Probabilistic Damage Tolerance Analysis Using Adaptive Multiple Importance Sampling

Juan D. Ocampo[1,2], Nathan Crosby[2], Harry Millwater[3], Michael Reyer[4], Sohrob Mottaghi[4], Marv Nuss, Beth Gamble, and Christopher Hurst [5]

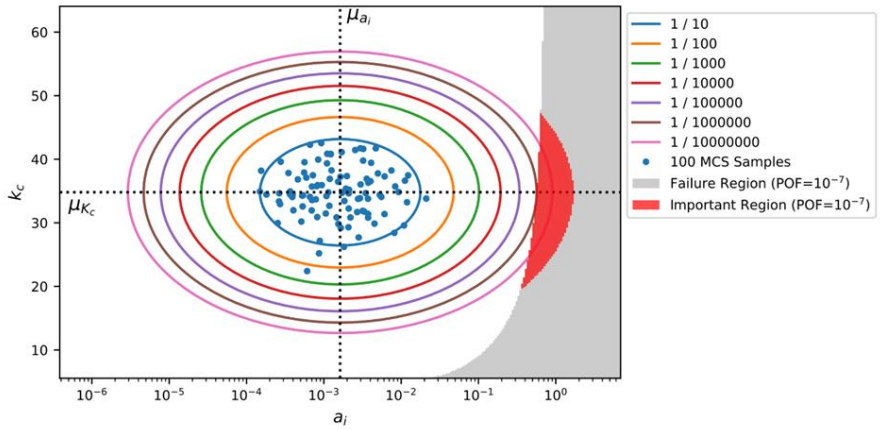
[1] St. Mary's University, [2] AeroMatter, LLC., [3] University of Texas at San Antonio,
[4] Federal Aviation Administration,[5] Textron Aviation

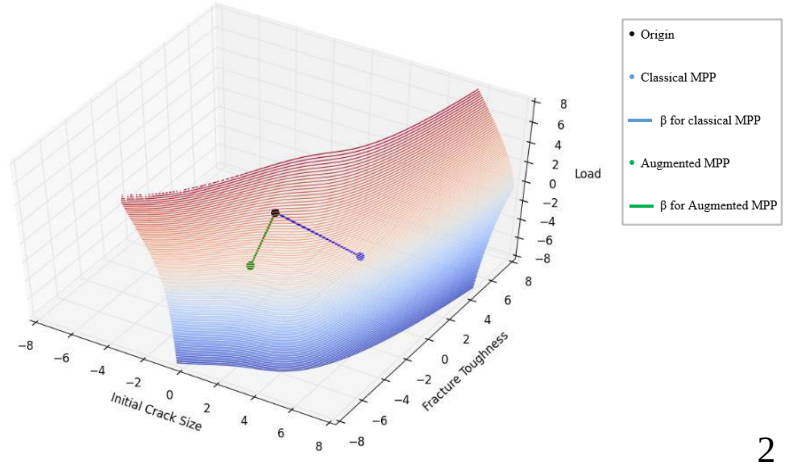
ICAF
International Committee on Aeronautical Fatigue and Structural Integrity

31st ICAF Symposium – Delft, 26-29 June 2023

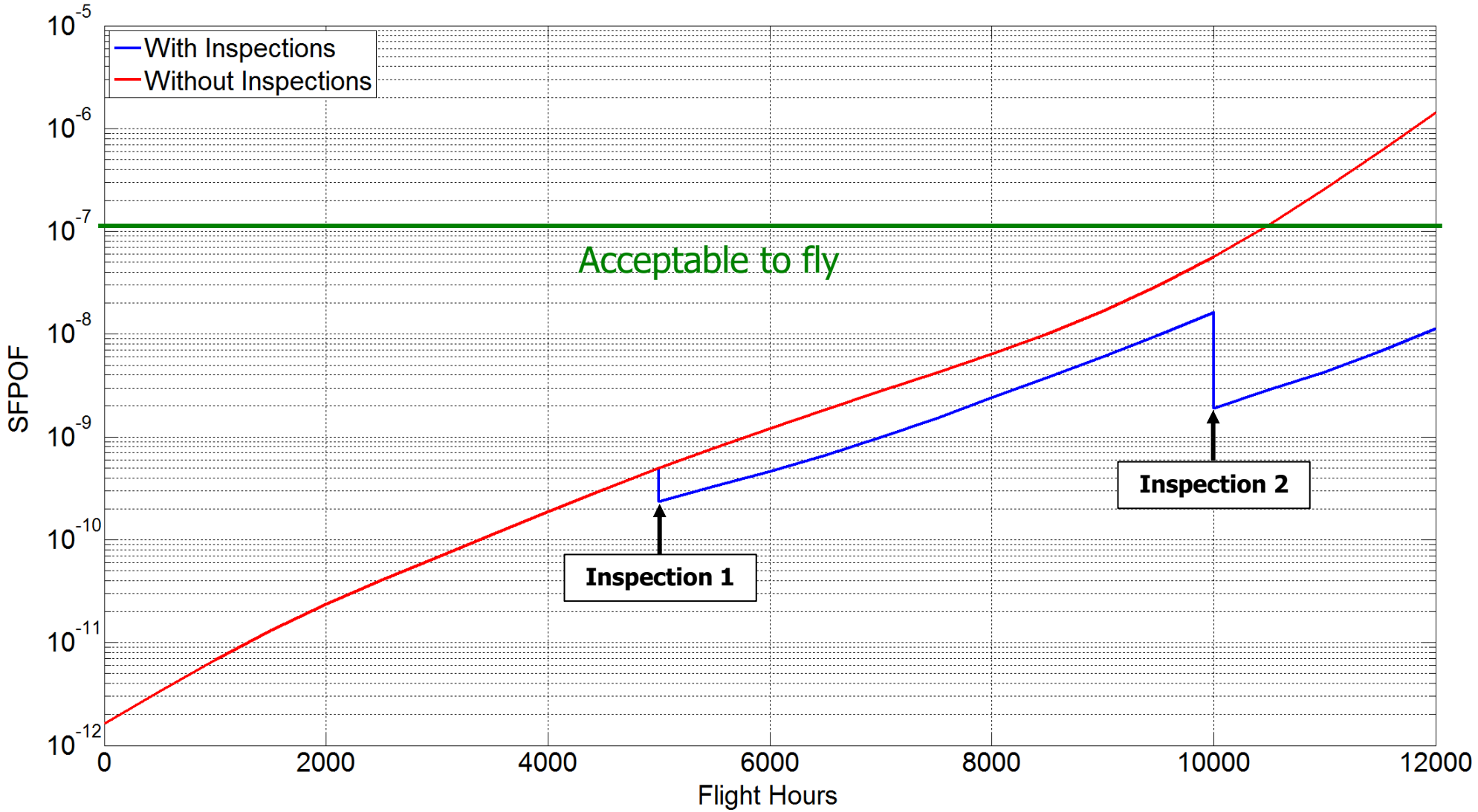
Probabilistic Risk Assessment

- Probabilistic Risk Assessment is an important tool for ensuring structural integrity of aircraft components.
- Based on the principles of probabilistic damage tolerance analysis.
- The Single Flight Probability-of-Failure is difficult to compute accurately and efficiently due to several challenges:
  - Very small probabilities, e.g., 1E-7 or smaller
    - Standard Monte Carlo sampling is impractical
  - Inspection and repair process results in multi-modal crack size distributions
    - FORM/SORM methods are impractical
  - Inspection optimization requires multiple analyses
    - Efficient reanalyses are required

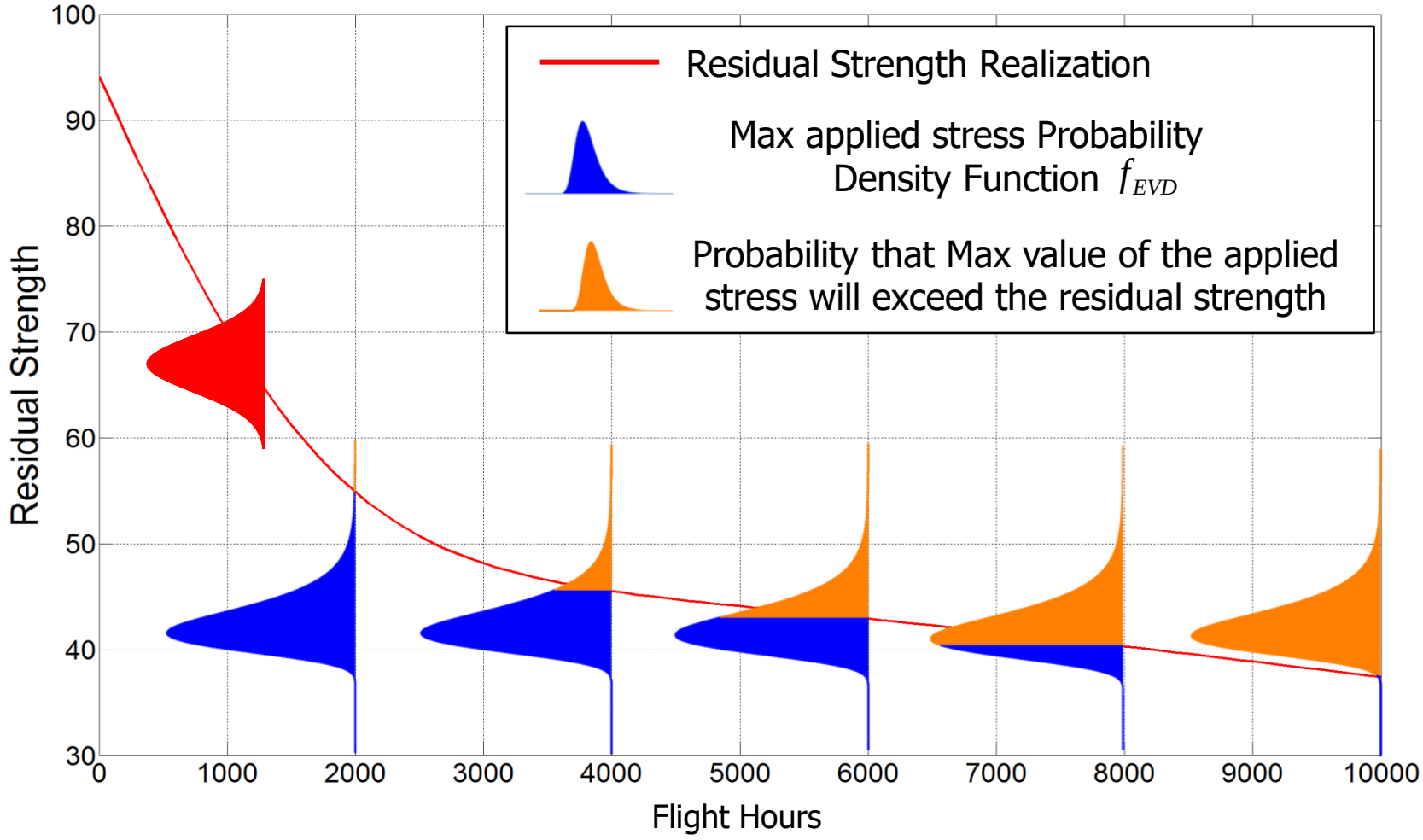




Probabilistic Risk Assessment



Probability of Failure Calculation



$$POF(t) = P[\sigma_{MAX} > \sigma_{RS}(t)]$$

Probability Equations

The probability-of-failure is the probability that maximum value of the applied stress (during the next flight) will exceed the residual strength σ_{RS} of the aircraft component.

$$POF_{Lincoln}(t) = P[\sigma_{Max} > \sigma_{RS}(t)] = \int_{\sigma_{Max} > \sigma_{RS}} [1 - F_{EVD}(\sigma_{RS}(\boldsymbol{x}, t_n))] f_X(\boldsymbol{x}) dx$$

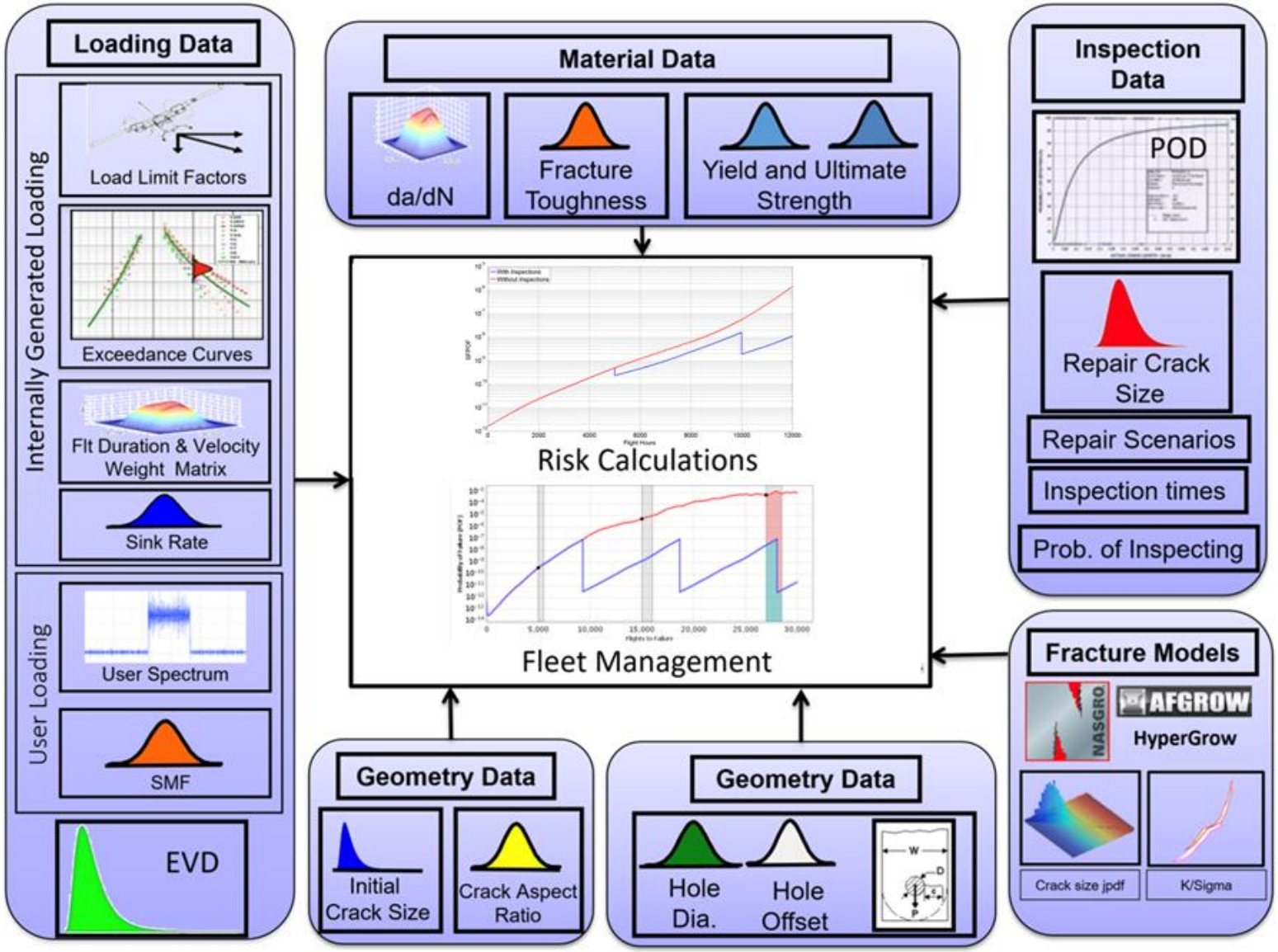
Other random variables

Survival to time t

$$POF_{Freudenthal}(t_n) = \frac{\int [\prod_{i=1}^{n-1} F_{EVD}(\sigma_{RS}(\boldsymbol{x}, t_i))][1 - F_{EVD}(\sigma_{RS}(\boldsymbol{x}, t_n)] f_X(\boldsymbol{x}) dx}{\int \prod_{i=1}^{n} F_{EVD}(\sigma_{RS}(\boldsymbol{x}, t_i)) f_X(\boldsymbol{x}) dx}$$

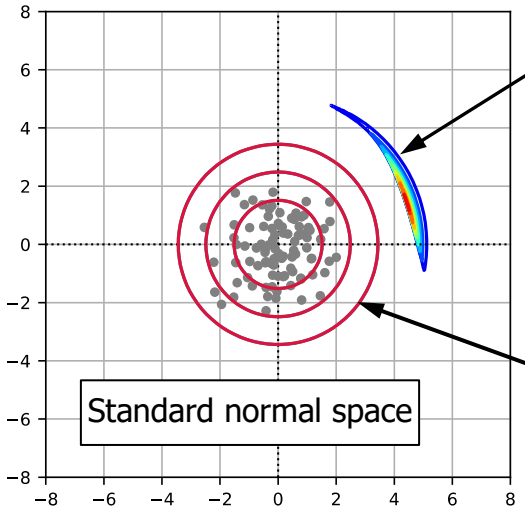
F_{EVD} – CDF of the maximum stress per flight (extreme value distribution)
σ_{RS} – residual strength

SMART |DT


Loading Data
Internally Generated Loading
Load Limit Factors
Exceedance Curves
Flt Duration & Velocity Weight Matrix
Sink Rate
User Loading
User Spectrum
SMF
EVD
Material Data
da/dN
Fracture Toughness
Yield and Ultimate Strength
Risk Calculations
Fleet Management
Geometry Data
Initial Crack Size
Crack Aspect Ratio
Geometry Data
Hole Dia.
Hole Offset
Inspection Data
POD
Repair Crack Size
Repair Scenarios
Inspection times
Prob. of Inspecting
Fracture Models
NASGRO
AFGROW
HyperGrow
Crack size jpdf
K/Sigma

Importance Sampling

Monte Carlo Sampling

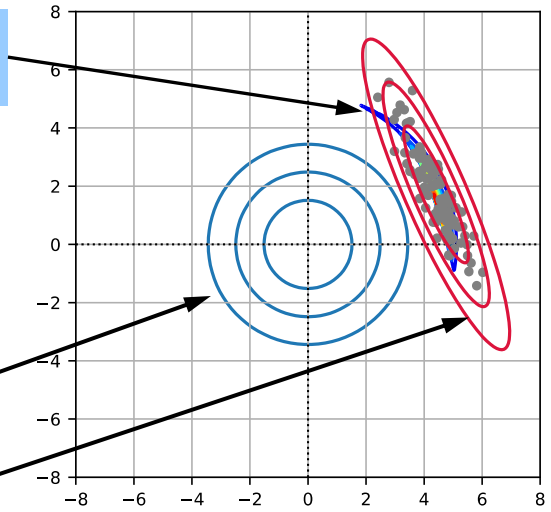


region of importance

$f(\boldsymbol{x})$

$q(\boldsymbol{x})$

Importance Sampling



- Define $H(\boldsymbol{x};t) = 1 - F_{\mathrm{EVD}}\left(\sigma_{\mathrm{RS}}(\boldsymbol{x};t)\right)$

$\boldsymbol{x}$ – random variables:

- initial defect size
- fracture toughness
- dadN variability
- geometric parameters
- etc.

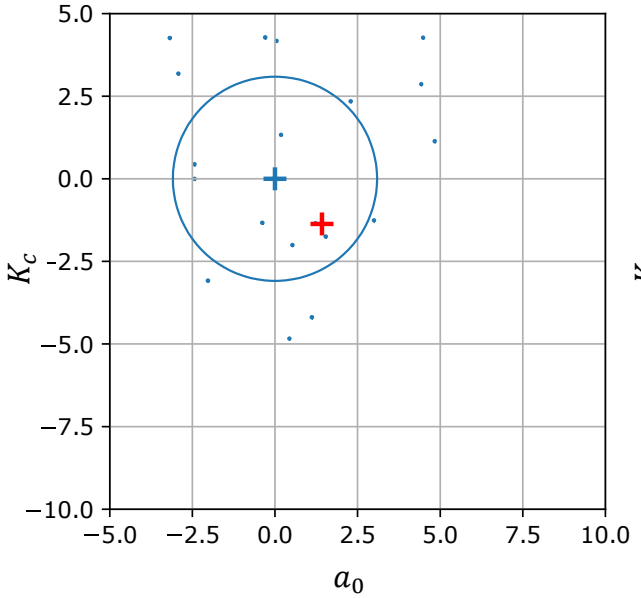
$$\mathrm{E}[H(\boldsymbol{x},t)] = \int H(\boldsymbol{x},t)\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \qquad \mathrm{E}[H(\boldsymbol{x},t)] = \int H(\boldsymbol{x},t)\, \frac{f(\boldsymbol{x})}{q(\boldsymbol{x})}\, q(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

$$\hat{E}[H(\boldsymbol{x},t)] = \frac{1}{N}\sum_i H(\boldsymbol{x}_i,t) \qquad \hat{E}[H(\boldsymbol{x},t)] = \frac{1}{N}\sum_i H(\boldsymbol{x}_i,t)\, w(\boldsymbol{x}_i)$$

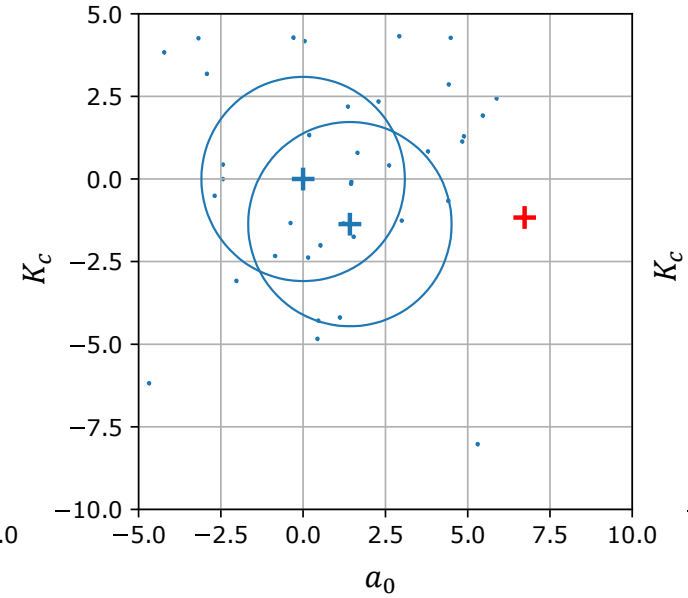
Draw samples from q

Importance weight
$w(\boldsymbol{x}_i) = f(\boldsymbol{x}_i) \,/\, q(\boldsymbol{x}_i)$

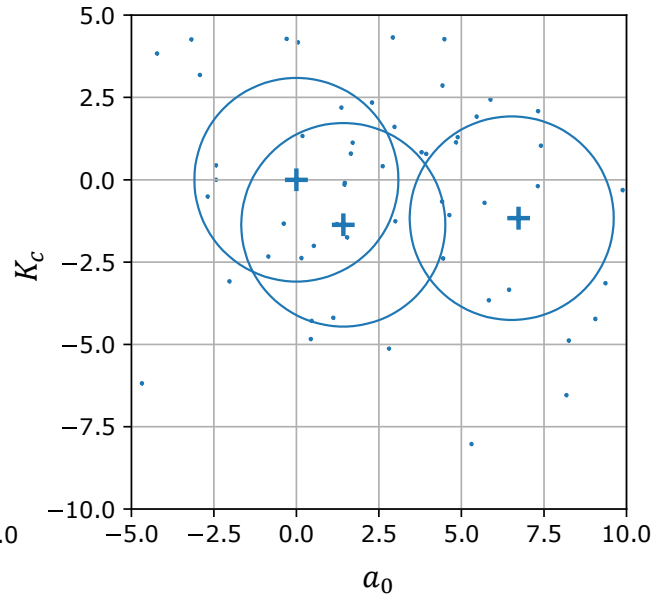
Phase I: Initialization



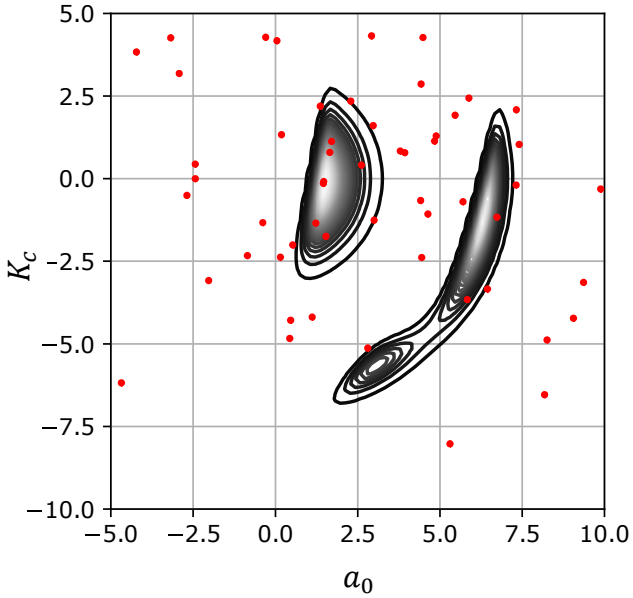
20 samples



40 samples



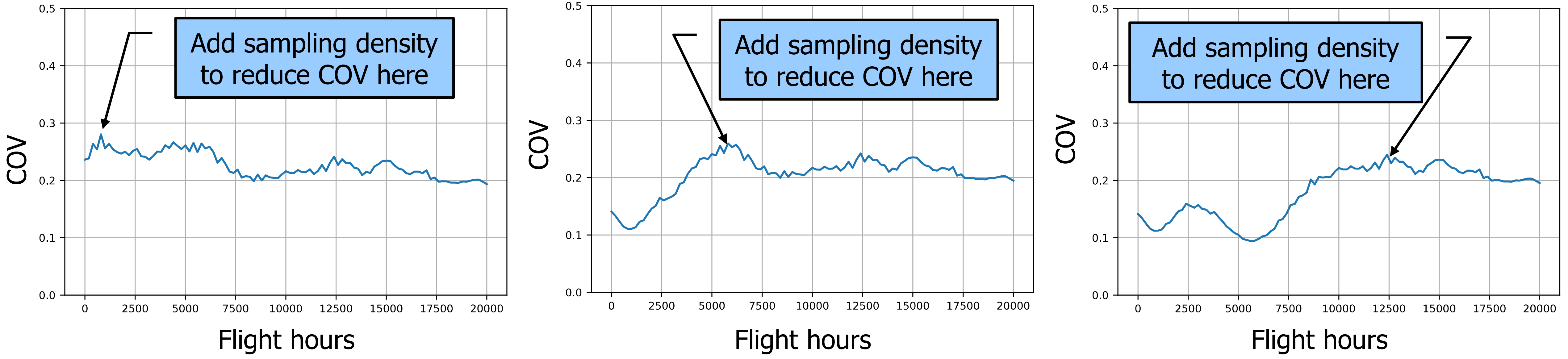
60 samples



Comparison w important regions

- Goal is to locate important regions in standard normal space.
- Generate samples near and around the important regions for all evaluation times.
- Performs exploration to find the location of important regions.
- The adaptation phase will focus on determining the scale and shape of important regions.

