Credible Intervals for Probability of Failure with Gaussian Processes

ILLINOIS INSTITUTE OF TECHNOLOGY

Aleksei Sorokin (asorokin@hawk.iit.edu), Vishwas Rao

Argonne NATIONAL LABORATORY

Background

- *Efficiently approximate the probability of system failure subject to random parameters.*
- A system *fails* when the output of an *expensive* simulation exceeds a given threshold.
- Computational simulations are seeing increased use in reliability certification.
- Applications in structural design, power grids, safety certification etc.

Formulation

- $U \sim \mathcal{U}[0,1]^d$, the parameter distribution, perhaps after a change of variables.
- $g : [0,1]^d \to \mathbb{R}$, the canonical stochastic Gaussian Process (GP) on $(\Omega, \mathcal{F}, \mathbb{G})$.
- $F(g) := \{u \in [0,1]^d : g(u) \geq 0\}$, the failure region for $g \in \Omega$.
- $S(g) := \{u \in [0,1]^d : g(u) < 0\}$, the success region for $g \in \Omega$.
- $P(g) := \mathbb{U}(F(g)) = \mathbb{E}_{\mathbb{U}}[1_{F(g)}(U)]$, the random *probability of failure* on $(\Omega, \mathcal{F}, \mathbb{G})$.

Binary Classification with GPs

- $p(U) := \mathbb{E}_{\mathbb{G}}[1_{F(g)}(U)] = \Phi\left(\frac{m(U)}{\sigma(U)}\right)$, the pointwise probability of failure where m and σ are the mean and variance of GP g under $\mathbb{G}$.
- $\hat{F} := \{u \in [0,1]^d : m(u) \geq 0\}$, the predicted failure region.
- $\hat{S} := \{u \in [0,1]^d : m(u) < 0\}$, the predicted success region.
- True Positive, False Positive, True Negative, False Negative regions for $g \in \Omega$:

$$\mathrm{TP}(g) = \hat{F} \cap F(g), \quad \mathrm{FP}(g) = \hat{F} \cap S(g), \quad \mathrm{TN}(g) = \hat{S} \cap S(g), \quad \mathrm{FN}(g) = \hat{S} \cap F(g).$$

- $\mathrm{ACC}(U) := \mathbb{E}_{\mathbb{G}}\left[1_{\mathrm{TP}(g)}(U) + 1_{\mathrm{TN}(g)}(U)\right] = \max\{p(U), 1-p(U)\}$, the expected accuracy.
- $\mathrm{ERR}(U) := 1 - \mathrm{ACC}(U) = \min\{p(U), 1-p(U)\}$, the *expected error.*

Estimator and Credible Interval

- $\hat{P} := \mathbb{E}_{\mathbb{U}}[1_{\hat{F}}(U)]$, the binary prediction based estimator.
- $\mathbb{G}(P \in [\underline{P}, \overline{P}]) \geq 1 - \alpha$, the credible interval where

$$\underline{P} := \max\left\{\hat{P} - \hat{\gamma}, 0\right\}, \qquad \overline{P} := \min\left\{\hat{P} + \hat{\gamma}, 1\right\}, \qquad \hat{\gamma} := \frac{\mathbb{E}_{\mathbb{U}}\left[\mathrm{ERR}(U)\right]}{\alpha}.$$

