AI for Muon Collider Design: progress and plans

Elena Fol

CERN

3rd ICFA Beam Dynamics Mini-Workshop
on Machine Learning Applications for Particle Accelerators
November 1-4, 2022

Outline

- Muon Collider overview
- Final Cooling: baseline and challenges
- Simulation tools and automatic optimization
- Supervised Learning for Final Cooling optimization:
  - Optimization speed-up
  - Finding approximated solution
  - Identification of most relevant design parameters
  - Classification of initial conditions
- Further potential ML applications

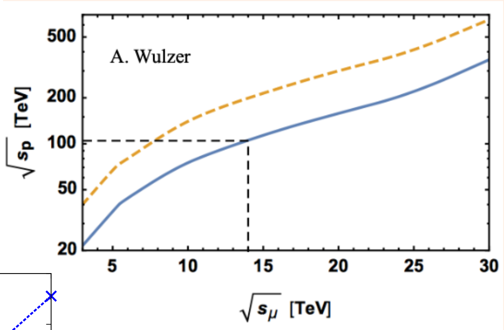
I. Muon Collider Overview

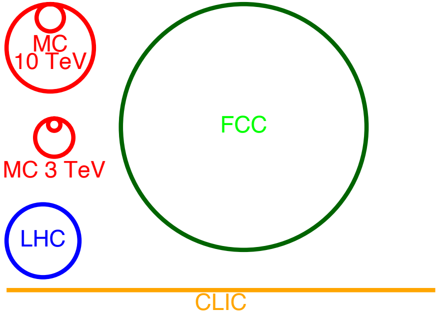
Why colliding muons?

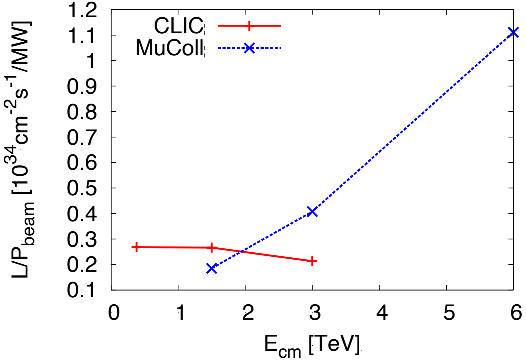
Physics potential:

14 TeV lepton collisions are comparable to 100 TeV proton collisions for production of heavy particle pairs

$$m_\mu = 200 \times m_e$$







D. Schulte

- **Luminosity per power** increases with energy: unique opportunity provided technology for the MC is available
- **Compact** => Expected to be **cost effective** and reducing power consumption in comparison to other options.

Muon Collider Collaboration

Muon collider is included into European Accelerator R&D Roadmap

✓ EU Design Study proposal has been successfully accepted in 2022

- ➡ provide a baseline concept of a muon collider
- ➡ Estimate performance and associated key challenges, cost and power consumption drivers
- ➡ Identify R&D path to demonstrate the feasibility

- Previous studies in the US and now: **US Snowmass recommended muon collider R&D** be considered a **high priority**
- Experimental programme in UK, alternatives studies by INFN
- **3 TeV collider** as the first exploration stage possible with technology accessible in **10-20 years**

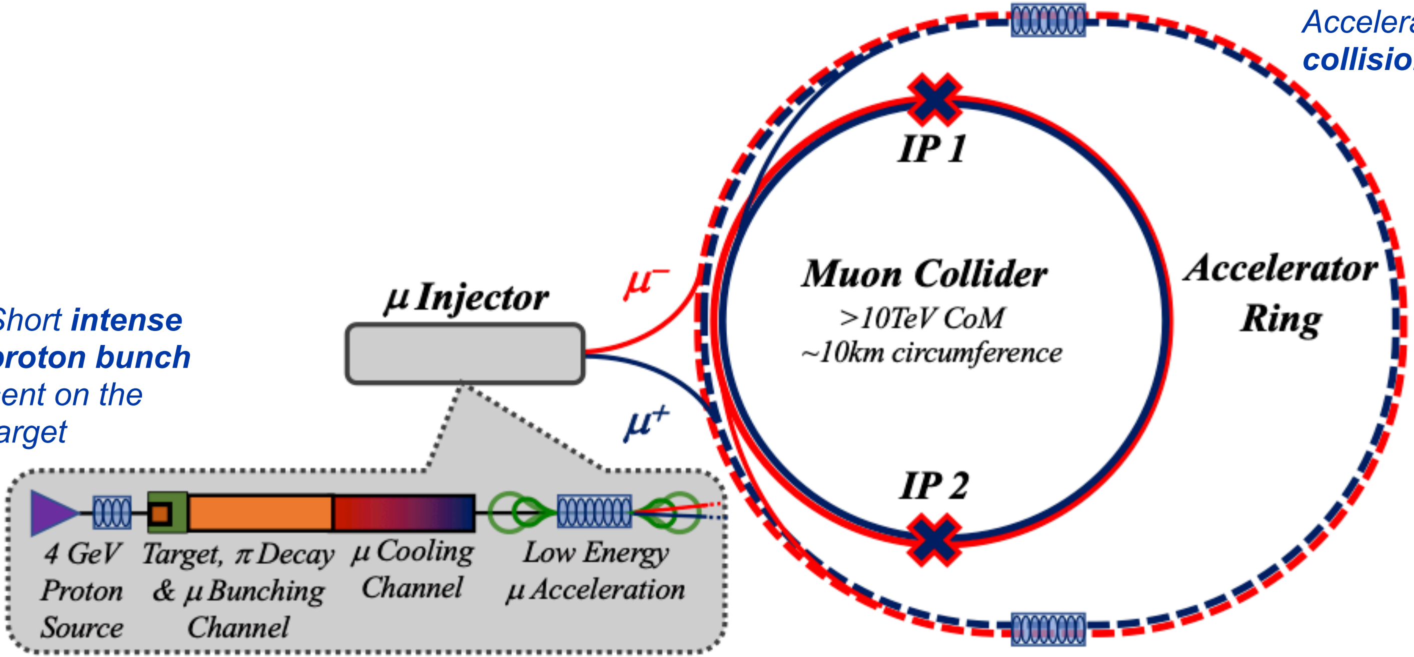
➡ unique promising option to reach **highest lepton energies with high luminosity**
➡ Roadmap process found muon collider challenging but **did not identify any showstopper**

