Integration of ML models into GEANT4 to simulate plasma acceleration

Use case : PALLAS experiment

ARNAUD HUBER[1]

A. SYTOV[2], K. CASSOU[3], V. KUBYTSKYI[3], M. LENIVENKO[3]

Context and Objectives

Context:

Laser-plasma accelerators are compact sources of high-energy electrons (**PALLAS experiment** & **LWFA**)

Simulations are needed to understand and optimize these systems

Problems:

Full-scale simulations are very resource-intensive (PIC models)

Objective:

Integrate ML-based models into Geant4 to efficiently simulate experiments

PALLAS use case

PALLAS is laser-plasma accelerator test facility under construction at IJClab.

- Aiming for *"high-quality beam"* **laser-plasma injector (LPI)** facility working in **Continuous 10 Hz operation**
- Achieve **reliability** and **control** comparable to conventional RF accelerators standards

LPI PALLAS beamline

Parameters	phase 1	phase 2	unit
energy	150	**200**	MeV
charge	15-30	**30**	pC
frep	1-5	10	Hz
energy spread	<5%	**< 3%**	rms
$\epsilon_{n,rms}$	1	**<1**	μm
stability	5%	**<3%**	-
reproductibility	5%	**3%**	-

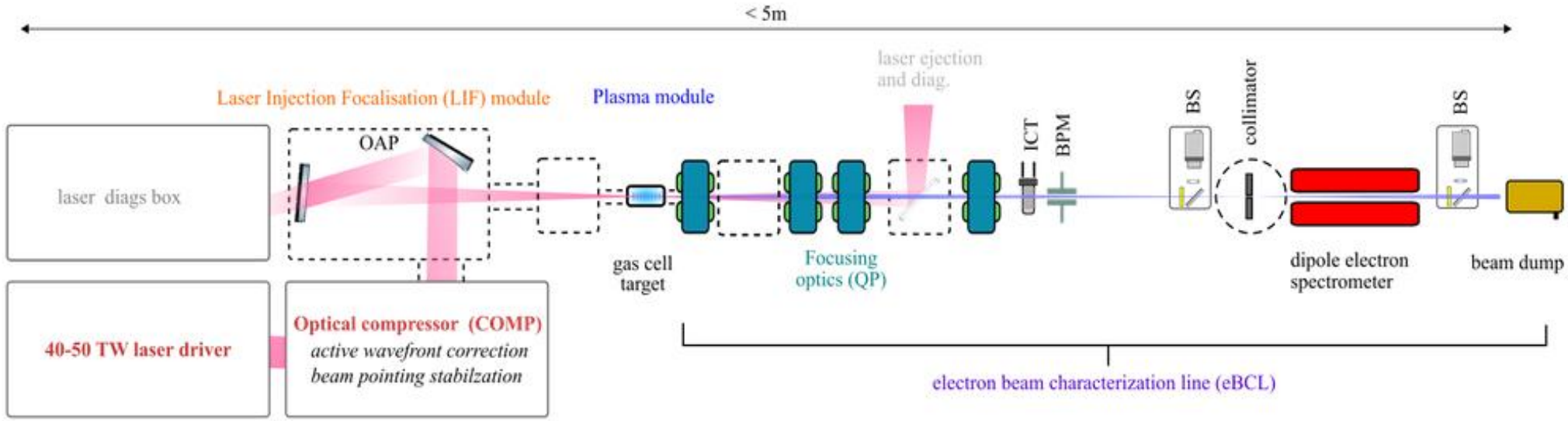
PALLAS use case

PALLAS is laser-plasma accelerator test facility under construction at IJClab.

- Aiming for *"high-quality beam"* **laser-plasma injector (LPI)** facility working in **Continuous 10 Hz operation**
- Achieve **reliability** and **control** comparable to conventional RF accelerators standards

LPI PALLAS beamline

Parameters	phase 1	phase 2	unit
energy	150	**200**	MeV
charge	15-30	**30**	pC
frep	1-5	10	Hz
energy spread	<5%	**< 3%**	rms
$\epsilon_{n,rms}$	1	**<1**	μm
stability	5%	**<3%**	-
reproductibility	5%	**3%**	-



Plasma target development :

- P.Drobniak et al. PRAB 26, 091302 (2023)

Beam Transport :

- Coline Guyot. PhD Thesis, Université Paris-Saclay, 2025.

e- beam compact characterization :

- Down to 500 keV energy
- Resolution emittance
- Longitudinal phase space

LWFA

Principle :

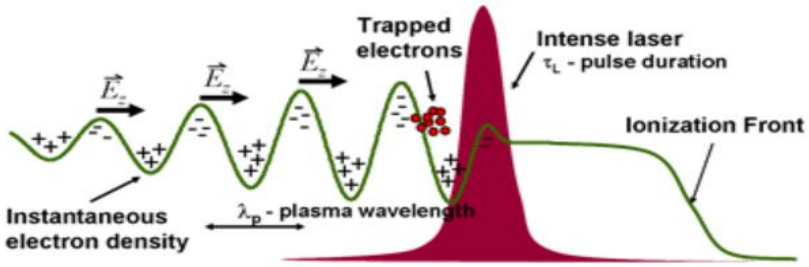
A laser travels through a gas → creating a plasma

- The laser pushes electrons aside, forming a wave
- Injected electrons "surfing" on this wave are accelerated

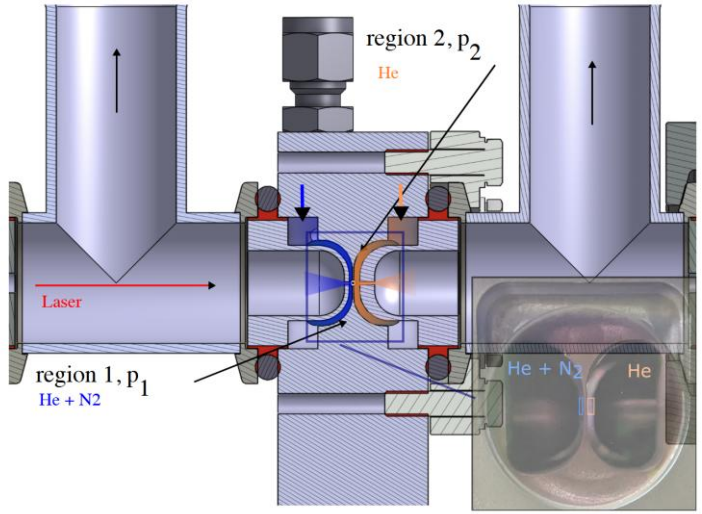
Key facts :

- Acceleration gradients > 100 GeV/m
- Experiments are **highly sensitive** to the **laser/plasma configuration**

Parameters	phase 1	phase 2	phase 3	unit
laser strengh, a_0	1.15-1.4	1.15-1.4	**>1.8**	
laser duration, t_L	40	**30**	**30**	fs (FWHM)
laser waist, w_0	18	18	18	um
Strehl ratio, S_r	**> 0.8**	**> 0.8**	**> 0.8**	-
beam pointing, δu_i	**<0.5**	**<0.5**	**<0.5**	urad
stability	1%	**<1%**	**<1%**	-
frep	10	10	10	Hz
target type	multi-cell	multi-cell	multi-cell	-
injection	ionisation	ionisation	ionisation	-
electron beamline	*CL1*	*CL1*	*TBD*	-



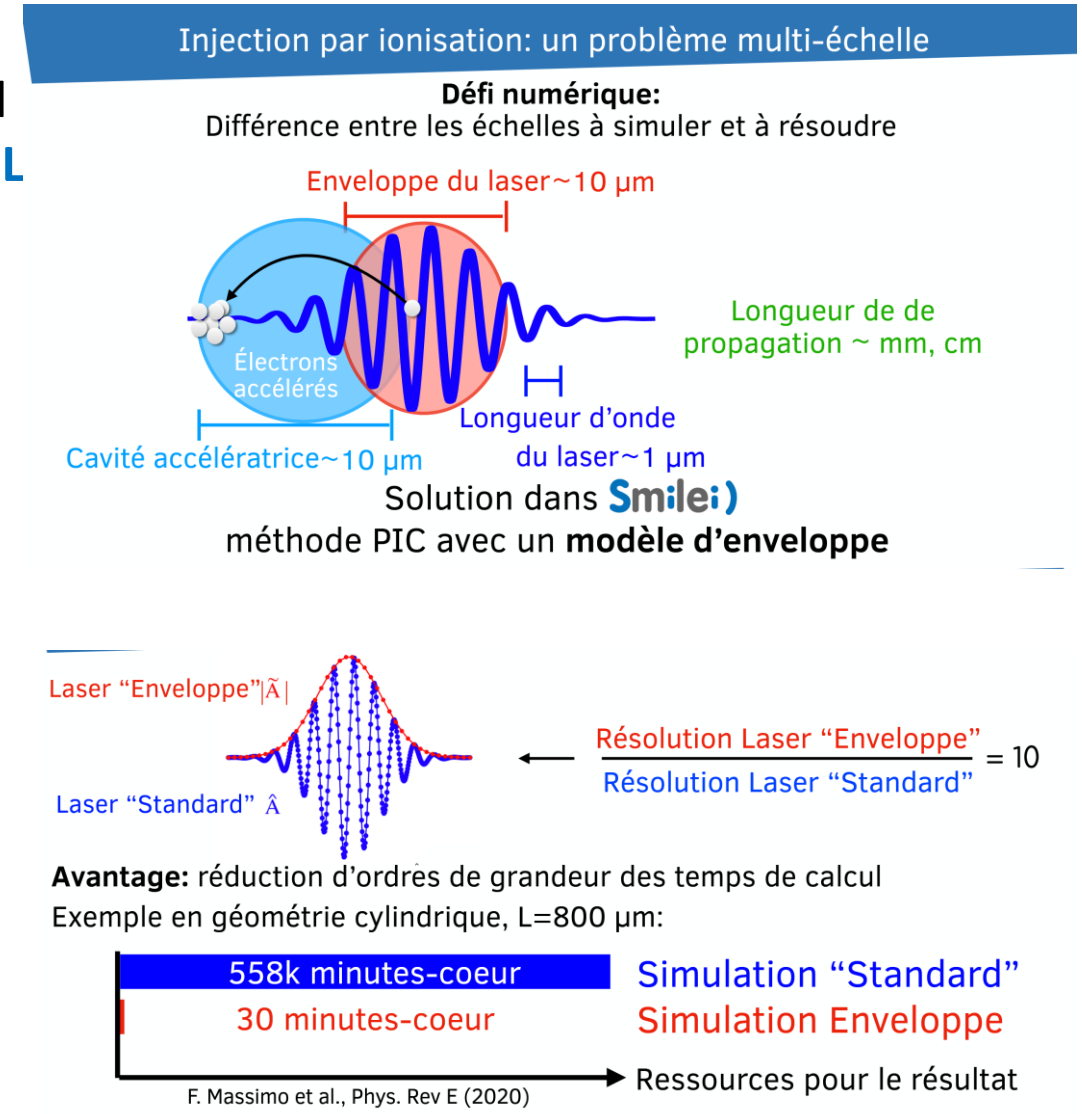
https://cuos.engin.umich.edu/researchgroups/hfs/research/laser-wakefield-acceleration/



