in image quality.

In our reconstruction variance analysis, where we required a perfectly Gaussian reconstruction PSF, this effect is revealed as a markedly decreased variance for reconstructions at a predefined resolution. In contrast, Toussaint *et al.* reported an unexpectedly large improvement of the resolution of the reconstructed image (Toussaint et al., 2020). We believe that this is because in their analysis, the shape of the reconstruction PSF was not constrained, and as a result, the increased amount of information provided by the data could show up as a very narrow, but non-Gaussian PSF, which could not be obtained from data with poorer TOF resolution.

In a recent publication, Ishikawa *et al.* (Ishikawa et al., 2025) analyzed the performance of a dual-panel PET system with ultrafast TOF. The focus of the paper is on the suppression of limited angle artifacts. However, figures 3 and 7 of that paper show that when everything

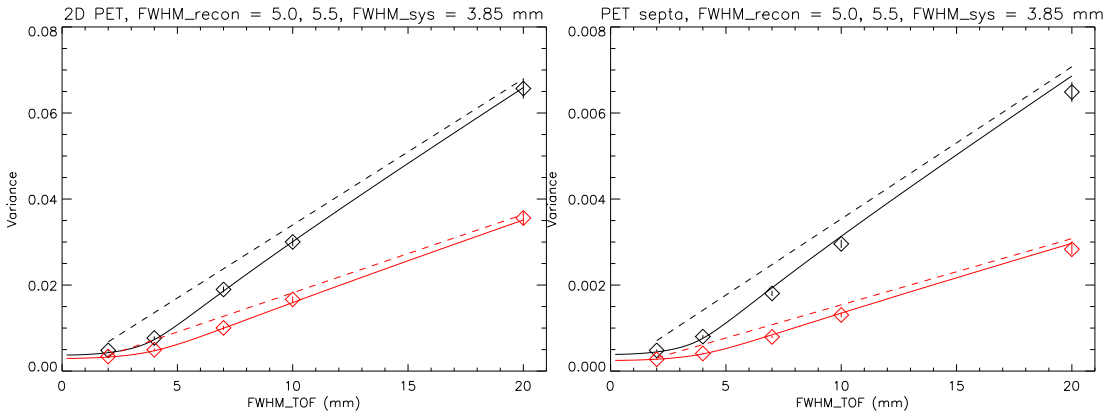

Figure 13: *The variance as a function of the TOF resolution in mm, for reconstruction PSFs of 5 mm (black) and 5.5 mm FWHM (red). The discrete detector resolution was 4 mm FWHM, but the analytical model used instead 3.85 mm FWHM. This reduces the discrepancies shown in fig. 10, which shows the same simulation results with the analytical model using 4 mm FWHM.*

else remains the same, the noise standard deviation in the reconstructed images decreases very fast with decreasing (ultrafast) TOF resolution, much faster than with the square root of the TOF resolution. This result is compatible with our variance analysis.
In the detection analysis, the effect of attenuation can be included in the parameter η, but in the variance analysis, we have ignored it. This is not really a limitation, because the symmetry would ensure that the attenuation were the same for all projections. It would simply decrease the number of measured counts by the same factor for all systems. The analysis focused on the center of a uniform sphere, which is an important limitation. For cylindrical PET systems, it seems reasonable to assume that if a change to detector, TOF and/or reconstruction resolution improves the variance in the center of a uniform sphere, then it will probably do the same with the variance in other, non-uniform objects. This same assumption justifies characterizing and comparing current PET systems with simple NEMA phantoms. However, this assumption may be less reliable for PET systems with unconventional geometries, possibly including very non-uniform angular coverage, because the features of such systems will be more position dependent. For such systems, analytical variance analysis will be more complicated, and it may be necessary to repeat it for multiple well-chosen points in the field of view.

7 Conclusion

Our theoretical analysis and our numerical simulations confirm the findings of Muehllehner (Muehllehner, 1985) and Surti *et al.* (Surti et al., 2013), who found that when the detector resolution of PET is improved, the same image quality can be obtained from data with fewer counts. In other words, improving the detector resolution increases the effective sensitivity of the PET system, even if that resolution was already better than the resolution required for the imaging task. The analysis also confirms and explains the findings of Toussaint *et al.*, who found that "ultra-fast" TOF markedly improves the achievable resolution of the reconstructed image, even if the TOF resolution is still poorer than the detector resolution (Toussaint et al., 2020). This effect is nicely explained by the analysis of the performance of the Hotelling observer for discriminating two different hot blobs. It is also confirmed by our reconstruction

variance analysis, where this benefit did not show up as increased resolution for the same variance, but as a decreased variance for the same resolution (imposed reconstruction PSF). In the experiments of (Toussaint et al., 2020), there was no constraint on the reconstruction PSF, such that the extra information due to improved TOF resolution could produce an image with improved sharpness for a similar variance.

8 Acknowledgment

The authors would like to thank Maxime Toussaint and Roger Lecomte for triggering this research with inspiring discussions about their unexpected simulation results (Toussaint et al., 2020), and Joel Karp for pointing our attention to Muehllehner's paper (Muehllehner, 1985). This work was supported in part by the Research Foundation - Flanders (FWO) under grant G062220N.

9 Appendices about detection SNR

In the computations below, we write the product of S and the SLoG as

$$S\,\mathrm{SLoG}(\vec{r},\sigma_o) \;=\; S\,\sigma_o^4\frac{d}{d\sigma_o}G_3(\vec{r},\sigma_o) \;=\; S_o\frac{d}{d\sigma_o}G_3(\vec{r},\sigma_o) \tag{46}$$

with $S_o = \sigma_o^4 S$.

9.A The SNR for a Gaussian TOF-kernel

The SNR^2 is computed for detecting presence or absence of the SLoG (with standard deviation σ_o) on top of a uniform activity distribution with concentration B. For that purpose, equation (8) is applied, with $\bar{m}_2 - \bar{m}_1$ equal to the SLoG blurred with the Gaussian resolution kernels in x, y and z direction, with standard deviations σ_t, σ_y and σ_z, respectively. Using $G_3(\vec{r},\sigma_o) = G_1(x,\sigma_o)G_1(y,\sigma_o)G_1(z,\sigma_o)$, one obtains

$$\begin{aligned}
&(\bar{m}_1 - \bar{m}_2)(x,y,z)\\
&\quad= \eta S_o\frac{d}{d\sigma_o}\left(G_1(x,\sigma_o)\otimes G_1(x,\sigma_t)\right)\left(G_1(y,\sigma_o)\otimes G_1(y,\sigma_y)\right)\left(G_1(z,\sigma_o)\otimes G_1(z,\sigma_z)\right)\\
&\quad= \eta S_o\frac{d}{d\sigma_o}G_1\left(x,\sqrt{\sigma_t^2+\sigma_o^2}\right)G_1\left(y,\sqrt{\sigma_y^2+\sigma_o^2}\right)G_1\left(z,\sqrt{\sigma_z^2+\sigma_o^2}\right)
\end{aligned} \tag{47}$$

where η is as defined just before (11). Assuming again that the intensities of the two blobs in the SLoG are small compared to the background activity, we set $\bar{m}$ in (8) to

$$\bar{m}(x,y,z) = \eta B(x,y,z)\otimes G_1(x,\sigma_t)\otimes G_1(y,\sigma_y)\otimes G_1(z,\sigma_z) \tag{48}$$

where $B(x,y,z) = B$ inside the uniform phantom and $B(x,y,z) = 0$ outside of it. In addition, we assume here that the size of the object is large compared to the region where the SLoG differs from zero and compared to the TOF resolution σ_t, such that effects at the boundary of the uniform phantom can be ignored. This allows us to set $\bar{m}(x,y,z) = \eta B$. Consequently,

(8) becomes:

$$\mathrm{SNR}^2 = \frac{\eta S_o^2}{B}\iiint dx\,dy\,dz \quad \frac{d}{d\sigma_o}G_1\Big(x,\sqrt{\sigma_t^2+\sigma_o^2}\Big)\,G_1\Big(y,\sqrt{\sigma_y^2+\sigma_o^2}\Big)\,G_1\Big(z,\sqrt{\sigma_z^2+\sigma_o^2}\Big)$$
$$\frac{d}{d\tilde{\sigma}_o}G_1\Big(x,\sqrt{\sigma_t^2+\tilde{\sigma}_o^2}\Big)\,G_1\Big(y,\sqrt{\sigma_y^2+\tilde{\sigma}_o^2}\Big)\,G_1\Big(z,\sqrt{\sigma_z^2+\tilde{\sigma}_o^2}\Big)$$
$$\text{subject to } \tilde{\sigma}_o = \sigma_o \tag{49}$$

