Multiobjective Optimization of Injectors

Christos F. Papadopoulos,
Lawrence Berkeley National Laboratory

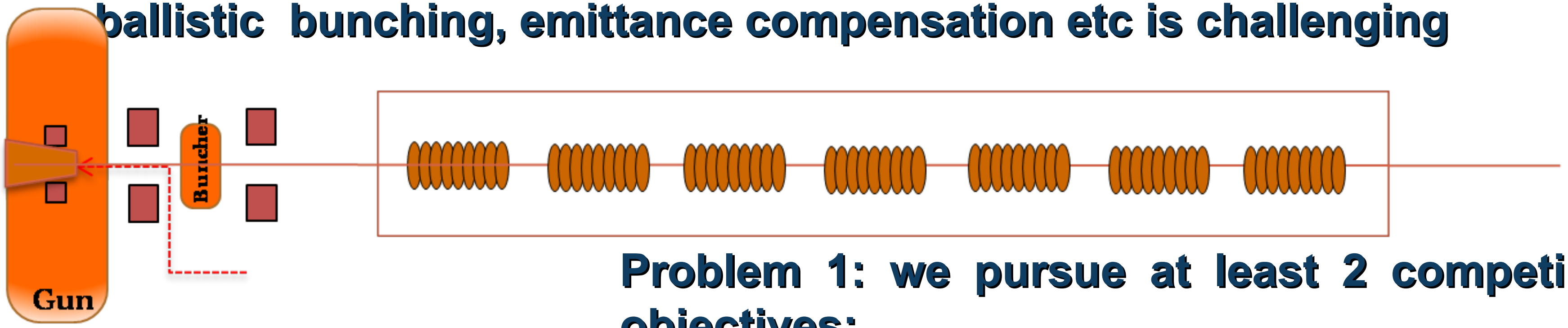
References

I. V. Bazarov, C. K. Sinclair, Phys. Rev. ST Accel. Beams 8, 034202 (2005):
Excellent resource on injector dynamics & multiobjective optimization

K. Deb, Multi-Objective Optimization Using Evolutionary Algorithms, (2002):
Thorough discussion of different optimization strategies

Our Motivation

Designing an injector that takes advantage of laser shaping, velocity and ballistic bunching, emittance compensation etc is challenging



🟧 **Buncher cavity:**
Control Phase, Amplitude
Used for compression

🟥 **Solenoids:**
Control Strengths
🟥 Used for emittance compensation

Multicell Accel. Cavities
Control Phase, Amplitude
Used for compression, acceleration

← **Laser Pulse:**
Control transverse and longitudinal profile
Defines initial bunch length and emittance

Problem 1: we pursue at least 2 competing objectives:
low emittance → requires low SC forces
High current → creates high SC forces

Problem 2: Components quickly add up.
In order to take advantage of the flexibility of the system, we need to optimize > 10 parameters.

Problem 3: Analytic models offer good qualitative understanding, but not quantitative design parameters, since they do not include nonlinear SC, detailed fields etc

Multiobjective Optimization

Definition of our 1st problem:

Maximize

$$f_m(x_1, x_{2,..}, x_n), m=1,\ldots, M\ ;$$

Subject to the constraints

$$g_j(x_1, x_{2,..}, x_n) \geq 0,\ j=1,\ldots, J\ ;$$

$$x_i^{(L)} \leq x_i \leq x_i^{(U)}, i=1,\ldots, N\ ;$$

In practical terms:

The objective functions f: The most important quantities of the beam, such as emittance, bunch length, energy spread. Ultimately, in start to end simulations, the goal is to optimize the quantities users want.

The constraints xu, xl: upper and lower values allowed in our "knobs", such as peak gradients of RF fields, minimum and maximum laser pulse length, etc

The constraints g: Constraints on calculated quantities, such as transverse beam size, number of particles lost, etc

The difference between the f's and the g's can be subtle, and deciding which is which depends on whether we have a gradual or a threshold effect.

Pareto Optimality

Definition of dominance:

Solution A is said to dominate solution C if A is not worse than C in all the objectives f, and is strictly better than C in at least one objective.

Definition of non-dominated subset:

The subset P' of solutions within a solution set P that are not dominated by another member of P.

