Double Harmonic Automatic Phasing

Authors: A. L. Pastinante, S. Albright, M. Marchi

Acknowledgements: A. Beeckman, S. F. Lauber, A. Lasheen, J. A. Wulff, BLonD Developers and PSB Operations

Outline

- Problem Statement
- Method
  - Objective Function
  - Bayesian Optimization
    - Scalar Optimization
    - Segmented Optimization
- PSB Tests
  - Scalar Optimization
  - Segmented Optimization
- Conclusion

Problem Statement

Double Harmonic Operation in the PSB:

Minimize space charge impedance by entering Bunch Lengthening Mode (BLM)

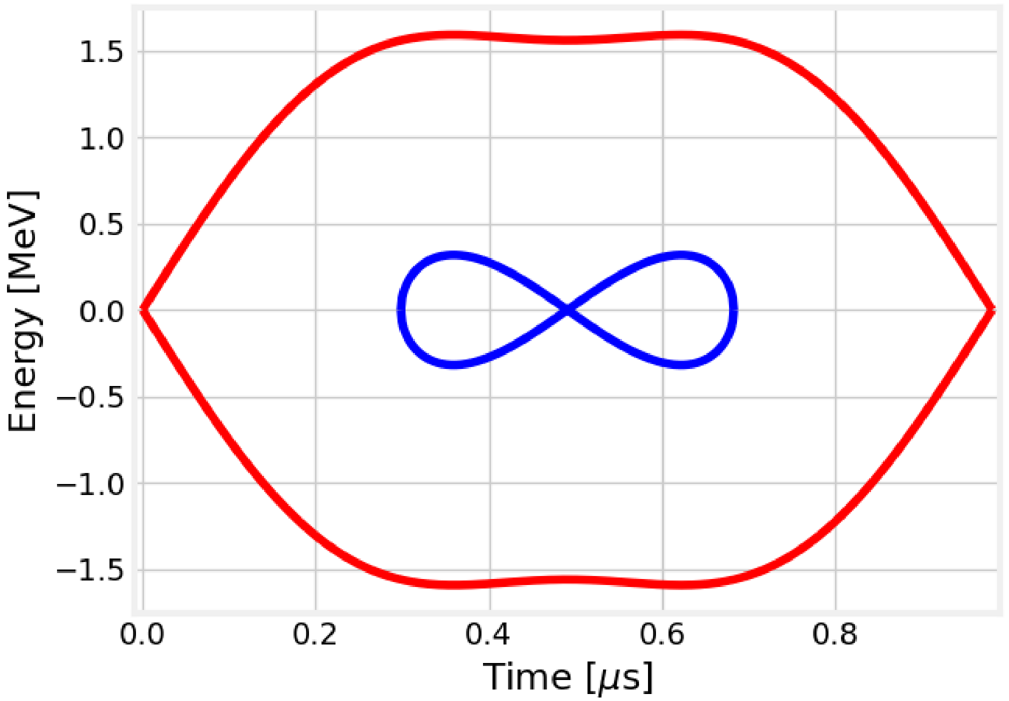
This requires correctly phasing the 2nd harmonic → Φ_2

Problem Statement

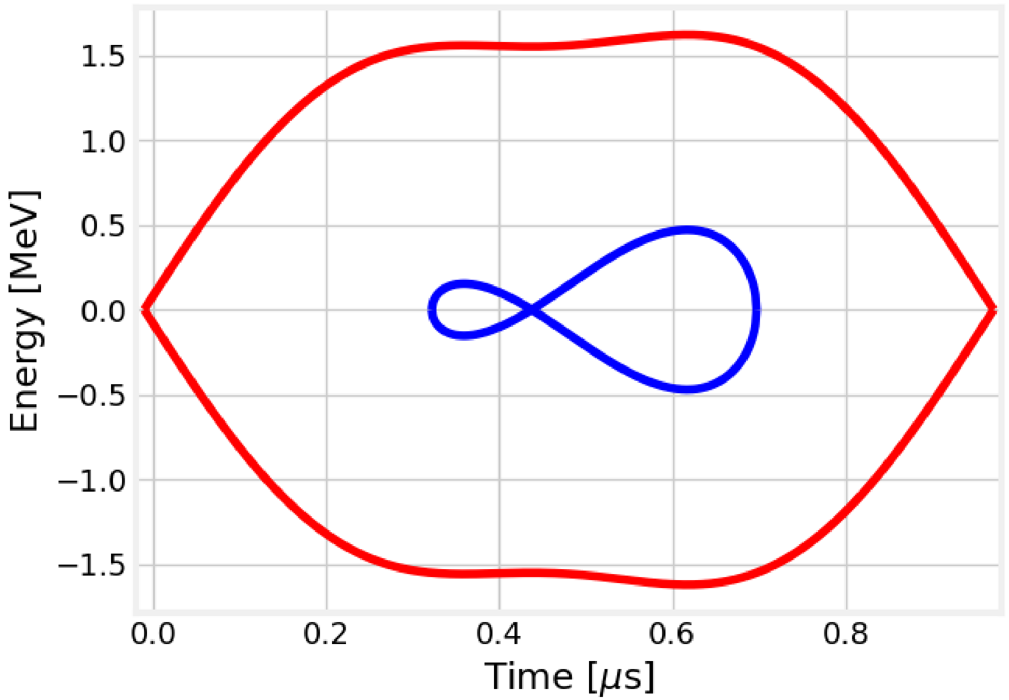
Double Harmonic Operation in the PSB:

Minimize space charge impedance by entering Bunch Lengthening Mode (BLM)

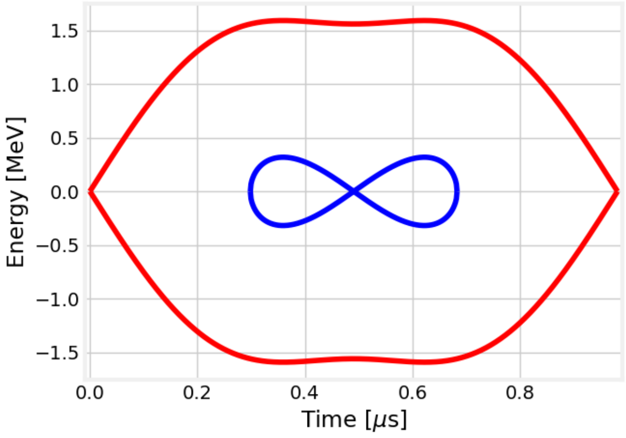
This requires correctly phasing the 2nd harmonic → Φ_2

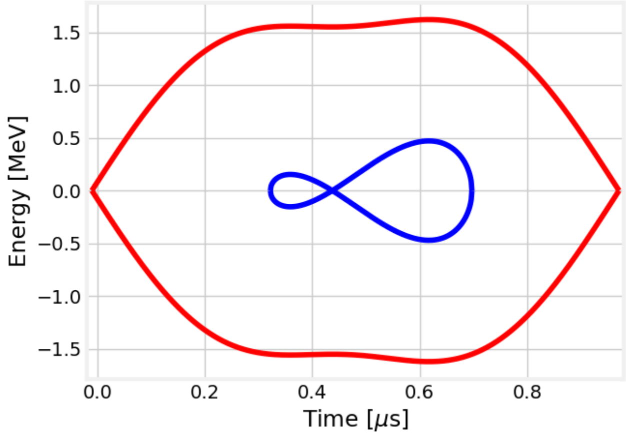


Separatrix Shape = $f(V_1,V_2,\Phi_2)$

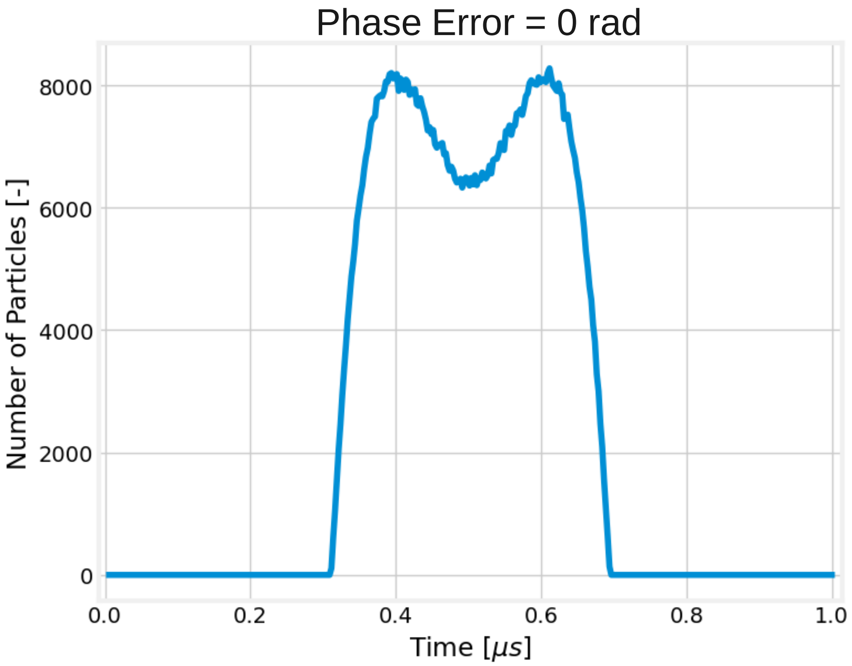


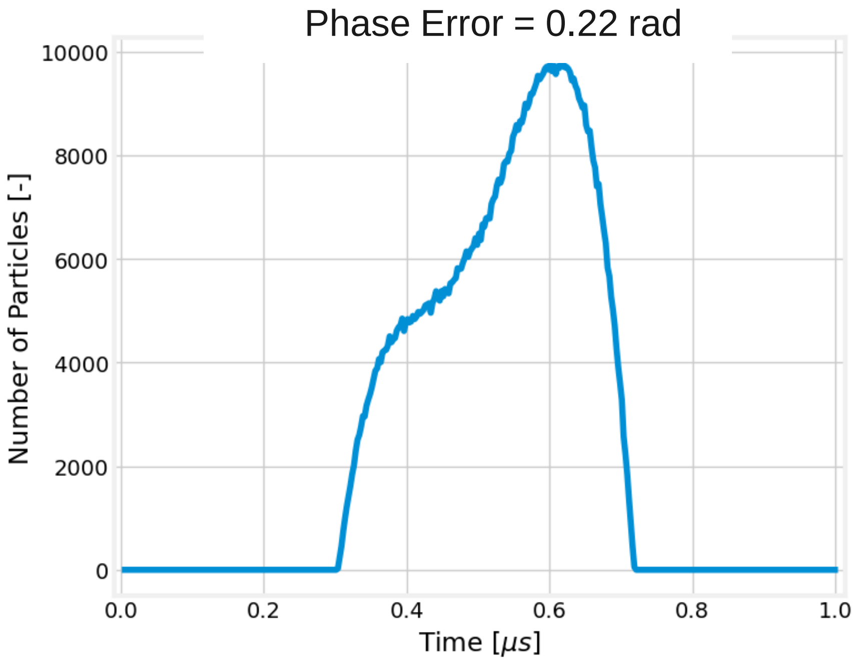
Problem Statement





Potential Well Shape
≈
Beam Profile Shape
=
$f(V_1, V_2, \Phi_2)$

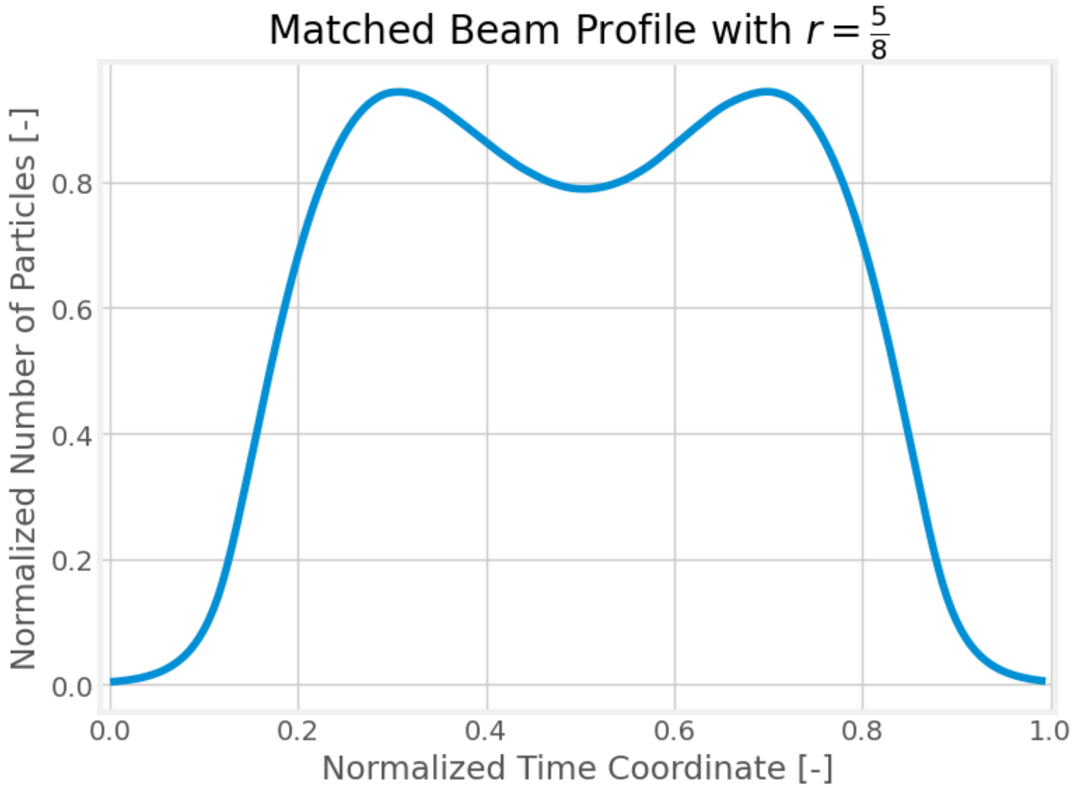




Method: Objective Function (minimization)

Method

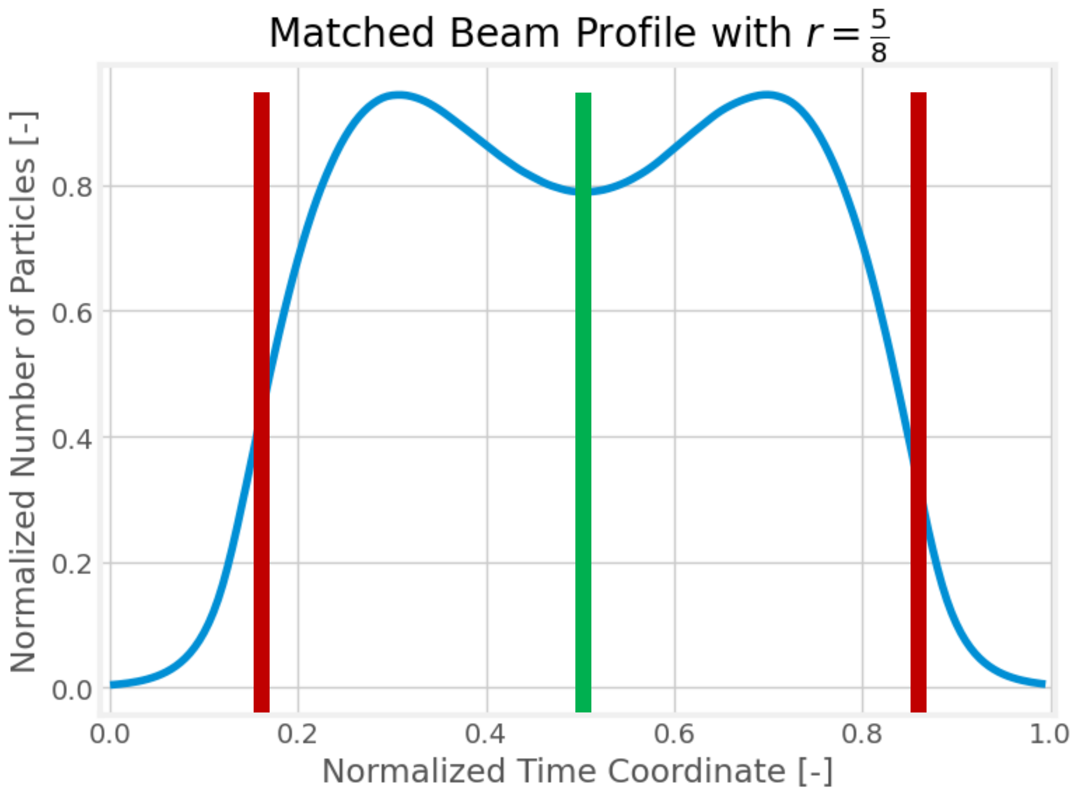
1. Filter and normalize profile



Method: Objective Function (minimization)

Method

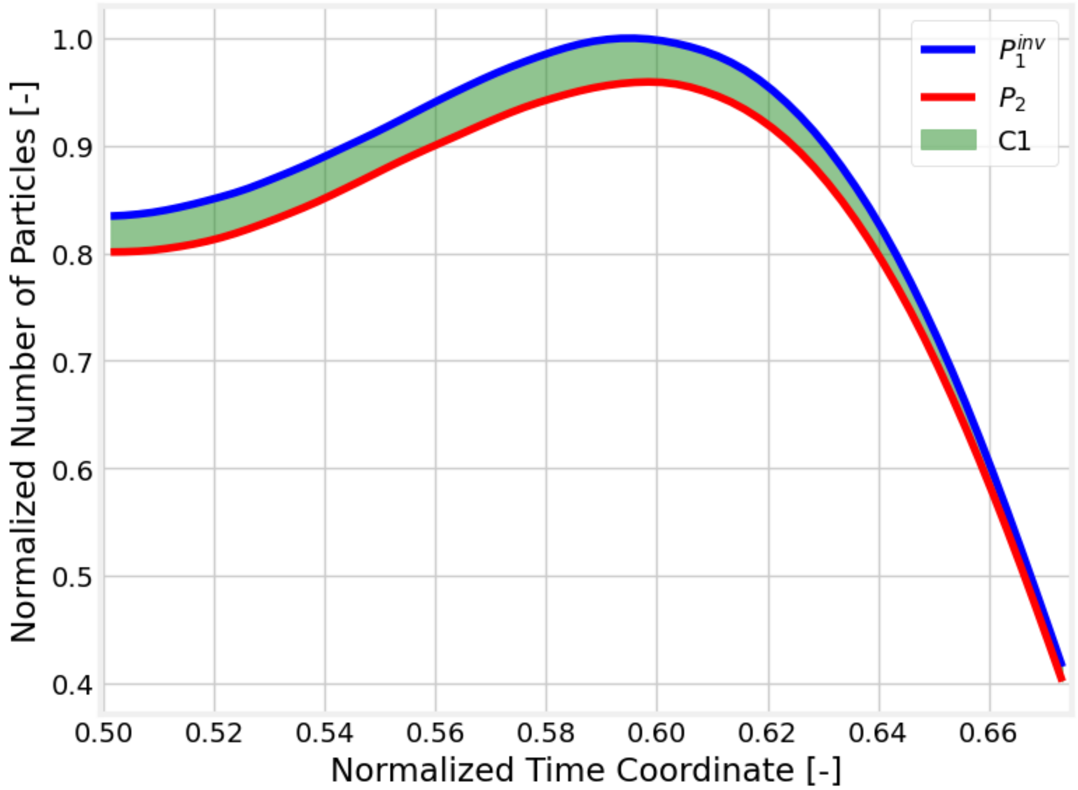
1. Filter and normalize profile
2. Cut at 95% of intensity and identify mid-point: part 1 (P_1) and part 2 (P_2)



Method: Objective Function (minimization)

