Overview of Reliability Analysis and Design Capabilities in DAKOTA

M. S. Eldred*
Sandia National Laboratories†*, Albuquerque, NM 87185*

B. J. Bichon‡
Vanderbilt University, Nashville, TN 37235

B. M. Adams§
Sandia National Laboratories, Albuquerque, NM 87185

Abstract. Reliability methods are probabilistic algorithms for quantifying the effect of uncertainties in simulation input on response metrics of interest. In particular, they compute approximate response function distribution statistics (probability, reliability, and response levels) based on specified probability distributions for input random variables. In this paper, recent algorithm research in first and second-order reliability methods is overviewed for both the forward reliability analysis of computing probabilities for specified response levels (the reliability index approach (RIA)) and the inverse reliability analysis of computing response levels for specified probabilities (the performance measure approach (PMA)). A number of algorithmic variations have been explored, and the effect of different limit state approximations, probability integrations, warm starting, most probable point search algorithms, and Hessian approximations is discussed. These reliability analysis capabilities provide the foundation for reliability-based design optimization (RBDO) methods, and bi-level and sequential formulations are presented. These RBDO formulations may employ analytic sensitivities of reliability metrics with respect to design variables that either augment or define distribution parameters for the uncertain variables. Relative performance of these reliability analysis and design algorithms is presented for a number of benchmark test problems using the DAKOTA software, and algorithm recommendations are given. These recommended algorithms are subsequently being applied to real-world applications in the probabilistic analysis and design of micro-electro-mechanical systems, and initial experiences with this deployment are provided.

Keywords: Uncertainty, Reliability, Reliability-based design optimization, Software, MEMS

---

* Principal Member of Technical Staff, Optimization and Uncertainty Estimation Department.

† Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

‡ NSF IGERT Fellow, Department of Civil and Environmental Engineering.

§ Limited Term Employee, Optimization and Uncertainty Estimation Department.

1. Introduction

Reliability methods are probabilistic algorithms for quantifying the effect of uncertainties in simulation input on response metrics of interest. In particular, they perform uncertainty quantification (UQ) by computing approximate response function distribution statistics based on specified probability distributions for input random variables. These response statistics include response mean, response standard deviation, and cumulative or complementary cumulative distribution function (CDF/CCDF) response level and probability/reliability level pairings. These methods are often more efficient at computing statistics in the tails of the response distributions (events with low probability) than sampling-based approaches since the number of samples required to resolve a low probability can be prohibitive. Thus, these methods, as their name implies, are often used in a reliability context for assessing the probability of failure of a system when confronted with an uncertain environment.

A number of classical reliability analysis methods are discussed in (Haldar and Mahadevan, 2000), including Mean-Value First-Order Second-Moment (MVFOSM), First-Order Reliability Method (FORM), and Second-Order Reliability Method (SORM). More recent methods which seek to improve the efficiency of FORM analysis through limit state approximations include the use of local and multipoint approximations in Advanced Mean Value methods (AMV/AMV+ (Wu et al., 1990)) and Two-point Adaptive Nonlinearity Approximation-based methods (TANA (Wang and Grandhi, 1994; Xu and Grandhi, 1998)), respectively. Each of the FORM-based methods can be employed for "forward" or "inverse" reliability analysis through the reliability index approach (RIA) or performance measure approach (PMA), respectively, as described in (Tu et al., 1999).

The capability to assess reliability is broadly useful within a design optimization context, and reliability-based design optimization (RBDO) methods are popular approaches for designing systems while accounting for uncertainty. RBDO approaches may be broadly characterized as bi-level (in which the reliability analysis is nested within the optimization, e.g. (Allen and Maute, 2004)), sequential (in which iteration occurs between optimization and reliability analysis, e.g. (Wu et al., 2001; Du and Chen, 2004)), or unilevel (in which the design and reliability searches are combined into a single optimization, e.g. (Agarwal et al., 2004)). Bi-level RBDO methods are simple and general-purpose, but can be computationally demanding. Sequential and unilevel methods seek to reduce computational expense by breaking the nested relationship through the use of iterated or simultaneous approaches.

In order to provide access to a variety of uncertainty quantification capabilities for analysis of large-scale engineering applications on high-performance parallel computers, the DAKOTA project (Eldred et al., 2003) at Sandia National Laboratories has developed a suite of algorithmic capabilities known as DAKOTA/UQ (Wojtkiewicz et al., 2001). This package contains the reliability analysis capabilities described in this paper and enables the RBDO approaches, and is freely available for download worldwide through an open source license.

This paper overviews recent algorithm research activities that have explored a variety of approaches for performing reliability analysis. In particular, forward and inverse reliability analyses have been explored using multiple limit state approximation, probability integration, warm starting, Hessian approximation, and optimization algorithm selections. These uncertainty quantification capabilities have also provided a foundation for exploring bi-level and sequential RBDO formulations.

Sections 2 and 3 describe these algorithmic components, Section 4 summarizes computational results for four benchmark test problems, Section 5 presents initial deployment of these methodologies to the probabilistic analysis and design of MEMS, and Section 6 provides concluding remarks.

2. Reliability Method Formulations

2.1. Mean Value

The Mean Value method (MV, also known as MVFOSM in (Haldar and Mahadevan, 2000)) is the simplest, least-expensive reliability method because it estimates the response means, response standard deviations, and all CDF/CCDF response-probability-reliability levels from a single evaluation of response functions and their gradients at the uncertain variable means. This approximation can have acceptable accuracy when the response functions are nearly linear and their distributions are approximately Gaussian, but can have poor accuracy in other situations. The expressions for approximate response mean μ_g, approximate response standard deviation σ_g, response target to approximate probability/reliability level mapping ($\bar{z} \to p, \beta$), and probability/reliability target to approximate response level mapping ($\bar{p}, \bar{\beta} \to z$) are

$$\begin{aligned}
\mu_g &= g(\mu_{\mathbf{x}}) &(1)\\
\sigma_g &= \sum_i \sum_j Cov(i,j) \frac{dg}{dx_i}(\mu_{\mathbf{x}}) \frac{dg}{dx_j}(\mu_{\mathbf{x}}) &(2)\\
\beta_{cdf} &= \frac{\mu_g - \bar{z}}{\sigma_g} &(3)\\
\beta_{ccdf} &= \frac{\bar{z} - \mu_g}{\sigma_g} &(4)\\
z &= \mu_g - \sigma_g \bar{\beta}_{cdf} &(5)\\
z &= \mu_g + \sigma_g \bar{\beta}_{ccdf} &(6)
\end{aligned}$$

respectively, where $\mathbf{x}$ are the uncertain values in the space of the original uncertain variables ("x-space"), $g(\mathbf{x})$ is the limit state function (the response function for which probability-response level pairs are needed), and β_{cdf} and β_{ccdf} are the CDF and CCDF reliability indices, respectively.

With the introduction of second-order limit state information, MVSOSM calculates a second-order mean as

$$\mu_g = g(\mu_{\mathbf{x}}) + \frac{1}{2} \sum_i \sum_j Cov(i,j) \frac{d^2 g}{dx_i dx_j}(\mu_{\mathbf{x}}) \tag{7}$$

This is commonly combined with a first-order variance (Eq. 2), since second-order variance involves higher order distribution moments (skewness, kurtosis) (Haldar and Mahadevan, 2000) which are often unavailable.

The first-order CDF probability $p(g \leq z)$, first-order CCDF probability $p(g > z)$, β_{cdf}, and β_{ccdf} are related to one another through

$$p(g \leq z) = \Phi(-\beta_{cdf}) \tag{8}$$

$$
\begin{aligned}
p(g > z) &= \Phi(-\beta_{ccdf}) && (9)\\
\beta_{cdf} &= -\Phi^{-1}(p(g \leq z)) && (10)\\
\beta_{ccdf} &= -\Phi^{-1}(p(g > z)) && (11)\\
\beta_{cdf} &= -\beta_{ccdf} && (12)\\
p(g \leq z) &= 1 - p(g > z) && (13)
\end{aligned}
$$

where $\Phi()$ is the standard normal cumulative distribution function. A common convention in the literature is to define g in such a way that the CDF probability for a response level z of zero (i.e., $p(g \leq 0)$) is the response metric of interest. The formulations in this paper are not restricted to this convention and are designed to support CDF or CCDF mappings for general response, probability, and reliability level sequences.

2.2. MPP Search Methods

All other reliability methods solve a nonlinear optimization problem to compute a most probable point (MPP) and then integrate about this point to compute probabilities. The MPP search is performed in uncorrelated standard normal space ("u-space") since it simplifies the probability integration: the distance of the MPP from the origin has the meaning of the number of input standard deviations separating the mean response from a particular response threshold. The transformation from correlated non-normal distributions (x-space) to uncorrelated standard normal distributions (u-space) is denoted as $\mathbf{u} = T(\mathbf{x})$ with the reverse transformation denoted as $\mathbf{x} = T^{-1}(\mathbf{u})$. These transformations are nonlinear in general, and possible approaches include the Rosenblatt (Rosenblatt, 1952), Nataf (Der Kiureghian and Liu, 1986), and Box-Cox (Box and Cox, 1964) transformations. The nonlinear transformations may also be linearized, and common approaches for this include the Rackwitz-Fiessler (Rackwitz and Fiessler, 1978) two-parameter equivalent normal and the Chen-Lind (Chen and Lind, 1983) and Wu-Wirsching (Wu and Wirsching, 1987) three-parameter equivalent normals. The results in this paper employ the Nataf nonlinear transformation which occurs in the following two steps. To transform between the original correlated x-space variables and correlated standard normals ("z-space"), the CDF matching condition is used:

$$\Phi(z_i) = F(x_i) \tag{14}$$

where $F()$ is the cumulative distribution function of the original probability distribution. Then, to transform between correlated z-space variables and uncorrelated u-space variables, the Cholesky factor $\mathbf{L}$ of a modified correlation matrix is used:

$$\mathbf{z} = \mathbf{L}\mathbf{u} \tag{15}$$

where the original correlation matrix for non-normals in x-space has been modified to represent the corresponding correlation in z-space (Der Kiureghian and Liu, 1986).

