Adaptive Surrogate Modeling for Simulation and Optimization of Dynamical Systems with Model Inexactness

Matthias Heinkenschloss

Department of Computational Applied Mathematics and Operations Research
Rice University, Houston, Texas
heinken@rice.edu

April 17, 2025

Joint work with Jonathan Cangelosi and John Steinman
Supported in part by AFOSR Grant FA9550-22-1-0004.

East Coast Optimization Meeting 2025
'Optimization and Digital Twins'
April 17 - 18, 2025
George Mason University, Arlington, Virginia (USA)

Outline

Motivation

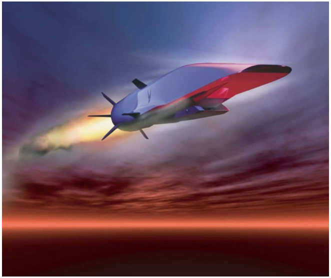
- From the 2024 National Academies Report: Foundational Research Gaps and Future Directions for Digital Twins

  'A digital twin is a set of virtual information constructs that mimics the structure, context, and behavior of a natural, engineered, or social system (or system of systems), is dynamically updated with data from its physical twin, has a predictive capability, and informs decisions that realize value. The bidirectional interaction between the virtual and the physical is central to the digital twin.'

- Consider a dynamical system twin.
  - Used to compute optimal controls for physical twin,
  - Dynamical system depends on quantities that are *estimated from physical twin data*.
  - What data are needed to reduce impact of model error?
  - Subproblem within digital twin contruction/update.

Motivation

- Dynamical systems in many applications depend on quantities that are *only known experimentally* or are *expensive to compute*.
  - Lift/drag/moment coefficients of a hypersonic vehicle.
  - Relative permeability and mobility in subsurface flows.
  - Constitutive laws.



Concept art for a Boeing X51 waverider vehicle.
Credit: U.S. Air Force. Accessed 3/27/2025.
`https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104467/x-51a-waverider/`

Motivation

- Cost of experiments makes many-query tasks like simulation and optimization impractical or intractable.
- Can approximate expensive (or high-fidelity) models by inexpensive surrogate models, but this introduces inexactness.
- Relationship between surrogate accuracy and solution accuracy is *problem-dependent*.
- Surrogate accuracy is *adjustable*, e.g., by interpolating high-fidelity model at user-specified points.
- Goal: select interpolation points such that the resulting surrogate yields an accurate solution.

Motivation

- Why not just pick a bunch of interpolation points?
  - Requires many expensive high-fidelity computations.
  - Can cause numerical issues in the surrogate.
  - Can make the surrogate itself too expensive.

- To avoid these issues, must prioritize “good” interpolation points, i.e., points where model error has greatest impact on trajectory.

- Key idea: define acquisition function that combines sensitivity information from simulation/optimization problem with surrogate error bounds to determine “good” interpolation points.

Literature review

- Many acquisition functions for adaptive surrogate-assisted optimization using Gaussian process (GP) surrogates:
  - Expected improvement [Jones et al., 1998]
  - U-function [Moustapha et al., 2016]
  - Gradient-enhanced inspection of local minima [Surmann et al., 2021]
  - Sobol sensitivity indices [Vohra et al., 2019]
  - Crowding distance [Sun et al., 2022]
  - Maximum posterior variance [Cangelosi et al., 2024]
- Existing approaches are mostly statistical/Bayesian + parametric.
- First sensitivity-driven + deterministic + parametric approach given in [Hart et al., 2023].
- ASMR approach is sensitivity-driven + deterministic + parameter-free and extends [Hart et al., 2023].

Outline

Problem formulation

- Consider the initial value problem

$$\mathbf{x}'(t) = \mathbf{f}\Big(t, \mathbf{x}(t), \mathbf{g}\big(t, \mathbf{x}(t)\big)\Big), \qquad \text{a.a. } t \in I,$$
$$\mathbf{x}(t_0) = x_0$$

  where $I := (t_0, t_f)$.

- Solution $\mathbf{x} : \overline{I} \to \mathbb{R}$ depends on *state-dependent* component function $\mathbf{g} : I \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_g}$.

- Let $\mathbf{g}_*$ be the "true" model, and let $\widehat{\mathbf{g}}$ be a surrogate.

- Computing $\mathbf{x}_* := \mathbf{x}(\cdot\,; \mathbf{g}_*)$ too expensive; can only compute $\widehat{\mathbf{x}} := \mathbf{x}(\cdot\,; \widehat{\mathbf{g}})$.

- Can estimate solution error using sensitivity of solution mapping $\mathbf{x}(\cdot\,; \mathbf{g})$ to perturbations in $\mathbf{g}$.

Differentiability of the dynamics

- To compute sensitivity of $\mathbf{x}(\cdot\,;\mathbf{g})$, need to set problem in appropriate function spaces and establish continuous Fréchet differentiability.
- Superposition/Nemytskii operator

$$\mathbf{F} : \big(L^\infty(I)\big)^{n_x} \times \big(\mathcal{G}^2(I)\big)^{n_g} \to \big(L^\infty(I)\big)^{n_x}$$

representing the right-hand side

$$\mathbf{F}(\mathbf{x},\mathbf{g}) := \mathbf{f}\Big(\cdot, \mathbf{x}(\cdot), \mathbf{g}\big(\cdot, \mathbf{x}(\cdot)\big)\Big)$$

is continuous Fréchet diff'ble under suitable conditions on $\mathbf{f}$, where

$$\begin{aligned}\big(\mathcal{G}^2(I)\big)^{n_g} := \big\{\mathbf{g} : I \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_g} \ :\ & \mathbf{g}(t,x) \text{ is twice continuously} \\ & \text{partially differentiable with respect to } x \in \mathbb{R}^{n_x} \text{ for a.a. } t \in I, \\ & \text{is measurable in } t \text{ for each } x \in \mathbb{R}^{n_x}, \text{ and } \|\mathbf{g}\|_{(\mathcal{G}^2(I))^{n_g}} < \infty\big\},\end{aligned}$$

$$\|\mathbf{g}\|_{(\mathcal{G}^2(I))^{n_g}} := \sum_{n=0}^{2} \operatorname*{ess\,sup}_{t \in I} \sup_{x \in \mathbb{R}^{n_x}} \left\| \frac{\partial^n}{\partial x^n} \mathbf{g}(t,x) \right\|.$$

[Cangelosi and Heinkenschloss, 2024]

Sensitivity of ODE solution

- Use Implicit Function Theorem to establish continuous Fréchet differentiability of solution mapping

$$\big(\mathcal{G}^2(I)\big)^{n_g} \ni \mathbf{g} \mapsto \mathbf{x}(\cdot\,;\mathbf{g}) \in \big(W^{1,\infty}(I)\big)^{n_x}$$

under suitable assumptions on $\mathbf{f}$.

- Sensitivity $\delta\mathbf{x} := \mathbf{x}_{\mathbf{g}}(\overline{\mathbf{g}})\delta\mathbf{g}$ given by solution of linear IVP

$$\begin{aligned} &\delta\mathbf{x}'(t) = \overline{\mathbf{A}}(t)\delta\mathbf{x}(t) + \overline{\mathbf{B}}(t)\delta\mathbf{g}\big(t, \overline{\mathbf{x}}(t)\big), && \text{a.a. } t \in I, \\ &\delta\mathbf{x}(t_0) = \mathbf{0} \end{aligned}$$

where $\overline{\mathbf{x}}(\cdot) := \mathbf{x}(\cdot\,;\overline{\mathbf{g}})$ and

$$\begin{aligned} \overline{\mathbf{A}}(\cdot) &:= \mathbf{f}_x\Big(\cdot, \overline{\mathbf{x}}(\cdot), \overline{\mathbf{g}}\big(\cdot, \overline{\mathbf{x}}(\cdot)\big)\Big) + \mathbf{f}_g\Big(\cdot, \overline{\mathbf{x}}(\cdot), \overline{\mathbf{g}}\big(\cdot, \overline{\mathbf{x}}(\cdot)\big)\Big)\overline{\mathbf{g}}_x\big(\cdot, \overline{\mathbf{x}}(\cdot)\big), \\ \overline{\mathbf{B}}(\cdot) &:= \mathbf{f}_g\Big(\cdot, \overline{\mathbf{x}}(\cdot), \overline{\mathbf{g}}\big(\cdot, \overline{\mathbf{x}}(\cdot)\big)\Big). \end{aligned}$$

