LAPP Annecy 08/10/2021

Artificial Intelligence approaches for Monte Carlo simulation

David Sarrut

CNRS - University of Lyon

CREATIS, CNRS UMR5220, Inserm U1044, INSA-Lyon, Université Lyon 1, France

CREATIS Research Lab

cnrs dépasser les frontières

CENTRE DE LUTTE CONTRE LE CANCER LEON BERARD Cancer center

ONLYLYON

- Image reconstruction
- MC simulations
- EBRT and Nuc Med

Monte Carlo simulations

- Born during WW2
  - Stanislaw Ulam, John von Neumann … (Manhattan Project)
  - Simulate radiation/particles transport



- MC in HEP and medical physics
  - Heavily used in High Energy Physics (CERN)
  - MedPhys: roots in the 70', **imaging systems** (SPECT, PET) and **Radiation Therapy**

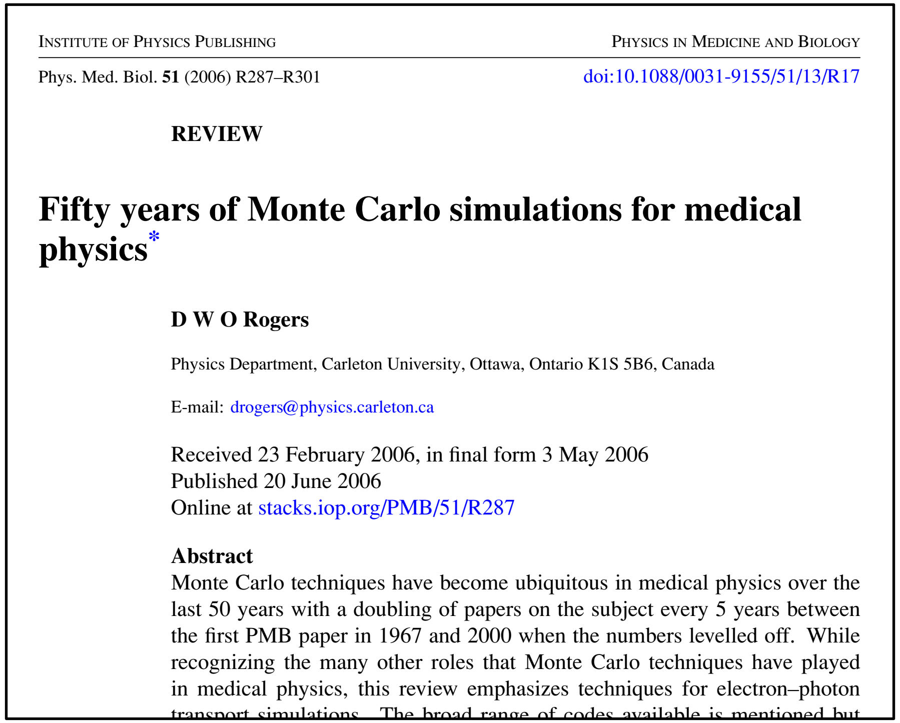
- Nowadays in medical physics
  - All TPS (Treatment Planning System)
  - All PET, SPECT ; Total-Body PET projects (Explorer, etc)

100%

Monte Carlo simulations evolution

- More than 60 years of evolution
  - More accurate physical databases
  - More generic codes
    (MCNPX, EGSNRC, Penelope, Geant4, Gate)
  - Faster algorithms
  - Use of powerful computing infrastructures
    (cluster, GPU)

- However
  - Increasing need for detailed and accurate physical processing (TOF, SiPM, CZT, etc)
  - Still long simulations times (need VRT)


INSTITUTE OF PHYSICS PUBLISHING — PHYSICS IN MEDICINE AND BIOLOGY
Phys. Med. Biol. **51** (2006) R287–R301 — doi:10.1088/0031-9155/51/13/R17

REVIEW

Fifty years of Monte Carlo simulations for medical physics*

D W O Rogers

Physics Department, Carleton University, Ottawa, Ontario K1S 5B6, Canada

E-mail: drogers@physics.carleton.ca

Received 23 February 2006, in final form 3 May 2006
Published 20 June 2006
Online at stacks.iop.org/PMB/51/R287

Abstract
Monte Carlo techniques have become ubiquitous in medical physics over the last 50 years with a doubling of papers on the subject every 5 years between the first PMB paper in 1967 and 2000 when the numbers levelled off. While recognizing the many other roles that Monte Carlo techniques have played in medical physics, this review emphasizes techniques for electron–photon transport simulations. The broad range of codes available is mentioned but

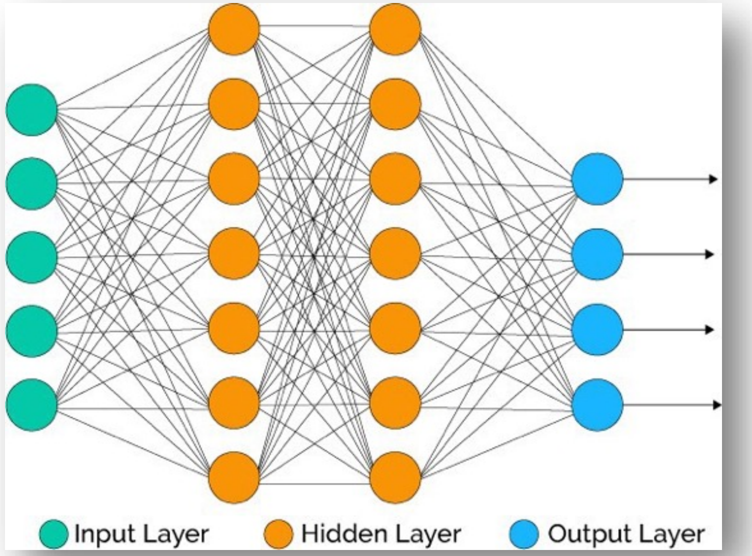
Artificial Intelligence (A.I)

- A.I. methods, image processing (photos, video)
- Deep Learning, neural network

- Medical physics:
  - Detection
  - Auto segmentation
  - Image generation (CT from MRI, CT from CBCT etc)
  - Image enhancement (remove artefacts)
  - Radiomics
  - etc …

Deep learning principle

- **Step1**: learn a model
  - Input training database (large), composed of numerous independent samples
  - Neural network architecture and learning methods

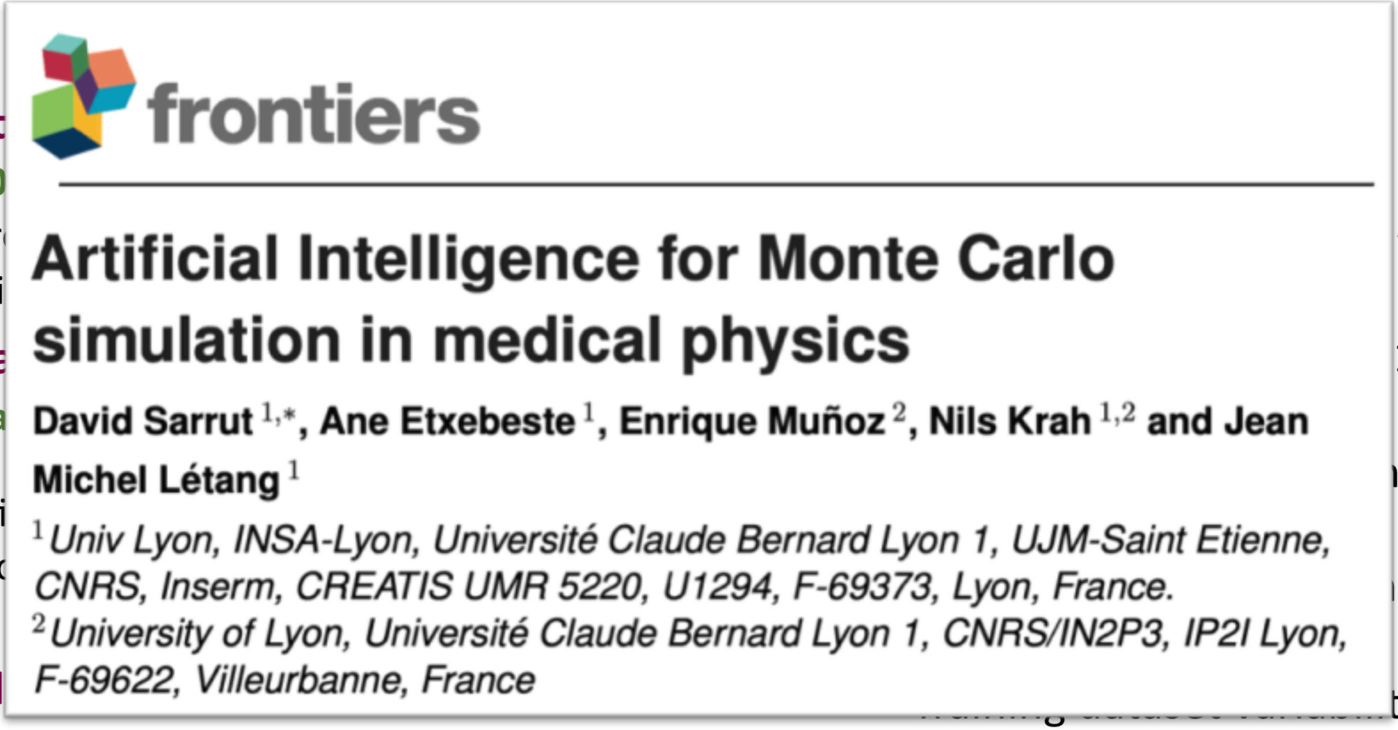


- **Step2**: use the model
  - Get input data, apply the NN

Could it be useful for MC ?

(Very short) literature review

- **DL and dose estimat**
  - [Lee2019, Götz20
  - U-Net architectur
  - Large dataset vari
- **DL for dose computa**
  - [Peng2019, Forna Madrigal2018]
  - Towards less parti
  - Photon, proton do
  - Towards GAN ?
- **DL for scatter model**
  - [B van der Heyden2020, Lee2019, Maier2018, Sharp2020]
  - U-Net, dense scatter estimation
- **DL for detector and source modelling / event selection**
  - [Sarrut2018, Sarrut2019, Zatcepin 2020, Sarrut2021]
  - Depth-of-interaction resolution in pixellated PET detectors



frontiers

Artificial Intelligence for Monte Carlo simulation in medical physics

David Sarrut [1,*], Ane Etxebeste [1], Enrique Muñoz [2], Nils Krah [1,2] and Jean Michel Létang [1]

[1] Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1294, F-69373, Lyon, France.
[2] University of Lyon, Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, F-69622, Villeurbanne, France

with **Monte Carlo** simulation

1

not be ready for clinic yet

s …

ty?

- Generalisation to other cases types?
- Robustness?

Examples of AI for Monte Carlo

- Example1: learning **Angular Response Function** for SPECT simulation
- Example2: learning **Phase-Space** for photon beam characterisation
- Example3: learning **Phase-Space** for SPECT imaging simulation

Deep learning **within** Monte Carlo simulation

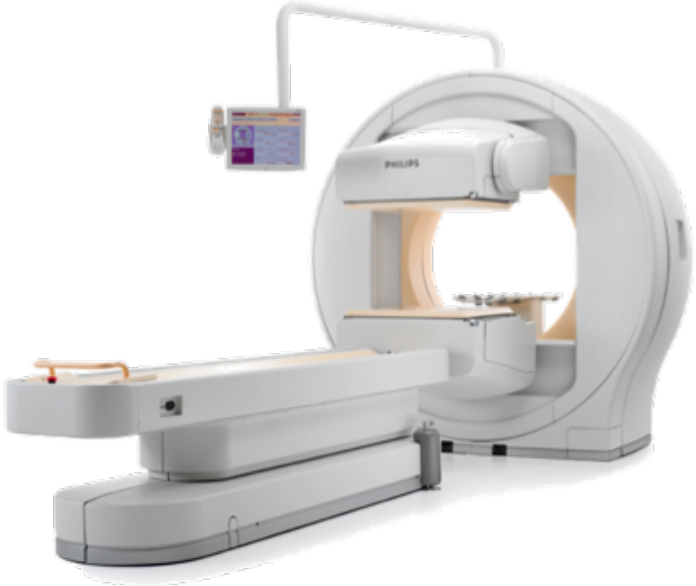
Example 1:
learning ARF for
SPECT simulation