The forward reliability analysis algorithm of computing CDF/CCDF probability/reliability levels for specified response levels is called the reliability index approach (RIA), and the inverse reliability analysis algorithm of computing response levels for specified CDF/CCDF probability/reliability levels is called the performance measure approach (PMA) (Tu et al., 1999). The differences between the

RIA and PMA formulations appear in the objective function and equality constraint formulations used in the MPP searches. For RIA, the MPP search for achieving the specified response level $\bar{z}$ is formulated as

$$\begin{aligned} &\text{minimize} \quad \mathbf{u}^T\mathbf{u} \\ &\text{subject to} \quad G(\mathbf{u}) = \bar{z} \end{aligned} \tag{16}$$

and for PMA, the MPP search for achieving the specified reliability/probability level $\bar{\beta}, \bar{p}$ is formulated as

$$\begin{aligned} &\text{minimize} \quad \pm G(\mathbf{u}) \\ &\text{subject to} \quad \mathbf{u}^T\mathbf{u} = \bar{\beta}^2 \end{aligned} \tag{17}$$

where $\mathbf{u}$ is a vector centered at the origin in u-space and $g(\mathbf{x}) \equiv G(\mathbf{u})$ by definition. In the RIA case, the optimal MPP solution $\mathbf{u}^*$ defines the reliability index from $\beta = \pm\|\mathbf{u}^*\|_2$, which in turn defines the CDF/CCDF probabilities (using Eqs. 8-9 in the case of first-order integration). The sign of β is defined by

$$G(\mathbf{u}^*) > G(\mathbf{0}) : \beta_{cdf} < 0, \beta_{ccdf} > 0 \tag{18}$$

$$G(\mathbf{u}^*) < G(\mathbf{0}) : \beta_{cdf} > 0, \beta_{ccdf} < 0 \tag{19}$$

where $G(\mathbf{0})$ is the median limit state response computed at the origin in u-space (where $\beta_{cdf} = \beta_{ccdf} = 0$ and first-order $p(g \leq z) = p(g > z) = 0.5$). In the PMA case, the sign applied to $G(\mathbf{u})$ (equivalent to minimizing or maximizing $G(\mathbf{u})$) is similarly defined by $\bar{\beta}$

$$\bar{\beta}_{cdf} < 0, \bar{\beta}_{ccdf} > 0 : \text{maximize } G(\mathbf{u}) \tag{20}$$

$$\bar{\beta}_{cdf} > 0, \bar{\beta}_{ccdf} < 0 : \text{minimize } G(\mathbf{u}) \tag{21}$$

and the limit state at the MPP ($G(\mathbf{u}^*)$) defines the desired response level result.

When performing PMA with specified $\bar{p}$, one must compute $\bar{\beta}$ to include in Eq. 17. While this is a straightforward one-time calculation for first-order integrations (Eqs. 10-11), the use of second-order integrations complicates matters since the $\bar{\beta}$ corresponding to the prescribed $\bar{p}$ is a function of the Hessian of G (see Eq. 38), which in turn is a function of location in u-space. A generalized reliability index (Eq. 50), which would allow a one-time calculation, may not be used since equality with $\mathbf{u}^T\mathbf{u}$ is not meaningful. The $\bar{\beta}$ target must therefore be updated in Eq. 17 as the minimization progresses (e.g., using Newton's method to solve Eq. 38 for $\bar{\beta}$ given $\bar{p}$ and κ_i). This works best when $\bar{\beta}$ can be fixed during the course of an approximate optimization, such as for the AMV^2+ and TANA methods described in Section 2.2.1. For second-order PMA without limit state approximation cycles (i.e., PMA SORM), the constraint must be continually updated and the constraint derivative should include $\nabla_{\mathbf{u}}\bar{\beta}$, which would require third-order information for the limit state to compute derivatives of the principal curvatures. This is impractical, so the PMA SORM constraint derivatives are only approximated analytically or estimated numerically. Potentially for this reason, PMA SORM has not been widely explored in the literature.

2.2.1. *Limit state approximations*

There are a variety of algorithmic variations that can be explored within RIA/PMA reliability analysis. First, one may select among several different limit state approximations that can be used to reduce computational expense during the MPP searches. Local, multipoint, and global approximations of the limit state are possible. (Eldred et al., 2005) investigated local first-order limit state approximations, and (Eldred et al., 2006) investigated local second-order and multipoint approximations. These techniques include:

1. a single Taylor series per response/reliability/probability level in x-space centered at the uncertain variable means. The first-order approach is commonly known as the Advanced Mean Value (AMV) method:
$$g(\mathbf{x}) \cong g(\mu_{\mathbf{x}}) + \nabla_x g(\mu_{\mathbf{x}})^T (\mathbf{x} - \mu_{\mathbf{x}}) \tag{22}$$
and the second-order approach has been named AMV^2:
$$g(\mathbf{x}) \cong g(\mu_{\mathbf{x}}) + \nabla_{\mathbf{x}} g(\mu_{\mathbf{x}})^T (\mathbf{x} - \mu_{\mathbf{x}}) + \frac{1}{2}(\mathbf{x} - \mu_{\mathbf{x}})^T \nabla^2_{\mathbf{x}} g(\mu_{\mathbf{x}})(\mathbf{x} - \mu_{\mathbf{x}}) \tag{23}$$

2. same as AMV/AMV^2, except that the Taylor series is expanded in u-space. The first-order option has been termed the u-space AMV method:
$$G(\mathbf{u}) \cong G(\mu_{\mathbf{u}}) + \nabla_u G(\mu_{\mathbf{u}})^T (\mathbf{u} - \mu_{\mathbf{u}}) \tag{24}$$
where $\mu_{\mathbf{u}} = T(\mu_{\mathbf{x}})$ and is nonzero in general, and the second-order option has been named the u-space AMV^2 method:
$$G(\mathbf{u}) \cong G(\mu_{\mathbf{u}}) + \nabla_{\mathbf{u}} G(\mu_{\mathbf{u}})^T (\mathbf{u} - \mu_{\mathbf{u}}) + \frac{1}{2}(\mathbf{u} - \mu_{\mathbf{u}})^T \nabla^2_{\mathbf{u}} G(\mu_{\mathbf{u}})(\mathbf{u} - \mu_{\mathbf{u}}) \tag{25}$$

3. an initial Taylor series approximation in x-space at the uncertain variable means, with iterative expansion updates at each MPP estimate ($\mathbf{x}^*$) until the MPP converges. The first-order option is commonly known as AMV+:
$$g(\mathbf{x}) \cong g(\mathbf{x}^*) + \nabla_x g(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) \tag{26}$$
and the second-order option has been named AMV^2+:
$$g(\mathbf{x}) \cong g(\mathbf{x}^*) + \nabla_{\mathbf{x}} g(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \nabla^2_{\mathbf{x}} g(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \tag{27}$$

4. same as AMV+/AMV^2+, except that the expansions are performed in u-space. The first-order option has been termed the u-space AMV+ method.
$$G(\mathbf{u}) \cong G(\mathbf{u}^*) + \nabla_u G(\mathbf{u}^*)^T (\mathbf{u} - \mathbf{u}^*) \tag{28}$$
and the second-order option has been named the u-space AMV^2+ method:
$$G(\mathbf{u}) \cong G(\mathbf{u}^*) + \nabla_{\mathbf{u}} G(\mathbf{u}^*)^T (\mathbf{u} - \mathbf{u}^*) + \frac{1}{2}(\mathbf{u} - \mathbf{u}^*)^T \nabla^2_{\mathbf{u}} G(\mathbf{u}^*)(\mathbf{u} - \mathbf{u}^*) \tag{29}$$

5. a multipoint approximation in x-space. This approach involves a Taylor series approximation in intermediate variables where the powers used for the intermediate variables are selected to match information at the current and previous expansion points. Based on the two-point exponential approximation concept (TPEA, (Fadel et al., 1990)), the two-point adaptive nonlinearity approximation (TANA-3, (Xu and Grandhi, 1998)) approximates the limit state as:

$$g(\mathbf{x}) \cong g(\mathbf{x}_2) + \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(\mathbf{x}_2) \frac{x_{i,2}^{1-p_i}}{p_i}(x_i^{p_i} - x_{i,2}^{p_i}) + \frac{1}{2}\epsilon(\mathbf{x}) \sum_{i=1}^{n}(x_i^{p_i} - x_{i,2}^{p_i})^2 \tag{30}$$

where n is the number of uncertain variables and:

$$p_i = 1 + \ln\left[\frac{\frac{\partial g}{\partial x_i}(\mathbf{x}_1)}{\frac{\partial g}{\partial x_i}(\mathbf{x}_2)}\right] \Bigg/ \ln\left[\frac{x_{i,1}}{x_{i,2}}\right] \tag{31}$$

$$\epsilon(\mathbf{x}) = \frac{H}{\sum_{i=1}^{n}(x_i^{p_i} - x_{i,1}^{p_i})^2 + \sum_{i=1}^{n}(x_i^{p_i} - x_{i,2}^{p_i})^2} \tag{32}$$

$$H = 2\left[g(\mathbf{x}_1) - g(\mathbf{x}_2) - \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(\mathbf{x}_2) \frac{x_{i,2}^{1-p_i}}{p_i}(x_{i,1}^{p_i} - x_{i,2}^{p_i})\right] \tag{33}$$

and $\mathbf{x}_2$ and $\mathbf{x}_1$ are the current and previous MPP estimates in x-space, respectively. Prior to the availability of two MPP estimates, x-space AMV+ is used.

