Optimization of Project Porttfolios

Senay Solak
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA30332
ssolak@isye.gatech.edu

John-Paul B. Clarke
School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA30332
john-paul.clarke@ae.gatech.edu

Ellis L. Johnson
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA30332
ejohnson@isye.gatech.edu

Abstract

The project portfolio optimization problem is a difficult practical problem, for which a comprehensive model and solution methodology has not been developed in the existing limited approaches in the literature. In this study, we fill this gap by formally defining and effectively modeling several complexities that are inherent in this problem, and developing efficient solution procedures. The proposed approach is the first such approach that models the endogenous uncertainty inherent in this decision process, and at the same time includes a computationally practical solution procedure. The implementation of the model by organizations can lead to significant increases in project returns. From a theoretical standpoint, a new and efficient formulation technique to model nonanticipativity in multistage stochastic programs with endogenous uncertainty is developed, enabling direct scenario based decomposition in such problems. In addition, a feasible dual conversion based tight lower bounding algorithm that can also be used for any similarly structured problem is developed.

1 Introduction

Project portfolio management involves research and development (R&D) projects aimed to design, test and improve a technology, or the process of building a technology. Technology development is often an essential part of the operational strategy of an organization, during which deployment or implementation decisions are made. In most cases, organizations have several potential technology projects with different characteristics that they can choose to invest in and develop using available resources. Selection of projects and allocation of resources to the selected projects are important decisions with huge economic implications for an organization.

Characteristics of technology projects involve resource levels required for research and development, and projected returns after deployment, which are unknown at the time of investment, but for which some information on the uncertainty is available. Given these uncertainties and resource limitations over a planning horizon, the project portfolio optimization problem deals with the selection of R&D projects and determination of optimal resource allocations for the current planning period such that the expected total discounted return or a function of this expectation for all projects over an infinite time horizon is maximized.

Depending on the application, the set of candidate projects may have several attributes. For instance, a technology project may require a deployment phase after development, which amounts to a delay in return realization. A technology can be developed over multiple years, however a fixed operating cost is incurred for each project that remains active, i.e. where development has started but is not complete. Furthermore, multi-way dependencies may exist between technologies, which implies that the joint return of two dependent technologies can be different from the sum of their individual returns.

Markowitz (1952) laid the background for modern financial portfolio theory, which has been studied extensively since then. Markowitz (1952) suggests that investors should select portfolios based on overall risk-reward characteristics of the securities, rather than investing on a single security with the best risk-reward characteristic. Tobin (1958) studies super efficient portfolios with risk free assets, while Sharpe (1964) develops the capital asset pricing models. Since then, many other modeling and optimization techniques have been proposed for financial portfolio optimization.

Although at first glance, it may seem that financial portfolio optimization theory could be directly applied to project portfolio management, there are clear differences between the two problems. One distinction is in the realization of returns. The realization time and the variance in the return of a project is dependent on the investment made on that project. However, for financial securities, both the risk and the time of return realization is independent of the amount of the security that is purchased. Assuming that no one investor will seek to make a single purchase of all or the vast majority of a company's stocks that will cause the price of the security to change by virtue of the purchase itself, the value of the security will solely be based on the performance of the company in question. A second difference between the two problems is about the correlation among project returns. In financial portfolio theory, the correlation in returns is assumed to be independent of the way in which resources are allocated. On the other hand, the correlation among the returns of projects is dependent on investment levels, because resources spent on one project are taken away from other projects, thus preventing early return realization in these projects. Finally, a third distinction is the dependencies of technology projects in terms of produced returns. In financial theory, the cumulative return from two purchased securities is assumed to be equal to the sum of the individual returns of the securities. However, as noted above, projects have dependencies which can have a positive or negative effect in the realization of cumulative joint returns.

Despite the importance and economic significance of project portfolio selection and the existence of several operations research models, the industrial use of these models has been limited. This is mainly due to the fact that none of the proposed models has been able to capture the full range of complexity that exists in project portfolios. Reyck et al. (2005) study the impact of project portfolio management techniques on the performance of projects and portfolios of projects. The authors identify certain key components required for an effective portfolio management approach, which include the following capabilities: capturing of financial returns and risks of assets, modeling interdependencies, determination of prioritization, alignment and selection of projects, modeling organizational constraints and ability to dynamically reassess the portfolio. Linton et al. (2002) provide a review of proposed project portfolio management methods, and categorize the existing methods into three groups. The first category contains approaches based on net present value (NPV) calculations, while the second group consists of scoring methods and the last group covers other decision analysis tools. However, none of the considered approaches are able to model and deliver the set of capabilities identified by Reyck et al. (2005).

The proposed models for project portfolio management include capital budgeting models, which typically use accounting based criteria, such as return on investment or internal rate of return. These models capture interdependencies between different projects, but fail to model the uncertainty in returns and required investments (Luenberger 1998). More recent project portfolio models capture both the uncertainty in returns and interdependencies. However, these models assume that the required cash flows for projects are known, and the investment decisions consist of binary starting or stopping decisions for projects (Ghasemzadeh et al. 1999, Gustafsson and Salo 2005). One example where the amount of resources allocated to each project is treated as a decision variable is given by Norkin et al. (1998). The example is formulated as a stochastic integer program, but the interdependencies between multiple projects are not modeled.

Other approaches to project portfolio management include real options based methods. Despite some disadvantages from an optimization perspective, these methods are superior to NPV based methods. Bardhan et al. (2006) propose a multi-period optimization model where the objective is based on real options values of the portfolio calculated according to the results from Bardhan et al. (2004). Campbell (2001) and Lee et al. (2001) model project contingencies as real options to determine optimal startup dates for the projects. Tralli (2004) devises a real options valuation architecture from a decision tree analysis structure and presents a case study. Similarly, Wu and Ong (2007) combines the mean-variance model of classical financial theory with real options, and describe a project selection methodology based on the developed framework. However, one major disadvantage of real options based approaches is that they require the estimation of cash flows for the projects. Given these estimates, these models try to determine the optimum starting, continuation or completion times for the projects in a portfolio. Thus, despite its significance, the option of rebalancing through allocation of resources in each planning period is not modeled (Cooper et al. 2001). Chan et al. (2007) emphasize this problem and suggest a dynamic methodology based on a two-phase model of project evolution. However, the model does not capture the interdependencies or resource allocation decisions discussed above.

There are also other somewhat more simplistic approaches to the technology project portfolio problem, which either contain deterministic models or include several restrictive assumptions. Dickinson et al. (2001) present a model developed to optimize a portfolio of product development improvement projects. Using a dependency matrix, which quantifies the interdependencies between projects, a deterministic nonlinear integer programming model is developed to optimize project selection. April et al. (2003) describe a simulation optimization tool for technology project portfolio management. The tool utilizes metaheuristics to search for good technology portfolios, and is limited in capturing the interdependencies among technologies. Elfes et al. (2005) address the problem of determining optimal technology investment portfolios that minimize mission risk and maximize the expected science return of space missions. The solution approach described in the study is based on a deterministic linear programming formulation and sensitivity analysis. Lincoln et al. (2006) develop a method for prioritization of technology investments using a deterministic linear programming formulation to maximize an objective function subject to overall cost constraints. Goldner and Borener (2006) describe a quantitative framework to evaluate the performance of research portfolios, where the developed tool is used to evaluate and explore independent investments strategies, but no numerical optimization techniques are described.

In addition to these models, most strategic planners and project portfolio managers rely on tools based on expert opinions, such as Analytical Hierarchy Process and Quality Function Deployment, in planning the funding of technology development (Thompson 2006). Similar systematic evaluation methods are also proposed by Sallie (2002) and Utturwar et al. (2002), where the authors propose bilevel approaches in selecting technologies to invest. The latter study also contains an optimization procedure based on a Genetic Algorithm implementation. Clearly, these tools are also very limited in their ability to fully quantify the complicated return and investment structure inherent in project portfolios.