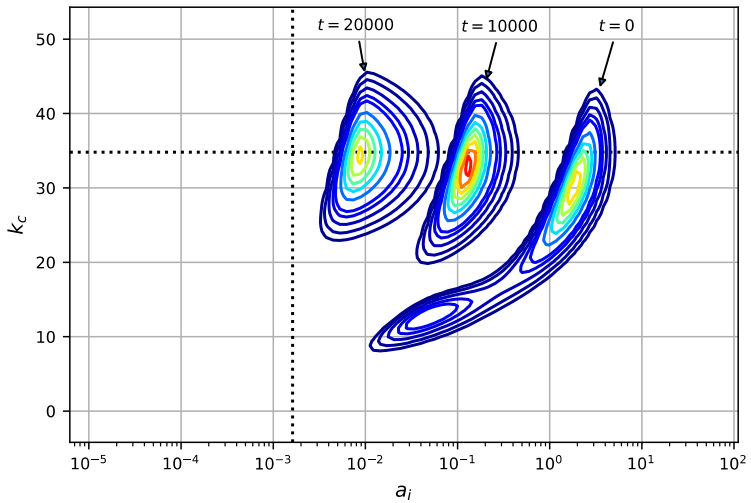
Phase II: Adaptation



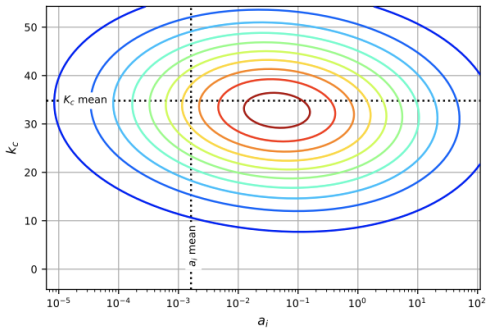
- Determine evaluation times at which to focus samples. Note, near-by times also obtain improved results.
- Use Coefficient of Variation (COV) which is a normalized error estimate.
- Ensures COV across all evaluation times is below a user-defined threshold.

Adaptive Multiple Importance Sampling Approach

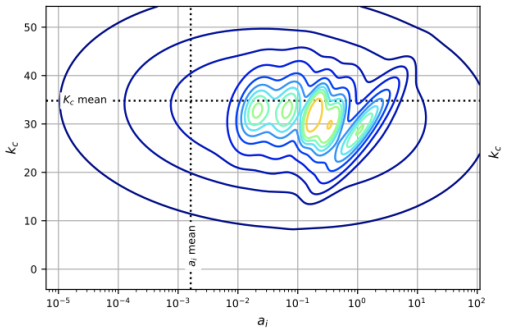
Individual evaluation time important regions



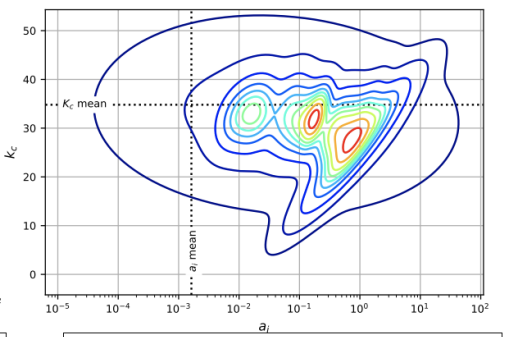
Combined important region



Mixture density (5 components)



Mixture density (10 components)



Mixture density (15 components)

- Approximate the averaged or combined important region using a mixture density composed of multivariate normal sampling densities optimized for individual evaluation times.
- Key advantage is that samples can be used for more than one important region where regions overlap.

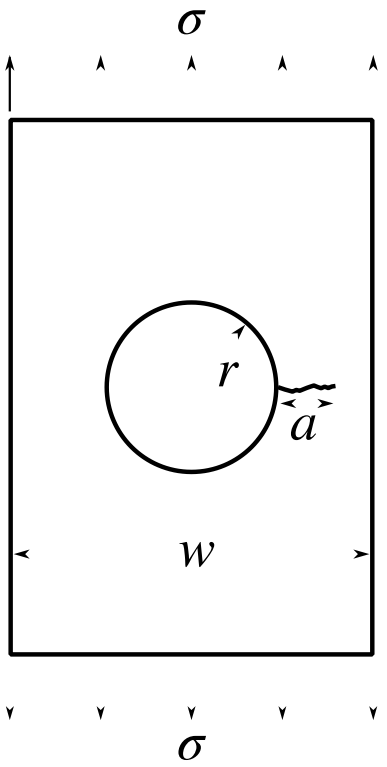
$$E[H(x,t)] = \frac{1}{N_{mix}\, N_{samp}} \sum_{j=1}^{N_{mix}} \sum_{i=1}^{N_{samp}} H(\boldsymbol{x}_{ij}, t)\, \frac{f(\boldsymbol{x}_{ij})}{(1/N_{mix}) \sum_{k=1}^{N_{mix}} q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_k)}$$

N.Crosby, "Efficient Adaptive Importance Sampling Estimation of Time Dependent Probability of Failure with Inspections for Damage Tolerant Aircraft Structures," PhD dissertation, University of Texas at San Antonio, 2021

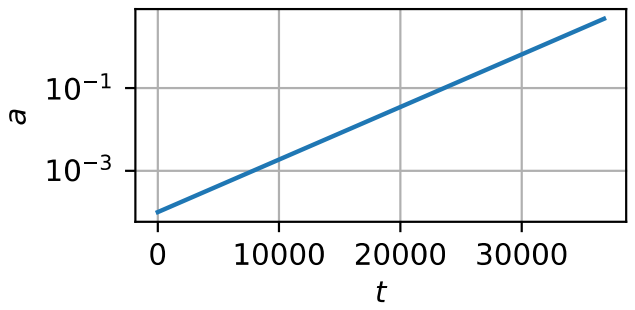
Example Problems

- Risk assessment handbook example using a closed-form crack growth equation.
- General aviation example with inspections.

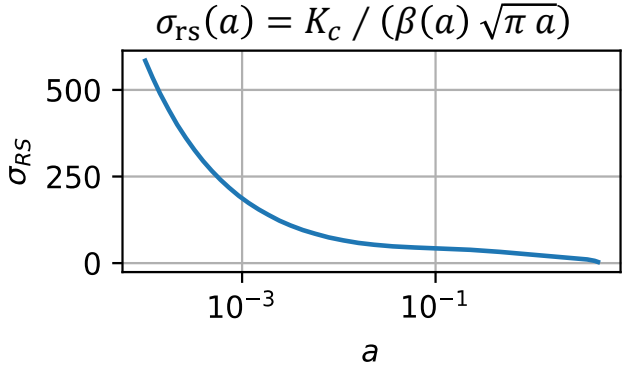
Risk Assessment Handbook Problem



$$a(t) = a_0 \cdot \exp(2.93 \times 10^{-4}\, t)$$



$$\sigma_{rs}(a) = K_c \,/\, (\beta(a)\sqrt{\pi\, a})$$



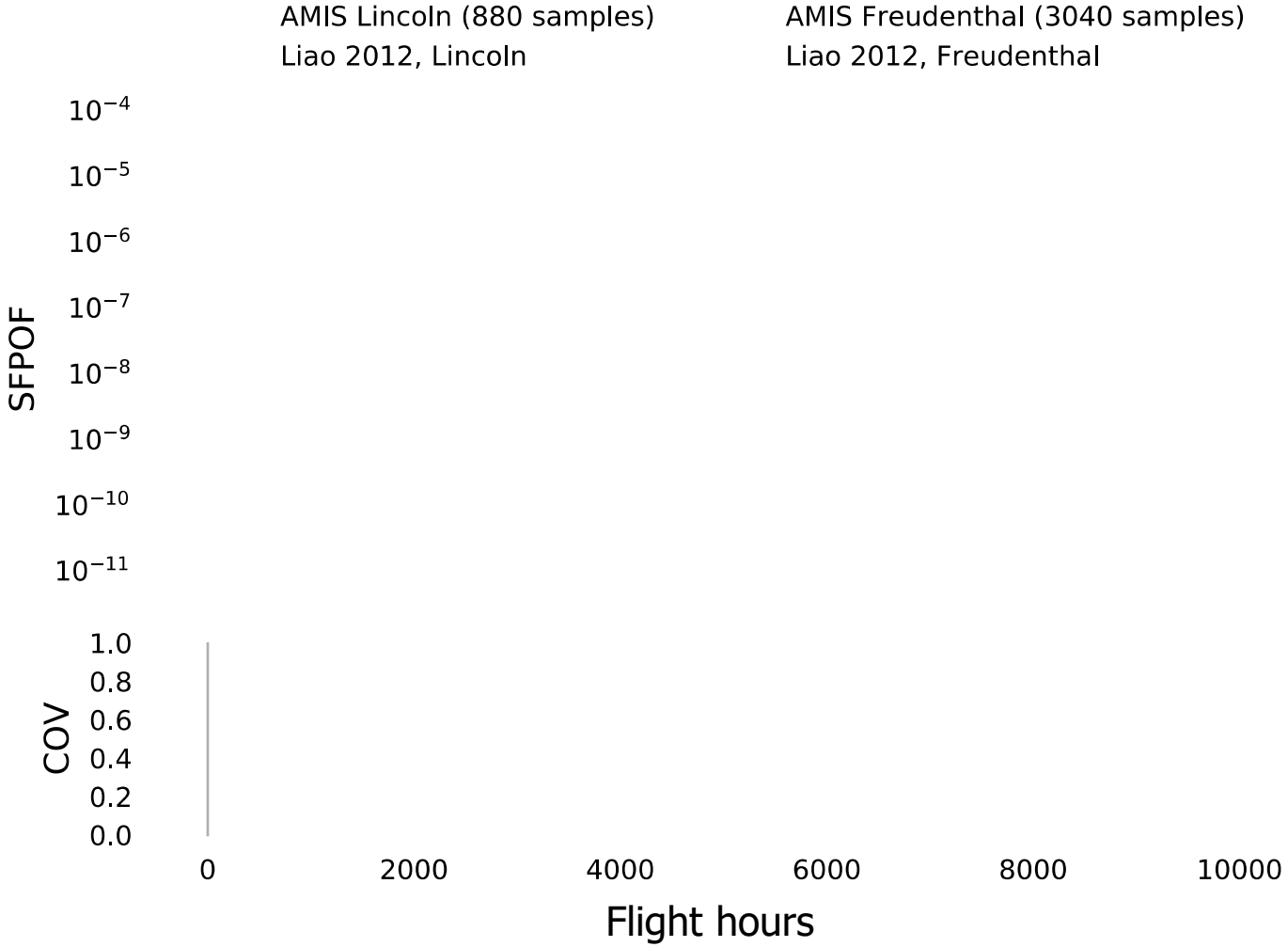
Parameter	Value
Width	Deterministic 10 in
Radius	Deterministic 0.125 in
Initial Crack Size	$LN(0.0032, 0.0047)$ in
Fracture Toughness	$N(34.8, 3.90)$ ksi $\sqrt{\text{in}}$
Maximum Stress per Flight	$W(5.0{,}10.0, 5.0)$ ksi

$$\beta(a) = \underbrace{\left(0.6762 + \frac{0.8734}{0.3254 + a\,/\,R}\right)}_{\beta_{\text{hole}}} \cdot \underbrace{\sqrt{\sec\left(\frac{\pi\,(R+a)}{W}\right)}}_{\beta_{\text{width}}}$$

Tuegel et al., Aircraft structural reliability and risk analysis handbook volume 1: Basic analysis methods., Technical report, Air Force Research Lab, Wright-Patterson AFB, OH, Aerospace Systems Dir, 2013

POF Results

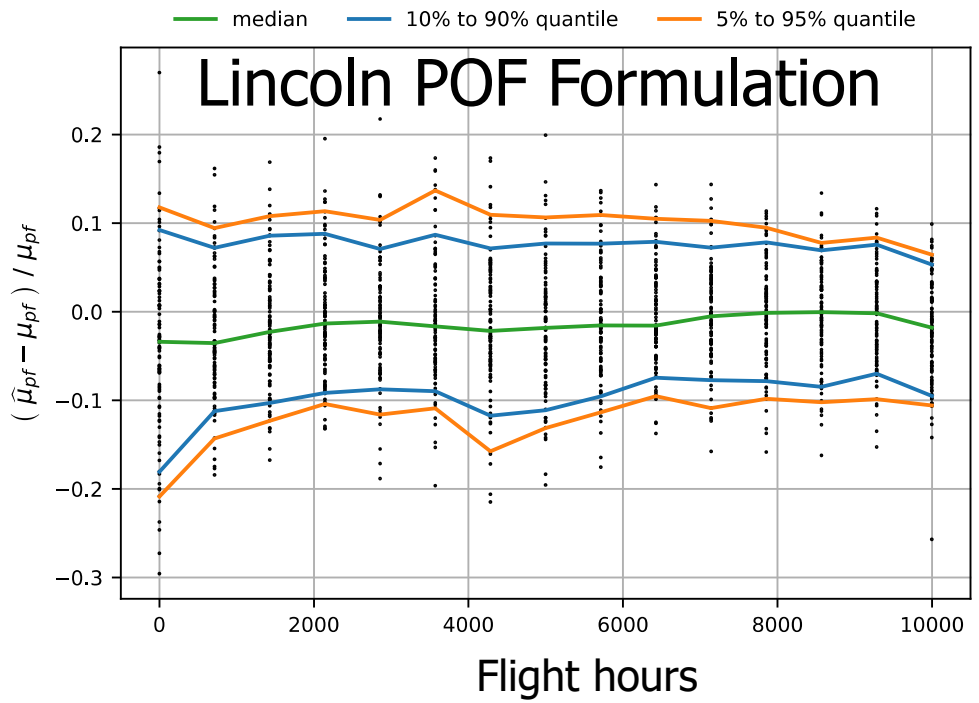
Independent verification

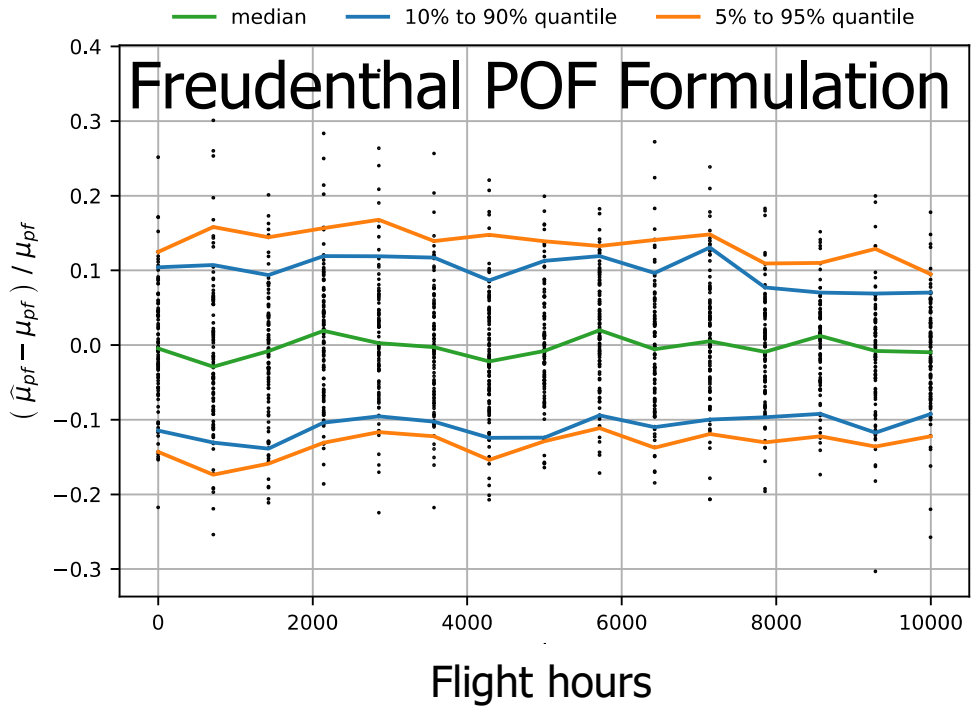


- 15 evaluation times
- COV threshold 0.1
- Lincoln Formulation
  - (assumes survival = 1 from flight 0 to flight t)
  - 80 samples per iteration
  - 11 iterations
  - 880 samples
- Freudenthal Formulation
  - (does not assume survival = 1 from flight 0 to flight t)
  - 160 samples per iteration
  - 19 iterations
  - 3040 samples

Liao M., Comparison of different single flight probability of failure (SFPOF) calculations for aircraft structural risk analysis. In Aircraft Airworthiness and Sustainment (AA&S) Conference, 2012

PDTA AMIS Accuracy of Error Estimates





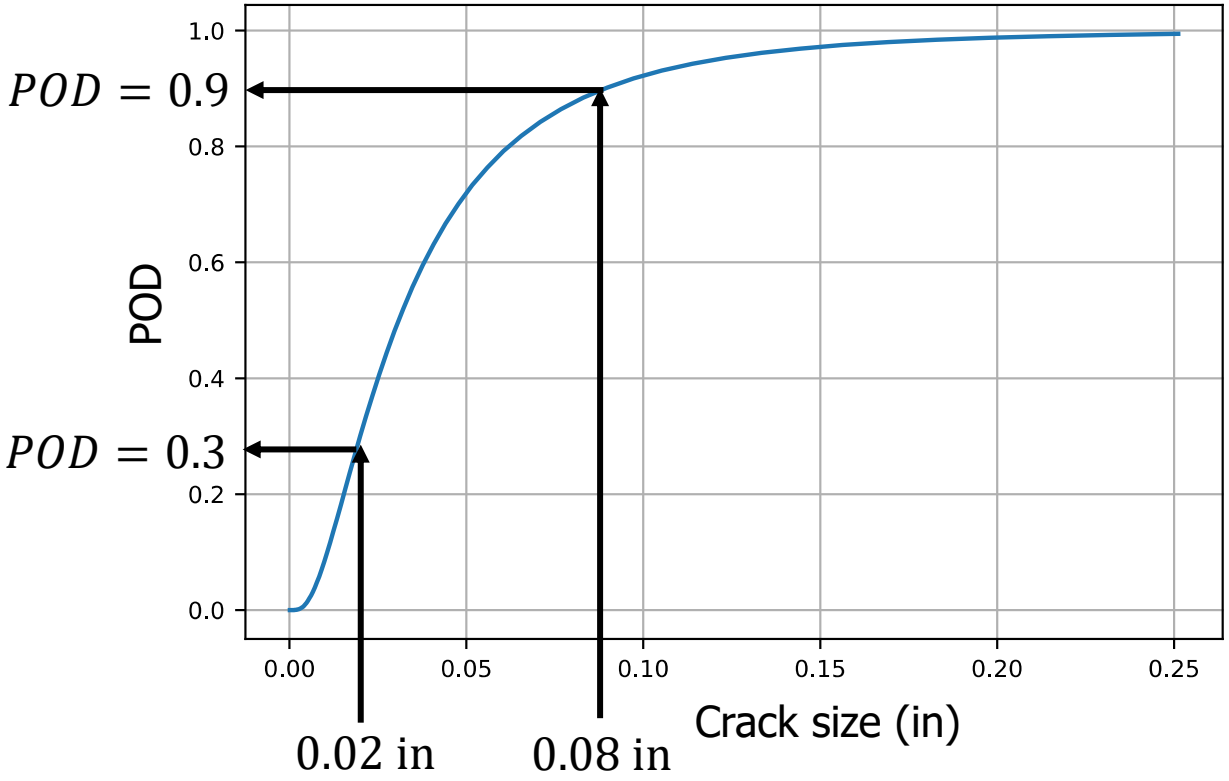
- Variations calculated for 100 PDTA AMIS runs.
- For both Lincoln and Freudenthal POF Formulations.
  - PDTA AMIS estimates are within the expected error bands, showing the sampling variance gives a good indication of estimator error.
  - PDTA AMIS median error is close to 0, showing the estimates are consistent.