6. a multipoint approximation in u-space. The u-space TANA-3 approximates the limit state as:

$$G(\mathbf{u}) \cong G(\mathbf{u}_2) + \sum_{i=1}^{n} \frac{\partial G}{\partial u_i}(\mathbf{u}_2) \frac{u_{i,2}^{1-p_i}}{p_i}(u_i^{p_i} - u_{i,2}^{p_i}) + \frac{1}{2}\epsilon(\mathbf{u}) \sum_{i=1}^{n}(u_i^{p_i} - u_{i,2}^{p_i})^2 \tag{34}$$

where:

$$p_i = 1 + \ln\left[\frac{\frac{\partial G}{\partial u_i}(\mathbf{u}_1)}{\frac{\partial G}{\partial u_i}(\mathbf{u}_2)}\right] \Bigg/ \ln\left[\frac{u_{i,1}}{u_{i,2}}\right] \tag{35}$$

$$\epsilon(\mathbf{u}) = \frac{H}{\sum_{i=1}^{n}(u_i^{p_i} - u_{i,1}^{p_i})^2 + \sum_{i=1}^{n}(u_i^{p_i} - u_{i,2}^{p_i})^2} \tag{36}$$

$$H = 2\left[G(\mathbf{u}_1) - G(\mathbf{u}_2) - \sum_{i=1}^{n} \frac{\partial G}{\partial u_i}(\mathbf{u}_2) \frac{u_{i,2}^{1-p_i}}{p_i}(u_{i,1}^{p_i} - u_{i,2}^{p_i})\right] \tag{37}$$

and $\mathbf{u}_2$ and $\mathbf{u}_1$ are the current and previous MPP estimates in u-space, respectively. Prior to the availability of two MPP estimates, u-space AMV+ is used.

7. the MPP search on the original response functions without the use of any approximations.

The Hessian matrices in AMV^2 and AMV^2+ may be available analytically, estimated numerically, or approximated through quasi-Newton updates. The quasi-Newton variant of AMV^2+ is conceptually similar to TANA in that both approximate curvature based on a sequence of gradient evaluations. TANA estimates curvature by matching values and gradients at two points and includes it through the use of exponential intermediate variables and a single-valued diagonal Hessian approximation. Quasi-Newton AMV^2+ accumulates curvature over a sequence of points and then uses it directly in a second-order series expansion. Therefore, these methods may be expected to exhibit similar performance.

The selection between x-space or u-space for performing approximations depends on where the approximation will be more accurate, since this will result in more accurate MPP estimates (AMV, AMV^2) or faster convergence (AMV+, AMV^2+, TANA). Since this relative accuracy depends on the forms of the limit state $g(x)$ and the transformation $T(x)$ and is therefore application dependent in general, DAKOTA/UQ supports both options. A concern with approximation-based iterative search methods (i.e., AMV+, AMV^2+ and TANA) is the robustness of their convergence to the MPP. It is possible for the MPP iterates to oscillate or even diverge. However, to date, this occurrence has been relatively rare, and DAKOTA/UQ contains checks that monitor for this behavior. Another concern with TANA is numerical safeguarding. First, there is the possibility of raising negative x_i or u_i values to nonintegral p_i exponents in Eqs. 30, 32-34, and 36-37. This is particularly likely for u-space. Safeguarding techniques include the use of linear bounds scaling for each x_i or u_i, offseting negative x_i or u_i, or promotion of p_i to integral values for negative x_i or u_i. In numerical experimentation, the offset approach has been the most effective in retaining the desired data matches without overly inflating the p_i exponents. Second, there are a number of potential numerical difficulties with the logarithm ratios in Eqs. 31 and 35. In this case, a safeguarding strategy is to revert to either the linear ($p_i = 1$) or reciprocal ($p_i = -1$) approximation based on which approximation has lower error in $\frac{\partial g}{\partial x_i}(\mathbf{x}_1)$ or $\frac{\partial G}{\partial u_i}(\mathbf{u}_1)$.

2.2.2. *Probability integrations*

The second algorithmic variation involves the integration approach for computing probabilities at the MPP, which can be selected to be first-order (Eqs. 8-9) or second-order integration. Second-order integration involves applying a curvature correction (Breitung, 1984; Hohenbichler and Rackwitz, 1988; Hong, 1999). Breitung applies a correction based on asymptotic analysis (Breitung, 1984):

$$p = \Phi(-\beta_p) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 + \beta_p \kappa_i}} \tag{38}$$

where κ_i are the principal curvatures of the limit state function (the eigenvalues of an orthonormal transformation of $\nabla^2_{\mathbf{u}} G$, taken positive for a convex limit state) and $\beta_p \geq 0$ (select CDF or CCDF probability correction to obtain correct sign for β_p). An alternate correction in (Hohenbichler and Rackwitz, 1988) is consistent in the asymptotic regime ($\beta_p \to \infty$) but does not collapse to first-order integration for $\beta_p = 0$:

$$p = \Phi(-\beta_p) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 + \psi(-\beta_p) \kappa_i}} \tag{39}$$

where $\psi() = \frac{\phi()}{\Phi()}$ and $\phi()$ is the standard normal density function. (Hong, 1999) applies further corrections to Eq. 39 based on point concentration methods.

To invert a second-order integration and compute β_p given p and κ_i (e.g., for second-order PMA as described in Section 2.2), Newton's method can be applied as described in (Eldred et al., 2006). Combining the no-approximation option of the MPP search with first-order and second-order integration approaches results in the traditional first-order and second-order reliability methods (FORM and SORM). Additional probability integration approaches can involve importance sampling in the vicinity of the MPP (Hohenbichler and Rackwitz, 1988; Wu, 1994), but are outside the scope of this paper. While second-order integrations could be performed anywhere a limit state Hessian has been computed, the additional computational effort is most warranted for fully converged MPPs from AMV+, AMV^2+, TANA, FORM, and SORM, and is of reduced value for MVFOSM, MVSOSM, AMV, or AMV^2.

2.2.3. *Hessian approximations*

To use a second-order Taylor series or a second-order integration when second-order information ($\nabla^2_{\mathbf{x}} g$, $\nabla^2_{\mathbf{u}} G$, and/or κ) is not directly available, one can estimate the missing information using finite differences or approximate it through use of quasi-Newton approximations. These procedures will often be needed to make second-order approaches practical for engineering applications.

In the finite difference case, numerical Hessians are commonly computed using either first-order forward differences of gradients using

$$\nabla^2 g(\mathbf{x}) \cong \frac{\nabla g(\mathbf{x} + h\mathbf{e}_i) - \nabla g(\mathbf{x})}{h} \tag{40}$$

to estimate the i^{th} Hessian column when gradients are analytically available, or second-order differences of function values using

$$\nabla^2 g(\mathbf{x}) \cong \frac{g(\mathbf{x}+h\mathbf{e}_i+h\mathbf{e}_j) - g(\mathbf{x}+h\mathbf{e}_i-h\mathbf{e}_j) - g(\mathbf{x}-h\mathbf{e}_i+h\mathbf{e}_j) + g(\mathbf{x}-h\mathbf{e}_i-h\mathbf{e}_j)}{4h^2} \tag{41}$$

to estimate the ij^{th} Hessian term when gradients are not directly available. This approach has the advantage of locally-accurate Hessians for each point of interest (which can lead to quadratic convergence rates in discrete Newton methods), but has the disadvantage that numerically estimating each of the matrix terms can be expensive.

Quasi-Newton approximations, on the other hand, do not reevaluate all of the second-order information for every point of interest. Rather, they accumulate approximate curvature information over time using secant updates. Since they utilize the existing gradient evaluations, they do not require any additional function evaluations for evaluating the Hessian terms. The quasi-Newton approximations of interest include the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \frac{\mathbf{B}_k \mathbf{s}_k \mathbf{s}_k^T \mathbf{B}_k}{\mathbf{s}_k^T \mathbf{B}_k \mathbf{s}_k} + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} \tag{42}$$

which yields a sequence of symmetric positive definite Hessian approximations, and the Symmetric Rank 1 (SR1) update

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{(\mathbf{y}_k - \mathbf{B}_k \mathbf{s}_k)(\mathbf{y}_k - \mathbf{B}_k \mathbf{s}_k)^T}{(\mathbf{y}_k - \mathbf{B}_k \mathbf{s}_k)^T \mathbf{s}_k} \tag{43}$$

which yields a sequence of symmetric, potentially indefinite, Hessian approximations. $\mathbf{B}_k$ is the k^{th} approximation to the Hessian $\nabla^2 g$, $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ is the step and $\mathbf{y}_k = \nabla g_{k+1} - \nabla g_k$ is the corresponding yield in the gradients. The selection of BFGS versus SR1 involves the importance of retaining positive definiteness in the Hessian approximations; if the procedure does not require it, then the SR1 update can be more accurate if the true Hessian is not positive definite. Initial scalings for $\mathbf{B}_0$ and numerical safeguarding techniques (damped BFGS, update skipping) are described in (Eldred et al., 2006).

2.2.4. *Optimization algorithms*

The next algorithmic variation involves the optimization algorithm selection for solving Eqs. 16 and 17. The Hasofer-Lind Rackwitz-Fissler (HL-RF) algorithm (Haldar and Mahadevan, 2000) is a classical approach that has been broadly applied. It is a Newton-based approach lacking line search/trust region globalization, and is generally regarded as computationally efficient but occasionally unreliable. DAKOTA/UQ takes the approach of employing robust, general-purpose optimization algorithms with provable convergence properties. This paper employs the sequential quadratic programming (SQP) and nonlinear interior-point (NIP) optimization algorithms from the NPSOL (Gill et al., 1998) and OPT++ (Meza, 1994) libraries, respectively.

2.2.5. *Warm Starting of MPP Searches*

The final algorithmic variation involves the use of warm starting approaches for improving computational efficiency. (Eldred et al., 2005) describes the acceleration of MPP searches through warm starting with approximate iteration increment, with $z/p/\beta$ level increment, and with design variable increment. Warm started data includes the expansion point and associated response values and the MPP optimizer initial guess. Projections are used when an increment in $z/p/\beta$ level or design variables occurs. Warm starts were consistently effective in (Eldred et al., 2005), with greater effectiveness for smaller parameter changes, and are used for all computational experiments presented in this paper.