In summary, the project portfolio optimization problem is a difficult practical problem, for which a comprehensive model and solution methodology has not been developed in the existing limited approaches in the literature. In this study, we fill this gap by formally defining and effectively modeling several complexities that are inherent in this problem, and developing efficient solution procedures. More specifically, contributions of this study include the following: A comprehensive model that captures all relevant concerns in project portfolio management has been developed. To the best of our knowledge, it is the first such approach that (i) provides an accurate representation of the stochastic decision process in project portfolio management, (ii) models the endogenous uncertainty inherent in this decision process, and at the same time (iii) includes a computationally practical solution procedure. In addition, from a theoretical standpoint, contributions are as follows: (i) a new and efficient formulation technique to model nonanticipativity in multistage stochastic programs with endogenous uncertainty is developed, (ii) the developed formulation enables scenario based decomposition in such problems, in addition to the application of other methods developed for classical multistage stochastic programs, and (iii) a tight lower bounding algorithm based on feasible dual conversion that can be used for any similarly structured problem is developed. As shown in Figure 1, our proposed methodology is able to capture all the important aspects required from a project portfolio optimization tool, as defined by Reyck et al. (2005), while all other existing methodologies fail to account for two or more of the complexities inherent in the project portfolio optimization problem.

The remainder of this paper is structured as follows. In Section 2 we describe a mathematical representation for the technology portfolio management problem and study its complexity. In Sections 3 and 4 we discuss multistage and two-stage stochastic programming formulations, while in Section 5 an efficient solution procedure for the resulting problems is described. In Section 6, we present some computational results and Section 7 is the conclusion with a discussion of possible extensions.

2 Mathematical Representation

The project portfolio optimization problem can formally be defined as follows. Assume a set $\mathcal{N}$ of projects with annual performance levels $Z_i \in \mathbb{R}^+$, implementation times $\Delta_i \in \mathbb{R}^+$, required investment levels $\theta_i \in \mathbb{R}^+$, annual fixed activity costs $f_i \in \mathbb{R}^+$ and a set of depending technology projects $\mathcal{D}_i \subset \mathcal{N}$, for each $i \in \mathcal{N}$. Although only two-way dependencies between technology projects are used in this study, the proposed models can be extended to handle multi-way dependencies in a similar fashion. We let $Z_{ij} \in \mathbb{R}$ be the joint annual performance level for technology $i \in \mathcal{N}$ and $j \in \mathcal{D}_i$, and define it as a function of Z_i and Z_j. Furthermore, a sequence of investment planning periods $t = 1, 2, ..., T$ with available resource levels, i.e. budgets $B_t \in \mathbb{R}^+$, are assumed. For presentation purposes, the models in the paper are described for a single resource application, however extension to multiple resources is trivial. The objective is to determine an investment schedule such that some function of the total discounted return over an infinite time horizon is maximized while total investment in a given period t does not exceed B_t. In typical applications, the decision maker is interested in the investment schedule for the current period only, which should take into account future realizations of the parameters. Hence, a realistic assumption is that the

	Stochastic	Inter-dependencies	Organizational constraints	Project Selection	Resource Allocation	Complete Dynamic Reassesment
April et al. (2003)	✓			✓	✓	
Bardhan et al. (2006)	✓	✓	✓	✓		
Campbell (2001)	✓	✓		✓		
Chan et al. (2007)	✓		✓	✓		✓
Dickinson et al. (2001)		✓	✓	✓		
Elfes et al. (2005)			✓	✓	✓	
Ghasemzadeh et al. (1999)		✓	✓	✓		
Gustaffson&Salo (2005)	✓	✓		✓		✓
Lee et al. (2001)	✓	✓		✓		
Luenberger (1998)		✓	✓	✓		
Norkin et al. (1998)	✓		✓	✓	✓	
Sallie (2002)		✓	✓	✓		
Utturwar et al. (2002)		✓	✓	✓		
PROPOSED APPROACH	✓	✓	✓	✓	✓	✓

Figure 1: Summary of the existing literature on project portfolio optimization and contributions of the proposed methodology

problem will be solved each planning period to determine the best investment policy for that period, considering the past and future investments.

In practice, almost all of the above parameters may contain a certain level of uncertainty. However, in most applications, the level of variance is significant only in two of the parameters, namely the returns Z_i and required investment levels θ_i . Note that Z_{ij} is defined as a function of Z_i and Z_j. Hence, for modeling purposes, we approximate all other parameters with their expected values, and assume that joint and marginal probability distributions of the returns and required investment levels for the technologies are known or well estimated. Once a mathematical model is developed that accounts for the stochasticity in these two parameters, uncertainty in other parameters can be captured through what-if analyses.

The following complexity analysis shows that even the simplest instances of the project portfolio optimization problem fall into the category of NP-hard optimization problems.

Proposition 1. *Project portfolio optimization is NP-hard.*

Proof. Proof We first show that the deterministic version of the problem is NP-hard. The proof of NP-hardness is by restriction to the bin packing problem. Consider an instance of the project portfolio management problem in which $B_t = B$, $\theta_i + f_i \leq B$, $\Delta_i = 0$, and $\mathcal{D}_i = \emptyset$ for all $i \in \mathcal{N}$,$t = 1, 2, ..., |\mathcal{N}|$. Let $\mathcal{S}^*$ be the optimal schedule for this instance and let t^* be the latest investment period in $\mathcal{S}^*$. It is easily seen that $\mathcal{S}^*$ is optimal if and only if the optimal solution for an instance of the bin packing problem with bin capacities B and item sizes $\theta_i + f_i$ is t^*. It follows that stochastic version of the project portfolio optimization problem is also NP-hard. □

Given the uncertainty in the problem parameters of the project portfolio optimization problem, it is natural to assume that the decision maker would be interested in maximizing the expected value -or a function of the expected value- of total return. For presentation purposes, we assume a risk-neutral objective function throughout the rest of this paper. However, several other objectives that capture the risk attitude of the decision maker can be modeled and solved using the methods described in this study. Given any such objective, the project portfolio management problem can be expressed as:

$$\max_{\mathbf{x} \in \mathcal{X}} \{g(\mathbf{x}) = \mathbb{E}[G(\mathbf{x}, \xi)]\} \tag{1}$$

where $\mathbf{x}$ and ξ represent the vectors of decision variables and uncertain parameters (θ_i, Z_i), respectively. In addition, $\mathcal{X} \subset \mathbb{R}^n$ is the set of feasible solutions and $G(\mathbf{x}, \xi)$ is the total return function. Optimization problem (1) is difficult to solve, since exact evaluation of the expected value function in the objective is not possible.

A natural temptation to solve (1) may involve replacing the uncertain parameters by their expected values, and then solving the resulting so-called mean value problem, which is

$$\max_{\mathbf{x} \in \mathcal{X}} \{G(\mathbf{x}, \bar{\xi})\} \tag{2}$$

where $\bar{\xi} = \mathbb{E}[\xi]$ is the expectation of the random vector ξ. If $\bar{x}$ represents the optimal solution to (2), and x^* is the true optimal solution to the stochastic optimization problem (1), then clearly

$$\mathbb{E}[G(\bar{x}, \xi)] \leq \mathbb{E}[G(x^*, \xi)] \tag{3}$$

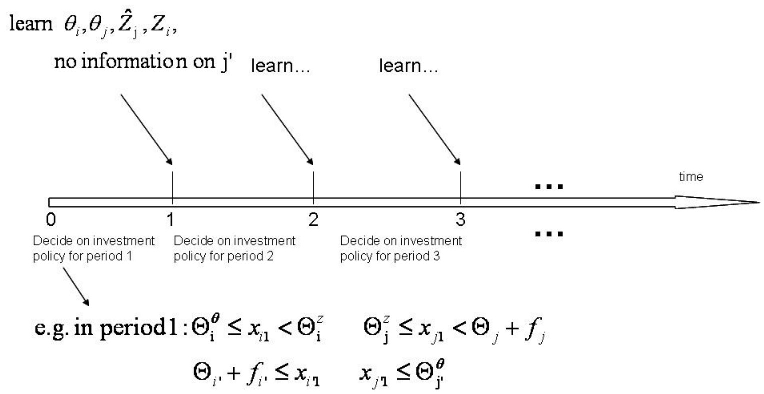

Figure 2: Decision process for the technology portfolio management problem, where realization of uncertainty is based on decisions made

The difference $\mathbb{E}[G(x^*, \xi)] - \mathbb{E}[G(\bar{x}, \xi)]$ measures how close the mean value solution is to the true solution, and is usually called the expected value of the stochastic solution (Birge and Louveaux 1997). However, the mean value problem usually does not reflect the decision process in a stochastic optimization problem correctly.

