Estimation of Risk Measures with Reduced-Order Models

Boris Kramer[1], T.Takhtaganov, M. Heinkenschloss, K. Willcox

Department of Aeronautics and Astronautics
Aerospace Computational Design Laboratory
Massachusetts Institute of Technology

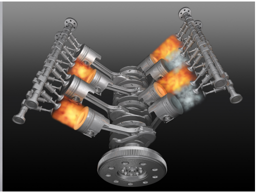
Model Reduction of Parametrized Systems IV
April 11, 2018

[1] Funded by: DARPA EQUiPS program award number UTA15-001067 and Air Force Center of Excellence on Multi-Fidelity Modeling of Rocket Combustor Dynamics, award FA9550-17-1-0195

Need efficient optimization under uncertainty methods for large-scale systems

- Optimization under uncertainty challenging with large-scale models
  - Cost functions require sampling from expensive solutions
  - Often $\mathcal{O}(100) - \mathcal{O}(10,000)$ solves needed

- Penalizing "tail risk" introduces nonlinear cost functions
  - How to sample efficiently (reduce # of samples) to compute cost function?

Risk-averse optimization

$$\min_{z \in \mathcal{Z}} \; J(z) := \mathcal{R}(s(y(\cdot; z))) + \frac{\alpha}{2}\|z\|_{\mathcal{Z}}^2,$$

where $y(\boldsymbol{\xi}, z)$ is a solution to the **parameterized high-fidelity model**

$$F(y(\boldsymbol{\xi}, z), \boldsymbol{\xi}, z) = 0, \qquad \forall \boldsymbol{\xi} \in \Xi.$$

- y: high-fidelity solution that depends on the parameter $\boldsymbol{\xi} = \boldsymbol{\xi}(\omega)$
- $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_M) : \Pi \to \Xi := \prod_{i=1}^{M} \subseteq \mathbb{R}^M$, density $\rho = \prod_{i=1}^{M} \rho_i$
- $\mathcal{R}(\cdot)$ is a risk measure
- z: deterministic control/design. One must decide on control before observing outcome $\boldsymbol{\xi} = \boldsymbol{\xi}(\omega)$.

Output quantity of interest:

$$X(\boldsymbol{\xi}) := s(y(\boldsymbol{\xi}; z))$$

Conditional-Value-at-Risk $=$ β-superquantile

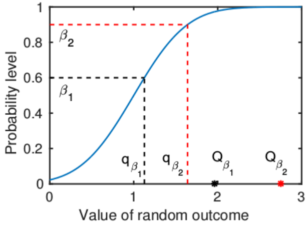
Value-at-Risk at level β ($=$ quantile level, often $\beta \geq 0.95$):

$$\mathrm{VaR}_\beta[X] = \inf\{t \in \mathbb{R} : \mathrm{Pr}[X > t] < 1 - \beta\}$$

Conditional-Value-at-Risk at level β**:** [Rockafellar and Uryasev, 2000]

$$\mathrm{CVaR}_\beta[X] = \mathrm{VaR}_\beta[X] + \frac{1}{1-\beta}\mathbb{E}\left[(X - \mathrm{VaR}_\beta[X])_+\right]$$

- Choose $\mathcal{R}(X) = \mathrm{CVaR}_\beta[X]$
- Penalizes rare outcomes, length of tail matters in CVaR_β
- CVaR_β in engineering design:
  - Royset et al., 2017
  - Yang/Gunzburger, 2017
  - Morio, 2012
  - Zou et al., 2017



CDF and quantiles of normal distribution with mean 1 and sd 0.5

How to use ROMs to estimate risk measures efficiently

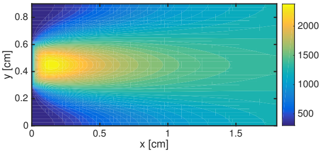
[Heinkenschloss/K./Takhtaganov/Willcox,'17]

1. Numerical test problem
2. Direct sampling from ROM when error bounds are available
3. Using ROMs/surrogates in importance sampling when error is not known

1) Numerical test problem

Test problem: Convection-diffusion-reaction [Buffoni and Willcox, 2010]

- Model includes a one-step reaction of the species $2H_2 + O_2 \rightarrow 2H_2O$
- Inflow of mixture at left boundary