3. Reliability-Based Design Optimization

Reliability-based design optimization (RBDO) methods are used to perform design optimization accounting for reliability metrics. The reliability analysis capabilities described in Section 2 provide a rich foundation for exploring a variety of RBDO formulations. (Eldred et al., 2005) investigated bi-level, fully-analytic bi-level, and first-order sequential RBDO approaches employing underlying first-order reliability assessments. (Eldred et al., 2006) investigated fully-analytic bi-level and second-order sequential RBDO approaches employing underlying second-order reliability assessments. These methods are overviewed in the following sections.

3.1. Bi-level RBDO

The simplest and most direct RBDO approach is the bi-level approach in which a full reliability analysis is performed for every optimization function evaluation. This involves a nesting of two distinct levels of optimization within each other, one at the design level and one at the MPP search level.

Since an RBDO problem will typically specify both the $\bar{z}$ level and the $\bar{p}/\bar{\beta}$ level, one can use either the RIA or the PMA formulation for the UQ portion and then constrain the result in the design optimization portion. In particular, RIA reliability analysis maps $\bar{z}$ to p/β, so RIA RBDO constrains p/β:

$$\begin{aligned} \text{minimize} \quad & f \\ \text{subject to} \quad & \beta \geq \bar{\beta} \\ \text{or} \quad & p \leq \bar{p} \end{aligned} \tag{44}$$

And PMA reliability analysis maps $\bar{p}/\bar{\beta}$ to z, so PMA RBDO constrains z:

$$\begin{aligned} \text{minimize} \quad & f \\ \text{subject to} \quad & z \geq \bar{z} \end{aligned} \tag{45}$$

where $z \geq \bar{z}$ is used as the RBDO constraint for a cumulative failure probability (failure defined as $z \leq \bar{z}$) but $z \leq \bar{z}$ would be used as the RBDO constraint for a complementary cumulative failure probability (failure defined as $z \geq \bar{z}$). It is worth noting that DAKOTA is not limited to these types of inequality-constrained RBDO formulations; rather, they are convenient examples. DAKOTA supports general optimization under uncertainty mappings (Eldred et al., 2002) which allow flexible use of statistics within multiple objectives, inequality constraints, and equality constraints.

An important performance enhancement for bi-level methods is the use of sensitivity analysis to analytically compute the design gradients of probability, reliability, and response levels. When design variables are separate from the uncertain variables (i.e., they are not distribution parameters), then the following first-order expressions may be used (Hohenbichler and Rackwitz, 1986; Karamchandani and Cornell, 1992; Allen and Maute, 2004):

$$\nabla_{\mathbf{d}} z = \nabla_{\mathbf{d}} g \tag{46}$$

$$\nabla_{\mathbf{d}} \beta_{cdf} = \frac{1}{\parallel \nabla_{\mathbf{u}} G \parallel} \nabla_{\mathbf{d}} g \tag{47}$$

$$\nabla_{\mathbf{d}} p_{cdf} = -\phi(-\beta_{cdf}) \nabla_{\mathbf{d}} \beta_{cdf} \tag{48}$$

where it is evident from Eqs. 12-13 that $\nabla_{\mathbf{d}} \beta_{ccdf} = -\nabla_{\mathbf{d}} \beta_{cdf}$ and $\nabla_{\mathbf{d}} p_{ccdf} = -\nabla_{\mathbf{d}} p_{cdf}$. In the case of second-order integrations, Eq. 48 must be expanded to include the curvature correction. For Breitung's correction (Eq. 38),

$$\nabla_{\mathbf{d}} p_{cdf} = \left[\Phi(-\beta_p) \sum_{i=1}^{n-1} \left(\frac{-\kappa_i}{2(1+\beta_p \kappa_i)^{\frac{3}{2}}} \prod_{\substack{j=1 \\ j \neq i}}^{n-1} \frac{1}{\sqrt{1+\beta_p \kappa_j}} \right) - \phi(-\beta_p) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1+\beta_p \kappa_i}} \right] \nabla_{\mathbf{d}} \beta_{cdf} \tag{49}$$

where $\nabla_{\mathbf{d}} \kappa_i$ has been neglected and $\beta_p \geq 0$ (see Section 2.2.2). Other approaches assume the curvature correction is nearly independent of the design variables (Rackwitz, 2002), which is equivalent to neglecting the first term in Eq. 49.

To capture second-order probability estimates within an RIA RBDO formulation using well-behaved β constraints, a generalized reliability index can be introduced where, similar to Eq. 10,

$$\beta^*_{cdf} = -\Phi^{-1}(p_{cdf}) \tag{50}$$

for second-order p_{cdf}. This reliability index is no longer equivalent to the magnitude of $\mathbf{u}$, but rather is a convenience metric for capturing the effect of more accurate probability estimates. The corresponding generalized reliability index sensitivity, similar to Eq. 48, is

$$\nabla_{\mathbf{d}} \beta^*_{cdf} = -\frac{1}{\phi(-\beta^*_{cdf})} \nabla_{\mathbf{d}} p_{cdf} \tag{51}$$

where $\nabla_{\mathbf{d}} p_{cdf}$ is defined from Eq. 49. Even when $\nabla_{\mathbf{d}} g$ is estimated numerically, Eqs. 46-51 can be used to avoid numerical differencing across full reliability analyses.

When the design variables are distribution parameters of the uncertain variables, $\nabla_{\mathbf{d}} g$ is expanded with the chain rule and Eqs. 46 and 47 become

$$\nabla_{\mathbf{d}} z = \nabla_{\mathbf{d}} \mathbf{x} \nabla_{\mathbf{x}} g \tag{52}$$

$$\nabla_{\mathbf{d}} \beta_{cdf} = \frac{1}{\| \nabla_{\mathbf{u}} G \|} \nabla_{\mathbf{d}} \mathbf{x} \nabla_{\mathbf{x}} g \tag{53}$$

where the design Jacobian of the transformation ($\nabla_{\mathbf{d}} \mathbf{x}$) may be obtained analytically for uncorrelated $\mathbf{x}$ or semi-analytically for correlated $\mathbf{x}$ ($\nabla_{\mathbf{d}} \mathbf{L}$ is evaluated numerically) by differentiating Eqs. 14 and 15 with respect to the distribution parameters. Eqs. 48-51 remain the same as before. For this design variable case, all required information for the sensitivities is available from the MPP search.

Since Eqs. 46-53 are derived using the Karush-Kuhn-Tucker optimality conditions for a converged MPP, they are appropriate for RBDO using AMV+, AMV^2+, TANA, FORM, and SORM, but not for RBDO using MVFOSM, MVSOSM, AMV, or AMV^2.

3.2. Sequential/Surrogate-based RBDO

An alternative RBDO approach is the sequential approach, in which additional efficiency is sought through breaking the nested relationship of the MPP and design searches. The general concept is to iterate between optimization and uncertainty quantification, updating the optimization goals based on the most recent probabilistic assessment results. This update may be based on safety factors (Wu et al., 2001) or other approximations (Du and Chen, 2004).

A particularly effective approach for updating the optimization goals is to use the $p/\beta/z$ sensitivity analysis of Eqs. 46-53 in combination with local surrogate models (Zou et al., 2004). In (Eldred et al., 2005) and (Eldred et al., 2006), first-order and second-order Taylor series approximations were employed within a trust-region model management framework (Giunta and Eldred, 2000) in order to adaptively manage the extent of the approximations and ensure convergence of the RBDO process. Surrogate models were used for both the objective function and the constraints, although the use of constraint surrogates alone is sufficient to remove the nesting.

In particular, RIA trust-region surrogate-based RBDO employs surrogate models of f and p/β within a trust region Δ^k centered at $\mathbf{d}_c$. For first-order surrogates:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{d}_c) + \nabla_d f(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) \\ \text{subject to} \quad & \beta(\mathbf{d}_c) + \nabla_d \beta(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) \geq \bar{\beta} \\ \text{or} \quad & p(\mathbf{d}_c) + \nabla_d p(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) \leq \bar{p} \\ & \parallel \mathbf{d} - \mathbf{d}_c \parallel_\infty \leq \Delta^k \end{aligned} \tag{54}$$

and for second-order surrogates:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{d}_c) + \nabla_{\mathbf{d}} f(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) + \frac{1}{2}(\mathbf{d} - \mathbf{d}_c)^T \nabla^2_{\mathbf{d}} f(\mathbf{d}_c)(\mathbf{d} - \mathbf{d}_c) \\ \text{subject to} \quad & \beta(\mathbf{d}_c) + \nabla_{\mathbf{d}} \beta(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) + \frac{1}{2}(\mathbf{d} - \mathbf{d}_c)^T \nabla^2_{\mathbf{d}} \beta(\mathbf{d}_c)(\mathbf{d} - \mathbf{d}_c) \geq \bar{\beta} \\ \text{or} \quad & p(\mathbf{d}_c) + \nabla_{\mathbf{d}} p(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) + \frac{1}{2}(\mathbf{d} - \mathbf{d}_c)^T \nabla^2_{\mathbf{d}} p(\mathbf{d}_c)(\mathbf{d} - \mathbf{d}_c) \leq \bar{p} \\ & \parallel \mathbf{d} - \mathbf{d}_c \parallel_\infty \leq \Delta^k \end{aligned} \tag{55}$$

For PMA trust-region surrogate-based RBDO, surrogate models of f and z are employed within a trust region Δ^k centered at $\mathbf{d}_c$. For first-order surrogates:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{d}_c) + \nabla_d f(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) \\ \text{subject to} \quad & z + \nabla_d z(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) \geq \bar{z} \\ & \parallel \mathbf{d} - \mathbf{d}_c \parallel_\infty \leq \Delta^k \end{aligned} \tag{56}$$

and for second-order surrogates:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{d}_c) + \nabla_{\mathbf{d}} f(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) + \frac{1}{2}(\mathbf{d} - \mathbf{d}_c)^T \nabla^2_{\mathbf{d}} f(\mathbf{d}_c)(\mathbf{d} - \mathbf{d}_c) \\ \text{subject to} \quad & z + \nabla_{\mathbf{d}} z(\mathbf{d}_c)^T(\mathbf{d} - \mathbf{d}_c) + \frac{1}{2}(\mathbf{d} - \mathbf{d}_c)^T \nabla^2_{\mathbf{d}} z(\mathbf{d}_c)(\mathbf{d} - \mathbf{d}_c) \geq \bar{z} \\ & \parallel \mathbf{d} - \mathbf{d}_c \parallel_\infty \leq \Delta^k \end{aligned} \tag{57}$$

where the sense of the z constraint may vary as described previously. The second-order information in Eqs. 55 and 57 will typically be approximated with quasi-Newton updates.