The decision process in the project portfolio management problem consists of recourse actions, by which the portfolio can be rebalanced at each period. Hence, an appropriate approach is to formulate problem (1) as a recourse problem, in which recourse actions can be taken after uncertainty is disclosed over the investment periods. In the following sections, we study two recourse models for the project portfolio management problem, and describe solution procedures for the two formulations.

3 The Multistage Stochastic Programming Model

The decision process in the project portfolio management problem consists of a multistage and multi-period structure, in which the goal is to determine an optimal allocation of the resources for the current planning period. However, the realization of uncertain parameters and the possibility of recourse actions in future periods must be accounted for in any optimal investment policy. Hence, resource allocations for the current period should position the decision maker in the best possible position against the uncertainties that will be realized in the future. The corresponding decision process for the project portfolio management problem can be described as follows, which is also represented in Figure 2, where examples of different investment levels leading to different information availability for projects i, i', j and j' are shown.

The resource requirement θ_i for each project i is known with certainty at the end of period t_θ^i, in which total investment in the project exceeds a threshold level Θ_i^θ, i.e. $t_\theta^i = \min_t\{t | \sum_{t' \le t} x_{it'} \ge \Theta_i^\theta\}$, where x_{it} represents the investment for project i in period t. Similarly, we assume that the uncertainty in the return of a project is revealed gradually over its development based on certain threshold levels. This process is modeled by assuming that an initial performance assessment $\hat{Z}_i$ will be available at the end of period $t_z^i = \min_t\{t | \sum_{t' \le t} x_{it'} \ge \Theta_i^z\}$ upon investing an amount of Θ_i^z in the project. As a result of this assessment, probabilities of different performance levels

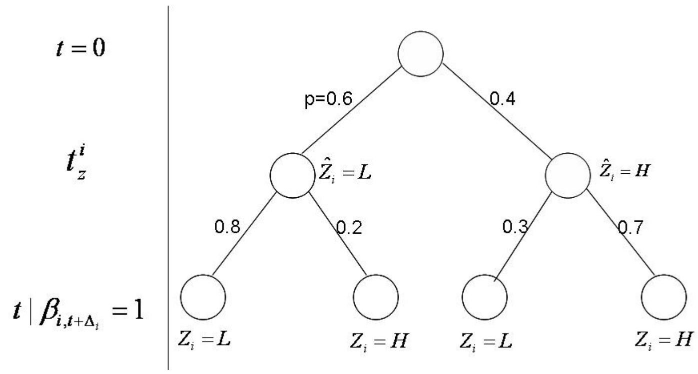

Figure 3: Tree showing gradual resolution of uncertainty in two phases

are updated. This assumption enables the modeling of the option of terminating a project if the initial assessment suggests that the probability of a high return is low for the project. Gradual resolution of uncertainty can be explained further as follows. Assume that Z_i can be realized at one of two levels: L, H with pre-development probabilities p_L and p_H, respectively. After investing an amount Θ_i^z in this project, an estimate $\hat{Z}_i$ is made, which can be seen as an intermediate realization of the uncertain parameter. If all uncertainty is resolved when technology development is over, then the probabilities for the actual realizations of the possible outcomes will be dependent on the intermediate realizations. This investment dependent probability distribution is described in Figure 3, where probabilities of possible Z_i values are updated according to the estimates $\hat{Z}_i$ which become available after investing Θ_i^z units of resources. If the development phase is continued, return Z_i will be known with certainty once all of the required resources are invested in project i. Multiple phases of gradual uncertainty resolution can be modeled by adding more layers to the described process, in the expense of adding more complexity to the stochastic problem.

The described process can be modeled as a multistage stochastic program, in which the uncertainty is in required investment levels, updated return estimates and final return levels. However, a complexity in this model is that the model contains endogenous uncertainty, i.e. realizations of the uncertain parameters are dependent on the investment decisions in current and future periods. Classical stochastic programming models assume that all stochastic processes in a given model are exogenous, which implies that the times of realizations of the uncertain parameters are not controlled by the decision maker, and the underlying scenario tree structure is known. However, this is not the case for the project portfolio optimization problem. Such problems are generally more difficult to formulate and solve than classical stochastic programming models, and there is very limited literature on such problems, which we discuss in Section 5.

As in many other stochastic programs, it is reasonable to assume for the project portfolio optimization problem that the random vector ξ has finite support or has a discrete distribution with K possible realizations, i.e. scenarios, $\xi^k := (\theta_i^k, \hat{Z}_i^k, Z_i^k)$, $k = 1, \dots, K$ with corresponding probabilities p_k. Then, it becomes possible to express problem (1) as one large mathematical program.

Before describing the mathematical model, we first introduce some new notation. In addition to the parameters described above, we let r be the discount factor throughout the planning period, $\mathcal{D}$ be the set of technologies that have a dependency relationship with another technology, i.e. $\mathcal{D} = \{i | i \in \mathcal{N}, \mathcal{D}_i \neq \emptyset\}$, and also set $\bar{\Delta} = \max_i\{\Delta_i\}$, $\bar{\Delta}_{ij} = \max\{\Delta_i, \Delta_j\}$. We also let $Y_{kk'}$ and $H_{kk'}$ be the set of technologies with different realizations of resource requirements and intermediate return estimates in scenarios k, k', respectively, i.e. $Y_{kk'} = \{i | \theta_i^k \neq \theta_i^{k'}\}$ and $H_{kk'} = \{i | \hat{Z}_i^k \neq \hat{Z}_i^{k'}\}$. Furthermore, we define the following decision variables for the problem, where the superscript k, which indicates that the variables are defined for each scenario, is omitted for clarity.

x_{it} :amount of investment in project i in period t, $t = 1, 2, ..., T$

τ_{it} :remaining required investment to complete the development of project i as of the end of period t, $t = 1, 2, ..., T$

y_{it} :1, if $t > t_\theta^i$, $t = 2, ..., T$; 0, otherwise

h_{it} :1, if $t > t_z^i$, $t = 2, ..., T$; 0, otherwise

α_{it} :1, if project i is started on or before period t, $t = 1, 2, ..., T$ 0, otherwise

β_{it} :1, if development and deployment of technology i are completed on or before period t, $t = \Delta_i, ..., T + \Delta_i$; 0, otherwise

γ_{it} :1, if project i is terminated prematurely in or before period t, $t = 2, ..., T$ 0, otherwise

δ_{ijt} :1, if development and deployment of dependent technologies i and j are completed on or before period t, $t = 1, 2, ..., T + \bar{\Delta}_{ij}$; 0, otherwise