Definition of Pareto-optimal set:

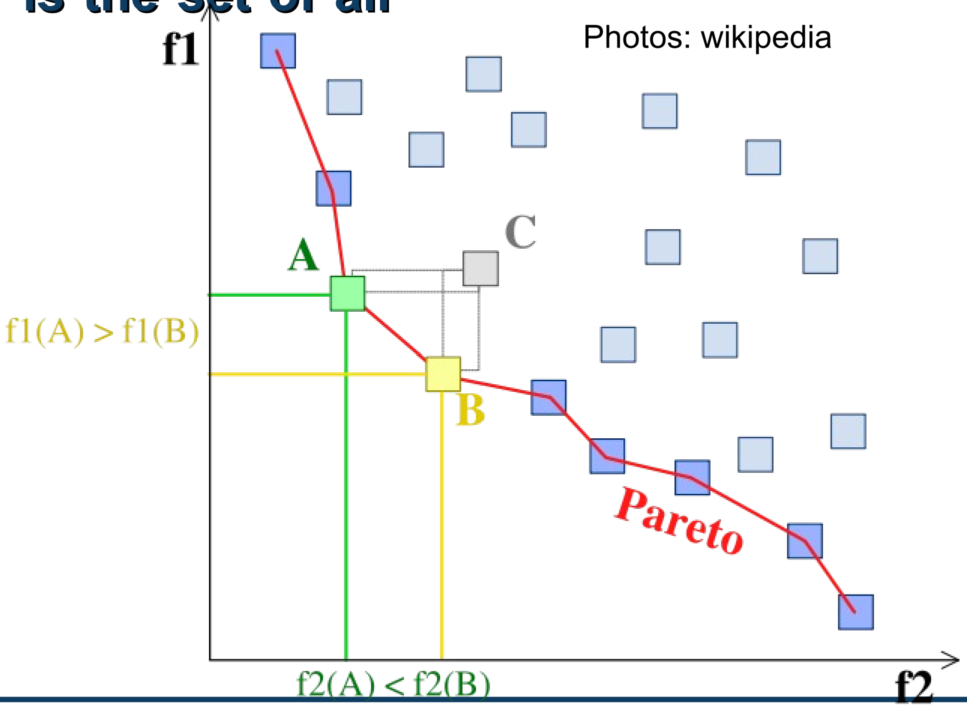
The non-dominated subset P' when P is the set of all the allowable search space

Vilfredo Pareto
1848-1923

Photos: wikipedia

Addresses problem 1:

The Pareto optimal set (or front) allows us to visualize the trade-off between the objectives.



Genetic Algorithms

A multi-stage process

1. Initialize population
2. Evaluate objective functions / constraints
3. Assign fitness to all individuals (convergence & diversity), non-dominated solutions are preferred.
4. Stochastically choose a subset for mating pool (higher fitness being preferred)
5. Apply crossing and mutation operators to generate offspring
   - Crossing: combine solutions to (hopefully) find a better one
   - Mutation: Introduce randomness to investigate larger volumes of parameter space
6. Evaluate objectives / constraints for the offspring
7. Repeat from step 3.

Addresses problem 2:

Global and efficient search in the multidimensional parameter space. The random aspect helps avoid local minima.

One caveat:

We can only approximate the Pareto optimal set, as we are not guaranteed that another solution which dominates all or some of the current ones does not exist.

ASTRA

Particle-in-Cell code, includes trans. and long. space charge, widely used and benchmarked for photoinjectors