Sensitivity of QoI

- Quantity of interest (QoI) that depends on the ODE solution $\mathbf{x}(\cdot\,;\mathbf{g})$:

$$q(\mathbf{x},\mathbf{g}) := \varphi\big(\mathbf{x}(t_f)\big) + \int_{t_0}^{t_f} l\Big(t, \mathbf{x}(t), \mathbf{g}\big(t, \mathbf{x}(t)\big)\Big)\, dt,$$

$$\widetilde{q}(\mathbf{g}) := q\big(\mathbf{x}(\cdot\,;\mathbf{g}), \mathbf{g}\big).$$

- QoI

$$\big(\mathcal{G}^2(I)\big)^{n_g} \ni \mathbf{g} \mapsto \widetilde{q}(\mathbf{g}) \in \mathbb{R}$$

is continuous Fréchet diff'ble under suitable assumptions on φ, l, $\mathbf{f}$. [Cangelosi and Heinkenschloss, 2024]

Sensitivity of QoI

- Shorthand: $\overline{\mathbf{x}}(\cdot) := \mathbf{x}(\cdot\,;\overline{\mathbf{g}})$, $\qquad \bar{l}[\cdot] := l\Big(\cdot, \overline{\mathbf{x}}(\cdot), \overline{\mathbf{g}}\big(\cdot, \overline{\mathbf{x}}(\cdot)\big)\Big)$
- QoI sensitivity $\delta q := \widetilde{q}_{\mathbf{g}}(\overline{\mathbf{g}})\delta\mathbf{g}$ given by

$$\int_{t_0}^{t_f} \Big[\overline{\mathbf{B}}(t)^T \overline{\boldsymbol{\lambda}}(t) + \nabla_g \bar{l}[t]\Big]^T \delta\mathbf{g}\big(t, \overline{\mathbf{x}}(t)\big)\, dt$$

  where $\overline{\boldsymbol{\lambda}} \in \big(W^{1,\infty}(I)\big)^{n_x}$ solves the linear adjoint equation

$$-\overline{\boldsymbol{\lambda}}'(t) = \overline{\mathbf{A}}(t)^T \overline{\boldsymbol{\lambda}}(t) + \nabla_x \bar{l}[t] + \overline{\mathbf{g}}_x\big(t, \overline{\mathbf{x}}(t)\big)^T \nabla_g \bar{l}[t], \quad \text{a.a. } t \in I,$$
$$\overline{\boldsymbol{\lambda}}(t_f) = \nabla_x \varphi\big(\overline{\mathbf{x}}(t_f)\big)$$

  with $\overline{\mathbf{A}}$, $\overline{\mathbf{B}}$ from sensitivity IVP.

- After one linear adjoint ODE solve, can compute for any $\delta\mathbf{g}$.

Classical Perturbation Analysis

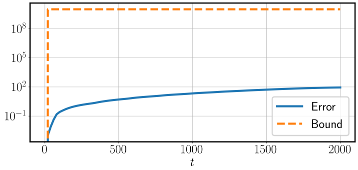
- Given $\mathbf{f} \in \mathcal{C}^1(\mathbb{R}^{n_x} \times \mathbb{R}^{n_g}, \mathbb{R}^{n_x})$, $\mathbf{g} \in \mathcal{C}^1(\mathbb{R}^{n_x}, \mathbb{R}^{n_g})$, consider ODE

$$\begin{aligned} \mathbf{x}'(t) &= \mathbf{f}\Big(t, \mathbf{x}(t), \mathbf{g}\big(t, \mathbf{x}(t)\big)\Big), \qquad\qquad t \in (t_0, t_f), \\ \mathbf{x}(t_0) &= \mathbf{x}_0. \end{aligned}$$

- $\mathbf{f}$ is known, $\mathbf{g}$ is approximated.
  True model $\mathbf{g}_*$ expensive to compute, or may not be exactly known.
  Can only compute with current surrogate model $\mathbf{g}_c$.

- ODE solution $\mathbf{x}(\cdot, \mathbf{g})$. Error bound for $\mathbf{x}(\cdot, \mathbf{g}_*) - \mathbf{x}(\cdot, \mathbf{g}_c)$?

- Existing ODE perturbation results, e.g., [Hairer et al., 1993], [Söderlind, 2006], extremely pessimistic (bound capped at 10^{10}). Applied to ODE for hypersonic vehicle simulation (specified later)



New Sensitivity-Based Error Bound

- Have proven Fréchet differentiability of $\mathbf{g} \mapsto \mathbf{x}(\cdot, \mathbf{g})$.
- Use sensitivity $\mathbf{x}(\mathbf{g}_c) - \mathbf{x}(\mathbf{g}_*) \approx \mathbf{x}_\mathbf{g}(\mathbf{g}_c)(\mathbf{g}_c - \mathbf{g}_*)$ and bound on model error along current trajectory

$$\left|\mathbf{g}_c(\mathbf{x}_c(t)) - \mathbf{g}_*(\mathbf{x}_c(t))\right| \leq c(\mathbf{g}_*)\,\boldsymbol{\epsilon}(\mathbf{x}_c(t)).$$

- Upper bound for weighted L^2-norm error computed via solution of linear-quadratic optimal control problem (LQOCP)

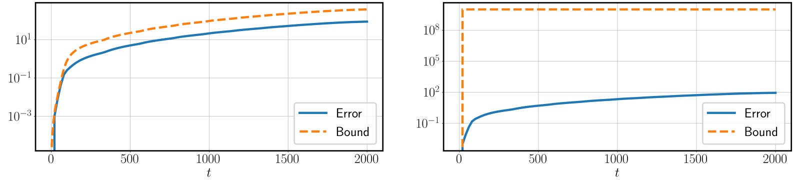
$$\begin{aligned}
\max_{\boldsymbol{\delta},\,\delta\mathbf{x}} \quad & \int_I \delta\mathbf{x}(t)^T \mathbf{Q}(t)\delta\mathbf{x}(t)\,dt && \text{(error norm)}\\
\text{s.t.} \quad & \delta\mathbf{x}'(t) = \mathbf{A}_c(t)\delta\mathbf{x}(t) + \mathbf{B}_c(t)\boldsymbol{\delta}(t), && \text{a.a. } t \in I, \text{ (sens. eqn.)}\\
& \delta\mathbf{x}(t_0) = \mathbf{0}, &&\\
& -\boldsymbol{\epsilon}\big(\mathbf{x}_c(t)\big) \leq \boldsymbol{\delta}(t) \leq \boldsymbol{\epsilon}\big(\mathbf{x}_c(t)\big), && \text{a.a. } t \in I, \text{ (model err. bd.)}
\end{aligned}$$

where

$$\mathbf{A}_c(\cdot) := \mathbf{f}_x\Big(\mathbf{x}_c(\cdot), \mathbf{g}_c\big(\mathbf{x}_c(\cdot)\big)\Big) + \mathbf{f}_g\Big(\mathbf{x}_c(\cdot), \mathbf{g}_c\big(\mathbf{x}_c(\cdot)\big)\Big)(\mathbf{g}_c)_x\big(\mathbf{x}_c(\cdot)\big),$$

$$\mathbf{B}_c(\cdot) := \mathbf{f}_g\Big(\mathbf{x}_c(\cdot), \mathbf{g}_c\big(\mathbf{x}_c(\cdot)\big)\Big).$$

New Sensitivity-Based Bound



- Sensitivity-based error bound tight (left) even when traditional ODE perturbation bound fails (right) (bound is capped at 10^{10}).

- Adjoint equation approach to avoid LQOCP when computing error bounds for quantity of interest $\widehat{q}(\mathbf{g}) := q\big(\mathbf{x}(\cdot;\mathbf{g})\big)$.

- Can use these bounds
  - to assess quality of current model $\mathbf{g}_c$ and, if needed,
  - to improve the model.

- Next construct $\mathbf{g}_c$ models using reproducing Kernel Hilbert spaces (RKHSs) and use new sensitivity-based bound to guide refinement.