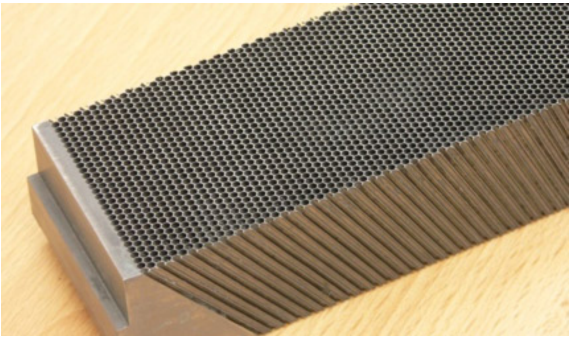
SPECT/CT imaging system

Radionuclide injection
^{99m}Tc, ^{177}Lu …

Emit gammas

Detect exiting gammas

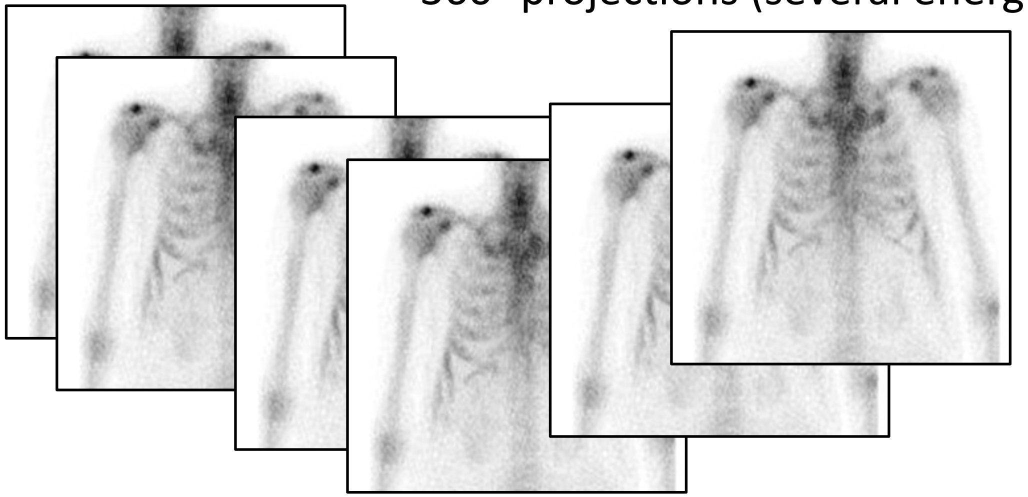


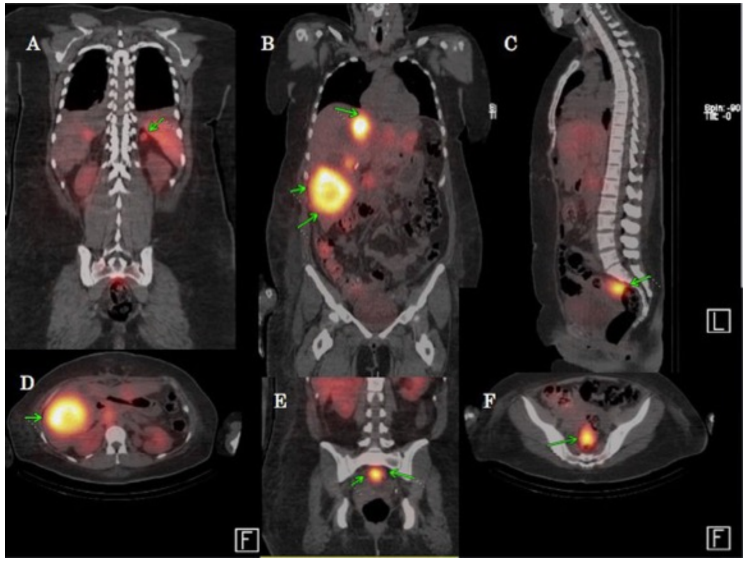


Collimator

+ scintillator detector (NaI, CsI, CZT)

360° projections (several energy windows)



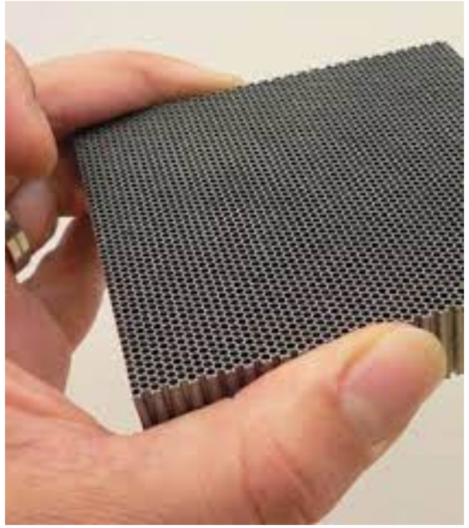


3D reconstruction (with CT)

SPECT Monte-Carlo simulation

- Long computation time
- Around 10^{-4} particles reaching detector
- Brute-force approach up to few days computation

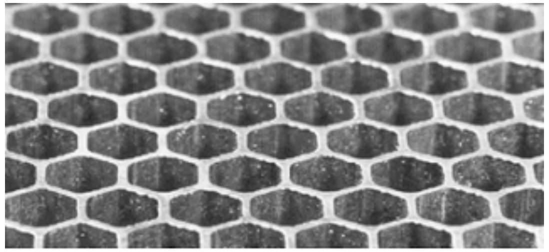
- Platforms:
  - SimSET [Harrison1993]
  - SIMIND [Ljungberg1989]
  - GATE/Geant4 [Sarrut2014]



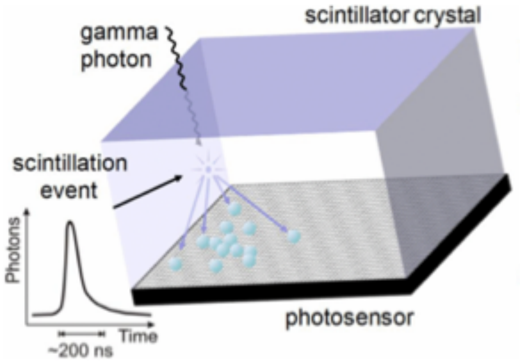
Collimator

SPECT Monte-Carlo simulation

- Several proposed Variance Reduction Techniques (VRT) :
  - GIS: Geometrical Importance Sampling [Beenhouwer2009]
  - ARF: Angular Response Function [Song2005, Descourt2010, Rydeen2018]
  - MPS: Multiple Projection Sampling [Beenhouwer2008, Liu2008]
  - CFD: Convolution Based Forced Detection [Liu2008]
  - FFD: Fixed Forced Detection [Cajgfinger2017]



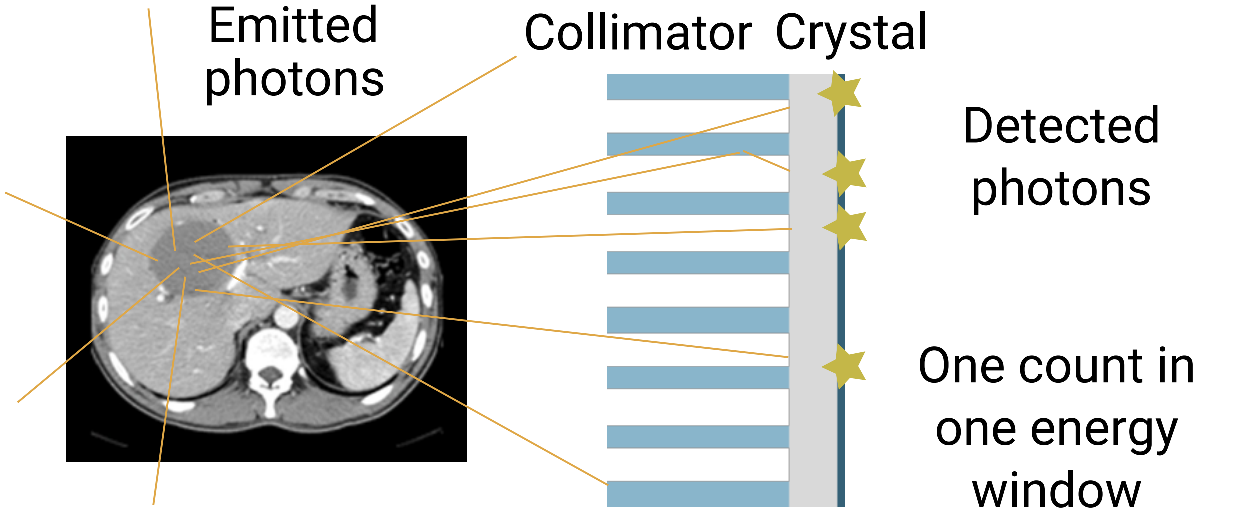
Collimator



[Braga2014]

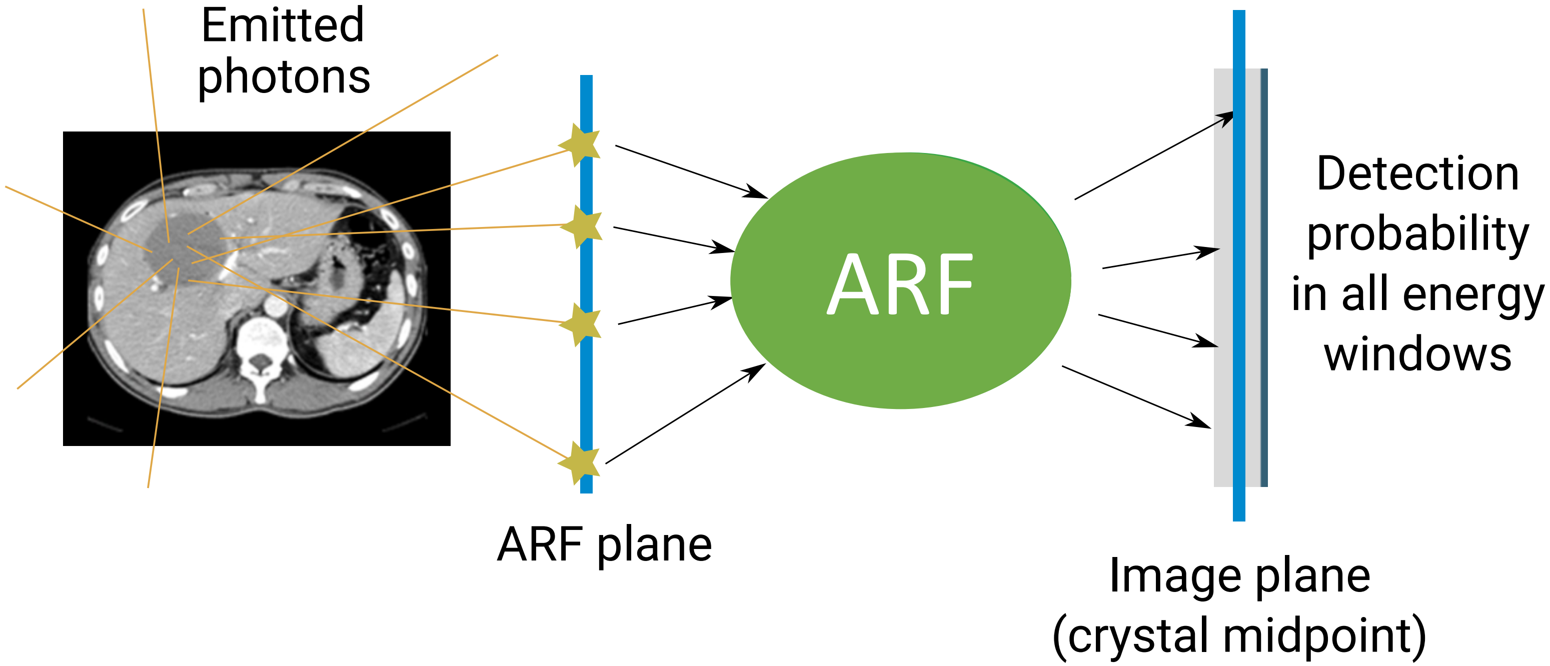
ARF: principles

Angular Response Function



- Replace SPECT head detection with tabulated response

  Incident particle at ARF plane
  use tables to get energy
  windows probabilities



- Assume:
  - Spatially invariant
  - Detection depends on direction + energy

ARF

Replace histogram tables by a neural network

- Advantages:
  - ARF tables needed to be computed only once
  - Variance reduction: probability instead of counts
  - Efficient, speedup x20-100 [Song2005, Descourt2010]

- Drawbacks:
  - ARF tables needed for every detector configurations
  - Large dataset needed to compute tables, 10^8 to 10^{11} [Rydeen2018]
  - Choice of table binning (3D histogram) not clear
  - Speedup not explicitly evaluated

Artificial neural network

- Learn a **predictive model** from a training dataset $h(\mathbf{x}) = \mathbf{y}$
- Input & output space
  - *in*: gamma energy and direction at the collimator entrance plane (3D)
  - *out*: probability the gamma is detected in the i^{th} energy window (nD)

$$h(E, \theta, \phi) = y_i$$

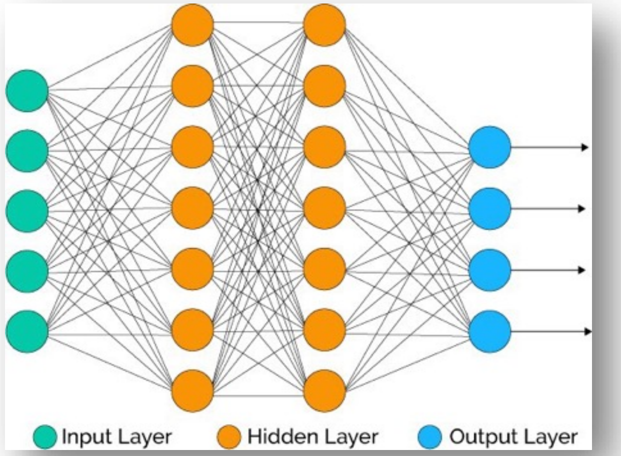
- Training dataset

  simulation, large source, complete energy spectra,
  complete detector (collimator/crystal)
  10^8 to 10^9 particles + **Russian Roulette**

Artificial neural network architecture

- 3 hidden linear fully connected layers
- 400 neurons by layer
- Activation function: ReLu
- Loss function: multiclass cross-entropy
- Optimisation: Adam [Kingma2014] (max 1000 iterations)
- Batch size: 5000 samples $\alpha = 0.0001$
- Adaptive learning rate

~ 10^5 weights (2 MB)