Method

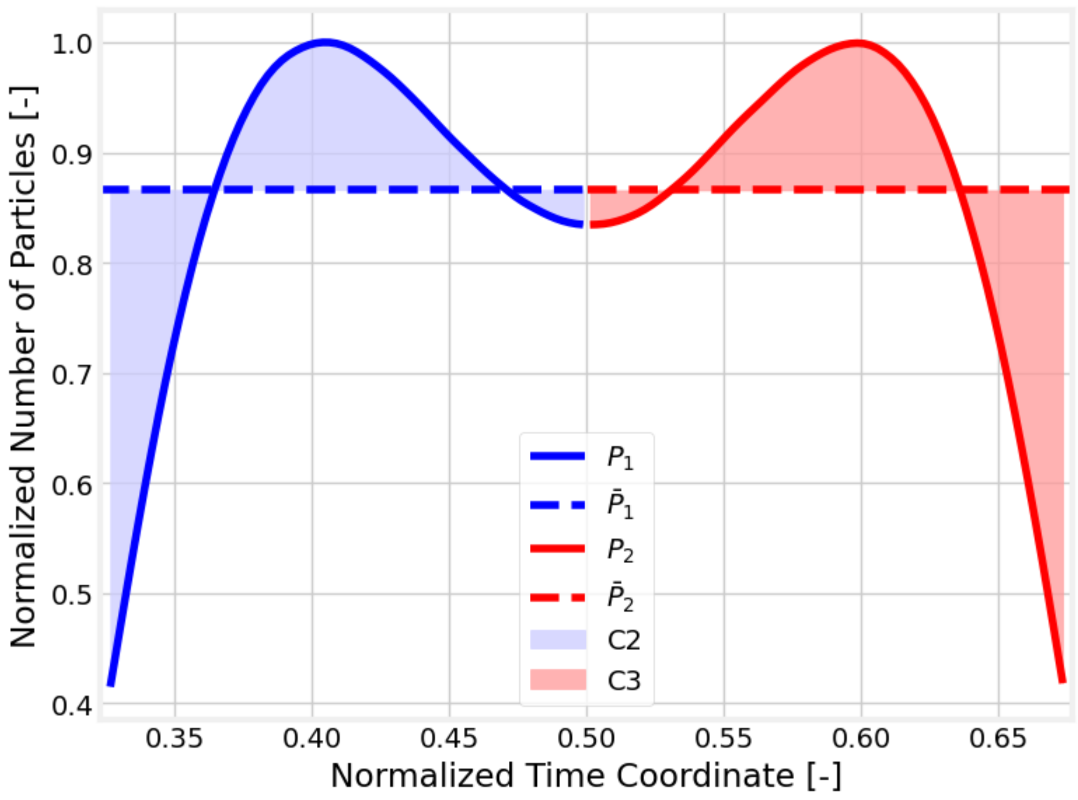
1. Filter and normalize profile
2. Cut at 95% of intensity and identify mid-point: part 1 (P_1) and part 2 (P_2)
3. Compare:
   - C1 = Flip P_1 and compute $|P_1^{inv} - P_2|$ (symmetry measure)



Method: Objective Function (minimization)

Method

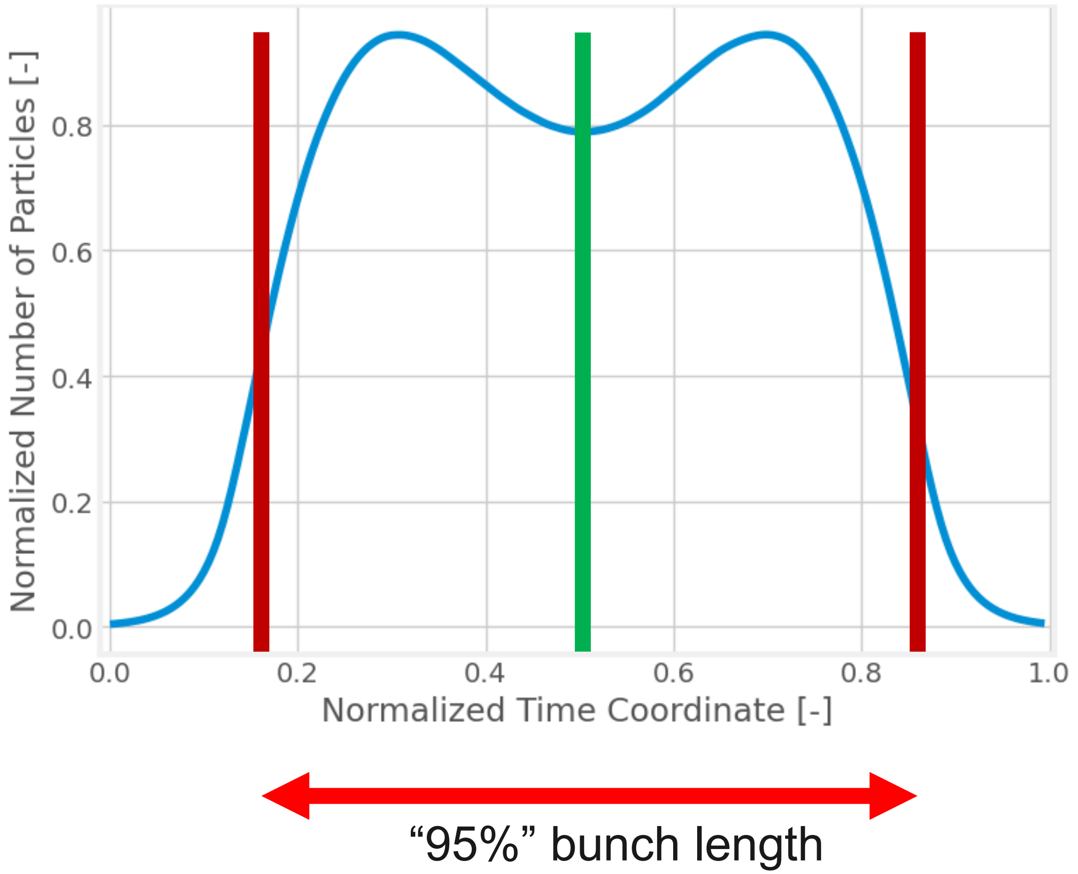
1. Filter and normalize profile
2. Cut at 95% of intensity and identify mid-point: part 1 (P_1) and part 2 (P_2)
3. Compare:
   - C1 = Flip P_1 and compute $|P_1^{inv} - P_2|$ (symmetry measure)
   - C2/3 = Compute the deviation of $P_{1/2}$ from its mean (measure of flatness)



Method: Objective Function (minimization)

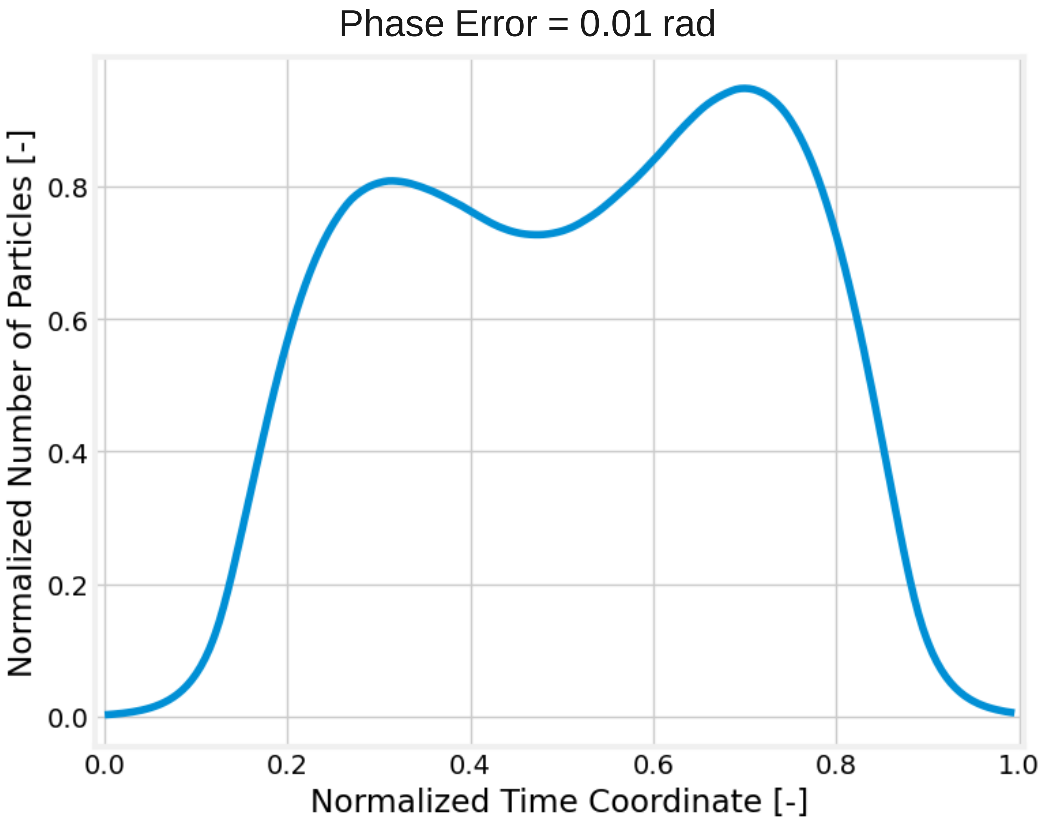
Method

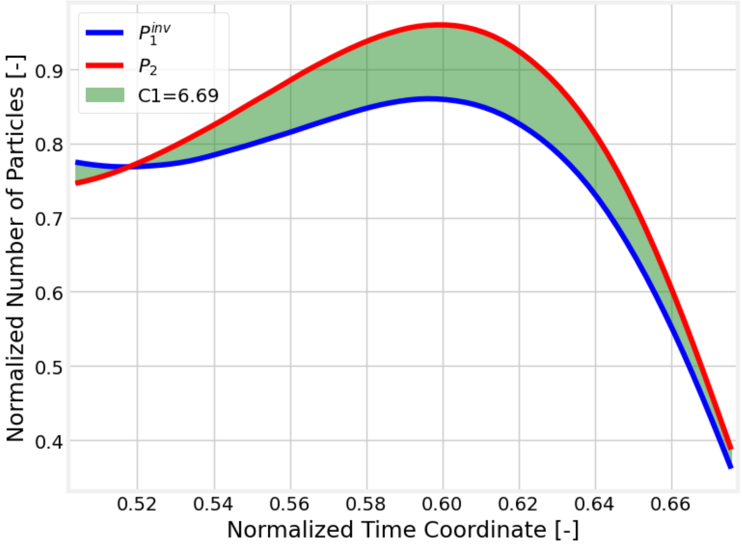
1. Filter and normalize profile
2. Cut at 95% of intensity and identify mid-point: part 1 (P_1) and part 2 (P_2)
3. Compare:
   - C1 = Flip P_1 and compute $|P_1^{inv} - P_2|$ (symmetry measure)
   - C2/3 = Compute the deviation of $P_{1/2}$ from its mean (measure of flatness)
   - BLF = Bunch length(95%)/ T_0
4. Compute:
   - $f(\Phi_2) = (C1 + C2 + C3)/BLF$

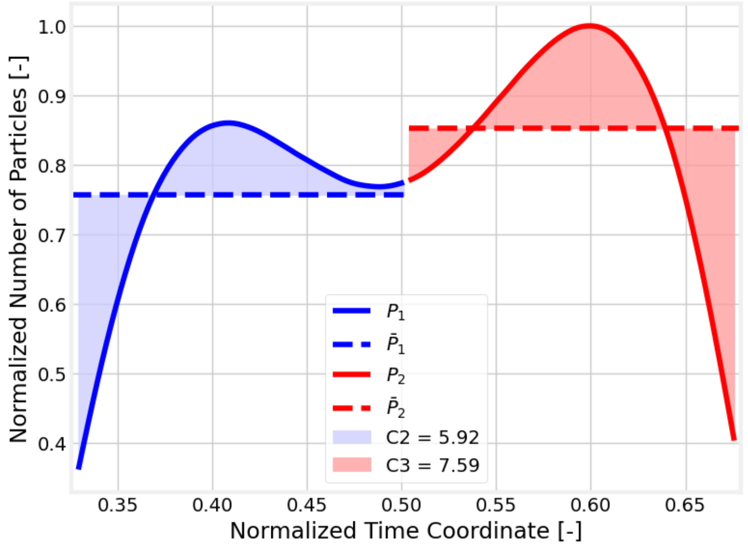


Method: Objective Function (Examples)

Wrong BLM Phasing

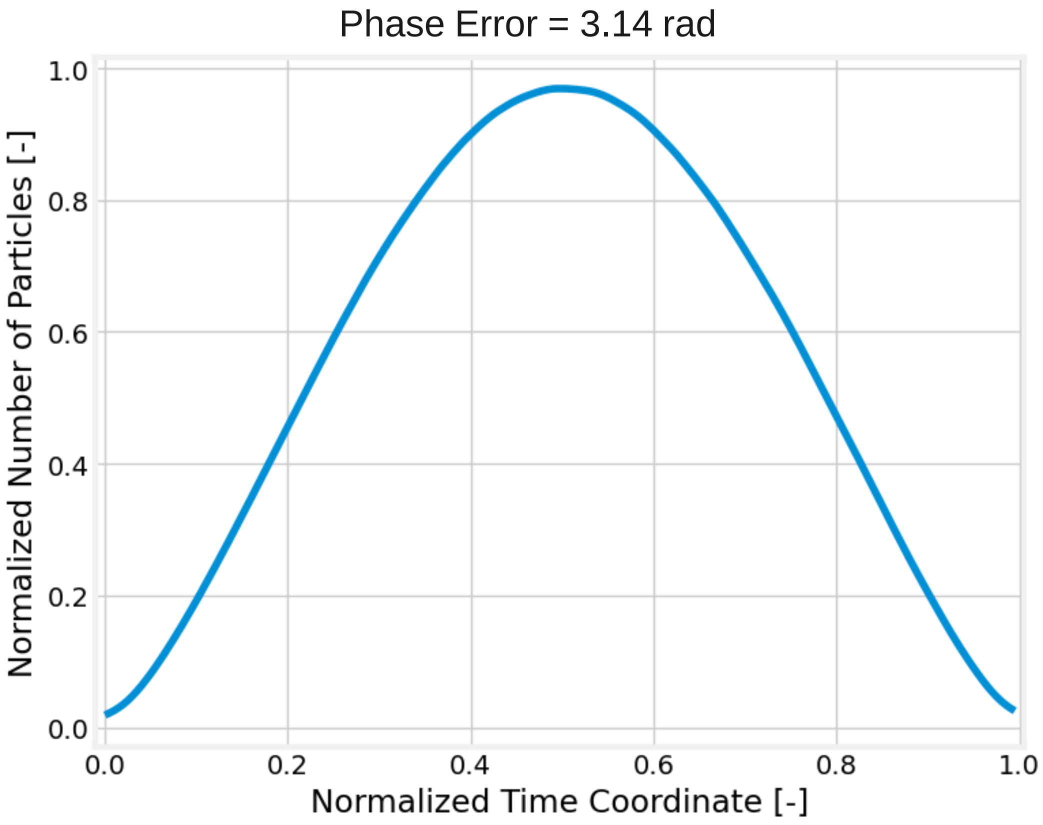


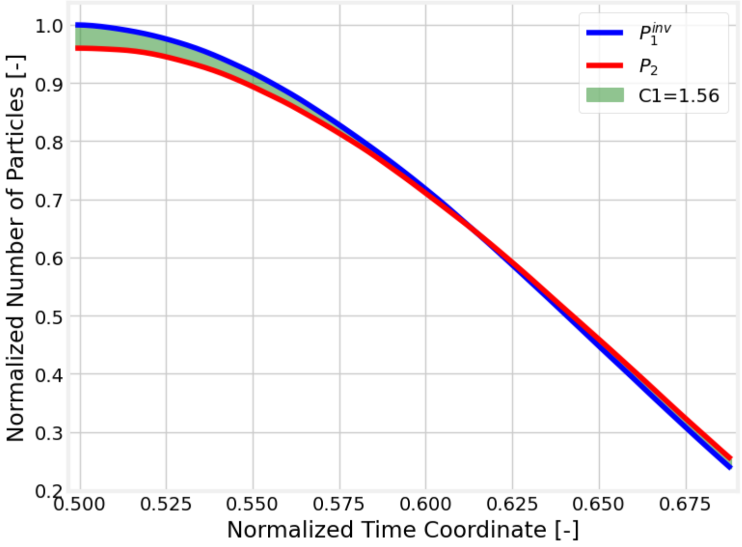


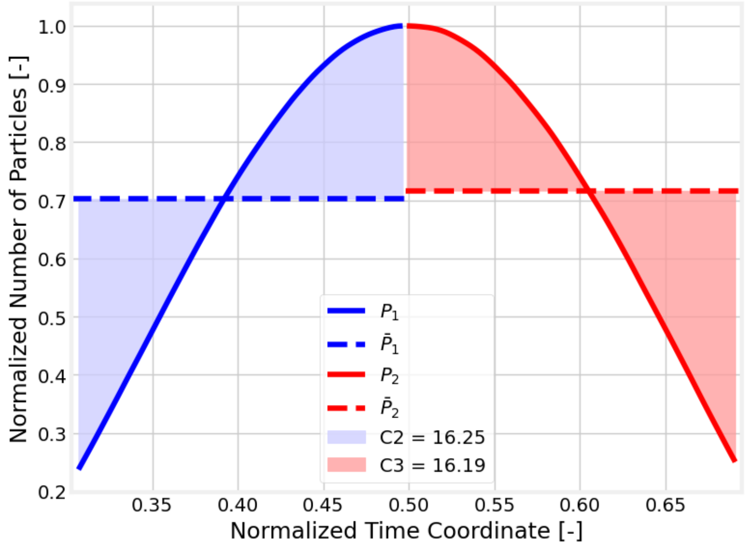


Method: Objective Function (Examples)

