Adding hypothesis testing to evolutionary RBDO with Monte Carlo simulations

D. Salazar1, R. Le Riche1,2 and X. Bay1

1 Ecole des Mines de Saint-Etienne, G2I/3MI, Saint-Etienne, France, {leriche,bay}@emse.fr, danielsalazaraponte@gmail.com.
2 CNRS laboratory Claude Goux, UMR 5146, Saint-Etienne, France.

Engineers have a long history with accounting for uncertainties in the design process [2]. They have first defined critical loadings and safety factors. More recently, uncertainty quantification is receiving a large attention from the engineering community because a finer characterization of the uncertainties is seen as an important performance reserve. In the presence of uncertainties, the performance of an individual system varies. Reliability based design optimization (RBDO) and robust optimization average out uncertainties by ultimately seeking solutions having the best statistical performance measure [1].

This study addresses reliability based optimization where the statistical performance criteria are estimated by Monte Carlo simulations. In particular, the objective function is expressed as a percentile: without loss of generality, let \( x \) denote the vector of deterministic optimization variables, \( U \) the random vector of uncertainties and \( f(x, U) \) the random objective function to minimize. We would like to minimize the 90% percentile of \( f \),

\[
\min_{x \in \mathbb{R}^n} q_{90}(x)
\]

\( q_{90}(x) \) is a performance measure such that 90% of the values of \( f(x, u) \) are below \( q_{90}(x) \). Typically, the percentiles of simulators outputs are not known analytically and are estimated by Monte Carlo (MC) simulations. If \( m \) simulations are performed, the empirical percentile, \( \hat{q}_{90}(x) \), is the \( \lfloor 0.9 \times m \rfloor \)th smallest value of \( f(x, u^1), \ldots, f(x, u^m) \). The percentile estimate, \( \hat{q}_{90}(x) \), is a noisy function whose average and variance are calculated by performing \( n_b \) batches of \( s_b \) independent simulations. \( s_b \) is kept constant at 20 evaluations, which is the smallest number of evaluations for which the distribution of \( \hat{q}_{90} \) can be assumed Gaussian.

The minimization of \( \hat{q}_{90}(x) \) is performed using an evolution strategy algorithm (ES, [1]) because it can be applied to noisy functions. To keep the number of parameters as low as possible, the simplest ES is studied, i.e., a (1+1)-ES with constant isotropic gaussian mutation: if \( x^t \) is the current optimization iterate, new candidate points are generated from \( x^t_j = x^t_j + \sigma N_j(0,1) \), \( j = 1, n \). In a “traditional” ES algorithm, a candidate point is accepted as new iterate if its performance is better than that of the previous iterate: \( x^{t+1} \leftarrow x^t \) if \( \hat{q}_{90}(x^t) < \hat{q}_{90}(x^t) \). An originality of this work is to take into account the random nature of \( \hat{q}_{90}(x) \) and replace the inequality by an hypothesis testing of given significance level \( \alpha \) [3]. The null (default) hypothesis \( H_0 \) is that the candidate point \( x^t \) is better than the current iterate \( x^t \). We will wrongly reject \( H_0 \) in at most \( \alpha \) percent of the cases. For \( \alpha < 0.5 \), the optimizer can be thought of as “exploratory” because it will tend to accept candidate points, and vice versa, \( \alpha > 0.5 \) yields a conservative optimizer in the sense that sufficient statistical evidence of the new point being better is required to accept the move. \( \alpha = 0.5 \) is the traditional ES algorithm with inequality test based on the empirical values.

An important obstacle to accounting for uncertainties in design is the cumulated computational cost
Figure 1: Comparison of the convergence of the traditional ($\alpha = 0.50$) and the conservative ($\alpha = 0.90$) optimizers on the 2D hyperbolic function, $n_b = 2$. The curve represents the average distance to the optimum, based on 30 independent runs, as a function of the number of calls to the simulator. The use of hypothesis testing with $\alpha = 0.9$ reduces the probability of divergence, hence improving the convergence during the first stage of the process (i.e., before 10000 analyses, see insert). Beyond 10000 analyses, no hypothesis testing is the best strategy.

The coupled effects of $\sigma$, $\alpha$ and $n_b$ are empirically studied here on two academic test functions and on a truss design problem. The academic test functions are noisy parabolic and hyperbolic functions, with the properties that their (average / variance) $\hat{q}$ ratios decrease and increase, respectively, as $x$ tends towards the optimum. Our principal findings, partly illustrated in Figure 1, are the following:

- Exploratory optimizers tend to diverge on noisy functions, therefore $\alpha \geq 0.5$ is recommended.
- Optimal step sizes $\sigma$ increase with the noise in the function (i.e., when $n_b$ decreases).
- The optimal $n_b$ increases when the signal-to-noise ratio or $\sigma$ decrease.
- The traditional optimizer outperforms the conservative optimizer (with hypothesis testing, $\alpha = 0.9$) excepted when there is a divergence risk (low signal-to-noise ratio, large $\sigma$’s, see Figure 1).

References