Challenges of plasma simulations

PIC simulations (Particle-In-Cell) :

- **Very accurate** but **substantial** computational resources needed
- Numerical challenge (**slides Francesco Massimo – GdR APPEL 2022**)



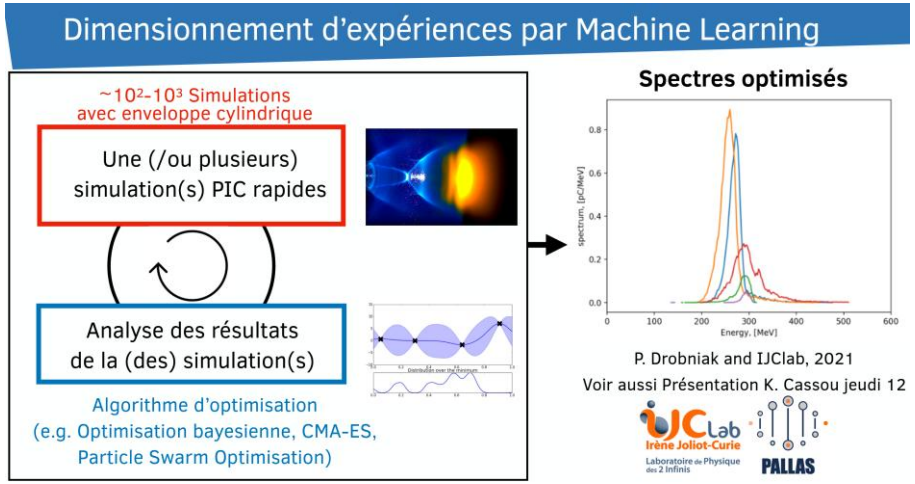
Challenges of plasma simulations

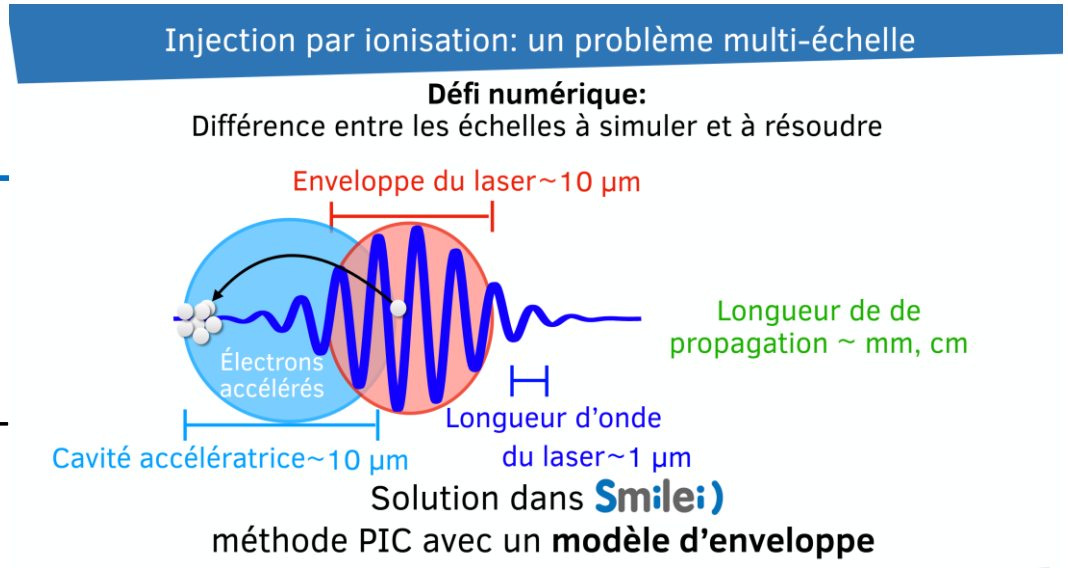
PIC simulations (Particle-In-Cell) :

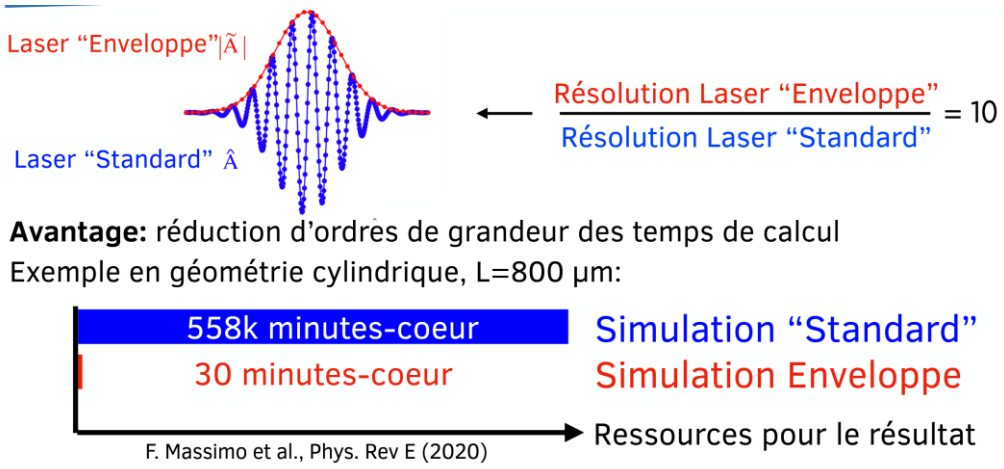
- **Very accurate** but **substantial** computational resources needed
- Numerical challenge (**slides Francesco Massimo – GdR APPEL 2022**)

Consequences :

- Difficult to explore **large parameter spaces** (nonlinear, multi-parametric, coupled physical processes)
- Not viable for full *start-to-end* simulations
- **Machine Learning** use case !!!







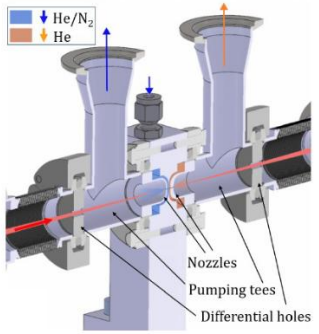
Solution : ML surrogate model

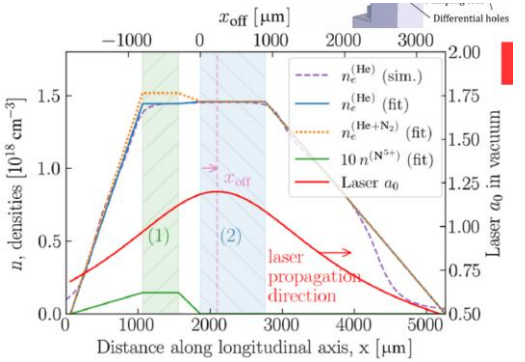
Idea :

- Train an ML model on the results of **Fast PIC simulations**
- **Predict** beam characteristics from input parameters

PALLAS gas cells[4]

Gas density profile generated[4] using **OpenFOAM[5]** code





1. V. Kubytskyi, Optimisation and surrogate models for PALLAS project, Journée modélisation LWFA GdR APPEL 2023
2. Gueladio Kane, Surrogate modelling for laser plasma acceleration, GdR SCIPAC 2025
3. A. Sytov et al. arXiv:2503.12154.
4. P.Drobniak et al. PRAB 26, 091302 (2023)
5. https://www.openfoam.com/
6. J.Derouillat et al. Smilei: CPC, 222:351-373, 2018; https://smileipic.github.io/Smilei/
7. G. Kane et al. arXiv2408.15845 (2024).

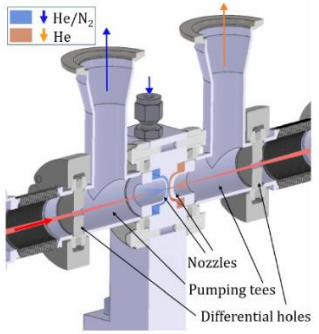
Solution : ML surrogate model

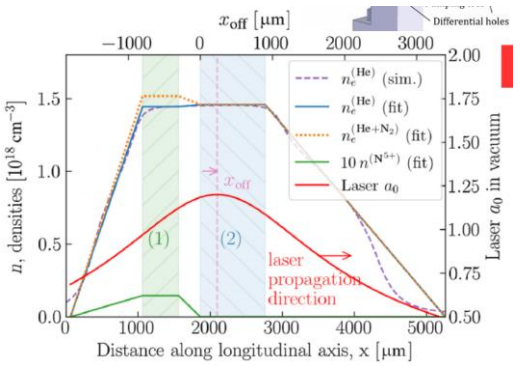
Idea :

- Train an ML model on the results of **Fast PIC simulations**
- **Predict** beam characteristics from input parameters

PALLAS gas cells[4]

Gas density profile generated[4] using **OpenFOAM**[5] code

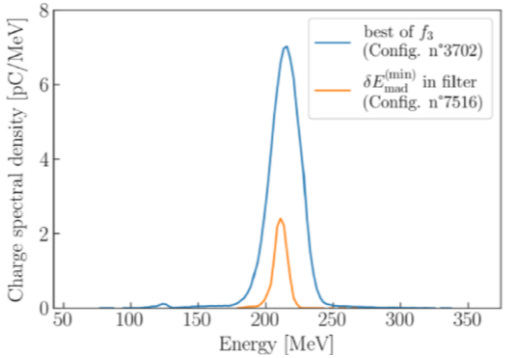




Numerical setup

- target configuration $x = [\, a_0, x_{of}, p_1 = p_2, c_{N_2} \,]$
- Smilei open source PIC code with envelop approximation coupled to ionization[1], azimuthal modes[2] and low ppc :
  ~ 12-24 h.cores/mm.(*running with 240 cores on 96k computing-core HPC SKL*)
- simplified profiles from CFD simulations and confined high-Z dopant.