BSM Phasing

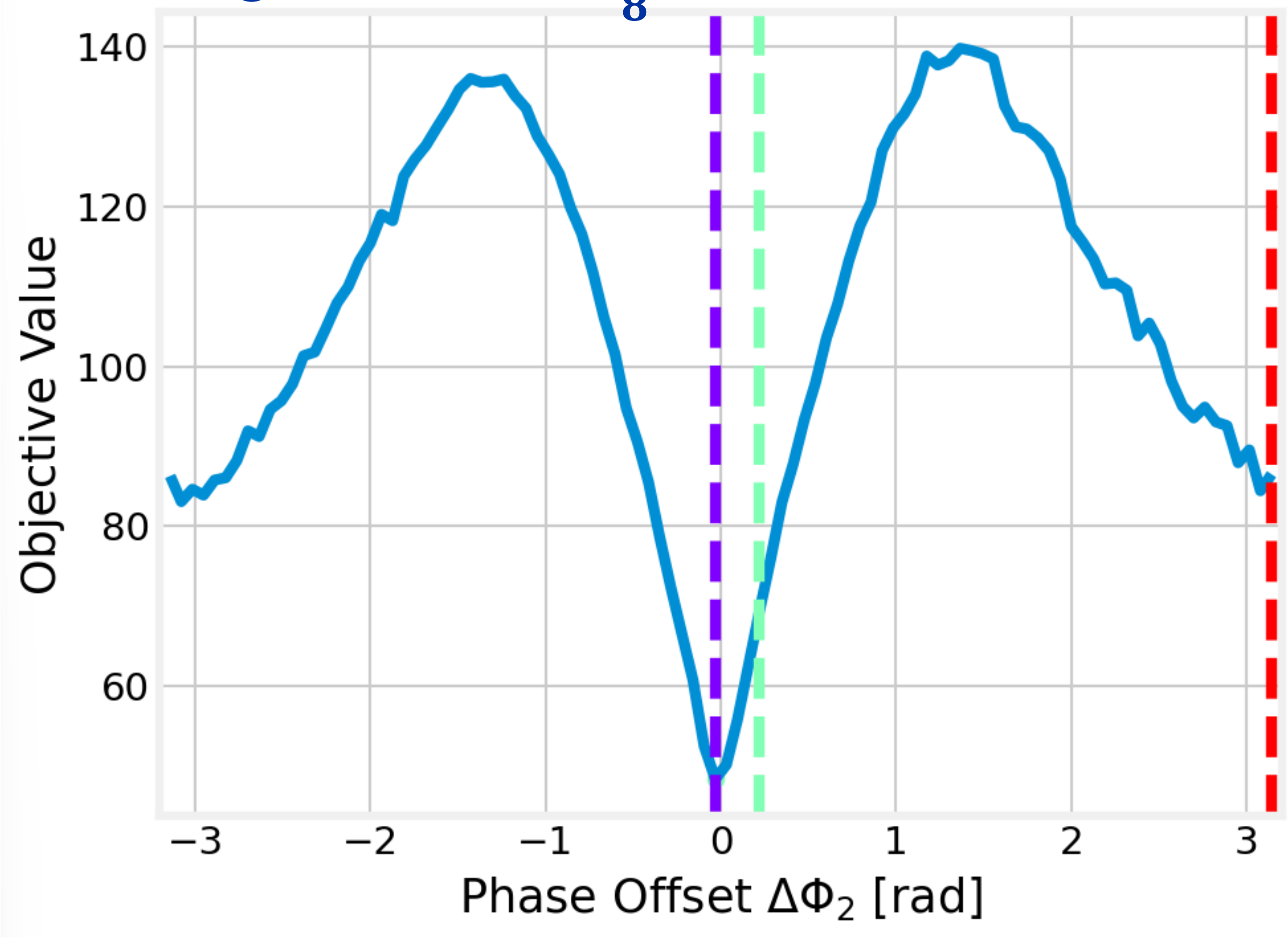






Method: Objective Function

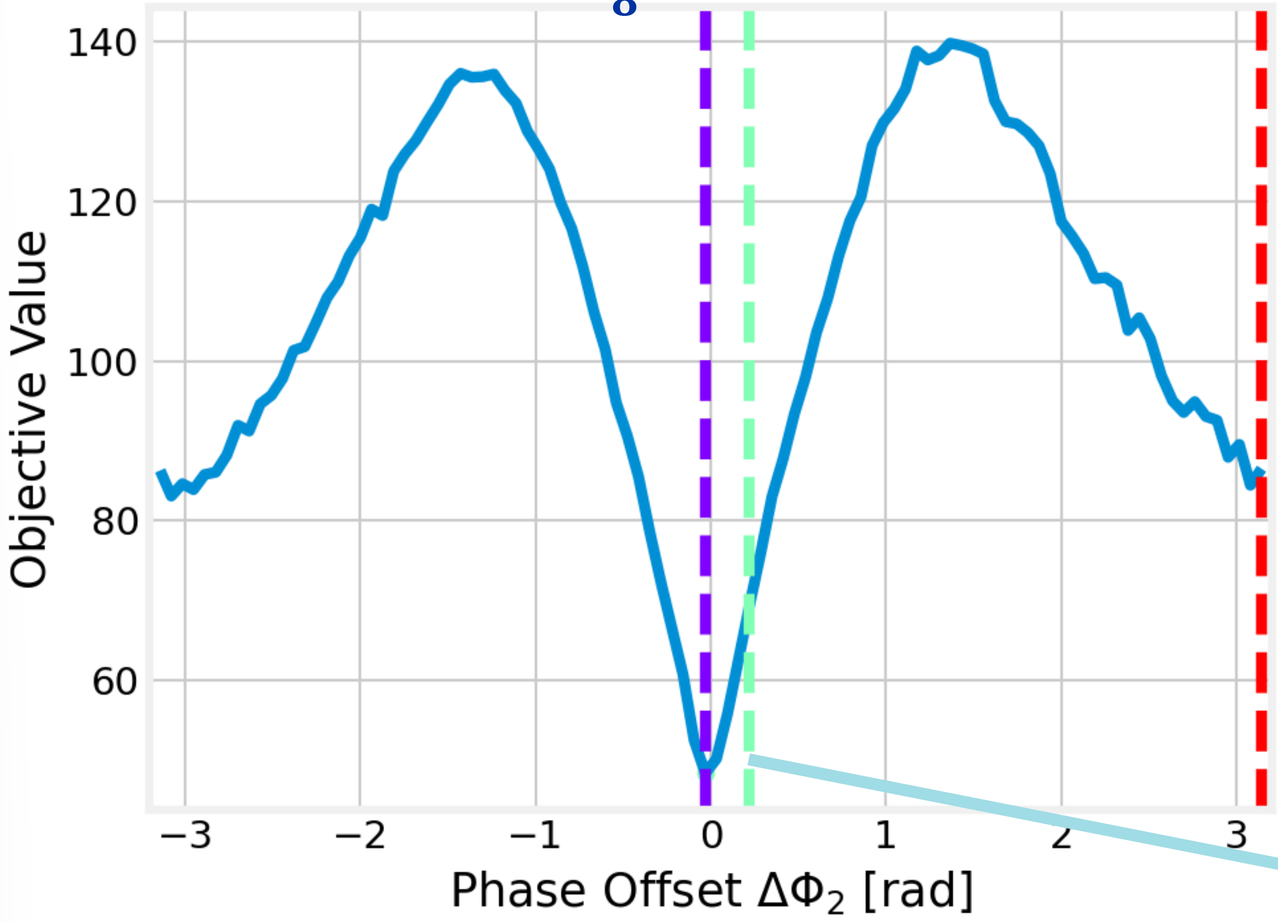
Voltage Ratio $r = \frac{5}{8}$

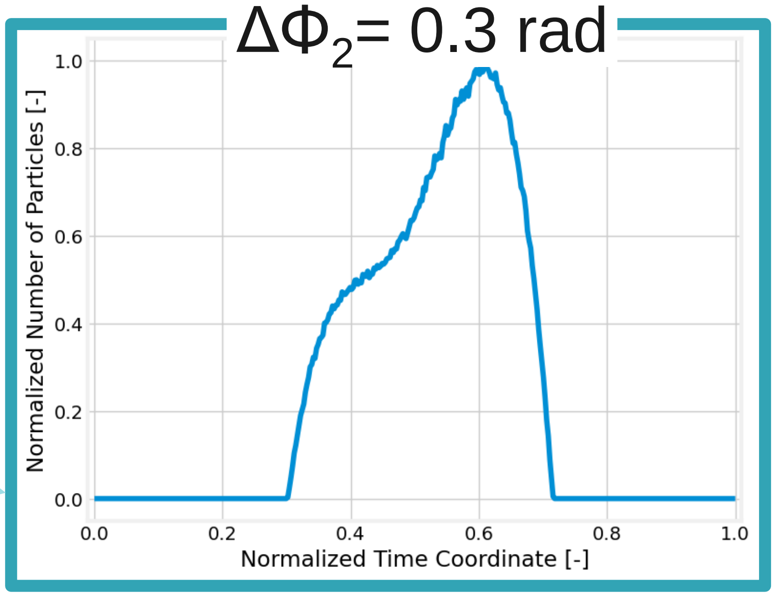


- Objective value as a function of phase offset from optimum

Method: Objective Function

Voltage Ratio $r = \frac{5}{8}$

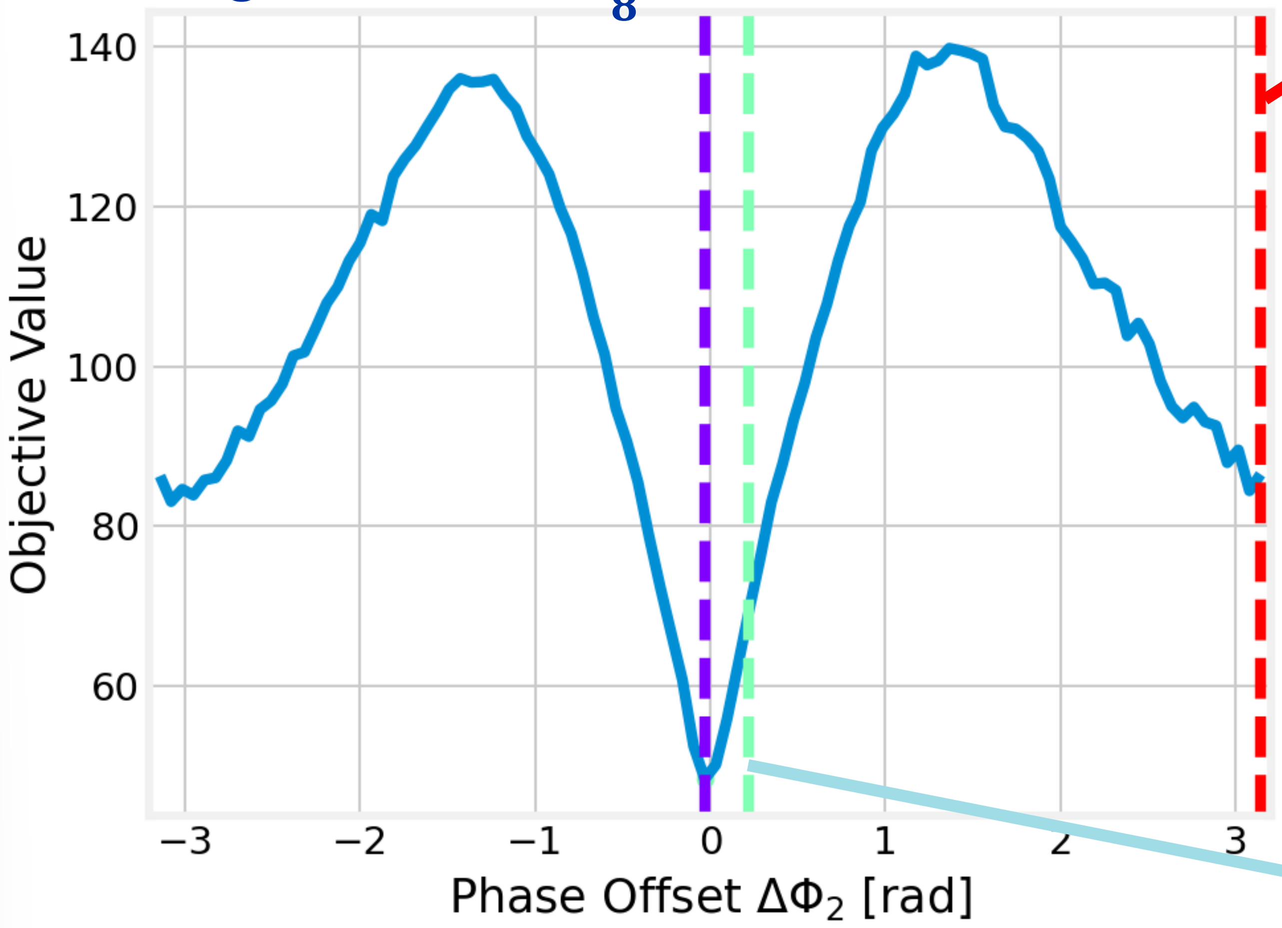


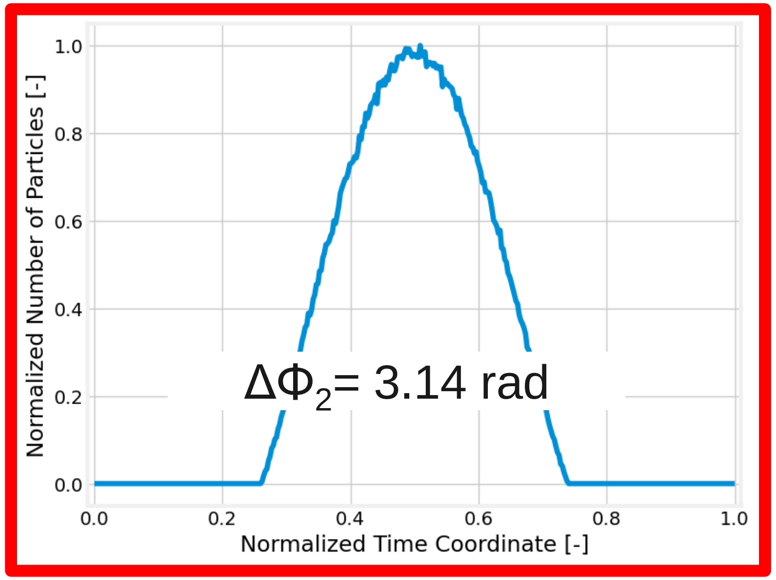


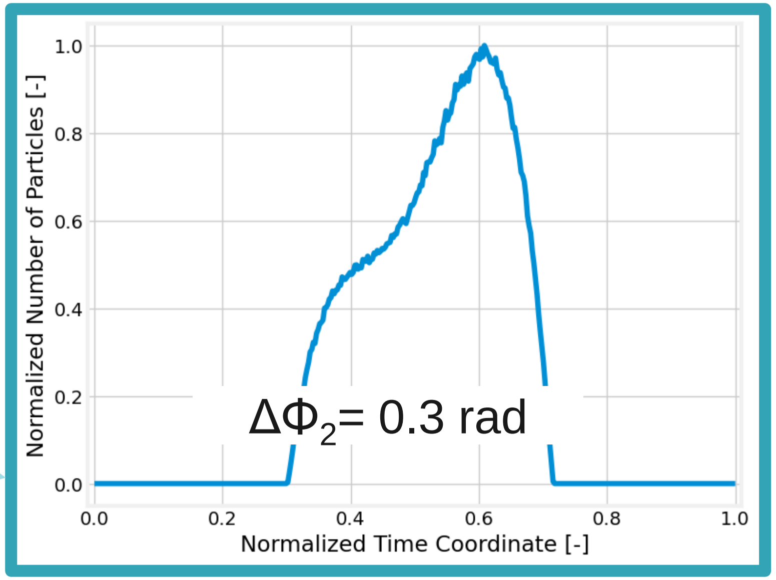
- Objective value as a function of phase offset from optimum

Method: Objective Function

Voltage Ratio $r = \frac{5}{8}$



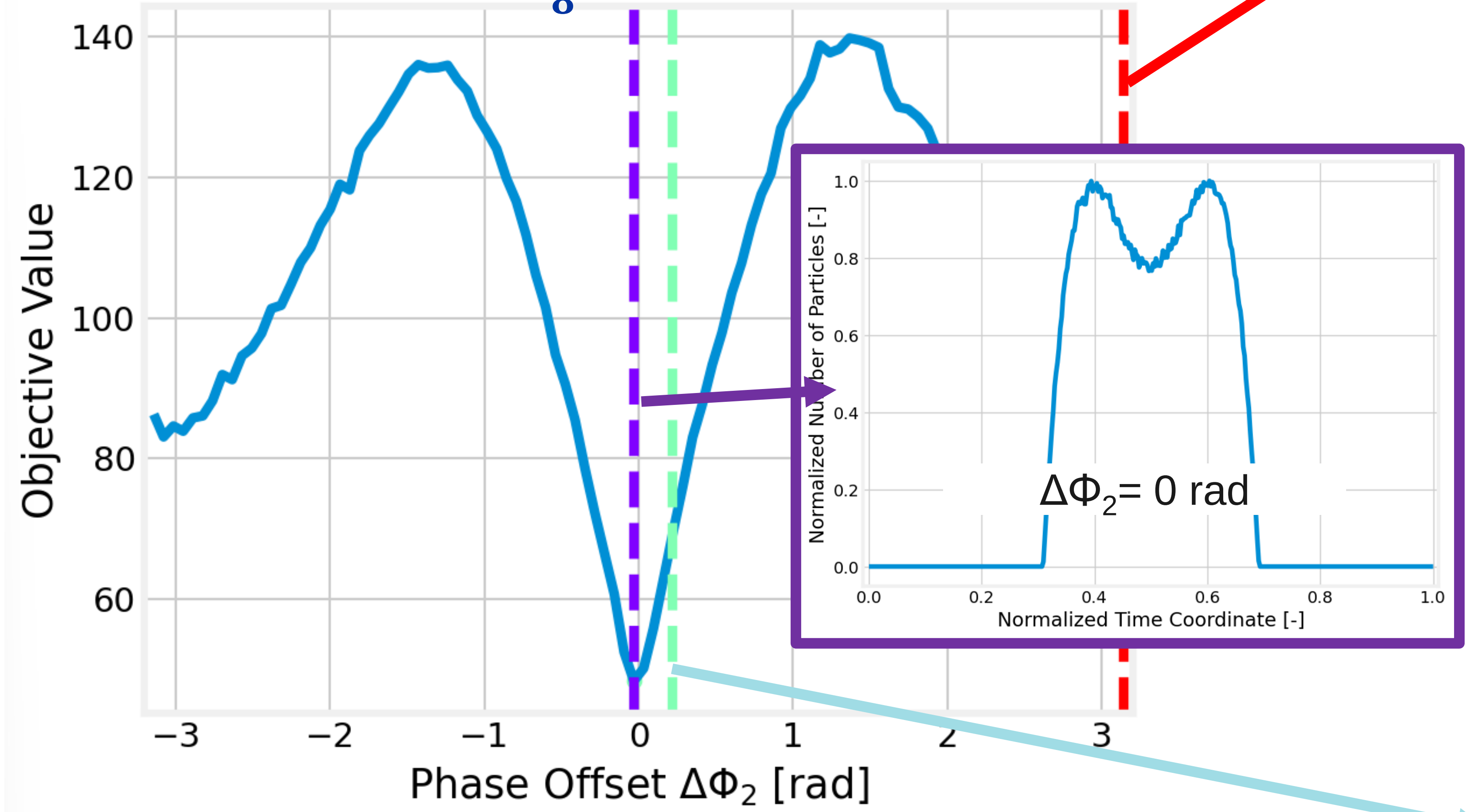


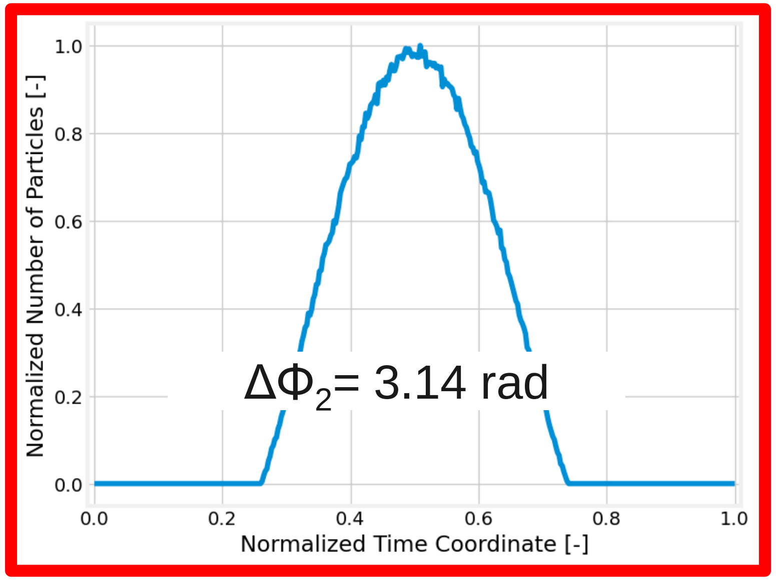


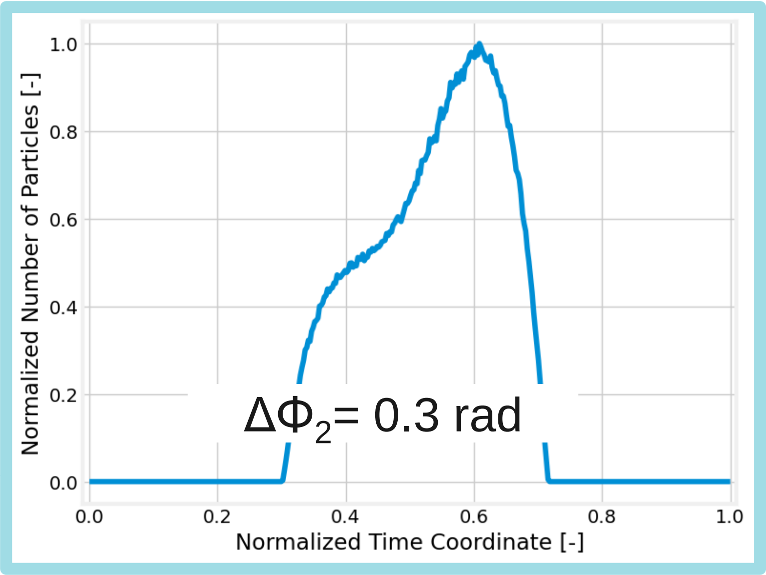
- Objective value as a function of phase offset from optimum