We can move the derivative signs in front of the integral signs (which is why we introduced the symbol $\tilde{\sigma}_o$ instead of simply taking the square). The three resulting integrals are independent and easy to evaluate, resulting in:

$$\mathrm{SNR}^2 = \frac{\eta}{(2\pi)^{3/2}}\frac{S_o^2}{B}\frac{d}{d\sigma_o}\frac{d}{d\tilde{\sigma}_o}\frac{1}{\sqrt{2\sigma_t^2+\sigma_o^2+\tilde{\sigma}_o^2}}\frac{1}{\sqrt{2\sigma_y^2+\sigma_o^2+\tilde{\sigma}_o^2}}\frac{1}{\sqrt{2\sigma_z^2+\sigma_o^2+\tilde{\sigma}_o^2}}\Bigg|_{\tilde{\sigma}_o=\sigma_o} \tag{50}$$

Computing the derivatives, setting $\tilde{\sigma}_o = \sigma_o$ and substituting $S_o = S\sigma_o^4$ produces equation (11).
For 2D-PET, the result is similar with two differences: the SLoG is scaled with σ_o^3 to make its maximum independent of the size, and the factor and terms involving σ_z must be dropped:

$$\mathrm{SNR}_{2D}^2 = \frac{\eta\,(S\sigma_o^3)^2\,\sigma_o^2}{8\pi B\sqrt{(\sigma_t^2+\sigma_o^2)(\sigma_y^2+\sigma_o^2)}}\left(\frac{1}{2}\left(\frac{1}{\sigma_t^2+\sigma_o^2}+\frac{1}{\sigma_y^2+\sigma_o^2}\right)^2+\frac{1}{(\sigma_t^2+\sigma_o^2)^2}+\frac{1}{(\sigma_y^2+\sigma_o^2)^2}\right) \tag{51}$$

9.B The SNR when the TOF-kernel is not much smaller than the object diameter

In appendix 9.A above, it was assumed that effects near the boundary of the uniform phantom could be ignored. If that assumption cannot be made, then equation (49) must be modified to

$$\mathrm{SNR}^2 = \eta S_o^2\frac{d}{d\sigma_o}\frac{d}{d\tilde{\sigma}_o}\left(\int\frac{1}{B(x)\otimes G_1(x,\sigma_t)}G_1\Big(x,\sqrt{\sigma_t^2+\sigma_o^2}\Big)\,G_1\Big(x,\sqrt{\sigma_t^2+\tilde{\sigma}_o^2}\Big)\,dx\right.$$
$$\int G_1\Big(y,\sqrt{\sigma_y^2+\sigma_o^2}\Big)\,G_1\Big(y,\sqrt{\sigma_y^2+\tilde{\sigma}_o^2}\Big)\,dy$$
$$\left.\int G_1\Big(z,\sqrt{\sigma_z^2+\sigma_o^2}\Big)\,G_1\Big(z,\sqrt{\sigma_z^2+\tilde{\sigma}_o^2}\Big)\,dz\right)\Bigg|_{\tilde{\sigma}_o=\sigma_o} \tag{52}$$

$$= \frac{\eta S_o^2}{4\pi}\frac{d}{d\sigma_o}\frac{d}{d\tilde{\sigma}_o}\left(\int\frac{G_1\Big(x,\sqrt{\sigma_t^2+\sigma_o^2}\Big)\,G_1\Big(x,\sqrt{\sigma_t^2+\tilde{\sigma}_o^2}\Big)}{B(x)\otimes G_1(x,\sigma_t)}dx\frac{1}{\sqrt{\sigma_y^2+\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2}}\sqrt{\sigma_z^2+\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2}}}\right)\Bigg|_{\tilde{\sigma}_o=\sigma_o}$$

Because σ_o and $\tilde{\sigma}_o$ play the same role, taking the derivatives of this product produces three terms, which can be written as

$$\begin{aligned}\mathrm{SNR}^2 =& \frac{\eta S_o^2}{4\pi}\int \frac{\mathrm{dG}_1^2\Big(x,\sqrt{\sigma_t^2+\sigma_o^2}\Big)}{B(x)\otimes G_1(x,\sigma_t)}dx\cdot \mathrm{IG}_2(\sigma_y^2,\sigma_z^2,\sigma_o^2)\\ &-\frac{\eta S_o^2}{4\pi}\int \frac{G_1\Big(x,\sqrt{\sigma_t^2+\sigma_o^2}\Big)\,\mathrm{dG}_1\Big(x,\sqrt{\sigma_t^2+\sigma_o^2}\Big)}{B(x)\otimes G_1(x,\sigma_t)}dx\cdot \mathrm{dIG}_2(\sigma_y^2,\sigma_z^2,\sigma_o^2)\\ &+\frac{\eta S_o^2}{8\pi}\int \frac{G_1^2\Big(x,\sqrt{\sigma_t^2+\sigma_o^2}\Big)}{B(x)\otimes G_1(x,\sigma_t)}dx\cdot \mathrm{ddIG}_2(\sigma_y^2,\sigma_z^2,\sigma_o^2)\end{aligned} \tag{53}$$

where we introduced the following functions:

$$\begin{aligned}\mathrm{dG}_1\left(x,\sqrt{\sigma_t^2+\sigma_o^2}\right) &= \frac{d}{d\sigma_o}G_1\left(x,\sqrt{\sigma_t^2+\sigma_o^2}\right) = \frac{\partial}{\partial\sigma}G_1\left(x,\sqrt{\sigma_t^2+\sigma_o^2}\right)\frac{d}{d\sigma_o}\sqrt{\sigma_t^2+\sigma_o^2}\\ &= \left(\frac{x^2}{(\sigma_t^2+\sigma_o^2)^{3/2}}-\frac{1}{\sqrt{\sigma_t^2+\sigma_o^2}}\right)G_1\left(x,\sqrt{\sigma_t^2+\sigma_o^2}\right)\frac{\sigma_o}{\sqrt{\sigma_t^2+\sigma_o^2}}\end{aligned} \tag{54}$$

$$\mathrm{IG}_2(v_y,v_z,v_o) = \frac{1}{\sqrt{v_y+v_o}\sqrt{v_z+v_o}} \tag{55}$$

$$\mathrm{dIG}_2(v_y,v_z,v_o) = \frac{\sqrt{v_o}}{\sqrt{v_y+v_o}\sqrt{v_z+v_0}}\left(\frac{1}{v_y+v_o}+\frac{1}{v_z+v_o}\right) \tag{56}$$

$$\mathrm{ddIG}_2(v_y,v_z,v_o) = \frac{v_o}{\sqrt{v_y+v_o}\sqrt{v_z+v_0}}\left(\frac{1}{2}\left(\frac{1}{v_y+v_o}+\frac{1}{v_z+v_o}\right)^2+\frac{1}{(v_y+v_o)^2}+\frac{1}{(v_z+v_o)^2}\right) \tag{57}$$

Note that function ddIG_2 can be obtained from function Φ defined in (18) by dropping all factors and terms involving TOF (i.e. v_t). Assuming that the object is large compared to the detector resolutions, but not compared to the TOF resolution, breaks the symmetry between the two.

If $B(x)$ is a profile through a uniform phantom with support in the interval $[-D/2, D/2]$, then the convolution of this background activity with a Gaussian TOF-kernel equals

$$B(x)\otimes G_1(x,\sigma_t) = \frac{B}{2}\left(\mathrm{erf}\left(\frac{D/2-x}{\sqrt{2}\sigma_t}\right)-\mathrm{erf}\left(\frac{-D/2-x}{\sqrt{2}\sigma_t}\right)\right) \tag{58}$$

Note that this assumption is not mandatory, (53) is also valid for a non-uniform background $B(x)$.

The three integrals in (53) must be evaluated numerically, which can be done efficiently because they are one-dimensional integrals of smooth functions.

