PHOTON DETECTION IN PET
USING ARTIFICIAL NEURAL NETWORKS

Michael Bronstein
Alexander Bronstein
Michael Zibulevsky
Yehoshua Y. Zeevi

WHAT IS PET ?

Positron emission tomography (PET) is a method of medical imaging which allows displaying metabolic activity in a slice of the body by means of detecting radiation, emitted from a radio-isotope injected into the patient's body.

Nuclear medical imaging methods allow studying physiological function of human organs, imaging the decay of radioisotopes bound to molecules with known biological properties. Unlike other methods of nuclear imaging, PET uses positron-emitting isotopes which occur naturally in many compounds of biological interest and can be incorporated in useful radiopharmaceuticals (e.g. glucose).

THE STAGES OF PET SCAN

PET scan begins when a radiopharmaceutical is injected into the patients’ body. After a certain period required for the isotope uptake, the body is placed into a radiation detector capable of registering the gamma quanta emitted from the body as the result of isotope decay. This information is collected and then converted into the slice image by a ***reconstruction algorithm***.

❶ **RADIOISOTOPE INJECTION**

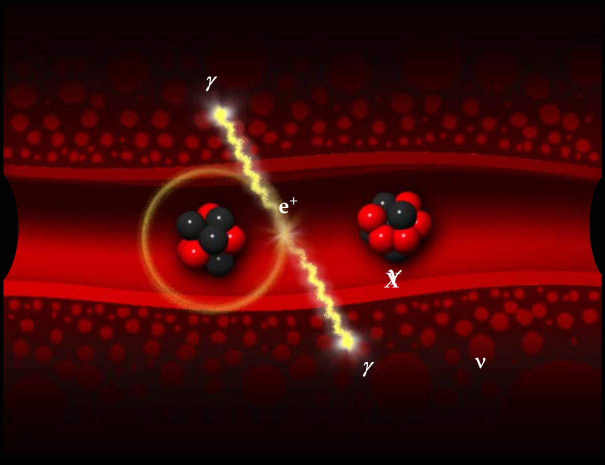
❷ **ACQUISITION**

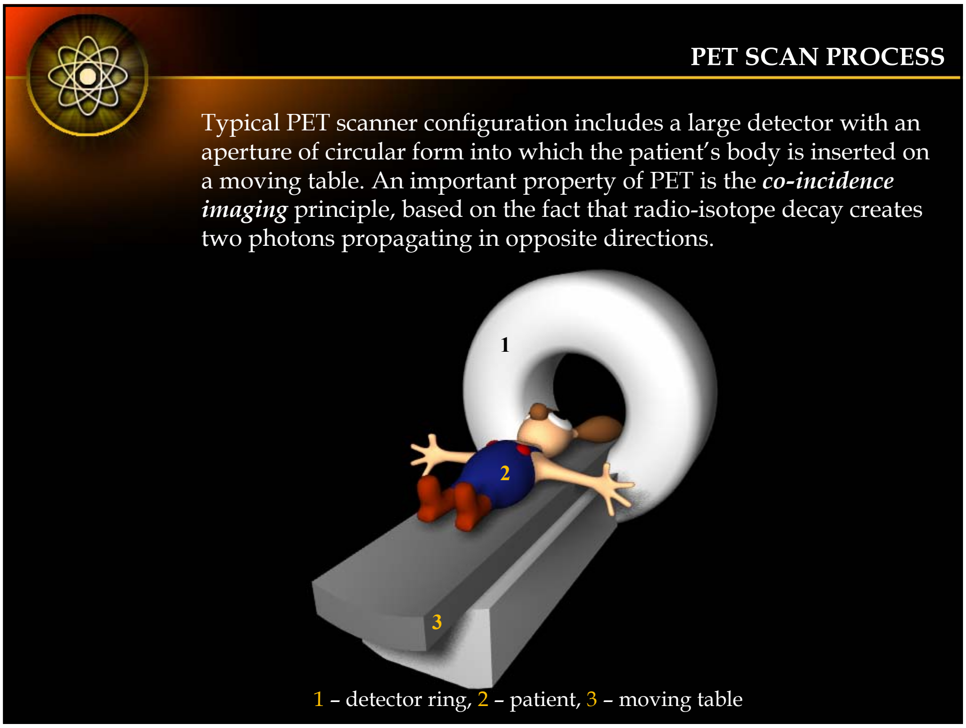
❸ **RECONSTRUCTION**

In this project the emphasis is made on the ***data acquisition*** stage.

PET SCAN PROCESS

The injected radio-isotope ***X*** undergoes ***positron decay***, resulting in emission of a positron $\mathbf{e^+}$ and a neutrino ν. During its propagation through the body tissues, the positron annihilates with an outer-shell electron producing a pair of high-energy photons propagating along almost collinear paths. These photons are detected during PET scan.