Method: Objective Function

Voltage Ratio $r = \frac{5}{8}$

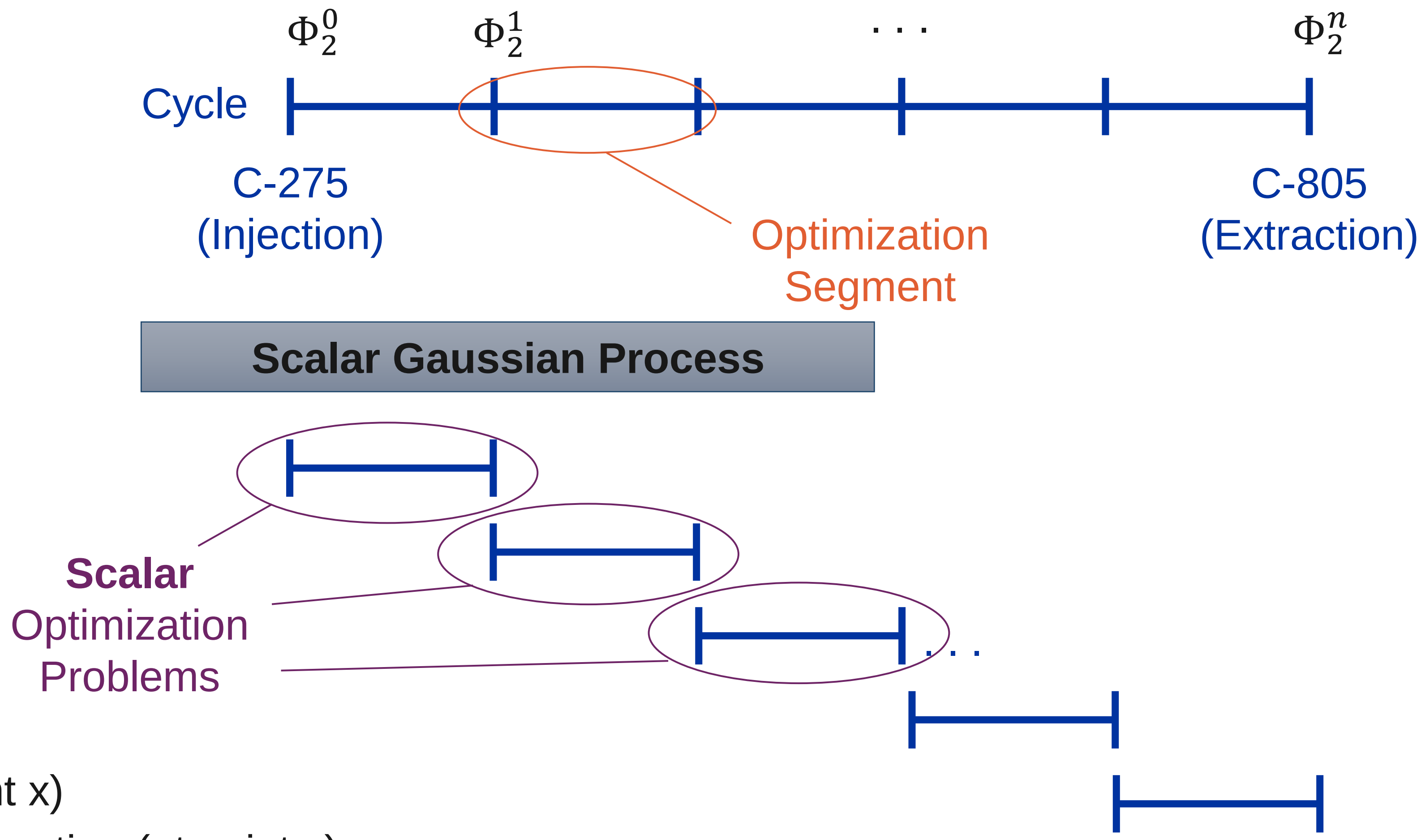






- Objective value as a function of phase offset from optimum

Method: Bayesian Optimization (scalar)



$f(\Phi_2^x)$ = Objective Function (at point x)
$\tilde{f}(\Phi_2^x)$ = Approximated Objective Function (at point x)

Method: Bayesian Optimization (scalar)

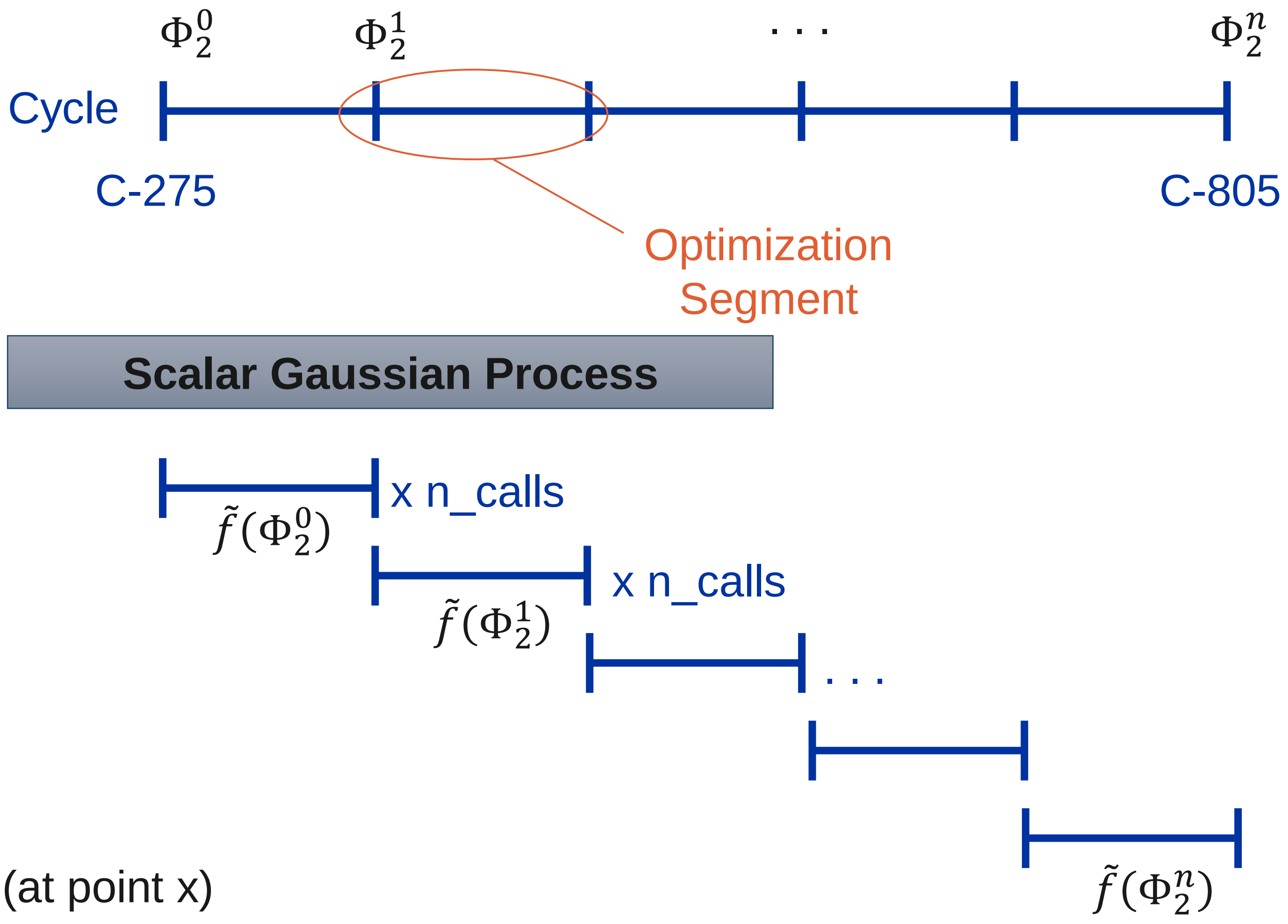
Modelling the objective function: Gaussian Process

+

Optimizing the objective function: Bayesian Optimization

=

Optimizer



$f(\Phi_2^x)$ = Objective Function (at point x)

$\tilde{f}(\Phi_2^x)$ = Approximated Objective Function (at point x)

Method: Bayesian Optimization (intervals)

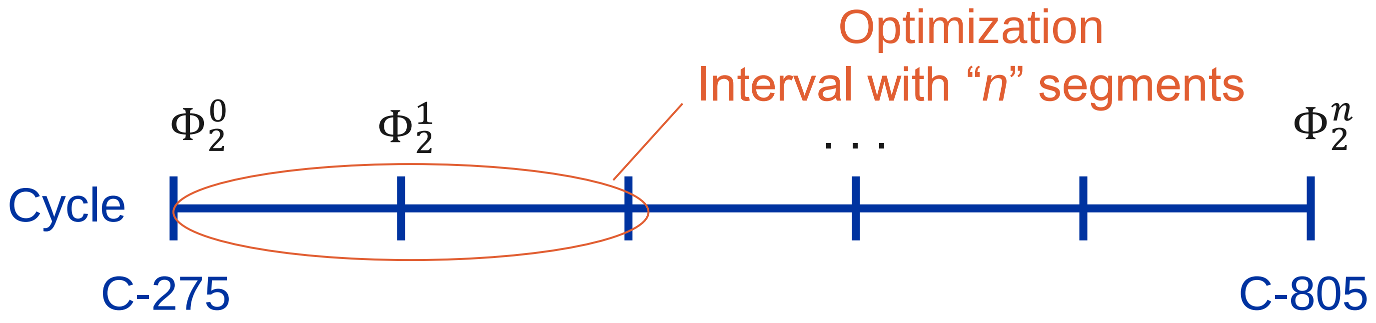
Modelling the objective function: Gaussian Process

\+

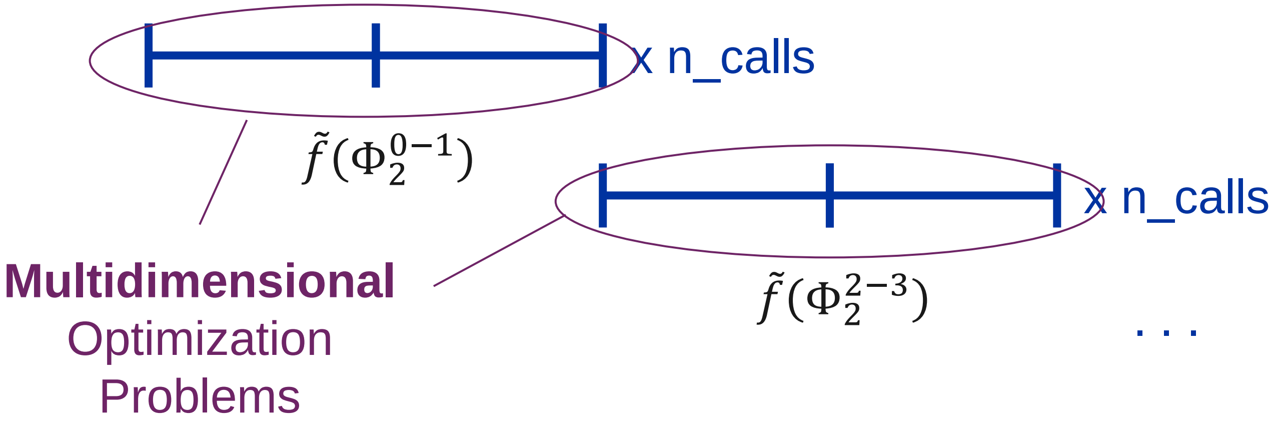
Optimizing the objective function: Bayesian Optimization

=

Optimizer



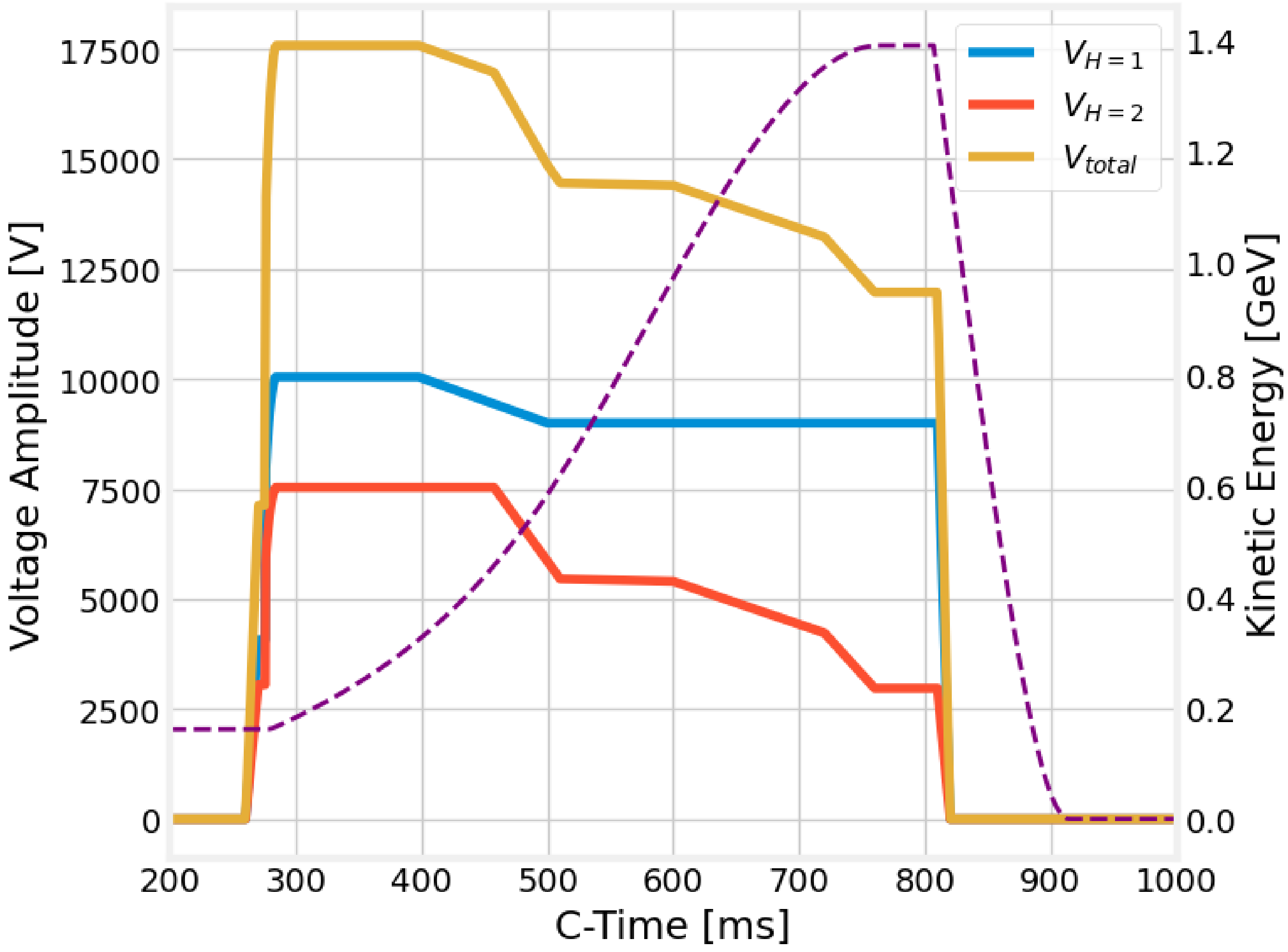
Multidimensional Gaussian Process



$f(\Phi_2^x)$ = Objective Function (at point x)

$\tilde{f}(\Phi_2^x)$ = Approximated Objective Function (at point x)

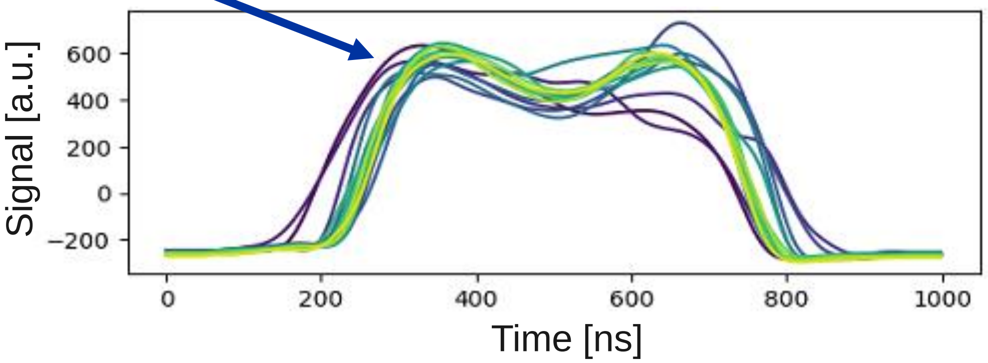
PSB Tests: Sequential Optimizer
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)



PSB Tests: Sequential Optimizer
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)

Injection makes optimization difficult

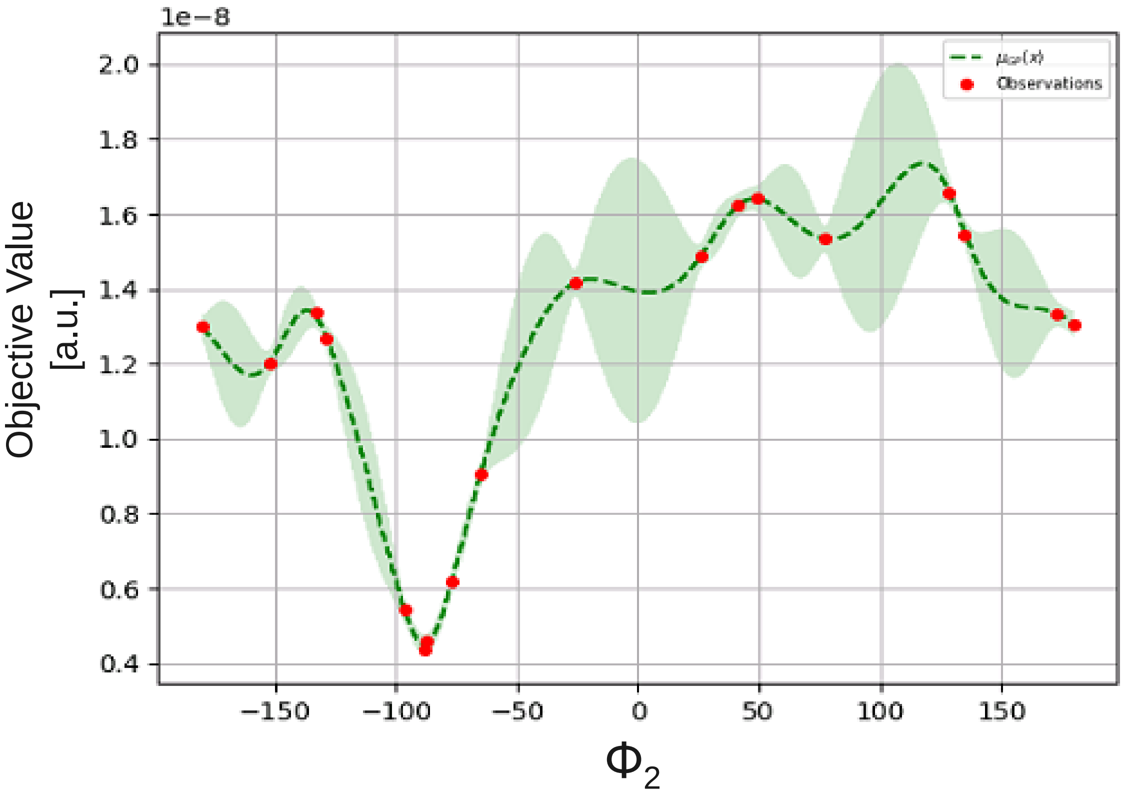
Phase set at -87º



C-278

C-275

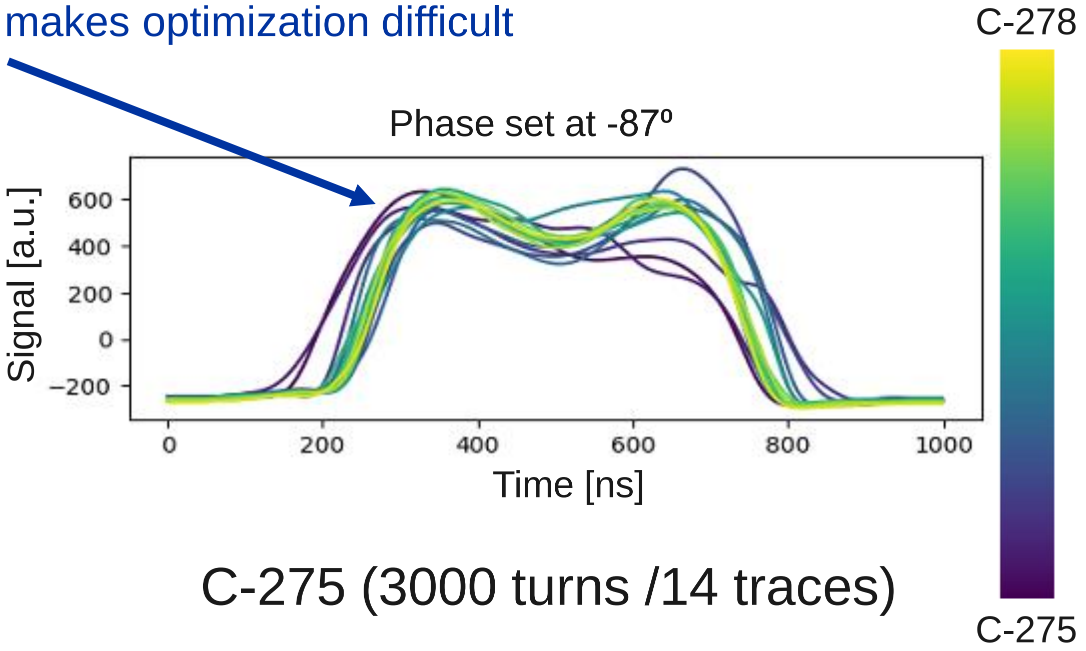
C-275 (3000 turns /14 traces)