4. Benchmark Problems

(Eldred et al., 2005) and (Eldred et al., 2006) have examined the performance of first and second-order reliability analysis and design methods for four analytic benchmark test problems: lognormal ratio, short column, cantilever beam, and steel column.

4.1. Reliability analysis results

Within the reliability analysis algorithms, various limit state approximation (MVFOSM, MVSOSM, x-/u-space AMV, x-/u-space AMV^2, x-/u-space AMV+, x-/u-space AMV^2+, x-/u-space TANA, FORM, and SORM), probability integration (first-order or second-order), warm starting, Hessian

Table I. RIA results for short column problem.

RIA Approach	SQP Function Evaluations	NIP Function Evaluations	CDF p Error Norm	Target z Offset Norm
MVFOSM	1	1	0.1548	0.0
MVSOSM	1	1	0.1127	0.0
x-space AMV	45	45	0.009275	18.28
u-space AMV	45	45	0.006408	18.81
x-space AMV^2	45	45	0.002063	2.482
u-space AMV^2	45	45	0.001410	2.031
x-space AMV+	192	192	0.0	0.0
u-space AMV+	207	207	0.0	0.0
x-space AMV^2+	125	131	0.0	0.0
u-space AMV^2+	122	130	0.0	0.0
x-space TANA	245	246	0.0	0.0
u-space TANA	296*	278*	6.982e-5	0.08014
FORM	626	176	0.0	0.0
SORM	669	219	0.0	0.0

approximation (finite difference, BFGS, or SR1), and MPP optimization algorithm (SQP or NIP) selections have been investigated. A sample comparison of reliability analysis performance, taken from the short column example, is shown in Tables I and II for RIA and PMA analysis, respectively, where "*" indicates that one or more levels failed to converge. Consistent with the employed probability integrations, the error norms are measured with respect to fully-converged first-order results for MV, AMV, AMV^2, AMV+, and FORM methods, and with respect to fully-converged second-order results for AMV^2+, TANA, and SORM methods. Also, it is important to note that the simple metric of "function evaluations" is imperfect, and (Eldred et al., 2006) provides more detailed reporting of individual response value, gradient, and Hessian evaluations.

Overall, reliability analysis results for the lognormal ratio, short column, and cantilever test problems indicate several trends. MVFOSM, MVSOSM, AMV, and AMV^2 are significantly less expensive than the fully-converged MPP methods, but come with corresponding reductions in accuracy. In combination, these methods provide a useful spectrum of accuracy and expense that allow the computational effort to be balanced with the statistical precision required for particular applications. In addition, support for forward and inverse mappings (RIA and PMA) provide the flexibility to support different UQ analysis needs.

Relative to FORM and SORM, AMV+ and AMV^2+ has been shown to have equal accuracy and consistent computational savings. For second-order PMA analysis with prescribed probability levels, AMV^2+ has additionally been shown to be more robust due to its ability to better manage $\bar{\beta}$ udpates. Analytic Hessians were highly effective in AMV^2+, but since they are often unavailable in practical applications, finite-difference numerical Hessians and quasi-Newton Hessian approximations were also demonstrated, with SR1 quasi-Newton updates being shown to be sufficiently

Table II. PMA results for short column problem.

PMA Approach	SQP Function Evaluations	NIP Function Evaluations	CDF z Error Norm	Target p Offset Norm
MVFOSM	1	1	7.454	0.0
MVSOSM	1	1	6.823	0.0
x-space AMV	45	45	0.9420	0.0
u-space AMV	45	45	0.5828	0.0
x-space AMV^2	45	45	2.730	0.0
u-space AMV^2	45	45	2.828	0.0
x-space AMV+	171	179	0.0	0.0
u-space AMV+	205	205	0.0	0.0
x-space AMV^2+	135	142	0.0	0.0
u-space AMV^2+	132	139	0.0	0.0
x-space TANA	293*	272	0.04259	1.598e-4
u-space TANA	325*	311*	2.208	5.600e-4
FORM	720	192	0.0	0.0
SORM	535	191*	2.410	6.522e-4

accurate and competitive with analytic Hessian performance. Relative to first-order AMV+ performance, AMV^2+ with analytic Hessians had consistently superior efficiency, and AMV^2+ with quasi-Newton Hessians had improved performance in most cases (it was more expensive than AMV+ only when a more challenging second-order $\bar{p}$ problem was being solved). In general, second-order reliability analyses appear to serve multiple synergistic needs. The same Hessian information that allows for more accurate probability integrations can also be applied to making MPP solutions more efficient and more robust. Conversely, limit state curvature information accumulated during an MPP search can be reused to improve the accuracy of probability estimates.

For nonapproximated limit states (FORM and SORM), NIP optimizers have shown promise in being less susceptible to PMA u-space excursions and in being more efficient than SQP optimizers in most cases. Warm starting with projections has been shown to be consistently effective for reliability analyses, with typical savings on the order of 25%. The x-space and u-space linearizations for AMV, AMV^2, AMV+, AMV^2+, and TANA were both effective, and the relative performance was strongly problem-dependent (u-space was more efficient for lognormal ratio, x-space was more efficient for short column, and x-space and u-space were equivalent for cantilever). Among all combinations tested, AMV^2+ (with analytic Hessians if available, or SR1 Hessians if not) is the recommended approach.

An important question is how Taylor-series based limit state approximations (such as AMV+ and AMV^2+) can frequently outperform the best general-purpose optimizers (such as SQP and NIP). The answer likely lies in the exploitation of the structure of the RIA and PMA MPP problems. By approximating the limit state but retaining $\mathbf{u}^T\mathbf{u}$ explicitly in Eqs. 16 and 17, specific problem structure knowledge is utilized in formulating a mixed surrogate/direct approach.

Table III. Analytic bi-level RBDO results, short column test problem.

RBDO Approach	Function Evaluations	Objective Function	Constraint Violation
RIA $\bar{z} \to p$ x-space AMV+	149	217.1	0.0
RIA $\bar{z} \to p$ x-space AMV^2+	129	217.1	0.0
RIA $\bar{z} \to p$ FORM	911	217.1	0.0
RIA $\bar{z} \to p$ SORM	1204	217.1	0.0
RIA $\bar{z} \to \beta$ x-space AMV+	72	216.7	0.0
RIA $\bar{z} \to \beta$ x-space AMV^2+	67	216.7	0.0
RIA $\bar{z} \to \beta$ FORM	612	216.7	0.0
RIA $\bar{z} \to \beta$ SORM	601	216.7	0.0
PMA $\bar{p}, \bar{\beta} \to z$ x-space AMV+	100	216.8	0.0
PMA $\bar{p} \to z$ x-space AMV^2+	98	216.8	0.0
PMA $\bar{\beta} \to z$ x-space AMV^2+	98	216.8	0.0
PMA $\bar{p}, \bar{\beta} \to z$ FORM	285	216.8	0.0
PMA $\bar{p} \to z$ SORM	306	217.2	0.0
PMA $\bar{\beta} \to z$ SORM	329	216.8	0.0

4.2. RBDO results

These reliability analysis capabilities provide a substantial foundation for RBDO formulations, and bi-level and sequential RBDO approaches have been investigated. Both approaches have utilized analytic gradients for z, β, and p with respect to augmented and inserted design variables, and sequential RBDO has additionally utilized a trust-region surrogate-based approach to manage the extent of the Taylor-series approximations. A sample comparison of RBDO performance, taken again from the short column example, is shown in Tables III and IV for bi-level and sequential surogate-based RBDO, respectively.

Overall, RBDO results for the short column, cantilever, and steel column test problems build on the reliability analysis trends. Basic first-order bi-level RBDO has been evaluated with up to 18 variants (RIA/PMA with different $p/\beta/z$ mappings for MV, x-/u-space AMV, x-/u-space AMV+, and FORM), and fully-analytic bi-level and sequential RBDO have each been evaluated with up to 21 variants (RIA/PMA with different $p/\beta/z$ mappings for x-/u-space AMV+, x-/u-space AMV^2+, FORM, and SORM). Bi-level RBDO with MV and AMV are inexpensive but give only approximate optima. These approaches may be useful for preliminary design or for warm-starting other RBDO methods. Bi-level RBDO with AMV+ was shown to have equal accuracy and robustness to bi-level FORM-based approaches and be significantly less expensive on average. In addition, usage of β in RIA RBDO constraints was preferred due to it being more well-behaved and more well-scaled than constraints on p. Warm starts in RBDO were most effective when the design changes were small, with the most benefit for basic bi-level RBDO (with numerical differencing at the design level), decreasing to marginal effectiveness for fully-analytic bi-level RBDO and to relative ineffectiveness

Table IV. Surrogate-based RBDO results, short column test problem.