Reproducing kernel Hilbert spaces (RKHSs)

- Surrogate models with pointwise error bounds may be constructed via standard kernel interpolation/optimal recovery methods. Similar to GP regression, but from a deterministic viewpoint. [**?**], [**?**]
- RKHS with kernel $\mathbf{k} : \Omega \times \Omega \to \mathbb{R}$, denoted $\mathcal{H}_{\mathbf{k}}(\Omega)$, is completion of

$$\operatorname{span}\{\mathbf{k}(y, \cdot) \mid y \in \Omega\}$$

with respect to the inner product

$$\left\langle \sum_{i=1}^{n} \alpha_i \, \mathbf{k}(y_i, \cdot), \, \sum_{j=1}^{m} \beta_j \, \mathbf{k}(y_j, \cdot) \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \mathbf{k}(y_i, y_j).$$

- Let $\mathbf{g}_* \in \mathcal{H}_{\mathbf{k}}(\Omega)$ with known evaluations

$$Y := (y_1, \ldots, y_N), \qquad G := \big(\mathbf{g}_*(y_1), \ldots, \mathbf{g}_*(y_N)\big).$$

- Want an interpolant of these points in $\mathcal{H}_{\mathbf{k}}(\Omega)$ and a pointwise error bound for the interpolant.

Kernel interpolation

- Kernel interpolation/optimal recovery problem:

$$\begin{aligned} \min \quad & \|\mathbf{g}\|_{\mathcal{H}_{\mathbf{k}}(\Omega)} \\ \text{s.t.} \quad & \mathbf{g}(y_i) = \mathbf{g}_*(y_i), \qquad\qquad i = 1, \dots, N. \end{aligned}$$

- Optimal solution given by

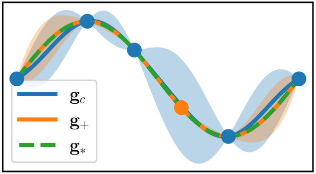
$$\mathbf{g}(y) = \mathbf{k}(y, Y)\, \mathbf{k}(Y, Y)^{-1}\, G$$

  where $\mathbf{k}(y, Y) := \begin{pmatrix} \mathbf{k}(y, y_1) & \cdots & \mathbf{k}(y, y_N) \end{pmatrix}$.
- Pointwise error bound given by

$$\begin{aligned} |\mathbf{g}(y) - \mathbf{g}_*(y)| &\le \|\mathbf{g}_*\|_{\mathcal{H}_{\mathbf{k}}(\Omega)}\, P(y; Y), \\ P(y; Y) &= \sqrt{\mathbf{k}(y, y) - \mathbf{k}(y, Y)\, \mathbf{k}(Y, Y)^{-1}\, \mathbf{k}(Y, y)}. \end{aligned}$$

- Note: error bound does not depend on outputs of $\mathbf{g}_*$, only the norm of $\mathbf{g}_*$ and its inputs.
- $\|\mathbf{g}_*\|_{\mathcal{H}_{\mathbf{k}}(\Omega)}$ not known in practice; must estimate somehow. Can use $\|\mathbf{g}\|_{\mathcal{H}_{\mathbf{k}}(\Omega)}$, but it will be an underestimate.

Adaptive surrogates



- $\mathbf{g}_c := \mathbf{g}(\cdot\,; Y_c, G_c)$ = current model
- $\mathbf{g}_+[y_+] := \mathbf{g}(\cdot\,; Y_+, G_+)$ = refined model
- $Y_+ := Y_c \cup \{y_+\}$

- Consider adding one point, but can add multiple points.

- Given any y_+, can compute magnitude of error bars for $\mathbf{g}_+[y_+]$ without computing the expensive $\mathbf{g}_*(y_+)$.

Workflow

1. Using current surrogate $\mathbf{g}_c$, solve ODE to obtain $\mathbf{x}_c := \mathbf{x}(\cdot\,; \mathbf{g}_c)$.
2. Use sensitivity + surrogate error bounds to choose an "optimal" y_+.
   - Details discussed next.
3. Compute $\mathbf{g}_*(y_+)$ and obtain refined model $\mathbf{g}_+$.
4. Solve ODE again to obtain better solution $\mathbf{x}_+ := \mathbf{x}(\cdot\,; \mathbf{g}_+)$.
5. Continue refining until some stopping criterion is met.

Reducing the solution error

- First-order Taylor approximation in Banach spaces gives

$$\underbrace{\mathbf{x}(\mathbf{g}_+[y_+]) - \mathbf{x}(\mathbf{g}_*)}_{\text{solution error}} \approx \underbrace{\mathbf{x}_\mathbf{g}(\mathbf{g}_+[y_+])}_{\text{sensitivity}} \underbrace{(\mathbf{g}_+[y_+] - \mathbf{g}_*)}_{\text{model error}}$$

  (can apply similar approximation for QoI error).
- Want to choose y_+ to make solution error small.
- Can't use $\mathbf{x}_\mathbf{g}(\mathbf{g}_+[y_+])$ since y_+ not chosen yet. Instead, approximate

$$\mathbf{x}(\mathbf{g}_+[y_+]) - \mathbf{x}(\mathbf{g}_*) \approx \mathbf{x}_\mathbf{g}(\mathbf{g}_c)(\mathbf{g}_+[y_+] - \mathbf{g}_*).$$

- Model error $\mathbf{g}_+[y_+] - \mathbf{g}_*$ is expensive to compute. Assume we have an error bound

$$|\mathbf{g}_+[y_+](y) - \mathbf{g}_*(y)| \leq c(\mathbf{g}_*)\,\boldsymbol{\epsilon}(y; Y_+, G_c)$$

  which is independent of $\mathbf{g}_*(y_+)$. ($\mathbf{g}$ assumed scalar)
- Idea: use pointwise error bound to obtain upper bound on $\|\mathbf{x}_\mathbf{g}(\mathbf{g}_c)(\mathbf{g}_+[y_+] - \mathbf{g}_*)\|$ (in some seminorm), and minimize over y_+.

Bound on solution error

- Upper bound on weighted L^2-(semi)norm of $\mathbf{x}_{\mathbf{g}}(\mathbf{g}_c)(\mathbf{g}_+[y_+] - \mathbf{g}_*)$ computed by finding solution

$$\delta\mathbf{x} \in \big(W^{1,\infty}(I)\big)^{n_x}, \qquad \delta\mathbf{g} \in \big(\mathcal{G}^2(I)\big)^{n_g} \qquad \text{of}$$

$$\begin{aligned}
\max_{\delta\mathbf{x},\delta\mathbf{g}} \quad & \frac{1}{2}\int_{t_0}^{t_f} \delta\mathbf{x}(t)^T \mathbf{Q}(t)\delta\mathbf{x}(t)\,dt \\
\text{s.t.} \quad & \delta\mathbf{x}'(t) = \mathbf{A}_c(t)\delta\mathbf{x}(t) + \mathbf{B}_c(t)\delta\mathbf{g}\big(\mathbf{y}_c(t)\big) \\
& \delta\mathbf{x}(t_0) = \mathbf{0} \\
& -c(\mathbf{g}_*)\,\boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \delta\mathbf{g}\big(\mathbf{y}_c(t)\big) \le c(\mathbf{g}_*)\,\boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big)
\end{aligned}$$

- Solution approximates worst-case trajectory error given error bound for refined model as a function of y_+.
  Problem ill-suited for discretization as δg is state-dependent.
  Can "relax" the problem to obtain something more tractable.

Bound on solution error

- Upper bound on weighted L^2-(semi)norm of $\mathbf{x_g}(\mathbf{g}_c)(\mathbf{g}_+[y_+] - \mathbf{g}_*)$ computed by finding solution

$$\delta\mathbf{x} \in \big(W^{1,\infty}(I)\big)^{n_x}, \qquad \boldsymbol{\delta} \in \big(L^\infty(I)\big)^{n_g} \qquad \text{of}$$

$$\begin{aligned} \max_{\delta\mathbf{x},\boldsymbol{\delta}} \quad & \frac{1}{2}\int_{t_0}^{t_f} \delta\mathbf{x}(t)^T\mathbf{Q}(t)\delta\mathbf{x}(t)\,dt \\ \text{s.t.} \quad & \delta\mathbf{x}'(t) = \mathbf{A}_c(t)\delta\mathbf{x}(t) + \mathbf{B}_c(t)\boldsymbol{\delta}(t) \\ & \delta\mathbf{x}(t_0) = \mathbf{0} \\ & -c(\mathbf{g}_*)\,\boldsymbol{\epsilon}\big(\mathbf{y}_c(t);Y_+,G_c\big) \le \boldsymbol{\delta}(t) \le c(\mathbf{g}_*)\,\boldsymbol{\epsilon}\big(\mathbf{y}_c(t);Y_+,G_c\big) \end{aligned}$$

- Convex maximization problem; global solution is NP-hard (assuming it exists).
  Solve using a tailored interior point method with zero as the initial point that leverages symmetry of problem.