POF Inspections

Number of inspections before t

$$PND(t) = \prod_{i=1}^{N_I(t)} \left[1 - POD\left(a(\tau_i)\right)\right]$$

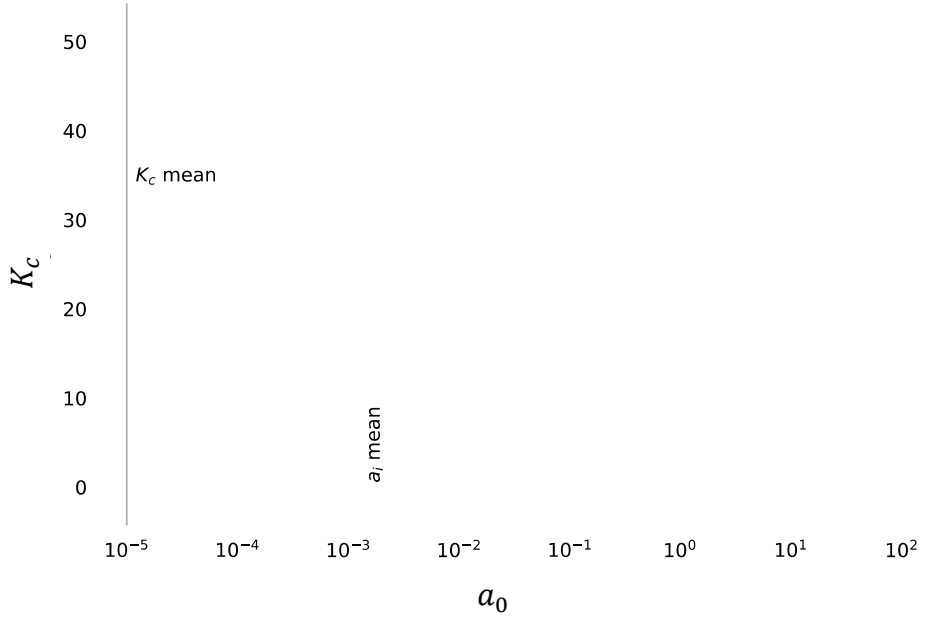
$$POF_{\text{no-surv}}(t) = \int PND\left(a(t)\right)\left[1 - F_{EVD}\left(\sigma_{RS}(t)\right)\right] f_{\boldsymbol{X}}\, \mathrm{d}\boldsymbol{x}$$



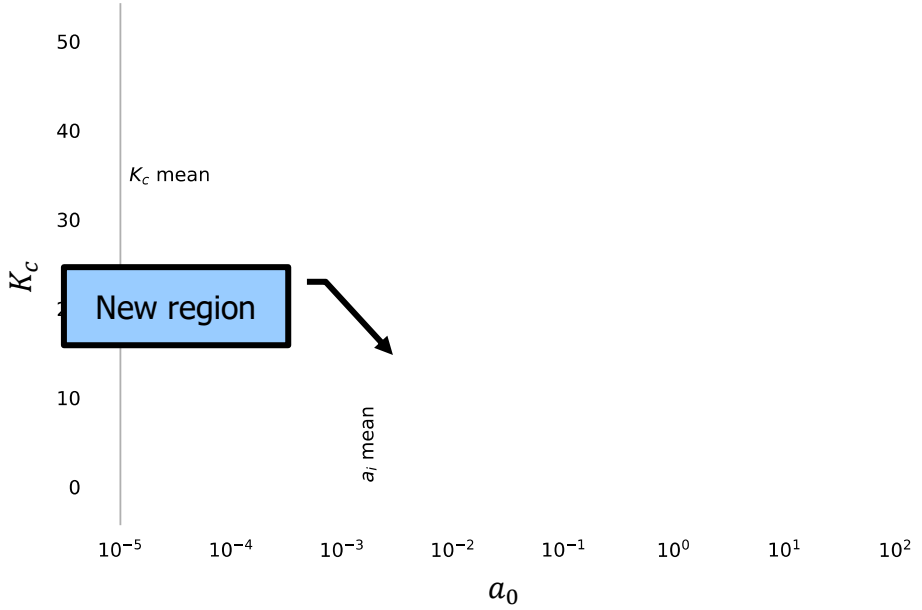
- Inspections are not deterministic – there is some probability of missing cracks
- In PDTA, this is modeled by reducing the probability of failure proportional to undetected cracks
- PND is the probability of not detecting a crack in any inspection(s) before t

Change in Combined Important Region Due to Inspection

Combined Important Region without inspection

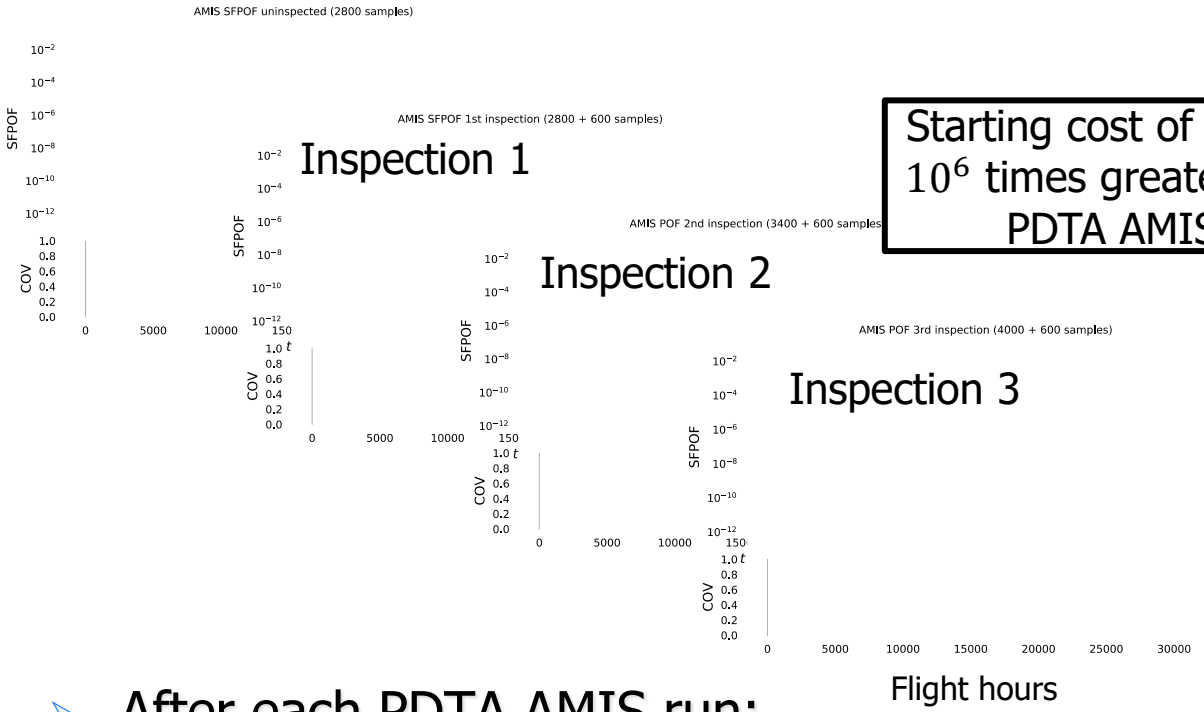


Combined Important Region with one Inspection

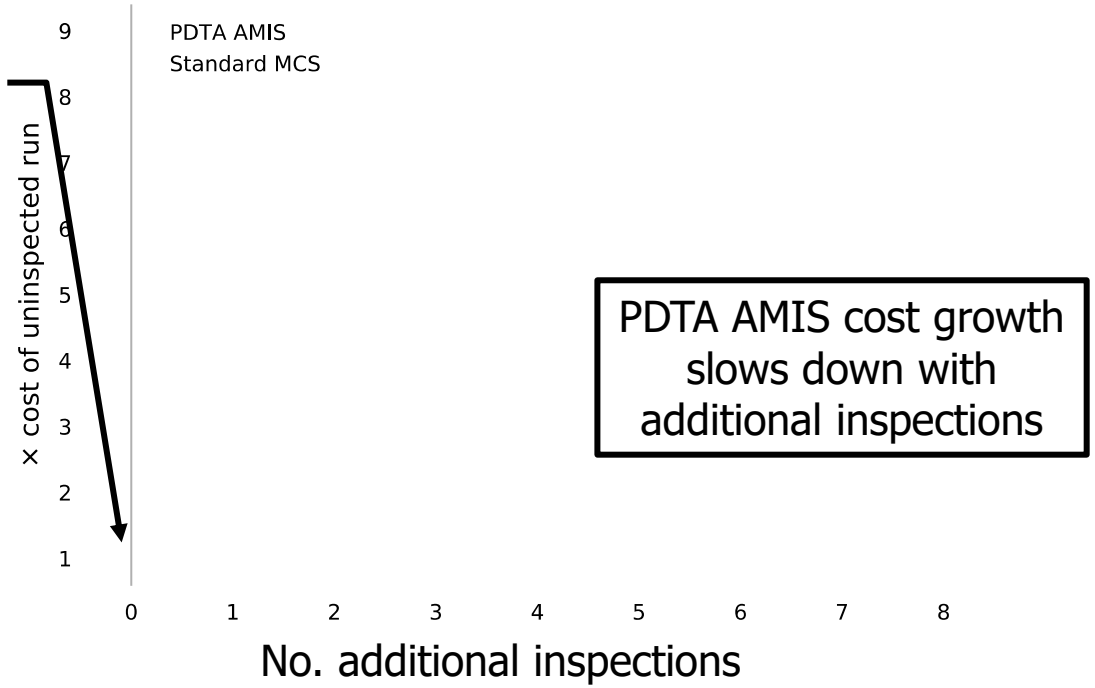


- Post-inspection, a new important region emerges around ($a_i = 0.007$, $k_c = 12.5$).
- Stored crack growth analyses reevaluated with the modified response function including an inspection provide a good general idea of the new important region location

Adding Inspections One-at-a-Time



Starting cost of SMC is 10^6 times greater than PDTA AMIS

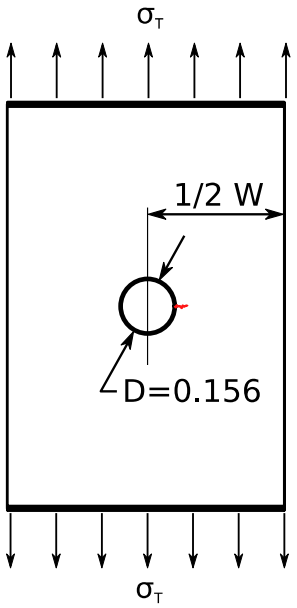


PDTA AMIS cost growth slows down with additional inspections

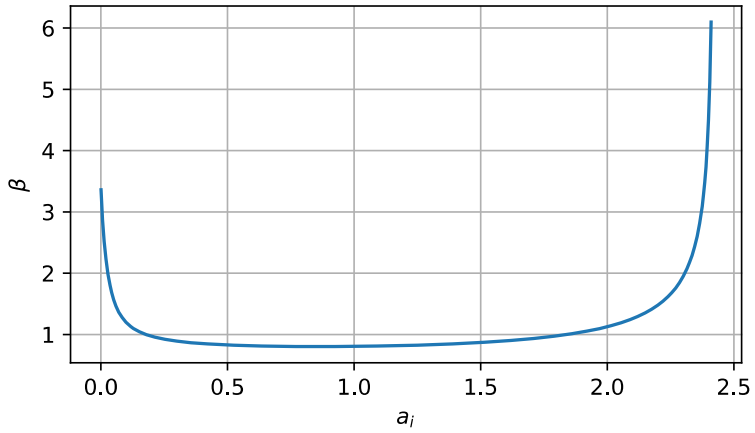
- After each PDTA AMIS run:
  - Update conditional POF, $H(\cdot)$, to include new inspection time in PND function
  - Recalculate $H(\cdot)$ for all samples over all times with existing crack growth evaluations
  - Re-run PDTA AMIS adaptation

- PDTA AMIS only has to add crack growth evaluations to adapt for the new inspection
- SMC must rerun all of the crack growth evaluations

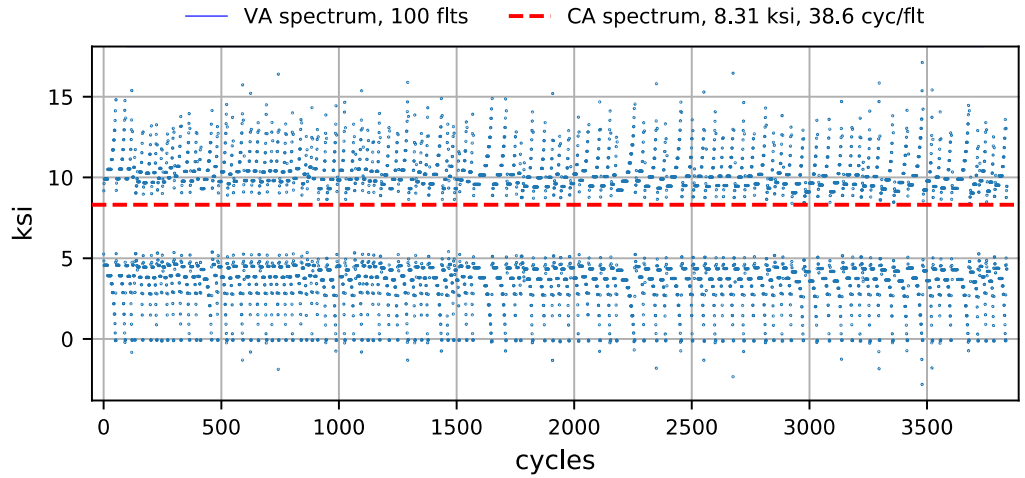
General Aviation Example Problem



Beta table

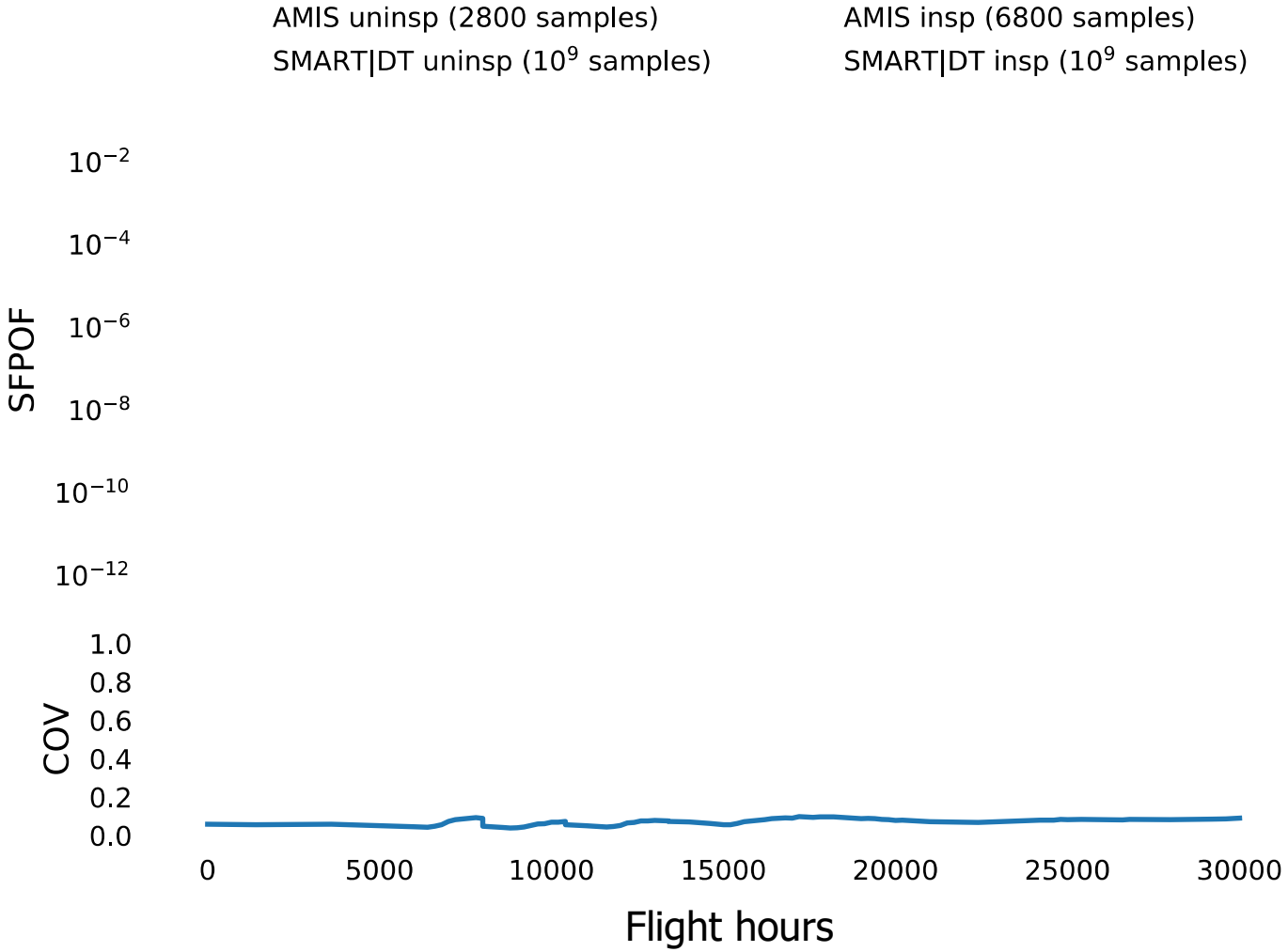


Spectrum



Parameter	Values
Width	Deterministic 5 in
Thickness	Deterministic 0.125 in
Log Paris Constant	$N(-9.0, 0.08)$
Paris Exponent	Deterministic 3.8
Initial Crack Size	$W(0.45, 4.17 \times 10^{-5})$ in
Fracture Toughness	$N(35.0, 3.5)$ ksi $\sqrt{\text{in}}$
Maximum Stress per Flight	$EVD(13.4, 1.3, 0.07)$ ksi
Probability of Detection	$LN(0.05, 0.065)$ in
Repair Quality (Crack Size)	Perfect

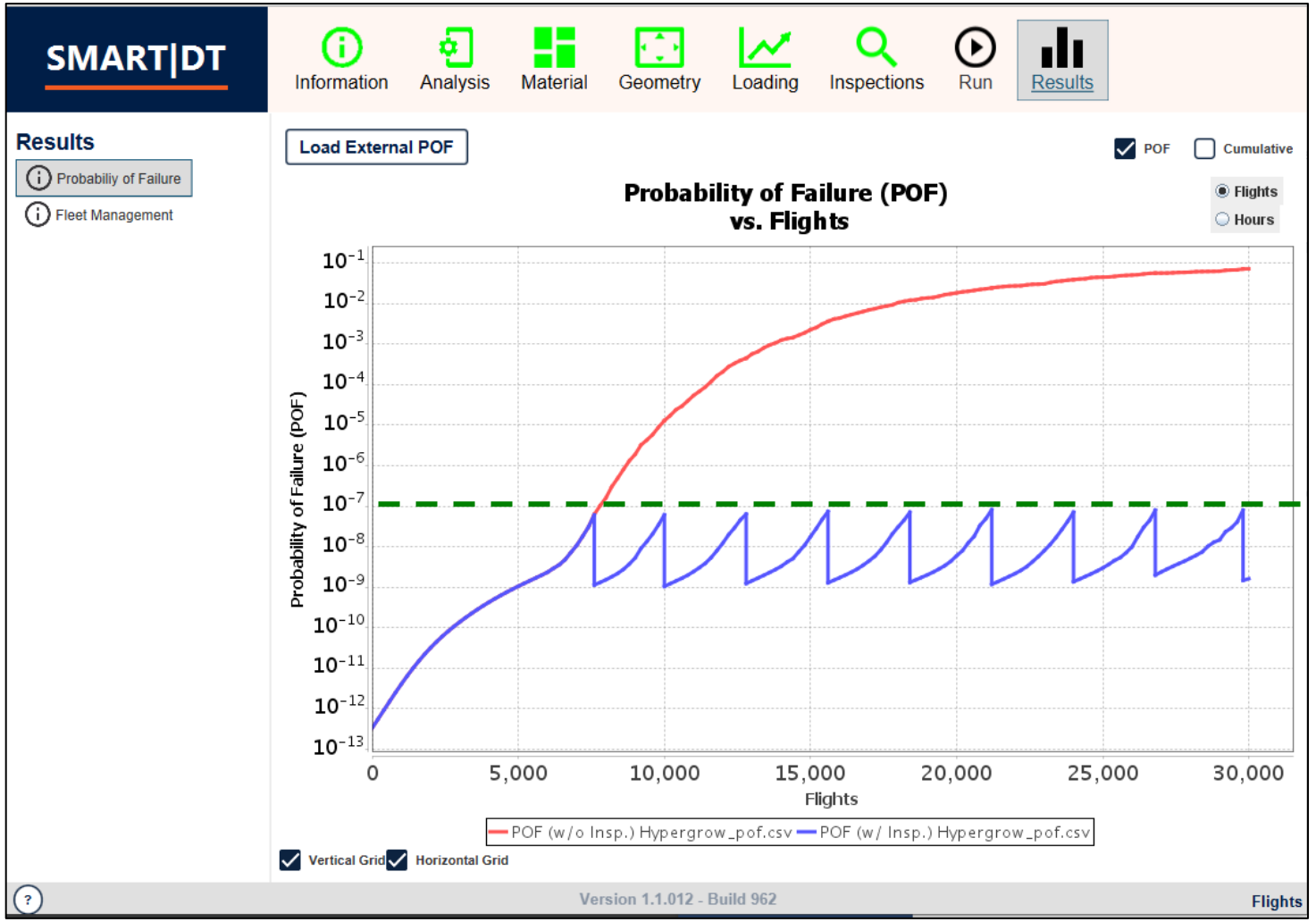
POF Results After Adding 8 Inspections



- PDTA AMIS
  - 2800 samples for uninspected POF
  - 6800 samples for inspected POF after adding 8 inspections
- PDTA AMIS in excellent agreement with SMC using 10^9 samples

POF Results with Opt. Inspections

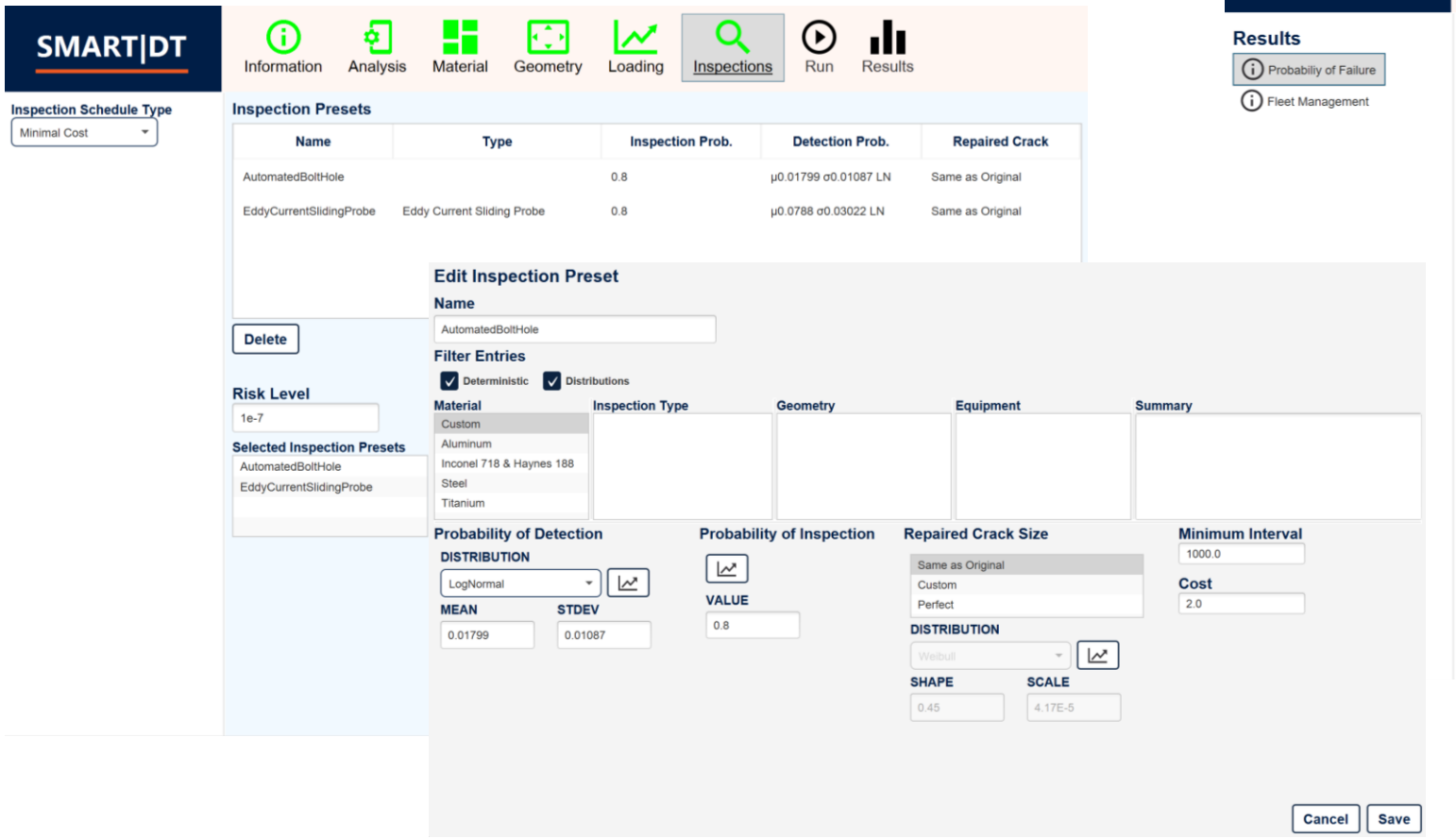
- Risk-Threshold – Find the minimal set of inspections that maintains risk below a given threshold using a single inspection type
  - Keep risk below 1E-7

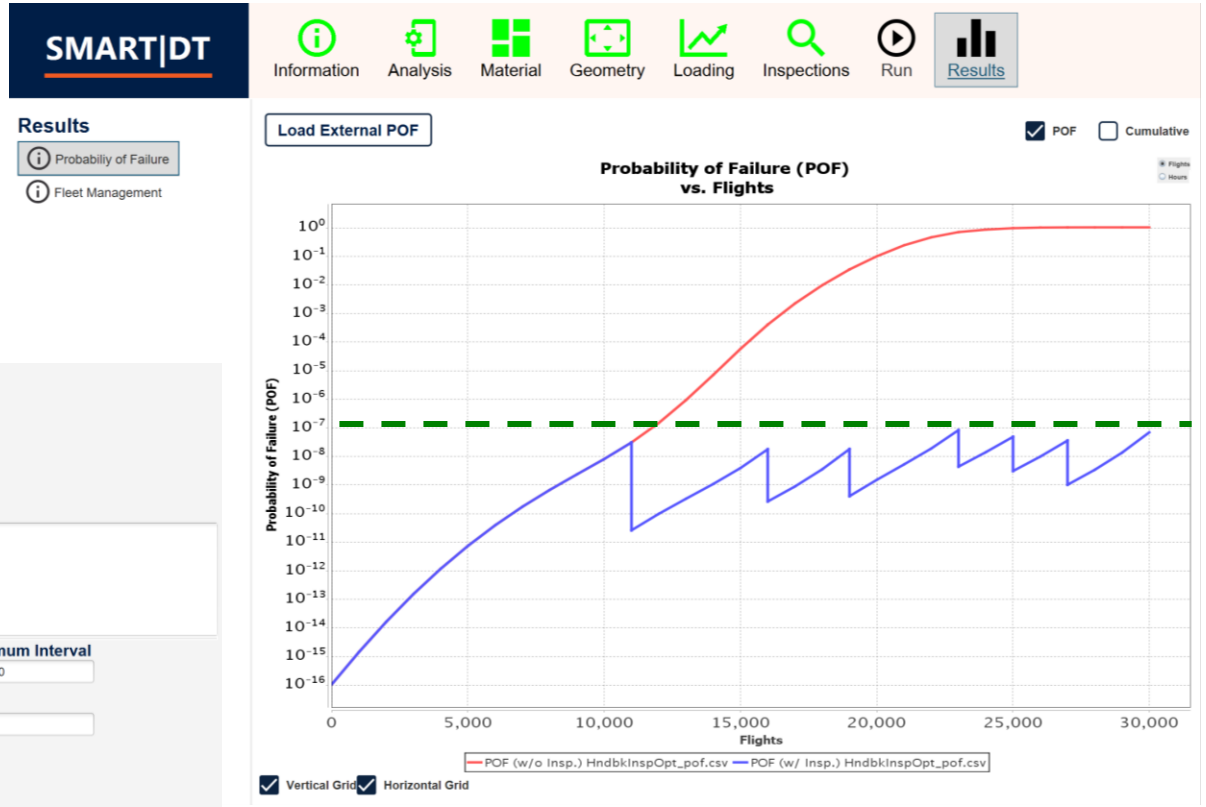


Juan Ocampo, Nathan Crosby, Harry Millwater, Chris Hurst, Beth Gamble, and Marv Nuss, "Fleet Management Considering Inspection Schedule Optimization", Aircraft Airworthiness & Sustainment Conf., Ponte Verde, FL, August 2022

POF Results with Opt. Inspections

- Minimal-Cost – Find the minimum cost set of inspections that maintains risk below a given threshold using multiple inspection types





Summary

- The AMIS algorithm estimates POF for risk assessment using 6 orders of magnitude fewer samples compared to standard Monte Carlo sampling for probabilities of 10^{-7} with COV of 0.1.

