

修士論文
Master thesis

SuperKEKB 加速器における機械学習を用いた
ビーム入射調整の高度化
(Upgrade of beam injection tuning using
machine learning at the SuperKEKB
accelerator)

令和 7 年 1 月
修士（理学）申請

東京大学大学院理学系研究科
物理学専攻
Department of Physics, Graduate School of Science,
The University of Tokyo

加藤臣之輔
Shinnosuke Kato

概要

Belle II 実験は 7 GeV の電子と 4 GeV の陽電子を衝突させる SuperKEKB 加速器を用いて B 中間子や τ 粒子をはじめとするさまざまな粒子を生成し、その崩壊を精密測定することによって新物理を探索することを目的とした実験である。SuperKEKB 加速器は 2024 年 12 月現在において世界最高の瞬間ルミノシティ記録を保持しており、さらに高いルミノシティを実現するための運転と改良が進められている。高い瞬間ルミノシティを維持するためには高いビーム電流を維持することが重要であり、そのためにはビーム入射部から蓄積リングへの入射効率、つまり入射した電流に対するリングに実際に蓄積された電流の割合を高い水準で維持することが要求される。しかし、SuperKEKB 加速器では、入射効率が数時間で低下することが多く、従来はオペレーターが手動で入射調整パラメータを操作することで効率を維持していた。この手法には高い専門技術を要するために対応可能な人材が限られており、また機器の不調からの復旧後やビーム調整の直後においては入射調整が難航することがあった。そこで本研究では、機械学習の一種であるベイズ最適化を用いて入射調整を自動化し、入射効率を改善するツールを開発した。このツールの運用により、入射効率が向上した事例が複数確認され、最も優れた運用例では入射効率を 32% 向上させることに成功した。

目次

第 1 章	序論	1
第 2 章	SuperKEKB 加速器	3
2.1	SuperKEKB 加速器	3
2.2	ルミノシティ	8
2.3	砂時計効果とナノビーム方式	10
2.4	ビーム入射部	13
第 3 章	研究の背景	16
3.1	目標ルミノシティと現在のルミノシティの比較	16
3.2	ルミノシティ向上へのアプローチ	18
3.3	入射調整に対する困難	19
第 4 章	入射調整における機械学習アルゴリズムの導入	22
4.1	入射調整パラメータ	22
4.2	入射効率	26
4.3	ベイズ最適化	27
4.4	入射調整に対する機械学習手法の導入	32
4.4.1	ベイズ最適化で用いる調整/評価パラメータの選定	32
4.4.2	本研究におけるベイズ最適化のセットアップ	35
4.4.3	GUI による入射調整ツール操作	38
第 5 章	機械学習を用いた入射調整の運用	40
5.1	運用例	40
5.2	再現性の確認	43
5.3	獲得関数、近傍探索度、ノイズ標準偏差の比較	53
5.4	運用実績	62

第 6 章	結論と展望	67
	謝辞	69
付録 A	付録	70
A.1	Belle II 測定器	70
A.2	メインリング	72
A.3	電流の寿命	74
A.4	ベイズ最適化の理論	75
A.4.1	多変量ガウス分布	75
A.4.2	線形回帰とガウス基底関数	76
A.4.3	ベイズ推定とベイズ線形回帰モデル	79
A.4.4	ガウス過程回帰法	81
A.4.5	カーネル関数	84
A.4.6	獲得関数	86
A.4.7	制限付き最適化	88
A.4.8	Optuna	89
付録 B	KEK Linac での陽電子収量最大化スタディ	90
B.1	スタディの目的と実験のセットアップ	90
B.1.1	スタディの目的	90
B.1.2	KEK Linac における手動調整手法	90
B.1.3	実験のセットアップ	91
B.2	スタディ結果	92
B.2.1	獲得関数の比較	92
B.2.2	初期化条件の比較	94
B.2.3	入射調整のために必要な機能の動作確認	96
B.3	まとめ	97
	参考文献	98

図目次

1.1	Belle 実験の積分ルミノシティと主な成果 [3]。ルミノシティが蓄積されるほど新しい物理成果が得られている。	1
-----	--	---

2.1	SuperKEKB 加速器の全体図 [7]。KEK Linac で 2 種類の電子ビームが生成され、その片方はタングステン標的に衝突して電磁シャワーを起こし、電子と陽電子が対生成される。陽電子を捕獲して 1.1 GeV まで加速させ、ダンピングリングで周回させることにより陽電子ビームの位置と運動量のばらつきを抑え、電子と陽電子の両ビームは KEK Linac でそれぞれ 7 GeV と 4 GeV まで加速され、ビーム輸送路を通過してメインリングへと入射する。	4
2.2	KEK Linac の全体図 [6]。A ~ C と 1~ 5 までの 8 つのセクターに分けられている。A セクターには電子銃が、1 セクターにはタングステン標的がある。2 セクターと 3 セクターの間にはダンピングリングがあり、5 セクター後部の射出部では SuperKEKB の電子、陽電子リングを含む 4 つの入射先がある。	5
2.3	ダンピングリングの全体図 [9]。2 セクターと 3 セクターの間に存在し、シンクロトロン放射を促すことによって陽電子のエミッタンスを減少させる役割を果たす。	5
2.4	ビーム輸送路の全体図 [10]。KEK Linac からビーム輸送路へと入射された電子と陽電子はビーム輸送路内で 90° 以上の急角度で曲げられ、メインリングへと入射する。この際にエミッタンスが増大しないようにモニターでビームの状態を確認し、マグネットでビームの軌道や広がりを調整している。	6
2.5	メインリングの概略図 [11]。富士実験棟付近から電子と陽電子を入射し、入射点の対面に位置する筑波実験棟に設置されている Belle II 測定器の中で粒子同士を衝突させる。直線部にはウィグラー電磁石と RF 空洞が設置されている。	7
2.6	衝突点付近の概略図と QCS-L/R クライオスタットの外観。 [12] (a) ではビームサイズを衝突点付近で絞り込むために QCS が Belle II 測定器の内部に挿入される設計となっていることが確認できる。(b) は建設途中の QCS と Belle II 測定器が切り離されている写真である。前方にはクライオスタットが、後方には Belle II 測定器が写っている。	7
2.7	メインリングにおけるロスモニターのアボート制御システム [13]。ロスモニターからの信号は、メインリング内に配置された 12 箇所の LCR のうち、最も近い LCR に集約される。その後、これらの信号は CCR に集約され、最終的にアボート制御のための処理が実行される。	8
2.8	砂時計効果の概略図 [14]。青線は式 (2.5) 左辺のベータ関数 $\beta_y(s)$ を表す。衝突点でのベータ関数 β_y^* をビーム進行方向のバンチ長 σ_z 以上に絞ると、 $\beta_y(s)$ は s の 2 乗に比例して膨らんでしまう。この形が砂時計に見えることから、砂時計効果と呼ばれている。	10
2.9	ナノビーム方式の説明 [15]。 $2\phi_c$ の衝突角で細長いビーム同士を衝突させることは、実効的なバンチ長 $\sigma_x / \sin \phi_c$ および実効的なバンチ幅 $\sigma_z \sin \phi_c$ を持つバンチが正面からすれ違う状況と等価である。	11
2.10	KEKB 加速器と SuperKEKB 加速器の衝突方式比較 [15]。SuperKEKB 加速器は前身となる KEBK 加速器と比較して、衝突角度を約 4 倍にしており、かつ衝突点でのベータ関数を $1/20$ まで絞ることを目標としている。	12

2.11	LER における衝突点付近でのベータ関数 [16]。衝突点でベータ関数を絞り込むためには、その手前でビームを膨らませなければならない。これによって力学口径が狭くなってしまい、結果としてビーム入射部に求められる入射要求が高くなる。	12
2.12	ビーム輸送路からメインリングへの入射調整概略図。赤線は陽電子ビームの軌道を表している。ビーム輸送路の軌道を入射バンチが通過し、メインリングの軌道を蓄積バンチが通過する。垂直ステアリングマグネットで垂直方向を、セプタムマグネット、四極マグネット、キッカーマグネットで水平方向を調整する。Stripline BPM は入射バンチの電流を、BCM と DCCT は蓄積電流を測定する。	13
2.13	入射調整で用いられるマグネットの外観	14
2.14	入射効率を測定するための蓄積電流測定装置	15
3.1	2035 年までの積分、瞬間ルミノシティの目標プロット [17]。赤線が瞬間ルミノシティ、青線が積分ルミノシティであり、2024 年 12 月の目標値は瞬間ルミノシティが $1.5 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ 、積分ルミノシティが 2 ab^{-1} である。	16
3.2	2024 年 12 月 31 日時点での積分ルミノシティのプロット [18]。赤線は積分ルミノシティの総計を、青棒は週ごとの積分ルミノシティを表している。	17
3.3	2022 年前期の運転における週ごとの最大蓄積電流 [19]。2022 年の 5 月 19 日から 26 日までの 1 週間に $\beta_{y\pm}^*$ を 0.8 mm に絞ったが、その前後の週と比べて電流を蓄積できていないことがわかる。	18
3.4	手動による入射調整事例。この事例ではオペレータが 2 時間ほど 3 つの入射調整パラメータを操作したが、入射効率は約 10% 付近で推移し改善は見られなかった。	20
3.5	1 ヶ月間の LER における、3 つの入射調整パラメータの 3 次元プロット。2024 年 11 月 16 日から 12 月 16 日までの入射調整パラメータ (y-position, y-angle, Septum Angle) を 3 次元のパラメータ空間にプロットした。(a) は時系列をカラーマップにしており、パラメータセットの変遷を追うことができる。(b) は入射効率をカラーマップにしており、高い部分と低い部分がまだらになっている。これは入射効率が高い点がパラメータ空間上で定まっていないことを表している。なお、入射調整パラメータの定義については 4 章で解説している。 . . .	21
4.1	HER での手動入射調整パネル。パネル右側では 7 つの入射調整パラメータを調整可能であり、パネル左側には設定したパラメータを基に入射の描画が表示されている。横軸は衝突点を原点としてビーム進行方向を正としたビーム進行方向の座標を表しており、単位は m である。縦軸はビームダクトの中心からの水平距離を表しており、単位は mm である。stored は蓄積バンチの軌道、injected は入射バンチの軌道を描写している。jump は入射バンチの縦軸の縮尺を蓄積バンチと合わせて描写し、二つのバンチの角度の差をわかりやすくしている。wall はセプタム隔壁の位置を、max height はビームダクトの位置を描写している。	22

4.2	ビーム垂直方向の入射説明図。2つのステアリングマグネットを用いて垂直方向のビーム軌道を操作する。ステアリングマグネット後部の入射地点での垂直距離が y-position、垂直角度が y-angle である。	24
4.3	ビーム水平方向の入射説明図 [9]。LER と HER の入射方法は共通しているため、この図では例として LER の入射部を示している。図の中心には赤色でセプタム隔壁が描写されており、隔壁の端を Injection point と定義している。これは図 4.2 の入射地点と同じ地点である。横軸は Injection point を原点とするビーム進行方向に沿った座標を表しており、縦軸はビームダクト中心を原点とした水平方向の座標を表している。青線は入射バンチと蓄積バンチの軌道を表しており、入射バンチの軌道はセプタムマグネットによってビームダクト中心と逆方向に曲げられる。これは入射バンチの水平角度を蓄積バンチの水平角度に近づけるためである。緑線で描写されているのは四極マグネットであり、両バンチの軌道はビームダクト中心方向に曲げられる。橙線で描写されている距離と角度は水平方向の入射調整パラメータである。Injection point での、ビームダクト中心からの入射バンチまでの水平距離が Septum Position、ビーム進行方向に対する水平角度が Septum Angle であり、四極マグネット中心でのビームダクト中心からの蓄積バンチまでの水平距離が Kicker Height である。	25
4.4	セプタムマグネットとビームダクトの関係図。入射バンチと蓄積バンチの水平距離が離れていると、バンチが合流して周回する際に水平振動が大きくなってしまうため、両バンチの距離はできる限りセプタム隔壁に近づけることが好ましい。	25
4.5	BCM と DCCT で測定した入射効率の比較。0 mA から入射を開始し、600 mA 程度まで蓄積させた時の BCM と DCCT での入射効率を表している。BCM では入射が始まってから 5 分後まで入射効率が表示されないことと、DCCT では蓄積電流がビーム寿命の影響を受けて徐々に低くなっていることがわかる。	27
4.6	ブラックボックス関数の概念図。関数全体の具体系は不明だが、測定を通して局所的な値は取り出すことができる。	27
4.7	1次元のガウス過程回帰の図。赤点は観測データを表しており、これらのデータに基づいて、定義域内の任意の点 ($0 \leq x^* \leq 10$) に対して正規分布の平均 (青線) および標準偏差 (水色の領域) を生成する。この x^* を定義域全体にわたって連続的に生成することで予測分布を構築している。ここで測定点の分散は $\sigma^2 = 1$ としており、ハイパーパラメータは Adam によって $\theta_1 = 23.3, \theta_2 = 1.52$ と最適化されている。	29
4.8	1次元のガウス過程回帰において、獲得関数での最大値予測を描写したプロット。UCB($\beta = 0.5$) では $x = 0.22$ が、EI では $x = 5.28$ が最も最大値らしいと予測しており、同じ予測分布でも獲得関数によって異なる最大値を取っている。	30
4.9	1次元でのベイズ最適化のデモンストレーション。初期条件を 5 点与え、予測分布を作成し、獲得関数が最大となった点の測定を繰り返す。測定点の情報が増えれば予測分布と獲得関数の両方の精度が向上するため、より精度良く最大値を探すことができる。	31

4.10	30 分間ビームの状態を変化させずに測定した入射効率。青線が入射効率、橙線が入射効率の 30 点の移動中央値を表している。橙線を見ると、入射効率の中央値自体がふらついていることがわかる。	33
4.11	30 分間ビームの状態を変化させずに測定した入射効率のヒストグラムとその正規分布フィッティング。測定誤差と入射効率自体のふらつきの効果が含まれている。フィッティングの結果、平均値は 67%、標準偏差は 2.7% であった。	34
4.12	二つの最適化手法。左図が小分け最適化を、右図が制限つき最適化を表しており、それぞれに図中に示したメリットとデメリットがある。	35
4.13	step by step 機能の概念図。本機能がないと、調整パラメータが急激に変化してしまい、ビームの振幅を増大させてしまう（赤点線）。本機能を導入することにより、調整パラメータの急激な変化を防ぎ、振幅の増大を緩和することができる（青線）。	37
4.14	近傍探索機能の概念図。2 次元の調整パラメータ空間において、近傍探索機能の有無に対する軌跡の違いを表している。左図ではトライアルごとにパラメータが大きく変動し、step by step 機能による待機時間が多く発生するが、右図ではトライアルあたりの変動が小さいために待機時間を削減することができる。	38
4.15	最適化ツールの GUI メインページ上部。Parameter settings では調整パラメータの使用の可否や定義域を設定できる。Initialization settings ではトライアル数や初期化の回数を設定することや、過去のコンフィギュレーションを参照することができる。Optimization buttons では最適化の開始、一時停止、再開などの操作ができる。Expert Settings には step by step 機能の幅や近傍探索機能のハイパーパラメータなど、普段の調整ではオペレーターが変更しないパラメータを変更することができる。	39
4.16	最適化ツールの GUI メインページ下部。テスト関数を用いて最適化を行っており左図はトライアル数に対する入射パラメータの変動を、右図ではトライアル数に対する入射効率（青点）とその最大値（橙点）を見ることができる。制限つき最適化では、ロスモニターで測定したビームロスレート（灰線）も表示される。制限は複数課すことができ、この図では TEST:C0 と TEST:C1 という 2 つのビームロスレートを描写している。	39
5.1	12 月 5 日の LER での 2 回にわたる小分け最適化結果。左図は各トライアルにおける調整パラメータの変動を表しており、右図はトライアルごとの入射効率を表している。青線が入射効率、橙線が最適化途中での入射効率の最大値である Peakhold を表している。また、紫点は最大値を記録したトライアルに付加されるマーカーである。y-position と y-angle を調整前の値から大きくし、Septum Angle が調整前の値から小さくすることで、入射効率が向上していることがわかる。	41

5.2	12 月 5 日の LER での 2 回にわたる小分け最適化での 3 次元プロット。調整パラメータ空間における最適化の軌跡を表しており、カラーマップは入射効率を表す。y-position と y-angle が調整前の値から上昇し、Septum Angle の値が調整前の値から減少するほど入射効率が向上していることが見て取れる。また、入射効率の分布が連続的であるため、局所的な最大値に捕捉されやすい小分け最適化を用いた場合でも入射効率が 100% に近いパラメータセットに到達できていることが分かる。	42
5.3	12 月 25 日の LER での小分け最適化結果。式 (4.5) で表される 2 バンチ用の入射効率を 55% から 68% まで向上させた結果、1 バンチ目の入射効率が 68% から 82% に、2 バンチ目の入射効率が 37% から 55% に向上した。	43
5.4	HER における、ビーム状態、定義域の条件を揃えた 3 回の測定結果。3 回の測定に共通して、調整前の値に対して y-angle と Septum Angle を大きくするほど入射効率が向上している。一方で y-position は 3 回の試行において決まった位置を取っていない。	46
5.5	HER での 3 回の最適化における全ての点と、各測定における上位 3 点の 3 次元プロット。(a) では y-position の変化に対して、入射効率があまり変わらないことが見て取れる。(b) では 1 回目と 3 回目の最適化では y-angle と Septum Angle の値が同じだが、y-position のみが異なることが確認でき、また 2 回目の最適化では 1 回目と 3 回目の最適化と比較して入射効率が低く、y-angle と Septum Angle が定義域の上端にたどり着く前に最適化が終了してしまっていることが見て取れる。	47
5.6	LER における、ビーム状態、定義域の条件を揃えた 2 回の測定結果。2 回の測定に共通して、調整前の値に対して y-angle と Septum Angle 小さくするほど入射効率が向上している。一方で y-position は 2 回の試行において決まった位置を取っていない。	49
5.7	LER での 2 回の最適化における全ての点と、各測定の上位 3 点の 3 次元プロット。図 5.5 と同様に、(a) では y-position の変化に対して、入射効率があまり変わらないことが見て取れ、(b) では 2 つの最適化の上位 3 点に対して y-angle と Septum Angle の値が同じだが、y-position のみが違うことが確認できる。	50
5.8	HER と LER の最適化における調整パラメータと Shapley 値のプロット。カラーマップがそれぞれのパラメータ値を示し、Shapley 値が入射効率の向上または悪化に寄与した度合いを表している。(a) では Septum Angle および y-angle の値が高い場合には入射効率が向上し、低い場合には効率が悪化していることが分かる。一方、(b) では Septum Angle および y-angle の値が低い場合に入射効率が向上していることが確認できる。また、両方のケースにおいて、y-position は入射効率の変動に対してあまり寄与していないこともわかる。	51
5.9	HER と LER の最適化における SHAP importance。Septum Angle および y-angle に比べて、y-position の寄与が相対的に小さいことが明らかとなった。また、HER では Septum Angle が、LER では y-angle が最も高い寄与を示していることがわかる。	52

5.10	獲得関数 EI、近傍探索度 5 に固定し、ノイズ標準偏差を 0 と 0.1% に設定した時の比較。(a) では最適化が活用的になりすぎたため、探索されない定義域内の領域が存在したまま 40 トライアルが終了している。(b) では探索的な成分が強く確認できる。	55
5.11	獲得関数 EI、近傍探索度 5 に固定し、ノイズ標準偏差を 1% と 3% に設定した時の比較。(a) では、初期化直後の y-position および y-angle に対して 0.1% の場合と同様に活用的な最適化が見られるものの、入射効率が向上すると確信した段階で探索的な挙動へ移行する傾向が見られた。(b) では初期化直後の活用的な傾向が 1% の場合よりも抑えられ、入射効率が向上した場合における探索的な動きも抑制されている。	56
5.12	近傍探索度 5、ノイズ標準偏差 0 の UCB($\beta = 2$)。図 5.10b と比較して、活用的な成分を重視した最適化が行われていることを確認できる。	57
5.13	近傍探索度 1、ノイズ標準偏差 0 の UCB ($\beta = 2$)。図 5.12 と比較すると、各トライアル間におけるパラメータの変化量が増加していることが確認できる。	58
5.14	ノイズ標準偏差 0、獲得関数を UCB ($\beta = 2$) に固定し、近傍探索度を 1 および 5 に設定した場合のトライアルごとのパラメータ変化量の絶対値。(a) は近傍探索度 1、(b) は近傍探索度 5 の結果であり、近傍探索度 1 の場合の方が、5 の場合と比較してパラメータの変化量が大きいことが確認できる。	59
5.15	ノイズ標準偏差 1、獲得関数を EI に固定し、近傍探索度を 0 および 5 に設定した場合の比較。(a) では入射効率の有意な改善は見られるが、定義域の端を繰り返し探索する挙動が確認され、適切な最適化が行われていない。(b) では正常に最適化が行われ、入射効率の有意な改善が見られる。	60
5.16	ノイズ標準偏差を 1、獲得関数を EI に固定し、近傍探索度を 0 および 5 に設定した場合のトライアルごとのパラメータ変化量の絶対値。(a) では、定義域の端から端まで大きく変化するトライアルが散見される。(b) ではパラメータの変動が抑えられていることが確認できる。	61
5.17	LER における入射効率が最大となった調整パラメータに対する 3 次元プロット (a) と共分散行列 (b)。(a) では各パラメータの定義域と比較して、y-angle の変動幅が最も小さい一方で、Septum Angle の変動幅が最も大きいことがわかる。(b) では y-position と y-angle の間に強い負の相関が存在し、y-position と Septum Angle の間に弱い負の相関があり、y-angle と Septum Angle の間には相関がほとんど見られないことがわかる。	64
5.18	2024 年 11 月 16 日から 12 月 16 日までの y-position, y-angle, Septum Angle の共分散行列。y-position と y-angle の間に強い負の相関が存在し、y-position, y-angle と Septum Angle の間には相関がほとんど見られないことがわかる。	65
5.19	ふらつき補正機能の概略図。この図は 2 次元のパラメータ空間を例としており、数トライアルに 1 回、基準点での入射効率を測定し、最適化開始前に測定した基準点の測定値をそのトライアルでの基準点の測定値で割った値を入射効率に乘じることで、ふらつきの影響を補正する仕組みを示している。図中では、各トライアルに対して補正を適用する過程が描かれている。	66

A.1	Belle II 測定器の全体図 [31]	70
A.2	コリメーターの概略図とメインリング内でのコリメーター配置 [33]	72
A.3	SBL によって損傷したコリメーター [34]	73
A.4	LER でのフィルパターンと各バンチの電流量	73
A.5	1 次元のガウス過程回帰の図。赤点が測定点であり、そこから平均 (青線) と分散 (水色の領域) を作成している。後述するカーネル関数は RBF カーネルを使用している。	83
A.6	制限付きベイズ最適化のフローチャート	89
B.1	KEK Linac における陽電子収量最大化のセットアップ	91
B.2	本研究でのスタディで用いられる装置の外観	92
B.3	EI の各トライアルに対する BPM の電荷量 (nC)	93
B.4	UCB ($\beta = 2$) の各トライアルに対する BPM の電荷量 (nC)	93
B.5	UCB ($\beta = 5$) の各トライアルに対する BPM の電荷量 (nC)	94
B.6	UCB ($\beta = 2$) と UCB ($\beta = 0.1$) の 5 ランに対する Peakhold の比較	94
B.7	UCB($\beta = 2$) の 1 ラン目のペアプロット	95
B.8	UCB($\beta = 2$) の 4 ラン目のペアプロット	95
B.9	初期化方法の説明	96
B.10	UCB ($\beta = 2$) での初期化条件に対する Peakhold の比較	96
B.11	近傍探索機能の実装の有無に対するパラメータのトライアルごとの差分比較	97

表目次

3.1	SuperKEKB の目標値と現在の値の比較	17
4.1	入射調整パラメータの説明	23
4.2	BCM と DCCT の比較	26
4.3	パラメータを固定した条件下で測定した入射効率の平均値と標準偏差	34
4.4	各パラメータについての定義域の幅	36
4.5	各パラメータについての 1 秒あたりの移動距離	37
5.1	12 月 5 日の LER 入射調整前後での入射効率、蓄積電流、瞬間ルミノシティの比較	42
5.2	12 月 25 日の LER 入射調整前後での入射効率、蓄積電流、瞬間ルミノシティの比較	44
5.3	HER における再現性確認スタディのセットアップ	44
5.4	HER における再現性確認のスタディ結果	45
5.5	LER における再現性確認スタディのセットアップ	48
5.6	LER における再現性確認のスタディ結果	49
5.7	最適化スタディにおける SHAP importance の解析結果	53
5.8	獲得関数、近傍探索度、ノイズ標準偏差の測定条件表	54
5.9	入射調整ツール運用における日ごとの最適化結果	62
5.10	入射調整ツール運用における日ごとのパラメータ重要度	63
5.11	LER における日ごとの入射効率が最大となった調整パラメータ	63
5.12	HER における日ごとの入射効率が最大となった調整パラメータ	64
B.1	KEK Linac における調整順序	90

第 1 章

序論

素粒子物理学における素粒子の振る舞いは標準模型によって記述されており、実験での観測結果はそのほとんどが標準模型と無矛盾である。一方で標準模型は万物を記述する究極の理論ではなく、現在のエネルギースケールで近似的に成り立っている理論であると考えられる。標準模型を超える新しい物理モデルについては確固たる結論が出ておらず、新物理の兆候を発見することが、現在の素粒子実験における重要な課題の一つである。Belle II 実験では、新粒子が量子効果によって反応に寄与し、物理量の測定値が標準模型の予測から微小にずれることを期待し、物理量を精密測定することによって高エネルギー領域の新粒子を間接探索している。

Belle II 実験の前身となる Belle 実験では、 $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$ 崩壊における時間に依存した CP 対称性の破れを測定することにより小林・益川理論 [1] の正当性を実証した [2]。また、Belle 実験では B 中間子崩壊の他にも、様々な物理成果を得ることに成功している。図 1.1 に、統計量に比例する物理量である積分ルミノシティに対する主な Belle 実験の成果を紹介する。積分ルミノシティの蓄積によって統計誤差が減少し、物理量

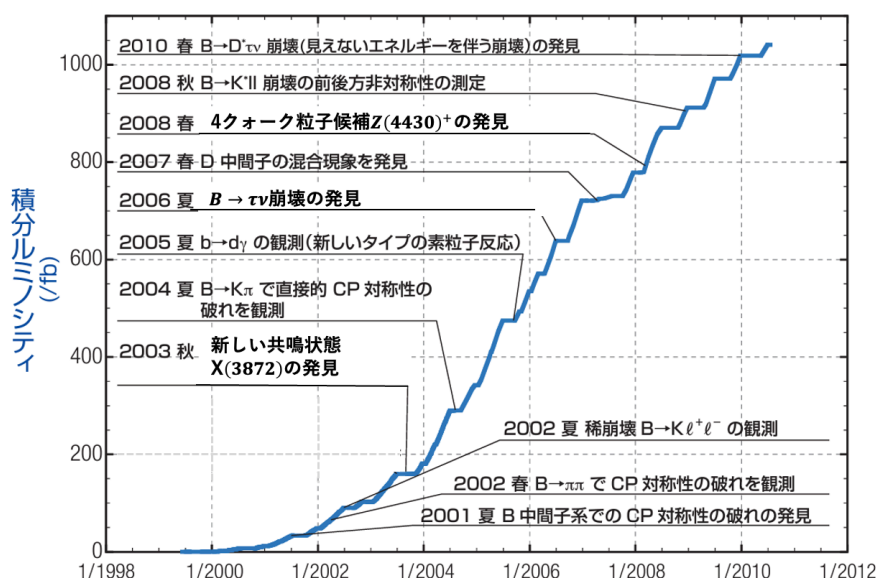


図 1.1: Belle 実験の積分ルミノシティと主な成果 [3]。ルミノシティが蓄積されるほど新しい物理成果が得られている。

の精密測定が可能になることで新しい物理成果が得られることが見て取れる。Belle II 実験では、積分ルミノシティを Belle 実験の 50 倍である 50 ab^{-1} まで蓄積することにより、Belle 実験よりも高度な精密測定が要求される物理事象を測定することができる。物理事象の具体例を以下に記す [4] [5]。

- ・ 時間依存する CP 対称性の破れの測定 ($B \rightarrow \phi K^0, \eta' K^0$ など)
- ・ レプトンフレーバーの破れの探索 ($\tau \rightarrow \mu\gamma, \mu\mu\mu$ など)
- ・ レプトンフレーバー普遍性の破れの探索 ($R(D^{(*)}) = \frac{\mathcal{B}(B \rightarrow D^{(*)}\tau\nu)}{\mathcal{B}(B \rightarrow D^{(*)}\ell\nu)}, (\ell = e, \mu)$ など)
- ・ 稀崩壊の測定 ($B \rightarrow \tau\nu$ など)
- ・ ダークマターの探索 ($e^+e^- \rightarrow \gamma A', \gamma \text{ALP}$ (A' : ダークフォトン, ALP: Axion-Like 粒子) など)
- ・ CKM 行列要素の精密測定 ($B \rightarrow D^{(*)}\ell\nu, \pi\ell\nu$ など)
- ・ QCD 理論の検証 ($B^0 \rightarrow \pi^+\pi^-$ など)

上記のように、Belle II 実験の長所はさまざまな物理対象を観測できることである。また、電子陽電子衝突では陽子陽子衝突の素粒子実験に比べてバックグラウンドが少なくクリーンな環境であるため、中性粒子を含む崩壊モードを陽子陽子衝突実験より精密に測定でき、Belle II 実験のみでしか精密測定できない固有の崩壊モードがある点も長所として挙げられる。なお、Belle II 測定器については付録 A.1 に記した。

Belle II 実験において電子と陽電子の衝突を担う SuperKEKB 加速器は 2024 年 12 月時点において世界最高の瞬間ルミノシティ記録を保持しており、さらに高いルミノシティを実現するための運転と改良が進められている。高い瞬間ルミノシティを維持するためには高いビーム電流を維持することが重要であり、そのためにはビーム入射部から蓄積リングへの入射効率、つまり入射した電流に対するリングに実際に蓄積された電流の割合を高い水準で維持することが要求される。しかし、SuperKEKB 加速器では、入射効率が数時間で低下することが多く、従来はオペレーターが手動で入射調整パラメータを操作することで効率を維持していた。この手法には高い専門技術を要するために対応可能な人材が限られており、また機器の不調からの復旧後やビーム調整の直後においては入射調整が難航することがあった。そこで本研究では、機械学習の一種であるベイズ最適化を用いて入射調整を自動化し、常に高い入射効率を維持するツールを開発した。

第 2 章では、本研究の前提知識となる SuperKEKB 加速器について解説する。第 3 章では研究の背景について述べる。第 4 章では、入射調整における機械学習アルゴリズムの導入方法について述べる。第 5 章では、機械学習を用いた入射調整の運用事例、スタディ結果、および運用実績について報告する。最後に、第 6 章では本研究の結論と展望をまとめる。

第 2 章

SuperKEKB 加速器

本章では、SuperKEKB 加速器およびその運転に関連する重要な要素について概説する。2.1 節では SuperKEKB 加速器の概要を解説し、2.2 節においてルミノシティの定義を説明する。2.3 節ではルミノシティを向上させるための衝突方式であるナノビーム方式について説明し、2.4 節では高ルミノシティ実現のために重要なコンポーネントであるビーム入射部を紹介する。

2.1 SuperKEKB 加速器

SuperKEKB 加速器は茨城県つくば市にある電子陽電子加速器であり、2024 年 12 月時点で世界最高の瞬間ルミノシティ記録 ($4.7 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$) を保持している。メインリングは周長 3 km の蓄積リングであり、7 GeV の電子を蓄積する HER (High Energy Ring) と 4 GeV の陽電子を蓄積する LER (Low Energy Ring) で構成されている。HER は時計回りに、LER は反時計回りにメインリングを周回し、Belle II 測定器内で衝突する。図 2.1 に SuperKEKB 加速器の全体図を記す。

電子と陽電子が生成されてから衝突するまでの概要を述べる。KEK Linac の最上流において、熱陰極 RF 電子銃で HER 周回用の電子が、熱陰極電子銃で陽電子生成用の電子が生成される [6]。その後生成された電子は加速管によって加速されていき、陽電子生成用の電子はタングステン標的に衝突し陽電子が生成される。その後、1.1 GeV まで加速された陽電子は、ダンピングリングを周回することによってビームの位置と運動量のばらつきが揃えられ、再び KEK Linac へと合流する。KEK Linac の射出部では電子は 7 GeV、陽電子は 4 GeV まで加速される。その後粒子はビーム輸送路を通過し、メインリングへと入射し、リング上を周回する。ビームを構成する粒子が密集している小さな塊のことをパンチと呼び、メインリングに入射する 1 パンチあたりの電荷は 4 nC を目標値としている。メインリング内に入射した粒子は周回を続けながら Belle II 測定器の中で衝突し、B 中間子や τ 粒子をはじめとするさまざまな粒子を生成する。

SuperKEKB 加速器では衝突点付近でビームサイズを絞り、KEKB 加速器の衝突角 (22 mrad) と比べて 4 倍程度大きい衝突角 (83 mrad) で衝突させるナノビーム方式によって高い瞬間ルミノシティを達成している。また、ビームを連続で入射し、蓄積電流を一定に維持したまま電子と陽電子を衝突させ続けるトップアップ入射方式を採用することにより、高い積分ルミノシティを獲得している。以降では SuperKEKB 加速器を構成するコンポーネントである、KEK Linac、ダンピングリング、ビーム輸送路、メインリングについて解説する。

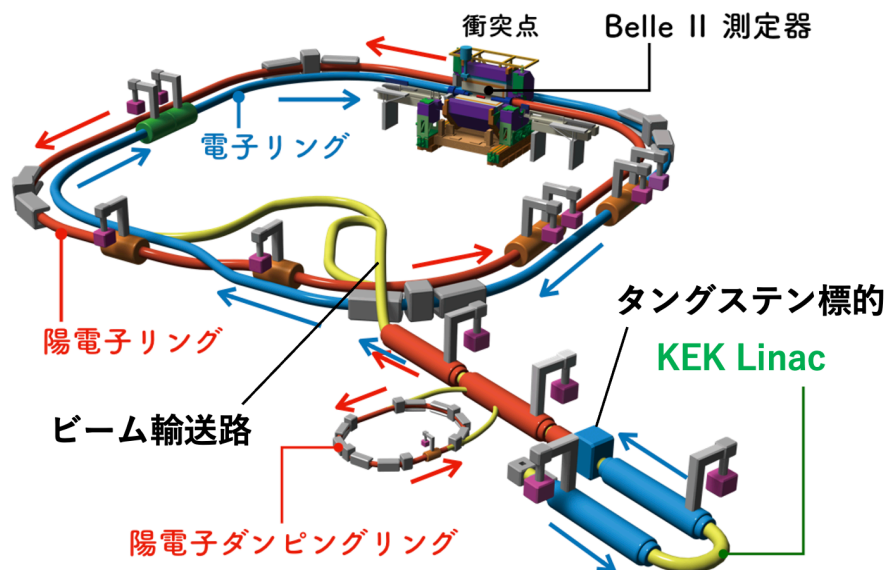


図 2.1: SuperKEKB 加速器の全体図 [7]。KEK Linac で 2 種類の電子ビームが生成され、その片方はタングステン標的に衝突して電磁シャワーを起こし、電子と陽電子が対生成される。陽電子を捕獲して 1.1 GeV まで加速させ、ダンピングリングで周回させることにより陽電子ビームの位置と運動量のばらつきを抑え、電子と陽電子の両ビームは KEK Linac でそれぞれ 7 GeV と 4 GeV まで加速され、ビーム輸送路を通過してメインリングへと入射する。

■KEK Linac KEK Linac の全体図を図 2.2 に記す。KEK Linac は 600 m 程度のビームラインであり、主に加速管、マグネット、測定器から構成されている。KEK Linac は A-C と 1-5 までの 8 つのセクターに分けられている。セクターは加速ユニットと呼ばれる構造を 8 つ並べて構成されており、加速ユニットは 4 つの加速管を並べて構成されている。例外として、A セクターでは電子銃が設置されているために加速ユニットが 4 つで構成されていたり、1 セクターでは陽電子生成用標的があるために標的の後部には陽電子を捕獲するための構造があったりと各セクターごとに細かな違いは存在する。B セクターと C セクターの間には R セクターと呼ばれるビームの向きを 180° 変化させる構造があり J-arc と呼ばれている。2 セクターと 3 セクターの間には陽電子用のダンピングリングがあり、これは次節で解説する。5 セクターの後部にある射出部では 4 つの入射先があり、入射タイミングを制御することによって、1 つの線型加速器で SuperKEKB の電子、陽電子リングを含む複数の施設の同時運転を可能にしている [8]。

■ダンピングリング ダンピングリングの全体図を 2.3 に記す。ダンピングリングの目的は、陽電子ビームを低エミッタンス化させることである。エミッタンスとはビームの位置と運動量のばらつきに係る物理量であり、エミッタンスが大きいほどビームが広がってってしまうため、一般にレプトンコライダーでは低エミッタンスなビームが好ましいとされる。しかし KEK Linac での陽電子ビームは電子ビームをタングステン標的に衝突させ、電磁シャワーによって生成された低エネルギーの陽電子を捕獲することによって生成する。そのため生成後の陽電子ビームは幅広い位置に幅広い運動量を持ったビーム、つまりエミッタンスが大きい

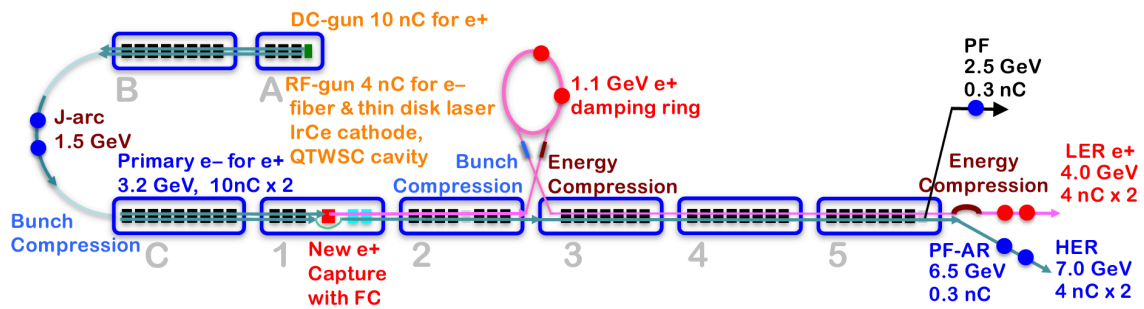


図 2.2: KEK Linac の全体図 [6]。A ～ C と 1～ 5 までの 8 つのセクターに分けられている。A セクターには電子銃が、1 セクターにはタングステン標的がある。2 セクターと 3 セクターの間にはダンピングリングがあり、5 セクター後部の射出部では SuperKEKB の電子、陽電子リングを含む 4 つの入射先がある。

ビームとなる。エミッタンスを減らすための方法としては、陽電子を曲げてシンクロトロン放射を発生させることが挙げられる。この放射が起こるとトランスヴァース方向 (ビーム進行方向に対して水平、垂直な 2 方向) の運動量を減らすことができ、また運動量が減少すれば粒子の位置も徐々に設計軌道に近づくため、トランスヴァース方向のエミッタンスを小さくすることができる。シンクロトロン放射は直線部では発生しないので、陽電子ビームをダンピングリングに入射して周回させることによって、シンクロトロン放射を発生させる。また、シンクロトロン放射で失ったエネルギーはダンピングリング内の RF 空洞でビーム進行方向に加速させることによって補う。これによりビーム進行方向のエミッタンスも小さくすることができる。

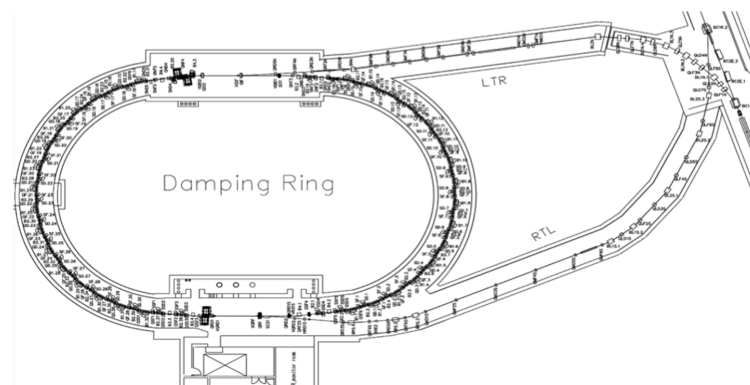


図 2.3: ダンピングリングの全体図 [9]。2 セクターと 3 セクターの間に存在し、シンクロトロン放射を促すことによって陽電子のエミッタンスを減少させる役割を果たす。

■ビーム輸送路 ビーム輸送路は KEK Linac とメインリングの間にある区間である。図 2.4 にビーム輸送路の全体図を示す。電子と陽電子の両方において、ビーム輸送路は 90° 以上の角度で曲げられており、一般にビーム軌道を大きく曲げるほどエミッタンスを維持するためには高度な調整が要求されるため、ビーム輸送路ではビームの位置や運動量を測定するビーム位置モニターやスクリーンモニターを用いてビームの状態を確認

し、マグネットでビームの軌道や広がりを見微調整しながら、エミッタンスの増大を抑えつつメインリングへとビームを入射している。

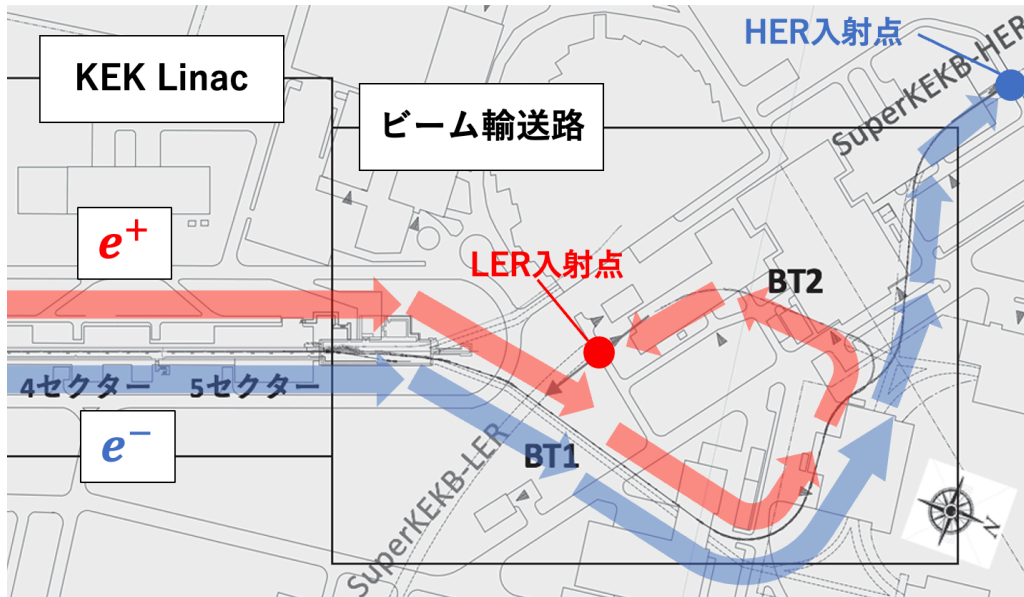


図 2.4: ビーム輸送路の全体図 [10]。KEK Linac からビーム輸送路へと入射された電子と陽電子はビーム輸送路内で 90° 以上の急角度で曲げられ、メインリングへと入射する。この際にエミッタンスが増大しないようにモニターでビームの状態を確認し、マグネットでビームの軌道や広がりを見調整している。

■**メインリング** メインリングの概略図を図 2.5 に示す。メインリングは、直線部とアーク部からなる構造を 4 つ繰り返して構成されており、各直線部の中心付近にはそれぞれ実験棟がある。Belle II 測定器は筑波実験棟に設置されており、衝突点の対面に位置する富士実験棟付近から電子と陽電子をそれぞれ入射する。大穂実験棟と日光実験棟にはウィグラー電磁石があり、シンクロトロン放射によってビームのエミッタンスを向上させる。また、ビーム周回で失ったエネルギーを回復するために、日光実験棟では超伝導加速空洞が、富士実験棟と大穂実験棟には常伝導加速空洞が設置されている。

以下では、後述するナノビーム方式に必要なコンポーネントである最終収束用超伝導電磁石と、3 章での入射調整に対する問題を理解するのに必要な知識であるアポートシステムについて述べる。また、その他のメインリングにおいて重要な要素については付録 A.2 に記した。

・最終集束用超伝導電磁石 (QCS)

QCS は、衝突点でビームサイズを絞り込むための超伝導電磁石であり、ナノビーム衝突方式の実現に欠かせない役割を果たす。2 台のクライオスタット（超伝導を維持するためにマグネットを冷却する装置）内に合計 55 台の超伝導電磁石が格納され、Belle II 測定器の両側面から内側に挿入されている。図 2.6 に衝突点付近の概略図と QCS クライオスタットの外観を示す。図 2.6a を見ると、ビームサイズを衝突点付近で絞り込むために QCS が Belle II 測定器の内部に挿入される設計となっていることが確認できる。図 2.6b は建設途中の写真であり、QCS と Belle II 測定器が切り離されている。前方にはクライオスタットが、後方には

