Sample-Efficient and Principled Aerospace Decision-Making with Small Data

Ashwin Renganathan, Ph.D.

Computational complex engineered Systems Design Laboratory (CSDL)
Penn State. sites.psu.edu/csdl

Advanced Modeling & Simulation (AMS) Seminar Series
NASA Ames Research Center, December 4th, 2025



Acknowledgements

Office of Naval Research
Science & Technology
USN
ONR
DEPARTMENT OF ENERGY
UNITED STATES OF AMERICA
PennState
Institute for Computational and Data Sciences

Acknowledgements



Kade Carlson – former graduate student; now at P1 AI

Outline

Introduction and overview

Kernel methods and Bayesian decision theory

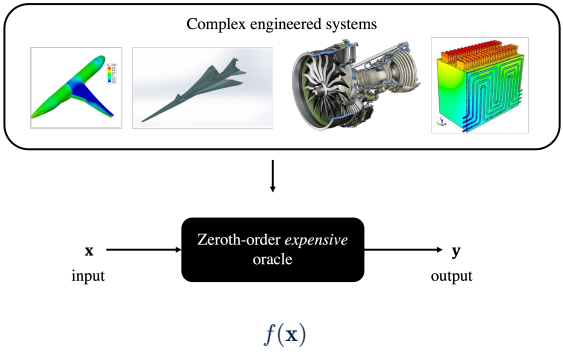
Multiobjective aerodynamic design optimization

Conclusions



CSDL overview

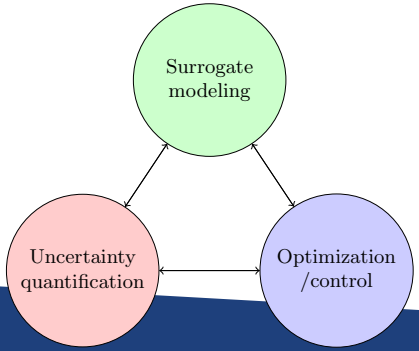
We develop *efficient and scalable computational techniques* toward the design complex engineered systems.



CSDL overview

- **What is design?** We interpret design $\equiv$ decision-making Bloebaum and McGowan (2012); we are interested in the **mathematics** of computational decision-making.

Figure 3: The "**decision-making triad**"



Motivation: certification by analysis (CbA)



Figure 4: FAA Part 25 ceritifcation: high angle of attack and cross-wind takeoff. Mauery et al. (2021)

Motivation: certification by analysis (CbA)

- Certification of aircraft is time-consuming (scheduling delays 3-5 years), expensive ($100M), risky, and occasionally unreliable.
- CbA: cut 50% flight tests; replace with computational analysis with rigorous UQ!
- Our methods from the decision-making triad play a key role.

Rigorous and *uncertainty quantified* computational decision-making techniques, leveraging *high fidelity* models of aircraft systems and subsystems, can drastically reduce certification times and cost.

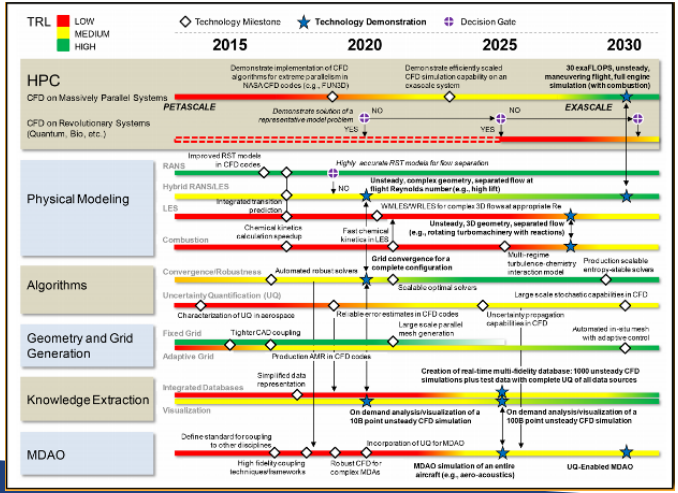
Motivation: certification by analysis (CbA)

- Certification of aircraft is time-consuming (scheduling delays 3-5 years), expensive ($100M), risky, and occasionally unreliable.
- CbA: cut 50% flight tests; replace with computational analysis with rigorous UQ!
- Our methods from the decision-making triad play a key role.

Challenge: decision-making (UQ, optimization/control, surrogates) with high-fidelity computer models is prohibitively expensive!

- For example, aerodynamic model of an aircraft costs ~ 2000 CPU hours.
- Evaluating the model more than a few dozen times is not practical.

Motivation: NASA CFD Vision 2030



Note 5: Aerospace community seeks to integrate physics-based simulations more with aircraft design, certification, and qualification process

CSDL research objectives



We are interested in developing decision-making methods in the presence of uncertainty, with expensive computer models, with the following attributes:

1. Sample efficiency (minimal queries of the expensive model).
2. Derivative-free (adapt to absence of derivatives).
3. Scalability (to high dimensions).
4. Theoretical guarantees (error bounds and/or consistency).

CSDL research objectives



We are interested in developing decision-making methods in the presence of uncertainty, with expensive computer models, with the following attributes:

1. **Sample efficiency** (minimal queries of the expensive model).
2. **Derivative-free** (adapt to absence of derivatives).
3. Scalability (to high dimensions).
4. Theoretical guarantees (error bounds and/or consistency).

What this talk is (not) about.

- Developing methods for the decision-making triad driven by **probabilistic AI/ML**.
- Development of **adaptive** techniques that enable sample efficiency.
- Development of techniques that can **accommodate uncertainty**.
- Development of methods that are *agnostic* to the underlying physics of the system.
- No neural networks.
- **Our approach**: kernel methods with Bayesian decision theory.

Kernel methods and Bayesian decision theory



Kernel methods

Our methodologies fall under the broad category of "kernel methods" Hofmann et al. (2008); Campbell (2001).

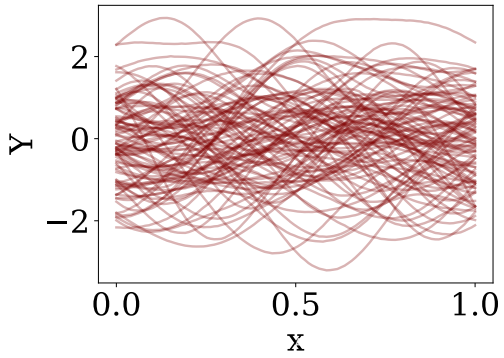
- A class of methods where data $\mathbf{x}$ is transformed into a higher dimensional feature space $\phi(\mathbf{x})$ for better expressibility.
- A *kernel* $k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^\top \Sigma \phi(\mathbf{x})$ can be thought of as a similarity function over all pairs of data points computed using inner products.
- **Kernel trick.** A way to directly specify $k(\mathbf{x}, \mathbf{x}')$ without the need to compute the feature space and its inner product Rasmussen and Williams (2006).

Gaussian process models

$$f(\mathbf{x}) \approx Y(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), k(\cdot,\cdot;\Omega)) \qquad \text{(Prior)}$$

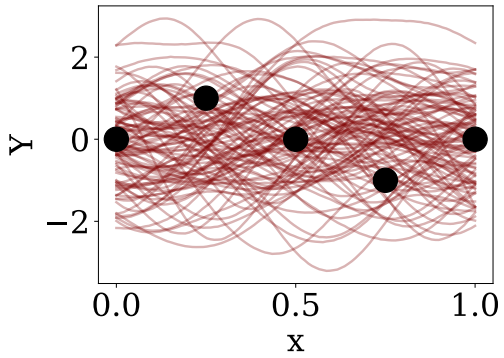
- Mean function $\mu(\mathbf{x})$, covariance function $k(\cdot,\cdot)$, hyperparameters Ω Rasmussen and Williams (2006).