- $c(\mathbf{g}_*)$ not known, but problem is scale-invariant.

Bound on QoI error

- Can similarly obtain approximate upper bound on QoI error by obtaining $\boldsymbol{\lambda}_c \in \big(W^{1,\infty}(I)\big)^{n_x}$ from adjoint equation and solving

$$\begin{aligned} \max_{\delta \mathbf{g}} \quad & \Big| \int_{t_0}^{t_f} \big(\mathbf{B}_c(t)^T \boldsymbol{\lambda}_c(t) + \nabla_g l_c[t]\big)^T \delta\mathbf{g}\big(\mathbf{y}_c(t)\big)\, dt \Big| \\ \text{s.t.} \quad & -c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \delta\mathbf{g}\big(\mathbf{y}_c(t)\big) \le c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \end{aligned}$$

- This is a linear program; much easier to solve. In fact, "L^∞ relaxation" has a simple analytical solution.

Bound on QoI error

- Can similarly obtain approximate upper bound on QoI error by obtaining $\boldsymbol{\lambda}_c \in \big(W^{1,\infty}(I)\big)^{n_x}$ from adjoint equation and solving

$$\begin{aligned} \max_{\boldsymbol{\delta}} \quad & \Big| \int_{t_0}^{t_f} \big(\mathbf{B}_c(t)^T \boldsymbol{\lambda}_c(t) + \nabla_g l_c[t]\big)^T \boldsymbol{\delta}(t)\, dt \Big| \\ \text{s.t.} \quad & -c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \leq \boldsymbol{\delta}(t) \leq c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big), \end{aligned}$$

Bound on QoI error

- Can similarly obtain approximate upper bound on QoI error by obtaining $\boldsymbol{\lambda}_c \in \big(W^{1,\infty}(I)\big)^{n_x}$ from adjoint equation and solving

$$\max_{\boldsymbol{\delta}} \quad \Big| \int_{t_0}^{t_f} \big(\mathbf{B}_c(t)^T \boldsymbol{\lambda}_c(t) + \nabla_g l_c[t]\big)^T \boldsymbol{\delta}(t)\, dt \Big|$$
$$\text{s.t.} \quad - c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}(t) \le c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big),$$

- Analytical solution given by

$$\boldsymbol{\delta}(t) = c(\mathbf{g}_*)\, \mathrm{sgn}\big(\mathbf{B}_c(t)^T \boldsymbol{\lambda}_c(t) + \nabla_g l_c[t]\big) \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big)$$

with objective value

$$\delta \widetilde{q}_{\mathrm{UB}}[y_+] = c(\mathbf{g}_*) \int_{t_0}^{t_f} \Big|\mathbf{B}_c(t)^T \boldsymbol{\lambda}_c(t) + \nabla_g l_c[t]\Big|^T \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big)\, dt.$$

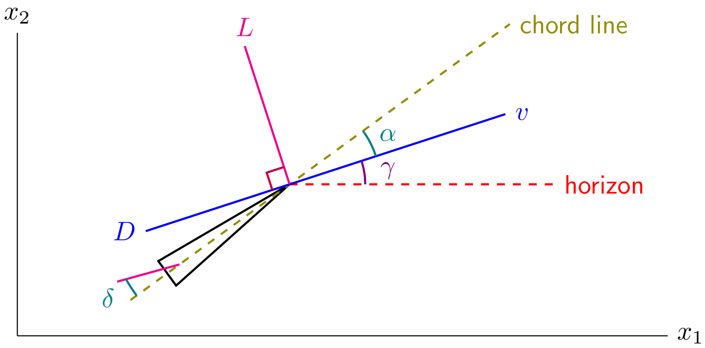
Refinement

- Optimization problem is parametrized by refinement point y_+, which affects the box constraints

$$-\, c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}(t) \le c(\mathbf{g}_*)\, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big)$$

- Choose y_+ such that the worst-case perturbation is minimized over a set of candidates, i.e. worst-case perturbation is acquisition function.
- Multiple possible strategies:
  - Randomly generate finitely many candidates and solve by brute force.
  - Consider points in a continuous range and use a numerical method.
- Currently use the first strategy; second approach part of future work.

Dynamic model



Dynamic model for longitudinal flight with flap deflection control.

- Elevator/flap angle $\delta \to$ pitch rate $q \to$ angle of attack α
  $\to$ lift L and drag $D \to$ speed v and flight path angle γ
  $\to$ downrange x_1 and altitude x_2

Setup

- Lift/drag/moment coefficients C_L, C_D, C_M modeled using surrogates:
$$\mathbf{g} = \begin{bmatrix} C_L(\alpha,\delta) \\ C_D(\alpha,\delta) \\ C_M(\alpha,\delta) \end{bmatrix}$$
each constructed in RKHS $\mathcal{H}_{\mathbf{k}}(\mathbb{R}^2)$ for Gaussian kernel.
Lengthscale adjusted based on average distance between data points.

- True model (synthetic):
$$\begin{aligned} C_L(\alpha,\delta) &= -0.04 + 0.8\alpha + 0.13\delta \\ C_D(\alpha,\delta) &= 0.012 - 0.01\alpha + 0.6\alpha^2 - 0.02\delta + 0.12\delta^2 \\ C_M(\alpha,\delta) &= 0.1745 - \alpha - \delta \end{aligned}$$

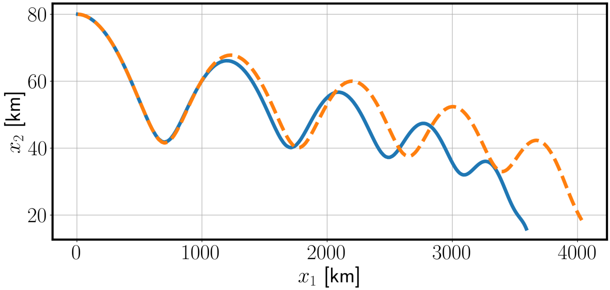
- In practice, would use CFD as high-fidelity model, which is much more expensive.

Initialization

- Initial models for aerodynamic coefficients C_L, C_D, C_M constructed from high-fidelity data at 3 Latin hypercube (LH) samples in range

$$\alpha \in [-5^\circ, 25^\circ], \qquad \delta \in [-15^\circ, 10^\circ]$$

- Simulated with initial surrogates to obtain nominal trajectory.



Nominal trajectory, true trajectory

Three model refinement approaches

- *ASMR approach*: Generate candidates y_+^i and select refinement point using sensitivity-based upper bound on solution error as acquisition function.

- *Max error bound (MEB) approach*: Generate candidates y_+^i and select refinement point using

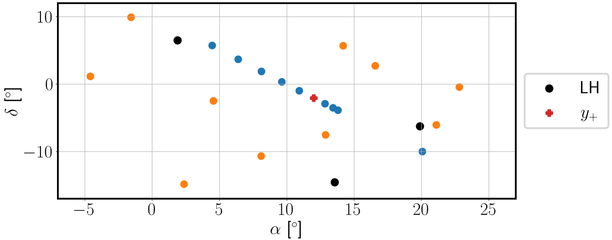
$$\arg\max_i \quad \left\| \begin{pmatrix} \|(C_L)_c\|_{\mathcal{H}_{\mathbf{k}}(\mathbb{R}^2)} P_L(y_+^i; Y_c) \\ \|(C_D)_c\|_{\mathcal{H}_{\mathbf{k}}(\mathbb{R}^2)} P_D(y_+^i; Y_c) \\ \|(C_M)_c\|_{\mathcal{H}_{\mathbf{k}}(\mathbb{R}^2)} P_M(y_+^i; Y_c) \end{pmatrix} \right\|_2$$

  as acquisition function.

