RC23456 (W0411-232) November 30, 2004
Mathematics

IBM Research Report

Planning Models for Component-Based Software Offerings under Uncertain Operational Profiles

Mary Helander, Bonnie Ray
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

CHAPTER 1

PLANNING MODELS FOR COMPONENT–BASED SOFTWARE OFFERINGS UNDER UNCERTAIN OPERATIONAL PROFILES

Mary Helander and Bonnie Ray

Department of Mathematical Sciences, IBM Watson Research Center
Yorktown Heights, NY 10598
E-mail: helandmus.ibm.com, bonnierus.ibm.com

We present a modeling framework for allocating development or support effort among software components to meet a specified cost or reliability objective. The approach is based on knowledge of the linkage between the operational profile and the architecture of the system, but allows for uncertainty through the use of probability distributions to characterize usage. An approach based on stochastic optimization is presented to obtain efficient solutions to the allocation problem. Results are demonstrated when uncertain usage is characterized by a Dirichlet distribution.

1. Introduction

Guidelines for achieving a specified reliability target under resource constraints play a key role in the software development planning process. In particular, methods that determine how to allocate resources among components of a software system to facilitate cost–efficient progress toward a quantified system reliability goal are essential. Here, a component is defined as a set of operations, a subsystem, a module, an object, or any other distinguishable software entity that can be assigned a failure intensity representing its reliability. If the intended usage of the system is given, then system reliability can be computed as a function of component failure intensities, the expected component utilizations, and the specified usage. Specification of usage through the assignment of occurrence probabilities to operations form the quantification known as the *operational profile* [1]. However for many systems, in particular commercial software systems, the usage of the system in a production setting may vary considerably from an expected usage, or the assumed usage as characterized for test phases of system development. This paper describes a modeling framework for cost and reliability planning in software system development that allows for random variation in the operational profile.

Previous authors have addressed the problem of software reliability allocation in the context of a specified operational profile. For instance, Poore, Mills, and Mutchler [2] used a spreadsheet approach to consider various strategies for allocating reliability to software modules. The paper of Helander, Zhao, Ohlsson [3] provides an analytical solution to the optimal allocation problem based on standard nonlinear optimization methods. A component utilization matrix is used to link system structure to *operations*, which are partitionings of the software system from a user's perspective. The paper of Leung [4] solved the reliability allocation problem in the case of an operational profile that was specified up to an ϵ allowable difference, where the ϵ uncertainty was the same across all operations. We build directly on the framework of [3], by allowing for very general uncertainty in the operational profile through the use of probability distributions on the operation occurrences. This is especially important as the high-level componentization of software systems becomes more common. By high-level, we mean that what were once considered as individual products are now bundled together in different ways and sold as software systems. An example is the componentization of software offerings from IBM, which bundles key elements of products such as DB2, Tivoli, and Websphere Application Server together so that they are quickly deployable in a customized setting. The usage of the components may vary widely from setting to setting, but reliability targets for the individual components need to be allocated so that overall reliability targets are achieved across a range of settings.

In Section 2, we provide the general model formulation for specification of component failure intensities to plan for a software system reliability target while minimizing costs. Section 3 gives the derivation of the stochastic optimization problem under a particular distributional assumption about usage. Section 4 presents an example. Section 5 concludes.

2. Mathematical Formulation as an Optimal Planning Problem

Let n denote the number of software components comprising a software system. In assembling components to form a software offering, we want the components to be reliable enough so that the probability of failure-free execution for a configured system has probability of at least ρ $(0 < \rho < 1)$ of achieving failure–free execution with respect to an execution time interval of length τ. The problem of interest is to determine what the failure intensities of the individual components should be to achieve this target in the most cost effective manner. This assumes a cost associated with achieving a specified performance, a common concept in software engineering economics.