Figure 6: Prior realizations [$\mu = 0$, $k(\mathbf{x},\cdot) = \exp(\|\mathbf{x}-\cdot\|^2/\ell_v^2)$]



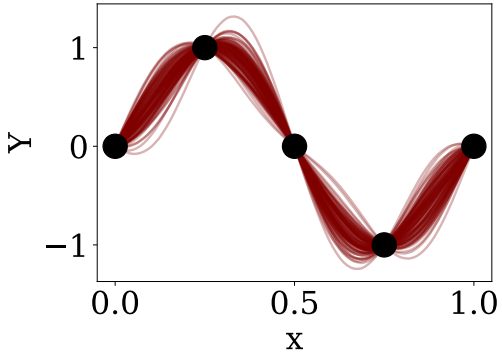
Gaussian process models

$$\mathcal{D} = \{\mathbf{x}_i, f(\mathbf{x}_i)\}_{i=1}^{n}$$ (Observations)

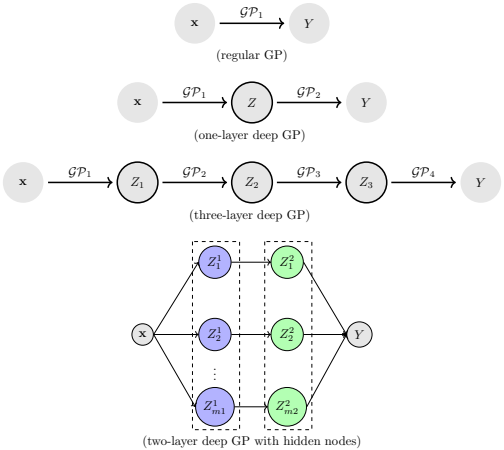


Gaussian process models - Bayes' law

$$Y(\mathbf{x})|\mathcal{D} \propto \mathcal{D}|\, Y(\mathbf{x}) \times\ Y(\mathbf{x}) = \mathcal{GP}(\mu_n, \sigma_n^2) \qquad \text{(Posterior)}$$

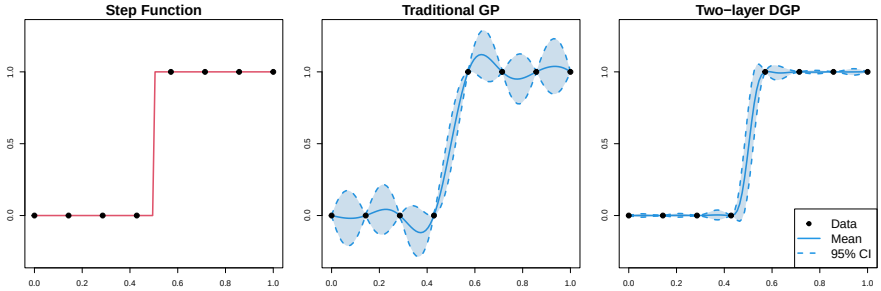


Deep Gaussian process models



Deep Gaussian process models

Figure 8: DGP Vs. traditional GP. Sauer et al. (2023)



Bayesian decision theory

- Let $u(\mathbf{x}) = u \circ f(\mathbf{x})$ quantify the *utility* of evaluating a candidate input $\mathbf{x}$ with f.
- **Surrogate estimate**: $u(\mathbf{x}) = u \circ Y(\mathbf{x})$
- Given observations $\boldsymbol{\mathcal{D}}_n$, new observations can be greedily chosen as

$$\mathbf{x}_{n+1} = \underset{\mathbf{x}}{\operatorname{argmax}} \ \mathbb{E}_{Y(\mathbf{x})}[u(\mathbf{x}) \mid \boldsymbol{\mathcal{D}}_n].$$

- $\alpha(\mathbf{x}) := \mathbb{E}_{Y(\mathbf{x})}[u(\mathbf{x}) \mid \boldsymbol{\mathcal{D}}_n]$ is called an **acquisition** function.
- Bayesian decision-theoretic optimization = **Bayesian optimization.**



Bayesian optimization

- Say the goal is to find $\mathbf{x}_* = \text{argmax}_{\mathbf{x}} f(\mathbf{x})$.
- And we fit a GP $Y_n(\mathbf{x})$ with observations $\{(\mathbf{x}_1, f(\mathbf{x}_1)), \ldots, (\mathbf{x}_n, f(\mathbf{x}_n))\}$.
- Denote as $f_+ = \max_{i \in [n]} f(\mathbf{x}_i)$ (that is, the *best* observed value).
- Then, a popular utility function to choose is $u(\mathbf{x}) = \max(0,\, Y(\mathbf{x}) - f_+)$.
  - The resulting acquisition function

$$\alpha(\mathbf{x}) = \mathbb{E}_{Y(\mathbf{x})}[\max(0,\, Y(\mathbf{x}) - f_+) \mid \boldsymbol{\mathcal{D}}_n] = \sigma_n(\mathbf{x})\phi\left(\frac{\mu_n(\mathbf{x}) - f_+}{\sigma_n(\mathbf{x})}\right) + (\mu_n(\mathbf{x}) - f_+)\Phi\left(\frac{\mu_n(\mathbf{x}) - f_+}{\sigma_n(\mathbf{x})}\right)$$

  - Is the well known **expected improvement (EI)** Jones et al. (1998)
- Choose $\mathbf{x}_{n+1} = \text{argmax}_{\mathbf{x}}\ \alpha(\mathbf{x})$; observe $(\mathbf{x}_{n+1}, f(\mathbf{x}_{n+1}))$, update GP, and repeat.

Our past work on kernel methods and Bayesian decision theory

- **Surrogate modeling for aerodynamics.**
  - **R**, Maulik, R., & Ahuja, J. (2021). Enhanced data efficiency using deep neural networks and Gaussian processes for aerodynamic design optimization. *Aerospace Science and Technology*, 111, 106522.
  - **R**, Maulik, R., & Rao, V. (2020). Machine learning for nonintrusive model order reduction of the parametric inviscid transonic flow past an airfoil. *Physics of Fluids*, 32(4).
  - **R**, Harada, K., & Mavris, D. N. (2020). Aerodynamic data fusion toward the digital twin paradigm. *AIAA Journal*, 58(9), 3902-3918.

- **Uncertainty quantification.**
  - Booth, A. S., **R**, & Gramacy, R. B. (2025). Contour Location for Reliability in Airfoil Simulation Experiments using Deep Gaussian Processes. *Annals of Applied Statistics*, 19(1), 191-211.
  - Booth, A. S., & **R** (2025). Two-stage design for failure probability estimation with Gaussian process surrogates. *Journal of Quality Technology*, 1-17.
  - **R**, Rao, V., & Navon, I. M. (2023). CAMERA: A method for cost-aware, adaptive, multifidelity, efficient reliability analysis. *Journal of Computational Physics*, 472, 111698.