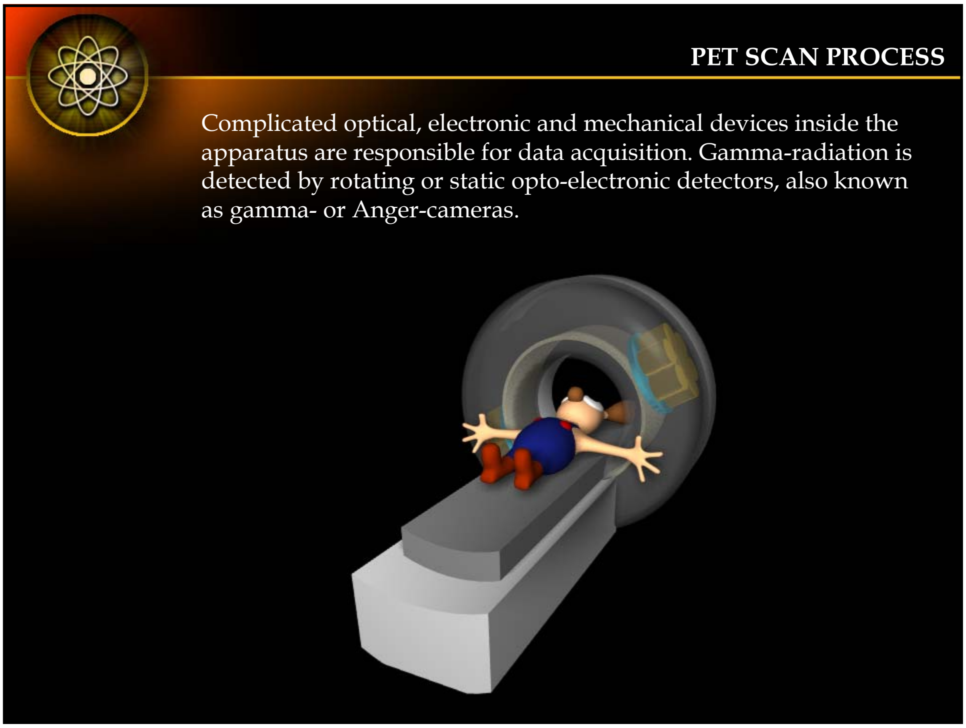




PET SCAN PROCESS

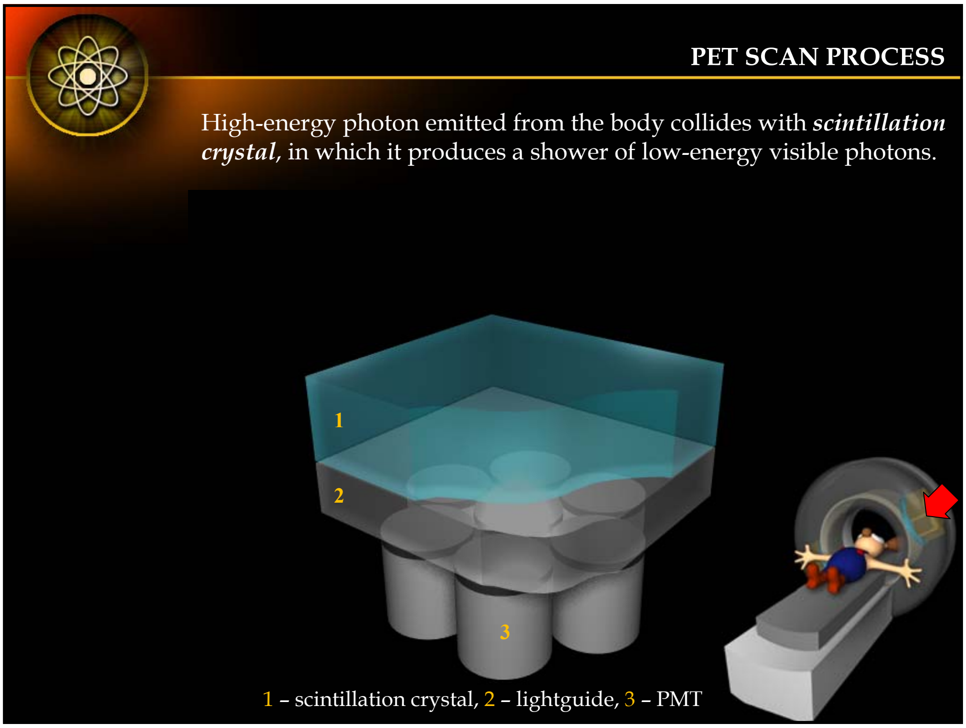
Typical PET scanner configuration includes a large detector with an aperture of circular form into which the patient's body is inserted on a moving table. An important property of PET is the ***co-incidence imaging*** principle, based on the fact that radio-isotope decay creates two photons propagating in opposite directions.

1 – detector ring, 2 – patient, 3 – moving table



PET SCAN PROCESS

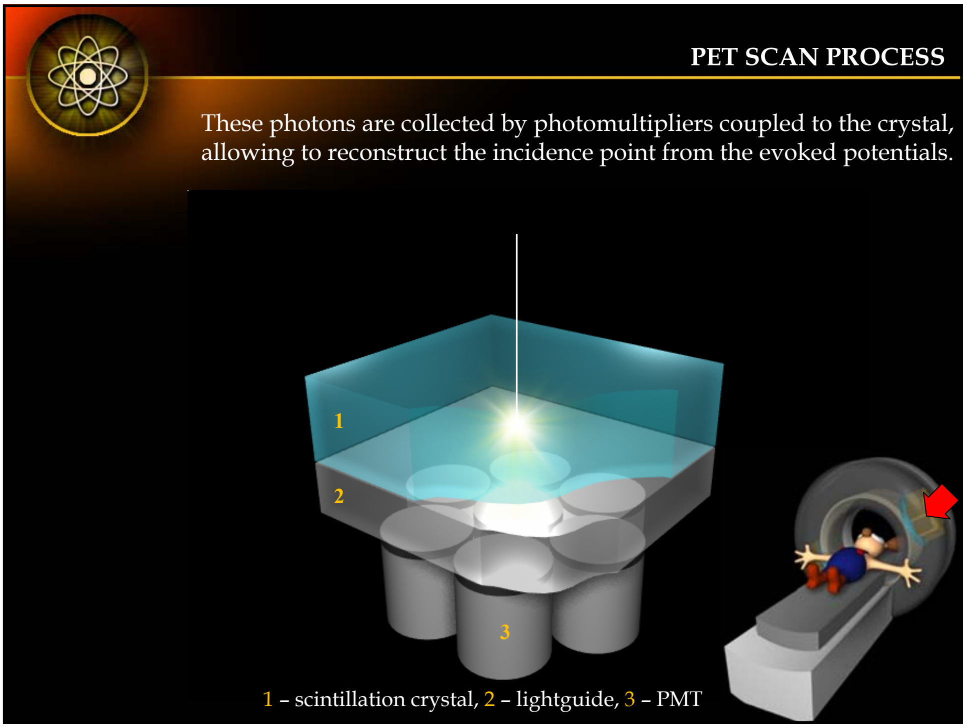
Complicated optical, electronic and mechanical devices inside the apparatus are responsible for data acquisition. Gamma-radiation is detected by rotating or static opto-electronic detectors, also known as gamma- or Anger-cameras.



PET SCAN PROCESS

High-energy photon emitted from the body collides with ***scintillation crystal***, in which it produces a shower of low-energy visible photons.

1 – scintillation crystal, 2 – lightguide, 3 – PMT



PET SCAN PROCESS

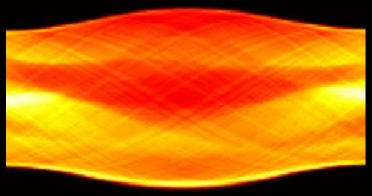
These photons are collected by photomultipliers coupled to the crystal, allowing to reconstruct the incidence point from the evoked potentials.

1 – scintillation crystal, 2 – lightguide, 3 – PMT

PET SCAN PROCESS

The detected events are binned into a radial distribution (***sinogram***). Only pairs of photons detected by opposite detectors are recorded, single events (***singles***) are rejected.

PHANTOM



SINOGRAM

RECONSTRUCTION

Image reconstruction is obtained by solving the inverse problem, i.e. obtaining the tracer distribution from its "projections".