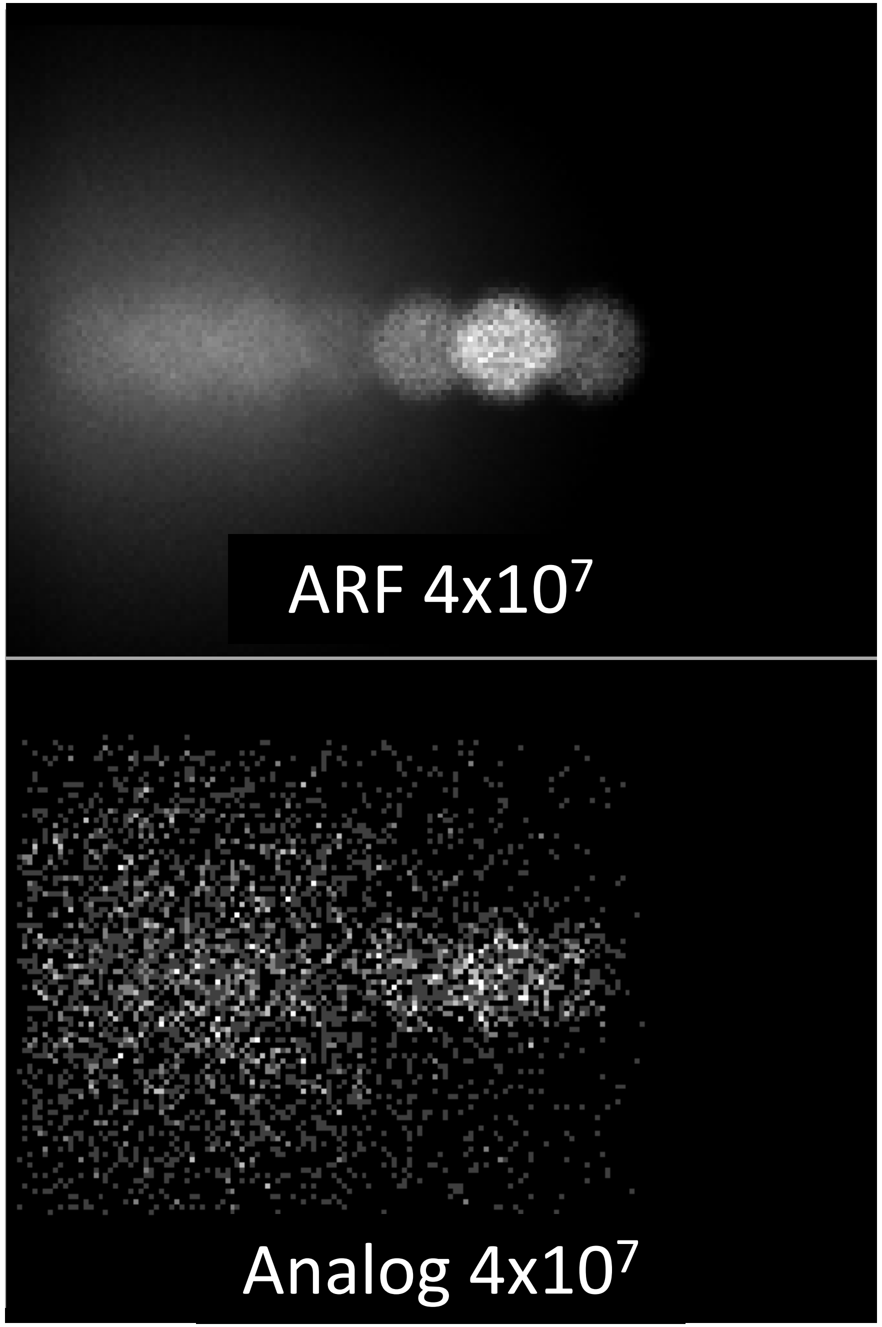
Results

- Simulation of 7 circular sources of different energies

- Efficiency

$$\varepsilon_k = \frac{1}{t \times \sigma_k^2}$$

Speedup: 20 – 1000



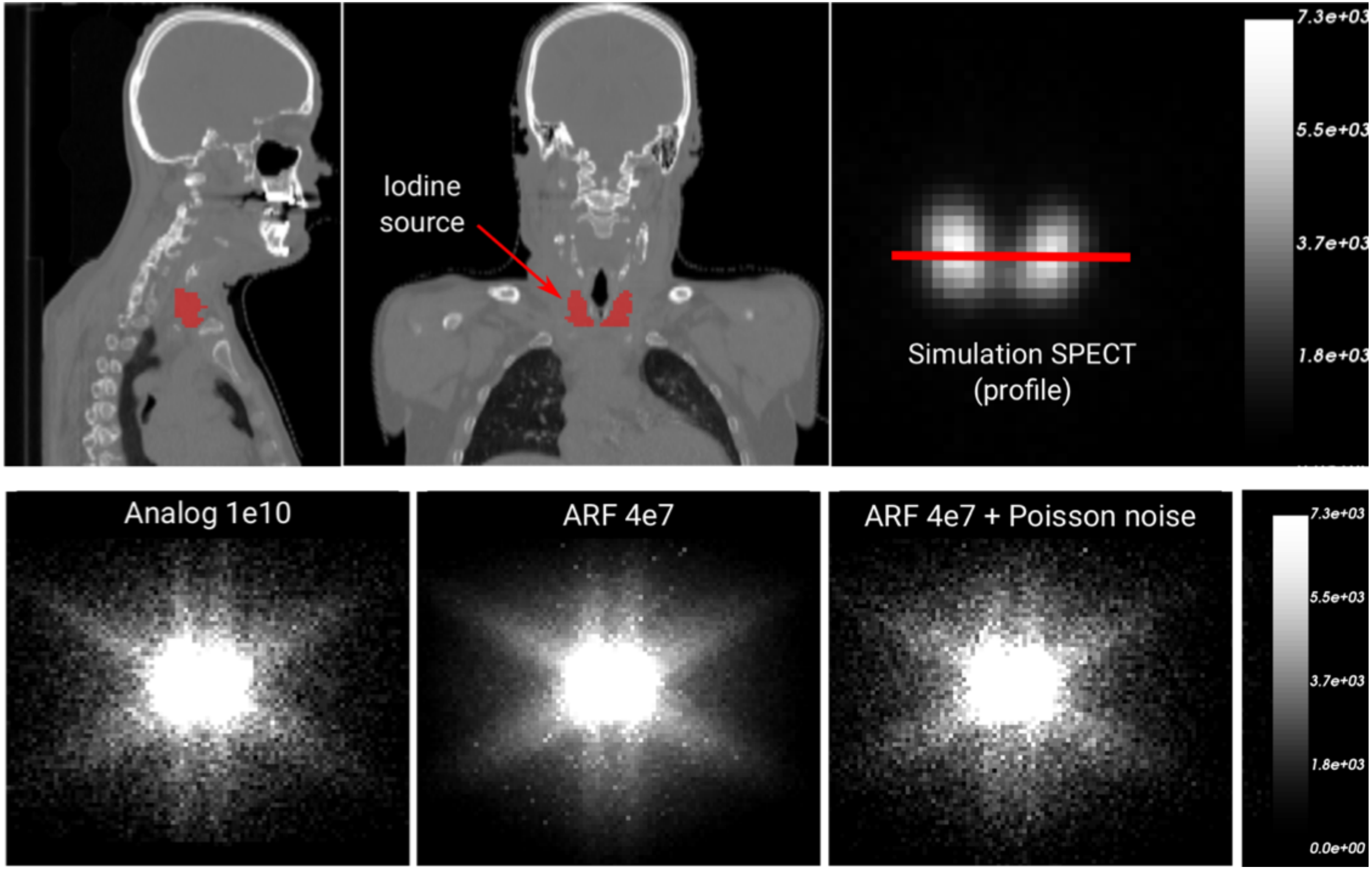
Results

25 days CPU time with 10^{10} particles

vs

2.5 hours with 4.10^7 particles

(> x200)



Example 1: conclusion

- Alternative approach to ARF by histogram using Artificial Neural Network
- Similar efficiency, require less data to build, more consistent (binning)
- Different noise distribution, need to add Poisson noise

- Available in GATE (open-source) www.opengatecollaboration.org

IOP Publishing | Phys. Med. Biol. 63 (2018) 205013 (12pp) | https://doi.org/10.1088/1361-6560/aae331

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

CrossMark

RECEIVED
11 June 2018

REVISED
17 September 2018

ACCEPTED FOR PUBLICATION
21 September 2018

PUBLISHED
17 October 2018

PAPER

Learning SPECT detector angular response function with neural network for accelerating Monte-Carlo simulations

D Sarrut[1], N Krah, J N Badel and J M Létang

Université de Lyon, CREATIS, CNRS UMR5220; Inserm U1044, INSA-Lyon, Université Lyon 1, Centre Léon Bérard, Lyon, France

[1] Author to whom any correspondence should be addressed.

E-mail: david.sarrut@creatis.insa-lyon.fr

Keywords: Monte-Carlo simulation, SPECT imaging, variance reduction technique, neural network

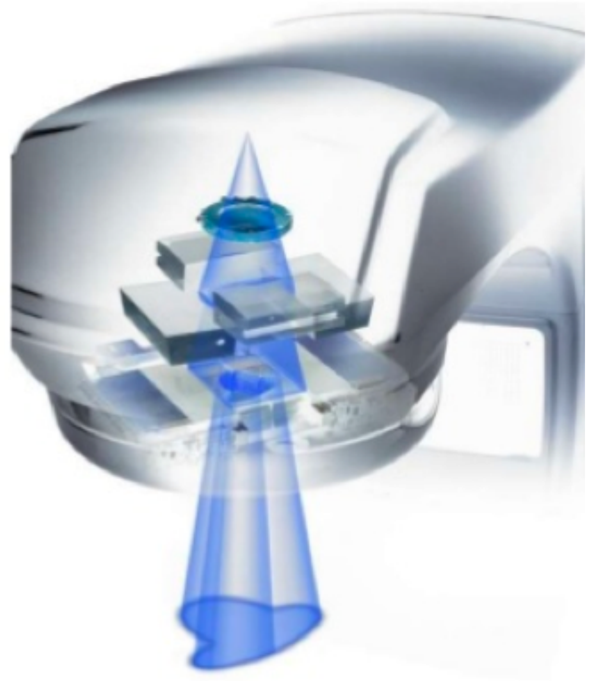
Abstract

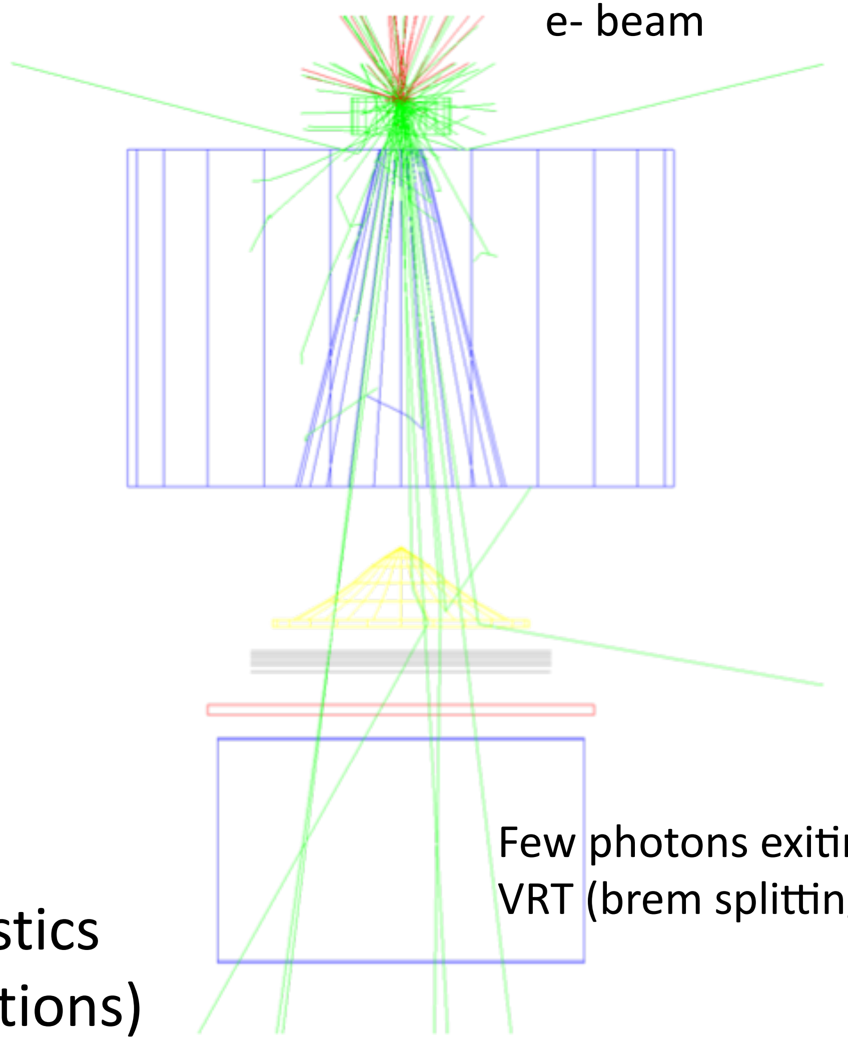
A method to speed up Monte-Carlo simulations of single photon emission computed tomography (SPECT) imaging is proposed. It uses an artificial neural network (ANN) to learn the angular response function (ARF) of a collimator–detector system. The ANN is trained once from a complete simulation including the complete detector head with collimator, crystal, and digitization process. In

