Single-event based TOF FBP image reconstruction in J-PET

Roman Shopa

National Centre for Nuclear Research,
Świerk Computing Centre, Otwock-Świerk, Poland

3rd Jagiellonian Symposium, June 27th, 2019

Outline

– **Motivation**
- Time-of-flight resolution of J-PET
- Real time image reconstruction
- Continuous geometry (no bins)

– **Single-event TOF FBP**
- Convolution vs sum: FBP in image space
- Adding time-of-flight
- 3D asymmetric kernel
- Kernel optimisation

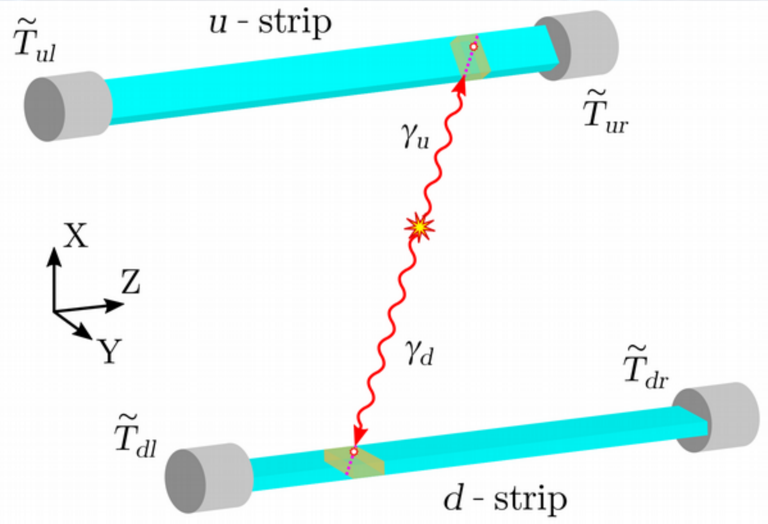
– **Results**
- Spatial resolution of “big barrel”
- Image quality for ideal J-PET

- **Summary and further plans**

Time-of-flight resolution of J-PET

In modern PET scanners, scintillation crystals (LSO:Ce, LYSO:Ce, $LaBr_3$:Ce) are capable of achieving coincidence resolving time **(CRT)** of ~ 100 ps. The lowest announced value is 249 ps, in fact **214 ps** (Biograph Vision scanner, Siemens) [ordineingegneripisa.it/obj/files/documenti/2018.2.23.10.46.26_834.pdf].

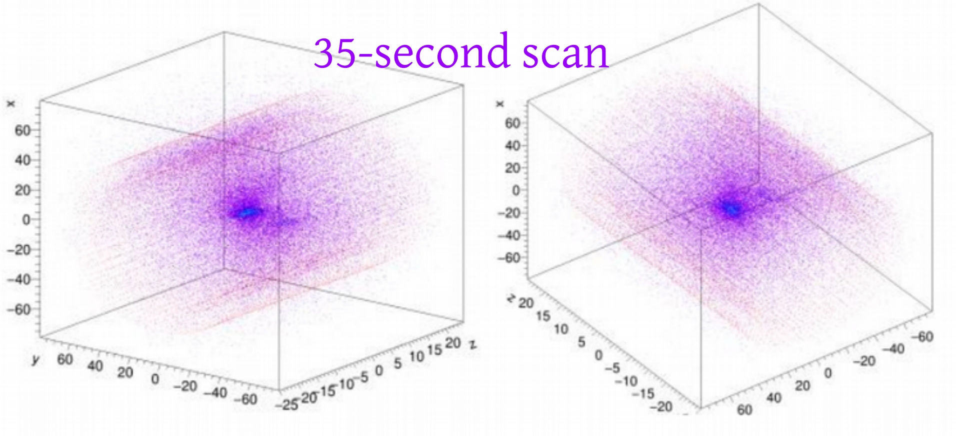
Plastic scintillators used in **Jagiellonian PET (J-PET)** are superior time-wise, despite the worse efficiency: CRT of **70 ps** – for 1-meter strips [Moskal P et al. PMB 2016]. The main factor is the readout – photomultipliers (PMs) attached at each end of the strips: silicon PM (SiPM) or tube PM (PMT).



For time-of-flight **(TOF)** available and CRT below 100 ps analytical reconstruction methods may outperform iterative ones [V Westerwoudt et al. IEEE Trans. Nucl. Sci. 2014].

Real time image reconstruction

TOF reconstructions produce comparable results for much lower statistics compared to non-TOF methods. Matched with small CRT, it substantiates image reconstruction on the fly during real time scans.



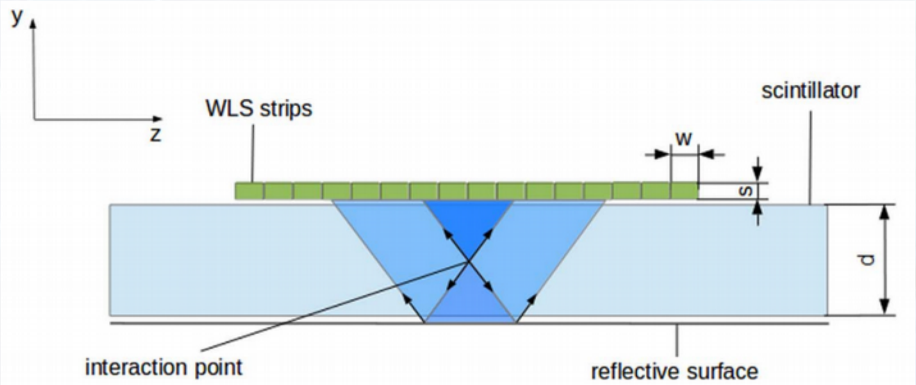
A platform based on Field Programmable Gate Array (**FPGA**) System-on-Chip (SoC) **has already been implemented** for J-PET [G Korcyl et al. IEEE Trans. Med. Im. 2018]. It performs event building, filtering, coincidence search and so-called Region-Of-Response **(ROR)** reconstruction, yet *without* filtered back projection **(FBP)** – only coordinates of the reconstructed points in 3D space. ***Newest solution (G.Korcyl presentation) – operate in projection space.***

ROR implies that *only small fraction* of field-of-view **(FOV)** is processed for each event, hence it might be possible to add Ramp/Hann/Hamming filters.

Continuous geometry

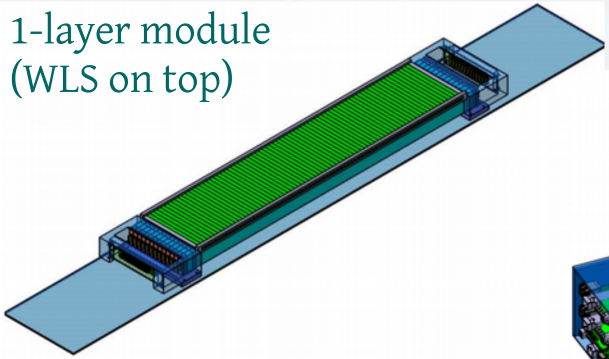
For non-TOF scanners, the measured 3D data of N detected emissions can be expressed as a set of projections: $\{\tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2, \ldots, \tilde{\mathbf{p}}_N\}, \quad \tilde{\mathbf{p}}_k = (s, \phi, \zeta, \theta)_k$.
The construction of a scanner defines discrete sets of s, ϕ, ζ, θ which form possible projection elements (bins). Adding TOF *would expand it* to $\mathbb{R}^5$.

However, *continuous strips in J-PET* do not fix ζ and θ. Besides, there are prospects for partial depth-of-interaction **(DOI)** information be extracted with the array of wavelength-shifting **(WLS)** strips [J Smyrski et al., N. Instr. Meth. Phys. Res. A 2017].

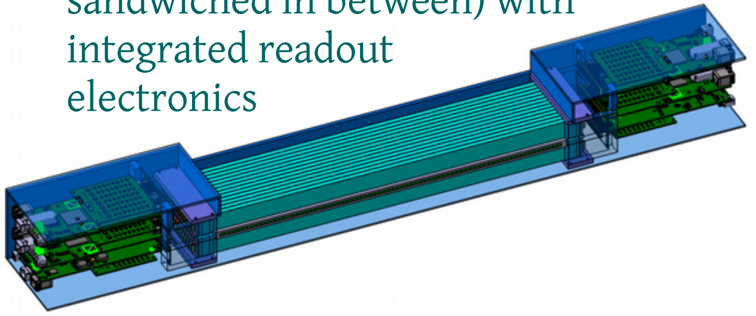


Using bins/projection space is unpractical!