One of the classical algorithms is the ***Filtered Backprojection***, based on the Radon transform; more complicated approaches, such as iterative reconstruction algorithms exist as well, but remain out of the scope of this project.

PHOTON DETECTION IN PET

Detection of high-energy photons in positron emission tomography is a basic, yet crucial stage prior to image reconstruction. The ability to precisely calculate the photon coordinates from electric pulses invoked in the photomultipliers (PMTs) as the result of scintillation implies high detection resolution and allows better reconstruction accuracy.

MOTIVATION

A significant part of this work was dedicated to developing such photon detection methods, which allow using a specific family of scintillation detectors based on thick ***potassium iodide*** crystals, since:

- NaI crystals have the largest light yield and the best energy resolution.
- NaI crystals can be made thick and may have almost 100% detection efficiency. However, till now the use of such crystals is limited due to the parallax effect.
- High detection efficiency implies lower radiopharmaceutical dose required for imaging.
- NaI crystals are cheap. Moreover, high detection efficiency implies less detecting devices required and therefore the PET scanner price can be reduced.

One of the goals was to overcome the problems related to application of such crystals in PET scanners.

PHOTON DETECTION IN PET

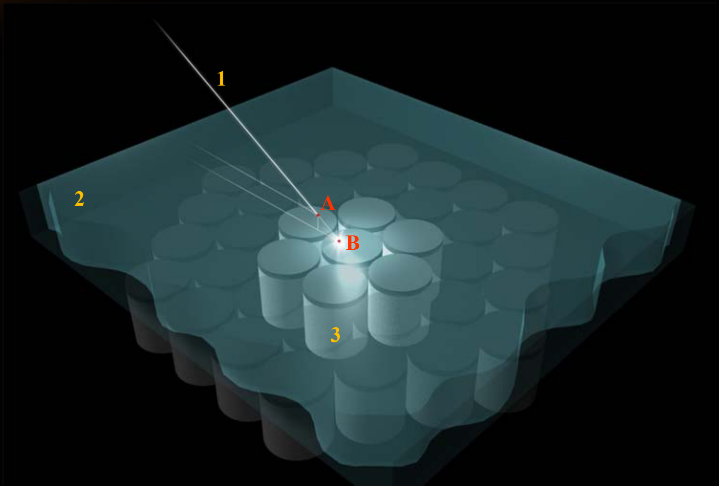
One of major problems in PET is the need to detect as much as possible of the emitted photons, in order to be able to reduce the patient’s exposure to radiation.

Scanners capable of detecting high percentage of photons should have scintillation detectors with the following properties:

- Thick NaI crystal
- Each scintillation event is treated by 7 active PMTs arranged in a hexagonal array
- Non-collimated Anger camera. Incident photons may come at different angles, far from the normal

PHOTON DETECTION IN PET

At large incidence angles, parallactic displacement of the scintillation point with respect to surface contact point may be very significant.



1 – incident photons, 2 – scintillation crystal, 3 – “active” PMTs
A – surface contact point, B – scintillation point

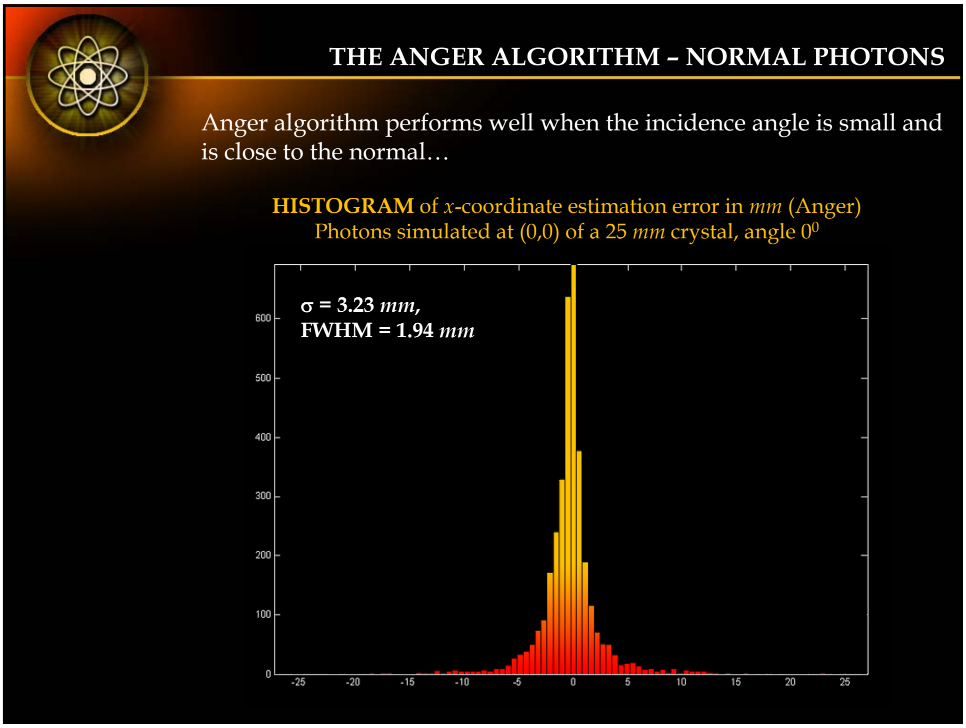
THE ANGER ALGORITHM

The most common method for scintillation point estimation is the ***Anger algorithm***.

Scintillation point coordinates are computed according to:

$$\hat{\boldsymbol{x}} = \frac{1}{n_{TOT}} \sum_i \boldsymbol{w}_i n_i$$

- Very simple and can be implemented even in analog hardware
- Used in the majority of modern scanners
- Appears precise only in collimated cameras or on thin crystals