This leads to the following multistage stochastic integer programming formulation:
Multistage Project Portfolio Optimization Problem(MPPM):

$$\max \sum_{k=1}^{K} p_k \sum_{i \in \mathcal{N}} \Big[\sum_{t \leq T-1} \beta^k_{i,t+\Delta_i} Z^k_i (1+r)^{-(t+\Delta_i)} + \beta^k_{i,T+\Delta_i} Z^k_i \big[\frac{(1+r)^{-(T+\bar{\Delta})}}{r} + \sum_{l=0}^{\bar{\Delta}-\Delta_i-1} (1+r)^{-(T+\Delta_i+l)} \big] + \sum_{\substack{j \in \mathcal{D}_i \\ j>i}} \Big[\sum_{t \leq T-1} \delta^k_{ij,t+\bar{\Delta}_{ij}} \widetilde{Z}^k_{ij} (1+r)^{-(t+\bar{\Delta}_{ij})} + \delta^k_{ij,T+\bar{\Delta}_{ij}} \widetilde{Z}^k_{ij} \big[\frac{(1+r)^{-(T+\bar{\Delta})}}{r} + \sum_{l=0}^{\bar{\Delta}-\bar{\Delta}_{ij}-1} (1+r)^{-(T+\bar{\Delta}_{ij}+l)} \big] \Big] \Big] \tag{4}$$

$$\alpha^k_{it} - \beta^k_{i,t+\Delta_i} \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{5}$$

$$\sum_{t' \leq t} x^k_{it'} - \Big(\max\{(\theta^k_i + t f_i), \max_{t' \leq t}\{B_{t'}\}\Big) \alpha^k_{it} \leq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{6}$$

$$\sum_{i \in \mathcal{N}} x^k_{it} \leq B_t \qquad \forall t \leq T, \forall k \tag{7}$$

$$x^k_{it} - B_t(\alpha^k_{it} - \beta^k_{i,t+\Delta_i-1} - \gamma^k_{it}) \leq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{8}$$

$$\beta^k_{i,t+\bar{\Delta}_{ij}} + \beta^k_{j,t+\bar{\Delta}_{ij}} - \delta^k_{ij,t+\bar{\Delta}_{ij}} \leq 1 \qquad \forall i \in D, \forall j \in D_i, j>i, \forall t \leq T, \forall k \tag{9}$$

$$\beta^k_{i,t+\bar{\Delta}_{ij}} + \beta^k_{j,t+\bar{\Delta}_{ij}} - 2\delta^k_{ij,t+\bar{\Delta}_{ij}} \geq 0 \qquad \forall i \in D, \forall j \in D_i, j>i, \forall t \leq T, \forall k \tag{10}$$

$$\tau^k_{it} - \tau^k_{i,t-1} + x^k_{it} - f_i(\alpha^k_{it} - \beta^k_{i,t+\Delta_i-1} - \gamma^k_{it}) \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{11}$$

$$x^k_{it} - f_i(\alpha^k_{it} - \beta^k_{i,t+\Delta_i-1} - \gamma^k_{it}) \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{12}$$

$$\tau^k_{it} + \theta^k_i \beta^k_{i,t+\Delta_i} \leq \theta^k_i \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{13}$$

$$\sum_{t'<t} x_{it'}^k - \Theta_{ik}^{\theta} y_{it}^k \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, t \neq 1, \forall k \tag{14}$$

$$\sum_{t'<t} x_{it'}^k - \left(\min\{\textstyle\sum_{t'<t} B_{t'}, (\theta_i^k + (t-1)f_i)\} - \Theta_{ik}^{\theta}\right) y_{it}^k \leq \Theta_{ik}^{\theta} \qquad \forall i \in \mathcal{N}, \forall t \leq T, t \neq 1, \forall k \tag{15}$$

$$\sum_{t'<t} x_{it'}^k - \Theta_{ik}^{z} h_{it}^k \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, t \neq 1, \forall k \tag{16}$$

$$\sum_{t'<t} x_{it'}^k - \left(\min\{\textstyle\sum_{t'<t} B_{t'}, (\theta_i^k + (t-1)f_i)\} - \Theta_{ik}^{z}\right) h_{it}^k \leq \Theta_{ik}^{z} \qquad \forall i \in \mathcal{N}, \forall t \leq T, t \neq 1, \forall k \tag{17}$$

$$\beta_{i,t+\Delta_i}^k + \gamma_{it}^k \leq 1 \qquad \forall i \in \mathcal{N}, \forall t \leq T, t \neq 1, \forall k \tag{18}$$

$$\alpha_{it}^k - \gamma_{i,t+1}^k \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{19}$$

$$x_{i1}^k - \sum_{k'=1}^{K} p_{k'} x_{i1}^{k'} = 0 \qquad \forall i \in \mathcal{N}, \forall k \tag{20}$$

$$x_{it}^k - x_{it}^{k'} + B_t \left[\sum_{j \in Y_{kk'}} (y_{jt}^k + y_{jt}^{k'}) + \sum_{j \in H_{kk'}} (h_{jt}^k + h_{jt}^{k'}) \right] \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, t \neq 1, \forall k, k' \tag{21}$$

$$x_{it}^k, \tau_{it}^k, \delta_{ijt}^k \geq 0 \qquad \forall i \in \mathcal{N}, \forall j \in D_i, j > i, \forall t, \forall k \tag{22}$$

$$\alpha_{it}^k, \beta_{it}^k, \gamma_{it}^k, h_{it}^k, y_{it}^k \in \{0,1\} \qquad \forall i \in \mathcal{N}, \forall t, \forall k \tag{23}$$

The objective function (4) in the above formulation assumes risk neutrality, and represents the expected total discounted return of the project portfolio. The total return is expressed as a function of the individual and joint returns depending on the deployment status of a developed technology. Joint return terms $\widetilde{Z}_{ij}$ are defined such that they represent the difference between the actual joint return contribution Z_{ij} and the sum of two individual returns. In other words, if two technologies are both implemented by period t, then the joint return contribution for that period is calculated as $Z_{ij} = Z_i + Z_j + \widetilde{Z}_{ij}$, where $\widetilde{Z}_{ij}$ can be positive or negative.

Constraint set (5) implies that project i must be started at least Δ_i periods before the corresponding technology is limited. Constraints (6) ensure that a positive investment must be made in order to start a project. Furthermore, (8) requires that an investment on a project can be made only if it is active, while (7) represents the resource limitations. Constraints (9)-(10) ensure that joint return from two dependent technologies is realized when the implementation of both technologies are complete. Constraints (11) calculate the required remaining investment for a technology development project in a given period, and (12) implies that the investment on a technology development project can not be less than the fixed cost incurred when the project is active. Constraints (13) ensure that a technology development project is complete only if the required remaining investment is 0. Furthermore, constraints (14)-(15) and (16)-(17) define indicator variables y_{it}^k and h_{it}^k, respectively. Constrains (18) state that a technology development project is either terminated successfully or unsuccessfully, while (19) ensures that a project is started before it is terminated.

In addition to the above, constraint set (20) represents the first stage nonanticipativity requirements, by ensuring that the decisions for the current period are the same for all scenarios. Notice that the nonanticipativity in other first stage variables are automatically satisfied if all first stage investment levels are the same. Since it is assumed that the uncertain variables are realized after certain levels of investment are made, a similar nonanticipativity structure must also be enforced between scenarios that share the same information history in later periods. In classical stochastic programming, nonanticipativity can explicitly be stated similar to (20), due to the exogenous nature of uncertainty in these problems. Since the uncertainty is endogenous in the project portfolio management problem, the nonanticipativity is conditional on the investment level decisions in each planning period. Constraints (21) capture this dependency by ensuring that a given pair of scenarios will be distinguished when one or more of the uncertain variables that distinguish them are revealed. The time of realization of the uncertainty is determined by the binary variables y_{it}^k and h_{it}^k. Notice that (21) are defined as inequalities for each possible pair of technologies so that if no distinguishing parameters are known, then the investment levels in the two technologies have to be equal. In addition, assuming independence of the corresponding probability distributions, any two scenarios that differ only in the realization of the final return values will have the same investment policy, since all investment decisions are made before these realizations. Hence, the return levels do not play a role in the nonanticipativity requirements. Representation of endogenous nonanticipativity in this compact way is distinct and more efficient than the existing models in the literature, since it enables the use of scenario decomposition methods as well as some other solution approaches proposed for classical multistage stochastic integer programming problems.

