Some (possible) ML applications for low-emittance storage rings

Ilya Agapov

2nd ICFA Workshop on ML for Particle Accelerators
PSI, 27 Feb 2019

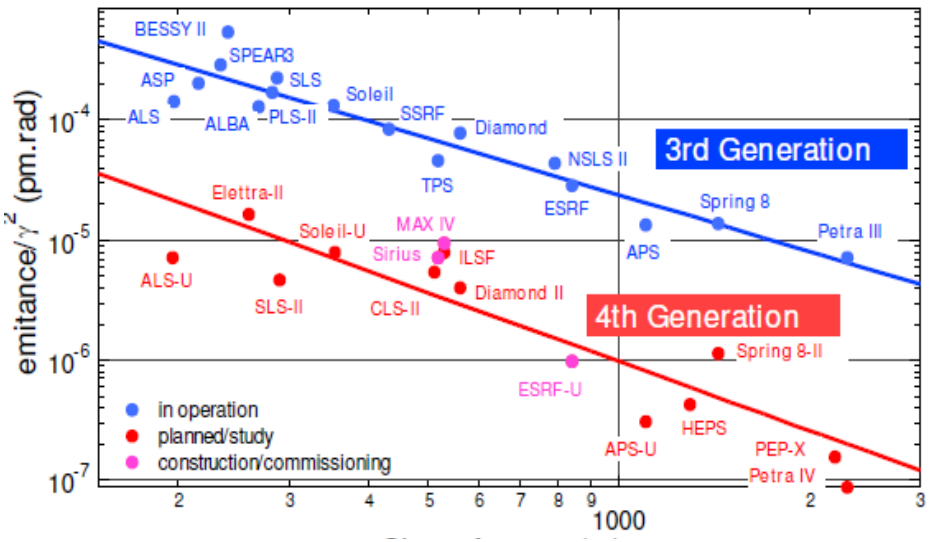
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

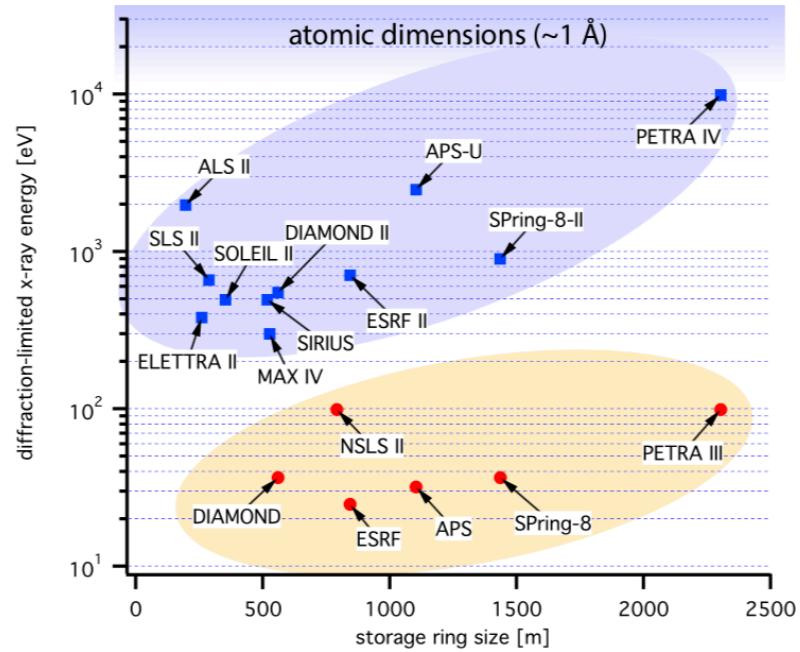
DESY.

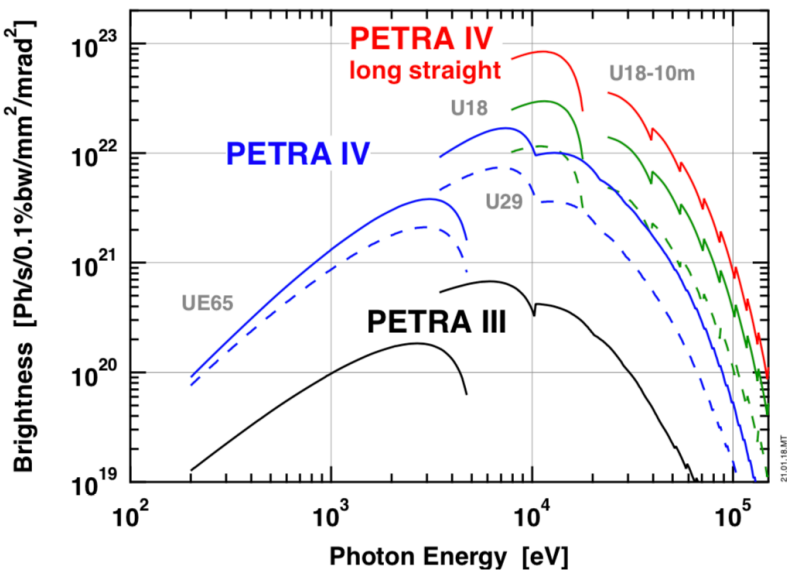
Content

- Introduction — needs of 4G light sources
- Possible ML applications
- Ongoing Big data relevant activities at DESY/PETRA