Denote $f(\lambda_1, \lambda_2, \ldots, \lambda_n)$ as the total cost of achieving the failure intensities $\lambda_1, \lambda_2, \ldots, \lambda_n$, which we assume to be a pseudoconvex, non–increasing function of the λ_j's. The function $\mathcal{R}(\lambda_1, \lambda_2, \ldots, \lambda_n; \tau)$ measures system reliability in terms of $\lambda_1, \lambda_2, \ldots, \lambda_n$, which will be used as the main decision variables in the reliability allocation cost–optimization problem. The following model simultaneously finds $\lambda_1, \lambda_2, \ldots, \lambda_n$:

$$\text{Minimize} \quad f(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{1}$$

$$\text{Subject to: } \mathcal{R}(\lambda_1, \lambda_2, \ldots, \lambda_n; \tau) \;=\; \exp\left\{ -\sum_{j=1}^{n} \sum_{i=1}^{m} p_i \mu_{ij} \lambda_j \tau \right\} \;\geq\; \rho \tag{2}$$

$$\lambda_j \;\geq\; 0 \quad \text{for } j = 1, \ldots, n \tag{3}$$

where $0 < p_i < 1$, $\sum_{i=1}^{m} p_i = 1$ are the operational profile parameters and μ_{ij} $(0 \leq \mu_{ij} \leq 1)$ are the component usage parameters for $i = 1, \ldots, m$ operations and $j = 1, \ldots, n$ components with $\sum_{j=1}^{n} u_{ij} = 1$ for each i. Note that the reliability function in (2) is derived under the assumption that failure events are statistically independent and the system architecture is such that a failure in one component causes failure of the entire system, *i.e.* the system is not fault-tolerant with respect to the components as identified in the planning model. Other forms for $\mathcal{R}(\cdot)$ could be derived under different assumptions concerning the system architecture, although the resulting optimization problems become more complex. Given positive association of failure events, it appears that the form based on (2) is helpful for planning conservatively since it should assure the system reliability target is a a lower bound.

This model specified by (1), (2), and (3) is the multivariate, nonlinear, constrained optimization model introduced in [3]. It can be equivalently restated as:

$$\text{Minimize} \quad f(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{4}$$

$$\text{Subject to: } g_0(\lambda_1, \lambda_2, \ldots, \lambda_n) \;=\; \rho \ln(\rho) \;+\; \rho\tau \sum_{i=1}^{m} p_i \sum_{j=1}^{n} \mu_{ij}\lambda_j \;\leq\; 0 \tag{5}$$

$$g_j(\lambda_1, \lambda_2, \ldots, \lambda_n) \;=\; -\lambda_j \;\leq\; 0 \quad \text{for } j = 1, \ldots, n \tag{6}$$

which follows a standard nonlinear programming model form involving minimization of a nonlinear objective function constrained by "≤ 0" inequalities. Conditions and solutions for this form are given in [3] for some common cost functions in software economics (see, *e.g.*, [5]).

In this framework, uncertainty in the operational profile can be introduced by treating the operational profile parameters as random variables, giving a *stochastic programming* formulation. This formulation and general solutions are discussed in the next section.

3. Stochastic Optimal Reliability Allocation

Let $\tilde{\xi} = \{\xi_i, \ldots, \xi_m\}$ denote the random vector representing random variable replacements for the deterministic operational profile parameters in constraint (5), *i.e.* for the p_i, $i = 1, \ldots, m$. The decision model stated by (4), (5) and (6) in the context of a random operation profile leads to a problem statement

$$\text{"Minimize"} \quad f(\lambda_1, \lambda_2, \ldots, \lambda_n; \tilde{\xi}) \tag{7}$$

$$\begin{aligned} \text{Subject to: } \; G_0(\lambda_1, \lambda_2, \ldots, \lambda_n; \tilde{\xi}) \;&=\; \rho \ln(\rho) \\ &+ \rho\tau \sum_{i=1}^{m} \xi_i \sum_{j=1}^{n} \mu_{ij}\lambda_j \;\leq\; 0 \end{aligned} \tag{8}$$

$$g_j(\lambda_1, \lambda_2, \ldots, \lambda_n) \;=\; -\lambda_j \;\leq\; 0 \quad \text{for } j = 1, \ldots, n \tag{9}$$

