AI-assisted design of Muon Collider Final Cooling Channel

Elena Fol

CERN, Beams Department- Accelerator Beam Physics Group

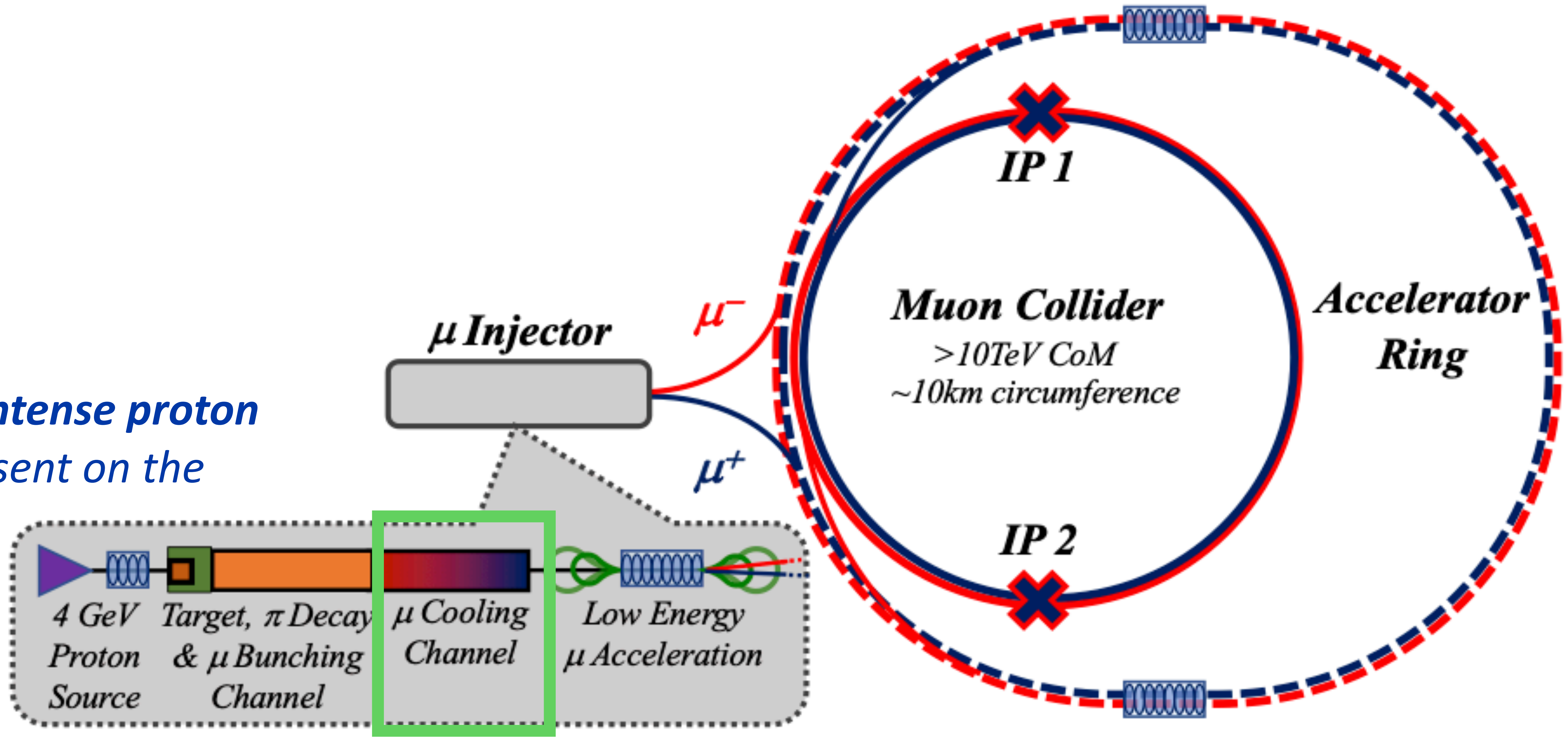
ICFA Workshop on ML for Accelerators

5-8 March 2024

Outline

- Muon Collider overview
- Concept and challenges of Final Cooling
- Lattice design optimization using ML
    - Surrogate models
    - Feature Importance Analysis with Decision Trees
    - Bayesian Optimization
    - Clustering and anomaly detection
- Current results
- Summary

Muon Collider: overview



Short ***intense proton bunch*** *sent on the target*

Interaction with the target produces ***pions***

➡ ***decay into muons***

Muons are captured and ***cooled to lower emittance***

$$\mathcal{L} \propto \gamma \langle B \rangle \sigma_\delta \frac{N_0}{\epsilon\epsilon_L} f_r N_0 \gamma$$

High energy

High field in collider ring

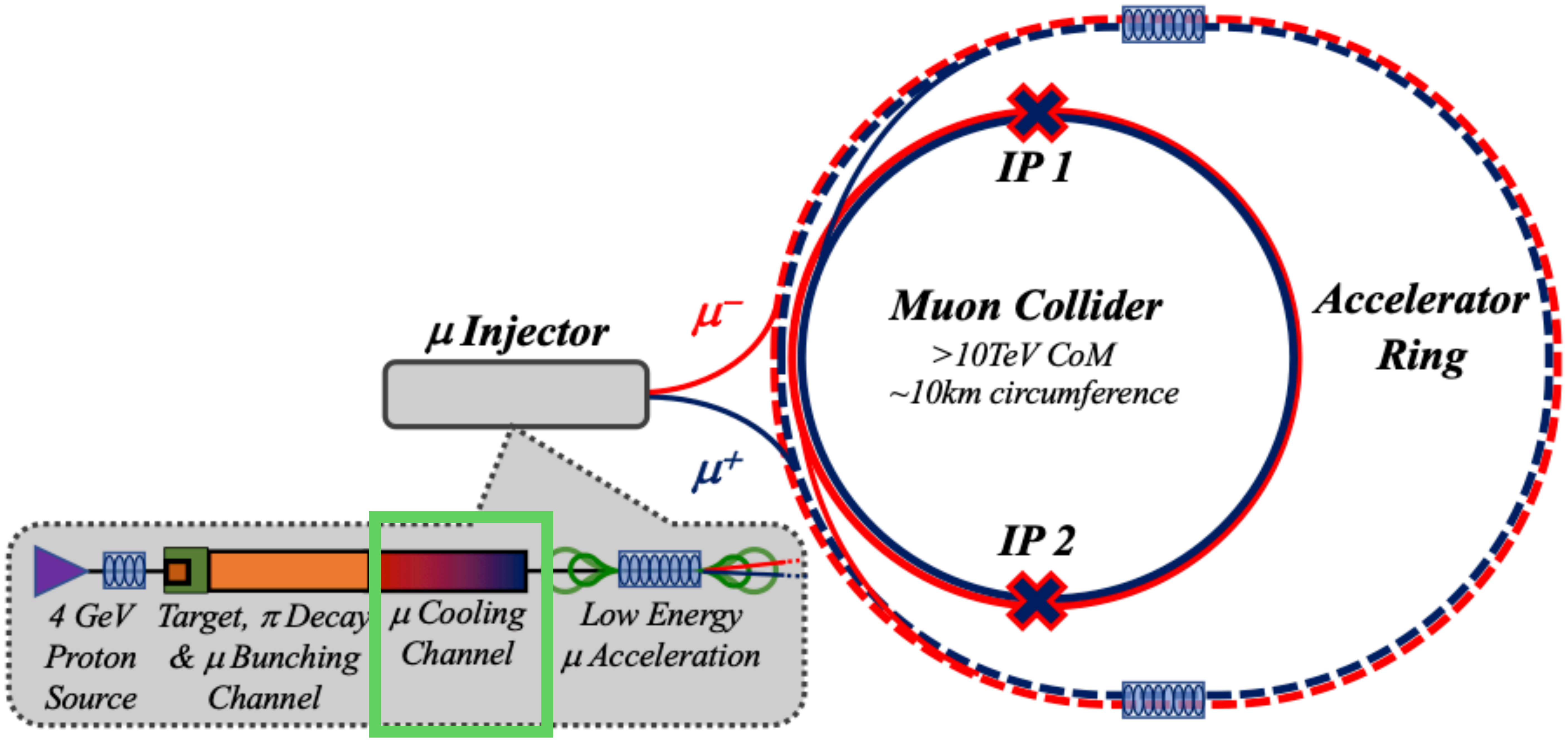
Large energy acceptance

Dense beam

High beam power

https://muoncollider.web.cern.ch

Muon Collider: overview



$$\mathcal{L} \propto \gamma \langle B \rangle \sigma_\delta \frac{N_0}{\epsilon \epsilon_L} f_r N_0 \gamma$$

High energy

High field in collider ring

Large energy acceptance

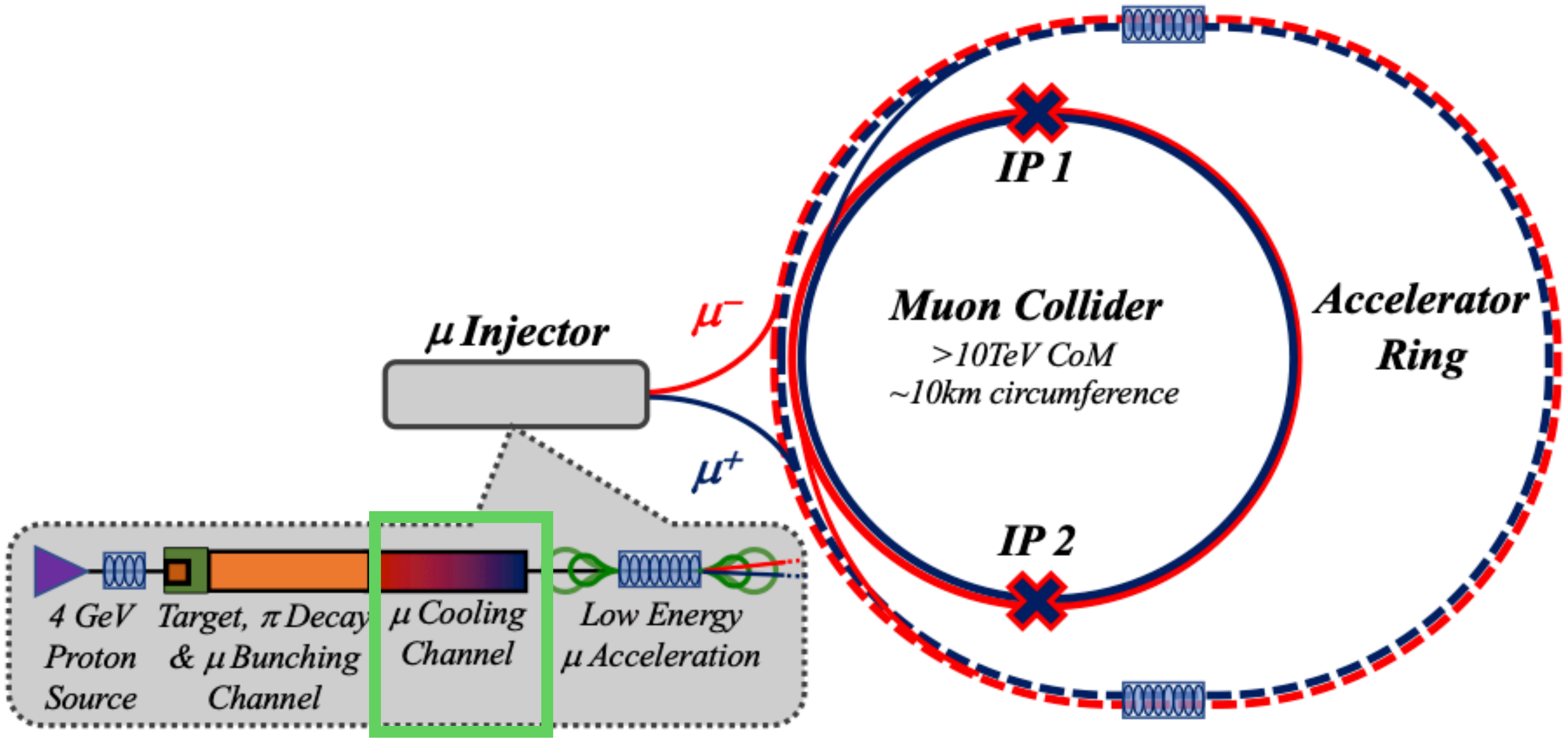
Dense beam

High beam power

Parameter	Unit	3 TeV	10 TeV	14 TeV
L	10^{34} cm^{-2}s^{-1}	1.8	20	40
N	10^{12}	2.2	1.8	1.8
f_r	Hz	5	5	5
P_{beam}	MW	5.3	14.4	20
C	km	4.5	10	14
<B>	T	7	10.5	10.5
ε_L	MeV m	7.5	7.5	7.5
σ_E / E	%	0.1	0.1	0.1
σ_z	mm	5	1.5	1.07
β	mm	5	1.5	1.07
ε	μm	25	25	25
$\sigma_{x,y}$	μm	3.0	0.9	0.63

https://muoncollider.web.cern.ch

Muon Collider: overview



$$\mathcal{L} \propto \gamma \langle B \rangle \sigma_\delta \frac{N_0}{\epsilon \epsilon_L} f_r N_0 \gamma$$

High energy
High field in collider ring
Large energy acceptance
Dense beam
High beam power

Parameter	Unit	3 TeV	10 TeV	14 TeV
L	10^{34} cm^{-2}s^{-1}	1.8	20	40
N	10^{12}	2.2	1.8	1.8
f_r	Hz	5	5	5
P_{beam}	MW	5.3	14.4	20
C	km	4.5	10	14
<B>	T	7	10.5	10.5
ε_L	MeV m	7.5	7.5	7.5
σ_E / E	%	0.1	0.1	0.1
σ_z	mm	5	1.5	1.07
β	mm	5	1.5	1.07
ε	μm	25	25	25
$\sigma_{x,y}$	μm	3.0	0.9	0.63

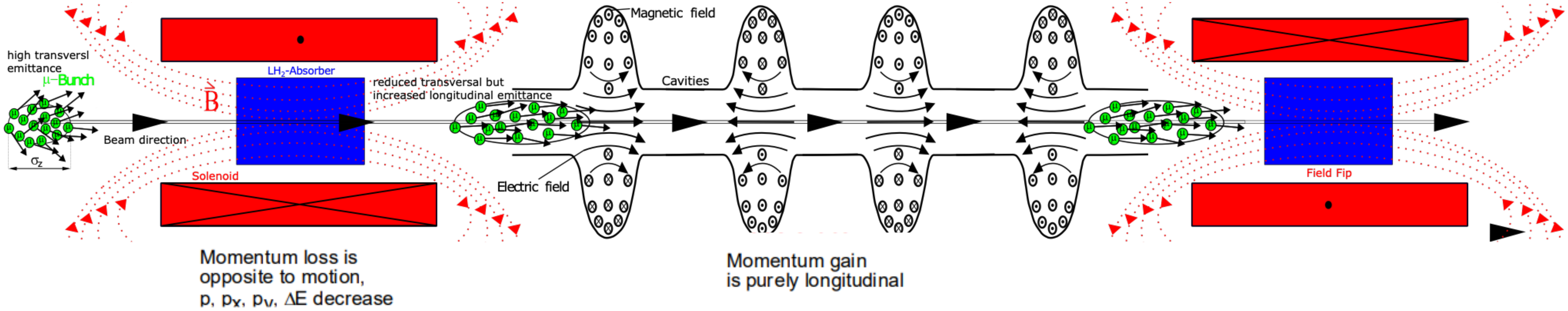
- **Ionisation cooling** (the reduction of occupied phase-space by muons)**: the only technique compatible with muon's lifetime (2.2 μs)**, demonstrated by MICE collaboration
- **Final Cooling Channel:** reduction of transverse emittance on the cost of longitudinal emittance growth

https://muoncollider.web.cern.ch

Technology and challenges of Final Cooling

Ionisation cooling: the only technique that works on the **timescale of the muon lifetime**

- Muons passing through a material —> energy loss due to the interaction with absorber material
- Reduction of normalised beam emittance
- Re-accelerating the beam to restore the longitudinal momentum

