Risk-Averse Optimization Formulation

- Recall, we want to solve this computationally expensive problem

\[
\min_{z \in Z, t \in \mathbb{R}} \hat{J}(t, z) \quad \text{where} \quad \hat{J}(t, z) := \left\{ t + \frac{1}{N} \sum_{j=1}^{N} v(G(S(\xi; z), \xi) - t) \right\}
\]
Inexact Trust Region Framework

Given an iterate $x_k = (t_k, z_k) \in X := \mathbb{R} \times Z$,

1. Model Update: Choose a new model $m_k \approx \hat{J}(x_k + s)$.

2. Step Computation: Approximate a solution, $s_k$, to the subproblem

$$\min_{s \in X} m_k(s) \quad \text{subject to} \quad \|s\|_X \leq \Delta_k.$$ 

3. Objective Update: Choose a new $\hat{J}_k(x) \approx \hat{J}(x)$

4. Step Acceptance: Compute

$$\rho_k = \frac{\hat{J}_k(x_k) - \hat{J}_k(x_k + s_k)}{m_k(0) - m_k(s_k)}$$

if $\rho_k \geq \eta \in (0, 1)$, then $x_{k+1} = x_k + s_k$ else $x_{k+1} = x_k$.

5. Trust Region Update: Choose a new trust region radius $\Delta_{k+1}$. 


Inexact Trust Region Framework

To obtain global convergence, we require

1. Inexact gradient condition

\[ \| \nabla m_k(0) - \nabla \tilde{J}(x_k) \| \leq \kappa \min \{ \| \nabla m_k(0) \|_x, \Delta_k \} \]

where \( \kappa > 0 \) is independent of \( k \).

2. Inexact objective condition

\[ |\tilde{J}(x_k) - \tilde{J}(x_k + s_k) - (\tilde{J}_k(x_k) - \tilde{J}_k(x_k + s_k))| \leq K \theta_k \]
\[ \theta_k^\omega \leq \eta \min \{ m_k(0) - m_k(s_k), r_k \} \]

for some \( K > 0 \), \( \eta \) is an algorithmic parameter and \( r_k \geq 0 \) with \( \lim_{k \to \infty} r_k = 0 \).

Observation: We cannot compute \( \tilde{J}(x_k) \) and \( \nabla \tilde{J}(x_k) \). We use adaptive risk-informed sampling.
Inexact TR and Adaptive RB

- From our local RB method, we obtain surrogates of $S(\xi; z)$, denoted by $S_{\text{mod}}(\xi; z)$ and $S_{\text{obj}}(\xi; z)$, to build $m_k$ and $\hat{J}_k$, respectively.

- Computable a posterior error indicators for the surrogate models:

\[
\|S_{\text{mod}}(\xi; z) - S(\xi; z)\|_U \lesssim \epsilon_{\text{mod}}^u(\xi; z) \\
\|S_{\text{obj}}(\xi; z) - S(\xi; z)\|_U \lesssim \epsilon_{\text{obj}}^u(\xi; z).
\]

- Local RB for the adjoint $(\Lambda_{\text{mod}}(\xi; z))$ and its computable a posterior error indicator:

\[
\|\Lambda_{\text{mod}}(\xi; z) - \Lambda(\xi; z)\|_U \lesssim \epsilon_{\text{mod}}^\lambda(\xi; z).
\]
Inexact TR and Adaptive RB (2)

- **Local Reduced Basis Models:** $S_{\text{mod}}(\xi; z)$, $S_{\text{obj}}(\xi; z)$, $\Lambda_{\text{mod}}(\xi; z)$

- **Inexact Gradient Condition:**

  $$
  \|\nabla \hat{J}(x_k) - \nabla m_k(0)\| \lesssim \sum_{j=1}^{m} p_j \delta_{\text{mod}}(\xi_j; x_k)
  $$

  $$
  \delta_{\text{mod}}(\xi_j; x_k) := \mathcal{F}(\epsilon_{\text{mod}}^u(\xi_j; x_k), \epsilon_{\text{mod}}^\lambda(\xi_j; x_k))
  $$

