LINAC Energy Management (LEM) Upgrade Path

He Zhang
Center for Advanced Studies of Accelerators (CASA), Jefferson Lab

OPS Stay Retreat, July 15th, 2015

Outline

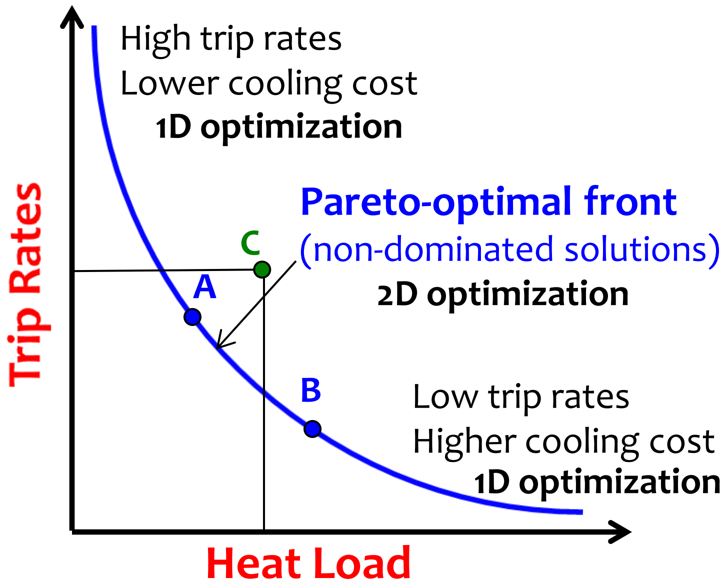
- Problem and goal
- One Objective minimization
- Multi-objective optimization
- Current and future work

This talk is based on the previous work by Balša Terzić, Alicia Hofler, Geoff Krafft, Jay Benesch, Arne Freyberger, Adam Carpenter, et al.

Background and Motivation

- Problem
  - Find the optimal set of cavity gradients to *simultaneously* minimize trip rates *and* minimize the dynamic heat load (electricity bill)
  - Monthly electricity bill for JLab is measured in millions of dollars – a large part of it is cryogenics
    Even modest improvements in cooling may translate into millions $ in savings
  - Dynamic heat load and trip rates are *competing objectives* – it is a multi-objective (2D) optimization problem
- Goal
  - Provide a set of feasible solutions (Pareto-optimal front) showing the *trade-offs between competing objectives heat load and trip rates*

A dominates C: C is not on the Pareto-optimal front



Model of the Problem

- The cavity power transfer to the liquid helium for CEBAF SRF cavities:

$$P(\mathbf{G}) = \sum_{i=1}^{N_c} \frac{G_i^2 L_i}{c_i Q_i(G_i)}$$

$\mathbf{G} = (G_1, G_2, ..., G_{N_c})$ cavity gradient,
L_i cavity length, $c_i = R_i/(Q_i L_i)$, R_i shunt impedance,
$Q_i(G_i)$ measured (unloaded) cavity quality factor

- The cavity trip rate:

$$T(\mathbf{G}) = 3600 \sum_{i=1}^{N_c} \exp[A + B_i(G_i - F_i)]$$

$A = -10.26813067$,
B_i the model trip slope,
F_i the fault gradient

- The constraint: the total energy gain in the linac is within 2 MeV of a prescribed energy E_{linac}.

$$\begin{aligned} &\text{Minimize} && P(\mathbf{G}), T(\mathbf{G}) \\ &\text{Subject to} && |E_{\text{linac}} - \textstyle\sum_{i=1}^{N_c} G_i L_i| < 2, \\ & && 3 \le G_i \le D_i,\ D_i \text{ max. gradient of a cavity.} \end{aligned}$$

1 Obj. Minimization Using Lagrange Multipliers

- Use Lagrange multipliers to minimize the *heat load only* or *trip rates only*
- Single-objective optimization problem:

Lagrangian:

$$\mathcal{L}(G_i, \lambda) = P(G_i) + \lambda(E - \sum_{i=1}^{N_c} G_i L_i)$$

N_c+1 equations:

$$\frac{\partial \mathcal{L}(G_i, \lambda)}{\partial (G_i, \lambda)} = 0$$

Solve for G_i and λ:

$$\lambda_p = \frac{2E_{\text{linac}}}{\sum_{i=1}^{N_c} c_i Q_i L_i}, \quad G_i = \frac{\lambda_p}{2} c_i Q_i$$

Conserved quantities:

$$\frac{G_i}{Q_i} = \frac{\lambda_p c_i}{2}$$

1 Obj. Minimization Using Lagrange Multipliers

- Use Lagrange multipliers to minimize the *heat load only* or *trip rates only*
- Single-objective optimization problem:

Lagrangian:

$$\mathcal{L}(G_i,\lambda) = P(G_i) + \lambda(E - \sum_{i=1}^{N_c} G_i L_i)$$

$$\mathcal{L}(G_i,\lambda) = T(G_i) + \lambda(E - \sum_{i=1}^{N_c} G_i L_i)$$

N_c+1 equations:

$$\frac{\partial \mathcal{L}(G_i,\lambda)}{\partial (G_i,\lambda)} = 0$$

Solve for G_i and λ:

$$\lambda_p = \frac{2E_{\text{linac}}}{\sum_{i=1}^{N_c} c_i Q_i L_i}, \quad G_i = \frac{\lambda_p}{2} c_i Q_i$$

$$\lambda_T = \frac{E_{\text{linac}} - \sum_{i=1}^{N_c} D_i L_i - \sum_{i=1}^{N_c}\left[\frac{L_i}{B_i}(\ln\frac{L_i}{3600 B_i} + A) - F_i L_i\right]}{\sum_{i=1}^{N_c}\frac{L_i}{B_i}},$$

$$G_i = \frac{\ln\frac{\lambda_T L_i}{3600 B_i} - A}{B_i} + F_i$$

Conserved quantities:

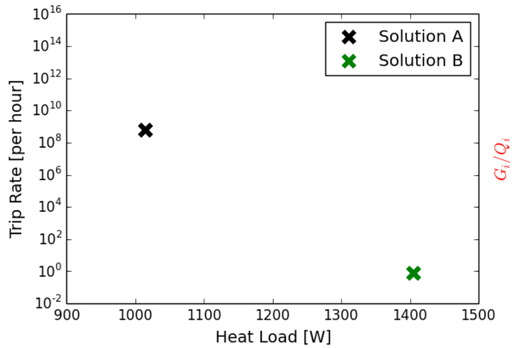
$$\frac{G_i}{Q_i} = \frac{\lambda_p c_i}{2} \qquad B_i \exp[A + B_i(G_i - F_i)] = \frac{\lambda_T L_i}{3600}$$

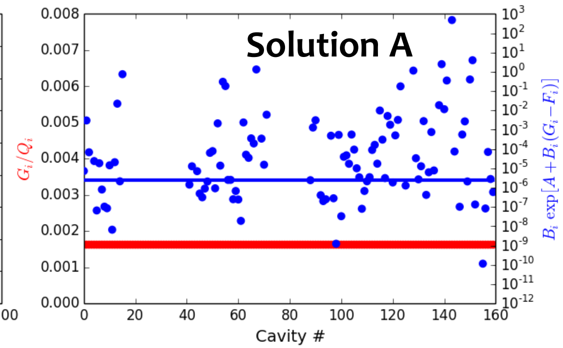
[Benesch *et al.* 2009 JL-TN-09-41]

1 Obj. Minimization Using Lagrange Multipliers

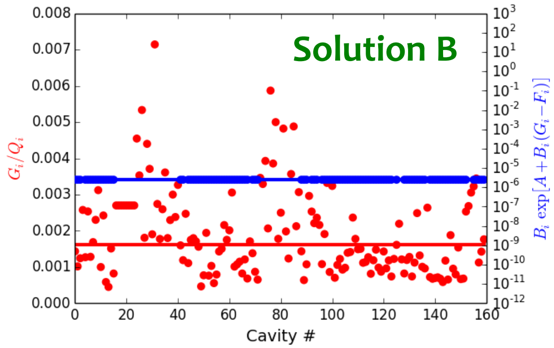
- Single-objective (1D) analytical solutions with Lagrange multipliers are pedagogic, but also somewhat useful
  - *Give us the limits of the optimization*

North Linac





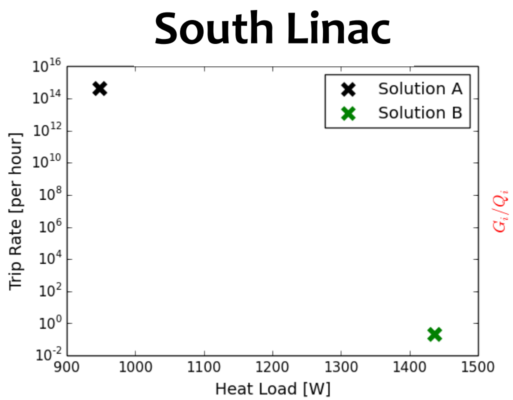
Solution A:
Minimize Heat Load
(Disregard Trip Rates)
Heat Load ~ 1015 W
Trip Rate ~ 6x10⁹ per hour

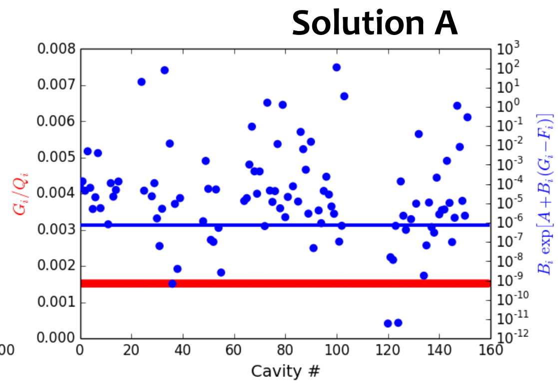


Solution B:
Minimize Trip Rates
(Disregard Heat Load)
Heat Load ~ 1405 W
Trip Rate ~ 0.74 per hour

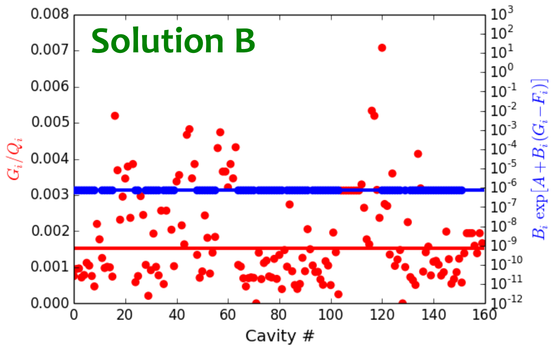
1 Obj Minimization Using Lagrange Multipliers

- Single-objective (1D) analytical solutions with Lagrange multipliers are pedagogic, but also somewhat useful
  - *Give us the limits of the optimization*





Solution A:
Minimize Heat Load
Disregard Trip Rates
Heat Load ~ 948 W
Trip Rate ~ 4x10^{14} per hour