Equation (53) is very general and can be readily extended to account also for the presence of random coincidences (randoms) and coincidences involving Compton scattered photons (scatters). This is done by simply replacing $B(x)$ with $P(x) = B(x)+R(x)+S(x)$, where $R(x)$ denotes the randoms profile and $S(x)$ the scatter profile. This implies the assumption that the detector resolution is good, such that the variation of the randoms and scatter contributions can be treated as independent of y and z. It is usually valid to assume that the randoms

$R(x)$ are independent of the TOF-index, and therefore $R(x) = R$. It follows that equation (53) can be extended to account for randoms by replacing everywhere $B(x) \otimes G_1(x, \sigma_t)$ with $B(x) \otimes G_1(x, \sigma_t) + R$. As also discussed in (Conti, 2005; Nuyts et al., 2022), the object profile $B(x)$ has a finite extent along the LOR, whereas the randoms profile R is unlimited. This is why the randoms do more damage when the TOF-kernel is wide compared to the object size, and in particular in non-TOF PET, where the TOF-kernel coincides with the coincidence window.

9.C The SNR for a TOF-kernel consisting of a sum of Gaussians

It is assumed that the TOF-kernel $k(x)$ consists of a sum of n possibly shifted Gaussians:

$$G_1(x, \sigma, a) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} \tag{59}$$

$$k(x) = \frac{1}{\sum_{i=1}^{n} A_i} \sum_{i=1}^{n} A_i \, G_1(x, \sigma_i, x_i) \tag{60}$$

This definition ensures that the integral of $k(x)$ equals unity. After convolving this TOF-kernel with a Gaussian with standard deviation σ_o, it becomes:

$$k_o(x) = \frac{1}{\sum_{i=1}^{n} A_i} \sum_{i=1}^{n} A_i \, G_1\left(x, \sqrt{\sigma_i^2 + \sigma_o^2}, x_i\right) \tag{61}$$

For the computations, it is useful to know that

$$G_1(x, \sigma_a, a) G_1(x, \sigma_b, b) = \frac{e^{-\frac{(a-b)^2}{2(\sigma_a^2+\sigma_b^2)}}}{\sqrt{2\pi}\sqrt{\sigma_a^2 + \sigma_b^2}} G_1\left(x, \frac{\sigma_a \sigma_b}{\sqrt{\sigma_a^2 + \sigma_b^2}}, \frac{\sigma_a^2 b + \sigma_b^2 a}{\sigma_a^2 + \sigma_b^2}\right) \tag{62}$$

Replacing the Gaussian TOF-kernel with $k_o(x)$ in equation (49) and integrating over the Gaussian detector resolution kernels results in

$$\mathrm{SNR}^2 = \frac{\eta S_o^2}{4\pi B} \frac{d}{\sigma_o} \frac{d}{\tilde{\sigma}_o} \int k_o(x) k_{\tilde{o}}(x) dx \cdot \frac{1}{\sqrt{\sigma_y^2 + \frac{\sigma_o^2 + \tilde{\sigma}_o^2}{2}}} \frac{1}{\sqrt{\sigma_z^2 + \frac{\sigma_o^2 + \tilde{\sigma}_o^2}{2}}} \Bigg|_{\tilde{\sigma}_o = \sigma_o} \tag{63}$$

$$k_o(x) k_{\tilde{o}}(x) = \frac{1}{\sqrt{2\pi}(\sum_i A_i)^2} \left(\sum_{i=1}^{n} \frac{A_i^2 \, G_1(x, ..)}{\sqrt{2\sigma_i^2 + \sigma_o^2 + \tilde{\sigma}_o^2}} + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2 A_i A_j e^{\frac{-(x_i - x_j)^2}{2(\sigma_i^2 + \sigma_j^2 + \sigma_o^2 + \tilde{\sigma}_o^2)}}}{\sqrt{\sigma_i^2 + \sigma_j^2 + \sigma_o^2 + \tilde{\sigma}_o^2}} G_1(x, ..) \right) \tag{64}$$

In the last equations we omitted the arguments of the Gaussians to improve the readability, they are given in (62). The integral of $k_o(x) k_{\tilde{o}}(x)$ is obtained by replacing the two Gaussians $G_1(x, ..)$ with 1, since they integrate to unity. This integral is inserted in (63) and the derivatives are computed; the computations are straightforward and produce equation (20). However, since they are also somewhat lengthy, a few intermediate results are given here. The calculations involve derivatives of the function ψ which is defined as

$$\psi(v_o, v_t, v_y, v_z) = \frac{1}{\sqrt{v_t + v_o}\sqrt{v_y + v_o}\sqrt{v_z + v_o}} \tag{65}$$

The derivatives are

$$\frac{d}{d\tilde{\sigma}_o}\psi\left(\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2},\sigma_t^2,\sigma_y^2,\sigma_z^2\right)$$
$$= -\frac{\tilde{\sigma}_o}{2}\ \psi\left(\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2},\sigma_t^2,\sigma_y^2,\sigma_z^2\right)\left(\frac{1}{\sigma_t^2+\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2}}+\frac{1}{\sigma_y^2+\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2}}+\frac{1}{\sigma_z^2+\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2}}\right) \tag{66}$$

$$\frac{d^2}{d\sigma_o d\tilde{\sigma}_o}\psi\left(\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2},\sigma_t^2,\sigma_y^2,\sigma_z^2\right)\Big|_{\tilde{\sigma}_o=\sigma_o} = \frac{1}{2}\Phi(\sigma_o^2,\sigma_t^2,\sigma_y^2,\sigma_z^2) \tag{67}$$

where Φ was defined in (18).
The computations also use the derivatives of ψ multiplied with an exponent:

$$\frac{d}{d\tilde{\sigma}_o}e^{-\frac{a^2}{2(2\sigma_t^2+\sigma_o^2+\tilde{\sigma}_o^2)}}\psi() = e^{-\frac{a^2}{2(2\sigma_t^2+\sigma_o^2+\tilde{\sigma}_o^2)}}\left(\frac{a^2\tilde{\sigma}_o}{(2\sigma_t^2+\sigma_o^2+\tilde{\sigma}_o^2)^2}\ \psi() + \frac{d\psi()}{d\tilde{\sigma}_o}\right) \tag{68}$$

$$\frac{d^2}{d\sigma_o d\tilde{\sigma}_o}e^{-\frac{a^2}{2(2\sigma_t^2+\sigma_o^2+\tilde{\sigma}_o^2)}}\psi()\Big|_{\tilde{\sigma}_o=\sigma_o}$$
$$= e^{-\frac{a^2}{4(\sigma_t^2+\sigma_o^2)}}\left(\psi()a^2\sigma_o^2\left(\frac{a^2}{16(\sigma_t^2+\sigma_o^2)^4}-\frac{1}{2(\sigma_t^2+\sigma_o^2)^3}\right.\right.$$
$$\left.\left.-\frac{1}{4(\sigma_t^2+\sigma_o^2)^2}\left(\frac{1}{\sigma_t^2+\sigma_o^2}+\frac{1}{\sigma_y^2+\sigma_o^2}+\frac{1}{\sigma_z^2+\sigma_o^2}\right)\right)+\frac{1}{2}\Phi()\right) \tag{69}$$

where the omitted arguments of ψ and Φ were $\left(\frac{\sigma_o^2+\tilde{\sigma}_o^2}{2},\sigma_t^2,\sigma_y^2,\sigma_z^2\right)$, with $\tilde{\sigma}_o$ replaced by σ_o in the right hand side of the second equation.

10 Appendices about the variance of the reconstructed image values

10.A Optimal variance of 3D TOF-PET activity reconstruction

The measured TOF-PET data can be written as equation (22), repeated here:

$$p(\hat{u},\vec{s}) = \int_{\mathbb{R}^3} f(\vec{r})h(\hat{u},\vec{s}-\vec{r})d\vec{r},\quad \hat{u}\in\Omega,\vec{s}\in\mathbb{R}^3 \tag{70}$$

where $f(\vec{r})$ is the activity distribution, $h(\hat{u},\vec{r})$ the system response for an LOR parallel to unit vector $\hat{u}$, and Ω the solid angle covered by the PET detectors. A 3D Fourier transform at fixed $\hat{u}$ produces

$$P(\hat{u},\vec{\nu}) = F(\vec{\nu})H(\hat{u},\vec{\nu}),\quad \hat{u}\in\Omega,\vec{\nu}\in\mathbb{R}^3 \tag{71}$$

Fourier transformed functions are denoted with an upper case. Equation (71) assumes that for every point $\vec{r}$ in the image, all LORs in Ω that cross the object have been measured. This is not the case for a realistic cylindrical PET system, but it is a reasonable approximation when focusing on the center of the field of view.