RBDO Approach	Function Evaluations	Objective Function	Constraint Violation
RIA $\bar{z} \rightarrow p$ x-space AMV+	75	216.9	0.0
RIA $\bar{z} \rightarrow p$ x-space AMV^2+	86	218.7	0.0
RIA $\bar{z} \rightarrow p$ FORM	577	216.9	0.0
RIA $\bar{z} \rightarrow p$ SORM	718	216.5	1.110e-4
RIA $\bar{z} \rightarrow \beta$ x-space AMV+	65	216.7	0.0
RIA $\bar{z} \rightarrow \beta$ x-space AMV^2+	51	216.7	0.0
RIA $\bar{z} \rightarrow \beta$ FORM	561	216.7	0.0
RIA $\bar{z} \rightarrow \beta$ SORM	560	216.7	0.0
PMA $\bar{p}, \bar{\beta} \rightarrow z$ x-space AMV+	76	216.7	2.1e-4
PMA $\bar{p} \rightarrow z$ x-space AMV^2+	58	216.8	0.0
PMA $\bar{\beta} \rightarrow z$ x-space AMV^2+	79	216.8	0.0
PMA $\bar{p}, \bar{\beta} \rightarrow z$ FORM	228	216.7	2.1e-4
PMA $\bar{p} \rightarrow z$ SORM	128	217.2	0.0
PMA $\bar{\beta} \rightarrow z$ SORM	171	216.8	0.0

for sequential RBDO. However, large design changes were desirable for overall RBDO efficiency and, compared to basic bi-level RBDO, fully-analytic RBDO and sequential RBDO were clearly superior.

In second-order bi-level and sequential RBDO, the AMV^2+ approaches were consistently more efficient than the SORM-based approaches. In general, sequential RBDO approaches demonstrated consistent computational savings over the corresponding bi-level RBDO approaches, and the combination of sequential RBDO using AMV^2+ was the most effective of all of the approaches. With initial trust region size tuning, sequential RBDO computational expense for these test problems was shown to be as low as approximately 40 function evaluations per limit state (35 for a single limit state in short column, 75 for two limit states in cantilever, and 45 for a single limit state in steel column). Finally, second-order RBDO with probability constraints was shown to be more challenging and expensive, but could be more precise in achieving the desired probabilistic performance.

5. Application to MEMS

In this section, we consider the application of DAKOTA's reliability algorithms to the design of micro-electro-mechanical systems (MEMS). In particular, we summarize initial results for one of the applications described in (Adams et al., 2006). These application studies provide essential feedback on the performance of algorithms for real-world design applications, which may contain computational challenges not well-represented in analytically defined test problems.

Pre-fabrication design optimization of microelectromechanical systems (MEMS) is an important emerging application of uncertainty quantification and reliability-based design optimization. Typically crafted of silicon, polymers, metals, or a combination thereof, MEMS serve as micro-scale sensors, actuators, switches, and machines with applications including robotics, biology and medicine, automobiles, RF electronics, and optical displays (Allen, 2005). Design optimization of these devices is crucial since fabrication costs, even for prototypes, can be prohibitive. There is considerable uncertainty in the micromachining and etching processes used to manufacture MEMS and consequently in the behavior of the finished products. RBDO coupled with computational mechanics models of MEMS offers a means to quantify this uncertainty and determine a priori the most reliable and/or robust design that meets performance criteria.

Of particular interest is the design of MEMS bistable mechanisms which toggle between two stable positions, making them useful as micro switches, relays, and nonvolatile memory. We focus on shape optimization of compliant bistable mechanisms, where instead of mechanical joints, material elasticity enables the bistability of the mechanism (Jensen et al., 2001). Figure 1 contains an electron micrograph of a MEMS compliant bistable mechanism in one of its stable positions. One achieves transfer between stable states by applying force to the center shuttle of the device via an electrostatic actuator, heat source, or other means to cause the flexible "legs" (horizontal beams) of the system to buckle through their instability and relax toward the other stable equilibrium.

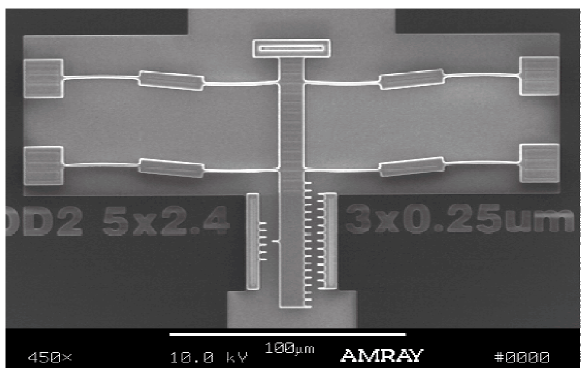


Figure 1. Electron micrograph of MEMS bistable mechanism. Source: J.W. Wittwer, Ph.D. dissertation.

Successful bistable switch actuation in this manner depends on the relationship between force and vertical displacement for the manufactured switch. In Figure 2 we present a schematic of a typical force–displacement curve for a bistable mechanism. The switch characterized by this curve has three equilibria: E_1 and E_3 are stable equilibria whereas E_2 is an unstable equilibrium (arrows indicate stability). A device with such a force–displacement curve could be used as a switch or actuator by setting it to position E_3 as shown in Figure 1 (requiring large force F_{max}) and then actuating by applying the small force F_{min} in the opposite direction to transfer through E_2 toward the equilibrium E_1. One could utilize this force profile to complete a circuit by placing a switch contact near the displaced position corresponding to maximum (closure) force as illustrated in Figure 2.

The design considered in this work is similar to the electron micrograph in Figure 1, for which design optimization has been considered in (Jensen et al., 2001) and design under uncertainty

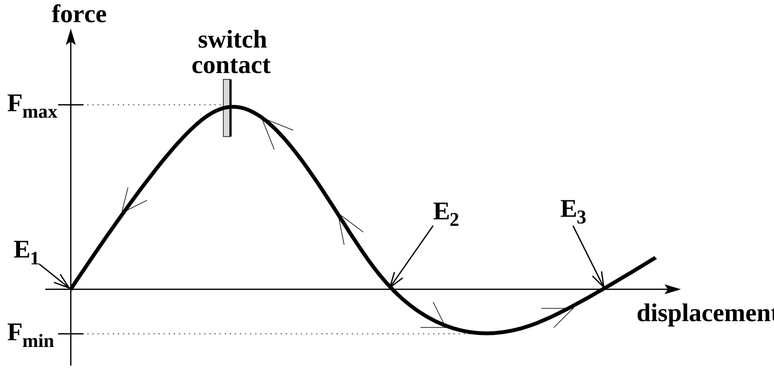


Figure 2. Schematic of force–displacement curve for bistable MEMS mechanism. Arrows indicate stability.

with mean value methods has been investigated in (Wittwer, 2005; Wittwer, 2006). The primary structural difference in (Adams et al., 2006) is in the shape of the legs, and Figure 3 shows a detail of the design of one of these legs.

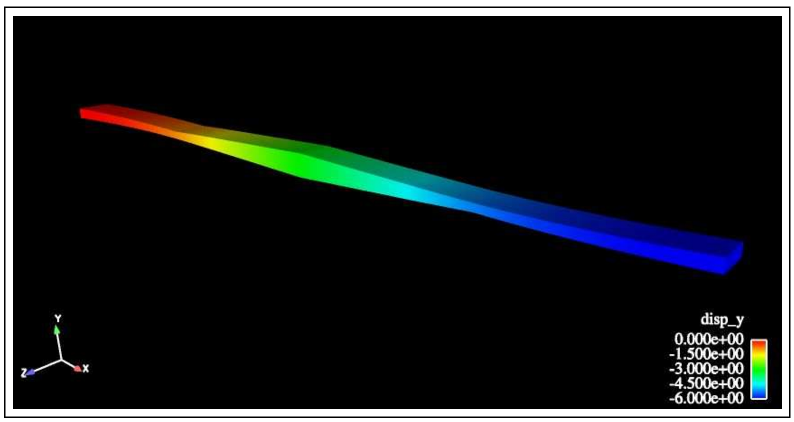


Figure 3. Sample of tapered beam leg for bistable mechanism.

The design criteria used for this bistable switch include

- minimize the magnitude of the force F_{min} required to actuate the switch (drive F_{min} toward zero), while maintaining its bistability ($F_{max} > 0$, $F_{min} < 0$)
- at least $50\mu N$ force at switch contact (to reliably attain closure), but no more than $150\mu N$ (to avoid contact damage)
- point of instability E_2 no more than $8\mu m$
- maximum stress no more than 1200 MPa

The force-displacement profile of bistable MEMS devices is highly sensitive to design geometry, so one can vary manufactured geometry in order to achieve various design criteria. However, due to manufacturing processes, fabricated geometry can deviate significantly from design-specified beam geometry. As a consequence of photo masks used in the process, fabricated in-plane geometry edges (contributing to widths and lengths) are $0.1 \pm 0.08\,\mu m$ less than specified. Uncertainty in the manufactured geometry can lead to substantial uncertainty in the positions of the stable equilibria and in the maximum and minimum force on the force–displacement curve. The manufactured thickness of the device is also uncertain, though this does not contribute as much to variability in the force–displacement behavior. Uncertain material properties such as Young's modulus and residual stress also influence the characteristics of the fabricated beam. For this application, we consider two uncertain variables: ΔW (edge bias on beam widths, which yields effective manufactured widths of $W_i + \Delta W, i = 0, \ldots, n$) and S_r (residual stress in the manufactured device), with distributions shown in Table V.