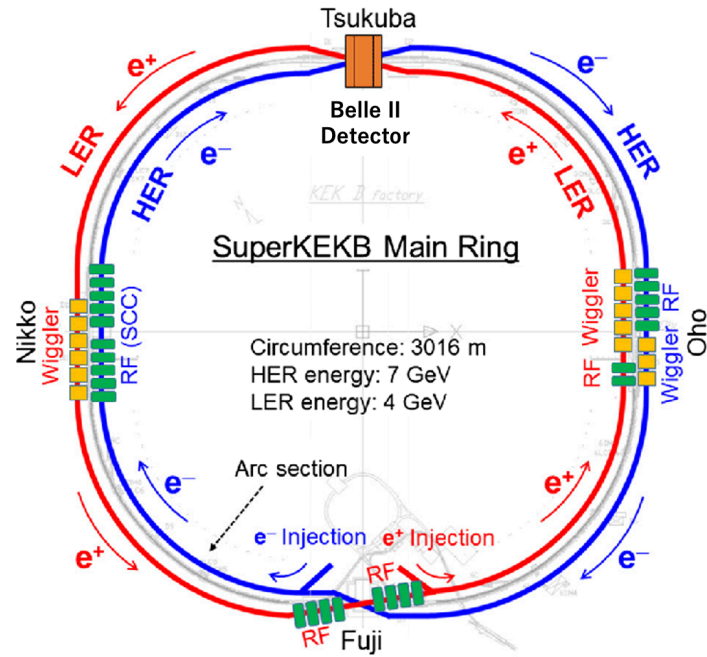
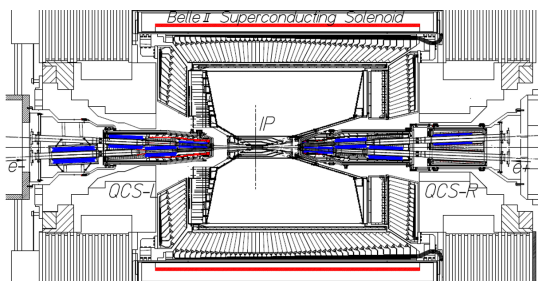
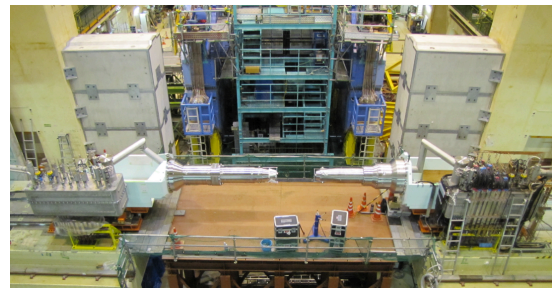


図 2.5: メインリングの概略図 [11]。富士実験棟付近から電子と陽電子を入射し、入射点の対面に位置する筑波実験棟に設置されている Belle II 測定器の中で粒子同士を衝突させる。直線部にはウィグラー電磁石と RF 空洞が設置されている。

Belle II 測定器が写っている。



(a) 概略図



(b) 外観

図 2.6: 衝突点付近の概略図と QCS-L/R クライオスタットの外観。 [12] (a) ではビームサイズを衝突点付近で絞込み込むために QCS が Belle II 測定器の内部に挿入される設計となっていることが確認できる。(b) は建設途中の QCS と Belle II 測定器が切り離されている写真である。前方にはクライオスタットが、後方には Belle II 測定器が写っている。

QCS 付近に大きなビームバックグラウンドが発生すると、電磁石の超伝導状態が一部解けてしまう。これにより抵抗が有限になることでジュール熱が発生し、電磁石が全て通常の導体になってしまうことをクエンチと呼ぶ。クエンチが発生してしまうと加速器運転の復旧には数時間を要してしまうため、コリメーター (付録 A.2) を用いて QCS 付近でバックグラウンドが大きくなるようにしている。

・アボートシステム

蓄積リング内において、ビームの一部が設計された軌道から外れ、加速器内の構造物に衝突して失われる現象をビームロスと呼ぶ。ビームロスが大きくなると装置への損傷や運転への影響が懸念されるため、事前にビームを意図的に加速器から排出する手法が採用されており、これをアボートと呼ぶ。メインリング内では、ビームロスを未然に防ぐためにロスモニターが設置されており、ビームロスが増大した際には迅速にアボートトリガー信号を出力する仕組みが構築されている。メインリングで使用するロスモニターの種類としては、ピンフォトダイオード、イオンチェンバー、光ファイバー、ダイヤモンドセンサー、シンチレーターを用いたセンサーなど多種多様なセンサーが導入されており、これらを用いて機器の損傷を防ぐための保護機構が構築されている。ロスモニターの信号はメインリング内に 12 箇所あるローカル制御室 (LCR) に集約され、その後中央制御室 (CCR) に集約される。図 2.7 にメインリングにおけるロスモニターのアボート制御システムを示す。

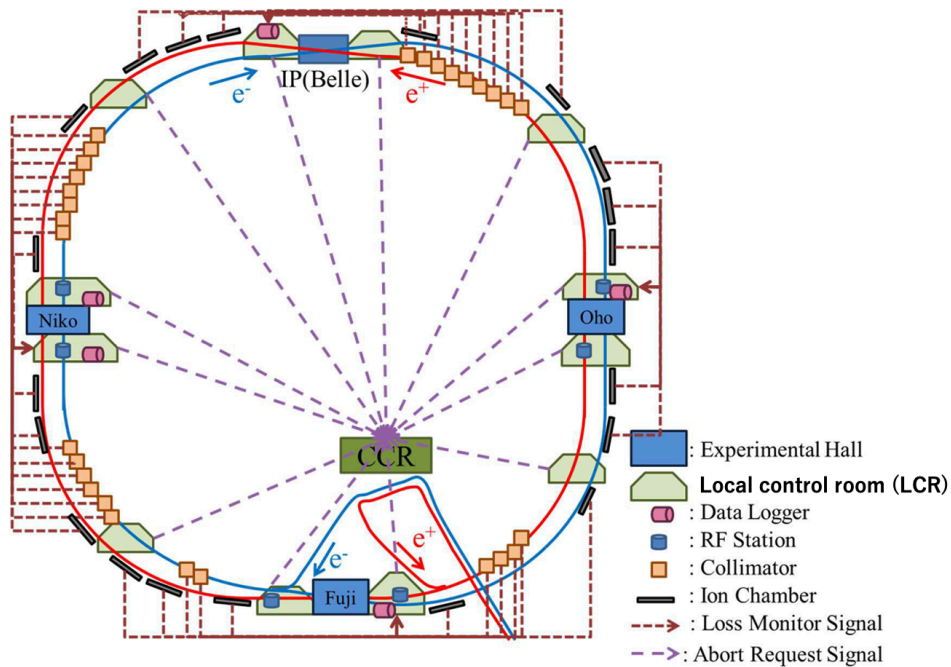


図 2.7: メインリングにおけるロスモニターのアボート制御システム [13]。ロスモニターからの信号は、メインリング内に配置された 12 箇所の LCR のうち、最も近い LCR に集約される。その後、これらの信号は CCR に集約され、最終的にアボート制御のための処理が実行される。

2.2 ルミノシティ

衝突型加速器実験において、反応断面積 σ の物理反応が起こる回数 N は次式で表される。

$$N = \sigma \int \mathcal{L} dt \quad (2.1)$$

ここで、 $\mathcal{L}$ は瞬間ルミノシティと呼ばれ、その時間積分が積分ルミノシティと定義される。統計量を増やすためには、瞬間ルミノシティを増加させるか、積分する時間を延ばす必要がある。本節では、ルミノシティの定義を説明するとともに、瞬間ルミノシティを向上させるために考案された衝突アプローチについて述べる。

衝突型加速器における瞬間ルミノシティは式 (2.2) のように表される。

$$\mathcal{L} = \frac{N_- N_+ n_b f_0}{4\pi \sigma_x^* \sigma_y^*} \quad (2.2)$$

ここで、 $+$ と $-$ はそれぞれ陽電子ビームと電子ビームのパラメータであることを表し、 N はバンチ内の粒子数、 n_b はバンチ数、 f_0 は周回周波数 (SuperKEKB 加速器では約 99.4 kHz)、 σ_x^* と σ_y^* は衝突点での水平方向と垂直方向のビームサイズである。以降は*の標識は衝突点における物理量であることを表す。また、式 (2.2) では陽電子と電子のビームサイズが同じことを仮定している。

ここで、式 (2.2) を変形して実験で使いやすい形にするために、陽電子ビームと電子ビームの間の相互作用をまとめたパラメータであるビームビームパラメータ ξ を導入する。陽電子ビームと電子ビームの間には電磁気的な引力がはたらいており、衝突点付近では互いのバンチが相手のバンチを引き寄せ、この力は電磁石の四極成分で近似でき、一般に四極成分があるときにはベータatron振動数がシフトするため、ビーム同士が衝突しているとき、その引力によってチューンシフトが生じる。この物理量をビームビームパラメータ ξ と呼び、具体的には式 (2.3) で表される。

$$\xi_{y\pm} = \frac{r_e}{2\pi\gamma_{\pm}} \cdot \frac{\beta_{y\pm}^* N_{\mp}}{\sigma_{y\mp}^* (\sigma_{x\pm}^* + \sigma_{y\pm}^*)} \quad (2.3)$$

ここで、 r_e は電子の古典半径、 $\gamma_{\pm}$ はローレンツ因子、 β はベータ関数を表す。2022 年 6 月の SuperKEKB におけるビームビームパラメータは $\xi_{y+} = 0.0407$ 、 $\xi_{y-} = 0.0279$ である。これは、例えば LER においてはビーム間の引力によって垂直方向のベータatron振動数が約 0.04 周期ずれることを意味している。

式 (2.2) を $\xi_{y\pm}$ が現れるように式変形すると、式 (2.4) のようになる。

$$\mathcal{L} = \frac{\gamma_{\pm}}{2er_e} (1 + a) \frac{\xi_{y\pm} I_{\pm}}{\beta_{y\pm}^*} \quad (2.4)$$

ここで、 I は総蓄積電流を表し、 a は衝突点付近でのビームサイズの水平方向と垂直方向の比である。現在の SuperKEKB 運転では 0.02 程度と 1 に対して非常に小さい。また、式 (2.4) では陽電子ビームと電子ビームの垂直と水平のビームサイズが等しいと仮定している。実際の加速器運転ではこの仮定は成り立たないことが多いが、簡単のためにこのような仮定を課している。

式 (2.4) を見ると、ルミノシティを増幅させるためには、総蓄積電流 $I_{\pm}$ 、ビームビームパラメータ $\xi_{y\pm}$ を増加させ、衝突点での垂直方向のベータ関数 $\beta_{y\pm}^*$ を減少させることが重要である。

2.3 砂時計効果とナノビーム方式

ルミノシティを向上させるためには、衝突点における垂直方向のベータ関数を絞ることが有効であることを前節で述べた。しかし、この方法には課題が存在する。実際のバンチは理想的な点状ではなく有限の長さを持つため、衝突点から離れた位置におけるベータ関数の値が式 (2.5) に従って変化するからである。ここで、 s 軸はビームの進行方向を表しており、 s は衝突点を原点とする位置を表す。

$$\beta_y(s) = \beta_y^* + \frac{s^2}{\beta_y^*} \quad (2.5)$$

例えば、衝突点でのベータ関数をビーム進行方向のバンチ長と同じ大きさに絞った場合、つまり、 $\beta_y^* = \sigma_z$ という条件では、バンチの端の $s = \sigma_z$ でベータ関数が 2 倍になってしまう。図 2.8 はビームサイズをバンチ長よりさらに絞った時の式 (2.5) を図示したものであり、この形が砂時計に見えることから砂時計効果と呼ばれている。

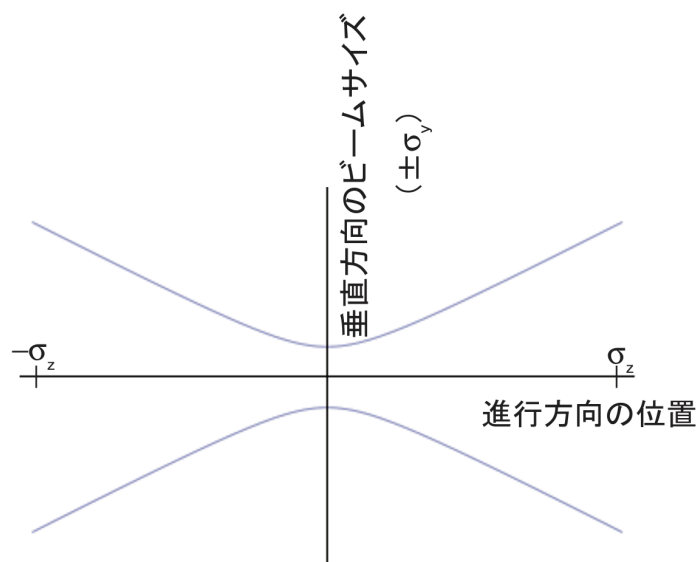


図 2.8: 砂時計効果の概略図 [14]。青線は式 (2.5) 左辺のベータ関数 $\beta_y(s)$ を表す。衝突点でのベータ関数 β_y^* をビーム進行方向のバンチ長 σ_z 以上に絞ると、 $\beta_y(s)$ は s の 2 乗に比例して膨らんでしまう。この形が砂時計に見えることから、砂時計効果と呼ばれている。

砂時計効果が発生すると、バンチの端におけるベータ関数が増大するため、衝突点でのベータ関数を減少させることによるルミノシティ向上の効果が鈍化する。SuperKEKB 加速器の前身である KEKB 加速器では、砂時計効果の影響により、バンチ長以下にベータ関数を絞ることが困難であった。一方で、SuperKEKB 加速器では砂時計効果を緩和し、ビーム同士の衝突条件を改善するため、ナノビーム方式を採用している。この方式の核心は、垂直方向のビームサイズを極めて小さくし、さらに比較的大きな交差角でビームを衝突させる点にある。図 2.9 を用いてナノビーム方式を説明する。まず、実験室系から赤矢印の方向にローレンツブースト

を行い、バンチが紙面上で水平方向に進行する系に移行する。この系において、バンチを赤線方向に射影すると、衝突条件は実効的なバンチ長 $\sigma_z^{\text{effective}} = \sigma_x / \sin \phi_c$ および実効的なバンチ幅 $\sigma_x^{\text{effective}} = \sigma_z \sin \phi_c$ を持つバンチが正面からすれ違う状況と等価になる。

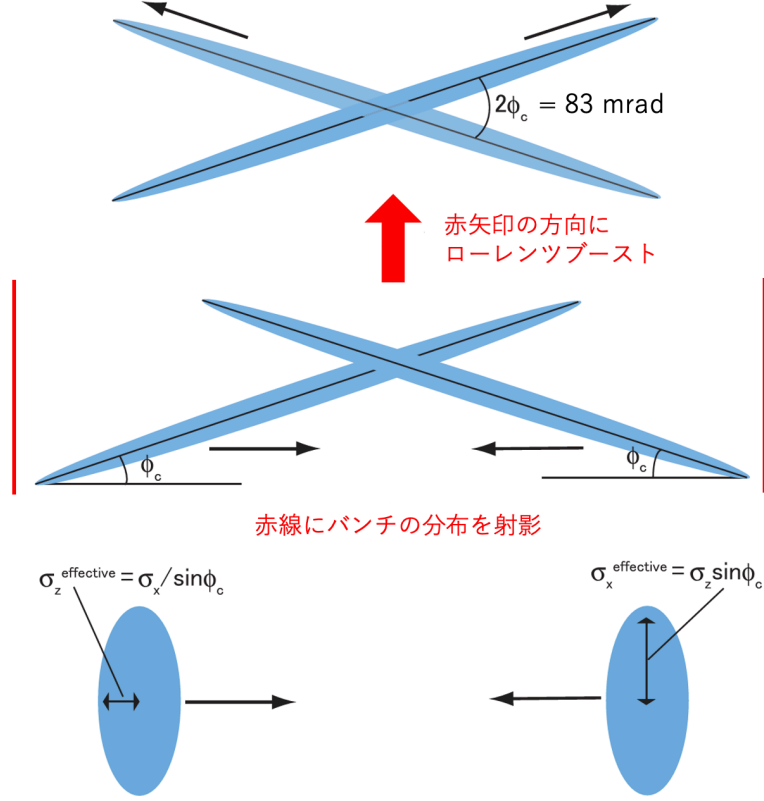


図 2.9: ナノビーム方式の説明 [15]。 $2\phi_c$ の衝突角で細長いビーム同士を衝突させることは、実効的なバンチ長 $\sigma_x / \sin \phi_c$ および実効的なバンチ幅 $\sigma_z \sin \phi_c$ を持つバンチが正面からすれ違う状況と等価である。

よってルミノシティの式 (2.2) は $\sigma_x \rightarrow \sigma_x^{\text{effective}}$ として

$$\mathcal{L}' = \frac{N_- N_+ n_b f_0}{4\pi \sigma_z \sin \phi_c \sigma_y^*} R_L \quad (2.6)$$

と表され、 $\sigma_z \sin \phi_c$ が σ_x より大きいときルミノシティを得る。このため、にナノビーム方式では角度 ϕ_c は大きい方が好ましい。ここで R_L は砂時計効果による幾何学的なロスを表す定数である。また、ビームビームパラメータについても同様に、

$$\xi'_{y\pm} = \frac{r_e}{2\pi\gamma_{\pm}} \cdot \frac{\beta_{y\pm}^* N_{\mp}}{\sigma_{y\mp}^* (\sigma_z \sin \phi_c + \sigma_{y\pm}^*)} R_{\xi\pm} \quad (2.7)$$

と表される。ここで $R_{\xi\pm}$ は砂時計効果による幾何学的な効果を表す定数である。式 (2.6) と式 (2.7) より、ルミノシティを新たなビームビームパラメータを用いて表すと、

$$\mathcal{L}' = \frac{\gamma_{\pm}}{2e r_e} (1 + a) \frac{\xi'_{y\pm} I_{\pm}}{\beta_{y\pm}^*} \frac{R_L}{R_{\xi\pm}} \quad (2.8)$$

と表される。

図 2.10 は SuperKEKB 加速器と、前身となる KEKB 加速器との衝突方法を比較している。

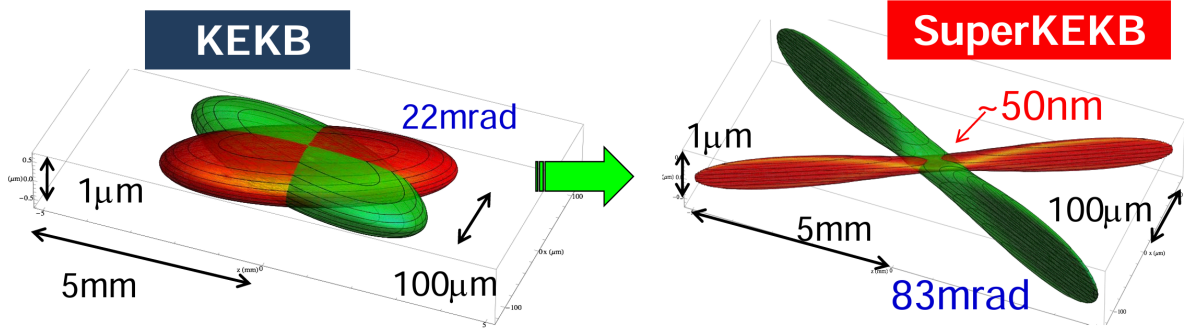


図 2.10: KEKB 加速器と SuperKEKB 加速器の衝突方式比較 [15]。SuperKEKB 加速器は前身となる KEKB 加速器と比較して、衝突角度を約 4 倍にしており、かつ衝突点でのベータ関数を 1/20 まで絞ることを目標としている。

まとめるとナノビーム方式とは、ビームに角度をつけて衝突させることにより、砂時計効果による幾何学的ロス以上に衝突点のベータ関数を絞る恩恵を得るという衝突方式である。

ここからは、ナノビーム方式の副作用について述べる。ベータ関数を絞る際には、その手前でベータ関数を広げる必要がある。図 2.11 は LER における衝突点付近のベータ関数を表しており、垂直方向では衝突点でビームを絞っているために、水平方向に比べてベータ関数が増大している部分が見て取れる。これにより前身の KEKB 加速器と比較して力学口径が小さくなってしまい、粒子が安定して周回できる領域が減少する。これによりビームの寿命が短くなり、後述するビーム入射部に対して求められる要求水準が高くなる。

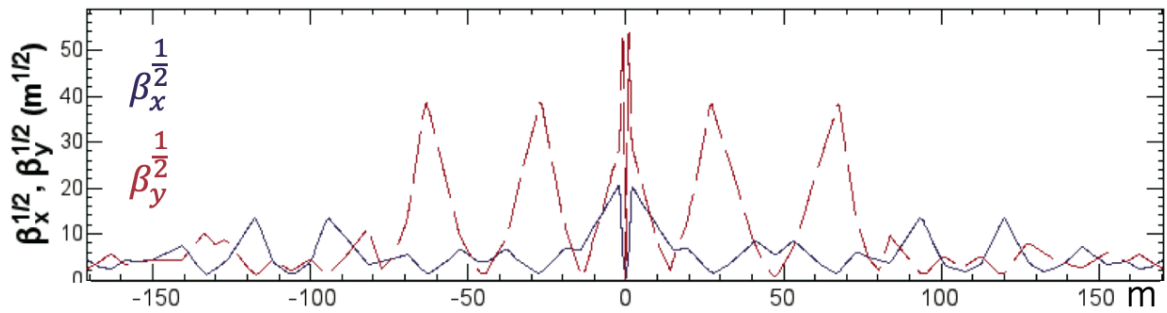


図 2.11: LER における衝突点付近でのベータ関数 [16]。衝突点でベータ関数を絞り込むためには、その手前でビームを膨らませなければならない。これによって力学口径が狭くなってしまい、結果としてビーム入射部に求められる入射要求が高くなる。

2.4 ビーム入射部

本節では、本研究で扱うビーム輸送路からメインリングへの入射調整に用いるコンポーネントについて解説する。入射調整の詳細については 4.1 節で詳しく述べる。本節では、機器の説明に重点を置くため、はじめに入射調整の概要を簡潔に述べ、その後に各コンポーネントについて解説する。

■入射調整の概要 図 2.4 において、ビーム輸送路から LER への入射点を赤点、HER への入射点を青点でそれぞれ示している。赤点で示した LER の入射点を拡大し、入射調整で使用するコンポーネントを示した詳細図が図 2.12 である。なお、HER と LER における入射調整で使用するマグネットの構成はほぼ同一であるため、図 2.12 では LER の入射調整セットアップを例として示している。ビーム輸送路から到達する入射バンチに対して、まず垂直ステアリングマグネットを用いて入射点での垂直方向の位置および角度を調整する。その後、セプタムマグネットを用いて、入射点での水平方向のビームダクト中心からの位置と角度を調整する。続いて、四極マグネットにより水平方向のビームダクト中心方向へと軌道を曲げ、最後にキッカーマグネットを用いて蓄積バンチの周回軌道へと軌道を合わせる。蓄積バンチに対しては、キッカーマグネットを用いて水平方向の入射バンチに向けて軌道を蹴り、さらに四極マグネットによって水平のビームダクト中心方向に軌道を曲げる。その後、再びキッカーマグネットを用いて蓄積リングの周回軌道に戻すことで入射が完了する。

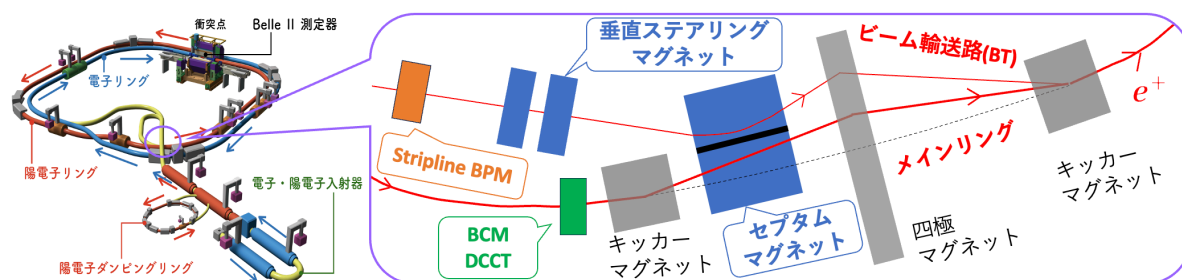


図 2.12: ビーム輸送路からメインリングへの入射調整概略図。赤線は陽電子ビームの軌道を表している。ビーム輸送路の軌道を入射バンチが通過し、メインリングの軌道を蓄積バンチが通過する。垂直ステアリングマグネット、セプタムマグネット、四極マグネット、キッカーマグネットで垂直方向を、セプタムマグネット、四極マグネット、キッカーマグネットで水平方向を調整する。Stripline BPM は入射バンチの電流を、BCM と DCCT は蓄積電流を測定する。

次に、入射効率を算出するための測定器について説明する。入射効率 E は定性的には式 (2.9) によって与えられる。

$$E = \frac{\text{単位時間の間にメインリングで増加した電流}}{\text{単位時間の間にメインリングへ入射される電流}} \quad (2.9)$$

式 (2.9) の分母にある、単位時間あたりにメインリングへ入射される電流は、ビーム輸送路に設置された Stripline BPM (Beam Position Monitor) によって測定される。一方、式 (2.9) の分子にある、単位時間あたりにメインリングで増加した電流については、BCM (Bunch Current Monitor) と DCCT (DC Current Transformer) の 2 種類の測定装置が存在するため、これに基づいて 2 つの異なる入射効率の定義が用いられ

ており、オペレータは状況に応じてこれら 2 種類の入射効率を使い分けている。BCM と DCCT はメインリング入射部の手前に設置されている。

■入射調整で用いられるマグネット 垂直ステアリングマグネット、セプタムマグネット、キッカーマグネットの外観を図 2.13 に示す。

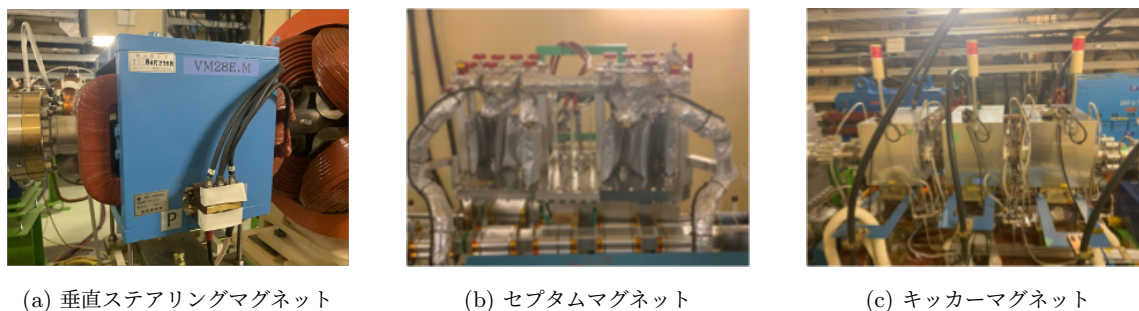
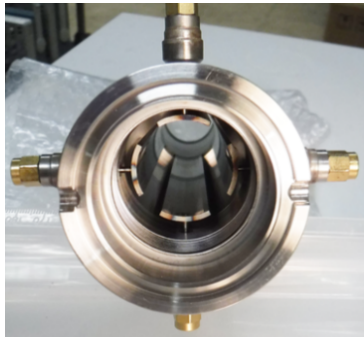


図 2.13: 入射調整で用いられるマグネットの外観

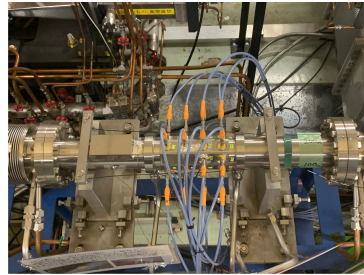
図 2.13a のステアリングマグネットは二極マグネットの一種でありビーム軌道を微調整する役割を担う。図 2.13b のセプタムマグネットは、ビーム輸送路とメインリングのビームダクト接合部に配置されたパルスマグネットであり、HER では 4 台、LER では 2 台の二極マグネットとセプタム隔壁と呼ばれる構造で構成されている。セプタム隔壁は、蓄積バンチに影響を与えることなく入射バンチを曲げる役割を果たす。図 2.13c のキッカーマグネットは、メインリング内の入射部の前後に設置された 2 台のパルスマグネットであり、蓄積ビームの軌道を大きく曲げる機能を持つ。1 台目のキッカーマグネットは、蓄積ビームをセプタム隔壁付近まで移動させる役割を担い、2 台目は、蓄積ビームと入射ビームを大きく曲げて周回軌道に戻す役割を果たす。蓄積ビームをセプタム隔壁付近まで寄せる理由は、2 台目のキッカーマグネットで蓄積ビームと入射ビームを蹴り返す際に、できる限り角度を揃えて蹴ることで、合流後の水平方向の振動を抑えるためである。この点については、第 4 章で入射調整パラメータを解説する際に詳しく説明する。セプタムマグネットおよびキッカーマグネットをパルスマグネットとして動作させる理由は、入射バンチに対応する蓄積バンチだけを選択的にセプタム隔壁付近に寄せる必要があるためである。

■入射調整で用いられる測定器 次に、入射調整で用いられる電流の測定器について説明する。図 2.14 にメインリングの電流を測定する 3 つの装置を記す。

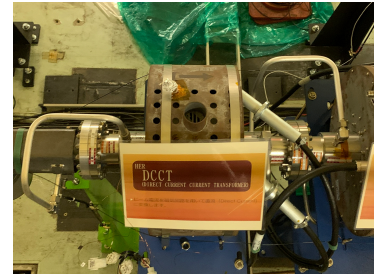
Stripline BPM は、ビームパイプ内壁に設置されたストリップ状の金属電極を利用し、入射バンチの通過によって生じる静電誘導信号を測定することで、入射バンチの位置および電荷（電流）を同定する。ボタン電極型 BPM も同様の測定原理を有するが、蓄積バンチは入射バンチに比べて静電誘導信号の強度が大きいため、ストリップ状の電極ではなく、ボタン電極を使用することが適している。ボタン電極型 BPM とその読み出し回路を組み合わせたシステムは、BCM と呼称されており、各バンチごとに蓄積バンチ電流の変化量を測定している。一方、DCCT は、ファラデーの電磁誘導の法則を利用して電流の変化を検出するが、直流成分を測定するためにフィードバック制御を行なっている。具体的には、磁心に巻かれたセンシングコイルが電流に対応する磁束を検出し、その磁束に基づいてフィードバックコイルが逆方向の磁束を生成する。これにより、磁



(a) Stripline BPM



(b) BCM 回路用ボタン電極型 BPM



(c) DCCT

図 2.14: 入射効率を測定するための蓄積電流測定装置

心内の磁束が常にゼロに保たれるように調整される。フィードバックコイルに流れる電流は元の直流電流に正確に比例しており、このフィードバック電流を測定することで、直流成分を間接的に高精度で検出することが可能となる。BCM と DCCT の使い分けについては、第 4 章で入射効率の定義で詳しく説明する。

第 3 章

研究の背景

本章では、SuperKEKB 加速器がルミノシティを向上させる際に生じている問題について議論する。3.1 節では目標ルミノシティと 2024 年 12 月時点のルミノシティを比較する。3.2 節ではルミノシティを向上させるアプローチを論じ、3.3 節で入射効率と入射調整について述べる。

3.1 目標ルミノシティと現在のルミノシティの比較

Belle II 実験における 2035 年までの積分、瞬間ルミノシティの目標プロットを図 3.1 に示す。図 3.1 における 2024 年 12 月の目標値は、積分ルミノシティが 2 ab^{-1} 、瞬間ルミノシティが $1.5 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ である。

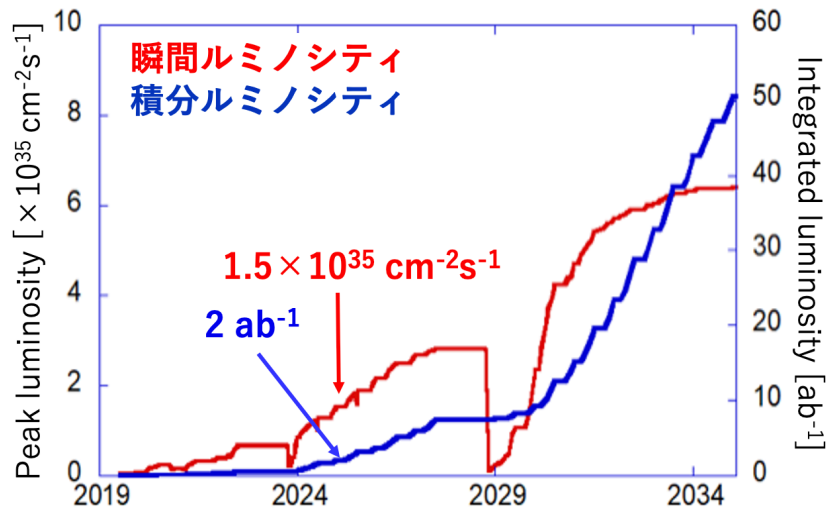


図 3.1: 2035 年までの積分、瞬間ルミノシティの目標プロット [17]。赤線が瞬間ルミノシティ、青線が積分ルミノシティであり、2024 年 12 月の目標値は瞬間ルミノシティが $1.5 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ 、積分ルミノシティが 2 ab^{-1} である。

これに対して、2024 年 12 月 31 日までのルミノシティのプロット図 3.2 を見ればわかるように、積分ルミノシティは 575 fb^{-1} であり、目標の 1/4 程度であることがわかる。また、最大の瞬間ルミノシティは

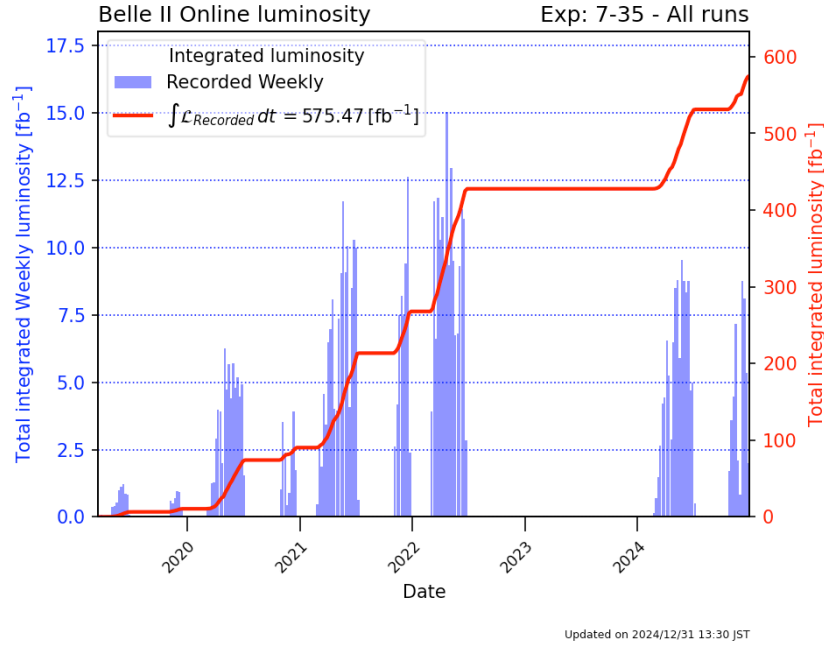


図 3.2: 2024 年 12 月 31 日時点での積分ルミノシティのプロット [18]。赤線は積分ルミノシティの総計を、青棒は週ごとの積分ルミノシティを表している。

$5.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ と、目標の $1/3$ 程度であり、積分、瞬間ルミノシティともに目標値を下回っていることがわかる。ここで、瞬間ルミノシティについてデザイン値と現在の最大値について、主要な物理量を比較したものが表 3.1 である。

表 3.1: SuperKEKB の目標値と現在の値の比較

Parameter	Unit	LER		HER	
		デザイン値	現在値	デザイン値	現在値
Energy (E)	GeV	4.000	4.000	7.007	7.007
Beam Current (I)	A	3.6	1.632	2.6	1.259
Number of Bunches		2500	2346	2500	2346
Bunch Current	mA	1.44	0.696	1.04	0.537
Beam-Beam Parameter ξ_y		0.27	0.0361	0.28	0.0267
Beta Function at IP β_x^*/β_y^*	mm	32 / 0.27	80 / 1.0	25 / 0.30	60 / 1.0

Peak Luminosity: デザイン値 : $8 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$, 現在値 : $5.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$

表 3.1 を見ると、式 (2.8) においてルミノシティを向上させる 3 つの要素である蓄積電流、衝突点でのベータ関数、ビームビームパラメータの全てに対して向上の余地がある。

3.2 ルミノシティ向上へのアプローチ

ルミノシティを向上させるためには、総蓄積電流 $I_{\pm}$ 、ビームビームパラメータ $\xi_{y\pm}$ を増加させ、衝突点での垂直方向のベータ関数 $\beta_{y\pm}^*$ を減少させることが重要であると述べた。ビームビームパラメータについては理論的にも発展の余地があり、向上させる手法が確立していないこともあるため、本節では蓄積電流と衝突点でのベータ関数について議論する。まず、ベータ関数を減少させるアプローチについて紹介する。図 3.3 は 2022 年前期の加速器運転での週ごとの最大蓄積電流を表している。5 月下旬に衝突点でのベータ関数を 1 mm から 0.8 mm に絞ったが、この時期は電流を蓄積できていないことがわかる。電流を蓄積できなかった要因としては、衝突点でのベータ関数を絞ったことによってビームの力学口径が狭くなり、入射効率が低下したことが挙げられる。

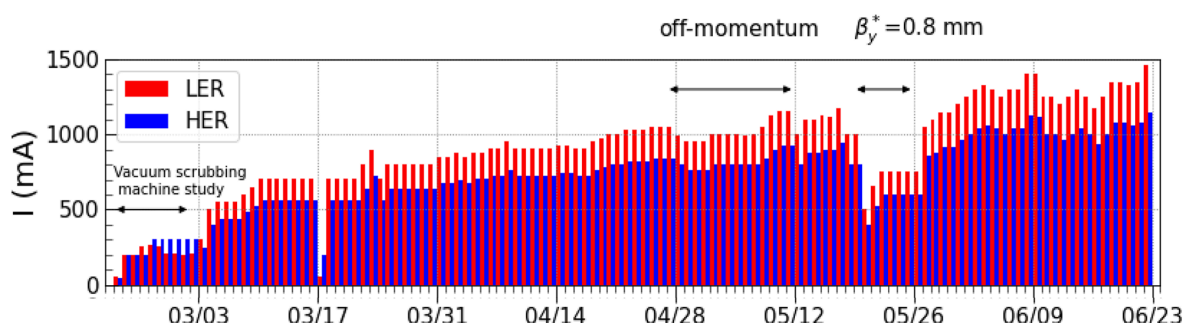


図 3.3: 2022 年前期の運転における週ごとの最大蓄積電流 [19]。2022 年の 5 月 19 日から 26 日までの 1 週間に $\beta_{y\pm}^*$ を 0.8 mm に絞ったが、その前後の週と比べて電流を蓄積できていないことがわかる。

次に、蓄積電流を上げるアプローチについて説明する。単位時間あたりに入射する電流量は、入射バンチの電流 i 、単位時間あたりの入射回数 f 、入射効率 E の積で表される。

$$\frac{dI_{\text{inj}}}{dt} = i \times f \times E \quad (3.1)$$

ここで i を大きくすると、入射バンチのエミッタンスが大きくなり、入射由来のバックグラウンドが増加する。そのため、SuperKEKB メインリングにビームを入射する手前でビームを調整することによって、エミッタンスをはじめとしたビームコンディションを向上させる必要がある。

次に、単位時間あたり入射回数 f に関して、2024 年 12 月時点での最大レートは LER で 23 Hz、HER で 21 Hz であり、これは入射制御システムの制約により増加させることが難しいため f 以外の要素で入射電流を向上させる必要がある。

入射効率に関しては次節で詳しく説明するが、高入射効率を維持できないことが現在の SuperKEKB 加速器における課題である。最適なパラメータセットが数時間のオーダーで変化していき、悪化を防ぐためにはエキスパートによる調整が 24 時間体制で必要であるため、調整を自動化するツールの開発が望まれている。

まとめると、瞬間ルミノシティを向上させるためには、ベータ関数を絞るか電流を上げるかの二つの選択肢

があり、2024 年の加速器運転では主に電流を上げる戦略をとっている。電流を上げるためには、入射バンチの電流 i を向上させるか、単位時間あたりの入射回数 f を上げるか、入射効率 E を向上させるかであるが、本研究では機械学習を用いた自動入射調整ツールを開発することにより、入射効率 E を高く維持することを目標とする。

蓄積電流の寿命に関しては付録 A.3 に記載した。蓄積電流の寿命は蓄積電流の 2 乗に比例するため、蓄積電流が増加するほど、その維持にはより高い入射性能が要求されることに留意する必要がある。

3.3 入射調整に対する困難

ここで、手動入射調整に対する困難を紹介する。SuperKEKB 加速器の運転は加速器制御棟にて行われており、KEK の加速器施設所属の職員と、業務委託先から派遣されたオペレーターにより、三交代のシフト体制で行われている。加速器の調整パラメータには自動化や半自動化されているもの^{*1}が多いが、入射調整は 2024 年時点で全く自動化されていない分野であり、職員やオペレーターが逐次入射効率の監視をしながら 24 時間体制で調整をしている。このような調整は、調整を行う担当者の技術に依存するために巧拙の差が出てしまう。そこでオペレーターの技量によらず、均一な最適化を行える自動調整ツールの実現が必要とされている。図 3.4 は、手動による入射調整の一例を示している。この事例では、オペレーターが 8 時 30 分から 10 時 20 分までの 2 時間弱にわたり、入射調整パラメータである y -position、 y -angle、Septum Angle を手動で調整したものの、入射効率は 10% 付近で推移し、大きな改善が見られなかったことを示している。

入射調整においてこのような事例が生じる理由は、最適な入射調整パラメータが時間により変動しやすいからである。図 3.5 は 2024 年 11 月 16 日から 12 月 16 日までの 1 ヶ月間の LER での y -position, y -angle, Septum angle を 3 次元プロットにした図である。図 3.5a は時系列を、図 3.5b は入射効率をカラープロットに反映している。図 3.5a をみると、パラメータセットが 1 ヶ月を通して刻一刻と変わっていることがわかる。また、図 3.5b で入射効率を見ると、高い部分と低い部分がまだらになっている。これは最適なパラメータセットが数日のうちに変化して低くなることを示している。

最適なパラメータセットが変動しやすい理由として、メインリングが KEK Linac およびビーム輸送路の下流に位置している点が挙げられる。この構造上、メインリングだけでなく上流のどこか一箇所でも変調が生じた場合、その影響を受けてしまうために最適なパラメータセットが変化しやすいと考えられる。

また、入射部について現実を再現するようなシミュレーションは 2024 年時点で存在しないため、入射調整部分は実質的にブラックボックス関数になっており、実際にデータを測定してみるまでどのパラメータセットが最適な入射効率を実現できるかがわからない。また、最適なパラメータセットが数時間単位で予測不能な時間変化をするために、過去に高い入射効率であったパラメータセットを再び設定しても改善しない場合も多い。さらに、入射効率の測定誤差が 3% 程度と大きく、機械学習を用いない古典的なブラックボックス最適化アルゴリズムである Downhill Simplex 法 [20] が機能しない [21]。その上、パラメータを大きく変化させると、ビームの安定周回条件を逸脱し、ビームアボートが促されてしまう。このような手動入射調整における課題を解決するため、自動入射調整ツールを開発することを本研究の目的とする。衝突型加速器の入射調整に機械学習を導入する事例は、これまでに例がなく、本研究はその初めての挑戦的な取り組みである。本研究で開

^{*1} tune 補正や光学 knob 調整など。

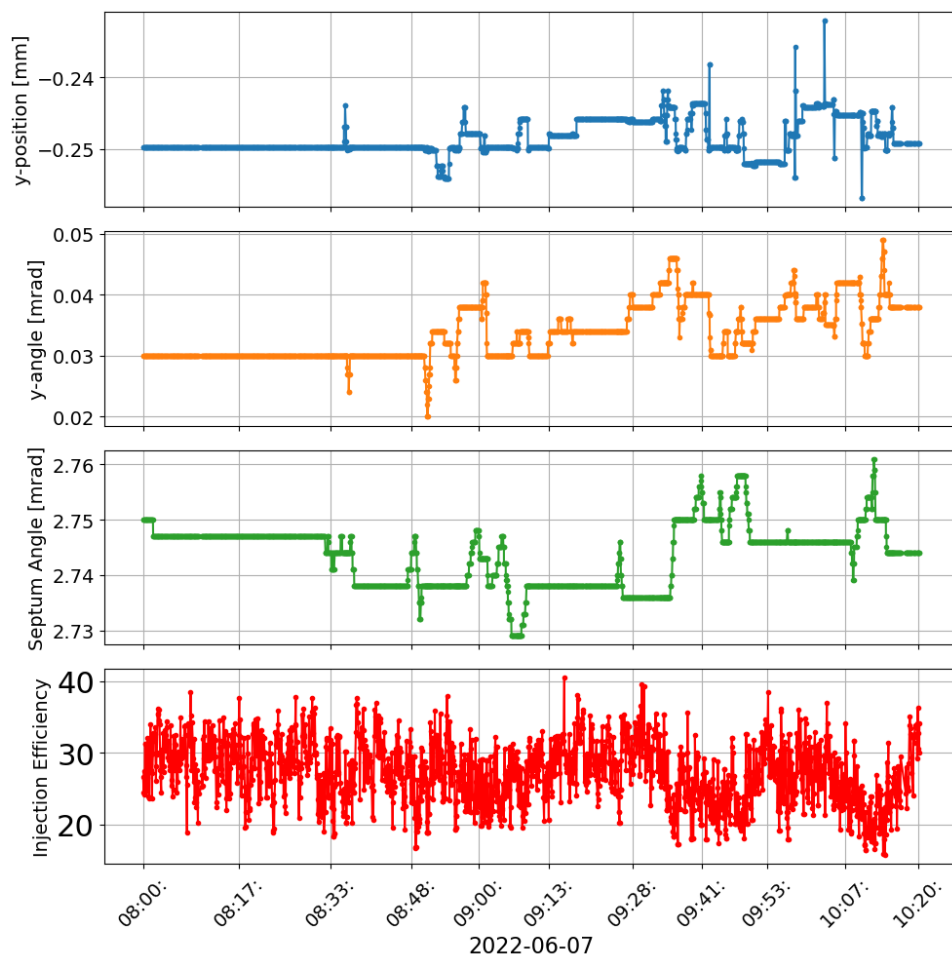
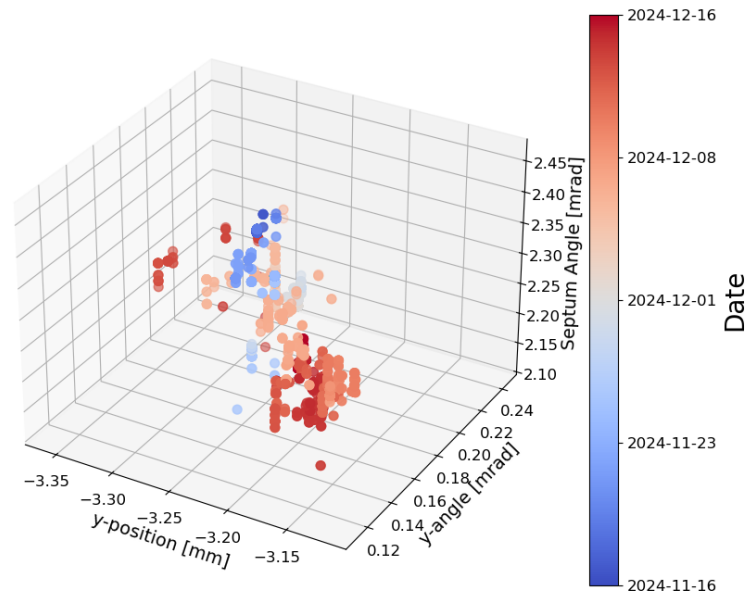