"Zoo" of acquisition functions

- Expected improvement (EI) Jones et al. (1998)
- Probability of improvement (PI) Mockus et al. (1978)
- GP upper confidence bound Srinivas et al. (2009).
- Entropy search (ES) Hennig and Schuler (2012), Predictive ES Hernández-Lobato et al. (2014), Max-value ES Wang and Jegelka (2017).
- Stepwise uncertainty reduction (SUR) Picheny (2014).
- Knowledge gradient Frazier et al. (2008).
- **Thompson sampling** Thompson (1933); Kandasamy et al. (2018)

Bayesian optimization with Thompson sampling

- Thompson sampling posits that we make decisions according to the probability that they are optimal.

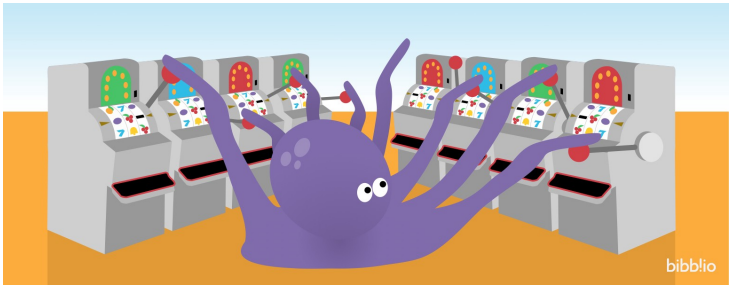


Figure 9: Multi-arm bandit. Source: towardsdatascience.com

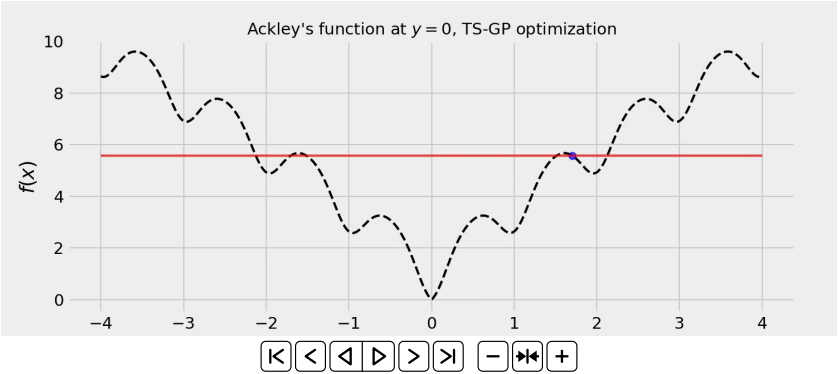
Bayesian optimization with Thompson sampling

- Thompson sampling posits that we make decisions according to the probability that they are optimal.
- That is, $\mathbf{x}_{n+1} \sim p_{\mathbf{x}_*}(\mathbf{x})$, where $p_{\mathbf{x}_*}(\mathbf{x})$ is the probability density of the maximizers of f.
- In the context of Bayesian optimization, we could use the surrogate Y to define this density as

$$p_{\mathbf{x}^*}(\mathbf{x}) = \int \delta\left(\mathbf{x} - \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}} \; Y(\mathbf{x})\right) p(Y|\boldsymbol{\mathcal{D}}_n)\, dY, \tag{1}$$

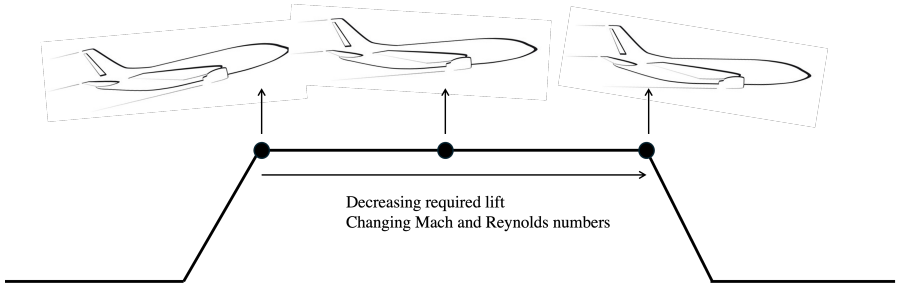
- Equivalently, draw a sample path from Y, find its maximizer, and repeat.

Bayesian optimization with Thompson sampling



Multiobjective aerodynamic design optimization

Aircraft aerodynamic design optimization



- Desining for a constant cruise condition can be suboptimal.
- Designing ***simultaneously for multiple points*** at cruise is the way to go.

AIAA Aircraft Design Optimization Discussion Group

Source: https://sites.google.com/view/mcgill-computational-aerogroup/adodg

Test Cases

1. Case 1: Drag Minimization of the NACA 0012 in Transonic Inviscid Flow.
2. Case 2: Drag Minimization of the RAE 2822 in Transonic Viscous Flow.
3. Case 3: Drag Minimization of a Rectangular Wing in Inviscid Subsonic Lifting Flow.
4. Case 4: Lift-Constrained Drag Minimization of the CRM Wing.
    1. Geometry File.
    2. Case 4.1: Lift-Constrained Drag Minimization of a Wing Allowing Section and Twist Variation.
    3. Case 4.2: Three-Point Wing Optimization to Minimize the Integral of C_D Over a Range of Aircraft Weights at a Fixed Mach Number and Altitude
    4. Case 4.3: Three-Point Wing Optimization to Minimize the Integral of C_D Over a Range of Mach numbers at a Fixed Aircraft Weight and Altitude
    5. Case 4.4: Three-Point Wing Optimization to Minimize the Integral of C_D Over a Range of Mach numbers at a Fixed Aircraft Weight with Altitude Varied to Maintain a Constant C_L Over The Range of Aircraft Mach Numbers
    6. Case 4.5: Nine-Point Wing Optimization to Minimize the Integral of C_D Over a Range of Aircraft Weights and Mach Numbers at a Fixed Altitude
5. Case 5: CRM Wing-Body-Tail Optimization at Flight Reynolds Number

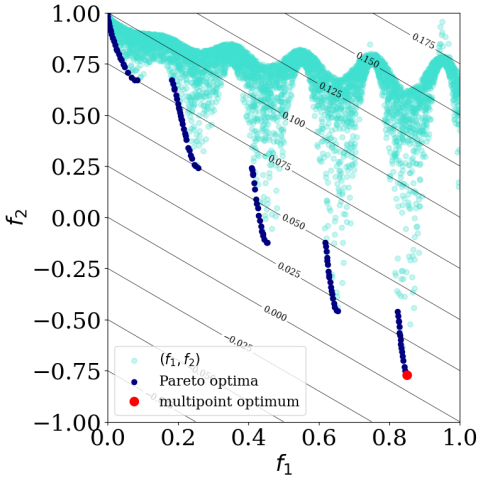
Multiobjective aerodynamic design optimization

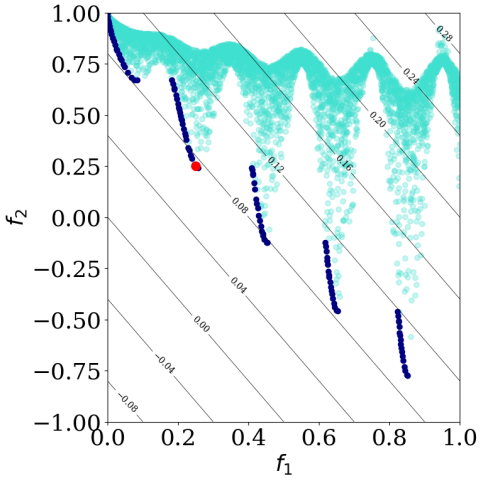
$$\max_{\mathbf{x}\in\mathcal{X}} \{-C_{D1}(\mathbf{x}), -C_{D2}(\mathbf{x})\} \tag{2}$$

- C_{D1} and C_{D2} are the aircraft drag at $M = 0.85,\ \alpha = 2.37^\circ$ and $M = 0.7,\ \alpha = 3.87^\circ$, respectively.
- $\mathbf{x}$ are the free-form deformation (FFD) control points and $\mathcal{X}$ is a compact domain.
- Euler simulations on the ***high speed NASA CRM***, via SU2 Economon et al. (2016), used to compute the aircraft drag.