PIC simulations with SMILEI[6]



1. V. Kubytskyi, Optimisation and surrogate models for PALLAS project, Journée modélisation LWFA GdR APPEL 2023
2. Gueladio Kane, Surrogate modelling for laser plasma acceleration, GdR SCIPAC 2025
3. A. Sytov et al. arXiv:2503.12154.
4. P.Drobniak et al. PRAB 26, 091302 (2023)
5. https://www.openfoam.com/
6. J.Derouillat et al. Smilei: CPC, 222:351-373, 2018; https://smileipic.github.io/Smilei/
7. G. Kane et al. arXiv2408.15845 (2024).

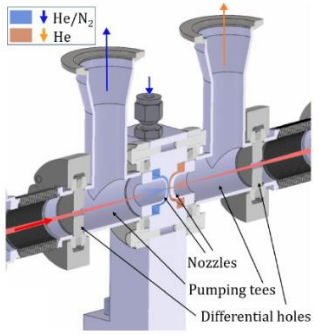
Solution : ML surrogate model

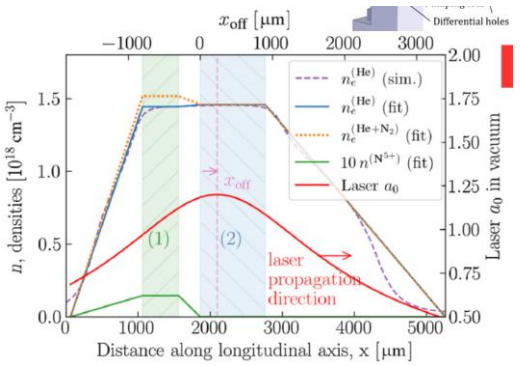
Idea :

- Train an ML model on the results of **Fast PIC simulations**
- **Predict** beam characteristics from input parameters
- (re)use knowledge of already computed configurations -> construct **surrogate model of the accelerator**[1,2]
- New configuration can be estimated from the surrogate model and validated with SMILEI.

PALLAS gas cells[4]

Gas density profile generated[4] using **OpenFOAM**[5] code

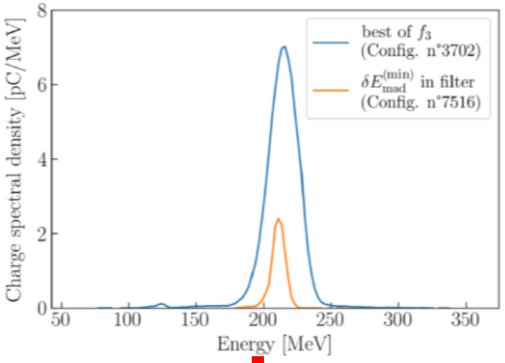


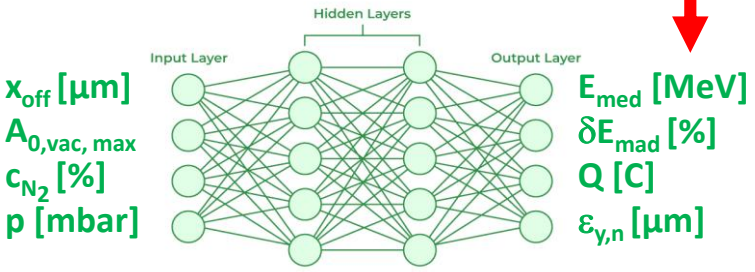


Numerical setup

- target configuration $x = [\, a_0,\, x_{of},\, p_1 = p_2,\, c_{N_2}\,]$
- Smilei) open source PIC code with envelop approximation coupled to ionization[1], azimuthal modes[2] and low ppc :
  ~ 12-24 h.cores/mm.(*running with 240 cores on 96k computing-core HPC SKL*)
- simplified profiles from CFD simulations and confined high-Z dopant.

PIC simulations with SMILEI[6]





ML surrogate model[2,4,7]

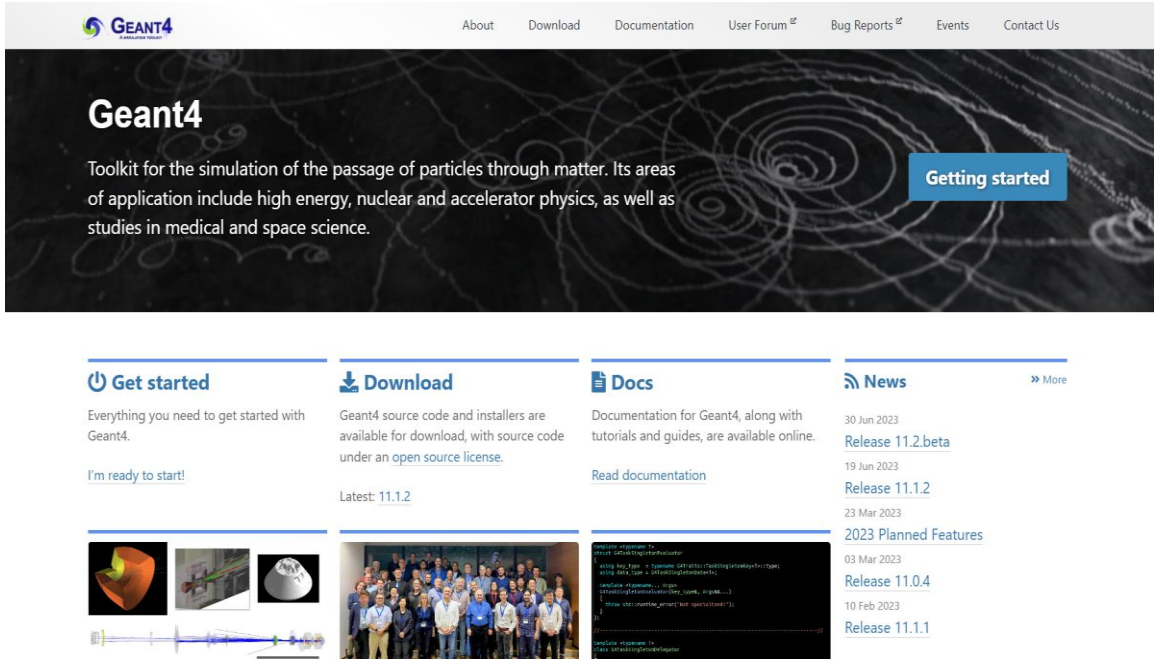
Layer (type)	Output Shape	Param #
Input Layer (InputLayer)	(None, 4)	0
Hidden Layer 1 (Dense)	(None, 100)	500
PReLU activation 1 (PReLU)	(None, 100)	100
Hidden Layer 2 (Dense)	(None, 100)	10,100
PReLU activation 2 (PReLU)	(None, 100)	100
Hidden Layer 3 (Dense)	(None, 100)	10,100
PReLU activation 3 (PReLU)	(None, 100)	100
Output Layer (Dense)	(None, 4)	404

1. V. Kubytskyi, Optimisation and surrogate models for PALLAS project, Journée modélisation LWFA GdR APPEL 2023
2. Gueladio Kane, Surrogate modelling for laser plasma acceleration, GdR SCIPAC 2025
3. A. Sytov et al. arXiv:2503.12154.
4. P.Drobniak et al. PRAB 26, 091302 (2023)
5. https://www.openfoam.com/
6. J.Derouillat et al. Smilei: CPC, 222:351-373, 2018; https://smileipic.github.io/Smilei/
7. G. Kane et al. arXiv2408.15845 (2024).

Overview :

- C++
- Object oriented
- Open Source
- **Toolkit** i.e. collection of tools
  - **Geant4 defaults model DOES NOT EXIST !!!**
  - **MUST** provide the **necessary information**, choose the GEANT4 tools

GEANT4 provides some features (building blocks)



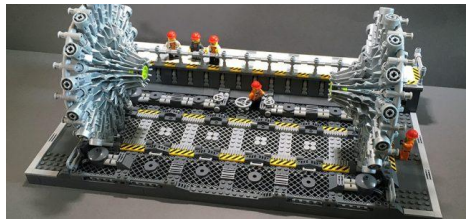
https://geant4.web.cern.ch/

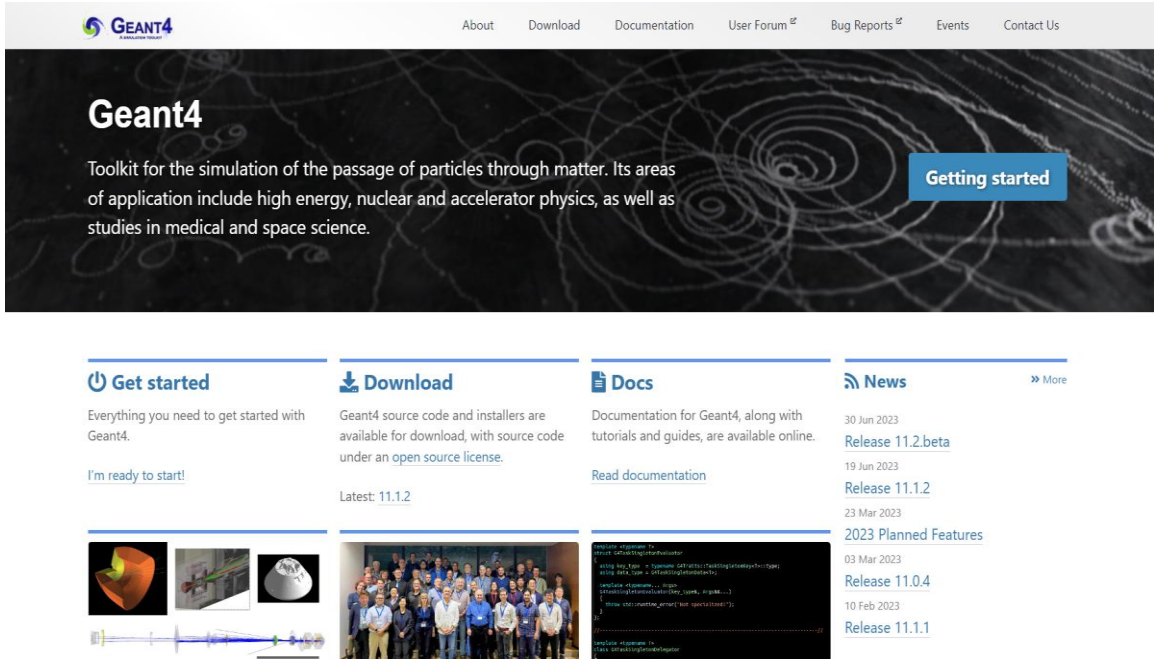
Overview :