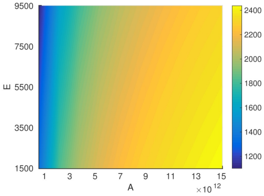


Uncertain parameters of the model relate to reaction terms:

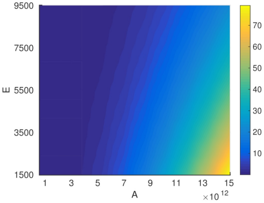
$$\boldsymbol{\xi} = [A, E] \in \Xi$$

Quantity of interest $X : \Xi \mapsto \mathbb{R}$ related to discretized temperature

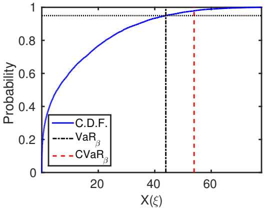
$$X(\boldsymbol{\xi}) = \exp\left(\frac{\|\mathbf{T}(\boldsymbol{\xi})\|_\infty - 2000}{100}\right).$$



Values of $\|\mathbf{T}(\xi)\|_\infty$



Values of $X(\xi)$

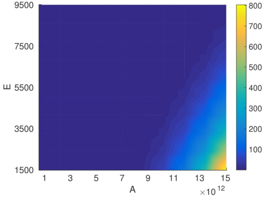


c.d.f. of $X(\boldsymbol{\xi})$, $n = 10^4$

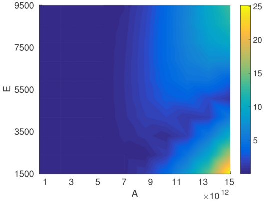
Reduced-order models via POD

- POD ROMs with (D)EIM for Arrhenius reaction terms gives approximate solutions $\mathbf{y} \approx V_r \mathbf{y}_r$
- Projection matrix V_r from $S = 100$ snapshots of HFM at 10×10 equally-spaced values A and E in Ξ
- Four different ROMs from $r = 1, 2, 3, 4$ POD basis vectors
- Surrogate models define a new random variable $X_r : \Xi \mapsto \mathbb{R}$ with

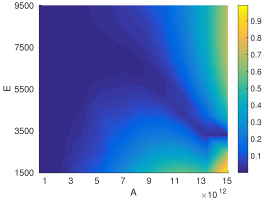
$$X_r(\boldsymbol{\xi}) = \exp\left(\frac{\|\mathbf{T}_r(\boldsymbol{\xi})\|_\infty - 2000}{100}\right)$$



Error of ROM 1, $\epsilon_1(\xi)$



Error of ROM 2, $\epsilon_2(\xi)$



Error of ROM 4, $\epsilon_4(\xi)$

2) CVaR$_\beta$ estimation via direct sampling from ROM

CVaR_β computation

CVaR_β requires sampling in the tail (=risk) region

- If the c.d.f. $H_X(x) = \Pr[X \leq x]$ is continuous at $x = \mathsf{VaR}_\beta[X]$, then $\Pr\big[X = \mathsf{VaR}_\beta[X]\big] = 0$, and

$$\mathsf{CVaR}_\beta[X] = \frac{1}{1-\beta}\mathbb{E}\big[X \cdot \mathbb{I}\left\{X \geq \mathsf{VaR}_\beta[X]\right\}\big].$$

Definition: The **risk region** corresponding to $\mathsf{CVaR}_\beta[X]$ is given by

$$\mathbb{G}_\beta[X] := \{\xi \mid X(\xi) \geq \mathsf{VaR}_\beta[X]\} \subset \Xi$$

and the corresponding indicator function of the risk region $\mathbb{G}_\beta[X]$ is

$$\mathbb{I}_{\mathbb{G}_\beta[X]}(\xi) := \mathbb{I}\left\{X(\xi) \geq \mathsf{VaR}_\beta[X]\right\}.$$

Reduced-order model (ROM) and risk region

Roadmap: Want bound on $|\mathbf{CVaR}_\beta[X] - \mathbf{CVaR}_\beta[X_r]|$

- Given is a HFM $X(\xi)$ and an approximate quantity of interest $X_r(\boldsymbol{\xi})$
- Assume the availability of a bound (will be relaxed later):

$$|X(\xi) - X_r(\xi)| \le \epsilon_r(\xi) \quad \text{for} \quad \xi \in \Xi.$$

Definition and Lemma: The ϵ-risk region corresponding to $\mathrm{CVaR}_\beta[X]$ is given by

$$\mathbb{G}^\epsilon_\beta[X_r] := \{\xi \,:\, X_r(\xi) + \epsilon_r(\xi) \ge \mathrm{VaR}_\beta[X_r - \epsilon_r]\},$$