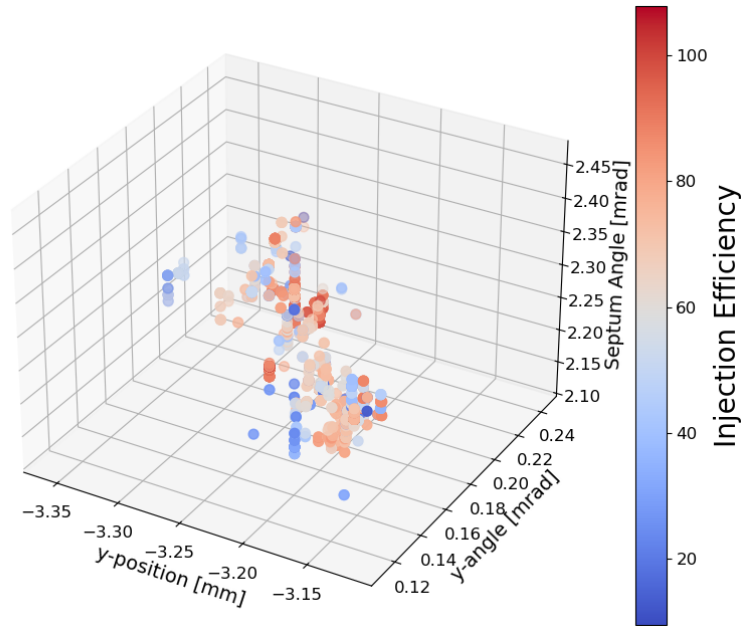
図 3.4: 手動による入射調整事例。この事例ではオペレータが2時間ほど3つの入射調整パラメータを操作したが、入射効率は約10%付近で推移し改善は見られなかった。

発する入射調整ツールには、以下の3つの性能を要求する。

- **入射効率の向上:** 過去のデータを用いずに入射効率を向上させる能力を有すること。また、最適化後に入射効率の向上しない場合であっても、悪化させないことを保証すること。
- **安全性の確保:** ビームアポートを引き起こさない設計であり、高い安全性が担保されていること。
- **利便性:** オペレータが直感的かつ簡便に操作できるユーザーインターフェースを備えていること。



(a) 日時



(b) 入射効率

図 3.5: 1 ヶ月間の LER における、3 つの入射調整パラメータの 3 次元プロット。2024 年 11 月 16 日から 12 月 16 日までの入射調整パラメータ (y-position, y-angle, Septum Angle) を 3 次元のパラメータ空間にプロットした。(a) は時系列をカラーマップにしており、パラメータセットの変遷を追うことができる。(b) は入射効率をカラーマップにしており、高い部分と低い部分がまだらになっている。これは入射効率が高い点がパラメータ空間上で定まっていないことを表している。なお、入射調整パラメータの定義については 4 章で解説している。

第 4 章

入射調整における機械学習アルゴリズムの導入

本章では、ビーム輸送路下流から SuperKEKB メインリングへの入射調整について扱う。4.1 節では入射調整のパラメータについて説明し、4.2 節で入射効率の定義について述べる。4.3 節では本研究で用いる機械学習手法であるベイズ最適化の概要を紹介し、4.4 節で入射調整に対して機械学習を導入する手法を述べる。

4.1 入射調整パラメータ

入射調整で用いられるパラメータについて説明する。図 4.1 は HER の SuperKEKB 手動入射調整パネルであり、現在は調整オペレーターが 7 個のパラメータを操作している。左側の図は図 2.12 と同様の描像を表しており、右側のパラメータの値を元に描画される。

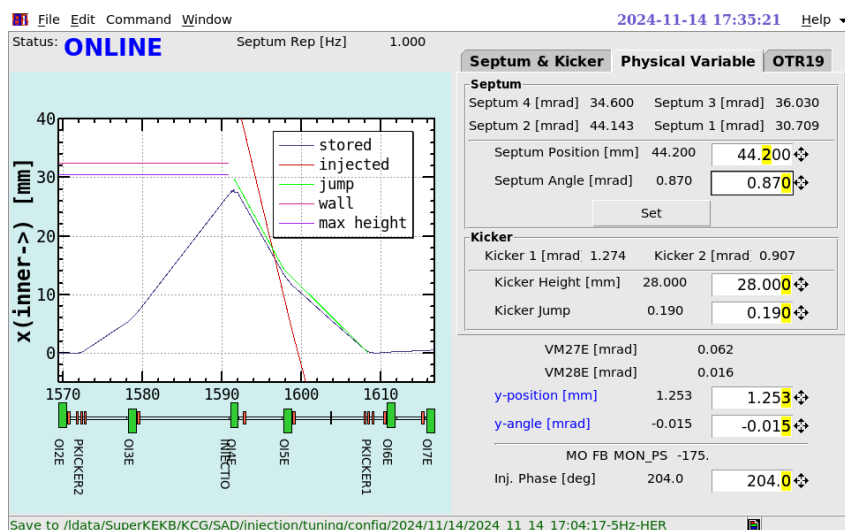


図 4.1: HER での手動入射調整パネル。パネル右側では 7 つの入射調整パラメータを調整可能であり、パネル左側には設定したパラメータを基に入射の描画が表示されている。横軸は衝突点を原点としてビーム進行方向を正としたビーム進行方向の座標を表しており、単位は m である。縦軸はビームダクトの中心からの水平距離を表しており、単位は mm である。stored は蓄積バンチの軌道、injected は入射バンチの軌道を描写している。jump は入射バンチの縦軸の縮尺を蓄積バンチと合わせて描写し、二つのバンチの角度の差をわかりやすくしている。wall はセプタム隔壁の位置を、max height はビームダクトの位置を描写している。

表 4.1 に、図 4.1 で使用される 7 つの入射調整パラメータの概要を示す。その後、個々のパラメータについて詳細に解説する。

表 4.1: 入射調整パラメータの説明

名称	パラメータの詳細
y-position [mm]	セプタムマグネット終端を通過する入射ビームのビームダクト中心からの垂直距離
y-angle [mrad]	セプタムマグネット終端を通過する入射ビームのビームダクト中心からの垂直角度
Septum Position [mm]	セプタムマグネット終端を通過する入射ビームのビームダクト中心からの水平距離
Septum Angle [mrad]	セプタムマグネット終端を通過する入射ビームのビーム進行方向からの水平角度
Kicker Height [mm]	キッカーマグネットによって蹴られた蓄積ビームが四極マグネットを通過する際のビームダクト中心からの距離
Kicker Jump	蓄積ビームと入射ビームの設計軌道からの振動の比率 (0~1 の連続値で、0 ほど入射ビームが振動する)
inj. Phase [deg]	メインリングのバケット位相 (付録 A.2) に対する入射器全体の RF 位相

上部 6 つのマグネットに関するパラメータ (Septum Position, Septum Angle, Kicker Height, Kicker Jump, y-position, y-angle) は、人間が調整する際に直感的に理解しやすい距離や角度といった物理量で定義されており、これらのパラメータを実現するために、複数のマグネットに印加すべき電流を算出して適用している。

■y-position, y-angle 垂直方向を操作するパラメータの概略図を図 4.2 に記した。ビーム垂直方向は図 2.12 や図 4.1 左図では紙面に対して垂直な方向を表している。ビーム輸送路下流には 2 つの垂直ステアリングマグネット対があり、このマグネットの印加電流を変化させることにより入射ビームの位置と角度を調整できる。垂直ステアリングマグネットの 7 m 後部にある入射地点に到達した時のビームダクト中心からの垂直距離を y-position、垂直角度を y-angle と定義する。また入射地点は後述する水平方向から見た図 4.3 の Injection point と同じ地点である。

■Septum Position, Septum Angle, Kicker height 水平方向を操作するパラメータの概略図を図 4.3 に記した。図 4.3(a) では HER の、(b) では LER の水平軌道が描かれている。入射バンチの軌道を操作するパラメータは Septum Position と Septum Angle の 2 種類であり、それぞれセプタムマグネット端の入射地点におけるビームダクト中心からの水平距離と水平角度に対応する。蓄積バンチを操作するパラメータは Kicker Height であり、キッカーマグネットによって蹴られた蓄積バンチが四極マグネットによって曲げ返される地

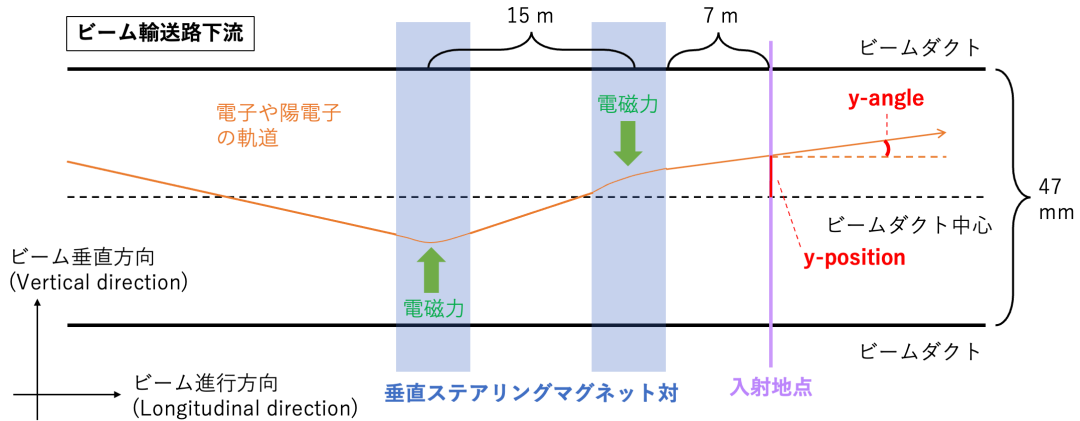


図 4.2: ビーム垂直方向の入射説明図。2つのステアリングマグネットを用いて垂直方向のビーム軌道を操作する。ステアリングマグネット後部の入射地点での垂直距離が y-position、垂直角度が y-angle である。

点を表す。通常、四極マグネットはその中心にビーム軌道を通してビームサイズを絞るために使用されるが、ここでは意図的にビーム軌道を四極マグネットの中心から外して使用している。この配置の目的は2点ある。1点目は、入射に関与しない蓄積バンチを四極マグネットの中心を通過させることで、電磁力を与えない点である。2点目は、四極マグネットの中心から離れるほど強く発生する二極成分を利用して、入射ビームを蓄積バンチよりも強くビームダクトの中心方向に曲げ、後方のキッカーマグネットで蹴り返す際に、両バンチが同じ位置で重なるようにする点である。1点目について詳細に解説する。蓄積ビームが2000個のバンチから構成されているとすると、1バンチは $10\text{ }\mu\text{s}$ でリングを一周するため、1秒間に蓄積バンチが入射部を通る回数は 2×10^8 回である。一方で、入射バンチが到達する頻度は最大で23 Hzであるため、キッカーマグネットによって蹴られ、入射バンチと合流する蓄積バンチの数は、全蓄積バンチの中ではごくわずかである。このため、四極マグネットはビーム軌道中心を通過する蓄積バンチには影響を与えず、キッカーマグネットで蹴られた蓄積バンチのみを曲げる役割を果たしている。

水平部のセプタムマグネットと、メインリングのビームダクトについての関係を記したものを図 4.4 に記す。Kicker Height と Septum Position の差を ΔX と定義する。入射した直後では、蓄積バンチと入射バンチはこの ΔX を振幅として振動する。その後、シンクロトロン放射によって徐々に振動は減衰していき、最終的に入射バンチは蓄積バンチと一体化する。 ΔX が大きいとバンチ間の振動が大きく、バンチが一体化する前に入射バンチが安定軌道から外れてしまったり、コリメータによって削られてしまうことが多くなる。そのため、Septum Angle と Kicker Height に関しては、できる限りセプタム隔壁に近づけることが好ましい。

■Kicker Jump Kicker Jump はバンチ間の振動において、どちらのバンチを基準として振動させるかを制御するパラメータであり、0 から 1 の値を取る。0 であれば蓄積バンチを中心に入射バンチが振動し、1 であれば入射バンチを中心に蓄積バンチが振動する。2024 年時点では 0-0.2 程度の値を採用しており、このパラメータは熟練のオペレータ以外が操作することは極めて少なく、エキスパートは 1-3 週間に 1 回の頻度で微調整を行っている。

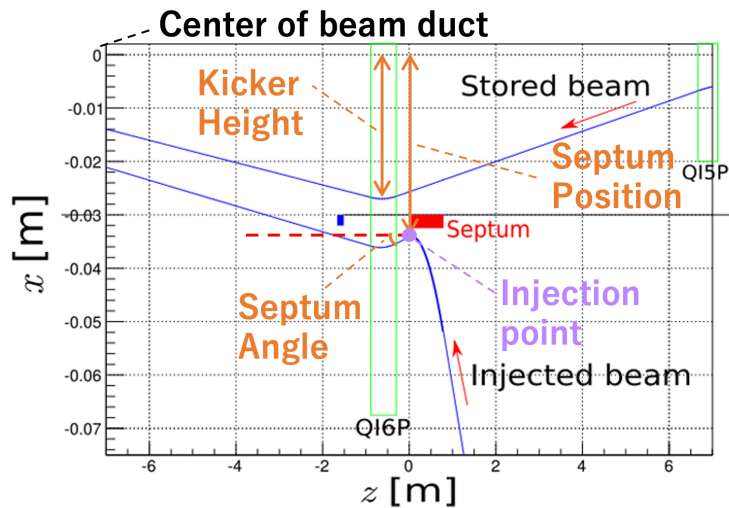
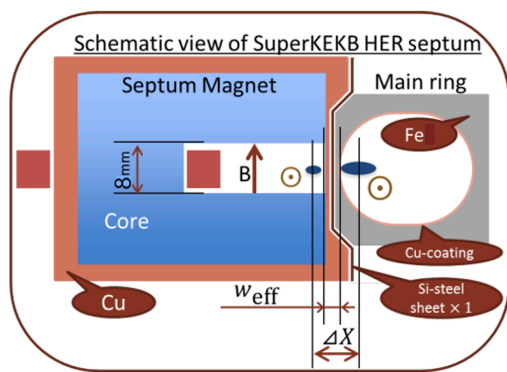
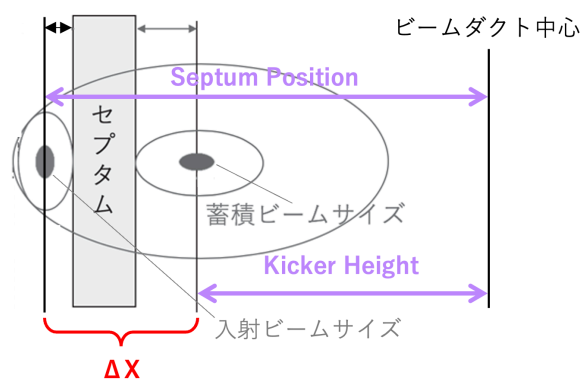


図 4.3: ビーム水平方向の入射説明図 [9]。LER と HER の入射方法は共通しているため、この図では例として LER の入射部を示している。図の中心には赤色でセプタム隔壁が描写されており、隔壁の端を Injection point と定義している。これは図 4.2 の入射地点と同じ地点である。横軸は Injection point を原点とするビーム進行方向に沿った座標を表しており、縦軸はビームダクト中心を原点とした水平方向の座標を表している。青線は入射バンチと蓄積バンチの軌道を表しており、入射バンチの軌道はセプタムマグネットによってビームダクト中心と逆方向に曲げられる。これは入射バンチの水平角度を蓄積バンチの水平角度に近づけるためである。緑線で描写されているのは四極マグネットであり、両バンチの軌道はビームダクト中心方向に曲げられる。橙線で描写されている距離と角度は水平方向の入射調整パラメータである。Injection point での、ビームダクト中心からの入射バンチまでの水平距離が Septum Position、ビーム進行方向に対する水平角度が Septum Angle であり、四極マグネット中心でのビームダクト中心からの蓄積バンチまでの水平距離が Kicker Height である。



(a) 機器に着目した図 [9]



(b) バンチ構造に着目した図 [10]

図 4.4: セプタムマグネットとビームダクトの関係図。入射バンチと蓄積バンチの水平距離が離れていると、バンチが合流して周回する際に水平振動が大きくなってしまうため、両バンチの距離はできる限りセプタム隔壁に近づけることが好ましい。

4.2 入射効率

次に、入射効率について説明する。現在 SuperKEKB 加速器には二つの定義の入射効率が存在し、オペレータは2種類の入射効率を状況に応じて使い分けしている。二つの入射効率についての測定対象とメリットをまとめたものが表 4.2 である。

表 4.2: BCM と DCCT の比較

種類	測定する電流	メリット
BCM	1 バンチあたりの電流	ビーム寿命の影響を受けにくい。 1 バンチ目と 2 バンチ目の入射効率を独立して測定できる。
DCCT	蓄積リング全体の電流	ビームの立ち上げ時でも値が測定できる。 100 ターン目以降のビームの振る舞いを測定できる。

BCM のメリットはビーム寿命の影響を受けにくいことである。BCM はバンチ単体の電流を測定するのに対して、DCCT は蓄積ビーム全体の電流を測定する。バンチ内の電子や陽電子はタウシェック散乱やビームガス散乱 (付録 A.3) によって減衰していく。2024 年時点の SuperKEKB 運転では蓄積ビームは約 2300 個のバンチから構成されているため、DCCT で測定する蓄積電流では、減衰の効果もバンチ単体を測定する BCM と比較して約 2300 倍となってしまうビーム寿命の影響を強く受けてしまう。

BCM のもう一つのメリットとしては2バンチ入射での入射効率を独立して測定できることである。2バンチ入射とは、1 回の入射においてバンチを2つ入れる入射方法のことである。この2つの入射バンチの入射効率を独立に測定することによって、1 バンチ目の入射効率はよいが2 バンチ目の入射効率が悪い場合などに2 バンチ目だけに注目して入射効率を上げるような最適化も可能になる。

DCCT のメリットとしてはビームの立ち上げ時でも電流の測定ができることである。BCM ではオフセットノイズをキャンセルするために、前回入射した際のバンチ電流と新しく入射した時のバンチ電流の差分を測定している。そのため、電流の立ち上げ時のようにバケットにバンチが入っていない場合には電流値を算出できない。BCM と DCCT での入射効率の比較を図 4.5 に示している。この図では 0 mA から 600 mA 程度まで同じ量のビームを入射し続けている。まずわかることは、DCCT が入射後すぐに入射効率を測定できるのに対して、BCM は測定を開始するまで5分程度かかっているということである。また、BCM で見た入射率は一定であり、実際の入射状況と一致するが、DCCT で見た入射率は蓄積電流に依存するビーム寿命の影響で低く見積もられていることも見て取れる。

さらに、DCCT の利点として、BCM では捉えられないような長期的な電流減衰現象を観測できる点が挙げられる。BCM では、入射バンチと蓄積バンチが合流した後、100 ターン後に前回入射時との電流の増加量を測定するため、数千ターンを要して徐々に進行するような現象を測定することはできない。たとえば、入射バンチの水平振動が大きい場合に、数千ターンをかけて蓄積バンチから粒子が徐々に溢れていくような現象は、BCM の測定では捉えられない。^{*1}

^{*1} 厳密には、トリガーされていないためデータとして記録されない。

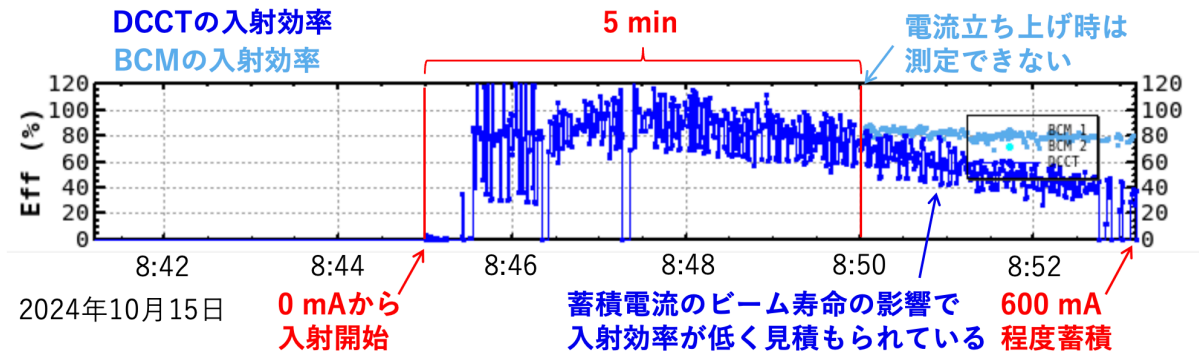


図 4.5: BCM と DCCT で測定した入射効率の比較。0 mA から入射を開始し、600 mA 程度まで蓄積させた時の BCM と DCCT での入射効率を表している。BCM では入射が始まってから 5 分後まで入射効率が表示されないことと、DCCT では蓄積電流がビーム寿命の影響を受けて徐々に低くなっていることがわかる。

4.3 ベイズ最適化

この節では本研究で用いる機械学習手法であるベイズ最適化の概要を紹介する。ベイズ最適化 [22] [23] はブラックボックス最適化の一種である。ブラックボックス最適化とは、数理的構造が明らかになっていない関数の最適化に用いられる手法であり、関数全体の具体形が一切わからないという前提のもとで、観測を通して効率よく最適化を行う手法のことである。ブラックボックスとは、調整可能なパラメータを格納したベクトルを $\mathbf{x}$ として、目的関数 $f(\mathbf{x})$ の具体的な形はわからないが、あるパラメータ設定 $\mathbf{x} = \mathbf{a}$ を測定したときに測定誤差を ϵ として $f(\mathbf{a}) + \epsilon$ を知ることができるような問題設定である。この概念を図示したものを図 4.6 に記す。

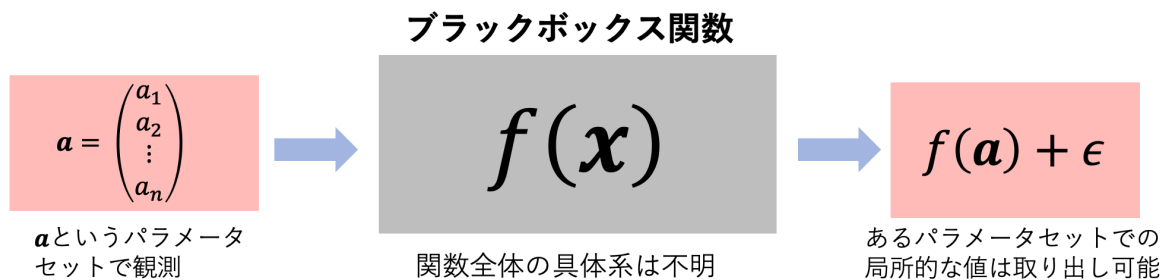


図 4.6: ブラックボックス関数の概念図。関数全体の具体系は不明だが、測定を通して局所的な値は取り出すことができる。

関数の測定が容易であれば、各パラメータの定義域を端から端まで測定し、領域全体に渡ってグリッドサーチをすることで関数の全体像を確実に把握することができる。しかし、関数の測定に時間やコストがかかる場合については、最適化に重要ではない領域までも測定してしまうため。効率が悪い。また、グリッドサーチは

パラメータの次元が増えるごとに測定点が指数関数的に増えてしまう欠点がある。例えば3パラメータを調整する際に各パラメータを10点測定するだけで、合計では1000点の測定が必要となり、1点の測定に1分を要する場合には全ての点を知るのに約17時間もかかってしまう。これは実際の加速器調整において現実的ではないため、少ない観測回数で目的関数の特徴を掴み、最適なパラメータを効率よく見つけるアルゴリズムが求められる。

ベイズ最適化はガウス過程回帰法 [24] という確率分布を作成する手法と、獲得関数という確率分布の中でどの点が最も最大値らしいかを表す関数を通じて、少数統計で関数の最大値の確率的予測を効率よく行うことができる最適化手法である。この手法は、リアルタイムで取得された測定データに基づいて関数の予測を行うため、事前に教師データを必要としない点で有用である。入射調整においては、第3章で説明したように、KEK Linac から供給されるビーム状態や、メインリングにおける各種パラメータ設定の変動により入射効率に再現性がない。このため、教師データを必要としない機械学習手法であるベイズ最適化を採用した。

ベイズ最適化を選定した理由について述べる。機械学習を用いない古典的なブラックボックス最適化アルゴリズムである Downhill Simplex 法は、測定誤差が大きい場合に局所的な最大値に捕捉され、全域的な最大値を探索できないことが多いことが知られている。さらに、KEK Linac の軌道を調整して電子ビームの収量を最大化する問題において、機械学習を用いた3つのブラックボックス最適化アルゴリズム（ベイズ最適化、TPE、CMA-ES）の結果を比較した先行研究 [25] では、ベイズ最適化が最も高い性能を示した。以上の理由から、本研究では入射調整に対してベイズ最適化を採用した。以下では、ガウス過程回帰法、獲得関数、ベイズ最適化について説明する。詳細な理論と実装方法に関しては付録 A.4 に記した。

■ガウス過程回帰法 ガウス過程とは、与えられた入力次元に対して、それと同じ出力次元の多変量ガウス分布（付録 A.4.1）を吐き出す”箱”のようなものである。ガウス過程回帰法とは、観測点によって作られた”箱”に、未知の点を入力し、その点における確率分布を出力させる回帰方法である。入力パラメータ $\mathbf{x}$ と、ブラックボックス関数の測定値 $y = f(\mathbf{x})$ があり、 N 回の測定によって $\mathbf{x}_n, y_n$ ($n = 1, 2, \dots, N$) が既知であるとする。ここで、測定点の情報をもとに、空間上の任意の点 $\mathbf{x}^*$ での y^* の確率分布は式 (4.1) のように表される（付録 A.4.4）。

$$p(y^*|\mathbf{y}, \mathbf{K}, \mathbf{x}^*) = \mathcal{N}(\mathbf{k}_*^T(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\mathbf{y}, k_{**} - \mathbf{k}_*^T(\mathbf{K} + \sigma^2\mathbf{I})^{-1}\mathbf{k}_*) \quad (4.1)$$

ここで $\mathcal{N}(a, b^2)$ は平均 a 、分散 b^2 の正規分布を表す。 $\mathbf{K}_{nn'} = k(\mathbf{x}_n, \mathbf{x}_{n'})$ ($n, n' = 1, 2, \dots, N$) を要素に持つ行列を $\mathbf{K}$ 、既知の点と未知の点からなる共分散ベクトルを $\mathbf{k}_* = (k(\mathbf{x}^*, \mathbf{x}_1), k(\mathbf{x}^*, \mathbf{x}_2), \dots, k(\mathbf{x}^*, \mathbf{x}_N))^T$ 、未知の点同士の分散を $k_{**} = k(\mathbf{x}^*, \mathbf{x}^*)$ 、ブラックボックス関数の測定値を $\mathbf{y} = (y_1, y_2, \dots, y_N)$ 、測定の分散を σ^2 で表す。 $k(\mathbf{x}, \mathbf{x}')$ の具体系については、どのような予測分布を回帰したいかによって異なる（付録 A.4.5）。本研究では、空間上の各点が無限回微分可能で連続であることを仮定した式 (4.2) を用いた。

$$k(\mathbf{x}, \mathbf{x}') = \theta_1 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{\theta_2}\right) \quad (4.2)$$

ここで、 θ_1, θ_2 はハイパーパラメータであり、本研究においては後述するベイズ最適化の際に、確率的勾配降下法を応用した手法である Adam [26] を用いて、式 (4.1) を θ の関数とみなした尤度 $p(\mathbf{x}, y|\theta_1, \theta_2)$ が最大となるような θ に逐次的に最適化されている。

式 (4.1) が意味することは、測定点 $\mathbf{x}_n, y_n$ ($n = 1, 2, \dots, N$) の情報が既知であり、式 (4.2) のハイパーパラメータが定められて $\mathbf{K}, \mathbf{y}$ も既知であるとき、空間上の任意の点 $\mathbf{x}^*$ でのブラックボックス関数の予想値 y^* は平均 $\mathbf{k}_*^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{y}$ 、分散 $k_{**} - \mathbf{k}_*^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{k}_*$ の正規分布に従うということである。図 4.7 は 1 次元でのガウス過程回帰のデモンストレーションである。この図では、7 個の既知の赤点の情報をもとに、測定していない点での平均と分散を回帰し、平均を青線で、標準偏差を水色の領域で表している。

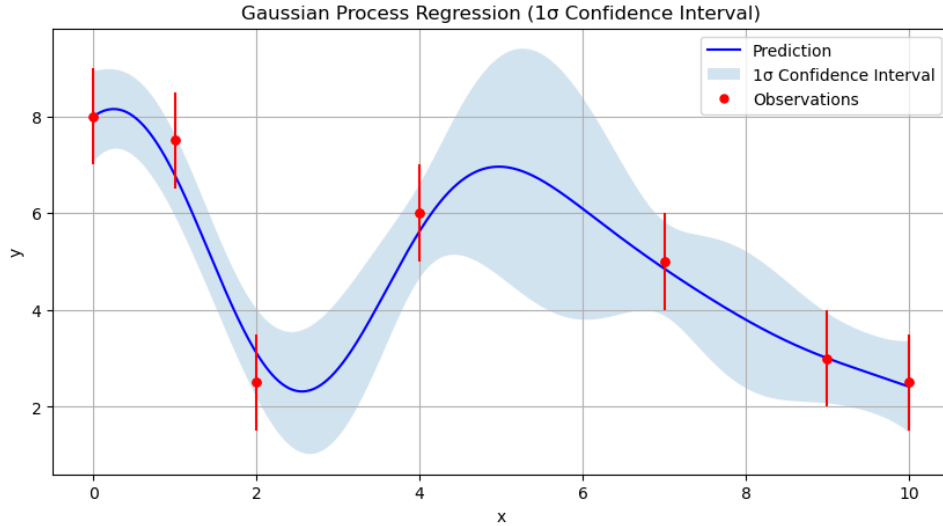


図 4.7: 1 次元のガウス過程回帰の図。赤点は観測データを表しており、これらのデータに基づいて、定義域内の任意の点 ($0 \leq x^* \leq 10$) に対して正規分布の平均（青線）および標準偏差（水色の領域）を生成する。この x^* を定義域全体にわたって連続的に生成することで予測分布を構築している。ここで測定点の分散は $\sigma^2 = 1$ としており、ハイパーパラメータは Adam によって $\theta_1 = 23.3, \theta_2 = 1.52$ と最適化されている。

■獲得関数 前節では、既知の測定点から未知の点の平均と分散を回帰する方法を説明した。次に、予測分布の中でどこが最も最大値らしい点であるかを議論する。例として、図 4.7 において最大値の候補を考えると、平均が一番大きいのは $x = 0.2$ 付近である。一方で $x = 5$ 付近では、周囲に測定点の情報が多く分散が大きいため、こちらに最大値が存在しているかもしれない。前者のような ”平均” を重視したアプローチを活用的な考え方といい、後者のような ”不確かさ (分散)” に重きを置くアプローチを探索的な考え方という。活用と探索はトレードオフな関係にあり、このバランスを取りつつ最も最大値らしい点を決める関数が獲得関数である。獲得関数の値の大きさは、予測分布の最大値らしさを表す。本研究では 2 つの獲得関数について述べる。1 つ目が UCB (Upper Confidence Bound) であり、2 つ目が EI (Expected Improvement) である。

(1) UCB (Upper Confidence Bound)

UCB は最も基本的な獲得関数の 1 つであり、定義は以下の通りである。ここで、 μ と σ は予測分布の平均値と標準偏差である。

$$\text{UCB}(\mathbf{x}) = \mu(\mathbf{x}) + \beta \sigma(\mathbf{x}) \quad (4.3)$$

UCB の特徴は獲得関数内にハイパーパラメータである β が存在していることであり、この値を調節することで活用と探索のバランスを調整することができる。反面、適切な β を設定する必要があるため、実装する前には予備試験によって β に対する応答が適切かを確かめる必要がある。

(2) EI (Expected Improvement)

EI も UCB と同様に獲得関数の中で最も基本的な獲得関数の内の 1 つであり、式 (4.4) で表される。EI の長所として、誤差が 0 である時には無限回の試行を繰り返せば必ず全域的最大値に到達することが保証されるという理論的な背景が挙げられる。式の導出に関しては付録 A.4.6 で説明している。

$$EI(\mathbf{x}) = (y_{\max} - \mu(\mathbf{x}))\Phi\left(\frac{y_{\max} - \mu(\mathbf{x})}{\sigma(\mathbf{x})}\right) + \sigma(\mathbf{x})\phi\left(\frac{y_{\max} - \mu(\mathbf{x})}{\sigma(\mathbf{x})}\right) \quad (4.4)$$

ここで、 ϕ と Φ はそれぞれ正規分布の確率密度分布と累積密度分布を表しており、 $y_{\max}$ はブラックボックス関数の測定値 y における最大値である。

1 次元でのガウス過程回帰を表した図 4.7 において、UCB ($\beta = 0.5$) と EI を用いて最大値をそれぞれ予測した結果が図 4.8 である。

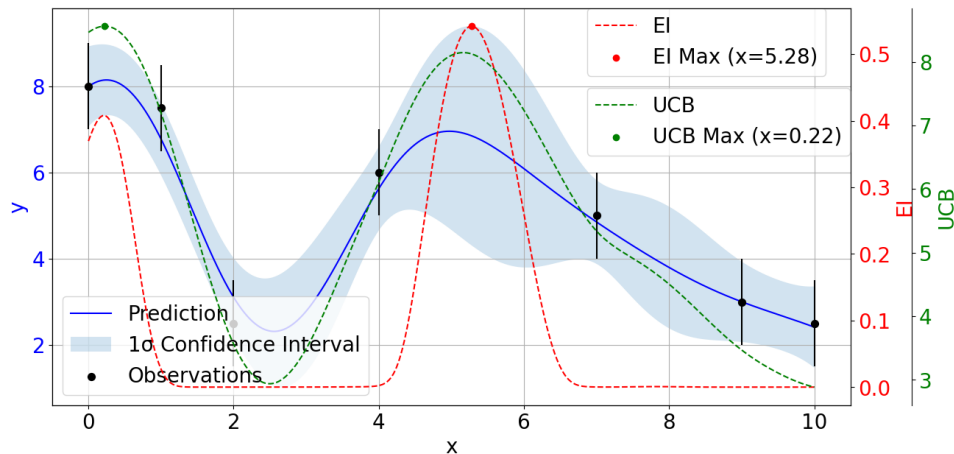
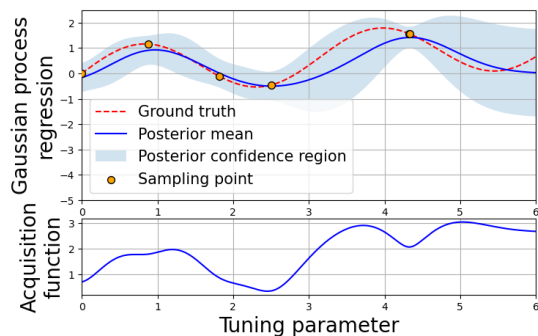
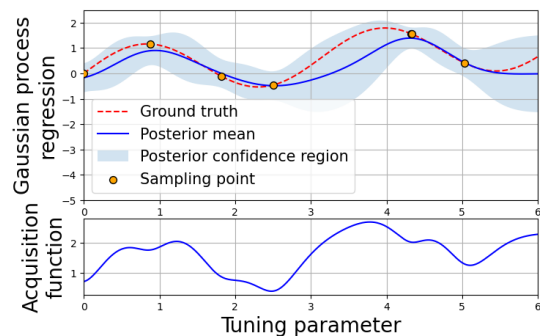


図 4.8: 1 次元のガウス過程回帰において、獲得関数での最大値予測を描写したプロット。UCB($\beta = 0.5$) では $x = 0.22$ が、EI では $x = 5.28$ が最も最大値らしいと予測しており、同じ予測分布でも獲得関数によって異なる最大値を取っている。

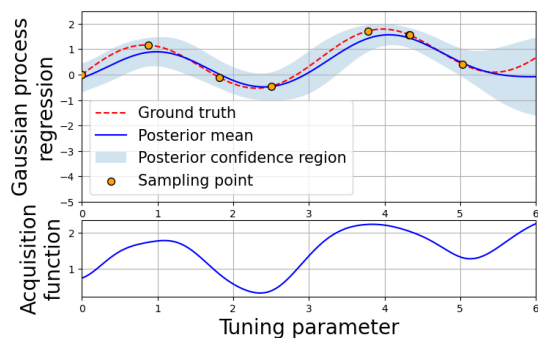
■ベイズ最適化 前節では、獲得関数によって確率分布における最大値の予測を行った。獲得関数の最大値を実際に測定することによって既知の点の情報が増え、ガウス過程回帰の予測分布が更新されることにより、獲得関数の最大値が更新される過程をトライアルと呼ぶ。ベイズ最適化とはトライアルを繰り返すことによって最大値らしい部分を逐次測定し、情報を更新することによって精度を向上させていく最適化手法のことである。図 4.9 ではテスト関数を用いて一次元のベイズ最適化のデモンストレーションをしている。実際の測定では未知の関数であるブラックボックス関数をテスト関数 $y = \sin(2x) + 0.2x$ と仮定し、赤の点線で表示した。定義域は $(0 \leq x \leq 6)$ とし、この中での最大値は $x = 3.98$ である。



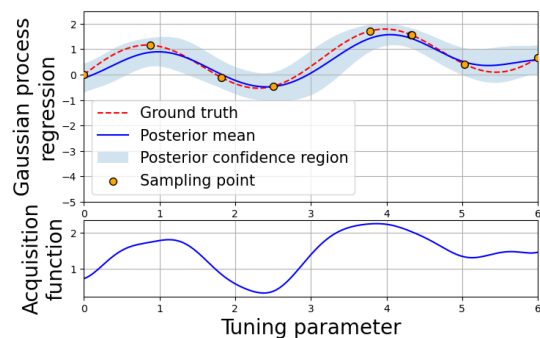
(a) 初期化後



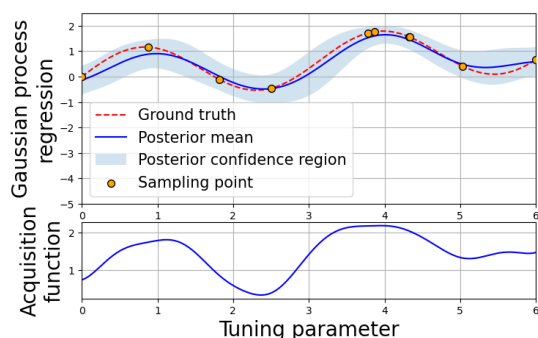
(b) 1 トライアル目



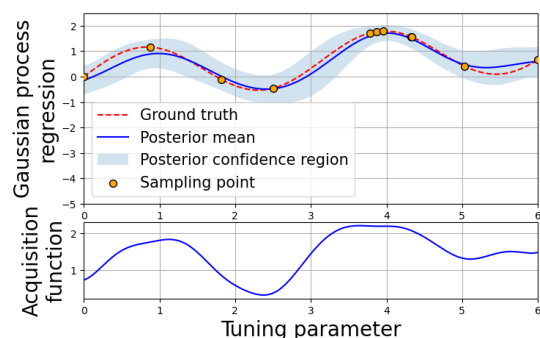
(c) 2 トライアル目



(d) 3 トライアル目



(e) 4 トライアル目



(f) 5 トライアル目

図 4.9: 1 次元でのベイズ最適化のデモンストレーション。初期条件を 5 点与え、予測分布を作成し、獲得関数が最大となった点の測定を繰り返す。測定点の情報が増えれば予測分布と獲得関数の両方の精度が向上するため、より精度良く最大値を探すことができる。

初期条件として、ランダムに選ばれた 5 点を測定した状態を考える (図 4.9a)。上段の図では、初期条件に基づいて作成された予測分布の平均を青線、標準偏差を水色の領域で示している。下段の図は獲得関数を示しており、獲得関数は UCB ($\beta = 2$) を用いている。図 4.9b から図 4.9f では、トライアルを繰り返すことで関数の最大値を探索している。5 回目のトライアルでは、真の最大値 ($x = 3.98$) に対して誤差が 0.01 の点 ($x = 3.97$) を発見できている。この精度で最大値に近い点をグリッドサーチによって確実に見つけるためには 600 回の測定が必要であるため、初期条件を含めてわずか 10 回の測定で達成可能なベイズ最適化の効率性が見て取れる。

4.4 入射調整に対する機械学習手法の導入

4.4.1 ベイズ最適化で用いる調整/評価パラメータの選定

本節ではベイズ最適化を入射調整に適用するためのパラメータ選定について述べる。はじめにベイズ最適化に使用する調整パラメータの選定をする。これは前節での x に該当する。次に、評価パラメータの選定をする。これは前節では y に該当し、本研究においてはブラックボックスとなっている入射効率を指す。

■調整パラメータの選定 入射調整パラメータは表 4.1 に示したように合計 7 個のパラメータで構成される。この中で、y-position、y-angle、Septum Angle の 3 パラメータは入射効率に対して大きな影響を及ぼすことが経験的に知られており、手動による入射調整においても他の 4 つのパラメータに比べて使用頻度が非常に高い。実際、オペレータは入射効率が低下した際、ほとんどの場合においてこれらの 3 つのパラメータを最初に調整する。

本研究は、メインリングへの入射調整に機械学習を適用する初めての試みであるため、まずは重要性が高く、かつ安全性が確保されている y-position、y-angle、Septum Angle の 3 パラメータを調整パラメータとして選定した。ここで、機械学習に使用しなかった 4 つのパラメータについて、導入を見送った理由を述べる。まず、Septum Position については急激なビームロスが発生してしまう可能性があり、かつスクリーンモニターなどで位置の補正が可能であるからである。4.1 節で述べたように、Septum Position はセプタム隔壁にできるだけ近く、かつ隔壁には当たらない位置に調整する方が好ましい。そしてこのパラメータはビーム軌道を補正するためのモニター装置であるスクリーンモニターを用いて一定期間ごとに補正されているため、ブラックボックス最適化に入れて探索させる類のパラメータではないと判断し採用を見送った。また、セプタム隔壁に当たった瞬間にビームロスが急激に増えてしまい、ビームアボートを起こしてしまう点も最適化に用いづらい。次に Kicker Height と Kicker Jump についてであるが、これは蓄積ビームの向きを変化させるパラメータであり、不適切なパラメータ設定をしてしまった場合に、加速器や Belle II 測定器に与える損傷の度合いが入射ビームに比べて非常に大きいため、今回は採用を見送った。inj. phase については、Linac を含む入射器全体の RF 位相を変化させるパラメータである。Linac は SuperKEKB 以外に PF-ring や PF-AR-ring にも入射を行っており、不適切なパラメータ設定によって大きく位相の値を変えてしまった場合に他のリングに与える影響が大きいため、今回は採用を見送った。

■評価パラメータの選定 評価パラメータとして、DCCT と BCM の両方を用いた入射効率が存在するが、今回は BCM を用いた入射効率を採用した。その理由は、DCCT を用いた入射効率の電流依存性が大きいことである。本研究では電流を一定に維持するトップアップ入射時に入射調整を実施しているが、ベイズ最適化の最中には入射効率が低くなるパラメータセットも探索する可能性があるため、一定電流が維持できなくなることがある。この際、電流が減少したことによって入射効率が高く見積もられてしまい、最適化がうまく機能しないことがあるため、電流依存性の少ない BCM を採用した。以降の記述において、入射効率とは特に断りのない限り BCM を用いて測定された入射効率を指すものとする。