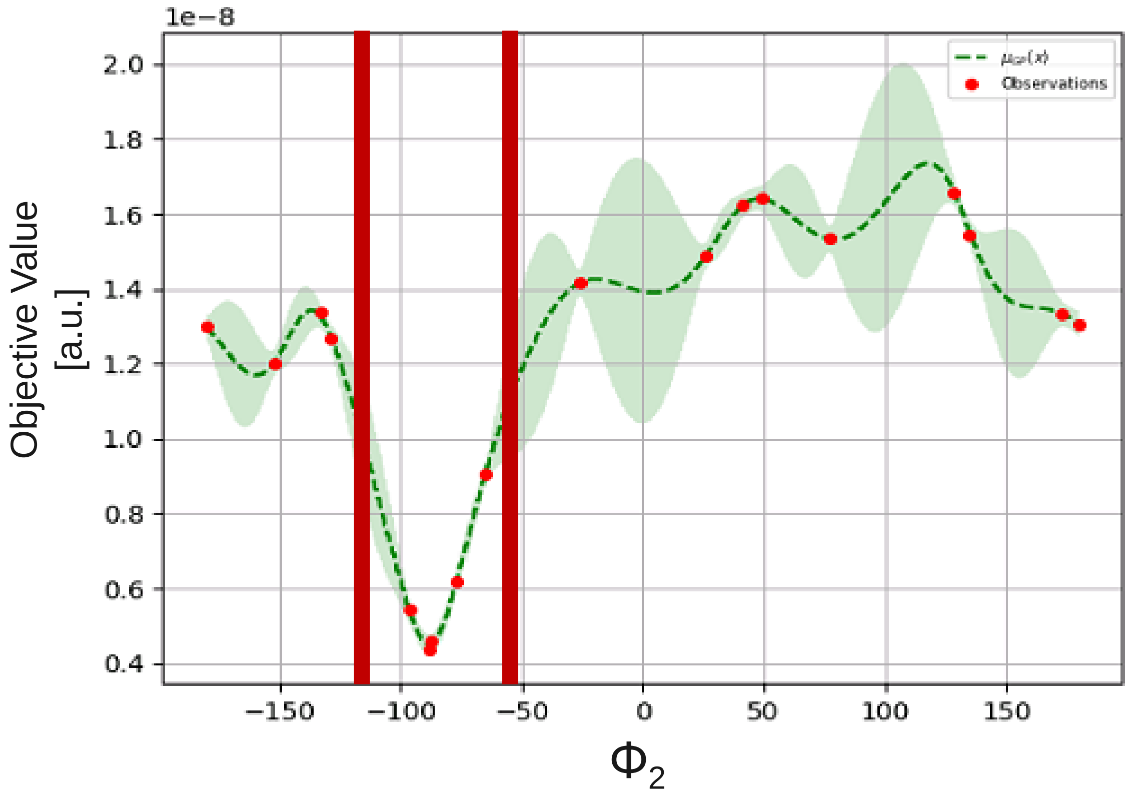
PSB Tests: Sequential Optimizer

(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)

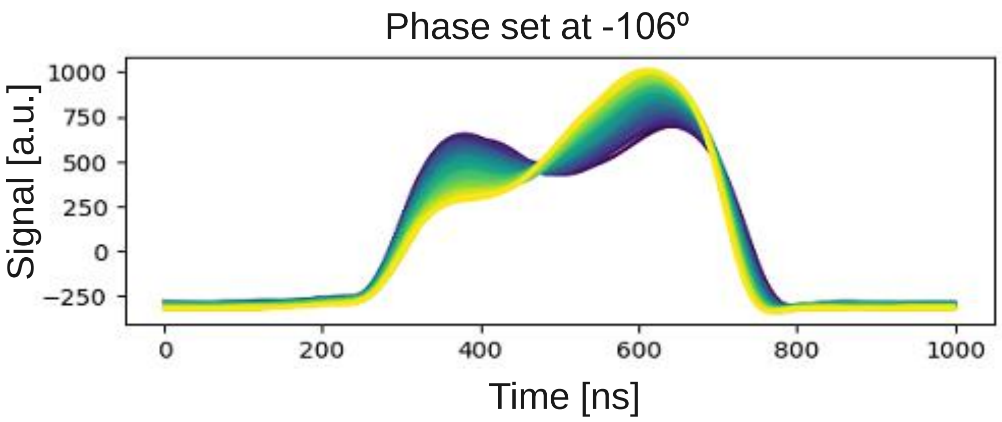
Injection makes optimization difficult

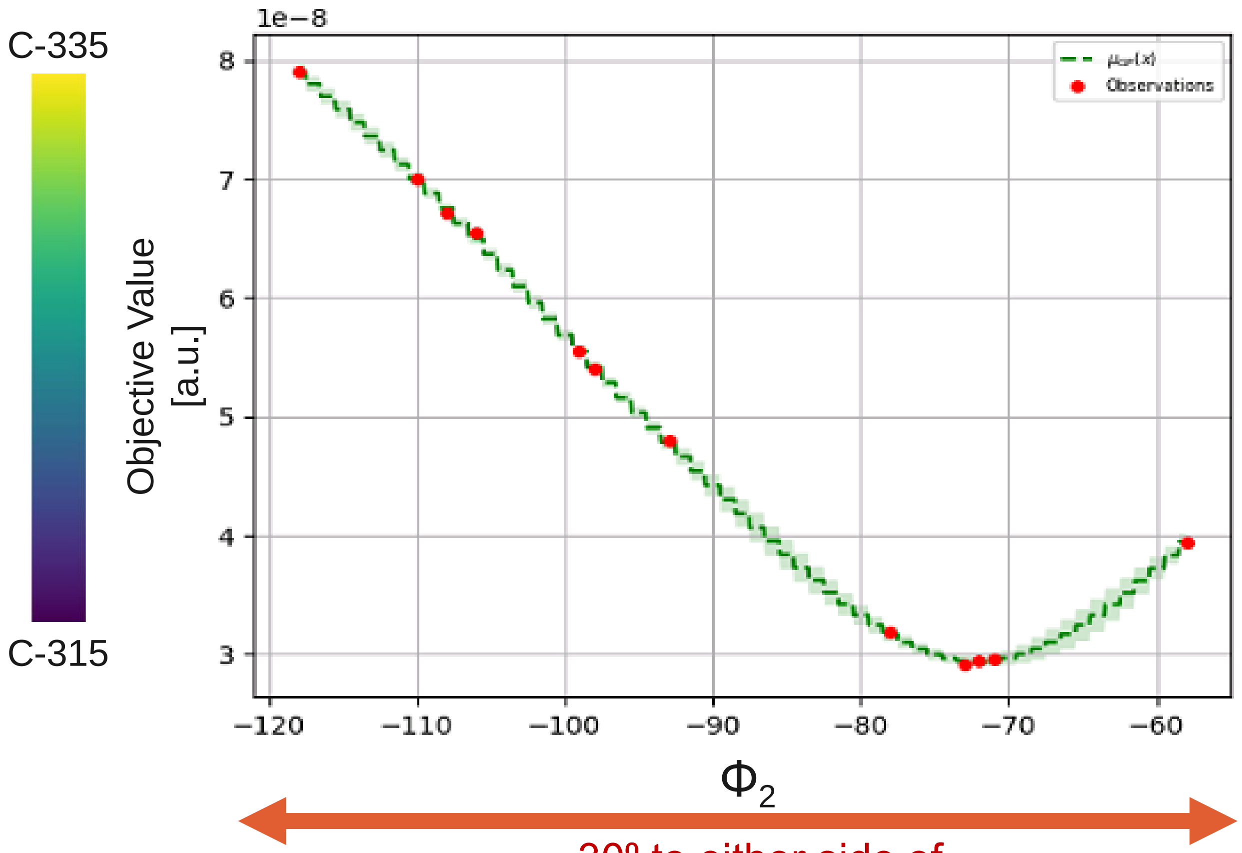


C-275 (3000 turns /14 traces)



PSB Tests: Sequential Optimizer
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)

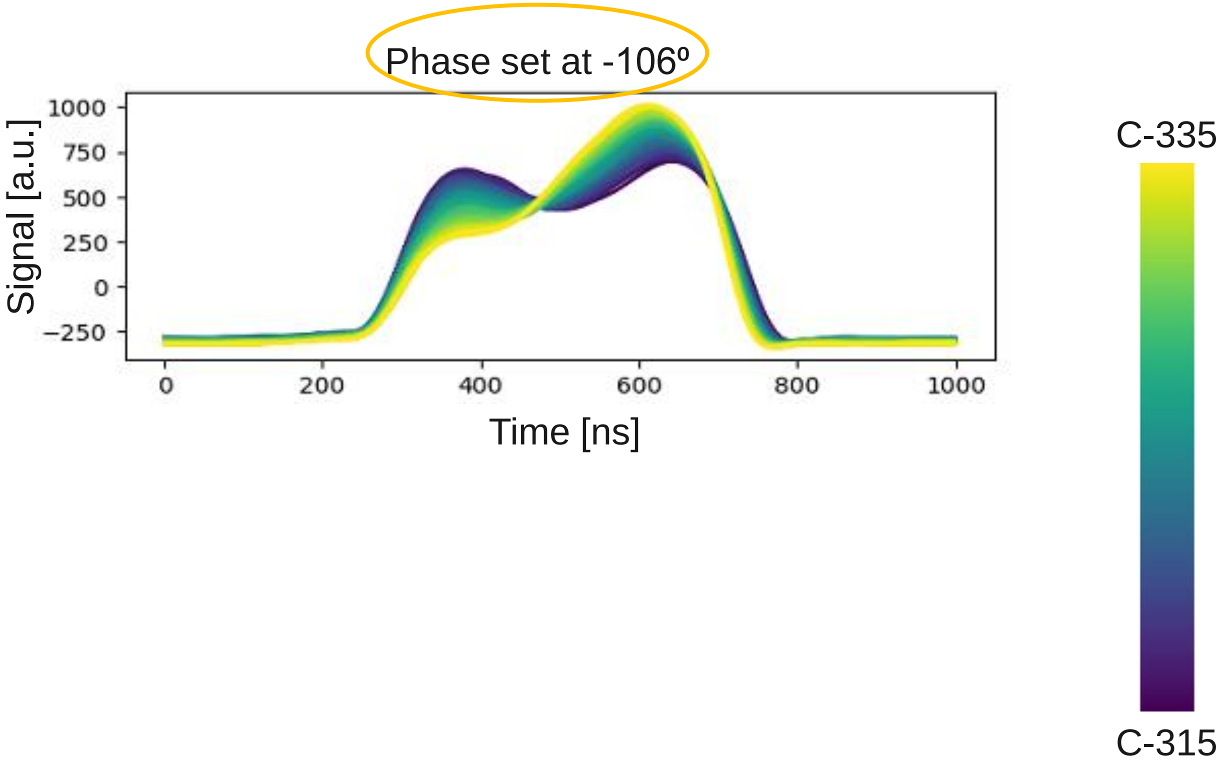


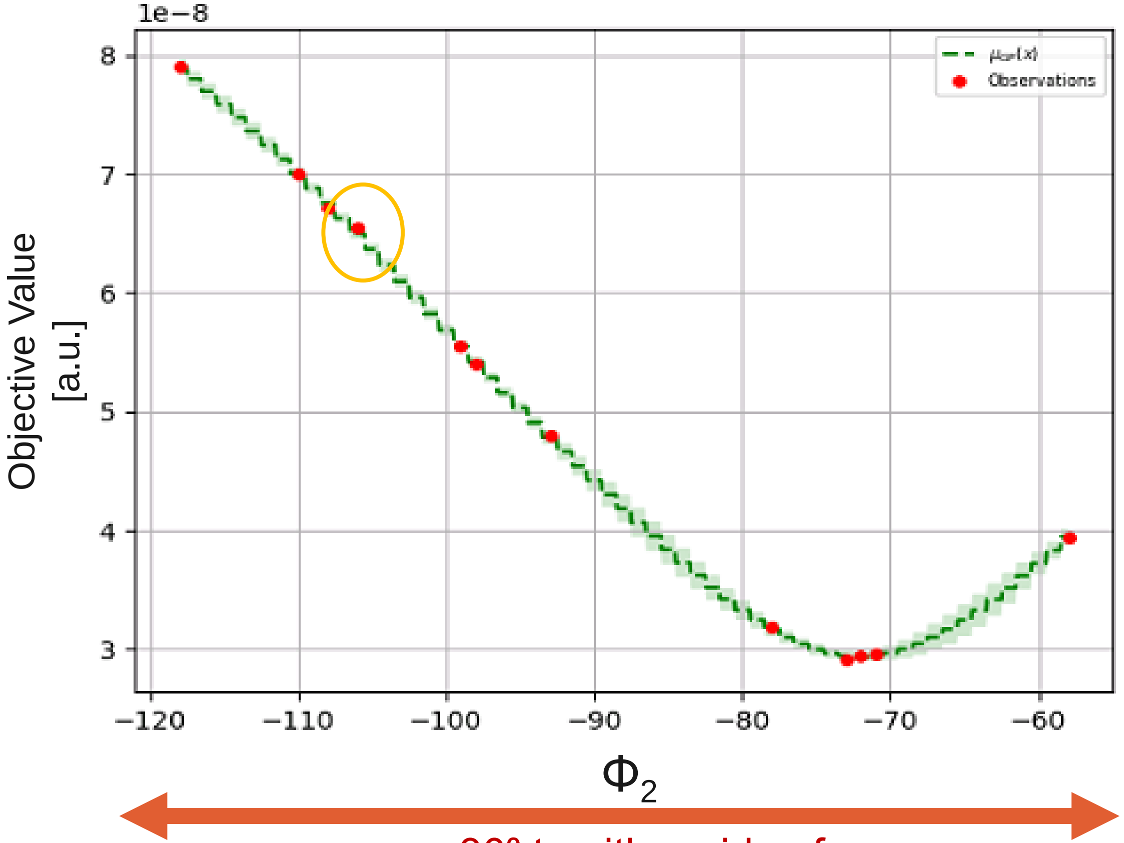


C-315 / C-335 (110 traces)

30º to either side of prev. solution

PSB Tests: Sequential Optimizer
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)



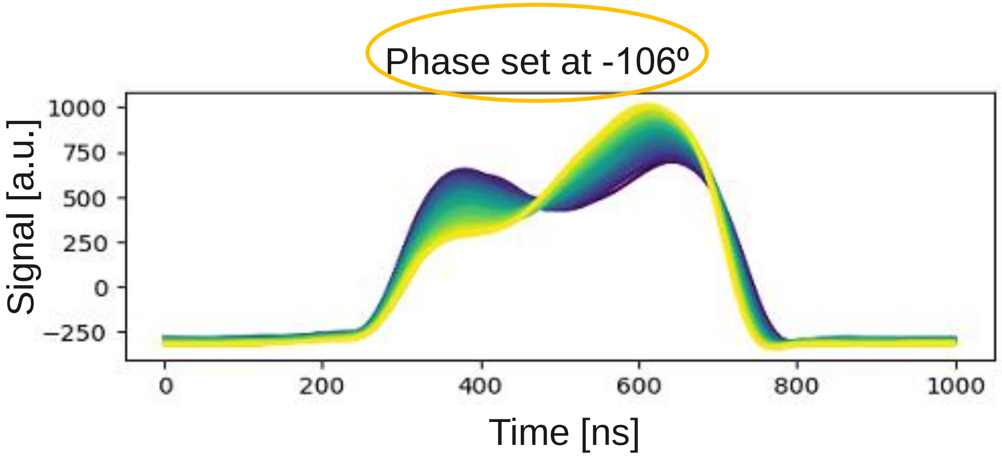


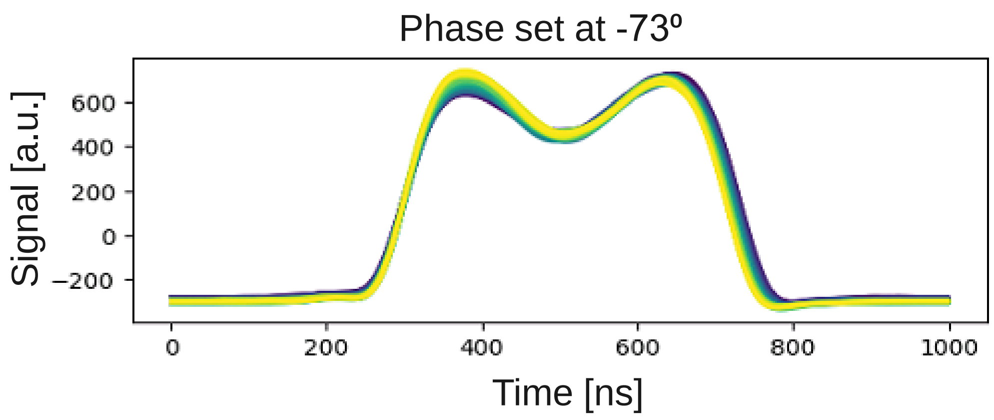
30º to either side of prev. solution

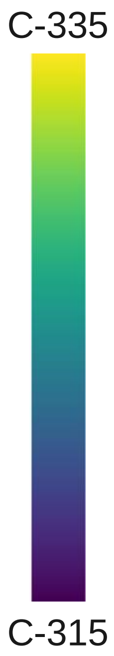
C-315 / C-335 (110 traces)

PSB Tests: Sequential Optimizer

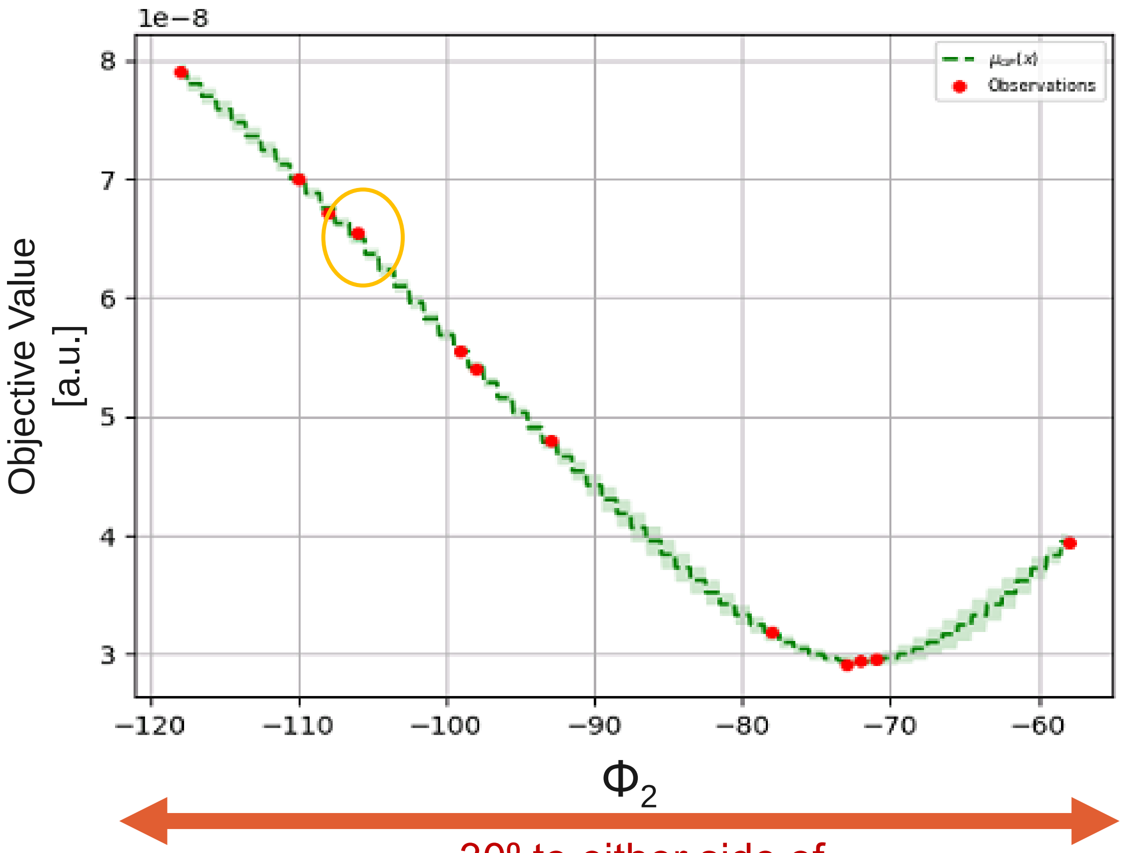
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)