- This efficiency allows for the incorporation of additional random variables into the problem and the use of more realistic fracture mechanics solutions.

- The SMART methodology in combination with AMIS is not limited to aircraft fleet COS and can also be applied to digital twin modeling, virtual testing, and other novel applications

Future Developments

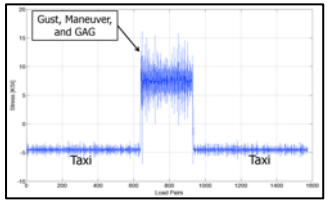
- Optimized inspection schedule
  - Determine the inspection times and inspection methods to keep the risk below a user-defined threshold with minimum cost.
- Bayesian Module Implementation for Sensor Integration
  - Update crack size simulations with sensor (POD) data.
- Probabilistic damage tolerance analysis of more realistic structures
  - Continuing damage, multisite damage, residual stresses, out-of-plane crack growth, etc.
- Approaches
  - NASGRO interface
  - Surrogate models
    - Machine learning approaches, e.g., Bingo software, BAMF, etc.

Smart|DT Software

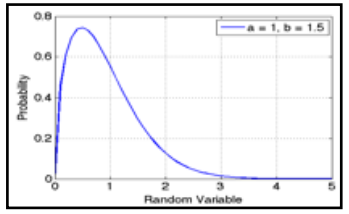
- Probabilistic risk assessment development has been funded by the US Federal Aviation Administration to develop the Smart|DT software.
- Available to the general public.
- Training presented annually and available online:
  - Aircraft Airworthiness Conference
  - https://smartdtsoftware.wixsite.com/smart

Web site link

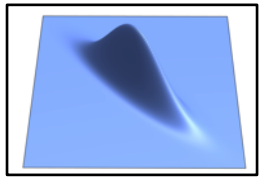
Loading Generation



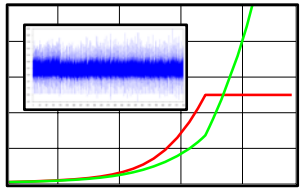
EVD Dist



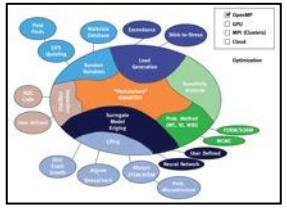
dadN variability



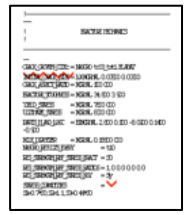
HyperGrow



Interoperability



Scriptable



Acknowledgements

- Probabilistic Damage Tolerance-Based Maintenance Planning for Small Airplanes, Federal Aviation Administration, Grant 09-G-016
- Probabilistic Fatigue Management Program for General Aviation, Federal Aviation Administration, Grant 12-G-012
- Probabilistic Modeling of Random Variables and K-Solution Developments for General Aviation Grant 16-G-005
- Advances in Probabilistic Damage Tolerance Analysis using the Smart/DT Software, Cooperative Agreement 692M152140011
  - **Sohrob Mattaghi (FAA Tech Center) – Program Manager**
  - **Michael Reyer (Kansas City Office) – Sponsor**
  - **Michael Gorelik – Chief Scientific and Technical Advisor for Fatigue and Damage Tolerance**

Thank you

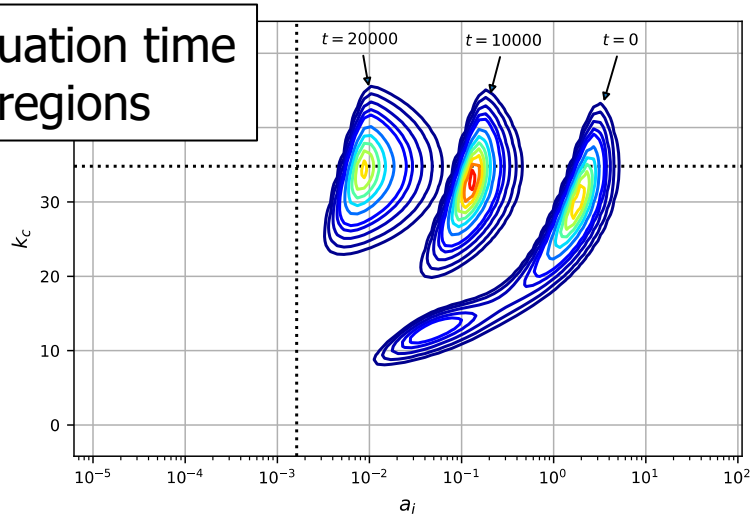
Web site link

Juan Ocampo
jocampo@stmarytx.edu
+1-210-995-5792

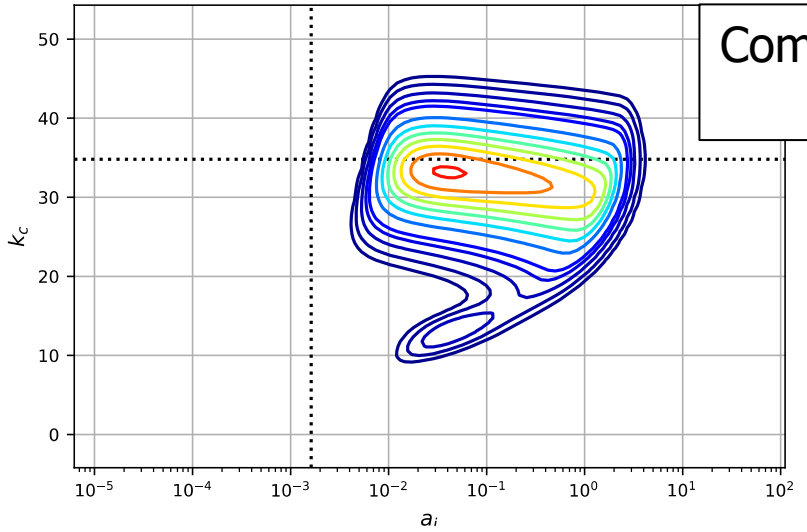
Backup Slides

Adaptive Multiple Importance Sampling Approach

Individual evaluation time important regions



Combined important region



$$E[H(x,t)] = \sum_{j=1}^{N_{mix}} \frac{1}{N_{samp}} \sum_{i=1}^{N_{samp}} H(\boldsymbol{x}_{ij}, t)\, \omega_j(\boldsymbol{x}_{ij})\, \frac{f(\boldsymbol{x}_{ij})}{g(\boldsymbol{x}_{ij}, \theta_j)}$$

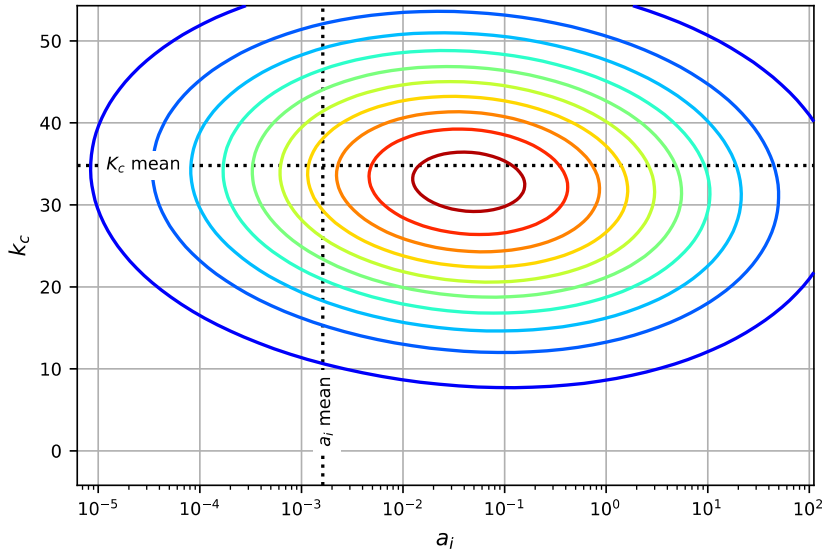
$$\omega_m^{std}(\boldsymbol{x}_{ij}) = \begin{cases} 1, & m = j \\ 0, & otherwise \end{cases}$$

$$\omega_m^{bh}(\boldsymbol{x}_{ij}) = \frac{n_m\, q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_m)}{\sum_{k=1}^{N_{mix}} n_k\, q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_k)}$$

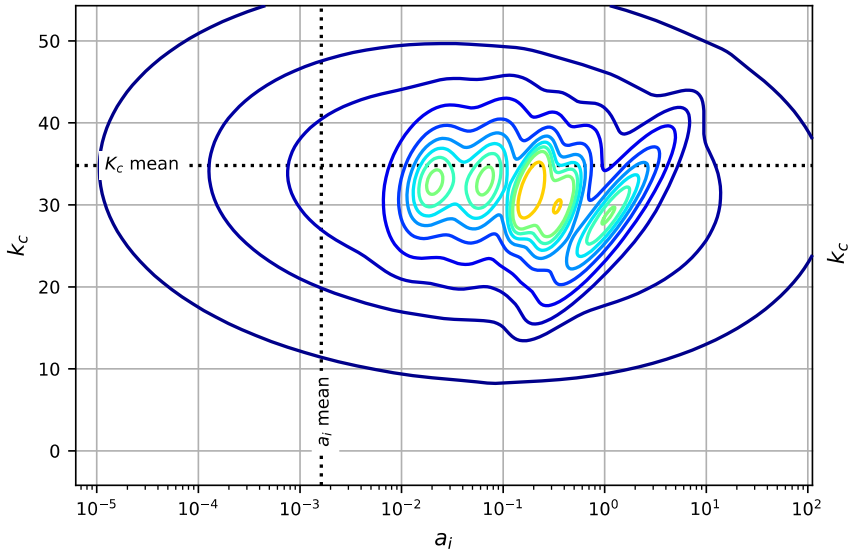
$$E[H(x,t)] = \frac{1}{N_{mix}\, N_{samp}} \sum_{j=1}^{N_{mix}} \sum_{i=1}^{N_{samp}} H(\boldsymbol{x}_{ij}, t)\, \frac{f(\boldsymbol{x}_{ij})}{(1/N_{mix}) \sum_{k=1}^{N_{mix}} q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_k)}$$

Adaptive Multiple Importance Sampling Approach

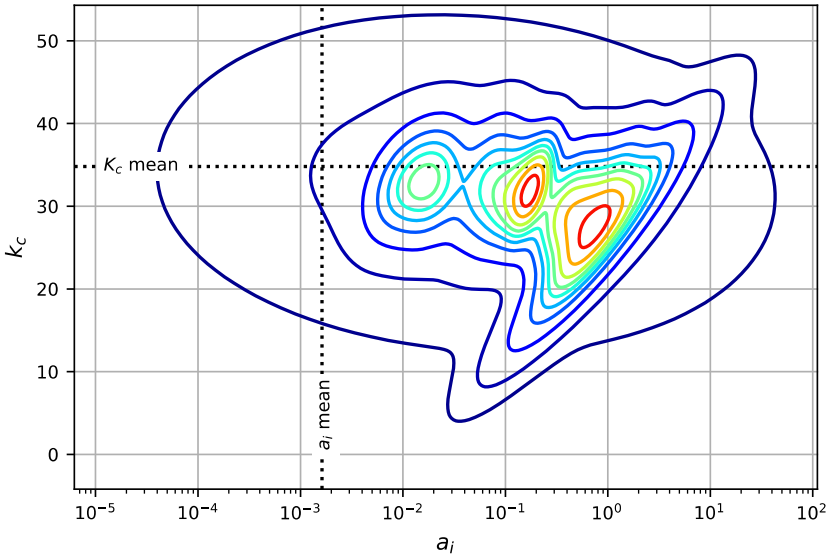
- Approximate the averaged or combined important region using a mixture density composed of multivariate normal sampling densities optimized for individual evaluation times.
- Key advantage is that samples can be used for more than one important region where regions overlap.



Mixture density (5 components)

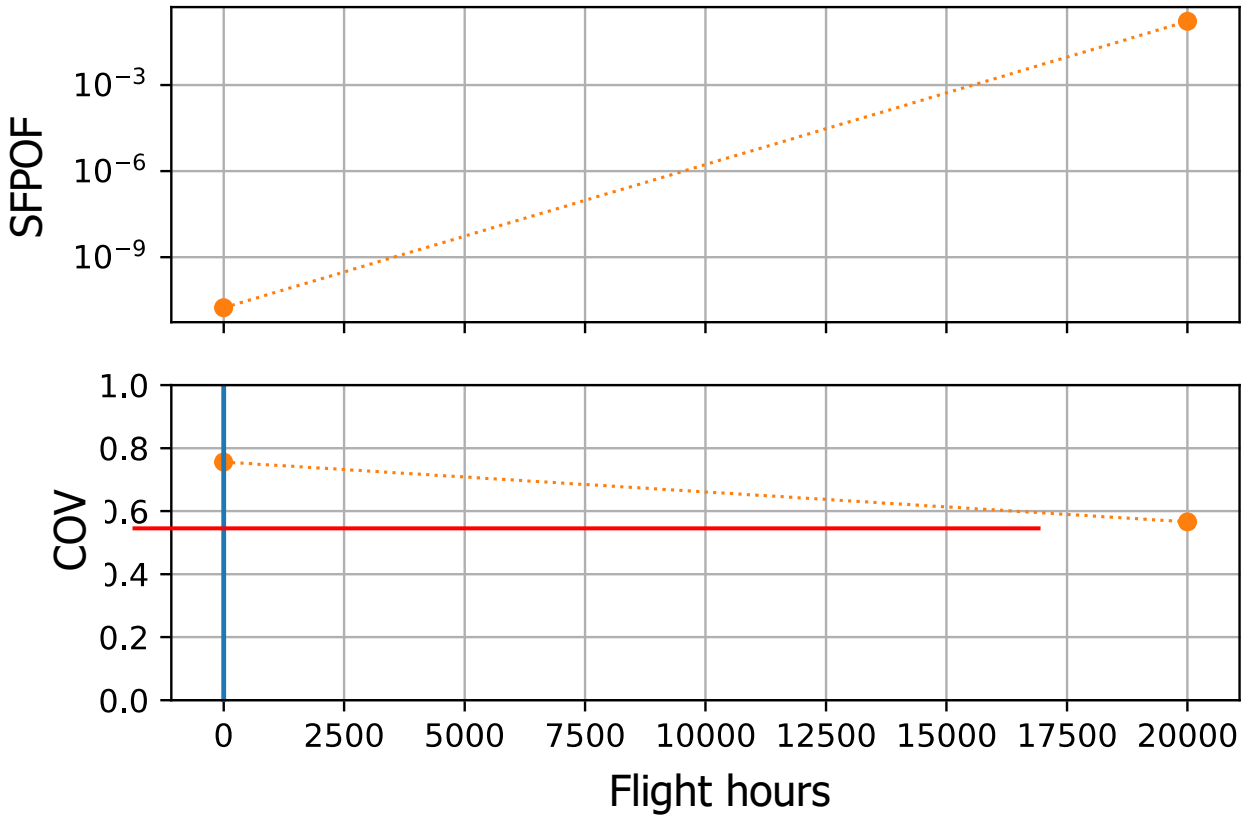


Mixture density (10 components)



Mixture density (15 components)

Academic Example



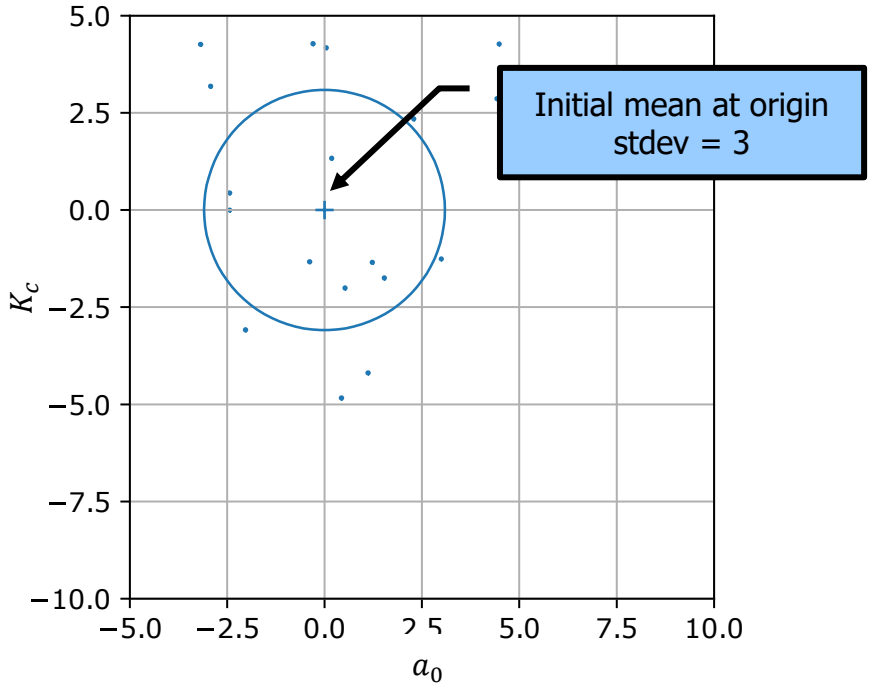
Random variables:

- Initial crack size
- Fracture toughness
- Max stress per flight

- 2 analysis times: t=0 and t=20000
- No inspections

Initialization 1/8

Initial focus @ Highest time, t=20000

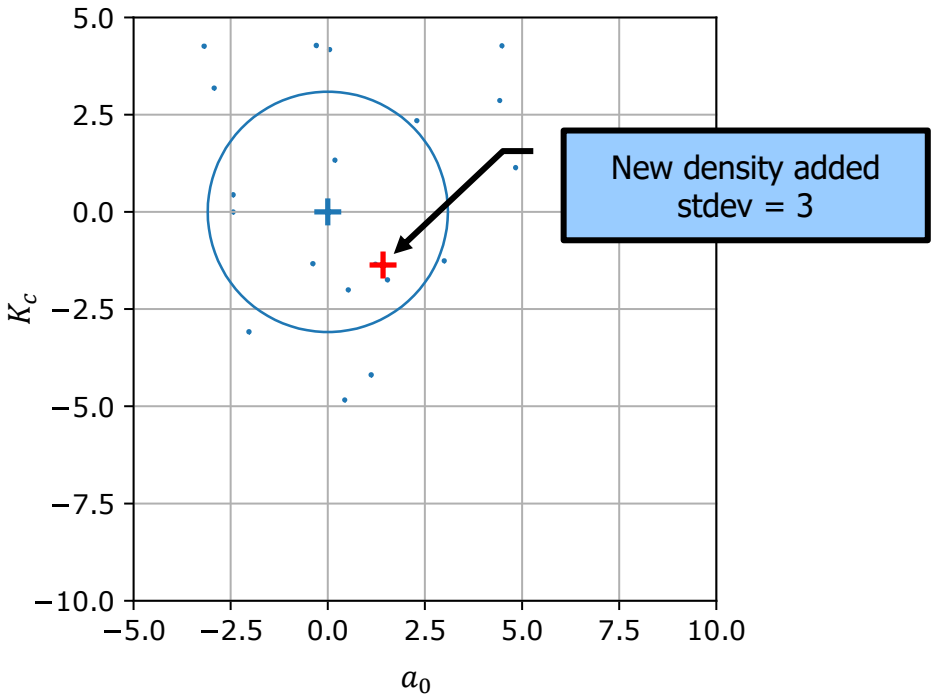


- $N_{mix} = 1$
- Initial mixture set at the mean with covariance matrix $c_\sigma^2 I$.
- 20 samples generated.

- Initialize empty mixture density, starting $\mu^\star$ at origin, $k^\star$ points to last time in list of evaluation times.
- Add $(\mu^\star, c_\sigma^2\, I)$ to $\{\theta_{mix}\}$, set $\epsilon^\star = \epsilon_{KL}/10$ to check that $\mu^\star$ is not changing before moving to next evaluation time.
- Generate samples from $N(\mu^\star, c_\sigma^2\, I)$ and evaluate crack growth and response functions.

Initialization 2/8

Focus time t=20000



$N_{mix} = 1$ prior density

$\epsilon^\star = 0.1$ (convergence threshold)

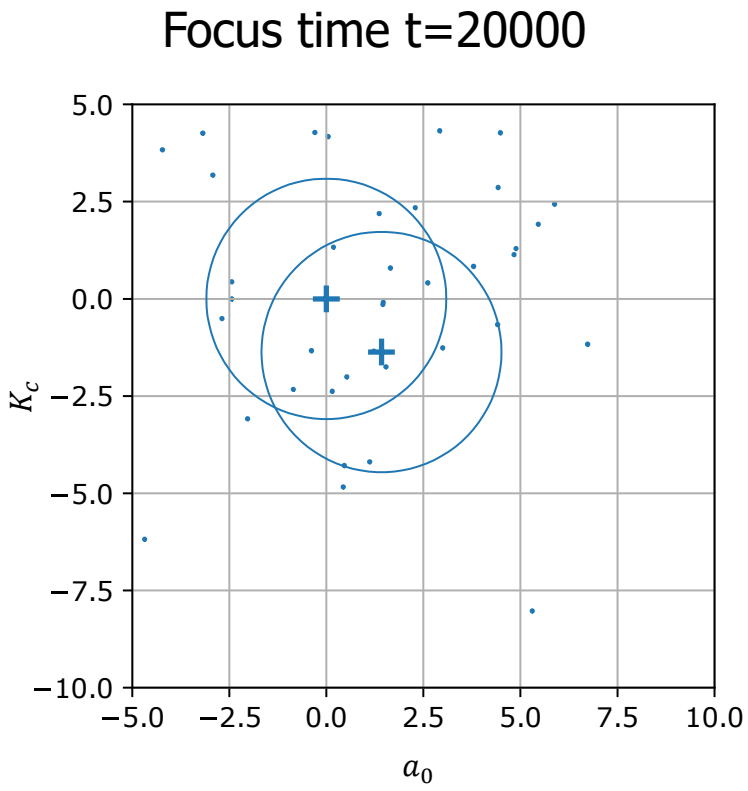
$\mu^\star = (1.4, -1.4)$

$D_{min} = 0.22$

- $D_{min} > \epsilon^*$ ∴ new density added.
- $N_{mix} = 2$

- Using the 20 samples, calculate $\mu^\star$ using the cross-entropy method.
- Evaluate D_{min} from $\mu^\star$ to all component densities in the mixture.

Initialization 3/8

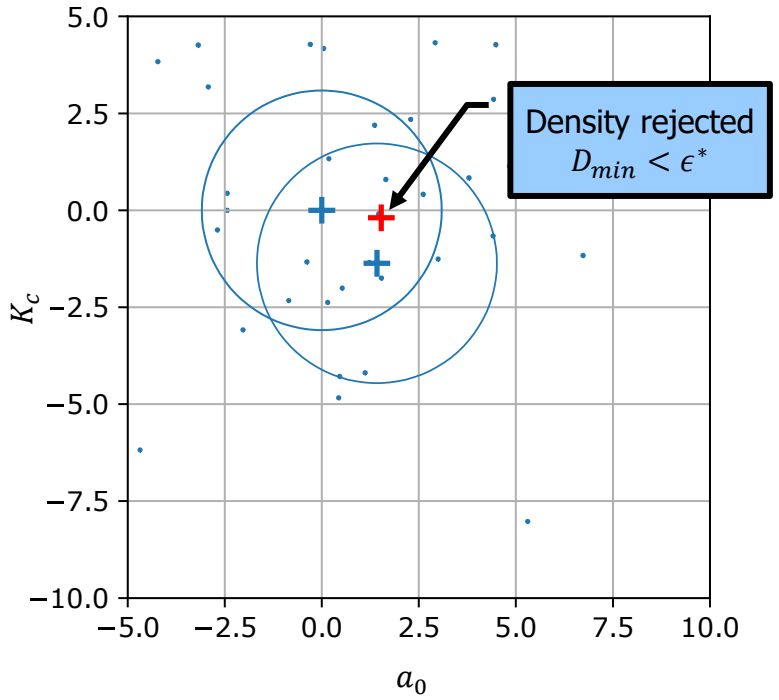


- 20 new samples drawn from density #2.
- 40 samples total.

- Add $(\mu^\star, c_\sigma^2\, I)$ to $\{\theta_{mix}\}$, set $\epsilon^\star = \epsilon_{KL}/10$ to check that $\mu^\star$ is not changing before moving to next evaluation time.
- Generate samples from $N(\mu^\star, c_\sigma^2\, I)$ and evaluate crack growth and response functions.