入射効率の測定誤差およびふらつきについて議論する。本研究では、ベイズ最適化における評価パラメータ

として入射効率を使用する際、測定誤差を低減する目的で 30 回の測定値の中央値を採用した。平均値ではなく中央値を採用した理由は、稀に入射効率が極端に低い外れ値となる場合があり、それが平均値に大きく影響を与えるためである。図 4.10 に、KEK Linac からメインリングまでビーム状態を変更せずに加速器運転を行った 30 分間の入射効率の変動と、30 回の移動中央値を適用した結果を示す。青線が入射効率、橙線が入射効率の 30 点の移動中央値を表しており、入射効率の中央値自体がふらついていることが見て取れる。

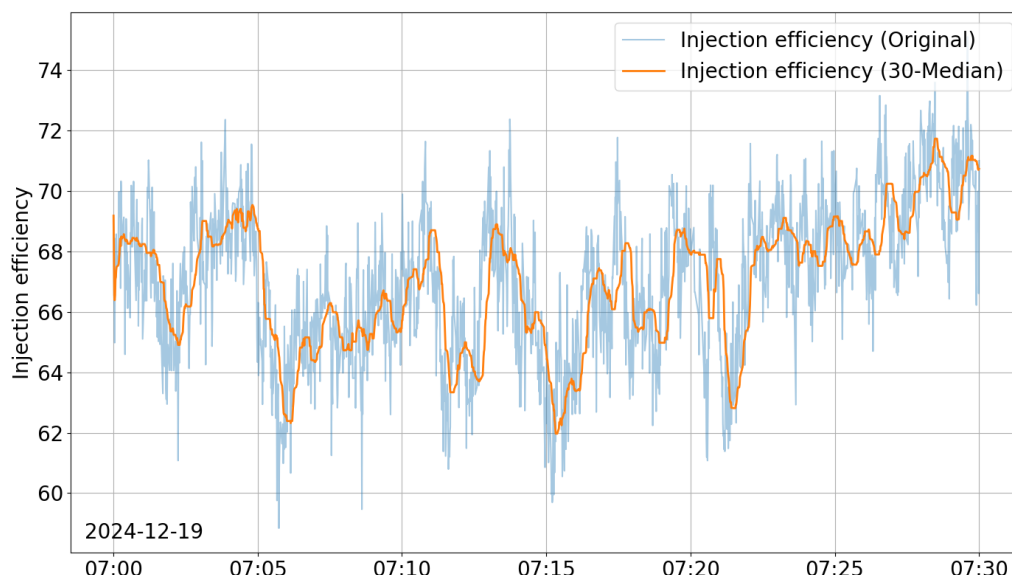


図 4.10: 30 分間ビームの状態を変化させずに測定した入射効率。青線が入射効率、橙線が入射効率の 30 点の移動中央値を表している。橙線を見ると、入射効率の中央値自体がふらついていることがわかる。

図 4.11 に 30 分の入射効率をヒストグラムにし、正規分布でフィッティングしたプロットを示す。この図には測定誤差と入射効率自体のふらつきの両方の効果が含まれており、フィッティングの結果、平均値は 67%、標準偏差は 2.7% であることがわかった。

図 4.10 および図 4.11 に示されている入射効率は、特定のビーム条件下での 30 分間のデータに基づくものである点に留意が必要である。同様の解析を別の日の異なるビーム条件において 2 回実施し、その結果を表 4.3 にまとめた。この表から、入射効率のばらつきは蓄積電流や入射効率に依存せず、時期や条件により異なる値を示すことが分かる。本研究では、入射効率の標準偏差を表 4.3 で得られた 3 日の標準偏差の平均値として設定し、LER では 2.32%、HER では 2.50% と定義した。また、ベイズ最適化による入射効率の向上が統計的に有意であるかを判定するため、標準偏差の 3 倍を閾値とし、それぞれ LER では 6.96%、HER では 7.50% 以上の向上を判定基準に採用した。

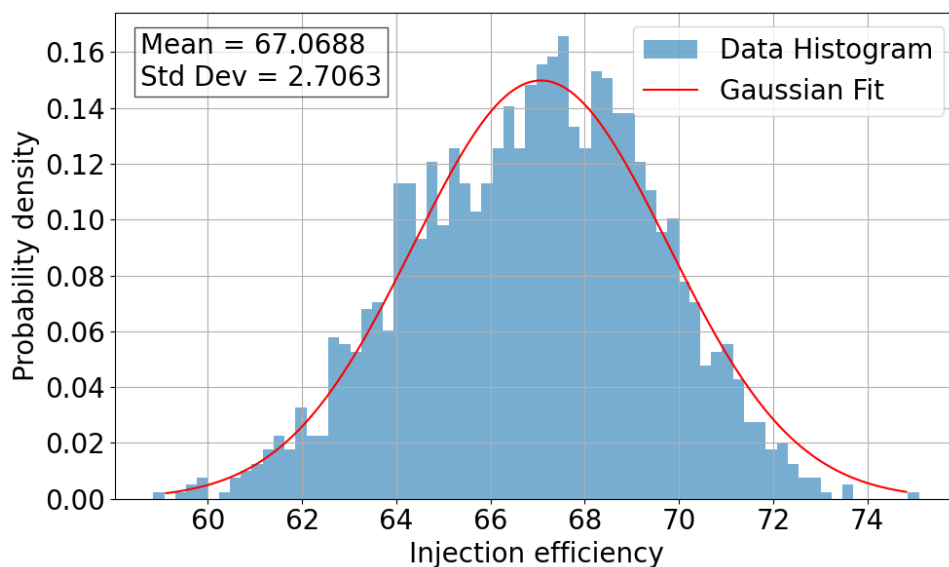


図 4.11: 30 分間ビームの状態を変化させずに測定した入射効率のヒストグラムとその正規分布フィッティング。測定誤差と入射効率自体のふらつきの効果が含まれている。フィッティングの結果、平均値は 67%、標準偏差は 2.7% であった。

表 4.3: パラメータを固定した条件下で測定した入射効率の平均値と標準偏差

測定日時	リングの種類	蓄積電流	入射効率の平均値	入射効率の標準偏差
12 月 5 日 0:30–1:00	LER	1500 mA	90.8%	1.86%
	HER	1200 mA	73.5%	2.54%
12 月 12 日 4:00–4:30	LER	1200 mA	54.8%	2.40%
	HER	960 mA	59.9%	2.70%
12 月 19 日 7:00–7:30	LER	1300 mA	67.1%	2.71%
	HER	1040 mA	66.9%	2.25%
3 回の測定の平均	LER			2.32%
	HER			2.50%

ここからは、2 バンチ入射を行っている際に、1 バンチ目や 2 バンチ目の入射効率を単独で最適化するのではなく、1 バンチ目と 2 バンチ目の入射効率を両方考慮したい場合の入射効率について述べる。この場合は式 (4.5) に示す $E_{2\text{bunch}}$ を調整パラメータとして最大化する。これはそれぞれのバンチの入射効率をそれぞれのバンチ電流で重みづけした入射効率であり、 $E_{2\text{bunch}}$ を最大化することはリング内に入る電流量を最大化することを意味する。

$$E_{2\text{bunch}} = \frac{(1 \text{ バンチ目の電流})(1 \text{ バンチ目の入射効率}) + (2 \text{ バンチ目の電流})(2 \text{ バンチ目の入射効率})}{(1 \text{ バンチ目の電流}) + (2 \text{ バンチ目の電流})} \quad (4.5)$$

4.4.2 本研究におけるベイズ最適化のセットアップ

ベイズ最適化を入射調整に適応させるためには、ビームアボートを考慮する必要がある。入射調整パラメータを大きな定義域の中で自由に変更すると、入射由来のビームバックグラウンドが大きくなってしまいう領域が生じ、パラメータがその領域に到達するとビームアボートが促され、蓄積電流を消失する結果となる。そのため、入射由来のビームバックグラウンドをある一定以下に抑え、安全性を確保した状況でベイズ最適化を実施する必要がある。本研究においてはこの条件を満たして最適化するための2つの手法を用意した。図4.12に2つの最適化手法の概略図を示した。

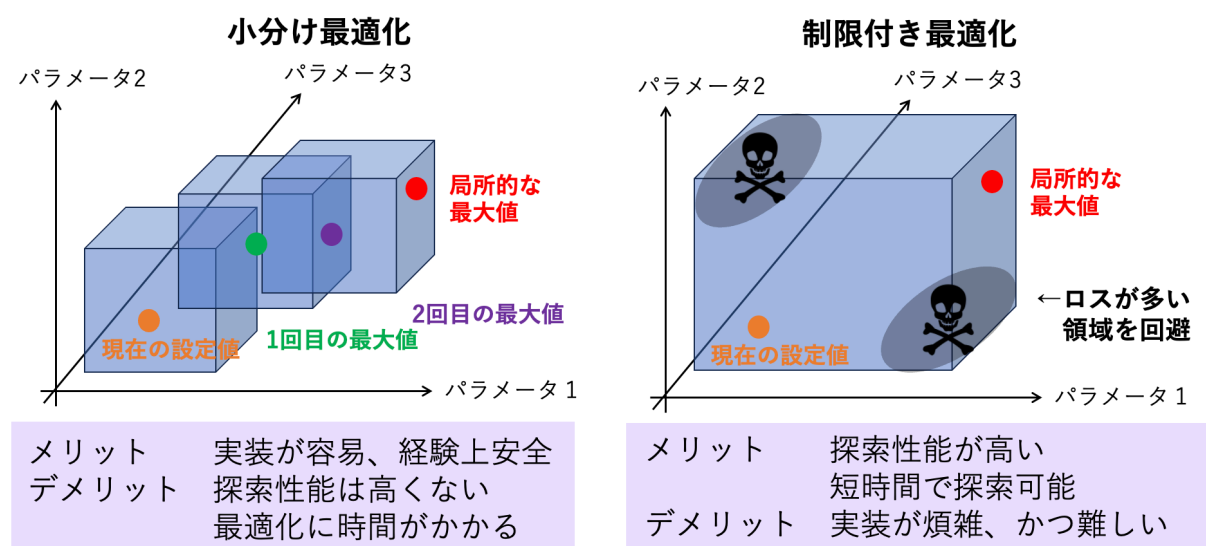


図 4.12: 二つの最適化手法。左図が小分け最適化を、右図が制限つき最適化を表しており、それぞれに図中に示したメリットとデメリットがある。

1つ目は小分け最適化であり、小さな定義域幅で最適化を行い、最適化の最大値を中心に再び小さな定義域幅での最適化を行う手法である。この手法は実装が容易で経験上安全であるが、後述する制限付き最適化に比べて探索性能は高くなく、かつ最適化に時間がかかってしまう。2つ目は制限付き最適化であり、入射効率の予測分布だけでなく、ロスモニターの測定値を用いることによってビームバックグラウンドの予測分布も同時に作成し、バックグラウンドが低い領域内において入射効率が高い点を最適値として選択する手法である。この手法は小分け最適化と比較して探索性能が高く、短時間で探索が可能であるが、実装が煩雑であり、制限をうまく課すことができなければアボートのリスクが高い手法でもある。本研究においては小分け最適化と制限つき最適化の両方を実装をしたが、スタディ時間の都合上、2024 後期の加速器運転では小分け最適化に注目してデータ取得を行なった。制約つき最適化の実装詳細については付録 A.4.7 に記した。

■**小分け最適化** 小分け最適化は、小さな定義域を設定し、その中で最適化を繰り返し実施する手法である。経験則として、ビームが安定的に周回している場合にはそのパラメータセットを中心とした十分に小さなパラ

メータ領域（表 4.4 参照）では急激なビームロスが発生しないことが知られている。^{*2}そのため、調整開始前のパラメータセットを中心に小さな定義域を設定し、その範囲内でベイズ最適化を行う手法が小分け最適化である。初期化条件としては、調整前のパラメータセットを基準に定義域の上下 10% の範囲内でランダムに点を生成し、それを初期点として使用する。本研究では、最適化開始前に 10 点の初期点を生成し、それを基にベイズ最適化を開始する。本研究では最適化において 30 トライアルを繰り返し、得られた最大値が定義域の端に位置する場合には、その最適値を中心とした新たな小さな定義域を設定し、最適化を 1～3 回程度繰り返す。この手法により、結果的に広い定義域に対しても効率的な探索を行うことが可能となる。一方、この手法には課題が 2 点存在する。1 点目は最適化を複数回繰り返すため、後述する制限付き最適化と比較して時間がかかる傾向にあることである。初期化条件の 10 点を便宜的にトライアル数に含めた場合、40 トライアルを 1 ランとすると、入射頻度や蓄積電流に依存するが、1 ランあたりおおむね 25 分程度を要する。そのため、最適化に 3 ランを必要とした場合、調整に 1 時間以上を要してしまう。2 点目は、定義域を小さく設定することから、局所的な最大値に捕獲されるリスクが高まり、広い定義域を探索可能な制限付き最適化と比較して最適化性能が低下する可能性があることである。

本研究で用いた調整パラメータに対する定義域の詳細を、表 4.4 に示す。この最適化には制限つき最適化が加わっていないため、ビームロスの高い領域が含まれている場合は高確率でアボートしてしまう。そのため、定義域の値を保守的に設定する必要がある。本研究では調整オペレーターが手動調整する際の経験を元に、1 回のランあたりの定義域の幅を定めた。

表 4.4: 各パラメータについての定義域の幅

調整パラメータ	y-position [mm]	y-angle [mrad]	Septum Angle [mrad]
1 回のランでの定義域の幅	0.020	0.020	0.020

以下では、本研究においてベイズ最適化を適応する際に実装した機能を 3 点紹介する。1 点目は step by step 機能であり、この機能がなければ、パラメータを動かしている最中にビームアボートを引き起こしてしまう。2 点目は近傍探索機能であり、最適化時間を短くするために導入した機能である。3 点目はノイズ標準偏差であり、予測分布がデータを過剰にフィットすることを防ぐパラメータである。

■Step by Step 機能 パラメーターを動かす速度が早すぎる場合、ビーム軌道を補正する軌道フィードバックより早く軌道を変化させてしまうため、ビームアボートを引き起こしてしまう。実験初期段階においてはこの点に留意しておらずアボートが多発した [27]。その具体的な原因は特定できていないが、考えられる要因として、パラメータの急激な変化によりビームの振幅が増大し、それを補正するための軌道フィードバックが十分に機能せず、軌道の修正が追いつかなくなることが挙げられる。step by step 機能を導入する目的は、入射調整パラメータを時間をかけて少しずつ動かし、振幅の増大を緩和することである。図 4.13 に step by step 機能の概念図を示した。

図 4.13 の赤点線は step by step 機能を実装しなかった場合の調整パラメータの変動であり、急激な変化が起こっているが、青線の step by step 機能を実装しなかった場合の調整パラメータの変動では急激な変化を

^{*2} ただし、調整パラメータに Septum Position や Kicker Height が含まれる場合、この仮定は適用できない。

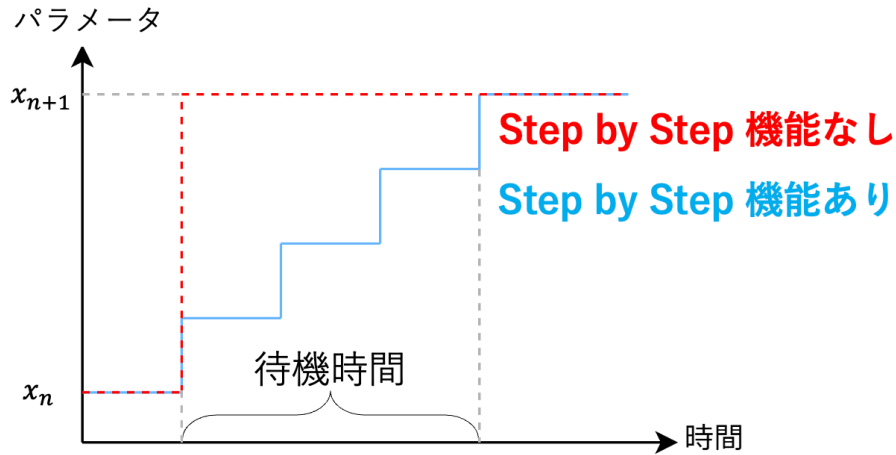


図 4.13: step by step 機能の概念図。本機能がないと、調整パラメータが急激に変化してしまい、ビームの振幅を増大させてしまう（赤点線）。本機能を導入することにより、調整パラメータの急激な変化を防ぎ、振幅の増大を緩和することができる（青線）。

防ぐことができる。ただし、パラメータを変化させるまでの待機時間が増え、最適化時間も増加するというデメリットが存在する。表 4.5 に今回のベイズ最適化で使用する各パラメータについての 1 秒あたりの移動距離を記す。この値はオペレータが調整する際の経験則を元に決定している。

表 4.5: 各パラメータについての 1 秒あたりの移動距離

調整パラメータ	y-position [mm]	y-angle [mrad]	Septum Angle [mrad]
1 秒あたりに動かす step 幅	0.0015	0.0015	0.0015

■**近傍探索機能** 近傍探索機能とは、獲得関数に対して距離の重み付けをして、現在地点からの距離が近い候補点ほど評価値が高くなるようにする機能である [28] [29]。この機能を導入する理由は、Step by Step 機能を導入してパラメータを大きく動かすと待機時間が増加してしまため、1 トライアルあたりのパラメータ変化量を減少させ、最適化時間を減らすためである。この機能を実装すると図 4.14 のように現在のパラメータセットの近傍のパラメータを優先して探索するようになる。

式 (4.6) は近傍探索の重みづけを表す式であり、 $\tilde{\alpha}(\mathbf{x})$ は重みづけ後の獲得関数、 $\alpha(\mathbf{x})$ が重みづけ前の獲得関数、 $\mathbf{x}_0$ が現時点でのパラメータセットである。探索点 $\mathbf{x}^*$ と $\mathbf{x}_0$ の距離が離れているほど $\tilde{\alpha}(\mathbf{x})$ が小さく見積もられる。 w は重み付けのハイパーパラメーターであり、近傍探索度と呼ぶ。近傍探索度が大きいほど現在のパラメータと距離が近いパラメータが優先されるようになる。第 5 章では近傍探索度の違いによるパラメータ変化の振る舞いと最適化時間の違いを議論している。

$$\tilde{\alpha}(\mathbf{x}^*) = \alpha(\mathbf{x}^*) \exp\left(-\frac{w^2(\mathbf{x}^* - \mathbf{x}_0)^2}{2}\right) \quad (4.6)$$

近傍探索機能のデメリットは、最適化性能が低下することである。式 (4.6) においては、距離による重み付

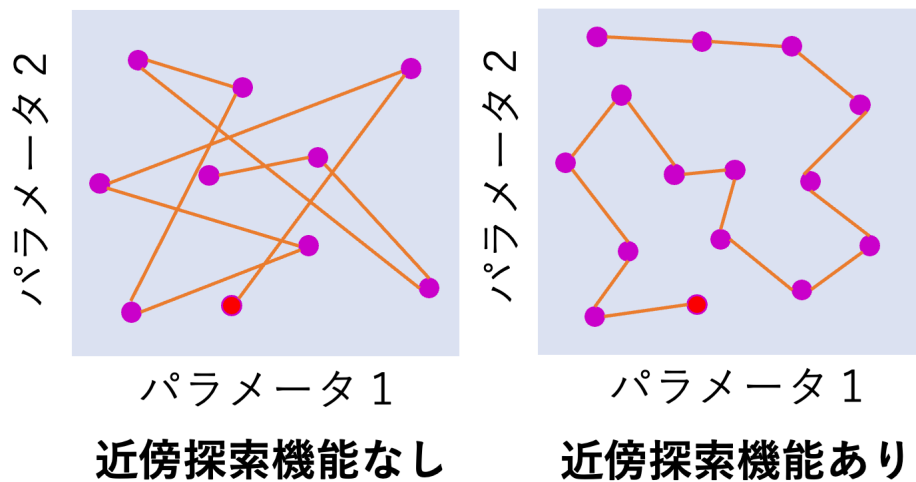


図 4.14: 近傍探索機能の概念図。2 次元の調整パラメータ空間において、近傍探索機能の有無に対する軌跡の違いを表している。左図ではトライアルごとにパラメータが大きく変動し、step by step 機能による待機時間が多く発生するが、右図ではトライアルあたりの変動が小さいために待機時間を削減することができる。

けによって現在位置と距離が近いパラメータが優先されるため、獲得関数が定めた最も最大値らしい値を反映しない。近傍探索度は最適化性能と最適化時間のトレードオフな関係を調整する役割を持つ。

■**ノイズ標準偏差** ノイズ標準偏差とは、式 (4.1) における σ である。ガウス過程回帰において σ を大きくすると、予測分布の分散が大きくなるため、モデルが測定誤差を考慮し、データを過剰にフィットすることを防ぐ。詳しくは付録 A.4.5 に記している。第 5 章ではノイズ標準偏差の違いによる調整パラメータの振る舞いの違いを議論している。

4.4.3 GUI による入射調整ツール操作

入射調整を誰でも円滑に行うことができるように、調整ツールの GUI を作成し、調整オペレータの要望に合わせてブラッシュアップを繰り返している。この GUI は既に SuperKEKB 制御棟に実装済みであり、後述するスタディ結果もこの GUI を用いて測定されたものである。図 4.15 は GUI のメイン操作画面である。このページはオペレータが操作することを想定しており、ベイズ最適化が Set ボタン → Start ボタンと簡易な操作で実行可能である。また、Set ボタンを押した後に定義域の範囲をずらしたい場合もスピンボックスを操作することで変更が可能である。過去の最適化セッティングを用いたい場合は Load ボタンから過去の設定ファイルを参照でき、Save ボタンを押せば現在の設定を保存できる。最適化を途中で止めたい場合には Pause ボタン、再開したい場合には Resume ボタン、最適化を途中で止めて現在までの一番入射効率が高いパラメータにセットするには Set the Best and Finish ボタン、GUI を閉じるには Quit ボタンを押せばよい。その他、オペレーターが普段操作しない設定に関しては Expert Setting という別の tab にまとめ、開発者や調整エキスパートのみが操作できるようにしている。

図 4.16 は GUI メインページの下部の、リアルタイムで最適化を監視できるモニターの画像であり、最適化

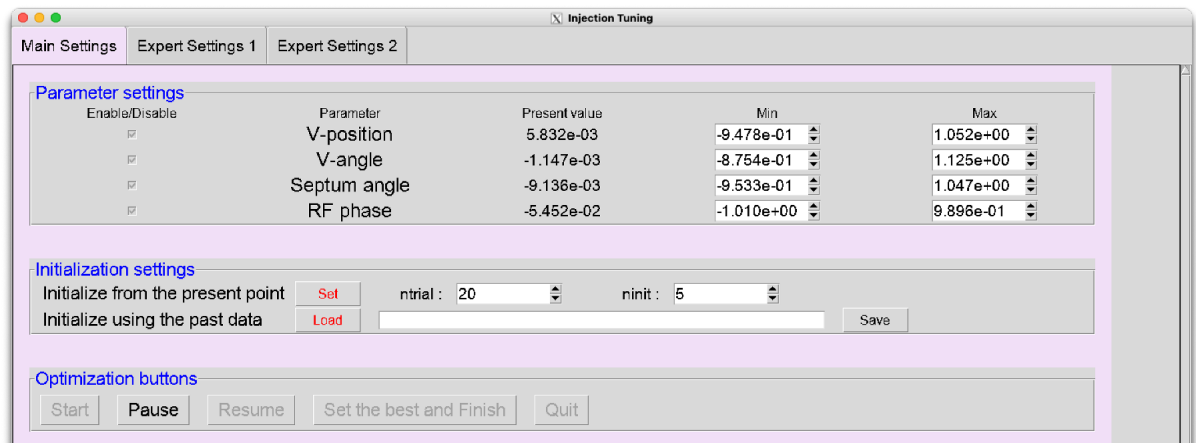


図 4.15: 最適化ツールの GUI メインページ上部。Parameter settings では調整パラメータの使用の可否や定義域を設定できる。Initialization settings ではトライアル数や初期化の回数を設定することや、過去のコンフィギュレーションを参照することができる。Optimization buttons では最適化の開始、一時停止、再開などの操作ができる。Expert Settings には step by step 機能の幅や近傍探索機能のハイパーパラメータなど、普段の調整ではオペレーターが変更しないパラメータを変更することができる。

途中でも入射効率やロスモニターの値を見ることができる。なお、図 4.16 はテスト関数を用いて最適化を行っている様子である。

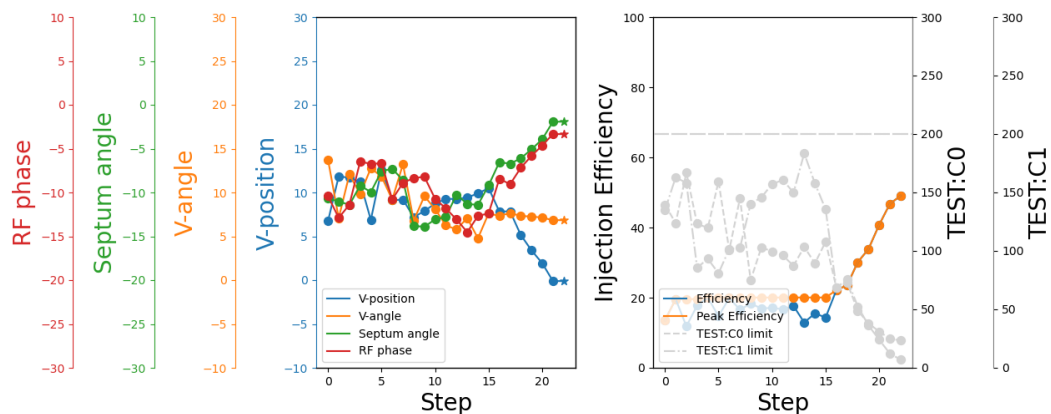


図 4.16: 最適化ツールの GUI メインページ下部。テスト関数を用いて最適化を行っており左図はトライアル数に対する入射パラメータの変動を、右図ではトライアル数に対する入射効率 (青点) とその最大値 (橙点) を見ることができる。制限つき最適化では、ロスモニターで測定したビームロスレート (灰線) も表示される。制限は複数課することができ、この図では TEST:C0 と TEST:C1 という 2 つのビームロスレートを描写している。

第 5 章

機械学習を用いた入射調整の運用

本章では、小分けベイズ最適化を用いた入射調整ツールの運用結果およびスタディ結果について報告する。5.1 節では小分けベイズ最適化を用いた入射調整ツールの具体的な運用例を示す。5.2 節では本手法の再現性を確認するため、1-2 時間程度の同一ビームコンディションおよび同一定義域から最適化を開始し、入射効率が一貫して向上するかを検証した。5.3 節では入射調整における獲得関数、ハイパーパラメータ、ノイズ標準偏差について比較を行った結果を述べる。5.4 節では、11 月から 12 月にかけて実施した入射調整ツールの運用結果を報告する。

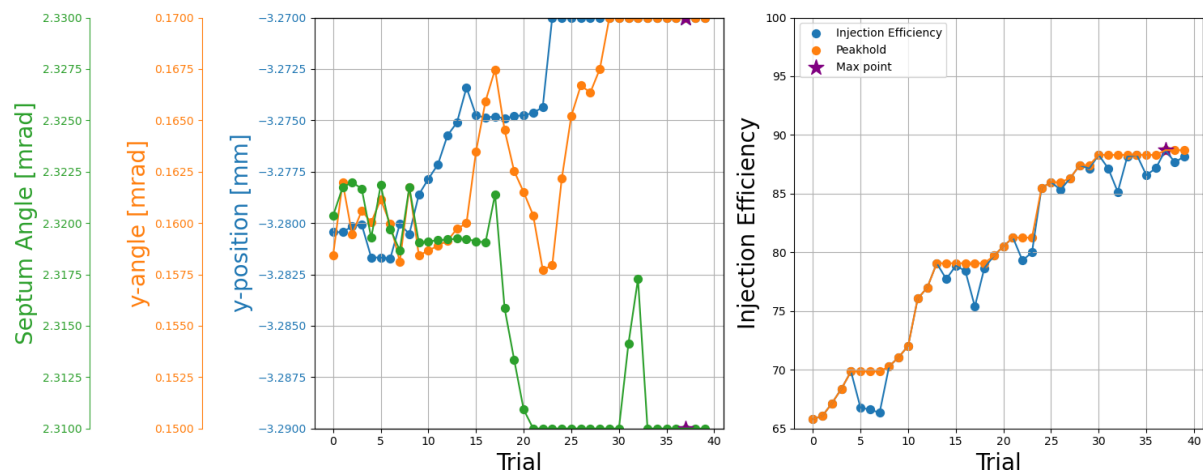
5.1 運用例

本節では、入射調整ツールの運用内容をより具体的に理解できるよう、ツールの使用により入射効率が改善され、顕著な成果が得られた事例を 2 点紹介する。

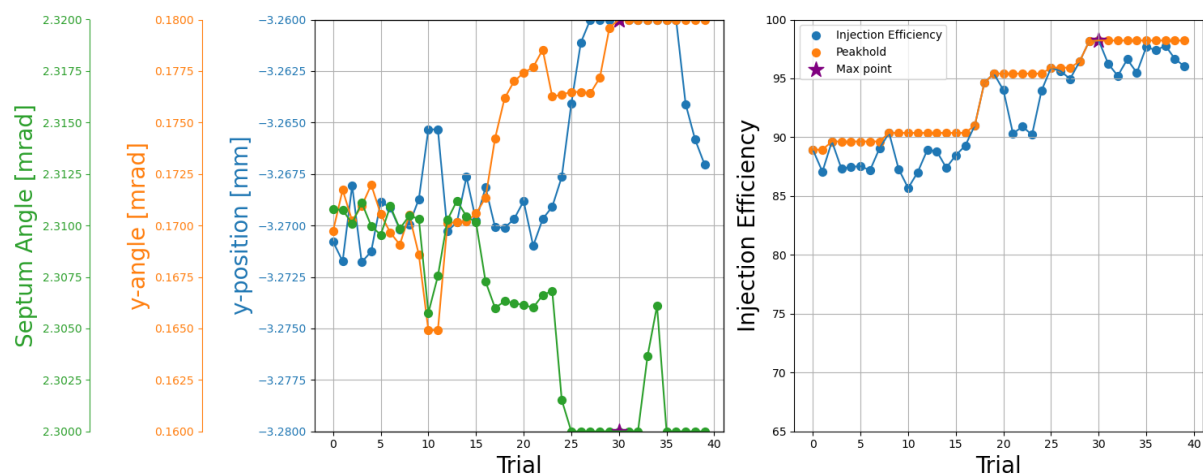
■運用例 1：入射効率改善によるルミノシティ向上への貢献 運用例 1 では、入射調整ツールを用いたことにより、LER の入射効率が 66% から 98% に改善され、蓄積電流を 1000 mA から 1300 mA まで増加させることに成功した例を紹介する。2024 年 12 月 5 日の 19 時時点における蓄積電流は LER が 1000 mA、HER が 800 mA であった。LER の入射効率が低いため、電流を 1000 mA 以上に積むことが困難な状況であり、LER の入射効率の低さが両リングにおける蓄積電流を制限していた。^{*1}最適化開始前のピークルミノシティは $2.01 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ であった。19:12 において、1 バンチ目の入射効率が 66%、2 バンチ目の入射効率が 52% の状態から最適化を開始した。本最適化では、目的関数として 1 バンチ目の入射効率を採用し、獲得関数を EI、近傍探索度を 5、ノイズ標準偏差は設定せず小分け最適化を実行した。最適化の結果を図 5.1 に示す。左図は各トライアルにおける調整パラメータの変動を示しており、定義域は調整前のパラメータ設定値 [y-position, y-angle, Septum Angle] : $[-32.80 \times 10^{-4} \text{ mm}, 1.60 \times 10^{-4} \text{ mrad}, 2.32 \text{ mrad}]$ を中心に、表 4.4 に示した幅で設定している。右図はトライアルごとの入射効率を示しており、青線が入射効率、橙線が最適化途中での入射効率の最大値である Peakhold を表している。また、紫点は最大値を記録したトライアルに付加されるマーカーである。最適化開始直後の 10 点は初期化条件に基づき生成されたものであり、それ以降

^{*1} メインリングでは、ビームビーム相互作用がルミノシティに及ぼす負の影響を緩和するため、LER と HER の電流比を 10:8 または 10:7 に設定している。

の 30 点はベイズ最適化によって選ばれている。本研究では、便宜上、初期化条件に基づく点もトライアル数に含めて扱うこととする。小分け最適化の 1 回目では、初期パラメータセットから y-position と y-angle の値を上げ、かつ Septum Angle を下げることににより入射効率が 90 % 近くまで向上した。その後、38 トライアル目の紫点のパラメータセットを中心に再び最適化を行ったところ、同様の傾向が見られた。2 回目が終わった時点で入射効率が 98 % になったため、最適化を終了した。この時は 2 パンチ目の入射効率も 65 % まで向上した。最適化が終了した時刻は 20:02 であり、2 回の最適化を合わせた経過時間は 50 分であった。



(a) 小分け 1 回目



(b) 小分け 2 回目

図 5.1: 12 月 5 日の LER での 2 回にわたる小分け最適化結果。左図は各トライアルにおける調整パラメータの変動を表しており、右図はトライアルごとの入射効率を表している。青線が入射効率、橙線が最適化途中での入射効率の最大値である Peakhold を表している。また、紫点は最大値を記録したトライアルに付加されるマーカーである。y-position と y-angle を調整前の値から大きくし、Septum Angle が調整前の値から小さくすることで、入射効率が向上していることがわかる。

次に、図 5.1 の 2 つの最適化をまとめて 3 次元プロットにした図 5.2 を記す。この図では、3 次元のパラメータ空間において y-position と y-angle が調整前の値から上昇し、Septum Angle の値が調整前の値から減

少するほど入射効率が向上していることが見て取れる。また、この例では、入射効率の分布が連続的であり、局所的な最大値に捕捉されやすい小分け最適化を用いた場合でも、入射効率が 100% に近いパラメータセットに到達できていることが分かる。

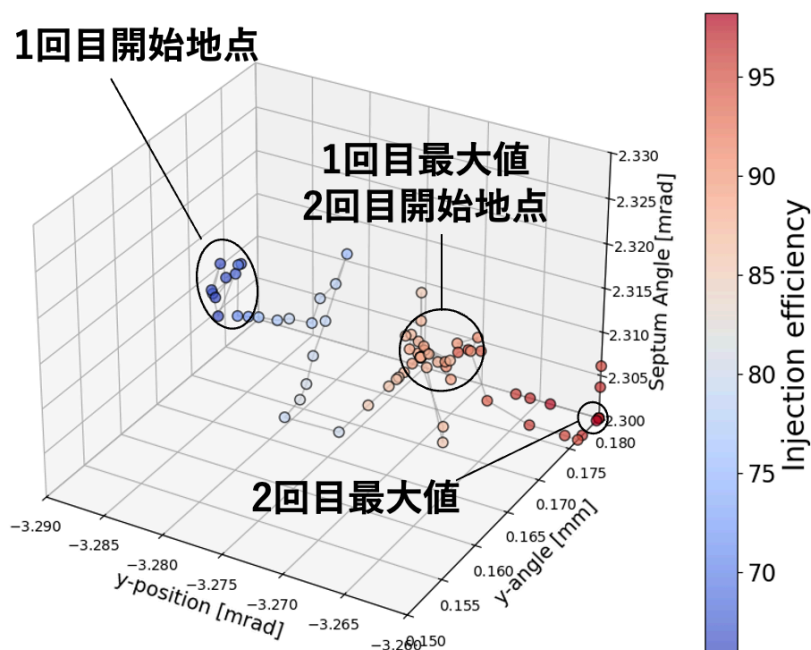


図 5.2: 12 月 5 日の LER での 2 回にわたる小分け最適化での 3 次元プロット。調整パラメータ空間における最適化の軌跡を表しており、カラーマップは入射効率を表す。y-position と y-angle が調整前の値から上昇し、Septum Angle の値が調整前の値から減少するほど入射効率が向上していることが見て取れる。また、入射効率の分布が連続的であるため、局所的な最大値に捕捉されやすい小分け最適化を用いた場合でも入射効率が 100% に近いパラメータセットに到達できていることが分かる。

この調整の結果、LER の入射に余裕ができ、蓄積電流を LER:1300 mA, HER:1140 mA まで引き上げることができた。この時、ピークルミノシティは $3.44 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ まで向上した。表 5.1 に、運用例 1 での入射効率、蓄積電流、瞬間ルミノシティの比較をまとめる。

表 5.1: 12 月 5 日の LER 入射調整前後での入射効率、蓄積電流、瞬間ルミノシティの比較

	1 バンチ目入射効率	2 バンチ目入射効率	LER 蓄積電流	HER 蓄積電流	瞬間ルミノシティ
最適化開始前	66%	52%	1000 mA	800 mA	$2.01 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
最適化終了後	98%	65%	1300 mA	1140 mA	$3.44 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$

■運用例 2：入射効率改善による LER の最大蓄積電流達成と瞬間ルミノシティの更新への貢献 運用例 2 では、入射調整ツールを用いた入射効率向上が要因の 1 つとなり、LER の最大蓄積電流である 1700 mA を達

成し、2022 年 6 月 22 日から約 2 年 6 か月ぶりに瞬間ルミノシティの記録が更新された例を紹介する。^{*2}12 月 25 日の 12 時 30 分時点での蓄積電流は LER が 1650 mA、HER が 1320 mA であり、瞬間ルミノシティは $4.70 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ であった。直後、HER においてセプタムマグネットの充電トリガーの誤動作により電流が蓄積できなくなり、復旧対応に 3 時間程度を要した。この間に LER の電流を 800 mA まで減少させ、ベイズ最適化による入射調整を行った。調整前は 1 バンチ目の入射効率が 68%、2 バンチ目の入射効率が 37% であった。ここから入射調整ツールを用いて、式 (4.5) で表される 2 バンチ入射用の入射効率 $E_{2\text{bunch}}$ が最大となるように入射調整を行った。獲得関数は UCB ($\beta = 2$)、近傍探索機能は使用せず、ノイズ標準偏差は 1% とした。最適化の結果を図 5.3 に記す。14 時 42 分から 15 時 09 分までの 27 分間の調整で 1 バンチ目の入射効率を 82%、2 バンチ目の入射効率を 55% まで向上させた。

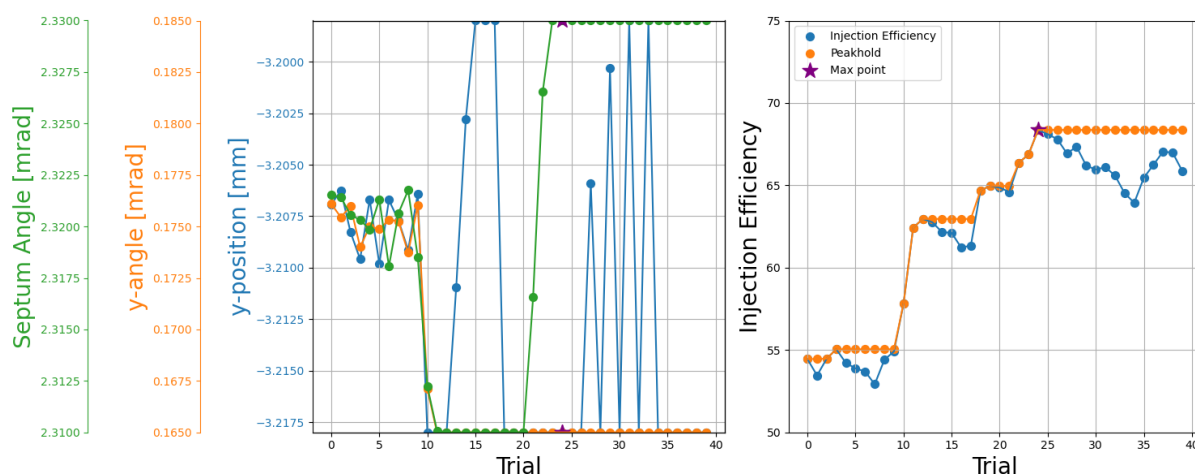


図 5.3: 12 月 25 日の LER での小分け最適化結果。式 (4.5) で表される 2 バンチ用の入射効率を 55% から 68% まで向上させた結果、1 バンチ目の入射効率が 68% から 82% に、2 バンチ目の入射効率が 37% から 55% に向上した。

その後、HER のセプタムマグネットが復旧し、両リングにおいて電流の蓄積が再開された。本ツールによる調整後、y-position および y-angle は固定されたまま、オペレーターが電流の増加率や衝突条件を監視しながら Septum Angle の調整を逐次実施した。そして、本ツールを用いた調整開始から約 3 時間半後の 18 時 38 分に、LER で蓄積電流の新記録である 1700 mA を達成し、HER の 1290 mA と合わせて瞬間ルミノシティが $4.74 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ に到達した。これにより、約 2 年 6 か月ぶりに瞬間ルミノシティの記録が更新された。表 5.2 に、運用例 2 での入射効率、蓄積電流、瞬間ルミノシティの比較をまとめる。

5.2 再現性の確認

この節では、最適化の再現性を確認するために、ビームコンディションが変わらないと思われる 2-3 時間の間に、同じ定義域、同じパラメータ設定で最適化を複数回繰り返し、最適化結果が同じ結果になるかを確認し

^{*2} ただし、この成果は他の様々な加速器調整の貢献が組み合わさった結果であり、入射調整ツールはその要因の一つであることに留意する必要がある。

表 5.2: 12 月 25 日の LER 入射調整前後での入射効率、蓄積電流、瞬間ルミノシティの比較

	1 バンチ目入射効率	2 バンチ目入射効率	LER 蓄積電流	HER 蓄積電流	瞬間ルミノシティ
12 時 36 分	68%	52%	1650 mA	1320 mA	$4.70 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
最適化開始前 (14 時 42 分)	68%	37%	800 mA	0 mA	$0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
最適化終了後 (15 時 09 分)	82%	55%	800 mA	0 mA	$0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$
18 時 38 分	79%	59%	1700 mA	1290 mA	$4.74 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$

た。ここで再現性とは、複数回の最適化において、入射効率が高くなる領域に規則性が見られることを指す。HER で 3 回、LER で 2 回データを取得した。

はじめに HER における再現性確認スタディについて論じる。このスタディのセットアップを表 5.3 に示す。測定は 12 月 14 日に実施され、蓄積電流は LER : 1350 mA、HER : 1080 mA、ビーム衝突あり、1 バンチ入射の条件下で行われた。パラメータの定義域は、y-position : $[9.00, 9.20] \times 10^{-4} \text{ mm}$ 、y-angle : $[-0.25, -0.05] \times 10^{-4} \text{ mrad}$ 、Septum Angle : $[0.862, 0.882] \text{ mrad}$ を設定した。

表 5.3: HER における再現性確認スタディのセットアップ

番号	測定開始時間	測定終了時間	最適化時間	獲得関数	近傍探索度	ノイズ標準偏差
1	12:14	12:38	24 分	EI	5	0
2	12:40	13:04	24 分	EI	5	0
3	13:05	13:29	24 分	EI	5	0

HER での 3 回の最適化に対して、トライアルごとのパラメータ変動とトライアルごとの入射効率をプロットしたものを図 5.4 に示す。1-3 回目の全ての最適化について、y-angle と Septum Angle は最適化前の値よりも大きい値で最大値を取っている。一方、y-position については、定義域の上端から下端まで自由に値を取っている。この振る舞いは 3 回目の最適化を見るとわかりやすい。10 トライアル付近で Septum Angle の値が上昇するとともに入射効率が 10% 程度向上し、17 トライアル付近で y-angle の値が減少したときには入射効率が 5% 程度悪化し、20 トライアル付近で y-angle がの値が上昇したときには入射効率が 10% 向上している。一方で y-position の値は 25-30 トライアルで定義域の上端から下端まで減少しているが、y-angle と Septum Angle を変化させた時と比較して入射効率に大きな変化は見られない。また、1 回目の 36 トライアル目と 3 回目の 39 トライアル目では、Septum Angle の値を定義域の上端から下端付近まで減少させた際に、入射効率が約 15% 程度大幅に減少する現象が観測されている。さらに、1 回目および 3 回目の最適化では、y-angle と Septum Angle が定義域の上端値を取っている場合に入射効率は 67% 程度を示しているが、2 回目の最適化では両パラメータが定義域の端まで探索されていない。その結果、入射効率の最大値は 65% を越えることがなかった。これらより、40 回のトライアルでは必ずしも定義域内の最大値に到達できない場合があることが示唆される。この問題を解決する方法として、以下の 2 点が考えられる。1 点目はトライアル数を単純に増加させ、探索回数を増やす方法である。この手法はより広範な探索が可能となる一方で、最適化時間の増加を招く点が課題である。2 点目はトライアル数を固定値として指定するのではなく、入射効率が 10 トライアル連続で改善しなくなった場合に最適化を終了するなど、終了条件を設けてトライアル数を動的に調整す

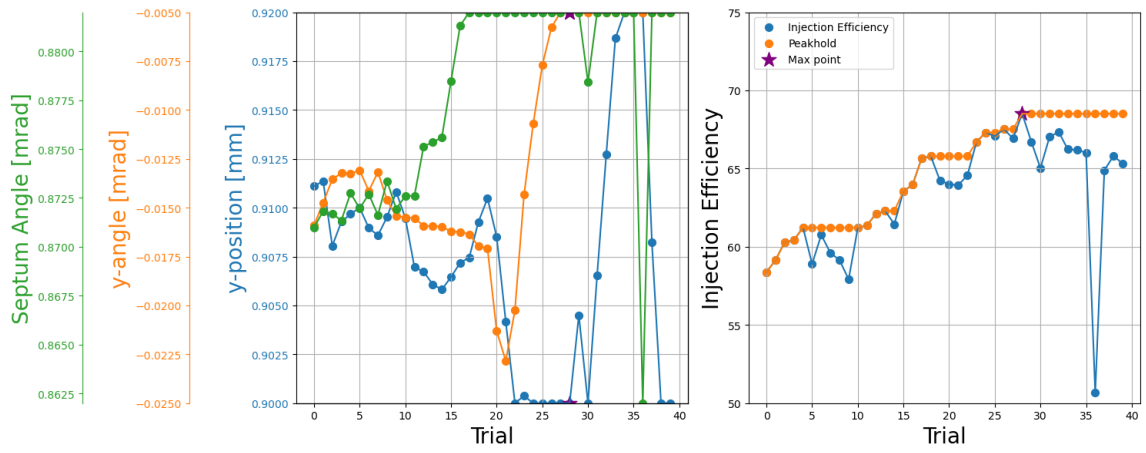
る手法が挙げられる。この方法は最適化時間を抑えつつ、効率的な探索を実現できる可能性がある。