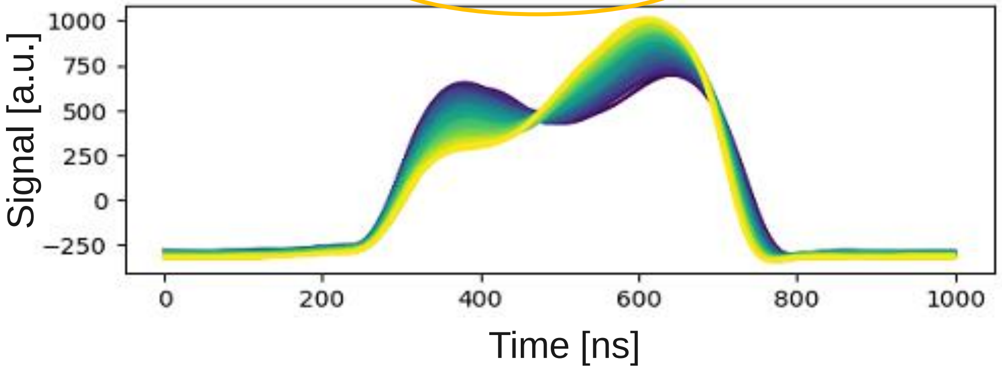
C-315 / C-335 (110 traces)



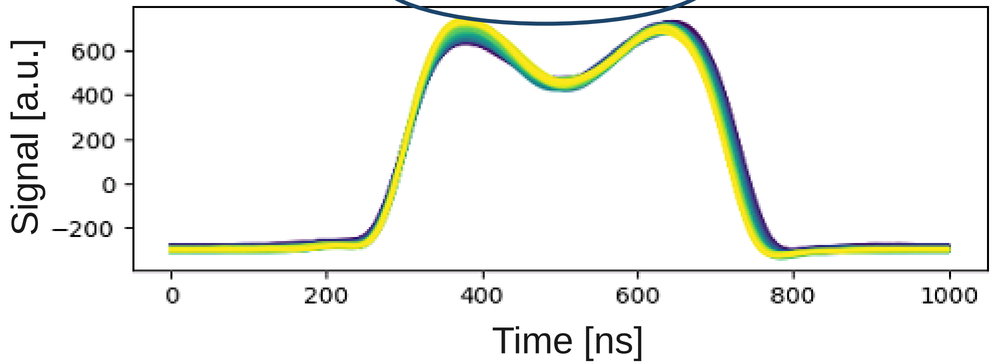
30º to either side of
prev. solution

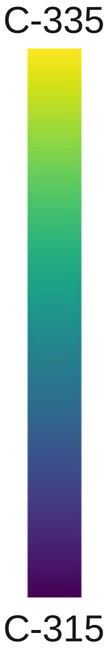
PSB Tests: Sequential Optimizer
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)

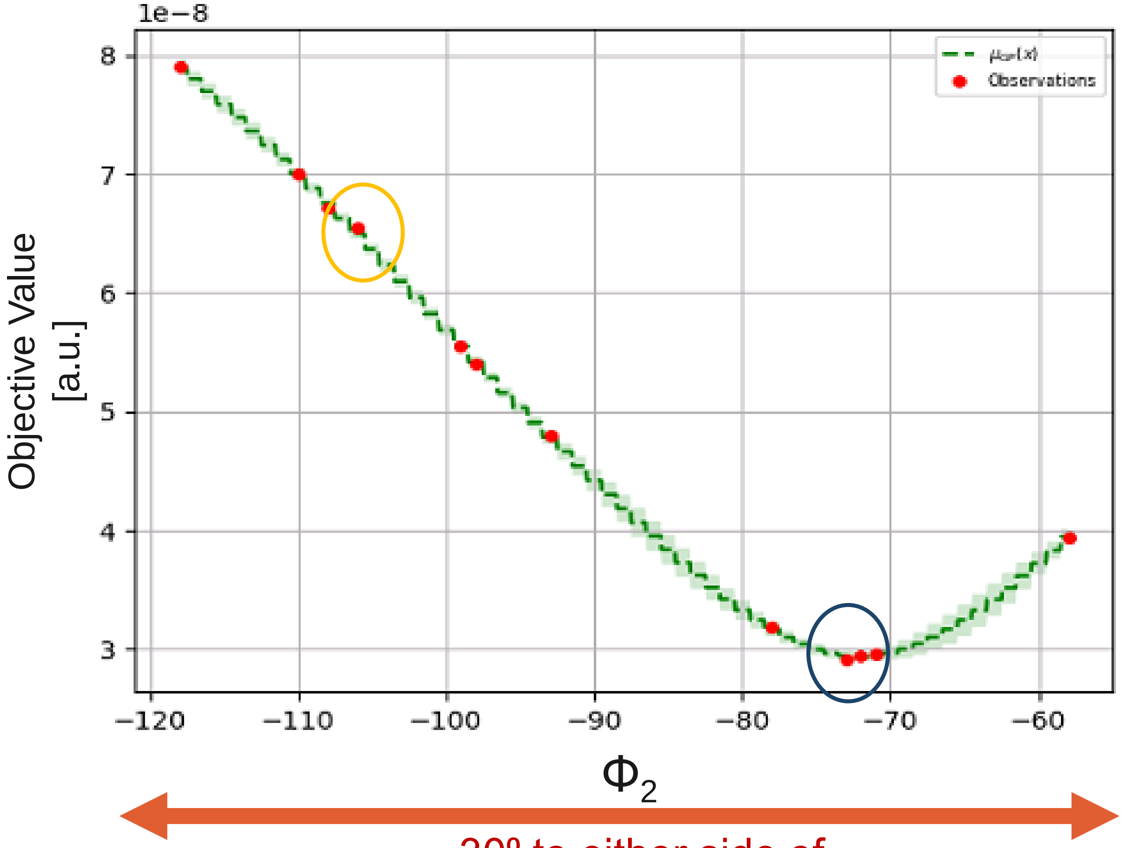
Phase set at -106º



Phase set at -73º





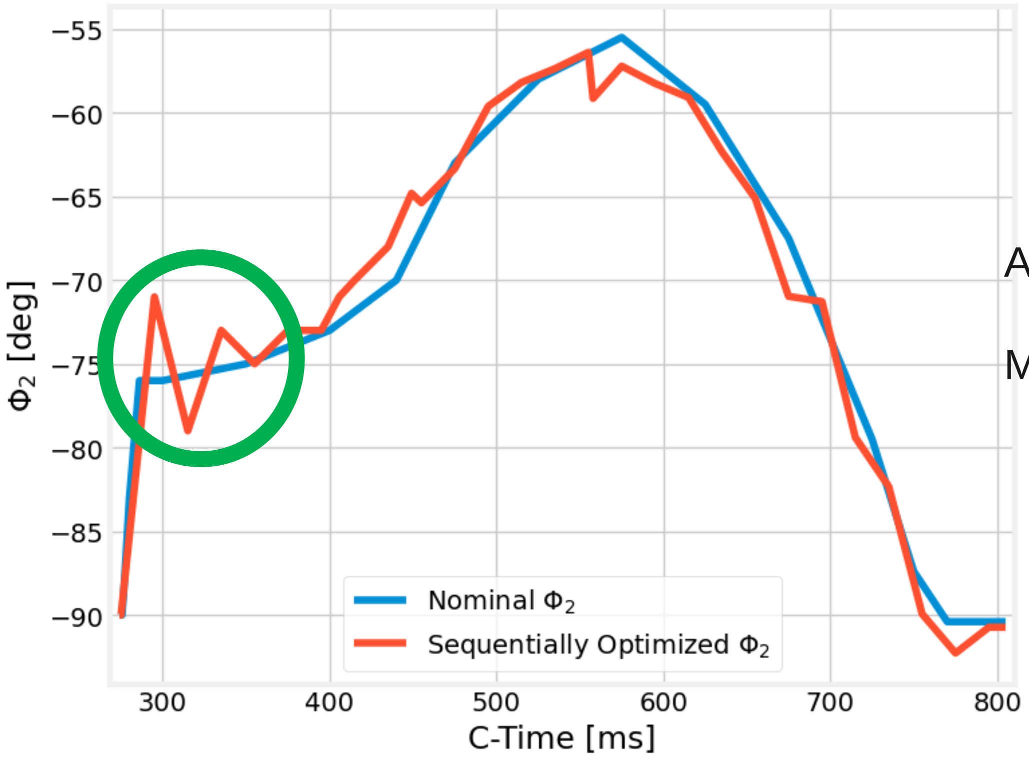


C-315 / C-335 (110 traces)

30º to either side of prev. solution

PSB Tests: Sequential Optimizer

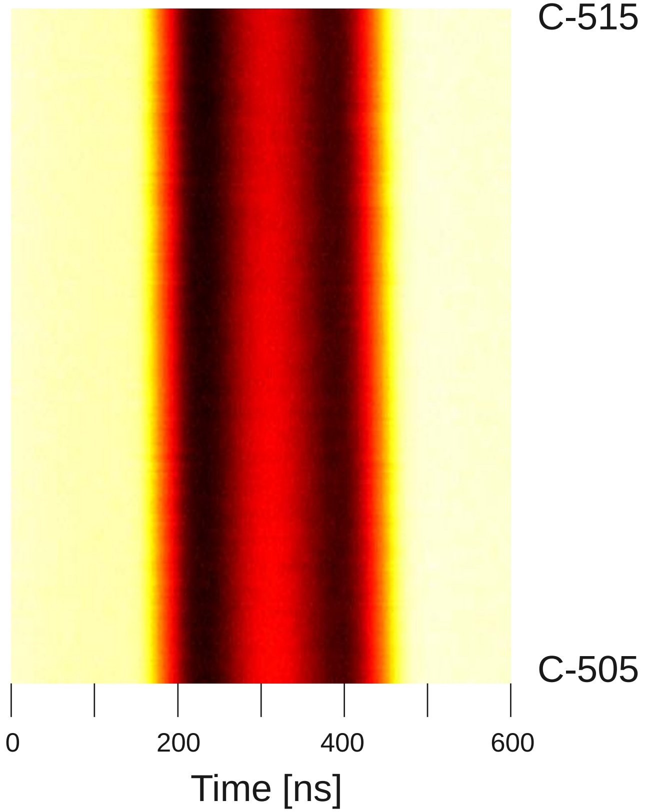
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)



Average difference: **1.42 º**

Maximum difference: **5 º**

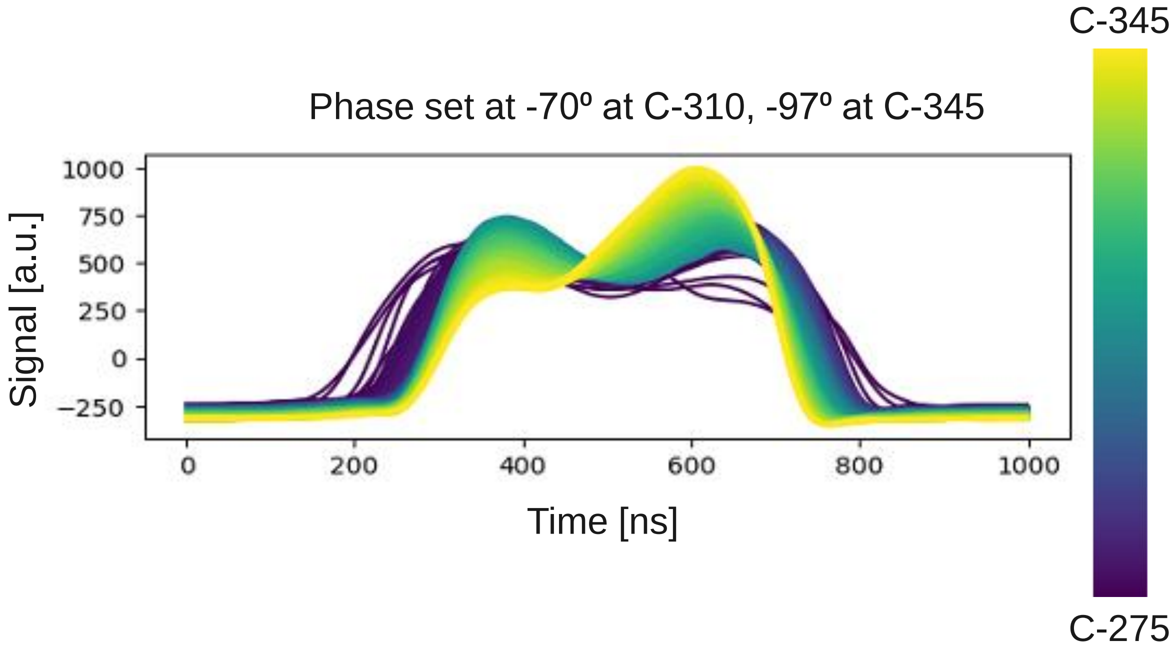
Tomoscope view



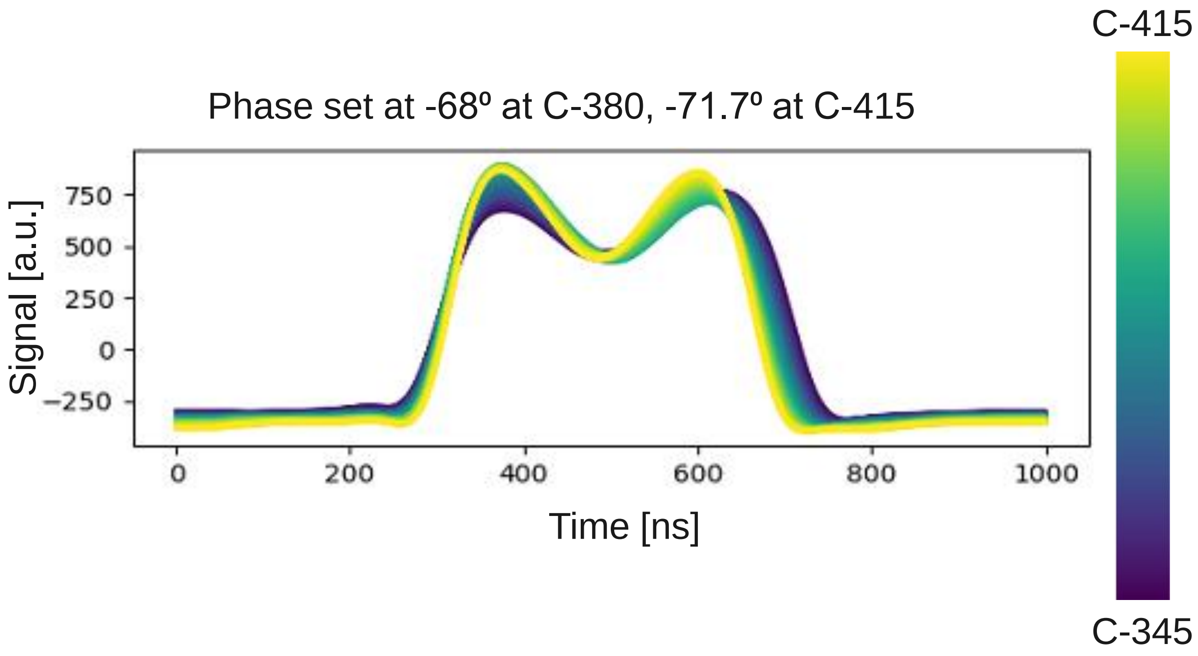
- Sequential optimization was partially successful (beam profile achieved & no loss)with no prior knowledge
- Slightly noisy, at the initial turns but can be fixed with filtering/averaging
- Time-step: 20 ms
- **Optimization time**: 282 cycles (50 min)

PSB Tests: Interval Optimizer
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)

2 phases/segments per interval

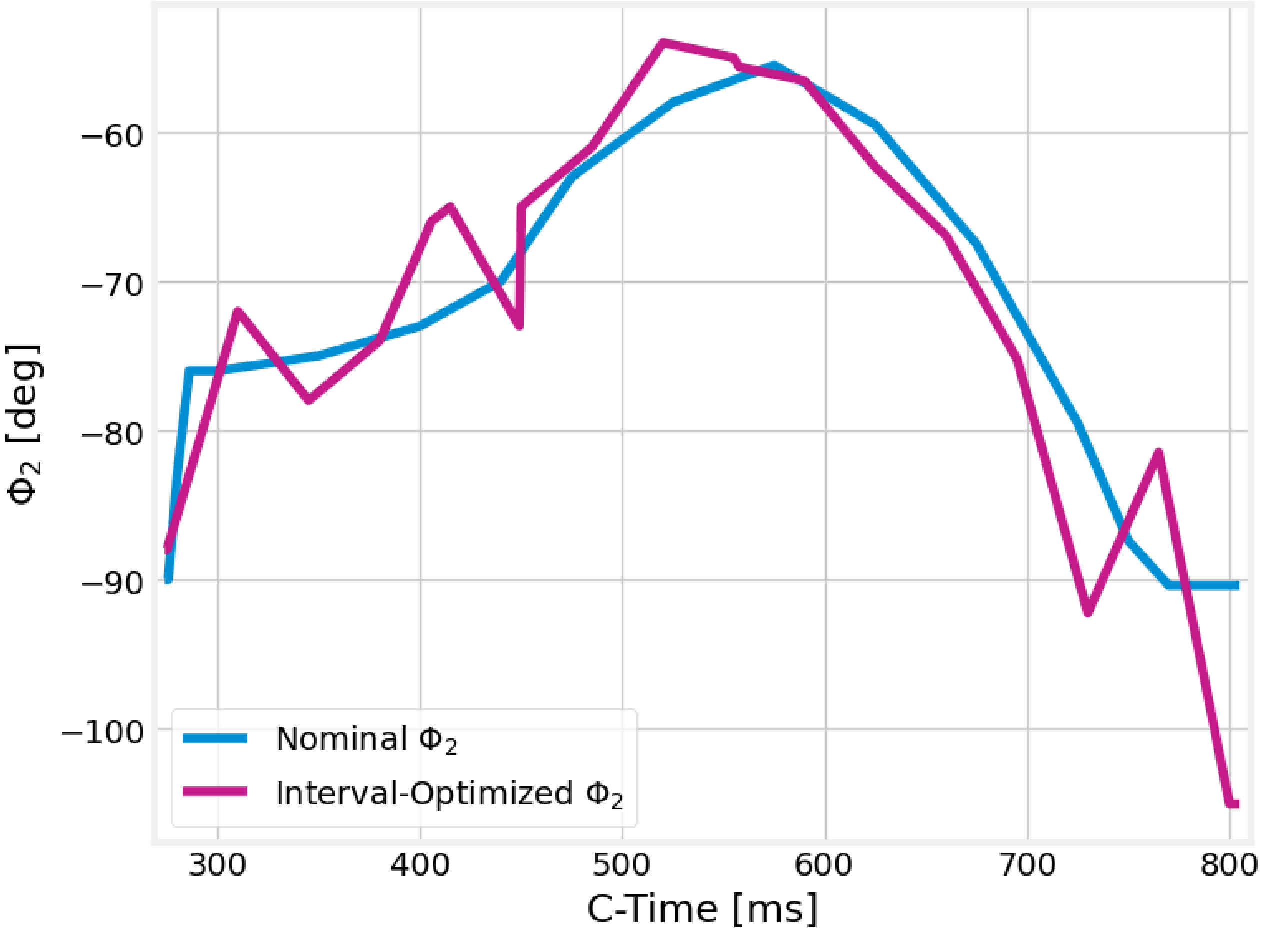


C-275 / C-345
(373 traces)