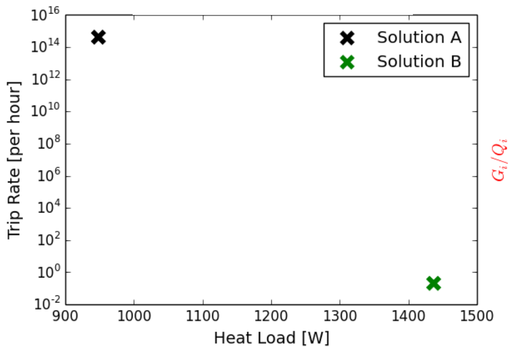


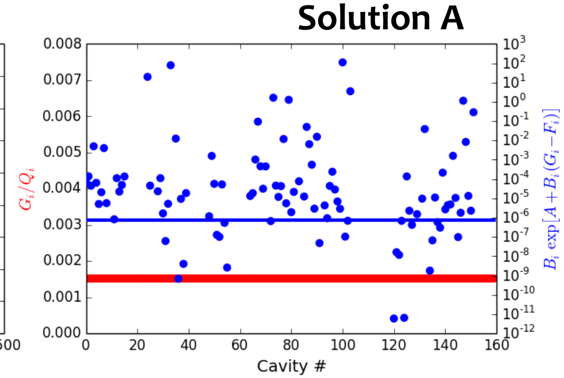
Solution B:
Minimize Trip Rates
(Disregard Heat Load)
Heat Load ~ 1437 W
Trip Rate ~ 0.2 per hour

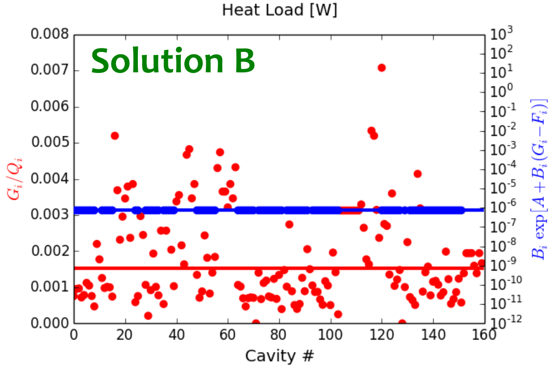
1 Obj Minimization Using Lagrange Multipliers

- Single-objective (1D) analytical solutions with Lagrange multipliers are pedagogic, but also somewhat useful
  - *Give us the limits of the optimization*

South Linac

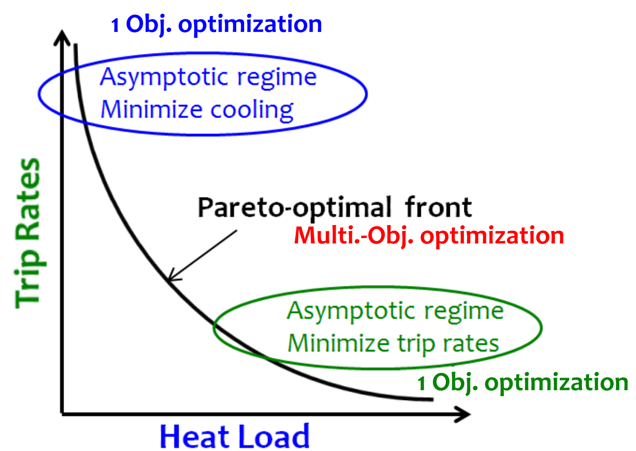






Solution A:
Minimize Heat Load
Disregard Trip Rates
Heat Load ~ 948 W
Trip Rate ~ 4×10^{14} per hour

Solution B:
Minimize Trip Rates
(Disregard Heat Load)
Heat Load ~ 1437 W
Trip Rate ~ 0.2 per hour



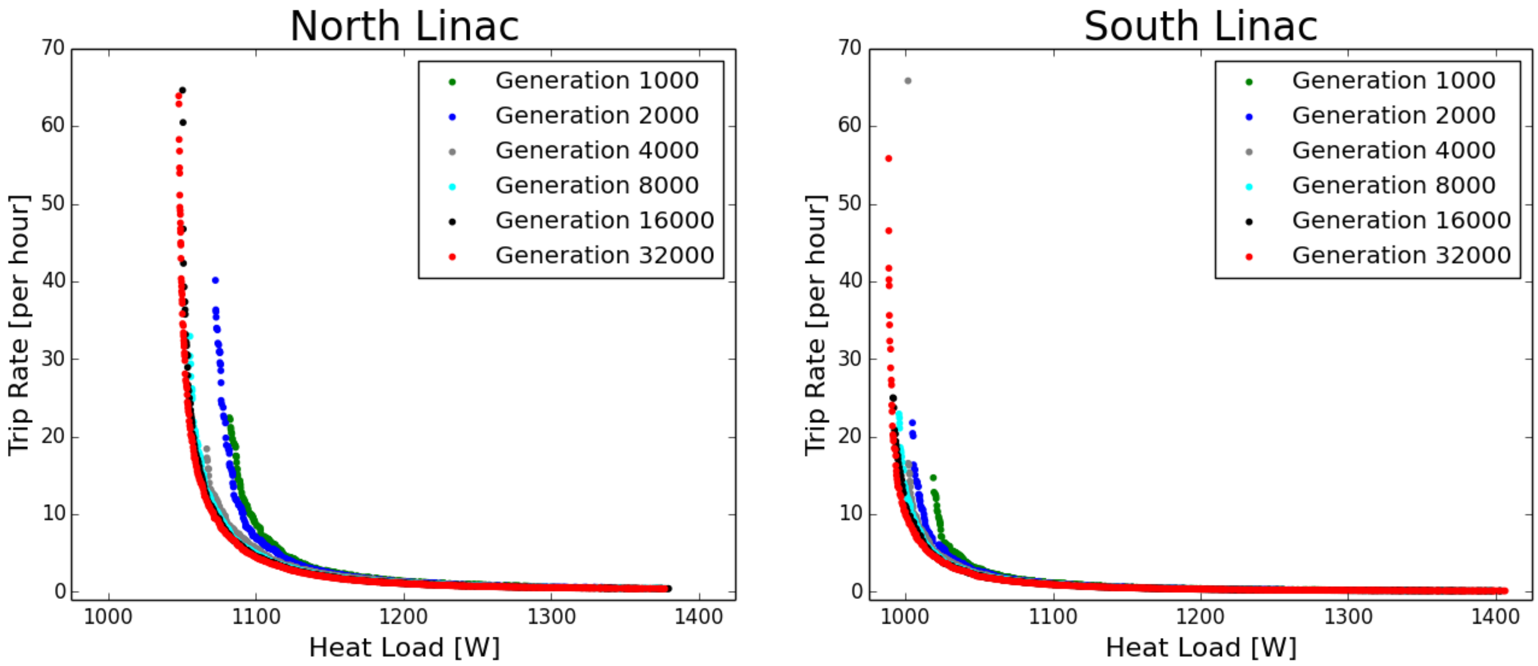
Numerical Optimization Method: Genetic Algorithm