The reconstruction operator computes the reconstruction value in every point $\vec{r}$ as a weighted integral of the measurements:

$$f_{\text{recon}}(\vec{r}) = \int_{\Omega} d\hat{u} \int_{\mathbb{R}^n} \mathscr{R}(\vec{r}, \hat{u}, \vec{s})\, p(\hat{u}, \vec{s})\, d\vec{s}, \quad \vec{r} \in \mathbb{R}^n \tag{72}$$

where $\mathscr{R}$ is the reconstruction kernel, which has dimension length^{-n}, and n is 2 for 2D and 3 for 3D images. We consider here reconstruction by filtered backprojection, using filter kernel $w(\hat{u}, \vec{s})$ and backprojection weight $h_b(\hat{u}, \vec{s})$, such that reconstruction can be rewritten as

$$p^F(\hat{u}, \vec{s}) = \int_{\mathbb{R}^3} p(\hat{u}, \vec{s'}) w(\hat{u}, \vec{s} - \vec{s'}) d\vec{s'}, \quad \hat{u} \in \Omega, \vec{s} \in \mathbb{R}^3 \tag{73}$$

$$f_{\text{recon}}(\vec{r}) = \int_{\Omega} d\hat{u} \int_{\mathbb{R}^3} p^F(\hat{u}, \vec{s}) h_b(\hat{u}, \vec{r} - \vec{s}) d\vec{s}. \tag{74}$$

Combining (70), (73) in (74) and applying the 3D Fourier transform gives:

$$F_{\text{recon}}(\vec{\nu}) = F(\vec{\nu}) \int_{\Omega} W(\hat{u}, \vec{\nu}) H_b(\hat{u}, \vec{\nu}) H(\hat{u}, \vec{\nu}) d\hat{u} \tag{75}$$

As mentioned above, we impose a predefined reconstruction PSF $r(\vec{r})$, or equivalently, modulation transfer function (MTF) $R(\vec{\nu})$. Therefore, filter W must satisfy for all $\vec{\nu} \in \mathbb{R}^3$ the filter equation

$$R(\vec{\nu}) = \int_{\Omega} W(\hat{u}, \vec{\nu}) H_b(\hat{u}, \vec{\nu}) H(\hat{u}, \vec{\nu}) d\hat{u} \tag{76}$$

A particular form of this equation, for perfect reconstruction resolution $R(\vec{\nu}) = 1$ and perfect detector resolution in $H(\hat{u}, \vec{\nu})$, is given by equation (17) in (Mallon & Grangeat, 1992).
Now the backprojection weight H_b and filter W must be designed to minimize the variance in the reconstructed image. Assuming uncorrelated measurement noise, the variance in the reconstructed image is

$$\text{Var}(f_{\text{recon}}(\vec{r})) = \int_{\Omega} \int_{\mathbb{R}^3} \text{Var}(p(\hat{u}, \vec{s}))\, |(w(\hat{u}, .) \otimes h_b(\hat{u}, .))(\vec{r} - \vec{s})|^2 d\vec{s}\, d\hat{u} \tag{77}$$

where $\text{Var}(p(\hat{u}, \vec{s}))$ is the variance of the data, caused by Poisson noise and possible corrections for attenuation, detector sensitivity etc.
To proceed we introduce **Approximation 1**: the convolution kernel $|(w(\hat{u}, .) \otimes h_b(\hat{u}, .))(\vec{s})|^2$ is very strongly peaked around $\vec{s} = 0$, and therefore, the data variance can be approximated by a constant, equal to the variance at the center of the kernel. Then, using Parseval's relation we have

$$\text{Var}(f_{\text{recon}}(\vec{r})) \simeq \int_{\Omega} \text{Var}(p(\hat{u}, \vec{r})) \left(\int_{\mathbb{R}^3} |W(\hat{u}, \vec{\nu})|^2 |H_b(\hat{u}, \vec{\nu})|^2 \, d\vec{\nu} \right) d\hat{u} \tag{78}$$

A further simplification is obtained by **Approximation 2**: $\text{Var}(p(\hat{u}, \vec{r})) \simeq \text{Var}(p(\vec{r}))$ is independent of the direction $\hat{u}$. This would be a good approximation for the center of a uniform

sphere [7]. Then,

$$\begin{aligned}\mathrm{Var}(f_{\mathrm{recon}}(\vec{r})) &\simeq \mathrm{Var}(p(\vec{r})) \int_{\mathbb{R}^3}\int_{\Omega} |W(\hat{u},\vec{\nu})H_b(\hat{u},\vec{\nu})|^2 \; d\hat{u}\; d\vec{\nu} \\ &= \mathrm{Var}(p(\vec{r})) \int_{\mathbb{R}^3}\int_{\Omega} |M(\hat{u},\vec{\nu})|^2 \; d\hat{u}\; d\vec{\nu} \\ &= \mathrm{Var}(p(\vec{r}))\; Q(\vec{r})\end{aligned} \tag{79}$$

where we defined the product $M(\hat{u},\vec{\nu}) = W(\hat{u},\vec{\nu})\; H_b(\hat{u},\vec{\nu})$.
We now seek the backprojection weight H_b and filter W, which minimize the image variance for a given prescribed resolution $R(\vec{\nu})$, as stated by the filter equation (76). Since (76) must be satisfied separately for each $\vec{\nu}$, we introduce a Lagrange multiplier $\zeta(\vec{\nu})$. The product $M(\hat{u},\vec{\nu})$ is then calculated by minimizing

$$\Psi(M) = \int_{\mathbb{R}^3}\int_{\Omega} |M(\hat{u},\vec{\nu})|^2 d\hat{u}\; d\vec{\nu} - \int_{\mathbb{R}^3} \zeta(\vec{\nu}) \left(\int_{\Omega} M(\hat{u},\vec{\nu})\; H(\hat{u},\vec{\nu})\; d\hat{u} - R(\vec{\nu}) \right) d\vec{\nu} \tag{80}$$

Taking the variation with respect to M, the solution of $\delta\Psi(M)/\delta M(\hat{u},\vec{\nu}) = 0$ is

$$M(\hat{u},\vec{\nu}) = W(\hat{u},\vec{\nu})\; H_b(\hat{u},\vec{\nu}) = \frac{1}{2}\zeta(\vec{\nu})\; H^*(\hat{u},\vec{\nu}) \tag{81}$$

where the asterisk denotes the adjoint. Next, ζ must be determined to satisfy (76), i.e. to satisfy for each $\vec{\nu}$

$$R(\vec{\nu}) = \frac{1}{2}\int_{\Omega} \zeta(\vec{\nu})\; H^*(\hat{u},\vec{\nu})\; H(\hat{u},\vec{\nu})\; d\hat{u} \tag{82}$$

Consequently, with our assumptions, the reconstruction pixel variance (79) is minimized when

$$W(\hat{u},\vec{\nu})\; H_b(\hat{u},\vec{\nu}) = \frac{R(\vec{\nu})\; H^*(\hat{u},\vec{\nu})}{\int_{\Omega} |H(\hat{u},\vec{\nu})|^2 d\hat{u}} \tag{83}$$

a well-known result in deconvolution problems with multiple kernels (redundant data). For perfect detector and reconstruction resolution, this equation corresponds to equation (14) for confidence weighting in (Watson, 2007). Inserting this in (79), the corresponding minimum variance is

$$\mathrm{Var}(f_{\mathrm{recon}}(\vec{r})) \simeq \mathrm{Var}(p(\vec{r})) \int_{\mathbb{R}^3} \frac{R^2(\vec{\nu})}{\int_{\Omega} |H(\hat{u},\vec{\nu})|^2 d\hat{u}}\; d\vec{\nu} \tag{84}$$

where we assumed that $R^*(\vec{\nu}) = R(\vec{\nu})$. The measurement MTF $H(\hat{u},\vec{\nu})$ combines the effects of the TOF-kernel, the finite spatial resolution of the detectors and other effects like the positron range and acollinearity of the emitted photon pair (Lecomte, 2009; Moses, 2011). We treat these effects as separable, locally shift invariant and independent of the LOR direction $\hat{u}$:

$$h(\hat{u},\vec{r}) = k(\vec{r}\cdot\hat{u})\; s(\vec{r}\perp\hat{u}) \tag{85}$$

where $k(x)$ is the 1D TOF-kernel along the LORs parallel to $\hat{u}$, $\vec{r}\perp\hat{u} = \vec{r} - (\vec{r}\cdot\hat{u})\hat{u}$ denotes the projection of the 3D vector $\vec{r}$ on the plane perpendicular to $\hat{u}$, and $s(\vec{r}\perp\hat{u})$ is the 2D