In the absence of derivative information, no methods exist today to solve this problem in an efficeint and scalable manner.

Multiobjective ($\min\ \{f_1, f_2\}$) versus multi-point ($\min\ \mathbf{w}^\top \mathbf{f}$) optimization





Theoretical statements

Every multi-point optimum is Pareto optimal. Carlson and Renganathan (2025).

Let the weights in the multi-point optimum $w \in \mathbb{R}^m$ satisfy $w_i > 0$ for every $i = 1, \ldots, m$. If

$$\mathbf{x}^\star \in \arg\min_{\mathbf{x} \in \mathcal{X}} w^\top f(\mathbf{x}),$$

then $\mathbf{x}^\star$ is Pareto optimal.

Converse not true: Every Pareto optimum is not necessarily multi-point optimal.

If the attainable objective set F is convex, then for every Pareto optimal point $f^\star \in F$ there exists $w \in \mathbb{R}^m$ with $w_i \geq 0$, $w \neq 0$, such that $f^\star$ minimizes $w^\top y$ over $y \in F$.

- We aim to devise *sample efficient* and *principled* approaches for derivative-free **multiobjective optimization** with constraints.

Constrained multiobjective optimization

We consider the following problem

$$\begin{aligned} \max_{\mathbf{x}\in\mathcal{X}} \; & \{f_1(\mathbf{x}), \ldots, f_K(\mathbf{x})\} \\ \text{s.t. } & c_i(\mathbf{x}) = 0, \; i \in \mathcal{E} \\ & c_i(\mathbf{x}) \geq 0, \; i \in \mathcal{I} \end{aligned} \tag{3}$$

- $f_i, c_i, \; \forall i$ are expensive stochastic zeroth-order oracles.
- We are interested in the solution to (3) in a **sample efficient** manner and with **theoretical guarantees**.

Pareto optimality

- Single objective optimization: first-order necessary (Karush-Kuhn-Tucker) conditions Nocedal and Wright (1999).
- Mulitobjective optimization: **Pareto optimality** is necessary.
- Let $\boldsymbol{f}(\mathbf{x}) = [f_1(\mathbf{x}), \ldots, f_K(\mathbf{x})]^\top$. When a solution $\boldsymbol{f}(\mathbf{x})$ Pareto **dominates** another solution $\boldsymbol{f}(\mathbf{x}')$,
  - then $f_k(\mathbf{x}) \geq f_k(\mathbf{x}'),\ \forall k = 1, \ldots, K$ and there exists $k \in [K]$ such that $f_k(\mathbf{x}) > f_k(\mathbf{x}')$.
- The set $\mathcal{Y}^* = \{\boldsymbol{f}(\mathbf{x}) \ : \ \mathbf{x}' \in \mathcal{X} : \boldsymbol{f}(\mathbf{x}') \succ \boldsymbol{f}(\mathbf{x})\}$ is called the *Pareto frontier* and the set $\mathcal{X}^* = \{\mathbf{x} \in \mathcal{X} \ : \ \boldsymbol{f}(\mathbf{x}) \in \mathcal{Y}^*\}$ is called the *Pareto set*.
- The ***hypervolume indicator*** is used to measure Pareto optimality.

Hypervolume indicator

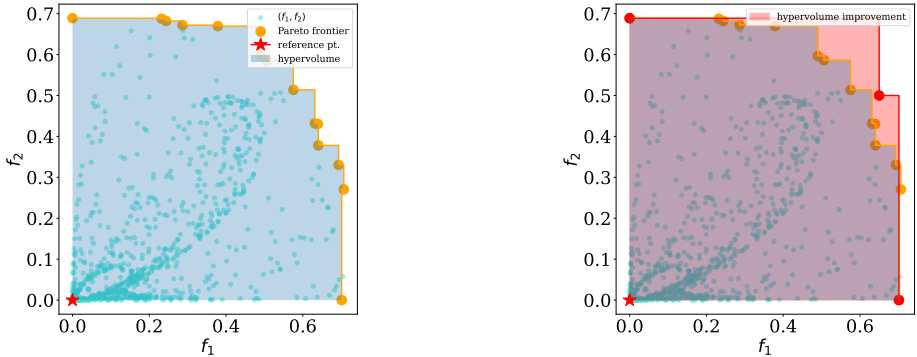


Figure 12: Illustrative two objective ($K = 2$) example.

Existing work

- De facto method is nondominated sorting genetic algorithm (NSGA) Deb et al. (2002, 2005).
- NSGA is **accurate** but **computationally expensive**.
- Bayesian optimization for multiobjective optimization = multiobjective Bayesian optimization (MOBO).
- Bayesian optimization is **less accurate** but **highly sample efficient**.
- Our goal: **get the best of both worlds!**



Existing work in MOBO

- **Scalarization.** Zhang et al. (2009); Zhou et al. (2012); Daulton et al. (2021); Paria et al. (2020)
- **Hypervolume based.** Couckuyt et al. (2014); Emmerich (2008); Yang et al. (2019); Daulton et al. (2020); Picheny (2015); Daulton et al. (2021)
- **Entropy based.** Hvarfner et al. (2022); Tu et al. (2022); Suzuki et al. (2020); Belakaria et al. (2019)

Existing work in MOBO suffers from these limitations:

- A hard (nonconvex, stochastic) acquisition function optimization.
- Inability to handle constraints/noisy objectives/*batch* acquisition.

We propose to address all of these limitations.



qPOTS: Batch Pareto optimal Thompson sampling

Renganathan and Carlson (2025) **AISTATS 2025**

- Our approach extends Thompson sampling Thompson (1933) to the multiobjective setting.
- We circumvent any hypervolume calculation or the intractable entropy calculation.
- Instead, we sample according to

$$p_{\mathcal{X}^*}(\mathbf{x}) = \int \delta\left(\mathbf{x} - \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}} \{Y_1(\mathbf{x}), \ldots, Y_K(\mathbf{x})\}\right) p(Y_1|\boldsymbol{\mathcal{D}}_n) \ldots p(Y_K|\boldsymbol{\mathcal{D}}_n) dY_1 \ldots dY_K, \tag{4}$$

- $p_{\mathcal{X}^*}(\mathbf{x})$ governs the distribution of the Pareto optimal solution set.

qPOTS: Batch Pareto optimal Thompson sampling

In practice,

- Fit K GP models: $\mathcal{M}_1, \ldots, \mathcal{M}_K$ for the K objectives, with observations $\boldsymbol{\mathcal{D}}_n^{1:K} \triangleq \{\boldsymbol{\mathcal{D}}_n^1, \ldots, \boldsymbol{\mathcal{D}}_n^K\}$.
- Draw K GP posterior sample paths (one from each GP).
  –
  $$Y(\cdot, \omega) = \mu_n(\cdot) + {\Sigma_n}^{1/2}(\cdot) \times Z(\omega),$$
  where Z is a standard multivariate normal random variable.
- Solve the following **cheap** multiobjective optimization problem using, e.g., **evolutionary approaches**
  $$\mathbf{x}_{n+1} \in X^* = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmax}} \{ Y_1(\mathbf{x}, \omega), \ldots, Y_K(\mathbf{x}, \omega) \}, \tag{5}$$