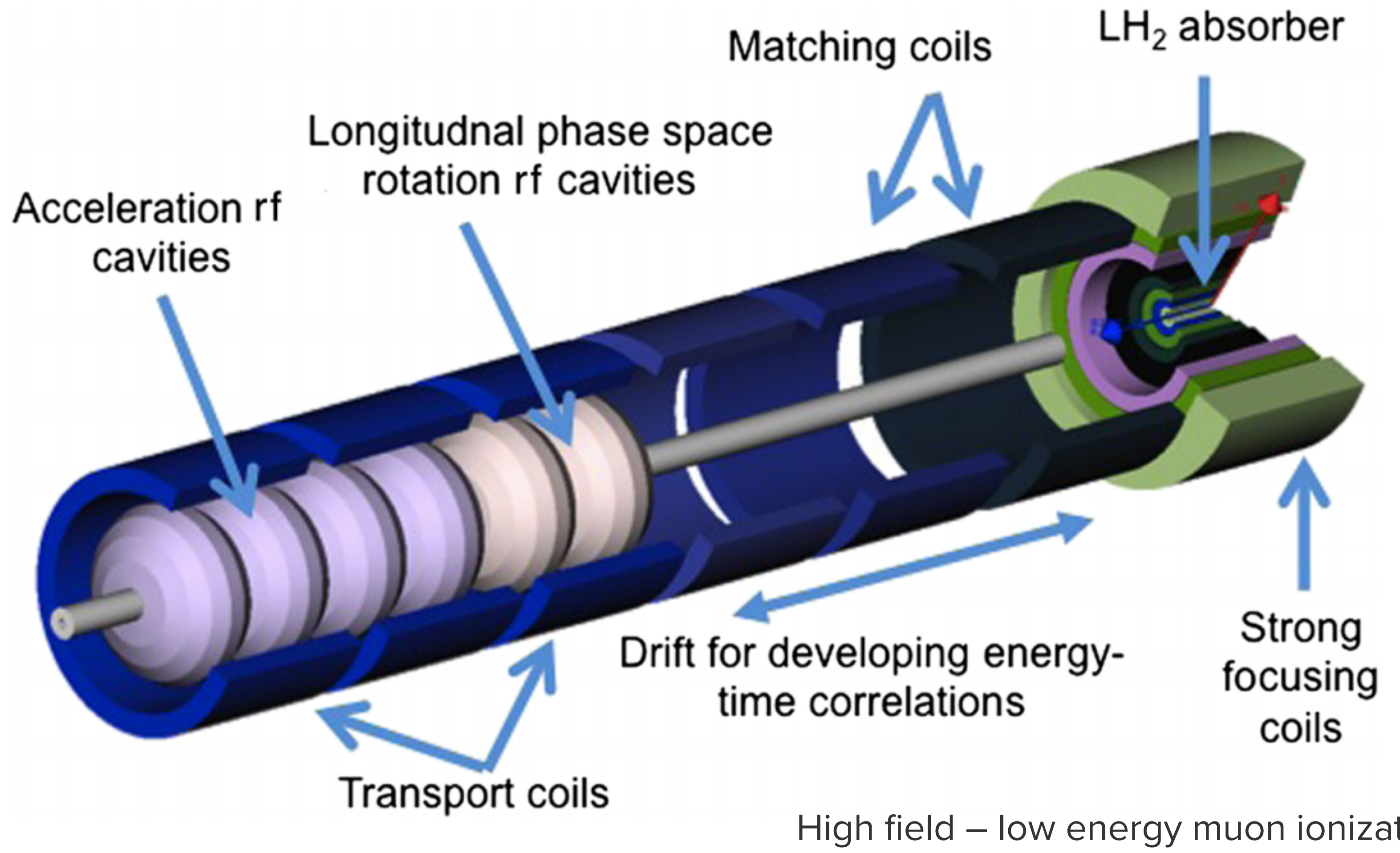


$$\boxed{\frac{d\varepsilon_T}{ds} = -\frac{1}{\beta^2 E}\frac{dE}{ds}\varepsilon_T + \frac{\beta\gamma\beta_T}{2}\frac{d\theta_0^2}{ds}}$$

Energy loss term (cooling)

Multiple scattering term (heating)

Challenges and objectives of Final Cooling



Re-accelerating, rotating the beam
=> Restore P_Z, reduce σE

energy loss due to the interaction with absorber material
=> Reduction of $\epsilon_\perp$

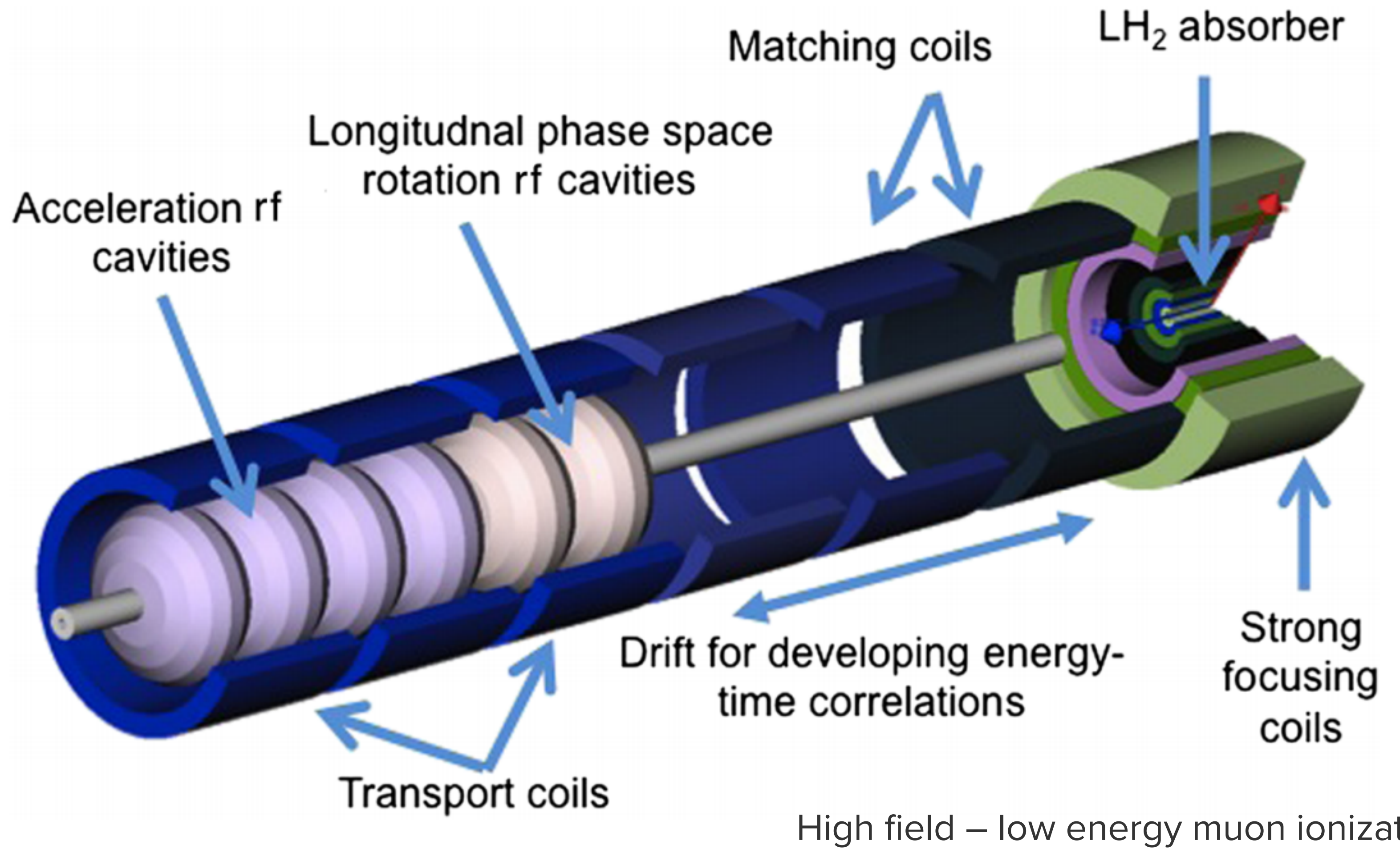
High field – low energy muon ionization cooling channel
Hisham Kamal Sayed, Robert B. Palmer, and David Neuffer
Phys. Rev. ST Accel. Beams **18**, 091001 – Published 4 September 2015

Challenges and objectives of Final Cooling

Lowering transverse emittance on the costs of :

- Longitudinal emittance growth
- Bunch length increasing: challenging RF set-up
- Energy spread
- Particle losses due to decays and energy loss

Re-accelerating, rotating the beam
=> Restore P_Z, reduce σE



High field – low energy muon ionization cooling channel
Hisham Kamal Sayed, Robert B. Palmer, and David Neuffer
Phys. Rev. ST Accel. Beams **18**, 091001 – Published 4 September 2015

energy loss due to the interaction with absorber material
=> Reduction of $\epsilon_\perp$

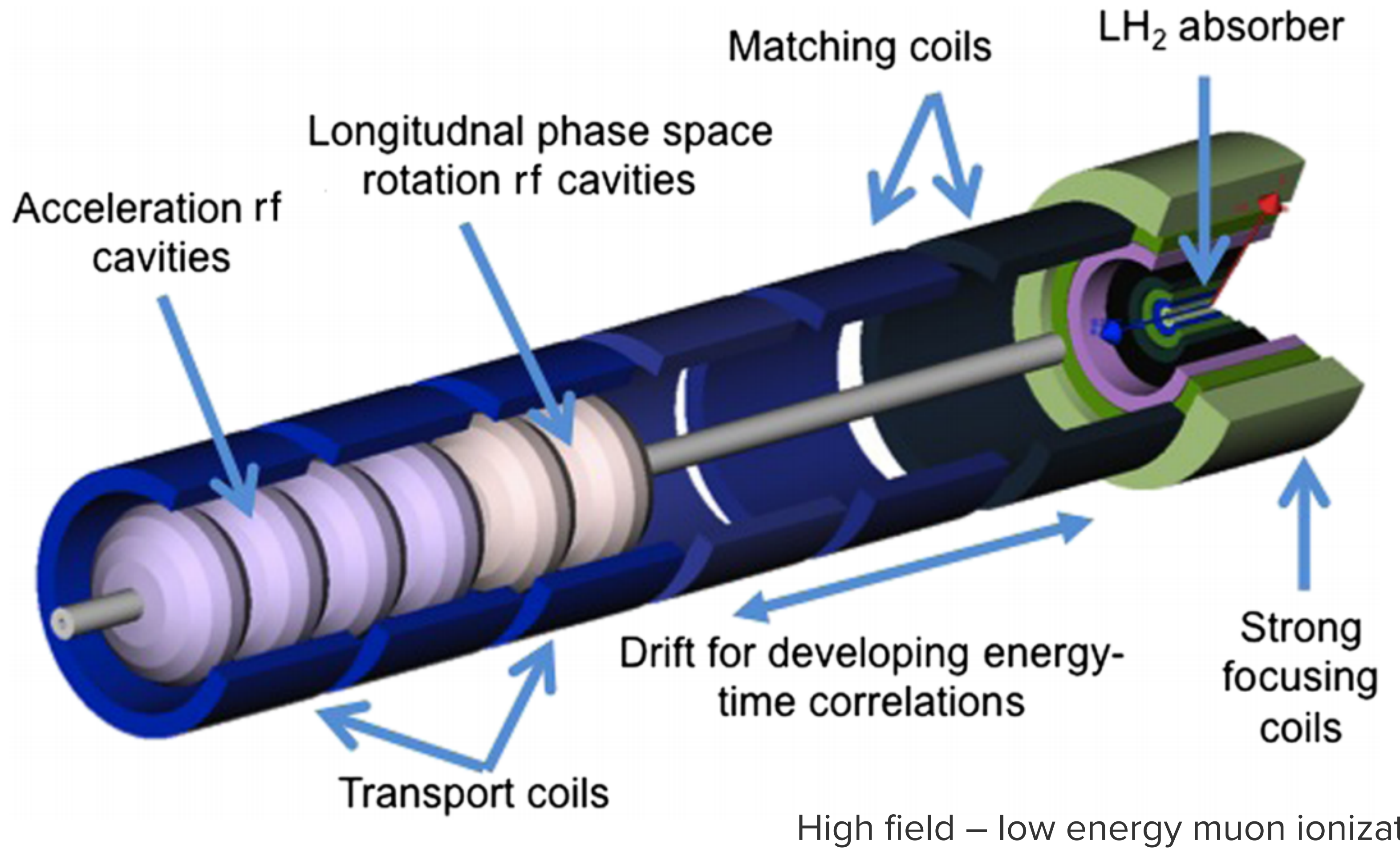
- Achieved in previous studies*: $\varepsilon_\perp$ = 55 µm, with $\varepsilon_\parallel$ = 76 mm, transmission 50%
- Target is $\varepsilon_\perp$ = 25µm => to be achieved using higher solenoid field, **optimization**

Challenges and objectives of Final Cooling

Lowering transverse emittance on the costs of :

- Longitudinal emittance growth
- Bunch length increasing: challenging RF set-up
- Energy spread
- Particle losses due to decays and energy loss

Re-accelerating, rotating the beam
=> Restore P_Z, reduce σE



energy loss due to the interaction with absorber material
=> Reduction of $\epsilon_{\perp}$

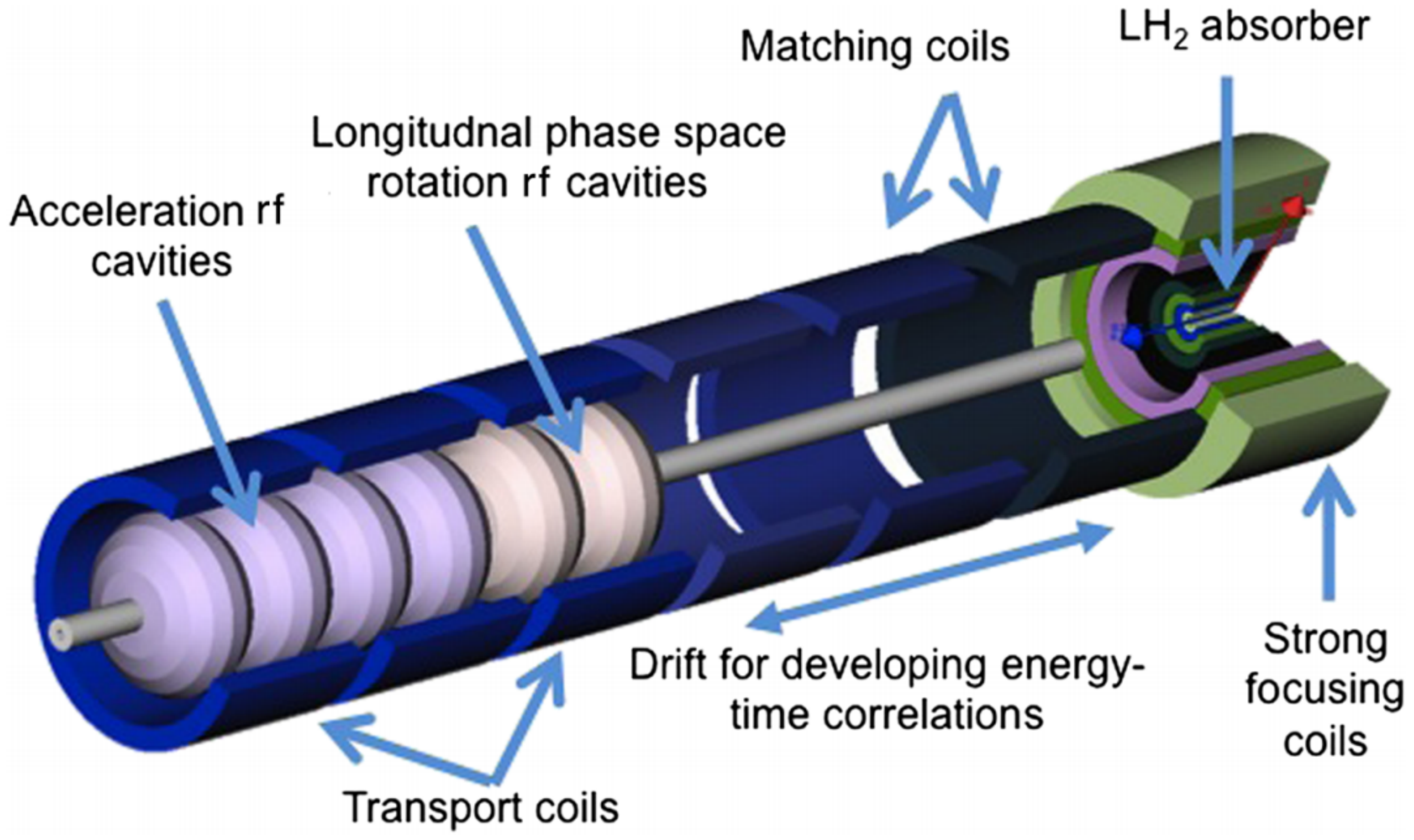
High field – low energy muon ionization cooling channel
Hisham Kamal Sayed, Robert B. Palmer, and David Neuffer
Phys. Rev. ST Accel. Beams **18**, 091001 – Published 4 September 2015

- Achieved in previous studies*: $\varepsilon_{\perp}$ = 55 µm, with $\varepsilon_{\parallel}$ = 76 mm, transmission 50%
- Target is $\varepsilon_{\perp}$ = 25µm => to be achieved using higher solenoid field, **optimization**

- How to **speed up** simulations-based design optimization?
- How to **estimate initial optimization parameters**?
- Robust **emittance estimation** during optimization?