1-layer module (WLS on top)



Prototype 2-layer module (WLS sandwiched in between) with integrated readout electronics



Single-event TOF FBP

Non-TOF FBP (2D): image $f(x,y)=(X^*_{R_F}\, p^F)(x,y)=\int_0^{\pi} d\phi\, p^F(s=x\cos\phi+y\sin\phi,\phi)$

Projection (sinogram): $p^F(s,\phi)=\int_{-R_F} ds'\, p(s',\phi)\, w(s-s')$

Convolution is used here only (w – Ram-Lak filter)

[DL Bailey et al., PET Basic Science, 2005]

3D TOF FBP (arbitrary voxel *v*): $f(v)=\int_\phi\int_\theta\int_\zeta\sum_{i=1}^{N}\mathcal{F}^{-1}\{W(\nu_s)\,\mathcal{F}[p_i(s,\phi,\theta,\zeta)]\}\cdot h(t-t_i)$

[Conti M et al. PMB 2005]

Forward and inverse Fourier transform, a filter in frequency domain $W(\nu_s)$, TOF kernel *h*(*t*)... ***too cumbersome!***

The alternative: *treat all lines-of-response (LORs) independently.*

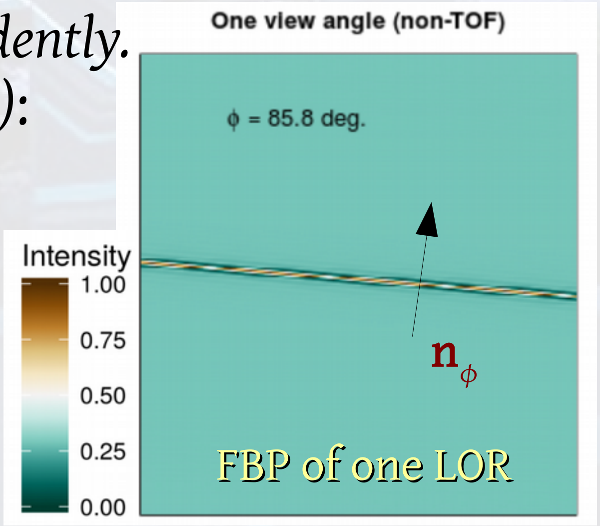
One LOR reflects ***one point*** on a sinogram, *filtered by w(s)*:

$p_i(s,\phi)=\mathbf{1}|_{s=s_i,\ \phi=\phi_i} \rightarrow p^F_i(s,\phi)=w(s-s_i^{(\phi)})$,

All points for a fixed ϕ: $p^F(s,\phi)=\Sigma_i\, p^F_i(s,\phi)$,

$f(x,y)=\Sigma_\phi\Sigma_i\,\mathbf{n}_\phi\, p^F_i(s=x\cos\phi+y\sin\phi,\phi)$,

$\mathbf{n}_\phi$ – vector that defines orientation.



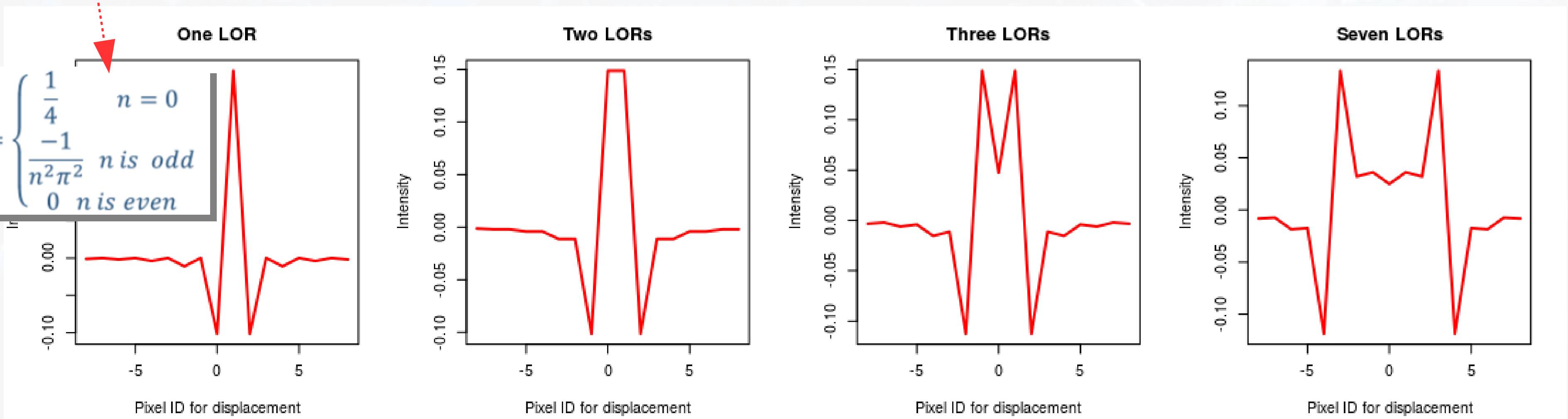
FBP in image space (2D)

Split the data into single-LOR backprojections first. We know how the reconstructed image looks like for one LOR. ***The sum of such images for all LORs will reflect the Filtered Back Projection (FBP)***:

$$f(x, y) = \Sigma_\phi \Sigma_i \mathbf{n}_\phi \, w(\, x \cos \phi + y \sin \phi - s_i^{(\phi)})$$

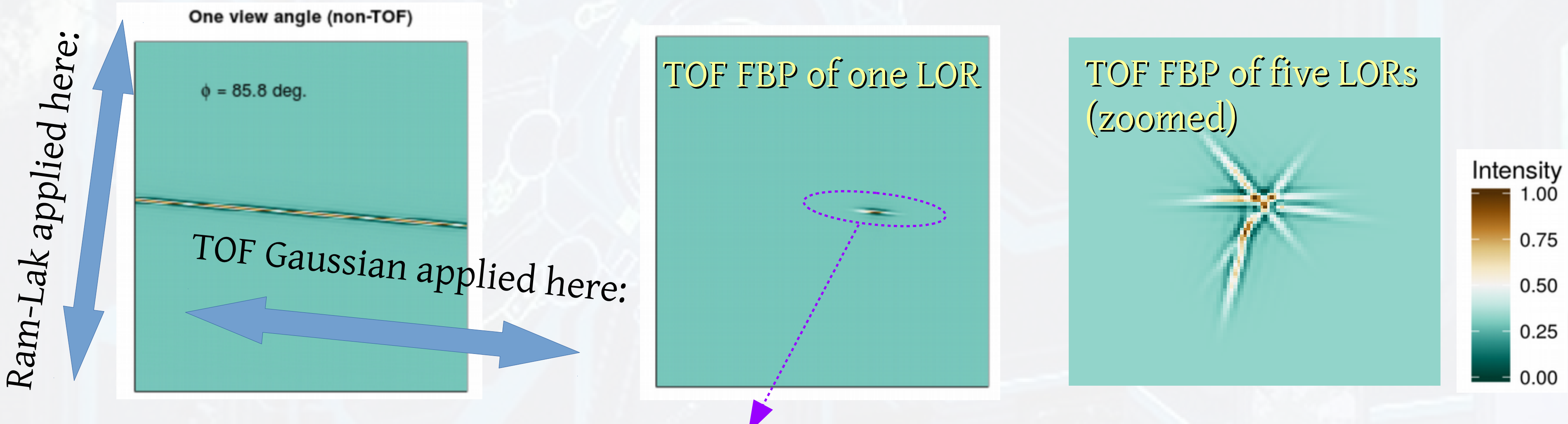
Options for filters: define analytically in image space (easy for Ram-Lak), use a table/polynome or create one image and move/rotate.

$$q(n) = \begin{cases} \frac{1}{4} & n = 0 \\ \frac{-1}{n^2\pi^2} & n \text{ is odd} \\ 0 & n \text{ is even} \end{cases}$$



Adding time-of-flight (2D)

Apply Gaussian kernel (with standard deviation σ_{TOF} calculated from coincidence resolving time, CRT), centered at the point estimated from TOF:



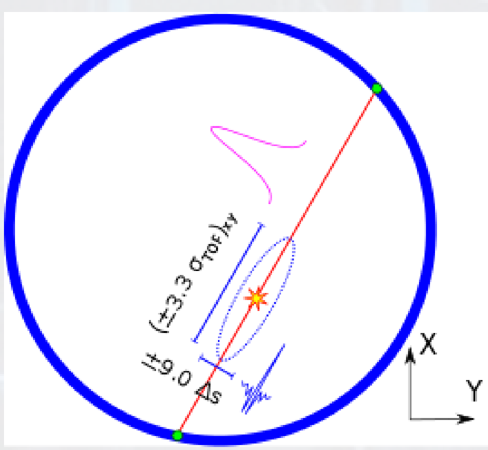
The resulting map of intensities will be ***sparse****, hence it is reasonable to restrict/truncate the known outcome – defined analytically or as an image.*