- **Inexact Objective Condition:**

  $$
  |\hat{J}_k(x_k) - \hat{J}(x_k)| \lesssim \sum_{j=1}^{m} p_j \delta_{\text{obj}}(\xi_j; x_k)
  $$

  $$
  \delta_{\text{obj}}(\xi_j; x_k) := \mathcal{G}(\epsilon_{\text{obj}}^u(\xi; x_k))
  $$

We can satisfy these conditions by controlling the Local RB errors.
NUMERICAL EXAMPLES
We consider the optimal control problem

$$\min_{z \in L^2(0,1)} \frac{1}{2} \text{CVaR}_\beta \left( kD S(z) - dk_Y^2 \right) + \frac{1}{2} \int_0^1 z^2 \, dx$$

where \( u = S(z) \) solves the weak form of

$$- \frac{\partial^2 u}{\partial x^2}(x, \omega) + b(x, \omega) \frac{\partial u}{\partial x}(x, \omega) = z(x), \quad x \in (0,1), \quad \text{a.s.}$$

$$u(0, \omega) = u(1, \omega) = 0, \quad \text{a.s.}$$

Here \( \xi \) are deterministic and \( b \) is the random field

$$b(x, \omega) = [0.5 + \epsilon_1(\omega)] 1_{[0,0.5]} + [0.8 + \epsilon_2(\omega)] 1_{[0.5,1]}.$$

Control Ansatz:

$$z = \sum_{i=1}^n z_i 1_{D_i}, \quad \text{where} \quad D_i = (0.1(i-1), 0.1i)$$
Results for 1D Advection-Diffusion

Optimal control

CDF of $G$ under optimal control

$G(\xi)$ under optimal control

$G(\xi)$ under optimal control
Results for 1D Advection-Diffusion
Evolution of Sampling over Optimization Run
Summary

- These are the main ingredients:
  - Local basis enriched with gradient information
  - A practical and effective error indicator
  - Inexact trust region framework
- The cost of the reduced bases solutions does not grow with the number of atoms added.
- We are in the process of extending this framework to nonlinear PDEs.
Current and Future Directions

• Treatment of physical and stochastic dimensions in a unified framework
• Model-form uncertainty
• Treatment of imperfect knowledge on underlying probability laws.
References


Algorithm 1: Adaptive algorithm to build $S_{mod}$ and $\Lambda_{mod}$ for $m_k$

If $k = 0$, model initialization:

- Let $\xi_0 = \mathbb{E}[\xi]$ and build the initial surrogate models $S_{mod}$ and $\Lambda_{mod}$ based on the solution $S(\xi_0; z_0)$, its gradient $\nabla_{\xi} S(\xi_0; z_0)$, the adjoint solution $\Lambda(\xi_0; S(\xi_0; z_0))$ and its gradient $\nabla_{\xi} \Lambda(\xi_0; S(\xi_0; z_0))$.

Model refinement:

Given $w_k$, $\Delta_k$ and $S_{mod}$, which is recycled from step $k - 1$,

- Build $m_k$ with $S_{mod}$, evaluate $\delta_{mod}(\xi_j; w_k)$ for all $j = 1, \ldots, N$, compute $E_{mod}(w_k)$ and $\|\nabla m_k(0)\|_Z$.
- While $E_{mod}(w_k) > \kappa \min \{\|\nabla m_k(0)\|_Z, \Delta_k\}$, do
  - Select $\xi_{max} = \arg \max_{j=1, \ldots, N} p_j \delta_{mod}(\xi_j; w_k)$.
  - Compute $S(\xi_{max}; z_k)$, $\Lambda(\xi_{max}; S(\xi_{max}; z_k))$, $\nabla_{\xi} S(\xi_{max}; z_k)$ and $\nabla_{\xi} \Lambda(\xi_{max}; S(\xi_{max}; z_k))$.
  - Incorporate the new information at $\xi_{max}$ into $S_{mod}$ and $\Lambda_{mod}$ using the local RB method.
  - Update $m_k$ with $S_{mod}$ and $\Lambda_{mod}$.
  - Update $\delta_{mod}(\xi_j; w_k)$ for $j = 1, \ldots, N$ and $E_{mod}(w_k)$, and recompute $\|\nabla m_k(0)\|_Z$.