- C++
- Object oriented
- Open Source
- **Toolkit** i.e. collection of tools
  - **Geant4 defaults model DOES NOT EXIST !!!**
  - **MUST** provide the **necessary information**, choose the GEANT4 tools

GEANT4 provides some features (building blocks)

Users construction in order to describe the problematic





https://geant4.web.cern.ch/

Many examples provided :

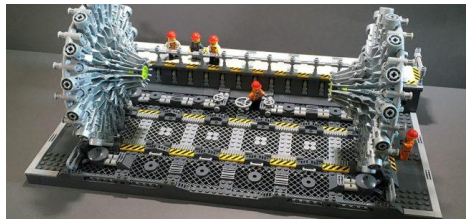
https://geant4-userdoc.web.cern.ch/Doxygen/examples_doc/html/index.html

Overview :

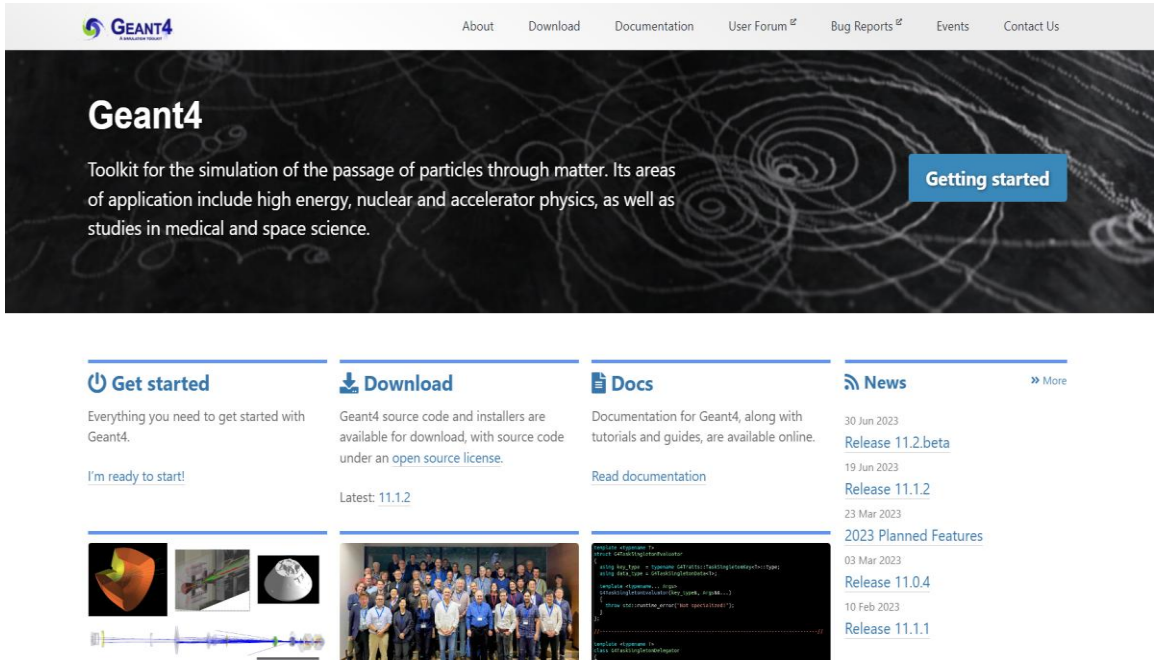
- C++
- Object oriented
- Open Source
- **Toolkit** i.e. collection of tools
  - **Geant4 defaults model DOES NOT EXIST !!!**
  - **MUST** provide the **necessary information**, choose the GEANT4 tools

GEANT4 provides some features (building blocks)

Users construction in order to describe the problematic



Many examples provided :
https://geant4-userdoc.web.cern.ch/Doxygen/examples_doc/html/index.html



https://geant4.web.cern.ch/

MC methods :

- Independant particles ✓
- No collective behaviour ✗

⇒ Does **not** model the plasma acceleration itself

⇒ **Link SMILEI – ML surrogate – GEANT4**

ML integration on GEANT4

Requirements :

- Full simulation of the experiment (from acceleration to detection)
- Adaptable to different experimental scenarios

Needs :

- Use of GEANT4's « ParticleGun »
- Data from the ML model (spectrum, direction, position, …)

```
/gun/SetStatusONNX true
/laser/SetOffsetLaserFocus 1680 ## um
/laser/SetNormVecPotential 1.23
/laser/SetFracDopTargetChamber 0.0617 # %(/100)
/laser/SetPressure 47.8 #mbar
```

```
G4double fXof_min = -399.8247;
G4double fXof_max = 1798.325;
G4double fXof_rescaled = fXof / (fXof_max - fXof_min) - fXof_min/(fXof_max - fXof_min);

G4double fA0_min = 1.100516;
G4double fA0_max = 1.849792;
G4double fA0_rescaled = fA0 / (fA0_max - fA0_min) - fA0_min/(fA0_max - fA0_min);

G4double fCN2_min = 0.002064164;
G4double fCN2_max = 0.1199827;
G4double fCN2_rescaled = fCN2 / (fCN2_max - fCN2_min) - fCN2_min/(fCN2_max - fCN2_min);

G4double fP1_min = 10.09451;
G4double fP1_max = 99.95741;
G4double fP1_rescaled = fP1 / (fP1_max - fP1_min) - fP1_min/(fP1_max - fP1_min);
```

```
#ifndef B1PrimaryGeneratorAction_h
#define B1PrimaryGeneratorAction_h 1

#include "G4VUserPrimaryGeneratorAction.hh"
#include "globals.hh"
#include "G4GeneralParticleSource.hh"
#include "G4ParticleGun.hh"
#include <memory>
#include <onnxruntime_c_api.h>
#include <onnxruntime_cxx_api.h>
class G4ParticleGun;
class G4Event;
```

Keras function (Python) -> **tf2onnx.convert.from_keras** -> ONNX file (C++)

```
PrimaryGeneratorAction::PrimaryGeneratorAction(): G4VUserPrimaryGeneratorAction(),
    fParticleGun(0),
    fEnvelopeBox(0)
{
    G4int n_particle = 1;
    fParticleGun  = new G4ParticleGun(n_particle);

    // default particle kinematic
    fParticleGun->SetParticleDefinition(
        G4ParticleTable::GetParticleTable()->FindParticle("e-"));

    //Neural network: create onnx session
    Ort::Env env(ORT_LOGGING_LEVEL_WARNING, "plasma");
    Ort::SessionOptions session_options;
    session_options.SetIntraOpNumThreads(1);
    auto sessionLocal =
        std::make_unique<Ort::Session>(env, "model2.onnx", session_options);
    fSession = std::move(sessionLocal);

    // Get input node information
    fMemory_info = Ort::MemoryInfo::CreateCpu(
        OrtAllocatorType::OrtArenaAllocator, OrtMemTypeDefault);
```

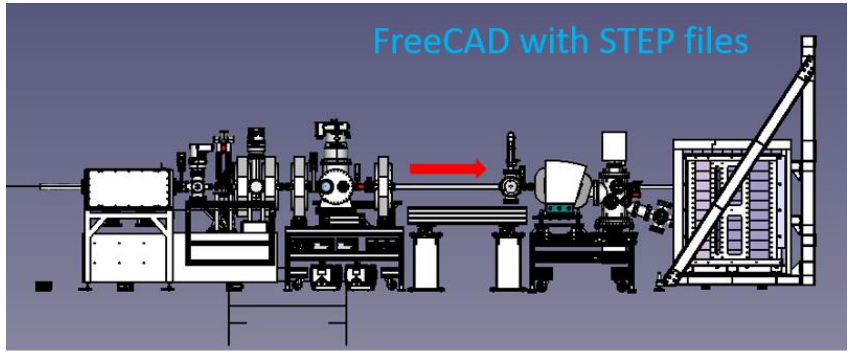
Results and benefics

Key points :

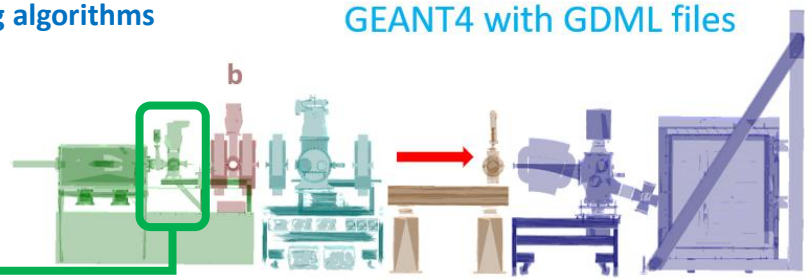
- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