4 The Two-stage Stochastic Programming Model

Although the actual decision making process for the project portfolio optimization problem contains multiple stages, a natural simplification is through a two stage approach, in which it is assumed that a realization of the random variables becomes known after investment decisions are made for the current period in the first stage. If $\mathbf{x}_1$ represents the first period decision variables and $\mathbf{x}_2$ is the vector of variables for the second stage which contains the remaining $T-1$ periods, then the corresponding two stage stochastic program can be written as follows:

$$\max_{\mathbf{x}_1} \mathbb{E}[G(\mathbf{x}_1, \xi)] \tag{24}$$

$$\text{s.t. } A\mathbf{x}_1 = b, \mathbf{x}_1 \in \mathcal{X}_1 \tag{25}$$

where $G(\mathbf{x}_1, \xi)$ is the optimal value of the second stage problem

$$\max_{\mathbf{x}_2} g(\omega)^T \mathbf{x}_2 \tag{26}$$

$$\text{s.t. } T(\omega)\mathbf{x}_1 + W(\omega)\mathbf{x}_2 = h(\omega), \mathbf{x}_2 \in \mathcal{X}_2 \tag{27}$$

In the above representation, the second stage problem (26)-(27) depends on the realization ω of the random vector ξ, which determines the values of g, T, W and h.

This leads to the following two-stage stochastic integer programming formulation for the project portfolio management problem:

Two-stage Project Portfolio Optimization Problem (2PPM):

$$\max\ (4) \tag{28}$$

$$(5), (7), (9), (10), (13), (20), (22) \tag{29}$$

$$\sum_{t' \leq t} x_{it'}^k - \left(\max\{(\theta_i^k + tf_i), B_1\}\right) \alpha_{it}^k \leq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{30}$$

$$x_{it}^k - B_t(\alpha_{it}^k - \beta_{i,t+\Delta_i-1}^k - \gamma_{i2}^k) \leq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{31}$$

$$\tau_{it}^k - \tau_{i,t-1}^k + x_{it}^k - f_i(\alpha_{it}^k - \beta_{i,t+\Delta_i-1}^k - \gamma_{i2}^k) \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{32}$$

$$x_{it}^k - f_i(\alpha_{it}^k - \beta_{i,t+\Delta_i-1}^k - \gamma_{i2}^k) \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{33}$$

$$\beta_{i,t+\Delta_i}^k + \gamma_{i2}^k \leq 1 \qquad \forall i \in \mathcal{N}, \forall t \leq T, t \neq 1, \forall k \tag{34}$$

$$\alpha_{i1}^k - \gamma_{i2}^k \geq 0 \qquad \forall i \in \mathcal{N}, \forall t \leq T, \forall k \tag{35}$$

$$\alpha_{it}^k, \beta_{it}^k, \gamma_{i2}^k \in \{0,1\} \qquad \forall i \in \mathcal{N}, \forall t, \forall k \tag{36}$$

In terms of formulation, the two-stage problem differs from the multistage model only in the definition of the constraints that involve the termination variables, since early termination decisions can only be made in the second stage. Similar to the $MPPM$, if the number of scenarios K is not large, problem (29)-(36) can be solved using standard integer programming methods. However, this is not possible for realistic instances of the project portfolio management problem, since they constitute much larger problems. A difference between $MPPM$ and $2PPM$ in terms of problem size is that, the cardinality of the scenario set is smaller in $2PPM$, since gradual revelation of uncertainty is not modeled.

The same solution procedure we describe in Section 5.3 in a multistage setting can efficiently be utilized for $2PPM$. Except that the nonanticipativity is only restricted to the first stage, so the Lagrangian is given as

$$L(\beta, \delta, x, \lambda) = \hat{g}_N(\beta, \delta) + \sum_{l=1}^{N} \left[\sum_{i \in \mathcal{N}} \left(\sum_{l'=1}^{N} \frac{\lambda_i^{l'} x_{i1}^l}{N} - x_{i1}^l \lambda_i^l \right) \right] \tag{37}$$

which can be expressed as

$$L(\beta, \delta, x) = \sum_{l=1}^{N} L_l(\beta_l, \delta_l, x_l,) \tag{38}$$

where

$$L_l(\beta_l, \delta_l, x_l) = \hat{g}_N^l(\beta, \delta) + \sum_{i \in \mathcal{N}} \left(\sum_{l'=1}^{N} \frac{\lambda_i^{l'} x_{i1}^l}{N} - x_{i1}^l \lambda_i^l \right) \tag{39}$$

The corresponding Lagrangian dual problem for problem (29)-(36) is then

$$\min_\lambda \{D(\lambda) = \max\{\textstyle\sum_{l=1}^{N} L_l(\beta_l, \delta_l, x_l, \lambda_l) : (29) - (36), \text{except}(20)\}\} \tag{40}$$

Computational results and the efficiency of the solution procedure for $2PPM$ are discussed in Section 6.

5 An Efficient Solution Procedure for $MPPM$ and $2PPM$

There are very few studies on stochastic programming problems with endogenous uncertainty. Jonsbraten et al. (1998) is the first to address such problems, in which an algorithmic procedure to solve this type of two-stage problems is described. The proposed method includes a branch and bound scheme to determine an optimal vector of decisions, each of which has a corresponding scenario tree. Goel and Grossmann (2004b) model the operational planning of offshore gas field developments as a multistage stochastic program with endogenous uncertainty. The stages of the problem contain decisions to install production and well platforms, which result with the realization of the uncertain parameters for the fields in which installations are performed. The problem is formulated using disjunctions, and an approximation algorithm based on decomposition and restriction of the search space is described. A similar formulation is also given in Goel and Grossmann (2004a), in which a Lagrangian duality based branch and bound procedure is proposed to solve the problem. Held and Woodruff (2005) consider a network interdiction problem where the endogenous uncertainty is in the structure of the network. Stages of the problem contains interdiction decisions followed by shortest path calculations in the interdicted network. Several problem specific heuristic solution methods are described and compared in the study. More recently, Goel and Grossmann (2006) generalize the disjunctive programming formulation in Goel and Grossmann (2004b) to problems containing both exogenous and endogenous certainty. The authors also discuss a set of theoretical properties that leads to a reduction in the problem size. However, these results are only applicable to small size problems, since they are valid only when all possible scenarios are included in the problem. Viswanath et al. (2004) and Tarhan and Grossman (2006) consider somewhat different versions of the above class of problems. Viswanath et al. (2004) address a two-stage network problem, where in the first stage survival probabilities of arcs can be changed by investment decisions. Tarhan and Grossman (2006) consider gradual uncertainty revelation over time in the synthesis of process networks.

None of the above studies contain efficient solution procedures to solve problems with endogenous uncertainty, and almost all computational studies are performed on small size problems. The general disjunctive programming formulation and the solution suggested by Goel and Grossmann (2006) does not contain a direct decomposition structure, which is typically used in solving classical stochastic programming problems. In this study, we aim to fill this gap by developing a formulation scheme that is amenable to scenario decomposition, and is applicable to the general class of such problems. In addition, effective solution procedures for the resulting subproblems are also developed.

The sample average approximation (SAA) method is a Monte Carlo sampling technique that approximates a stochastic program by a smaller problem based on a random sample from the set of possible scenarios. Let $\xi^1, ..., \xi^N$ be an i.i.d. random sample of N realizations of the random vector ξ. Then the SAA problem for (1) is:

$$\max_{\mathbf{x}\in\mathcal{X}}\{\hat{g}_N(\mathbf{x}) = \frac{1}{N}\sum_{l=1}^{N} G(\mathbf{x}, \xi^l)\} \tag{41}$$

If v^* and $\hat{v}_N$ represent the optimal values of the "true" and SAA problems respectively, it is well known that $\hat{v}_N$ is a valid upper statistical bound for v^* . Furthermore, Shapiro (2003) shows that for multistage stochastic programming problems $\hat{v}_N$ converges to v^* with probability 1 as $N \to \infty$, although no result is available on the rate of convergence. Hence, the choice of large values of N will lead to better approximations of the true objective function. However, since the computational complexity of the SAA problem increases exponentially with the value of N, it is more efficient to select a smaller sample size N, and solve several SAA problems with i.i.d. samples.