- Collaboration across large number on universities and laboratories
  - Web site: https://muoncollider.web.cern.ch
- First Muon Collider Collaboration Meeting (October 11-14, 2022)

Muon Collider Overview

Acceleration to **high collision energy**

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

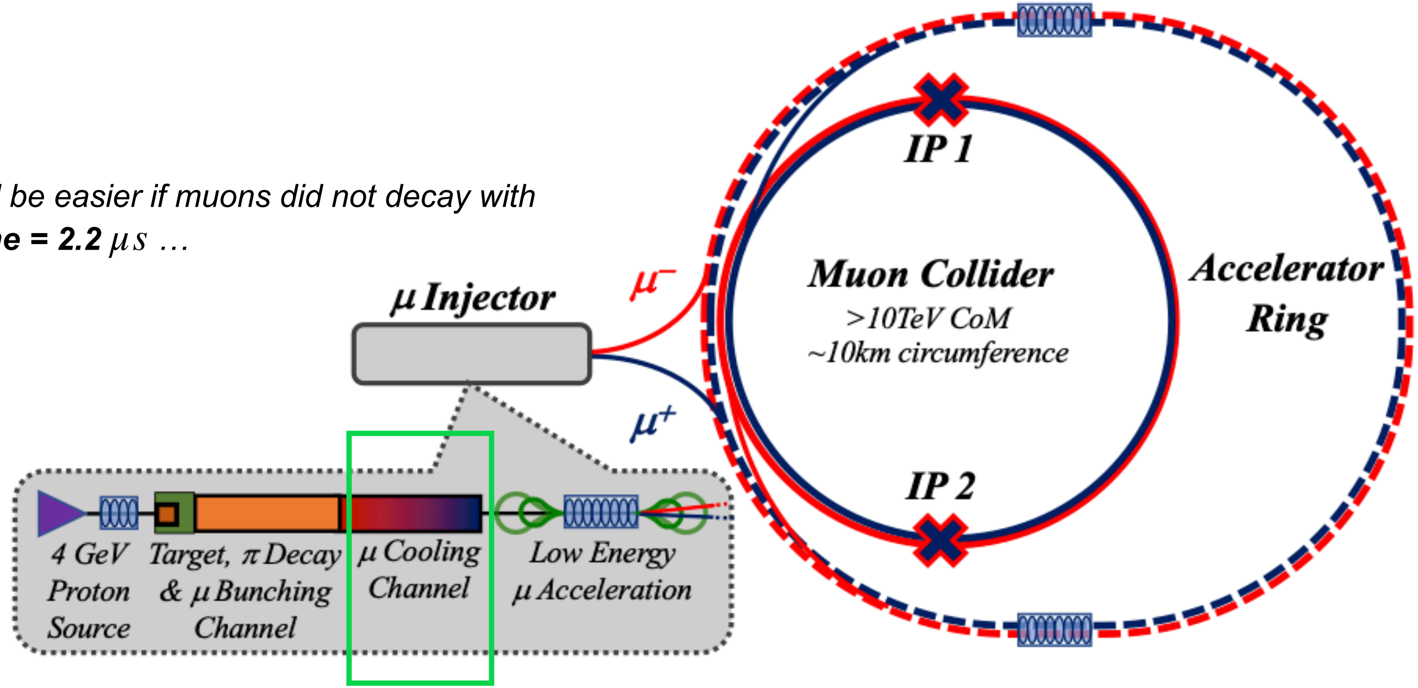


Interaction with the target produces pions
➡ **decay into muons**

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

Muon Collider Overview

Would be easier if muons did not decay with
***lifetime = 2.2** μs …*



Muons are created as pions decay products and form a beam with a **huge emittance**

- **Cooling** (the reduction of occupied phase-space by muons) is required

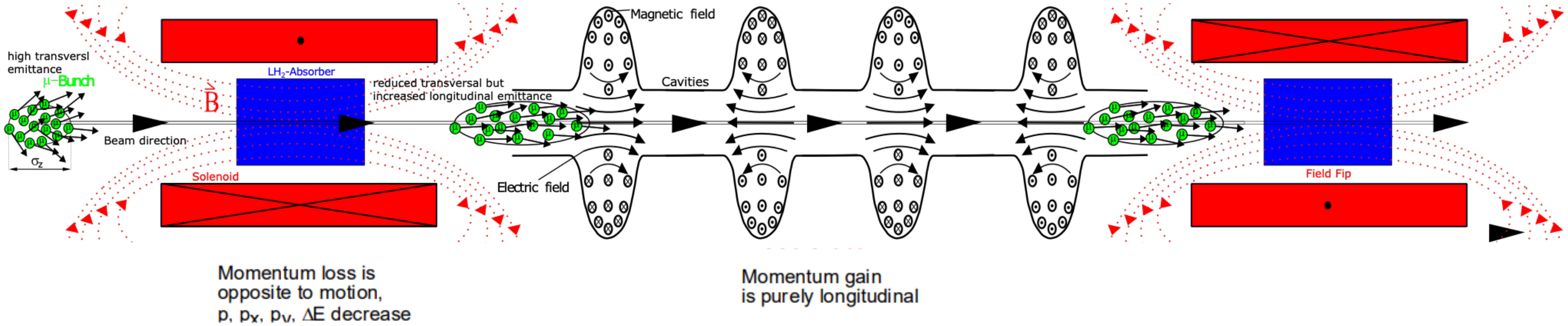
Traditional cooling techniques are not suitable due to **muons lifetime**