4G light sources

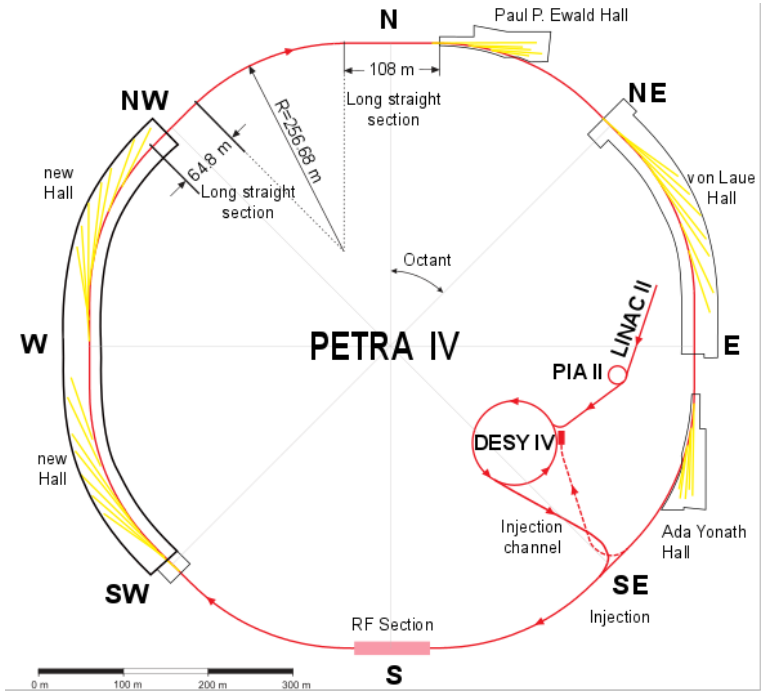






Changing the lattice from DBA to MBA
In a storage ring gives up to two orders of magnitude in figures of merit such as Brightness

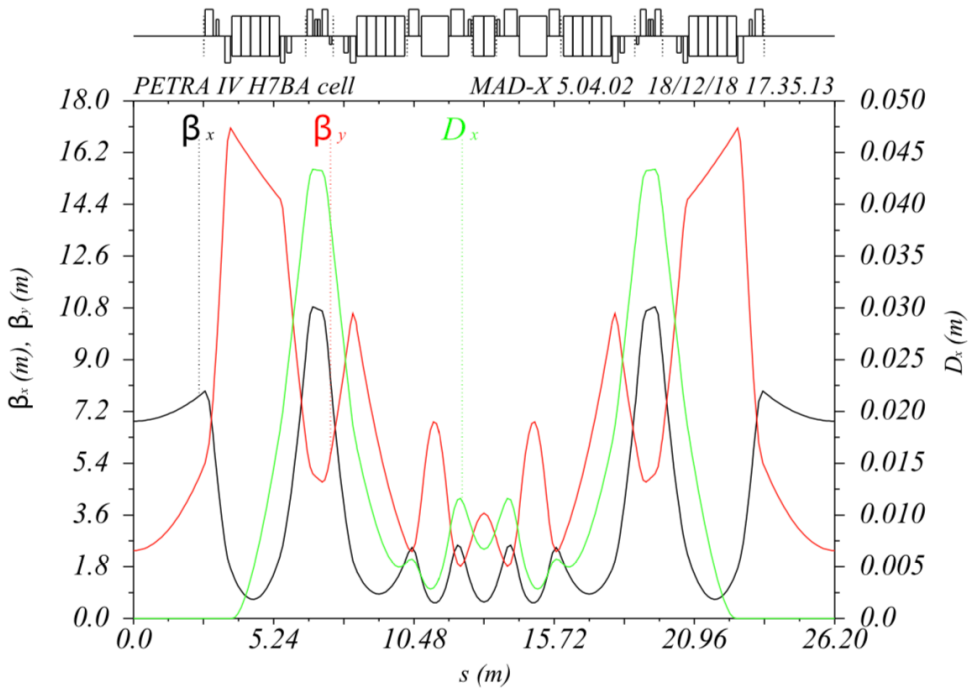
PETRA IV



Design parameter	PETRA III		PETRA IV	
Energy / GeV	6		6	
Circumference / m	2304		2304	
Emittance (horz. / vert.) / pm rad	1300 / 10		$<$ 20 / 4	$<$ 50 / 10
Total current / mA	100		200 [1]	80 [2]
Number of bunches	960	40	1600 *	80
Bunch population / 10^{10}	0.5	12	0.6	4.8
Bunch current / mA	0.1	2.5	0.125	1.0
Bunch separation / ns	8	192	4 / 20 (gaps)	96

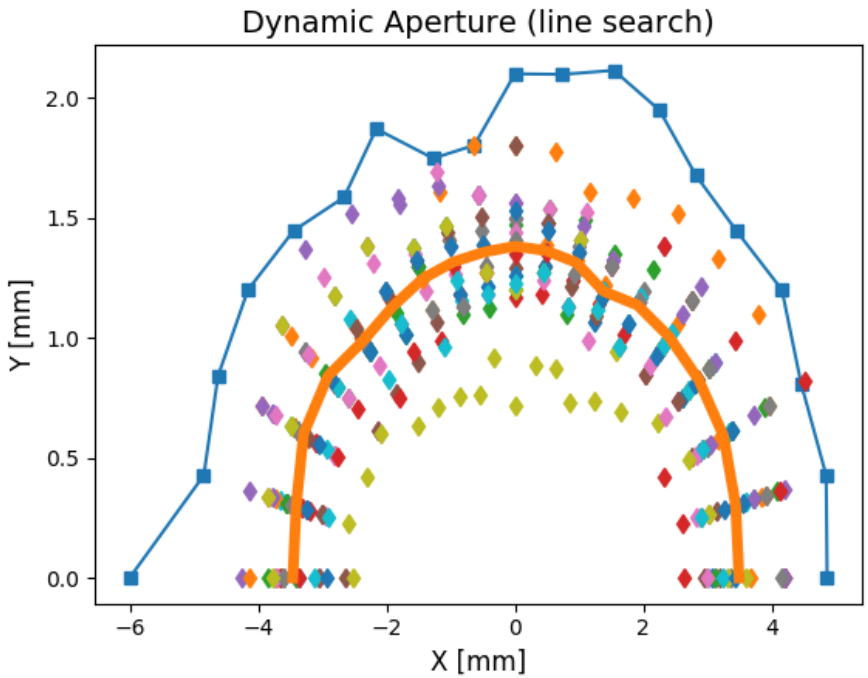
Challenges of 4G light source design

- Many tradeoffs to be made (vacuum chamber sizes, magnet strength, etc…) to stay technically feasible
- Nonlinear dynamics dominating (many options have close to 0 MA and DA)
- IBS and Touschek effects very prominent, need to operate with a lengthening cavity and round beams in high-intensity modes
- More demanding injection, extraction, and beam dumps