300 pC charge

10k-100k particles

Variable step size, Variable grid

Not enough to resolve microbunching

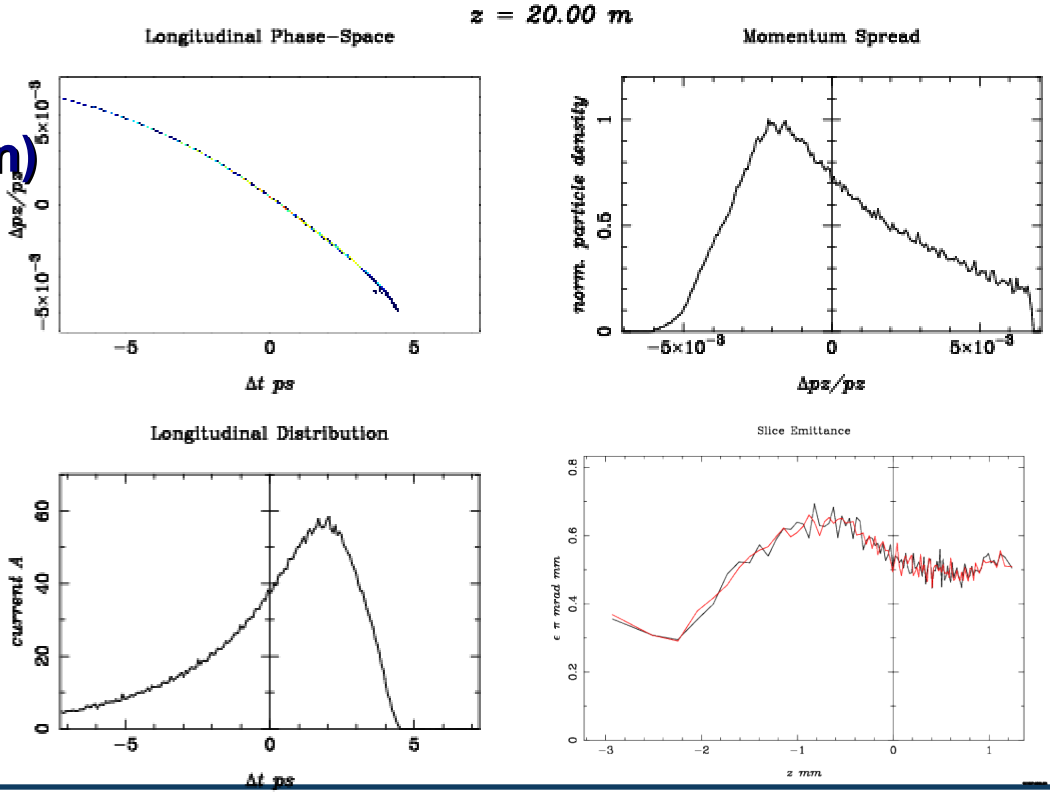
No CSR or wakefields

Good enough for core properties, emittance growth

FAST (10 mins-1hr for a single run)

Addresses Problem 3:

Nonlinear SC forces, realistic fields



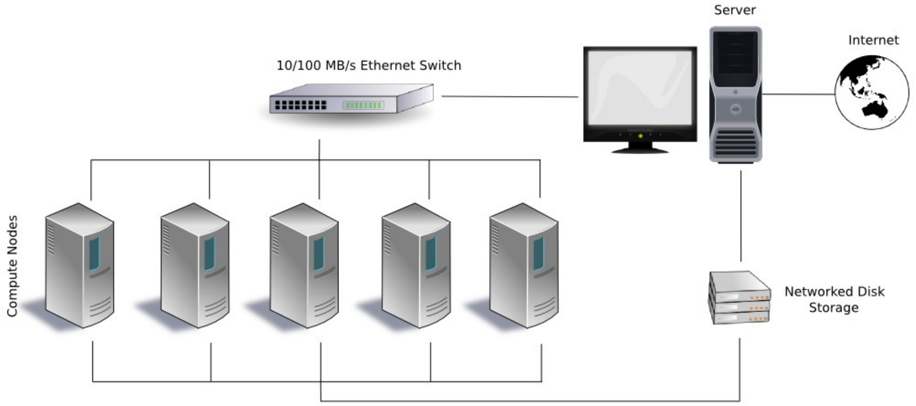
Parallelism

Use a cluster of linux machines

Computes 100s of solutions at once, usually corresponding to 1 generation

“Embarrassingly” parallel, since each solution is calculated on 1 processor

Few cycles are lost, since each solution takes comparable time, and overall computation scales well with the number of processors