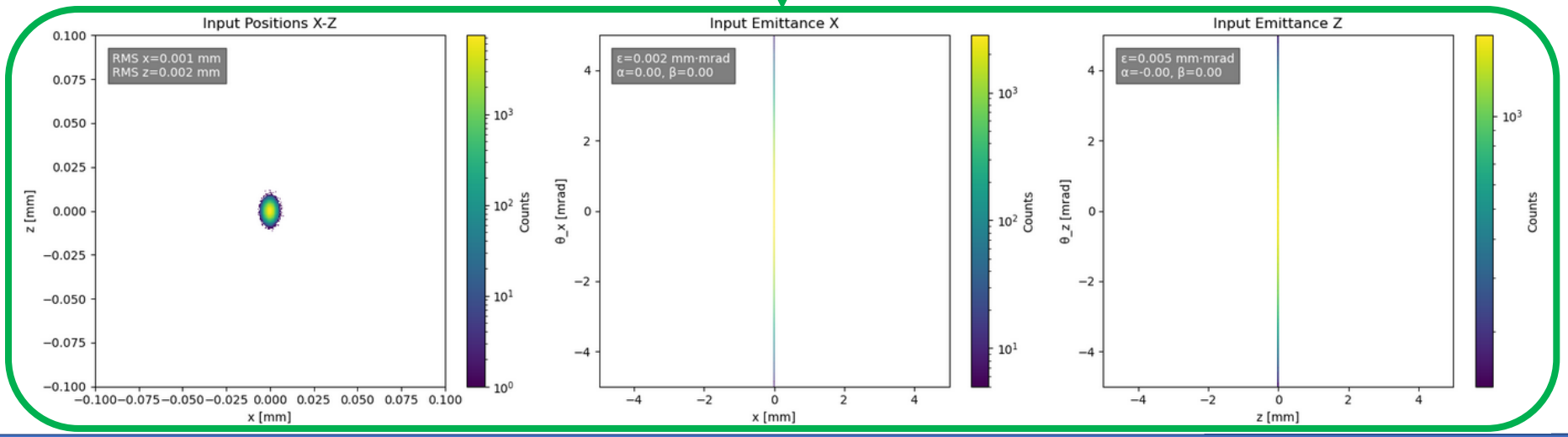
Results :

- **Particle generation consistent with PIC model**
- Complete implementation of beam transport in GEANT4 (quadrupoles)



Meshing algorithms





Particle generation

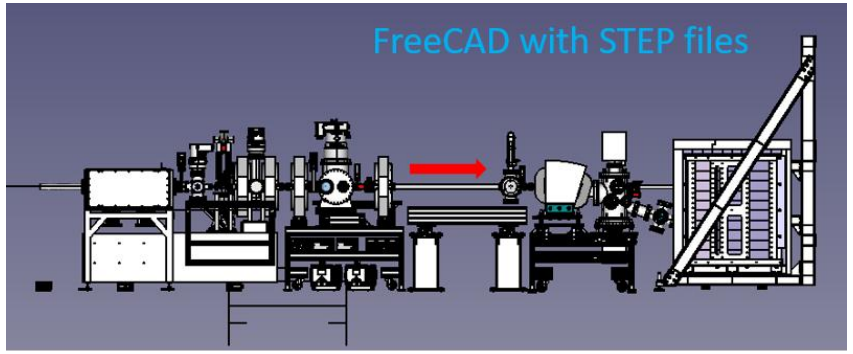
Results and benefics

Key points :

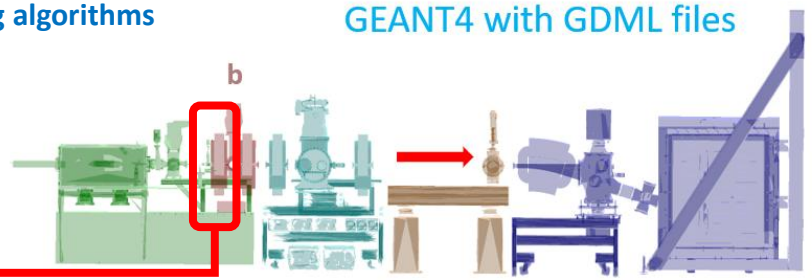
- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

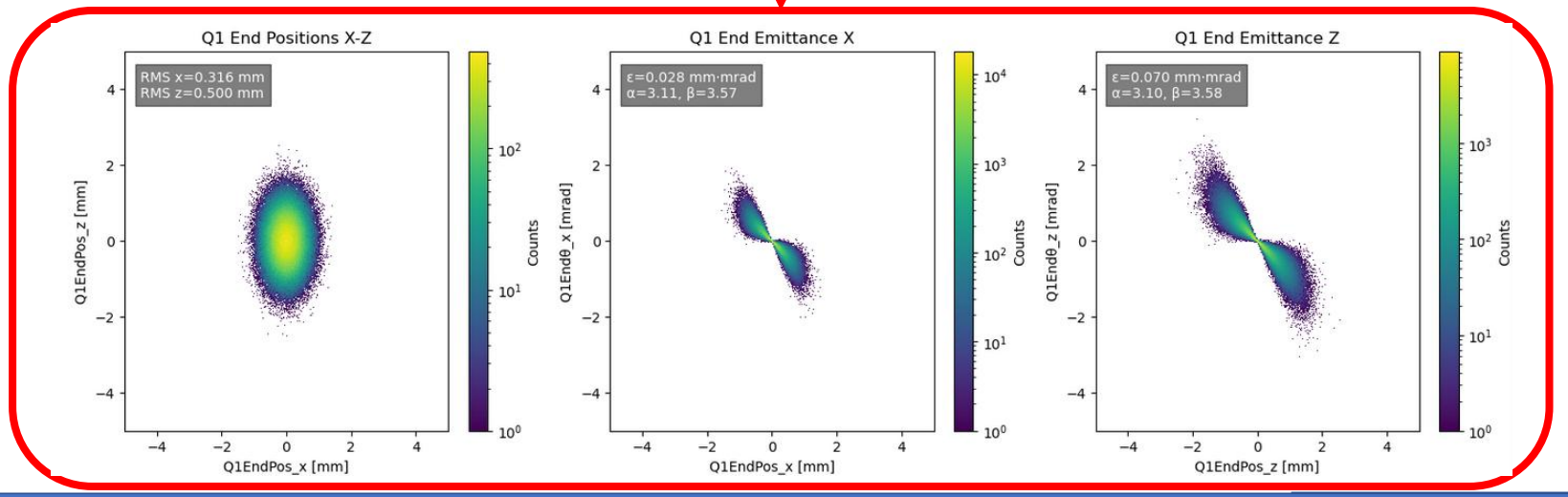
Results :

- Particle generation consistent with PIC model
- **Complete implementation of beam transport in GEANT4 (quadrupoles)**



Meshing algorithms





Q1 exit

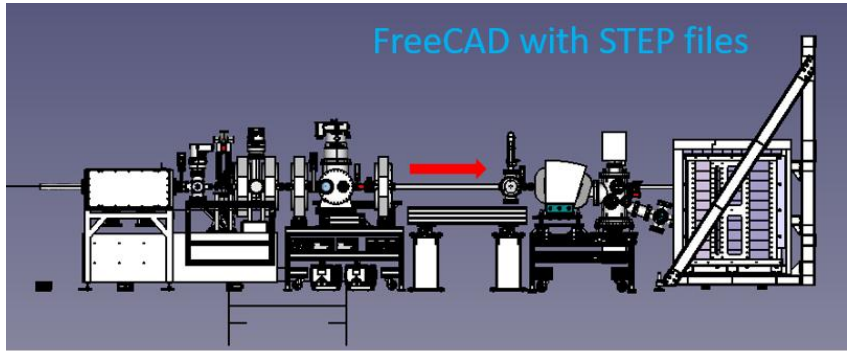
Results and benefics

Key points :

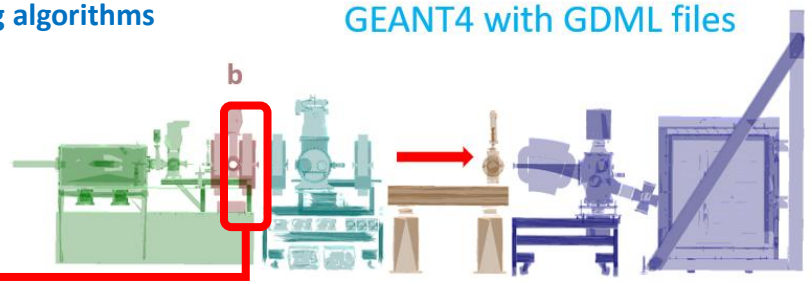
- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

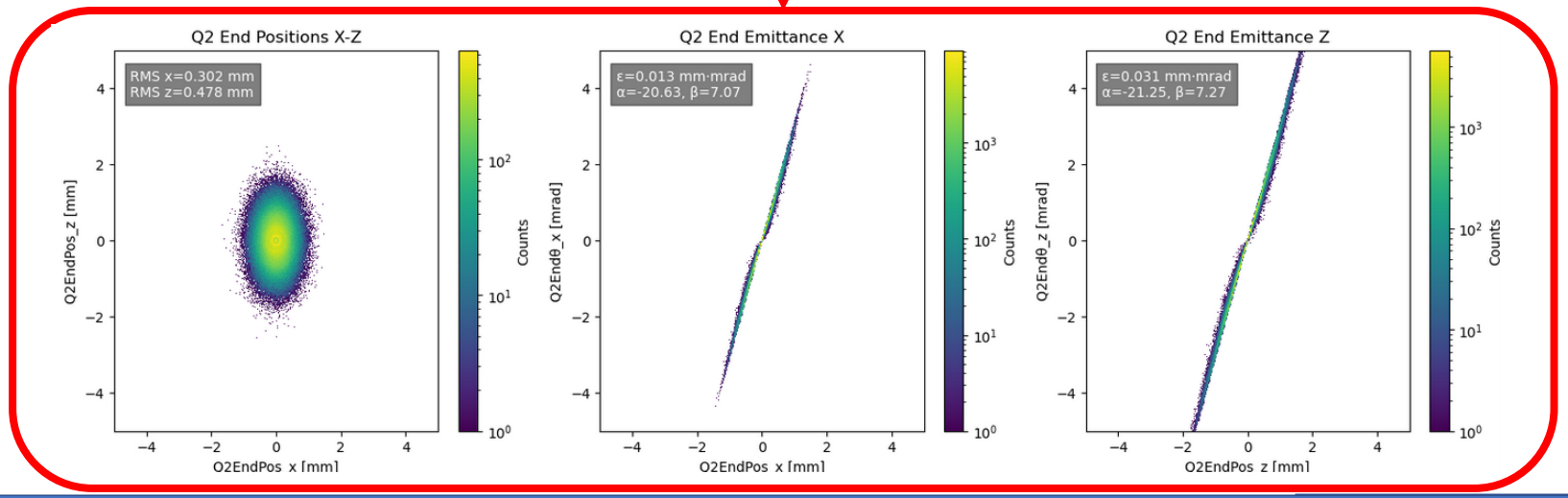
Results :