Essentially, we leverage the **best of both worlds**: accuracy of evolutionary approaches and sample-efficiency of MOBO.

qPOTS: Batch Pareto optimal Thompson sampling

- **Balancing exploration and exploitation.** We choose $\mathbf{x}_{n+1} \in X^*$ that maximize the minimum distance to the previously observed sites X_n.
- Let $\gamma(\mathbf{x}, \mathbf{x}') : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ denote the Euclidean distance between $\mathbf{x}$ and $\mathbf{x}'$, then the maximin point is defined as

$$\mathbf{x}_{n+1} \in \underset{\mathbf{x}^* \in X^*}{\operatorname{argmax}} \min_{\mathbf{x}_i \in X_n} \gamma(\mathbf{x}^*, \mathbf{x}_i).$$

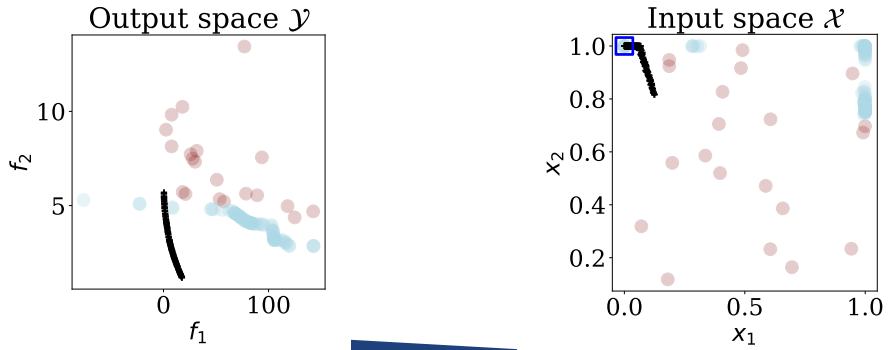
- Easy extension to **batch acquisition**:

$$\begin{aligned}
\mathbf{x}_{n+1} &\in \underset{\mathbf{x}^* \in X^*}{\operatorname{argmax}} \min_{\mathbf{x}_i \in X_n} \gamma(\mathbf{x}^*, \mathbf{x}_i) \\
\mathbf{x}_{n+2} &\in \underset{\mathbf{x}^* \in X^*}{\operatorname{argmax}} \min_{\mathbf{x}_i \in X_n \bigcup \mathbf{x}_{n+1}} \gamma(\mathbf{x}^*, \mathbf{x}_i) \\
&\vdots \\
\mathbf{x}_{n+q} &\in \underset{\mathbf{x}^* \in X^*}{\operatorname{argmax}} \min_{\mathbf{x}_i \in X_n \bigcup \{\mathbf{x}_{n+1}, \ldots, \mathbf{x}_{n+q-1}\}} \gamma(\mathbf{x}^*, \mathbf{x}_i).
\end{aligned} \tag{6}$$

qPOTS demo

Two objective $K = 2$ Branin-Currin function $(n = 20)$.

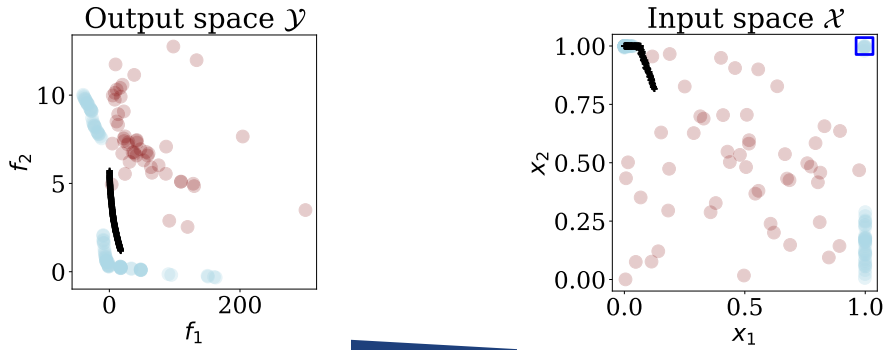
Figure 13: Maroon: f_i observations, blue circles: GP posterior $\mathcal{Y}^*/\mathcal{X}^*$, black: *true* $\mathcal{Y}^*/\mathcal{X}^*$, blue square: qPOTS acquisition.



qPOTS demo

Two objective $K = 2$ Branin-Currin function $(n = 50)$.

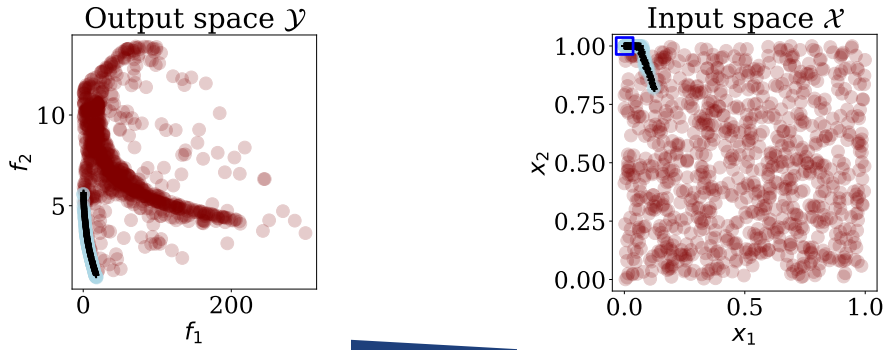
Figure 14: Maroon: f_i observations, blue circles: GP posterior $\mathcal{Y}^*/\mathcal{X}^*$, black: *true* $\mathcal{Y}^*/\mathcal{X}^*$, blue square: qPOTS acquisition.



qPOTS demo

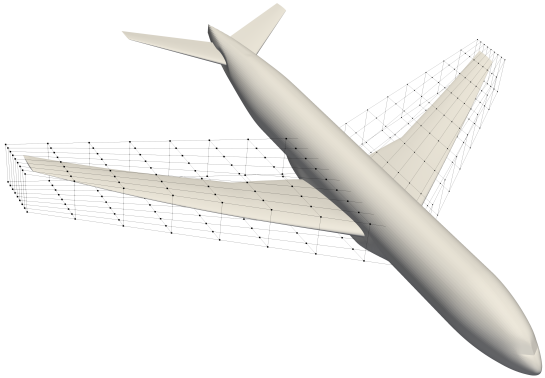
Two objective $K = 2$ Branin-Currin function $(n = 900)$.

Figure 15: Maroon: f_i observations, blue circles: GP posterior $\mathcal{Y}^*/\mathcal{X}^*$, black: *true* $\mathcal{Y}^*/\mathcal{X}^*$, blue square: qPOTS acquisition.