In this formulation, we have replaced g_0 by G_0 to indicate that g_0 is now a random variable. As noted in Kall and Wallace [6], this problem as a whole, and the constraint (8), are not well defined when trying to make a decision for setting the $\lambda_1, \lambda_2, \ldots, \lambda_n$ values, prior to knowing a realization of $\tilde{\xi}$. To address this, we may consider a deterministic equivalent of the model specified by (7), (8) and (9) by replacing (8) with a *probabilistic* *constraint* such as:

$$P[G_0(\lambda_1, \lambda_2, \ldots, \lambda_n; \tilde{\xi}) \leq 0] \geq \psi, \tag{10}$$

where $0 < \psi < 1$. This constraints says that we want the chance that the overall reliability target is met to be at least ψ.

Note that when a realization of $\tilde{\xi}$ is specified, *e.g.*, as $p_1, \ldots, p_m$, then the values of p_i should sum to one, and each value must be between zero and one. A convenient probability distribution that achieves these properties is a Dirichlet distribution. A random variable $\tilde{\xi}$ is said to follow a Dirichlet distribution if its probability distribution function has the form

$$p(\tilde{\xi}) = Dirichlet(\tilde{\xi};\ \mathbf{v}) = \frac{1}{Z(\mathbf{v})} \prod_{i=1}^{m} \xi_i^{v_i - 1}, \tag{11}$$

when $\xi_1, \ldots, \xi_m \geq 0, \sum_{i=1}^m \xi_i = 1$ and $v_1, \ldots, v_m > 0$. The normalization constant is

$$Z(\mathbf{v}) = \frac{\prod_{i=1}^m \Gamma(v_i)}{\Gamma(\sum_{i=1}^m v_i)} \quad . \tag{12}$$

A univariate Dirichlet distribution reduces to a standard Beta distribution. The values $\mathbf{v} = (v_1, \ldots, v_m)$ determine the shape of the distribution. See [7] for additional details on Dirichlet distributions.

Standard results in stochastic optimization, for example [6], show that a solution involving (10) is obtained by applying the same non–linear programming techniques used to solve the deterministic formulations. For example, when (10) is quasiconvex and differentiable with respect to $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the Kuhn-Tucker conditions may be used to characterize and identify an optimal solution. Application of such techniques, including validation of properties such as quasiconvexity for (10), requires derivation of the distribution of G_0. The next subsection explores this derivation.

3.1. *Derivation of the Distribution for G_0*

Upon examination of expression (10), we see that G_0 is basically a shifted and scaled linear combination of Dirichlet random variables. Recent results by Provost and Cheong [8] provide an expression for the cumulative distribution function of a linear combination of Dirichlet random variables in integral form. Let $a = \rho \log(\rho)$, $b = \rho\tau$ and $c_i = \sum_{j=1}^n \mu_{ij}\lambda_j$ in (8). Then the distribution function is given by

$$F_{G_0}(z) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{sin[\frac{1}{2}\sum_{i=1}^m 2v_i \tan^{-1}\{(c_i - \frac{(z-a)}{b})w\}]}{w \prod_{i=1}^m \{1 + (c_i - \frac{(z-a)}{b})^2 w^2\}^{(2v_i)/4}}\, dw, \tag{13}$$

for $a + b\min(c_i) < z < a + b\max(c_i)$.

Note that $-\frac{a}{b} = -\log(\rho)/\tau = \lambda_s$, the system failure intensity target. Then it follows from (7)–(10), that the optimization problem can now be restated as

$$\text{Minimize} \quad f(\lambda_1, \lambda_2, \ldots, \lambda_n; \tilde{\xi}) \tag{14}$$

$$\text{Subject to: } h(\lambda_1, \lambda_2, \ldots, \lambda_n) = -F_{G_0}(0; \lambda_1, \lambda_2, \ldots, \lambda_n) + \psi \leq 0 \tag{15}$$

$$-\lambda_1, -\lambda_2, \ldots, -\lambda_n \leq 0 \quad \text{for } j = 1, \ldots, n, \tag{16}$$

where

$$F_{G_0}(0; \lambda_1, \ldots, \lambda_n) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{sin[\frac{1}{2}\sum_{i=1}^m 2v_i \tan^{-1}\{(c_i - \lambda_s)w\}]}{w\prod_{i=1}^m \{1 + (c_i - \lambda_s)^2 w^2\}^{(v_i)/2}}\, dw, \tag{17}$$

from (13).