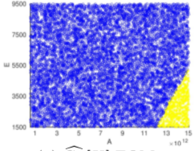
and define

$$\max_{\xi \in \mathbb{G}_\beta[X] \cup \mathbb{G}_\beta[X_r]} \epsilon_r(\xi) \le \max_{\xi \in \mathbb{G}^\epsilon_\beta[X_r]} \epsilon_r(\xi) =: \epsilon^G_r.$$

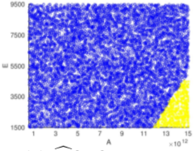
The ϵ-risk region covers the true risk region:

$$\mathbb{G}_\beta[X] \subseteq \mathbb{G}^\epsilon_\beta[X_r] \quad \text{and} \quad \mathbb{G}_\beta[X_r] \subseteq \mathbb{G}^\epsilon_\beta[X_r]$$

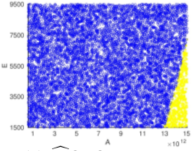
Risk regions for four different ROMs



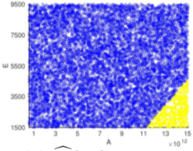
(a) $\widehat{\mathbb{G}_\beta}[X]$ FOM.



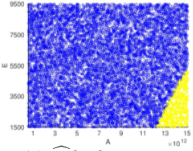
(b) $\widehat{\mathbb{G}_\beta}[X_r]$ ROM 1.



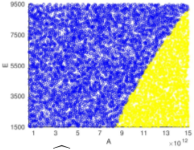
(c) $\widehat{\mathbb{G}_\beta}[X_r]$ ROM 2.



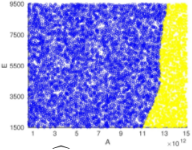
(d) $\widehat{\mathbb{G}_\beta}[X_r]$ ROM 3.



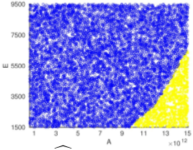
(e) $\widehat{\mathbb{G}_\beta}[X_r]$ ROM 4.



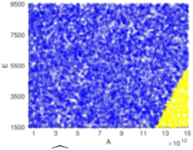
(f) $\widehat{\mathbb{G}^\epsilon_\beta}[X_r]$ ROM 1.



(g) $\widehat{\mathbb{G}^\epsilon_\beta}[X_r]$ ROM 2.



(h) $\widehat{\mathbb{G}^\epsilon_\beta}[X_r]$ ROM 3.



(i) $\widehat{\mathbb{G}^\epsilon_\beta}[X_r]$ ROM 4.

New error bound for ROM-estimated CVaR_β

Theorem: The error between CVaR_β of the full-order model X and CVaR_β of the reduced-order model X_r is bounded as

$$\begin{aligned}
&\left|\mathsf{CVaR}_\beta[X] - \mathsf{CVaR}_\beta[X_r]\right| \\
&\le \left(1 + \frac{\max\left\{\Pr\left[\{X = \mathsf{VaR}_\beta[X]\}\right],\ \Pr\left[\{X_r = \mathsf{VaR}_\beta[X_r]\}\right]\right\}}{1-\beta}\right)\ \epsilon_r^G \\
&\le \left(1 + \frac{1}{1-\beta}\right)\ \epsilon_r^G.
\end{aligned}$$

If X and X_r have c.d.f.'s that are continuous at $\mathsf{VaR}_\beta[X]$ and at $\mathsf{VaR}_\beta[X_r]$, respectively, then

$$\left|\mathsf{CVaR}_\beta[X] - \mathsf{CVaR}_\beta[X_r]\right| \le \epsilon_r^G.$$

- **Only need error ϵ_r^G in the ϵ-risk region $\mathbb{G}_\beta^\epsilon[X_r]$**
- We do not need the error function $\epsilon_r(\xi)$ in all of Ξ

Error bound guides model selection

- Estimates of CVaR_β at level $\beta = 0.95$
- Maximum error in ROM ϵ-risk region $\widehat{\mathbb{G}}_\beta^\epsilon[X_r]$: $\widehat{\epsilon}_r^G$