▶ Surrogate models
▶ Feature Importance Analysis with Decision Trees
▶ Bayesian Optimization
▶ Clustering and anomaly detection

Final Cooling: Optimization Strategy



- Global optimization:
  would have **14 parameters** to optimize in **each cell**
- Expected to need **~16 cells in total**

▶ Step-by-step approach, testing different optimization algorithms

Final Cooling: Optimization Strategy

I. Estimate **optimal momenta and absorber lengths** in every cell, with objective $\epsilon_{\perp} = 25\mu m$.

- Nelder – Mead
- Using cooling equations* as objective function

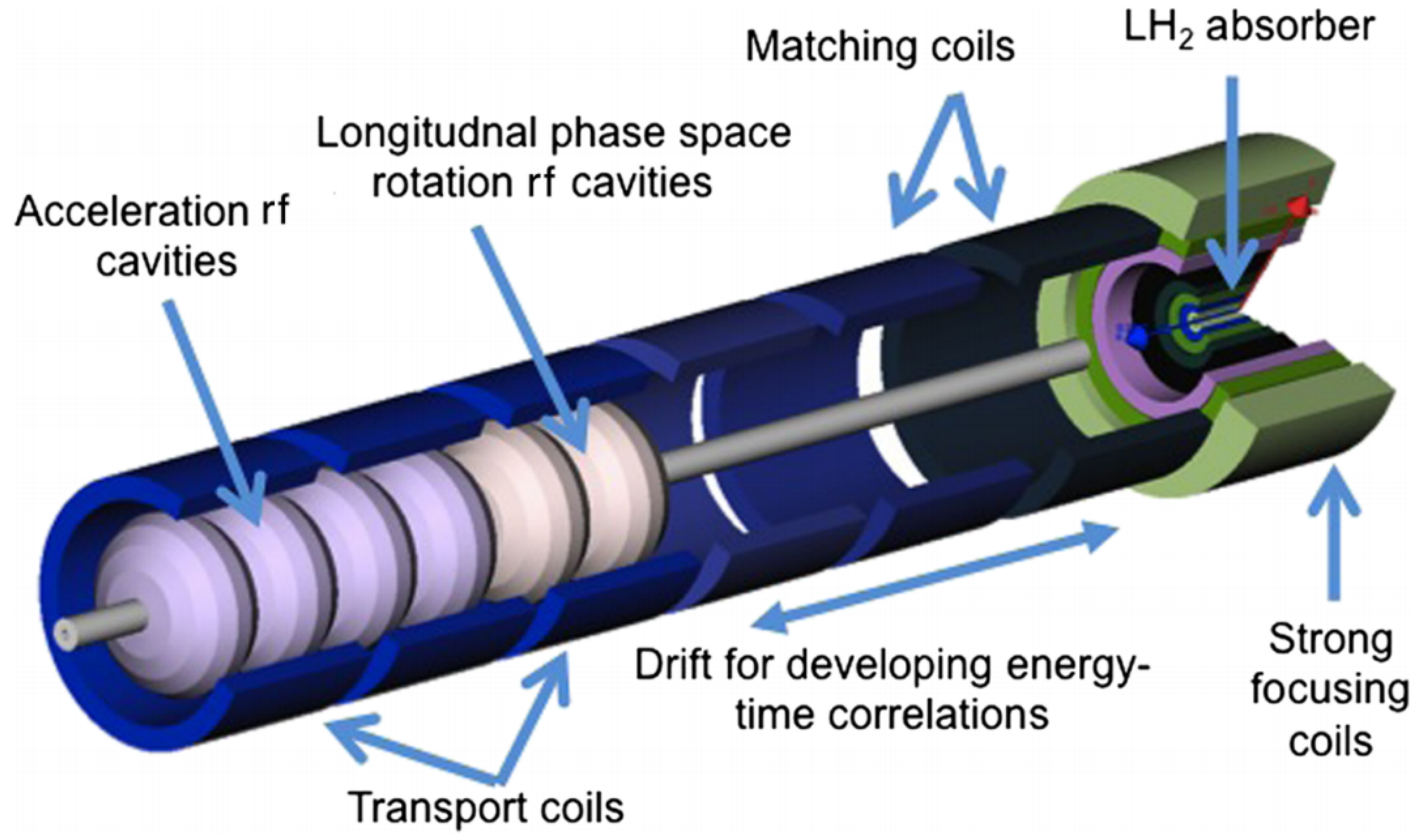
II. **Optics control**, ensure low beta-function in absorber by **optimizing solenoid field and matching coils**

- Numerical optimization, simulations
- Surrogate model (Decision-tree based)

III. **Optimize acceleration and rotation** of the bunch after absorber (simplified RF model)

IV. **Optimize a realistic RF system**: frequencies, phases, gradients to **control the longitudinal dynamics**

- Bayesian Optimization, BOBYQA
- Clustering for robust emittance estimation



- Global optimization: would have **14 parameters** to optimize in **each cell**
- Expected to need **~16 cells in total**

▸ Step-by-step approach, testing different optimization algorithms

* D. Neuffer, Introduction to muon cooling, Nucl. Instrum. Methods Phys. Res., Sect. A 532, 26 (2004).

Initial beam momenta and absorber thickness

1. Estimate **optimal momenta and absorber lengths** in every cell, with objective $\epsilon_\perp = 25\mu m$.

$$\frac{d\epsilon_\perp}{ds} = -\frac{\epsilon_\perp}{\beta^2 E}\frac{dE}{ds} + \frac{\beta_\perp E_s^2}{2\beta^3 mc^2 L_R E}$$

$$\frac{dE}{ds} = 4\pi N_A \rho r_e^2 m_e c^2 \frac{Z}{A}\left[\frac{1}{\beta^2}\ln\left(\frac{2m_e c^2\gamma^2\beta^2}{I(Z)}\right) - 1 - \frac{\delta}{2\beta^2}\right]$$

➡ Provides **starting momenta** and **absorber lengths** for all cells

- 40 T, Liquid hydrogen absorber, initial beam:
  $P_z = 135MeV/c, \epsilon_\perp = 300\mu m, \epsilon_\parallel = 1.5mm, \sigma t = 50mm, \sigma E = 3.2MeV$

Cell	P_z [MeV/c]	Absorber [cm]	$\epsilon_{\perp,start}[\mu m]$	$\epsilon_{\perp,end}$	$P_{z,end}$
14	65	14	40	**24.5**	10
13	70	15	50	40	55.5
12	76	13	70	50	40
11	75	15	85	70	53.5
10	89.2	22	100	85	67.5
9	92.6	21	115	100	74
8	110	25	125	114.6	93.6
7	115	34	140	124.7	93.4
6	124.5	37	155	140	103.4
5	120	36	175	155	98.5
4	127.5	43	200	175	102.4
3	130	40	225	200	108.5
2	125	45	260	220	99
1	135	55	**300**	250	106

- *Note: this assumes ideal optics matching and control of longitudinal parameters*
- *Transmission is not included*

Initial beam momenta and absorber thickness

I. Estimate **optimal momenta and absorber lengths** in every cell, with objective $\epsilon_\perp = 25\mu m$.

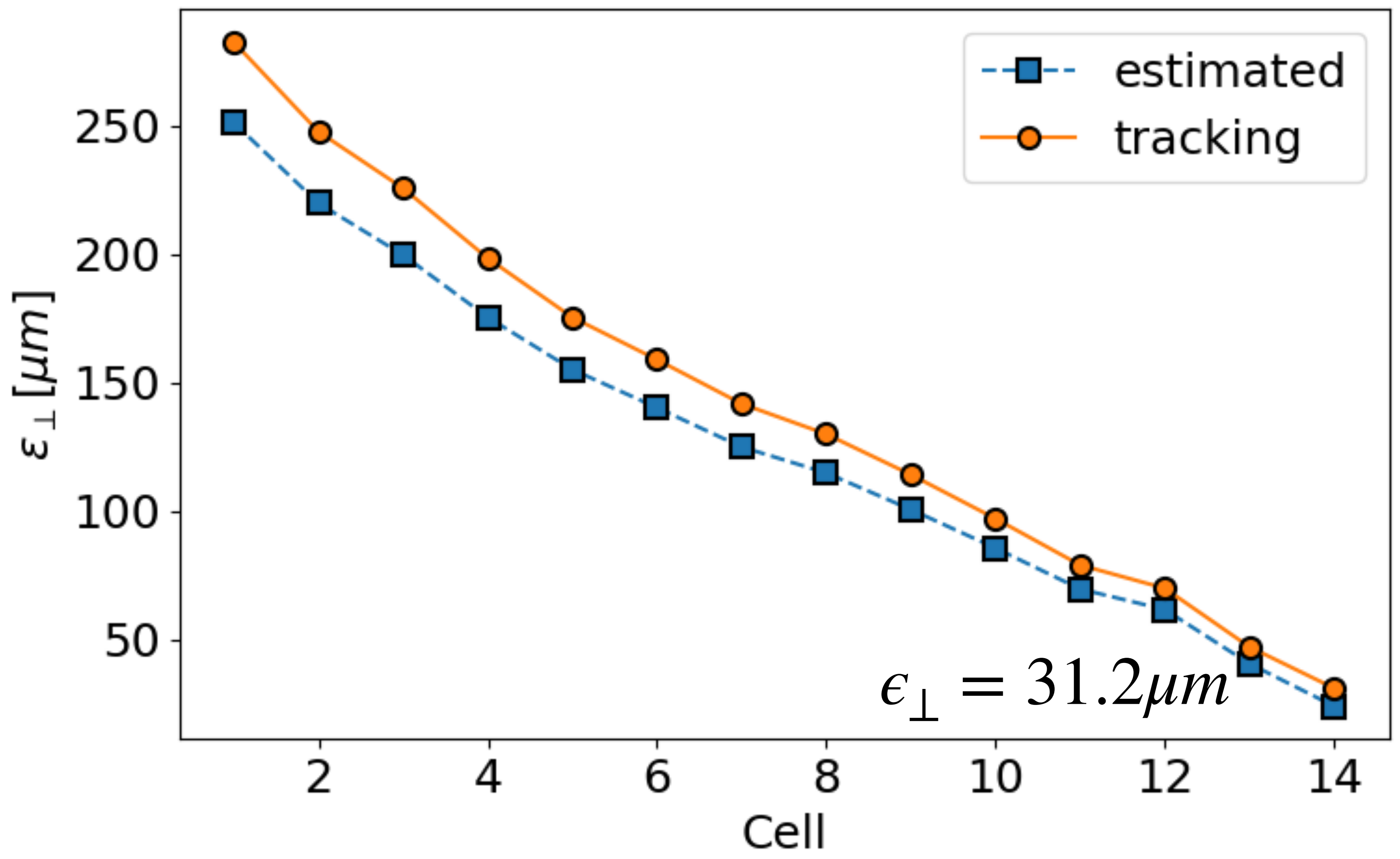
$$\frac{d\epsilon_\perp}{ds} = -\frac{\epsilon_\perp}{\beta^2 E}\frac{dE}{ds} + \frac{\beta_\perp E_s^2}{2\beta^3 mc^2 L_R E}$$

$$\frac{dE}{ds} = 4\pi N_A \rho r_e^2 m_e c^2 \frac{Z}{A}\left[\frac{1}{\beta^2}\ln\left(\frac{2m_e c^2\gamma^2\beta^2}{I(Z)}\right) - 1 - \frac{\delta}{2\beta^2}\right]$$

➡ Provides **starting momenta** and **absorber lengths** for all cells

- 40 T, Liquid hydrogen absorber, initial beam:
  $P_z = 135MeV/c, \epsilon_\perp = 300\mu m, \epsilon_\parallel = 1.5mm, \sigma t = 50mm, \sigma E = 3.2MeV$

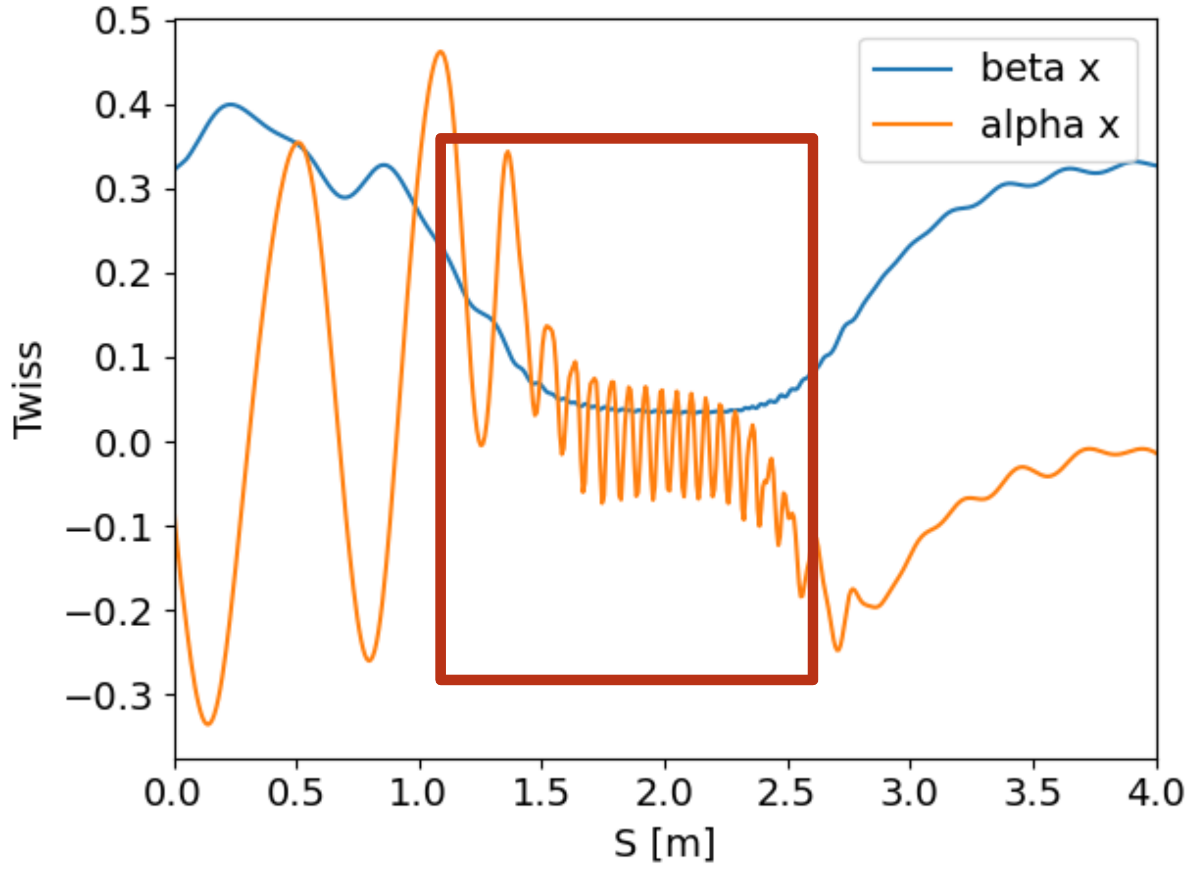
Cell	P_z [MeV/c]	Absorber [cm]	$\epsilon_{\perp,start}[\mu m]$	$\epsilon_{\perp,end}$	$P_{z,end}$
14	65	14	40	**24.5**	10
13	70	15	50	40	55.5
12	76	13	70	50	40
11	75	15	85	70	53.5
10	89.2	22	100	85	67.5
9	92.6	21	115	100	74
8	110	25	125	114.6	93.6
7	115	34	140	124.7	93.4
6	124.5	37	155	140	103.4
5	120	36	175	155	98.5
4	127.5	43	200	175	102.4
3	130	40	225	200	108.5
2	125	45	260	220	99
1	135	55	**300**	250	106