Physics in Medicine and Biology, 2018

Example 2:
learning Linac phase-space

Radiation Therapy Linac head simulation





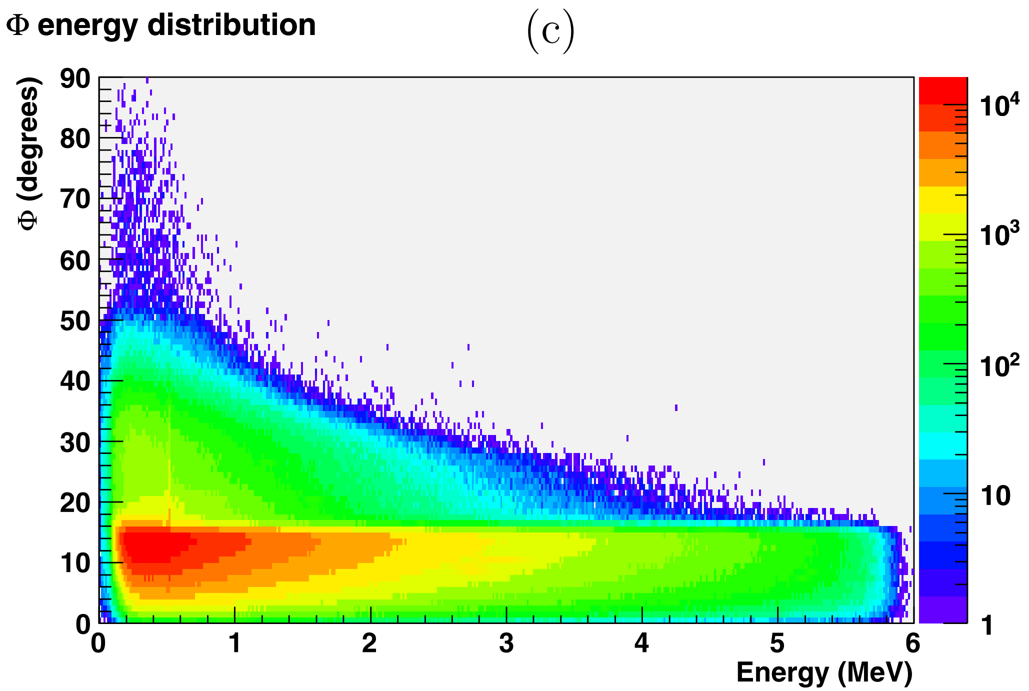
Goal: determine beam characteristics
(energy, position, direction distributions)

Phase Space (PHSP)

- Store beam properties as **Phase Space**
  - A PHSP is a list of particles (around 10^8, 10^9)
  - Properties: E, x, y, z, dx, dy, dz, w, (time)

- Advantages:
  - Computed only once
  - Fast to use
  - Can be shared

- Drawback
  - Several GB
  - When a cluster is used, should be shared among workers
  - Limited number of particles



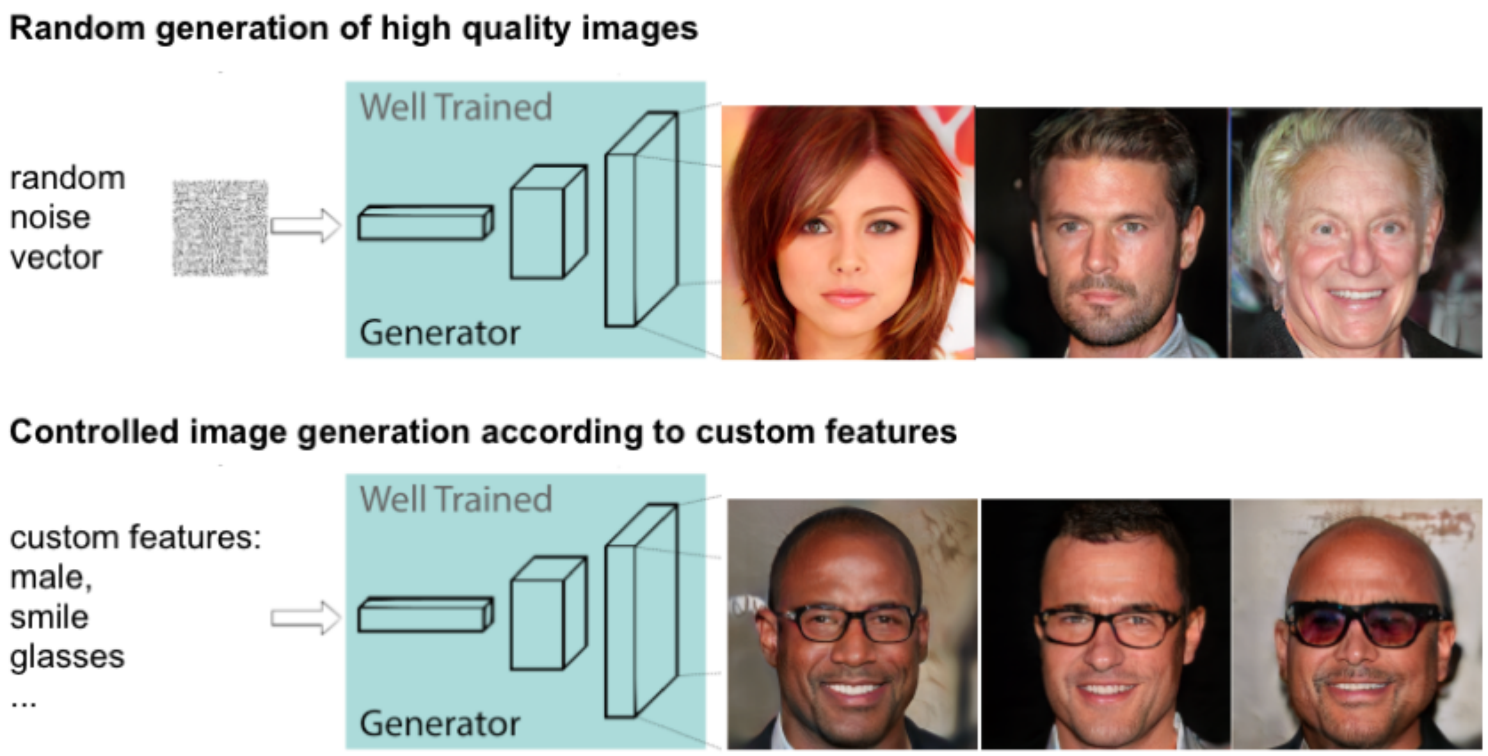
Example of dependence of direction ϕ and energy

GAN: Generative Adversarial Network

[Goodfellow, 2014]

Goal: « learn » a multidimensional probability distribution

Initial application :
artificial images generation



https://www.thispersondoesnotexist.com
https://www.thiscatdoesnotexist.com

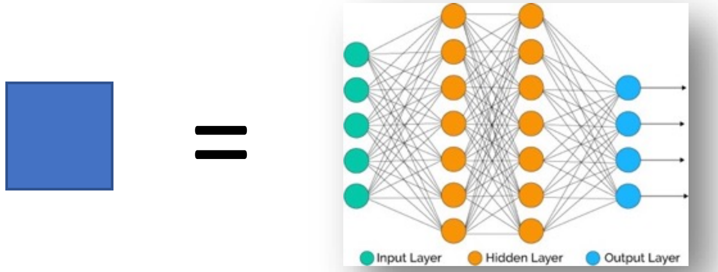
GAN: Generative Adversarial Network

- Training dataset $\boldsymbol{x} \in \mathbb{R}^d$
  - Dimension d=7 (E, X, Y, Z, dX, dY, dZ)
  - Samples of an unknown distribution p_{real}
- Generator $G(\boldsymbol{z}; \boldsymbol{\theta}_G)$

- Discriminator $D(\boldsymbol{x}; \boldsymbol{\theta}_D)$

GAN: Generative Adversarial Network

- Training dataset $\boldsymbol{x} \in \mathbb{R}^d$
  - Dimension d=7 (E, X, Y, Z, dX, dY, dZ)
  - Samples of an unknown distribution p_{real}



Neural network

$$\boldsymbol{z} \rightarrow \mathrm{G} \rightarrow \tilde{\boldsymbol{x}}$$

$$\boldsymbol{x} \rightarrow \mathrm{D} \rightarrow 1|0$$

Alternate G and D optimisation updates

Loss function

- GAN notoriously difficult to train
- GAN zoo … https://github.com/hindupuravinash/the-gan-zoo
- Alternative formulations: Wasserstein GAN [Arjovsky 2017]

- “Earth-mover” distance (EMD) : cost of the optimal transport
- Un-tracktable in practice, but approximated:

$$J_D(\boldsymbol{\theta}_D, \boldsymbol{\theta}_G) = \mathbb{E}_{\boldsymbol{z}}[D(G(\boldsymbol{z}))] - \mathbb{E}_{\boldsymbol{x}}[D(\boldsymbol{x})]$$
$$J_G(\boldsymbol{\theta}_D, \boldsymbol{\theta}_G) = -\mathbb{E}_{\boldsymbol{z}}[D(G(\boldsymbol{z}))]$$

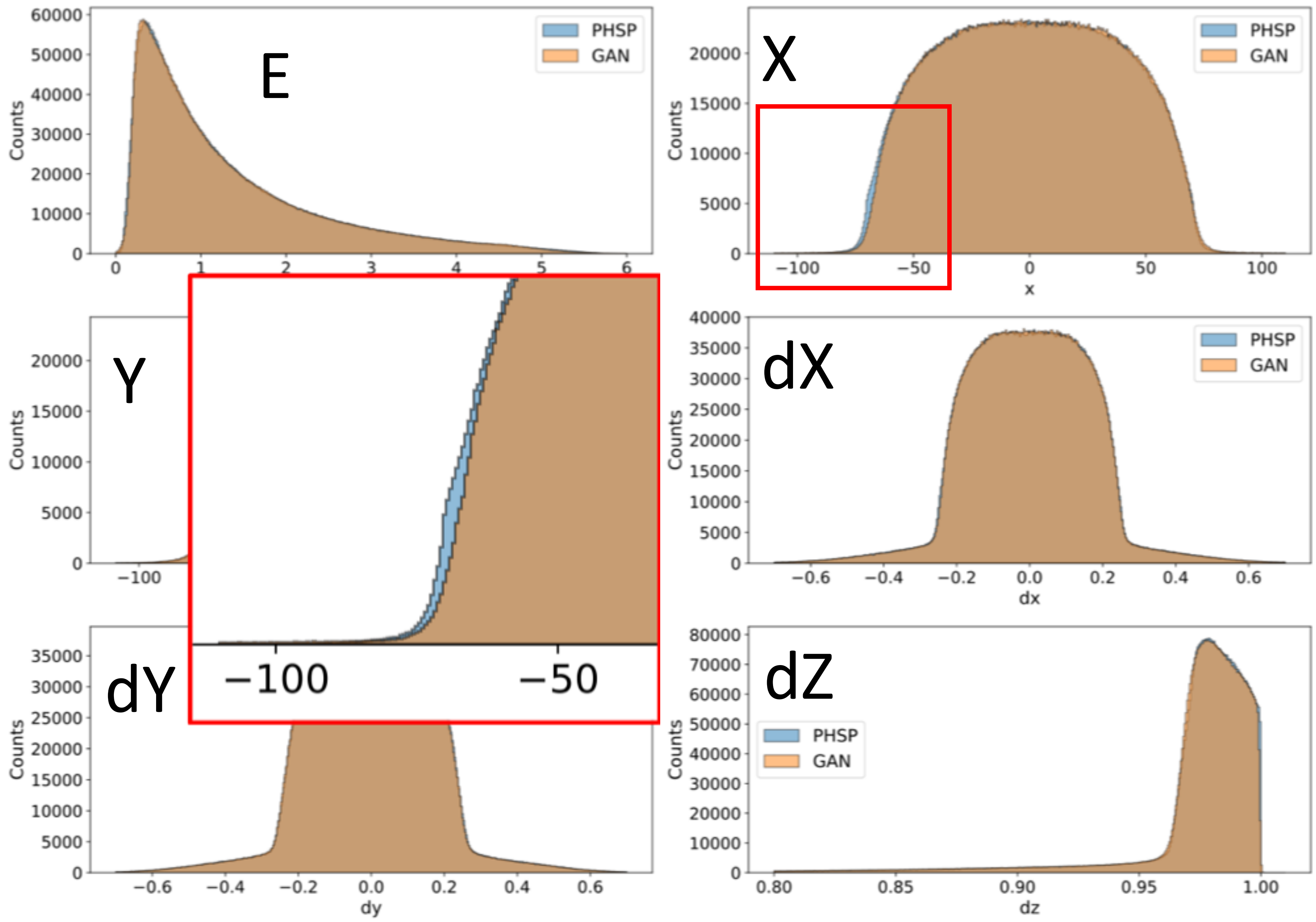
Experiments

PHSP from IAEA web site

PHSP	Size	Nb of particles
Elekta PRECISE 6MV	2 files of 3.9 GB	1.3×10^8 photons each file
CyberKnife IRIS 60mm	2 files of 1.6 GB	5.8×10^7 photons each file

Results

Marginal distributions of the 6 parameters obtained from the reference PHSP and from the GAN, for Elekta 6MV linac.