- This is a *high-dimensional*, *non-linear*, *multi-objective* optimization problem
- Traditional, gradient-based methods (Newton, conjugate-gradient, steepest descent, etc…) are *not globally convergent*:
    - Get stuck in a local minimum and never come out
    - Final solution depends on the initial guess
- Genetic algorithm (GA) is what is needed here: *globally-convergent, multidimensional, multi-objective, robust, non-linear optimization*
- Platform and Programming Language Independent Interface for Search Algorithms (*PISA*) from ETH Zürich and Alternate *PISA* (*APISA*) from Cornell
- We used GAs before on a number of problems in accelerator physics [Hofler, Terzić, Kramer, Zvezdin, Morozov, Roblin, Lin & Jarvis 2013, *PR STAB* 16, 010101]
- Heat load & trip rate optimization by GA is published

  [Terzić, Hofler, Reeves, Khan, Krafft, Benesch, Freyberger & Ranjan 2014, *PR STAB* 17, 101003]

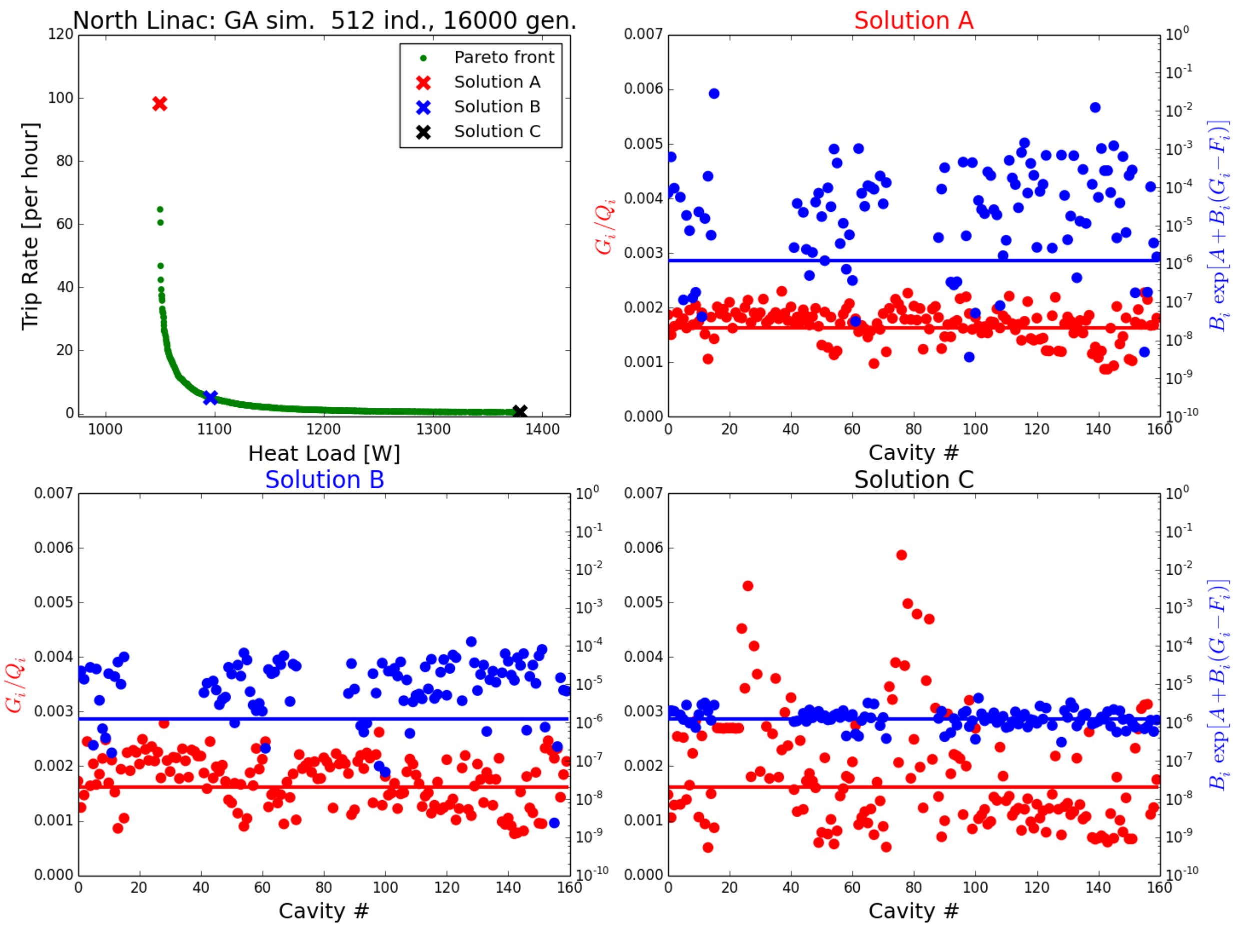
Multi-Objective GA Minimization: Results

- GA simulation: 512 ind. per gen. on MacBook Pro 2.7 GHz Intel Core i7
- Pareto-optimal front – textbook behavior
  - Longer simulation, more generations – better results (front creeps left)
  - Execution time rough estimates: 3 minutes per 4000 generations