- **Ionisation cooling**: fast novel technique, principle is demonstrated by MICE collaboration

Technology and challenges of Final Cooling

- Energy loss due to the interaction with absorber material
- Reduction of transverse beam emittance
- Re-accelerating the beam to restore the longitudinal momentum

B. Stechauner



Lowering transverse emittance on the costs of :

➡ Longitudinal emittance growth

➡ Bunch length increasing: challenging RF set-up

➡ Energy spread (needs to be kept within the accelerator acceptance)

➡ Number of survived particles (length of the channel vs. muon lifetime)

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

Energy loss term

Multiple scattering term

II. Final Cooling Channel: optimization and surrogate models

Extending simulation framework

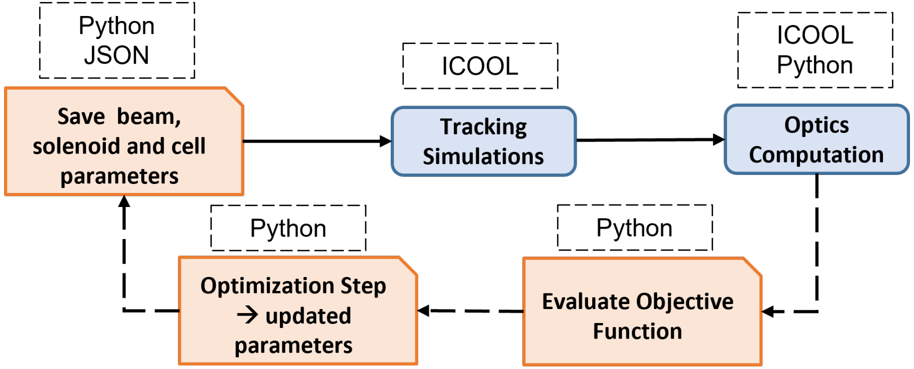
International MUON Collider Collaboration

Simulation tools for ionisation cooling

- ICOOL: developed for 3D tracking of particles in ionisation cooling channels

  Problems:

- Modification of text file-like input decks
- Evaluation of tracking results

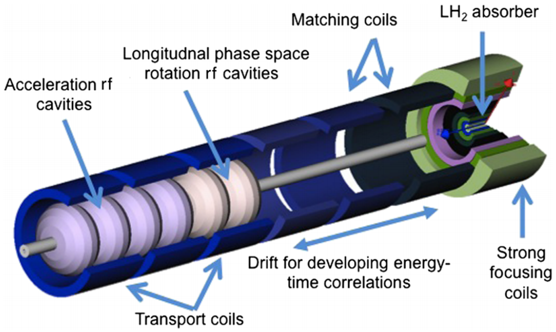


- Python "wrapper" for launching ICOOL
- Automatic computation of initial beam distribution, **generation of ICOOL code**
  - ✓ Additional analysis in Python
  - ✓ **Storing input and output of simulation** in a structured format (JSON)

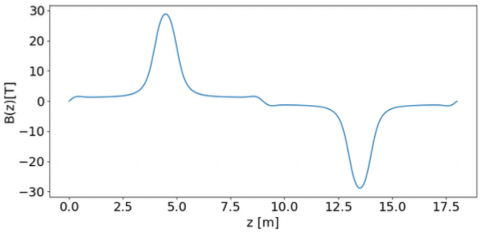
✓ Simplified optimization set-up
✓ Easily extendable
✓ Easy integration of optimization methods
✓ **Enables to use the simulations as training dataset**

First attempt: simplified lattice, optics matching

- **Efficient transverse emittance reduction:** beam satisfying **optical constraints** $\alpha = 0$, $\beta = 2p/qB$
- Objective function: **minimize** $\bar{\alpha} + |\beta_{ideal} - \beta_{sim}|$
- Free parameters: radii of solenoid coils, maximum field



Simplified lattice: 2 cooling cells, peak B(z) = 30 T

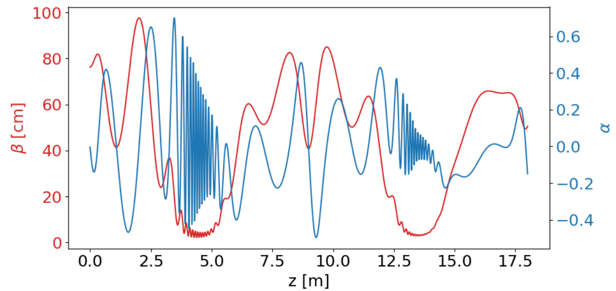


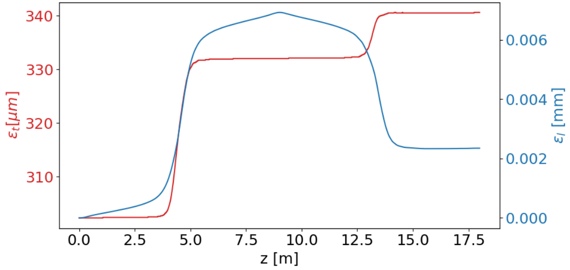
Applied optimizations methods:

- Nelder-Mead
- Differential Evolution: stochastic population-based method, allows parallelization
- Extremum Seeking:
  A. Scheinker and D. Scheinker, “Constrained extremum seeking stabilization of systems not affine in control,” International Journal of Robust and Nonlinear Control 28, 568–581 (2018)

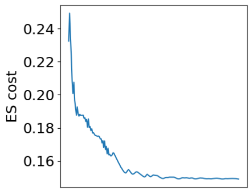
First attempt: simplified lattice, optics matching

✦ Objective function: minimize $\bar{\alpha} + |\beta_{ideal} - \beta_{sim}|$, with $\beta = 2p/qB$

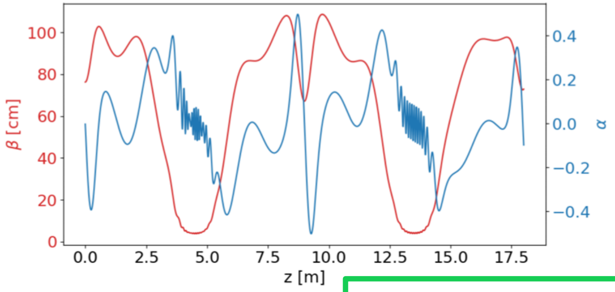


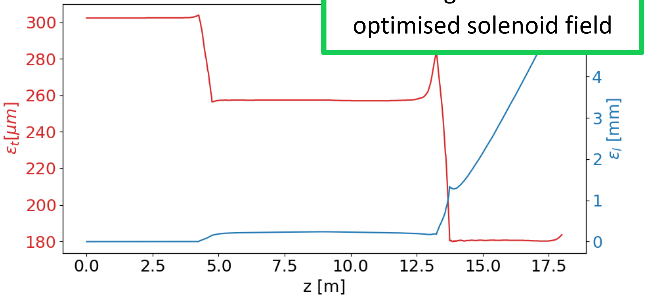


Optimizing magnet field using ES



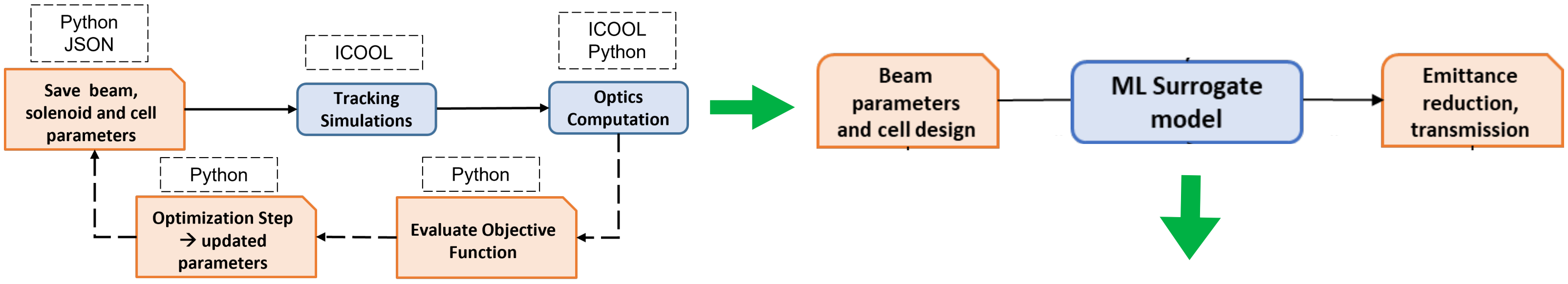
- ✓ Eliminating emittance blow up due to unmatched optics.
- ✓ Transverse cooling using liquid hydrogen absorber in the centre of optimised solenoid field



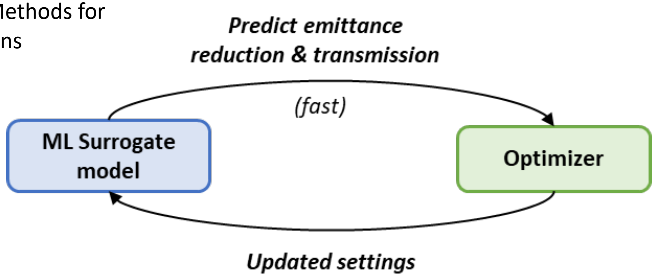


Placing absorber in optimised solenoid field

Training models to predict simulation output



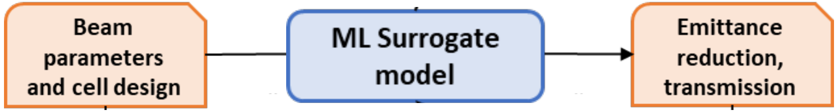
- A. Edelen et al. „Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems" ,Phys. Rev. Accel. Beams 23, 044601, 2020
- E. Fol „Evaluation of Machine Learning Methods for LHC Optics Measurements and Corrections Software", CERN-THESIS-2017-336, 2017



<u>Strategy for optimization speed-up:</u>

1. Train a surrogate model
   - **Using input-output pairs collected from tracking simulations during ES- optimization**
2. Continue optimization, but skip tracking in evaluation step: replace with ML model prediction

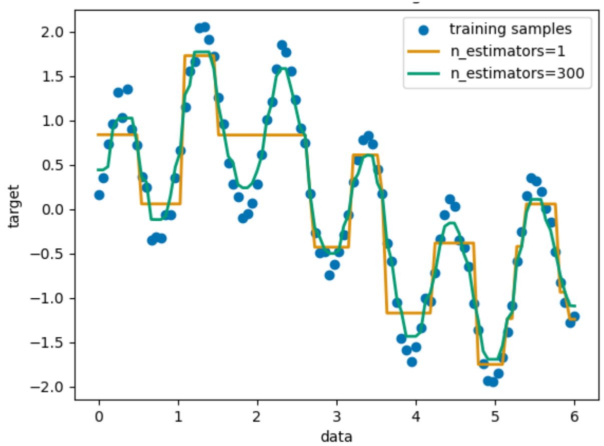
Training models to predict simulation output