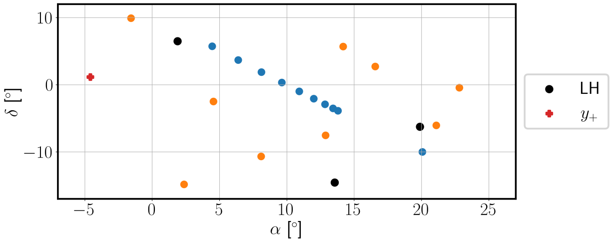
- *LH approach*: Simply generate $3 + r$ LH samples (not adaptive).
  - Candidate generation strategy: selected (α, δ) values at 10 equispaced times in $[t_0, t_f]$ along the current trajectory, plus 10 additional LH candidates (20 total).

Refinement point

ASMR:

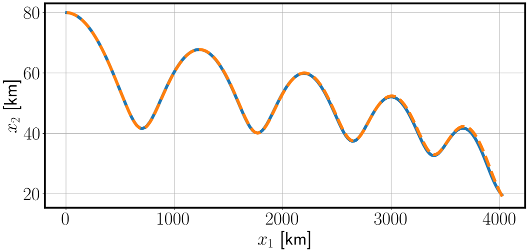


MEB:

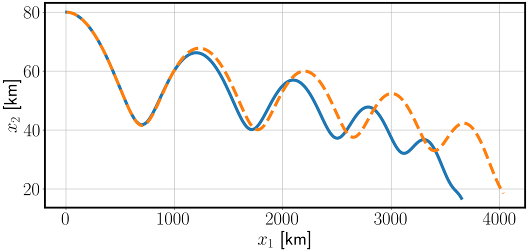


Trajectories ($r = 1$)

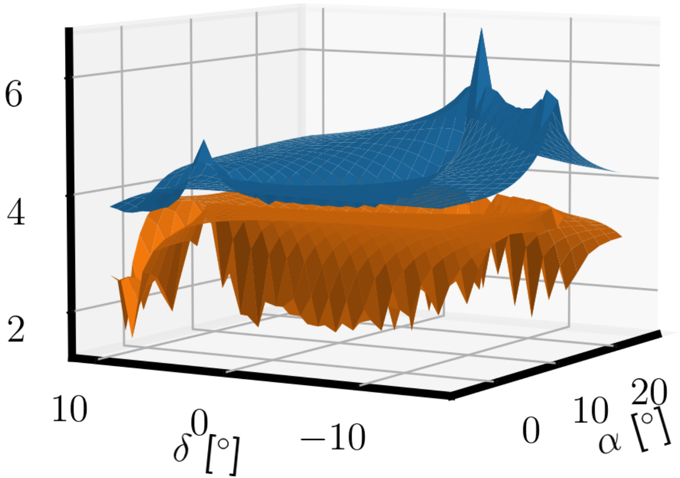
ASMR:



MEB:

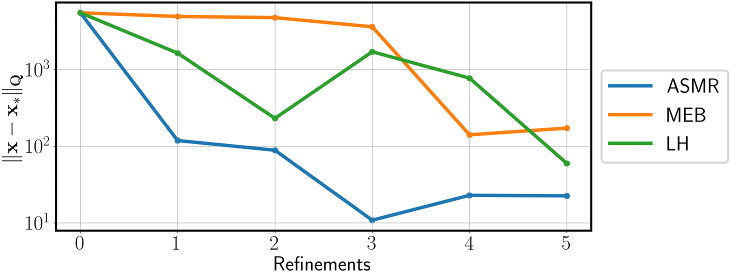


ASMR acquisition function vs. post-refinement error



ASMR acquisition function, trajectory error after refinement
(log scale)

Trajectory error comparison



- ASMR outperforms both approaches, but convergence quickly flattens out due to inability to reduce interpolation error as increased density of samples worsens conditioning of $\mathbf{k}(Y, Y)$.
- May observe better performance with hierarchical modeling approaches or compactly supported kernels, e.g., Wendland kernels.

Outline

Problem formulation

- Consider the optimal control problem

$$
\begin{aligned}
&\min \quad \varphi\big(\mathbf{x}(t_f)\big) + \int_{t_0}^{t_f} l\Big(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big)\Big)\, dt \\
&\text{s.t.} \quad \mathbf{x}'(t) = \mathbf{f}\Big(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big)\Big), \qquad \text{a.a. } t \in I, \\
&\qquad\ \ \mathbf{x}(t_0) = x_0
\end{aligned}
$$

- Optimization variable: time-dependent control input $\mathbf{u} \in \big(L^\infty(I)\big)^{n_u}$.

- To obtain a sensitivity result, apply Implicit Function Theorem to first-order necessary optimality conditions.

- Need strong second-order sufficient conditions to satisfy hypotheses of IFT and handle two-norm discrepancy.

Differentiability of the dynamics

- One additional derivative of $\mathbf{g}$ needed for IFT in optimization setting:

$$\big(\mathcal{G}^3(\Omega)\big)^{n_g} := \big\{\mathbf{g} : \Omega \to \mathbb{R}^{n_g} \ : \ \mathbf{g}(y) \text{ is three times continuously differentiable with respect to } y \in \Omega, \text{ and } \|\mathbf{g}\|_{(\mathcal{G}^3(\Omega))^{n_g}} < \infty\big\},$$

$$\|\mathbf{g}\|_{(\mathcal{G}^3(\Omega))^{n_g}} := \sum_{n=0}^{3} \sup_{y\in\Omega} \left\| \frac{\partial^n}{\partial y^n} \mathbf{g}(y) \right\|.$$

- Right-hand side operator

$$\mathbf{F} : \Big(\big(W^{1,\infty}(I)\big)^{n_x} \times \big(L^\infty(I)\big)^{n_u}\Big) \times \big(\mathcal{G}^3(\Omega)\big)^{n_g} \to \big(L^\infty(I)\big)^{n_x}$$

representing the right-hand side

$$\mathbf{F}(\mathbf{x}, \mathbf{u}; \mathbf{g}) := \mathbf{f}\Big(\cdot, \mathbf{x}(\cdot), \mathbf{u}(\cdot), \mathbf{g}\big(\mathbf{x}(\cdot), \mathbf{u}(\cdot)\big)\Big)$$

is continuously Fréchet differentiable due to continuous embedding of $\mathcal{G}^3(\Omega)$ in $\mathcal{G}^2(\Omega)$ and $W^{1,\infty}(I)$ in $L^\infty(I)$.

(Recall differentiability was established in L^∞ spaces.)

Optimality conditions

- Let $(\overline{\mathbf{x}}, \overline{\mathbf{u}}) \in \big(W^{1,\infty}(I)\big)^{n_x} \times \big(L^{\infty}(I)\big)^{n_u}$ be a local minimum. Then there exists $\overline{\boldsymbol{\lambda}} \in \big(W^{1,\infty}(I)\big)^{n_x}$ such that

$$\begin{aligned}
&\boldsymbol{\psi}(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\boldsymbol{\lambda}}; \overline{\mathbf{g}}) \\
&:= \begin{pmatrix} \overline{\mathbf{f}}[\cdot] - \overline{\mathbf{x}}'(\cdot) \\ x_0 - \overline{\mathbf{x}}(t_0) \\ \overline{\boldsymbol{\lambda}}'(\cdot) + \overline{\mathbf{A}}[\cdot]^T \overline{\boldsymbol{\lambda}}(\cdot) + \big(\nabla_x \bar{l}[\cdot] + \overline{\mathbf{g}}_x[\cdot]^T \nabla_g \bar{l}[\cdot]\big) \\ \nabla_x \overline{\varphi}[t_f] - \overline{\boldsymbol{\lambda}}(t_f) \\ \overline{\mathbf{B}}[\cdot]^T \overline{\boldsymbol{\lambda}}(\cdot) + \big(\nabla_u \bar{l}[\cdot] + \overline{\mathbf{g}}_u[\cdot]^T \nabla_g \bar{l}[\cdot]\big) \end{pmatrix} \\
&= 0
\end{aligned}$$

  where

$$\overline{\mathbf{A}}[\cdot] = \overline{\mathbf{f}}_x[\cdot] + \overline{\mathbf{f}}_g[\cdot]\overline{\mathbf{g}}_x[\cdot], \qquad \overline{\mathbf{B}}[\cdot] = \overline{\mathbf{f}}_u[\cdot] + \overline{\mathbf{f}}_g[\cdot]\overline{\mathbf{g}}_u[\cdot].$$

- $\boldsymbol{\psi}$ continuously Fréchet differentiable and $\boldsymbol{\psi}_{(\mathbf{x},\mathbf{u},\boldsymbol{\lambda})}$ bijective $\Rightarrow \mathbf{z}(\mathbf{g}) := \big(\mathbf{x}(\mathbf{g}), \mathbf{u}(\mathbf{g}), \boldsymbol{\lambda}(\mathbf{g})\big)$ is continuously Fréchet diff'ble by IFT.