[7] Writing $\mathrm{Var}(p(\vec{r}))$ is abuse of notation, but partly justified by noting that in the center of a sphere, the variance $\mathrm{Var}(p(\hat{u},\vec{r}))$ does not depend on $\hat{u}$ because the expectation of the measurement $p(\hat{u},\vec{r})$ does not depend on $\hat{u}$.

blurring kernel in the plane perpendicular to the LOR. We will use the same notation for 2D PET, in which case $\vec{r} \perp \hat{u}$ becomes a scalar. Inserting the Fourier transform of $h(\hat{u}, \vec{r})$ in (84) produces equation (24). The optimal image variance ($\mathrm{Var}(f_{\mathrm{recon}})$) does not depend on the reconstruction filter W nor on the weighted backprojection H_b, it only depends on the measurement kernel H and the imposed reconstruction PSF R.

10.B Spherical PET with full angular coverage

Inserting Gaussians for the TOF-kernel, the 2D detector blurring and the reconstruction PSF, equation (24) is rewritten as

$$\begin{aligned}\frac{\mathrm{Var}(f_{3\mathrm{D}}(\vec{r}))}{\mathrm{Var}(p_{3\mathrm{D}}(\vec{r}))} &\simeq \int_0^\infty 4\pi\nu^2 d\nu \frac{e^{-4\pi^2\nu^2\sigma^2_{\mathrm{recon}}}}{\int_0^\pi 2\pi \sin\theta \; d\theta \; e^{-4\pi^2\nu^2\sigma_t^2\cos^2\theta - 4\pi^2\nu^2\sigma^2_{\mathrm{sys}}\sin^2\theta}} \\ &= 2\int_0^\infty \nu^2 d\nu \frac{e^{-4\pi^2\nu^2\sigma^2_{\mathrm{recon}}}}{\int_{-1}^1 d(\cos\theta) e^{-4\pi^2(\sigma_t^2\nu^2\cos^2\theta + \sigma^2_{\mathrm{sys}}(\nu^2 - \nu^2\cos^2\theta))}} \\ &= \int_0^\infty \nu^3 d\nu \frac{e^{-4\pi^2\nu^2\sigma^2_{\mathrm{recon}}}}{\int_0^\nu d\xi \; e^{-4\pi^2(\sigma_t^2\xi^2 + \sigma^2_{\mathrm{sys}}(\nu^2-\xi^2))}} \end{aligned} \tag{86}$$

where the last equation (reproducing (25)) was obtained by substituting $\xi = \nu\cos\theta$. Note that for the integration over Ω, we used the full solid angle of 4π, counting every LOR twice. In eq. (41) it is stated how to compute $\mathrm{Var}(p_{3\mathrm{D}}(\vec{r}))$ to obtain quantitative predictions for actual TOF-PET experiments.
In the special case where $\sigma_t = \sigma_{\mathrm{sys}} < \sigma_{\mathrm{recon}}$, equation (25) reduces to [8]

$$\begin{aligned}\left.\frac{\mathrm{Var}(f_{3\mathrm{D}}(\vec{r}))}{\mathrm{Var}(p_{3\mathrm{D}}(\vec{r}))}\right|_{\sigma_t=\sigma_{\mathrm{sys}}} &\simeq \int_0^\infty \nu^3 d\nu \frac{e^{-4\pi^2\nu^2\sigma^2_{\mathrm{recon}}}}{\int_0^\nu d\xi \; e^{-4\pi^2\sigma^2_{\mathrm{sys}}\nu^2}} \quad = \quad \int_0^\infty \nu^2 d\nu \; e^{-4\pi^2\nu^2(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})} \\ &= \frac{\sqrt{\pi}}{32\pi^3} \frac{1}{(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})^{3/2}}. \end{aligned} \tag{87}$$

For the other cases, we have no closed form solutions for the integrals; numerical integration will be used to evaluate them.
In case $\sigma_{\mathrm{sys}} < \sigma_{\mathrm{recon}}$ and $\sigma_{\mathrm{sys}} < \sigma_t$, one can continue from (25) as follows:

$$\begin{aligned}\left.\frac{\mathrm{Var}(f_{3\mathrm{D}}(\vec{r}))}{\mathrm{Var}(p_{3\mathrm{D}}(\vec{r}))}\right|_{\sigma_{\mathrm{sys}}<\sigma_{\mathrm{recon}},\sigma_t} &\simeq \int_0^\infty \nu^3 d\nu \frac{e^{-4\pi^2\nu^2(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})}}{\int_0^\nu d\xi \; e^{-4\pi^2(\sigma_t^2 - \sigma^2_{\mathrm{sys}})\xi^2}} \\ &= 4\sqrt{\pi}\sqrt{\sigma_t^2 - \sigma^2_{\mathrm{sys}}} \int_0^\infty \nu^3 d\nu \frac{e^{-4\pi^2\nu^2(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})}}{\mathrm{erf}\left(2\pi\nu\sqrt{(\sigma_t^2 - \sigma^2_{\mathrm{sys}})}\right)} \end{aligned} \tag{88}$$

In case $\sigma_t < \sigma_{\mathrm{recon}}$ and $\sigma_t < \sigma_{\mathrm{sys}}$, one could use erfi instead of erf, but that was found to be numerically unstable. Instead we used:

$$\left.\frac{\mathrm{Var}(f_{3\mathrm{D}}(\vec{r}))}{\mathrm{Var}(p_{3\mathrm{D}}(\vec{r}))}\right|_{\sigma_t<\sigma_{\mathrm{sys}},\sigma_{\mathrm{recon}}} \simeq \int_0^\infty \nu^3 d\nu \frac{e^{-4\pi^2\nu^2(\sigma^2_{\mathrm{recon}} - \sigma_t^2)}}{\int_0^\nu d\xi \; e^{-4\pi^2(\sigma^2_{\mathrm{sys}} - \sigma_t^2)(\nu^2-\xi^2)}} \tag{89}$$

[8] We used $\int_0^\infty x^2 e^{-ax^2} dx = -\frac{d}{da}\int_0^\infty e^{-ax^2} dx = -\frac{d}{da}\frac{1}{2}\sqrt{\frac{\pi}{a}} = \frac{\sqrt{\pi}}{4}\frac{1}{a^{3/2}}$.

If $\sigma_t \gg \sigma_{\text{recon}}$ and $\sigma_t \gg \sigma_{\text{sys}}$, then a good approximation can be obtained by replacing the integration to ν by integration to ∞ in the denominator of (88). This is justified because for large σ_t, the integrand in the denominator decreases much faster than the numerator. As a result, the denominator reaches its maximum value very quickly as a function of ν, so the error caused by replacing the denominator by its maximum value everywhere is small. This produces without further approximations

$$\begin{aligned}\left.\frac{\text{Var}(f_{\text{3D}}(\vec{r}))}{\text{Var}(p_{\text{3D}}(\vec{r}))}\right|_{\sigma_t \gg \sigma_{\text{sys}},\sigma_{\text{recon}}} &\simeq \frac{\sqrt{\pi}}{8\pi^4}\frac{\sqrt{\sigma_t^2-\sigma_{\text{sys}}^2}}{(\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2)^2} \\ &\simeq \frac{\sqrt{\pi}}{8\pi^4}\frac{\sigma_t}{(\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2)^2}\end{aligned} \tag{90}$$

predicting that the variance is proportional to the width of the TOF-kernel, and inversely proportional to the fourth power of the width of the post-reconstruction blurring kernel. Figure 14 shows which equations are applied to compute the variance ratio, as a function of $\sigma_t/\sigma_{\text{recon}}$ and $\sigma_{\text{sys}}/\sigma_{\text{recon}}$.

