Dynamic Aperture Optimization of storage ring based colliders with Multi-Objective Algorithm

ZHANG, Yuan  WANG, Yiwei  WANG, Dou  GENG, Huiping

ICFA Mini-Workshop on Dynamic Aperture of Circular Accelerator, 1-3 November 2017, IHEP, Beijing, China

Many Thanks: J. Qiang(LBNL), K. Oide(KEK), D. Zhou(KEK)
Outline

• Introduction
• Algorithm
• Application
• Summary
Multi-objective genetic algorithm (MOGA)

- Application in storage ring based light source is very popular and successful
  - APS/DLS, ELEGANT, M. Borland, in 48th ICFA Beam Dynamics Workshop on Future Light Sources
  - NSLSII, L. Yang, Y. Li, W. Guo and S. Krinsky, PRST-AB, 14, 054001 (2011)
  - SLS, BMAD, M. Ehrlichman, arXiv: 1603.02459
  - HEPS, Accelerator Toolbox, Y. Jiao and G. Xu,
  - ...
Different Algorithm

• Particle Swarm, SPEAR3, X. Huang, J. Safranek, NIMA 757, 48, 2014
• Differential Evolution, J. Qiang et al., IPAC’13
• Downhill Simplex, SuperKEKB, FCC, K. Oide et al.
• .......
Excitation

- K. Oide, “A design of beam optics for FCC-ee”, Sep. 2015 @IHEP

255 sextupole pairs per half ring

Resulting dynamic aperture almost satisfies the requirements

- Basically a 90 degree FODO cell.
- QFs are longer (3 m) than QDs (1.3 m) to mitigate the radiation, as discussed later.
- All sextupoles are paired with f transformation.

- 255 sextupole pairs per half ring

E = 175 GeV, $\beta^*_y = 2$ mm
E = 45.6 GeV, $\beta^*_y = 1$ mm

* The dynamic aperture was optimized with element-by-element radiation damping, automatic tapering, and crab waist.
Why we did the job?

• We need to optimize the DA of CEPC
• We want to try the direct DA optimization in collider, just as the community has done in light source
• Different optimization algorithm is worth to be tried
• SAD(http://acc-physics.kek.jp/sad/) is used for the DA determination. It is a parallel code, but the scalability is not very good. A MPI-based parallel code to call SAD will be much more efficient.
Differential Evolution Algorithm (single objective)

- The “DE community” has been growing since the early DE years of 1994 – 1996
- DE is a very simple population based, stochastic function minimizer which is very powerful at the same time.
- There are a few strategies, we choose ‘rand-to-best’. Attempts a balance between robustness and fast convergence.

\[
v(i,j) = \begin{cases} 
  x(i,j) + F \times [x(b,j) - x(i,j)] + F \times [x(r1,j) - x(r2,j)], & \text{If } \text{rand}(j) < CR \\
  x(i,j), & \text{Otherwise}
\end{cases}
\]

- Different problems often require different settings for NP, F and CR

http://www.hindawi.com/journals/ijap/2013/713680.fig.003.jpg
Multi-objective Optimization

• Most problems in nature have several (possibly conflicting) objectives to be satisfied.

• Many of these problems are frequently treated as single-objective optimization problems by transforming all but one objective into constraints.

• The term optimize means finding such a solution which would give the values of all the objective functions acceptable to the decision maker.

Kung et al., J. ACM 22, 4 (Oct. 1975), 469-476

MODE: Multi-Objective optimization by Differential Evolution

The parallel algorithm is referencing to J. Qiang (IPAC’13)

1. Initialize the population of parameter vectors
2. Generate the offspring population using the above differential evolution algorithm
3. Find the non-dominated population, which are treated as the best solutions in DE to generate offspring
4. Sorting all the population, select the best NP solution as the parents
5. Return to step 2, if stopping condition not met
MODE: Scalable Enough at 1000-nodes farm?