Multi-Objective GA Minimization: Results

- GA simulation: North Linac, 512 ind. per gen., 16000 generations



Solution A (1D):
Minimize Heat Load

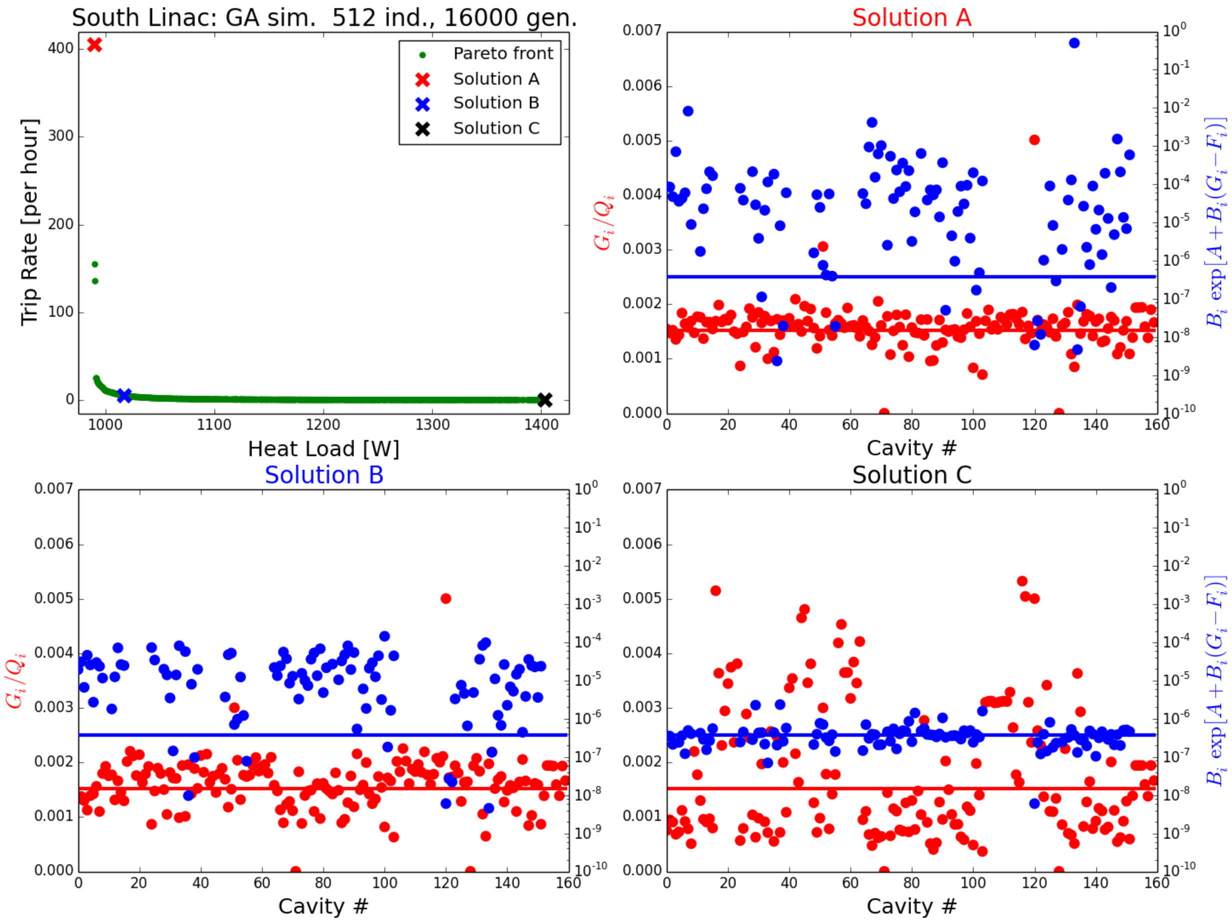
$$G_i/Q_i = const. = 1.62 \times 10^{-3}$$

Solution C (1D):
Minimize Trip Rates

$$B_i \exp\left[A + B_i\left(G_i - F_i\right)\right] = const.$$
$$= 1.24 \times 10^{-6}$$

Multi-Objective GA Minimization: Results

- GA simulation: South Linac, 512 ind. per gen., 16000 generations



Solution A (1D):

Minimize Heat Load

$$G_i/Q_i = const. = 1.52 \times 10^{-3}$$

Solution C (1D):

Minimize Trip Rates

$$B_i \exp\left[A + B_i\left(G_i - F_i\right)\right] = const. = 3.68 \times 10^{-7}$$