Initialization 4/8

Focus time t=20000



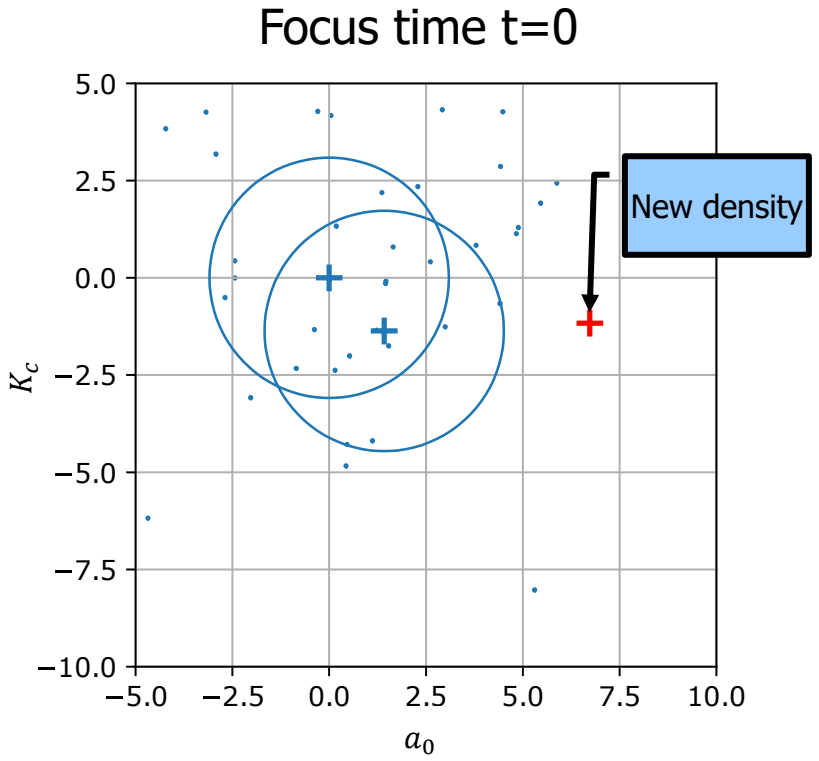
- $N_{mix} = 2$ prior densities.
- All 40 samples used to compute new location μ^*.

$$\epsilon^\star = 0.1$$
$$\mu^\star = (1.5, -0.2)$$
$$D_{min} = 0.08$$

- $D_{min} < \epsilon^*$
- Evaluation time satisfied. Moving to next time value.
- No samples generated since the new point (red) is close to another density ($D_{min} = 0.08$).

- Calculate $\mu^\star$ using standard weights.
- Evaluate D_{min} from $\mu^\star$ to all component densities in the mixture.

Initialization 5/8



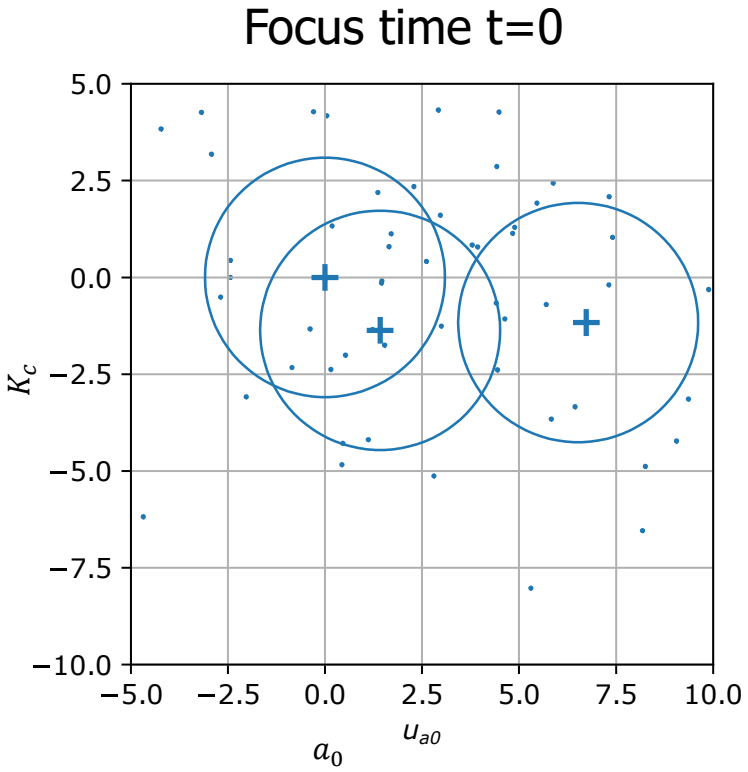
- $N_{mix} = 2$ prior densities.
- All 40 samples used to compute new location μ^*.

$$\epsilon^\star = 1.0$$
$$\mu^\star = (6.7, -1.2)$$
$$D_{min} = 1.6$$

- New density added centered at red cross.
- $N_{mix} = 3$

- The value H changes for the new time point.
- Calculate $\mu^\star$ using standard weights.
- Evaluate D_{min} from $\mu^\star$ to all component densities in the mixture.
- $D_{min} > \epsilon^\star$ so focus on the next evaluation time and locate its important region.

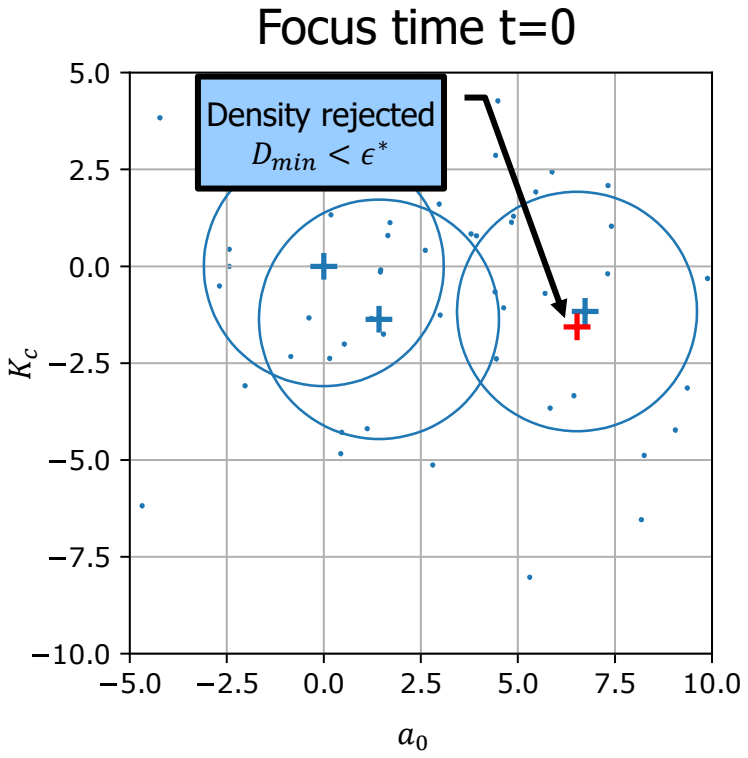
Initialization 6/8



- $N_{mix} = 3$ prior densities.
- Generate 20 new samples from density #3.
- Now 60 samples total.

- Add $(\mu^\star, c_\sigma^2\, I)$ to $\{\theta_{mix}\}$, set $\epsilon^\star = \epsilon_{KL}/10$ to check that $\mu^\star$ is not changing before moving to next evaluation time.
- Generate samples from $N(\mu^\star, c_\sigma^2\, I)$ and evaluate crack growth and response functions.

Initialization 7/8



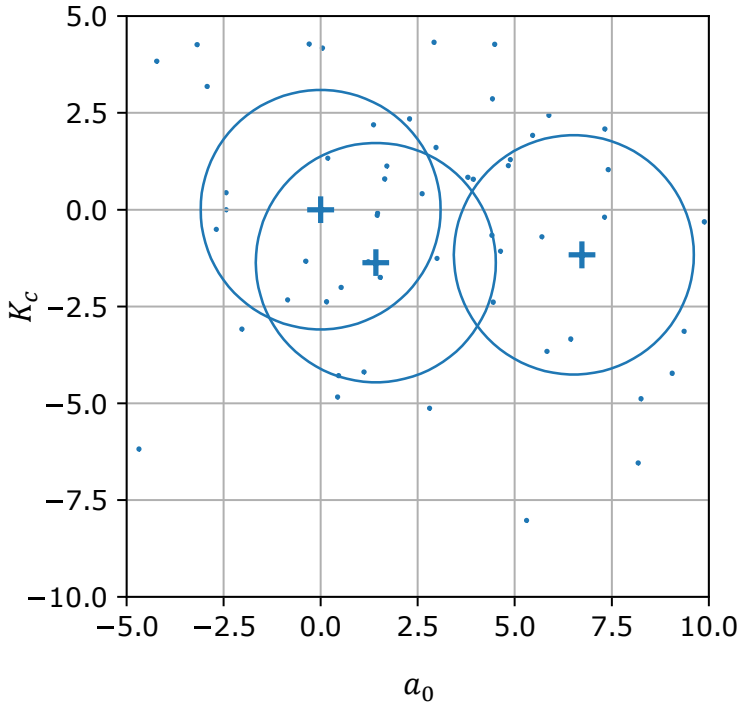
- $N_{mix} = 3$ prior densities.
- 60 samples used to compute new location μ^*.

$$\epsilon^\star = 0.1$$
$$\mu^\star = (6.5, -1.6)$$
$$D_{min} = 0.01$$

- $D_{min} < \epsilon^*$. Evaluation time satisfied. Moving to next time value.
- No samples generated since the new point (red) is close to another density ($D_{min} = 0.01$)

- Calculate $\mu^\star$ using standard weights.
- Evaluate D_{min} from $\mu^\star$ to all component densities in the mixture.
- $D_{min} < \epsilon^\star$ so focus on the next evaluation time and locate its important region

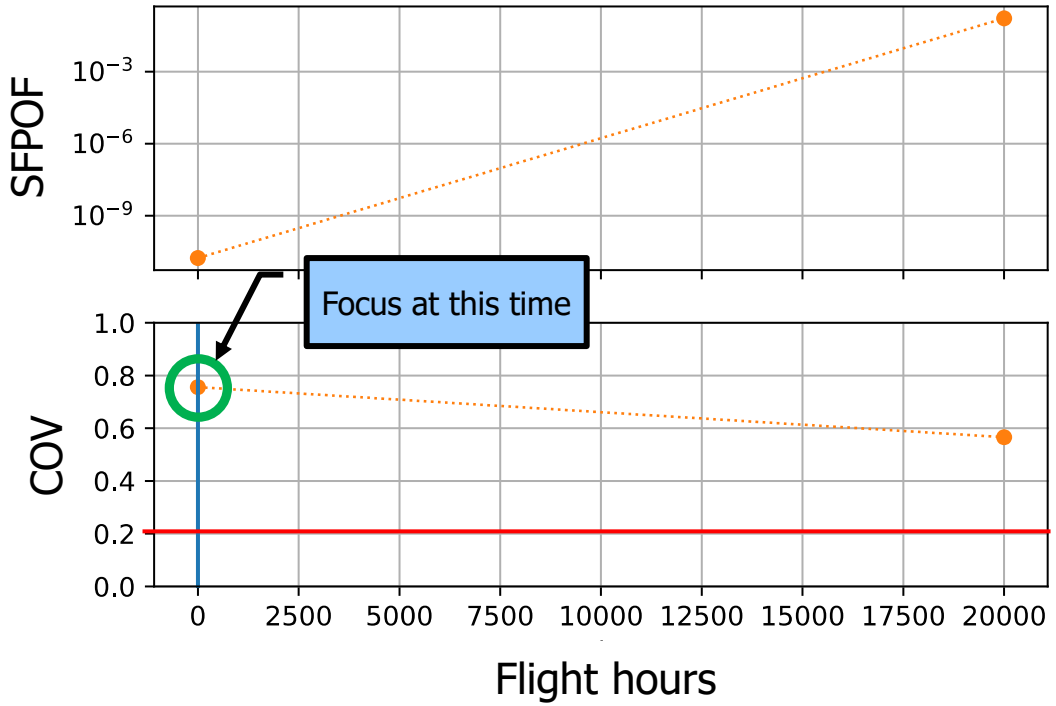
Initialization 8/8



- Initialization complete.
- 3 densities sufficient for initialization for 2 time points.
- Total of 60 samples.
- Crack growth and POF values saved for every sample.

- $k^\star = 0$, so initialization routine is finished.
- Return mixture density, realizations, crack growth evaluations and response function evaluations.

Adaptation Iteration 1/6



These equations are used to compute the POF and COV.

$$w_{bh}(x_{ij}) = \frac{q(x_{ij}, \theta_j)}{\sum_{l=1}^{N_m} (1 / N_m)\, q(x_{ij}, \theta_l)}$$

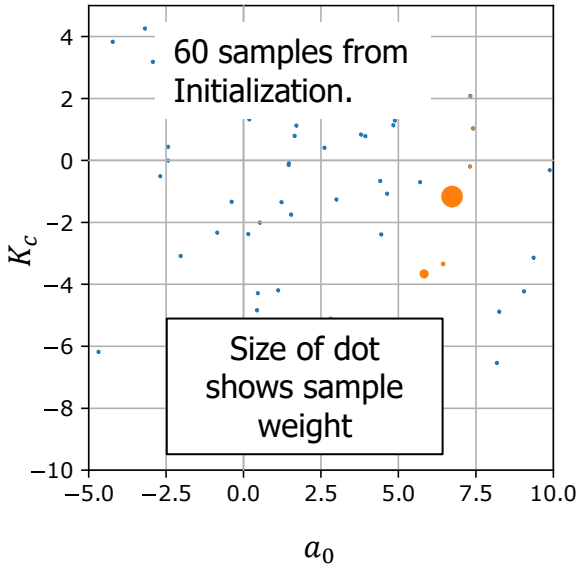
$$P_f(t_k) = \frac{1}{N} \sum_j \sum_i h_{ijk}\, w_{bh}(x_{ij})$$

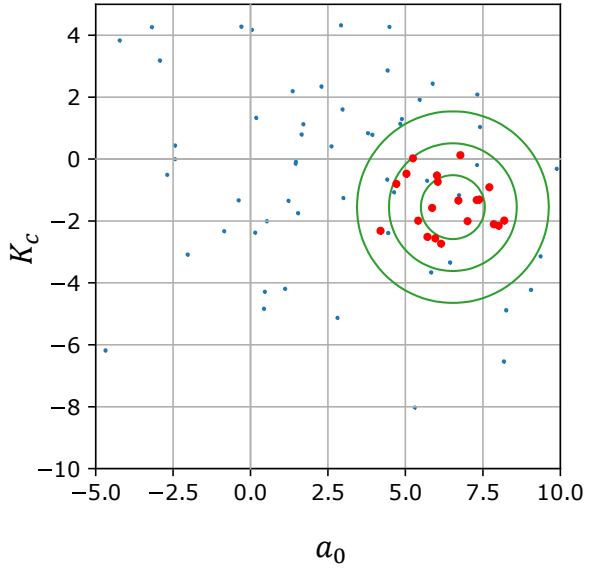
$$\mathrm{var}[P_f(t_k)] = \frac{1}{N} \sum_j \sum_i \left(h_{ijk}\, w_{bh}(x_{ij}) - P_f(t_k) \right)^2$$

$$\mathrm{cov}[P_f(t_k)] = \frac{\sqrt{\mathrm{var}[P_f(t_k)]/N}}{P_f(t_k)}$$

- Update balance heuristic importance weights
- Calculate estimates, estimator variances, and COVs
- Check exit condition (all COVs < ϵ_{cov}) or max iterations reached
- Select time with the highest COV

Adaptation Iteration 1/6





Focus time t=0

- 60 samples used to compute a new density.

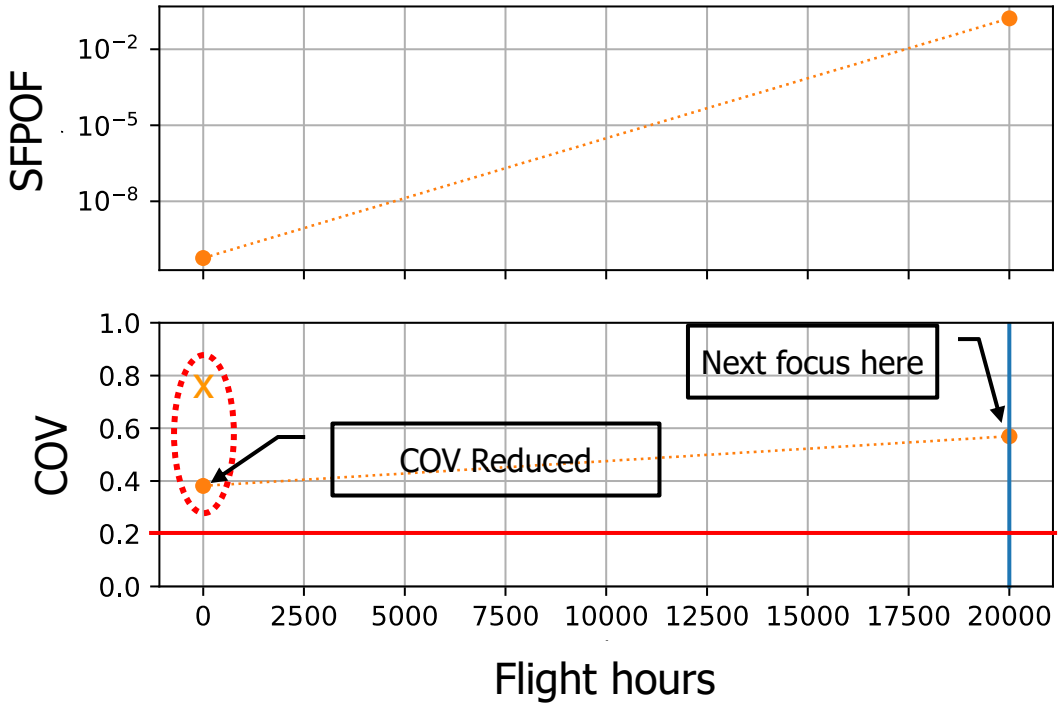
$$\mu^\star = \frac{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\boldsymbol{x}_{ij}}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$
$$\Sigma^\star = I$$

- New density added, $N_{mix} = 4$.
- 20 new samples generated (80 total).

$$w_{std}(\boldsymbol{x}_{ij}) = h_{ijk}\frac{f(\boldsymbol{x}_{ij})}{q(\boldsymbol{x}_{ij};\boldsymbol{\theta}_j)}$$
$$n_{eff} = \frac{\left(\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}{\sum_j \sum_i \left(h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}$$

- Calculate standard weights
- Effective sample size is 1, insufficient to update covariance matrix. $\Sigma^\star = I$.
- New density is $N(\mu^\star, I)$, generate samples and evaluate crack growth and response function

Adaptation Iteration 2/6



- POF and COV's computed from 80 samples.
- Weights updated using all 80 samples.
- All pofs and covs updated.
- COV reduced at t=0, moving to t=20,000.

$$w_{bh}(x_{ij}) = \frac{q(x_{ij}, \theta_j)}{\sum_{l=1}^{N_m} (1 / N_m)\, q(x_{ij}, \theta_l)}$$

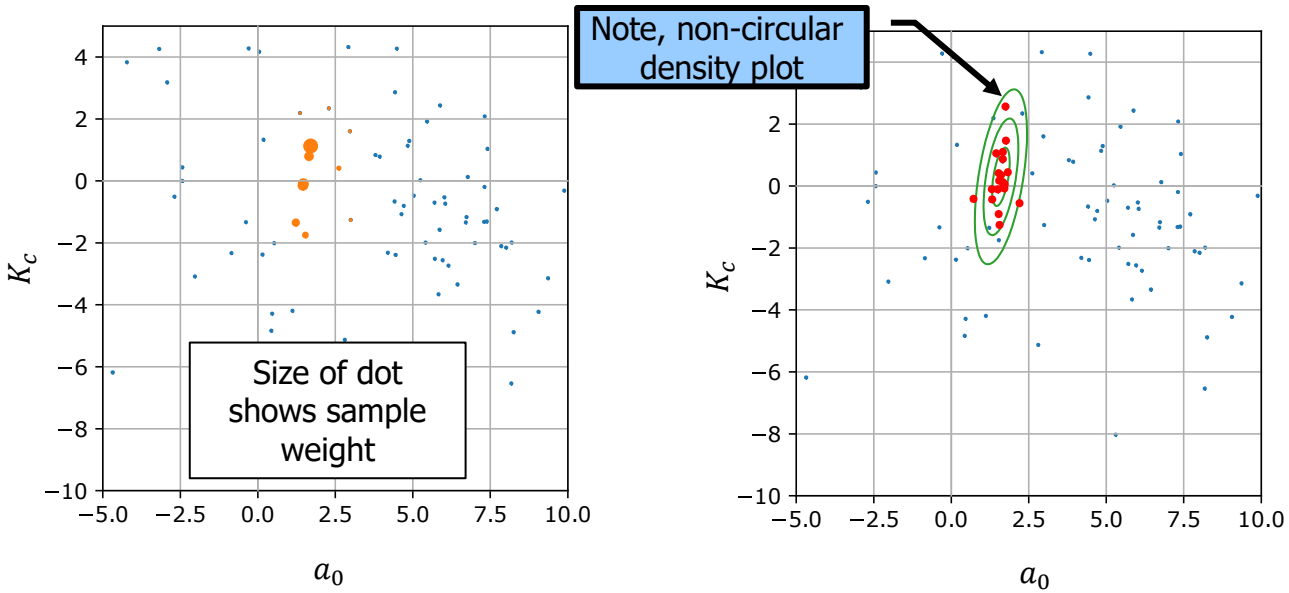
$$P_f(t_k) = \frac{1}{N} \sum_j \sum_i h_{ijk}\, w_{bh}(x_{ij})$$

$$\mathrm{var}[P_f(t_k)] = \frac{1}{N} \sum_j \sum_i \left(h_{ijk}\, w_{bh}(x_{ij}) - P_f(t_k) \right)^2$$

$$\mathrm{cov}[P_f(t_k)] = \frac{\sqrt{\mathrm{var}[P_f(t_k)]/N}}{P_f(t_k)}$$

- Update balance heuristic importance weights
- Calculate estimates, estimator variances, and COVs
- Check exit condition (all COVs < ϵ_{cov}) or max iterations reached
- Select time with the highest COV

Adaptation Iteration 2/6



Focus time t=20000

- New density added, $N_{mix} = 5$.
- 20 new samples generated.
- 100 samples total.

$$\mu^{\star} = \frac{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\boldsymbol{x}_{ij}}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$

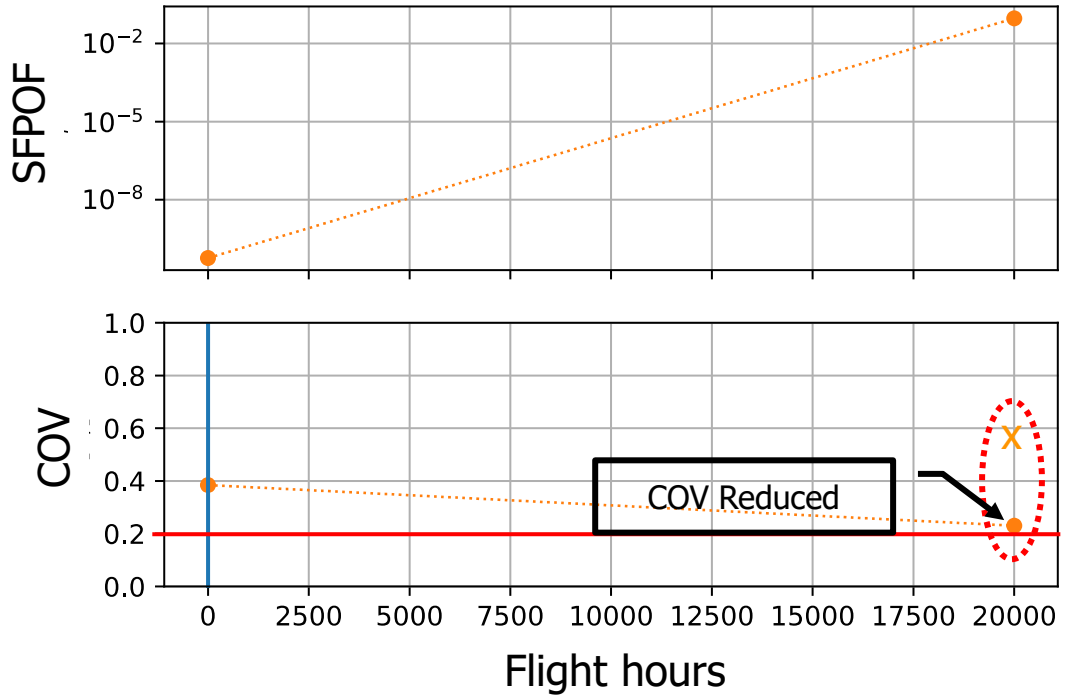
$$\Sigma^{\star} = \frac{\sum_j \sum_i (\boldsymbol{x}_{ij} - \boldsymbol{\mu}_{\star})^{\mathsf{T}} h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})(\boldsymbol{x}_{ij} - \boldsymbol{\mu}_{\star})}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$

$$w_{std}(\boldsymbol{x}_{ij}) = h_{ijk} \frac{f(\boldsymbol{x}_{ij})}{q(\boldsymbol{x}_{ij}; \boldsymbol{\theta}_j)}$$

$$n_{eff} = \frac{\left(\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}{\sum_j \sum_i \left(h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}$$

- Calculate standard weights.
- Effective sample size is 4, sufficient to update covariance matrix.
- New density is $N(\mu^{\star}, \Sigma^{\star})$, generate samples and evaluate crack growth and response function.