Sensitivity computation

- Partial Fréchet derivatives:

$$
\begin{aligned}
&\boldsymbol{\psi}_{(\mathbf{x},\mathbf{u},\boldsymbol{\lambda})}(\overline{\mathbf{x}},\overline{\mathbf{u}},\overline{\boldsymbol{\lambda}};\overline{\mathbf{g}})(\delta\mathbf{x},\delta\mathbf{u},\delta\boldsymbol{\lambda}) \\
&\quad = \begin{pmatrix}
\overline{\mathbf{A}}[\cdot]\delta\mathbf{x}(\cdot) + \overline{\mathbf{B}}[\cdot]\delta\mathbf{u}(\cdot) - \delta\mathbf{x}'(\cdot) \\
-\delta\mathbf{x}(t_0) \\
\delta\boldsymbol{\lambda}'(\cdot) + \overline{\mathbf{H}}_{xx}[\cdot]\delta\mathbf{x}(\cdot) + \overline{\mathbf{H}}_{xu}[\cdot]\delta\mathbf{u}(\cdot) + \overline{\mathbf{A}}[\cdot]^T\delta\boldsymbol{\lambda}(\cdot) \\
\nabla^2_{xx}\overline{\varphi}[t_f]\delta\mathbf{x}(t_f) - \delta\boldsymbol{\lambda}(t_f) \\
\overline{\mathbf{H}}_{ux}[\cdot]\delta\mathbf{x}(\cdot) + \overline{\mathbf{H}}_{uu}[\cdot]\delta\mathbf{u}(\cdot) + \overline{\mathbf{B}}[\cdot]^T\delta\boldsymbol{\lambda}(\cdot)
\end{pmatrix}
\end{aligned}
$$

and

$$
\boldsymbol{\psi}_{\mathbf{g}}(\overline{\mathbf{x}},\overline{\mathbf{u}},\overline{\boldsymbol{\lambda}};\overline{\mathbf{g}})\delta\mathbf{g} = \begin{pmatrix}
\overline{\mathbf{f}}_g[\cdot]\delta\mathbf{g}[\cdot] \\
0 \\
\overline{\mathbf{H}}_{xg}[\cdot]\delta\mathbf{g}[\cdot] + \delta\mathbf{g}_x[\cdot]^T\overline{\mathbf{d}}[\cdot] \\
0 \\
\overline{\mathbf{H}}_{ug}[\cdot]\delta\mathbf{g}[\cdot] + \delta\mathbf{g}_u[\cdot]^T\overline{\mathbf{d}}[\cdot]
\end{pmatrix}
$$

where

$$
\overline{\mathbf{H}}[\cdot] = \overline{l}[\cdot] + \overline{\boldsymbol{\lambda}}(\cdot)^T\overline{\mathbf{f}}[\cdot], \qquad\qquad \overline{\mathbf{d}}[\cdot] = \overline{\mathbf{f}}_g[\cdot]^T\overline{\boldsymbol{\lambda}}(\cdot).
$$

Sensitivity computation

- Sensitivity given by $\delta \mathbf{z} = -\boldsymbol{\psi}_{(\mathbf{x},\mathbf{u},\boldsymbol{\lambda})}^{-1} \boldsymbol{\psi}_{\mathbf{g}} \delta \mathbf{g}$, which is solution of the linear quadratic optimal control problem (LQOCP)

$$
\begin{aligned}
\min_{\delta \mathbf{x}, \delta \mathbf{u}} \quad & \int_{t_0}^{t_f} \begin{pmatrix} \mathbf{c}_x(t) \\ \mathbf{c}_u(t) \end{pmatrix}^T \begin{pmatrix} \delta \mathbf{x}(t) \\ \delta \mathbf{u}(t) \end{pmatrix} dt \\
& + \frac{1}{2} \int_{t_0}^{t_f} \begin{pmatrix} \delta \mathbf{x}(t) \\ \delta \mathbf{u}(t) \end{pmatrix}^T \begin{pmatrix} \overline{\mathbf{H}}_{xx}[t] & \overline{\mathbf{H}}_{xu}[t] \\ \overline{\mathbf{H}}_{ux}[t] & \overline{\mathbf{H}}_{uu}[t] \end{pmatrix} \begin{pmatrix} \delta \mathbf{x}(t) \\ \delta \mathbf{u}(t) \end{pmatrix} dt \\
& + \sigma_f^T \delta \mathbf{x}(t_f) + \frac{1}{2} \delta \mathbf{x}(t_f)^T \nabla_{xx}^2 \overline{\varphi}[t_f] \delta \mathbf{x}(t_f) \\
\text{s.t.} \quad & \delta \mathbf{x}'(t) = \overline{\mathbf{A}}[t] \delta \mathbf{x}(t) + \overline{\mathbf{B}}[t] \delta \mathbf{u}(t) + \mathbf{r}(t), \qquad \text{a.a. } t \in I, \\
& \delta \mathbf{x}(t_0) = r_0
\end{aligned}
$$

with

$$
\mathbf{r}(t) = \overline{\mathbf{f}}_g[t] \delta \mathbf{g}[t], \qquad \begin{pmatrix} \mathbf{c}_x(t) \\ \mathbf{c}_u(t) \end{pmatrix} = \begin{pmatrix} \overline{\mathbf{H}}_{xg}[t] \delta \mathbf{g}[t] + \delta \mathbf{g}_x[t]^T \overline{\mathbf{d}}[t] \\ \overline{\mathbf{H}}_{ug}[t] \delta \mathbf{g}[t] + \delta \mathbf{g}_u[t]^T \overline{\mathbf{d}}[t] \end{pmatrix},
$$

$$
r_0 = \sigma_f = 0.
$$

Sensitivity computation

- Consider the QoI

$$q(\mathbf{x}, \mathbf{u}; \mathbf{g}) := \phi\big(\mathbf{x}(t_f)\big) + \int_{t_0}^{t_f} \ell\Big(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big)\Big)\, dt.$$

(May be different from objective of OCP, hence ϕ, ℓ instead of φ, l.)

- QoI as a function of OCP solution given by

$$\widetilde{q}(\mathbf{g}) = q\big(\mathbf{x}(\mathbf{g}), \mathbf{u}(\mathbf{g}); \mathbf{g}\big).$$

- QoI sensitivity computed by solving "adjoint OCP" given by same LQOCP as before, but with

$$\mathbf{r}(t) \equiv 0, \qquad \begin{pmatrix} \mathbf{c}_x(t) \\ \mathbf{c}_u(t) \end{pmatrix} = \begin{pmatrix} \nabla_x \overline{\ell}[t] + \overline{\mathbf{g}}_x[t]^T \nabla_g \overline{\ell}[t] \\ \nabla_u \overline{\ell}[t] + \overline{\mathbf{g}}_u[t]^T \nabla_g \overline{\ell}[t] \end{pmatrix},$$

$$r_0 = 0, \qquad \sigma_f = \nabla_x \overline{\phi}[t_f].$$

Sensitivity computation

- QoI sensitivity given by

$$\widetilde{q}_{\mathbf{g}}(\overline{\mathbf{g}})\delta\mathbf{g} = \int_{t_0}^{t_f} \Big(\overline{\mathbf{H}}_{gx}[t]\widetilde{\delta\mathbf{x}}(t) + \overline{\mathbf{H}}_{gu}[t]\widetilde{\delta\mathbf{u}}(t) + \overline{\mathbf{f}}_g[t]^T\widetilde{\delta\boldsymbol{\lambda}}(t) + \nabla_g\overline{\ell}[t] \Big)^T \delta\mathbf{g}[t] \\ + \overline{\mathbf{d}}[t]^T\delta\mathbf{g}_x[t]\widetilde{\delta\mathbf{x}}(t) + \overline{\mathbf{d}}[t]^T\delta\mathbf{g}_u[t]\widetilde{\delta\mathbf{u}}(t)\, dt$$

where $\widetilde{\delta\mathbf{x}}$, $\widetilde{\delta\mathbf{u}}$ solve adjoint OCP with adjoint $\widetilde{\delta\boldsymbol{\lambda}}$.