<u>Decision Trees:</u>

- Partition data based on a sequence of thresholds
- Continuous target y, in region estimate:
- Mean Square Error

$$H(X_m) = \frac{1}{N_m} \sum_{i \in N_m} (y_i - c_m)^2$$



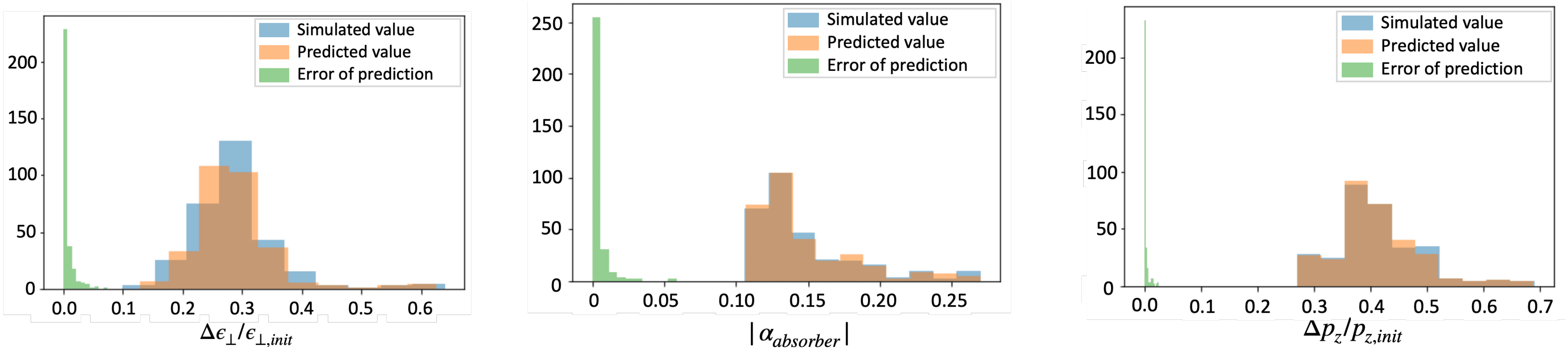
<u>Random Forest:</u>

- Random subset of examples, train separate model on each subset
- Only random subset of features is used at each split
- Increases variance, tend not to overfit

Training models to predict simulation output

- ✓ Random Forest regressor, 1200 simulations
- ✓ **98.3%** accuracy on a test set (300 simulations)

Predicting beam properties included in objective function:



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

II. Inverse Models towards complete Final Cooling design

General Idea

- Muon Collider design: increase luminosity and overall efficiency
  => Final Cooling needs to be **optimised as integrated part of the entire complex**
  => Continuously integration of **changing requirements** and constraints
  => **flexible optimization strategy** is needed

<u>Possible strategies</u>

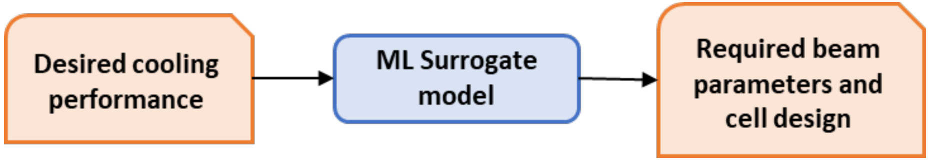
- Push transverse emittance minimization
  - assuming fixed initial beam parameters coming from previous muon production stage
- **"Backwards" optimization**
  - starting from the downstream requirements (final emittance, beam energy, etc.)

General Idea

- Muon Collider design: increase luminosity and overall efficiency
  => Final Cooling needs to be **optimised as integrated part of the entire complex**
  => Continuously integration of **changing requirements** and constraints
  => **flexible optimization strategy** is needed

<u>Possible strategies</u>

- Push transverse emittance minimization
  - assuming fixed initial beam parameters coming from previous muon production stage
- **"Backwards" optimization**
  - starting from the downstream requirements (final emittance, beam energy, etc.)



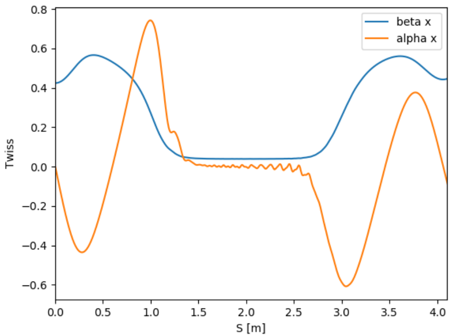
- **Combined dataset to train a simplified model including parameters of interest**
- **Starting point providing fast estimates**

Full cooling cell optimization: absorbers and RF

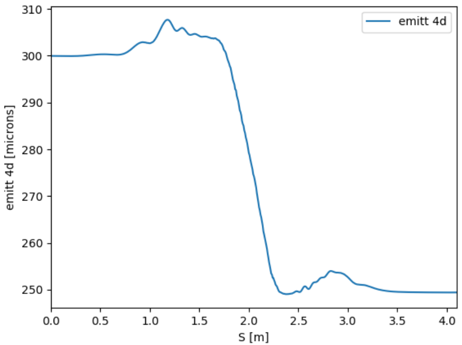
RF-Track (developed by A. Latina):

- User interfaces in **Python and Octave**
- Possibility to include collective effects
- Parallelisation, **fast executable**
- Used for tuning of full cell structure, including RF
- Available at: https://gitlab.cern.ch/rf-track/download

I. Linear optics matching



II. Minimize transverse emittance using absorber



III. **Optimise towards target values** for transverse and longitudinal emittance, **include re-acceleration**

- Energy loss and re-acceleration: What are **optimal beam energies** at the end of each cell?
- Trade-off between transverse cooling and longitudinal emittance increase?