As stated above, the Kuhn-Tucker sufficient conditions can be used to find a global optimal solution to the stochastic optimization problem provided (14) is pseudoconvex and $h(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is quasiconvex and differentiable with respect to the λ_i in the feasible region.

The Kuhn-Tucker sufficient conditions are as follows:

$$\left[\tfrac{\partial TC}{\partial \lambda_1}\, ,\; \tfrac{\partial TC}{\partial \lambda_2}\, ,\; \ldots\, ,\; \tfrac{\partial TC}{\partial \lambda_n}\right]^T + d\left[\tfrac{\partial h}{\partial \lambda_1}\, ,\; \tfrac{\partial h}{\partial \lambda_2}\, , \ldots\, ,\; \tfrac{\partial h}{\partial \lambda_n}\right]^T - [\gamma_1\, ,\; \gamma_2\, , \ldots\, , \gamma_n]^T = 0 \tag{18}$$

$$d\, h(\lambda_1, \lambda_2, \ldots, \lambda_n) \;=\; 0 \tag{19}$$

$$\gamma_j \lambda_j \;=\; 0 \qquad \text{for } j = 1, \ldots, n \tag{20}$$

where d and $\gamma_1, \gamma_2, \ldots, \gamma_n$ are Lagrangian multipliers associated with inequality constraints (5) and (6) respectively.

From (15), we have

$$h(\lambda_1, \lambda_2, \ldots, \lambda_n) = -F_{G_0}(0; \lambda_1, \lambda_2, \ldots, \lambda_n) + \psi. \tag{21}$$

After some algebraic manipulations, the partial derivative of

$F_{G_0}(0; \lambda_1, \ldots, \lambda_n)$ with respect to λ_j needed for (18) has the following form:

$$\frac{\partial F_{G_0}}{\partial \lambda_j} = \frac{1}{\pi} \int_0^\infty \frac{\cos[\sum_{i=1}^m v_i \tan^{-1}\{(c_i - \lambda_s)w\}] \sum_{i=1}^m v_i \mu_{ij} [\frac{1}{1+(c_i-\lambda_s)^2 w^2}]}{\prod_{i=1}^m \{1 + (c_i - \lambda_s)^2 w^2\}^{v_i/2}} \, dw$$
$$- \frac{1}{\pi} \int_0^\infty \frac{\sin[\sum_{i=1}^m v_i \tan^{-1}\{(c_i - \lambda_s)w\}]}{\prod_{i=1}^m \{1 + (c_i - \lambda_s)^2 w^2\}^{v_i/2}} \sum_{i=1}^m \frac{v_i w (c_i - \lambda_s)\mu_{ij}}{[1 + (c_i - \lambda_s)^2 w^2]} \, dw. \tag{22}$$

Looking again at (17), it is not obvious to show see that $h(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is quasiconvex, given its complicated form. However, inspection of feasible region plots in two-dimensions suggests that for $\psi > 0.5$, the boundary of the feasible region is convex, and appears to be approximately piecewise linear. When $\psi = 0.5$, the boundary of the feasible region coincides with that for the deterministic solution. Interestingly, the feasible region appears to be concave for $\psi < 0.5$ However, recall that ψ denotes the chance that the overall reliability target will be met. Thus, from a practical standpoint, restriction to $\psi > 0.5$ to achieve a convex feasible region does not unduly limit the formulation.