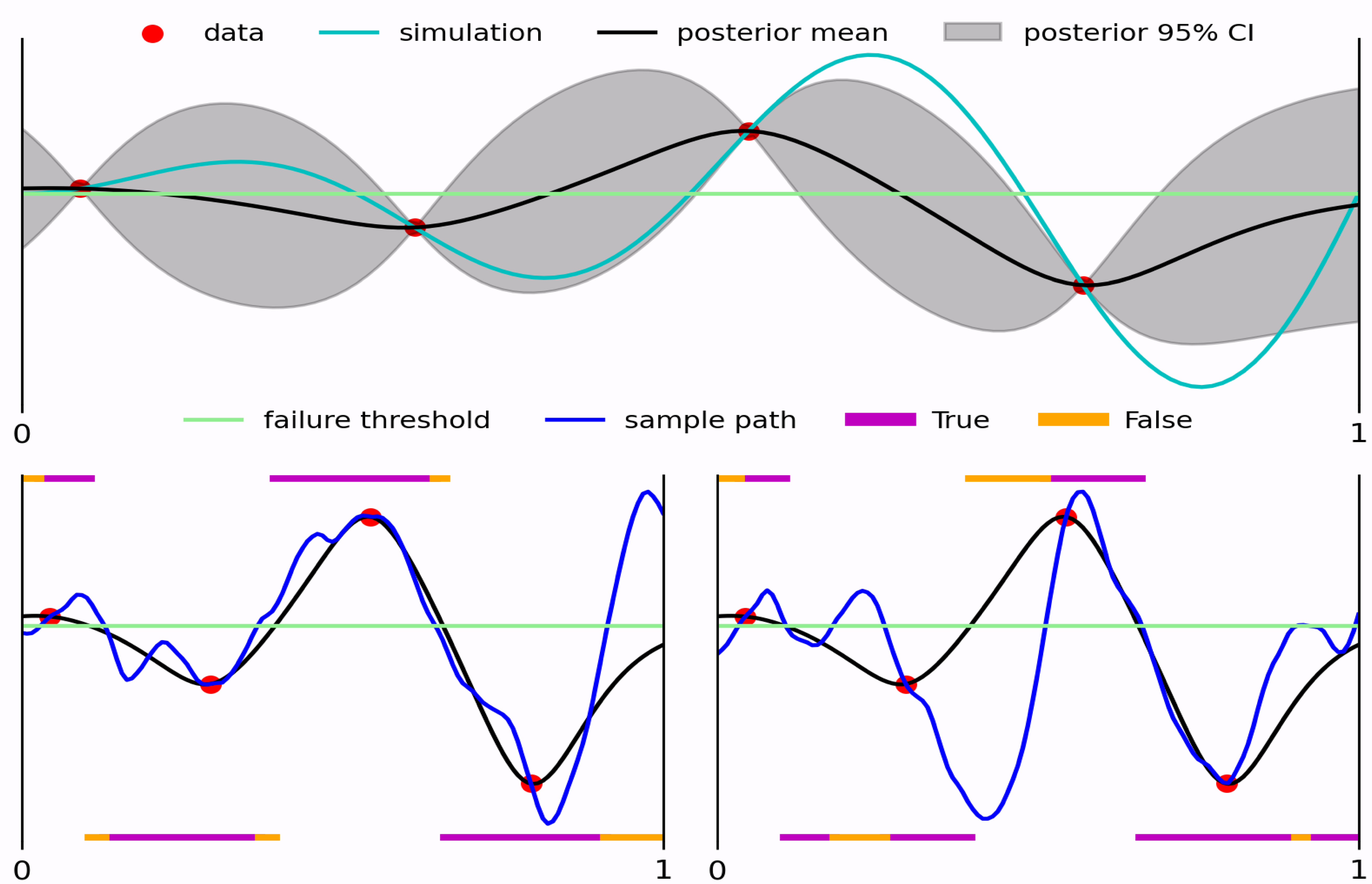
Algorithm

Input: a GP prior distribution $\mathbb{G}_0$, a deterministic simulation $g : [0,1]^d \to \mathbb{R}$, and $n \to 0$

1. Evaluate g at some batch of nodes $X_1, \dots, X_b$ and set $n \to n + b$.
2. Update the GP distribution to $\mathbb{G}_n$ based on all previous evaluations of g.
3. Compute (Quasi-)Monte Carlo approximates $\hat{P}^{\mathrm{QMC}_n}, \hat{\gamma}^{\mathrm{QMC}_n}$ of $\hat{P}_n, \hat{\gamma}_n$ from $\mathbb{G}_n$.
4. If the approximate $1 - \alpha$ credible interval

$$[\underline{P}^{\mathrm{QMC}_n}, \overline{P}^{\mathrm{QMC}_n}] := [\max\{\hat{P}^{\mathrm{QMC}_n} - \hat{\gamma}^{\mathrm{QMC}_n}, 0\}, \min\{\hat{P}^{\mathrm{QMC}_n} + \hat{\gamma}^{\mathrm{QMC}_n}, 1\}]$$

is too wide and the sample budget it not expired, go to step 1.



Above: The top figure shows the true simulation and a posterior Gaussian process fit to a few data points. The bottom two plots show some Gaussian process sample paths $g_1, g_2 \in \Omega$ with their corresponding TP, FP, TN, and FN regions. The predicted failure TP and FP regions are shown along the top of each plot while the predicted success TN and FN regions are shown along the bottom of each plot.
Below: Moving left to right traverses algorithm iterations. The plots in the top row show the posterior mean and predicted failure boundary in green. The new samples in each iteration are shown in red while samples from previous iterations are in blue. The bottom row shows convergence of the credible interval interval in gray with half width in pink. The true error is visualized in blue. Notice the credible interval captures the true probability of failure shown in black.