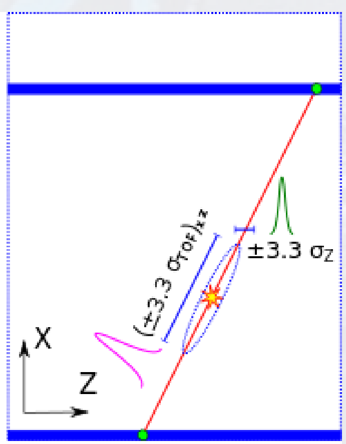
The narrower the TOF kernel is, the less memory is needed to update the result: the fastest will be the combination of SiPM/WLS, the slowest – PMT readout.
“Old” LORs ***could be dropped*** – ***a time window for real time imaging is possible!***

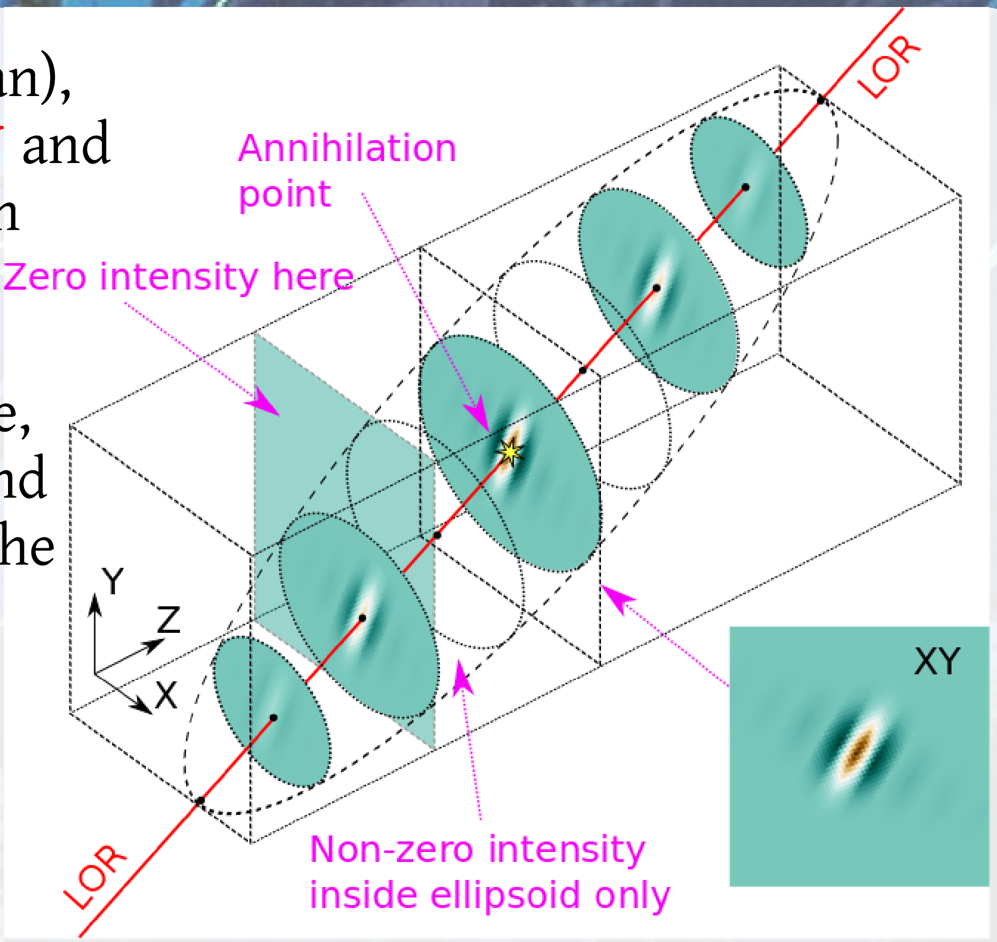
3D asymmetric kernel

Apply **TOF kernel along LOR** (Gaussian), **Ram-Lak filter normal to LOR in XY** and a **small Gaussian along Z** (depends on the distance between slices).

Update intensity within a small volume, limited by at least ±3.3σ for Gaussian and ±9.0Δ*s* for Ram-Lak (Δ*s* – sampling for the displacement *s* in projection space).







The volume of the ellipsoid is much smaller than the whole FOV, resembles ROR in FPGA solution.

Kernel optimisation

There are distinct similarities with multivariate kernel density estimation (KDE), ***applied to annihilation positions, estimated directly from TOF.***

For a d-dimensional dataset $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ of the size n, the KDE is:

$$\hat{f}_{n\boldsymbol{H}}(\boldsymbol{x}) = n^{-1} \sum_{i=1}^{n} |\boldsymbol{H}|^{-1/2} K\left[\boldsymbol{H}^{-1/2}(\boldsymbol{x} - \mathbf{X}_i)\right]$$

$\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$, $K(\cdot)$ – ***spherically symmetric*** multivariate kernel (e.g. Gaussian),

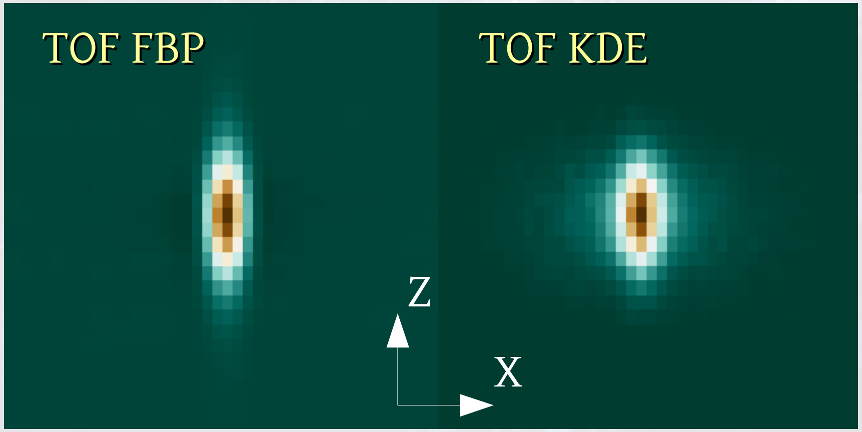
$\boldsymbol{H}$ – ***bandwidth matrix***, symmetric and positive definite. Its choice is crucial!

There are lots of algorithms for bandwidth selection, e.g.:

- asymptotic approximation mean integrated squared error (AMISE)
- plug-in bandwidth selector (multistage) [Chacon JE et al. Test 2010]

Elements of matrix $\mathbf{H} \ll \sigma_{TOF} < \sigma_z$!

It is reasonable to optimise Gaussian kernels to smaller sigmas, otherwise it imposes additional smearing along Z. (example for 1-mm source, ideal scanner, [Kowalski P et al. PMB 2018])



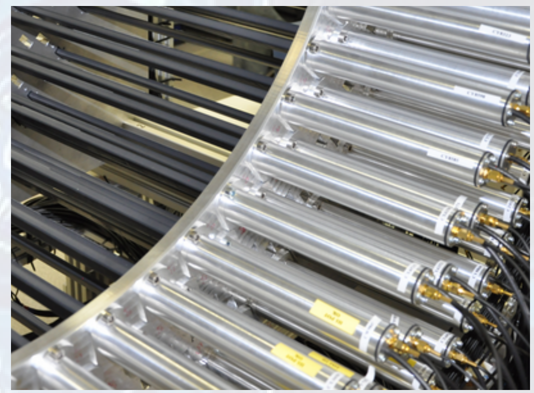
Results: spatial resolution

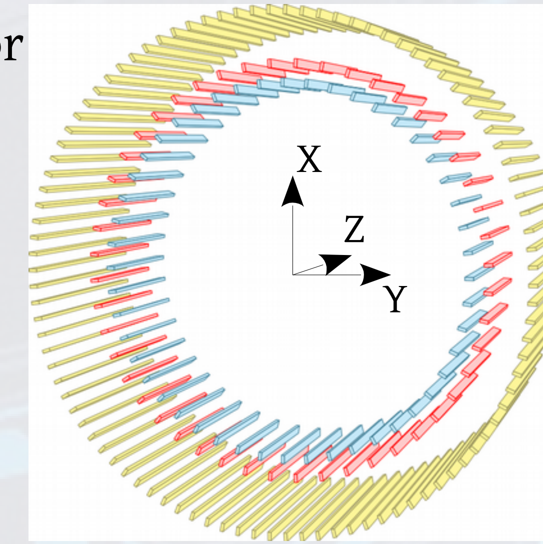
Jagiellonian PET (“big barrel”): 3 layers of plastic scintillator strips, 192 detector strips of the size: 7 mm × 19 mm × **500 mm**

Radii:

- 425.0 mm (48 strips)
- 467.5 mm (48 strips)
- 575.0 mm (96 strips)

Gaps between strips are dictated by the size of **PMT** readouts.