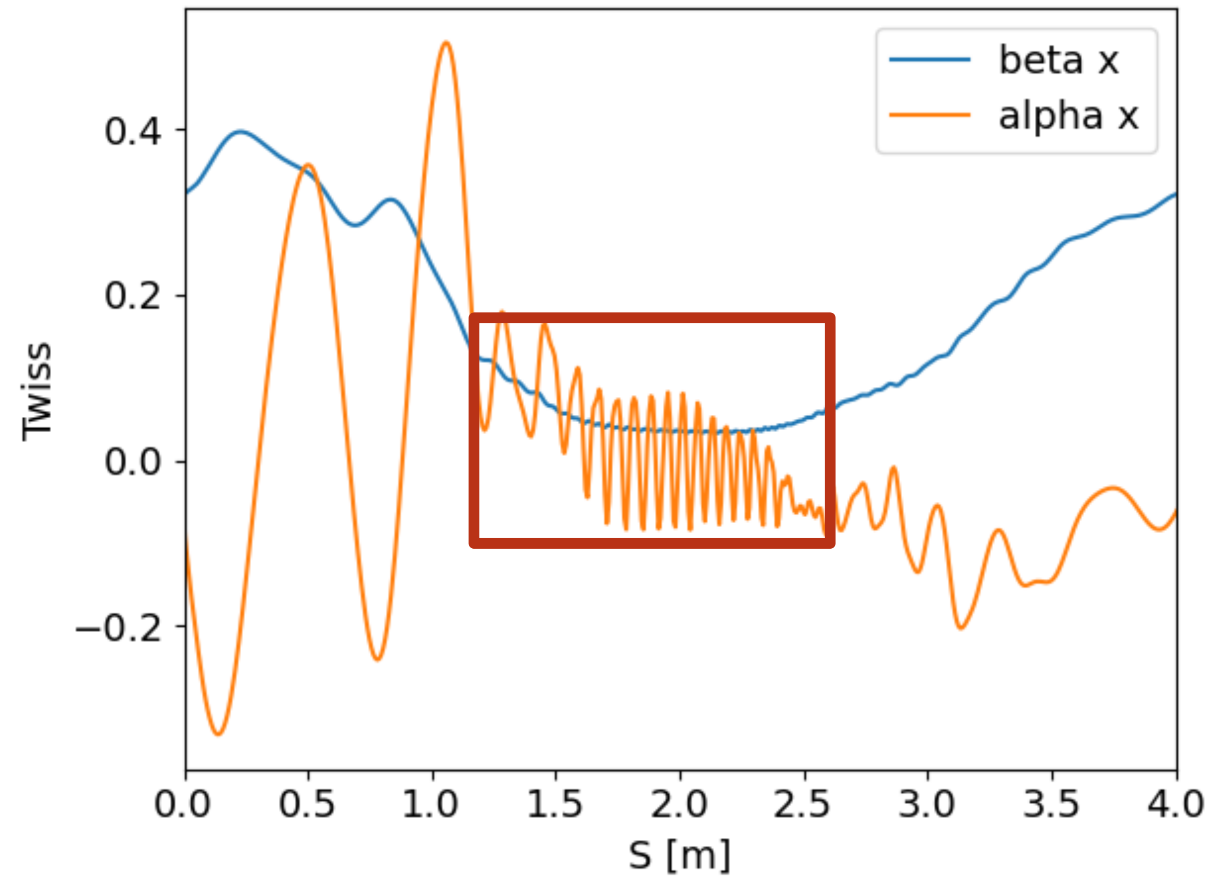


✓ Tracking simulations using **optimised parameters** confirm the potential for **lower emittance** (compared to the baseline studies)

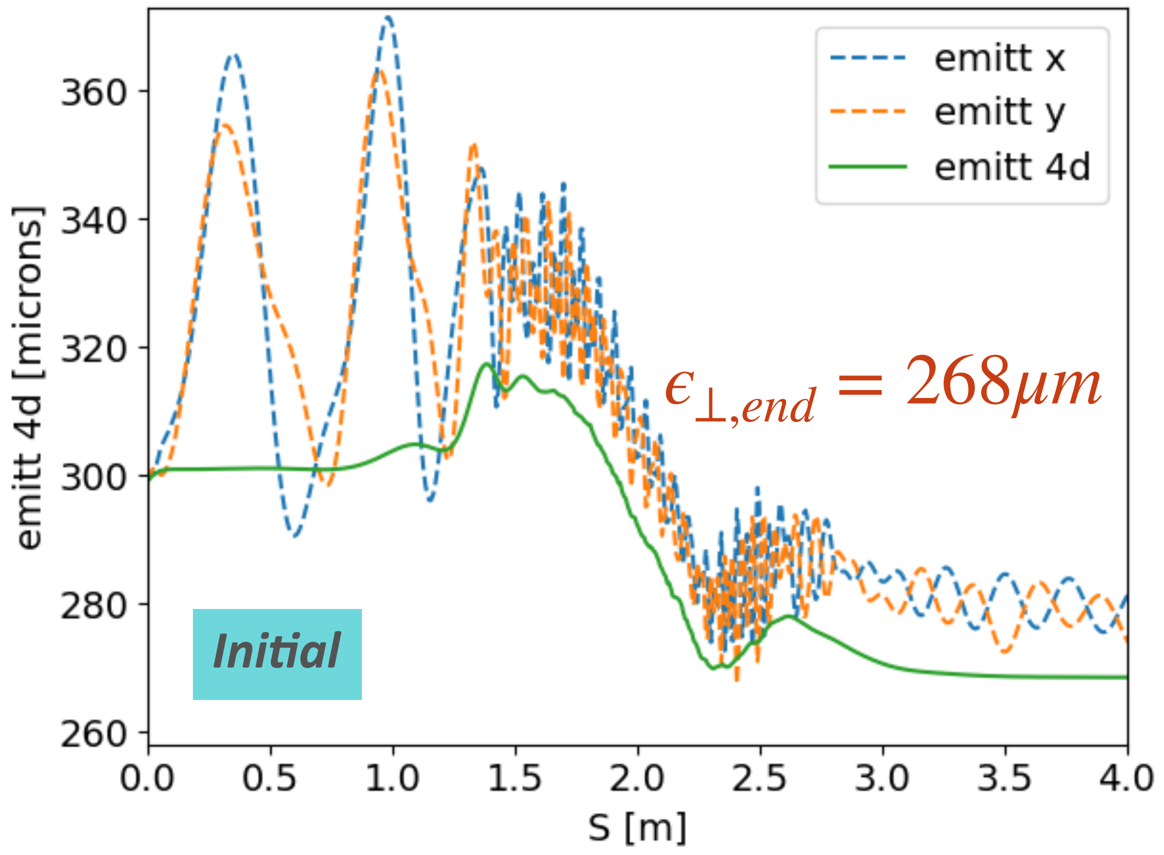
Optics matching

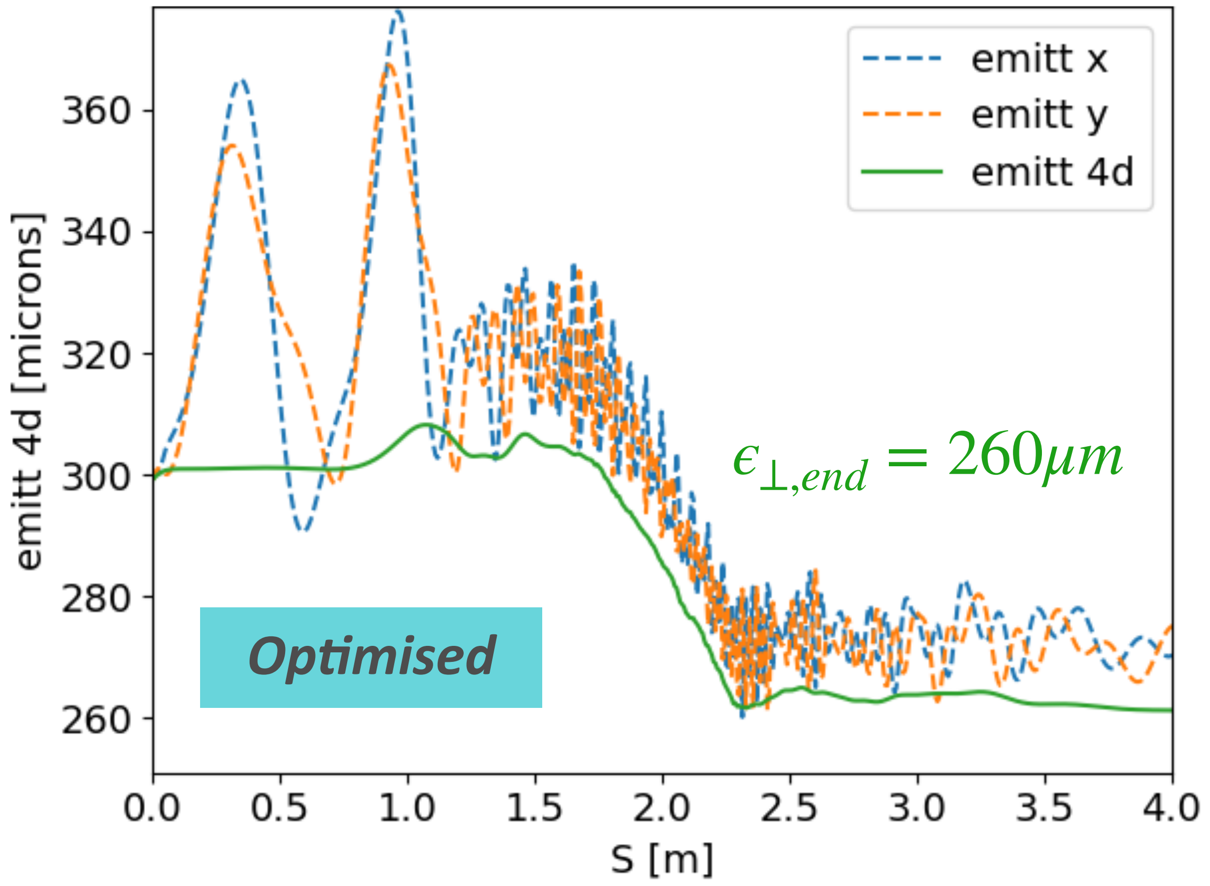
II. **Optics control**, ensure low beta-function in absorber by **optimizing solenoid field and matching coils**





➡ **Mitigates emittance blow up** in the fridge fields and **controls the optics in absorber** region





- Optimization parameters:
  - *- solenoid length,*
  - *- Strength of the low B-field between the cells*
  - *- Matching coils placed left and right from the absorber*
- *Method: **BOBYQA** : fast-executable, gradient-free*

Coralia Cartis, Jan Fiala, Benjamin Marteau and Lindon Roberts, Improving the Flexibility and Robustness of Model-Based Derivative-Free Optimization Solvers, *ACM Transactions on Mathematical Software*, 45:3 (2019), pp. 32:1-32:41

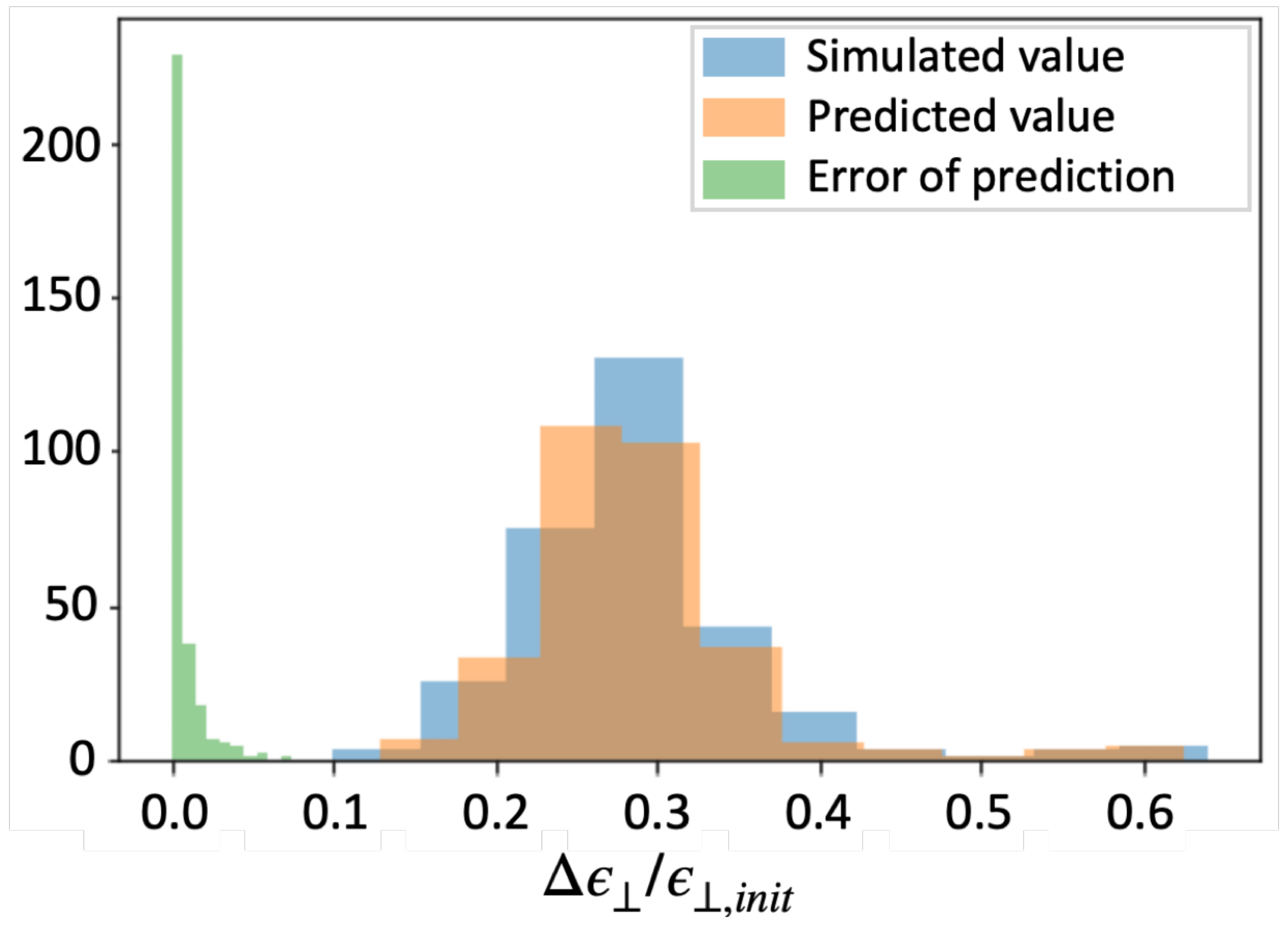
Optimizing solenoid fields: Surrogate Modeling