Addresses Problem 3

Allow for faster computation of multiple solutions

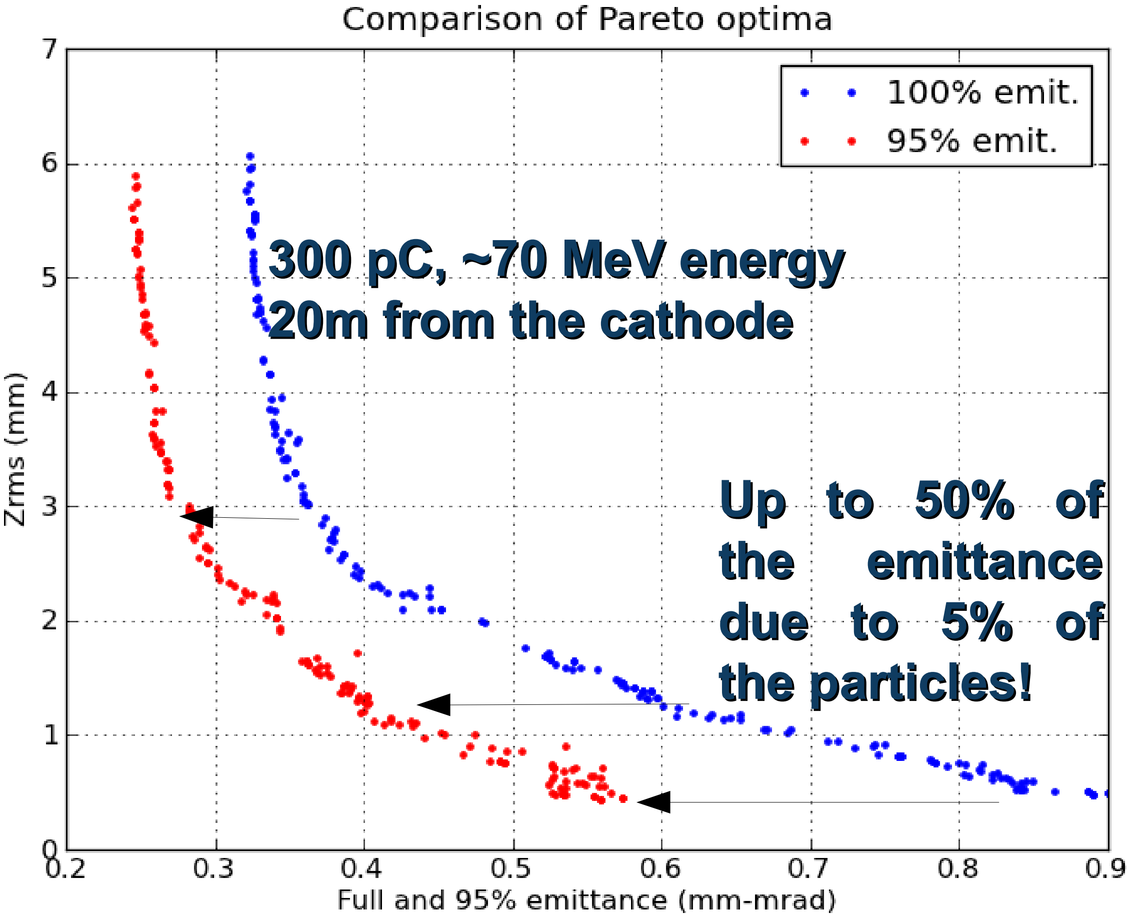
An example of the process

Choose objectives at the injector exit

Start with emittance and bunch length

Also interested in 95% emittance to estimate halo

100's of generations, a few days running in a cluster (computer time, not people's time)



300 pC, ~70 MeV energy
20m from the cathode

Up to 50% of the emittance due to 5% of the particles!

Requirements from linac and FEL simulations:

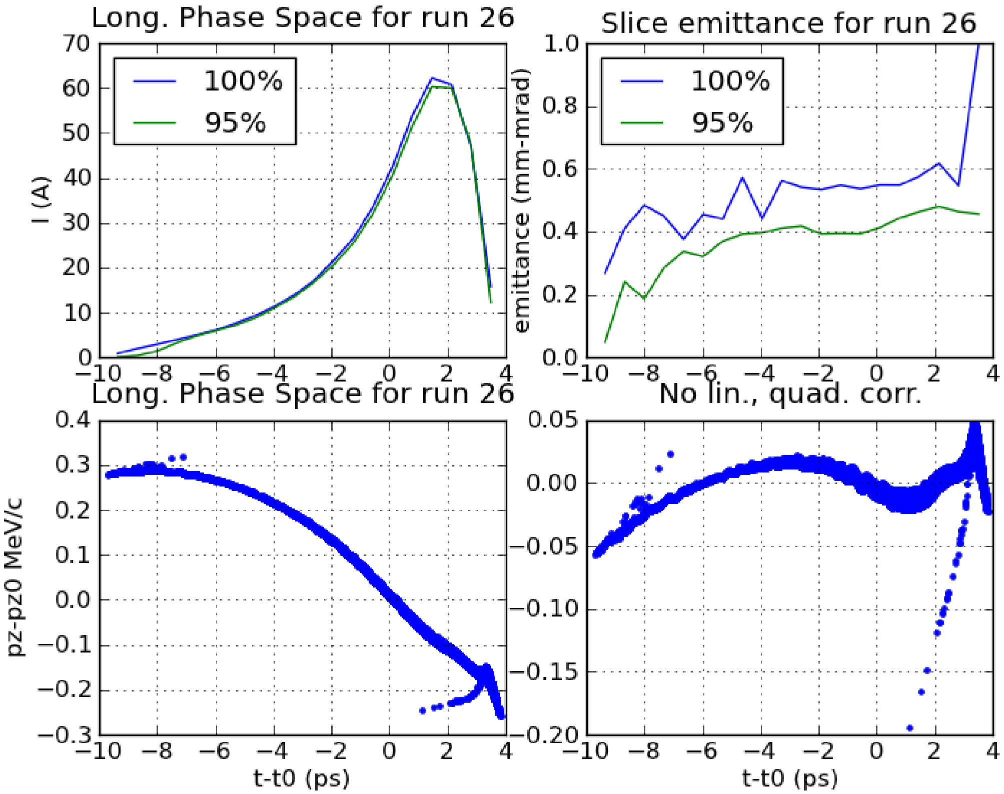
Ipeak >50 A
εn < 0.6 μm
ΔE < 4 keV

Pick one solution: Run 26

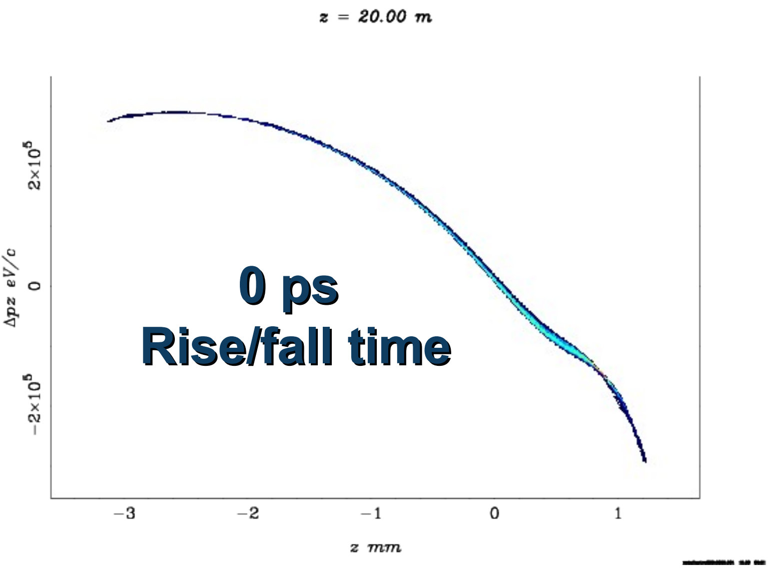
Within specs for peak current, slice emittance (95%)

But the tails and and the high order momentum-time correlations were problematic during the linac transport.