Let M represent the number of SAA problems solved, and let $\hat{v}_N^m$ and $\hat{\mathbf{x}}_N^m$, $m = 1, \ldots, M$, denote the optimal objective value and solution of the mth replication, respectively. Since generally only the first stage investment decisions have practical importance for the project portfolio management problem, we assume that $\hat{\mathbf{x}}_N^m$ represents these first stage decisions. Once a feasible solution $\hat{\mathbf{x}}_N^m \in \mathcal{X}$ is obtained by solving the SAA problem, the objective value $g(\hat{\mathbf{x}}_N^m)$ can be approximated by the unbiased estimator

$$\hat{g}_{N'}(\hat{\mathbf{x}}_N^m) = \frac{1}{N'} \sum_{l=1}^{N'} G(\hat{\mathbf{x}}_N^m, \xi^l) \tag{42}$$

where N' is typically larger than N, since the computational effort required to estimate the objective value for a given solution is generally less than that required to solve the SAA problem. On the other hand, this phase may also be difficult for multistage problems, since it requires solving a multistage problem with endogenous uncertainty where only first stage decisions are known. Hence, any solution procedure must especially be efficient in calculating $\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)$. One would also want to estimate the quality of the solution $\hat{\mathbf{x}}_N^m$. This can be done by computing an estimate of the optimality gap $v^* - g(\hat{\mathbf{x}}_N^m)$, where $g(\hat{\mathbf{x}}_N^m)$ can be estimated by (42), and v^* can be approximated by

$$\bar{v}_N^M = \frac{1}{M} \sum_{m=1}^{M} \hat{v}_N^m \tag{43}$$

The sampling procedure can be terminated once the optimality gap estimate is sufficiently small or after performing all M replications, and the best solution among the SAA solutions can be selected using an appropriate criterion. However, the variance of the optimality gap estimator is also important, and must be taken into account in determining the quality of a solution. One option is to add a multiple z_α of the estimated standard deviation of the gap estimator to the gap estimator, where $z_\alpha = \Phi^{-1}(1-\alpha)$ and $\Phi(z)$ is the cumulative distribution function of the standard normal distribution (Kleywegt et al. 2002). If the sample sizes are not large, then z_α can be replaced by $t_{\alpha,\nu}$ from the t-distribution, where ν is the corresponding degrees of freedom. Then, an adjusted optimality gap estimator can be calculated by

$$\bar{v}_N^M - \hat{g}_{N'}(\hat{\mathbf{x}}_N^m) + z_\alpha \left(\hat{\sigma}^2_{\bar{v}_N^M} + \hat{\sigma}^2_{\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)} \right)^{1/2} \tag{44}$$

where $\hat{\sigma}^2_{\bar{v}_N^M}$ and $\hat{\sigma}^2_{\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)}$ are the estimates of the variances for the estimators of v^* and $g(\hat{\mathbf{x}}_N^m)$, respectively, and are calculated as

$$\hat{\sigma}^2_{\bar{v}_N^M} = \frac{1}{(M-1)M} \sum_{m=1}^{M} (\hat{v}_N^m - \bar{v}_N^M)^2 \tag{45}$$

$$\hat{\sigma}^2_{\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)} = \frac{1}{(N'-1)N'} \sum_{l=1}^{N'} \left(G(\hat{\mathbf{x}}_N^m, \xi^l) - \hat{g}_{N'}(\hat{\mathbf{x}}_N^m) \right)^2 \tag{46}$$

Effective implementation of the above sampling procedure requires that the SAA problems can be solved efficiently for relatively large values of the sample size N. For a given set of scenarios, (4)-(23) is a mixed integer programming problem and applications of standard solution methods fail to produce a solution even when N is set to values less than 10. As an efficient solution procedure for the SAA problem, we propose a Lagrangian relaxation and decomposition scheme coupled with an efficient lower bounding heuristic, which we name as the feasible dual conversion algorithm. The development of such a procedure is especially important, since for most multistage stochastic problems, even finding a feasible solution to serve as a lower bound is difficult. We show in Section 5.2 that the minimum feasible dual conversion heuristic is an effective procedure in calculating tight lower bounds for the technology portfolio management problem.

5.1 A Lagrangian Relaxation and Decomposition Scheme

Model (4)-(23) is linked in scenarios through the nonanticipativity constraints (20) and (21). Let $\hat{g}_N(\beta, \delta)$ represent the objective function (4) with K and p_k replaced by N and $1/N$, respectively. Then by subjecting the nonanticipativity conditions to Lagrangian relaxation, we form the following Lagrangian

$$L(\beta, \delta, x, y, h, \lambda, \mu) = \hat{g}_N(\beta, \delta) + \sum_{l=1}^{N} \sum_{i \in \mathcal{N}} \lambda_i^l \left[\sum_{l'=1}^{N} \frac{1}{N} x_{i1}^{l'} - x_{i1}^l \right] + \frac{1}{N} \sum_{l=1}^{N} \sum_{l' \neq l} \sum_{i \in \mathcal{N}} \sum_{1<t \leq T} \mu_{it}^{ll'} \left[x_{it}^l - x_{it}^{l'} + B_t \left[\sum_{j \in Y_{ll'}} \left(y_{jt}^l + y_{jt}^{l'} \right) + \sum_{j \in H_{ll'}} \left(h_{jt}^l + h_{jt}^{l'} \right) \right] \right] \tag{47}$$

where λ_i^l and $\mu_{it}^{ll'}$ are the Lagrange multipliers. Notice that the formulation of the nonanticipativity constraints (20) and the multiplication of the relaxed constraints (21) by $\frac{1}{N}$ in the above Lagrangian account for the scenario probabilities, and prevent the ill-conditioning in the Lagrangian dual as discussed by Louveaux and Schultz (2003). A major advantage of the described formulation of the nonanticipativity constraints is that when they are relaxed, the Lagrangian (47) can be decomposed by scenarios for given dual vectors λ and μ , and can be expressed as

$$L(\beta, \delta, x, y, h) = \sum_{l=1}^{N} L_l(\beta_l, \delta_l, x_l, y_l, h_l) \tag{48}$$

where

$$L_l(\beta_l, \delta_l, x_l, y_l, h_l) = \hat{g}_N^l(\beta, \delta) + \sum_{i \in \mathcal{N}} \left[\sum_{l'=1}^{N} \frac{\lambda_i^{l'} x_{i1}^l}{N} - x_{i1}^l \lambda_i^l \right] + \frac{1}{N} \sum_{i \in \mathcal{N}} \sum_{1<t \leq T} \left[x_{it}^l \sum_{l' \neq l} \left(\mu_{it}^{ll'} - \mu_{it}^{l'l} \right) + B_t \left[\sum_{l' \neq l} \sum_{j \in Y_{ll'}} y_{jt}^l \left(\mu_{it}^{ll'} + \mu_{it}^{l'l} \right) + \sum_{l' \neq l} \sum_{j \in H_{ll'}} h_{jt}^l \left(\mu_{it}^{ll'} + \mu_{it}^{l'l} \right) \right] \right] \tag{49}$$

The corresponding Lagrangian dual problem for problem (4)-(23) is then

$$\min_{\lambda, \mu} \{ D(\lambda, \mu) = \max \{ \textstyle\sum_{l=1}^{N} L_l(\beta_l, \delta_l, x_l, y_l, h_l, \lambda_l, \mu_l) : (5) - (19), (22), (23), \mu_l \geq 0 \} \} \tag{50}$$

Problem (50) is a nonsmooth convex minimization problem which can be solved by subgradient optimization methods (Hiriart-Urruty and Lemarechal 1993). At each iteration of these methods, the solution of $D(\lambda, \mu)$ is required to obtain a subgradient. Notice that $D(\lambda, \mu)$ is separable, and reduces to the solving N problems of manageable size, each of which corresponds to a single scenario. Components of the subgradient vector are then given by $\frac{\lambda_i^{l'} x_{i1}^l}{N} - x_{i1}^l \lambda_i^l$ and $x_{it}^l \sum_{l' \neq l} \Big(\mu_{it}^{ll'} - \mu_{it}^{l'l} \Big) + B_t \Big[\sum_{l' \neq l} \sum_{j \in Y_{ll'}} y_{jt}^l \left(\mu_{it}^{ll'} + \mu_{it}^{l'l} \right) + \sum_{l' \neq l} \sum_{j \in H_{ll'}} h_{jt}^l \left(\mu_{it}^{ll'} + \mu_{it}^{l'l} \right) \Big]$, where x_{i1}^l, y_{it}^l and h_{it}^l are the optimal solutions to the scenario subproblems.