Challenges of 4G light source commissioning

- Users demand 1 max 2 years dark time, tight schedule
- Machine extremely sensitive to errors, ~30 µm alignment required
- Involved (automated) startup procedures required to store beam and reach target emittance



Dynamic aperture after simulated startup procedure for PETRA IV

Uncorrected machine is typically unstable

Challenges of 4G light source operation

- Reliability demands grow (95% -> 99%)
- But machines are more sensitive with larger number of components
- We would like to meet availability goals and provide required beam-hours to users
- But at the same time we would like to keep doing accelerator physics and spend dedicated machine time on studies rather than machine setup
- This could only be successful if all standard procedures are highly automated
  - Startup of components (magnet cycling etc.)
  - Orbit correction
  - BBA
  - Optics measurement and correction
- Methods are well understood but (high-level) controls software not designed for autonomous operation

Where does ML come in?

ML-assisted optimisation

Promise of faster and more efficient hyperparameter tuning

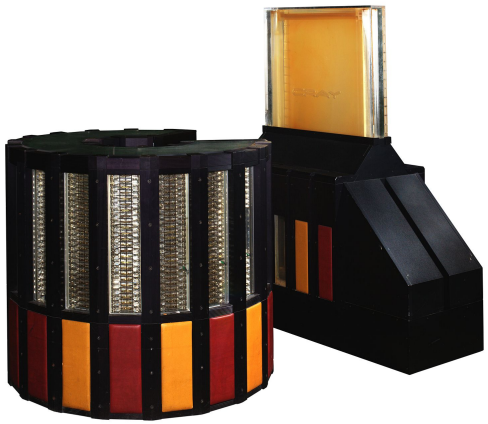
Relevant for storage rings but not of primary importance for operation

- ML first looked at in the FEL optimisation and failure prediction context
  - FEL tuning has a direct impact on user operation and plays central role
  - FELs are newer machines, more adequate diagnostics and DAQ
- 3G storage rings much more stable in operation than FELs. Some optimisation is required but effort incomparable to an FEL. Most optimisation work (lifetime, Dynamic aperture, injection efficiency) has little direct impact on users
- Optimizers such as OCELOT, RCDS etc., used at storage rings, but not our biggest. concern
- Focus on automation

Infrastructure – computing power and GPUs

5 orders of magnitude between most powerful computer in 1985 and a 5K consumer card (5 min training -> 1 year)

Cray-2 1.9 GFLOPS (1985)



	Tesla V100 for NVLink
PERFORMANCE with NVIDIA GPU Boost™	DOUBLE-PRECISION 7.8 teraFLOPS
	SINGLE-PRECISION 15.7 teraFLOPS
	DEEP LEARNING 125 teraFLOPS

Intel Pentium III (2000)
2 GFLOPS

Infrastructure at DESY

- Maxwell cluster @ DESY
- Low-medium-end HPC (480 nodes, 26488 cores) but very useful
- Plans for an ‘interdisciplinary data analysis facility’ (IDAF) and Data Science Center

Bright future

Platforms

	NV V100	NV P100	GTX 1080Ti	K40M	K20M	Vega 10 Instinct MI25	Quadro M6000
Maxwell	10	76	(4)	7 (+5)	12	-	14

	NV V100	NV P100	GTX 1080Ti	K40M	K20M	Vega 10	Quadro M6000
SP	14899 Gflop/s	9000 Gflop/s	10600 Gflop/s	4700 Gflop/s	3900 Gflop/s	12300 Gflop/s	6070 Gflop/s

c/o Frank Schluenzen DESY IT

ML and storage ring beam dynamics

- Although single-particle beam dynamics well developed, several directions still considered art, such as:
  - Multi-parameter matching in high dimensions (e.g. MBA. cell design)
  - Nonlinear aberrations (beyond simple ideas like low-order achromats, reducing sextuple strength, -I)
- A common line of reasoning — build fast “surrogate models” based on NN trained on simulated data

Sanity check — FODO

In [2]:
```
# try to learn FODO stability
import os,sys
from ocelot import *
from pylab import *
QF = Quadrupole(l=0.1, k1=0.1)
QD = Quadrupole(l=0.1, k1=-0.14)
D = Drift(l=1.0)
fodo = (QF,D,QD,QD,D,QF)
lat = MagneticLattice(fodo)
```

```
n_train = 10000

x_train = np.zeros([n_train,3])
y_train = np.zeros([n_train,1])


for i in range(n_train):
    QF.k1 = np.random.rand()
    QD.k1 = -np.random.rand()
    D.l = 2.0 * np.random.rand()

    x_train[i,0] = QF.k1
    x_train[i,1] = QD.k1
    x_train[i,2] = D.l
    lat.update_transfer_maps()
    tws = twiss(lat, Twiss())
    if tws is None:
        y_train[i,0] = 0
    else:
        y_train[i,0] = 1
```

```
import keras
from keras.models import Sequential
from keras.layers import Input, Dense, Dropout
from keras.utils import to_categorical
from pylab import *

#fixing a duplicate openmp dylib on mac???
import os
os.environ['KMP_DUPLICATE_LIB_OK']='True'

model = Sequential()
model.add(Dense(64, input_dim=3, activation='relu'))
#model.add(Dropout(0.5))
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(1, activation='sigmoid'))
model.compile(loss='binary_crossentropy', optimizer='sgd',metrics=['accura
model.fit(x_train,y_train, epochs=30, batch_size=16)
```

```
Using TensorFlow backend.
```