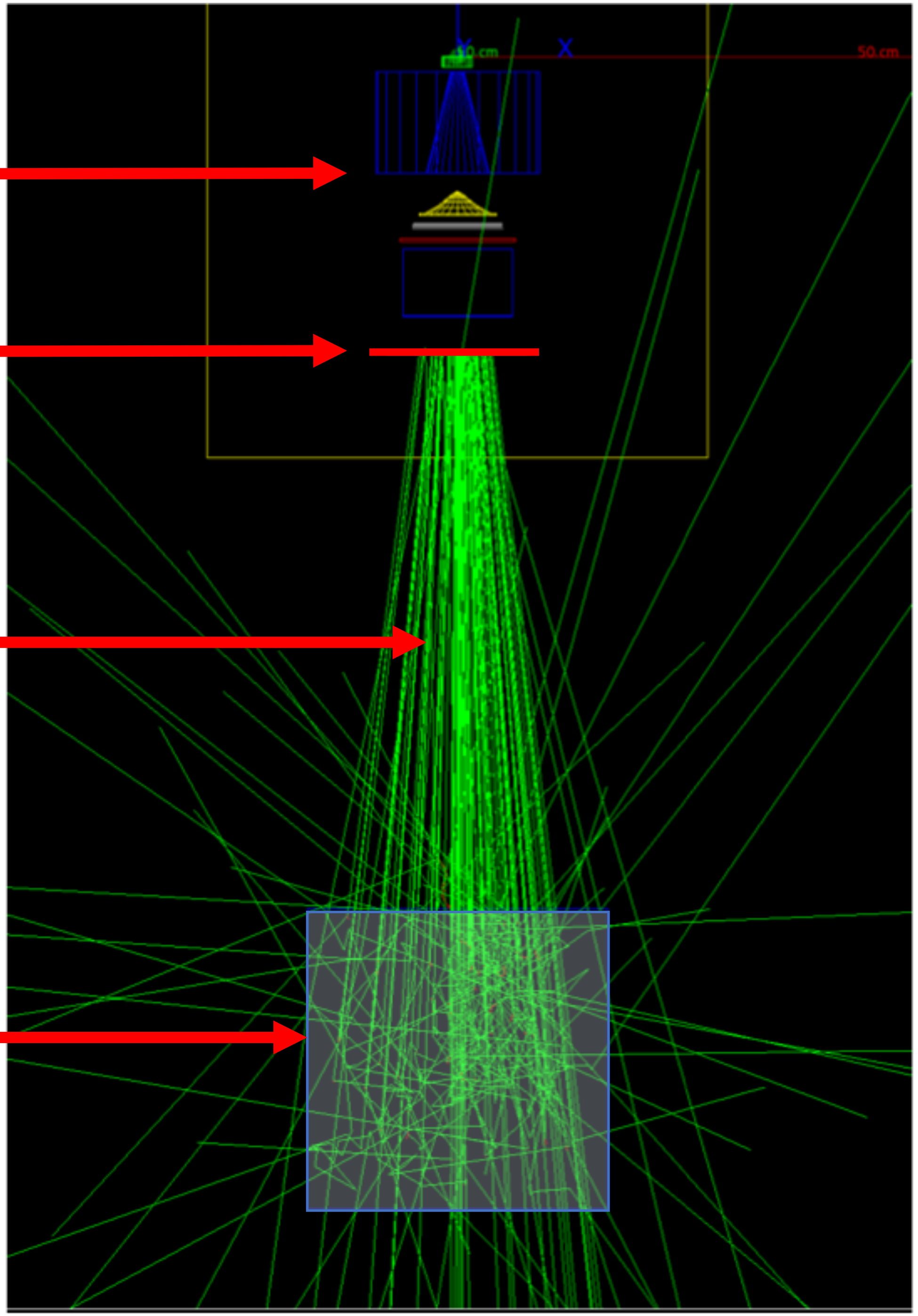


Results

- Dose distribution in water from PHSP
  10^8 primary photons

- Compare dose between:
  1. PHSP1 vs PHSP2
  2. PHSP1 vs GAN

- Voxel by voxel dose comparison



LINAC head

PHSP plane

Beam

Waterbox

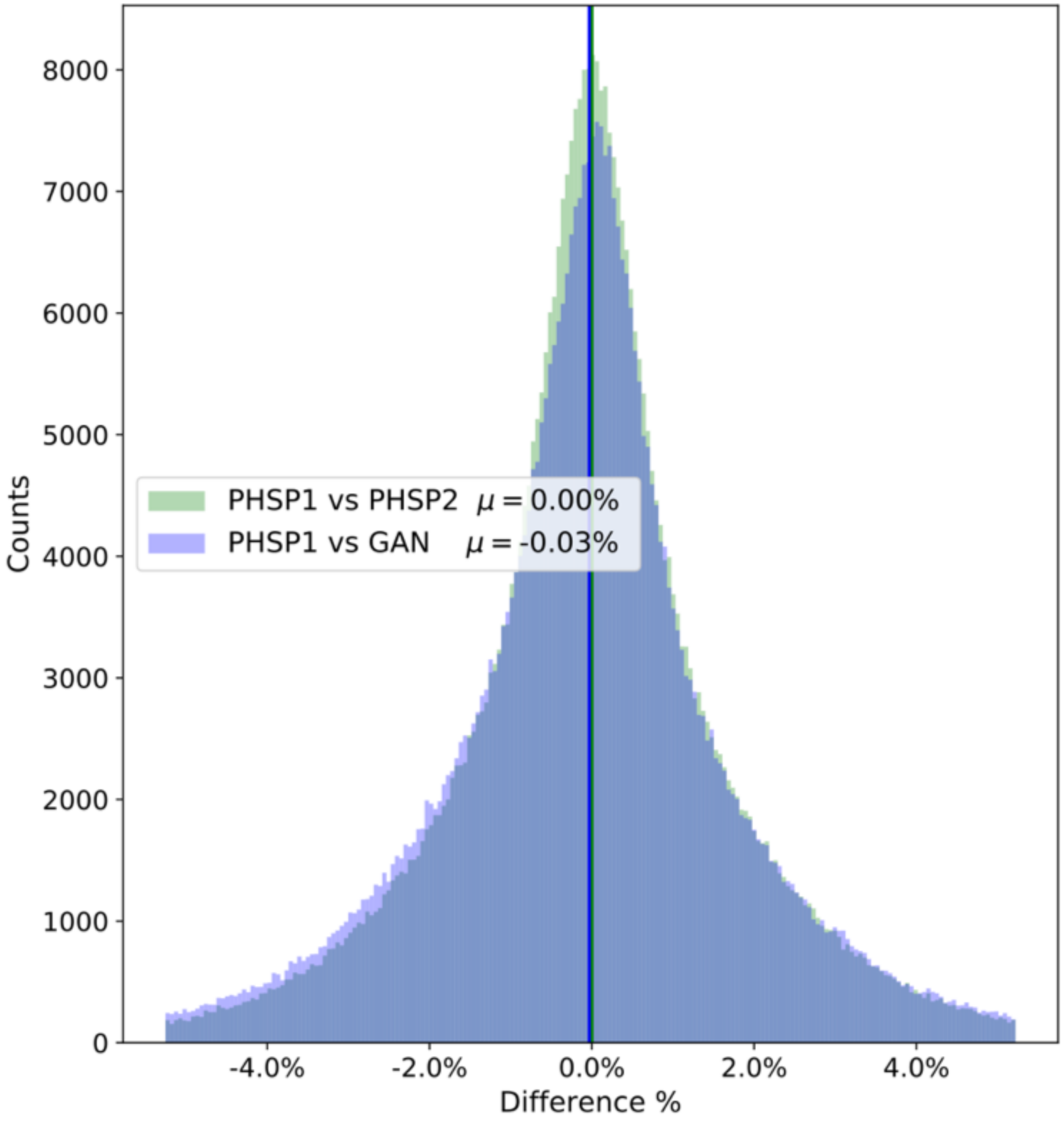
Results

Distributions of relative differences between

- PHSP1 and PHSP2
- PHSP1 and GAN

Vertical lines indicate the mean differences

Difference relative to the prescribed dose



Example 2: conclusion

- Using GAN to represent a Phase-Space is feasible
- Final GAN model: few MB (vs PHSP = 4 GB)
- Sufficient for dose computation
- Training is difficult: hyperparameters, 511 keV peak, …

- Available in GATE www.opengatecollaboration.org
- Perspectives :
  - Could it be learned from less particles ?
  - Detailed statistical analysis in progress
  - Other applications of GAN within MC simulations

IOP Publishing Phys. Med. Biol. 64 (2019) 215004 (11pp) https://doi.org/10.1088/1361-6560/ab3fc1

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

CrossMark

RECEIVED
23 April 2019

REVISED
27 August 2019

ACCEPTED FOR PUBLICATION
30 August 2019

PUBLISHED
23 October 2019

PAPER

Generative adversarial networks (GAN) for compact beam source modelling in Monte Carlo simulations

D Sarrut[1], N Krah[1,2] and J M Létang[1]

[1] Université de Lyon, CREATIS; CNRS UMR5220, Inserm U1044, INSA-Lyon, Université Lyon 1, Centre Léon Bérard, France

[2] Université Lyon, CNRS/IN2P3 UMR 5822, IP2I Lyon, Villeurbanne, France

E-mail: david.sarrut@creatis.insa-lyon.fr

Keywords: Monte-Carlo simulation, generative adversarial network, phase-space, linac

Abstract

A method is proposed and evaluated to model large and inconvenient phase space files used in Monte Carlo simulations by a compact generative adversarial network (GAN). The GAN is trained based on a phase space dataset to create a neural network, called Generator (G), allowing G to mimic the multidimensional data distribution of the phase space. At the end of the training process, G is stored with about 0.5 million weights, around 10 MB, instead of a few GB of the initial file. Particles are then generated with G to replace the phase space dataset.

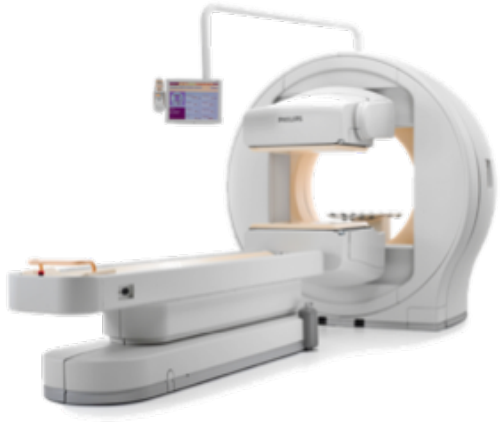
This concept is applied to beam models from linear accelerators (linacs) and from brachytherapy seed models. Simulations using particles from the reference phase space on one hand and those generated by the GAN on the other hand were compared. 3D distributions of deposited energy