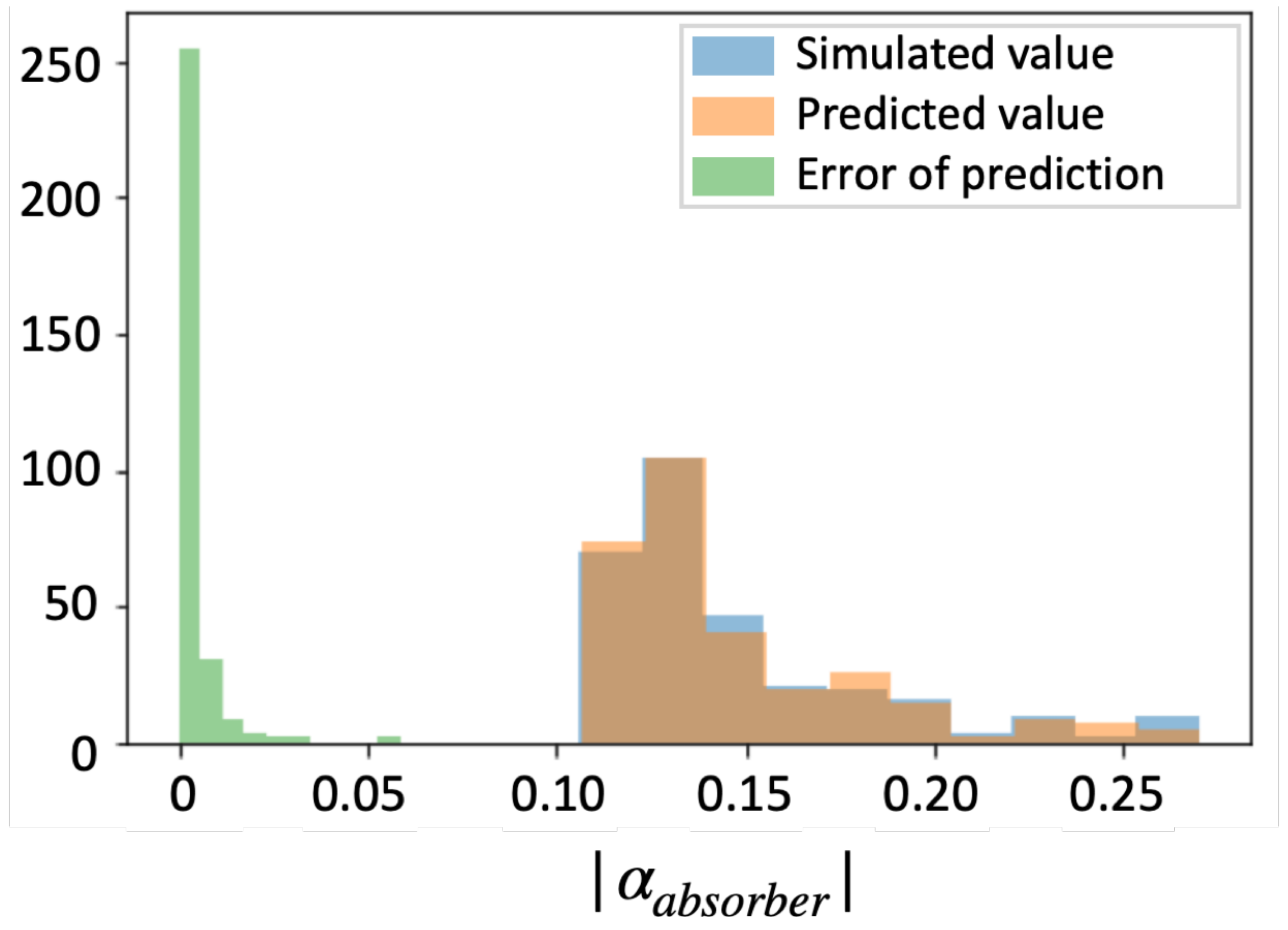
Proof of concept:

1. Run numerical optimisers, **systematically saving the data** (results of tracking simulations using ICOOL)

2. Train a **surrogate model** (Random Forest Regressor):
   - input = parameters of the solenoid field in a cooling cell
   - output = optics observables

3. **Replace time-costly simulations** with ML model, find optimal parameters





✓ Compute optimization function from ML-model prediction
✓ Optimization in a **few minutes instead of ~1.5 hours** for 200 steps using tracking simulations

Longitudinal phase-space optimization: Bayesian Optimization

- **Objective function** : $\frac{\epsilon_{\perp}\epsilon_{||}}{N_\mu}$,

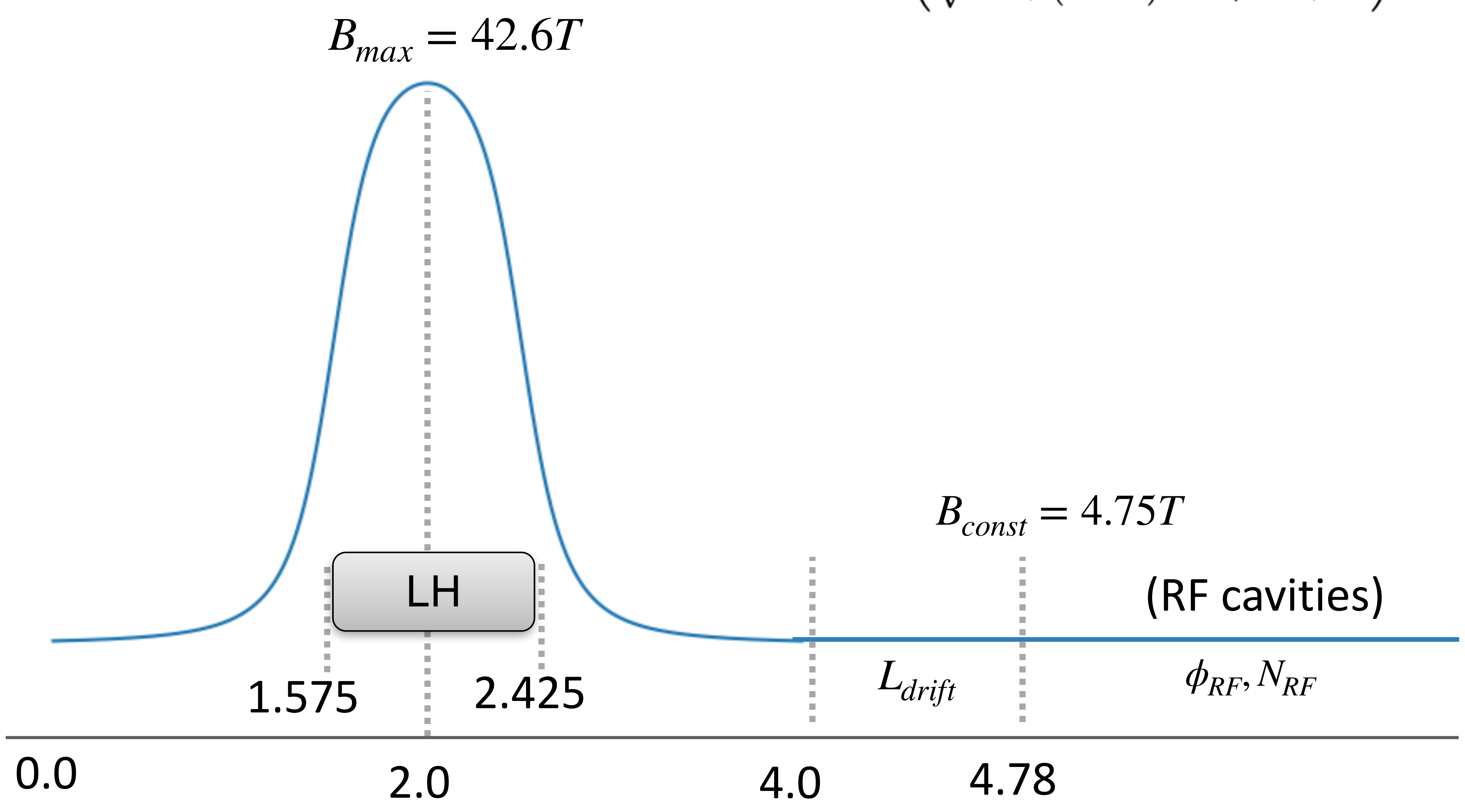
  obtained using RF-Track simulation code

  (developed by A. Latina https://gitlab.cern.ch/rf-track)

Example for cell 1:
Absorber thickness: 0.85 m
Solenoid length = 1.48 m

Solenoid field in RF-Track:

$$B(z) = 0.5 \cdot B_0 \left(\frac{L - z}{\sqrt{R^2 + (L - z)^2}} + \frac{z}{\sqrt{R^2 + z^2}} \right)$$



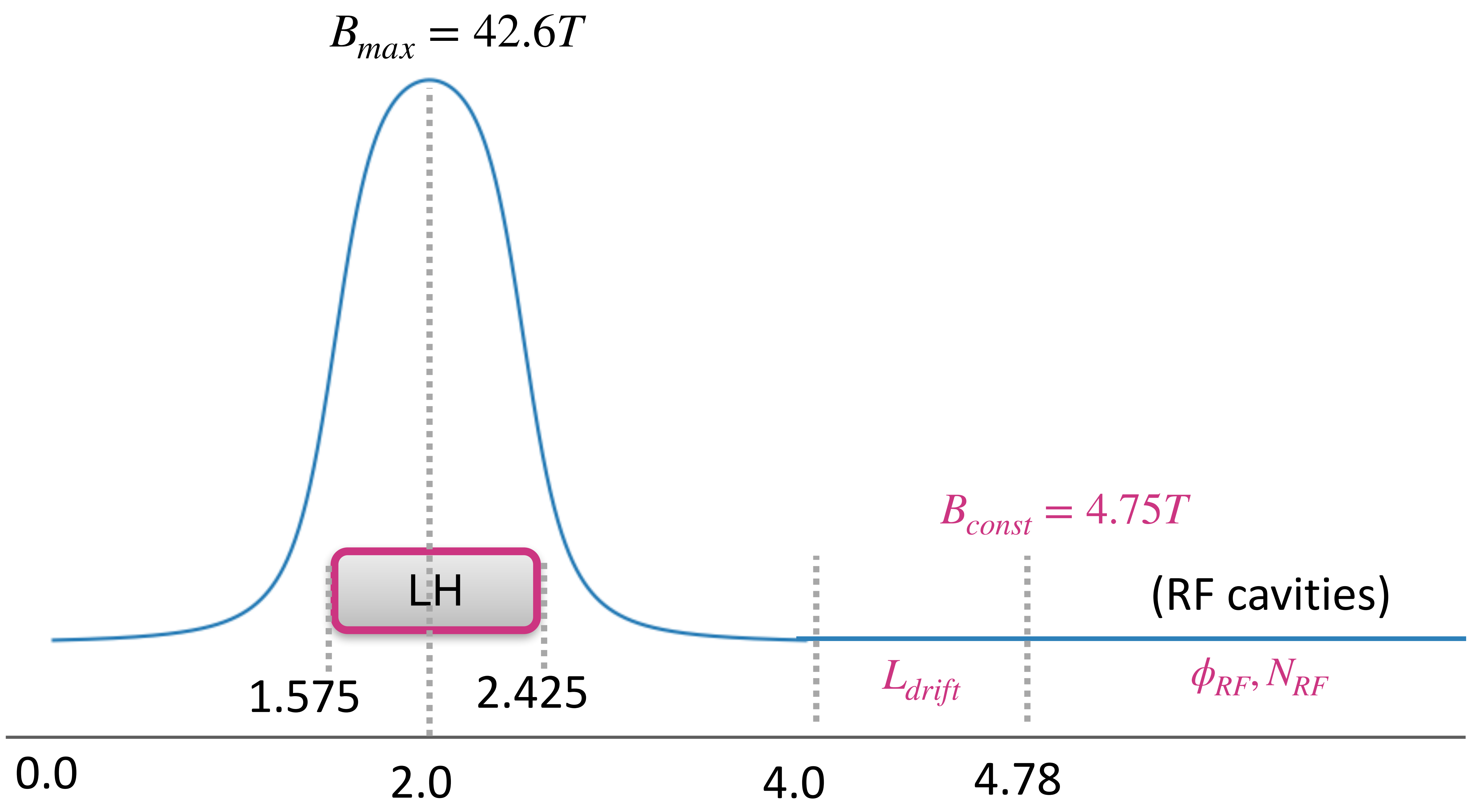
Longitudinal phase-space optimization: Bayesian Optimization

- **Objective function** : $\frac{\epsilon_{\perp}\epsilon_{||}}{N_{\mu}}$,

  obtained using RF-Track simulation code

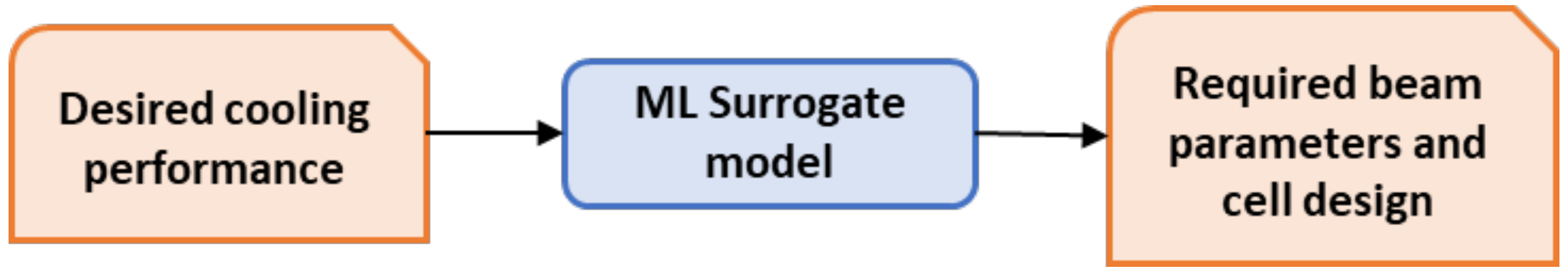
  (developed by A. Latina https://gitlab.cern.ch/rf-track)

- **Free parameters:**
  - Absorber (liquid hydrogen) thickness
  - Drift length
  - Number of accelerating RF cavities, rf phase
  - Number of rotating RF cavities, rf phase
  - B-field in RF region to match the field in the cooling cell and the change in momentum



Longitudinal phase-space optimization: Bayesian Optimization

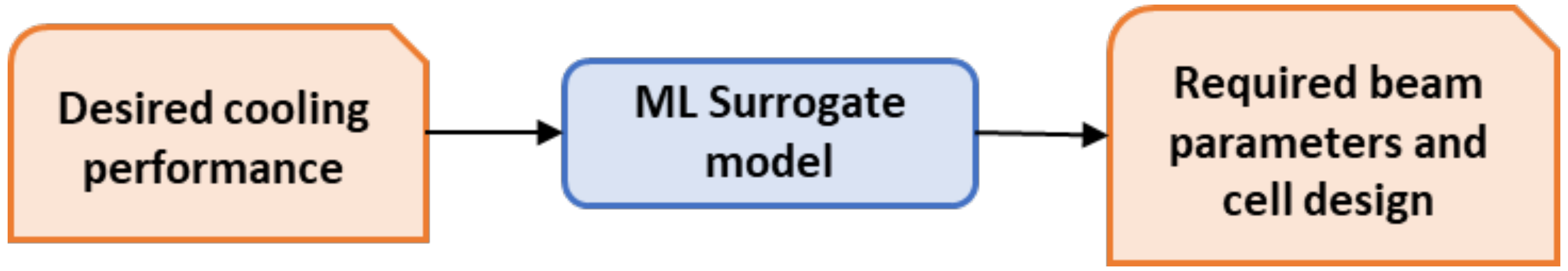
- **Optimization procedure:**
  - Run optimization for each cell, a few iterations
  - Create a surrogate model to estimate the initial parameters
  - Bayesian Optimization*, BOBYQA
- * **Update probabilistic model** based on function evaluation
- Optimise an acquisition function (e.g. expected improvement) for sampling the new optimisation step
- Balance exploration and exploitation by controlling parameters of acquisition function
- Surrogate Model: Random Forest
- Skopt implementation (https://scikit-optimize.github.io)