- Particle generation consistent with PIC model
- **Complete implementation of beam transport in GEANT4 (quadrupoles)**



Meshing algorithms





Q2 exit

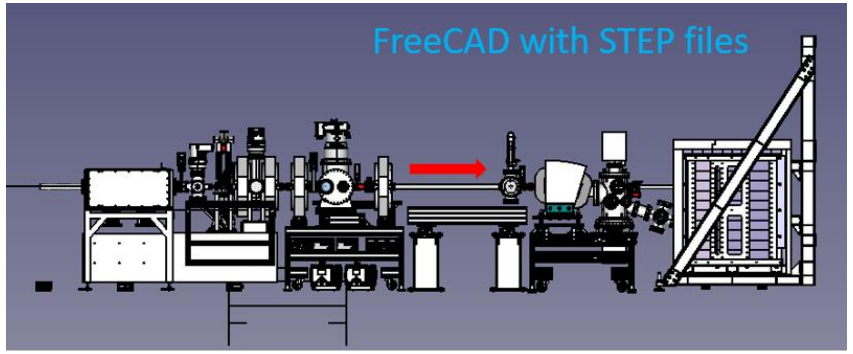
Results and benefics

Key points :

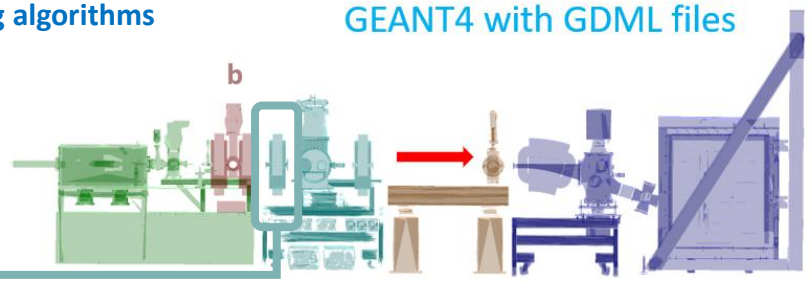
- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

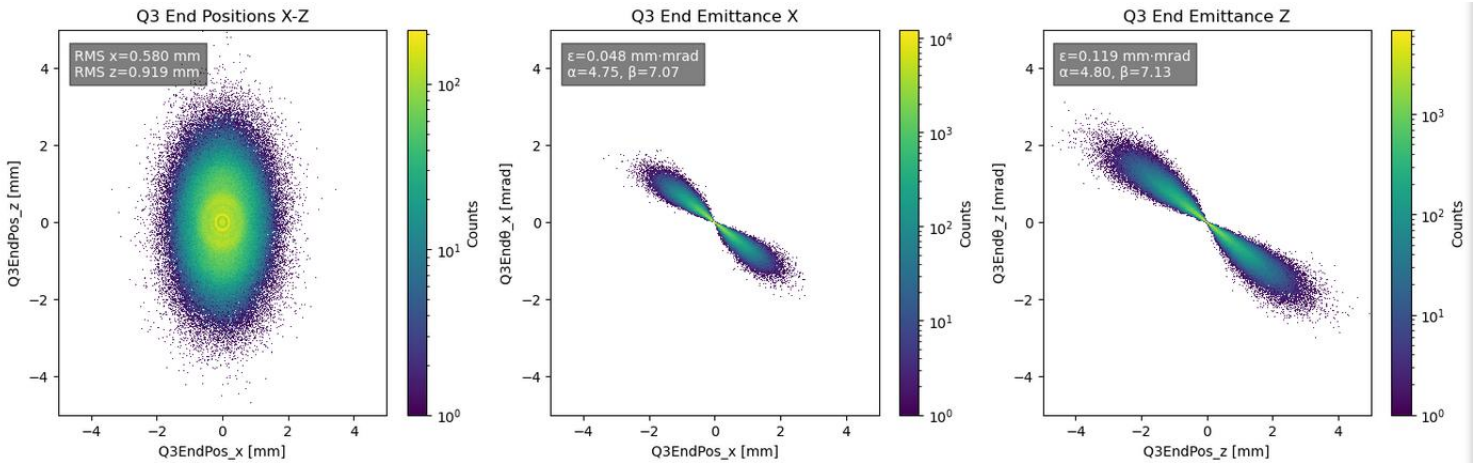
Results :

- Particle generation consistent with PIC model
- **Complete implementation of beam transport in GEANT4 (quadrupoles)**



Meshing algorithms





Q3 exit

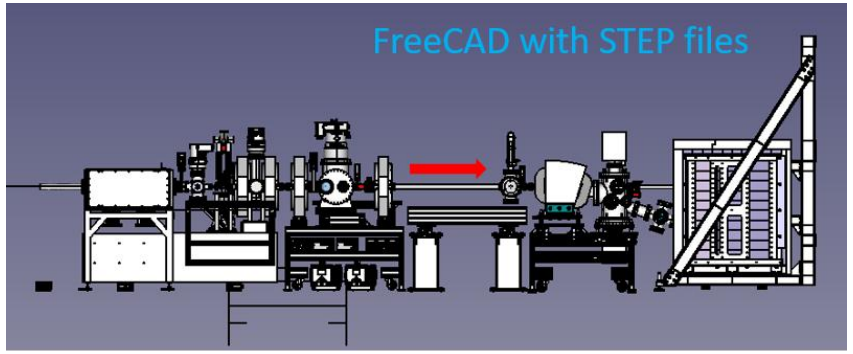
Results and benefics

Key points :

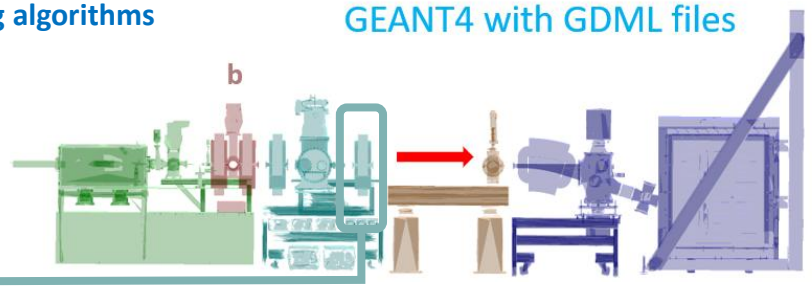
- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

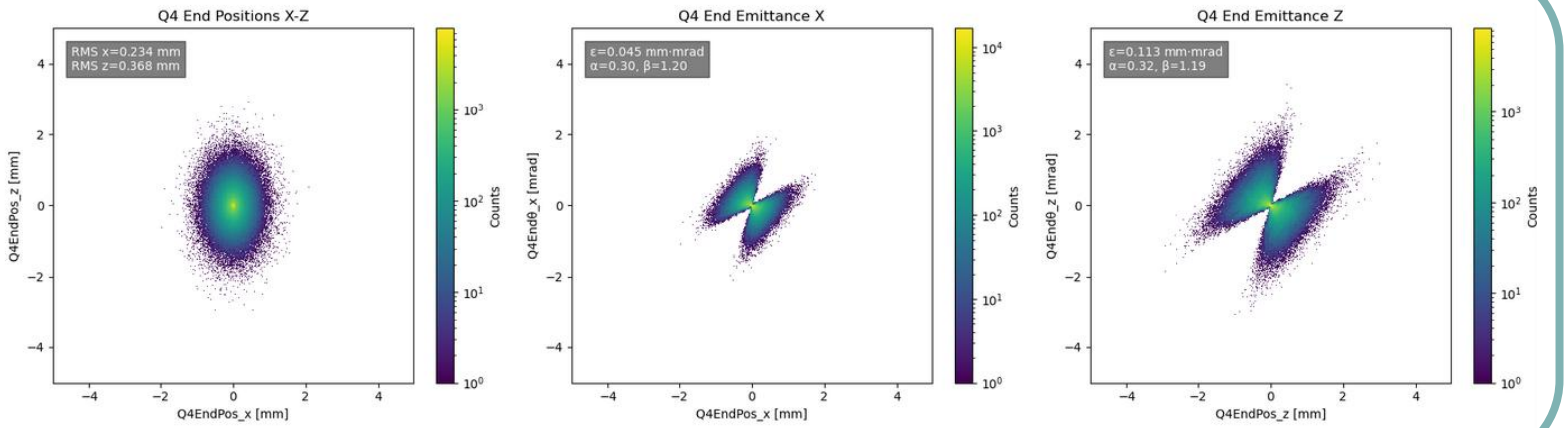
Results :

- Particle generation consistent with PIC model
- **Complete implementation of beam transport in GEANT4 (quadrupoles)**



Meshing algorithms





Q4 exit

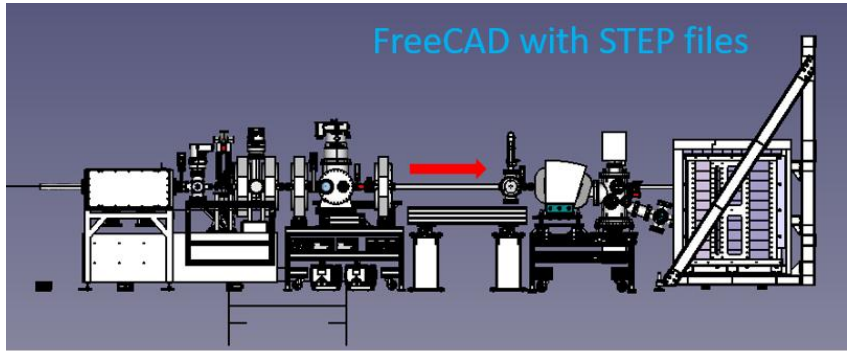
Results and benefics

Key points :