For the project portfolio optimization problem, we propose a modified subgradient algorithm, in which step sizes in updating the dual variables are determined according to a weighted combination of the subgradients from previous iterations. More specifically, a new step direction at iteration j is determined by

$$\hat{\Gamma}^j = \pi_0\Gamma^j + \pi_1\Gamma^{j-1} + \pi_2\Gamma^{j-2} + \pi_3\Gamma^{j-3} \tag{51}$$

where Γ terms represent the subgradients and π terms are weights such that $\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1$. Individual π values can be selected according to an experimental analysis based on the problem considered. Updates of the multipliers are then performed using the following combined dynamic step size rule:

$$\lambda^{j+1} = \lambda^j - \max\{\frac{\phi}{j}, \frac{\kappa(\bar{L}^j - \underline{L}^j)}{||\hat{\Gamma}^j||}\}\hat{\Gamma}^j \tag{52}$$

$$\mu^{j+1} = \max\{0, \mu^j - \max\{\frac{\phi}{j}, \frac{\kappa(\bar{L}^j - \underline{L}^j)}{||\hat{\Gamma}^j||}\}\hat{\Gamma}^j\} \tag{53}$$

where ϕ and κ, $\kappa < 2$, are constants that can be modified during the algorithm. Above rule, which has been verified through computational studies, ensures that initial step sizes are large enough to prevent early convergence to a non-optimal solution. The implementation of the overall solution algorithm includes frequent lower bound calculations during the iterations of the subgradient method, and the convergence rate of the subgradient algorithm is especially important from an overall computational perspective. Hence, the stepsizes are determined as efficiently as possible to improve the convergence rate of the algorithm. Despite the large size of the dual vector for realistic instances of the problem, computational studies have shown that the convergence of the subgradient algorithm is relatively fast. Results of the tested models are discussed in Section 6.

It is well known that, due to the integrality requirements, the optimal solution of the Lagrangian dual gives an upper bound for the objective value of (4)-(23), which is at least as tight as the bound obtained from the LP relaxation of the problem. Furthermore, any Lagrangian dual solution is an upperbound for the original problem. However, a major difficulty in solving multistage stochastic programming problems is to determine good feasible solutions for tight lower bounds. Clearly, except in rare cases, the solutions of the Lagrangian dual will not satisfy the nonanticipativity constraints.

We present a heuristic procedure that uses the Lagrangian dual solutions in subgradient iterations to search for a feasible solution to the primal problem, which provides a lower bound for the optimal objective value. Given a Lagrangian dual solution, the method looks for a primal solution with minimum deviation from the dual solution. The search, which has produced very tight bounds in the computational studies described in Section 6, is implemented using integer programming models of manageable size. To ease the computational difficulty, the procedure is implemented gradually using subsets of scenarios, which are determined by the variable values and the objective value contributions of the scenarios in the dual solution. This procedure, which can also be applied as a bounding procedure in similar stochastic programming problems, is described in detail below.

5.2 The Feasible Dual Conversion Algorithm

The objective function for the project portfolio optimization problem is defined by the values of the binary variables β_{it}, which represent the periods that the return realizations begin. Hence, the corresponding values in a given Lagrangian dual solution describe some infeasible investment policy in which nonanticipativity constraints are not enforced but are only penalized. Clearly, the optimal objective value of the primal problem is expected to be as close as possible or comparable to that of this infeasible policy. Although, due to the combinatorial nature of the problem, the optimal investment policy in the presence of nonanticipativity can be significantly different than the policy suggested by the given dual solution, one can obtain a "good" investment policy by converting the dual solution into a feasible solution by a minimal change in the β_{it} values in the Lagrangian dual solution. We present below an algorithm to achieve this, as well as a bound on the quality of the solution obtained through the algorithm. The feasible dual conversion algorithm performs such conversions in a systematic way that ensures the quality of the resulting solution as well as computational efficiency.

Algorithm 1 (Feasible Dual Conversion)**.** *The steps of the algorithm are as follows:*

Step 1. Initialization : Let β^j *represent the vector of corresponding values in a solution to the Lagrangian dual problem* (50) *at iteration* j *of the subgradient algorithm for dual variables* λ^j *and* μ^j*. Let* $\underline{\beta}^l_{it}$*,* $\underline{\hat{g}}_N$*,* $\underline{L}_l$ *be the lowerbounds on* β^l_{it}*,* $\hat{g}_N$ *and* L_l *for scenario* l*. Choose a scenario subset size* S*. Set* $\underline{\beta}^l_{it} = 0$ *for all* i, t, l*,* $\mathbb{S} = \emptyset$*,* $\mathbb{S}' = \emptyset$*,* $\mathbb{N} = \{l_1, l_2, \ldots, l_N\}$*.*

Step 2. Scenario subset selection : Rank all $s \in \mathbb{N}$ *according to scenario objectives* L^j_s*, and form subset* $\mathbb{S}$ *by selecting the first* S *scenarios among the ranked scenarios in* $\mathbb{N}$*. Let* $\mathbb{S}' = \mathbb{S}' \cup \mathbb{S}$ *and* $\mathbb{N} = \mathbb{N} \backslash \mathbb{S}$*.*

Step 3. Variable fixing : For each $s \in \mathbb{S}$*, determine period* t^s_o *in which* s *becomes distinguishable from all other scenarios according to scenario solutions* β^s_{it}*, i.e.*

$$t^s_o = \min_t \{t | \min_{s' \neq s} \{ \sum_{j \in Y_{ss'}} (\beta^s_{j,t+\Delta_j} + \beta^{s'}_{j,t+\Delta_j}) + \sum_{j \in H_{ss'}} (\beta^s_{j,t+\Delta_j} + \beta^{s'}_{j,t+\Delta_j}) \} \geq 1 \} \tag{54}$$

For each $i \in \mathcal{N}$ *such that* $\beta^s_{i,t+\Delta_i} = 1$*, and* $t \leq t^s_o$*; if* $\beta^s_{i,t+\Delta_i} - \beta^s_{i,t+\Delta_i - 1} = 1$*, then set* $\underline{\beta}^s_{i,t+\Delta_i} = 1$*.*

Step 4. Feasibility determination: Check feasibility of (4)-(23) *with the lower bounds on* β^s_{it} *for the scenario set* $\mathbb{S}'$*. If feasible, let* $\dot{\beta}^s_{it}$ *represent the corresponding values in this solution, and fix* $\beta^s_{it} = \dot{\beta}^s_{it}$*. If* $\mathbb{N} \neq \emptyset$*, go to Step 2.*

Step 5. Minimum dual conversion : If (4)-(23) *is infeasible, determine the minimum number of relaxations* r_o *required on* $\underline{\beta}^s_{it} = 1$ *for* $s \in \mathbb{S}$ *to obtain a feasible solution. Find the best possible feasible solution that can be achieved by relaxing at most* r_o *of the bounds* $\underline{\beta}^s_{it}$*. Fix* $\beta^s_{it} = \dot{\beta}^s_{it}$*. If* $\mathbb{N} \neq \emptyset$*, go to Step 2.*