Courtesy of Yongjun Li (BNL)
New Parallel Paradigm

*High Parallel + High Scalability*

- Even the time taken by different task is different
- Even some node is very busy
Definition of Objective Cost Value

• DA Boundary:
  - \( \frac{x^2}{20^2} + \frac{z^2}{16^2} = 1 \) (example)
  - \( z \) for energy deviation in unit of \( \sigma_p \)
  - \( x \) for transverse amplitude in unit of \( \sigma \)
  - transverse coupling: designed value

• The difference between the DA boundary and real DA is defined the objective cost value

• Less cost value is better
SuperKEKB:
dynamic aperture is a serious issue


Difficulty in the Nano-Beam scheme

w/o beam-beam

LER

$\frac{2J_y}{2J_x} = 0.27 \%$

HER

$\frac{2J_y}{2J_x} = 0.25 \%$

with beam-beam (W-S)

LER

aperture lost

Transverse aperture is reduced significantly.

HER

aperture lost
DA Optimization of LER (SuperKEKB)

• Objectives:
  • \( \nu_x \in (0.53, 0.66), \nu_y \in (0.55, 0.66), \) for \( \delta_p \in (-0.019, 0.019) \)
  • \( \frac{x^2}{50^2} + \frac{z^2}{26^2} = 1, \) for \( z=\text{Range}[-24,24,3], \)
    \( \epsilon_{x,0} = 1.89 \text{ nmrad}, \delta_{p,0} = 7.7e^{-4} \)
• Variables: 68
  • 2 Octupoles
  • 54 sextupole pairs
  • 12 skew sextupole pairs
CEPC DA Optimization Knobs

50 knobs in total

- IR sextupoles: (10)
- Arc Sextupole (32)
- Phase advance (8)
IR knobs

- Main Chromaticity Sextupoles (2)
- Neighbor weaker sextupole to correct finite length effect (2)

A. Bogomyagkov, ArXiv:0909.4872
- Sextupole to correct higher order chromaticity in vertical direction (1)
Y. Cai, PRAB.19.111006

Different strength between Upstream and Downstream of IP.
10 knobs in IR.
Arc sextupole

- 90/90 FODO
- Non-interleave sextupole scheme
- 4 SF + 4 SD sextupole configurations in one arc section
- 7 sub-period in one arc section
- 4 arc section in half ring

Total knobs: 32
Phase Advance Tuning

• 10 knobs in x/y direction
• Keep tune fixed
• Only 8 free knobs
Contribution from phase advance
Some Try of Speed-up/Optimization Method

Brute-force dynamic aperture tracking is very time consuming

• More strategy of DE algorithm are randomly selected used

• The objective is first eased, for example only track 50 turns instead of 100 or 200 turns.

• The time consuming cost function is calculated only when the necessary constraints be satisfied. [arXiv: 1603.02459]

• First try to optimize with less variables, then more variables.

• Iteration with non-dominated solution
Optimization in one-step vs multiple steps (32 arc sextupole family)

**STEP-BY-STEP:**

- **4 Arc Sextupoles:** SFO/SDO/SFI/SDI
- **16 Arc Sextupoles:** 4*(SFO/SDO)+4*(SFI/SDI)
- **32 Arc Sextupoles:** 8*(SFO/SDO)+8*(SFI/SDI)

If computing resources are enough, it’s better to optimize all the variables in one step.
Iteration with Pareto Front Solution

- Non-dominated solution achieved + Other initialized solutions are randomly generated
- No very clear further optimization
- We usually do the iteration if time permits

$B_x = 0.37m$

$B_x = 0.50m$
Optimize with different population size

$N_p=1200\text{~}1800$ (20-30 times variable number) is good enough
Different Chromaticity Constraint

Half ring tune control:
Qx in (0.52, 0.58)
Qy in (0.58, 0.64)

In our normal optimization, we prefer to control the tune in the maximum momentum offset (0.006~0.008)

The final optimum momentum acceptance is about 0.016.