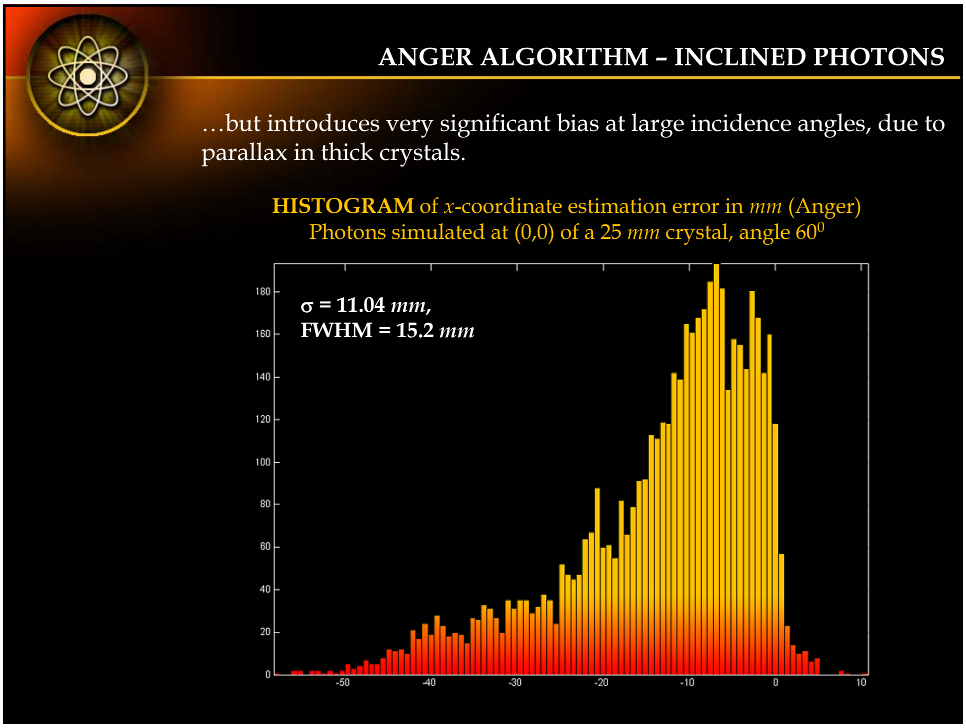


THE ANGER ALGORITHM – NORMAL PHOTONS

Anger algorithm performs well when the incidence angle is small and is close to the normal…

HISTOGRAM of *x*-coordinate estimation error in *mm* (Anger)
Photons simulated at (0,0) of a 25 *mm* crystal, angle 0^0

σ = 3.23 *mm*,
FWHM = 1.94 *mm*



ANGER ALGORITHM – INCLINED PHOTONS

…but introduces very significant bias at large incidence angles, due to parallax in thick crystals.

HISTOGRAM of *x*-coordinate estimation error in *mm* (Anger)
Photons simulated at (0,0) of a 25 *mm* crystal, angle 60^0

σ = 11.04 *mm*,
FWHM = 15.2 *mm*

PERFORMANCE OF THE ANGER ALGORITHM

Simulation shows that:

- Anger algorithm has strong bias at large angles
- Anger algorithm has large standard deviation
- Bias correction can remove the bias, but increases the standard deviation

As the result, Anger algorithm produces imprecise estimation of declined photons, which is liable to limit its application in scanners that use non–collimated scintillation cameras, particularly in fully-3D PET.

DEPTH-DETECTING ALGORITHMS

In order to compensate for the parallax effect, the detection algorithm should estimate either

- the surface contact point and the incidence angle, or
- the scintillation coordinates (x,y,z) , including the penetration depth.

However, simulations show that for 512*KeV* photons in thick crystals the average number of interactions is about 2 due to Compton scattering and, as the result, ***depth of interaction*** is not properly defined!

Hence, depth-detecting algorithm are unable to provide a sufficiently good estimation.

NEURAL NETWORKS FOR PHOTON DETECTION

- Scintillation detector may be considered a stochastic non-linear system, which maps scintillation position coordinates to PMT responses:

$$f : \tilde{N}^2 \to \tilde{N}^n$$

- It is very hard (or even impossible) to derive an analytical form of such mapping
- On the other hand, the mapping function can be easily sampled by creating incident photons at some point and recording the PMT responses
- We are looking for a non-linear estimator of the inverse mapping, which minimizes the expected quadratic error

---

NEURAL NETWORKS FOR PHOTON DETECTION

All this leads to a natural formulation of the optimal parametric estimator:

Given an unknown function $f : \tilde{N}^2 \to \tilde{N}^n$***, sampled on a set of N points,***

$$T = \left\{ \boldsymbol{x}^i = \left(x_1^i, .., x_n^i \right), y^i = f\left(x^i \right) + \xi_i \right\}_{i=1}^{N}$$

find an optimal (minimal MSE) parametric estimator of Y, $\Phi\left(\boldsymbol{x}; \mathbf{W} \right)$

In practice, the following criterion is used:

$$\min_{\mathbf{W}} \tfrac{1}{N} \sum_{i=1}^{N} \left(\Phi\left(\boldsymbol{x}^i; \mathbf{W} \right) - y^i \right)^2$$

PRINCIPLES OF NEURAL NETWORKS

A common way to solve non-linear estimation problems is based on Artificial Neural Networks (ANN). This is an information processing paradigm inspired by biological neural systems, and usually neural networks have the following common properties:

- The network is a structure of interconnected computational units (***neurons***), acting independently and simultaneously
- The processing capability of the network is usually stored in inter-connection strengths (***synaptic weights***)
- The network learns from a ***training set***, by a supervised or an unsupervised learning algorithm
- After learning, the network is supposed to generalize the given function
- Theoretically, under certain conditions, any "sufficiently good" function can be approximated to any degree of precision

---

DIFFERENT TYPES OF NEURAL NETWORKS

Different types of neural networks differ in their architecture and learning algorithms. The most popular types of neural networks used in regression problems include:

- Multilayer perceptron (MLP) feed-forward networks
- Radial-basis function (RBF) networks
- Neuro-fuzzy systems

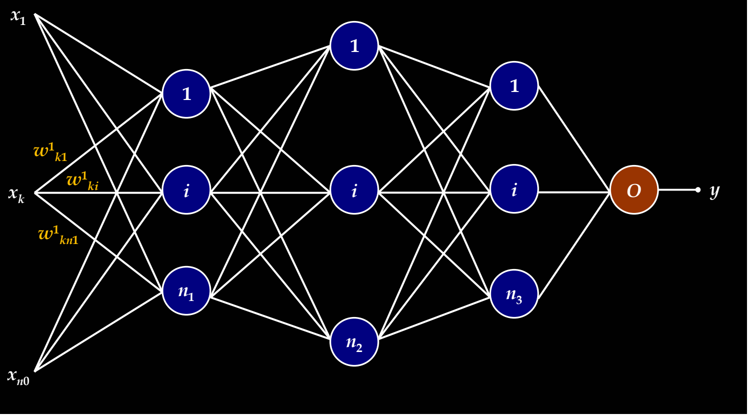
Learning is usually performed by some efficient optimization algorithm, including:

- Conjugate Gradients (CG) algorithm
- BFGS quasi-Newton algorithm
- Levenberg-Marquardt algorithm