Table V. Uncertain variables used in RBDO.

variable	mean	std. dev.	distribution
Δw	-0.2 μm	0.08	normal
S_r	-11 Mpa	4.13	normal

Given 13 geometric design variables $\mathbf{d}$ describing lengths, widths, and orientations of the legs and the two specified uncertain variables $\mathbf{x}$, we perform a reliability-based design optimization to compute a design that is reliably bistable, but requires minimum force to actuate. The limit state for this problem is

$$g(\mathbf{x}) = F_{min}(\mathbf{x}) \tag{58}$$

and we define failure to be lack of bistability ($F_{min} \geq 0$) and require a reliability index $\beta_{ccdf} \geq 2$. The RBDO problem utilizes an RIA $\bar{z} \rightarrow \beta$ approach:

$$\begin{array}{lrcl} \max & & & F_{min}(\mathbf{d}, \mathbf{x}) \\ \text{s.t.} & 2 & \leq & \beta_{ccdf}(\mathbf{d}, \mathbf{x}) \\ & 50 & \leq & F_{max}(\mathbf{d}, \mathbf{x}) \ \leq \ 150 \\ & & & E_2(\mathbf{d}, \mathbf{x}) \ \leq \ 8 \\ & & & S_{max}(\mathbf{d}, \mathbf{x}) \ \leq \ 1200 \end{array} \tag{59}$$

although a PMA $\bar{\beta} \rightarrow z$ approach could also be used. The use of the F_{min} metric in both the objective function and the reliability constraint results in a powerful problem formulation since, in addition to yielding a design with specified reliability, it also produces a robust design. By forcing F_{min} toward zero while requiring two standard deviations of surety, the optimization problem favors designs with less variability in F_{min}. This renders the design performance less sensitive to the uncertainties in the problem.

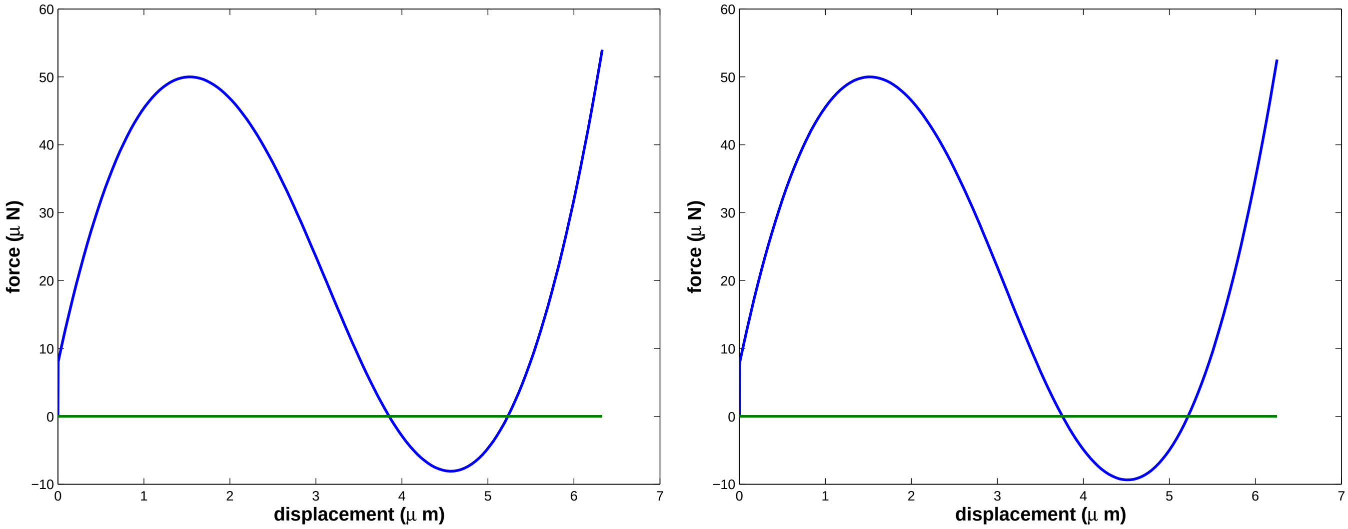
We solve the optimization problem by applying DAKOTA's bi-level RBDO approach in combination with mesh generation using CUBIT and finite element analysis using Adagio. Adagio is a quasi-static nonlinear mechanics code, implemented in Sandia National Laboratories' SIERRA framework of multiphysics codes (Edwards, 2004), that is used to simulate the elastic deformation of the device through discrete displacement steps to produce a force–displacement curve. We compare three reliability analysis methods for this MEMS application: (1) MVFOSM (no MPP search), (2) AMV+, and (3) FORM. The latter two are advantaged by their ability to provide (semi)analytic derivatives of reliability metrics with respect to design variables for the optimizer (see Section 3.1), whereas the former is much less expensive per reliability analysis but must resort to numerical design derivatives due to the use of σ_g (analytic derivatives of Eq. 2 with respect to $\mathbf{d}$ are impractical to evaluate).

Results for the three methods are presented in Table VI and the optimal force–displacement curves are shown in Figure 4. Optimization with MVFOSM offers substantial improvement over the initial design, yielding a design with a substantially smaller minimum force and tighter reliability constraint β. However, since mean value analyses estimate reliability based solely on evaluations at the means of the uncertain variables, they can yield inaccurate reliability metrics in cases of nonlinearity or nonnormality. In this example, the actual verified reliability of the optimal MVFOSM-based design is only 1.75, less than the prescribed reliability of $\beta = 2$. The optimal designs for the AMV+ and FORM-based RBDO methods were indistinguishable from each other, but relative to MVFOSM-based RBDO, yield a more conservative value of F_{min} due to the improved estimation of β. In each of the three cases, the variability in F_{min} has been reduced from approximately 5.7 to 4.6 μN per (verified) input standard deviation, resulting in designs that are less sensitive to the input uncertainties.

Table VI. RBDO results (MVFOSM and first-order MPP methods) for MEMS bistable mechanism.

lower bound	RBDO metric	upper bound	MVFOSM initial	MVFOSM optimal	AMV+/FORM initial	AMV+/FORM optimal
	F_{min} (μN)		-23.03	-8.08	-23.03	-9.37
2	β		5.66	2.00	4.02	2.00
50	F_{max} (μN)	150	67.35	50.0	67.35	50.0
	E_2 (μm)	8	4.06	3.85	4.06	3.76
	S_{max} (MPa)	1200	396	313	396	323
	Verified β		4.02	1.75		

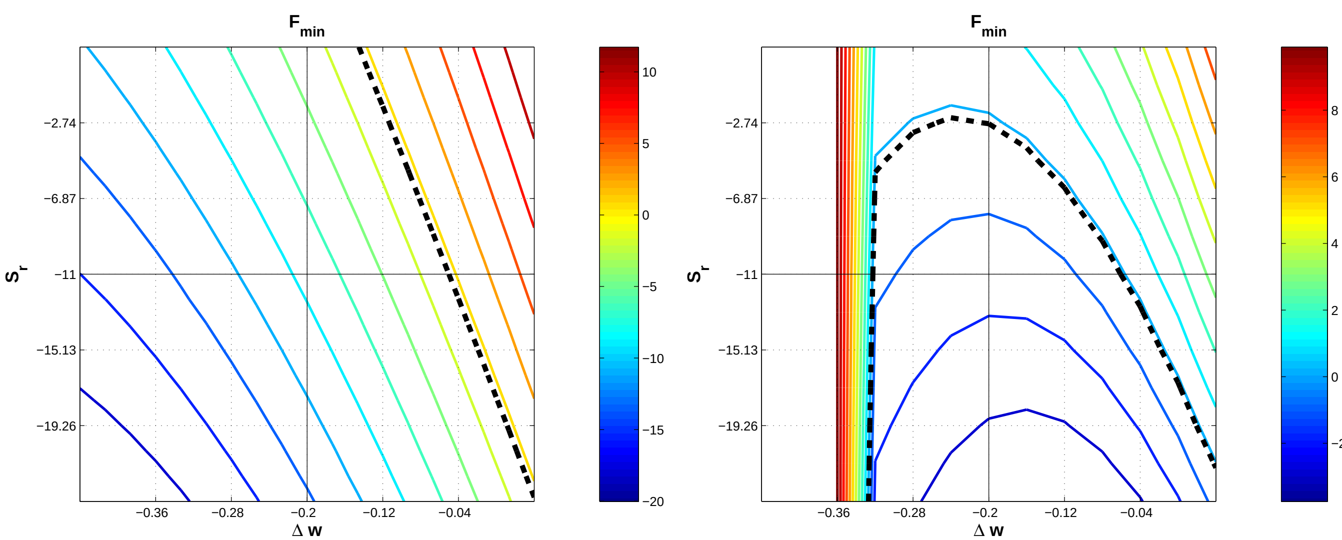
In Figure 5,we see the results of parameter studies for the metric $F_{min}(\mathbf{d}, \mathbf{x})$ as a function of the uncertain variables $\mathbf{x}$ for two different sets of design variables $\mathbf{d}$. Since the uncertain variables are both normal, the transformation to u-space used by AMV+ and FORM is linear. The former design variable set corresponds to the optimal values obtained from MVFOSM-based RBDO, and in this case the limit state is relatively linear and well-behaved in the range of interest. First-order probability integrations should be sufficiently accurate. For the second design variable set, however, multiple computational challenges are evident. In this case, the limit state has significant nonlinearity (requiring more sophisticated probability integrations) and its simulation can be seen to be unreliable in the left tail of the edge bias (resulting from too flimsy a structure). This highlights a number of difficulties common in engineering applications: highly nonlinear limit states, nonsmooth and multimodal limit states, and simulation failures caused by, e.g., evaluations in the tails of input



(a) MVFOSM (b) AMV+/FORM

Figure 4. Optimal force-displacement curves resulting from RBDO of MEMS bistable mechanism.

distributions. These difficulties must be mitigated through a combination of algorithm research, problem formulation, and simulation refinement.



(a) Optimal MVFOSM $\mathbf{d}$ (b) Problematic $\mathbf{d}$

Figure 5. Contour plot of $F_{min}(\mathbf{d}, \mathbf{x})$ as a function of uncertain variables for different design variable sets. Dashed line shows where limit state $F_{min} = 0$.

6. Conclusions

This paper has overviewed recent algorithm research in first and second-order reliability methods. A number of algorithmic variations have been presented, and the effect of different limit state approximations, probability integrations, warm starting, most probable point search algorithms, and Hessian approximations has been discussed. These reliability analysis capabilities provide the foundation for reliability-based design optimization (RBDO) methods, and bi-level and sequential formulations have been presented. These RBDO formulations employ analytic sensitivities of reliability metrics with respect to design variables that either augment or define distribution parameters for the uncertain variables.