	$\widehat{\text{CVaR}}_\beta^{\text{MC}}$	Abs error	Rel error (%)	$\widehat{\epsilon}_r^G$
HFM	53.94	—	—	—
ROM 1	361.40	307.47	570.05	776.00
ROM 2	44.80	9.14	16.94	24.47
ROM 3	49.91	4.02	7.46	9.04
ROM 4	53.87	0.07	0.13	0.96

- Note that $\left| \widehat{\text{CVaR}}_\beta^{\text{MC}}[X] - \widehat{\text{CVaR}}_\beta^{\text{MC}}[X_r] \right| \leq \widehat{\epsilon}_r^G$
- Choice of ROM to sample from can be guided through error bound

What if we don't have a rigorous error bound $|X(\xi) - X_r(\xi)| \leq \epsilon_r(\xi)$?

4) CVaR$_\beta$ estimation using ROMs + importance sampling

Importance sampling: A change of measure

- Define $\operatorname{supp}(\rho) := \{\xi \in \Xi \mid \rho(\xi) > 0\}$.
- Let φ be another density with $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\varphi)$.
- For any integrable function $g : \Xi \to \mathbb{R}$ and $w(\xi) := \frac{\rho(\xi)}{\varphi(\xi)}$ we have

$$\mathbb{E}_\rho[g] = \int_\Xi g(\xi)\ \rho(\xi)\ d\xi = \int_\Xi g(\xi) w(\xi)\ \varphi(\xi)\ d\xi = \mathbb{E}_\varphi[gw].$$

- For CVaR_β, perform change of measure, and account for the change by re-weighting:

$$\mathsf{CVaR}_\beta[X] = \frac{1}{1-\beta} \int_{\widetilde{\Xi}} \mathbb{I}_{\mathbb{G}_\beta[X]}(\xi) X(\xi) w(\xi) \varphi(\xi) d\xi$$

- **Assumption:** The support $\widetilde{\Xi}$ of the biasing density φ satisfies

$$\mathbb{G}_\beta[X] \subset \widetilde{\Xi}.$$

Optimal biasing density gives zero variance

Lemma: Under certain conditions, $\widehat{\mathsf{CVaR}}^{\mathsf{IS}}_\beta[X] \to \mathsf{CVaR}_\beta[X]$ w.p. 1 as $n \to \infty$ and

$$\sqrt{n}\left(\widehat{\mathsf{CVaR}}^{\mathsf{IS}}_\beta[X] - \mathsf{CVaR}_\beta[X]\right) \Rightarrow \frac{\left(\mathbb{V}_\varphi[(X(\cdot) - \mathsf{VaR}_\beta[X])_+ \, w(\cdot)]\right)^{1/2}}{1-\beta}\mathcal{N}(0,1).$$

Theorem: The optimal biasing density

$$\varphi^*(\xi) = \frac{\mathbb{I}_{\mathbb{G}_\beta[X]}(\xi)\,(X(\xi) - \mathsf{VaR}_\beta[X])\,\rho(\xi)}{(1-\beta)\,(\mathsf{CVaR}_\beta[X] - \mathsf{VaR}_\beta[X])}$$

results in zero "variance", i.e.,

$$\frac{\mathbb{V}_\varphi[(X(\cdot) - \mathsf{VaR}_\beta[X])_+ \, w(\cdot)]}{n(1-\beta)^2} = 0.$$

- Usual problem: Optimal biasing density depends on $\mathsf{CVaR}_\beta, \mathsf{VaR}_\beta$ which we want to compute. $\Rightarrow$ But helps in finding a good biasing density which results in low variance

Approximation of the optimal IS density via ROMs possible

- Optimal biasing density φ^* motivates the initial choice

$$\varphi_1(\xi) = \frac{\mathbb{I}_{\mathbb{G}_\beta[X]}(\xi)\ \rho(\xi)}{1-\beta}.$$

- Still depends in the risk region of the expensive X
- Use a ROM and ϵ-risk region to get

$$\varphi(\xi) := \frac{\mathbb{I}_{\mathbb{G}^\epsilon_\beta[X_r]}(\xi)\ \rho(\xi)}{\Pr[\mathbb{G}^\epsilon_\beta[X_r]]}.$$