K. Oide (2016):
Less chromaticity ≠ Better dynamic aperture
Contribution of $\beta^*_x$
Contribution of $\beta_\chi^*$ (2)
More sextupole configurations

Enlarge the DA about 1-σ on average
Optimize with DAMPONLY model

Damp at each element: No diffusion coming from synchrotron radiation.

200 turns,
10 samples
Fluctuation Effect on the vertical dynamic aperture

- On-Momentum DA with damping at each element and radiation fluctuation at each element.

Only serious DA loss in vertical direction.
Effect of Synchrotron Radiation in Quadrupoles

FODO Arc - Horizontal
• K. Oide, PRAB 19.111005
• Maximum momentum deviation
  \[ \Delta p = \frac{\alpha_z}{2 \pi v_s J_z} R_Q n^2 \epsilon_x \exp(-\frac{\alpha_z}{4v_s}) \]
  \[ n \equiv \Delta x/\sigma_x \]
  \[ R_Q = \frac{2 \sqrt{2}}{\theta_c^2} \left( \frac{\sqrt{2}+1}{l_{QF}} + \frac{\sqrt{2}-1}{l_{QD}} \right) \]
  \[ \alpha_z \text{ and } J_z, \text{ the synchrotron damping rate} \]
  \[ \text{and longitudinal damping partition number. } l_{QF,QD} \text{ the lengths of the quadrupoles.} \]

IR QD0 – Vertical
• A. Bogomyagkov (BINP), FCC-ee, Z

\[
\begin{align*}
J_y &= 0 \\
J_y &= \frac{U_0}{U_0(\sigma_y)} \frac{4 \sigma_{\gamma,y}^2 k_y^2}{\cos(2\psi_q) \left( \frac{\nu_y}{\nu_y} \right)^2 \left( \frac{-1+(j_y/2)^2}{j_y} \right) \propto \frac{K_0^2}{K_L^2 L_d}} \\
J_y &\approx 50 \sigma_y
\end{align*}
\]
Momentum deviation versus amplitude

- This could explain why the DA with (FLUC) is flat versus $\delta$.
- For example, if we want to achieve $20\sigma@\delta = 0$, the DA should be also about $20\sigma@\delta = 0.01$.
Suppress noise of DA result with radiation fluctuation

- DA is tracked with different initial phase: 
  \( \left( 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2} \right) \) for different energy
- 10 more times survey for on-momentum particle is tracked, and the minimum value is treated as the on-momentum DA
- Tracked DA result will be clipped to ensure DA at large momentum deviation will be less than that at small deviation
- Only two objective: min-DA of \((0, \pi)\) and min-DA of \(\left( \frac{\pi}{2}, \frac{3\pi}{2} \right)\)
On Momentum Dynamic Aperture (CEPC-Higgs)

Initial Phase: 0

Initial Phase: π/2

100 samples are tracked. 200 turns are tracked.

Synchrotron motion, synchrotron radiation in dipoles, quads and sextupoles, tapering, Maxwellian fringes, kinematical terms, crab waist are included.
Off momentum Dynamic Aperture

Momentum Acceptance: 0.017

Error bar means min and max

90% survival

100 samples. Radiation fluctuation is included. 0.3% emittance coupling. 200 turns are tracked.
Chromaticity
Crab Waist

\[ K_2 = \frac{(-1)^m}{2\theta} \sqrt{\frac{\beta_{x,IP}}{\beta_{x,S} \beta_{y,S} \beta_{y,IP}}} \]

- 90% survival

IP Downstream: \( K_2 = 0.69 \, \text{m}^{-2} \)

IP Upstream: \( K_2 = -0.69 \, \text{m}^{-2} \)
Summary

• Mode is developed for CEPC DA optimization
• The normal procedure of optimization is established
• All effects (exception: beam-beam, error, solenoid) is included in the dynamic aperture survey
• DA achieved: $20\sigma_x * 20\sigma_y$, 1.7% momentum acceptance
• backup
## CEPC Parameters