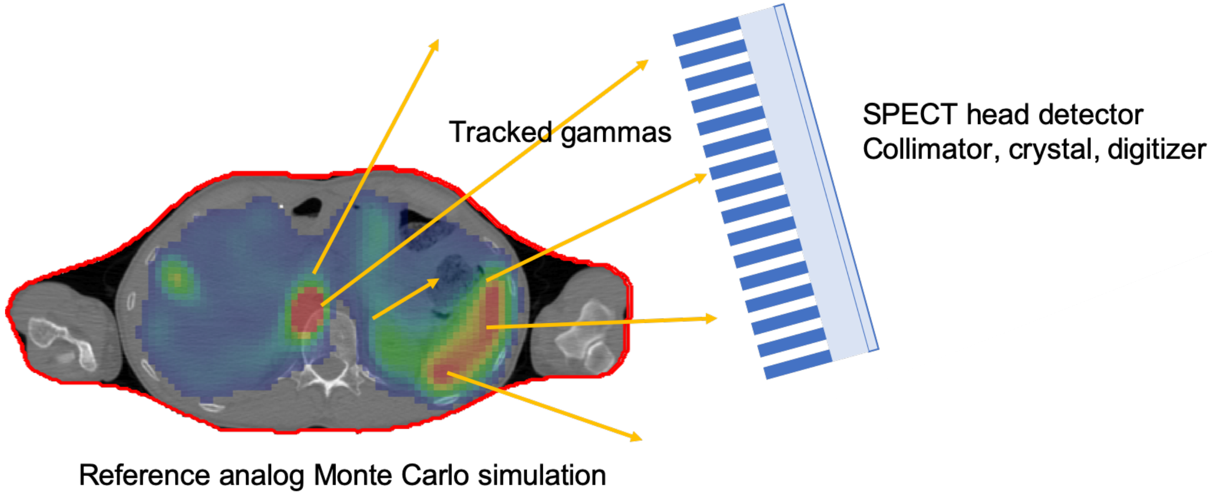
Physics in Medicine and Biology, 2019

Example 3:
learning phase-space
for SPECT

SPECT simulation

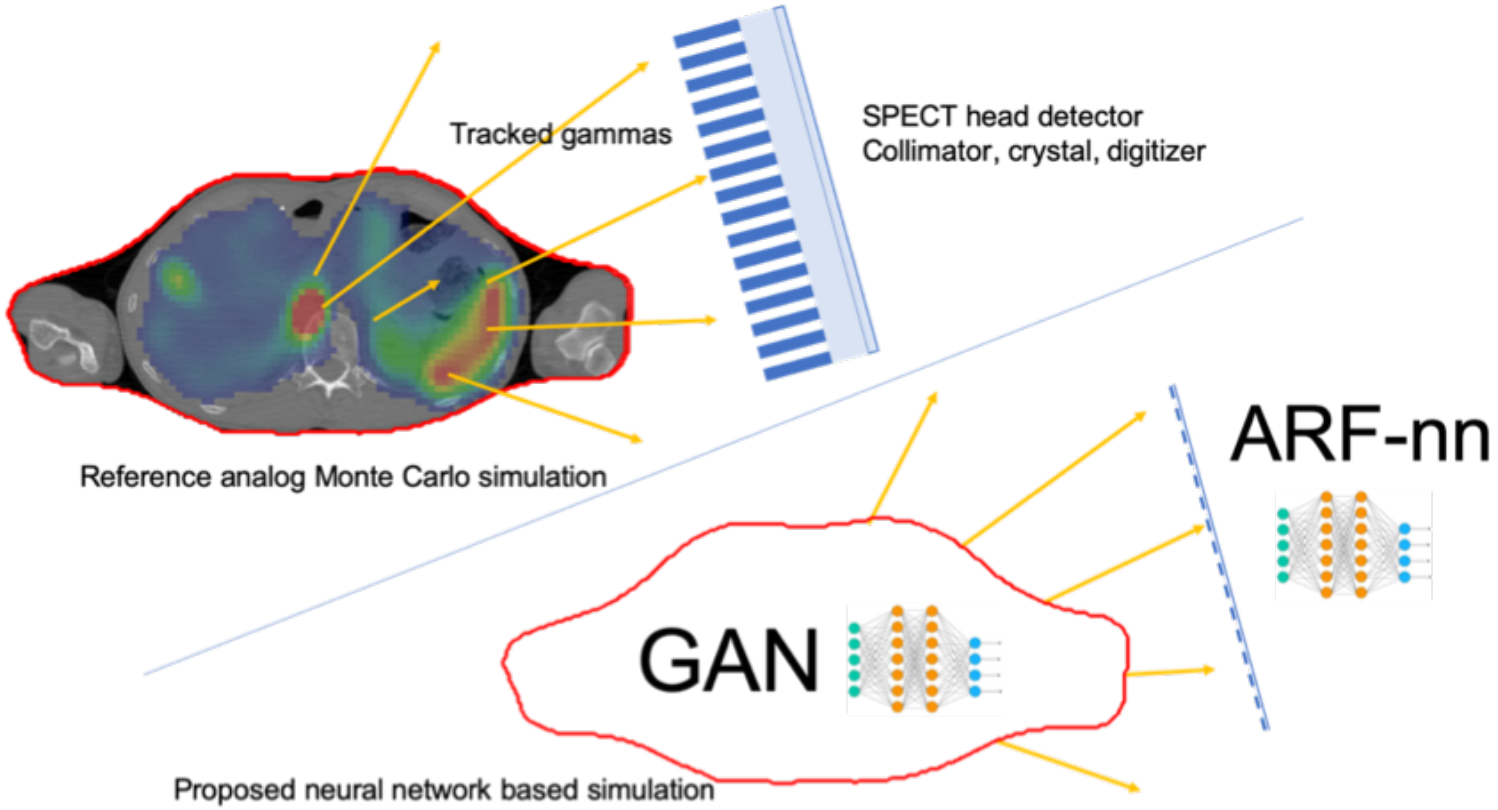
- Part1: previously, detector response (ARF)
- Part2: from emission to patient exiting gamma





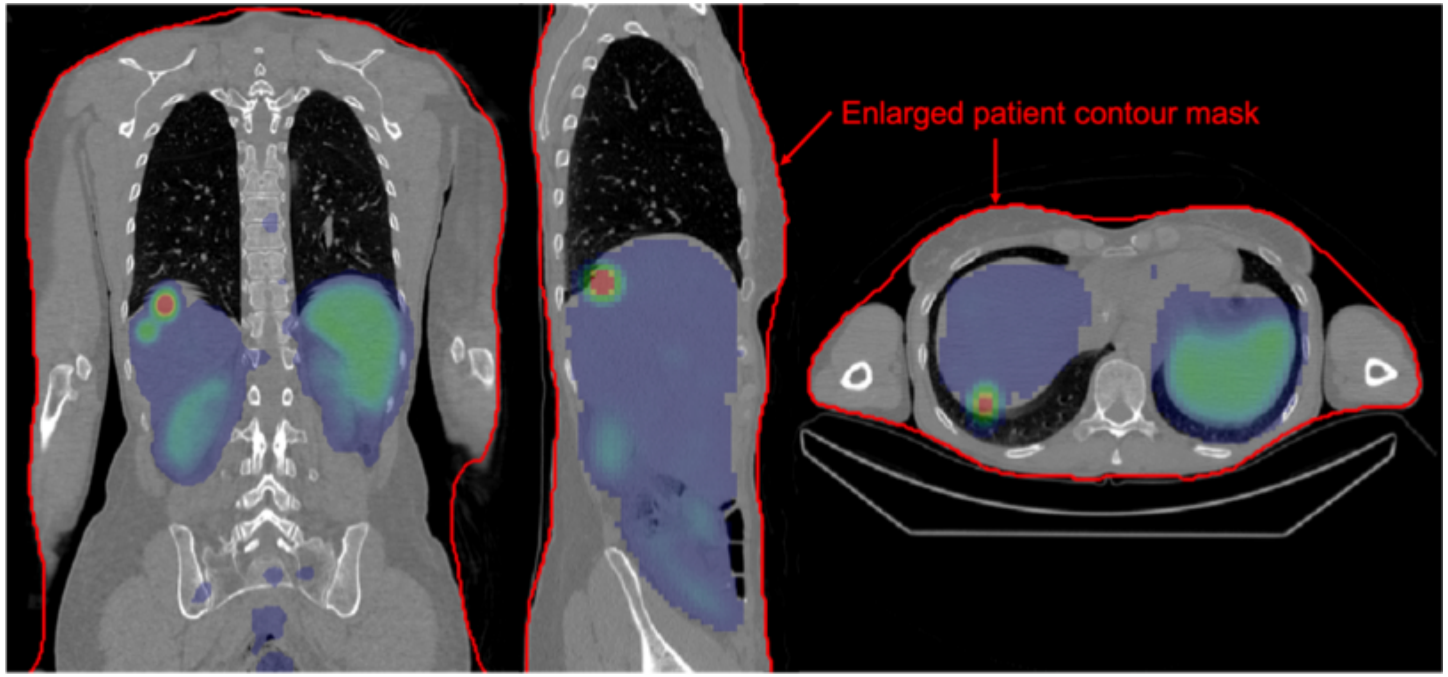
Training dataset

- Step1: generate low stat dataset
- Step2: train a GAN
- Step3:
  - Use GAN a source
  - Use ARF as detector



Learning exiting gamma ?

- Track particles:
  - From activity source …
  - … to patient skin
- Allows to
  - Consider scatter
  - Consider complex source
  - Consider pharmakocinetic
- Store
  - E, position, direction



GAN

- Wasserstein GAN
- Several Gradient Penalties

$$\text{WGAN Loss} = \mathop{\mathbb{E}}_{\tilde{\mathbf{x}}\sim\mathbb{P}_g}[D(\tilde{\mathbf{x}})] - \mathop{\mathbb{E}}_{\mathbf{x}\sim\mathbb{P}_r}[D(\mathbf{x})] + \lambda \mathop{\mathbb{E}}_{\hat{\mathbf{x}}\sim\mathbb{P}_{\hat{\mathbf{x}}}}[\text{GP}]$$

- 4 hidden layers
- 700 neurons / layers
- $2\text{x}10^6$ parameters
- 10^5 epoch

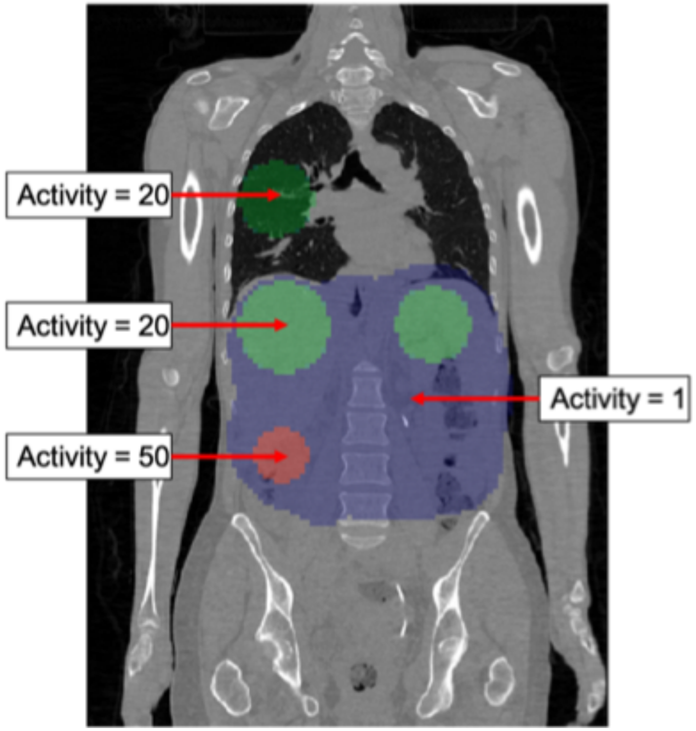
Grad. Pen.	Least Square	Hinge
L1	$(\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_1 - 1)^2$	$\max\{0, (\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_1 - 1)\}$
L2	$(\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_2 - 1)^2$ [18]	$\max\{0, (\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_2 - 1)\}$
L^∞	$(\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_\infty - 1)^2$	$\max\{0, (\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_\infty - 1)\}$
Square Hinge	$(\max\{0, (\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_2 - 1)\})^2$ [19]	
0-GP	$(\|\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|\|_2)^2$ [20]	

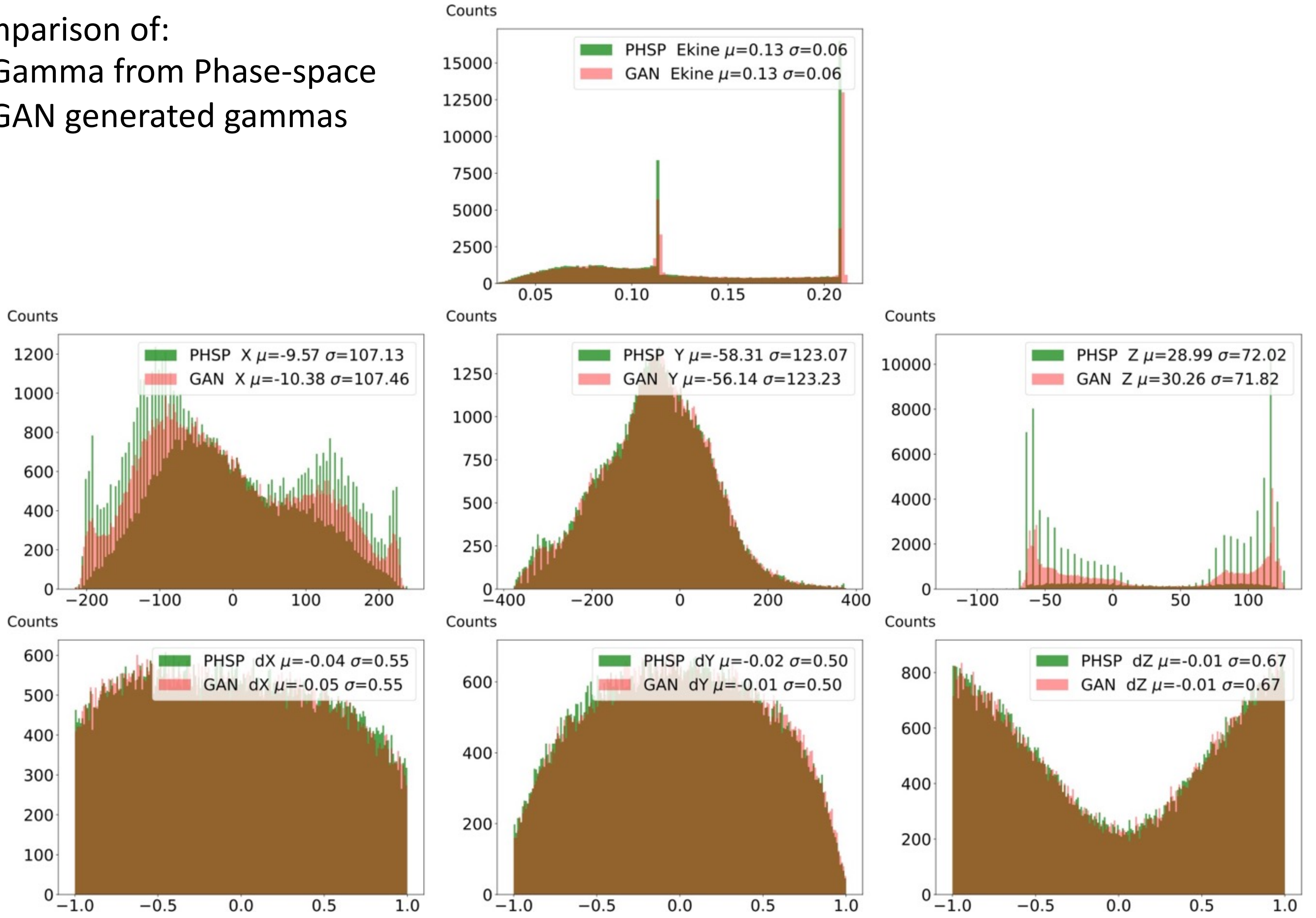
Table 1. Gradient penalties according to [21, 18, 20, 19]. In the equations, $\hat{\mathbf{x}} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$, with $\mathbf{x}$ sampled from $\mathbb{P}_r$ the *real* probability distribution of the gammas from the training dataset, and $\mathbf{y}$ is sampled from $\mathbb{P}_g$ the *generated* gamma distribution. $\alpha \sim \mathcal{U}(0,1)$ is sampled from the unit hyperball (following notation of [18]).