MULTILAYER PERCEPTRONS (MLP) NETWORKS

Multilayer perceptron networks are feedforward neural networks. This model consists of an input layer, one or more hidden layers and one output layer. Training minimizes the quadratic error between the network output and the desired output:

$$E^k(W)=\frac{1}{2}\left\|\Phi_{MLP}\left(\boldsymbol{x}^k;W\right)-\boldsymbol{d}^k\right\|_2^2\to\min$$

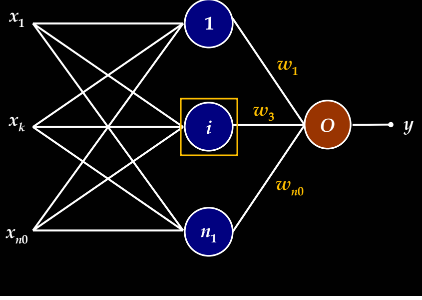


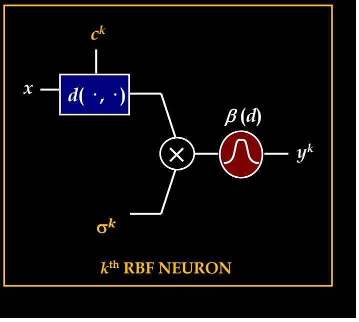
RADIAL-BASIS FUNCTION (RBF) NETWORKS

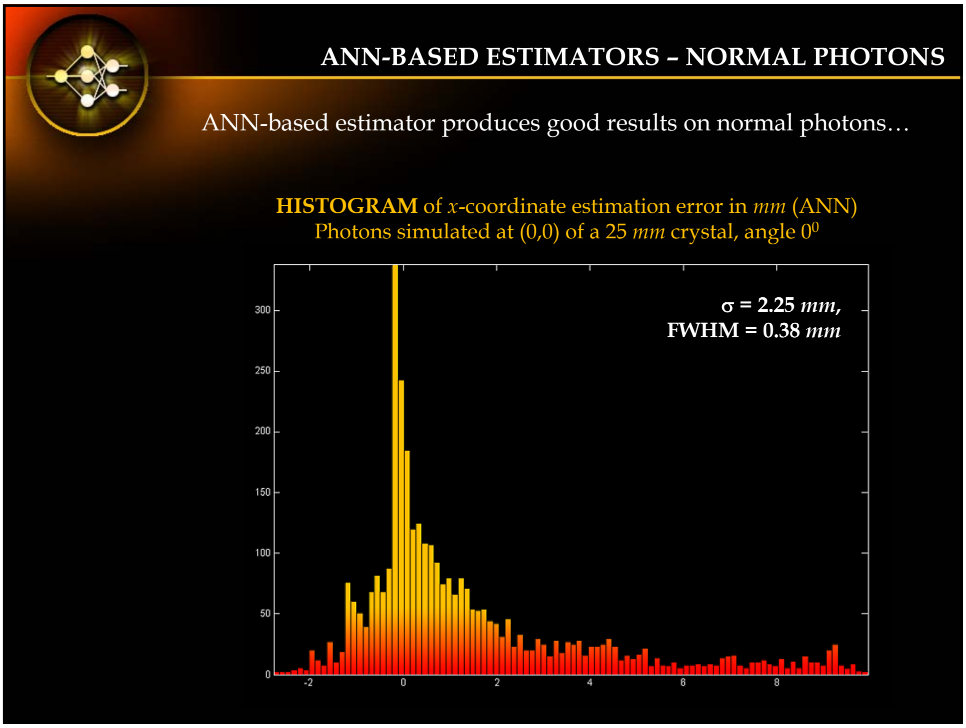
Radial-basis function networks are a popular alternative to multilayer perceptron networks for different function approximation problems. The advantage of RBF networks is the fact that their neurons act as local estimators, i.e. only several ***active*** neurons produce non-zero output.

During training the following functional is minimized:

$$E(W)=\sum_{k=1}^{M}\left(d^k-\Phi_{RBF}\left(\boldsymbol{x}^k;W\right)\right)^2+\lambda\left\|P\Phi_{RBF}\right\|^2\to\min$$





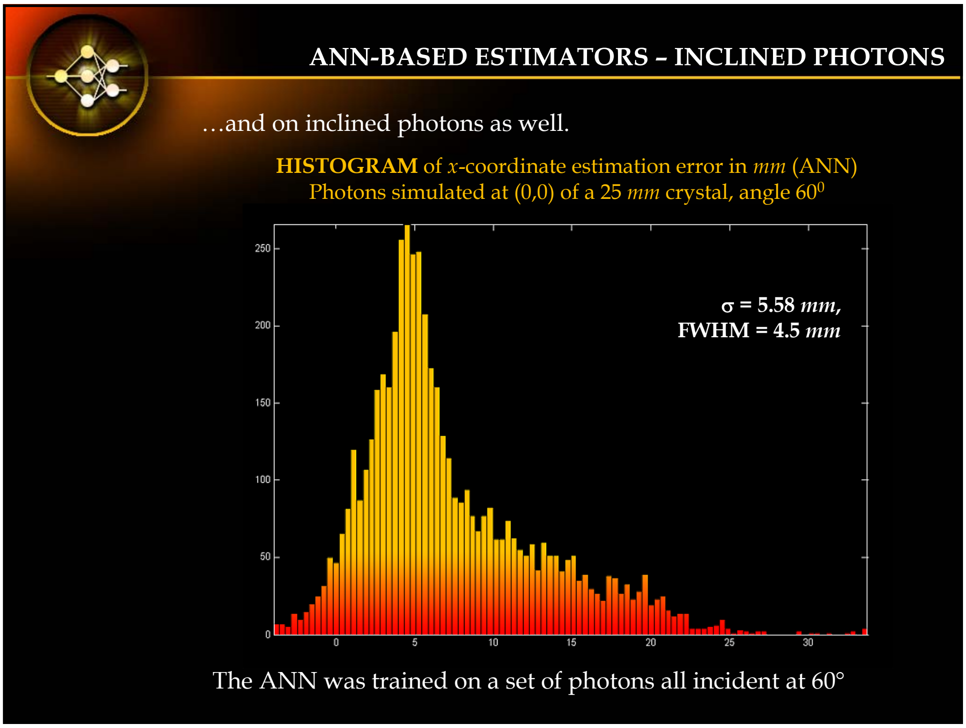


ANN-BASED ESTIMATORS – NORMAL PHOTONS

ANN-based estimator produces good results on normal photons…

HISTOGRAM of *x*-coordinate estimation error in *mm* (ANN)
Photons simulated at (0,0) of a 25 *mm* crystal, angle 0^0

σ = 2.25 *mm*,
FWHM = 0.38 *mm*



ANN-BASED ESTIMATORS – INCLINED PHOTONS

…and on inclined photons as well.