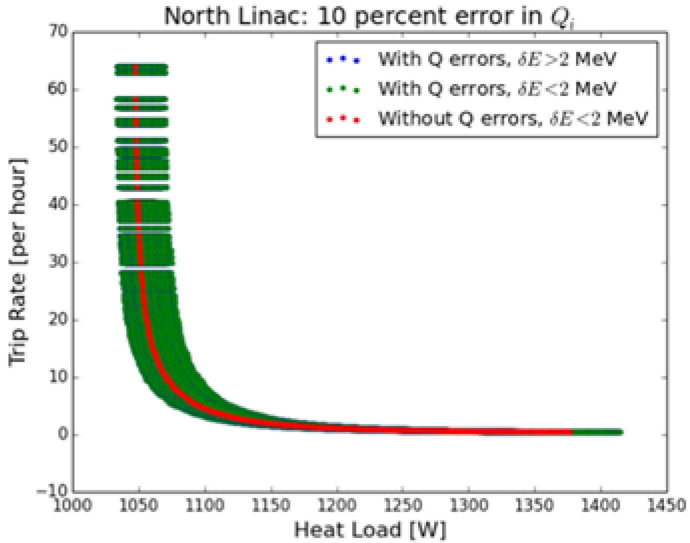
Summary of Previous Work

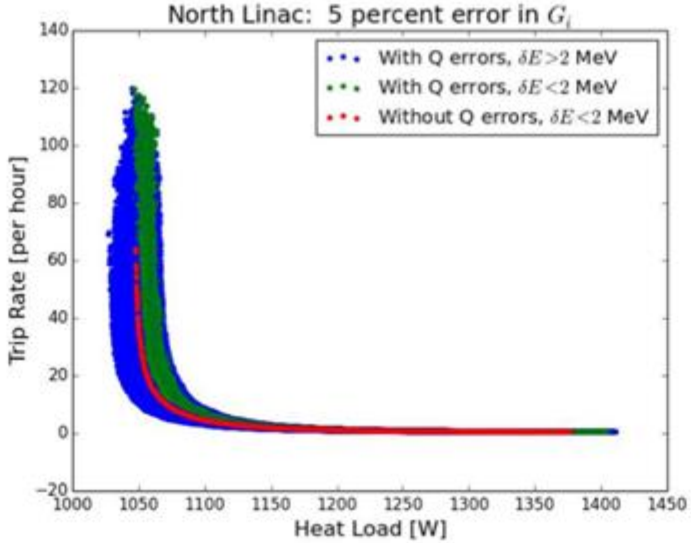
•Simultaneous minimization of the heat load and trip rates using GA

- Provides an *entire Pareto-optimal front of solutions*
- Performance of C++ prototype:

  Full simulation (32k gen.): < 30 min. "Quick peek" (4k gen.): ~ 3 min.
- Made contact with 1D minimization using Lagrange multipliers
- Made contact with Arne's first GA implementation (fix TR, minimize HL)

  For TR=5/hour, 13% lower heat load by multi-objective optimization
- Robust for errors in Q_i and G_i





Current & Future Work

- Current work
  - Developing user friendly GA package in C++ (A. Holfler & A. Carpenter)

Problems (Previous GA system)	Solution (Standalone GA library)
Suitable for Propotyping: • Originally developed as GA test bed • Inefficient process management Cumbersome to maintain and use • Multiple versions with different capabilities • Not well documented GA processing entwined in the system • Not easily extracted or repurposed • GAs not available for general use outside the system	• Available for studies and control room applications • Software development cycle: Written requirements, system design, design review, and user documentation • Support GAs most often used in accelerator physics applications: SPEA2&NSGA_II* • Easy to configure and use • Option to support particle swarm

* Strength Pareto Evolutionary Algorithm 2 (SPEA2)
Nondominated Sorting Genetic Algorithm II (NSGA-II)

  - Investigating particle swarm (H. Zhang)
- Future work
  - $Q_i(G_i)$ for all cavities
  - Compare GA with particle swarm
  - Increase the efficiency by parallelization on modern hardware


谢谢
DANKSCHEEN
BIYAN
SHUKRIA
TINGKI
YAQHANYELAY
CHALTU
NUHUN
SPASSIBO
SNACHALHUYA
TASHAKKUR ATU
THANK
YOU
GRACIAS
SUKSAMA
EKHMET
ARIGATO
SHUKURIA
GRAZIE
MEHRBANI
MAAKE
KOMAPSUMNIDA
GOZAIMASHITA
EFCHARISTO
JUSPAXAR
PALDIES
BOLZÏN
MERCI
MINMONCHAR
SIKOMO
SPASIBO
DENKAUJA
ATTO
MERASTAWHY
GAEJTHO
TAVTAPUCH
MEDAWAGSE
AGUYJE
FAKAAUE
LAH
WABEEJA
MAITEKA
HUI
GUI

Backup Slides

* Strength Pareto Evolutionary Algorithm 2 (SPEA2)
Nondominated Sorting Genetic Algorithm II (NSGA-II)

6 GeV-Era Simulation with 12 GeV Consequences

- We model PVDIS Run from 2009 to make contact with earlier work
- This approach is not tied to a particular configuration
- Model for trips in old cavities given in Benesch *et al.* 2009 JL-TN-09-41
- lem.dat file provides all information needed for the simulation

No trip model — Parameters used in the simulation

160 rows, 1 per cavity

Name	Loaded Q	$DRVH_i$	PASKsigma	F_i [MV/m]	B_i	Q_i	L_i [m]
NL02-1	6.23	12	0.060	11.37	1.52	6.20E+009	0.5
NL02-2	6.61	8.7	0.052	7.64	2.26	5.70E+009	0.5
NL02-3	5.55	8.5	0.063	8	2.15	4.90E+009	0.5
NL02-4	5.41	8.5	0.050	99	0	3.30E+009	0.5
NL02-5	6.84	8.5	0.056	7.68	1.91	4.40E+009	0.5
NL02-6	8.22	11.5	0.071	11.1	1.04	3.10E+009	0.5
NL02-7	5.67	8.5	0.065	7.66	1.84	4.30E+009	0.5
NL02-8	5.7	9.25	0.057	8.4	1.83	3.70E+009	0.5
NL03-1	8	10.5	0.054	7.07	1.74	2.10E+009	0.5
NL03-2	5.91	9.8	0.057	7.63	1.19	1.50E+009	0.5

- Same formalism will be used for the 12 GeV configuration whenever new Qs, DRVHs and *B*s become available for the new cavities

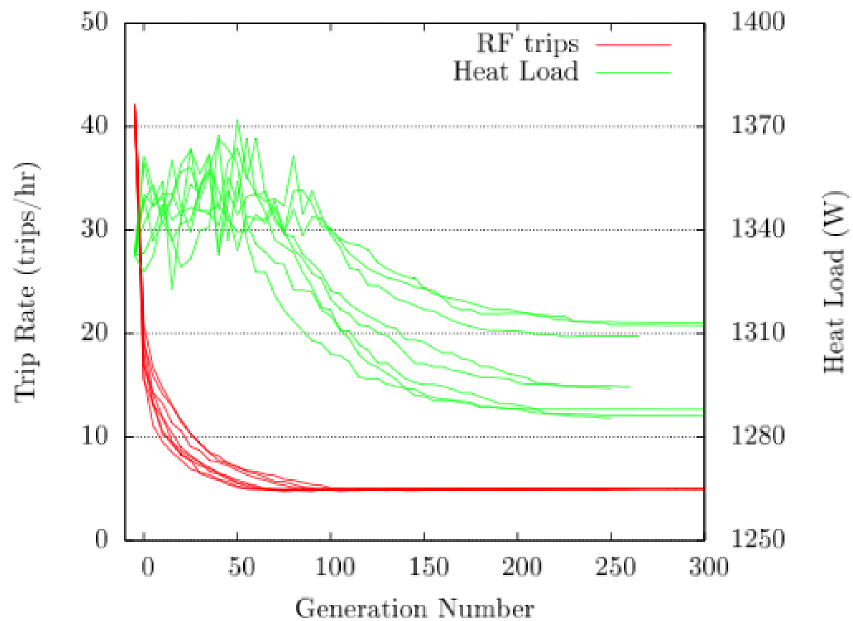
Earlier Work on the Subject: Arne's GA Simulation