➡ Fast design estimate

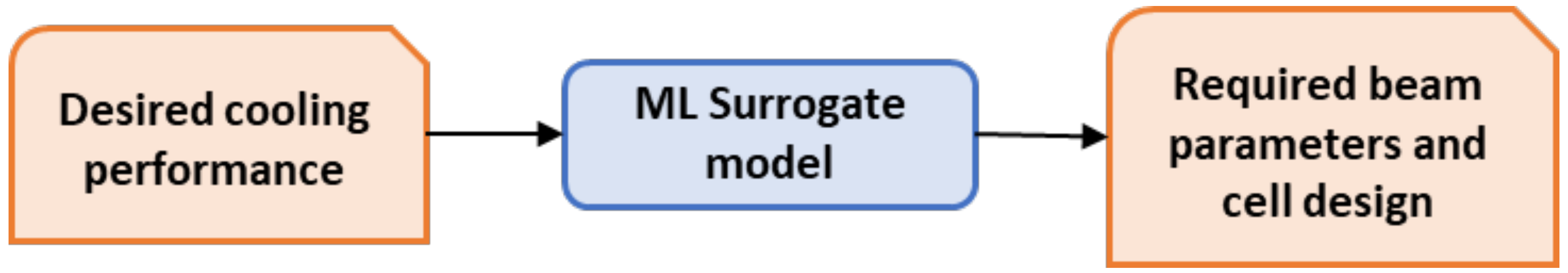
➡ **Use as initial guess** for optimisation algorithms (optimal solution is found within fewer steps)

Surrogate model from simulation data



- **Optimization procedure:**
  - Run optimization for a cell, a few iterations
    => increase exploration rate to create a diverse dataset
  - Create a surrogate model **from valid simulation set-ups** to estimate the initial parameters
  - Update initial parameters in optimiser
    => decrease the exploration rate: look for solutions around (sub-optimal) SM prediction

Surrogate model from simulation data



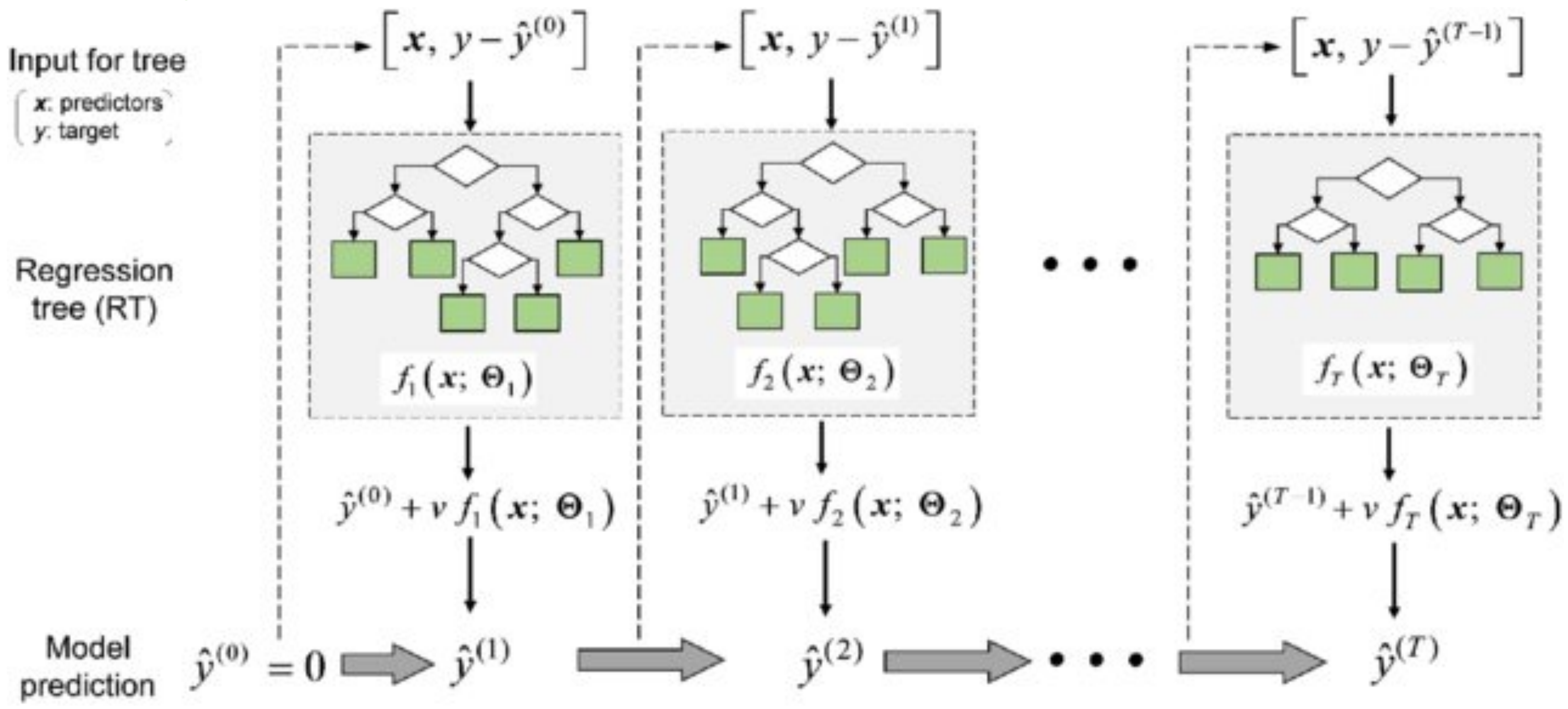
- Input: $\epsilon_{\perp start}, P_{z,start}, \epsilon_{\perp}, \sigma_t, \sigma E, N_{\mu}$
- Output: $L_{drift}, N_{rot}, N_{cav}, \phi_{RF}, L_{absorber}, L_{sol}$
- **Train/Test R2: 0.97 / 0.9 (~5000 samples)**

XGBoost:

- Extreme **Gradient boosting** = ensemble ML algorithms based on decision tree models
- Trees are added one at a time to the ensemble
  => Fit to correct the prediction errors made by prior models
  => Using Gradient Descent
  => Combing “weak learners” into a single strong learner iteratively

Greedy Function Approximation: A Gradient Boosting Machine, Friedman
https://github.com/tqchen/xgboost

- **Optimization procedure:**
  - Run optimization for a cell, a few iterations
    => increase exploration rate to create a diverse dataset
  - Create a surrogate model **from valid simulation set-ups** to estimate the initial parameters
  - Update initial parameters in optimiser
    => decrease the exploration rate: look for solutions around (sub-optimal) SM prediction



Wang, Mao-Xin et. al(2020). SS-XGBoost: A Machine Learning Framework for Predicting Newmark Sliding Displacements of Slopes.

Surrogate model from simulation data

Desired cooling performance → ML Surrogate model → Required beam parameters and cell design

- Input: $\epsilon_{\perp start}, P_{z,start}, \epsilon_{\perp}, \sigma_t, \sigma E, N_\mu$
- Output: $L_{drift}, N_{rot}, N_{cav}, \phi_{RF}, L_{absorber}, L_{sol}$
- Train/Test R2: 0.97 / 0.9 (~5000 samples)

Example, cell 4: $\epsilon_{\perp,start} = 170\mu m$

- Training on the full data set, collected from cell 1-12, ~5000 samples

  Target:

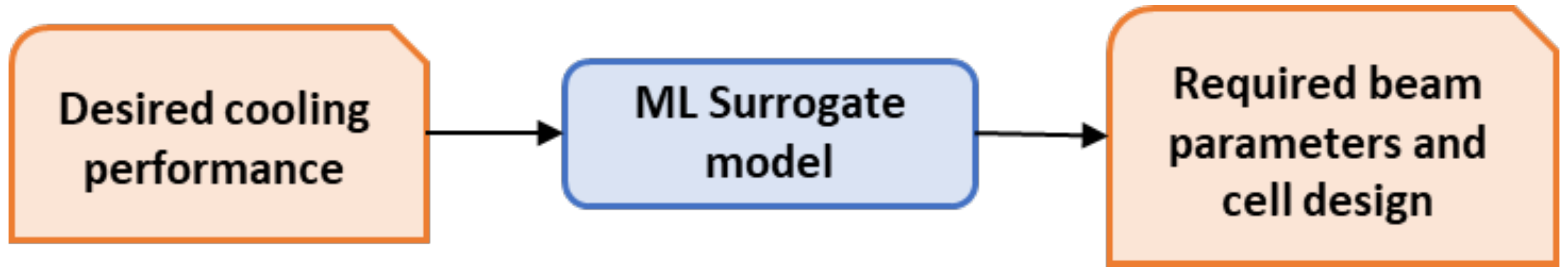
  $$\epsilon_{\perp} = 150\mu m, \sigma_t = 400mm, \sigma E = 2.0MeV, N_\mu = 75\,\%$$

- **Optimization procedure:**
  - Run optimization for a cell, a few iterations
    => increase exploration rate to create a diverse dataset
  - Create a surrogate model **from valid simulation set-ups** to estimate the initial parameters
  - Update initial parameters in optimiser
    => decrease the exploration rate: look for solutions around (sub-optimal) SM prediction

International MUON Collider Collaboration

CERN

Surrogate model from simulation data



- Input: $\epsilon_{\perp start}, P_{z,start}, \epsilon_{\perp}, \sigma_t, \sigma E, N_{\mu}$
- Output: $L_{drift}, N_{rot}, N_{cav}, \phi_{RF}, L_{absorber}, L_{sol}$
- Train/Test R2: 0.97 / 0.9 (~5000 samples)

Example, cell 4: $\epsilon_{\perp,start} = 170 \mu m$

- Training on the full data set, collected from cell 1-12, ~5000 samples

  Target:

  $$\epsilon_{\perp} = 150 \mu m, \sigma_t = 400 mm, \sigma E = 2.0 MeV, N_{\mu} = 75\,\%$$

  Simulated with parameters predicted by ML-model:

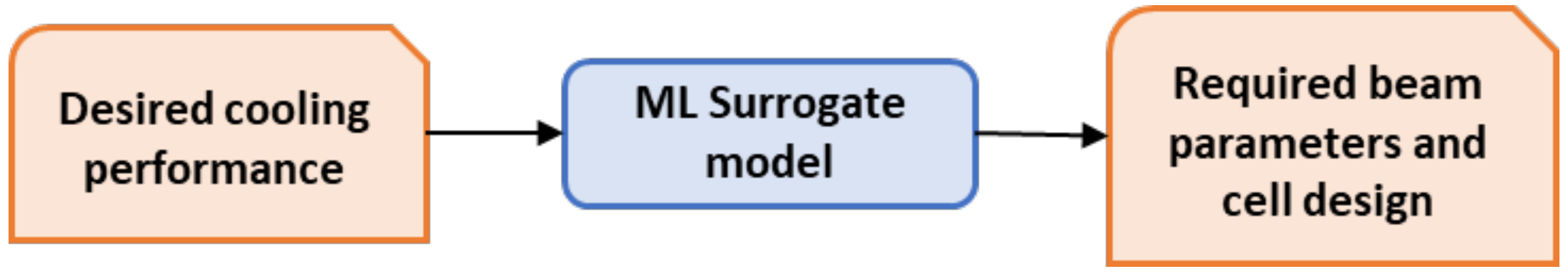
  $$\epsilon_{\perp} = 149 \mu m, \sigma_t = 404 mm, \sigma E = 3.5 MeV, N_{\mu} = 69\,\%$$

- **Optimization procedure:**
  - Run optimization for a cell, a few iterations

    => increase exploration rate to create a diverse dataset
  - Create a surrogate model **from valid simulation set-ups** to estimate the initial parameters
  - Update initial parameters in optimiser

    => decrease the exploration rate: look for solutions around (sub-optimal) SM prediction

Surrogate model from simulation data



- Input: $\epsilon_{\perp start}, P_{z,start}, \epsilon_{\perp}, \sigma_t, \sigma E, N_\mu$
- Output: $L_{drift}, N_{rot}, N_{cav}, \phi_{RF}, L_{absorber}, L_{sol}$
- Train/Test R2: 0.97 / 0.9 (~5000 samples)

Example, cell 4: $\epsilon_{\perp,start} = 170 \mu m$

- Training on the full data set, collected from cell 1-12, ~5000 samples

  Target:

  $$\epsilon_{\perp} = 150 \mu m, \sigma_t = 400 mm, \sigma E = 2.0 MeV, N_\mu = 75\,\%$$

  Simulated with parameters predicted by ML-model:

  $$\epsilon_{\perp} = 149 \mu m, \sigma_t = 404 mm, \sigma E = 3.5 MeV, N_\mu = 69\,\%$$

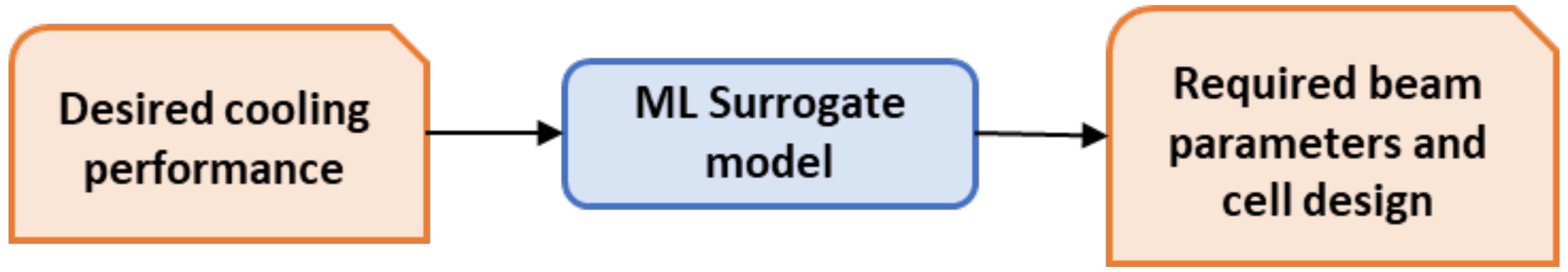
  Optimiser, 150 steps, starting with predicted parameters:

  $$\epsilon_{\perp} = 150 \mu m, \sigma_t = 280 mm, \sigma E = 2.1 MeV, N_\mu = 71\,\%$$

- **Optimization procedure:**
  - Run optimization for a cell, a few iterations
    => increase exploration rate to create a diverse dataset
  - Create a surrogate model **from valid simulation set-ups** to estimate the initial parameters
  - Update initial parameters in optimiser
    => decrease the exploration rate: look for solutions around (sub-optimal) SM prediction

International MUON Collider Collaboration

Surrogate model from simulation data



- Input: $\epsilon_{\perp start}, P_{z,start}, \epsilon_{\perp}, \sigma_t, \sigma E, N_{\mu}$
- Output: $L_{drift}, N_{rot}, N_{cav}, \phi_{RF}, L_{absorber}, L_{sol}$
- Train/Test R2: 0.97 / 0.9 (~5000 samples)

Example, cell 4: $\epsilon_{\perp,start} = 170\mu m$

- Training on the full data set, collected from cell 1-12, ~5000 samples

  Target:

  $\epsilon_{\perp} = 150\mu m, \sigma_t = 400mm, \sigma E = 2.0MeV, N_{\mu} = 75\,\%$

  Simulated with parameters predicted by ML-model:

  $\epsilon_{\perp} = 149\mu m, \sigma_t = 404mm, \sigma E = 3.5MeV, N_{\mu} = 69\,\%$

  Optimiser, 150 steps, starting with predicted parameters:

  $\epsilon_{\perp} = 150\mu m, \sigma_t = 280mm, \sigma E = 2.1MeV, N_{\mu} = 71\,\%$

- **Optimization procedure:**
  - Run optimization for a cell, a few iterations

    => increase exploration rate to create a diverse dataset
  - Create a surrogate model **from valid simulation set-ups** to estimate the initial parameters
  - Update initial parameters in optimiser

    => decrease the exploration rate: look for solutions around (sub-optimal) SM prediction

- Training **only** on data, collected for the **current cell**

  Note: here $\epsilon_{\perp start}, P_{z,start}$ ***are not included, 320 samples***

  Target:

  $\epsilon_{\perp} = 140\mu m, \sigma_t = 400mm, \sigma E = 2.0MeV, N_{\mu} = 70\,\%$

  Simulated with parameters predicted by ML-model:

  $\epsilon_{\perp} = 140\mu m, \sigma_t = 600mm, \sigma E = 3.4MeV, N_{\mu} = 66\,\%$