HISTOGRAM of *x*-coordinate estimation error in *mm* (ANN)
Photons simulated at (0,0) of a 25 *mm* crystal, angle 60^0

σ = 5.58 *mm*,
FWHM = 4.5 *mm*

The ANN was trained on a set of photons all incident at 60°

PERFORMANCE OF ANN-BASED ESTIMATORS

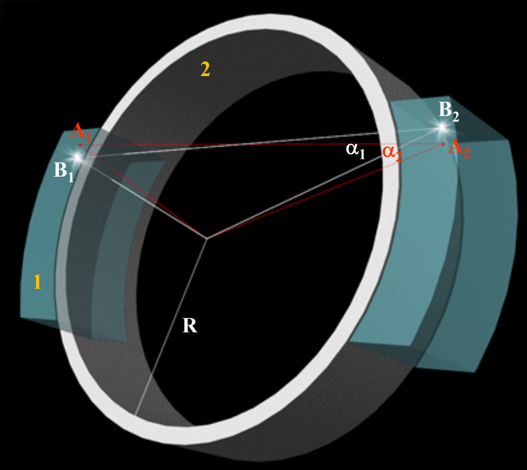
In comparison to Anger algorithm,

- ANN-based estimator has small bias (which can be practically removed by extensive training), and thus does not require bias correction
- ANN-based estimator has small standard deviation and a small FWHM:σ ratio
- ANN-based estimator performs well at different angles if appropriately trained

Simulation results prove these conclusions.

COMPLIMENTARY COINCIDENCE INFORMATION

Complimentary information from coincident events can serve for incidence angle estimation and used in the detection algorithm.



1 – scintillator, 2 – detector ring, B$_1$ – incident event, B$_2$ – coincident event, A$_1$ – incident estimation, A$_2$ – coincident estimation

PHOTON DETECTION ALGORITHM

The idea of utilizing complimentary information allows design a photon detection algorithm based on local regression:

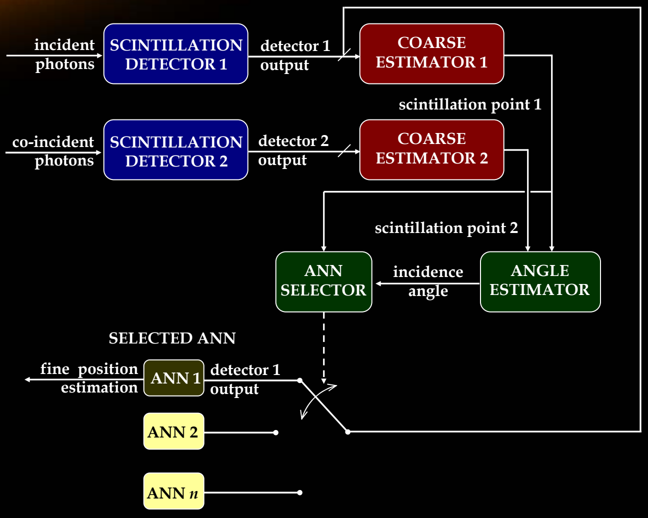
- The core of the detection algorithm is a set of neural networks, trained on scintillation events in different regions and at ***different incidence angles***
- The number of regions and angles depends on the desired accuracy and on practical computational limitation
- The sets of neural networks are trained independently on PMT responses resulting from scintillation events in appropriate regions and angles
- Incidence angle is estimated using the coincident event

PHOTON DETECTION ALGORITHM

The algorithm works as follows:

- Compute the scintillation planar coordinates using some coarse estimator, e.g. Anger algorithm
- Estimate a region in which scintillation occurred
- Compute the incidence angle using information from coincident event. Anger algorithm may be used to estimate the angle with sufficiently high precision
- Select a fine estimator trained at appropriate region and angle
- Compute fine estimation of the scintillation coordinates

PHOTON DETECTION ALGORITHM


incident photons

SCINTILLATION DETECTOR 1

detector 1 output

COARSE ESTIMATOR 1

scintillation point 1

co-incident photons

SCINTILLATION DETECTOR 2

detector 2 output

COARSE ESTIMATOR 2

scintillation point 2

ANN SELECTOR

incidence angle

ANGLE ESTIMATOR

SELECTED ANN

fine position estimation

ANN 1

detector 1 output

ANN 2

ANN *n*

PHOTON DETECTION ALGORITHM COMPLEXITY

ANN-based estimator has low computational complexity:

- Since each time only a small number of PMTs is active, the input vector is small, which implies simple neural network
- The number of operations performed by the algorithm is proportional to the number of synaptic weights, and since the networks are small, the computational complexity of the algorithm is low
- As the result, the algorithm can be implemented in hardware and used in real-time
- ANN-based estimators can be combined with Anger algorithm
- Training of the fine estimator can be performed on scintillation events, distributed according to the error of a coarse estimator

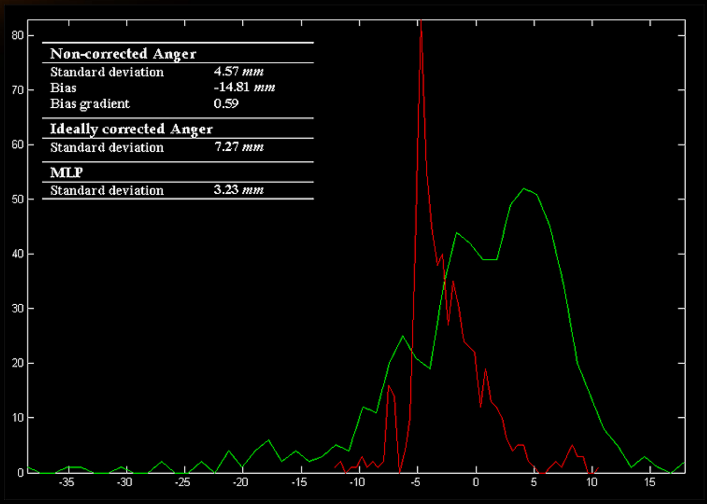
SIMULATION RESULTS

Simulation studies were performed on NaI crystals of different thickness.

- For thick crystals (45 *mm*) ANN-based estimators significantly outperform the classical approach
- The difference both in standard deviation and bias is especially significant at large incidence angles, which proves the suitability of the proposed methods for parallax distortion compensation.