Adaptation Iteration 3/6



- 100 samples used.
- All pofs and covs updated.

$$w_{bh}(x_{ij}) = \frac{q(x_{ij}, \theta_j)}{\sum_{l=1}^{N_m} (1 / N_m)\, q(x_{ij}, \theta_l)}$$

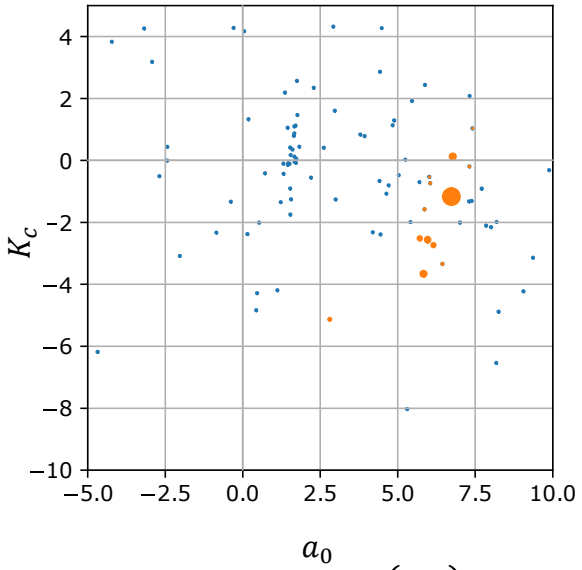
$$P_f(t_k) = \frac{1}{N} \sum_j \sum_i h_{ijk}\, w_{bh}(x_{ij})$$

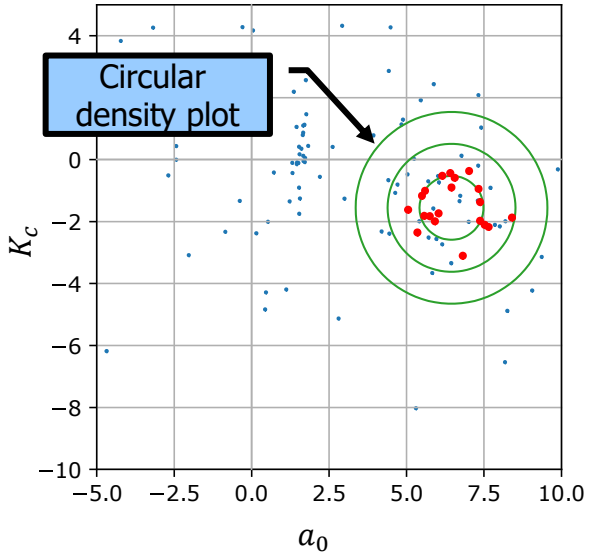
$$\mathrm{var}[P_f(t_k)] = \frac{1}{N} \sum_j \sum_i \left(h_{ijk}\, w_{bh}(x_{ij}) - P_f(t_k) \right)^2$$

$$\mathrm{cov}[P_f(t_k)] = \frac{\sqrt{\mathrm{var}[P_f(t_k)]/N}}{P_f(t_k)}$$

- Update balance heuristic importance weights.
- Calculate estimates, estimator variances, and COVs.
- Check exit condition (all COVs < ϵ_{cov}) or max iterations reached.
- Select time with the highest COV.

Adaptation Iteration 3/6





Focus time t=0

- New density added, $N_{mix} = 6$.
- 20 new samples generated.
- 120 samples total.

$$\mu^{\star} = \frac{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\boldsymbol{x}_{ij}}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$

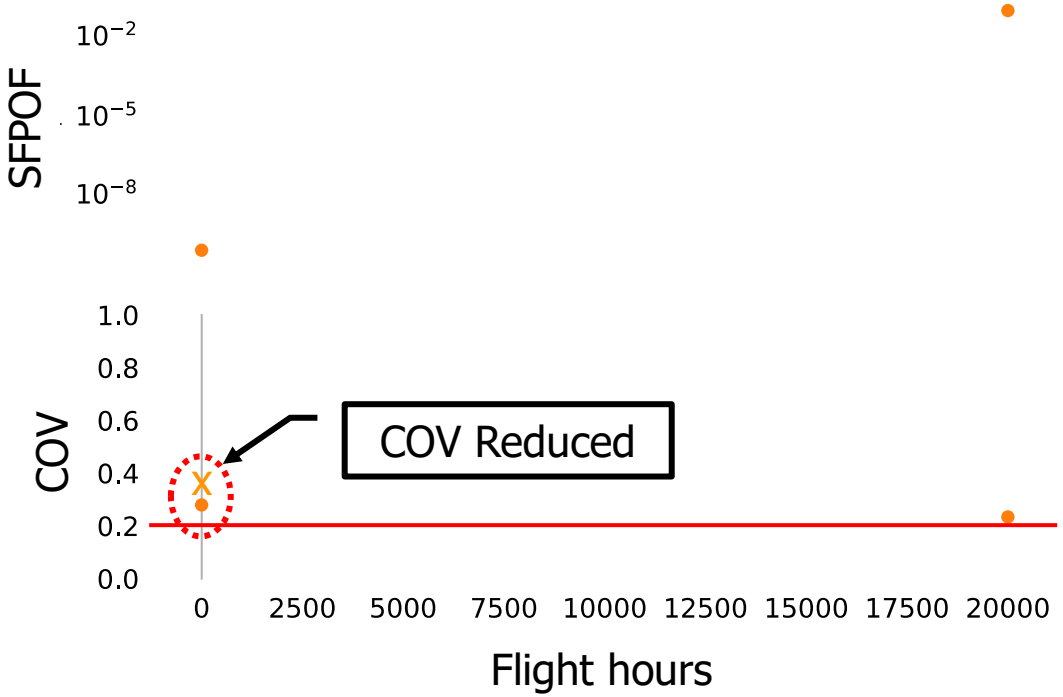
$$\Sigma^{\star} = I$$

$$w_{std}(\boldsymbol{x}_{ij}) = h_{ijk}\frac{f(\boldsymbol{x}_{ij})}{q(\boldsymbol{x}_{ij};\boldsymbol{\theta}_j)}$$

$$n_{eff} = \frac{\left(\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}{\sum_j \sum_i \left(h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}$$

- Calculate standard weights.
- Effective sample size is 2, insufficient to update covariance matrix.
- New density is $N(\mu^{\star}, I)$, generate samples and evaluate crack growth and response function.

Adaptation Iteration 4/6



- 120 samples used.
- All pofs and covs updated.

$$w_{bh}(x_{ij}) = \frac{q(x_{ij}, \theta_j)}{\sum_{l=1}^{N_m} (1 / N_m)\, q(x_{ij}, \theta_l)}$$

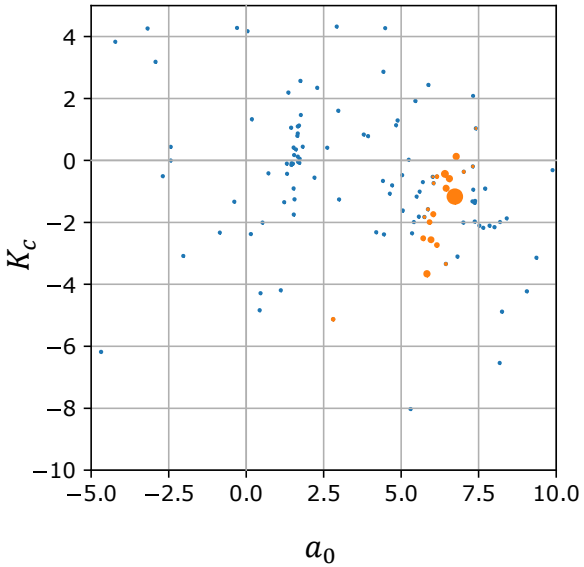
$$P_f(t_k) = \frac{1}{N} \sum_j \sum_i h_{ijk}\, w_{bh}(x_{ij})$$

$$\mathrm{var}[P_f(t_k)] = \frac{1}{N} \sum_j \sum_i \left(h_{ijk}\, w_{bh}(x_{ij}) - P_f(t_k) \right)^2$$

$$\mathrm{cov}[P_f(t_k)] = \frac{\sqrt{\mathrm{var}[P_f(t_k)]/N}}{P_f(t_k)}$$

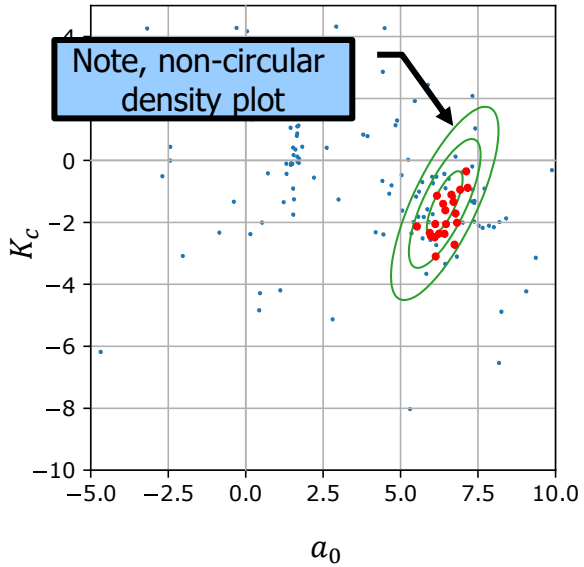
- Update balance heuristic importance weights
- Calculate estimates, estimator variances, and COVs
- Check exit condition (all COVs < ϵ_{cov}) or max iterations reached
- Select time with the highest COV

Adaptation Iteration 4/6



$$w_{std}(\boldsymbol{x}_{ij}) = h_{ijk}\frac{f(\boldsymbol{x}_{ij})}{q(\boldsymbol{x}_{ij};\boldsymbol{\theta}_j)}$$

$$n_{eff} = \frac{\left(\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}{\sum_j \sum_i \left(h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}$$



Focus time t=0

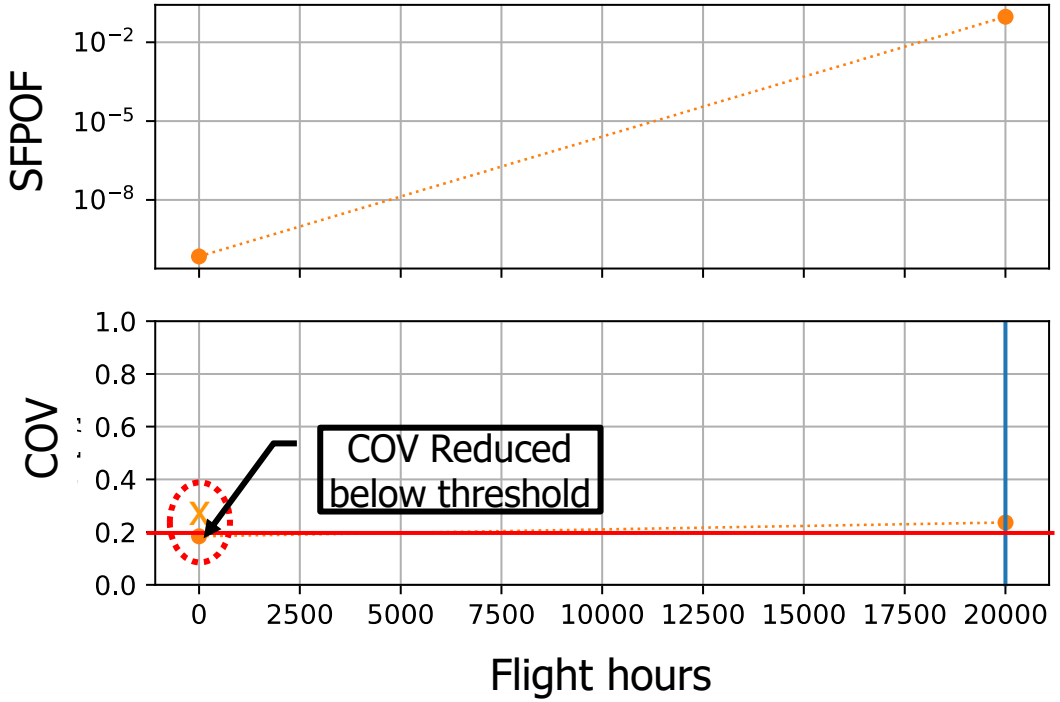
- New density added, $N_{mix} = 7$.
- 20 new samples generated.
- 140 samples total.

$$\mu^\star = \frac{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\boldsymbol{x}_{ij}}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$

$$\Sigma^\star = \frac{\sum_j \sum_i (\boldsymbol{x}_{ij} - \boldsymbol{\mu}_\star)^\top h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})(\boldsymbol{x}_{ij} - \boldsymbol{\mu}_\star)}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$

- Calculate standard weights
- Effective sample size is 4, sufficient to update covariance matrix
- New density is $N(\mu^\star, \Sigma^\star)$, generate samples and evaluate crack growth and response function

Adaptation Iteration 5/6



- 140 samples used.
- All pofs and covs updated.
- COV(t=0) now below threshold.

$$w_{bh}(x_{ij}) = \frac{q(x_{ij}, \theta_j)}{\sum_{l=1}^{N_m} (1 / N_m)\, q(x_{ij}, \theta_l)}$$

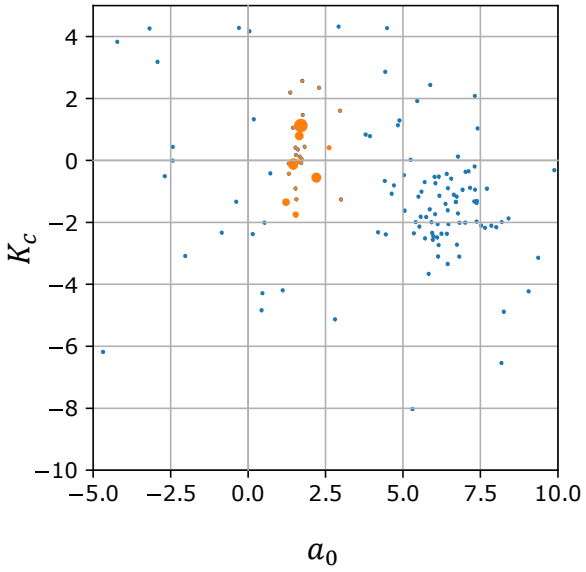
$$P_f(t_k) = \frac{1}{N} \sum_j \sum_i h_{ijk}\, w_{bh}(x_{ij})$$

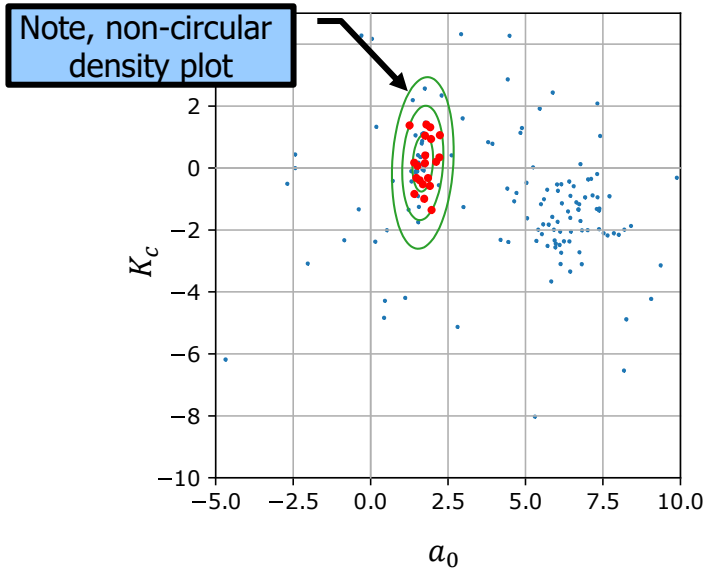
$$\mathrm{var}[P_f(t_k)] = \frac{1}{N} \sum_j \sum_i \left(h_{ijk}\, w_{bh}(x_{ij}) - P_f(t_k) \right)^2$$

$$\mathrm{cov}[P_f(t_k)] = \frac{\sqrt{\mathrm{var}[P_f(t_k)]/N}}{P_f(t_k)}$$

- Update balance heuristic importance weights.
- Calculate estimates, estimator variances, and COVs.
- Check exit condition (all COVs < ϵ_{cov}) or max iterations reached.
- Select time with the highest COV.

Adaptation Iteration 5/6





Focus time t=20000

- $N_{mix} = 8.$
- 20 new samples added.
- 160 samples total.
- All pofs and covs updated.

$$\mu^{\star} = \frac{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\boldsymbol{x}_{ij}}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$

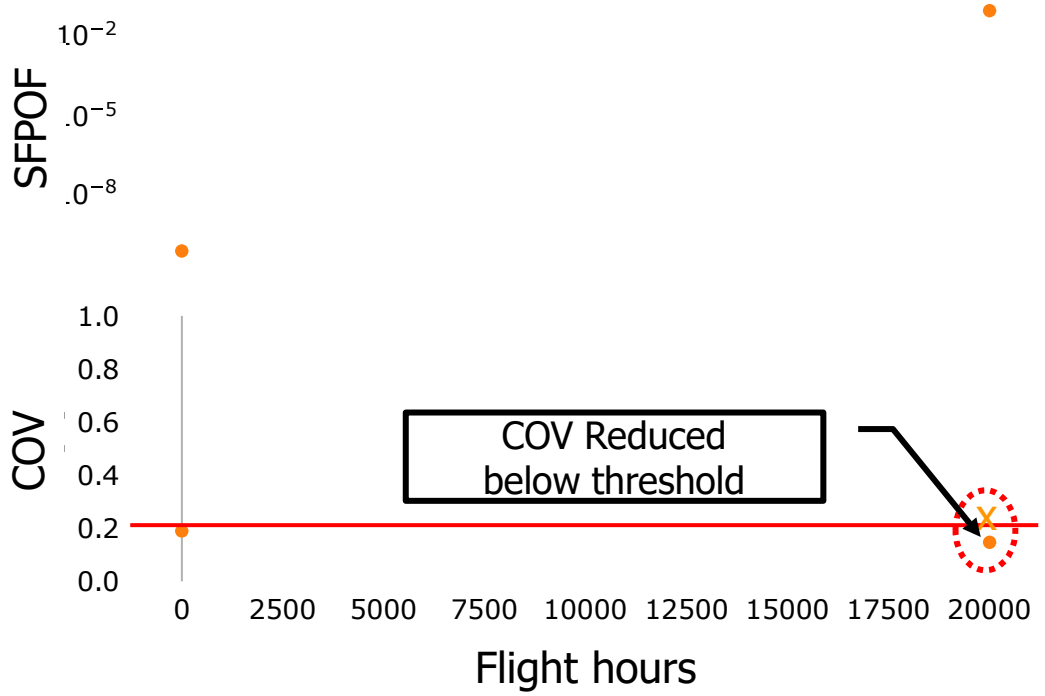
$$\Sigma^{\star} = \frac{\sum_j \sum_i (\boldsymbol{x}_{ij} - \boldsymbol{\mu}_{\star})^{\mathsf{T}} h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})(\boldsymbol{x}_{ij} - \boldsymbol{\mu}_{\star})}{\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})}$$

$$w_{std}(\boldsymbol{x}_{ij}) = h_{ijk} \frac{f(\boldsymbol{x}_{ij})}{q(\boldsymbol{x}_{ij}; \boldsymbol{\theta}_j)}$$

$$n_{eff} = \frac{\left(\sum_j \sum_i h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}{\sum_j \sum_i \left(h_{ijk}\, w_{std}(\boldsymbol{x}_{ij})\right)^2}$$

- Effective sample size is 6, sufficient to update covariance matrix.
- New density is $N(\mu^{\star}, \Sigma^{\star})$, generate samples and evaluate crack growth and response function.

Adaptation Iteration 6/6



- 160 samples used.
- All COVs below threshold.
- CG results saved for all 160 samples. samples.

$$w_{bh}(x_{ij}) = \frac{q(x_{ij}, \theta_j)}{\sum_{l=1}^{N_m} (1 / N_m)\, q(x_{ij}, \theta_l)}$$

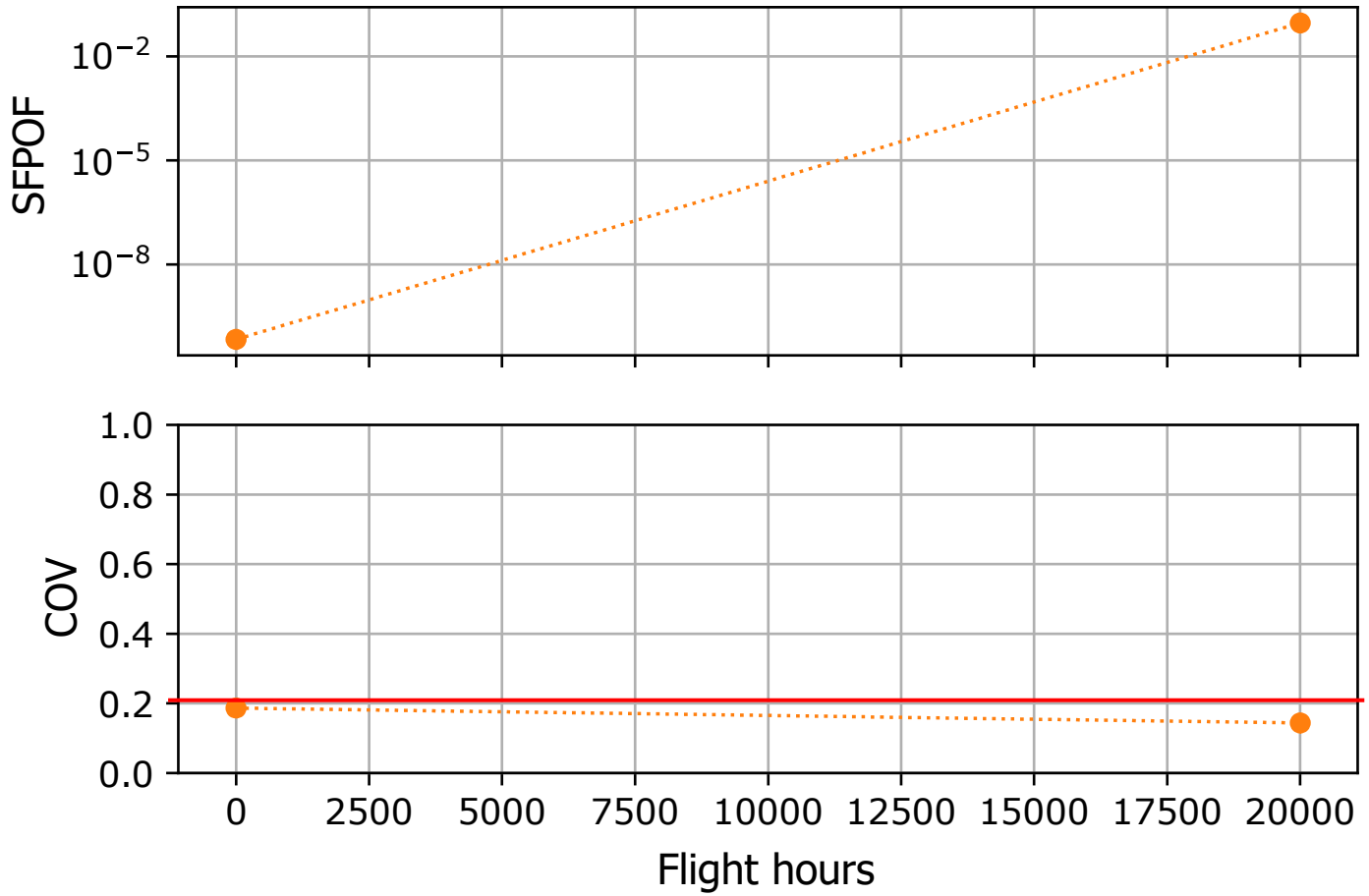
$$P_f(t_k) = \frac{1}{N} \sum_j \sum_i h_{ijk}\, w_{bh}(x_{ij})$$

$$\mathrm{var}[P_f(t_k)] = \frac{1}{N} \sum_j \sum_i \left(h_{ijk}\, w_{bh}(x_{ij}) - P_f(t_k) \right)^2$$

$$\mathrm{cov}[P_f(t_k)] = \frac{\sqrt{\mathrm{var}[P_f(t_k)]/N}}{P_f(t_k)}$$

- Update balance heuristic importance weights.
- Calculate estimates, estimator variances, and COVs.
- Check exit condition (all COVs < ϵ_{cov}) or max iterations reached
- All COVs < ϵ_{cov}.