- After one adjoint LQOCP solve, can compute for any $\delta\mathbf{g}$.

- Sensitivity of solution/QoI depends on derivatives of $\delta\mathbf{g}$; need error bounds for these as well.

Derivative error bound

- Need error bound for model and its first derivatives to define acquisition function in optimization setting.
- Notation: D_i^α denotes partial differentiation of the i-th argument of a function with respect to the multi-index α.
- If $D_1^\alpha D_2^\alpha \mathbf{k}(\cdot,\cdot)$ exists and is continuous, derivatives of kernel interpolants in $\mathcal{H}_{\mathbf{k}}(\Omega)$ also have pointwise error bounds:

$$|D^\alpha \mathbf{g}(y) - D^\alpha \mathbf{g}_*(y)| \le \|\mathbf{g}_*\|_{\mathcal{H}_{\mathbf{k}}(\Omega)} \, P^\alpha(y;Y),$$
$$P^\alpha(y;Y) = \sqrt{D_1^\alpha D_2^\alpha \mathbf{k}(y,y) - D_1^\alpha \mathbf{k}(y,Y)\, \mathbf{k}(Y,Y)^{-1}\, D_2^\alpha \mathbf{k}(Y,y)}.$$

Bound on QoI sensitivity

- Worst-case bound on QoI sensitivity obtained by finding solution

$$\delta \mathbf{g} \in \big(\mathcal{G}^3(\Omega)\big)^{n_g} \qquad \text{of}$$

$$\begin{aligned}
\max_{\delta \mathbf{g}} \quad & \Big| \int_{t_0}^{t_f} \Big(\overline{\mathbf{H}}_{gx}[t] \widetilde{\delta \mathbf{x}}(t) + \overline{\mathbf{H}}_{gu}[t] \widetilde{\delta \mathbf{u}}(t) + \overline{\mathbf{f}}_g[t]^T \widetilde{\delta \boldsymbol{\lambda}}(t) + \nabla_g \overline{\ell}[t] \Big)^T \delta \mathbf{g}[t] \\
& + \overline{\mathbf{d}}[t]^T \delta \mathbf{g}_x[t] \widetilde{\delta \mathbf{x}}(t) + \overline{\mathbf{d}}[t]^T \delta \mathbf{g}_u[t] \widetilde{\delta \mathbf{u}}(t) \, dt \Big| \\
\text{s.t.} \quad & - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \delta \mathbf{g}\big(\mathbf{y}_c(t)\big) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big), \\
& - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^x\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \delta \mathbf{g}_x\big(\mathbf{y}_c(t)\big) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^x\big(\mathbf{y}_c(t); Y_+, G_c\big), \\
& - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^u\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \delta \mathbf{g}_u\big(\mathbf{y}_c(t)\big) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^u\big(\mathbf{y}_c(t); Y_+, G_c\big)
\end{aligned}$$

- Once again, perform "L^∞ relaxation" to obtain optimal control problem.

Bound on QoI sensitivity

- Worst-case bound on QoI sensitivity obtained by finding solution

$$\boldsymbol{\delta} \in \big(L^\infty(I)\big)^{n_g}, \quad \boldsymbol{\delta}^x \in \big(L^\infty(I)\big)^{n_g \times n_x}, \quad \boldsymbol{\delta}^u \in \big(L^\infty(I)\big)^{n_g \times n_u} \qquad \text{of}$$

$$\begin{aligned}
\max_{\delta \mathbf{g}} \quad & \Big| \int_{t_0}^{t_f} \Big(\overline{\mathbf{H}}_{gx}[t] \widetilde{\delta \mathbf{x}}(t) + \overline{\mathbf{H}}_{gu}[t] \widetilde{\delta \mathbf{u}}(t) + \overline{\mathbf{f}}_g[t]^T \widetilde{\delta \boldsymbol{\lambda}}(t) + \nabla_g \overline{\ell}[t] \Big)^T \boldsymbol{\delta}(t) \\
& + \overline{\mathbf{d}}[t]^T \boldsymbol{\delta}^x(t) \widetilde{\delta \mathbf{x}}(t) + \overline{\mathbf{d}}[t]^T \boldsymbol{\delta}^u(t) \widetilde{\delta \mathbf{u}}(t) \, dt \Big| \\
\text{s.t.} \quad & - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}(t) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big), \\
& - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^x\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}^x(t) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^x\big(\mathbf{y}_c(t); Y_+, G_c\big), \\
& - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^u\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}^u(t) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^u\big(\mathbf{y}_c(t); Y_+, G_c\big)
\end{aligned}$$

- Analytical solution easily obtainable for this problem.

Bound on QoI sensitivity

$$\max_{\delta \mathbf{g}} \quad \Big| \int_{t_0}^{t_f} \Big(\overline{\mathbf{H}}_{gx}[t] \widetilde{\delta \mathbf{x}}(t) + \overline{\mathbf{H}}_{gu}[t] \widetilde{\delta \mathbf{u}}(t) + \overline{\mathbf{f}}_g[t]^T \widetilde{\delta \boldsymbol{\lambda}}(t) + \nabla_g \overline{\ell}[t] \Big)^T \boldsymbol{\delta}(t)$$

$$+ \overline{\mathbf{d}}[t]^T \boldsymbol{\delta}^x(t) \widetilde{\delta \mathbf{x}}(t) + \overline{\mathbf{d}}[t]^T \boldsymbol{\delta}^u(t) \widetilde{\delta \mathbf{u}}(t) \, dt \Big|$$

s.t. $\quad - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}(t) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big),$

$\quad - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^x\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}^x(t) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^x\big(\mathbf{y}_c(t); Y_+, G_c\big),$

$\quad - c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^u\big(\mathbf{y}_c(t); Y_+, G_c\big) \le \boldsymbol{\delta}^u(t) \le c(\mathbf{g}_*) \, \boldsymbol{\epsilon}^u\big(\mathbf{y}_c(t); Y_+, G_c\big)$ solved by

$$\boldsymbol{\delta}(t) = c(\mathbf{g}_*) \, \mathrm{sgn}\Big(\overline{\mathbf{H}}_{gx}[t] \widetilde{\delta \mathbf{x}}(t) + \cdots \Big) \boldsymbol{\epsilon}\big(\mathbf{y}_c(t); Y_+, G_c\big),$$

$$\boldsymbol{\delta}_i^x(t) = c(\mathbf{g}_*) \, \mathrm{sgn}\big(\overline{\mathbf{d}}[t] \widetilde{\delta \mathbf{x}}_i(t) \big) \boldsymbol{\epsilon}_i^x\big(\mathbf{y}_c(t); Y_+, G_c\big), \quad i = 1, \ldots, n_x,$$

$$\boldsymbol{\delta}_i^u(t) = c(\mathbf{g}_*) \, \mathrm{sgn}\big(\overline{\mathbf{d}}[t] \widetilde{\delta \mathbf{u}}_i(t) \big) \boldsymbol{\epsilon}_i^u\big(\mathbf{y}_c(t); Y_+, G_c\big), \quad i = 1, \ldots, n_u$$

with objective value

$$\delta \widetilde{q}_{\mathrm{UB}}[y_+] = c(\mathbf{g}_*) \int_{t_0}^{t_f} \Big| \overline{\mathbf{H}}_{gx}[t] \widetilde{\delta \mathbf{x}}(t) + \overline{\mathbf{H}}_{gu}[t] \widetilde{\delta \mathbf{u}}(t) + \overline{\mathbf{f}}_g[t]^T \widetilde{\delta \boldsymbol{\lambda}}(t) + \nabla_g \bar{l}[t] \Big|^T \boldsymbol{\epsilon}[t]$$

$$+ \; \big| \overline{\mathbf{d}}[t] \big|^T \boldsymbol{\epsilon}^x[t] \, \big| \widetilde{\delta \mathbf{x}}(t) \big| \; + \; \big| \overline{\mathbf{d}}[t] \big|^T \boldsymbol{\epsilon}^u[t] \, \big| \widetilde{\delta \mathbf{u}}(t) \big| \, dt$$