C-345 / C-415
(371 traces)

PSB Tests: Interval Optimizer
(MD14683) - 1.4 GeV Cycle: Low Intensity (40e10 ppb)



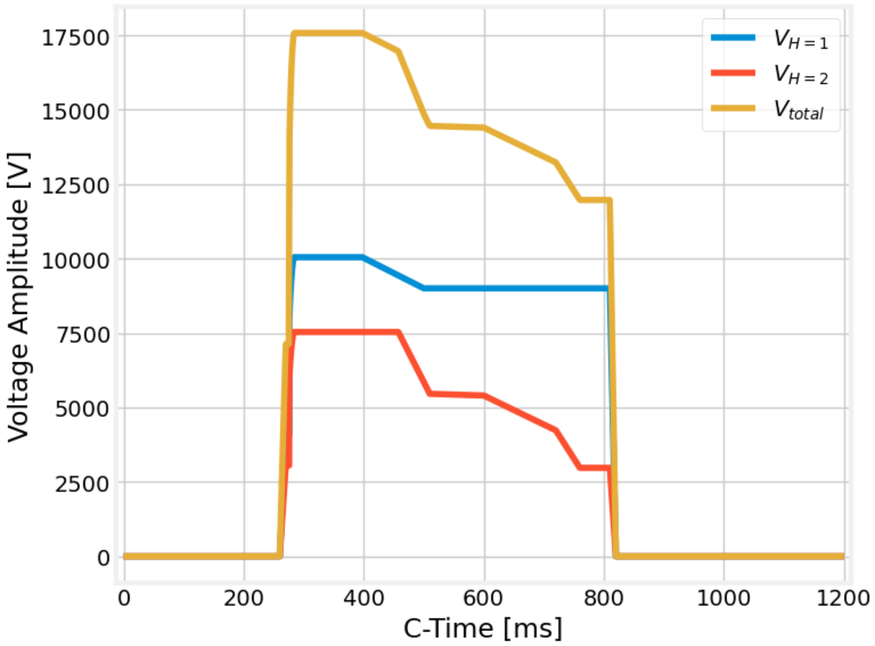
2 points per interval

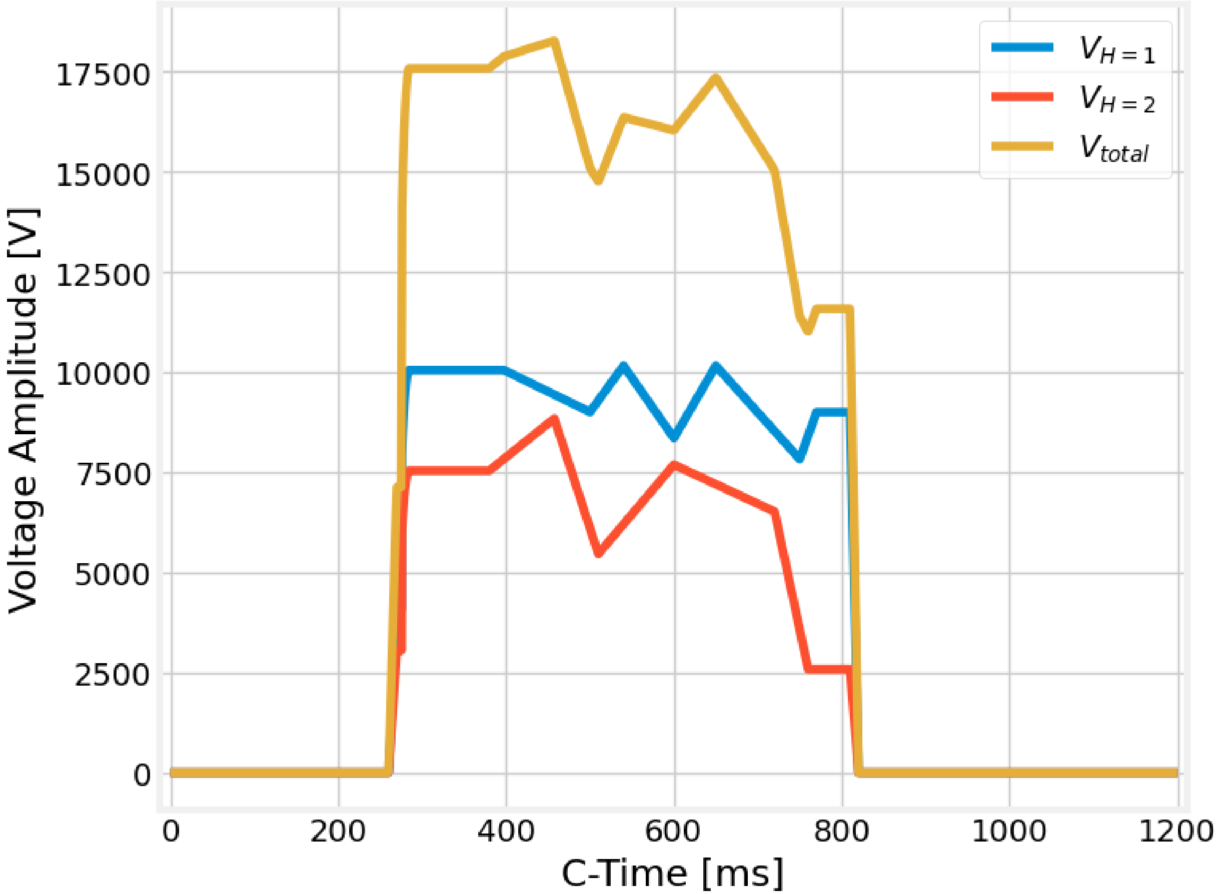
Average difference: **3.28 º**

Maximum difference: **11.28 º**

- Desired beam profile shape maintained, but not optimally
- More points = more noise
- Time-step: 35 ms
- **Optimization time:** 157 Cycles (17 min 30s)

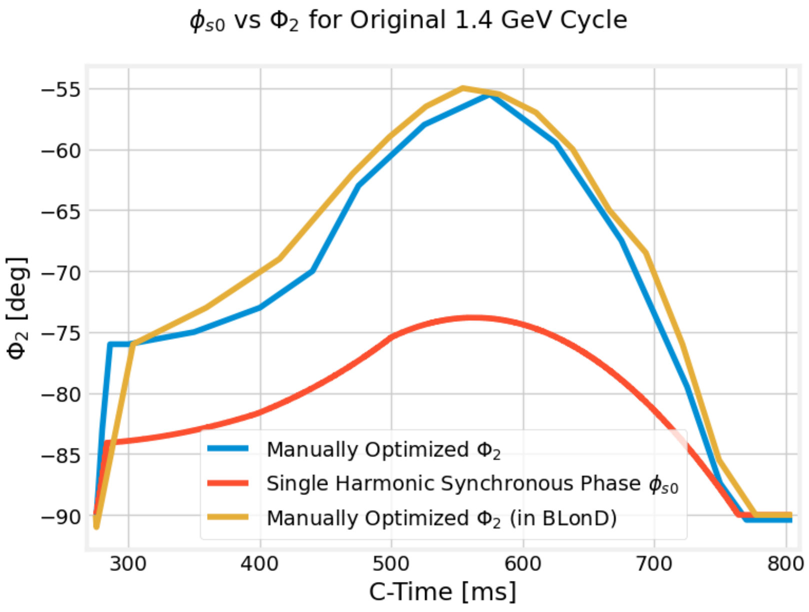
PSB Tests: Interval Optimizer
(MD14683) – Modified 1.4 GeV Cycle: Low Intensity (40e10 ppb)

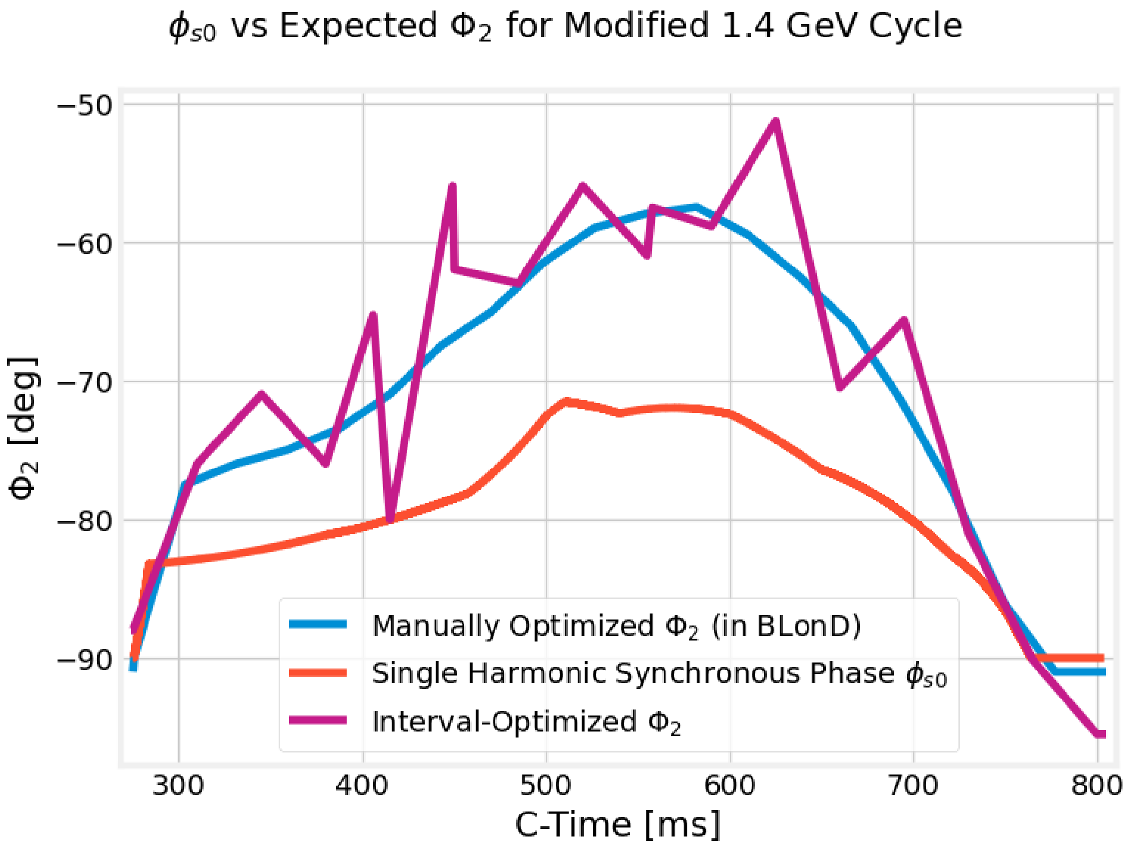




PSB Tests: Interval Optimizer

(MD14683) – Modified 1.4 GeV Cycle: Low Intensity (40e10 ppb)





Conclusions

- Proof-of-principle study for automatic phase setting
- The optimizer approximately computes the correct phasing, but this is not robust to all operational conditions (especially not high B-dots due to bunch profile perturbations, as was seen in simulation).
- Need to complement more complex voltage programs with a manual optimization to be able to compare
- Next steps:
  - Train a Neural Network to compute the objective function with labelled data that can classify and compute best profile under **all** circumstances
  - Find a framework which can compute phases throughout the cycle for faster convergence and better performance

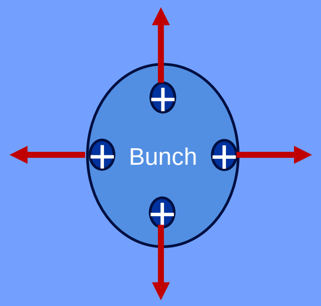
Thank you for listening!
Questions?

Future work

- Triple Harmonic Operation:
  - GP to compute the 2nd and 3rd harmonic phase → much like current segmented approach
  - Could also be extended to compute amplitudes of all voltages at an optimization interval (if KPI fusion goes well)
- AI approach to attempt to compute amplitudes and phases over the entire cycle (time series forecasting) and possibly interpret the data:
  - Interpretation could be through mapping the input-output space or by using a model which is inherently interpretable (like Temporal KANs)
  - Explore & Trade-Off : Temporal Fusion Transformers / Bi-GRU / VCformer / etc.

Supporting Slides: Space Charge Effect

Space Charge = Defocusing Effect on Longitudinal Bunch Profile in PSB*



$$\gamma < \gamma_T = \textit{de-focusing} \text{ (PSB)}$$

$$F_{SC} \propto \frac{I}{\beta\gamma^2}$$

$$\Delta U_{SC} \propto \frac{d\lambda_\phi}{dt}\frac{1}{\beta\gamma^2}$$

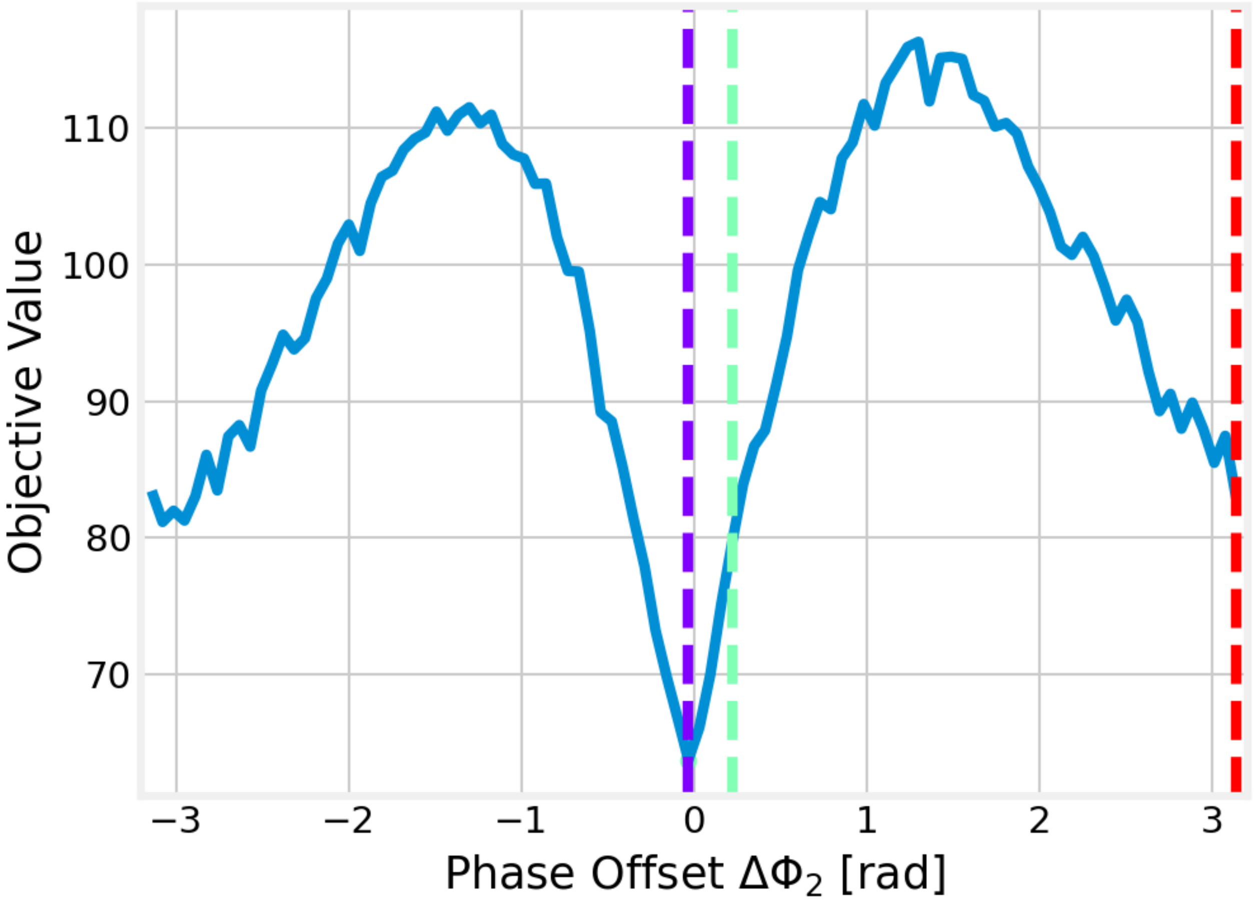
Minimize this

Minimize bunch profile gradient by altering 2nd harmonic phase Φ_2

(+ compensating beam loading: small shift in Φ_2)

Supporting Slides: Objective Function

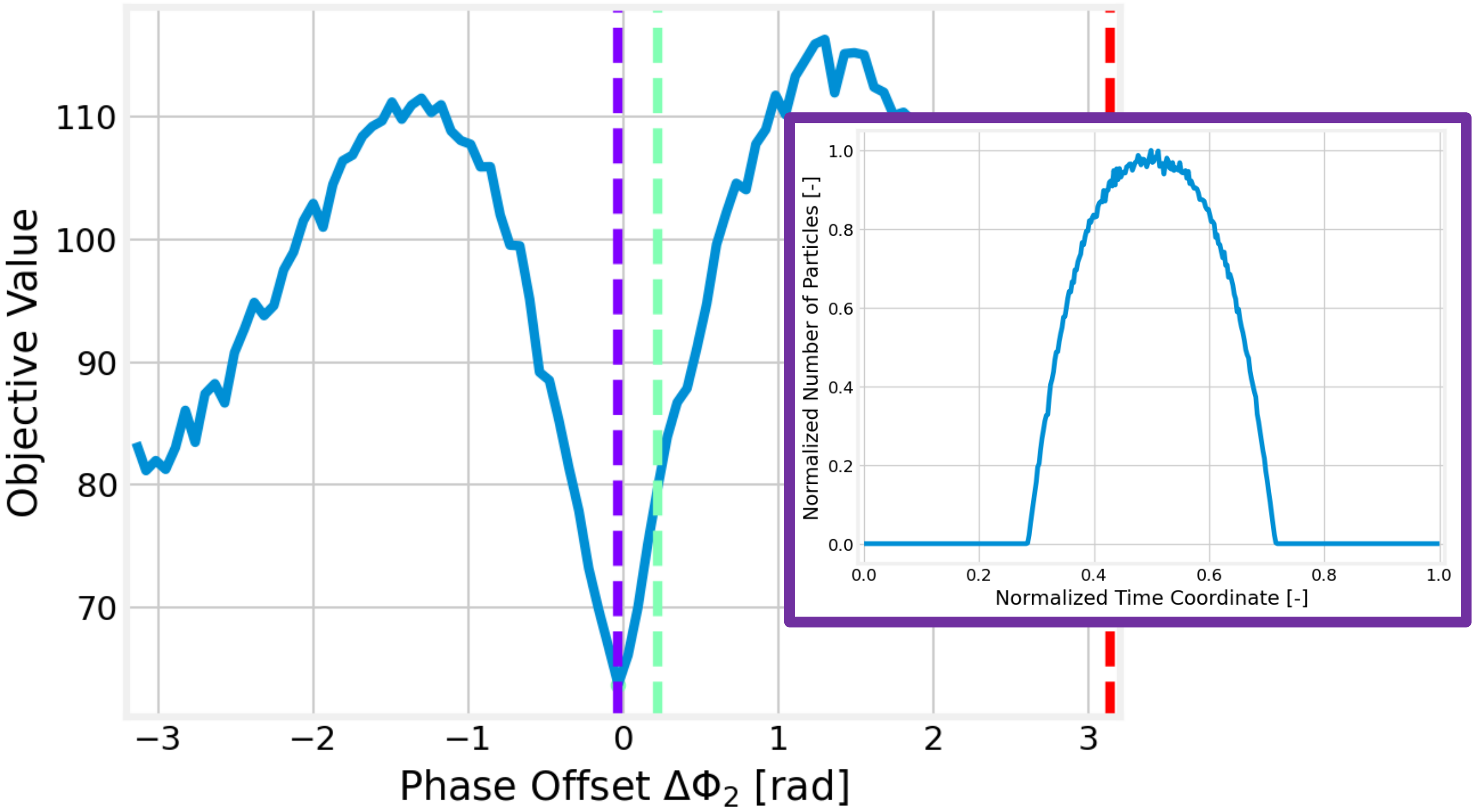
Change in Voltage Ratio $r = \frac{3}{8}$