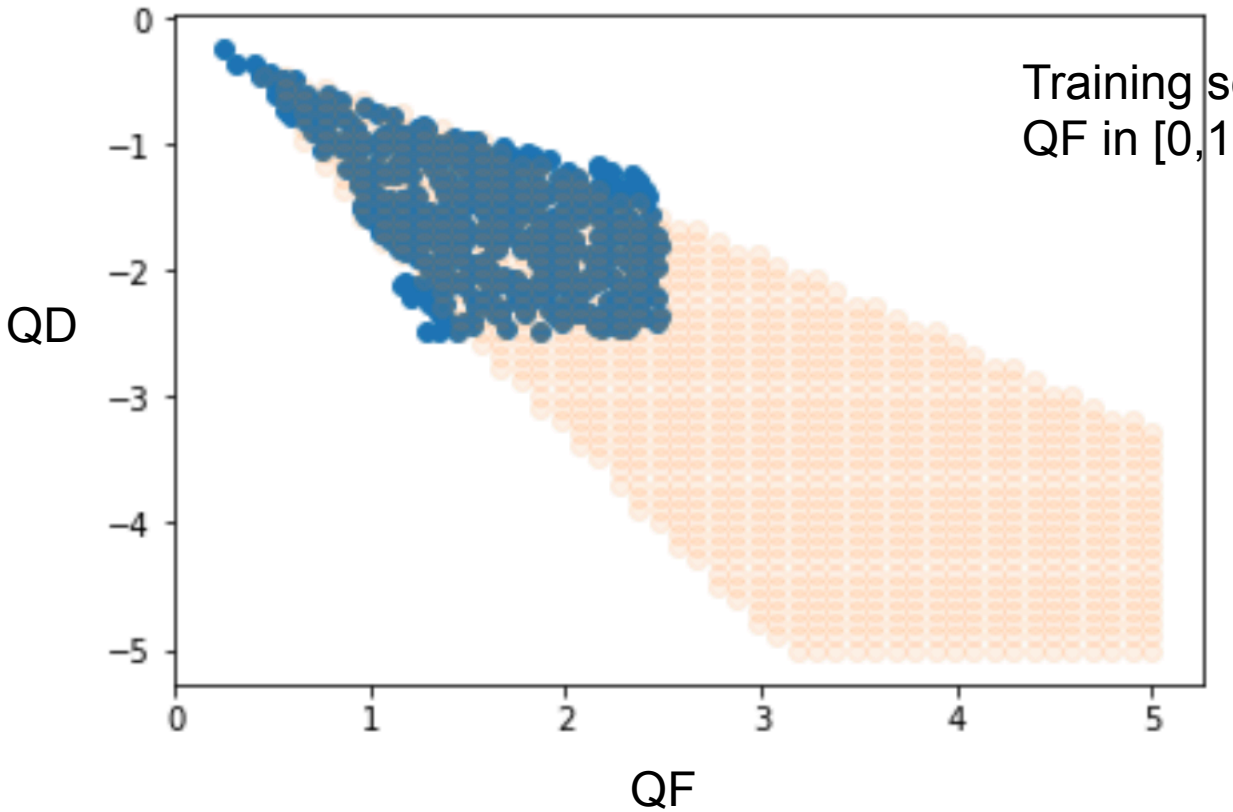
```
Epoch 1/30
10000/10000 [==============================] - 2s 211us/step - loss: 0.325
Epoch 2/30
10000/10000 [==============================] - 2s 153us/step - loss: 0.2748 - acc: 0.9313
Epoch 3/30
10000/10000 [==============================] - 2s 152us/step - loss: 0.2614 - acc: 0.9313
```

```
n_test = 5000

x_test = np.zeros([n_test,2])
y_test = np.zeros([n_test,1])

D.l = 3.9

for i in range(n_test):
    QF.k1 = 5*np.random.rand()
    QD.k1 = -5*np.random.rand()

    x_test[i,0] = QF.k1
    x_test[i,1] = QD.k1
    lat.update_transfer_maps()
    tws = twiss(lat, Twiss())
    if tws is None:
        y_test[i,0] = 0
    else:
        y_test[i,0] = 1
```

Sanity check — FODO

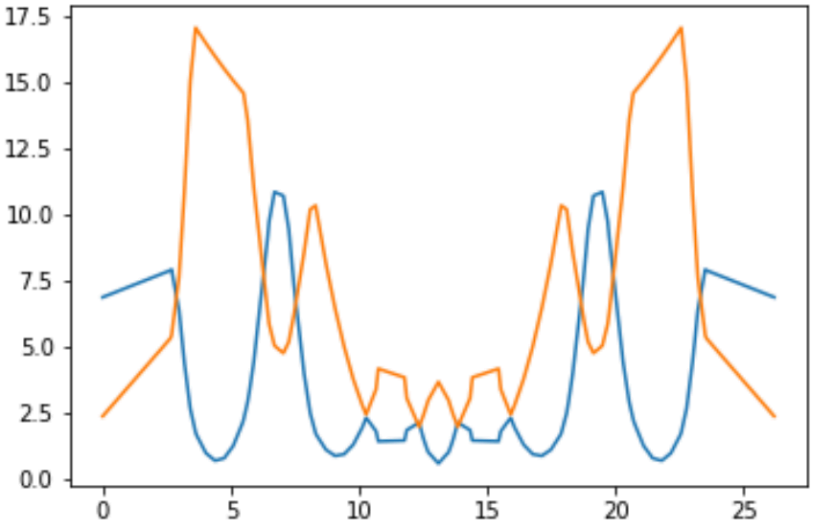
Stability diagram check on the test set
QF in [0,5] QD in [-5,0] L=3.9

Training set was
QF in [0,1] QD in [-1,0] L in [0,2]