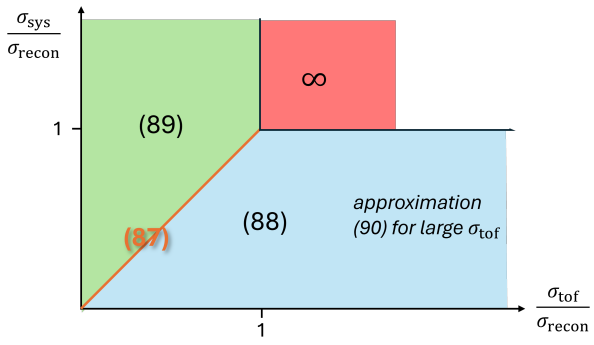

Figure 14: *The plot shows which equations are used to compute the variance ratio as a function of* $\sigma_t/\sigma_{\text{recon}}$ *and* $\sigma_{\text{sys}}/\sigma_{\text{recon}}$. *See also equation (26).*

10.C Multi-ring PET with septa

Inserting the corresponding Gaussians in (24) one obtains (29) as follows:

$$\begin{aligned}\frac{\text{Var}(f_{\text{septa}}(\vec{r}))}{\text{Var}(p_{\text{septa}}(\vec{r}))} &\simeq \int_0^\infty 2\pi\nu_r d\nu_r \int_{-\infty}^\infty d\nu_z \frac{e^{-4\pi^2\sigma_{\text{recon}}^2(\nu_r^2+\nu_z^2)}}{\int_0^{2\pi} d\theta e^{-4\pi^2(\nu_r^2\sigma_t^2\cos^2\theta+\nu_r^2\sigma_{\text{sys}}^2\sin^2\theta+\nu_z^2\sigma_{\text{sys}}^2)}} \\ &= 2\pi\int_0^\infty \nu_r d\nu_r \frac{e^{-4\pi^2\sigma_{\text{recon}}^2\nu_r^2}}{\int_0^{2\pi} d\theta e^{-4\pi^2\nu_r^2(\sigma_t^2\cos^2\theta+\sigma_{\text{sys}}^2\sin^2\theta)}} \int_{-\infty}^\infty d\nu_z e^{-4\pi^2(\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2)\nu_z^2} \\ &= \frac{\sqrt{\pi}}{\sqrt{\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2}}\int_0^\infty \nu_r d\nu_r \frac{e^{-4\pi^2\sigma_{\text{recon}}^2\nu_r^2}}{\int_0^{2\pi} d\theta e^{-4\pi^2\nu_r^2(\sigma_t^2\cos^2\theta+\sigma_{\text{sys}}^2\sin^2\theta)}}\end{aligned} \tag{91}$$

Setting $\sigma_{\text{sys}} = \sigma_t < \sigma_{\text{recon}}$, (91) reduces to

$$\left.\frac{\text{Var}(f_{\text{septa}}(\vec{r}))}{\text{Var}(p_{\text{septa}}(\vec{r}))}\right|_{\sigma_t=\sigma_{\text{sys}}} \simeq \frac{\sqrt{\pi}}{\sqrt{\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2}}\int_0^\infty \nu_r d\nu_r \frac{e^{-4\pi^2(\sigma_{\text{recon}}^2-\sigma_{\text{sys}}^2)\nu_r^2}}{2\pi} \tag{92}$$

This integral produces (31).
For the case of moderate TOF resolution, we rewrite equation (91) as

$$\frac{\mathrm{Var}(f_{\mathrm{septa}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{septa}}(\vec{r}))} \simeq \frac{\sqrt{\pi}}{\sqrt{\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}}}} \int_0^\infty \nu_r d\nu_r \frac{e^{-4\pi^2(\sigma^2_{\mathrm{recon}}-\sigma^2_{\mathrm{sys}})\nu_r^2}}{2\int_0^\pi d\theta e^{-4\pi^2\nu_r^2(\sigma_t^2-\sigma^2_{\mathrm{sys}})\cos^2\theta}}. \tag{93}$$

The integrand in the denominator is a Gaussian, which will only differ significantly from zero if its argument is small. Assuming that ν_r is large and $\sigma_t \gg \sigma_{\mathrm{recon}} > \sigma_{\mathrm{sys}}$, this is only the case if $\cos\theta$ is close to zero. Near that zero crossing, we approximate $\cos\theta$ with the linear function $\cos\theta \simeq \theta - \pi/2$, as was also done in (Nuyts et al., 2022). In addition, since for large values of $|\theta|$ the integrand vanishes, the integration range can be extended to $[-\infty, \infty]$. This produces

$$\begin{aligned}\left.\frac{\mathrm{Var}(f_{\mathrm{septa}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{septa}}(\vec{r}))}\right|_{\sigma_t \gg \sigma_{\mathrm{sys}},\sigma_{\mathrm{recon}}} &\simeq \frac{\sqrt{\pi}}{\sqrt{\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}}}} \int_0^\infty \nu_r d\nu_r \frac{e^{-4\pi^2(\sigma^2_{\mathrm{recon}}-\sigma^2_{\mathrm{sys}})\nu_r^2}}{2\int_{-\infty}^\infty d\theta e^{-4\pi^2\nu_r^2(\sigma_t^2-\sigma^2_{\mathrm{sys}})\theta^2}} \\ &= \frac{\sqrt{\pi}}{32\pi^2} \frac{\sqrt{\sigma_t^2 - \sigma^2_{\mathrm{sys}}}}{(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})^2} \simeq \frac{\sqrt{\pi}}{32\pi^2} \frac{\sigma_t}{(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})^2}\end{aligned} \tag{94}$$

10.D Circular 2D PET

Inserting Gaussians in (24) and doing away with the z dimension results in

$$\begin{aligned}\frac{\mathrm{Var}(f_{\mathrm{2D}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{2D}}(\vec{r}))} &\simeq \int_0^\infty 2\pi\nu d\nu \frac{e^{-4\pi^2\nu^2\sigma^2_{\mathrm{recon}}}}{\int_0^{2\pi} d\theta e^{-4\pi^2\nu^2(\sigma_t^2\cos^2\theta + \sigma^2_{\mathrm{sys}}\sin^2\theta)}} \\ &= 2\pi \int_0^\infty \nu d\nu \frac{e^{-4\pi^2\nu^2(\sigma^2_{\mathrm{recon}}-\sigma^2_{\mathrm{sys}})}}{\int_0^{2\pi} d\theta e^{-4\pi^2\nu^2(\sigma_t^2-\sigma^2_{\mathrm{sys}})\sin^2\theta}}\end{aligned} \tag{95}$$

When $\sigma_{\mathrm{sys}} = \sigma_t < \sigma_{\mathrm{recon}}$, equation (95) reduces to

$$\begin{aligned}\left.\frac{\mathrm{Var}(f_{\mathrm{2D}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{2D}}(\vec{r}))}\right|_{\sigma_t=\sigma_{\mathrm{sys}}} &= \frac{1}{2}\int_0^\infty d(\nu^2) e^{-4\pi^2\nu^2(\sigma^2_{\mathrm{recon}}-\sigma^2_{\mathrm{sys}})} \\ &= \frac{1}{8\pi^2(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})}\end{aligned} \tag{96}$$

If $\sigma_t \gg \sigma_{\mathrm{recon}} > \sigma_{\mathrm{sys}}$, then a good approximation can be obtained in the same way as for (32), i.e. by replacing $\sin\theta$ with its linear approximation θ near zero in (95).

$$\begin{aligned}\left.\frac{\mathrm{Var}(f_{\mathrm{2D}}(\vec{r}))}{\mathrm{Var}(p_{\mathrm{2D}}(\vec{r}))}\right|_{\sigma_t \gg \sigma_{\mathrm{sys}},\sigma_{\mathrm{recon}}} &\simeq 2\pi \int_0^\infty \nu d\nu \frac{e^{-4\pi^2\nu^2(\sigma^2_{\mathrm{recon}}-\sigma^2_{\mathrm{sys}})}}{2\int_{-\infty}^\infty d\theta e^{-4\pi^2\nu^2(\sigma_t^2-\sigma^2_{\mathrm{sys}})\theta^2}} \\ &= \frac{1}{16\pi} \frac{\sqrt{\sigma_t^2 - \sigma^2_{\mathrm{sys}}}}{(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})^{3/2}} \simeq \frac{1}{16\pi} \frac{\sigma_t}{(\sigma^2_{\mathrm{recon}} - \sigma^2_{\mathrm{sys}})^{3/2}}\end{aligned} \tag{97}$$