From (17), we also see that $F_{G_0}(0; \lambda_1, \ldots, \lambda_n)$ takes value 1/2 when $\lambda_i = \lambda_s$ for each $i = 1, \ldots, n$. This implies that, for any ψ, the boundary of the feasible region always goes through the point $\lambda_i = \lambda_s$, which forms an inflection point (under varying ψ) for the boundry of the feasible region corresponding to the system reliability constraint.

Given that we have not definitively determined the quasi-convexity of (17), it is possible that the solution to (14) obtained using numerical optimization techniques may not provide the global optimal solution. While this is not entirely satisfying, a "close-to-optimal" solution may still provide significant gains over the deterministic solution when uncertainty in the usage profile is large. We provide examples to illustrate the effects of allowance for uncertainty in Section 4.

3.2. *Solution Implementation*

To obtain a solution to (14)—(17), the `fmincon` function of Matlab's Optimization Toolbox (v. 6.2, Release 13) was used to directly minimize (14) subject to the constraints (15)—(17). The function (17) was evaluated numerically using adaptive Simpson quadrature methods, with the Matlab function `quad`. We note that several parameters of `fmincon` and `quad` affect the quality of the obtained solution. First, the precision tolerance (`tol`) of the `quad` function can result in differences in the optimal solution obtained. Decreasing the tolerance to a very small value usually provided good results, at the cost of increasing the solution computation time or exceeding the maximum number of allowed function evaluations. However, making the tolerance value too small sometimes resulted in singularity and non-convergence.

Second, several parameters of the `fmincon` function affect the quality of the solution. These are a) `tolx`, the required change in λ_i for continued iteration, b) `tolfun`, the required change in the objective function for continued iteration, and c) `tolconstr`, the required distance of the solution from the active constraint. Matlab documentation recommends staying with the default values, although we found it necessary to experiment with smaller values in order to achieve algorithm convergence.

While these algorithm parameters impacted the achieved solution, the choice of starting point for finding the solution seemed to have the greatest effect. We found that starting at a value a small step into the feasible region from the deterministic solution generally proved an effective strategy.

4. Examples

In this section, we graphically illustrate the effect of uncertainty on the allocation problem specified by (1), (2), and (3), by extending the $n = 2$ component example from [3] to allow variation in assumed usage. Specifically, the input parameters of the model in the deterministic case are:

$$n = 2 \qquad \rho = .99 \qquad \vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$$

$$m = 2 \qquad \tau = 1 \qquad \vec{\mu} = \begin{bmatrix} \mu_{11} \ \mu_{12} \\ \mu_{21} \ \mu_{22} \end{bmatrix} = \begin{bmatrix} 0.6 \ 0.4 \\ 0.0 \ 1.0 \end{bmatrix}.$$

The cost is assumed to be inversely proportional to the achieved reliability, following an inverse power law of the form $C(\lambda) = \frac{\beta}{(\lambda-\delta)^\alpha}$, $\lambda > \delta$. The **InvPow** cost parameters are taken as

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 320 \\ 410 \end{bmatrix} \qquad \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0.30 \\ 0.25 \end{bmatrix} \qquad \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Figure 1 shows the solution for the deterministic version of the model, as in [3], which corresponds to the stochastic solution for any choice of Dirichlet parameters v_i when the chance of meeting the reliability target, ψ, is

0.5. Using a Dirichlet distribution with $v_1 = 7, v_2 = 3$ to characterize uncertainty in p_i, Figure 2 shows the effect on the feasible region of increasing ψ, while Figure 3 shows the cost contours and resulting solution, respectively. Table 1 gives the precise values of λ_1, λ_2 and the resulting Total Cost to achieve the specified reliability under the Inverse Power cost function given above. As expected, the Total Cost increases with ψ.

The four plots of Figure 4 show the effect of varying v_i over the values $v_1 = v_2 = 0.5$ (top left), 1 (top right), 2 (bottom left), and 16 (bottom right) for values of ψ varying from 0.2 to 0.8, where the Dirichlet distributions corresponding to the v_1 values are shown in Figure 5. We see that decreased uncertainty in the usage, as specified through decreased variability in p_i, results in less variation from the $\psi = 0.5$ case (also corresponding to the deterministic boundary).