<table>
<thead>
<tr>
<th></th>
<th>Higgs</th>
<th>W</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of IPs</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy (GeV)</td>
<td>120</td>
<td>80</td>
<td>45.5</td>
</tr>
<tr>
<td>Circumference (km)</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR loss/turn (GeV)</td>
<td>1.68</td>
<td>0.33</td>
<td>0.035</td>
</tr>
<tr>
<td>Half crossing angle (mrad)</td>
<td>16.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piwinski angle</td>
<td>2.75</td>
<td>4.39</td>
<td>10.8</td>
</tr>
<tr>
<td>$N_e$/bunch ($10^{10}$)</td>
<td>12.9</td>
<td>3.6</td>
<td>1.6</td>
</tr>
<tr>
<td>Bunch number</td>
<td>286</td>
<td>5220</td>
<td>10900</td>
</tr>
<tr>
<td>Beam current (mA)</td>
<td>17.7</td>
<td>90.3</td>
<td>83.8</td>
</tr>
<tr>
<td>SR power/beam (MW)</td>
<td>30</td>
<td>30</td>
<td>2.9</td>
</tr>
<tr>
<td>Bending radius (km)</td>
<td>10.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Momentum compaction (10^{-5})</td>
<td>1.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_IP$ x/y (m)</td>
<td>0.36/0.002</td>
<td>0.54/0.0018</td>
<td>0.17/0.0029</td>
</tr>
<tr>
<td>Emittance x/y (nm)</td>
<td>1.21/0.0036</td>
<td>0.54/0.0018</td>
<td>0.17/0.0029</td>
</tr>
<tr>
<td>Transverse $\sigma_p$ (um)</td>
<td>20.9/0.086</td>
<td>13.9/0.060</td>
<td>7.91/0.076</td>
</tr>
<tr>
<td>$\xi/IP$</td>
<td>0.024/0.094</td>
<td>0.009/0.055</td>
<td>0.005/0.0165</td>
</tr>
<tr>
<td>RF Phase (degree)</td>
<td>128</td>
<td>134.4</td>
<td>138.6</td>
</tr>
<tr>
<td>$V_{RF}$(GV)</td>
<td>2.14</td>
<td>0.465</td>
<td>0.053</td>
</tr>
<tr>
<td>$f_{RF}$ (MHz) (harmonic)</td>
<td>650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nature bunch length $\sigma_z$ (mm)</td>
<td>2.72</td>
<td>2.98</td>
<td>3.67</td>
</tr>
<tr>
<td>Bunch length $\sigma_z$ (mm)</td>
<td>3.48</td>
<td>3.7</td>
<td>5.18</td>
</tr>
<tr>
<td>HOM power/cavity (kw)</td>
<td>0.46 (2cell)</td>
<td>0.32 (2cell)</td>
<td>0.11 (2cell)</td>
</tr>
<tr>
<td>Energy spread (%)</td>
<td>0.098</td>
<td>0.066</td>
<td>0.037</td>
</tr>
<tr>
<td>Energy acceptance requirement (%)</td>
<td>1.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy acceptance by RF (%)</td>
<td>2.06</td>
<td>1.48</td>
<td>0.75</td>
</tr>
<tr>
<td>Photon number due to beamstrahlung</td>
<td>0.25</td>
<td>0.11</td>
<td>0.08</td>
</tr>
<tr>
<td>Lifetime due to beamstrahlung (hour)</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F (hour glass)</td>
<td>0.93</td>
<td>0.96</td>
<td>0.986</td>
</tr>
<tr>
<td>$L_{amp}$/IP ($10^{14}$cm$^2$s$^{-1}$)</td>
<td>2.0</td>
<td>4.1</td>
<td>1.0</td>
</tr>
</tbody>
</table>