Relative performance of these reliability analysis and design algorithms has been measured for a number of benchmark test problems using the DAKOTA software. The most effective techniques in these computational experiments have been AMV^2+ for reliability analysis and sequential/surrogate-based approaches for RBDO. Continuing efforts in algorithm research will build on these successful methods through investigation of sequential RBDO with mixed surrogate and direct models (for probabilistic and deterministic components, respectively) and second-order RIA RBDO formulations employing generalized reliability indices.

These reliability analysis and design algorithms are now being applied to real-world applications in the shape optimization of micro-electro-mechanical systems, and initial experiences with this deployment are presented. Issues identified in deploying reliability methods to complex engineering applications include highly nonlinear, nonsmooth/noisy, and multimodal limit states, and potential simulation failures when evaluating parameter sets in the tails of input distributions. To mitigate these difficulties, a combination of continuing algorithm research, enhancements in problem formulation, and refinements to modeling and simulation capabilities is recommended.

7. Acknowledgments

The authors would like to express their thanks to Jonathan Wittwer and Jordan Massad for their assistance in formulating design problems for MEMS and to the Sandia Computer Science Research Institute (CSRI) for support of this collaborative work between Sandia National Laboratories and Vanderbilt University.

References

Adams, B.M., Eldred, M.S., Wittwer, J.W., and Massad, J.E., Reliability-Based Design Optimization for Shape Design of Compliant Micro-Electro-Mechanical Systems, abstract for *11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Portsmouth, VA, Sept. 6-8, 2006.

Agarwal, H., Renaud, J.E., Lee, J.C., and Watson, L.T., A Unilevel Method for Reliability Based Design Optimization, paper AIAA-2004-2029 in *Proceedings of the 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Palm Springs, CA, April 19-22, 2004.

Allen, J.J., Micro Electro Mechanical System Design, Taylor and Francis, Boca Raton, 2005.

Allen, M. and Maute, K., Reliability-based design optimization of aeroelastic structures, *Struct. Multidiscip. O.*, Vol. 27, 2004, pp. 228-242.

Box, G.E.P. and Cox, D.R., An Analysis of Transformations, *J. Royal Stat. Soc.*, Series B, Vol. 26, 1964, pp. 211-252.

Breitung, K., Asymptotic approximation for multinormal integrals, *J. Eng. Mech., ASCE*, Vol. 110, No. 3, 1984, pp. 357-366.

Chen, X., and Lind, N.C., Fast Probability Integration by Three-Parameter Normal Tail Approximation, *Struct. Saf.*, Vol. 1, 1983, pp. 269-276.

Der Kiureghian, A. and Liu, P.L., Structural Reliability Under Incomplete Probability Information, *J. Eng. Mech., ASCE*, Vol. 112, No. 1, 1986, pp. 85-104.

Du, X. and Chen, W., Sequential Optimization and Reliability Assessment Method for Efficient Probabilistic Design, *J. Mech. Design*, Vol. 126, 2004, pp.225-233.

Edwards, H.C., Sierra Framework for Massively Parallel Adaptive Multiphysics Applications Sandia Technical Report SAND2004-6277C, April 2003, Sandia National Laboratories, July 2005, Albuquerque, NM.

Eldred, M.S., Giunta, A.A., Wojtkiewicz, S.F., Jr., and Trucano, T.G., Formulations for Surrogate-Based Optimization Under Uncertainty, paper AIAA-2002-5585 in *Proceedings of the 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Atlanta, GA, Sept. 4-6, 2002.

Eldred, M.S., Giunta, A.A., van Bloemen Waanders, B.G., Wojtkiewicz, S.F., Jr., Hart, W.E., and Alleva, M.P., DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis. Version 3.1 Users Manual. Sandia Technical Report SAND2001-3796, Revised April 2003, Sandia National Laboratories, Albuquerque, NM.

Eldred, M.S., Agarwal, H., Perez, V.M., Wojtkiewicz, S.F., Jr., and Renaud, J.E., Investigation of Reliability Method Formulations in DAKOTA/UQ, (to appear) *Structure & Infrastructure Engineering: Maintenance, Management, Life-Cycle Design & Performance*, Taylor & Francis Group.

Eldred, M.S., Bichon, B.J., and Wojtkiewicz, S.F., Jr., Second-Order Reliability Formulations in DAKOTA/UQ, (in review) *Structure & Infrastructure Engineering: Maintenance, Management, Life-Cycle Design & Performance*, special issue on reliability analysis in aerospace systems, Taylor & Francis Group.

Fadel, G.M., Riley, M.F., and Barthelemy, J.-F.M., Two Point Exponential Approximation Method for Structural Optimization, *Structural Optimization*, Vol. 2, No. 2, 1990, pp. 117-124.

Gill, P.E., Murray, W., Saunders, M.A., and Wright, M.H., User's Guide for NPSOL 5.0: A Fortran Package for Nonlinear Programming, System Optimization Laboratory, Technical Report SOL 86-1, Revised July 1998, Stanford University, Stanford, CA.

Giunta, A.A. and Eldred, M.S., Implementation of a Trust Region Model Management Strategy in the DAKOTA Optimization Toolkit, paper AIAA-2000-4935 in *Proceedings of the 8th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, Long Beach, CA, September 6-8, 2000.

Haldar, A. and Mahadevan, S., Probability, Reliability, and Statistical Methods in Engineering Design, 2000 (Wiley: New York).

Hohenbichler, M. and Rackwitz, R., Sensitivity and importance measures in structural reliability, *Civil Eng. Syst.*, Vol. 3, 1986, pp. 203-209.

Hohenbichler, M. and Rackwitz, R., Improvement of second-order reliability estimates by importance sampling, *J. Eng. Mech., ASCE*, Vol. 114, No. 12, 1988, pp. 2195-2199.

Hong, H.P., Simple Approximations for Improving Second-Order Reliability Estimates, *J. Eng. Mech., ASCE*, Vol. 125, No. 5, 1999, pp. 592-595.

Jensen, B.D., Parkinson, M.B., Kurabayashi, K., Kowell, L.L., and Baker, M.S., Design Optimization of a Fully-Compliant Bistable Micro-Mechanism, *Proc. 2001 ASME Intl. Mech. Eng. Congress and Exposition*, New York, NY, Nov. 11-16, 2001.

Karamchandani, A. and Cornell, C.A., Sensitivity estimation within first and second order reliability methods, *Struct. Saf.*, Vol. 11, 1992, pp. 95-107.

Kuschel, N. and Rackwitz, R., Two Basic Problems in Reliability-Based Structural Optimization, *Math. Method Oper. Res.*, Vol. 46, 1997, pp.309-333.

Meza, J.C., OPT++: An Object-Oriented Class Library for Nonlinear Optimization, Sandia Technical Report SAND94-8225, Sandia National Laboratories, Livermore, CA, March 1994.

Nocedal, J., and Wright, S.J., Numerical Optimization, Springer, New York, 1999.

Rackwitz, R., and Fiessler, B., Structural Reliability under Combined Random Load Sequences, *Comput. Struct.*, Vol. 9, 1978, pp. 489-494.

Rackwitz, R., Optimization and risk acceptability based on the Life Quality Index, *Struct. Saf.*, Vol. 24, 2002, pp. 297-331.

Rosenblatt, M., Remarks on a Multivariate Transformation, *Ann. Math. Stat.*, Vol. 23, No. 3, 1952, pp. 470-472.

Sues, R., Aminpour, M. and Shin, Y., Reliability-Based Multidisciplinary Optimization for Aerospace Systems, paper AIAA-2001-1521 in *Proceedings of the 42rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Seattle, WA, April 16-19, 2001.

Tu, J., Choi, K.K., and Park, Y.H., A New Study on Reliability-Based Design Optimization, *J. Mech. Design*, Vol. 121, 1999, pp.557-564.

Wang, L. and Grandhi, R.V., Efficient Safety Index Calculation for Structural Reliability Analysis, *Comput. Struct.*, Vol. 52, No. 1, 1994, pp. 103-111.

Wittwer, J.W., Simulation-Based Design Under Uncertainty for Compliant Microelectromechanical Systems, Ph.D. dissertation, Brigham Young University, April, 2005.

Wittwer, J.W., Baker, M.S., and Howell, L.L., Robust design and model validation of nonlinear compliant micromechanisms, *J. Microelectromechanical Sys.*, Vol. 15, No. 1, 2006, *to appear.*

Wojtkiewicz, S.F., Jr., Eldred, M.S., Field, R.V., Jr., Urbina, A., and Red-Horse, J.R., A Toolkit For Uncertainty Quantification In Large Computational Engineering Models, paper AIAA-2001-1455 in *Proceedings of the 42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Seattle, WA, April 16-19, 2001.

Wu, Y.-T., and Wirsching, P.H., A new algorithm for structural reliability estimation, *J. Eng. Mech., ASCE*, Vol. 113, 1987, pp. 1319-1336.

Wu, Y.-T., Millwater, H.R., and Cruse, T.A., Advanced Probabilistic Structural Analysis Method for Implicit Performance Functions, *AIAA J.*, Vol. 28, No. 9, 1990, pp. 1663-1669.

Wu, Y.-T., Computational Methods for Efficient Structural Reliability and Reliability Sensitivity Analysis, *AIAA J.*, Vol. 32, No. 8, 1994, pp. 1717-1723.

Wu, Y.-T., Shin, Y., Sues, R., and Cesare, M., Safety-Factor Based Approach for Probability-Based Design Optimization, paper AIAA-2001-1522 in *Proceedings of the 42rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Seattle, WA, April 16-19, 2001.

Xu, S., and Grandhi, R.V., Effective Two-Point Function Approximation for Design Optimization, *AIAA J.*, Vol. 36, No. 12, 1998, pp. 2269-2275.

Zou, T., Mahadevan, S., and Rebba, R., Computational Efficiency in Reliability-Based Optimization, *Proceedings of the 9th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability*, Albuquerque, NM, July 26-28, 2004.