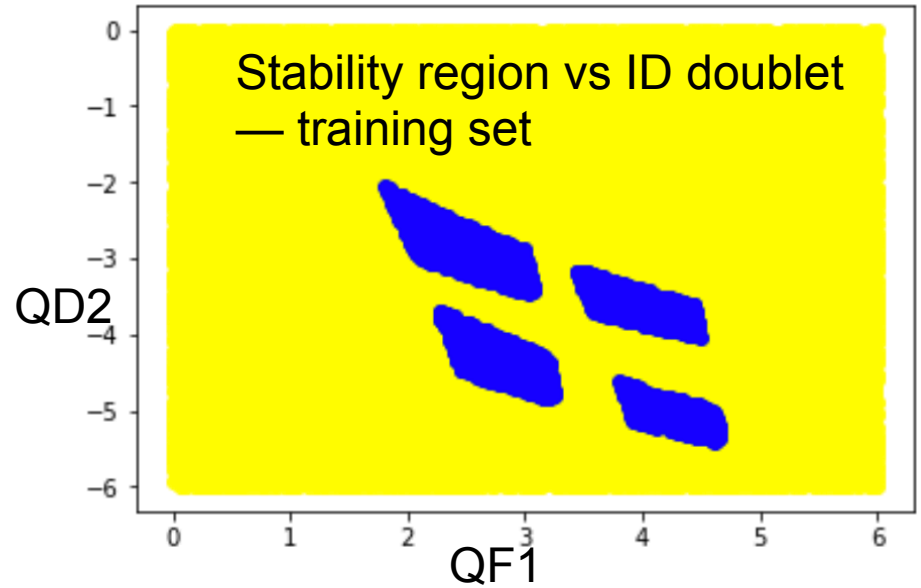


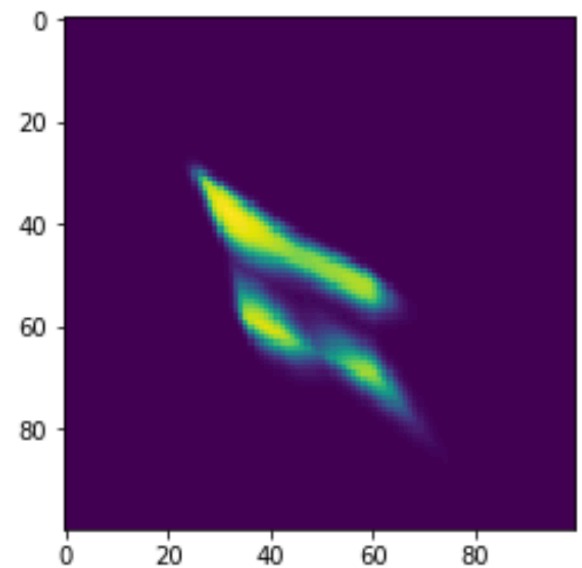
- The trained NN generalises surprisingly well beyond training set parameters!
- However only within a certain range, and validity range needs checks
- Practical application not clear — some real-time FPGA-based response matrix calculation for non-linear, time-dependent situations etc. ?

A more complex linear optics example



Cell of a 4G storage ring ('HMBA')



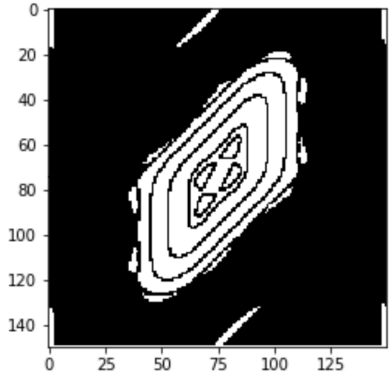


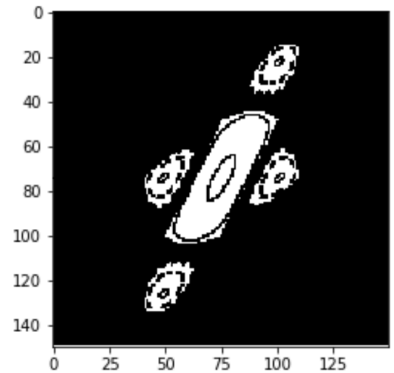
Learned
By NN

- Straightforward application of NN tends to inefficiently represent such data
- Bottleneck is the training set generation: In reality we need at least 10 dimensions to work with!!!
- More work needed

Sanity check — Hénon nmap

2-layer NN trained to generate Hénon map based on nonlinearity parameter as input

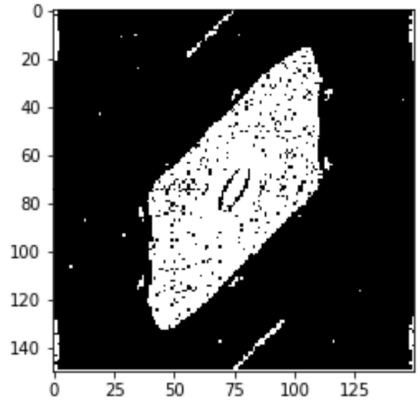


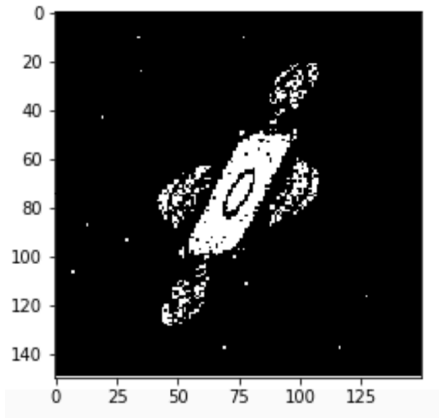


Phase portraits

NN can generate images which look like phase portraits also capturing rough dependency on the non-linearity parameter
It is however very hard to train the NN to reproduce fine features
Practical application not clear (unless sold as art)

Generated by NN





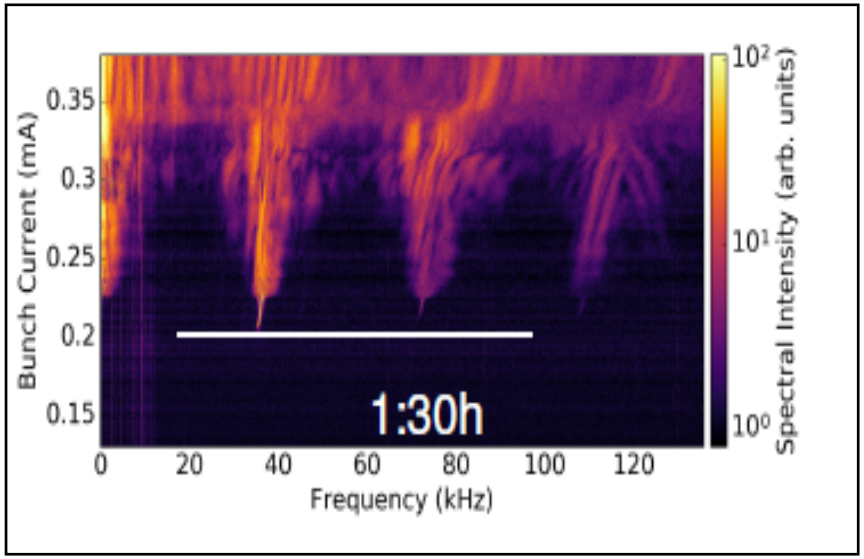
Portrait of Edmond Belamy
Generated by GAN
Sold at Christies for $432,500, 2018