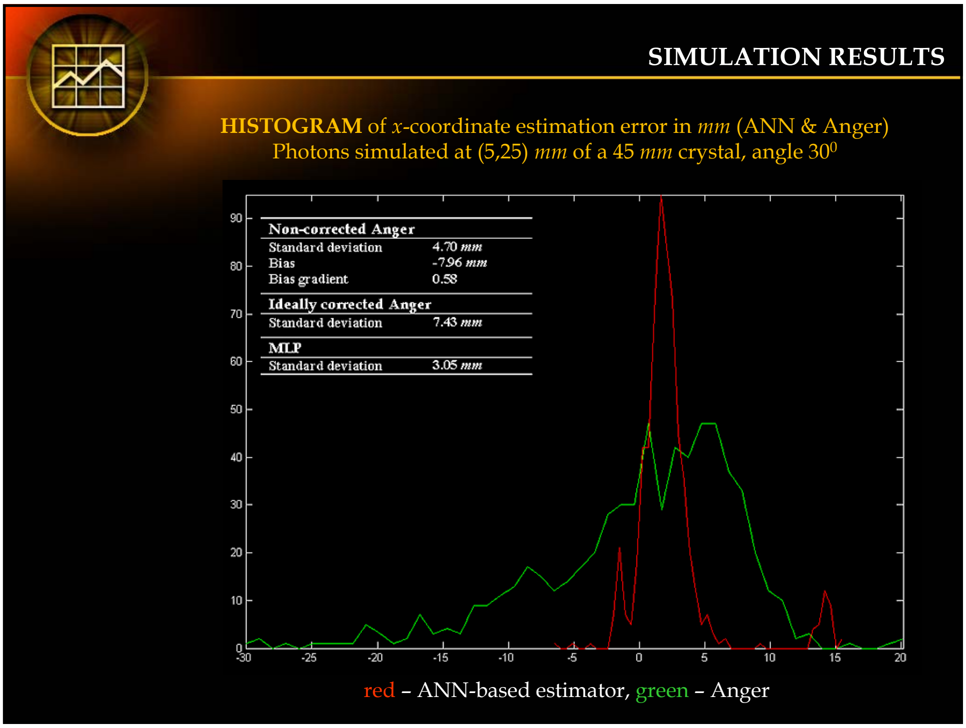
SIMULATION RESULTS

HISTOGRAM of *x*-coordinate estimation error in *mm* (ANN & Anger)
Photons simulated at (17,17) *mm* of a 45 *mm* crystal, angle 30^0



Non-corrected Anger	
Standard deviation	4.57 *mm*
Bias	-14.81 *mm*
Bias gradient	0.59
Ideally corrected Anger	
Standard deviation	7.27 *mm*
MLP	
Standard deviation	3.23 *mm*

red – ANN-based estimator, green – Anger

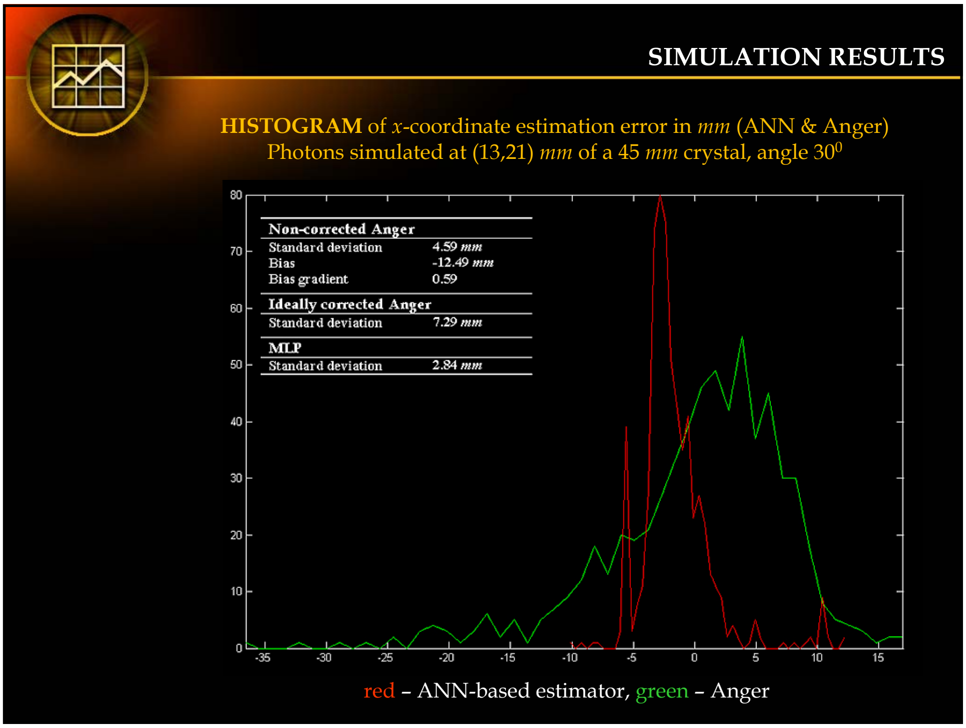


SIMULATION RESULTS

HISTOGRAM of x-coordinate estimation error in *mm* (ANN & Anger)
Photons simulated at (5,25) *mm* of a 45 *mm* crystal, angle 30^0

Non-corrected Anger	
Standard deviation	4.70 *mm*
Bias	-7.96 *mm*
Bias gradient	0.58
Ideally corrected Anger	
Standard deviation	7.43 *mm*
MLP	
Standard deviation	3.05 *mm*

red – ANN-based estimator, green – Anger



SIMULATION RESULTS

HISTOGRAM of x-coordinate estimation error in *mm* (ANN & Anger)
Photons simulated at (13,21) *mm* of a 45 *mm* crystal, angle 30^0

Non-corrected Anger	
Standard deviation	4.59 *mm*
Bias	-12.49 *mm*
Bias gradient	0.59
Ideally corrected Anger	
Standard deviation	7.29 *mm*
MLP	
Standard deviation	2.84 *mm*

red – ANN-based estimator, green – Anger

SIMULATION RESULTS

Algorithms were compared on a thick (45 *mm*) NaI crystal, at central region and at different incidence angles. Along with MLP, RBF networks and neuro-fuzzy function approximation (NEFPROX) algorithm were tested.