- Since $\mathbb{G}_\beta[X] \subseteq \mathbb{G}^\epsilon_\beta[X_r]$: supp$(\rho) = \widetilde{\Xi} = \mathbb{G}^\epsilon_\beta[X_r] \subseteq$ supp$(\varphi) = \Xi$

Theorem: IS with φ reduces variance compared to MC sampling with ρ by

$$\frac{\mathbb{V}_\varphi\left[\mathbb{I}_{\mathbb{G}_\beta[X]}(\cdot)\,(X(\cdot) - \mathsf{VaR}_\beta[X])\,w(\cdot)\right]}{\mathbb{V}_\rho\left[\mathbb{I}_{\mathbb{G}_\beta[X]}(\cdot)\,(X(\cdot) - \mathsf{VaR}_\beta[X])\right]} \leq \Pr\big[\mathbb{G}^\epsilon_\beta[X_r]\big].$$

Importance sampling gives accurate CVaR_β estimates

- $m = 10^4$ ROM evaluations to explore risk region $\mathbb{G}_\beta^\epsilon[X_r]$
- Acceptance-rejection algorithm to get samples from biasing distribution
- All $\widehat{\text{CVaR}}_\beta^{\text{IS}}[X]$ estimates use $n = 100$ HFM samples, averaged over $K = 100$ runs
- Reference $\text{CVaR}_\beta^{\text{ref}} = \widehat{\text{CVaR}}_\beta^{\text{MC}}[X] = 53.94$ from 10^4 HFM samples

$$\text{MAE} = \frac{1}{K}\sum_{k=1}^{K}\left|\widehat{\text{CVaR}}_\beta^{\text{IS}^{(k)}}[X] - \text{CVaR}_\beta^{\text{ref}}[X]\right|, \quad \text{MRE} = \frac{\text{MAE}}{\left|\text{CVaR}_\beta^{\text{ref}}[X]\right|}\times 100\%$$

- When used with IS, even inaccurate ROMs can give good estimates

	Av $\widehat{\text{CVaR}}_\beta^{\text{IS}}[X]$	MAE	MRE (%)
IS 1	54.02	1.99	3.70
IS 2	54.39	1.59	2.96
IS 3	53.74	1.20	2.23
IS 4	53.94	0.66	1.22

As expected, IS reduces the variance

Estimated variance reduction computed with 100 samples for IS densities $r = 1, 2, 3, 4$.

	$\widehat{\mathbb{V}_\varphi}[\widehat{\mathsf{CVaR}}_\beta^{\mathsf{IS}}[X]]/\widehat{\mathbb{V}_\rho}[\widehat{\mathsf{CVaR}}_\beta^{\mathsf{MC}}[X]]$	$\widehat{\Pr}[\mathbb{G}_\beta^\epsilon[X_r]]$
IS 1	0.2258	0.2463
IS 2	0.1519	0.1771
IS 3	0.0691	0.0967
IS 4	0.0214	0.0519

- **Recall theorem:**

$$\frac{\mathbb{V}_\varphi\left[\,\mathbb{I}_{\mathbb{G}_\beta[X]}(\cdot)\,(X(\cdot) - \mathsf{VaR}_\beta[X])\,w(\cdot)\right]}{\mathbb{V}_\rho\left[\,\mathbb{I}_{\mathbb{G}_\beta[X]}(\cdot)\,(X(\cdot) - \mathsf{VaR}_\beta[X])\,\right]} \leq \Pr\big[\mathbb{G}_\beta^\epsilon[X_r]\big].$$

- As ROMs become more accurate, $\epsilon_r \to 0$, and $\widehat{\Pr}[\mathbb{G}_\beta^\epsilon[X_r]] \to 1 - \beta = 0.05$.

Review and conclusion

Review:

- Showed that using ROMs to sample from for CVaR_β computations is practical + established error bound
- Using ROMs together with importance sampling can produce efficient estimates (ROM only needs to identify risk-region correctly)

Conclusions:

- ROM only has to be accurate in ϵ-risk region $\Rightarrow$ currently working on adaptive ROM construction to make method more efficient
- Computationally, investing in a ROM pays off (shown in paper)

M. Heinkenschloss, B. Kramer, T. Takhtaganov, K. Willcox.
Conditional-Value-at-Risk estimation via Reduced-Order Models.
Submitted. Available as ACDL Technical Report TR 2017-04.