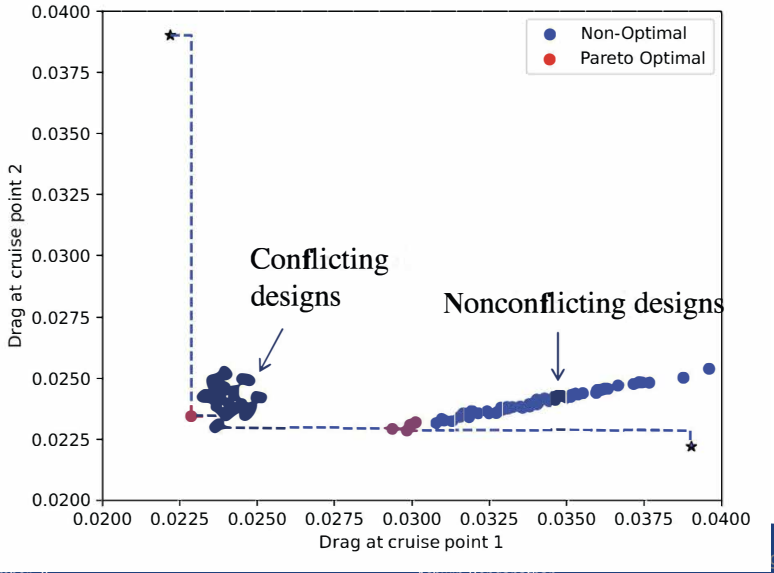


Multiobjective aerodynamic design optimization of the NASA CRM

Carlson and Renganathan (2025) – (in Aerospace Science and Tech.)

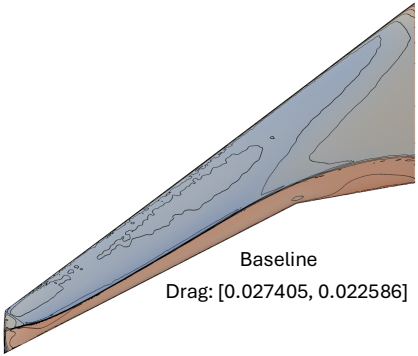


Design space



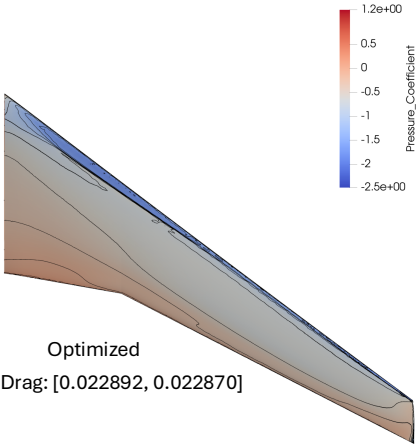
Pareto optimal design 1

Baseline



Baseline
Drag: [0.027405, 0.022586]

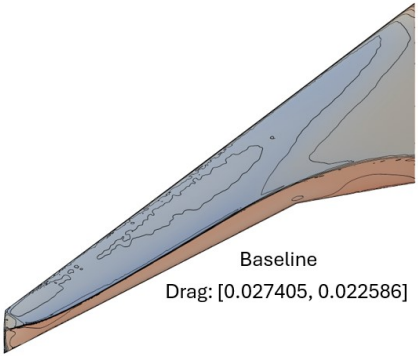
Pareto optimized



Optimized
Drag: [0.022892, 0.022870]

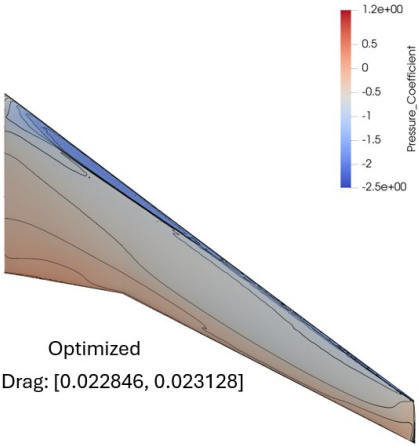
Pareto optimal design 2

Baseline



Baseline
Drag: [0.027405, 0.022586]

Pareto optimized

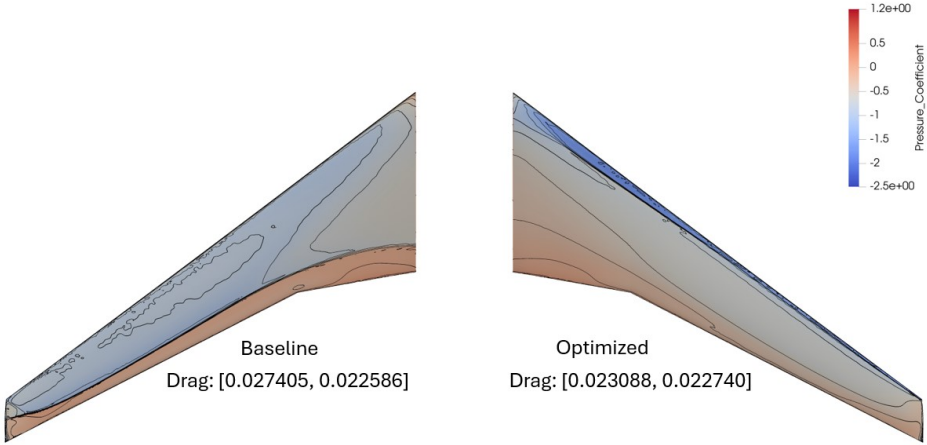


Optimized
Drag: [0.022846, 0.023128]

Pareto optimal design 3

Baseline

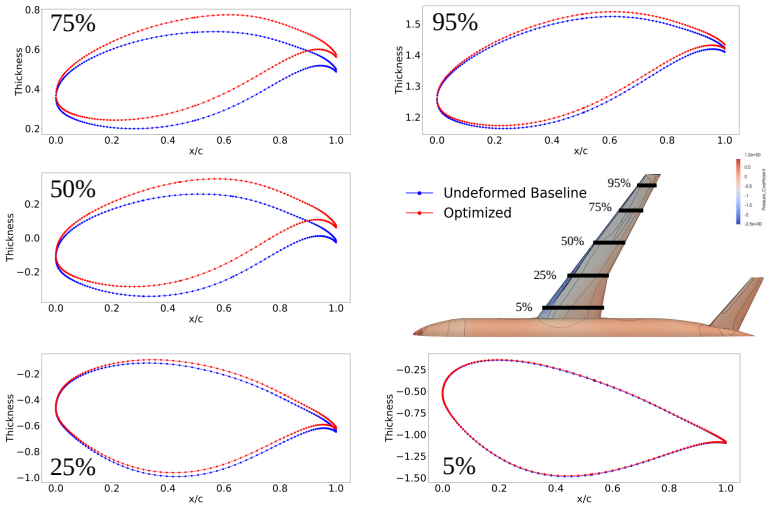
Pareto optimized



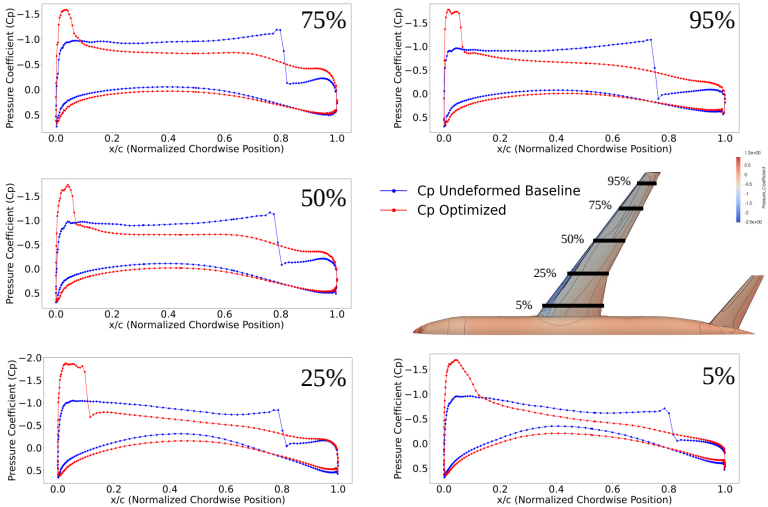
Baseline
Drag: [0.027405, 0.022586]