TABLE of MSE root (*mm*) in a central region, angle 0^0

	Minimal error	Maximal error	Average error
Anger [unbiased]	4.9759	5.1251	5.0492
Linear regression	2.7248	5.2052	3.8609
MLP	2.8911	3.9501	3.4359
RBF	2.8994	3.7875	3.3093
NEFPROX	2.8478	4.4031	3.5772

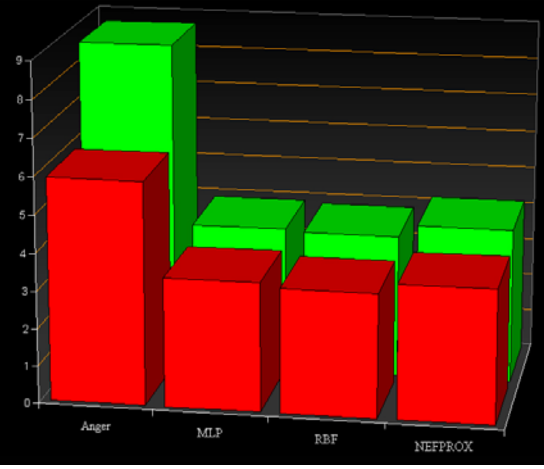
TABLE of MSE root (*mm*) in a central region, angle 30^0

	Minimal error	Maximal error	Average error
Anger [unbiased]	7.1340	7.4158	7.2878
Linear regression	3.2015	6.3351	4.7618
MLP	3.3570	4.4614	3.8643
RBF	3.4071	4.4825	3.7918
NEFPROX	3.2437	5.0703	4.1138

SIMULATION RESULTS

Non-linear regression methods appear more precise by as much as 50% with respect to Anger !

COMPARATIVE CHART of average MSE root (*mm*) in a central region
Red – angle 0^0, Green – angle 30^0



CONCLUSIONS

We believe that our work has proven the possibility of implementing ANNs and other non-linear parametric estimation methods for photon detection in positron emission tomography.

The suggested approach:

- allows improve spatial resolution of Anger cameras
- allows compensate for non-linearity distortions and intrinsic imperfections of the detector
- allows use cheap NaI crystals
- allows use thick crystals with detection efficiency close to 100%. As the result, the patient’s exposure to radiation can be reduced

---

CONCLUSIONS

Radiopharmaceutical dose is the major factor limiting the frequency and the number of PET scans a patient can undergo. Reducing the dose is therefore a significant achievement and one of the most important goals in PET technology development.

FURTHER RESEARCH DIRECTIONS

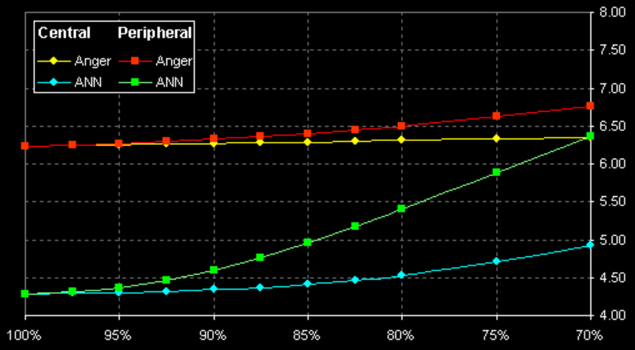
Further research is needed in order to study the behavior of the proposed methods in real conditions. The following aspects have been preliminarily studied and require additional research:

- Sensitivity to the PMT aging effect, i.e. to reduction of the PMT gain factor with time
- Sensitivity to the pile-up effect, i.e. to time overlap of subsequent scintillation events
- Sensitivity to incidence angle variation

FURTHER RESEARCH DIRECTIONS

Real photomultipliers tend to "age", i.e. their gain decreases with time. The sensitivity of a photon detection algorithm to "aging" of any of the detector PMTs must be as low as possible.

CHART of average MSE root (*mm*) of Anger and ANN estimators vs. relative PMT gain for central and peripheral PMTs.

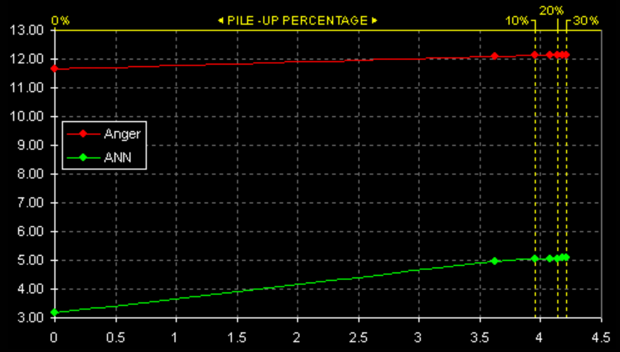


Preliminary tests show good performance of ANN-based algorithms. However, further research of this subject is necessary.

FURTHER RESEARCH DIRECTIONS

In real detectors, scintillation events are liable to "pile-up", i.e. overlap in time, thus introducing random noise. This effect is critical for high-count scans and limits the intensity of the radioactive tracer, which can be properly treated by the scanner.

CHART of average MSE root (*mm*) of Anger and ANN estimators vs. average scintillation event rate (*events/msec*).

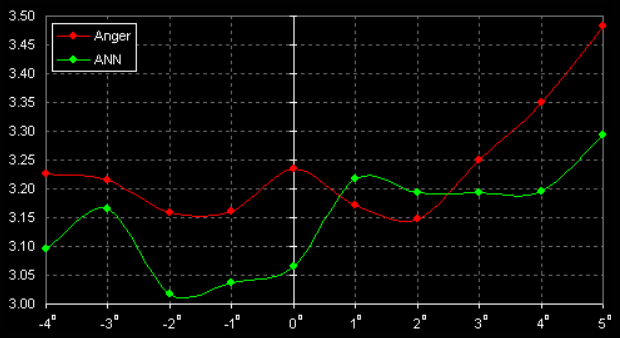


Preliminary tests show low sensitivity of ANN-based algorithms to the pile-up effect. However, further research is desirable.

FURTHER RESEARCH DIRECTIONS

Since the proposed ANN estimators are angle-dependent, and photon incidence angles can be approximated up to some finite precision, it is important to study the sensitivity of the algorithm to incidence angle variation and thus determine the required angular resolution.

CHART of average MSE root (*mm*) of Anger and ANN estimators trained for normal photons vs. incidence angle (*degrees*).



Preliminary tests show sufficiently low sensitivity of ANN-based algorithms to incidence angle variations. Further research is required.

FURTHER RESEARCH DIRECTIONS

In addition, other aspects should be studied as well prior to physically implementing the proposed methods in PET scanners.

- Research the stability of ANN-based estimators under camera parameters variation and conclude whether it is possible to train the networks on a specific camera or simulation and apply them to other cameras
- Research the possibility to disqualify non-collinear events (***randoms***)
- Create a faithful algorithm for low-energy events discrimination
- Research the influence of detector surface properties, coloring and crystal thickness on the performance of ANN-based estimators