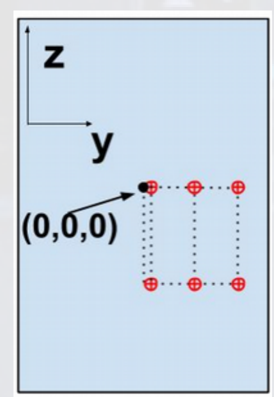


The data:

– ***simulations*** made in GATE for 1-mm spherical NEMA source (370 kBq), at six positions (x = 1 cm/10 cm/20 cm, z = 0/18.75 cm) , 100,000 coincidences per one set.

– ***early experiment (Run-4)***, the source size may differ from NEMA + higher activity, placed at (y = 1 cm/10 cm/20 cm, z = 0/–18.75 cm), 150,000 events taken from each measurement.

Courtesy of Monika Pawlik-Niedźwiecka



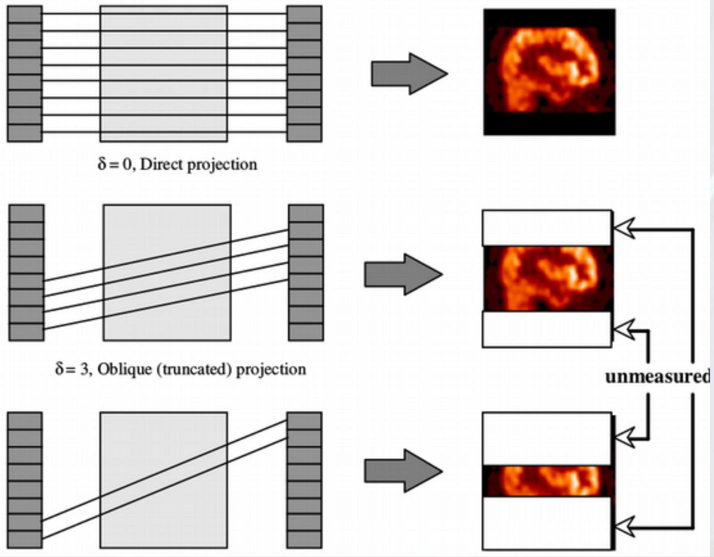
Data: single ^{22}Na source placed in positions according to NEMA

Time of measurement for single position: 3 hours

Activity of source: ~1134 kBq

Results: spatial resolution

Truncation problem and scanner sensitivity: sparse multilayer transverse geometry of "big barrel" implies non-uniform sensitivity, while ***"total body"*** size (50 cm) diminishes the truncation effects.

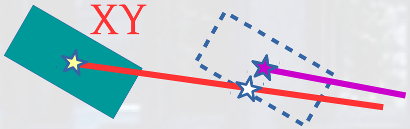


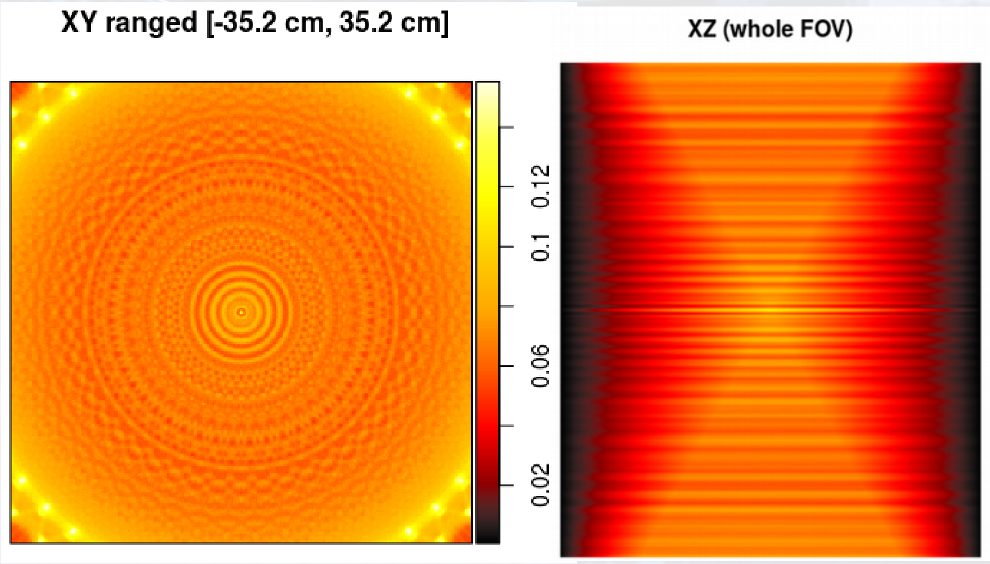
Re-projection was not used as in ***STIR*** *framework* [K Thielemans et al., PMB 2012], but sensitivity map was generated using hybrid 2D+2D approach: Monte Carlo simulation for XY plane and analytical estimation for XZ plane, based on the work [A Strzelecki Ph.D. dissertation PAN, Warsaw 2016].

Reconstructed image (for a voxel v)

$$f_{\mathrm{true}}(v) = f(v)/s(v),$$

$s(v)$ – sensitivity matrix.





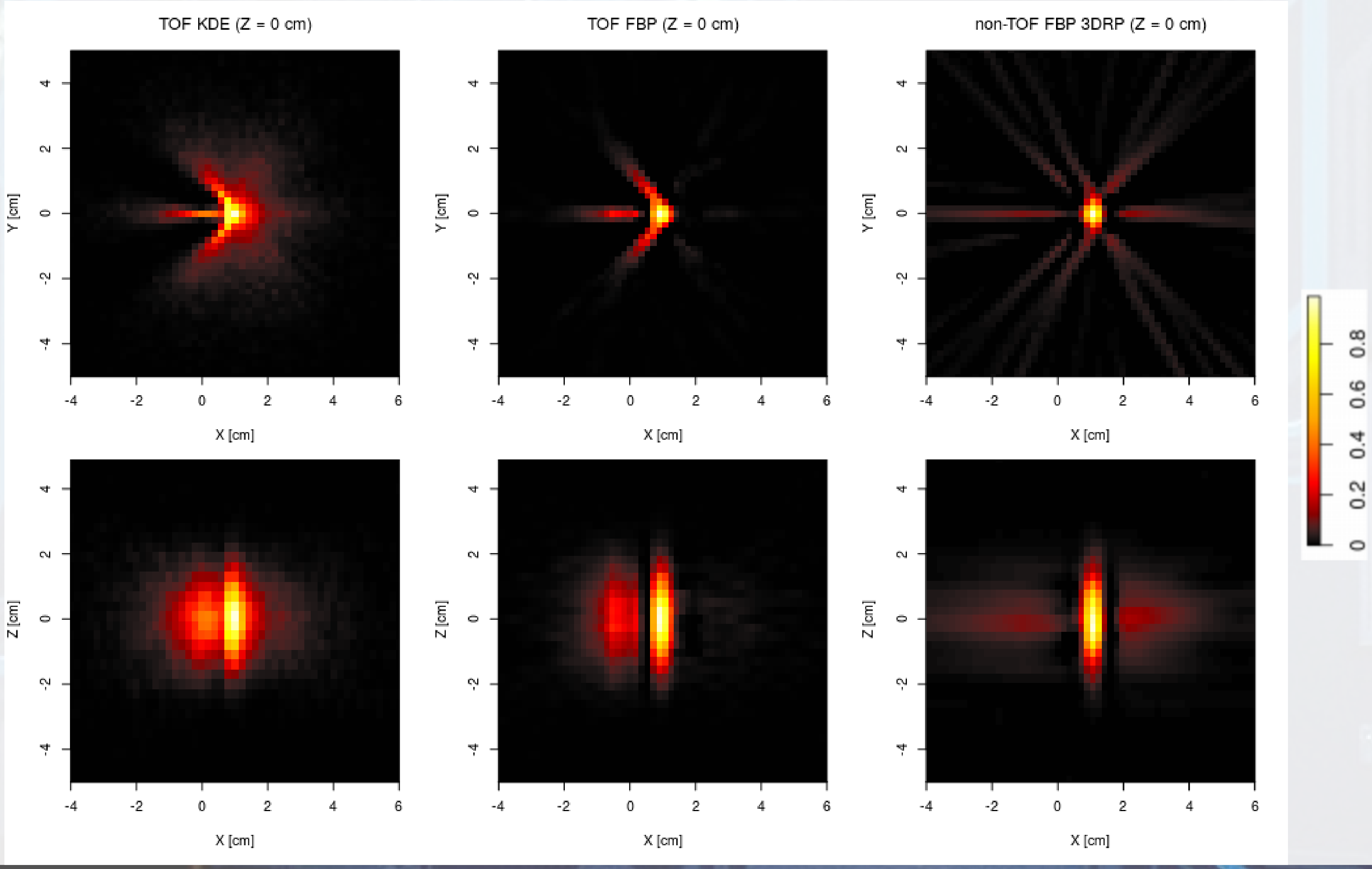
Reference images:

FBP 3DRP (STIR) – needs hit remapping onto a single layer (R = 43.73 cm, 384 strips)

TOF KDE – no filters, symmetric 3D kernel

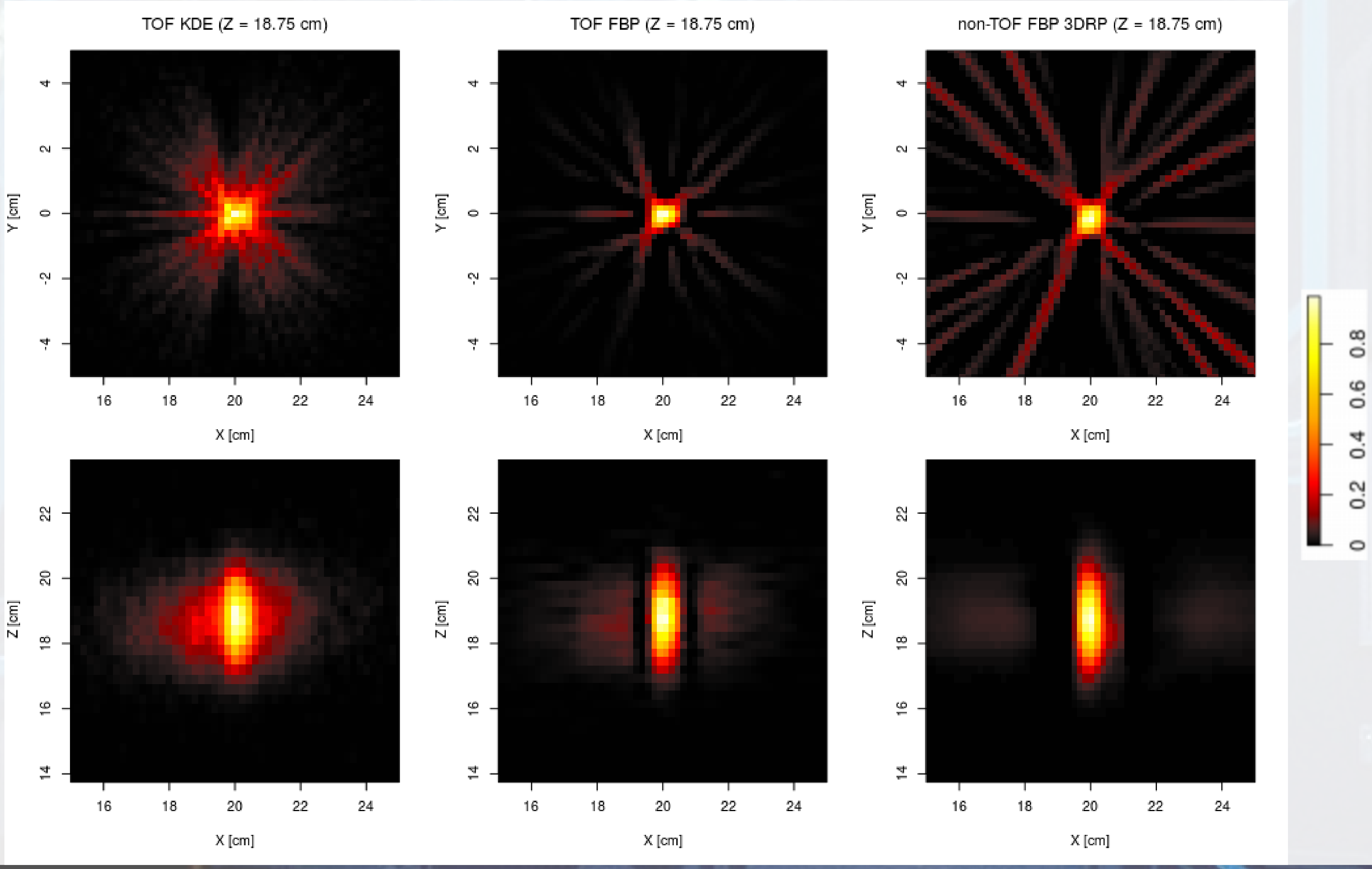
Results: spatial resolution

3D reconstructions by TOF KDE, TOF FBP and FBP 3DRP (STIR) for the ***simulated*** source at (1 cm, 0 cm, 0 cm), PMT readouts, ~100,000 events:



Results: spatial resolution

3D reconstructions by TOF KDE, TOF FBP and FBP 3DRP (STIR) for the ***simulated*** source at (20 cm, 0 cm, 18.75 cm), PMT readouts, ~100,000 events:



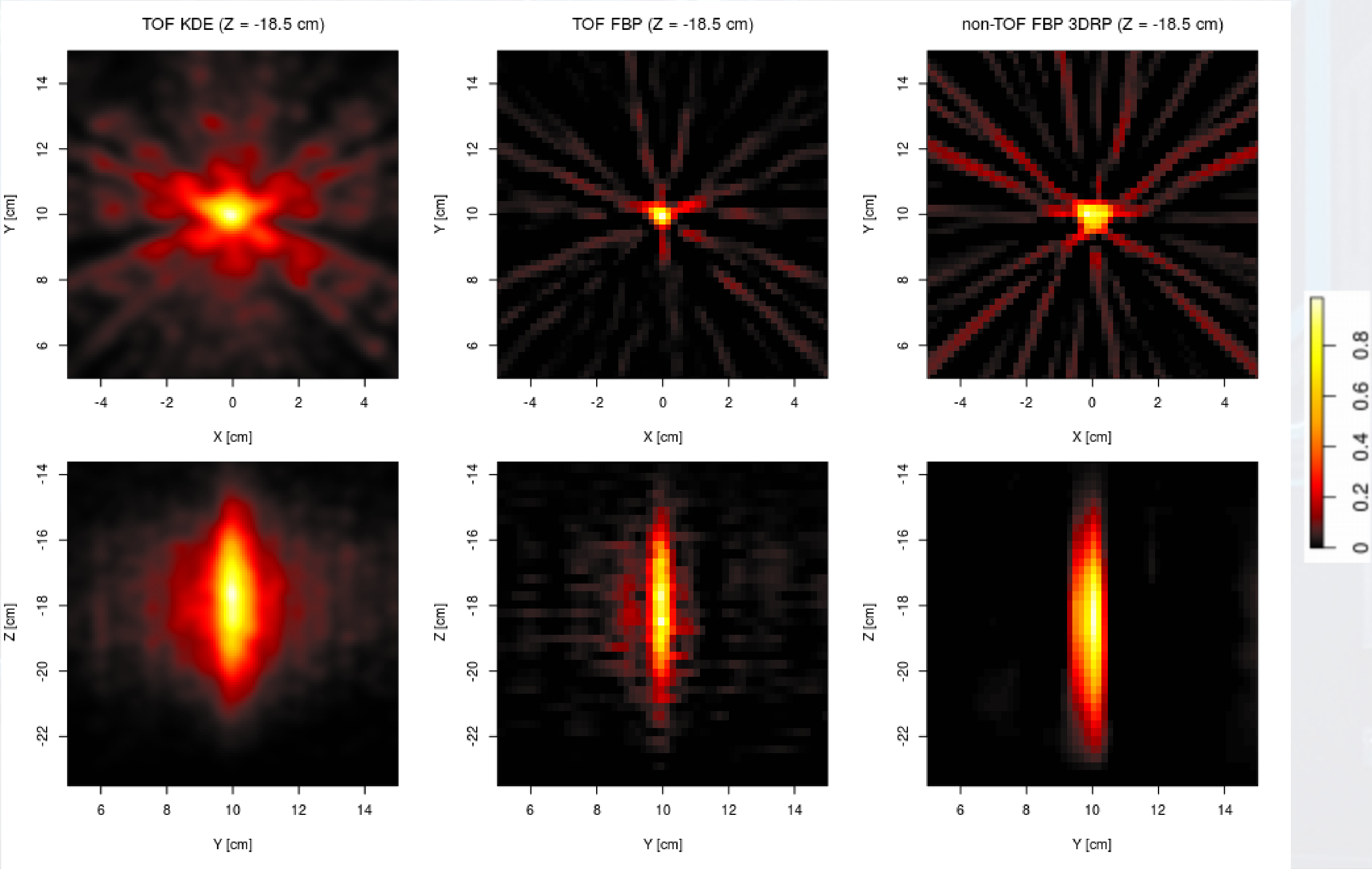
Results: spatial resolution

GATE simulations of 1-mm spherical source (370 kBq) – NEMA standard, PMT readout, for the accuracy of 0.5 mm. Voxel size 1.8 mm × 1.8 mm × 2.6 mm