Academic Example Summary

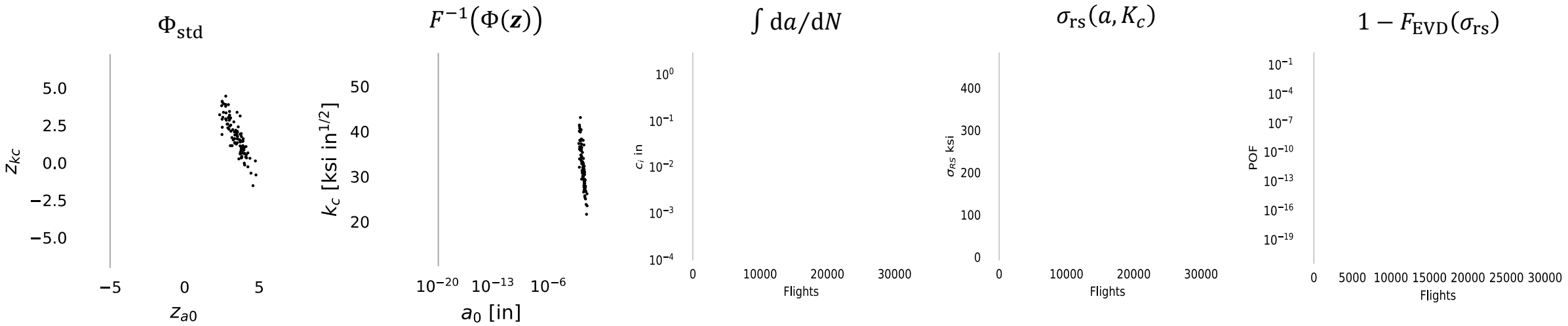


- SFPOF computed using 160 samples.
- (60 initialization, 100 adaptation).
- All COVs below user-defined threshold of 20%.

AMIS Method

	Typical Approach	AMIS Method
Probabilistic method	Numerical integration	Adaptive multiple importance sampling
Random variables	Only 3 random variables (toughness, initial crack size, max load)	Up to ~20 (toughness, initial crack size, max load, dadn variability, geometric factors)
Accuracy	Limit user control/information	Adaptivity ensures convergence to user-defined threshold
Reanalyses	Requires complete reanalysis	Mixture densities ensure efficient reanalysis
Efficiency	Efficient even for small POFs	Efficient even for small POFs

Transformation to Normal Space



- Transformation to standard normal space simplifies working with multiple distributions
  - Samples are generated in standard multivariate normal space
  - Nominal density is independent standard normal
  - Estimated optimal sampling density, likelihood ratio, and parameter updates use normal space
  - Samples are transformed to original space using inverse CDF transform (Nataf for correlated variables) for crack growth evaluation

Cross Entropy Parameter Calculation

Parameters $\hat{\mu}^{\star}(t)$, and $\hat{\Sigma}^{\star}(t)$ maximize the cross entropy

$$\boldsymbol{\theta}_{\text{opt}} = \underset{\boldsymbol{\theta}}{\text{argmax}} \left(\sum_i \sum_j H(\boldsymbol{x}_{ij}, t) w(\boldsymbol{x}_{ij}) \ln \left(q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}) \right) \right)$$

Closed form solution

$$\hat{\mu}^{\star}(t) = \frac{\sum_{i}^{N_{samp}} \sum_{j}^{N_{mix}} H(\boldsymbol{x}_{ij}, t)\, w_{std}(\boldsymbol{x}_{ij})\, \boldsymbol{x}_{ij}}{\sum_i \sum_j H(\boldsymbol{x}_{ij}, t)\, w_{std}(\boldsymbol{x}_{ij})}$$

Mean of mixture densities

$$\hat{\Sigma}^{\star}(t) = \frac{\sum_i \sum_j (\boldsymbol{x}_{ij} - \hat{\mu}^{\star}) H(\boldsymbol{x}_{ij}, t)\, w_{std}(\boldsymbol{x}_{ij})\, (\boldsymbol{x}_{ij} - \hat{\mu}^{\star})^T}{\sum_i \sum_j H(\boldsymbol{x}_{ij}, t)\, w_{std}(\boldsymbol{x}_{ij})}$$

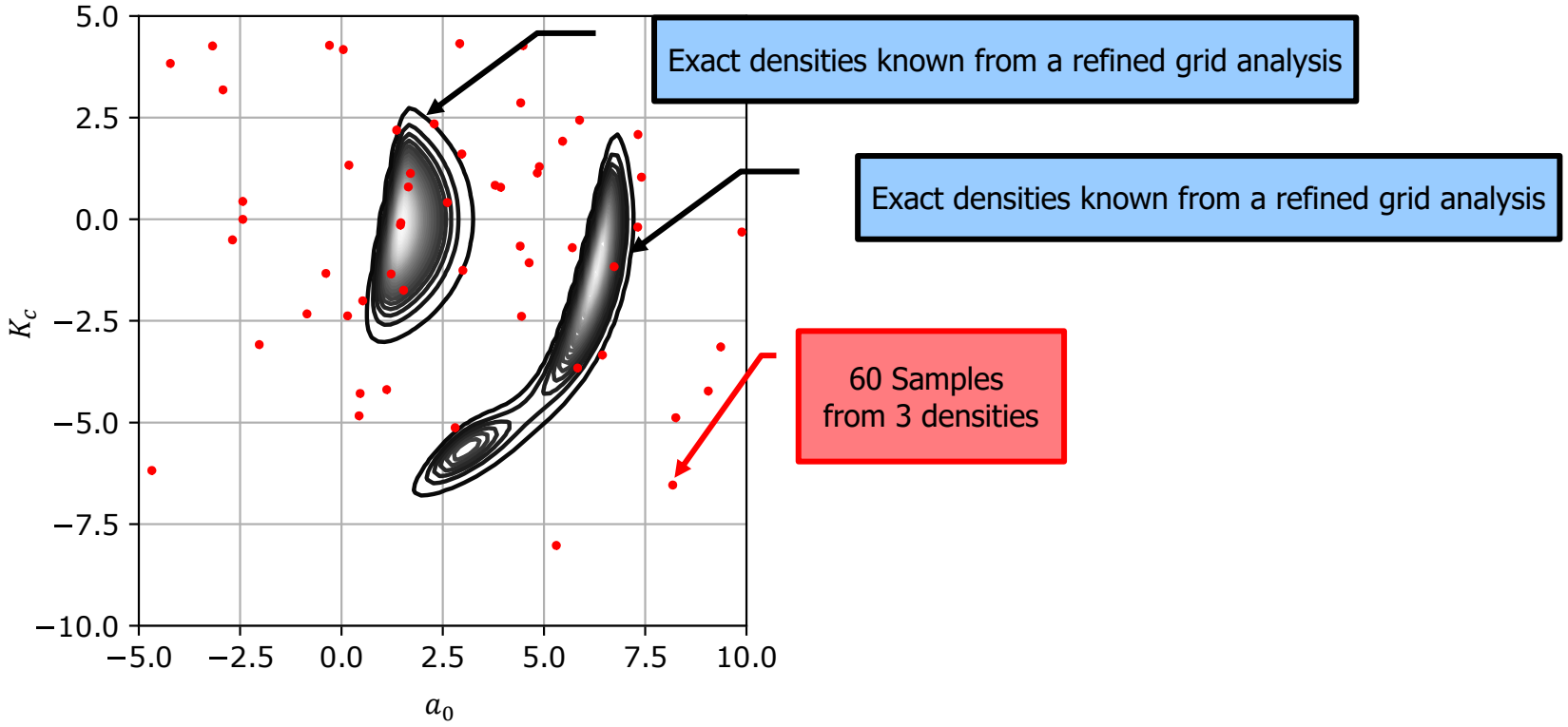
Covariance of mixture densities

- H depends on the Lincoln or Freudenthal formulation, e.g.,$1 - F_{EVD}(\sigma_{RS}(\boldsymbol{x}, t_n))$

Initialization Overview

- Initialization algorithm searches for important regions for each time and adds component densities to the mixture
- $N_{maxiter}$ and N_s set the maximum number of iterations and samples generated per iteration
- $\{t_{pf}\}$, $CG(\cdot)$ and $H(\cdot)$ define the evaluation times, crack growth function and response function
- c_σ controls the variance of the initialization sampling densities
- ϵ_{KL} controls distance between initialization densities
  - Multiple steps will be taken for each evaluation time if the sampling density location changes with additional samples

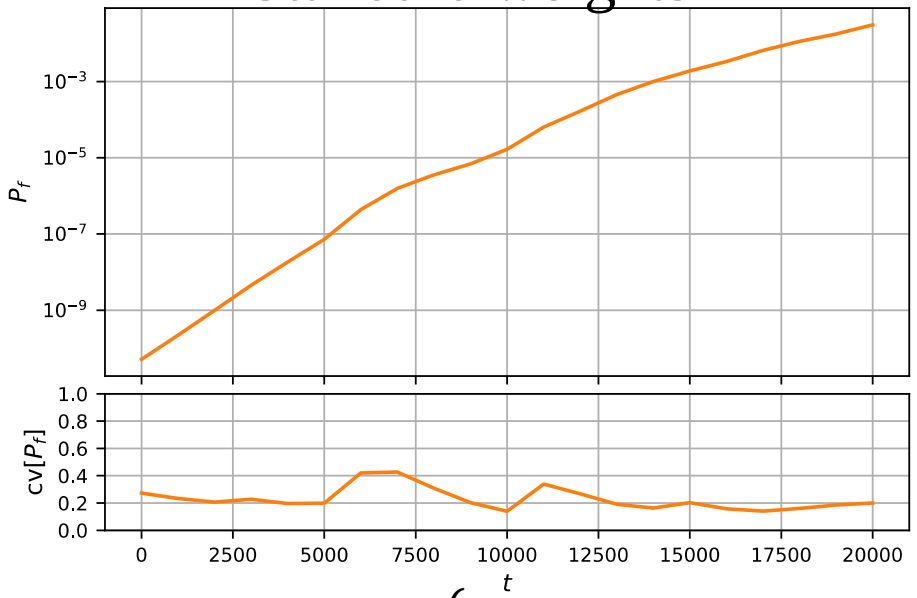
Initialization Summary



- Samples have been generated in and around the important regions
- Exploration by the initialization is very important – the adaptation phase does not search for new important regions that are not found during the initialization phase

Mixture density weights – multiple definitions

Standard weights

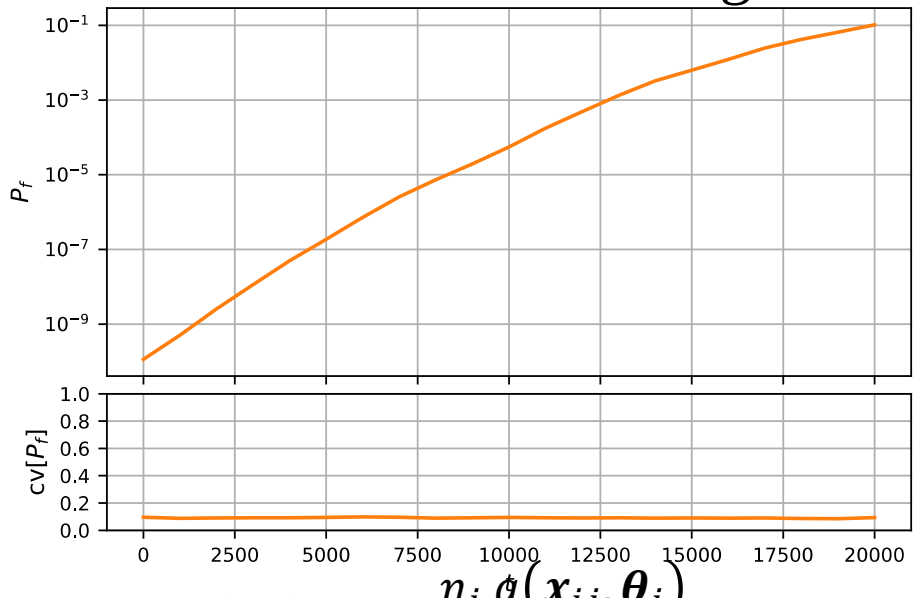


$$\omega_m^{std}(\boldsymbol{x}_{ij}) = \begin{cases} 1 & j = m \\ 0 & \text{otherwise} \end{cases}$$

$$w_{std}(\boldsymbol{x}_{ij}) = \frac{f(\boldsymbol{x}_{ij})}{q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_j)}$$

- Large weights will not decrease as component densities are added

Balance heuristic weights



$$\omega_m^{BH}(\boldsymbol{x}_{ij}) = \frac{n_j\, q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_j)}{\sum_{l=1}^{N_{mix}} n_l\, q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_l)}$$

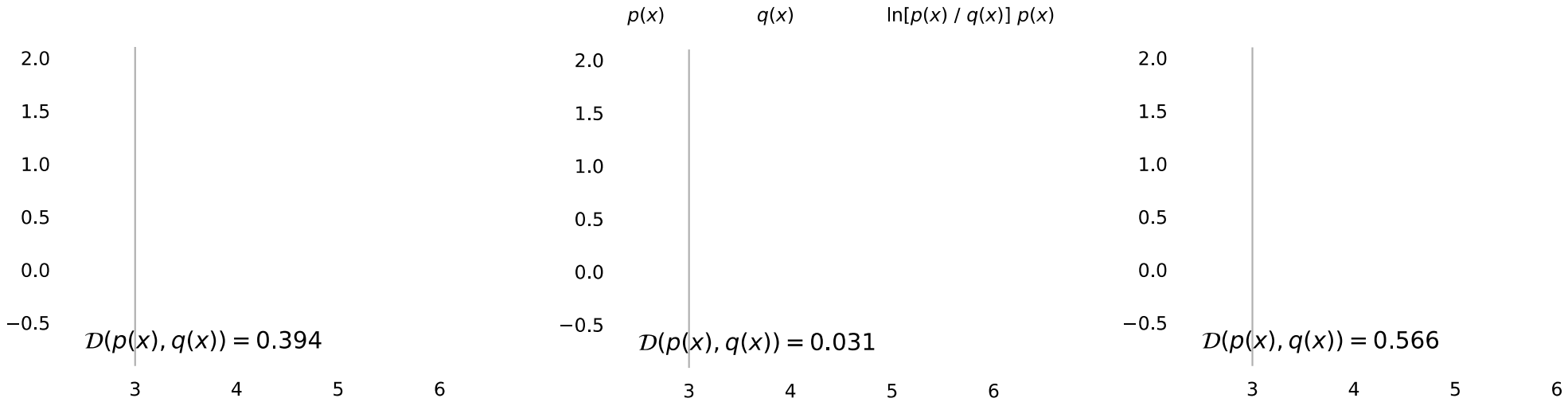
$$w_{BH}(\boldsymbol{x}_{ij}) = \frac{f(\boldsymbol{x}_{ij})}{\sum_{l=1}^{N_{mix}} (1 / n_{mix})\, q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_l)}$$

- Give the best weight out of all component densities for each sample
- If the mixture includes the nominal density, this weighting has an upper bound

Cross Entropy – Kullback-Leibler Divergence

Kullback-Leibler Divergence

$$\mathcal{D}\big(p(\boldsymbol{x}), q(\boldsymbol{x})\big) = \int \ln\left(\frac{p(\boldsymbol{x})}{q(\boldsymbol{x})}\right) p(\boldsymbol{x})\, dx = \int \ln\big(p(\boldsymbol{x})\big)\, p(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} - \int \ln\big(q(\boldsymbol{x})\big)\, p(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$



- $\mathcal{D}$ is a metric – a measure of difference between two PDFs
  - $\mathcal{D} \geq 0$, $\mathcal{D} = 0$ when the PDFs are identical

Mixture density

$$\widehat{\mathrm{POF}}(t) = \sum_{j=1}^{N_{mix}} \frac{1}{N_{samp}} \sum_{i=1}^{N_{samp}} H(\boldsymbol{x}_{ij}, t)\, \omega_j(\boldsymbol{x}_{ij})\, \frac{f(\boldsymbol{x}_{ij})}{g(\boldsymbol{x}_{ij}, \theta_j)}$$

Mixture importance weight for sample $\boldsymbol{x}_{ij}$

Nominal density

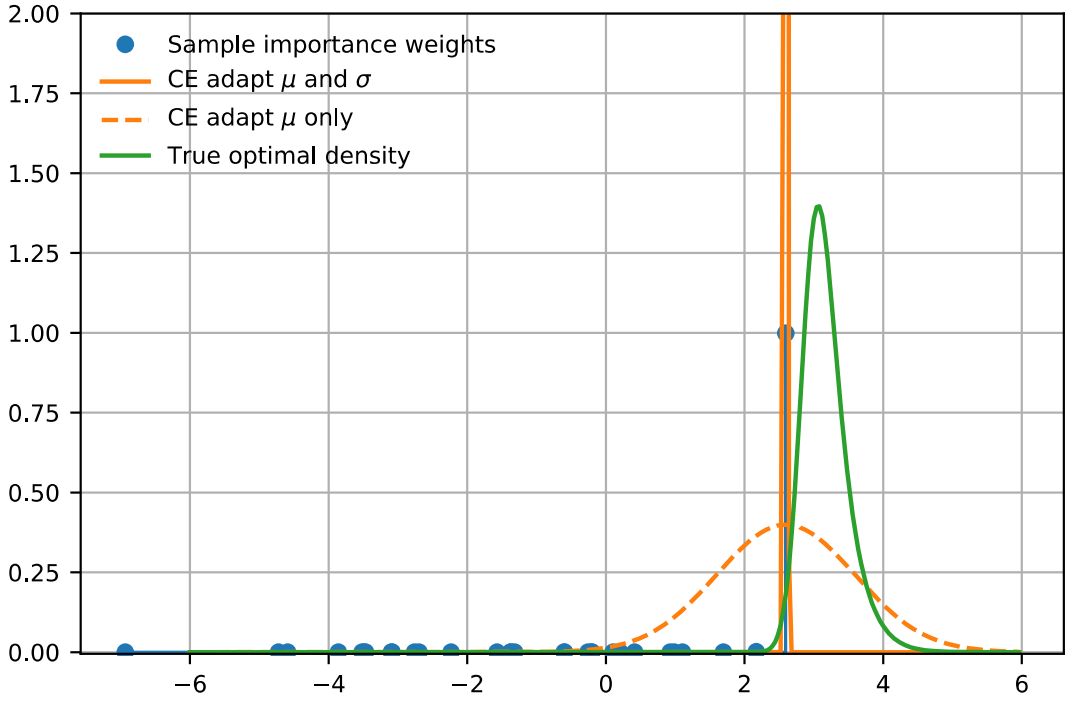
$$\omega_m^{bk}(\boldsymbol{x}_{ij}) = \frac{n_m\, q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_m)}{\sum_{k=1}^{N_{mix}} n_k\, q(\boldsymbol{x}_{ij}, \boldsymbol{\theta}_k)} \quad \forall\ i, j$$

Mixture weights for sample $\boldsymbol{x}_{ij}$, $\sum_{m=1}^{N_{mix}} \omega_m(\boldsymbol{x}_{ij}) = 1$

j-th component density with parameters $\boldsymbol{\theta}_j$

- Mixture density is composed of component sampling densities optimal for individual times
- Mixture importance weight is a weighted sum of importance weights from all component densities
- The estimator is a double summation over component densities and samples from each density

Effective Sample Size

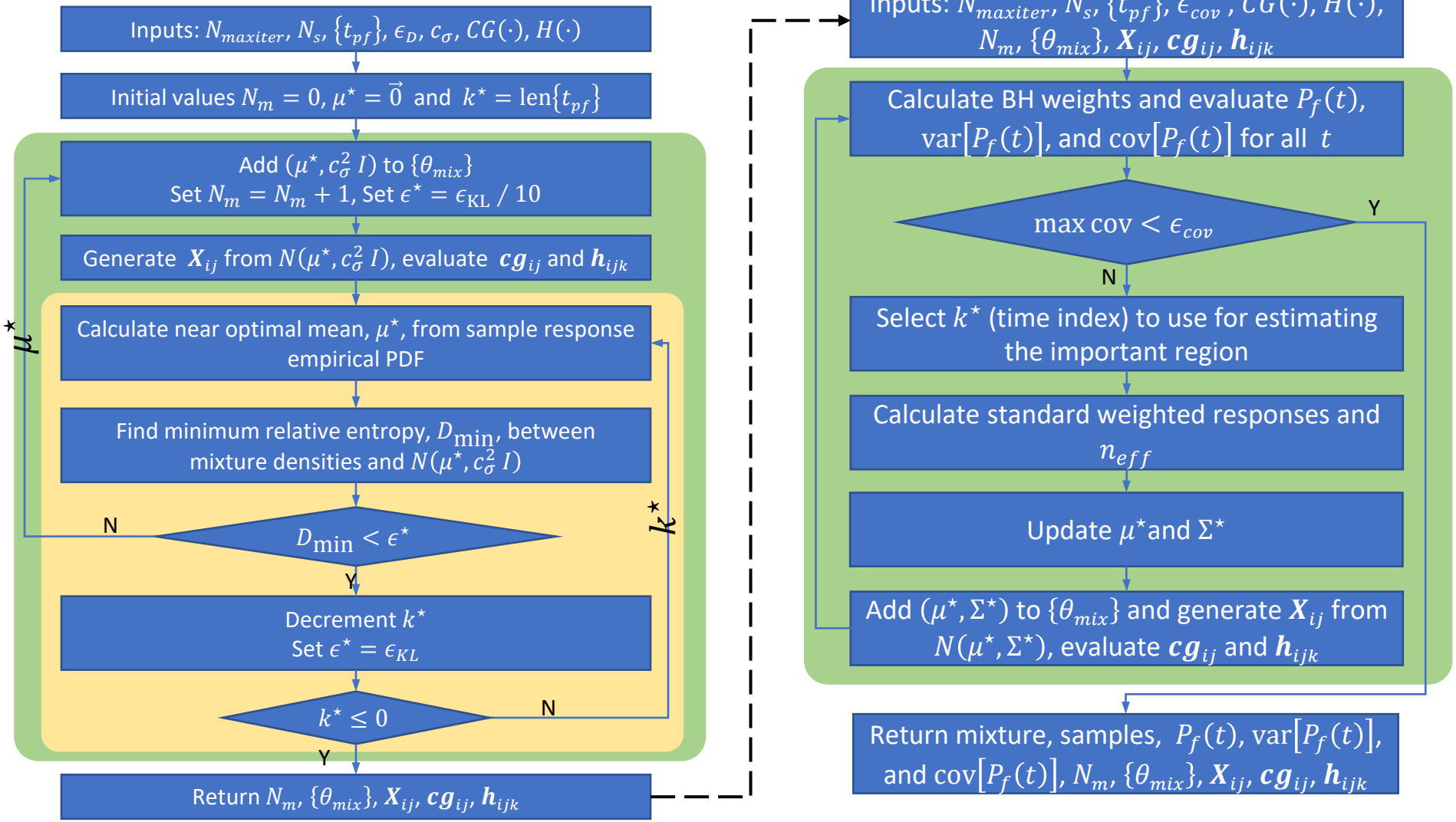


Backup

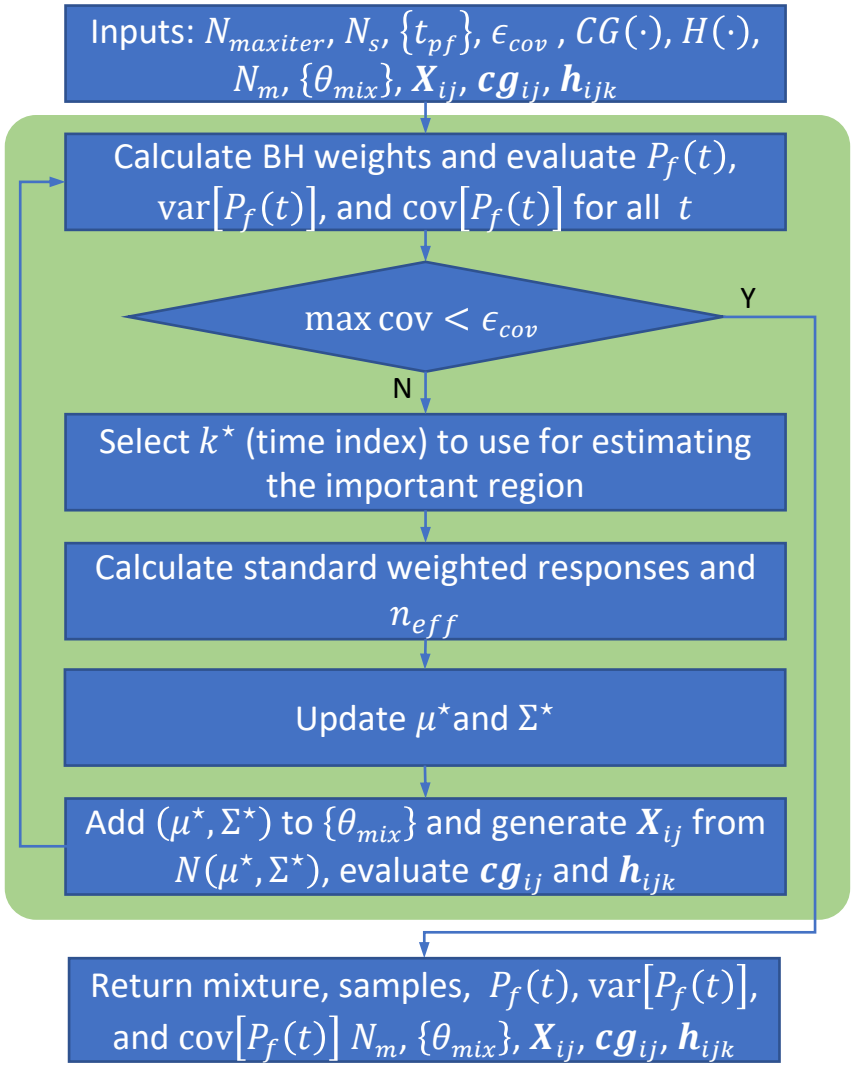
$$n_{eff} = \frac{\left(\sum_j \sum_i H(\boldsymbol{x}_{ij}, t)\, w_{std}(\boldsymbol{x}_{ij})\right)^2}{\sum_j \sum_i \left(H(\boldsymbol{x}_{ij}, t) w_{std}(\boldsymbol{x}_{ij})\right)^2}$$