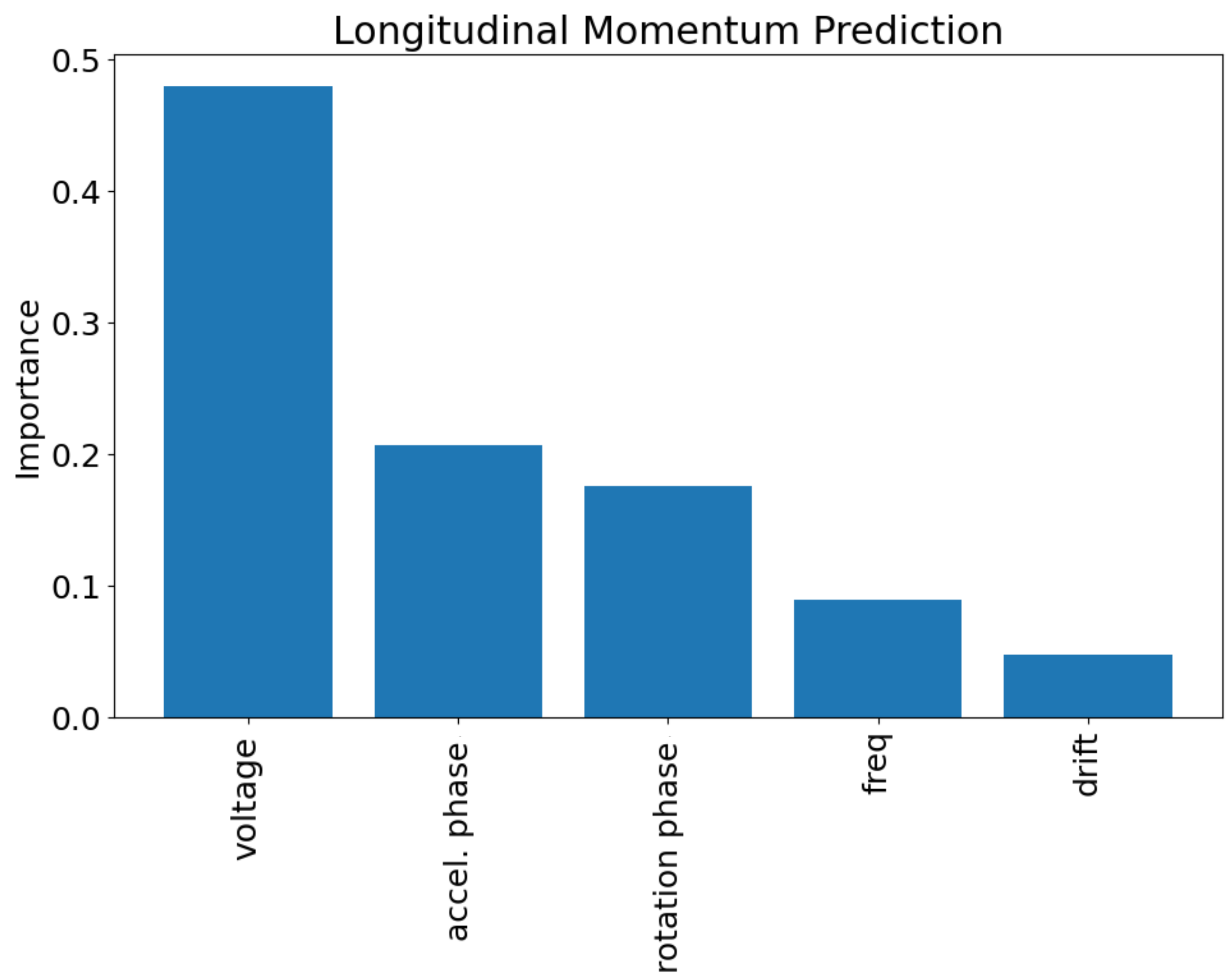
Model interpretability: permutation features importance

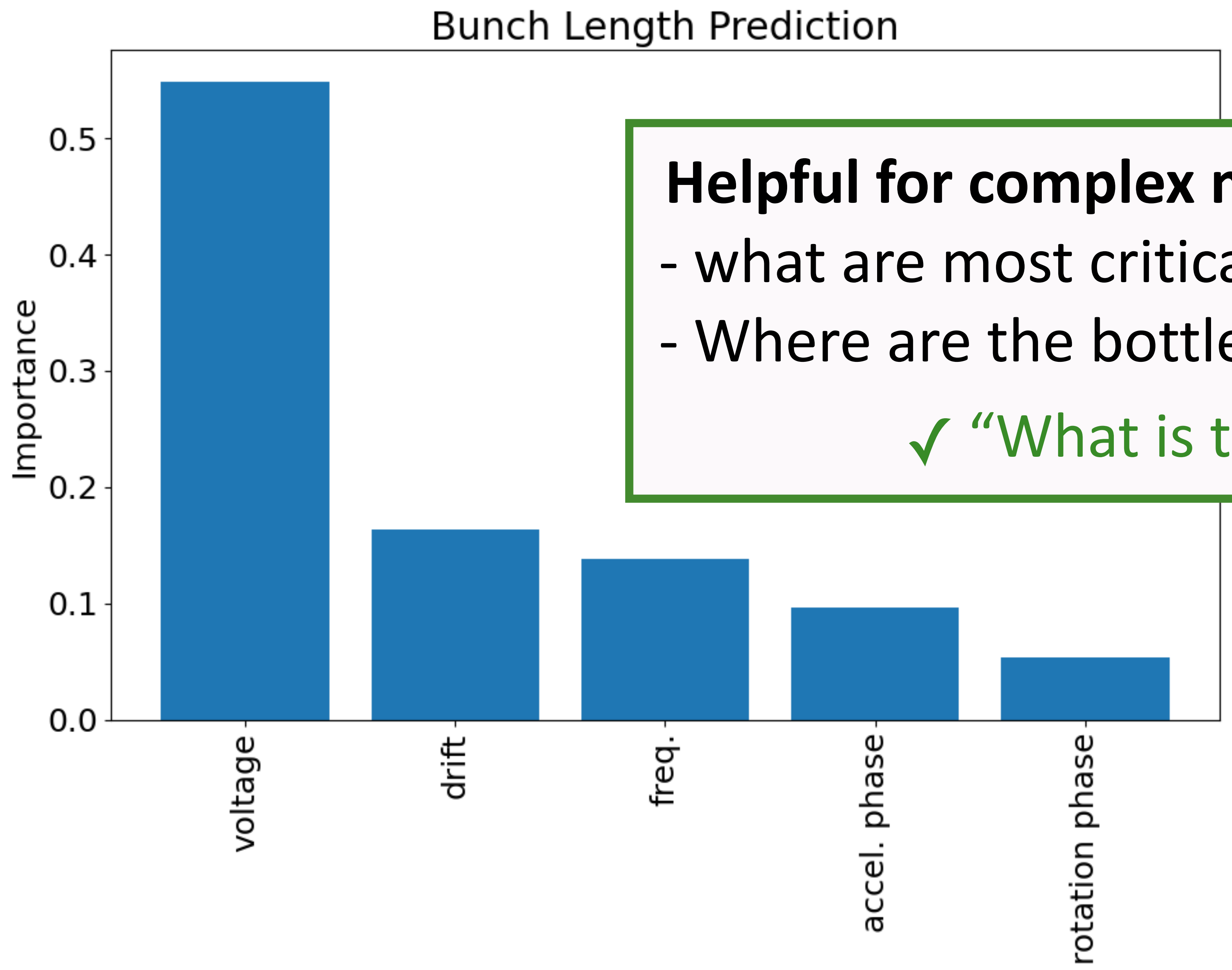
Feature permutation

- Measuring how much **model's performance decreases** when each **feature is randomly shuffled**
- Identify **which features have greatest impact** on model's output

Example: optimization of RF in cooling cells:

- Model created from optimization data: Input: **cell set up**, output: **beam parameters** at the end of a cooling cell





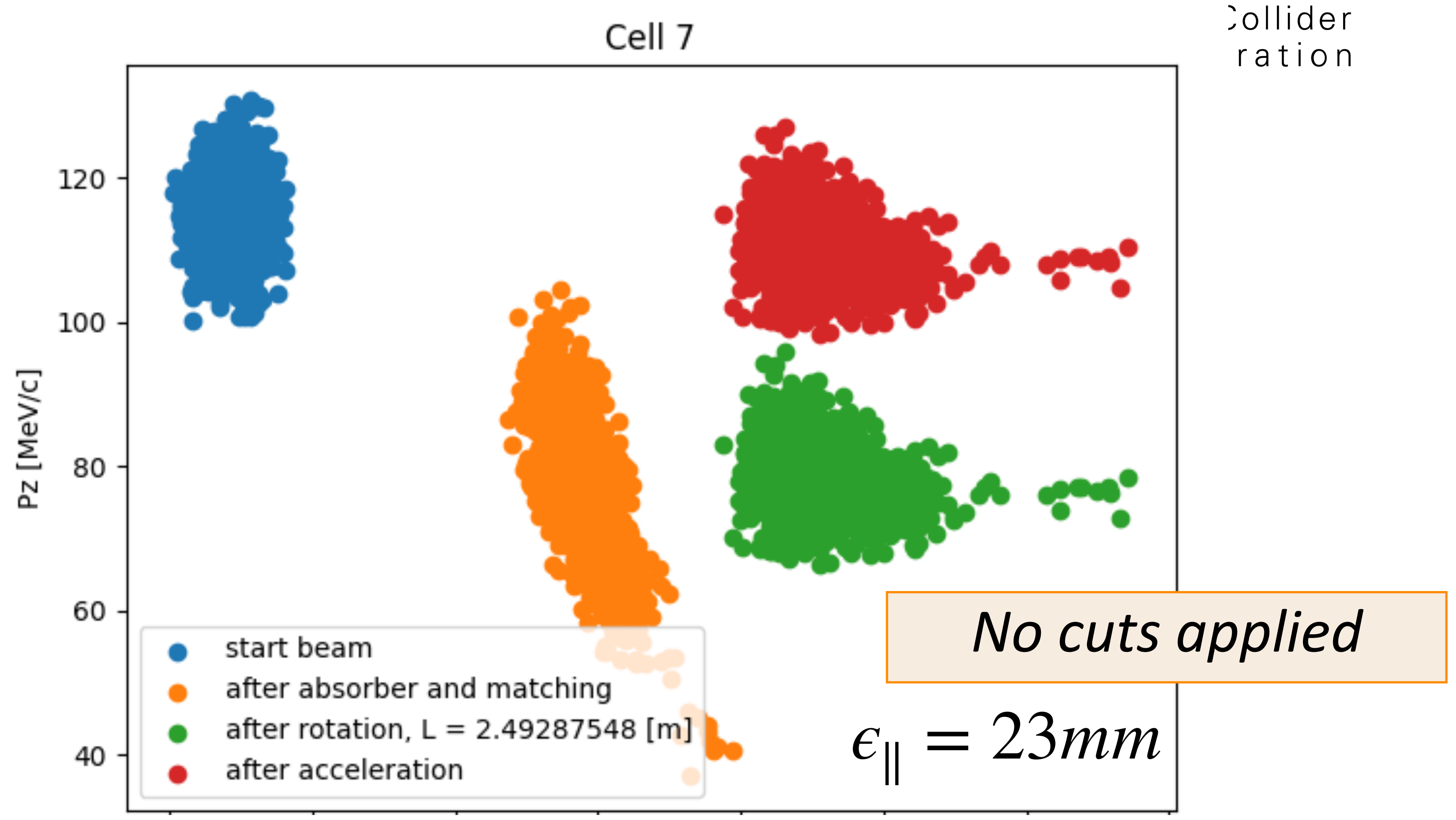
Helpful for complex models:
- what are most critical parameters to be optimised?
- Where are the bottle necks?

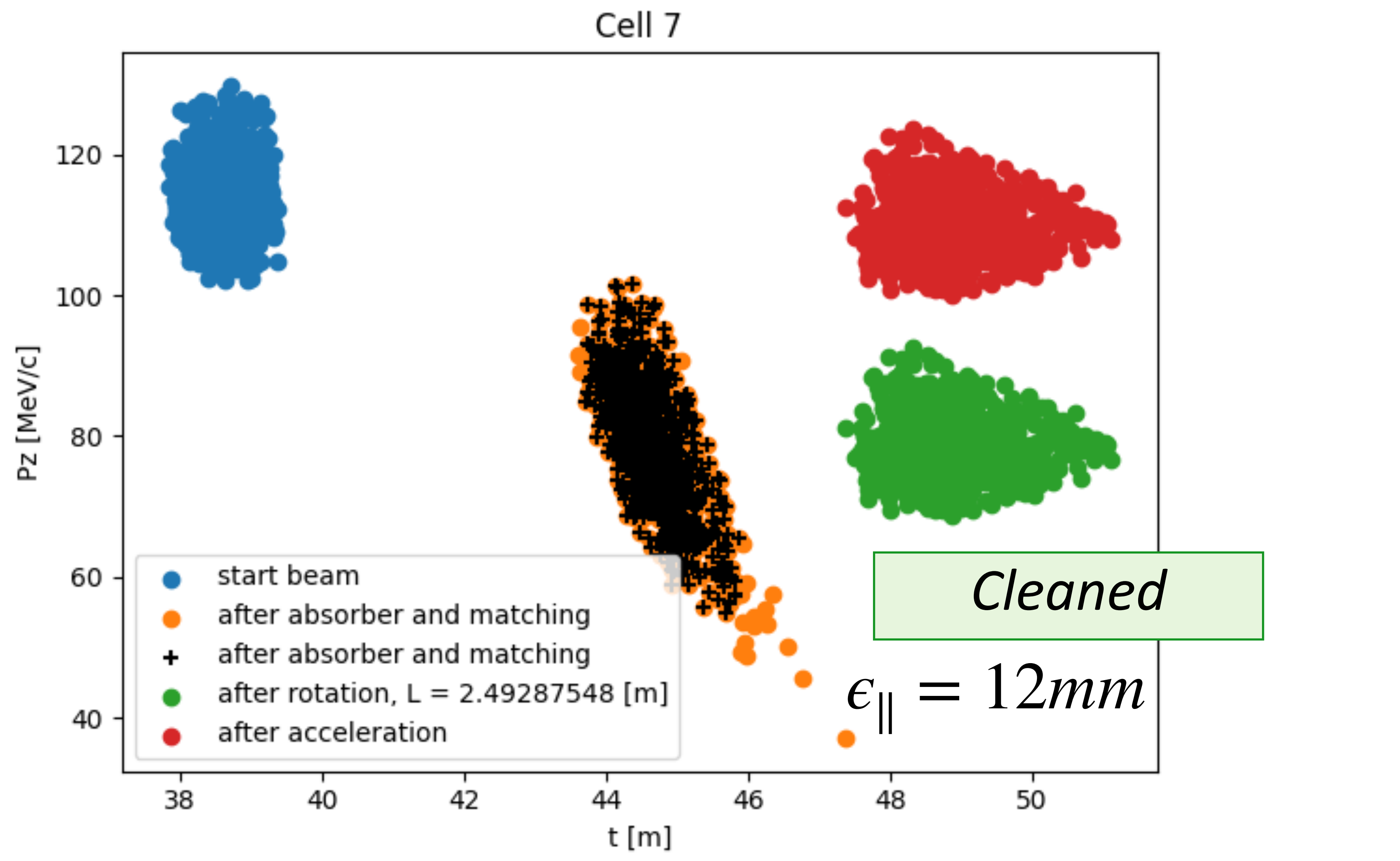
✓ "What is this model actually learning?"

Final cooling optimization: robust emittance estimation

Objective function : $\dfrac{\epsilon_{\perp}\epsilon_{||}}{N_{\mu}}$

- Too high emittance can be caused by a few “outliers”
- Traditional “3 sigma-cut” not always reliable, especially towards the end of the channel
- Robust algorithm to exclude the outliers before evaluating the emittances?