	Readout:	**PMT**		
Source at: (y_{src} = 0 cm)	Algorithm	FWHM (in mm) along axis		
		X	Y	Z
3-layer scanner ("big barrel"): *R* = 42.5/46.75/57.5 cm, 48/48/96 strips of the dimension 7 × 19 × 500 mm, 100,000 events per simulation				
x_{src} = 1 cm, z_{src} = 0 cm	FBP 3DRP	4.5	7.0	20.0
	TOF KDE	5.5	6.0	20.0
	TOF FBP	**5.0**	**6.5**	**20.5**
x_{src} = 10 cm, z_{src} = 0 cm	FBP 3DRP	5.0	7.0	20.0
	TOF KDE	–	–	–
	TOF FBP	**5.5**	**5.0**	**20.0**
x_{src} = 10 cm, Z_{src} = 18.75 cm	FBP 3DRP	5.5	7.5	20.5
	TOF KDE	–	–	–
	TOF FBP	**5.5**	**5.0**	**20.0**
x_{src} = 20 cm, z_{src} = 18.75 cm	FBP 3DRP	6.5	7.5	21.0
	TOF KDE	7.5	6.0	22.0
	TOF FBP	**7.0**	**5.5**	**18.0**

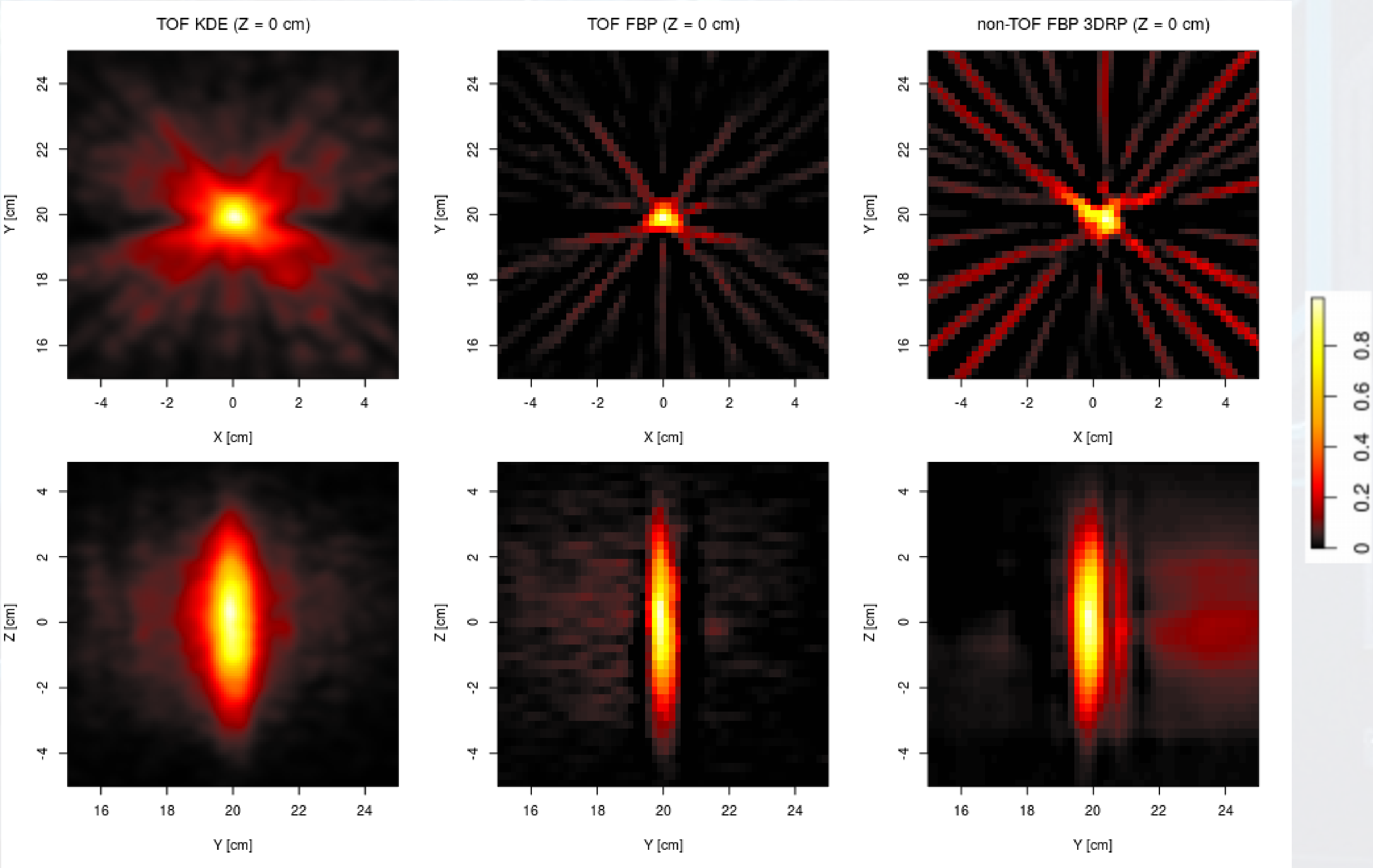
Results: spatial resolution

3D reconstructions by TOF KDE, TOF FBP and FBP 3DRP (STIR) for the ***measured*** source (0 cm, 10 cm, -18.5 cm?), PMT readouts, ~150,000 events:



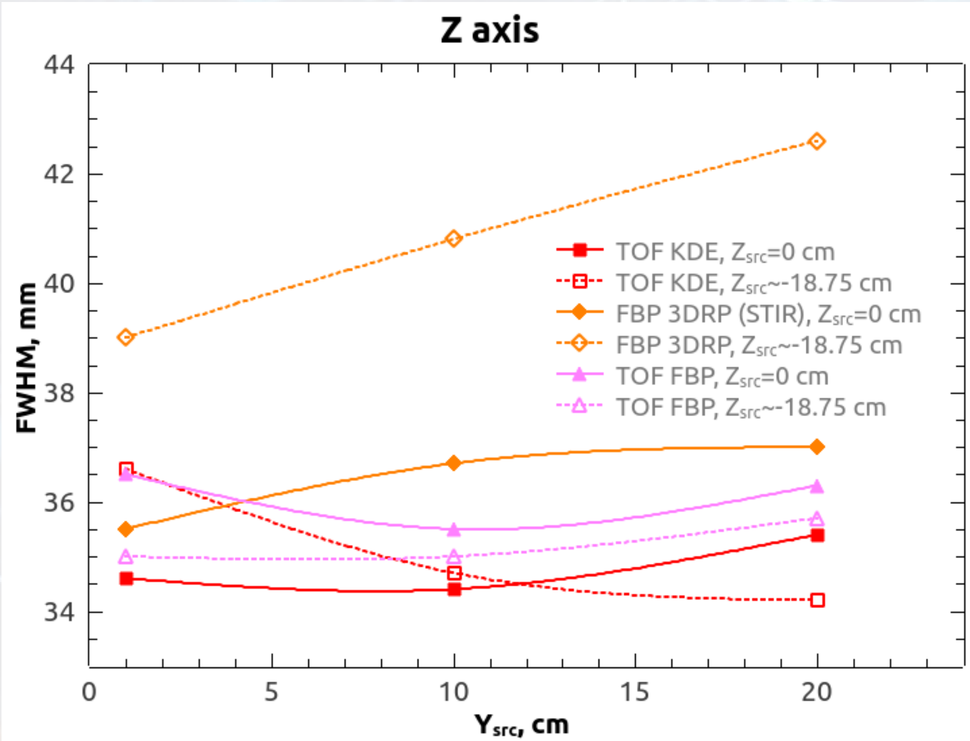
Results: spatial resolution

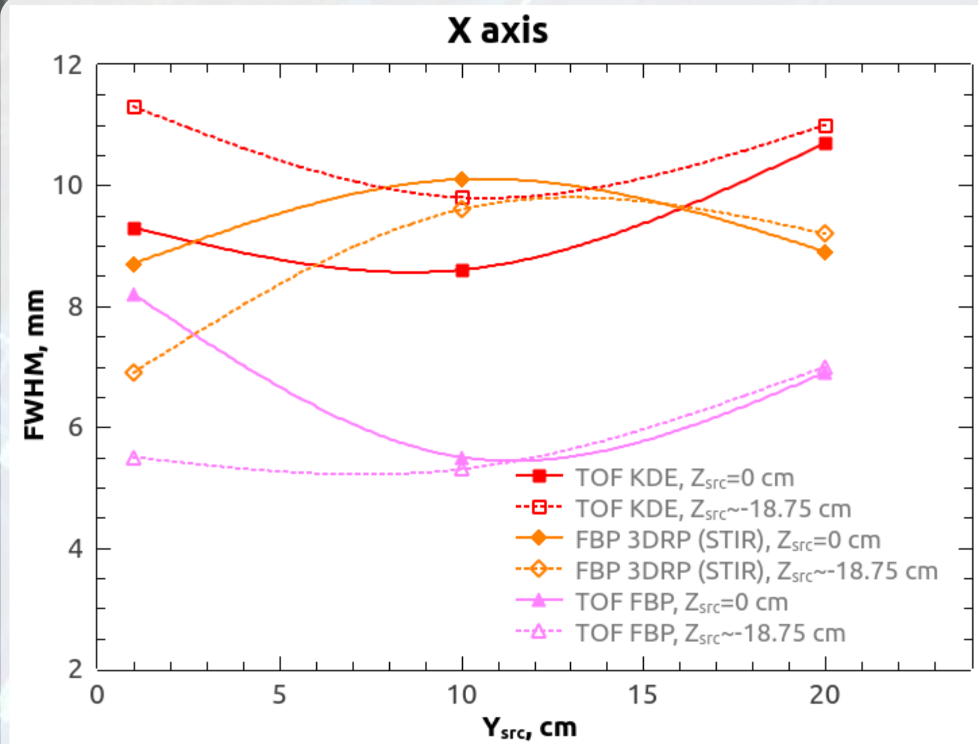
3D reconstructions by TOF KDE, TOF FBP and FBP 3DRP (STIR) for the ***measured*** source (0 cm, 20 cm, 0 cm), PMT readouts, ~150,000 events:

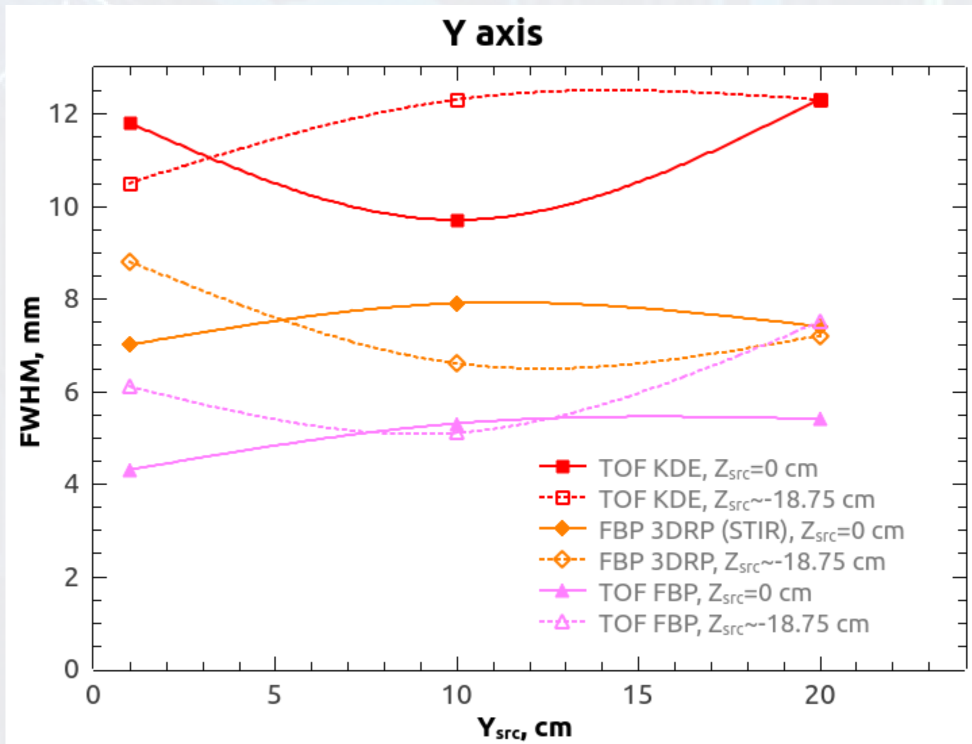


Results: spatial resolution

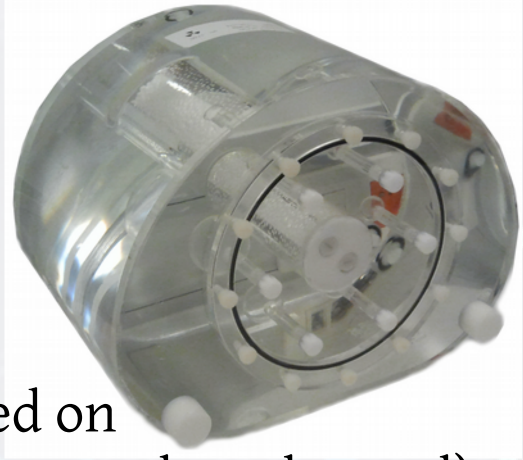
PSF (FWHM) values for Z axis are systematically lower if compared with the simulated data (***early experiment***)
TOF FBP is better than FBP 3DRP, but in order to use STIR, all hits were remapped onto 1-layer (384 strips)







Results: image quality



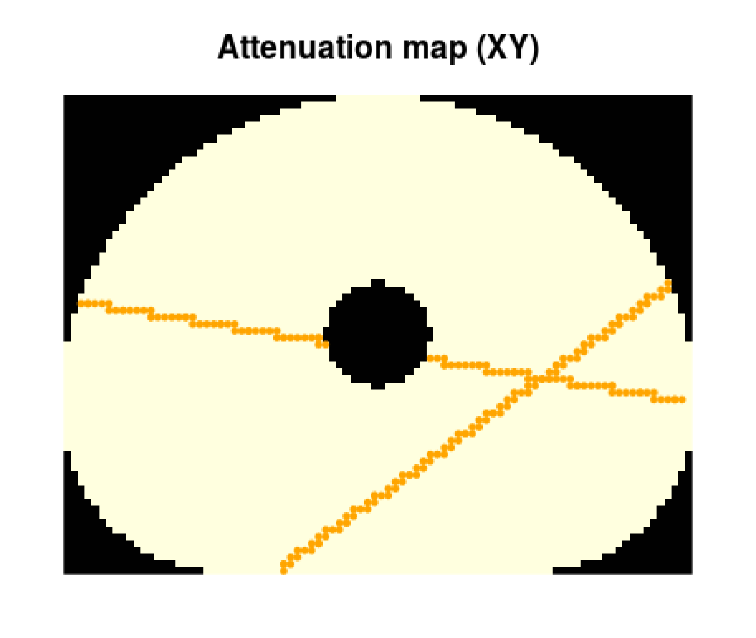
IEC NEMA phantom, simulated in GATE (at the centre of the scanner, one long measurement (3000 s), filtered by ***true coincidences*** only (data size 10-20 mln.)
Ideal geometry: 384 strips, R=43.73 cm, **SiPM** (CRT=235 ps)

Attenuation correction was added to TOF FBP: each LOR is treated as a projector, attenuation path is estimated based on Siddon algorithm (computing the intersecting length of a ray with each voxel) [R Li et al., Journ. Comp. Sci. 2010]:

Update intensity for each LOR as

$$I = I_0 \exp(-\mu x), \text{ where } \mu^{PET}(H_2O) = 0.096\ \text{cm}^{-1}.$$

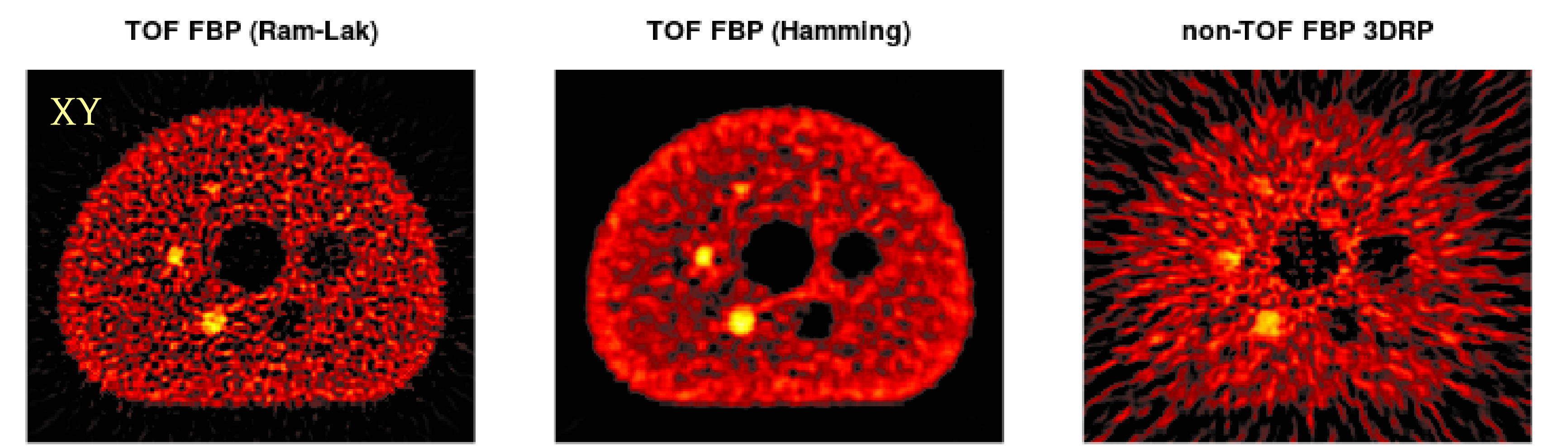
Attenuation map was created comprising all phantom volume filled with radioactive liquid, but without cold spheres and capillaries.