Single-particle dynamics perspectives

- Deep learning might be useful for single-particle dynamics analysis
- The key is however in computation speed, not directly related to ML
- Questions of calculating DA, MA, low-dimensional sextuple strength optimisation etc. are now routinely solved within minutes to hours on ~1000-core machines (required lengthy dedicated studies in the 1990s)
- We could gain another 1-2 orders of magnitude in performance with tracking on GPU — e.g. GPU version of *Sixtrack*
- With super-massive computations NN might be a way to represent simulated results, along with other statistical tools
- Letting ‘AI’ do lots of calculations and then pick its brain, same way you would do with an RA?

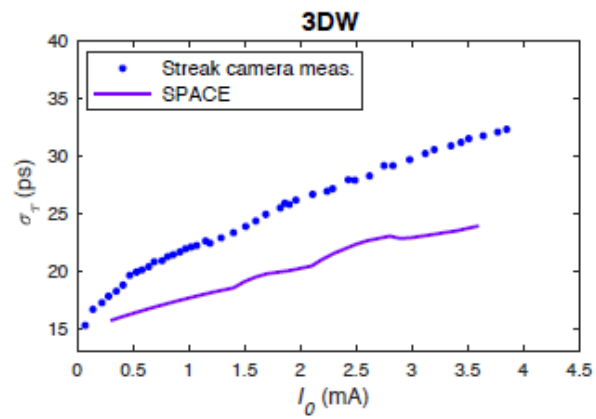
ML and collective effects

- One of less understood areas of storage rings is high-charge limitations
- Relevant for PETRA IV timing mode
- Relevant for many lower-energy machines pushing for higher current (500 mA NSLS-II, MAX-IV) or coherent radiation modes
- Large-scale simulation/measurement campaigns + model fitting could be useful (some work done at KIT)



Experimental microbunching instability studies at KARA/KIT

NLSL-II bunch lengths vs. intensity
from A. Blednych, BCTLESR 2019



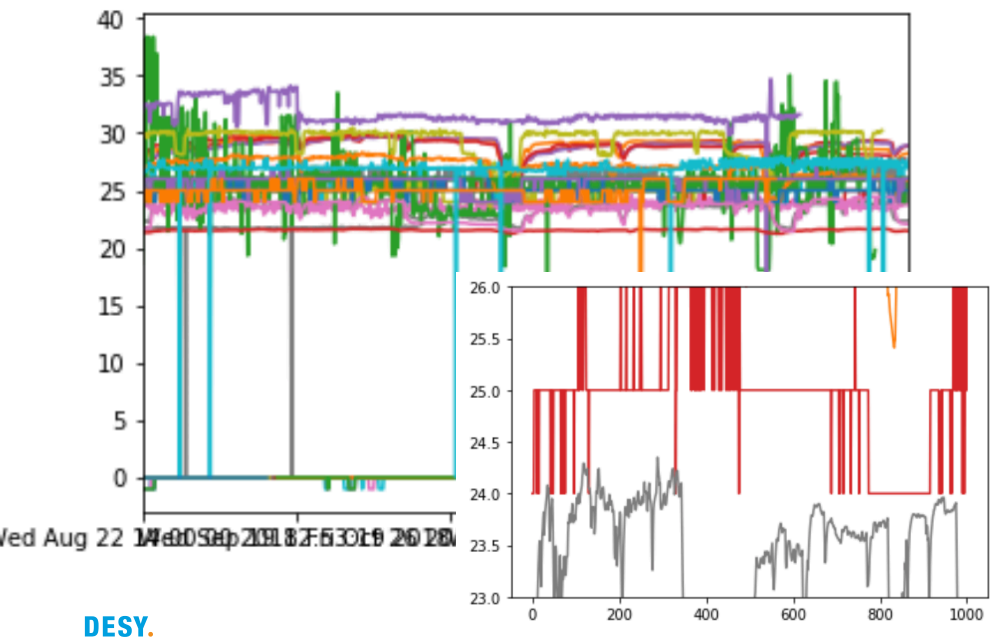
- We believe, the instability thresholds can be detected from the bunch length measurements with a streak camera resolution better than $0.5ps$.

ML and optics measurement

- Straightforward applications of optics measurement (especially all quick methods like fast LOCO, pinged beam and AC dipole) often performs poorly
- Need some expertise to do the measurement (e.g. properly select the current) and process data (e.g. remove bad BPM readings)
- ML techniques are looked at to speed that up
- For a light source its hard to beat a properly done LOCO

Big data and operational statistics

- A typical activity is to figure out what correlates to what through archive mining and try to improve things
- Typical problems:
  - Unprocessed data from PETRA III archive not easy to deal with (G. in g. out for any algorithm).
  - Lengthy manual processing required to arrive at a meaningful dataset
  - Amount and quality of data after processing could be surprisingly low in comparison to the raw data