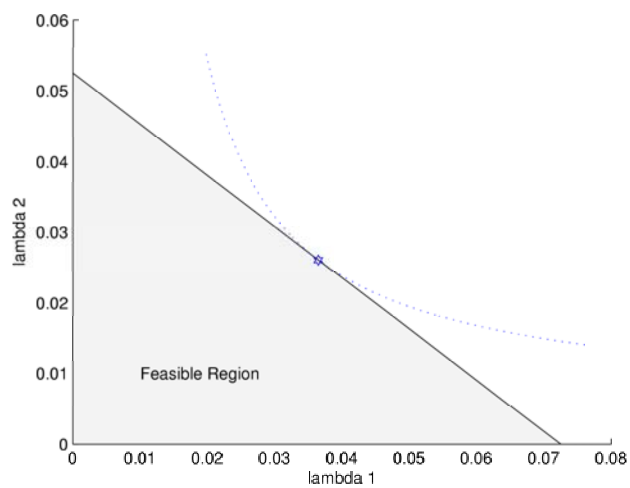

Fig. 1. Solution to the deterministic model for the 2 component example

5. Conclusions

In this paper, we have formulated and solved the optimal reliability allocation problem for a system of software components when uncertainty in the operational profile is quantified using probability distributions. In future research, we plan to make more explicit the relationship between the

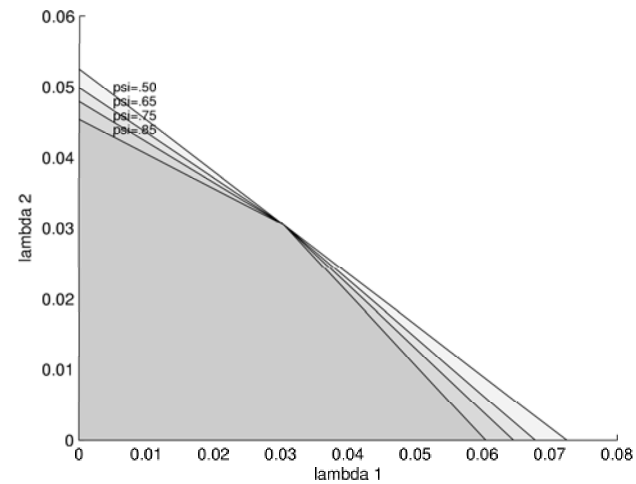

Fig. 2. Feasible region changes with varying ψ.

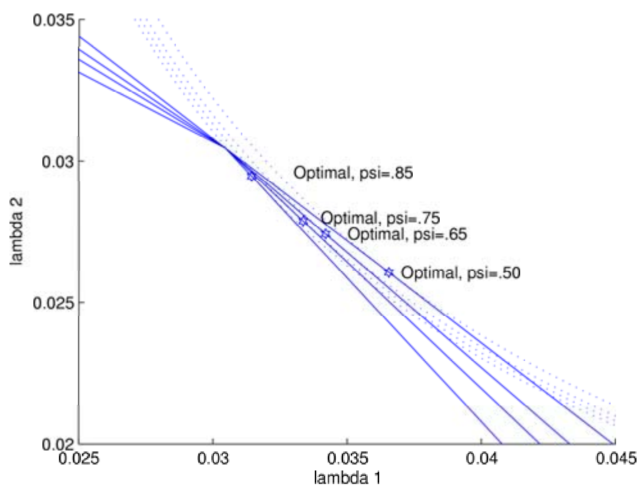

Fig. 3. Solutions to the stochastic model for the two component example with varying ψ

Table 1. Component reliability solutions for two-component system under varying probabilistic constraints and resulting total cost

	ψ	λ_1	λ_2	Cost
Deterministic	—-	0.03654	0.02606	1884.10
Stochastic	0.50	0.03654	0.02606	1884.10
Stochastic	0.65	0.03419	0.02742	1888.58
Stochastic	0.75	0.03335	0.02788	1890.98
Stochastic	0.85	0.03144	0.02946	1893.06