10.E Cylindrical fully 3D PET: moderate TOF resolution and small axial aperture

A good approximation can be obtained for wide TOF-kernels and PET systems with short axial length, i.e. $\sigma_t \gg \sigma_{\text{recon}} > \sigma_{\text{sys}}$ and $|\beta| \ll \pi/2$. Inserting (38) in (37) one obtains:

$$\frac{\text{Var}(f_{\text{cyl}}(\vec{r}))}{\text{Var}(p_{\text{cyl}}(\vec{r}))} \simeq 2\pi \int_0^\infty d\nu \int_0^{\pi/2} d\varphi \frac{\nu^2 \cos\varphi \, e^{-4\pi^2(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})\nu^2}}{\int_{-\beta}^{\beta} \cos\varphi' d\varphi' \int_0^\pi d\theta' e^{-4\pi^2(\cos\varphi'\cos\varphi\cos\theta'+\sin\varphi'\sin\varphi)^2(\sigma_t^2-\sigma^2_{\text{sys}})\nu^2}}.$$

Assuming that $|\beta|$ and therefore φ' are small, we introduce the approximation that $\sin\varphi' \simeq 0$. In addition, as before, we replace $\cos\theta'$ with its linear approximation $\theta' - \pi/2$, because only small values of $\cos\theta'$ produce integrands significantly different from zero. For the same reason, the integration range can be extended from $[0,\pi]$ to $[-\infty,\infty]$. This produces an expression for which a closed form solution is obtained without further approximations:

$$\begin{aligned}\frac{\text{Var}(f_{\text{cyl}}(\vec{r}))}{\text{Var}(p_{\text{cyl}}(\vec{r}))}\bigg|_{\sigma_t \gg \sigma_{\text{recon}} > \sigma_{\text{sys}}} &\simeq 2\pi \int_0^\infty d\nu \int_0^{\pi/2} d\varphi \frac{\nu^2 \cos\varphi \, e^{-4\pi^2(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})\nu^2}}{\int_{-\beta}^{\beta} \cos\varphi' d\varphi' \int_{-\infty}^{\infty} d\theta' e^{-4\pi^2\cos^2\varphi'\cos^2\varphi\,\theta'^2(\sigma_t^2-\sigma^2_{\text{sys}})\nu^2}} \\ &= \frac{1}{2\beta}\frac{\sqrt{\pi}}{32\pi^2}\frac{\sqrt{\sigma_t^2-\sigma^2_{\text{sys}}}}{(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})^2} \simeq \frac{1}{2\beta}\frac{\sqrt{\pi}}{32\pi^2}\frac{\sigma_t}{(\sigma^2_{\text{recon}}-\sigma^2_{\text{sys}})^2} \end{aligned} \tag{98}$$

The last approximation is based on the assumption $\sigma_t \gg \sigma_{\text{sys}}$. As expected, this is identical to the corresponding equation (32) for a PET with septa, except for the factor $1/(2\beta)$. The equation is also similar to equation (28) for a complete spherical PET system. Setting β to $\pi/2$ (although this violates the assumption that axial acceptance angle β is small), equation (98) predicts a variance which is smaller by a factor $\pi/4 \simeq 0.785$ than that of the more accurate equation (28). For smaller β values, the error should be smaller than that.

10.F Application to discrete data

To explain the discretization, it is useful to briefly summarize what has been done so far, and repeat the derivation for the discrete formulation.
In the continuous model, the reconstruction can be written as in (72), repeated here

$$f_{\text{recon}}(\vec{r}) = \int_\Omega d\hat{u} \int_{\mathbb{R}^n} \mathscr{R}(\vec{r},\hat{u},\vec{s})\, p(\hat{u},\vec{s})\, d\vec{s}, \quad \vec{r} \in \mathbb{R}^n. \tag{99}$$

The reconstruction image $f_{\text{recon}}(\vec{r})$ has dimension length^{-n}, where n is 2 for 2D and 3 for 3D. Because we consider TOF-PET, the projection $p(\hat{u},\vec{s})$ has the same dimension as the image. Consequently, reconstruction kernel $\mathscr{R}(\vec{r},\hat{u},\vec{s})$ has dimension length^{-n} too.
Consider now discretization by sampling the image and the data. Approximating the integral with the usual trapezoidal rule, equation (99) becomes

$$f_{\text{recon}}(\vec{r}_k) \simeq f^s_{\text{recon}}(\vec{r}_k) = \Delta_s\Delta_u \sum_{i=1}^{N_{\text{proj}}} \sum_{j=1}^{J} \mathscr{R}(\vec{r}_k,\hat{u}_i,\vec{s}_j) p(\hat{u}_i,\vec{s}_j), \quad k = 1,\ldots,K \tag{100}$$

where f^s_{recon} denotes the discrete approximation, N_{proj} is the number of parallel projections and J is the number of data bins in each projection. Δ_s is the data voxel volume (or pixel

area) and has dimension length^n, and Δ_u is the discretization of the solid angle when $n = 3$, and of the angle when $n = 2$. Assuming non-correlated noise on the data samples, the variance of the reconstruction at sampled point k is

$$\mathrm{Var}(f^s_{\mathrm{recon}}(\vec{r}_k)) = \Delta_s^2 \Delta_u^2 \sum_{i=1}^{N_{\mathrm{proj}}} \sum_{j=1}^{J} \mathscr{R}^2(\vec{r}_k, \hat{u}_i, \vec{s}_j) \mathrm{Var}(p(\hat{u}_i, \vec{s}_j)) \tag{101}$$

Similar to what was done in the continuous derivation, we assume that for a central image pixel k, the variance is almost constant and can be moved out of the summations:

$$\begin{aligned} \mathrm{Var}(f^s_{\mathrm{recon}}(\vec{r}_k)) &\simeq \Delta_s^2 \Delta_u^2 \ \mathrm{Var}(p(\vec{r}_k)) \sum_{i=1}^{N_{\mathrm{proj}}} \sum_{j=1}^{J} \mathscr{R}^2(\vec{r}_k, \hat{u}_i, \vec{s}_j) \\ &\simeq \Delta_s \Delta_u \ \mathrm{Var}(p(\vec{r}_k)) \int_{\Omega} \int_{\mathbb{R}^n} \mathscr{R}^2(\vec{r}_k, \hat{u}, \vec{s}) \, d\hat{u} \, d\vec{s} \\ &= \Delta_s \Delta_u \ \mathrm{Var}(p(\vec{r}_k)) \ Q(\vec{r}_k) \end{aligned} \tag{102}$$

where Q is the same function that was introduced in (79). In practice, it is customary to work with binned image samples and binned data samples, corresponding to counts per reconstructed voxel and to counts per data bin:

$$\begin{aligned} f^b_{\mathrm{recon}}(\vec{r}_k) &= f^s_{\mathrm{recon}}(\vec{r}_k) \Delta_r \\ p^b(\hat{u}_i, \vec{s}_j) &= p(\hat{u}_i, \vec{s}_j) \Delta_s \end{aligned} \tag{103}$$

where Δ_r is the volume (or area) of the image voxel (or pixel), with dimension length^n. Note that each projection is binned, but there is no binning per angle: in our simulations we use a unit sensitivity, such that the total count in the projection equals the total activity in the image, independent of the number of projections that we simulate. Inserting this in equation (102) produces:

$$\mathrm{Var}(f^b_{\mathrm{recon}}(\vec{r}_k)) \simeq \frac{\Delta_r^2 \Delta_u}{\Delta_s} \mathrm{Var}(p^b(\vec{r}_k)) \ Q(\vec{r}_k) \tag{104}$$

Equation (104) is equivalent to (79), with the only difference that the former is for continuous images and data, whereas the latter is for binned image and data samples. It shows that this discretization procedure introduces a factor $\Delta_r^2 \Delta_u / \Delta_s$, which must be accounted for in the discrete implementation.