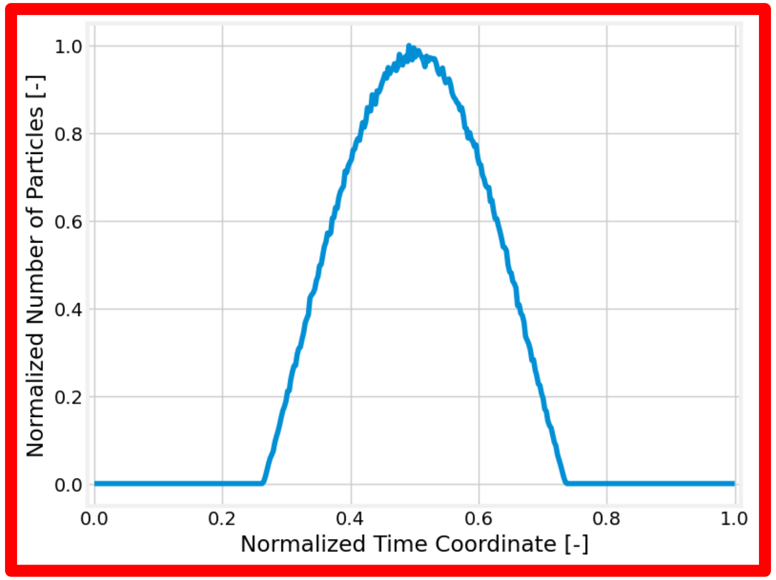


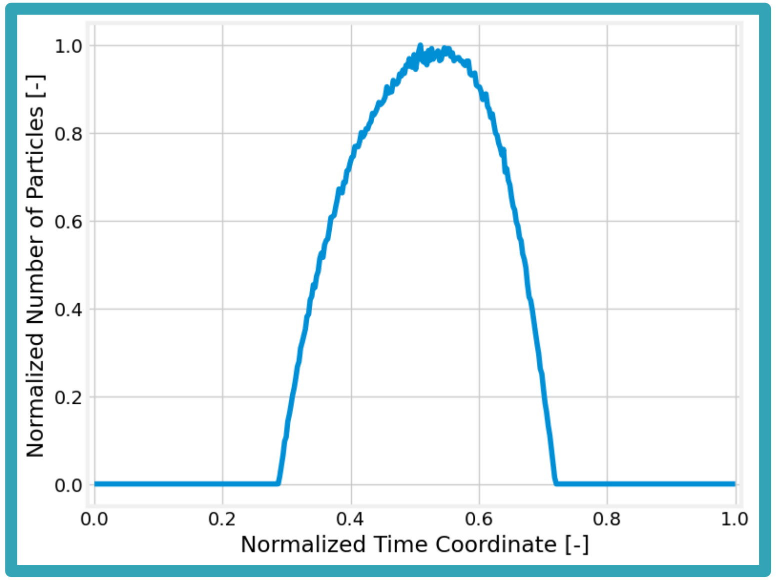
- Objective function mapped against phases for a non-accelerating bucket

Supporting Slides: Objective Function

Change in Voltage Ratio $r = \frac{3}{8}$





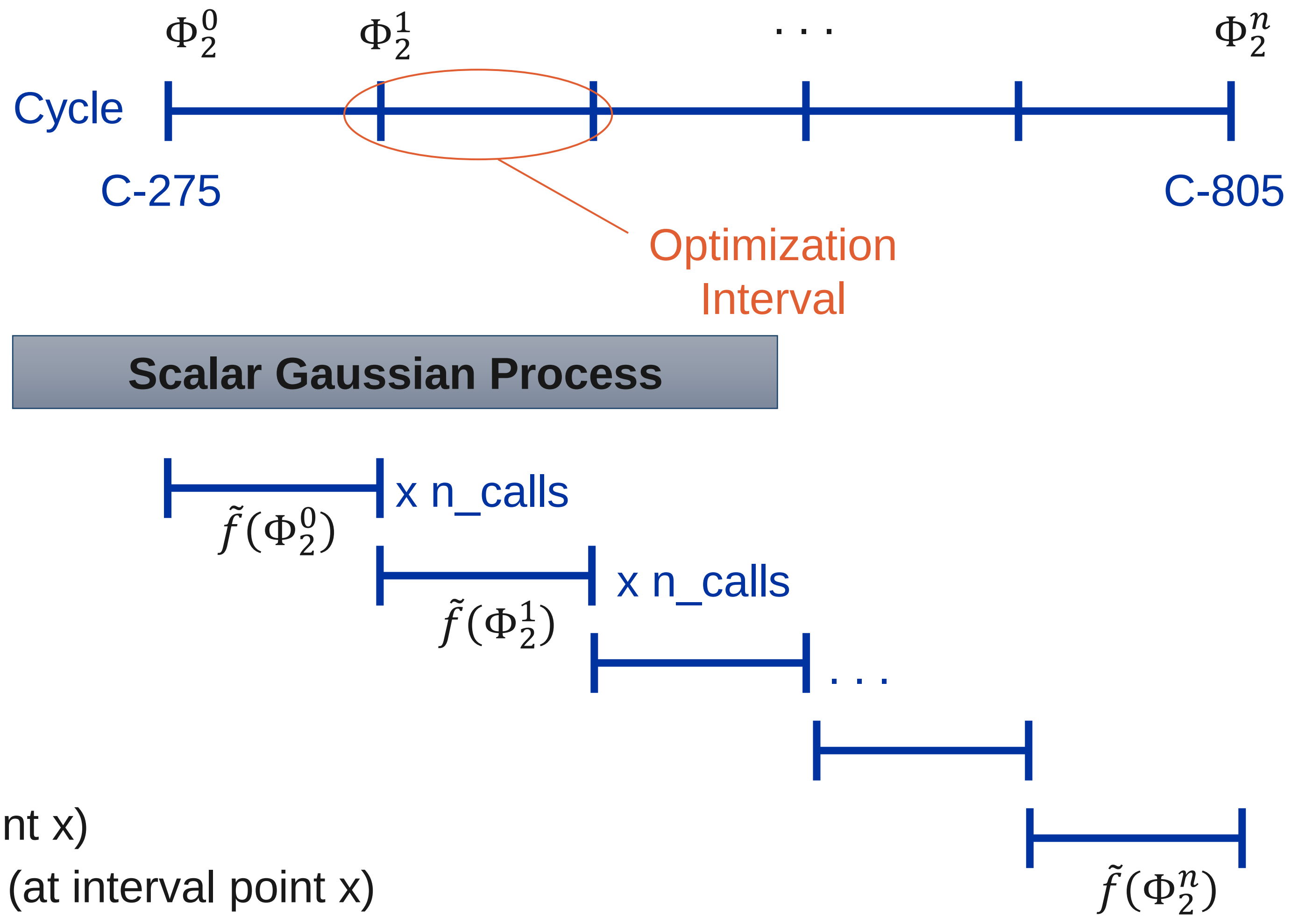


- Objective function mapped against phases for a non-accelerating bucket

Supporting Slides: Bayesian Optimization (scalar)

Modelling the objective function: Gaussian Process

- Objective function is modelled by a collection of functions and the probability distribution over those functions.
- The relation between different samples of the objective function and the variable (phase at a point) is captured by a *covariance kernel* which defines the collection of functions
- Objective function approximated at each *optimization interval*



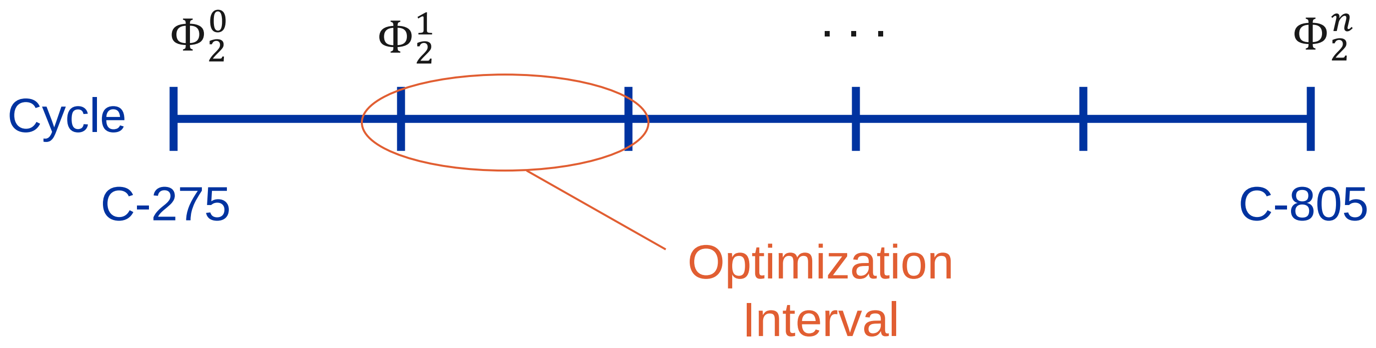
$f(\Phi_2^x)$ = Objective Function (at interval point x)

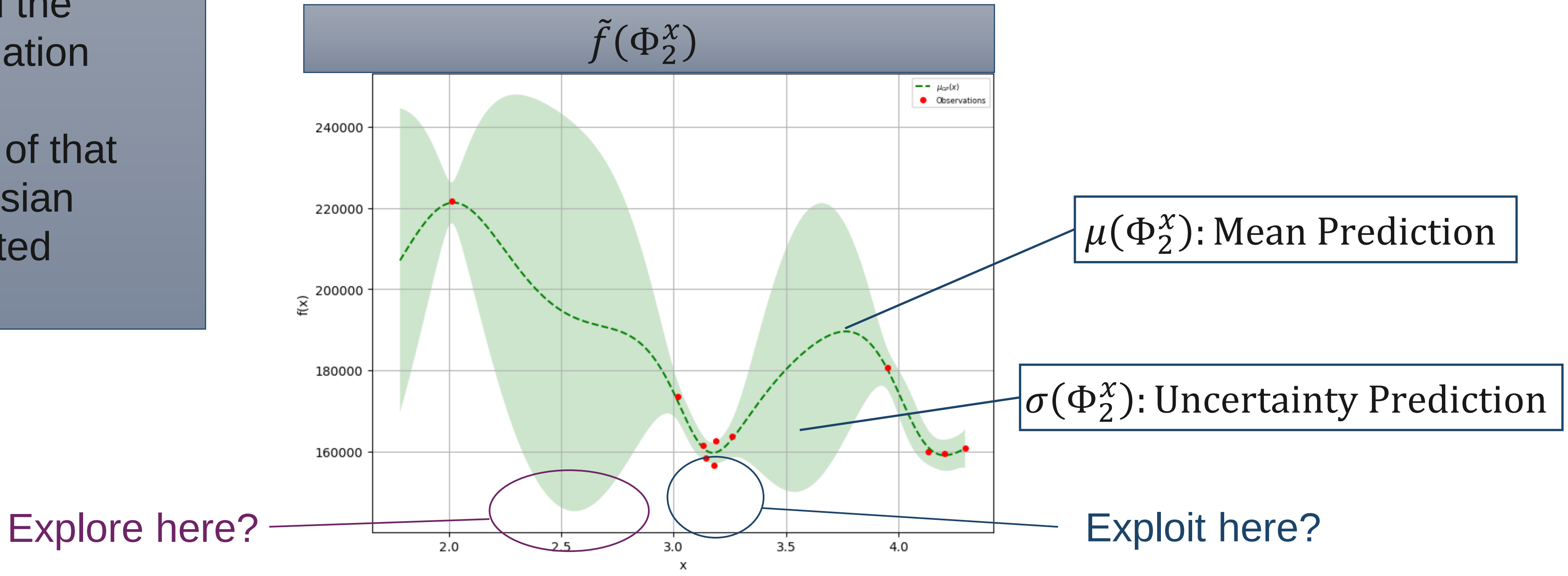
$\tilde{f}(\Phi_2^x)$ = Approximated Objective Function (at interval point x)

Supporting Slides: Bayesian Optimization (scalar)

Optimizing the objective function: Bayesian Optimization

- An *acquisition function* (Expected Improvement) determines the best sampling point by weighting the minimum of the predicted mean (exploitation) and the predicted standard deviation (exploration)
- Then the sample taken of that point updates the Gaussian Process (i.e. the predicted objective function)

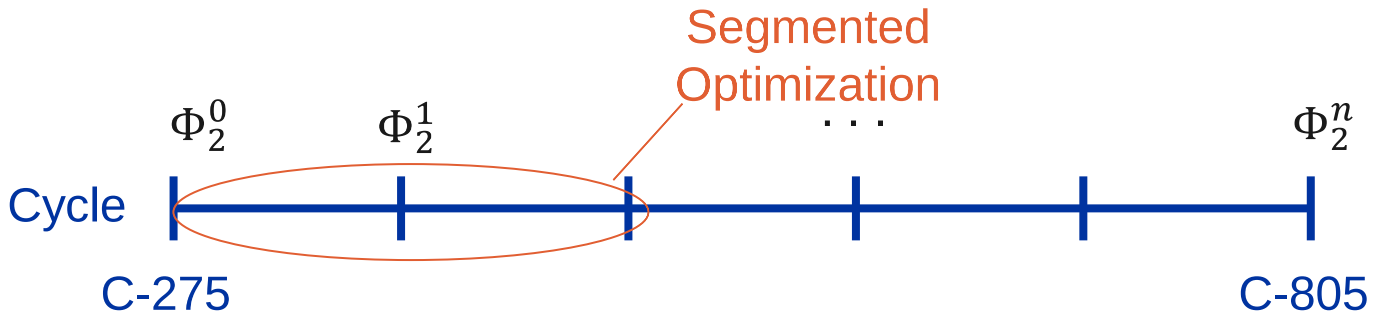




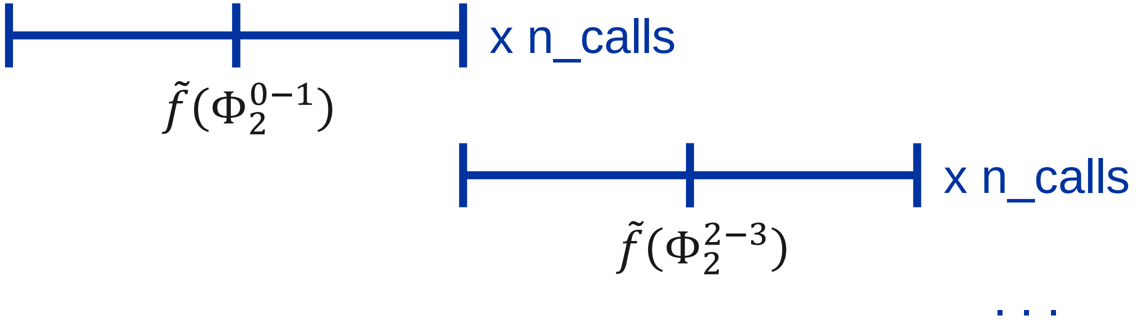
Supporting Slides: Bayesian Optimization

Modelling the objective function: Gaussian Process

- Objective function is modelled by a collection of functions and the probability distribution over those functions.
- The relation between different samples of the objective function and the variables (**2, 3, etc. phases** at points) is captured by a *covariance kernel* which defines the collection of functions
- Objective function approximated at each *segmented optimization interval*



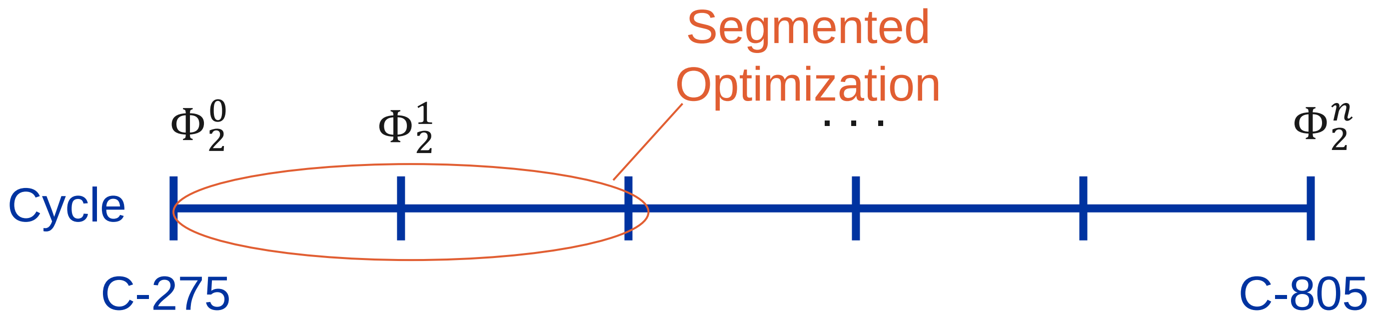
Multidimensional Gaussian Process



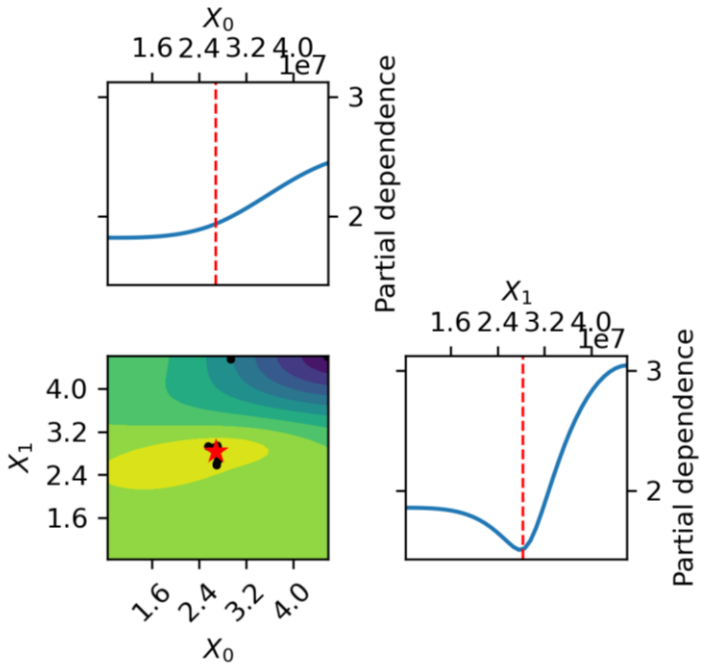
Supporting Slides: Bayesian Optimization (segmented)

Optimizing the objective function: Bayesian Optimization

- An *acquisition function* (Expected Improvement) determines the best sampling point by weighting the minimum of the predicted mean (exploitation) and the predicted standard deviation (exploration)
- Then the sample taken of that point updates the Gaussian Process (i.e. the predicted objective function)



$\tilde{f}(\Phi_2^{0-1})$



Generated by averaging out the other variable

Supporting Slides: Covariance Kernel

- Matern Kernel (scalar and multidimensional variants)

$$k_M(x, x') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}|x - x'|}{\ell} \right)^{\nu} K_{\nu} \left(\frac{\sqrt{2\nu}|x - x'|}{\ell} \right)$$

- It is a generalized version of the RBF kernel, with smoothness incorporated

Supporting Slides: Acquisition Function

- Expected improvement: we want to maximize a utility function given by:

$$u(x) = \max\big(0, f' - f(x)\big)$$

$x = \Phi_2^x$
f' = Minimum objective values **observed**
$f(x)$ = Approximated Objective Function mean

- The expected improvement is:

$$a_{\text{EI}}(x) = \mathbb{E}\big[u(x) \mid x, \mathcal{D}\big] = \int_{-\infty}^{f'} (f' - f)\,\mathcal{N}\big(f; \mu(x), K(x,x)\big)\,\mathrm{d}f$$

$$= \big(f' - \mu(x)\big)\Phi\big(f'; \mu(x), K(x,x)\big) + K(x,x)\mathcal{N}\big(f'; \mu(x), K(x,x)\big)$$

Exploration term

Exploitation term