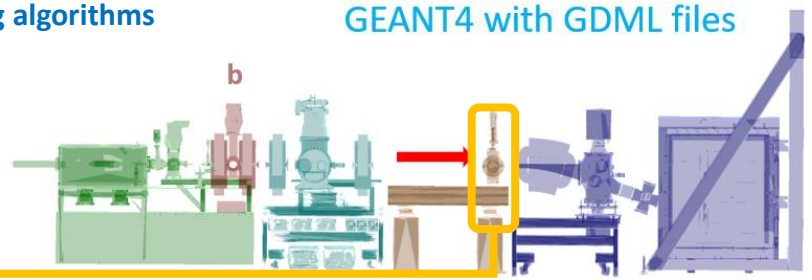
- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

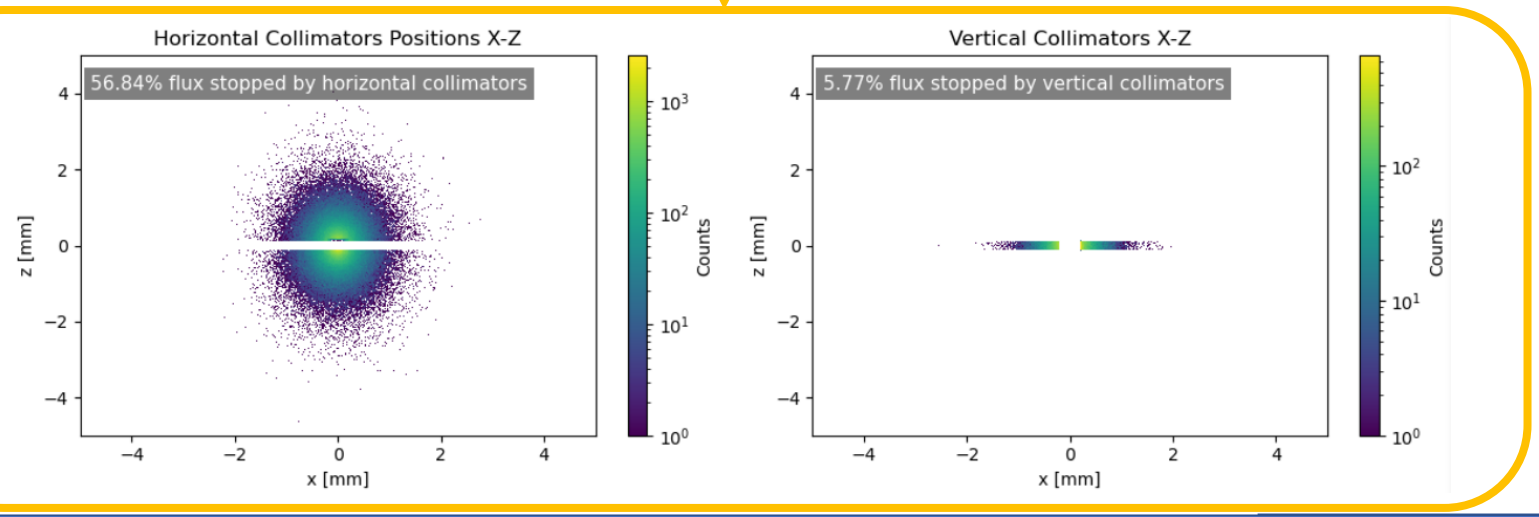
Results :

- Particle generation consistent with PIC model
- **Complete implementation of beam transport in GEANT4 (quadrupoles)**



Meshing algorithms





Collimators

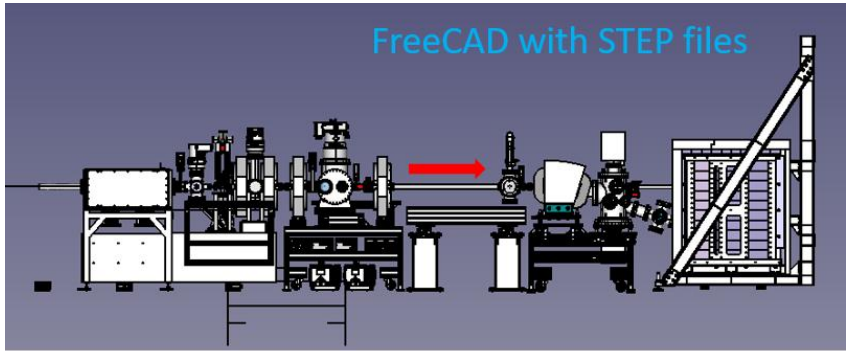
Results and benefics

Key points :

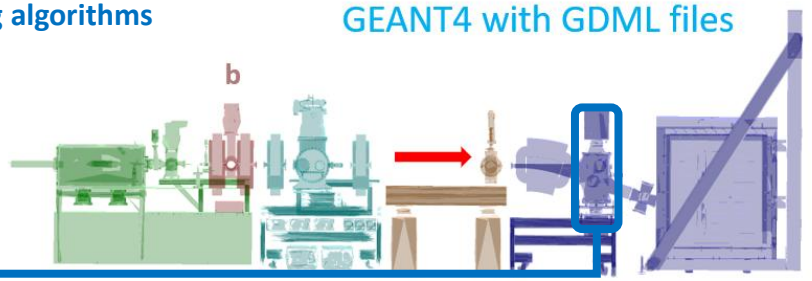
- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

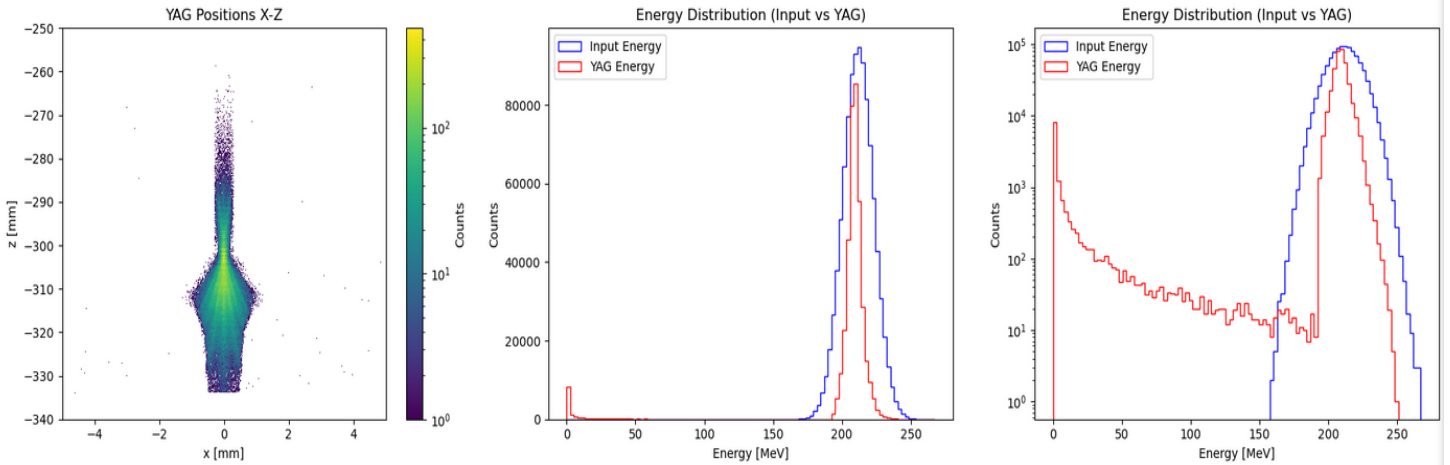
Results :

- Particle generation consistent with PIC model
- **Complete implementation of beam transport in GEANT4 (quadrupoles)**