Global optimization of cooling performance

Proof of concept: optimising cells "backwards" starting from final target values

- First "stage": 4 cooling cells
- Free parameter: $P_z, \epsilon_\perp, \epsilon_\parallel, \sigma_z$ at the start of the channel, absorber length, drift length, number of RF cavities, RF frequency, voltage, phase in each cell (24 in total)
- Target parameters of cell n = initial parameters of cell n+1: $P_z, \epsilon_\perp, \epsilon_\parallel, \sigma_z$

- Desired beam parameters: transverse emittance: 230 mm mrad, longitudinal emittance < 10 mm, pz = 110 MeV/c (based on previous design) "High field – low energy muon ionization cooling channel", H. Sayed, Robert B. Palmer, D. Neuffer Phys. Rev. ST Accel. Beams **18**, 091001 – Published 4 September 2015

Global optimization of cooling performance

International MUON Collider Collaboration

Proof of concept: optimising cells "backwards" starting from final target values

- First "stage": 4 cooling cells
- Free parameter: $P_z, \epsilon_\perp, \epsilon_{||}, \sigma_z$ at the start of the channel,
  absorber length, drift length, number of RF cavities, RF frequency, voltage, phase in each cell (24 in total)
- Target parameters of cell n = initial parameters of cell n+1: $P_z, \epsilon_\perp, \epsilon_{||}, \sigma_z$

➢ Desired beam parameters: transverse emittance: 230 mm mrad, longitudinal emittance < 10 mm, pz = 110 MeV/c (based on previous design) "High field – low energy muon ionization cooling channel", H. Sayed, Robert B. Palmer, D. Neuffer Phys. Rev. ST Accel. Beams **18**, 091001 – Published 4 September 2015

Limitations of traditional tracking

- Optimization propagating requirement from the end of the channel, individually for every cell:
  => Tracking: generates a new gaussian beam for every cell, however correlations are expected to develop throughout the channel!
- How to obtain (nearly) optimal parameter without tracking simulations?

ML - assisted Optimization

$P_z, \epsilon_\perp, \epsilon_\parallel, \sigma_z$ after final cell

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

$P_z, \epsilon_\perp, \epsilon_\parallel, \sigma_z$ at the start of each cell + cell parameter

1. Train a model on simulations (saved from optimization runs or simple parameter scans)
2. Predict parameters for several consecutive cells starting from final target values.
3. **Tracking using initial beam** predicted for the 1st cell and **parameters of each cell.**

Result is as required?

Fast design estimate

Further optimization needed?

Use as initial guess for optimisation algorithm
-> optimal solution is found within fewer steps

First results: beam evolution in optimised cooling cells

- Target: transverse emittance: 230 mm mrad, longitudinal emittance < 10 mm, pz = 110 MeV/c

Differential Evolution Algorithm

Cell	Beam parameters (end of the cell)				
	Emittance Tr. [mm mrad]	Emittance Long. [mm]	Bunch length	Pz [MeV/c]	Pz spread
	300.0	1.5	50.0	135	3.5
1	295	1.8	72	123	3.6
2	285	2.4	92	112	3.7
3	274	9.3	260	104	6.5
4	**260**	**16.5**	**715**	**93**	**7.1**

Beam Tracking with predicted parameters from Random Forest model

Cell	Beam parameters (end of the cell)				
	Emittance Tr. [mm mrad]	Emittance Long. [mm]	Bunch length	Pz [MeV/c]	Pz spread
	300.0	1.5	50.0	135	3.5
1	295	1.7	79	125	3.6
2	283	2.2	61	118	4.6
3	270	2.3	128	105	2.4
4	**255**	**4.8**	**210**	**95**	**4.1**

First results: beam evolution in optimised cooling cells

International MUON Collider Collaboration

- Target: transverse emittance: 230 mm mrad, longitudinal emittance < 10 mm, pz = 110 MeV/c

Differential Evolution Algorithm

Cell	Beam parameters (end of the cell)				
	Emittance Tr. [mm mrad]	Emittance Long. [mm]	Bunch length	Pz [MeV/c]	Pz spread
	300.0	1.5	50.0	135	3.5
1	295	1.8	72	123	3.6
2	285	2.4	92	112	3.7
3	274	9.3	260	104	6.5
4	**260**	**16.5**	**715**	**93**	**7.1**

***Note**: numbers are preliminary, ionisation cooling in RF-Track is still under development*

Beam Tracking with predicted parameters from Random Forest model

Cell	Beam parameters (end of the cell)				
	Emittance Tr. [mm mrad]	Emittance Long. [mm]	Bunch length	Pz [MeV/c]	Pz spread
	300.0	1.5	50.0	135	3.5
1	295	1.7	79	125	3.6
2	283	2.2	61	118	4.6
3	270	2.3	128	105	2.4
4	**255**	**4.8**	**210**	**95**	**4.1**

***Note**: the model achieved only 80% accuracy on a test set, improvements are still possible*

✓ Better trade-off between longitudinal and transverse emittance
✓ Demonstrated proposed optimization strategy
✓ Flexible automatic optimization framework

III. Identifying most relevant parameters

Feature importance analysis

☑ Comes for free when building surrogate models using Random Forest algorithm

- Prediction **loss with/without permutation** of each variable:

- Decrease in the model score is indicative of how much the model depends on the feature:
  how important this feature is for a particular model?
  https://scikit-learn.org/stable/modules/permutation_importance