HER における再現性確認のスタディ結果を表 5.4 に示す。

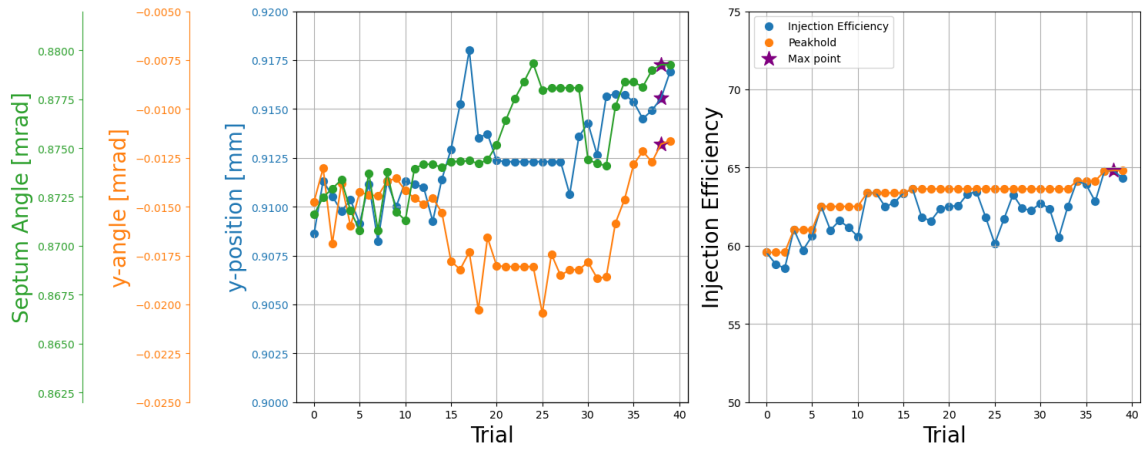
表 5.4: HER における再現性確認のスタディ結果

番号	最大入射効率	y-position [mm]	y-angle [mrad]	Septum Angle [mrad]
1	68.5%	0.900	-0.005	0.882
2	64.8%	0.916	-0.013	0.879
3	67.4%	0.920	-0.009	0.882

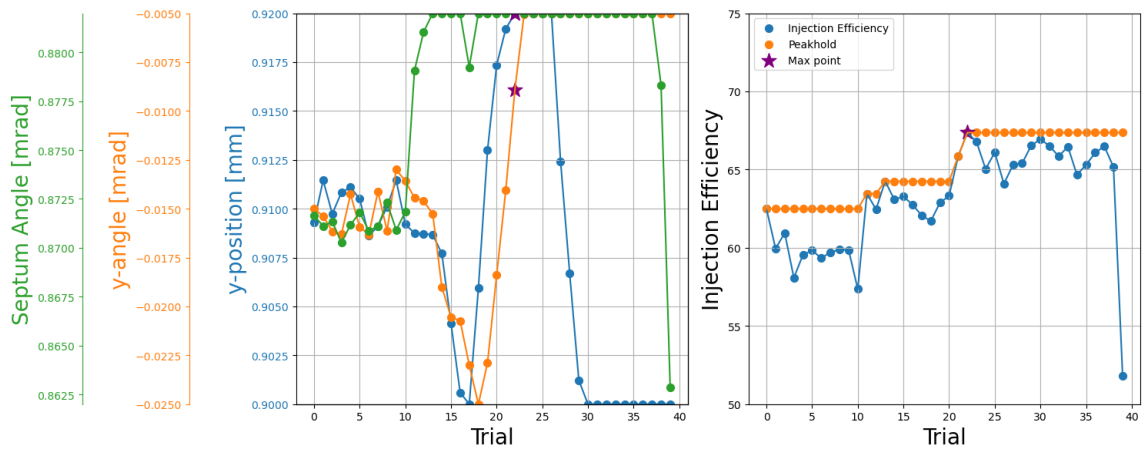
3 回の最適化における全ての点と、上位 3 点の 3 次元プロットをまとめたものを図 5.5 に示す。図 5.5a では y-position の変化に対して、入射効率があまり変わらないことが見て取れる。図 5.5b では 1 回目と 3 回目の最適化では y-angle と Septum Angle の値が同じだが、y-position のみが異なることが確認でき、また 2 回目の最適化では 1 回目と 3 回目の最適化と比較して入射効率が低く、y-angle と Septum Angle が定義域の上端にたどり着く前に最適化が終了してしまっていることが見て取れる。



(a) 1 回目

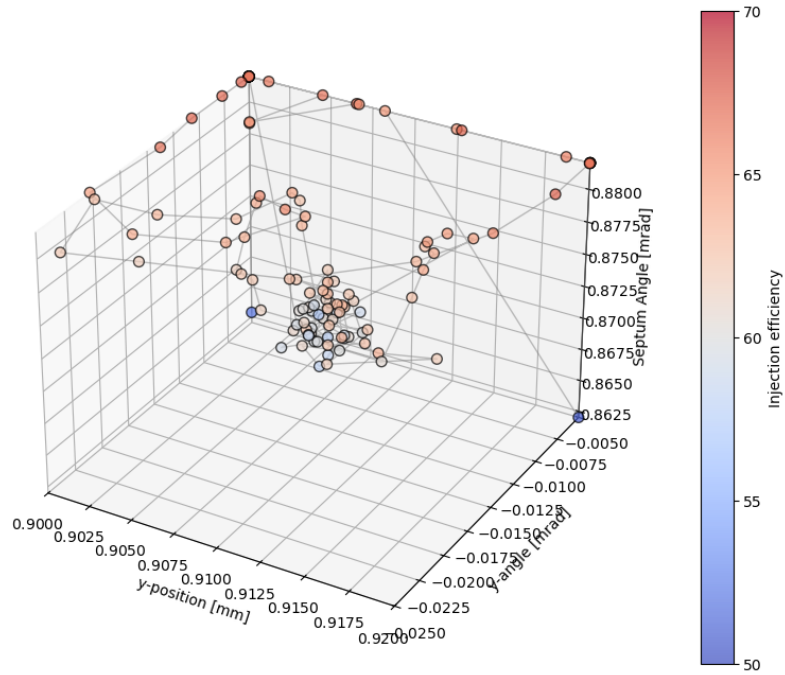


(b) 2 回目

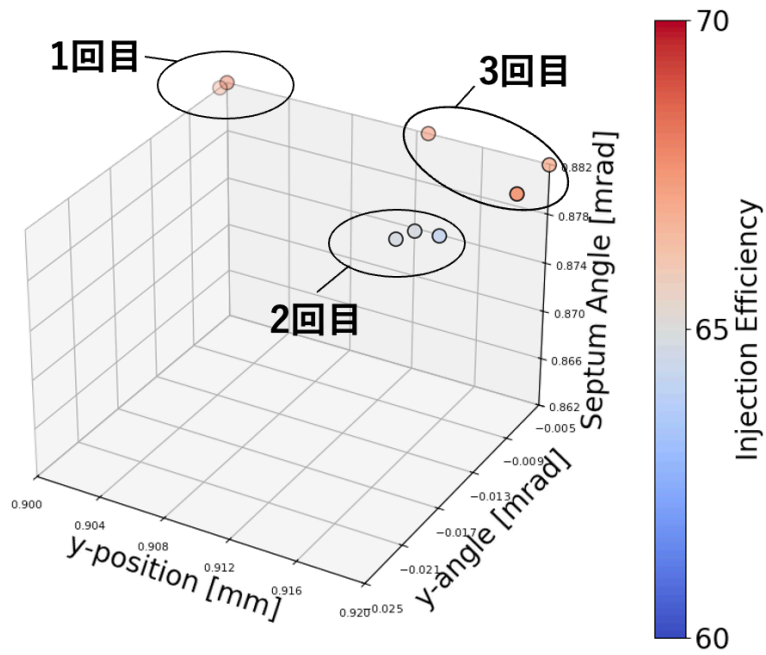


(c) 3 回目

図 5.4: HER における、ビーム状態、定義域の条件を揃えた 3 回の測定結果。3 回の測定に共通して、調整前の値に対して y-angle と Septum Angle を大きくするほど入射効率が向上している。一方で y-position は 3 回の試行において決まった位置を取っていない。



(a) 全ての点



(b) 上位3点

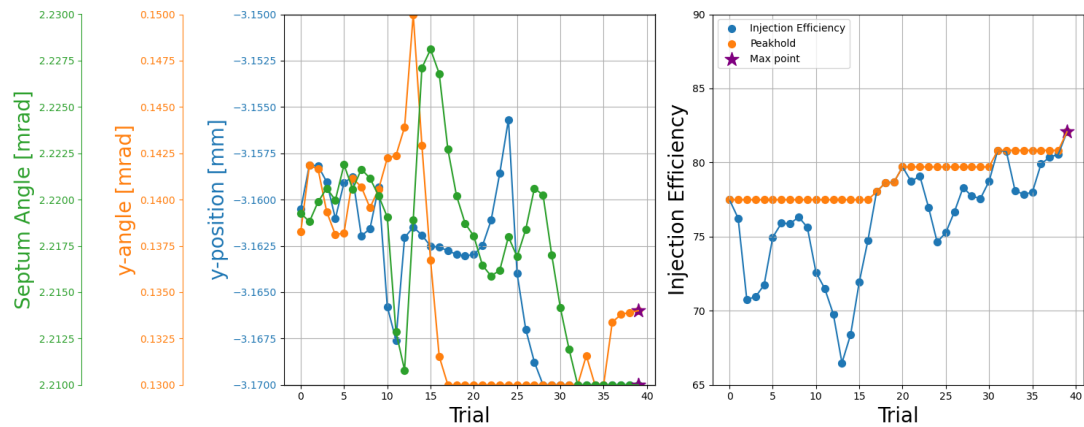
図 5.5: HER での 3 回の最適化における全ての点と、各測定における上位 3 点の 3 次元プロット。(a) では y-position の変化に対して、入射効率があまり変わらないことが見て取れる。(b) では 1 回目と 3 回目の最適化では y-angle と Septum Angle の値が同じだが、y-position のみが異なることが確認でき、また 2 回目の最適化では 1 回目と 3 回目の最適化と比較して入射効率が低く、y-angle と Septum Angle が定義域の上端にたどり着く前に最適化が終了してしまっていることが見て取れる。

次に、LER における再現性確認スタディのセットアップを表 5.3 にまとめた。測定は 2024 年 12 月 13 日に実施され、蓄積電流は LER : 1000 mA、HER : 800 mA、ビーム衝突あり、2 バンチ入射の条件下で行われた。パラメータの定義域は、y-position : $[-31.70, -31.50] \times 10^{-4}$ mm、y-angle : $[1.30, 1.50] \times 10^{-4}$ mrad、Septum Angle : $[2.21, 2.23]$ mrad を設定した。

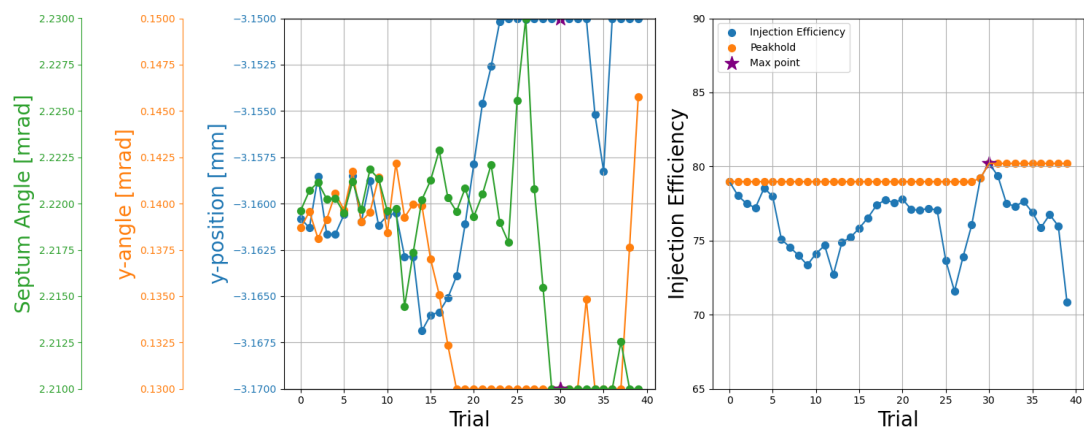
表 5.5: LER における再現性確認スタディのセットアップ

番号	測定開始時間	測定終了時間	最適化時間	獲得関数	近傍探索度	ノイズ標準偏差
1	21:41	22:05	24 分	EI	5	0
2	22:07	22:32	25 分	EI	5	0

2 回の最適化に対して、トライアルごとのパラメータ変動と入射効率を図 5.6 に記す。1、2 回目の最適化において y-angle と Septum Angle はどちらも最適化前の値より小さい値において入射効率の最大値を取っている。それに対して y-position は 1 回目では下端の、2 回目では上端の値を取った。また、1 回目の 10-20 トライアル目を見ると、y-angle と Septum Angle の値が大きくなったときに入射効率が改悪し、逆に小さくなったときには向上していることがわかる。



(a) 1 回目



(b) 2 回目

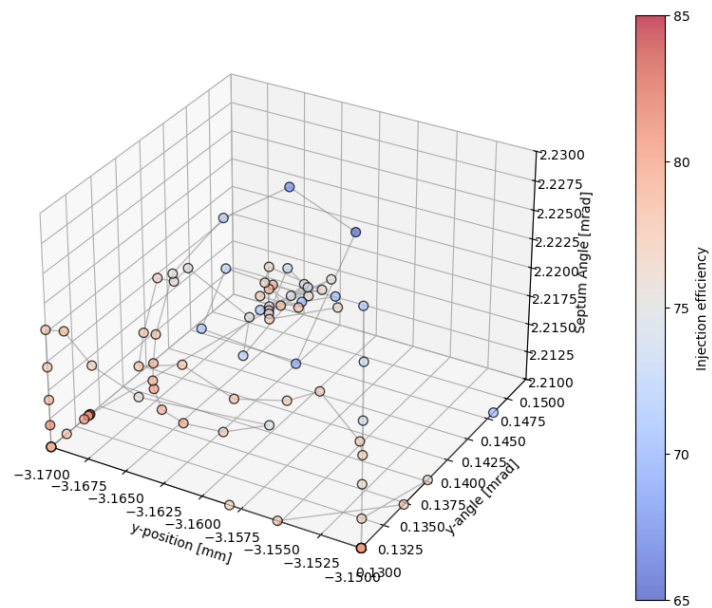
図 5.6: LER における、ビーム状態、定義域の条件を揃えた 2 回の測定結果。2 回の測定に共通して、調整前の値に対して y-angle と Septum Angle 小さくするほど入射効率が向上している。一方で y-position は 2 回の試行において決まった位置を取っていない。

LER における再現性確認のスタディ結果を表 5.6 に示す。

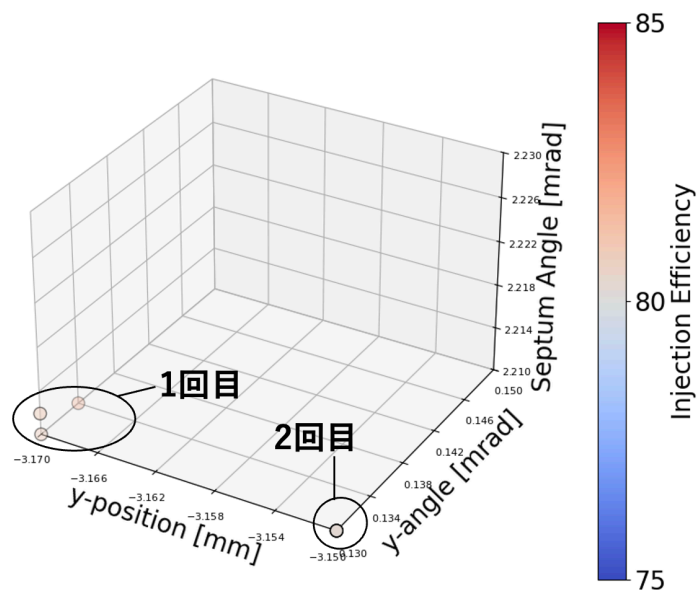
表 5.6: LER における再現性確認のスタディ結果

番号	最大入射効率	y-position [mm]	y-angle [mrad]	Septum Angle [mrad]
1	82.1%	-3.170	0.134	2.210
2	80.2%	-3.150	0.130	2.210

LER での 2 回の最適化に対して、トライアルごとのパラメータ変動とトライアルごとの入射効率をプロットしたものを図 5.7 に示す。図 5.7a では y-position の変化に対して、入射効率があまり変わらないことが見て取れ、図 5.7b では 2 つの最適化の上位 3 点に対して y-angle と Septum Angle の値が同じだが、y-position のみが違うことが確認できる。



(a) 全ての点



(b) 上位3点

図 5.7: LER での 2 回の最適化における全ての点と、各測定の上位 3 点の 3 次元プロット。図 5.5 と同様に、(a) では y -position の変化に対して、入射効率があまり変わらないことが見て取れ、(b) では 2 つの最適化の上位 3 点に対して y -angle と Septum Angle の値が同じだが、 y -position のみが違うことが確認できる。

これらの結果をまとめると、y-angle および Septum Angle に関しては再現性が確認されたが、y-position については再現性が確認できなかった。この理由として、y-position が入射効率に対する寄与が小さいパラメータである可能性や、設定した定義域の幅が狭すぎる可能性が考えられる。

3つの入射調整パラメータが入射効率にどのように寄与しているかを評価するために、SHAP (SHapley Additive exPlanations) [30] を用いて解析を行う。SHAP とは、機械学習モデルの予測結果を解釈可能にする手法であり、各パラメータがモデルの予測値に与える影響を Shapley 値として定量的に評価することができる。例として、図 5.4c (HER:3 回目) および図 5.6a (LER:1 回目) の最適化結果に基づいて計算された Shapley 値を、図 5.8 に示す。この図では、カラーマップがそれぞれのパラメータ値を示し、Shapley 値が入射効率の向上または悪化に寄与した度合いを表している。図 5.8a (HER) を見ると、Septum Angle および y-angle の値が高い場合には入射効率が向上し、低い場合には効率が悪化していることが分かる。一方、図 5.8b (LER) では、これとは対照的に、Septum Angle および y-angle の値が低い場合に入射効率が向上していることが確認できる。また、両方のケースにおいて、y-position は入射効率の変動に対してあまり寄与していないことも見て取れる。

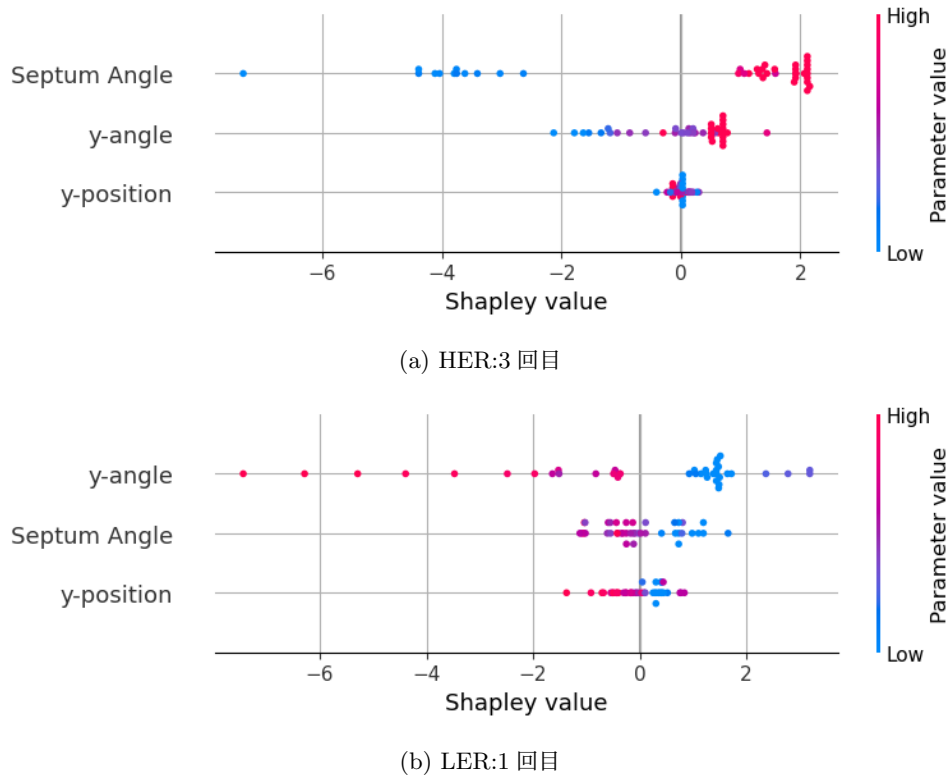
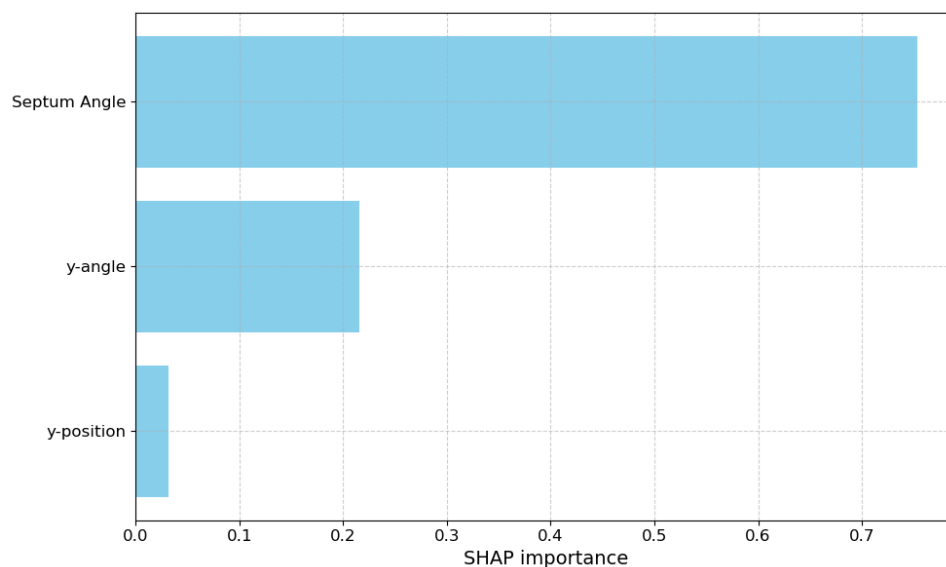
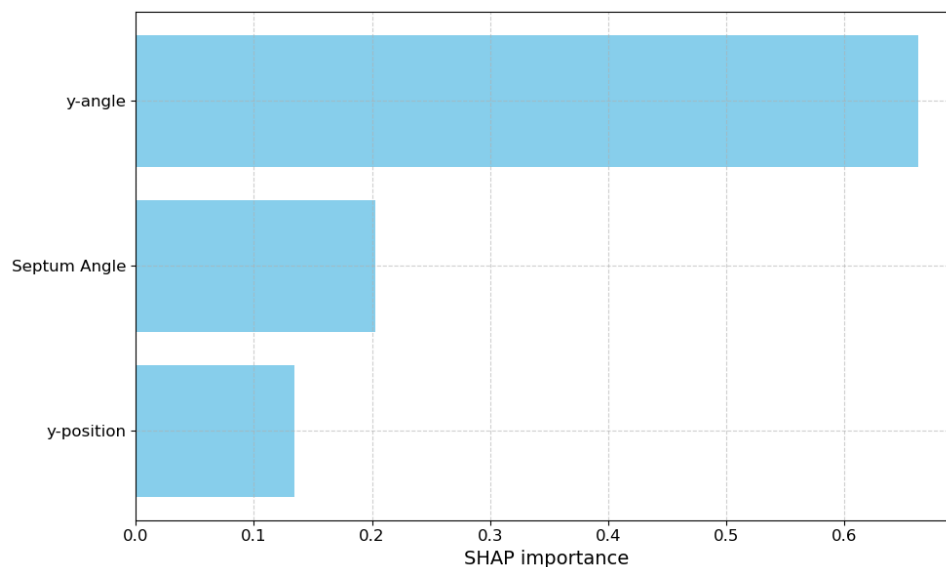


図 5.8: HER と LER の最適化における調整パラメータと Shapley 値のプロット。カラーマップがそれぞれのパラメータ値を示し、Shapley 値が入射効率の向上または悪化に寄与した度合いを表している。(a) では Septum Angle および y-angle の値が高い場合には入射効率が向上し、低い場合には効率が悪化していることが分かる。一方、(b) では Septum Angle および y-angle の値が低い場合に入射効率が向上していることが確認できる。また、両方のケースにおいて、y-position は入射効率の変動に対してあまり寄与していないこともわかる。

次に、SHAP importance を用いた解析を行った。SHAP importance は、各パラメータの Shapley 値の絶対値を合計し、それを全体で規格化して合計が 1 となるよう調整した指標であり、最適化における各パラメータの寄与度を定量的に示すものである。図 5.9 に、HER の 3 回目および LER の 1 回目の最適化における SHAP importance を示す。この結果から、Septum Angle および y-angle に比べて、y-position の入射効率への寄与が小さいことが明らかとなった。また、HER では Septum Angle が、LER では y-angle が最も高い寄与を示していることがわかる。



(a) HER:3 回目



(b) LER:1 回目

図 5.9: HER と LER の最適化における SHAP importance。Septum Angle および y-angle に比べて、y-position の寄与が相対的に小さいことが明らかとなった。また、HER では Septum Angle が、LER では y-angle が最も高い寄与を示していることがわかる。

本節で取り上げた 3 回の HER における最適化と 2 回の LER における最適化に関する SHAP importance の解析結果を表 5.7 に示す。この結果から、全 5 回のランにおいて、y-position の寄与が一貫して Septum Angle および y-angle に比べて相対的に小さいことが確認された。また、HER では Septum Angle が、LER では y-angle が最も高い寄与を示していることが明らかとなった。各パラメータが入射効率にどの程度寄与するかを定量的に評価し、その効果を明確に示したことは、特筆すべき成果であるといえる。

表 5.7: 最適化スタディにおける SHAP importance の解析結果

番号	y-position	y-angle	Septum Angle
HER1 回目	0.089	0.117	0.794
HER2 回目	0.073	0.226	0.701
HER3 回目	0.031	0.215	0.753
LER1 回目	0.134	0.663	0.203
LER2 回目	0.150	0.464	0.386

5.3 獲得関数、近傍探索度、ノイズ標準偏差の比較

本節では、獲得関数、近傍探索度、ノイズ標準偏差を変化させた際の振る舞いを比較するために実施した入射調整のスタディについて述べる。本スタディの測定条件を表 5.8 に示す。

比較に入る前に、まず入射効率の向上について結果を確認する。12 月 7 日に実施した 9 回の条件設定変更において、入射効率の向上量は 0–2.9% であり、HER のふらつきの標準偏差 (2.5%) 程度に留まり、有意な向上を示したとは言えない。この理由として、測定はオペレータによる手動調整後に行われたため、3 つの調整パラメータが既に局所的な最大値付近に設定されていたことが考えられる。ただし、入射効率が向上しない場合でも、悪化することは一度もなかった。12 月 14 日に実施した最適化では、入射効率の向上量が 4.9–5.9% と、HER のふらつきの標準偏差 (2.5%) の約 2 倍に相当する結果が得られた。しかし、本研究では有意性の基準を標準偏差の 3 倍以上と定義しているため、有意な向上とは判断しない。12 月 25 日の入射効率向上量は 8.1–13.0% であり、LER のふらつきの標準偏差 (2.3%) の 3 倍以上に相当する。このため、この日における入射効率の向上は統計的に有意であると判断できる。以降では、入射効率の向上量そのものではなく、条件を変更した際の調整パラメータの挙動や最適化に要する時間について議論する。

■**ノイズ標準偏差** 本項では、ノイズ標準偏差が探索と活用のバランスに与える影響を検討する。はじめに、表 5.8 に示す条件でノイズ標準偏差を 0.1% に設定した場合と、設定しない場合の比較を行う。ノイズ標準偏差を 0.1% に設定するということは、0.1% 以下の変化を誤差として扱い、向上と見なさないことを意味する。このため、ノイズ標準偏差の有無を比較しやすいようにはじめは 0.1% で測定を実施した。図 5.10 には、獲得関数を EI、近傍探索度を 5 に固定した状態で、ノイズ標準偏差を 0.1% 課した場合 (番号 5) と課さない場合 (番号 6) の比較結果を示している。その結果、図 5.10a では最適化が過度に活用的となり、定義域内の一部が探索されないまま 40 トライアルが終了している。一方で、図 5.10b では探索的な成分が強く現れ、定義域全体を幅広く探索している様子が確認できた。この結果は、獲得関数を UCB ($\beta = 2$) や UCB ($\beta = 5$) に

表 5.8: 獲得関数、近傍探索度、ノイズ標準偏差の測定条件表

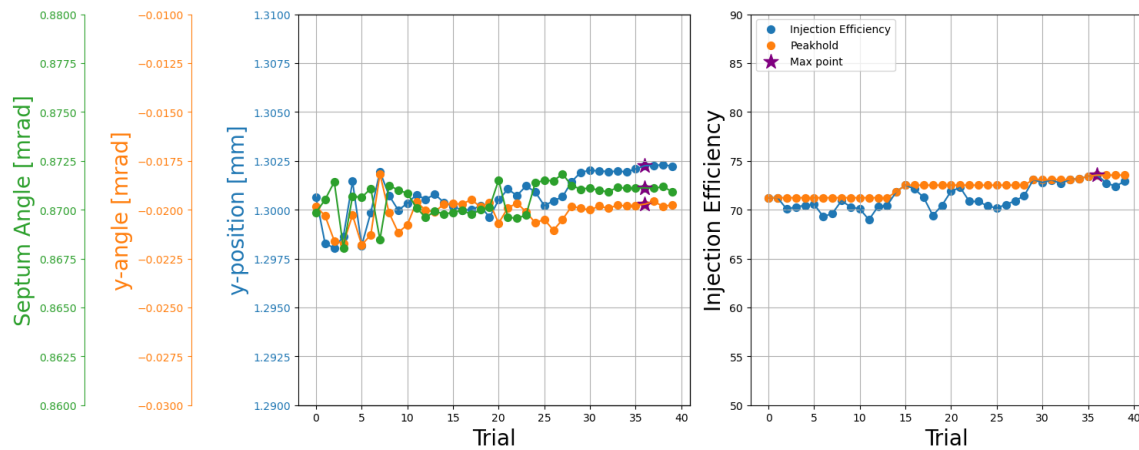
番号	測定日時	獲得関数	近傍探索度	ノイズ標準偏差	調整前/後入射効率
HER	2024 年 12 月 7 日				
1	12:50–13:14	UCB ($\beta = 2$)	5	0.1%	71.7% / 71.7%
2	13:17–13:41	UCB ($\beta = 2$)	1	0.1%	68.2% / 71.1%
3	13:43–14:07	UCB ($\beta = 2$)	0	0.1%	71.3% / 73.3%
4	14:10–14:35	UCB ($\beta = 5$)	0	0.1%	71.2% / 73.3%
5	14:38–15:02	EI	5	0.1%	71.2% / 73.5%
6	15:34–16:00	EI	5	0	74.3% / 74.3%
7	16:22–16:49	UCB ($\beta = 5$)	0	0	70.7% / 72.2%
8	16:56–17:22	UCB ($\beta = 2$)	1	0	70.7% / 70.9%
9	17:31–17:55	UCB ($\beta = 2$)	5	0	69.6% / 70.4%
HER	2024 年 12 月 14 日				
10	13:05–13:29	EI	5	0	62.5% / 67.4%
11	13:38–14:02	EI	5	1%	62.5% / 68.4%
12	14:07–14:31	EI	5	3%	64.5% / 69.4%
LER	2024 年 12 月 25 日				
13	13:40–14:09	EI	0	1%	56.9% / 69.9%
14	14:12–14:36	EI	5	1%	64.9% / 73.0%

変更した場合や、近傍探索度を 0 や 1 に設定した場合でも、ノイズ標準偏差を 0.1% に設定すると、図 5.10a に示すような活用的な最適化が行われることが確認された（番号 1–4）。

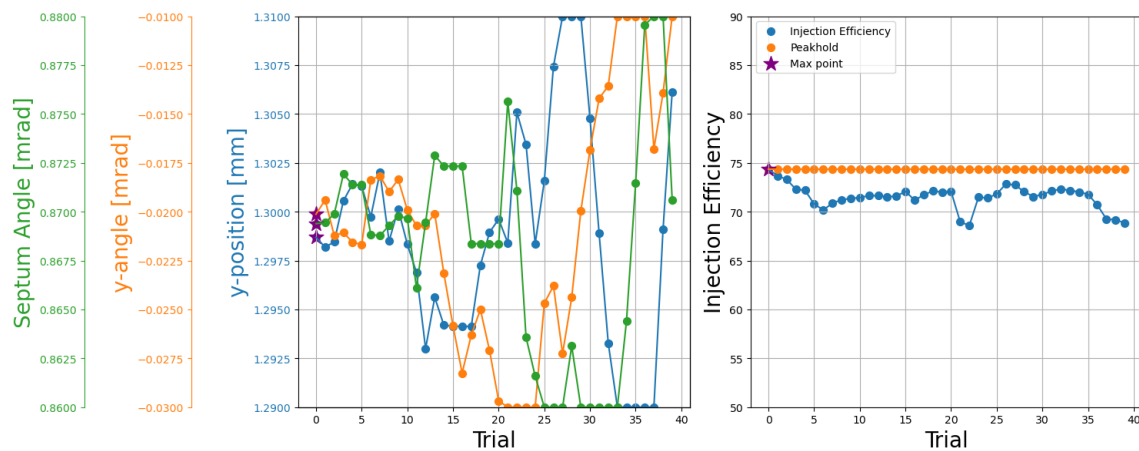
ノイズ標準偏差を 0.1% に設定した際に、最適化が過度に活用的になる理由について考察する。ノイズ標準偏差を 0.1% に設定すると、入射効率にわずか 0.1% の向上が観測された場合であっても、その結果が過度に重視される傾向が生じる。この結果、探索成分が抑制され、活用成分が優先される傾向が強まると考えられる。一方で、ノイズ標準偏差を設定しない場合は、入射効率に誤差がないものとして扱うため、0.1% 以上の向上があってもその影響を過度に評価せず、探索と活用のバランスが維持され则认为られる。

次に、獲得関数を EI、近傍探索度を 5 に固定し、ノイズ標準偏差を設定しない場合、1%、および 3% に設定した場合について最適化を実施した（番号 10–12）。これらの値を選定した理由は、ノイズ標準偏差を 0 から HER のふらつきの標準偏差（2.5%）程度まで段階的に増加させることで、最適化挙動の変化を理解するためである。ノイズ標準偏差を設定しない場合の結果については、5.2 節の図 5.4c で示されている。ノイズ標準偏差を 1% および 3% に設定した場合の結果を図 5.11 に示す。

ノイズ標準偏差を課さなかった場合の最適化挙動（図 5.4c）は、1% や 3% のノイズ標準偏差を課した場合と比較して、初期化条件終了直後に探索を重視した動きが確認される。一方、ノイズ標準偏差を 1% に設定した場合には、初期化直後の y-position および y-angle に対して 0.1% の場合と同様に活用的な最適化が見られ



(a) ノイズ標準偏差 0.1%

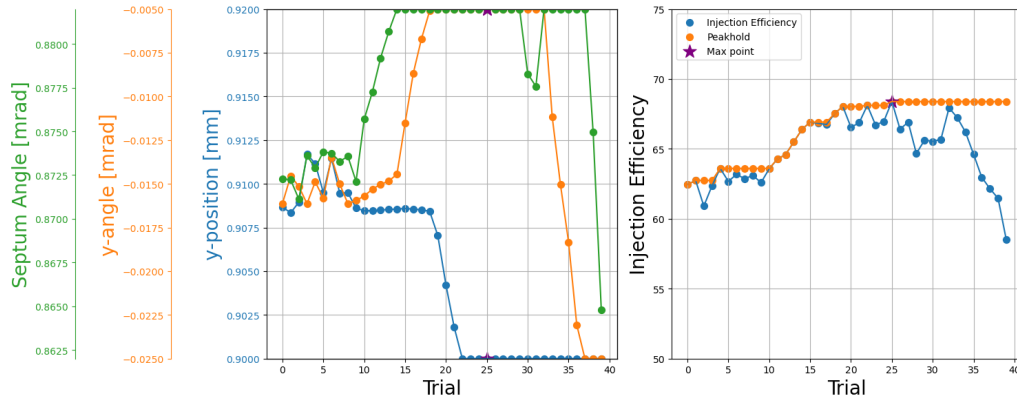


(b) ノイズ標準偏差 0

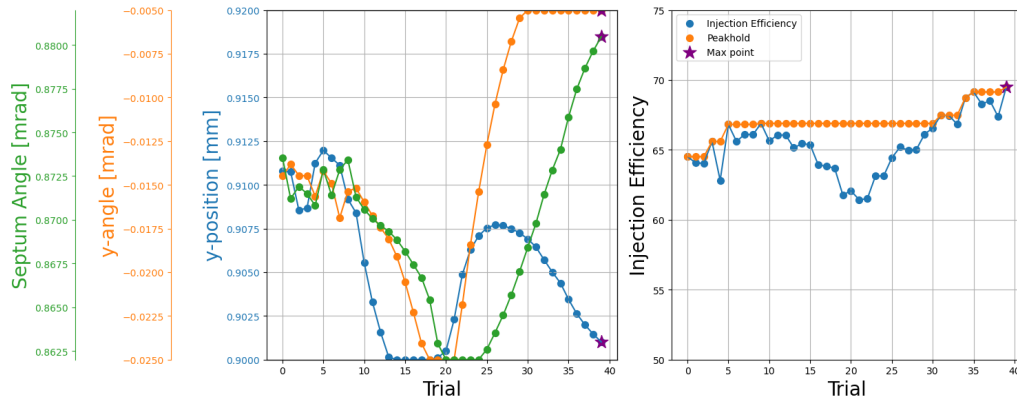
図 5.10: 獲得関数 EI、近傍探索度 5 に固定し、ノイズ標準偏差を 0 と 0.1% に設定した時の比較。(a) では最適化が活用的になりすぎたため、探索されない定義域内の領域が存在したまま 40 トライアルが終了している。(b) では探索的な成分が強く確認できる。

るものの、入射効率が向上すると確信した段階で探索的な挙動へ移行する傾向が見られた。3% に設定した場合では、初期化直後の活用的な傾向が 1% の場合よりも抑えられ、入射効率が向上した場合における探索的な動きも抑制されている。この結果は、ノイズ標準偏差を大きく設定するほど誤差に対するロバスト性が向上し、大幅な変化を妨げるという特性に基づいており、妥当な挙動と考えられる。最適化時間については全て 24 分と違いが見られなかった。

ノイズ標準偏差の適切な設定値について考察する。5.2 節で議論したように、12 月 14 日の番号 10-12 を測定した際のビーム条件では、y-angle と Septum Angle が上限値に達した際に入射効率が最大になることが確認されている。ノイズ標準偏差を 0、1%、3% のいずれに設定してもこれらの領域に到達していることから、小分け最適化においては、ノイズ標準偏差を 0.1% のような極端に小さい値に設定しない限り、十分な性能を発揮することが示唆される。一方で、制限つき最適化においては、入射効率のふらつきによる上振れを入射効率の向上と誤認し、急激に探索的な挙動を示すことは、安全性の観点から回避すべきである。この問題を防ぐ



(a) ノイズ標準偏差 1%



(b) ノイズ標準偏差 3%

図 5.11: 獲得関数 EI、近傍探索度 5 に固定し、ノイズ標準偏差を 1% と 3% に設定した時の比較。(a) では、初期化直後の y-position および y-angle に対して 0.1% の場合と同様に活用的な最適化が見られるものの、入射効率が向上すると確信した段階で探索的な挙動へ移行する傾向が見られた。(b) では初期化直後の活動的な傾向が 1% の場合よりも抑えられ、入射効率が向上した場合における探索的な動きも抑制されている。

ためには、ノイズ標準偏差を入射効率のふらつきの標準偏差に相当する 2-3% 程度に設定することが適切であると考えられる。

■獲得関数 本項では、獲得関数が探索と活用のバランスに与える影響について検討する。近傍探索度を 5、ノイズ標準偏差を 0 に固定した状態で、獲得関数を EI に設定した場合（番号 5）と UCB ($\beta = 2$) に設定した場合（番号 9）を比較する。EI の結果については 5.10b で既に示されている。UCB ($\beta = 2$) のプロットを図 5.10 に示す。この結果から、UCB ($\beta = 2$) は EI に比べて定義域内で探索する領域が狭く、活動的な特性を持つことが示された。この傾向は、先行研究である KEK Linac での陽電子ビームの電荷収量最大化スタディ（付録 B）における実験結果とも一致している。また、最適化時間については、EI と UCB ($\beta = 2$) の間で差は見られず、どちらも 24 分であった。

獲得関数の最適化条件に対する適合性について考察する。KEK Linac におけるビーム軌道調整の先行研究では、アボートのリスクがないため定義域を幅広く設定することが可能であり、その結果、定義域内に局所的

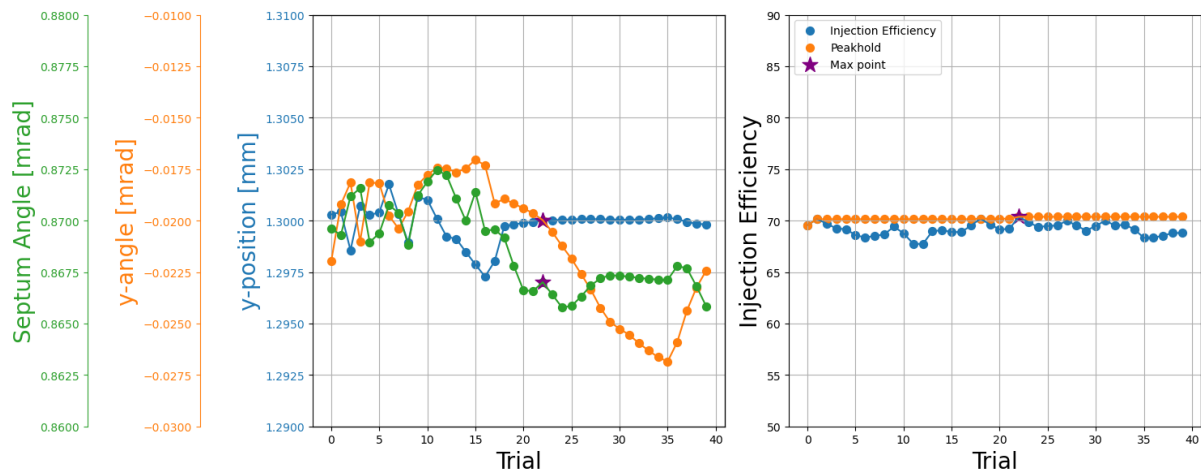


図 5.12: 近傍探索度 5、ノイズ標準偏差 0 の UCB($\beta = 2$)。図 5.10b と比較して、活用的な成分を重視した最適化が行われていることを確認できる。

な最大値が存在するが多かった。この条件下では、獲得関数として EI よりも活用的な探索を行う UCB ($\beta = 2$) の方が、最適化速度および性能の両面で優れていたことが確認されている。一方で、小分け最適化では定義域が狭いため、定義域内に局所的な最大値が存在するケースよりも、定義域外に局所的な最大値が存在するケースが多く見られる。その結果、定義域の端が最も高い値を示す傾向がある。このような条件下では、定義域内を活用的に探索する UCB ($\beta = 2$) よりも、定義域全体を広範囲に探索する EI の方が適合性が高いと考えられる。さらに、ビームロスが多い領域を回避しながら最適化を進める必要がある制限つき最適化では、安全性を優先し、探索に慎重なアプローチが求められる。この観点から、制限つき最適化においては EI よりも活用的な性質を持つ UCB ($\beta = 2$) が適合性が高いと考えられる。

■近傍探索度 本項では、近傍探索度が最適化速度に与える影響について検討する。はじめに、獲得関数を UCB ($\beta = 2$)、ノイズ標準偏差を 0 に固定した状態で、近傍探索度を 1 に設定した場合（番号 8）と 5 に設定した場合（番号 9）を比較する。近傍探索度 1 の結果を図 5.13 に示す。近傍探索度 5 の結果については、図 5.12 で既に表示されている。これらの 2 つのプロットを比較すると、近傍探索度を 1 に設定した場合では、5 に設定した場合と比較してトライアル間のパラメータの変化量が増加していることが確認できる。また、表 5.8 に示されているように、近傍探索度を 1 に設定した際の最適化時間は 26 分、5 に設定した際は 24 分であり、近傍探索度を高く設定することで最適化時間が 2 分短縮されていることが確認された。

各トライアルにおけるパラメータの変化量を可視化するため、パラメータの差分の絶対値を縦軸に取ったプロットを図 5.14 に示す。このプロットから、近傍探索度を 1 に設定した場合の方が、5 に設定した場合と比較してパラメータの変化量が多いことが確認できる。また、3 つのパラメータに対して、1 秒あたりに動かすステップ幅は表 4.5 に示されている通り 0.0015 mm または mrad であることから、近傍探索度 1 の場合、パラメータの移動に要する時間が長くなり、それが蓄積して 40 トライアル全体で最適化時間に約 2 分の差が生じたと考えられる。