Optimized
Drag: [0.023088, 0.022740]

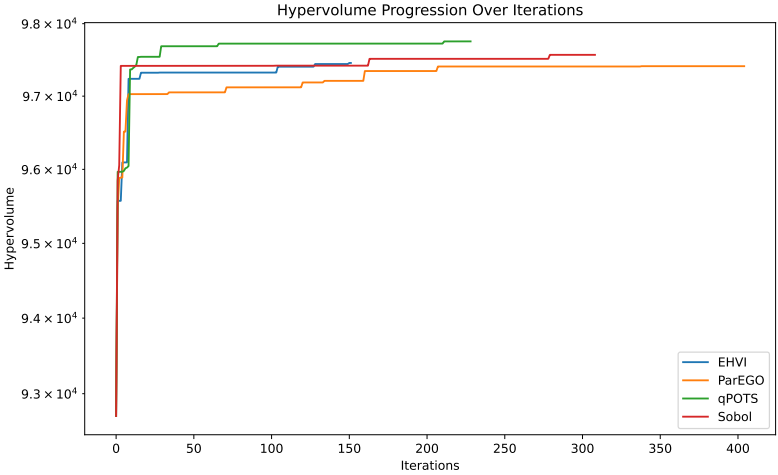
Pareto optimal wing shapes



Pareto optimal pressure distributions



Hypervolume plot (higher the better)



Summary

- qPOTS is a fast, accurate, derivative-free, and sample efficient multiobjective optimization method improving on state of the art.
- Strong empirical performance with theoretical guarantee.
- Multiobjective optimization of the NASA CRM:
  - 288 cores, 3 days of wall clock time.
  - Outperforms all competing methods.
- Read our full paper for details on other experiments.



Software

```
>> pip install qpots
```

---

csdlpsu/**qpots**

Batch Pareto optimal Thompson sampling



2 Contributors | 0 Issues | 8 Stars | 2 Forks



Conclusions

Conclusions

- CSDL develops scalable computational techniques toward the decision-making for complex engineered systems.
- Foundational contributions to the *mathematics* of computational design.
- Novel methods, algorithms, and software toward the **decision-making triad:** surrogate modeling, uncertainty quantification, and optimization.
- We predominantly work with **probabilistic machine learning** techniques – kernel methods with Bayesian decision theory.
- Follow our work at `sites.psu.edu/csdl`.

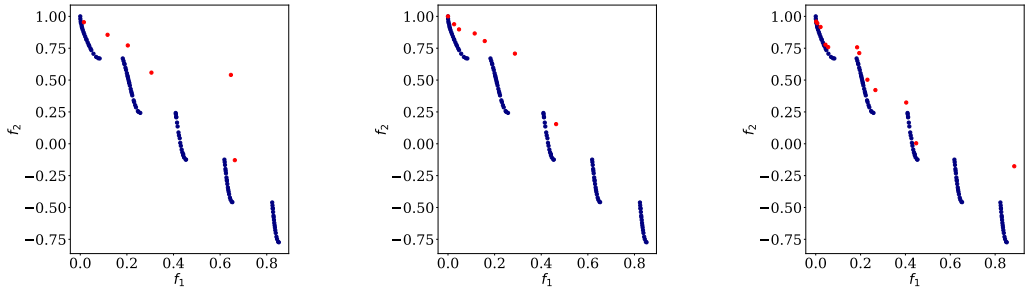


Thank you

(a) Paper

(b) Software

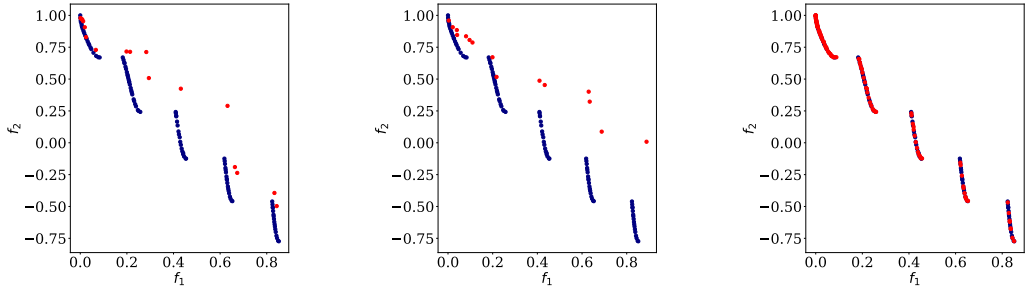
A quick demo (ZDT3; $K = 2$)



(a) qNEHVI (b) qNPAREGO (c) Sobol

Figure 18: Demonstration on the ZDT3 ($K = 2$) test function that has disjoint Pareto frontiers. Navy are the true Pareto frontier and red are the predicted Pareto frontier. With 20 seed points and only an additional 204 acquisitions, in batches of $q = 4$, notice that qPOTS is quickly able to resolve the true Pareto frontier while other methods in the state of the art struggle to come any close.

A quick demo (ZDT3; $K = 2$)



(a) PESMO (b) MESMO (c) qPOTS (Ours)

Figure 19: Demonstration on the ZDT3 ($K = 2$) test function that has disjoint Pareto frontiers. Navy are the true Pareto frontier and red are the predicted Pareto frontier. With 20 seed points and only an additional 204 acquisitions, in batches of $q = 4$, notice that qPOTS is quickly able to resolve the true Pareto frontier while other methods in the state of the art struggle to come any close.

Extension to constrained problems

Handing constraints is straightforward in qPOTS.

- Let there be a total of C constraints $\{c_1, \ldots, c_C\}$, and $\mathcal{I} \cup \mathcal{E} = [C]$.
- We fit a total of $K + C$ posterior GPs $Y_1, \ldots, Y_K, Y_{K+1}, \ldots, Y_{K+C}$.
- For all $i \in [C]$, we define the constrained qPOTS to choose points according to

$$p_{\mathcal{X}^*}(\mathbf{x}) = \int \delta\left(\mathbf{x} - \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmax}} \left[\{Y_1(\mathbf{x}), \ldots, Y_K(\mathbf{x})\} \times \mathbb{1}_{\{\bigcap_i Y_{k+i\in\mathcal{E}}(\mathbf{x})=0\}} \times \mathbb{1}_{\{\bigcap_i Y_{k+i\in\mathcal{I}}(\mathbf{x})\geq 0\}}\right]\right) p(Y_1|\boldsymbol{\mathcal{D}}_n) \ldots p(Y_K|\boldsymbol{\mathcal{D}}_n) dY_1 \ldots dY_K, \tag{7}$$

Constrained qPOTS demo

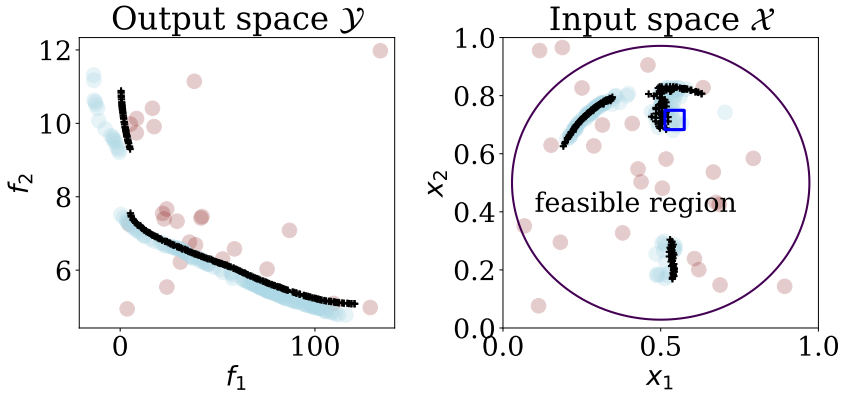


Figure 20: Illustration of qPOTS on the constrained Branin-Currin with $C = 1$ nonlinear constraint. Black circle on the right is constraint boundary and region inside that is feasible. Notice that qPOTS chooses a new candidate from the *constrained* random Pareto set $\mathcal{X}^*$

qPOTS theoretical guarantee

Theorem (Convergence with high probability)