- When one or two weights have much larger weight, updating the covariance matrix parameters can give a point-like or degenerate sampling density
- n_{eff} estimates the effective number of samples (between 0 and N) based on how evenly the weights are distributed
- When n_{eff} is low, only the location parameter is updated

PDTA AMIS Algorithm



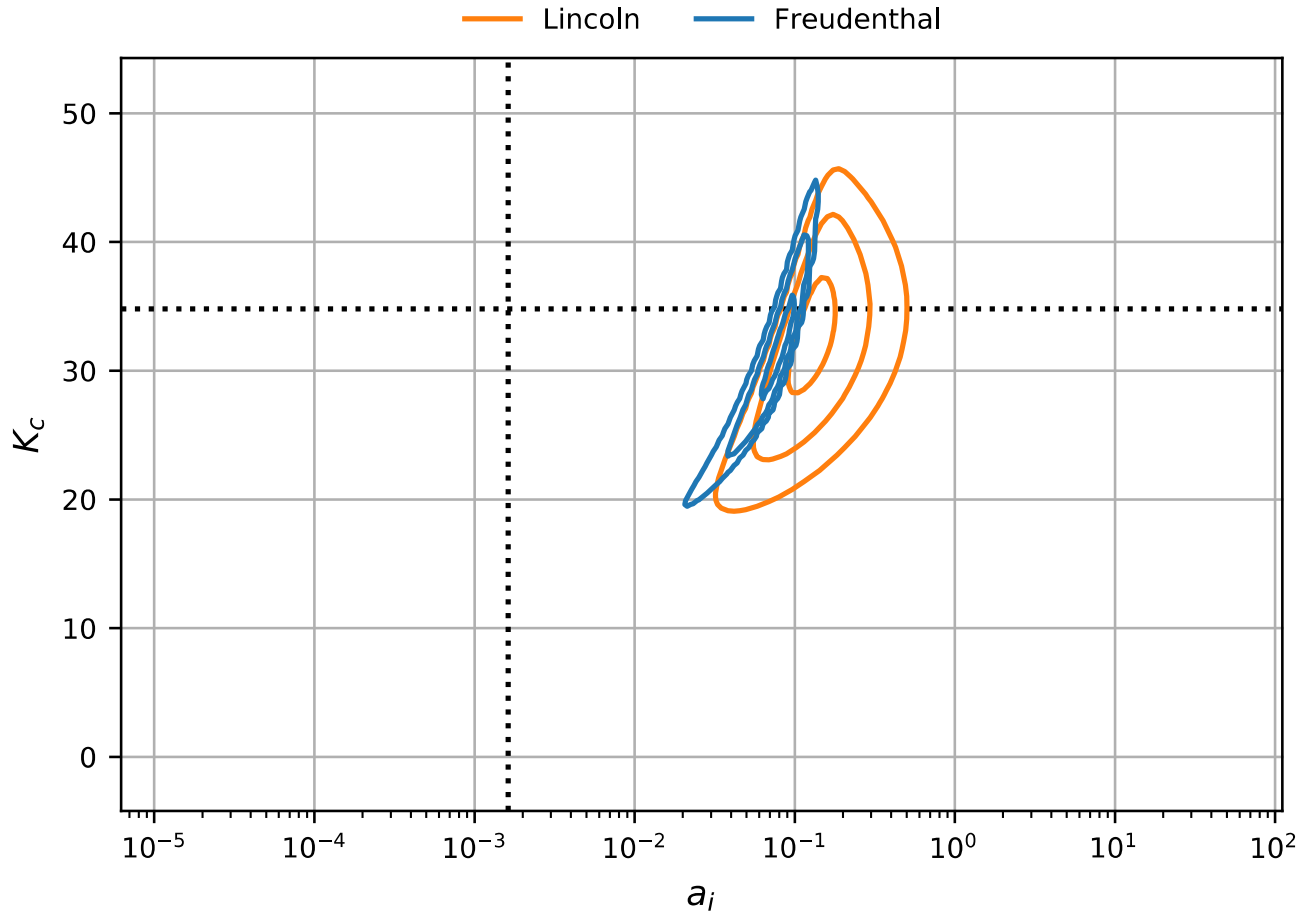
Adaptation Overview



- Adaptation algorithm adds sampling densities to the mixture until coefficient of variations (COVs) for all estimation times are below the user threshold, ϵ_{cov}
- Balance heuristic weights are used for to estimate POF and COV
- Standard weights are used for calculating new component density parameters
- Inputs include outputs from the initialization stage

Add text why the 2 different weights are used
Proof of consistency??

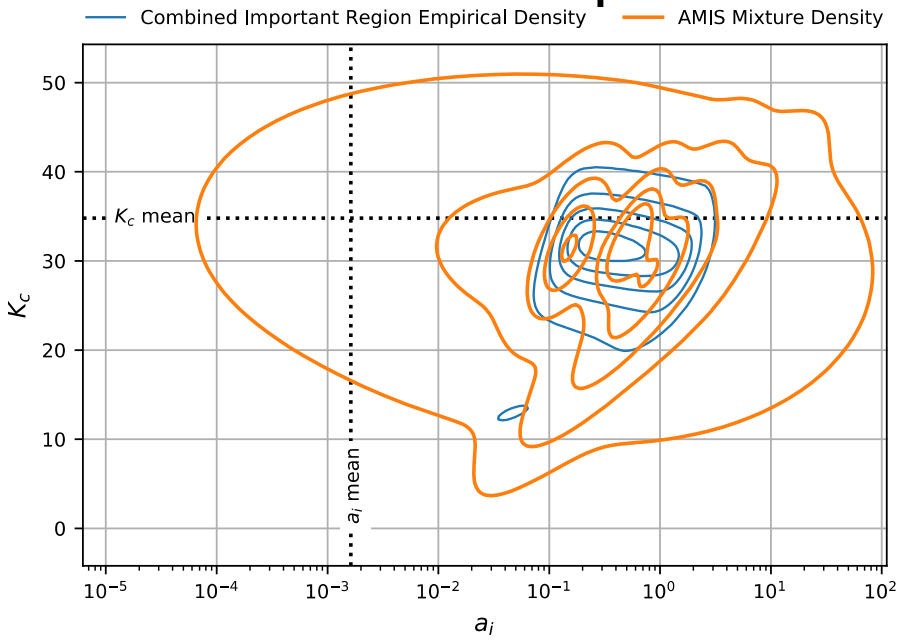
Single Evaluation Time Important Region Contour Comparison



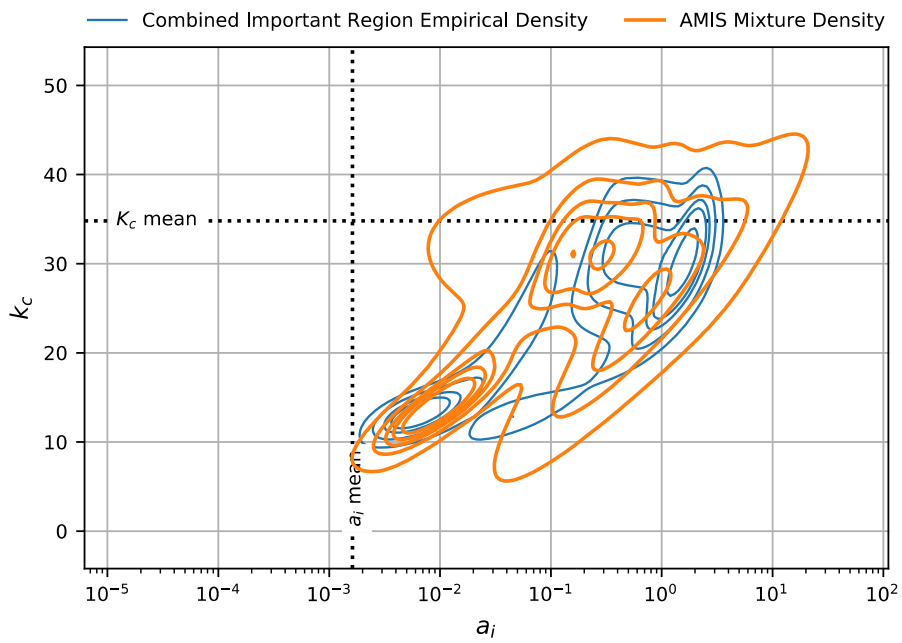
- Freudenthal important regions are much thinner
  - Reduced overlap which reduces potential for samples to contribute in more than one region

PDTA AMIS Reuse

POF without inspection

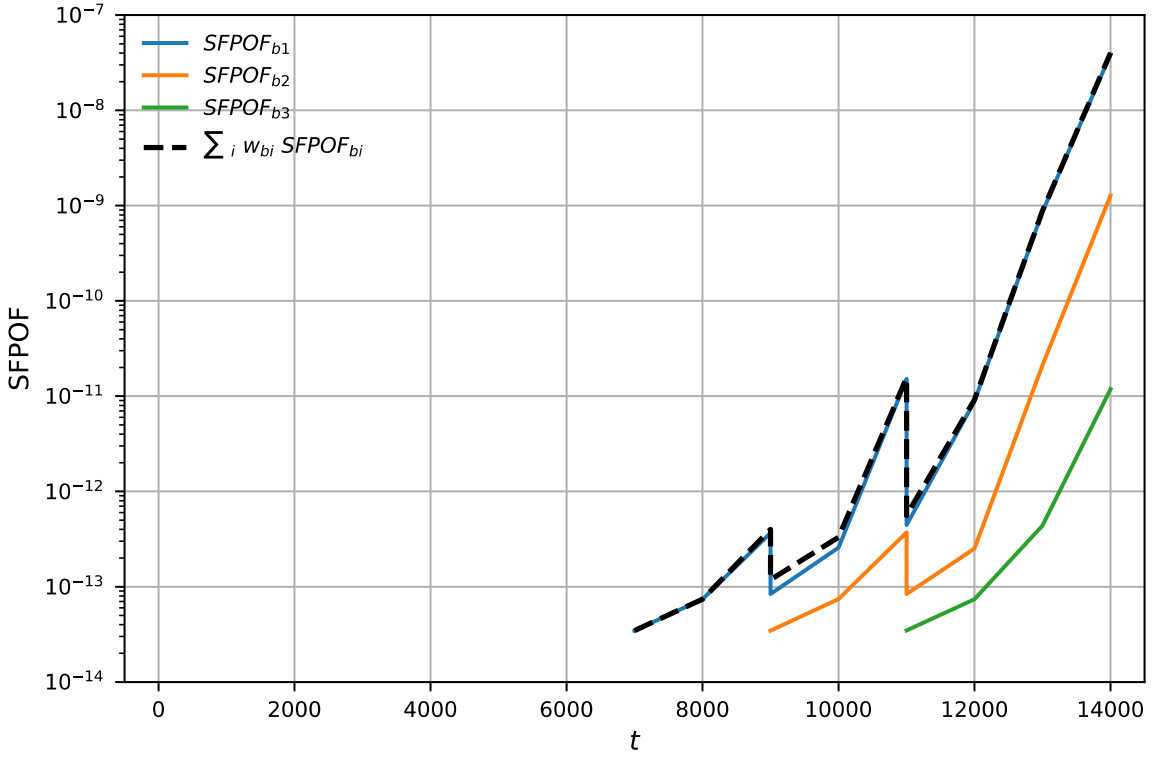


POF with one Inspection



- Initial run (left plot) not including any inspections, completed with 880 samples
- The stored crack growth analyses were reevaluated with a modified response function including an inspection
- After re-running the adaptation algorithm the mixture density has been re-adapted using 320 additional samples to include the new important region near (0.01, 10)

Modeling Repair Quality using Weighted Branch Integration



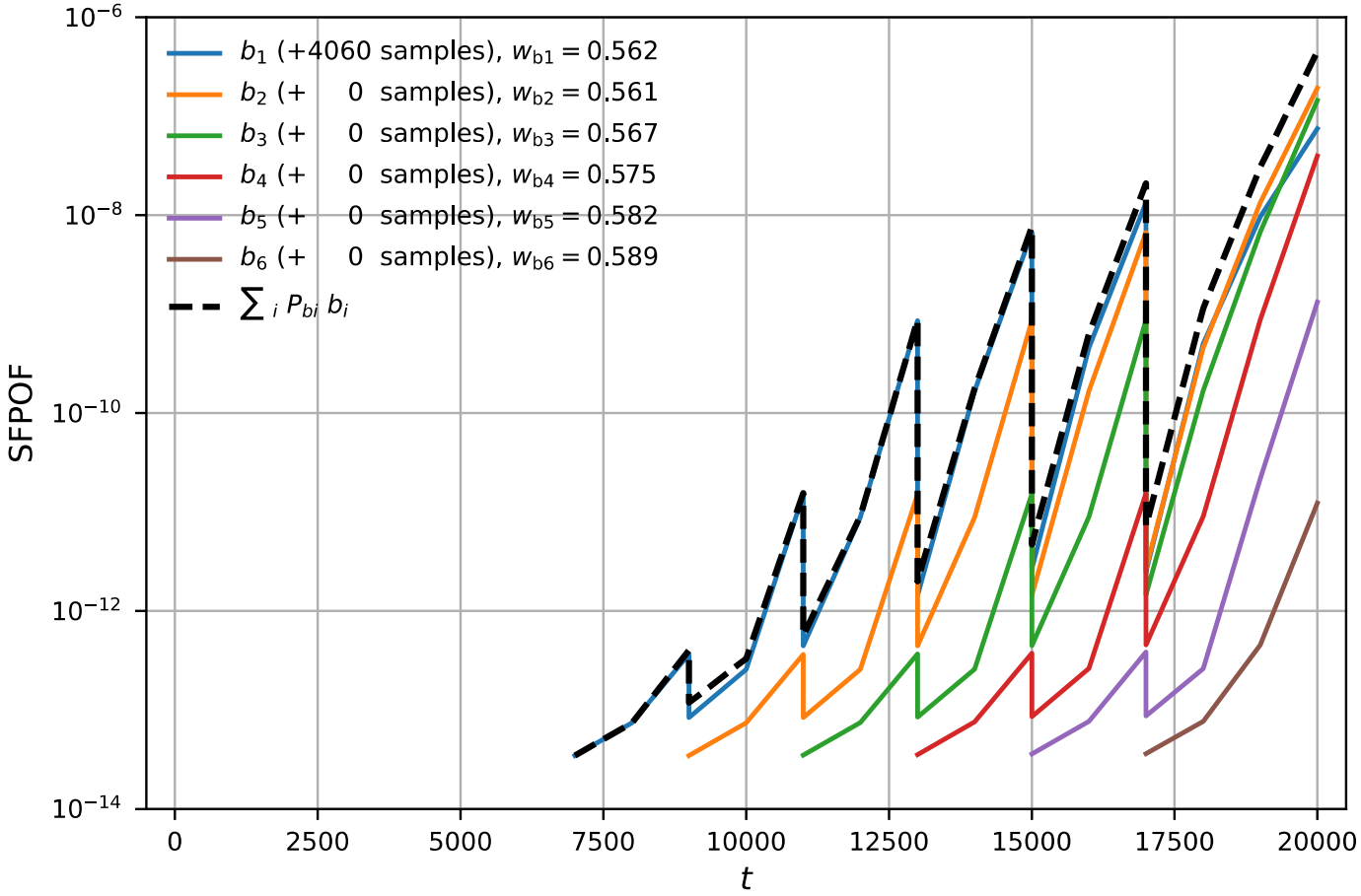
	Percentage of Cracks Detected			
	7000	9000	11000	$\mathbf{w_{branch}}$
Trunk	$PDET_{0,1}$	$PDET_{0,2}$	$PDET_{0,3}$	$w_0 = 1$
B 1	-	$PDET_{1,2}$	$PDET_{1,3}$	$w_1 = PDET_{0,1}$
B 2		-	$PDET_{2,3}$	$w_2 = PDET_{0,2} + w_1\, PDET_{1,2}$
B 3			-	$w_3 = PDET_{0,3} + w_1\, PDET_{1,3} + w_2\, PDET_{2,3}$

$$PDET_{i,j}(\tau_j) = \int PND(\tau_j)\, POD_j\left(a(\tau_j)\right)\, f_X\, dx$$

$$w_i = \sum_k^{i-1} w_k\, PDET_{k,j}$$

- Repair cracks account for the possibility that a new crack forms in the same component
- The weighted branch integration method includes repairs as independent PDTA analyses
  - Each branch is weighted according to the percentage of cracks detected in the trunk and other branches

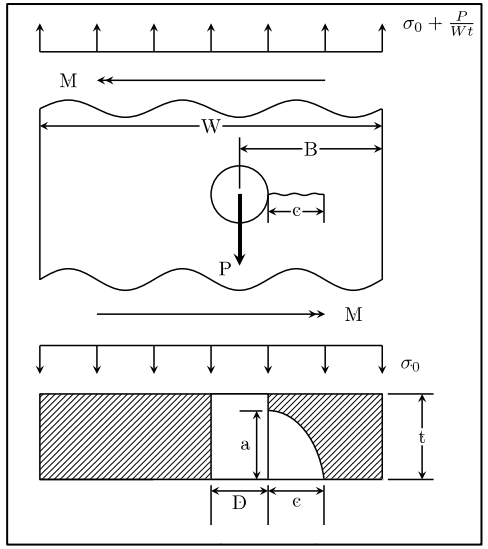
Repair Branch Analyses

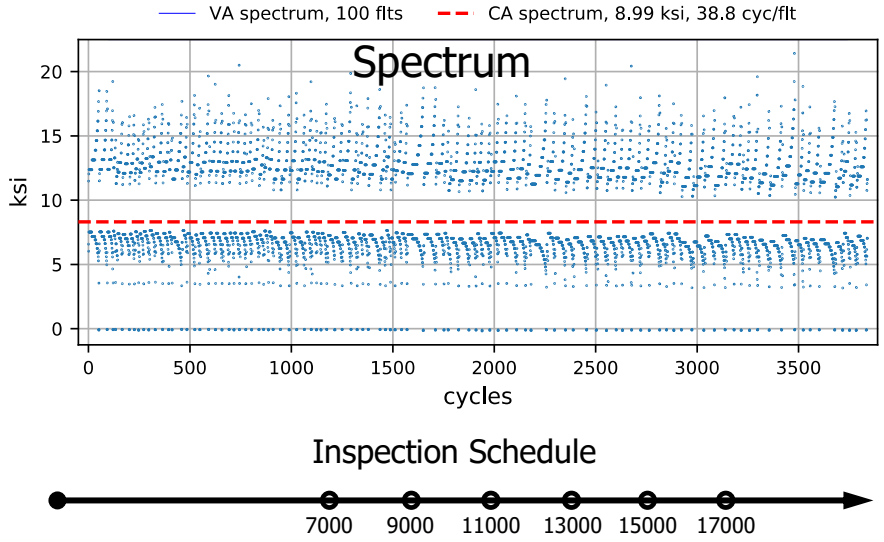


	Percentage of Cracks Detected						
	7000	9000	11000	13000	15000	17000	w_{branch}
Trunk	0.562	0.268	0.106	0.035	0.010	0.003	**1.000**
B 1	-	0.521	0.301	0.135	0.047	0.013	**0.562**
B 2		-	0.521	0.301	0.135	0.047	**0.561**
B 3			-	0.521	0.301	0.135	**0.567**
B 4				-	0.521	0.301	**0.575**
B 5					-	0.521	**0.582**
B 6						-	**0.589**

- Branches are identical analyses except for the part from $t = 17000$ to $t = 20000$
- The PDTA AMIS algorithm is able to estimate POFs and PDETs for all branches from 1st branch samples

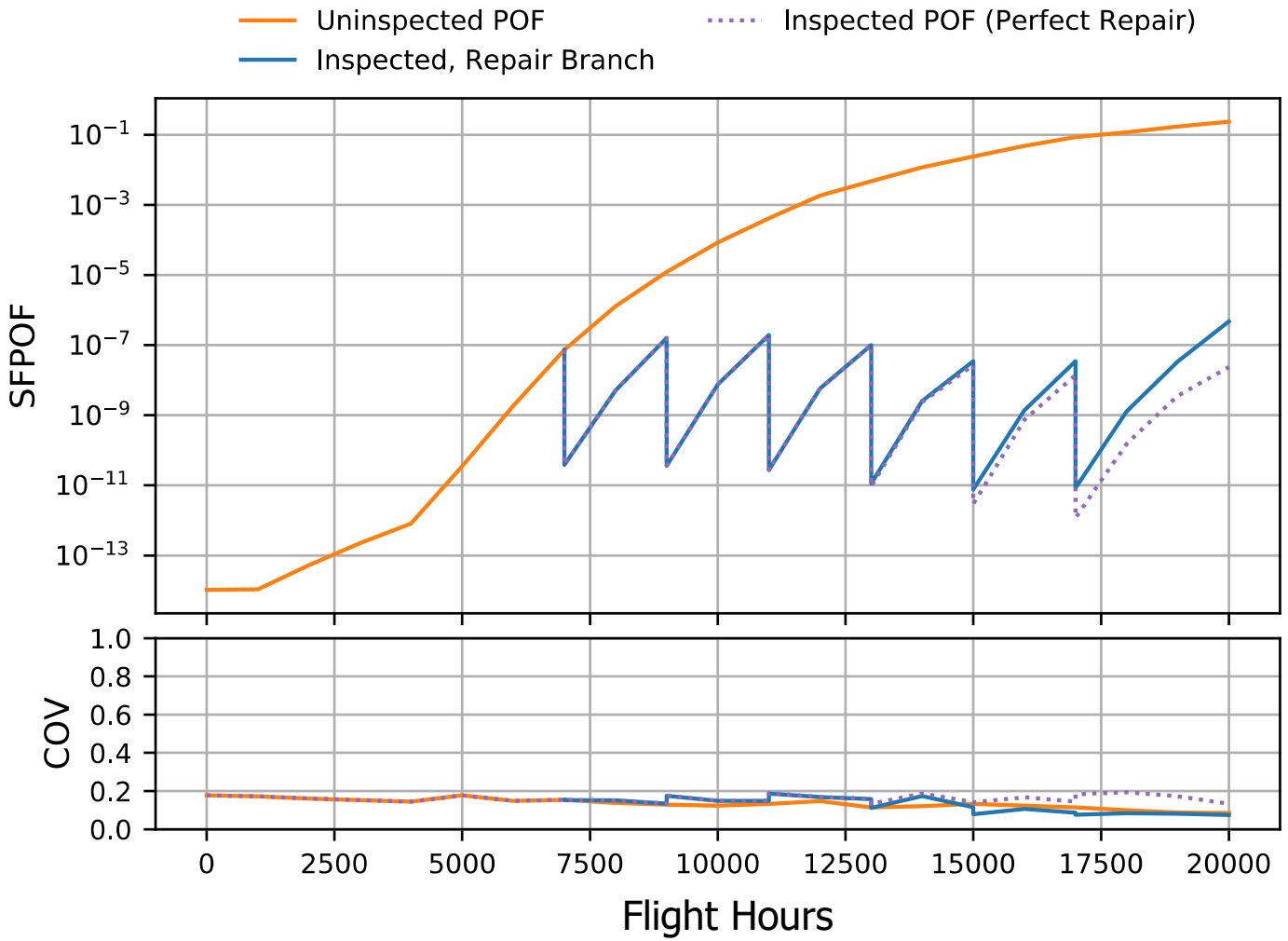
NASGRO Example with Inspections and Repairs





Parameter	Value
Width	Deterministic 2.5 in
Thickness	Deterministic 0.25 in
Initial Crack Size	$LN(0.005, 0.002)$ in
Aspect Ratio (A/C) [1]	$N(1.5, 0.14)$
Fracture Toughness	$N(34.8, 3.90)$ ksi $\sqrt{\text{in}}$
Log Paris Constant	$N(-8.777, 0.08)$
Paris Exponent	Deterministic 3.273
Hole Diameter	Deterministic 0.1562 in
Hole Offset [2]	$N(0.5, 0.05)$ in
Maximum Stress per Flight	$EVD(16.74, 2.08, 0.0)$ ksi
Probability of Detection	$LN(0.021, 0.028)$ in
Repair Quality (crack size)	$LN(0.01, 0.004)$ in

POF Results with Repairs



- PDTA AMIS
  - Trunk inspected POF: 4060 samples
  - Trunk uninspected POF: +0 samples
  - Trunk Percent Cracks Det: +140 samples
  - Repair crack Branch POFs: 4060 samples
- 8260 total samples