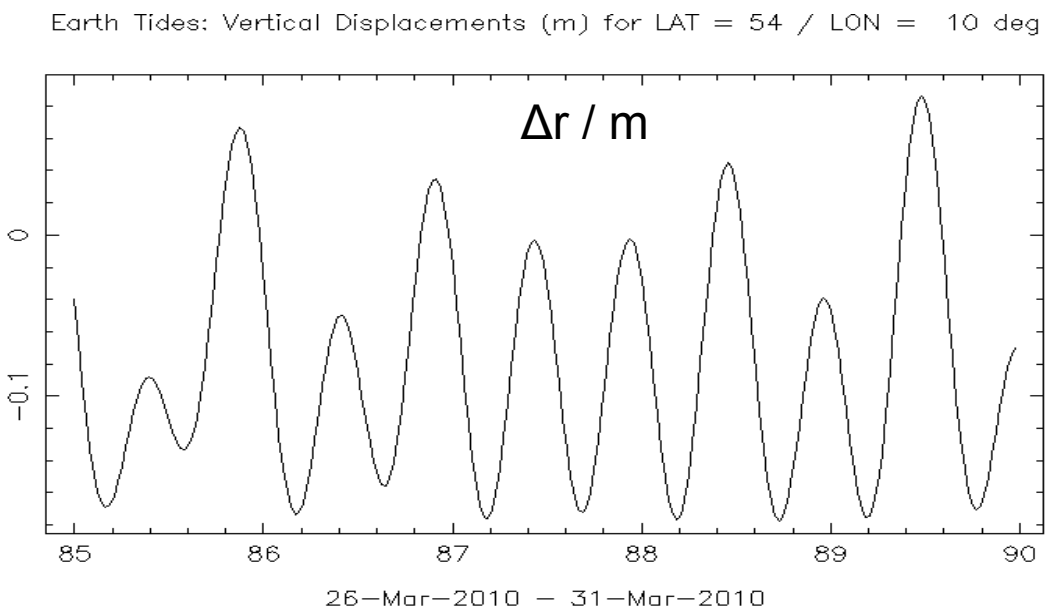
Typical pull from archive

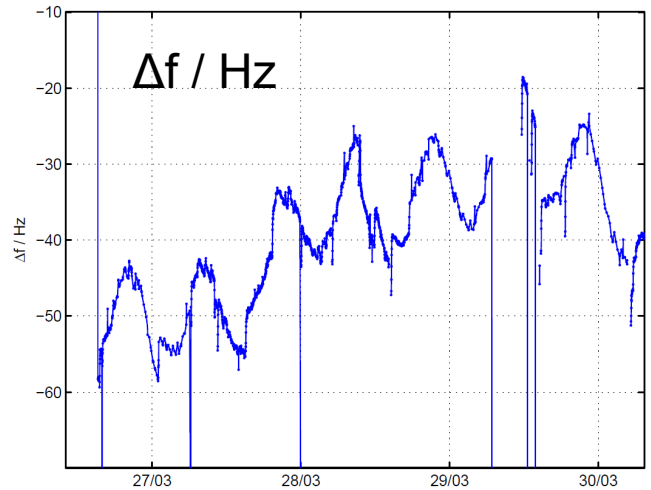
- All temperature sensors, Aug 22-Nov30 2018
- Fraction of channel data missing (25%)
- Unmatched timestamps
- Data archived on thresholds
- Bad readings/datapoints

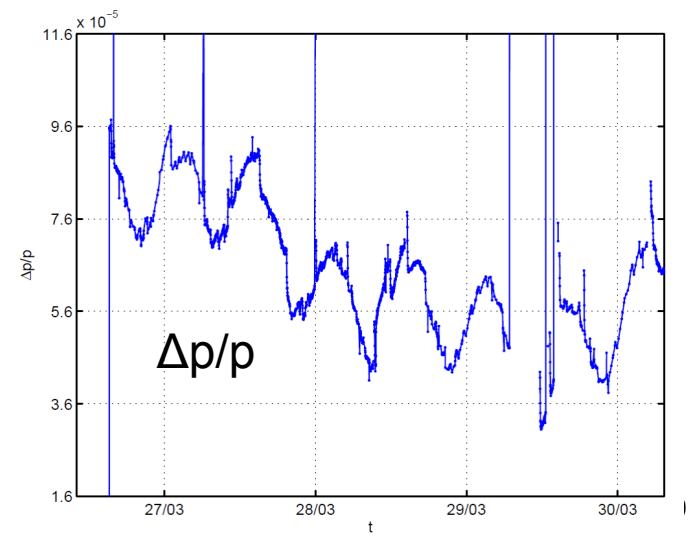
Operational statistics examples

- Required RF frequency adjustment found out to be correlated to tides

Joachim Keil (2010)



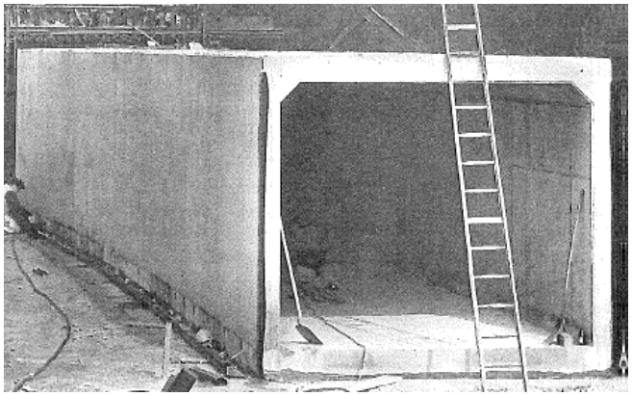


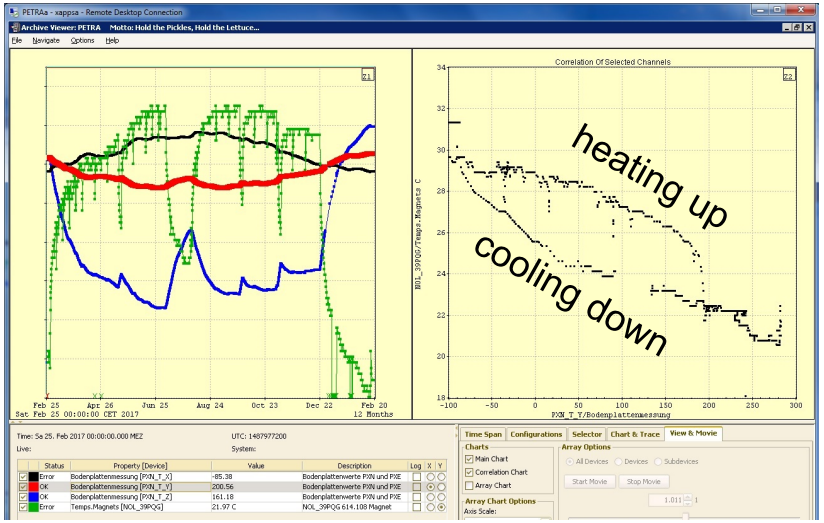


Operational statistics examples

- Movement of tunnel segments with temperature
- Correlation analysis shows ~40 µm/K vertical, ~300 µm/K long. displacement
- >10 deg. yearly temperature variation
- Stabilisation work on critical path for the PETRA IV project