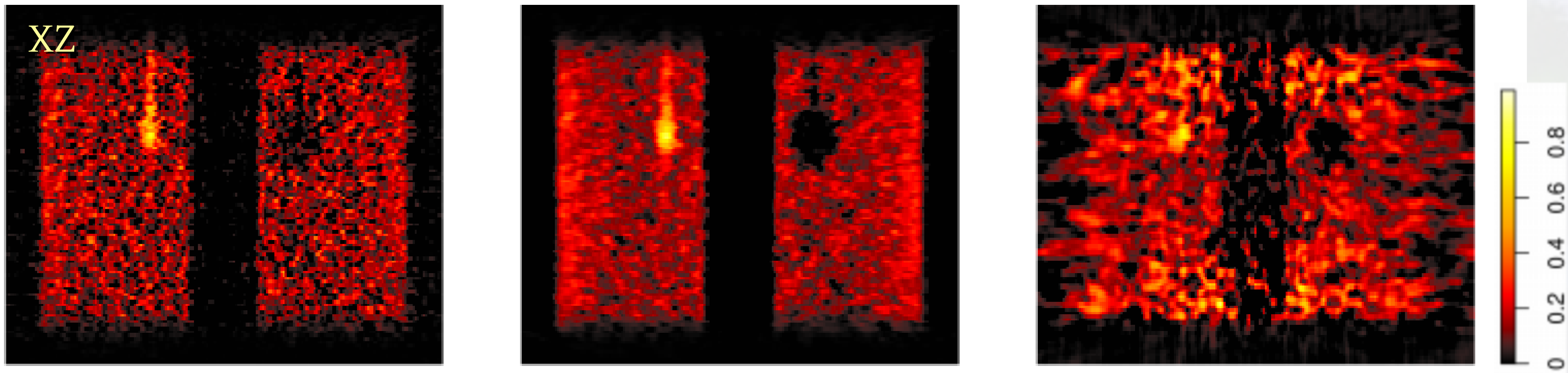
This upgrade extends reconstruction time by ***less than 10%*** (still possible for real time).

Results: image quality

10 mln. true coincidences, results for FBP 3DRP are **obtained by P. Kopka** (see poster). Contrast recovery coefficient **(CRC)**, background variation **(BV)** and signal-to-noise ratio **(SNR)** were estimated for 13-mm and 22-mm spheres.

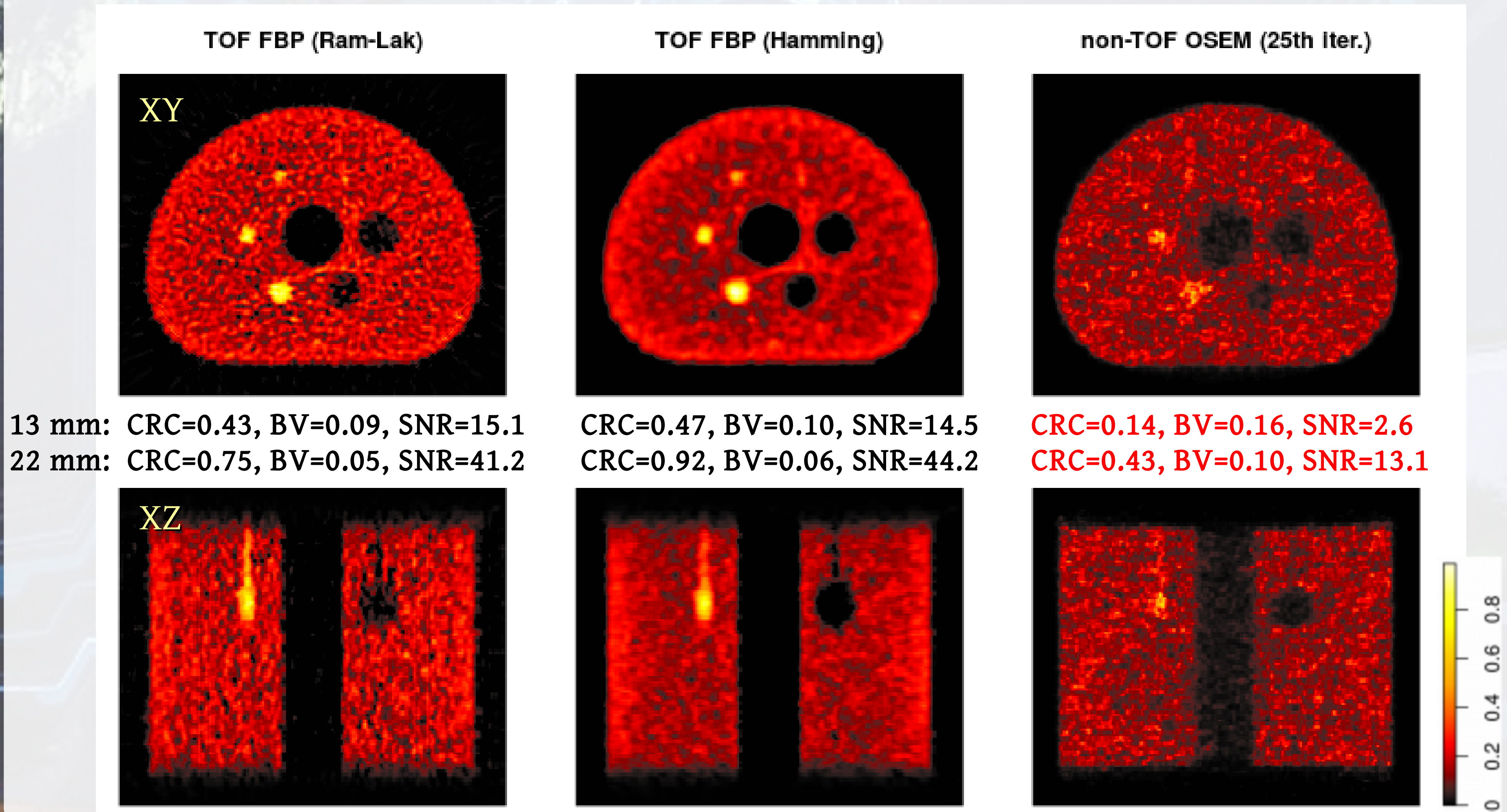


	TOF FBP (Ram-Lak)	TOF FBP (Hamming)	non-TOF FBP 3DRP
13 mm:	CRC=0.49, BV=0.17, SNR=9.0	CRC=0.48, BV=0.16, SNR=9.1	CRC=0.32, BV=0.27, SNR=3.5
22 mm:	CRC=0.82, BV=0.09, SNR=26.6	CRC=0.94, BV=0.10, SNR=29.8	CRC=0.77, BV=0.18, SNR=13.0



Results: image quality

20 mln. true coincidences, results for OSEM (STIR) are **obtained by P. Kopka**.
Full-sized TOF kernel was used for all images.



Summary and further plans

– TOF based reconstructions are promising solution for J-PET due to the excellent temporal resolution and shorter scan times if compared to non-TOF.

– There is a need for the specific algorithms due to the complex geometries of J-PET (sensitivity map is essential), the continuous character of strips (hence TOF), along with the eventual DOI information, estimated by WLS. Using bins in projection space is unpractical.

– TOF FBP could be employed using filters defined in image space, applied directly to each LOR as three separate kernels in event-by-event way. Scalability of this process opens up a possibility for real time imaging (already built and tested for non-filtered reconstructions using FPGA), since the intensity should be updated only for the small fraction of voxels.

– Imposing Gaussian kernels along LOR and Z-axis would blur the image thus reducing spatial resolution. The process of optimisation may differ from the bandwidth selection for TOF KDE, because 3D kernel is not symmetric.

– Single-event based TOF FBP achieve similar or better results for spatial resolution and image quality, compared to non-TOF reconstructions from STIR and non-filtered TOF KDE.

Yet to resolve:

– Explore the ways to optimise the parameters for asymmetric TOF FBP kernel, as well as the choice of optimal filter/cut-off frequency (apodisation).
– Compare the results for TOF FBP with other TOF based algorithms (MLEM, TV etc).
– Analyse performance benchmarks

Thank You for Your attention!