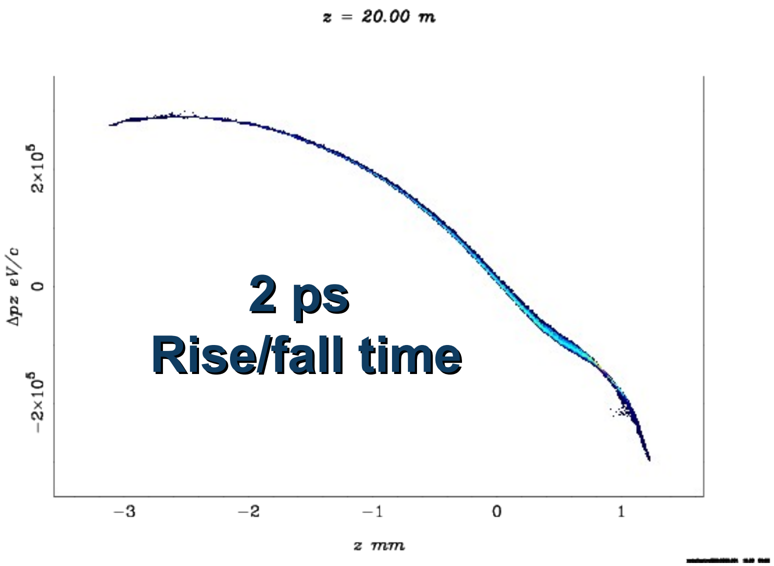
Specifically, after removing 1st and 2nd order t-pz correlation, the remainder rms e-spread is 12.95 keV



Tail management

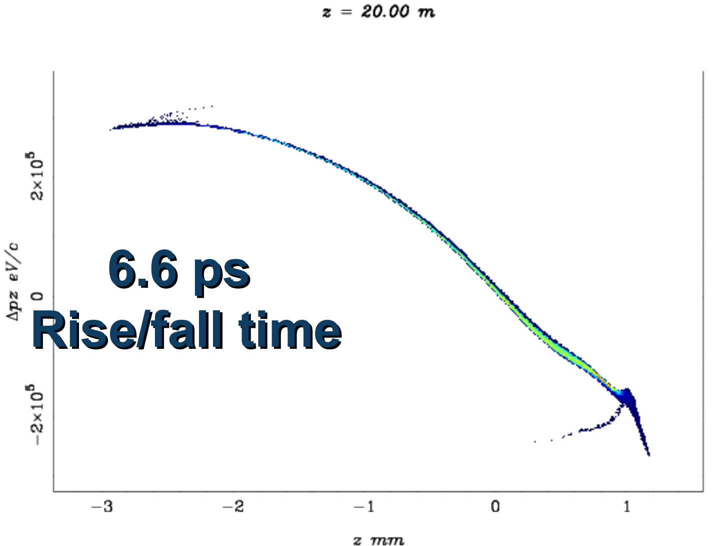


0 ps
Rise/fall time



2 ps
Rise/fall time

Improve the laser pulse by reducing the rise/fall time
For run 26:
66 ps plateau,
6.6 ps rise/fall time
We should be able to do better than that