Results

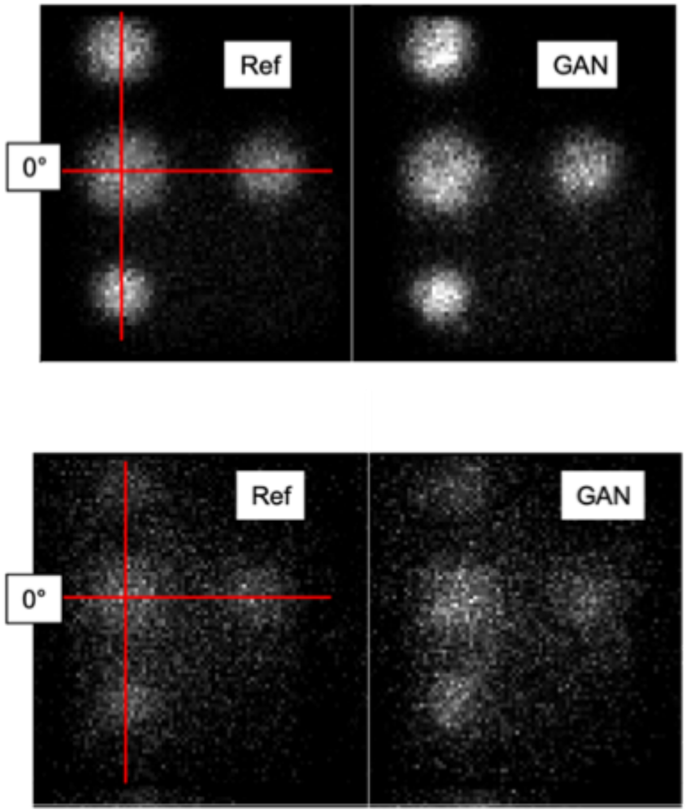
Comparison of:

- Gamma from Phase-space
- GAN generated gammas

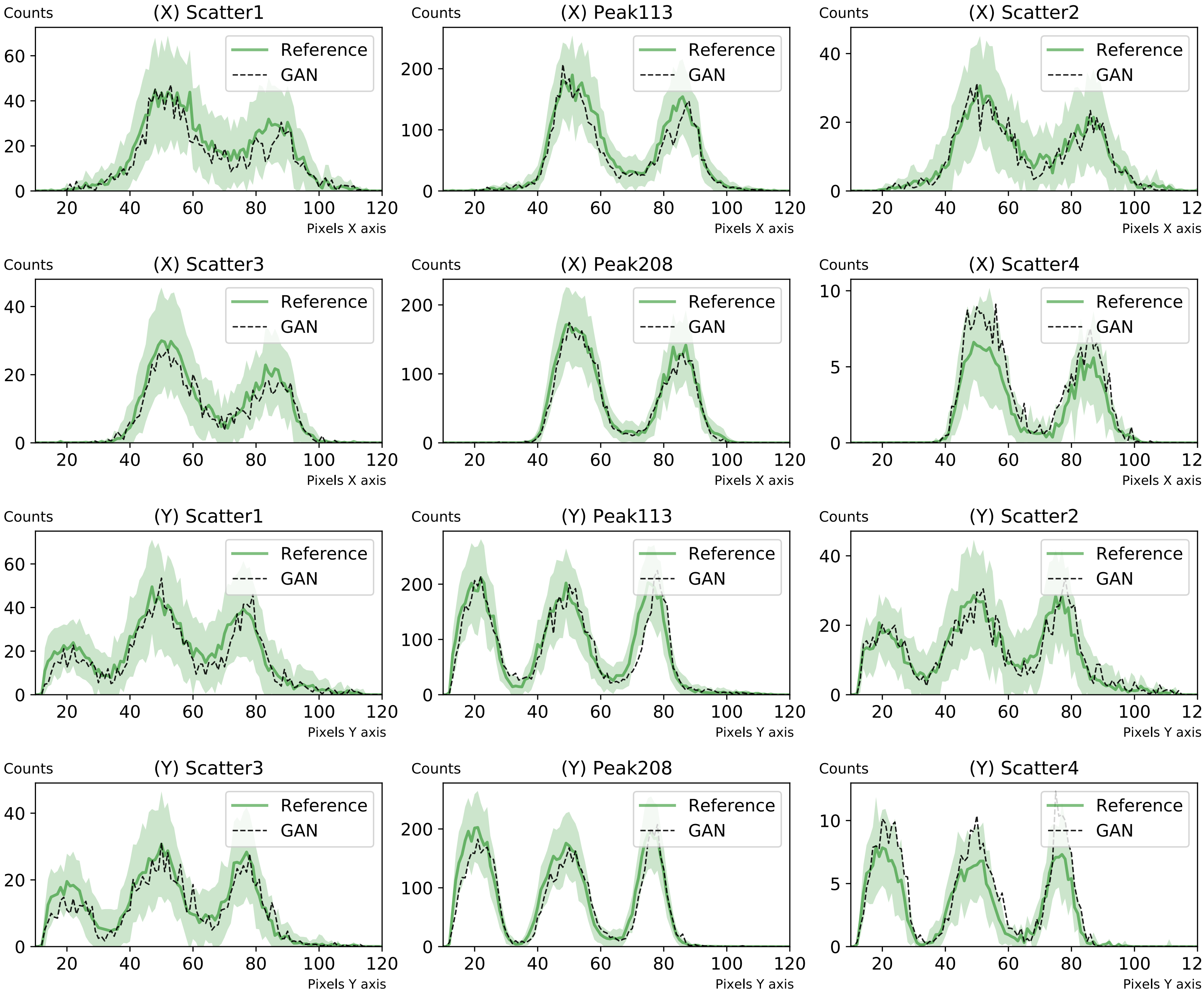




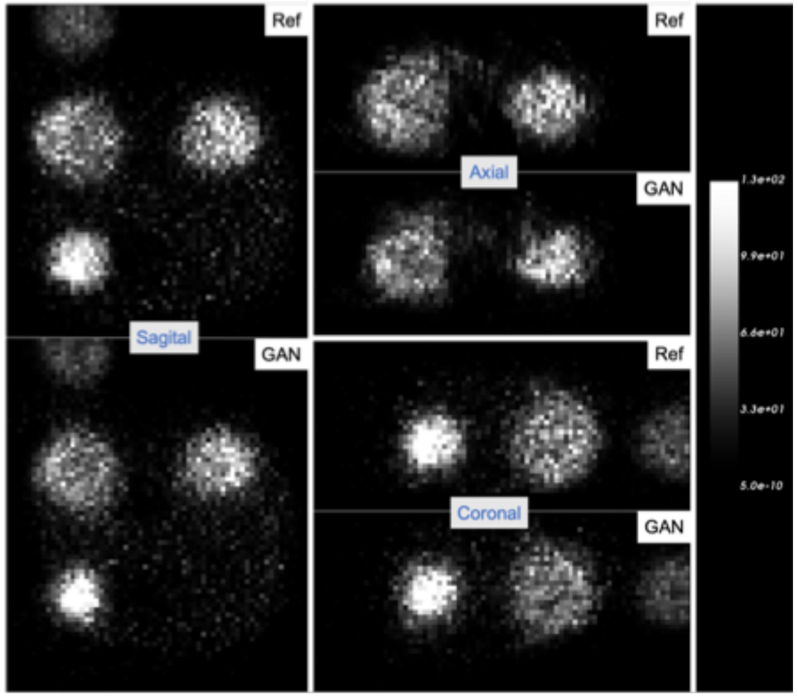
Results



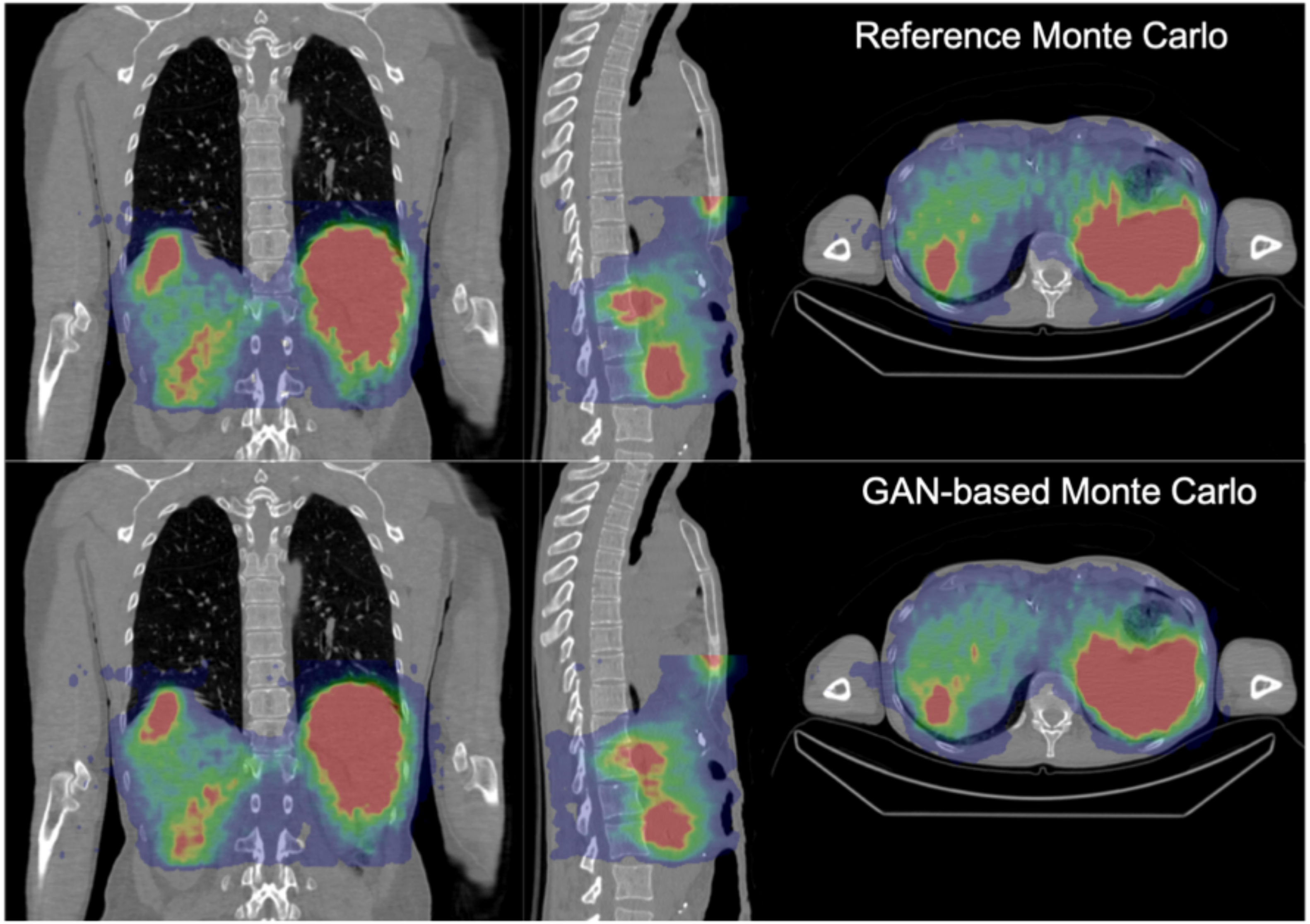
2D projections



Results



Test1: 3D reconstruction



Test2: 3D reconstruction

Discussion

- Full ^{177}Lu treatment (4 GBq)
- Analog computation time
  - Around PPS = 5k (Particle Per Second)
  - Around 1 day CPU for 5e8 particles per projection x 60 projections
- Computation time
  - Low stats Monte Carlo simulation for training dataset: 10 h (4 GB)
  - Training GAN: 4h (GPU)
  - Using GAN: PPS = 600k

Discussion

- Feasible
- Accuracy: is it sufficient ?
- Time gain ? To compare to other VRT
- Available in GATE www.opengatecollaboration.org

- GAN training
  - Size of training dataset ?
  - Gradient penalty ?
  - How to optimize learning ?
  - Transfert learning ?
  - How about one patient to the other ?

IOP Publishing   Phys. Med. Biol. 66 (2021) 055014   https://doi.org/10.1088/1361-6560/abde9a

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

CrossMark

RECEIVED
5 October 2020

REVISED
14 January 2021

ACCEPTED FOR PUBLICATION
21 January 2021

PUBLISHED
19 February 2021

PAPER

Modeling complex particles phase space with GAN for Monte Carlo SPECT simulations: a proof of concept

D Sarrut[1], A Etxebeste[1], N Krah[1,2] and JM Létang[1]

[1] Université de Lyon, CREATIS, CNRS UMR5220, Inserm U1044, INSA-Lyon, Université Lyon 1, Centre Léon Bérard 69373, France
[2] University of Lyon, Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, F-69622, Villeurbanne, France

E-mail: david.sarrut@creatis.insa-lyon.fr

Keywords: Monte-Carlo simulation, generative adversarial network, SPECT, phase space

Abstract

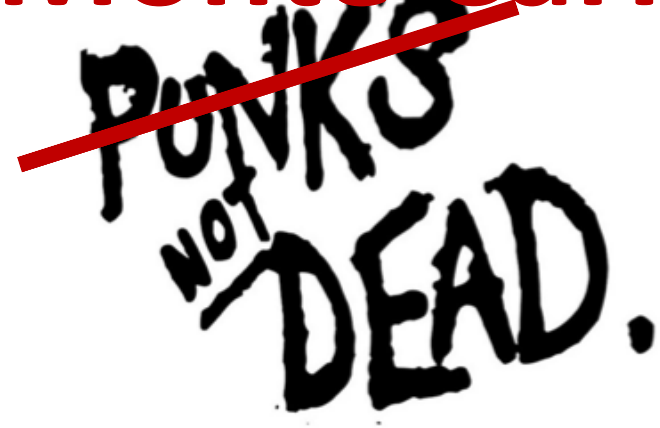
A method is proposed to model by a generative adversarial network the distribution of particles exiting a patient during Monte Carlo simulation of emission tomography imaging devices. The resulting compact neural network is then able to generate particles exiting the patient, going towards the detectors, avoiding costly particle tracking within the patient. As a proof of concept, the method is evaluated for single photon emission computed tomography (SPECT) imaging and combined with another neural network modeling the detector response function (ARF-nn). A complete rotating SPECT acquisition can be simulated with reduced computation time compared to conventional Monte Carlo simulation. It also allows the user to perform simulations with several imaging systems or parameters, which is useful for imaging system design.