次に、ノイズ標準偏差を 1、獲得関数を EI に固定した状態で、近傍探索度を 0 に設定した場合（番号 13）

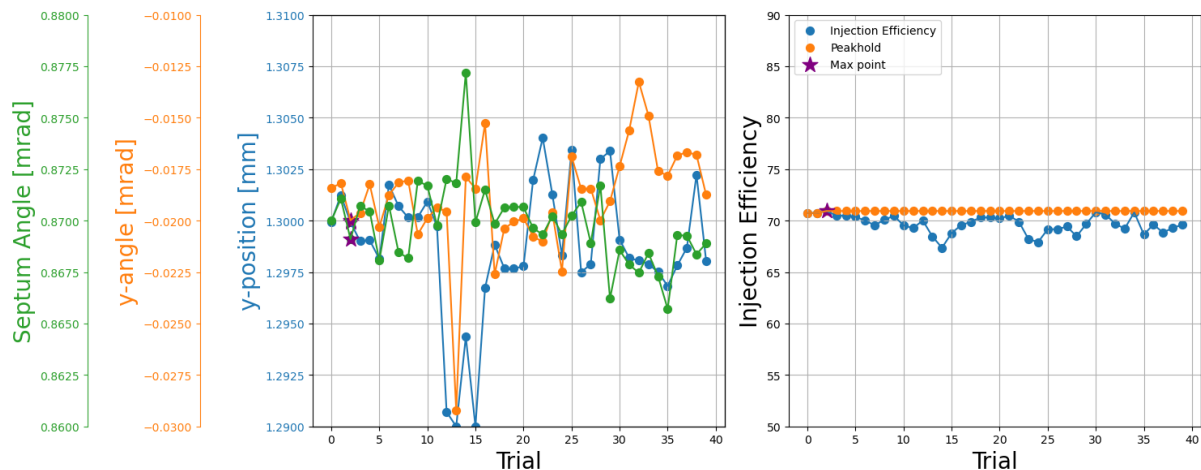


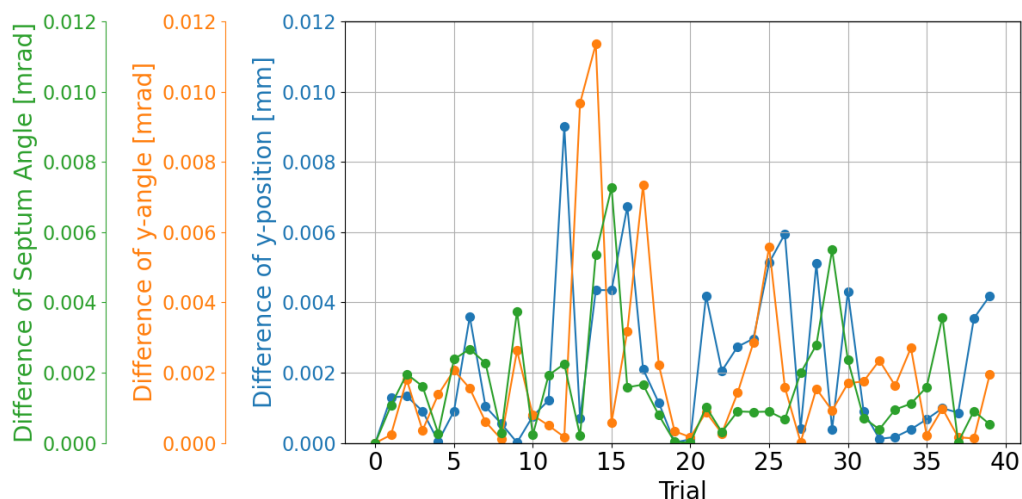
図 5.13: 近傍探索度 1、ノイズ標準偏差 0 の UCB ($\beta = 2$)。図 5.12 と比較すると、各トライアル間におけるパラメータの変化量が増加していることが確認できる。

と 5 に設定した場合 (番号 14) を比較する。これらの結果を図 5.15 に示す。近傍探索度を 0 に設定した場合 (図 5.15a) では、入射効率に有意な改善が見られたものの、初期化後に定義域の端を繰り返し探索する挙動が確認された。この挙動により定義域内部の探索が十分に行われず、正常な最適化が達成されていないことが見て取れる。一方で、近傍探索度を 5 に設定した場合 (図 5.15b) では、正常な最適化が確認され、入射効率にも有意な改善が見られた。また、表 5.8 に示されているように、近傍探索度を 0 に設定した際の最適化時間は 29 分、5 に設定した際は 24 分であり、近傍探索度を高く設定することで最適化時間が 5 分短縮されたことが確認された。

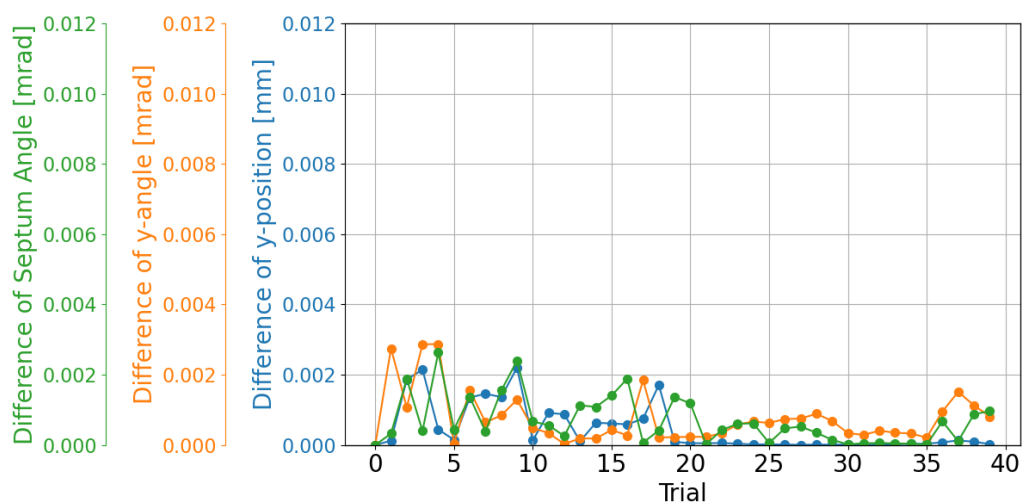
図 5.15a において確認された、定義域の端を繰り返し探索する挙動について考察する。前項で述べたように、EI は UCB と比較して探索的な傾向が強く、加えて、定義域の端を高く評価する傾向もある。さらに、ノイズ標準偏差を 1% に設定したことで、入射効率が 1% と比較して大幅に改善した場合、獲得関数の評価が過大となり探索的な挙動が強調されたと考えられる。この結果、探索が定義域の端付近で偏り、定義域の壁を繰り返し行き来するような挙動が引き起こされたと考えられる。

図 5.15 に示したパラメータの差分の絶対値を縦軸に取ったプロットを図 5.16 に示す。これらのプロットを比較すると、近傍探索機能の効果が顕著に表れていることが分かる。図 5.16a では、パラメータの変化量が大きく、定義域の端から端まで移動するトライアルが複数確認される。このような大きな移動には 13 秒を要するため、最適化全体の時間が延び、その結果として最適化時間に 5 分の差が生じている。一方で、図 5.16b では近傍探索機能の効果によりパラメータの変動が抑制されていることが確認できる。

最適化における適切な近傍探索度について考察する。近傍探索度は、最適化性能と最適化時間のトレードオフを調整する役割を持つ。近傍探索度を 5 に設定した場合について、図 5.15b では最大のパラメータ変化量が 0.003 mrad であり、これは約 2 秒で到達可能な距離である。また、図 5.16b に示す最大のパラメータ変化量である 0.065 mrad に関しても、約 5 秒で到達可能である。これらの移動時間は、30 回の入射効率測定に要する約 30 秒という時間に比べて支配的ではないため、近傍探索度を 5 以上に強める必要性は低いと考えられる。さらに、近傍探索度を 5 に設定した際に最適化性能が低下する事例は確認されていないため、小分け最適



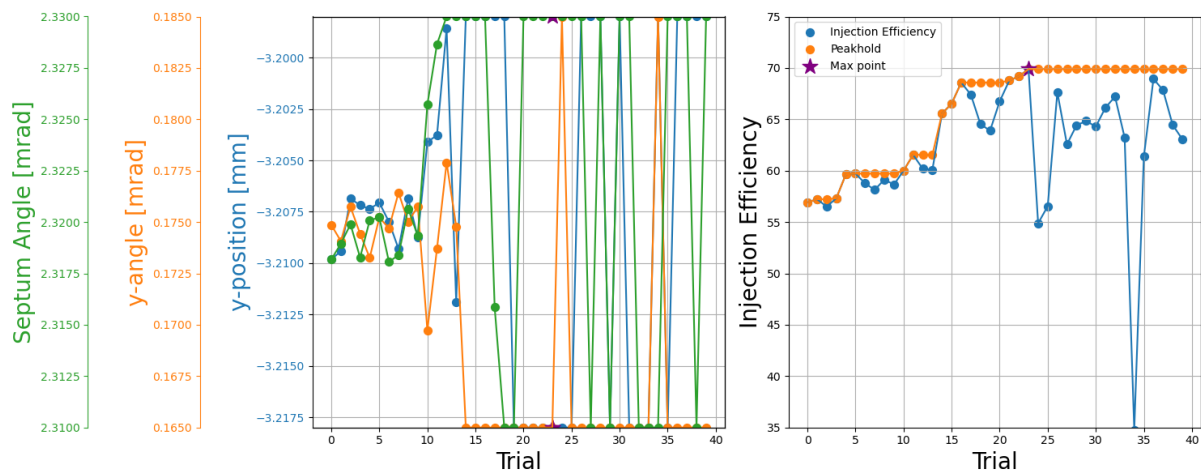
(a) 近傍探索度 1



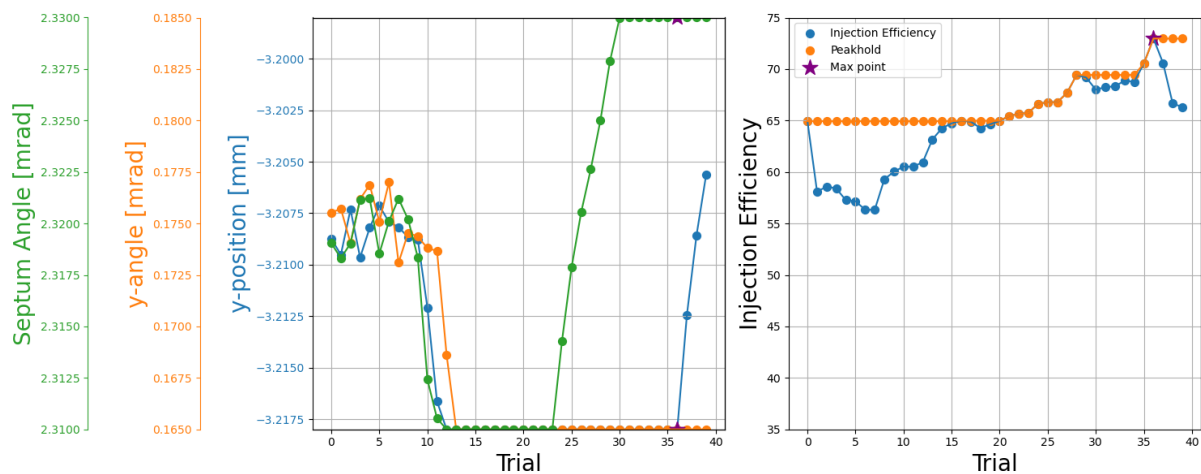
(b) 近傍探索度 5

図 5.14: ノイズ標準偏差 0、獲得関数を UCB ($\beta = 2$) に固定し、近傍探索度を 1 および 5 に設定した場合のトライアルごとのパラメータ変化量の絶対値。(a) は近傍探索度 1、(b) は近傍探索度 5 の結果であり、近傍探索度 1 の場合の方が、5 の場合と比較してパラメータの変化量大きいことが確認できる。

化においては近傍探索度 5 が適切な設定であると考えられる。制限付き最適化においても、近傍探索度 5 は有効であると考えられるが、パラメータ変動に伴うビームロスリスクをより慎重に管理する必要がある場合には、さらに高い近傍探索度を設定することも考えられる。

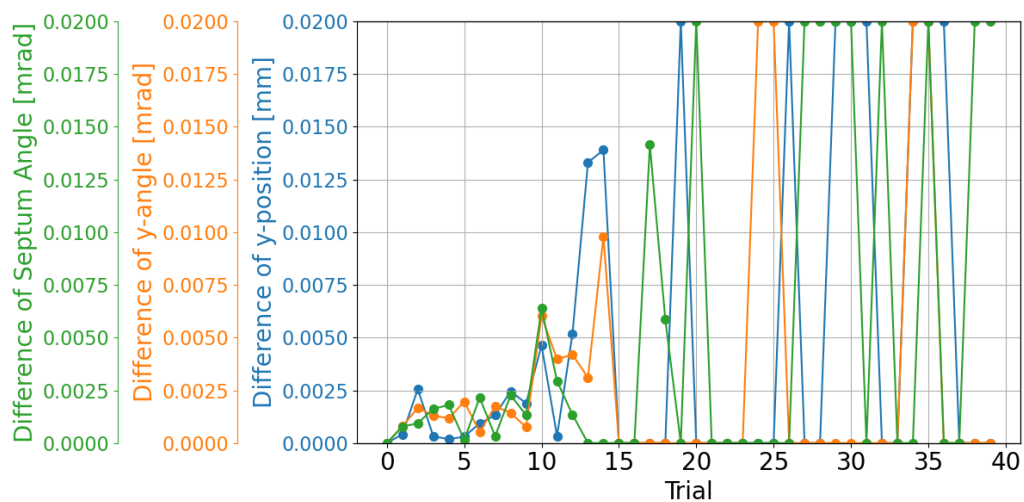


(a) 近傍探索度 0

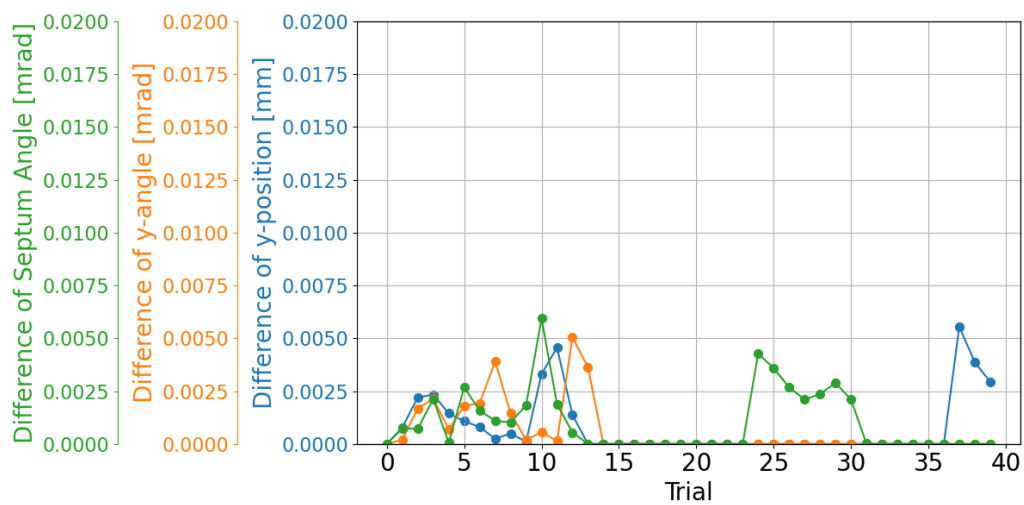


(b) 近傍探索度 5

図 5.15: ノイズ標準偏差 1、獲得関数を EI に固定し、近傍探索度を 0 および 5 に設定した場合の比較。(a) では入射効率の有意な改善は見られるが、定義域の端を繰り返し探索する挙動が確認され、適切な最適化が行われていない。(b) では正常に最適化が行われ、入射効率の有意な改善が見られる。



(a) 近傍探索度 0



(b) 近傍探索度 5

図 5.16: ノイズ標準偏差を 1、獲得関数を EI に固定し、近傍探索度を 0 および 5 に設定した場合のトライアルごとのパラメータ変化量の絶対値。(a) では、定義域の端から端まで大きく変化するトライアルが散見される。(b) ではパラメータの変動が抑えられていることが確認できる。

5.4 運用実績

2024年11月14日から12月25日の期間に、LERでは21ラン、HERでは19ラン、合計40ランが運用された。その結果、日ごとの入射調整運用における入射効率の向上をまとめたものが表5.9である。機械学習を活用したツールがメインリングの定常的な運転に適用されたことは前例がなく、本研究における重要な成果の1つである。以下に備考について説明する。まず、番号2および4以外のLERでの最適化では、2バンチ調整用の入射効率を使用し、番号2および4のLERと全てのHERでの最適化では1バンチ目の入射効率を最適化している。また、番号3以外の最適化はトップアップ入射を用いて実施した。番号3の最適化では、途中でLERの機器故障により入射が滞り、電流が減少しながらの最適化となった。この結果、ビームサイズが小さくなり、コリメーターによるビーム損失が減少したことで入射効率が向上した。したがって、トップアップ入射時の最適化と比較して入射効率の向上量が高めに見積もられる点に留意が必要である。さらに、番号12以外のデータはビーム衝突が行われている物理運転中に取得されている。一方、番号12ではHERのセプトマグネットの充電トリガーが誤動作した影響により、ビーム衝突が行われない片リング運転中に最適化を実施した。

表 5.9: 入射調整ツール運用における日ごとの最適化結果

番号	測定日	リングの種類	最適化回数	調整前/後入射効率	入射効率向上量	備考
1	11月14日	HER	3	59.4% / 70.2%	10.8%	
2	11月15日	LER	2	86.3% / 90.8%	4.5%	1バンチ目の入射効率
3	11月17日	HER	2	20.7% / 26.8%	6.1%	LERの電流が減少中
4	12月05日	LER	3	65.8% / 98.2%	32.4%	運用例1 1バンチ目の入射効率
5	12月06日	LER	2	71.9% / 76.9%	5.0%	
6	12月07日	HER	9	68.2% / 71.1%	2.9%	パラメータ比較スタディ
7	12月11日	LER	3	61.7% / 64.4%	2.7%	
8	12月13日	LER	5	77.5% / 82.0%	4.5%	再現性の確認スタディ
9	12月14日	HER	5	58.4% / 68.5%	10.1%	パラメータ比較スタディ 再現性の確認スタディ
10	12月14日	LER	2	69.0% / 77.8%	8.8%	
11	12月17日	LER	1	59.8% / 69.1%	9.3%	
12	12月25日	LER	3	54.5% / 68.4%	13.9%	運用例2 パラメータ比較スタディ 片リング運転中

はじめに、本ツールの最適化性能について議論する。第4章において、ベイズ最適化による入射効率の向上が統計的に有意である条件を、LERでは6.96%、HERでは7.50%以上の向上と定義した。表5.9に示された結果から、LERでは8回中4回（番号4, 10, 11, 12）、HERでは4回中2回（番号1, 9）の最適化が統計的に有意であることが確認された。この結果は、3つの重要なパラメータのみを操作し、図4.4に示された狭い定義域内での調整にもかかわらず、試行の半数で入射効率を有意に向上させた点で、重要な成果であると言える。

次に、パラメータの重要度について議論する。入射調整ツールの運用における各日のパラメータ重要度を示したものが表 5.10 である。3 つのパラメータの中で最も重要度が高いものを太字で示している。この表から、日によって入射効率の向上に寄与するパラメータが異なることがわかる。特に、入射効率が有意に向上した 6 日間では、3 つのパラメータそれぞれが最大の重要度を示すケースが確認された。また、LER と HER の両リングにおいて y-angle および Septum Angle が y-position に比べて入射効率向上に寄与する回数が多いことが読み取れる。さらに、最も入射効率の改善量が低かった日（番号 7）では、3 つのパラメータの重要度がほぼ均等（1/3 程度）であり、いずれのパラメータも入射効率の改善に大きく寄与していないことがわかる。

表 5.10: 入射調整ツール運用における日ごとのパラメータ重要度

番号	測定日	リングの種類	入射効率向上量	y-position	y-angle	Septum Angle
1	11 月 14 日	HER	10.8%	0.12	0.15	0.73
2	11 月 15 日	LER	4.5%	0.12	0.60	0.28
3	11 月 17 日	HER	6.1%	0.33	0.41	0.27
4	12 月 05 日	LER	32.4%	0.62	0.23	0.15
5	12 月 06 日	LER	5.0%	0.13	0.31	0.56
6	12 月 07 日	HER	2.9%	0.22	0.41	0.37
7	12 月 11 日	LER	2.7%	0.31	0.37	0.32
8	12 月 13 日	LER	4.5%	0.19	0.54	0.27
9	12 月 14 日	HER	10.1%	0.05	0.16	0.79
10	12 月 14 日	LER	8.8%	0.22	0.18	0.60
11	12 月 17 日	LER	9.3%	0.12	0.43	0.45
12	12 月 25 日	LER	13.9%	0.10	0.72	0.18

次に、LER と HER における入射効率が最大となった調整パラメータを、それぞれ表 5.11 および表 5.12 に示す。なお、HER の番号 1 における Septum Angle の値が番号 3 以降と大きく異なっているのは、2024 年 11 月 16 日に Septum Angle のオフセットが変更されたためである。

表 5.11: LER における日ごとの入射効率が最大となった調整パラメータ

番号	測定日	リングの種類	入射効率向上量	y-position [mm]	y-angle [mrad]	Septum Angle [mrad]
2	11 月 15 日	LER	4.5%	-3.219	0.143	2.454
4	12 月 05 日	LER	32.4%	-3.260	0.180	2.300
5	12 月 06 日	LER	5.0%	-3.174	0.138	2.300
7	12 月 11 日	LER	2.7%	-3.179	0.141	2.275
8	12 月 13 日	LER	4.5%	-3.170	0.134	2.210
10	12 月 14 日	LER	8.8%	-3.178	0.140	2.252
11	12 月 17 日	LER	9.3%	-3.275	0.169	2.241
12	12 月 25 日	LER	13.9%	-3.218	0.165	2.330

表 5.12: HER における日ごとの入射効率が最大となった調整パラメータ

番号	測定日	リングの種類	入射効率向上量	y-position [mm]	y-angle [mrad]	Septum Angle [mrad]
1	11 月 14 日	HER	10.8%	1.253	-0.015	0.272
3	11 月 17 日	HER	6.1%	1.219	-0.019	0.929
6	12 月 07 日	HER	2.9%	1.300	-0.020	0.869
9	12 月 14 日	HER	10.1%	0.900	-0.005	0.882

LER における入射効率が最大となった調整パラメータに対し、3 次元プロットおよび共分散行列を作成した結果を図 5.17 に示す。図 5.17a では、y-position, y-angle, Septum Angle の変化幅がそれぞれ 0.1 mm, 0.04 mrad, 0.25 mrad であることが確認された。これらの変化幅を各パラメータの定義域 (0.01 mm, 0.01 mrad, 0.01 mrad) と比較すると、y-angle の変動幅が最も小さい一方で、Septum Angle の変動幅が最も大きいことが明らかとなった。図 5.17b に示す共分散行列からは、y-position と y-angle の間に強い負の相関が存在し、y-position と Septum Angle の間に弱い負の相関があることが分かった。また、y-angle と Septum Angle の間には相関がほとんど見られないことが示された。

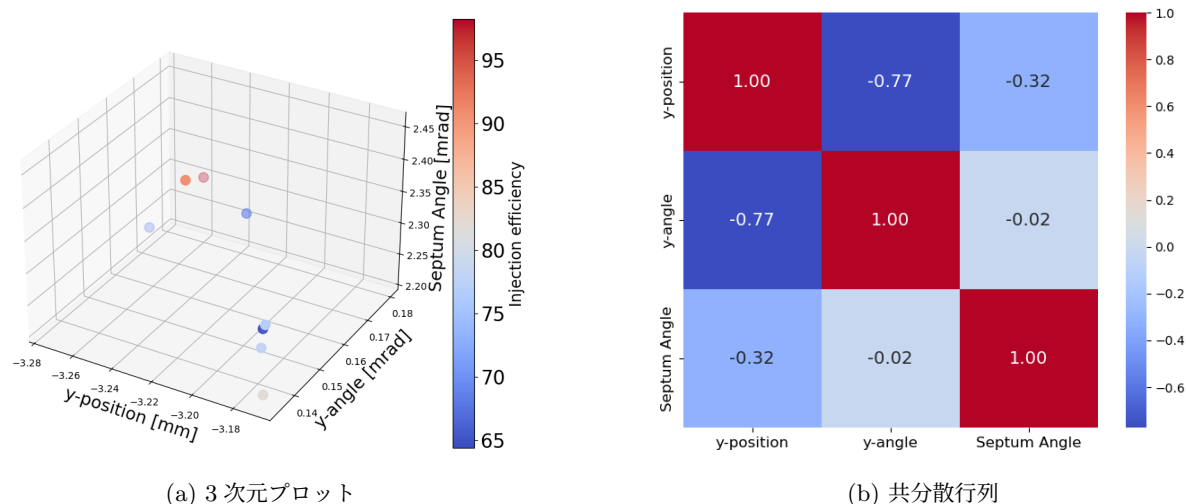


図 5.17: LER における入射効率が最大となった調整パラメータに対する 3 次元プロット (a) と共分散行列 (b)。(a) では各パラメータの定義域と比較して、y-angle の変動幅が最も小さい一方で、Septum Angle の変動幅が最も大きいことがわかる。(b) では y-position と y-angle の間に強い負の相関が存在し、y-position と Septum Angle の間に弱い負の相関があり、y-angle と Septum Angle の間には相関がほとんど見られないことがわかる。

3 つの調整パラメータに対する 1 ヶ月間の長期的な入射効率の変化を表した図 3.5 において共分散行列を作成した結果を図 5.18 に示す。この図から、y-position と y-angle の間に強い負の相関が存在し、y-position および y-angle と Septum Angle の間にはほとんど相関が見られないことが明らかとなった。y-position と y-angle の間に強い負の相関が見られる点は、図 4.2 に示されるパラメータの定義を考慮すると妥当である。一方で、垂直方向のパラメータ (y-position, y-angle) と水平方向のパラメータ (Septum Angle) は独立していると考えられるため、これらの間に相関がないという結果も整合的である。図 5.18 と図 5.17b を比較する

と、両者は概ね一致しているものの、図 5.17b では y-position と Septum Angle の間に弱い負の相関が見られる点が異なる。この相違の理由として、オペレータによる手動調整と機械学習を用いた調整の性質の違いが影響している可能性が考えられる。ただし、図 5.17b は 8 点のデータに基づいて生成された共分散行列であり、少数統計の結果である。そのため、この議論を深めるには、より多くの機械学習による最適化を行い、統計的信頼性を高める必要がある。

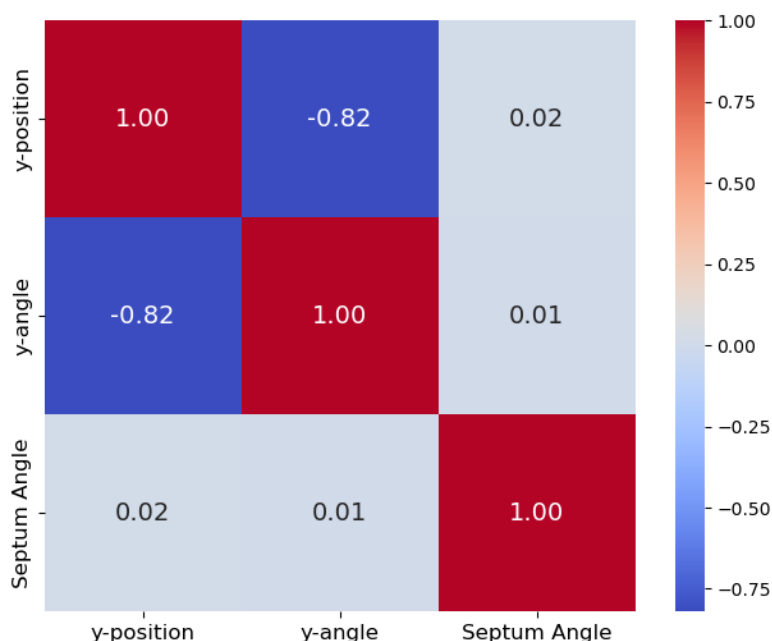


図 5.18: 2024 年 11 月 16 日から 12 月 16 日までの y-position, y-angle, Septum Angle の共分散行列。y-position と y-angle の間に強い負の相関が存在し、y-position, y-angle と Septum Angle の間には相関がほとんど見られないことがわかる。

調整ツールの運用を通じて、最適化性能をさらに向上させるために必要だと感じた機能について論じる。本研究では、入射効率のふらつきが原因でベイズ最適化の最適化性能が低下しており、この課題を解決するためにはふらつきを補正する機能の導入が求められる。図 5.19 にふらつき補正機能の概略図を示す。この機能では、パラメータ空間上に基準点を設定し、最適化開始前にその基準点の入射効率を測定する。その後、数トライアルに 1 回、基準点での入射効率を測定し、最適化開始前に測定した基準点の測定値をそのトライアルでの基準点の測定値で割った値を入射効率に乗じることで、ふらつきの影響を補正する仕組みである。このアプローチにより、ベイズ最適化に輸入される評価値の精度が向上し、最適化性能の改善が期待される。ただし、基準点を参照する頻度を増加させることは、最適化全体の実行時間を延長する要因となるため、参照頻度を適切に設定する必要がある。

最後に、本ツールのビームアポートに対する安全性について議論する。入射由来のビームバックグラウンドが原因でビームがアポートされる時は、Belle II 測定器付近に設置されたシンチレータを用いたロスモニターまたはダイヤモンドセンサーを用いたロスモニター^{*3}がアポートトリガーを出すことが多いが、今回の 40 回

^{*3} それぞれ Belle II CLAWS と VXD ダイヤモンドセンサー

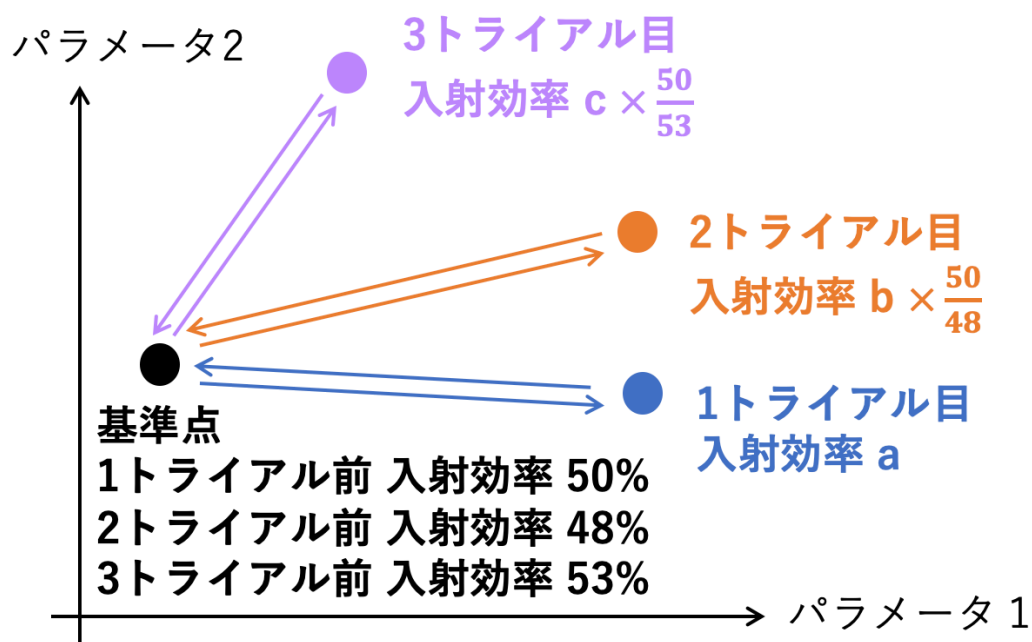


図 5.19: ふらつき補正機能の概略図。この図は 2 次元のパラメータ空間を例としており、数トライアルに 1 回、基準点での入射効率を測定し、最適化開始前に測定した基準点の測定値をそのトライアルでの基準点の測定値で割った値を入射効率に乘じることで、ふらつきの影響を補正する仕組みを示している。図中では、各トライアルに対して補正を適用する過程が描かれている。

のランでは、この 2 つのロスモニターがアボートトリガーを出すことは 1 度もなかった。この結果から、小分け最適化手法は、ビームアボートに対して高い安全性を有する最適化手法として有効であると考ええる。

第 6 章

結論と展望

Belle II 実験では新物理探索のためには大量の統計が必要であり、SuperKEKB 加速器ではルミノシティの向上が要求されている。瞬間ルミノシティはビームの蓄積電流に比例し、単位時間あたりの蓄積電流の増加量はメインリングへの入射効率に比例する。したがって、高い入射効率を維持することがルミノシティ向上に直結する。しかし、運転中には入射効率が数時間で低下することが多く、これが蓄積電流を妨げる要因となっている。この課題を解決するため、機械学習手法の一種であるベイズ最適化を用いて入射調整ツールを開発した。このツールの目的は入射効率をリアルタイムで最適化し、高い効率を継続的に維持することである。本ツールには以下の 3 点の性能が求められる。1 点目は最適化性能を備え、入射効率の向上が見込めない場合でも効率を悪化させない信頼性を有すること。2 点目はビームアポートを引き起こさない高い安全性を担保していること。3 点目はオペレーターが容易に操作可能な利便性を有することである。

本ツールを LER で 8 日間、HER で 4 日間運用した結果、LER では 4 日間、HER では 2 日間において、入射効率が測定値のふらつきを上回る統計的に有意な向上を示した。また、入射効率が向上しなかった場合でも効率が悪化することは一度も観測されなかった。これにより、1 点目の性能要件を満たしていることが確認された。さらに、40 回のランの全期間を通じて入射調整に起因するビームアポートが一度も発生しなかったことから、2 点目の性能要件も満たしていることが確認された。加えて、オペレーターが簡単に操作できるユーザーインターフェースを構築し、データ取得において筆者の引率なしで職員やオペレーターが本ツールを操作した例もあることから、3 点目の性能要件も満たしていることが確認された。機械学習を用いた自動調整ツールがメインリングの定常運転に適用されたことは、これまでに前例がなく、本研究の重要な意義の一つである。また、衝突型加速器における入射調整に機械学習を導入した事例としても、本研究が初めての試みであり、その先駆的意義は大きい。

効果的であった運用例として、1 バンチ目の入射効率を 66 % から 98 % に向上させた結果、両リングの蓄積電流を最適化前の 1.3 倍まで増加させることに成功した。また、LER における最大蓄積電流である 1700 mA の達成に寄与した。SHAP を用いたパラメータ重要度の解析では、日ごとに入射効率に寄与する調整パラメータが異なることを定量的に示した。入射調整ツールを用いたスタディでは、パラメータ重要度の高い入射調整パラメータに対しては再現性が確認された。さらに、獲得関数の比較では、小分け最適化において EI が UCB ($\beta = 2$) よりも適していることが確認された。また、近傍探索度を導入することで、最適化時間を最大で 5 分短縮することに成功した。

本ツールの最適化性能をさらに向上させるためには、小分け最適化における定義域幅、step by step 機能のステップ幅、近傍探索度、ノイズ標準偏差、獲得関数の選択といった要素について、さらなる詳細な検討が必要である。また、現在は調整パラメータが3つに限定されているが、inj. phase やビーム輸送路におけるマグネット調整パラメータを追加することで、最適化性能が一層向上する可能性がある。加えて、最適化性能を飛躍的に向上させるためには制限付き最適化の運用が重要である。小分け最適化では、現在のパラメータセットの近傍に存在する局所的な最大値を探索することが主な役割であるが、制限付き最適化ではより広い定義域を探索することが可能である。このアプローチにより、従来の手動調整では見つけられなかった新たな最適パラメータセットを発見できる可能性が大きく高まる。このような制限付き最適化の導入は、最適化性能をさらに向上させるための重要なステップである。

最後に、本ツールの将来的な展望について論じる。本ツールの小分け最適化は、ビームアバートを引き起こさない高い安全性を備えており、物理データ取得中でも最適化を実施できるという利点を有している。また、入射効率が向上するパラメータセットを見つけれなかった場合でも、調整前のパラメータセットに戻すことで入射効率の低下を防ぐことが可能である。したがって、本ツールの運用にはリスクが低く、初期の選択肢として気軽に実施することができる。この特性を活用することで、入射効率が低下した場合、まず本ツールを使用して最適化を試行し、その後必要に応じて手動調整に移行するという運用が可能となる。この運用方法が実現すれば、従来の属人的な調整に依存することなく、一定の品質と安定性を備えた入射調整が可能になる。また、ベイズ最適化は汎用性の高い手法であり、入射調整に限らず、ビーム光学調整やコリメータの調整など、現在自動化されていない他の加速器調整にも適用可能である。そのため、本ツールの開発は加速器全体の調整効率をさらに向上させる基盤となりうる。

謝辞

人と環境に恵まれた、何ごとにも代えがたい、本当に貴重な2年間であったと、心からそう感じます。指導教官である後田裕教授には、研究内容に関するご指導のみならず、研究者としての取り組むべき姿勢について多くのことを教えていただきました。日々の研究において特にお世話になり、加速器研究の面白さと奥深さを教えていただいた三塚岳准教授にも心より感謝申し上げます。

次いで、ゼミでの建設的なご指摘や、発表スライドおよび修士論文の添削で多大なご助力をいただいた宇野健太助教、古賀太一朗助教、中沢遊助教にも深く御礼申し上げます。さらに、加速器の入射部分や入射調整に関して多くのことを教えてくださった KEK 加速器の飯田直子准教授、ベイズ最適化の実装において有益なアドバイスをいただいた夏井拓也准教授、Spring-8 主幹研究員の岩井瑛人氏にも心より感謝いたします。また、加速器ゼミを開催して基礎を教えてくださった小関研 OB の安居孝晃助教、浅見高史氏にも大変お世話になりました。このほかにも、東大後田研および総研大 B ファクトリー研のスタッフの皆様、Belle II 秘書の皆様、KEK 加速器施設の関係者の皆様、加速器オペレーターの皆様をはじめ、私の研究を支えてくださった多くの方々に深く感謝申し上げます。また、奨学金を通じて金銭的な支援をしていただいた服部奨学財団の皆様にも、心より御礼申し上げます。

さらに、日々の研究生生活をともに過ごし、公私にわたってお世話になりました、後田研の能丸くん、王世傑くん、王子涵さん、朴さん、許さん、総研大 B ファクトリー研の劉さん、曾さん、小関研の有馬くん、樋口研の男谷さん、島崎さん、そして Belle II に携わる学生の皆さんをはじめ、素粒子実験や加速器実験の中で一緒にさせていただいた学生の皆さん、あなた方のおかげでより一層楽しく研究ができました。ありがとうございました。

そして最後に、これまで私を育ててくれ、常に研究を見守り支えてくれた両親に、心から感謝いたします。本当にありがとうございました。

付録 A

付録

A.1 Belle II 測定器

Belle II 測定器の全体図を図 A.1 に記す。Belle II 測定器は高エネルギー加速器研究機構の筑波実験棟の地下 4 階に存在し、高さ約 8 m、厚み約 8 m、重さ約 1,400 トンの測定器である。

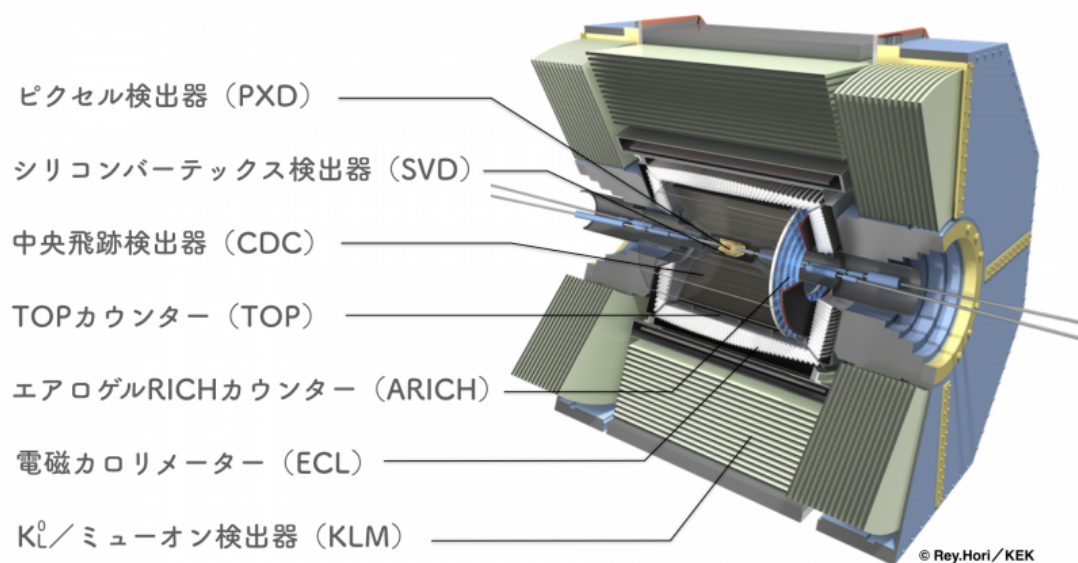


図 A.1: Belle II 測定器の全体図 [31]

各検出器について個別に解説する [32][5]。

・崩壊点検出器

Pixel detector (PXD)

PXD は崩壊点の最内層に存在する 2 層からなる半導体検出器であり、高ルミノシティ化に伴う高ヒット占有率に対応するため、ピクセル型の検出器となっている。このため、後述するストリップセンサー型検出器の SVD と比較して、単位面積あたりのチャンネル数が多くヒット占有率を大幅に低下させることができる。ま

た、センサーは Depleted Field-Effect Transistor (DEPFET) 技術に基づいて製造されており、測定器の厚みを 50 μm まで薄くすることに成功している。

Silicon vertex detector (SVD)

SVD は 2 層ある PXD の外層に位置しており、4 層から構成される。SVD を全てピクセル型にしてしまうと膨大なチャンネル数になってしまうため、高い崩壊点検出能力を維持しながら、チャンネル数を抑えたストリップセンサー型を採用している。SVD は後述する CDC と比較して高精度な位置分解能を有し、低運動量粒子や長寿命粒子の軌跡と崩壊点を再構成する重要な役割を担う。

・スペクトロメーター

Central drift chamber (CDC)

CDC は 14,336 本のワイヤーが 56 層に配置され、ビーム進行方向に対して並行なアクシアルワイヤーと、進行方向に対して傾いたステレオワイヤーから構成されている。これにより、飛跡の 3 次元再構成が可能となる。CDC の役割は 3 点あり、荷電粒子の軌跡再構成と運動量測定、低運動量粒子の粒子識別、トリガー信号の生成である。

Electromagnetic calorimeter (ECL)

ECL は、タリウム付加ヨウ化セシウム (CsI(Tl)) をシンチレーション結晶としたエネルギー全吸収型のカロリメータである。B 中間子から崩壊する粒子の約 1/3 は中性粒子であるが、PXD、SVD、CDC では荷電粒子のみしか検出ができない。そのため、ECL は中性粒子の位置やエネルギーの情報を得る重要な役割を担っている。また、電子の精密測定や、トリガー生成の役割も担っている。

・粒子識別装置

Time of propagation counter (TOP)

TOP はバレル部の粒子識別装置であり、主に $K^\pm$ 粒子と $\pi^\pm$ 粒子を識別する。石英反射体内部において粒子が通過した際に発生するチェレンコフ光を全反射させることで光検出器まで伝播させ、その伝播時間によって粒子を識別する。光検出器にはマイクロチャンネルプレート光電子増倍管 (MCP-PMT) を使用している。

Aerogel ring imaging Cherenkov detector (ARICH)

ARICH はエンドキャップ部の粒子識別装置であり、TOP と同じく主に $K^\pm$ 粒子と $\pi^\pm$ 粒子を識別する。この検出器では石英と同じ組成式 (SiO_2) でありながら、非常に低密度の個体であるシリカエアロゲルを輻射体として用い、密度を変化させながらレイヤー状に重ねる。そうすると $K^\pm$ 粒子と $\pi^\pm$ 粒子が通過した時のチェレンコフ光による円の半径の違いを精度よく区別することができ、単レイヤーと比べて精度のよい粒子識別が可能になる。チェレンコフ光を検出する光検出器には HAPD と呼ばれる、光電子増倍管とアバランシェフォトダイオードを組み合わせた光子検出器を用いている。

$K_L - \mu$ detector (KLM)

KLM は K_L 粒子と μ 粒子を識別する検出器であり、厚さ 4.7 cm の鉄板と厚さ 4.4 cm の抵抗板チェンバーまたはプラスチックシンチレータが交互に配置されている。 K_L 粒子は強い相互作用により鉄板の中でハドロンシャワーを出す μ 粒子は強い相互作用をしないため、より深く鉄板を通過する。この違いにより 2 つの粒子を識別している。

A.2 メインリング

この節では本編で説明しなかったメインリングのコンポーネントについて解説する。

■**コリメーター** SuperKEKB メインリングでは、入射由来のバックグラウンドや、安定周回領域から外れた蓄積ビーム由来のバックグラウンドが発生している。このバックグラウンドは Belle II 測定器におけるデータ取得の質を下げるだけでなく、測定器に放射性由来の損害を与える。また、QCS のクエンチを引き起こして加速器運転を中断させてしまうこともある。そのため、コリメーターで物理的にビームを削ることにより、上記の問題を軽減している。図 A.2a にコリメーターの概略図を、図 A.2b に SuperKEKB メインリングでのコリメーターの配置を記す。