Let $\delta \in (0,1)$ and Assumption 1 hold with $L^2 = 2\log\left(\frac{2}{\delta}\right)$. Further, let g be some affine function that operates on f. Let $\tilde{\mathcal{X}}$ denote a discretization of $\mathcal{X}$ with $|\tilde{\mathcal{X}}| < \infty$, where we evaluate the oracle f. Let $\epsilon \geq 0$, and $a_n = \sqrt{2\log(b_n|\tilde{\mathcal{X}}|)/\delta}$, where $\sum_{n=1}^{\infty} 1/b_n = 1$. Then the vector of posterior GP objectives converges to $\mathbf{f}$ with high probability; that is, the following holds:

$$\mathbb{P}r\left[\forall \mathbf{x} \in \mathcal{X},\ \lim_{n\to\infty} \|\mathbf{f}(\mathbf{x}) - \mathbf{Y}^{n-1}(\mathbf{x})\|_\infty \leq \epsilon\right] > 1 - \delta.$$

Our approach, qPOTS, is **guaranteed** to find the true Pareto frontier under mild assumptions.



References I

Belakaria, S., Deshwal, A., and Doppa, J. R. (2019). Max-value Entropy Search for Multi-Objective Bayesian Optimization. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

Bloebaum, C. L. and McGowan, A.-M. R. (2012). The design of large-scale complex engineered systems: present challenges and future promise. In *12th aiaa atio conference and 14th aiaa/issmo ma&o conference, aiaa paper*, number 2012-5571.

Campbell, C. (2001). An introduction to kernel methods. *Studies in Fuzziness and Soft Computing*, 66:155–192.

Carlson, K. and Renganathan, A. (2025). Multiobjective aerodynamic design optimization of the nasa common research model. *Aerospace Science and Technology*, page 111120.

Couckuyt, I., Deschrijver, D., and Dhaene, T. (2014). Fast calculation of multiobjective probability of improvement and expected improvement criteria for Pareto optimization. *Journal of Global Optimization*, 60(3):575–594.

Daulton, S., Balandat, M., and Bakshy, E. (2020). Differentiable expected hypervolume improvement for parallel multi-objective bayesian optimization. *Advances in Neural Information Processing Systems*, 33:9851–9864.

Daulton, S., Balandat, M., and Bakshy, E. (2021). Parallel Bayesian Optimization of Multiple Noisy Objectives with Expected Hypervolume Improvement. In *Advances in Neural Information Processing Systems*, volume 34, pages 2187–2200. Curran Associates, Inc.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197.

Deb, K., Thiele, L., Laumanns, M., and Zitzler, E. (2005). Scalable test problems for evolutionary multiobjective optimization. In *Evolutionary multiobjective optimization: theoretical advances and applications*, pages 105–145. Springer.

Economon, T. D., Palacios, F., Copeland, S. R., Lukaczyk, T. W., and Alonso, J. J. (2016). SU2: An open-source suite for multiphysics simulation and design. *Aiaa Journal*, 54(3):828–846.

Emmerich, M. (2008). The computation of the expected improvement in dominated hypervolume of Pareto front approximations.

Frazier, P. I., Powell, W. B., and Dayanik, S. (2008). A knowledge-gradient policy for sequential information collection. *SIAM Journal on Control and Optimization*, 47(5):2410–2439.

Hennig, P. and Schuler, C. J. (2012). Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13(Jun):1809–1837.

Hernández-Lobato, J. M., Hoffman, M. W., and Ghahramani, Z. (2014). Predictive entropy search for efficient global optimization of black-box functions. In *Advances in Neural Information Processing Systems*, pages 918–926.

References II

Hofmann, T., Schölkopf, B., and Smola, A. J. (2008). Kernel methods in machine learning.

Hvarfner, C., Hutter, F., and Nardi, L. (2022). Joint entropy search for maximally-informed bayesian optimization. *Advances in Neural Information Processing Systems*, 35:11494–11506.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492.

Kandasamy, K., Krishnamurthy, A., Schneider, J., and Póczos, B. (2018). Parallelised bayesian optimisation via thompson sampling. In *International conference on artificial intelligence and statistics*, pages 133–142. PMLR.

Mauery, T., Alonso, J., Cary, A., Lee, V., Malecki, R., Mavriplis, D., Medic, G., Schaefer, J., and Slotnick, J. (2021). A guide for aircraft certification by analysis. Technical report.

Mockus, J., Tiešis, V., and Žilinskas, A. (1978). The application of Bayesian methods for seeking the extremum. *Towards Global Optimization*, 2:117–129.

Nocedal, J. and Wright, S. J. (1999). *Numerical optimization*. Springer.

Paria, B., Kandasamy, K., and Póczos, B. (2020). A flexible framework for multi-objective bayesian optimization using random scalarizations. In *Uncertainty in Artificial Intelligence*, pages 766–776. PMLR.

Picheny, V. (2014). A stepwise uncertainty reduction approach to constrained global optimization. In *Artificial intelligence and statistics*, pages 787–795. PMLR.

Picheny, V. (2015). Multiobjective optimization using gaussian process emulators via stepwise uncertainty reduction. *Statistics and Computing*, 25(6):1265–1280.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.

Renganathan, A. and Carlson, K. (2025). qpots: Efficient batch multiobjective bayesian optimization via pareto optimal thompson sampling. In *International Conference on Artificial Intelligence and Statistics*, pages 4051–4059. PMLR.

Sauer, A., Gramacy, R. B., and Higdon, D. (2023). Active learning for deep gaussian process surrogates. *Technometrics*, 65(1):4–18.

Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. (2009). Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the International Conference on Machine Learning*.

Suzuki, S., Takeno, S., Tamura, T., Shitara, K., and Karasuyama, M. (2020). Multi-objective bayesian optimization using pareto-frontier entropy. In *International Conference on Machine Learning*, pages 9279–9288. PMLR.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294.

References III

Tu, B., Gandy, A., Kantas, N., and Shafei, B. (2022). Joint entropy search for multi-objective bayesian optimization. *Advances in Neural Information Processing Systems*, 35:9922–9938.

Wang, Z. and Jegelka, S. (2017). Max-value entropy search for efficient Bayesian optimization. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pages 3627–3635.

Yang, K., Emmerich, M., Deutz, A., and Bäck, T. (2019). Multi-objective bayesian global optimization using expected hypervolume improvement gradient. *Swarm and evolutionary computation*, 44:945–956.

Zhang, Q., Liu, W., Tsang, E., and Virginas, B. (2009). Expensive multiobjective optimization by moea/d with gaussian process model. *IEEE Transactions on Evolutionary Computation*, 14(3):456–474.

Zhou, A., Zhang, Q., and Zhang, G. (2012). A multiobjective evolutionary algorithm based on decomposition and probability model. In *2012 IEEE Congress on Evolutionary Computation*, pages 1–8. IEEE.