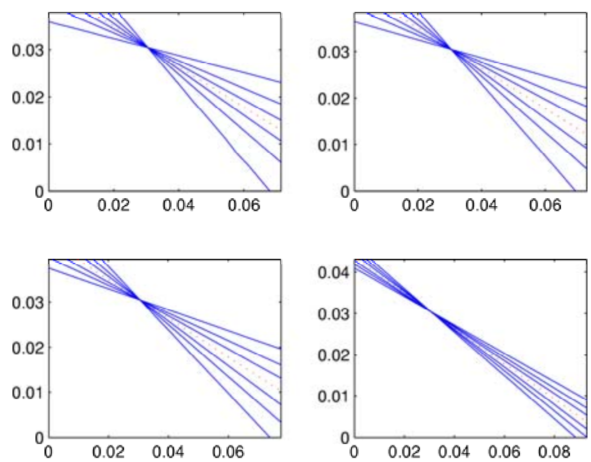

Fig. 4. Behavior of the feasible region boundary formed by the system reliability for the two component example with ψ varying from $0.2, \ldots, 0.8$ and $v_1 = 0.5$ (top left), 1 (top right), 2 (bottom left), 16 (bottom right).

stochastic formulation of the problem, as outlined in this paper, and other "robust" deterministic approaches to the same problem, such as that given in [4]. Additionally, we will consider how the present formulation can be made relevant to in application areas having attributes other than reliability as the system attribute of interest.

Acknowledgments

The authors wish to thank Samer Takriti and Guiseppe Paleologo for extensive discussions concerning stochastic optimization methods and tools.

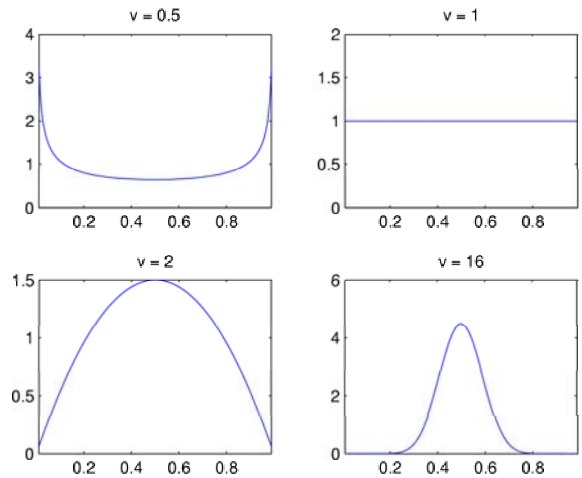

Fig. 5. Illustration of the Dirichlet distribution for $v_1 = 0.5$ (top left), 1 (top right), 2 (bottom left), 16 (bottom right).

References

1. J. D. Musa, Operational Profiles in Software Reliability Engineering, *IEEE Software*, vol. 10, no. 2, 14–32 (1993).
2. J.H.Poore, H.D. Mills, and D. Mutchler, "Planning and certifying software system reliability,", *IEEE Software*, vol.10, no. 1, 88–99 (1993).
3. M. E. Helander, M. Zhao, and N. Ohlsson, Planning models for software reliability and cost. *IEEE Transactions on Software Engineering*, vol. 24, no. 6, 420–434 (1998).
4. Y.-W. Leung, Software reliability planing models under an uncertain operational profile. *Journal of the Operational Research Society* vol. 48, 401–411 (1997).
5. B. W. Boehm, *Software Engineering Economics*. Prentice–Hall: New Jersey (1981).
6. P. Kall and S. W. Wallace, *Stochastic Programming*. John Wiley and Sons: New York (1994).
7. S. Kotz, N., Balakrishnan, and N. Johnson, Continuous Multivarite Distributions, Volume 1: Models and Applications, John Wiley and Sons: New York (2000).
8. S.B. Provost and Y.-H. Cheong, On the distribution of linear combinations of the components of a Dirichlet random vector, *The Canadian Journal of Statistics*, vol. 28, 417–425 (2000).