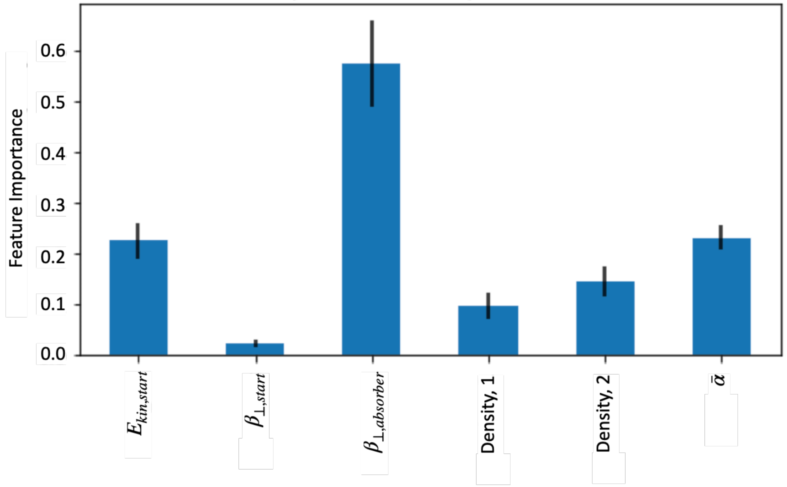
Important considerations:

- First, make sure that model's scores are sufficiently high (e.g. through cross-validation):
  features demonstrating l**ow importance for a bad model could be very important for a good model**.
- **Correlated features**: one of the features is permuted
  - ◉ the model still has access to the feature through its correlated feature
  - ◉ **lower importance** for both features, where they **might actually be important**.

Does the model “understand” the physics behind training data?

International
UON Collider
Collaboration

Predicting Transverse emittance reduction



<u>Simple model</u>

- **energy loss in absorber and optics matching**
- 2 cooling cells, only B-field and absorbers
- Varying initial beam energy, solenoid coils and absorber density

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

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

… obvious to an (experienced) physicist
—> Big achievement for a decision tree
✓ “what is this model actually learning?”

Further potential ML applications in Muon Collider Design

Sample-efficient optimization: limiting initial conditions

- Optimization of last cooling cells becomes more challenging:
  - lower energies, longer bunches, more energy spread
- Certain combinations of initial beam conditions and cell parameters can lead to tracking failure
  - Classify a few simulation set-ups based on tracking results
  - Find a **stable conditions boundary**
  - Run optimization exploring parameter space within this boundary

Sample-efficient optimization: limiting initial conditions

- Optimization of last cooling cells becomes more challenging:
  - lower energies, longer bunches, more energy spread
- Certain combinations of initial beam conditions and cell parameters can lead to tracking failure
  - Classify a few simulation set-ups based on tracking results
  - Find a **stable conditions boundary**
  - Run optimization exploring parameter space within this boundary

Example of applying this ML-approach to DA optimization in HL-LHC:

Sample-efficient optimization: limiting initial conditions

- Optimization of last cooling cells becomes more challenging:
  - lower energies, longer bunches, more energy spread
- Certain combinations of initial beam conditions and cell parameters can lead to tracking failure
  - Classify a few simulation set-ups based on tracking results
  - Find a **stable conditions boundary**
  - Run optimization exploring parameter space within this boundary

Example of applying this ML-approach to DA optimization in HL-LHC:

Dynamics Aperture (DA) needs to be estimated in **numerical simulations:**
- Excludes disconnected stable islands from the calculation of the volume
- Very **computationally expensive** to sample the phase-space with 6D scans

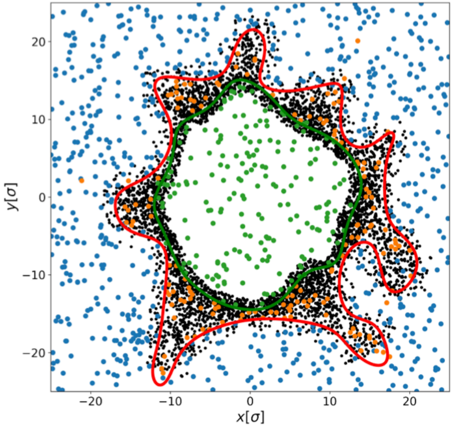
Solution: **estimate the border of stable phase-space region with supervised learning**

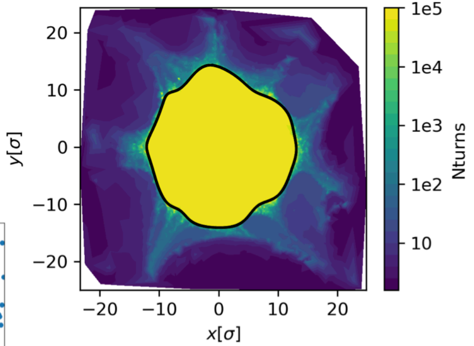
Sample-efficient optimization: limiting initial conditions

<u>ML-approach to Dynamic Aperture optimization in HL-LHC:</u>

- Support-Vector machine algorithm, binary classification
- Linear classifier, learning non-linear decision function using kernel-transformation (here: Radial Basis function)
- Training data: label particles into survived N turns/ not survived
- Learn the decision boundary => adaptively sample the phase-space

➢ **Fewer simulations to find valid phase space parameters**





F.F. Van der Veken, et al., "Determination of the Phase-Space stability border with ML", IPAC'22

Beyond final cooling optimization

- **Ionisation cooling simulation:**
  - Limited fidelity due to the lack of experimental data
  - MICE experiment: comparison between experimental data and Geant4 simulations
    ➡ *Combining simulations and experimental data to build high-fidelity models?*