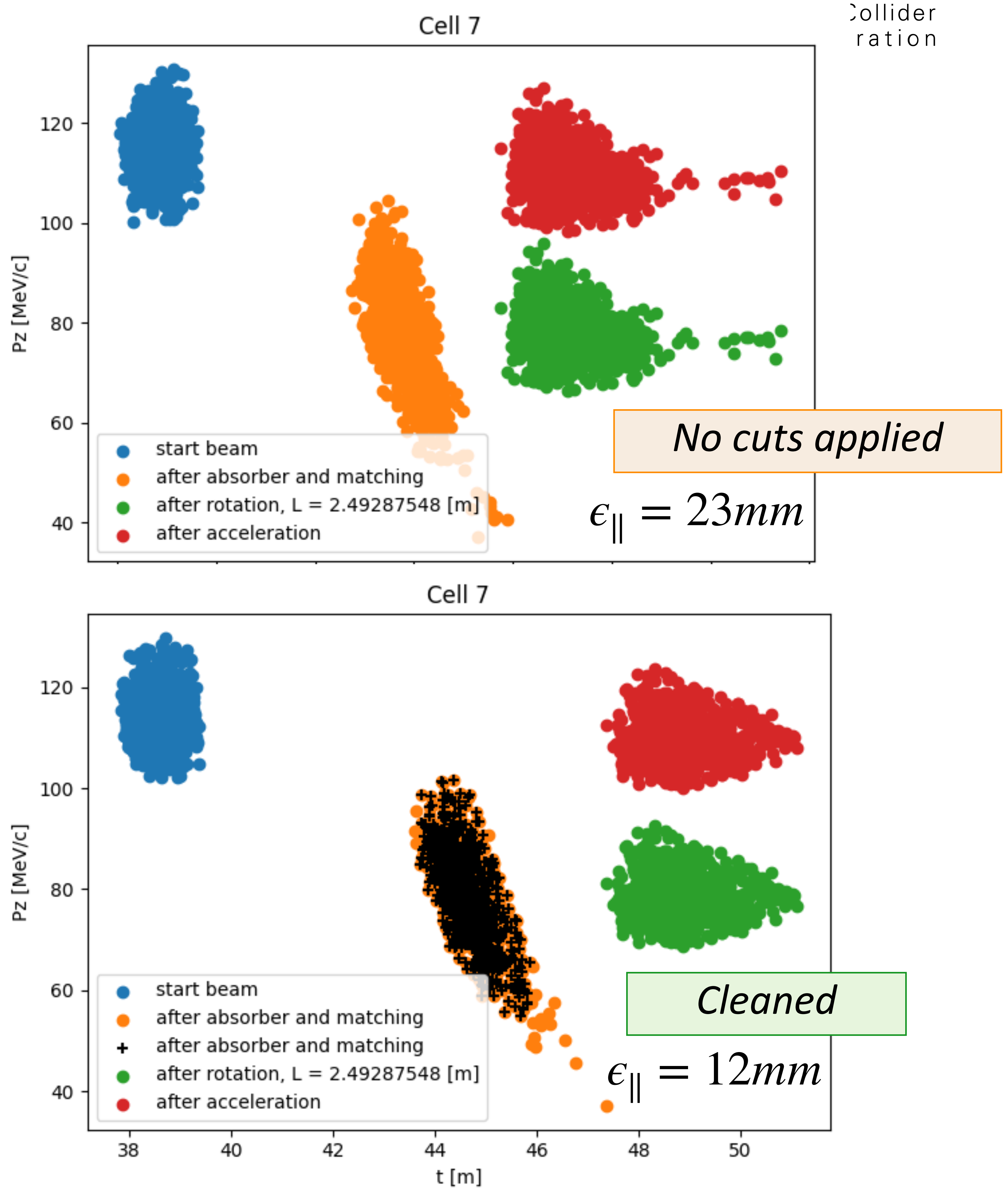




Final cooling optimization: robust emittance estimation

Objective function : $\dfrac{\epsilon_{\perp}\epsilon_{||}}{N_{\mu}}$

- Too high emittance can be caused by a few “outliers”
- Traditional “3 sigma-cut” not always reliable, especially towards the end of the channel
- Robust algorithm to exclude the outliers before evaluating the emittances?
- Comparing anomaly detection techniques, density-based clustering
- **Unsupervised Learning (no data, no training needed), fast-executable**



Suitable methods for robust emittance estimation

- **Local Outlier Factor**
  + **outliers** as points with significantly **lower density** compared to their neighbours
  + effective for **high-dimensional data sets**
  - requires threshold specification

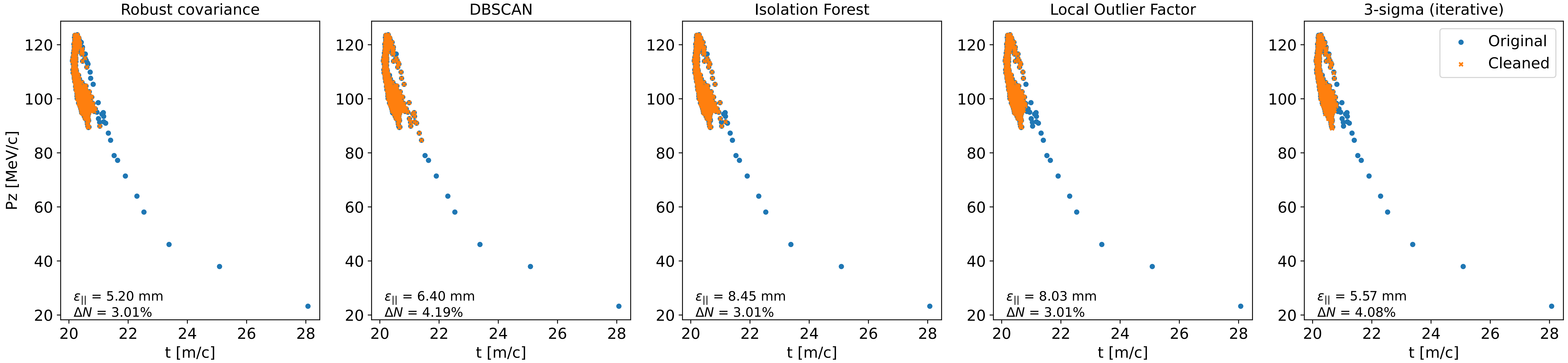
- **DBSCAN**
  + separates regions based on the **density**, can identify **noise**
  + can handle clusters of **arbitrary shape and size**
  - requires a threshold on minimum N samples and distance

- **Isolation Forest**
  + specifically designed for **outliers detection**
  + **robust**, based on ensembles of decision trees
  - requires expected outliers rate in the dataset

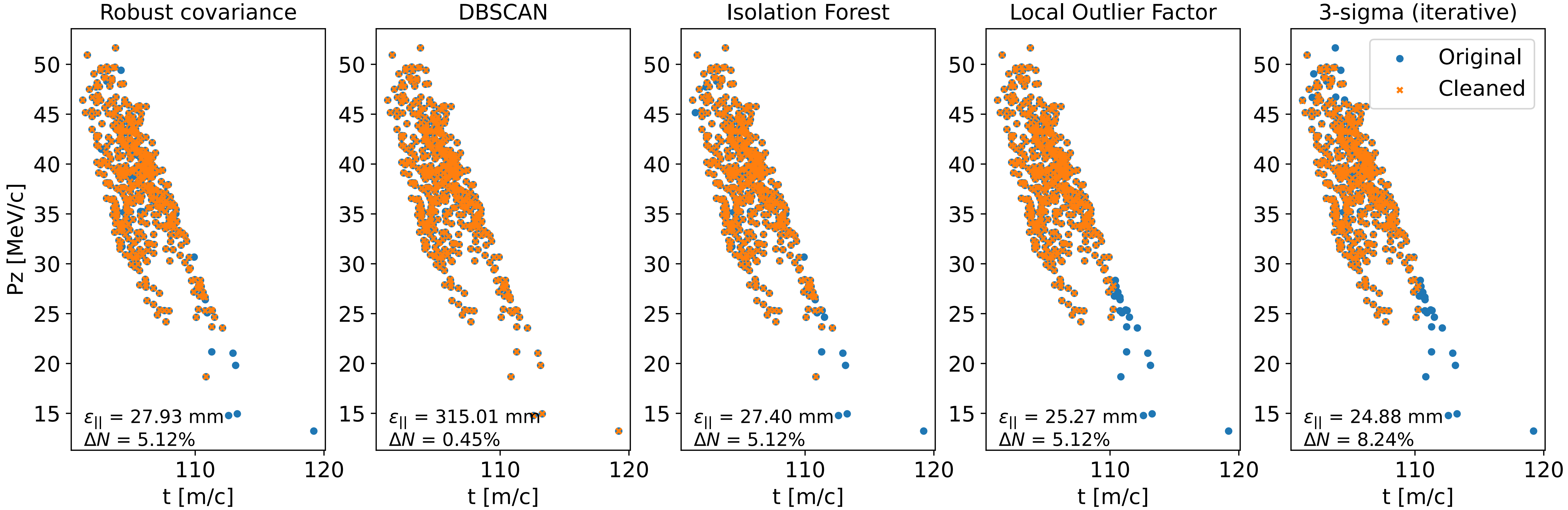
- **Minimum Covariance Determinant**
  + aims to find the **subset with the smallest determinant of the covariance matrix**
  => samples from the same distribution
  + no thresholds to be defined
  + direct output is "clean" 6D covariance matrix
  => rms properties of the particle distribution

Comparison of clustering techniques for emittance estimation

Example: cell 3



Example: cell 6



Cooling performance, full channel

Preliminary

	$\epsilon_{\perp}[\mu m]$	$\epsilon_{\|\|}$ [mm]	N [%]
3σ-cut	39	85	35
IF	33	82	33
$\boxed{\text{MCD}}$	35	80	38
No cuts	46	106	42

Start-to-end FC channel: current results

Cell	LH_2 [m]	Drift [m]	N_{RF} rot.	N_{RF} accel.	f_{RF} [MHz]	G [MV/m]	$\phi_{RF,rot.}$ degrees
2	0.466	0.3238	5	5	111.06	19.81	-180
3	0.46958	1.363	10	7	56.85	14.17	90
4	0.4	2.5	9	8	40.13	11.9	51
5	0.3	1.8358	7	2	34.91	11.11	-10
6	0.25	2.0	5	10	30.61	10.4	-54
7	0.3	0.984	5	14	11.637	6.823	-82
8	0.1	3.6464	2	7	16.17	8.04	67
9	0.17	3.64	2	11	13.38	7.32	67
10	0.08	2.555	11	2	8.226	5.39	-6
11	0.0541	2.895	11	4	5.676	4.48	-96

$P_{z,start}$ [MeV/m]	σE_{start} [MeV]	σt_{start} [mm]	$P_{z,end}$ [MeV/m]	σE_{end} [MeV]	σt_{end} [mm]	$\epsilon_{\|\|}$ [mm]	$\epsilon_{\perp}$ [μm]	N [%]
145.0	3.2	50.0	100.0	4.3	125.2	2.1	221.2	95.5
118.8	2.0	201.9	89.0	2.4	130.7	2.9	177.2	87.3
118.9	2.7	192.8	89.2	3.0	268.2	4.0	151.0	81.5
114.5	2.5	399.7	87.4	3.4	173.7	5.0	137.2	71.9
92.9	2.9	209.5	62.0	4.4	592.6	9.2	109.9	65.6
84.9	3.7	1625.8	57.4	1.6	911.6	13.1	93.2	56.3
89.8	1.5	916.6	55.2	2.7	926.9	22.3	69.3	52.5
71.8	2.4	1354.7	57.7	2.9	1365.3	28.0	63.8	48.2
77.2	2.2	1774.2	53.5	3.1	1695.2	40.5	51.0	43.2
61.5	1.8	2561.5	43.5	2.8	2398.1	59.3	42.2	39.0
60.5	2.2	3101.1	49.2	2.8	2954.3	77.3	37.2	36.4

- Already after 9 cells better performance is achieved compared to the baseline:

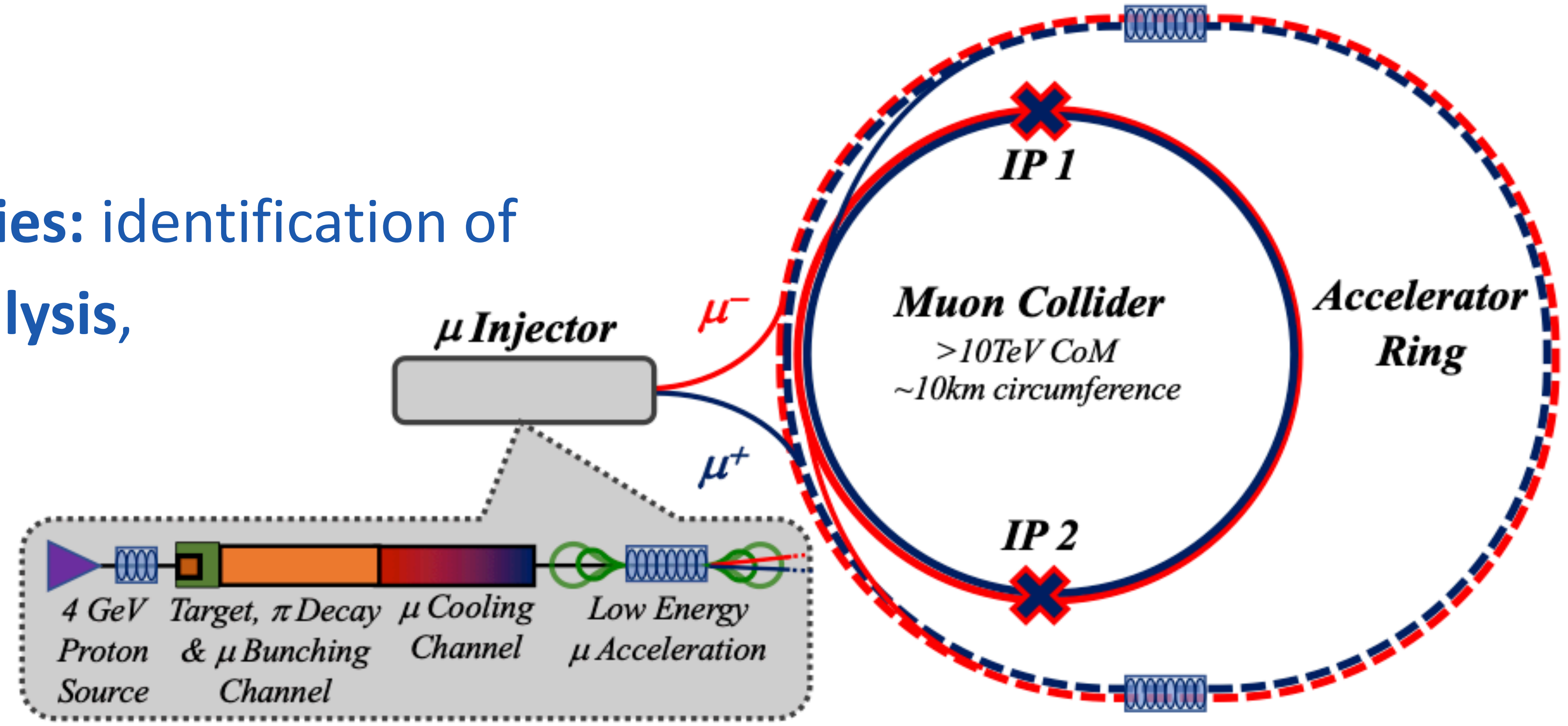
  (9 cells, $\epsilon_\perp = 40\mu m, \epsilon_\perp = 51mm$)

  (16 cells,$\epsilon_\perp = 55\mu m, \epsilon_\perp = 76mm$)
- Potential to improve the transmission by minimising the relative energy spread
- Potential to combine with other cooling techniques

Summary

Muon Collider Design (Final Cooling Channel)

- **Surrogate models** for both, fast objective function evaluation and estimation of initial parameter
- **Bayesian Optimization** combining modelling and optimization
- **Anomaly detection techniques** for robust emittance analysis
- "Proof-of-concept": **Opening several opportunities for accelerator design studies:** identification of most critical parameters for collider performance (e.g. **feature importance analysis**, but also dimensionality reduction techniques)
- Start-to-end **optimisation framework** utilizing **fast-executable methods** for changing requirements as design evolves.



https://muoncollider.web.cern.ch

Practical Advice

- Start with **simpler models** - they are **easier to tune and interpret**.
  Neural Networks are not always the perfect solution!
- Numerical Optimisers are powerful tools and can be made even more efficient using **surrogate models** - **save and structure your data!**
- Not all ML algorithms need large amount of data - consider translating your problem as **Unsupervised Learning** task (e.g. anomaly detection)

Thanks a lot for your attention!