- Used a *perl-based* GA algorithm (for details see JLAB-TN-12-057)
  - *perl* is an interpreted language → slow (> 1 day for 150 generations)
    - From the footnote – acknowledgement that we can do better: "Improvements in execution speed of the GA would be possible utilizing a compiled programming language."
  - *Arne's work provides an important proof-of-concept*
- Key differences between Arne's and this implementation

Arne's	this
• 1D optimization (minimize HL, TR fixed)	• 2D optimization (minimize both HL, TR)
• 90% of initial population of gradients is ±2 MV/m from initial value	• Unbiased sampling of the entire allowed search space [3, $DRVH_i$]
• Focused on the *premier individual* from each generation (top fitness)	• Provide a Pareto-optimal front of feasible solutions (enable trade-off)
• Interpreted perl	• Compiled C++

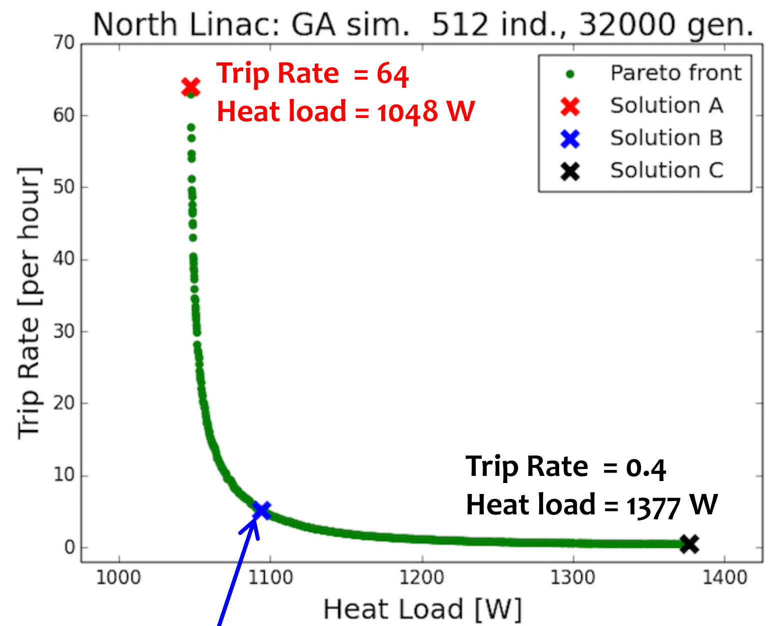
Comparison to Arne's Results: North Linac

Arne's Tech Note (Fig. 2)



Trip Rate = 5
Heat load = 1285 W

Our Study

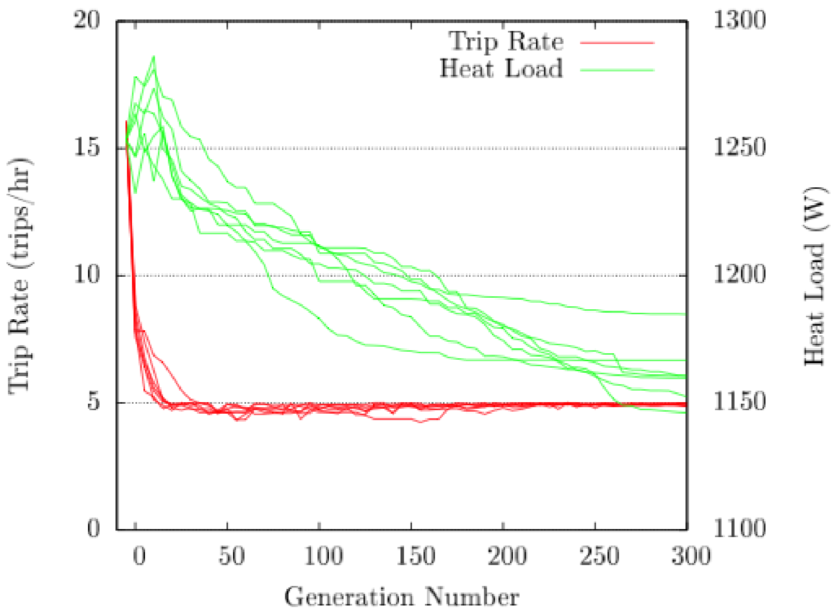


Trip Rate = 5
Heat load = 1094 W
(~4% from the minimum of 1048 W)