Meshing algorithms





YAG screen

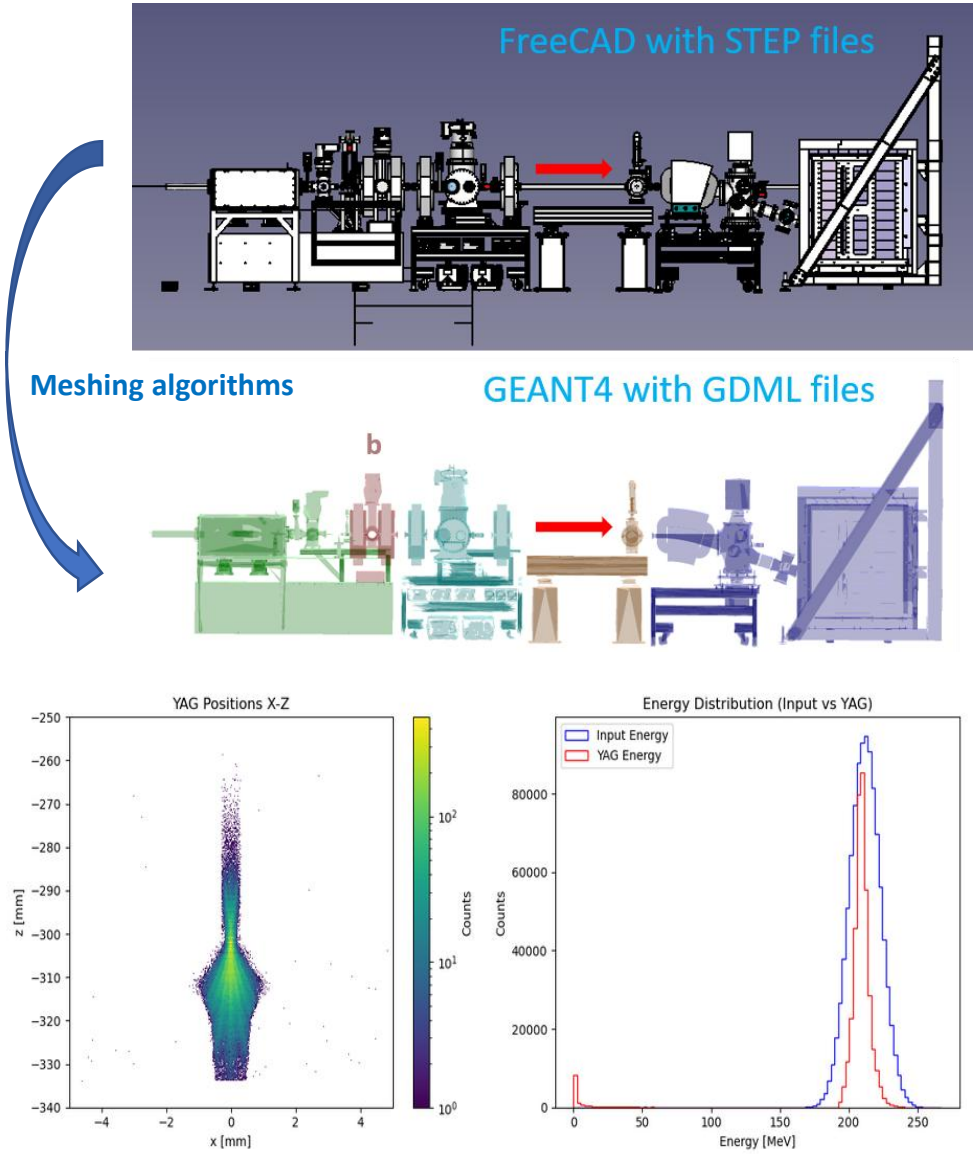
Results and benefics

Key points :

- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

Results :

- Particle generation consistent with PIC model
- Complete implementation of beam transport in GEANT4 (quadrupoles)
- **ML analysis with (Laser parameters, Qx values, Qx distances, slits openings) in order to maximise Charge and/or Energy beam**



Results and benefics

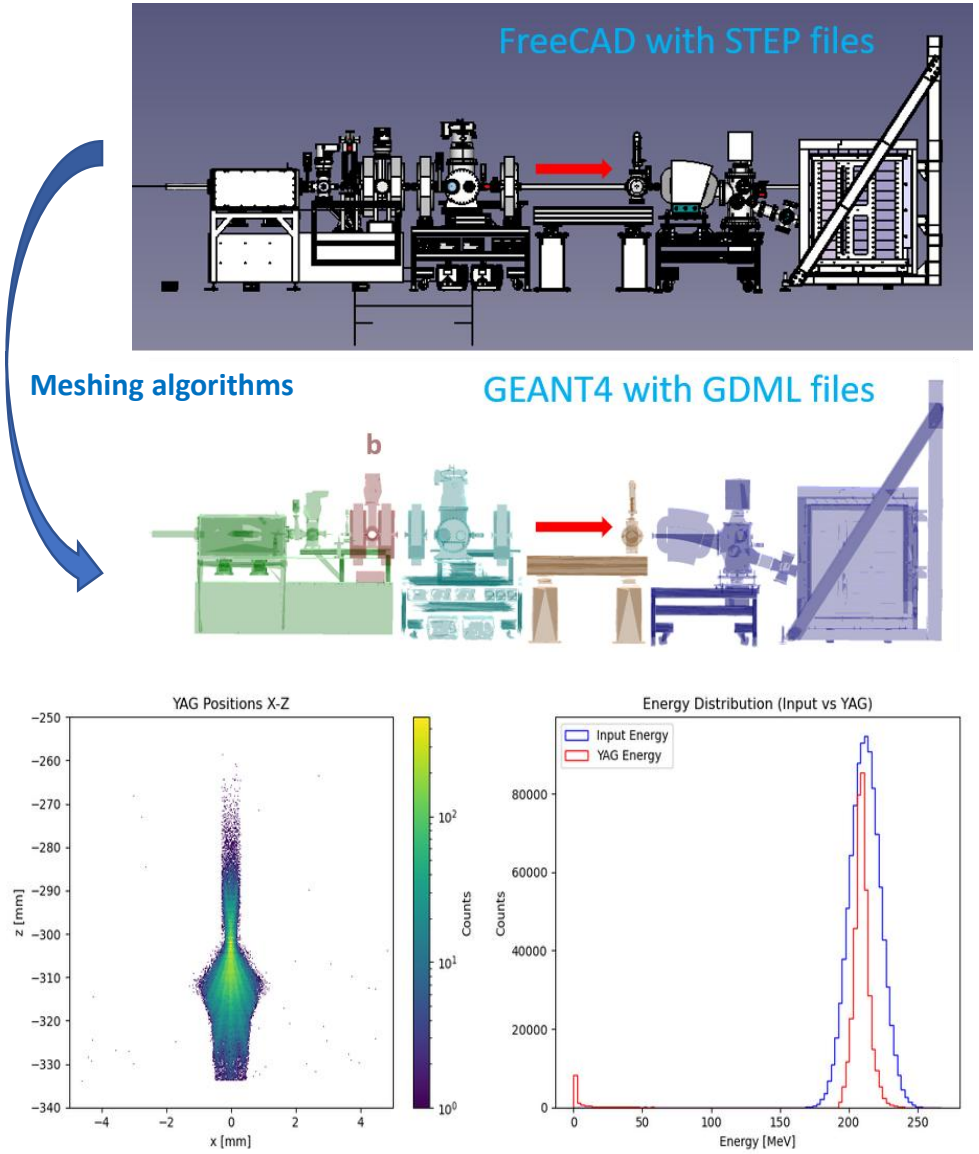
Key points :

- **Full geometry** of the PALLAS beamline using **GDML**
- **MT** on CPU possible
- Easy exploration of **varied scenarios**

Results :

- Particle generation consistent with PIC model
- Complete implementation of beam transport in GEANT4 (quadrupoles)
- ML analysis with (Laser parameters, Qx values, Qx distances, slits openings) in order to maximise Charge and/or Energy beam

- Soon available as an “advanced example”
  - **ML model trained on particle-in-cell simulations of laser-driven plasma wakefield acceleration as a particle source and entire experimental setup implemented using GDML**



Conclusion & Perspectives

- ✓ **Reduced complexity and computational cost**, while maintaining high-fidelity simulation accuracy.
- ✓ **ML model used**: multilayer perceptron (MLP) trained on PIC simulation data of laser-plasma interactions
- ✓ **ONNX framework** enabled seamless translation of the ML model from Python to C++ for integration into Geant4.
- ✓ **Implementation in Geant4**, centered on the PALLAS facility, enabling start-to-end simulations combining plasma acceleration physics with Monte Carlo modeling.
- ✓ **Full experimental setup of PALLAS** implemented in Geant4 using GDML files, including beam transport and detector systems.


Official example of GEANT4 collaboration

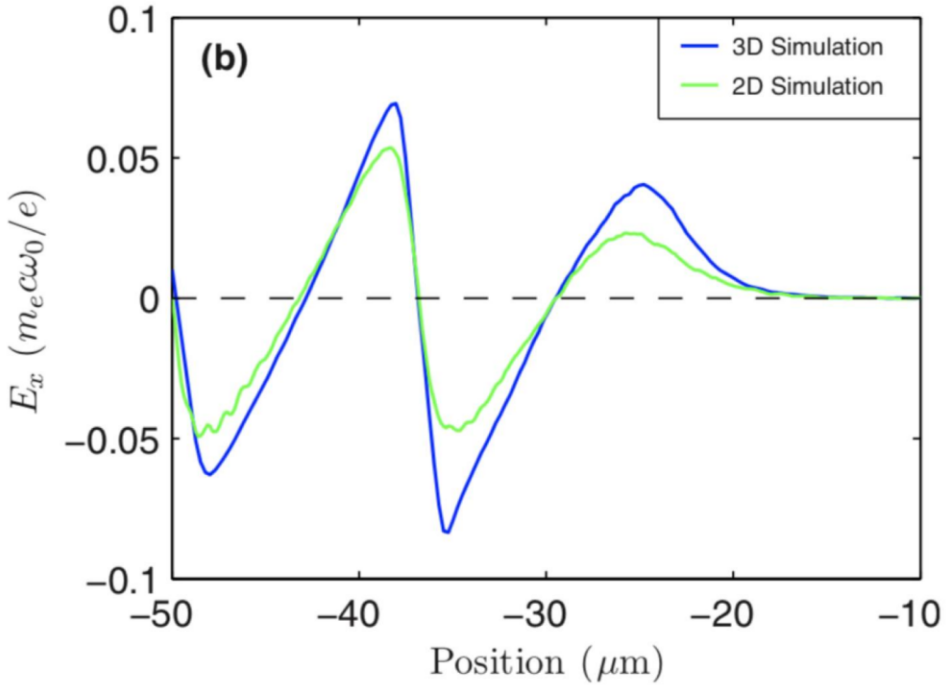

Scalable and adaptable framework for future experiments and applications.


Potential applications: synchrotron light sources, free electron lasers, nuclear physics, and advanced radiotherapy.

The LWFA 3D PIC simulation problem is enormous

!!! SLIDE FRANCESO MASSIMO !!!

2D cartesian simulations:
Not accurate enough



X. Davoine et al., Phys. Plasmas 15, 113102 (2008)

Example of 3D simulation scale
with "classic" electromagnetic PIC loop
and "classic" numerical schemes

- Laser wavelength **λ_0=0.8 µm** -> 0.016 µm , Δx=0.016 µm, cΔt=0.99Δx
- Laser duration: 30 fs -> Window size Lx = 40 µm = 2500 * Δx
- w_0=12 µm -> Δy=Δz=0.5µm , Window size Ly = Lz = 125 µm = 192 * Δy
- Acceleration length = **1 mm** ~ 60000 c* Δt
- 8 macro-particles per cell -> ~2500*250*250*8 =

10^9 macro-particles, pushed for 60k timesteps!

Smallest scale to sample [space and time]

Largest scale to simulate [space and time] (Note: current LWFA studies require several mm, up to 10s of cm, or even m)