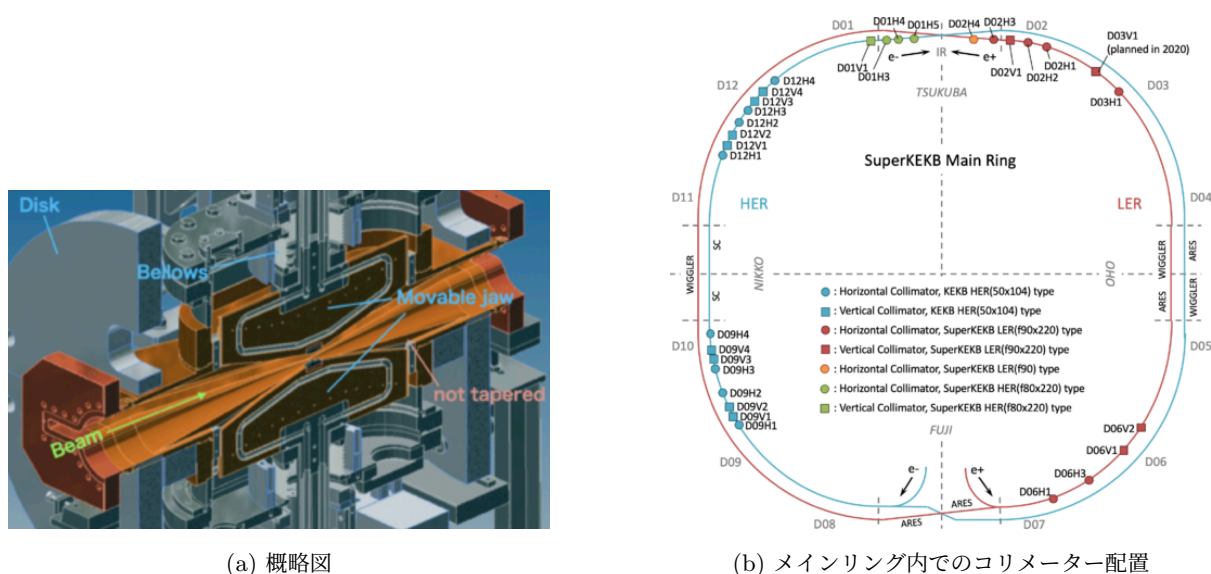


図 A.2: コリメーターの概略図とメインリング内でのコリメーター配置 [33]

コリメーターのデメリットとしては、ビームが通過できる空間を遮断することで力学口径が小さくなってしまい、蓄積ビームの寿命を減少させてしまうことが挙げられる。入射由来のバックグラウンドを抑えながら蓄積ビームの寿命を減らさない適度な口径にコリメーターを設定することが理想である。

また、2024 年 12 月現在では、突発ビームロス (SBL) と呼ばれる、従来のビームロスよりも早いタイムスケールでビーム不安定性が増大し、アボートトリガー信号によってビームがダンプする前に機器に損傷を与える現象が発生しており、発生メカニズムは解明されていない。そのため、機器保護の観点から早期の原因解明が期待されている。図 A.3 は SBL によって損傷したコリメーターの写真である。

■**バケットとフィルパターン** バンチを入射することができる単位をバケットと呼び、SuperKEKB メインリングでは 1 周あたりに 5120 バケットと配置すると取り決めている。理論上は全てのバケットにバンチを入射することが可能だが、前方のバンチが作った電磁場によって後方のバンチが乱されてしまう Wake field 現象が生じてしまうために、実際には 2~3 バケット毎にバンチを格納する。また、後述するアボートシステムに

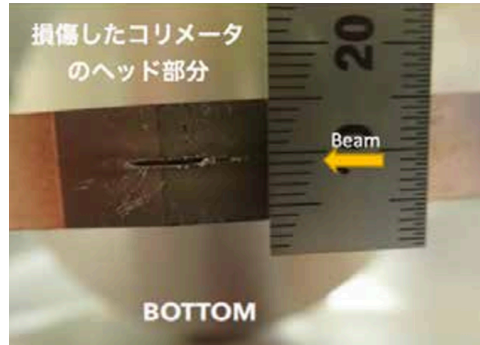


図 A.3: SBL によって損傷したコリメーター [34]

よってビームアボートが促された際に、ビームをダンプすることができるように、150 バケット程度の空白が開いたアボートギャップと呼ばれる配置が 2 箇所存在する。このような、バンチを入射するバケットやアボートギャップなどの配置をフィルパターンと呼ぶ。図 A.4 に 2024 年の加速器運転における LER フィルパターンと各バンチの蓄積電流量の一例を記す。2~3 バケット毎にバンチが格納され、またアボートギャップが 2 つあることも確認できる。

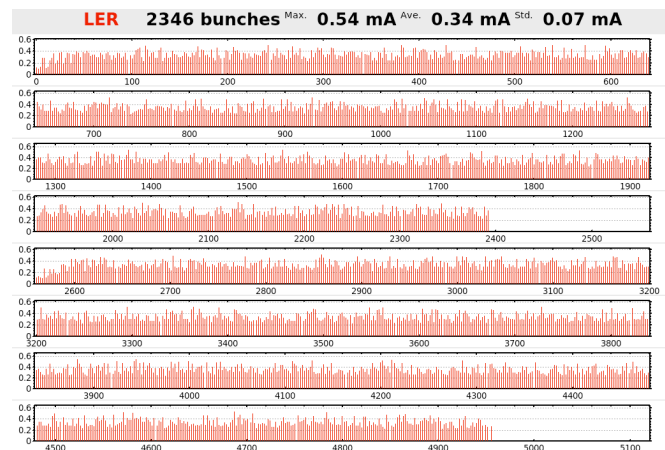


図 A.4: LER でのフィルパターンと各バンチの電流量

A.3 電流の寿命

この節では電流の寿命について説明する [35]。ビーム同士の衝突が起きていない場合には、電流が減衰する要因として以下の 2 点が挙げられる。1 点目はビーム内の粒子がビームダクト中のガスと散乱衝突するビームガス散乱効果であり、この効果由来の寿命 $\tau_{\text{beam-gas}}$ は式 (1.1) で表される。ここで、 $\bar{P}_{\text{eff}}$ は有効的なビームダクト内の圧力、 B はビームガス散乱の効率に関するスケール因子である。

$$\frac{I}{\tau_{\text{beam-gas}}} = B \times I \bar{P}_{\text{eff}}, \quad (1.1)$$

ここで $\bar{P}_{\text{eff}}$ について、電流に比例する項と比例しない項を合わせて式 (1.2) のように表せる。

$$\bar{P}_{\text{eff}} = P_0 + I P_1 \quad (1.2)$$

2 点目は、散乱運動をしながら確率的に運動量を変化させて周回しているバンチ内の粒子の一部が安定的に周回できる運動量領域から外れてビームがロスしてしまう効果であり、このことをタウシェック効果と呼ぶ。この効果由来の寿命 τ_{Touschek} は式 (1.3) で表される。ここで、 n_b はバンチ数、 $\sigma_x, \sigma_y, \sigma_z$ はそれぞれの軸方向のビームサイズ、 T はタウシェック散乱の効率に関するスケール因子である。

$$\frac{I}{\tau_{\text{Touschek}}} = T \times \frac{I^2}{n_b \sigma_x \sigma_y \sigma_z}, \quad (1.3)$$

ビーム同士の衝突がない場合のビーム寿命は以下のように表せる。ここで τ_{other} はビームガス散乱とタウシェック散乱以外に由来する寿命を表す。

$$\frac{I}{\tau(I)} = \frac{I}{\tau_{\text{beam-gas}}(I)} + \frac{I}{\tau_{\text{Touschek}}(I)} + \frac{I}{\tau_{\text{other}}(I)} \quad (1.4)$$

よってビーム同士の衝突がない場合の電流の寿命による減衰は、式 (1.5) のように表せる。

$$\frac{dI_{\text{decay}}(t)}{dt} = -\frac{I(t)}{\tau(I)} \quad (1.5)$$

最終的に、単位時間あたりの蓄積電流の変化量は以下の式 (1.6) で表される。

$$\frac{dI_{\text{total}}(t)}{dt} = \frac{dI_{\text{inj}}(t)}{dt} + \frac{dI_{\text{decay}}(t)}{dt} \quad (1.6)$$

ここで、式 (1.5) で表される減衰項は電流の 2 乗に比例して増大することがわかる。そのため、大電流を蓄積しようとする、式 (3.1) で要求される入射電流も多く電流が必要になる。また、実際の運転ではビーム同士が衝突することによる相互作用や、コリメーターのコンフィギュレーションによっても変化することには注意が必要である。

A.4 ベイズ最適化の理論

この節では、ベイズ最適化に必要な数学的知識からはじめ、ベイズ最適化の理解に必要な要諦を論じる。

A.4.1 多変量ガウス分布

ベイズ最適化で用いられるガウス過程回帰法では多変量ガウス分布を扱うため、本節ではその基本的な性質を論じる。まず、変数 x が平均 μ 、分散 σ^2 の 1 次元ガウス分布に従うときの確率密度分布は以下で表される。

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1.7)$$

このガウス分布を多次元化へと一般化した分布を考える。具体的には、ベクトル $\mathbf{z}$ が D 次元の独立した標準正規分布に従うときの同時確率密度分布 $\mathcal{N}(\mathbf{z}|0, \mathbf{I}_D)$ を以下のように定義し、

$$\begin{aligned} \mathcal{N}(\mathbf{z}|0, \mathbf{I}_D) &\equiv \prod_{i=1}^D \mathcal{N}(z_i|0, 1) \\ &= \frac{1}{(2\pi)^{D/2}} \exp\left(-\frac{1}{2} \sum_{d=1}^D z_d^2\right) \\ &= \frac{1}{(2\pi)^{D/2}} \exp\left(-\frac{1}{2} \mathbf{z}^T \mathbf{z}\right) \end{aligned} \quad (1.8)$$

$\mathbf{z}$ に対して以下のような任意の線形変換

$$\mathbf{x} = A\mathbf{z} + \mu \quad (1.9)$$

を施し変換された $\mathbf{x}$ が満たす確率密度分布を多変量ガウス分布 $\mathcal{N}(\mathbf{x})$ と定義する。 $\mathcal{N}(\mathbf{x})$ を通常のガウス分布の自然な拡張として考えると、全空間で積分したときに 1 であることを要請するので、規格化定数を C と置くと

$$C \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x}) dx_1 \dots dx_D = 1 \quad (1.10)$$

を満たす。また、左辺を式変形すると、

$$\begin{aligned} C \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathcal{N}(\mathbf{x}) dx_1 \dots dx_D &= C \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathcal{N}(z|0, \mathbf{I}_D) |A| dz_1 \dots dz_D \\ &= C|A| \end{aligned} \quad (1.11)$$

となり、 $C = 1/|A|$ であることがわかる。ここで、 $\mathbf{x} \rightarrow \mathbf{z}$ でのヤコビアンは線形変換式 (1.9) の行列式となることを用いた。これにより $\mathcal{N}(\mathbf{x})$ は

$$\begin{aligned}
\mathcal{N}(\mathbf{x}) &= \frac{1}{|A|} \mathcal{N}(A^{-1}(\mathbf{x} - \boldsymbol{\mu}) | 0, \mathbf{I}_D) \\
&= \frac{1}{(2\pi)^{D/2} |A|} \exp \left(-\frac{1}{2} (\mathbf{A}^{-1}(\mathbf{x} - \boldsymbol{\mu}))^T \mathbf{A}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \\
&= \frac{1}{(2\pi)^{D/2} |A|} \exp \left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T (\mathbf{A} \mathbf{A}^T)^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right)
\end{aligned} \tag{1.12}$$

となる。次に $\mathbf{x} = [x_1, \dots, x_D]^T$ の要素 x_i における分散 σ_{ii} 、及び要素 x_i, x_j の共分散 σ_{ij} からなる共分散行列 $\boldsymbol{\Sigma}$ について考えると、定義より

$$\begin{aligned}
\boldsymbol{\Sigma} &= \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1D} \\ \vdots & \ddots & \vdots \\ \sigma_{D1} & \cdots & \sigma_{DD} \end{bmatrix} \\
&= \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T] \\
&= \mathbb{E}[(A\mathbf{z})(A\mathbf{z})^T] \\
&= A \mathbb{E}[\mathbf{z}\mathbf{z}^T] A^T \\
&= A A^T
\end{aligned} \tag{1.13}$$

となる。ここで式 (1.9) と、 A の要素が定数であるため期待値の外に出せること、及び $\mathcal{N}(\mathbf{z} | 0, \mathbf{I}_D)$ であることを用いた。また、両辺に絶対値を取ると

$$|\boldsymbol{\Sigma}| = |A A^T| = |A| |A^T| = |A| |A| = |A|^2 \tag{1.14}$$

と表せるので、最終的に、平均ベクトル $\boldsymbol{\mu}$ 、共分散行列 $\boldsymbol{\Sigma}$ の多変量ガウス分布は以下で表される。

$$\mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2} \sqrt{|\boldsymbol{\Sigma}|}} \exp \left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T (\boldsymbol{\Sigma})^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \tag{1.15}$$

この分布は $\mathbf{x}$ の各成分の平均と分散の情報のみならず、共分散の情報も有している。つまり、各成分の正規分布と、その相関の強さについての情報を持っており、この性質が後述するガウス過程回帰法において応用されている。

A.4.2 線形回帰とガウス基底関数

この節では、ベイズ最適化に用いられるガウス過程回帰法の前段階として、ベイズ線形回帰モデルについて論じるが、その前段階として、まずは一般的な線形回帰について論じる。パラメータ $\mathbf{x}$ と具体系の不明なブラックボックス関数 $f(\mathbf{x})$ と測定誤差 ϵ によって y が観測されていると仮定する。

$$y = f(\mathbf{x}) + \epsilon \tag{1.16}$$

このとき、線形回帰によって $f(\mathbf{x})$ を予測することを考える。線形回帰では観測データの組 $(\mathbf{x}, y)$ が得られた場合に、各次元の重み $\mathbf{w}$ をパラメータとして用いて以下のような線形モデルを作成する。モデルにより回帰予測された値を $\hat{y}$ とすると、切片項が w_0 であることに注意して、

$$\hat{y} = w_0 + \sum_{i=1}^D w_i x_i = w_0 + \mathbf{w}^T \mathbf{x} \quad (1.17)$$

と表される。 $(\mathbf{x}, y)$ を N 回測定したとき、回帰予測された $\hat{y}$ をベクトル $\hat{\mathbf{y}}$ として並べると、以下のように行列表現でまとめることができる。 x_{ij} は i 回目の測定の j 番目のパラメータを表している。

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1D} \\ 1 & x_{21} & x_{22} & \cdots & x_{2D} \\ 1 & x_{31} & x_{32} & \cdots & x_{3D} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \cdots & x_{ND} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_D \end{bmatrix} \quad (1.18)$$

ここで、中央の各測定値が格納された行列は計画行列と呼ばれ、これを $\mathbf{X}$ で表すと

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{w} \quad (1.19)$$

となる。観測データから重みの値を適切に定める操作を学習と呼び、その標準的な方法としては最小二乗法が挙げられる。最小二乗法はモデルと実測値の差が一番小さくなる重み w_i の組を求める手続きであり、

$$E = \sum_{i=1}^D (y - \hat{y}) = \sum_{i=1}^D (y - \mathbf{w}^T \mathbf{x}) \quad (1.20)$$

で表される E が最小となる重み w_i 、つまり $\frac{\partial E}{\partial \mathbf{w}} = \mathbf{0}$ を満たす $\mathbf{w}$ を見つければよく、導出は省略するがこの解は最終的に、

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (1.21)$$

と求めることができる。

さて、ここまでは多次元空間に対して、超平面を用いて線形回帰する方法を考えてきたが、1次元において複雑な回帰を行う場合にもここまでの議論が流用できる。式 (1.17) において、 $\mathbf{x}$ を回帰に使用したい関数を格納したベクトル $\phi(x)$ とすると $\phi_0(x) = 1$ とすれば

$$\hat{y} = \mathbf{w}^T \phi(x) \quad (1.22)$$

と表せる。これは直線以外の複雑な回帰が可能であり、例えば $\phi_1(x) = x^2$ や $\phi_2(x) = \sin x$ など非線形な関数を用いて、より柔軟なモデリングが可能になる。式 (1.19) において計画行列を

$$\Phi = \begin{bmatrix} 1 & \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_D(x_1) \\ 1 & \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_D(x_2) \\ 1 & \phi_1(x_3) & \phi_2(x_3) & \cdots & \phi_D(x_3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \phi_1(x_N) & \phi_2(x_N) & \cdots & \phi_D(x_N) \end{bmatrix} \quad (1.23)$$

に置き換えた式

$$\hat{\mathbf{y}} = \Phi \mathbf{w} \quad (1.24)$$

の解は以下のように与えられる。

$$\mathbf{w} = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{y} \quad (1.25)$$

ここで問題になるのは、どのような $\phi(x)$ を用いればより柔軟な回帰ができるかであるが、その有望な候補としてガウス基底関数が挙げられる。ガウス基底関数は

$$\phi_i(x) = \exp\left(-\frac{(x - \mu_i)^2}{\sigma^2}\right) \quad (1.26)$$

で表される基底関数であり、 y の定義域が $[-a, a]$ であるとき、定義域内の μ_i に $\phi_i(x) = \exp\left(-\frac{(x - \mu_i)^2}{\sigma^2}\right)$ を配置し、各 $\phi_i(x)$ を重み w_i で重ね合わせた線形基底関数モデルは

$$\hat{y} = \sum_{i=-a}^a w_i \exp\left(-\frac{(x - \mu_i)^2}{\sigma^2}\right) \quad (1.27)$$

で表される。このモデルは正規分布を細かく重ね合わせたものであるため、重みを適切に最適化することで回帰関数として非常に高い表現力を持つことが直感的にも理解できる。実際、ガウス基底関数を用いた線形基底関数モデルは万能近似性という関数近似器として良い性質を持つことが理論的にも示されている。

次に、この議論を多次元に拡張することを考える。しかしながら、少し考えればこれはうまくいかないことがわかる。実際に式 (1.27) を多次元にしてみると、

$$\hat{y} = \sum_{i=-a_i}^{a_i} \sum_{j=-a_j}^{a_j} \cdots \sum_{k=-a_k}^{a_k} w_{ij\cdots k} \exp\left(-\frac{(x - \mu_{ij\cdots k})^2}{\sigma^2}\right) \quad (1.28)$$

となり、次元の増加に伴って最適化する重みの数が指数関数的に増えてしまい、最適化にかかる計算コストが現実的な値ではなくなってしまう。例えば $a=10$ の場合について考えると、1次元では重みは 21 個であるが、2次元になっただけで 441 個になり、4次元では約 200000 個の重みを最適化しなくてはならず、これは現実的ではない。このように、パラメータの次元が増えるほど最適化する重みが指数関数的に増える現象のことを一般に“次元の呪い”と呼び、多次元回帰問題においてはこの問題にどう対処するかが鍵になる。

A.4.3 ベイズ推定とベイズ線形回帰モデル

次節で論じるガウス過程回帰法は、前節の”次元の呪い”をうまく回避しながら柔軟な回帰を行う手法である。これまでの議論では、重み $\mathbf{w}$ を決定論的に、つまりはある定数として扱っていた。しかしこれを多変量ガウス分布として確率的に扱うことにより、次元の呪いを緩和することが可能になる。また、これまでは重みを定数として扱っていたので $\hat{y}$ も定数であったが、重みが確率分布となったとき $\hat{y}$ も確率分布として扱う必要が出てくる。本節は確率分布に対して回帰問題を扱う際に必要な知識であるベイズ推定と、ベイズ推定を使って線形回帰をするベイズ線形回帰モデルについても論じる。

まずはじめに、ベイズの定理について説明する。ある 2 つの確率密度関数 X, Y について、それらを同時に満たす確率分布を $p(X, Y)$ が既知であるとする。ここから、 X 単独の確率密度関数 $p(X)$ を得るには、全ての Y に対して $p(X, Y)$ を足し上げれば良いので、

$$p(X) = \int p(X, Y) dY \quad (1.29)$$

のように定義できる。次に、 X が既知であるときの Y の条件付き確率分布を $p(Y|X)$ で表す。このとき、以下のような等式が成り立つ。

$$p(X, Y) = p(Y|X)p(X) = p(X|Y)p(Y) \quad (1.30)$$

これは、 X と Y が同時に起こる確率は、 X が起こる確率と X が起こったときに Y が起こる確率の積と等しいことを表しており、 X と Y を入れ替えても同様の議論が成り立つ。この式を式変形したものがベイズの定理であり、以下で表される。

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)} \quad (1.31)$$

右辺の $p(Y|X)$ を事後分布と呼び、左辺の $p(X|Y)$ を尤度関数、 $p(Y)$ を事前分布、 $p(X)$ を正規化定数と呼ぶ。名称の意味については、ベイズ推定を理解するとわかる。ベイズ推定とは事前分布と尤度関数から事後分布を算出する手法であり、具体的には以下の手続きで表される。

1. 尤度関数のモデリングをする。パラメータ θ に対して Y が生成される過程を $p(Y|\theta)$ で仮定する。
2. 事前分布のモデリングをする。パラメータ θ を満たす確率を確率分布 $p(Y)$ で仮定する。
3. 事後分布を計算する。ベイズの定理を用いることによって、観測結果が Y であったときに、パラメータが θ であるときの確率分布 $p(\theta|Y)$ を求める。

これをこれまで扱ってきた回帰問題に適用し、事後分布を実際に求める。パラメータを重み $\mathbf{w}$ とし、

1. 尤度関数 $p(\mathbf{y}|\Phi, \mathbf{w})$ をモデル化する。 $\mathbf{y}$ の要素である y は式 (1.16) で表されている。この式の中の測定

誤差 ϵ について、平均 $\mathbf{0}$ 、共分散行列 $\sigma^2 \mathbf{I}$ の正規分布に従うと仮定する。つまり複数回の測定誤差をベクトルとして重ねた ϵ に対して $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ であると仮定する。また、ブラックボックス関数 $f(x)$ は式 (1.24) でモデル化されているため、最終的に Φ と $\mathbf{w}$ のもとで y が実現される確率 $p(y|\Phi, \mathbf{w})$ は以下のようにモデル化できる。

$$p(y|\Phi, \mathbf{w}) = \mathcal{N}(\Phi \mathbf{w}, \sigma^2 \mathbf{I}) \quad (1.32)$$

2. 事前分布 $p(\mathbf{w})$ をモデル化する。これを決定論的ではなく平均 $\mathbf{0}$ 、分散 $\lambda^2 \mathbf{I}$ に従う多変量ガウス分布と仮定して以下のようにモデル化する。

$$p(\mathbf{w}) = \mathcal{N}(\mathbf{0}, \lambda^2 \mathbf{I}) \quad (1.33)$$

3. 事後分布はベイズの定理と上記の仮定より以下のように計算される。

$$\begin{aligned} p(\mathbf{w}|\Phi, \mathbf{y}) &= \frac{p(\mathbf{y}|\Phi, \mathbf{w})p(\mathbf{w})}{\int p(\mathbf{y}|\Phi, \mathbf{w})p(\mathbf{w}) d\mathbf{w}} \\ &\propto \mathcal{N}(\Phi \mathbf{w}, \sigma^2 \mathbf{I}) \times \mathcal{N}(\mathbf{0}, \lambda^2 \mathbf{I}) \\ &\propto \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y} - \Phi \mathbf{w})^T(\mathbf{y} - \Phi \mathbf{w}) - \frac{1}{2\lambda^2}\mathbf{w}^T \mathbf{w}\right) \\ &\propto \exp\left(-\frac{1}{2}\left(\mathbf{w} - \left(\frac{\Phi^T \Phi}{\sigma^2} + \frac{I}{\lambda^2}\right)^{-1} \frac{\Phi^T \mathbf{y}}{\sigma^2}\right)^T \left(\frac{\Phi^T \Phi}{\sigma^2} + \frac{I}{\lambda^2}\right)^{-1} \left(\mathbf{w} - \left(\frac{\Phi^T \Phi}{\sigma^2} + \frac{I}{\lambda^2}\right)^{-1} \frac{\Phi^T \mathbf{y}}{\sigma^2}\right)\right) \end{aligned} \quad (1.34)$$

$$(1.35)$$

これは平均 $\bar{\boldsymbol{\mu}} = \left(\frac{\Phi^T \Phi}{\sigma^2} + \frac{I}{\lambda^2}\right)^{-1} \frac{\Phi^T \mathbf{y}}{\sigma^2}$ 、共分散行列 $\bar{\boldsymbol{\Sigma}} = \left(\frac{\Phi^T \Phi}{\sigma^2} + \frac{I}{\lambda^2}\right)$ で表される多変量ガウス分布となっている。

ここで事後分布とは、現在の観測データから得られた情報 $(\Phi, \mathbf{y})$ を元に、重み $\mathbf{w}$ が実現される確率であるので、この重みの下で観測していない未知の入力 $\mathbf{x}^*$ が与えられたときに観測結果 y^* が実現する確率を表す分布 $p(y^*|\phi(\mathbf{x}^*), \mathbf{w})$ を全ての重み $\mathbf{w}$ について足し上げれば、観測データ $(\Phi, \mathbf{y})$ のみから任意の $\mathbf{x}^*$ についての任意の y^* の出力を表す予測分布 $p(y^*|\Phi, \mathbf{y}, \mathbf{x}^*)$ を計算することができる。実際に計算すると、

$$\begin{aligned} p(y^*|\Phi, \mathbf{y}, \mathbf{x}^*) &= \int p(y^*|\phi(\mathbf{x}^*), \mathbf{w})p(\mathbf{w}|\Phi, \mathbf{y}) d\mathbf{w} \\ &= \int \mathcal{N}(\mathbf{w}^T \mathbf{x}^*, \sigma^2 \mathbf{I}) \times \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}^{-1}) d\mathbf{w} \\ &= \mathcal{N}(\bar{\boldsymbol{\mu}}^T \phi(\mathbf{x}^*), \sigma^2 \mathbf{I} + \phi(\mathbf{x}^*)^T \bar{\boldsymbol{\Sigma}}^{-1} \phi(\mathbf{x}^*)) \end{aligned} \quad (1.36)$$

となる。これはまさに、測定点の数値を元に測定していない点の期待値と不確かさを回帰予測する確率分布となっている。式 (1.36) において特筆すべきことは、線形回帰パラメータである重み $\mathbf{w}$ が積分消去され、どのようなパラメータも仮定しないノンパラメトリックな分布となっていることである。これにより、ブラックボックス関数 $f(x)$ に対して、具体的な関数形をパラメータで特徴付けない非常に柔軟な回帰を行うことができる。(厳密には、式 (1.36) では確率分布の各点が正規分布に従うと言う仮定や、実際の計算では Φ の具体形

を仮定しなければいけないと言う特徴付けがある。しかし一般にはパラメータによって特徴付けされない分布をノンパラメトリックな分布と呼ぶ。)

また、式 (1.36) では重み $\mathbf{w}$ が積分消去されていることは”次元の呪い”を克服する上でも非常に重要である。ベイズ推定では事後分布を出す際に $\mathbf{w}$ に依存した式が出てしまっているが、事後分布を仲介せずに初めから予測分布を直接出すことができれば、”次元の呪い”を回避することができる。次節で扱うガウス過程回帰法がまさにその答えとなっている。

A.4.4 ガウス過程回帰法

ガウス過程とは例えると、与えられた入力次元に対して、それと同じ出力次元の多変量ガウス分布を吐き出す”箱”のようなものである。ガウス過程回帰法とは、観測点によって作られた”箱”に、未知の点を入力し、予測分布を出力させる回帰方法である。ここで重要なのは、この箱を回帰する際に、重み $\mathbf{w}$ に依存しない手続きができることであり、”次元の呪い”の影響を受けないことである。本節ではガウス過程の定義について述べ、その後ガウス過程回帰法の定義とその応用方法について述べる。

まず、ガウス過程の定義について述べる。関数 f の任意の各点が入力に依存した多変量ガウス分布で表されるときに f はガウス過程に従うと言う。より厳密な定義は以下の通りである。

任意 n 個のベクトル $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{X}$ に対して、その関数値ベクトル $\mathbf{f} = (f(\mathbf{x}_1), \dots, f(\mathbf{x}_n))$ が $\mu: \mathcal{X} \rightarrow \mathbb{R}$ を満たす要素で構成されたベクトル $\boldsymbol{\mu} = (\mu(\mathbf{x}_1), \dots, \mu(\mathbf{x}_n))$ を平均に持ち、 $k: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ を満たす要素 $K_{nn'} = k(\mathbf{x}_n, \mathbf{x}_{n'})$ で構成された行列 K を共分散行列とする多変量ガウス分布 $\mathcal{N}(\boldsymbol{\mu}, K)$ に従うとき、関数 $f: \mathcal{X} \rightarrow \mathbb{R}$ はガウス過程に従うと定義する。

また、以降は関数 f がガウス過程に従うことを以下のように表す。

$$f \sim \mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (1.37)$$

これまで扱ってきた式 (1.24) について、重み $\mathbf{w}$ が式 (1.33) のように正規分布で表されるとすると y がガウス過程に従うことを記す。はじめは簡単のため、誤差 ϵ がなく、正確に回帰される状態、つまり $y = \hat{y}$ を考える。(誤差がある場合は後に議論する。) 式 (1.24) の計画行列 Φ は観測点が既知のときには定数行列であり、重み $\mathbf{w}$ は正規分布で仮定される。よって y は正規分布を線形変換したものとなり多変量ガウス分布に従うことがわかる。この分布は具体的に計算することができ、まず平均について、

$$\boldsymbol{\mu} = \mathbb{E}(\mathbf{y}) = \mathbb{E}(\Phi \mathbf{w}) = \Phi \mathbb{E}(\mathbf{w}) = \mathbf{0} \quad (1.38)$$

次に共分散行列について、

$$\Sigma = \mathbb{E}(\mathbf{y}\mathbf{y}^T) - \mathbb{E}(\mathbf{y})\mathbb{E}(\mathbf{y}^T) = \mathbb{E}(\Phi \mathbf{w}(\Phi \mathbf{w})^T) - \mathbf{0} = \Phi \mathbb{E}(\mathbf{w}\mathbf{w}^T) \Phi^T = \lambda^2 \Phi \Phi^T \quad (1.39)$$

よって、 $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \lambda^2 \Phi \Phi^T)$ と表せる。次に、これがガウス過程に従っていることを示すためには、共分散

行列が入力ベクトルを用いて表されなければならない。そこで $\lambda^2 \Phi \Phi^T$ について考えると、

$$\Sigma = \lambda^2 \Phi \Phi^T = \lambda^2 \begin{bmatrix} \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_D(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \cdots & \phi_D(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \cdots & \phi_D(x_N) \end{bmatrix} \begin{bmatrix} \phi_0(x_1) & \phi_0(x_2) & \cdots & \phi_0(x_N) \\ \phi_1(x_1) & \phi_1(x_2) & \cdots & \phi_1(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_D(x_1) & \phi_D(x_2) & \cdots & \phi_D(x_N) \end{bmatrix} \quad (1.40)$$

ここで、 Σ の n 行 n' 列目の要素を $k_{nn'}$ とすると、 $\phi(\mathbf{x}) = (\phi_0(x), \phi_1(x), \dots, \phi_D(x))$ として、

$$k_{nn'} = \lambda^2 \phi(\mathbf{x}_n)^T \phi(\mathbf{x}_{n'}) \quad (1.41)$$

と表せる。ここで、共分散行列の要素が入力ベクトルによって表されているため、 y はガウス過程

$$y \sim \mathcal{GP}(\mathbf{0}, k_{nn'}(\mathbf{x}_n, \mathbf{x}_{n'})) \quad (1.42)$$

に従っていることが示された。重み $\mathbf{w}$ は式 (1.39) で共分散行列を計算するときの期待値計算で消去されており、最終的な式 (1.42) は重みに依存しない式となっていることがわかる。また、式 (1.41) は関数 $\phi(\mathbf{x}_n)$ と $\phi(\mathbf{x}'_n)$ の内積を表しており、測定点 $\mathbf{x}_n$ と $\mathbf{x}'_n$ の値が近ければ近いほど値が大きくなる。多変量ガウス分布において共分散が大きいということは、分布の間に相関が強いということであるから、出力される y_n と $y_{n'}$ も似た値を取りやすくなるというガウス過程の定性的な性質も見取れる。

ここから、ガウス過程を用いて回帰することを考える。式 (1.42) において、未知の点 $\mathbf{x}^*$ を増やしたときの y^* の値を考えると、これら全体がまたガウス分布に従うので、入力点からなる共分散行列を $\mathbf{K}$ とし、既知の点と未知の点からなる共分散ベクトルを $\mathbf{k}_* = (k(\mathbf{x}^*, \mathbf{x}_1), k(\mathbf{x}^*, \mathbf{x}_2), \dots, k(\mathbf{x}^*, \mathbf{x}_N))^T$ 、未知の点同士の分散を $k_{**} = k(\mathbf{x}^*, \mathbf{x}^*)$ として、

$$\begin{pmatrix} y \\ y^* \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{pmatrix} \mathbf{K} & \mathbf{k}_* \\ \mathbf{k}_*^T & k_{**} \end{pmatrix} \right) \quad (1.43)$$

と表される。ここで求めたいのは y が既知であるときに y^* が満たされる条件付き確率 $p(y^* | y, \mathbf{K}, \mathbf{x}^*)$ である。この具体形を求めるためには多変量ガウス分布の条件付き分布を与える以下の定理を用いる。

確率変数ベクトル $(\mathbf{a}, \mathbf{b})^T$ が以下のような多変量ガウス分布に従うとする。

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{pmatrix}, \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \right) \quad (1.44)$$

このとき、 $\mathbf{a}$ を既知だと見なした時の $\mathbf{b}$ の確率分布 $p(\mathbf{a}, \mathbf{b})$ は

$$\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_b + \Sigma_{ba} \Sigma_{aa}^{-1} (\mathbf{a} - \boldsymbol{\mu}_a) \quad (1.45)$$

$$\bar{\Sigma} = \Sigma_{bb} - \Sigma_{ba} \Sigma_{aa}^{-1} \Sigma_{ab} \quad (1.46)$$

で表される多変量ガウス分布 $\mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\Sigma})$ に従う。

証明は付録に記す。この定理を用いることで、ガウス過程の予測分布は以下のように表される。

$$p(y^*|\mathbf{y}, \mathbf{K}, \mathbf{x}^*) = \mathcal{N}(\mathbf{k}_*^T \mathbf{K}^{-1} \mathbf{y}, k_{**} - \mathbf{k}_*^T \mathbf{K}^{-1} \mathbf{k}_*) \quad (1.47)$$

次に誤差を考慮して議論する。式 (1.16) と同様に観測データ y は回帰値 $\mathbf{f} = (f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_N))$ に誤差 $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ が加えられたものと解釈すると、測定点の分散項に誤差が加わるので以下のように表せる。

$$\begin{pmatrix} \mathbf{y} \\ y^* \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{pmatrix} \mathbf{K} + \sigma^2 \mathbf{I} & \mathbf{k}_* \\ \mathbf{k}_*^T & k_{**} \end{pmatrix} \right) \quad (1.48)$$

この結果、式 (1.47) において $\mathbf{K} \rightarrow \mathbf{K} + \sigma^2 \mathbf{I}$ と置き換えれば良いので、最終的に誤差を含めたガウス過程の予測分布は以下のように表される。

$$p(y^*|\mathbf{y}, \mathbf{K}, \mathbf{x}^*) = \mathcal{N}(\mathbf{k}_*^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{y}, k_{**} - \mathbf{k}_*^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{k}_*) \quad (1.49)$$

まとめると、ガウス過程回帰は事前分布 $\mathcal{GP}(0, \mathbf{K})$ から事後分布の予測分布 $\mathcal{GP}(\mathbf{k}_*^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{y}, k_{**} - \mathbf{k}_*^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{k}_*)$ を作成する手続きである。

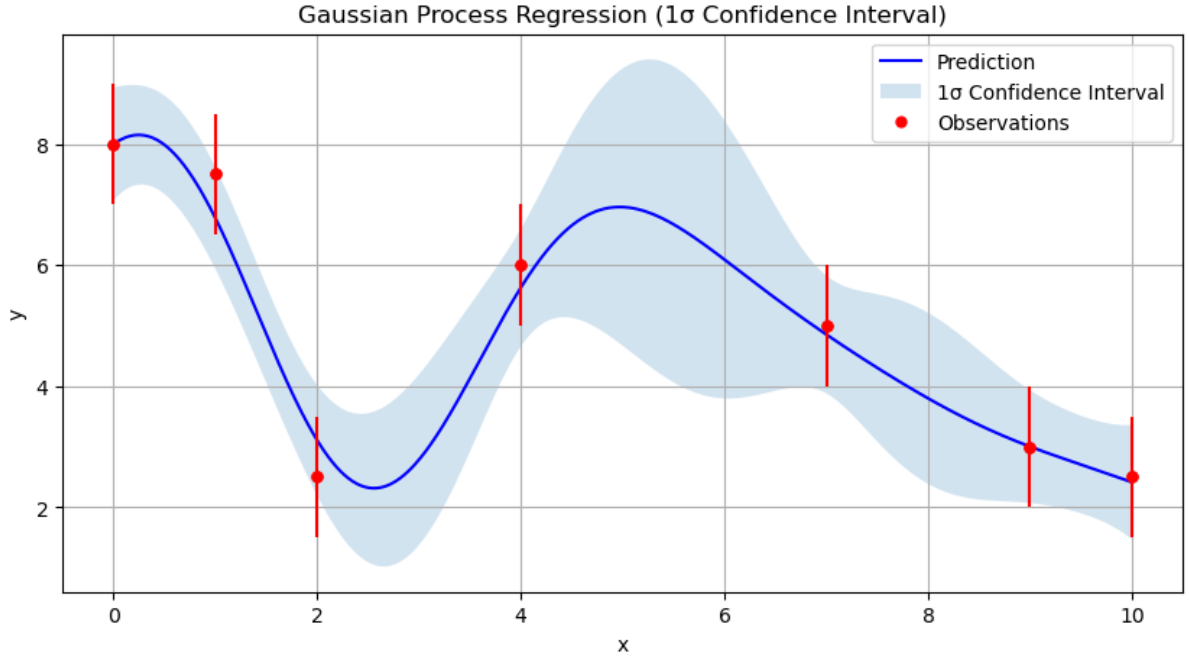


図 A.5: 1次元のガウス過程回帰の図。赤点が測定点であり、そこから平均 (青線) と分散 (水色の領域) を作成している。後述するカーネル関数は RBF カーネルを使用している。

A.4.5 カーネル関数

さて、これまでの議論では共分散行列 $\mathbf{K}$ について具体形を与えていなかった。ガウス過程を満たすためには共分散行列の要素が式 (1.41) を満たすことが必要である。ここで回帰問題を扱う際は $\phi(\mathbf{x})$ の具体形を決めることで回帰される関数に特徴付けをしていた。例えば、 $\phi(\mathbf{x}) = \mathbf{x}$ であれば線形回帰であり、式 (1.26) ではガウス基底回帰となる。ガウス基底回帰では、 $\phi(\mathbf{x})$ から $k(\mathbf{x}, \mathbf{x}')$ を計算するのではなく、 $k(\mathbf{x}, \mathbf{x}')$ を直接関数として定めてしまう。この関数のことをカーネル関数といい、カーネル関数を要素とする共分散行列をカーネル行列と呼ぶ。この手法の優れている点は、 $\phi(\mathbf{x})$ の具体形が煩雑になる場合や無限次元になって具体的に書き下せない場合にも共分散行列を定義することができることであり、計算量の節約にもなる。このように直接 $\phi(\mathbf{x})$ の内積計算を行わずカーネル関数で内積計算を行うことをカーネルトリックと呼ぶ。

カーネル関数の中で一番感覚的にわかりやすいものは線形カーネルであり、以下のように表される。

$$k(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{x}' \quad (1.50)$$

式 (1.50) を用いて作成されたカーネル行列は式 (1.19) の X と同じになるため、このカーネルを用いたガウス過程回帰の平均は線形回帰と一致する。(厳密にはガウス過程回帰では誤差を取り扱っているので、L2 正則化を施した線形回帰であるリッジ回帰と等しくなる。)

次に、式 (1.26) のガウス基底関数をカーネル関数として扱うことを考える。回帰問題においては、有限のガウス基底関数を配置し、それを重みでパラメタライズした。今回はガウス基底関数が空間上に連続的に分布していることを表すようなカーネルを表現したい。そのために、まずある x 軸上の $-H < x < H$ の範囲に、点 h/H を $2H^2 + 1$ 個の $h = -H^2, -H^2 + 1, \dots, -1, 0, 1, \dots, H^2 - 1, H^2$ 上に以下のガウス基底関数を配置したと考える。

$$\phi_h(x) = \tau \exp\left(-\frac{(x - \frac{h}{H})^2}{r^2}\right) \quad (1.51)$$

カーネル関数は定義より、

$$k(x, x') = \sum_{h=-H^2}^{H^2} \phi_h(x) \phi_h(x') \quad (1.52)$$

このグリッド上に配置したガウス基底関数を無限に細かくすることを考えると、 $H \rightarrow \infty$ として

$$\begin{aligned}
k(x, x') &= \lim_{H \rightarrow \infty} \sum_{h=-H^2}^{H^2} \phi_h(x) \phi_h(x') \\
&= \int_{-\infty}^{\infty} \tau^2 \exp\left(-\frac{(x-h)^2}{r^2}\right) \exp\left(-\frac{(x'-h)^2}{r^2}\right) dh \\
&= \tau^2 \exp\left(-\frac{1}{2r^2}(x-x')^2\right) \int_{-\infty}^{\infty} \exp\left(-\frac{2}{r^2}\left(h - \frac{x+x'}{2}\right)^2\right) dh \\
&= \tau^2 \sqrt{\frac{\pi r^2}{2}} \exp\left(-\frac{1}{2r^2}(x-x')^2\right)
\end{aligned} \tag{1.53}$$

$$= \tau^2 \sqrt{\frac{\pi r^2}{2}} \exp\left(-\frac{1}{2r^2}(x-x')^2\right) \tag{1.54}$$

となる。このカーネルは x と x' の距離にのみ依存しているので、 $\mathbf{x}$ が複数の観測点を格納したベクトルであるときにも成り立つ。よって $\theta_1 = \tau^2 \sqrt{\frac{\pi r^2}{2}}, \theta_2 = 2r^2$ とおくと最終的に、

$$k(\mathbf{x}, \mathbf{x}') = \theta_1 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{\theta_2}\right) \tag{1.55}$$

となる。これを RBF カーネルと呼び、後述するベイズ最適化においては非常にスタンダードなカーネル関数である。RBF カーネルは無限次元のガウス基底関数 ϕ を無限次元の重み $\mathbf{w}$ で重み付けした予測分布を表しており、これは前述のベイズ推定を用いた回帰手法では実現不可能である。また、RBF カーネルは非常に強力で汎用的なカーネル関数であり、後述するベイズ最適化のカーネルとして最もスタンダードなカーネル関数の一つである。

次に、誤差を表すカーネルについて考える。これはホワイトノイズカーネルと呼ばれ、以下のように表される。これは共分散行列の分散項に誤差を付け加えるカーネルである。

$$k(\mathbf{x}, \mathbf{x}') = \sigma^2 \delta(n, n') \tag{1.56}$$

ガウス過程回帰にカーネルを用いる上での大きなメリットはカーネル同士を線型結合により組み合わせることができることであり、本研究では RBF カーネルとホワイトノイズカーネルを見合わせたカーネル

$$k(\mathbf{x}, \mathbf{x}') = \theta_1 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{\theta_2}\right) + \sigma^2 \delta(n, n') \tag{1.57}$$

を採用している。ここで、 θ_1, θ_2 はカーネル関数のハイパーパラメータであり、特に θ_2 を length scale という。これは関数の距離にかかる係数であり、length scale が小さいと共分散が距離の差に敏感になるため、より細かい特徴を持つ回帰分布を作成し、逆に大きいと、広範囲な入力空間の相関を捉えるため全体として激しく変動しない回帰分布を作成する。本研究ではハイパーパラメータの最適化に関して、ベイズ最適化の最適化途中に Adam という勾配降下法を応用した手法を用いて最適化することにより、最適化の途中でハイパーパラメータを柔軟に変更している。

また、本研究の改善点として、カーネルをより高度なものにすることが挙げられる。その二つを以下に紹介する。

(1) Matérn カーネル

RF カーネルは正規分布を連続的に重ねて作成された関数であるため、関数が滑らかな、つまり任意の点で微分可能であるという数学的に強い性質がある。回帰する関数として入射効率やロスモニター値を考えたときに、ビームがコリメータやセプタム障壁にあたってパラメータの滑らかでない急激な変化が起こった場合に RF カーネルではうまく対応できない可能性がある。Matérn カーネルは滑らかではない回帰関数を生成することができ、より実際の入射効率やロスモニターの値を再現するような回帰関数が作れるのではないのかと考えている。Matérn カーネルの具体形は以下で表される。

$$k_\nu(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{\|\mathbf{x} - \mathbf{x}'\|}{\theta} \right)^\nu K_\nu \left(\sqrt{2\nu} \frac{\|\mathbf{x} - \mathbf{x}'\|}{\theta} \right) \quad (1.58)$$

ここで、 $\Gamma(\nu)$ は第二種変形ベッセル関数、 θ と ν はハイパーパラメータである。特に ν は関数の滑らかさを表すパラメータであり、このカーネルで回帰された関数は $\lfloor \nu \rfloor$ 回微分可能である。解析的な扱いやすさから通常は半整数 $\nu = 1/2, 3/2, 5/2$ や ∞ が採用されることが多い。