Reference tracking

- Trajectory optimization problem:

$$
\begin{aligned}
\min \quad & \varphi\big(\mathbf{x}(t_f)\big) + \int_{t_0}^{t_f} l\Big(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big)\Big)\, dt \\
\text{s.t.} \quad & \mathbf{x}'(t) = \mathbf{f}\Big(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big)\Big), \qquad \text{a.a. } t \in I, \\
& \mathbf{x}(t_0) = x_0, \\
& \mathbf{b}\big(\mathbf{x}(t_f)\big) = 0, \\
& \mathbf{C}\Big(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{g}\big(t, \mathbf{x}(t), \mathbf{u}(t)\big)\Big) \le 0, \qquad \text{a.a. } t \in I
\end{aligned}
$$

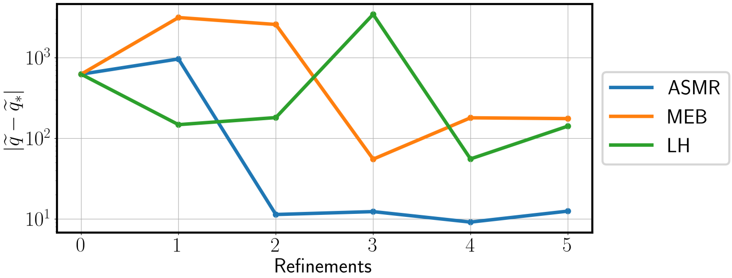
- Not in the form used for sensitivity analysis; however, can solve for reference trajectory $\overline{\mathbf{x}}$, $\overline{\mathbf{u}}$ and consider reference tracking problem

$$
\begin{aligned}
\min \quad & \frac{1}{2}\int_{t_0}^{t_f} \big(\mathbf{x}(t) - \overline{\mathbf{x}}(t)\big)^T \mathbf{Q}(t)\big(\mathbf{x}(t) - \overline{\mathbf{x}}(t)\big)\, dt \\
& \quad + \frac{1}{2}\int_{t_0}^{t_f} \big(\mathbf{u}(t) - \overline{\mathbf{u}}(t)\big)^T \mathbf{R}(t)\big(\mathbf{u}(t) - \overline{\mathbf{u}}(t)\big)\, dt \\
\text{s.t.} \quad & \mathbf{x}'(t) = \mathbf{f}\Big(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big)\Big), \qquad \text{a.a. } t \in I, \\
& \mathbf{x}(t_0) = x_0
\end{aligned}
$$

Reference tracking

- Optimal solution of reference tracking problem with $\overline{\mathbf{g}}$ is $(\overline{\mathbf{x}}, \overline{\mathbf{u}})$; no need to re-solve.
- In this scenario, model refinement improves ability to track reference trajectory in real time when viewing surrogate errors as "disturbance," easing demands on a feedback controller.
- Use final downrange $\mathbf{x}_1(T)$ as QoI.
- Re-solve original OCP for new reference trajectory after each refinement.

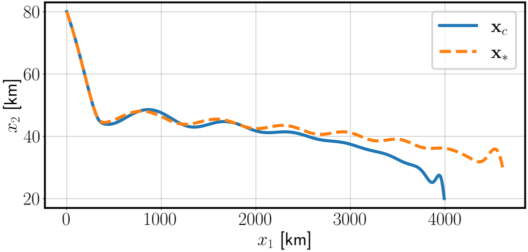
QoI error comparison



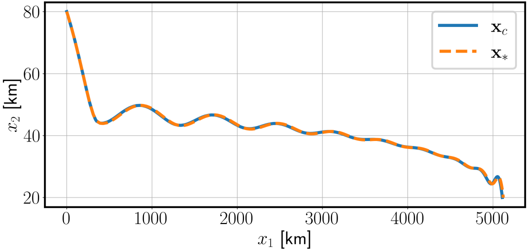
- Difference in downrange between reference trajectory and optimal tracking trajectory with true model.
- Similar convergence behavior as before, except it takes 2 refinements to reduce the error.
- Better model construction approaches would help reduce error further.

Trajectories

Initial:



ASMR $(r = 2)$:



Outline

Conclusions

- Derived sensitivity analysis results for ODEs and OCPs with dynamics involving Nemytskii operators that generalizes parametric sensitivity analysis to state-dependent component functions.

- Developed and implemented the ASMR framework for trajectory simulation and optimization problems with kernel-based surrogate models; outperformed other data acquisition approaches in numerical examples.

- This is a *sensitivity-driven, deterministic, parameter-free* approach to surrogate-assisted optimization and control.

Literature I

[Cangelosi and Heinkenschloss, 2024] Cangelosi, J. R. and Heinkenschloss, M. (2024).

Sensitivity of ODE solutions with respect to component functions in the dynamics.
arXiv:2411.09655v2, DOI: 10.48550/arXiv.2411.09655,
https://doi.org/10.48550/arXiv.2411.09655.

[Cangelosi et al., 2024] Cangelosi, J. R., Heinkenschloss, M., Needels, J. T., and Alonso, J. J. (2024).

Simultaneous design and trajectory optimization for boosted hypersonic glide vehicles.

In *Paper AIAA 2024-0375. 2024 AIAA Science and Technology Forum and Exposition (AIAA SciTech Forum)*. DOI: 10.2514/6.2024-0375,
https://doi.org/10.2514/6.2024-0375.

[Hairer et al., 1993] Hairer, E., Nørsett, S. P., and Wanner, G. (1993).

Solving Ordinary Differential Equations I. Nonstiff Problems, volume 8 of *Springer Series in Computational Mathematics*.

Springer-Verlag, Berlin, second edition, DOI: 10.1007/978-3-540-78862-1,
https://doi.org/10.1007/978-3-540-78862-1.

Literature II

[Hart et al., 2023] Hart, J., van Bloemen Waanders, B., Hood, L., and Parish, J. (2023).

Sensitivity-driven experimental design to facilitate control of dynamical systems.

J. Optim. Theory Appl., 196(3):855–881, DOI: 10.1007/s10957-023-02172-w, https://doi.org/10.1007/s10957-023-02172-w.

[Jones et al., 1998] Jones, D. R., Schonlau, M., and Welch, W. J. (1998).

Efficient global optimization of expensive black-box functions.

J. Global Optim., 13(4):455–492, DOI: 10.1023/A:1008306431147, https://doi.org/10.1023/A:1008306431147.

[Moustapha et al., 2016] Moustapha, M., Sudret, B., Bourinet, J.-M., and Guillaume, B. (2016).

Quantile-based optimization under uncertainties using adaptive kriging surrogate models.

Structural and Multidisciplinary Optimization, 54(6):1403–1421, DOI: 10.1007/s00158-016-1504-4, http://dx.doi.org/10.1007/s00158-016-1504-4.

Literature III

[Söderlind, 2006] Söderlind, G. (2006).
The logarithmic norm. History and modern theory.
BIT, 46(3):631–652, DOI: 10.1007/s10543-006-0069-9,
http://dx.doi.org/10.1007/s10543-006-0069-9.

[Sun et al., 2022] Sun, R., Duan, Q., and Mao, X. (2022).
A multi-objective adaptive surrogate modelling-based optimization algorithm for constrained hybrid problems.
Environmental Modelling & Software, 148:105272, DOI: 10.1016/j.envsoft.2021.105272,
https://doi.org/10.1016/j.envsoft.2021.105272.

[Surmann et al., 2021] Surmann, D., Ligges, U., and Weihs, C. (2021).
Infill criterion for multimodal model-based optimisation.
arXiv:1810.02118v1, DOI: 10.48550/arXiv.1810.02118,
https://doi.org/10.48550/arXiv.1810.02118.

[Vohra et al., 2019] Vohra, M., Alexanderian, A., Safta, C., and Mahadevan, S. (2019).
Sensitivity-driven adaptive construction of reduced-space surrogates.
J. Sci. Comput., 79(2):1335–1359, DOI: 10.1007/s10915-018-0894-4,
https://doi.org/10.1007/s10915-018-0894-4.