- **Integrated model of muon collider complex:**
  - Optimization routines is a typical instrument across different collider sub-systems
  - Systematically saving the data
    ➡ *Collecting data from otherwise non-compatible simulations tools*

✓ **Opens several opportunities:** identification of most critical parameters for collider performance (e.g. feature importance analysis, but also dimensionality reduction techniques)

✓ **Fast-executable model** for changing requirements as design evolves

Conclusions

Conclusions

- Early stages of design study: a lot of progress has been made
  (First Muon Collider Collaboration Meeting (October 11-14, 2022))
  - rapid changing requirements
  - integration of different system components

- How can ML help?
  - Focus on simulation studies: speed-up, sample-efficiency
  - Numerical data-driven models: fast estimates adapting to new requirements

- First steps of integrating ML tools into final cooling design
  - Possibility to speed-up optimization
  - Benefit from data collected during optimization: Surrogate models providing fast estimates

- Optimization frameworks progressing together with the design study: new ideas are welcome!

International
MUON Collider
Collaboration

Thanks a lot for your attention!

Muon Cooling: required emittances

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

High energy

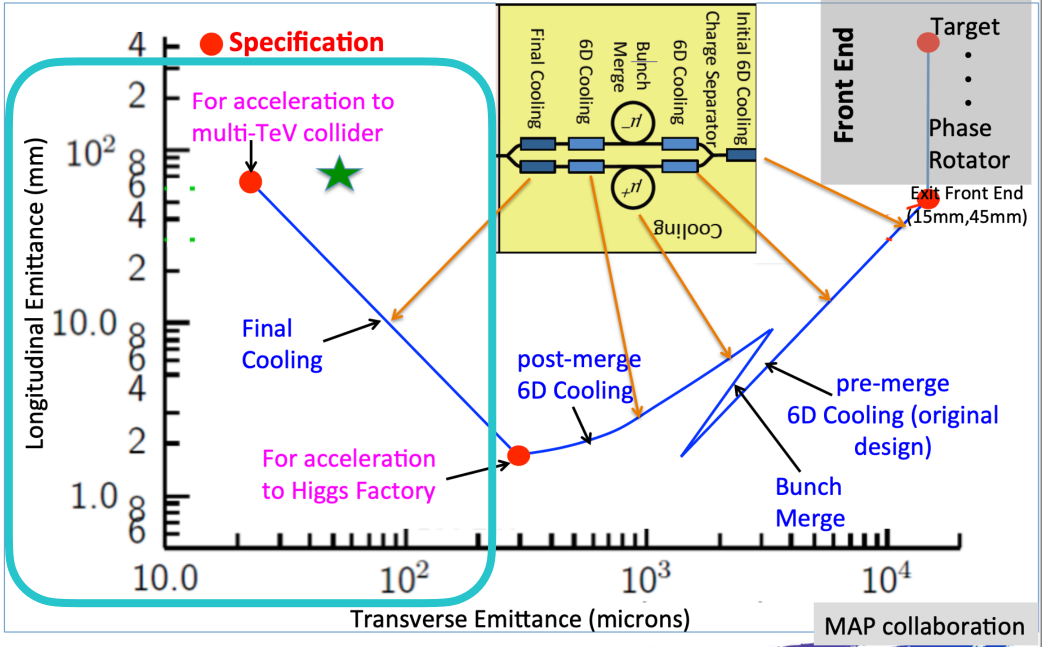
High field in collider ring

Large energy acceptance

Dense beam

High beam power

- Constant current for required luminosity scaling
- **Emittance preservation**
- Advance lattice design
- High field magnets



<u>Current focus of the study:</u>

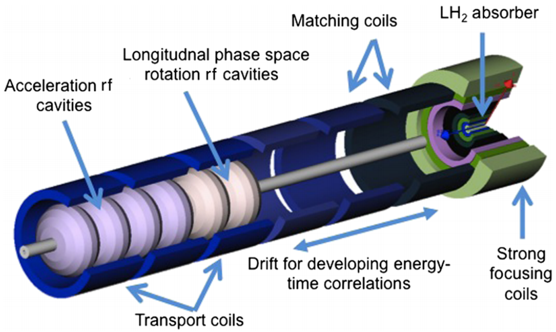
Design of muon collider to satisfy the target performance

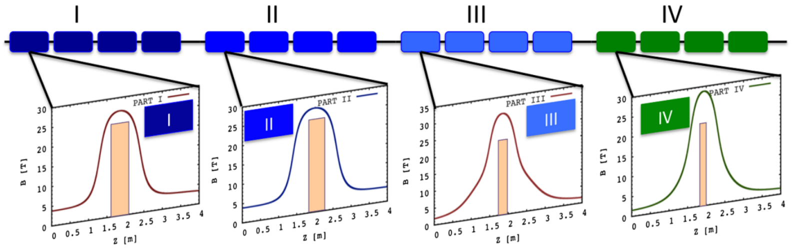
=> Final Cooling: Integration and optimisation of overall cooling design

Final Cooling baseline

- A Gaussian beam with $\varepsilon_{\perp}$=300 µm and $\varepsilon_{\parallel}$ = 1.5mm
- Beam momentum is reduced initially to 135 MeV/c
- High-field magnets 25—32 T, beam momenta ranged from 135 MeV/c to 70 MeV/c

- **Achieved in previous studies: $\varepsilon_{\perp}$ = 55 µm, with $\varepsilon_{\parallel}$ = 70 mm, transmission of 50%**
- **Preferred $\varepsilon_{\perp}$ = 25µm**
- should be possible to achieve with **stronger focusing fields, alternative absorber configuration**
- **Challenge:** trade off optimising longitudinal momentum, emittance and energy spread for efficient transverse emittance reduction





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