Michael Bieler, Jens Klute





Ongoing campaign

- Mechanical stabilisation of primary importance
- Temperature variation main deformation cause
- System too complex for reliable modelling (e.g. with ANSYS)
- Need 3D temperature mapping to draw further conclusions
- Step 1: lots of resistive sensors in a short section
- Step 2: mapping with IR cameras, tests
- Step 3: if successful, more mapping with IR cameras
- In parallel: mechanical stability with stretched wires (in tunnel)

Bjoern Lemcke

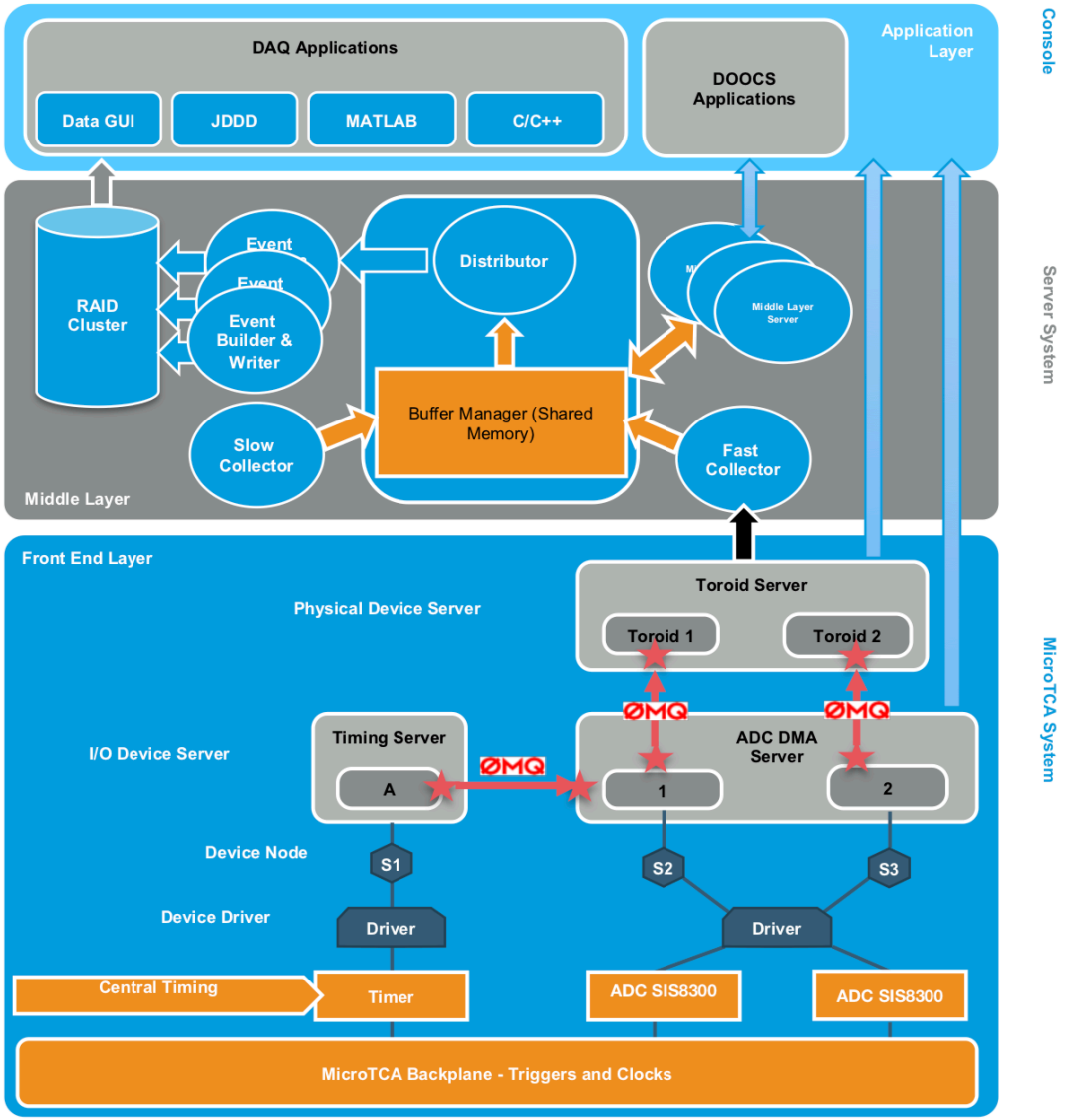
Foils
S103941PEX36A

PT100

High level controls and DAQ

Tim Wilksen

- High throughput DAQ required to archive parameters
- Experience from XFEL
- Work ongoing to refactor high level controls



Prototyping ML workflow — software stack
Results at the 3rd ML workshop
Happy to collaborate!!!

NVIDIA CUDA

python™

TensorFlow

Keras

Pandas

django

mongoDB

PyTine

Conclusion and Outlook

- Role to be played by ML in 4G light sources to be seen. No low-hanging fruit. Within the PETRA IV project ML is a part of a larger simulation and automation campaign
- Goals for simulation:
  - Ultra-fast codes capable of scanning large parameter spaces
- Goals for controls/operation/automation
  - Rework data acquisition and curation
  - Rework high level controls
  - Boost diagnostics and monitoring
  - Free personnel and machine time for studies given high pressure on availability and reliability
- Activities being launched to address that (developments starting supported by several funding vehicles within the Helmholtz Association)
- In ML-centric programming paradigm data is the most important asset, and we are focusing on data collection/production first
- Will have more in a year!