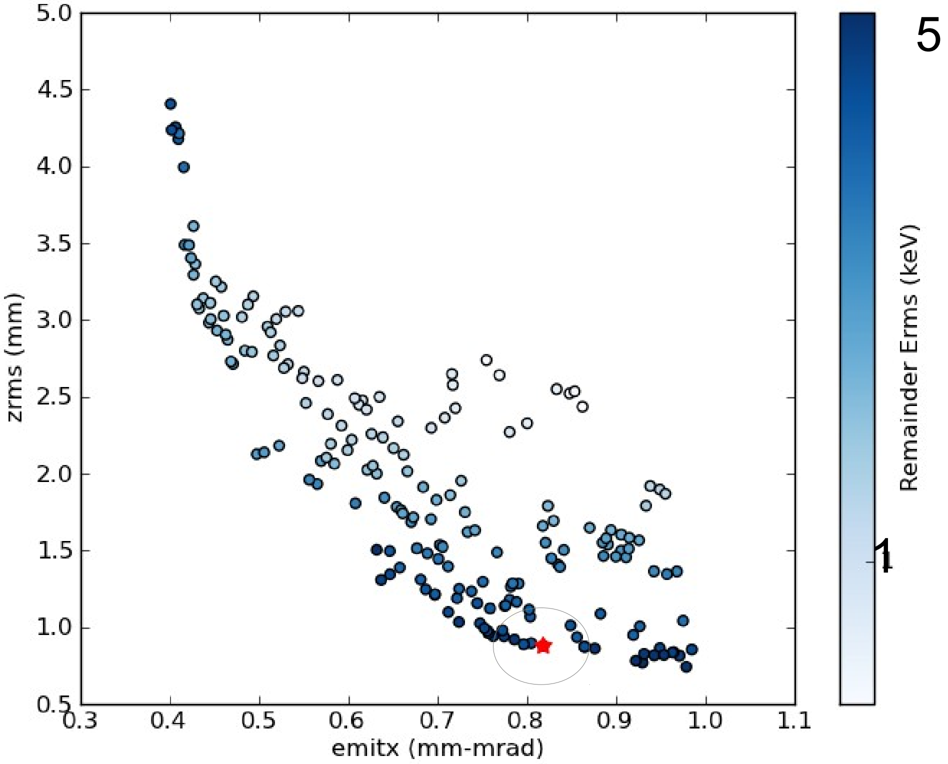


6.6 ps
Rise/fall time

High order energy-time correlations

Add a 3rd objective: The energy spread after removing <t*pz> (can use the downstream linac for this) and <t^2*pz> (use a 3rd harmonic cavity)

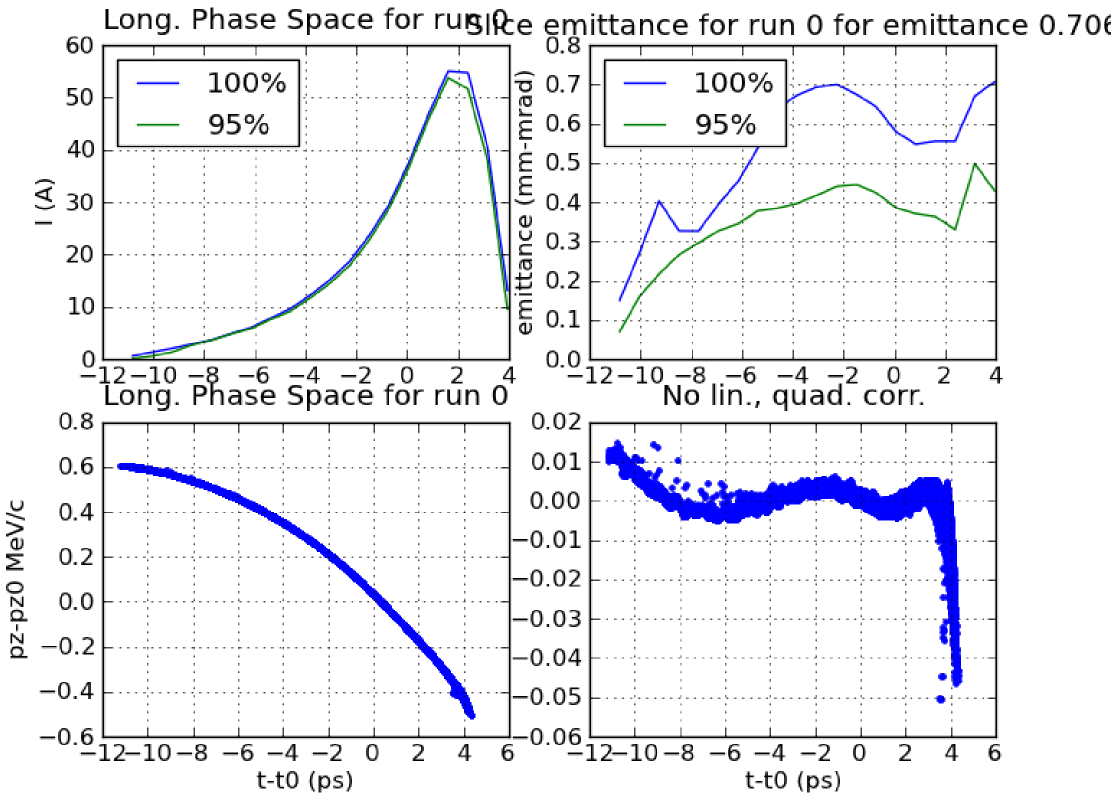
The pareto front is no longer a 1D curve



New solution: Run 184

We lose a bit in current and emittance, but reduce the remainder energy spread by a factor of more than 3

New remainder rms e spread = 3.325 keV



What changed

Reducing the laser tails improves the emittance
In turn this allows for shorter bunches at the cathode

But this intuition only came after the optimizer had found the new solution.
Tweaking by hand would require a recalculation of the whole beamline