Physics in Medicine and Biology, 2021

opengatecollaboration.org

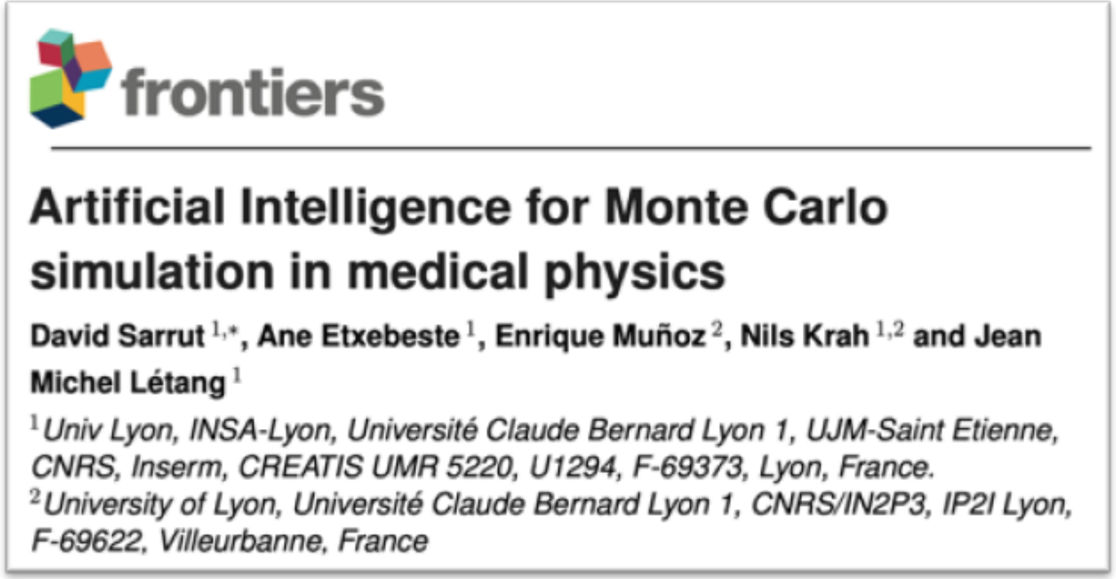
General conclusion

Monte Carlo



- **AI** may (also) be useful to **MC simulations**
  - ARF, GAN for phase-space, …
  - Faster, smoother, stronger
- Still experimental
- Available in GATE www.opengatecollaboration.org

- New challenges
  - Learning dataset size ?
  - Learning time ?
  - Transfer learning ?
  - Conditional learning ?
  - Convergence guarantee ?
  - Final Accuracy ?


frontiers

Artificial Intelligence for Monte Carlo simulation in medical physics

David Sarrut[1,*], Ane Etxebeste[1], Enrique Muñoz[2], Nils Krah[1,2] and Jean Michel Létang[1]

[1]*Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1294, F-69373, Lyon, France.*
[2]*University of Lyon, Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, F-69622, Villeurbanne, France*

Thanks for your attention !

CREATIS Lyon, France

GATE
opengatecollaboration.org

LyriCAN

cnrs
dépasser les frontières

PRIMES
UNIVERSITE DE LYON

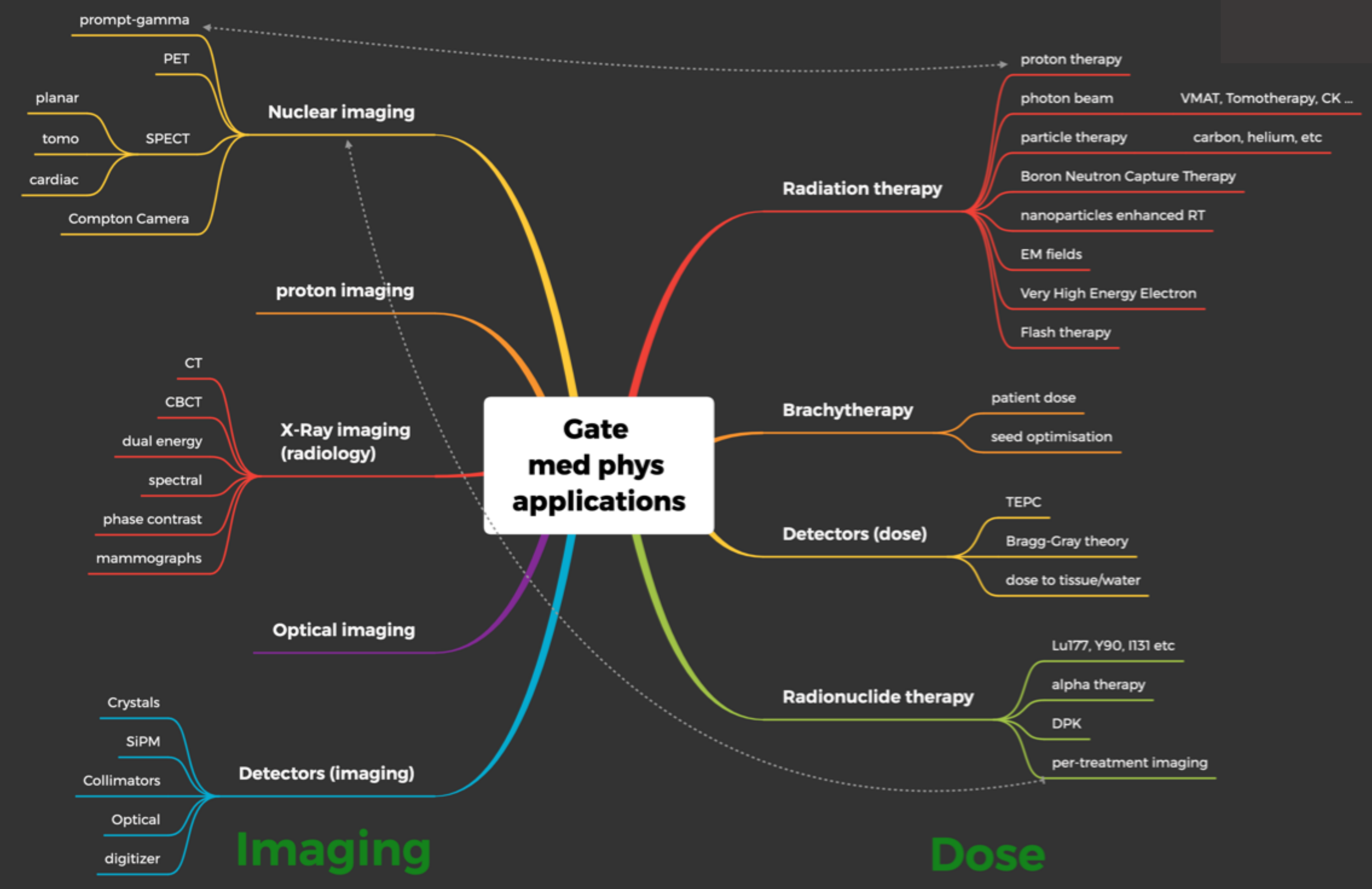
CENTRE DE LUTTE CONTRE LE CANCER LEON BERARD

GATE
opengatecollaboration.org

GEANT4
A SIMULATION TOOLKIT

https://github.com/OpenGATE/Gate

https://opengate.readthedocs.io


Gate
med phys
applications
Nuclear imaging
prompt-gamma
PET
SPECT
planar
tomo
cardiac
Compton Camera
proton imaging
X-Ray imaging
(radiology)
CT
CBCT
dual energy
spectral
phase contrast
mammographs
Optical imaging
Detectors (imaging)
Crystals
SiPM
Collimators
Optical
digitizer
Imaging
Radiation therapy
proton therapy
photon beam
VMAT, Tomotherapy, CK ...
particle therapy
carbon, helium, etc
Boron Neutron Capture Therapy
nanoparticles enhanced RT
EM fields
Very High Energy Electron
Flash therapy
Brachytherapy
patient dose
seed optimisation
Detectors (dose)
TEPC
Bragg-Gray theory
dose to tissue/water
Radionuclide therapy
Lu177, Y90, I131 etc
alpha therapy
DPK
per-treatment imaging
Dose

Results: correlation

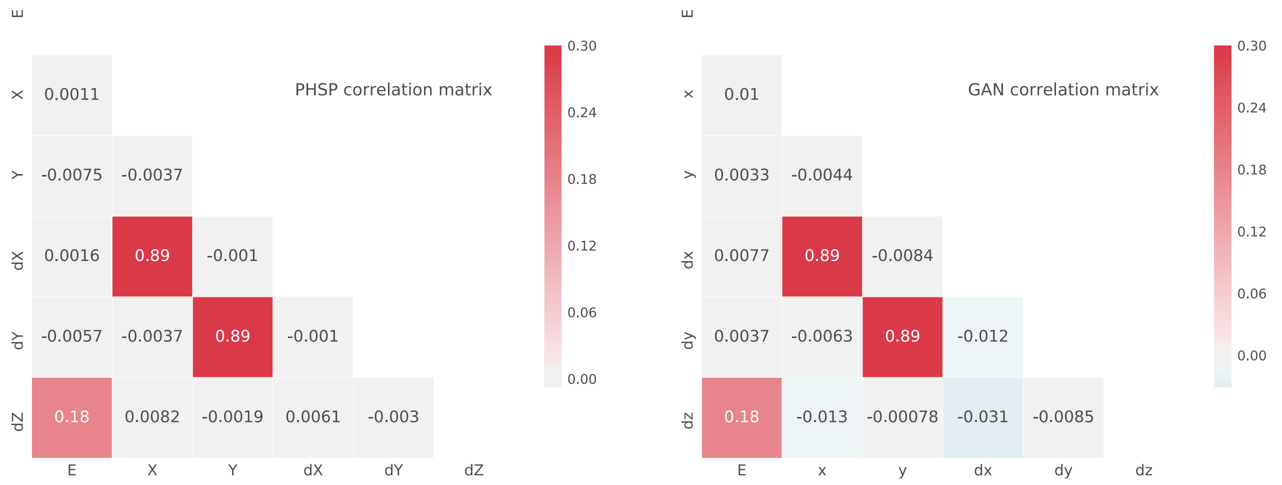


Figure 3. Correlation matrices between all six parameters for phase space file (PHSP, left) and GAN (right), for Elekta 6 MV linac.

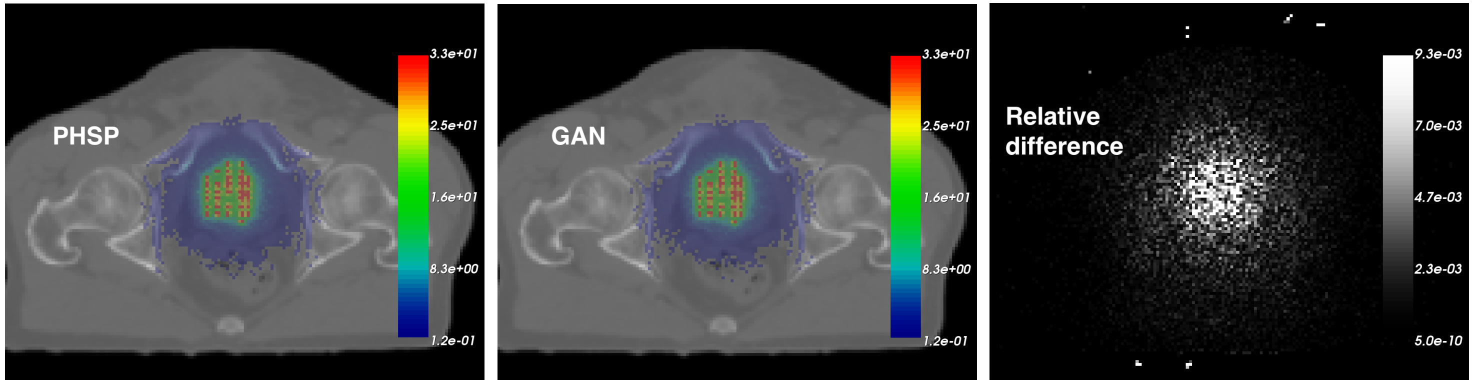


Figure 9. Slices of CT prostate image with deposited energy overlay (in MeV), computed by phase space (PHSP, left) and GAN generated (center) particles. The right image shows the dose difference Δ_{GAN} relative to the maximum dose (the maximum difference was below 4%).

Virtual Source Models

- Several VSM have been proposed
  - [Fix2001] [Grevillot2011] [Chabert2016], …
  - Histograms-based description (6D !): correlations bw variables
  - Analytical function model, adapted sampling procedures
  - Correlated-histograms with adaptive binning schemes, Kernel-Density Estimator (KDE)

  …
- May be efficient but
  - Simplification specific to one Linac type
  - Not a unique standardized method
  - Not easily generalisable to other Linac types (Cyberknife, Tomotherapy, FFF, etc)