End

Return $S_{mod}$ and $\Lambda_{mod}$.
Algorithm 2: Adaptive algorithm to build $S_{\text{obj}}$ for $\hat{J}_k$

If $k = 0$, model initialization:
- Let $\xi_0 = \mathbb{E}[\xi]$ and build the initial surrogate model $S_{\text{mod}}$ based on the solution $S(\xi_0, z_0)$, the gradient $\nabla_\xi S(\xi_0, z_0)$.

Else, model refinement:
Given $w_k$, $s_k$, $\text{pred}_k$, $r_k$ and $S_{\text{obj}}$, which is recycled from step $k - 1$,
- Compute $\gamma_k = K \left( \eta \min \{\text{pred}_k, r_k\} \right)^{1/\alpha}$, evaluate $\delta_{\text{obj}}(\xi_j; w_k)$, $\delta_{\text{obj}}(\xi_j; w_k + s_k)$ for $j = 1, \ldots, N$ and compute $E_{\text{obj}}(w_k)$ and $E_{\text{obj}}(w_k + s_k)$.
- While $E_{\text{obj}}(w_k) + E_{\text{obj}}(w_k + s_k) > \gamma_k$, do
  - Select $\xi_{\text{max}} = \arg\max_{j=1,\ldots,N} p_j \left( \delta_{\text{obj}}(\xi_j; w_k) + \delta_{\text{obj}}(\xi_j; w_k + s_k) \right)$.
  - Compute $S(\xi_{\text{max}}, z_k)$, $\nabla_\xi S(\xi_{\text{max}}, z_k)$, $S(\xi_{\text{max}}, z_k + \xi_k)$, and $\nabla_\xi S(\xi_{\text{max}}, z_k + \xi_k)$.
  - Incorporate the new information at $\xi_{\text{max}}$ into $S_{\text{obj}}$ using the local RB method.
  - Update $\delta_{\text{mod}}(\xi_j; w_k)$ and $\delta_{\text{obj}}(\xi_j; w_k + s_k)$ for $j = 1, \ldots, N$, update $E_{\text{obj}}(w_k)$ and $E_{\text{obj}}(w_k + s_k)$.
- End
End
Return $S_{\text{obj}}$. 
Basic Definitions

Consider a function $F(X, z)$, where $X$ is random and $z$ is deterministic.

- The cumulative distribution of $F$ is defined as
  \[
  \psi(x, z) = P(F(X, z) \leq x)
  \]

- The $\beta$ Value-at-Risk ($\beta$-VaR)
  \[
  x_{VaR}^\beta := \min \{x : \psi(x, z) \geq \beta\}
  \]

- The $\beta$ Conditional-Value-at-Risk ($\beta$-CVaR)
  \[
  x_{CVaR}^\beta := E \left[ F(X, z) | F(X, z) \geq x_{VaR}^\beta \right]
  \]
Conventional Trust Region Algorithm

Given $\bar{\Delta} > 0$, $\Delta_0 \in (0, \bar{\Delta})$, and $\eta \in [0, \frac{1}{4})$:

for $k = 0, 1, 2, \ldots$

Obtain $p_k$ by (approximately) solving (4.3);
Evaluate $\rho_k$ from (4.4);
if $\rho_k < \frac{1}{4}$

$\Delta_{k+1} = \frac{1}{4} \| p_k \|$ 

else

if $\rho_k > \frac{3}{4}$ and $\| p_k \| = \Delta_k$

$\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$

else

$\Delta_{k+1} = \Delta_k$;

if $\rho_k > \eta$

$x_{k+1} = x_k + p_k$

else

$x_{k+1} = x_k$;

end (for).

From Numerical Optimization, Nocedal and Wright, Second ed.
Numerical Example

- The regularization term is taken as
  \[ \varphi(z) = \alpha \|z\|^2_z, \; \alpha > 0 \]

- The control \( z(x) \) is assumed to be piecewise constants, i.e.,
  \[ z(x) = \sum_{i=1}^{10} z_i 1_{I_i}(x) \]
  where \( I_i = (0.1(i - 1), 0.1i), \; i = 1, \ldots, 10 \).

- We used the following TR parameters: \( \eta_1 = 0.05, \; \eta_2 = 0.75, \; \gamma = 0.5, \; \omega = 0.75, \; \kappa = 0.1, \; K = 1 \) and \( \Delta_0 = 1 \).

- We set the penalty parameter to \( \alpha = 0.1 \) and the number of Monte Carlo samples to \( N = 4000 \).