Reduced heat load by 15% in the North Linac

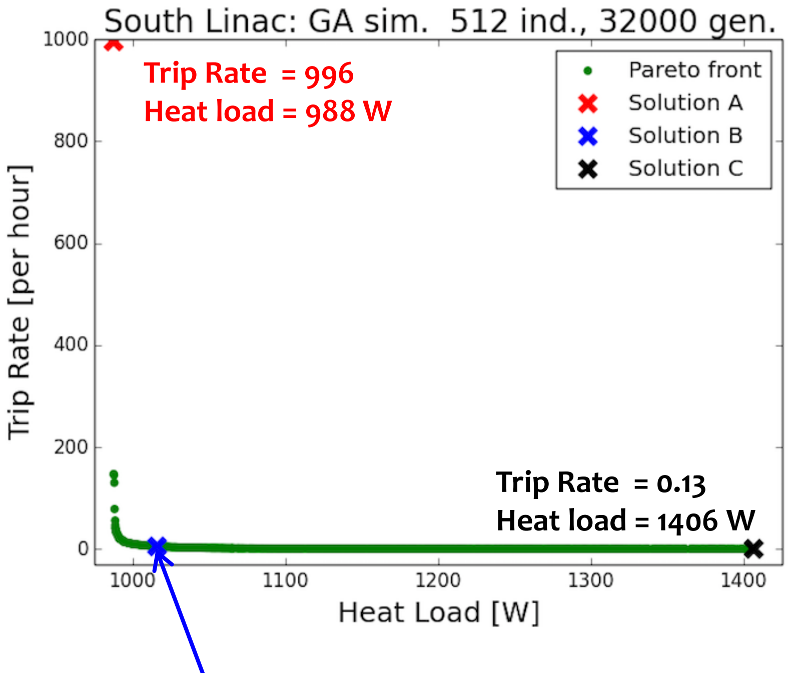
Comparison to Arne's Results: South Linac

Arne's Tech Note (Fig. 5)



Trip Rate = 5
Heat load = 1150 W

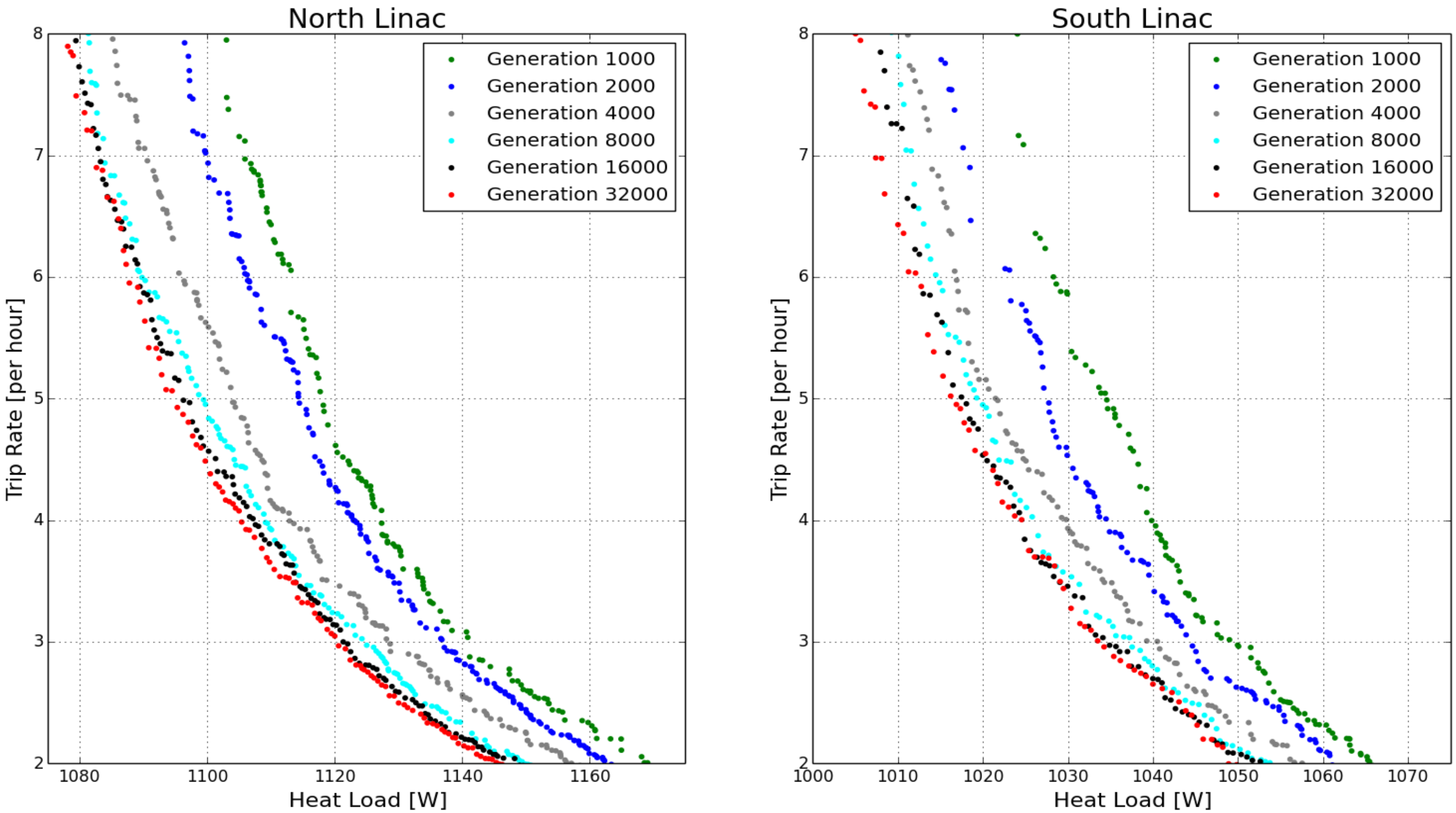
Our Study



Trip Rate = 5
Heat load = 1016 W
(~3% from the minimum of 996 W)

Reduced heat load by 12% in the South Linac

Convergence of the Pareto-Optimal Front



32000 Vs. 4000 NL ~ 1% (~10 W)
32000 Vs. 8000: NL ~ 0.5% (~5 W)
32000Vs. 16000: NL < 0.2% (<2 W)

32000 Vs. 4000: SL < 1% (<10 W)
32000 Vs. 8000: SL < 0.5% (<5 W)
32000 Vs. 16000: SL < 0.2% (<2 W)

Sensitivity to Measurement Error