In our experiments, we have used isotropic pixels or voxels $\Delta_r = d^n$. Further, we used PET detectors of the same size d and a TOF bin which was T times larger than d. When simulating a total of N_{proj} projections distributed over Ω, we have $\Delta_u = \Omega / N_{\mathrm{proj}}$. Finally, because the discrete measurements are subject to Poisson noise, we have $\mathrm{Var}(p^b(\vec{r}_k)) = p^b(\vec{r}_k)$. Consequently, $\Delta_s = d^n T$, and we obtain for the variance in the discrete reconstruction:

$$\mathrm{Var}(f^b_{\mathrm{recon}}(\vec{r}_k)) \simeq p^b(\vec{r}_k) \ \frac{d^n}{T} \ \frac{\Omega}{N_{\mathrm{proj}}} \ Q(\vec{r}_k). \tag{105}$$

We chose to set the pixel sizes to $d = 1$, implying that $p^b(\hat{u}_i, \vec{s}_j) = p(\hat{u}_i, \vec{s}_j)$ in (103). Consequently, using $d^n = 1$, the analytical results can be verified quantitatively with PET simulation experiments by applying the rules given in (41).

11 References

Barrett, H. H., & Myers, K. J. (2013). *Foundations of image science*. John Wiley & Sons.

Barrett, H. H., Yao, J., Rolland, J. P., & Myers, K. J. (1993). Model observers for assessment of image quality. *Proc Natl Acad Sci USA*, *90*(21), 9758–9765.

Budinger, T. F. (1983). Time-of-flight positron emission tomography: Status relative to conventional PET. *Journal of nuclear medicine*, *24*(1), 73–78.

Conti, M. (2005). Effect of randoms on signal-to-noise-ratio in TOF PET, In *Ieee nuclear science symposium conference record, 2005*. IEEE.

Conti, M., Eriksson, L., & Westerwoudt, V. (2013). Estimating image quality for future generations of TOF PET scanners. *IEEE Transactions on Nuclear Science*, *60*(1), 87–94.

Defrise, M., Townsend, D., & Deconinck, F. (1990). Statistical noise in three-dimensional positron tomography. *Physics in Medicine & Biology*, *35*(1), 131.

Efthimiou, N., Kratochwil, N., Gundacker, S., Polesel, A., Salomoni, M., Auffray, E., & Pizzichemi, M. (2020). TOF-PET image reconstruction with multiple timing kernels applied on Cherenkov radiation in BGO. *IEEE transactions on radiation and plasma medical sciences*, *5*(5), 703–711.

Gundacker, S., Turtos, R. M., Auffray, E., Paganoni, M., & Lecoq, P. (2019). High-frequency SiPM readout advances measured coincidence time resolution limits in TOF-PET. *Phys Med Biol*, *64*(5), 055012.

Ishikawa, T., Akamatsu, G., Tashima, H., Nishikido, F., Hashimoto, F., Ota, R., Haneishi, H., Kwon, S. I., Cherry, S. R., & Yamaya, T. (2025). Imaging simulation of a dual-panel PET geometry with ultrafast TOF detectors. *arXiv preprint arXiv:2502.01006*.

Konstantinou, G., Lecoq, P., Benlloch, J., & Gonzalez, A. (2021). Metascintillators for ultrafast gamma detectors: A review of current state and future perspectives. *IEEE Trans Radiat Plasma Med Sci*, *6*(1), 5–15.

Kwon, S. I., Ota, R., Berg, E., Hashimoto, F., Nakajima, K., Ogawa, I., Tamagawa, Y., Omura, T., Hasegawa, T., & Cherry, S. R. (2021). Ultrafast timing enables reconstruction-free positron emission imaging. *Nature photonics*, *15*(12), 914–918.

Lecomte, R. (2009). Novel detector technology for clinical PET. *European journal of nuclear medicine and molecular imaging*, *36*(Suppl 1), 69–85.

Lecoq, P., Morel, C., Prior, J. O., Visvikis, D., Gundacker, S., Auffray, E., Križan, P., Turtos, R. M., Thers, D., Charbon, E., Varela, J., de La Taille, C., Rivetti, A., Breton, D., Pratte, J.-F., Nuyts, J., Surti, S., Vandenberghe, S., Marsden, P., ... Benoit, M. (2020). Roadmap toward the 10 ps time-of-flight PET challenge. *Physics in Medicine & Biology*, *65*(21), 21RM01.

Maas, M. C., Schaart, D. R., van der Laan, D. J., Bruyndonckx, P., Lemaître, C., Beekman, F. J., & van Eijk, C. W. (2009). Monolithic scintillator PET detectors with intrinsic depth-of-interaction correction. *Physics in Medicine & Biology*, *54*(7), 1893.

Maebe, J., & Vandenberghe, S. (2023). Effect of detector geometry and surface finish on Cerenkov based time estimation in monolithic BGO detectors. *Physics in Medicine & Biology*, *68*(2), 025009.

Mallon, A., & Grangeat, P. (1992). Three-dimensional PET reconstruction with time-of-flight measurement. *Physics in Medicine & Biology*, *37*(3), 717.

Matej, S., Surti, S., Jayanthi, S., Daube-Witherspoon, M. E., Lewitt, R. M., & Karp, J. S. (2009). Efficient 3-D TOF PET reconstruction using view-grouped histo-images: DIRECT–

direct image reconstruction for TOF. *IEEE Transactions on medical imaging*, *28*(5), 739–751.

Merlin, T., & Visvikis, D. (2022). Assessment of gaussian and multi-gaussian kernel models for the reconstruction of TOF-PET datasets with event-by-event variable time resolution, In *2022 ieee nuclear science symposium and medical imaging conference (nss/mic)*. IEEE.

Moses, W. W. (2011). Fundamental limits of spatial resolution in PET. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *648*, S236–S240.

Muehllehner, G. (1985). Effect of resolution improvement on required count density in ECT imaging: A computer simulation. *Physics in Medicine & Biology*, *30*(2), 163.

Nuyts, J., Defrise, M., Gundacker, S., Roncali, E., & Lecoq, P. (2022). The SNR of positron emission data with gaussian and non-gaussian time-of-flight kernels, with application to prompt photon coincidence. *IEEE Transactions on Medical Imaging*, *42*(5), 1254–1264.

Nuyts, J., Defrise, M., Morel, C., & Lecoq, P. (2023). The SNR of time-of-flight positron emission tomography data for joint reconstruction of the activity and attenuation images. *Physics in Medicine & Biology*, *69*(1), 015011.

Ota, R., Nakajima, K., Ogawa, I., Tamagawa, Y., Shimoi, H., Suyama, M., & Hasegawa, T. (2019). Coincidence time resolution of 30 ps fwhm using a pair of Cherenkov-radiator-integrated mcp-pmts. *Physics in Medicine & Biology*, *64*(7), 07LT01.

Schaefferkoetter, J., Casey, M., Townsend, D., & El Fakhri, G. (2013). Clinical impact of time-of-flight and point response modeling in PET reconstructions: A lesion detection study. *Physics in Medicine & Biology*, *58*(5), 1465.

Shultzman, A., Schütz, R., Kurman, Y., Lahav, N., Dosovitskiy, G., Roques-Carmes, C., Bekenstein, Y., Konstantinou, G., Latella, R., Zhang, L., Et al. (2024). Towards a second generation of metascintillators using the Purcell effect. *arXiv preprint arXiv:2406.15058*.

Surti, S., Shore, A. R., & Karp, J. S. (2013). Design study of a whole-body PET scanner with improved spatial and timing resolution. *IEEE transactions on nuclear science*, *60*(5), 3220–3226.

Tomitani, T. (1981). Image reconstruction and noise evaluation in photon time-of-flight assisted positron emission tomography. *IEEE Trans Nucl Science*, *NS-28*(6), 4582–4589.

Toussaint, M., Lecomte, R., & Dussault, J.-P. (2020). Improvement of spatial resolution with iterative PET reconstruction using ultrafast TOF. *IEEE transactions on radiation and plasma medical sciences*, *5*(5), 729–737.

Toussaint, M., Lecomte, R., & Dussault, J.-P. (2025). Annihilation photon acolinearity with ultra-fast ToF-PET. *IEEE Nuclear Science Symposium and Medical Imaging Conference*.

Turtos, R. M., Gundacker, S., Auffray, E., & Lecoq, P. (2019). Towards a metamaterial approach for fast timing in PET: Experimental proof-of-concept. *Phys Med Biol*, *64*(18), 185018.

Watson, C. C. (2007). An evaluation of image noise variance for time-of-flight PET. *IEEE Transactions on Nuclear Science*, *54*(5), 1639–1647.