(2) ARD カーネル

RF カーネルにおいて length scale は 1 つであった。しかしこれでは各パラメータ同士で特徴の質が違う場合に対応できない。例えば、本研究で用いられる入力パラメータにはマグネットの位置や角度、入射位相など、様々なパラメータがあり、これらが作るパラメータ空間内での回帰関数は各次元において距離の差に対しての振る舞いが違うと考えるのが妥当である。このため、複数の length scale を用意した関数である ARD カーネルを用いることで回帰関数の性能が向上する可能性がある。ARD カーネルの具体形は以下で表される。

$$k(\mathbf{x}, \mathbf{x}') = \exp \left(-\frac{1}{2} \sum_{i=1}^d \frac{(x_i - x'_i)^2}{\theta_i^2} \right) \quad (1.59)$$

また、Matérn カーネルと ARD カーネルを合わせた以下のようなカーネルを作成することも可能である。

$$k_\nu(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \sqrt{\sum_{i=1}^d \frac{(x_i - x'_i)^2}{\theta_i^2}} \right)^\nu K_\nu \left(\sqrt{2\nu} \sqrt{\sum_{i=1}^d \frac{(x_i - x'_i)^2}{\theta_i^2}} \right) \quad (1.60)$$

ただしカーネルを複雑にしすぎると計算量が増加してしまい、また表現力が高くなりすぎて過学習を起こしてしまうこともあるため、これらのカーネル関数を導入する際は上記の事項を勘案する必要がある。

A.4.6 獲得関数

ブラックボックス最適化とは少ない測定回数で最大値を効率よく見つける手法のことであった。そしてここまではブラックボックス関数の最大値がどこかを予測するために、ガウス過程回帰法とカーネル法を用いて回帰予測をする方法について述べた。次に考えることとしては、この分布の中でどこが最も最大値らしい点であるかということである。ガウス過程回帰法は任意の各点において平均と分散を持っているので、以下の二つの

考え方ができる。一つ目は関数の中で平均が大きい部分の近くにもっと大きな値があるのではないかと
 考え方であり、これを活用的であるという。二つ目は点の情報がありなく分散が大きい領域にはもっと大
 きな値が眠っているのではないかとという考え方であり、これを探索的であるという。この活用と探索という
 概念はトレードオフな関係にあり、このバランスを取りつつ最も最大値らしい点を決める関数が獲得関数で
 ある。本研究では2つの獲得関数について述べる。1つ目がEI(Expected Improvement)であり、2つ目が
 UCB(Upper Confidence Bound)である。

■EI (Expected Improvement) EIは獲得関数の中で最も基本的な獲得関数の内の一つである。ベイズ最適化
 を行う際には、まずEIを用いて試してみることが多い。その理由としては、関数値やその勾配が解析的に解
 けるため比較的高速に計算ができることと、誤差が0である時には無限回の試行を繰り返せば必ずグローバル
 マキシマムに到達することが保証されるという理論的な背景が挙げられる。EIはこれまでの測定点での最適
 な測定値に対して、次点の測定候補である $\mathbf{x}^*$ がどの程度改善するかという量を、その改善量の期待値で評価
 したものである。具体的には、以下の式で表される $EI(\mathbf{x})$ が一番高い候補点 $\mathbf{x}^* = \operatorname{argmax}(EI(\mathbf{x}^*))$ を採用
 し、実際に測定する。

$$EI(\mathbf{x}) = \mathbb{E}[\max(0, y(\mathbf{x}) - y_{\max})] \quad (1.61)$$

$y_{\max}$ は現在までの最適化の中での最大値を取ったパラメータである。ガウス過程回帰の性質より、関数上
 のどの候補点 $\mathbf{x}^*$ においても y は正規分布をしている。 $\mathbb{E}[\max(0, y(\mathbf{x}) - y_{\max})]$ というのは、ある $\mathbf{x}^*$ を決めた
 際の y の正規分布において、 $y_{\max}$ 以上の部分を $(y - y_{\max})$ という重み付けをして積分した面積に等しい。こ
 の $EI(\mathbf{x})$ を期待改善量という。期待改善量を解析的に計算できるような形に変形する。なお、 $\mu_n(\mathbf{x}), \sigma_n(\mathbf{x})$
 はそれぞれ n 回目の最適化における事後分布の平均と分散である。

$$\begin{aligned} EI(\mathbf{x}) &= \mathbb{E}[\max(0, y(\mathbf{x}) - y_{\max})] \\ &= \int_{y_{\max}}^{\infty} (y(\mathbf{x}) - y_{\max}) \mathcal{N}(\mu_n(\mathbf{x}), \sigma_n(\mathbf{x})) dy \\ &= \int_{y_{\max}}^{\infty} (y(\mathbf{x}) - y_{\max}) \frac{1}{\sqrt{2\pi}\sigma_n(\mathbf{x})} \exp\left(-\frac{(y_{\max} - \mu_n(\mathbf{x}))^2}{2\sigma_n(\mathbf{x})^2}\right) dy \end{aligned}$$

ここで、正規分布の標準変換 $Z = \frac{y - \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}$ を施すと、 $Z_{\max} = \frac{y_{\max} - \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}$ として、

$$\begin{aligned} EI(\mathbf{x}) &= \int_{Z_{\max}}^{\infty} (\sigma_n(\mathbf{x})Z + \mu_n(\mathbf{x}) - y_{\max}) \frac{1}{\sqrt{2\pi}\sigma_n(\mathbf{x})} \exp\left(-\frac{Z^2}{2}\right) \sigma_n(\mathbf{x}) dZ \\ &= (\mu_n(\mathbf{x}) - y_{\max}) \int_{Z_{\max}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right) dZ + \sigma_n(\mathbf{x}) \int_{Z_{\max}}^{\infty} Z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right) dZ \\ &= (\mu_n(\mathbf{x}) - y_{\max}) \Phi(Z) + \sigma_n(\mathbf{x}) \left[-\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right) \right]_{Z_{\max}}^{\infty} \\ &= (\mu_n(\mathbf{x}) - y_{\max}) \Phi(Z) + \sigma_n(\mathbf{x}) \phi(Z) \end{aligned} \quad (1.62)$$

となる。ここで、 $\phi(Z)$ と $\Phi(Z)$ はそれぞれ正規分布の確率密度分布と累積密度分布を表し、これは解析的
 に解けることがわかる。最終的にEIの獲得関数は

$$EI(\mathbf{x}) = (y_{\max} - \mu_n(\mathbf{x}))\Phi\left(\frac{y_{\max} - \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) + \sigma_n(\mathbf{x})\phi\left(\frac{y_{\max} - \mu_n(\mathbf{x})}{\sigma_n(\mathbf{x})}\right) \quad (1.63)$$

であり、この獲得関数を最大化するような候補点 $\mathbf{x}^*$ が選ばれる。

EI の欠点として、誤差が大きい場合には最適化がうまくいかないということが一般的に知られており、今回の実験で使用するためには入射効率を 30 回測定し、その中央値を採用することで系統誤差を減らすという工夫が必要となった。

■UCB (Upper Confidence Bound) UCB も EI と同じく最も基本的な獲得関数の 1 つであり、関数値や勾配が解析的に計算できる性質も EI と同様に持ち合わせている。定義は以下の通りであり、EI に比べて非常にシンプルに記述される。

$$UCB(\mathbf{x}) = \mu_n(\mathbf{x}) + \beta\sigma_n(\mathbf{x}) \quad (1.64)$$

UCB の特徴は獲得関数内にハイパーパラメータである β が存在していることであり、この値を調節することで活用と探索のバランスを変更することができる。反面、ハイパーパラメータがあるということは適切な値を自身で設定する必要があり、これはデメリットにもなりうる。

A.4.7 制限付き最適化

制限付き最適化は基本的には通常のベイズ最適化と同様に、確率分布を作成し、獲得関数を作成し、獲得関数の最大値を実際に測定するという手続きを行うが、最適化の途中にて何かしらの制約条件を課し、制約を満たさない候補に対してはペナルティを与えるという工程が加わった最適化である。本研究における制約とはアポートしない領域であり、これはビームロスモニターの値が高いか低いかで判断することができる。ロスモニターで実際に測定していない部分については、入射効率と同じく測定点の情報をもとに確率分布を挿入することによって見積もることができる。

本研究において制限つき最適化を実装するための具体的な手法について述べる。ベイズ最適化の各トライアルについて、入射効率の予測分布だけでなく、ビームロス値の予測分布も同時に作成し、予測分布をもとにリスクを以下のように定義する。ここで $\hat{\mu}_n(\mathbf{x}), \hat{\sigma}_n(\mathbf{x})$ はビームロスの予測分布の平均と分散である。

$$\mathbf{Risk}(\mathbf{x}) = \hat{\mu}_n(\mathbf{x}) + R\hat{\sigma}_n(\mathbf{x}) \quad (1.65)$$

この Risk がある閾値 T を下回る点のみを採用し、そのほかの値は採用しないようにする。^{*1}ここで、 R は Risk のハイパーパラメータである。予測分布の各点は正規分布に従っているので、 $R = 1$ であれば 1σ の片側信頼区間である 84 % の確率で T を下回ることを担保し、同様に $R = 2$ であれば 97.5 % の確率で T を下回ることを担保する。また、制限は複数存在していてもよく、別々のロスモニターの値を同時に監視することも可能である。図 A.6 に制限付きベイズ最適化のフローチャートを示す。

^{*1} 内部処理としては、獲得関数の値に $-\infty$ の減点をする。

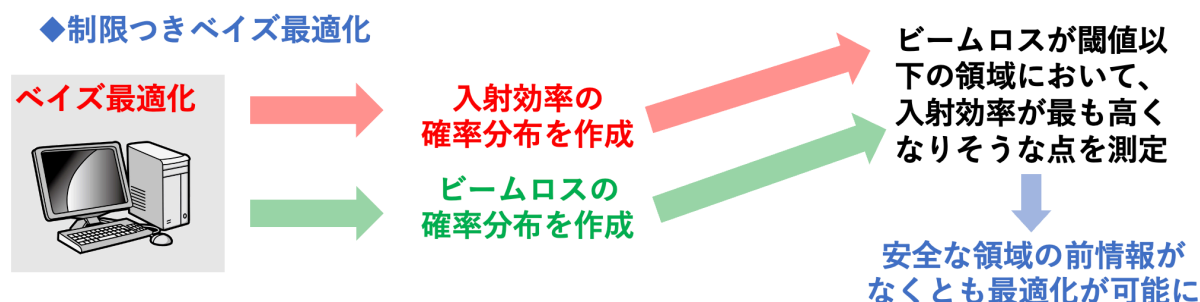


図 A.6: 制限付きベイズ最適化のフローチャート

制限付き最適化のデメリットとしては、設定すべきパラメータが多く、制限を適切に課することが難しいことが挙げられる。例えば式 (1.65) での閾値 T をどの値にすればいいのか、リスク R を定める際、どの程度の信頼区間を担保すればいいのか、また、SuperKEKB の中でどのビームロスモニターを用いればいいのかなど設定項目は幅広い。これらの 1 つでもうまくいかなければビームをアポートしてしまう危険性があるため、小分け最適化と比較して実装までに多くのスタディ時間とビームコンディションに応じた調整が必要になる。

ベイズ最適化には SafeOpt という、このように安全な領域内での最適化を想定した制約つき最適化パッケージがあるが、これは GpyOpt というライブラリをベースに作成されている。しかし GPyOpt の開発は終了しており、公式 GitHub リポジトリではメンテナンスの終了とリポジトリのアーカイブが宣言されているため、今後の更新やサポートは期待できない。このような事情から、GpyOpt ではなく BoTorch という PyTorch ベースのモジュールを使用し、自身で制限つき最適化のコードを作成した。

A.4.8 Optuna

Optuna はオープンソースとして開発されたブラックボックス最適化ツールであり、ベイズ最適化をはじめとして、様々な最適化手法を比較的簡易なコーディングによって実装可能である。Optuna を使用するメリットとしては以下の項目があげられる。

1. 多様なアルゴリズムを最小限の変更で使用可能であり、カスタマイズも容易にできること。
2. 様々な可視化機能が備わっており、データを取得した際にその最適化結果の多角面からの解析がデフォルトで用意されていること。
3. ユーザーが多く、情報が他の最適化ツールに比べて容易に手に入ること。

本研究では、Optuna に実装されている BoTorch のライブラリをもとにベイズ最適化を実装している。

付録 B

KEK Linac での陽電子収量最大化スタディ

本章では SuperKEKB 入射調整の予備実験として KEK Linac で実施された陽電子収量最大化スタディについて論じる。B.1 節では KEK Linac で行われているビーム調整の方法と今回の実験で用いる機器について紹介し、B.2 節でスタディ結果について報告し、B.3 節でまとめを述べる。

B.1 スタディの目的と実験のセットアップ

B.1.1 スタディの目的

本章の目的は、次章で扱う SuperKEKB 入射調整に機械学習の一種であるベイズ最適化を適応させるための予備実験として、KEK Linac における陽電子収量最大化スタディにおいてベイズ最適化が機能するかを確認することであった。KEK Linac で予備実験を行った理由は 2 点挙げられる。1 点目は KEK Linac がビーム上流に存在するため、マグネットや RF 空洞の異変によるビーム変調の影響を受けることがビーム下流に比べて少なく、ビーム状態が時間によって変化しにくい点である。2 点目は蓄積リングとは異なりビーム軌道を多く変化させてもビームアポートを促すことがなく、電流を再び蓄積するまでの待機時間がないため、スタディが迅速に行える点である。試験結果では、ベイズ最適化による最適化性能、獲得関数の性能評価とハイパーパラメータの選定、入射調整試験で必要となる機能の動作確認について報告する。また、Linac スタディを通じて明らかになった特筆すべき性質についても論じる。

B.1.2 KEK Linac における手動調整手法

KEK Linac でのビーム調整は調整オペレーターによって行われており、その種類は多岐に渡る。調整順序は以下の通りである。

表 B.1: KEK Linac における調整順序

順序	調整目的	調整内容
1	ビーム進行方向の軌道中心	バンチャー位相や RF 位相の調整
2	垂直、水平方向の軌道中心	ステアリングマグネットの調整

3	ビーム進行方向の位相空間	バンチャー位相の調整
	垂直、水平方向の位相空間	四極マグネットの調整
4	Dispersion	上記含む様々な機器の微調整

本研究では表 B.1 の順序 2 に対してベイズ最適化を導入した。6 つのステアリングマグネットを用いて陽電子生成用の電子ビームを操作し、タングステン標的の後部の陽電子の収量を最大化させるという調整である。タングステン標的より上流の電子ビームの軌道と、電子とタングステン標的が衝突して発生した電磁シャワーによって生成された低エネルギーの陽電子ビームの電荷量の関係は自明ではなく、ブラックボックスになっている。この調整は機械学習を導入する良い応用例となる。

B.1.3 実験のセットアップ

実験のセットアップは以下の通りである。

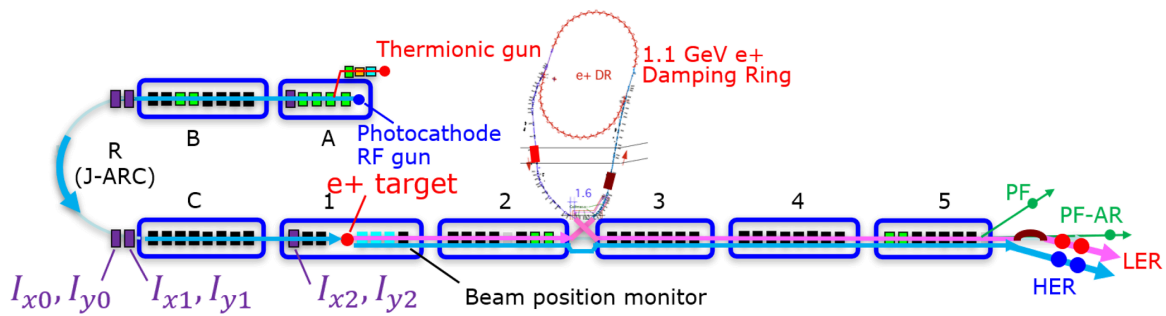


図 B.1: KEK Linac における陽電子収量最大化のセットアップ

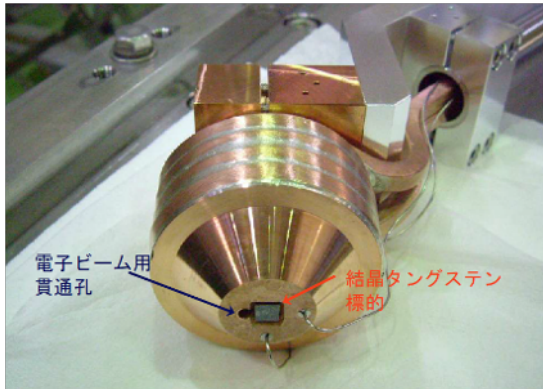
調整パラメータはステアリングマグネットの印加電流である。x 軸と y 軸に独立してビーム軌道を蹴れるステアリングマグネット対を 3 対操作するので、合計パラメータは $2 \times 3 = 6$ パラメータである。2 対は J-ARC の最後部に、1 対はセクション 1 の最前部に配置されている。セクション 1 の中心付近にタングステンターゲットがあり、電子ビームをターゲットに当てることで陽電子を生成する。評価パラメータとしてはセクション 1 後部のビーム位置モニター (BPM) で測定された陽電子電荷を用いる。本研究でのスタディで用いられる装置の外観の外観を図 B.2 に示す。



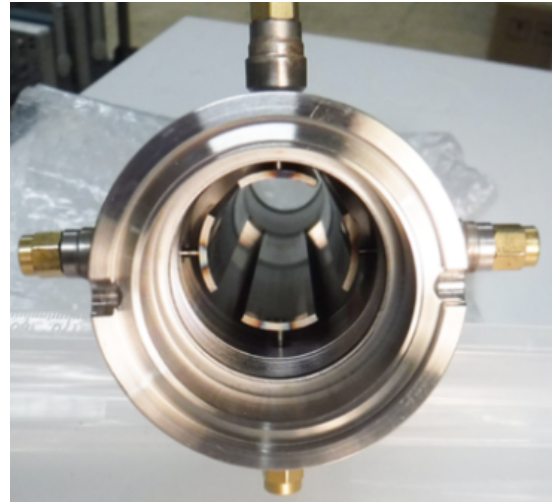
(a) 陽電子生成用熱電子銃



(b) パルスステアリングマグネット



(c) 陽電子生成用タングステンターゲット



(d) Stipline BPM

図 B.2: 本研究でのスタディで用いられる装置の外観

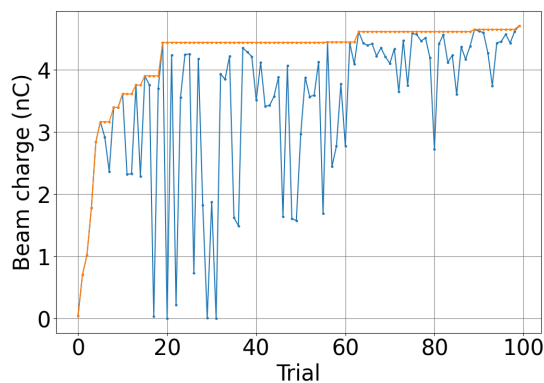
B.2 スタディ結果

3つの項目に注目する。獲得関数の比較、初期化条件の比較、diff per trial である。

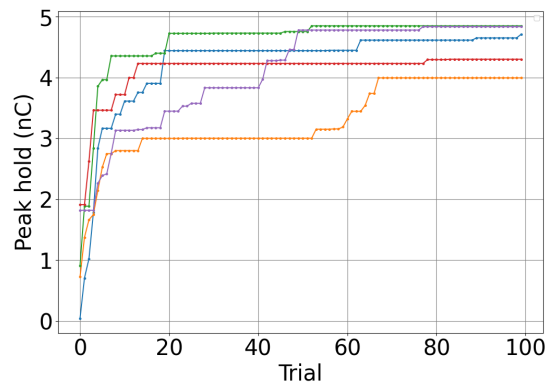
B.2.1 獲得関数の比較

2023 年 6 月 14 日に入射器制御棟にてスタディを行った。この日のオペレータの調整での陽電子収量は 4.6nC であった。ベイズ最適化によって 1 ランにかかる時間は 10 分程度であり、初期化方法は最初の 10 点をパラメータ空間内にランダムで生成する方式を採用した。パラメータの定義域は調整オペレータの経験に基づいて設定した。獲得関数について、EI と UCB ($\beta = 2$) と UCB ($\beta = 5$) を比較する。EI の 1 回のランの結果は以下である。

ここで青線は各トライアルに対する BPM の電荷量であり、橙線はその最大値を保持した線であり、



(a) シングルラン

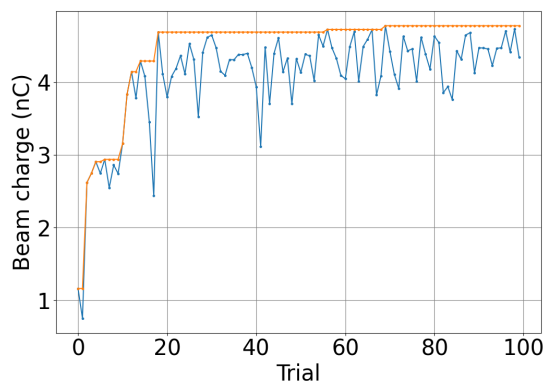


(b) 5 回のランの Peakhold

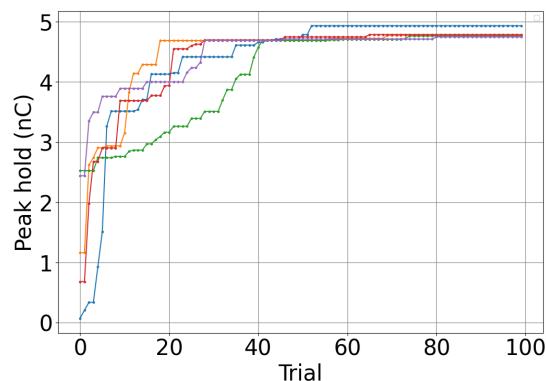
図 B.3: EI の各トライアルに対する BPM の電荷量 (nC)

Peakhold と呼ばれる。EI の 5 ランに対して各 Peakhold を重ねた図が図 B.3b である。

同様に UCB ($\beta = 2$) のシングルランと Peakhold は以下。



(a) シングルラン



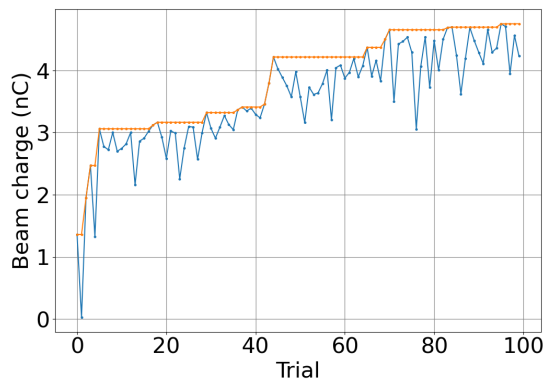
(b) 5 回のランの Peakhold

図 B.4: UCB ($\beta = 2$) の各トライアルに対する BPM の電荷量 (nC)

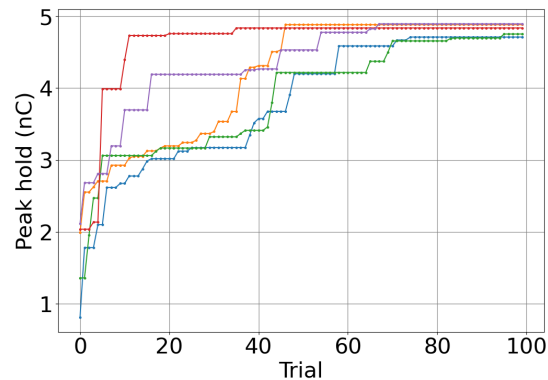
同様に UCB ($\beta = 5$) のシングルランと Peakhold は以下。

まず、シングルランの図で EI と UCB を比較すると、EI ではビーム電荷が低い部分を探索する傾向があるが UCB はビーム電荷が全体的に高い傾向がある。このことから EI と UCB の振る舞いの違いが定性的に見て取れる。次に、ピークホールドに関しては EI はオペレータの調整値を越していないランが 2 ラン見受けられるが、UCB は全てのランでオペレータを超える最適化が確認できた。また、UCB の中では $\beta = 2$ の方が最適化にかかるトライアル数が少ないことから、三つの獲得関数の中では UCB ($\beta = 2$) が最適化性能と最適化速度の両方において優れていることがわかった。

次に、 β が 2 より小さい場合を検証していなかったため、UCB ($\beta = 2$) と UCB ($\beta = 0.1$) についても 2023 年 12 月 12 日に追実験を行った。その結果が以下。なおこの日のオペレータの調整値は 4.9 nC である。また、5 回の平均をベイズ最適化の評価値として採用した。

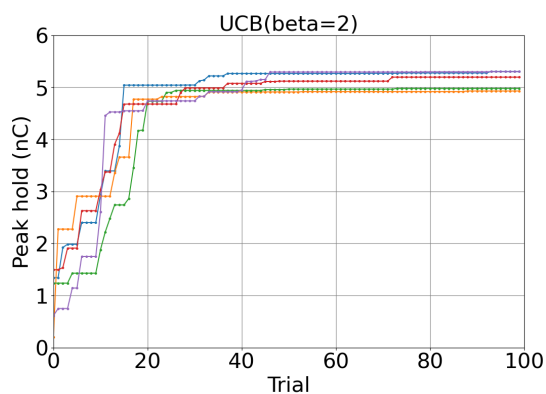


(a) シングルラン

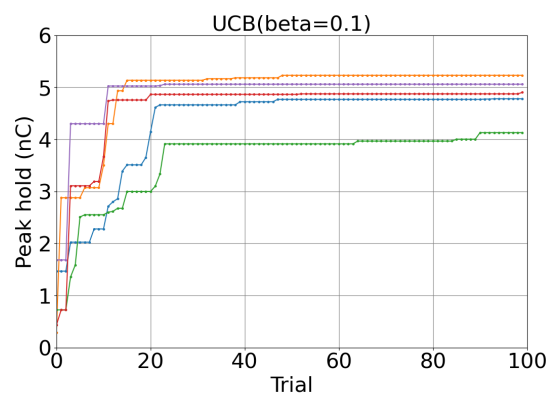


(b) 5 回のランの Peakhold

図 B.5: UCB ($\beta = 5$) の各トライアルに対する BPM の電荷量 (nC)



(a) UCB ($\beta = 2$)



(b) UCB ($\beta = 0.1$)

図 B.6: UCB ($\beta = 2$) と UCB ($\beta = 0.1$) の 5 ランに対する Peakhold の比較

それぞれの Peakhold を比較すると、 $\beta = 2$ では全てのランにおいてオペレータの調整値を越しているが、 $\beta = 0.1$ では 2 ランのみしか調整値を越していないため、UCB ($\beta = 2$) が最適化性能において優れていることがわかる。

また特筆すべきこととして、最適化においては全ての最適化で特定のパラメーターセットに辿り着いていくわけではなく、別のそれぞれ別の最適なパラメーターセットに辿り着いていることが挙げられる。例えば UCB($\beta = 2$) の 2 つのランでのペアプロットを見比べると、最大値は同程度であるが別のパラメータセットが選択されていることが見て取れる。

B.2.2 初期化条件の比較

初期化条件を 3 つの条件で比較した。1 つ目は上記の方法と同様に最初の 10 点を乱数で生成する場合、2 つ目は定義域の端から始める場合、3 つ目は 2 日前の最適化のデータを参照し、上位 10 点を初期値とした場合である。

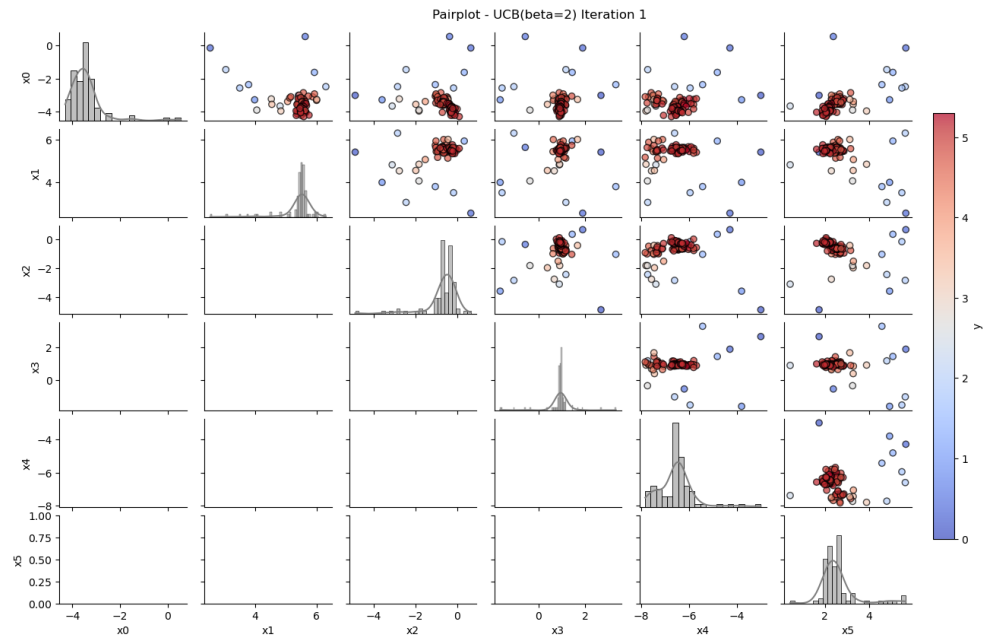


図 B.7: $UCB(\beta = 2)$ の 1 ラン目のペアプロット

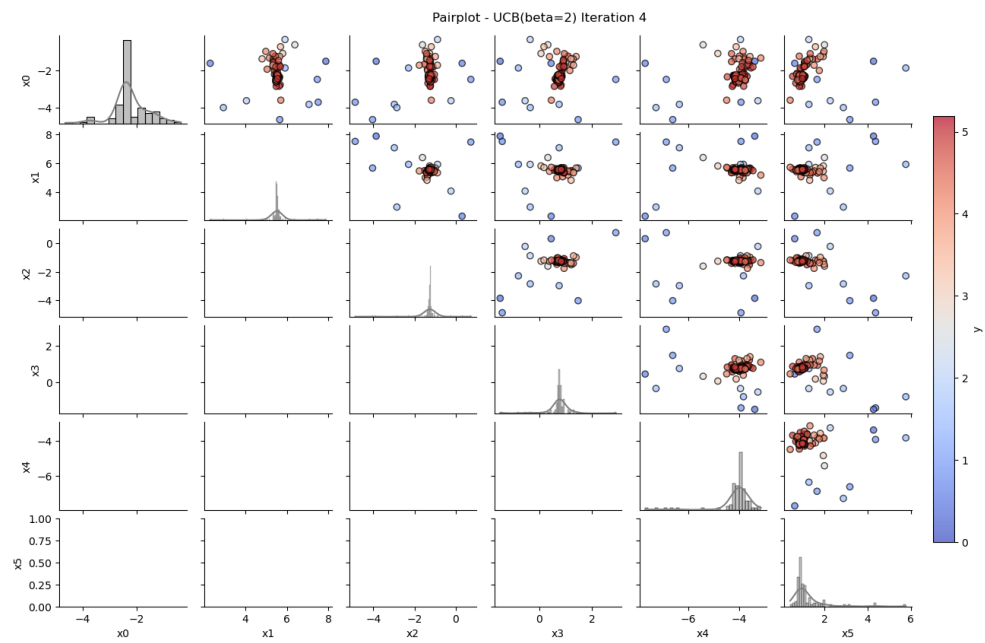


図 B.8: $UCB(\beta = 2)$ の 4 ラン目のペアプロット

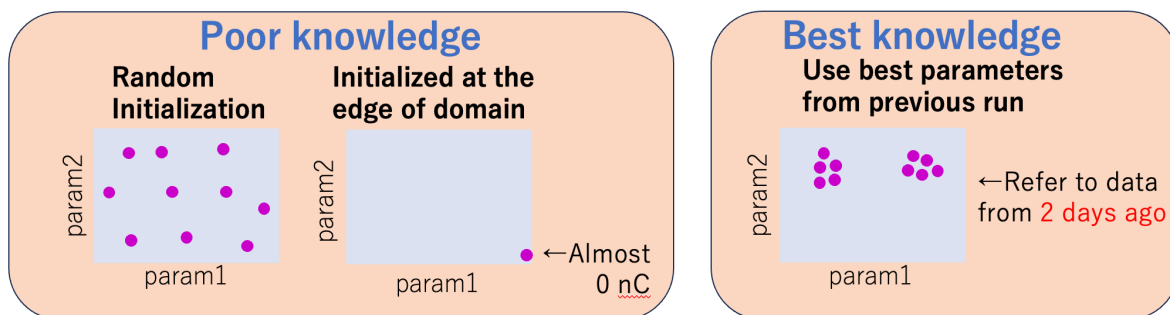


図 B.9: 初期化方法の説明

10 点をランダムに生成した場合は図 B.6a で既に示した。残りの二つの初期化条件においては 2023 年 12 月 14 日に 3 ランずつ測定した。測定日の調整オペレータの Peakhold の結果を比べた結果が以下。なおこの日のオペレータの調整値は 5.2 nC である。

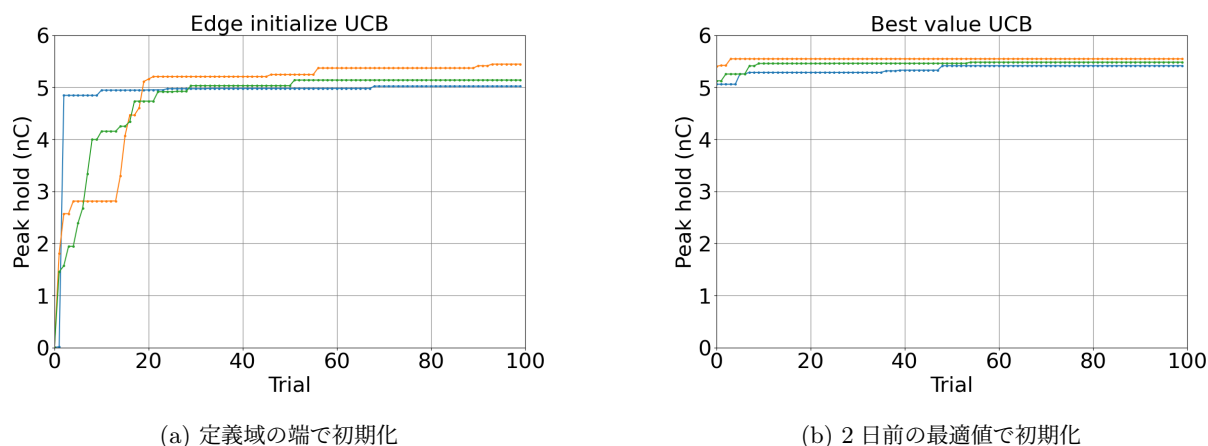
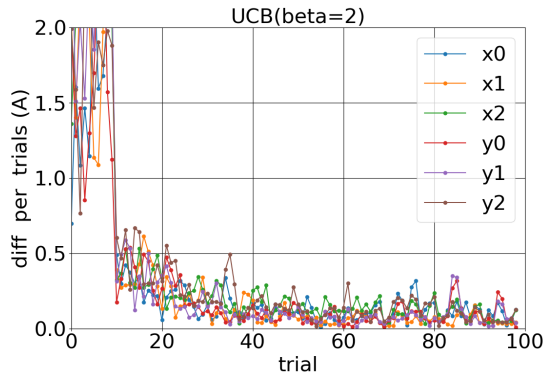


図 B.10: UCB ($\beta = 2$) での初期化条件に対する Peakhold の比較

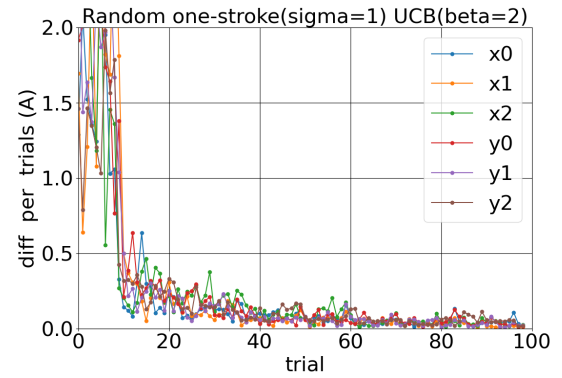
この結果から、以下の 2 点が見て取れる。1 点目は乱数生成や定義域の端からのコールドスタートであっても、100 トライアル内でオペレータの調整値を超える最適化が可能であるということである。2 点目は過去の最適値のデータを用いるウォームスタートではほぼ数十秒でオペレータの調整値を越す最適化が可能であるということである。これらより、バイズ最適化は情報が無い時点から最適値を探索する機能と、過去の情報を用いて最適値を微調整する機能を状況に応じて使い分けられることがわかった。

B.2.3 入射調整のために必要な機能の動作確認

この節では、本編で使用した近傍探索機能の動作確認を行った。近傍探索機能を実装した場合としていない場合での各パラメータについて 1 つ前のトライアル値からの差分を取ったプロットを作成した。なお、10 トライアル目まではパラメータの変化幅が大きくなっているが、これは定義域内のランダムなパラメータで初期化しているためである。



(a) 近傍探索機能実装なし



(b) 近傍探索機能実装あり

図 B.11: 近傍探索機能の実装の有無に対するパラメータのトライアルごとの差分比較

2つのグラフを見比べると、導入後に各パラメータの移動距離が抑制されていることがわかる。

B.3 まとめ

KEK Linac において、陽電子の収量を最大化するために6つのステアリングマグネットをベイズ最適化によって調整した。その結果、オペレーターと同等の最適化性能を達成できることが確認された。特に、獲得関数としてUCB ($\beta = 2$) を採用した場合、最適化性能および最適化速度の両面で優れていることが示された。また、初期条件の比較では、コールドスタートから最適化を開始した場合でも、オペレーターによる調整値と同程度の性能を達成できることが確認された。さらに、数日前のデータを参照して最適化を開始するウォームスタートでは、最適化時間が大幅に短縮されることが示された。加えて、近傍探索機能についても動作確認を行い、その適切な機能が確認された。

参考文献

- [1] Makoto Kobayashi and Toshihide Maskawa. “CP Violation in the Renormalizable Theory of Weak Interaction”. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.
- [2] K Abe et al. “Observation of large CP violation in the neutral B meson system.” In: *Physical review letters* 87 (Sept. 2001), p. 091802.
- [3] *Belle II* 実験 パンフレット. 2018. URL: https://www.kek.jp/proffice/Pamphlets/IPNS/BELLE%20II_J.pdf.
- [4] *SuperKEKB Design Report, Physics Motivation*. 2014. URL: <https://kds.kek.jp/event/15914/contributions/28485/attachments/23460/28033/Chapter1PhysicsMotivation.pdf>.
- [5] W. Altmannshofer et al. “The Belle II Physics Book”. In: *PTEP* 2019.12 (2019). Ed. by E. Kou and P. Urquijo. [Erratum: *PTEP* 2020, 029201 (2020)], p. 123C01. DOI: 10.1093/ptep/ptz106. arXiv: 1808.10567 [hep-ex].
- [6] *SuperKEKB Design Report, Injector Linac*. 2014. URL: https://www-linac.kek.jp/linac-com/report/skb-tdr/11_Linac-201023.pdf.
- [7] *SuperKEKB プロジェクト全体図*. 2018. URL: <https://www.kek.jp/ja/about/pr/image/12548>.
- [8] 夏井 拓也. *SuperKEKB へのビーム輸送路とダンピングリング*. OHO 2019. 2019. URL: http://accwww2.kek.jp/oho/OH0txt/OH0-2019/03_Natsui_Takuya.pdf.
- [9] *SuperKEKB Design Report, Beam Transport*. 2014. URL: https://kds.kek.jp/event/15914/contributions/28479/attachments/136391/166653/BT_Revised2020_10_20.pdf.
- [10] 飯田 直子. *SuperKEKB へのビーム輸送路とダンピングリング*. OHO 2019. 2019. URL: http://accwww2.kek.jp/oho/OH0txt/OH0-2019/07_Iida_Naoko.pdf.
- [11] Y. Suetsugu et al. “Mitigating the electron cloud effect in the SuperKEKB positron ring”. In: *Physical Review Accelerators and Beams* 22 (Feb. 2019). DOI: 10.1103/PhysRevAccelBeams.22.023201.
- [12] 大内 徳人. *SuperKEKB ビーム衝突点用超伝導電磁石システムの建設*. pasj. 2017. URL: https://www.pasj.jp/web_publish/pasj2017/proceedings/PDF/TUOM/TUOM05.pdf.
- [13] 池田 仁美. *BEAM LOSS MONITOR AT SUPERKEKB*. IBIC. 2014. URL: <https://accelconf.web.cern.ch/ibic2014/papers/tupd22.pdf>.
- [14] 船越 義裕. *SuperKEKB のマシンパラメータ*. 2011. URL: http://accwww2.kek.jp/oho/oho11/oho11_txt/1_Funakoshi_20110823.pdf.

- [15] 後田 裕. *SuperKEKB と Belle II*. 2016. URL: https://indico.cern.ch/event/492899/contributions/1171468/attachments/1249245/1841123/Belle_II_1.pdf.
- [16] *SuperKEKB Design Report, Lattice Design*. 2014. URL: https://kds.kek.jp/event/15914/contributions/28486/attachments/136910/166608/LatticeDesign2020_10_19.pdf.
- [17] *Luminosity projection*. 2024. URL: https://www-superkekb.kek.jp/Luminosity_projection/.
- [18] *Luminosity status*. 2024. URL: <https://www.belle2.org/research/luminosity/>.
- [19] 大西 幸喜. *26th KEKB Accelerator Review Committee*. 2022. URL: https://www-kekb.kek.jp/MAC/2022/Report/Ohnishi_Machine.pdf.
- [20] J. A. Nelder and R. Mead. “A Simplex Method for Function Minimization”. In: *The Computer Journal* 7.4 (Jan. 1965), pp. 308–313. ISSN: 0010-4620. DOI: 10.1093/comjnl/7.4.308. eprint: <https://academic.oup.com/comjnl/article-pdf/7/4/308/1013182/7-4-308.pdf>. URL: <https://doi.org/10.1093/comjnl/7.4.308>.
- [21] 夏井 拓也. *SuperKEKB positron beam tuning using ML*. LCWS 2024. 2024. URL: https://agenda.linearcollider.org/event/10134/contributions/54875/attachments/39670/62645/LCWS_Natsui.pdf.
- [22] J. Mockus, V. Tiesis, and A. Zilinskas. “The application of Bayesian methods for seeking the extremum”. In: *Towards Global Optimization*. Ed. by L. C. W. Dixon and G. P. Szego. Vol. 2. Elsevier, 1978, pp. 117–129.
- [23] J. Mockus. *Bayesian Approach to Global Optimization: Theory and Applications*. Dordrecht, Netherlands: Kluwer Academic Publishers, 1989. ISBN: 978-94-015-7793-1. DOI: 10.1007/978-94-015-7793-1.
- [24] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006. ISBN: 9780262182539. URL: <http://www.gaussianprocess.org/gpml/>.
- [25] Gaku Mitsuka et al. “Machine-learning approach for operating electron beam at KEK electron/positron injector linac”. In: *Phys. Rev. Accel. Beams* 27 (8 2024), p. 084601. DOI: 10.1103/PhysRevAccelBeams.27.084601. URL: <https://link.aps.org/doi/10.1103/PhysRevAccelBeams.27.084601>.
- [26] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: 1412.6980 [cs.LG]. URL: <https://arxiv.org/abs/1412.6980>.
- [27] 加藤 臣之輔. 機械学習を用いた *SuperKEKB* 加速器入射調整ツールの実用化に向けて. JPS 2024. 2024. URL: <https://kds.kek.jp/event/52312/contributions/275395/attachments/182974/245320/16aWA202-13.pdf>.
- [28] Auralee Edelen Ryan Roussel. *Proximal Biasing for Bayesian Optimization and Characterization of Physical Systems*. NeurIPS Workshop on Machine Learning and the Physical Sciences. 2021. URL: https://ml4physicalsciences.github.io/2021/files/NeurIPS_ML4PS_2021_82.pdf.

- [29] Ryan Roussel et al. “Bayesian optimization algorithms for accelerator physics”. In: *Phys. Rev. Accel. Beams* 27.8 (2024), p. 084801. DOI: 10.1103/PhysRevAccelBeams.27.084801. arXiv: 2312.05667 [physics.acc-ph].
- [30] Scott M Lundberg and Su-In Lee. “A Unified Approach to Interpreting Model Predictions”. In: *Advances in Neural Information Processing Systems*. Vol. 30. 2017.
- [31] 「*SuperKEKB* プロジェクト」 *Belle II* 実験の本格的な物理解析のためのデータ取得 (フェイズ 3) がいよいよ始まる. 2018. URL: <https://www.ipmu.jp/ja/20190312-Belle2-phase3>.
- [32] T. Abe. “Belle II Technical Design Report”. In: (2010). arXiv: 1011.0352 [physics.ins-det].
- [33] T. Ishibashi et al. “Movable collimator system for SuperKEKB”. In: *Phys. Rev. Accel. Beams* 23 (5 2020), p. 053501. DOI: 10.1103/PhysRevAccelBeams.23.053501. URL: <https://link.aps.org/doi/10.1103/PhysRevAccelBeams.23.053501>.
- [34] 梶裕志 吉原圭亮. *SuperKEKB* でのビームロスとの戦い. 2023. URL: <https://www.jahep.org/hepnews/2023/42-1-3-kekb.pdf>.
- [35] A. Natochii et al. “Measured and projected beam backgrounds in the Belle II experiment at the SuperKEKB collider”. In: *Nucl. Instrum. Meth. A* 1055 (2023), p. 168550. DOI: 10.1016/j.nima.2023.168550. arXiv: 2302.01566 [hep-ex].