Step 6. Bound calculation : Let $\dot{\mathbf{x}}$ *and* $\dot{g}_N$ *represent the final solution vector and objective function value. If* $\dot{g}_N > \underline{\hat{g}}_N$*, set* $\underline{\hat{g}}_N = \dot{g}_N$*. For each scenario* l*, calculate* $\dot{L}_l(\dot{\mathbf{x}}, \lambda^{j+1}, \mu^{j+1})$*. If* $\dot{L}_l > \underline{L}^{j+1}_l$*, set* $\underline{L}^{j+1}_l = \dot{L}_l$*.*

After the initialization of the algorithm in Step 1 according to a Lagrangian dual solution obtained in a subgradient iteration, Step 2 identifies the scenarios with the maximum possible contribution to the total expected return. In Step 3, projects that determine nonanticipativity relationships and that are also likely to deviate from the Lagrangian solution are identified. The β_{it} variables for these projects are fixed so that they are completed on or before the time suggested by the ideal policy from the dual solution. Almost in all cases, this will lead to an infeasible solution, which is checked in Step 4. Then, a conversion procedure is implemented in Step 5. In this phase, first the minimum number of relaxations on the fixed β_{it} variables required to obtain a feasible solution is determined by solving an integer programming problem, which is assumed to be easily solvable for scenario subset size S. Note that such a feasible solution always exists. Another option is to minimize a weighted sum of the relaxations, where the weights are determined by the contribution of each technology into the overall objective function. Then, given this minimum requirement for feasibility, an optimization is performed to determine the best possible solution by performing at most that many relaxations on fixed β_{it} variables. Again, it is assumed that such an optimization can be performed efficiently for S scenarios. The procedure is repeated $\frac{N}{S}$ times, which results with a feasible solution for the primal problem. In Step 6, bounds on the objective values are updated to simplify the solution process in later iterations. Indeed, in the overall implementation, a history of all such solutions are maintained, and used to determine the best possible lowerbound on scenario subproblems at each iteration. Despite the additional memory requirement, it has been observed that this significantly reduces the solution times for the scenario subproblems.

One may think that a better approach would be such that all β_{it} values in the Lagrangian dual solution are fixed in Step 3. However, this may significantly increase the computational complexity of the optimization problems solved in Step 5. Also, by minimizing the number of required relaxations, Step 6 minimizes the computational difficulty of the subsequent optimization problem, and the deviation from the dual solution is kept minimal with respect to technologies with the highest return levels. The following propositions define a bound on the quality of the solution produced by the feasible dual conversion algorithm, which translates to an upper bound on the duality gap.

Proposition 2. *Let i^s represent a project i in scenario s, and let I_β be the set containing all i^s such that $\underline{\beta^s_{it}} = 1$ for some t. For a given set $\mathbb{S}$ of scenarios, group i^s according to the order of completion in the dual scenario solutions, i.e. projects completed first in each scenario represent a group, as well as those completed second, third, etc. In case of ties, assign groups arbitrarily. Let R^n, $n \leq |\mathcal{N}|$, represent the cardinality of the largest compatibility set in group n, where projects i^s and $j^{s'}$ are defined to be in the same compatibility set if $\underline{\beta^s_{i,t+\Delta_i}} = \underline{\beta^{s'}_{j,t+\Delta_j}} = 1$, $\beta^s_{i,t+\Delta_i} = 1$ does not imply $\beta^{s'}_{j,t+\Delta_j} = 0$ or vice versa, and if they are compatible with the projects in the maximum cardinality compatibility set in group $n-1$. Then, for any application of the feasible dual conversion algorithm on $\mathbb{S}$,*

$$r_o \leq |I_\beta| - \sum_{n \leq |\mathcal{N}|} R^n$$

Proof. Proof Clearly, an upper bound on r_o is $|I_\beta|$. Note that, to obtain feasibility, a relaxation of the lower bound on β^s_{it} or $\beta^{s'}_{jt}$ is required if i^s and $j^{s'}$ are not compatible. Hence, required number of relaxations for each group will be minimum if β^s_{it} is set to 1 for all members of the maximum cardinality compatibility set, and the variables corresponding to the remaining projects in the group are relaxed. By the definition of compatibility, a feasible solution can always be obtained by fixing $\sum_{n \leq |\mathcal{N}|} R^n$ of the β^s_{it} variables, where $i^s \in I_\beta$, at their lowerbounds. Hence an upper bound on the number of relaxations required for feasibility is $|I_\beta| - \sum_{n \leq |\mathcal{N}|} R^n$. □

The above bound on the number of relaxations is easy to calculate, since the size of the groups formed in the bound calculation procedure is in the order of S. The procedure requires the identification of the maximum cardinality compatibility set, which is equivalent to solving the NP-hard maximum clique problem on a compatibility graph. As noted, the size of the groups enable easy determination of this set. On the other hand, less tight bounds can be obtained by using bounds known for the maximum clique problem and selecting a clique arbitrarily to fix some of the variables. Proposition 3 uses the bound on r_o to develop a bound for the quality of the solutions obtained by the feasible dual conversion algorithm.

Proposition 3. *Consider a ranking of projects $i^s \in I_\beta$, i.e. $< i^s_{(1)}, i^s_{(2)}, \cdots >$ such that $z_{i^s_{(1)}} \geq z_{i^s_{(2)}} \geq \ldots$, where z^s_i is the contribution of project i to the scenario objective in the dual solution. Define $r^U_o = |I_\beta| - \sum_{n \leq |\mathcal{N}|} R^n$, and let $t^s_{o'}$ represent the period that scenario s becomes distinguishable from all other scenarios according to a modified dual solution obtained by assuming no investment is made in project i^s prior to period $t+1$, if $\underline{\beta^s_{it}} = 1$ and $i^s \in \{i^s_{(1)}, i^s_{(2)}, \ldots, i^s_{(r^U_o)}\}$ or $\underline{\beta^s_{it}} = 0$. Furthermore, assume that z^s_c represents the return in scenario s from the optimum single-scenario investment schedule over periods $t^s_{o'}, \ldots, T$ for all projects that are not completed by $t^s_{o'}$ according to the modified dual solution. For the optimum partial schedule calculations, assume that for i^s such that $\underline{\beta^s_{it}} = 0$ for all t, $\theta_i = \tau_{i,t^s_{o'}}$, if $\tau_{i,t^s_{o'}} < \theta_i$ in the modified dual solution and all x^s_{it} satisfy modified nonanticipativity for $t \leq t^s_{o'}$. If $F^*_N(x)$ is the optimal objective function value for the SAA problem with N scenarios, and $F_N(\dot{x})$ is the objective value of a solution generated by the feasible dual conversion algorithm,then*

$$F^*_N(x) - F_N(\dot{x}) \leq \sum_{k=1}^{N/S} \sum_{s \in \mathbb{S}^k} \{-z^s_c + \sum_{i=i^s_{(1)}}^{i^s_{(r^U_o)}} z^s_i\}$$

Proof. Proof Consider the first iteration of the feasible dual conversion algorithm, and assume that S scenarios with highest scenario objectives are selected. Notice that an upperbound for the contribution of these scenarios in the optimal solution is given by $\sum_{s=1}^{S} L_s(x, \lambda, \mu)$. Let ΔZ represent the total change in the objective value of the feasible solution for scenario s compared with the dual solution. Clearly, $\Delta Z \leq \sum_{i=i^s_{(1)}}^{i^s_{(r_o)}} z^s_i$, since a feasible solution always exists with r_o relaxations on the bounds $\underline{\beta^s_{it}} = 1$. Without loss of generality, assume that these relaxations correspond to i^s with the highest contributions to the objective function. We show that the modified dual solution described above is feasible. Suppose this solution is not feasible, which implies that the corresponding investment schedule does not satisfy the modified nonanticipativity requirements. Since the modified dual solution consists only of projects with $\underline{\beta^s_{it}} = 1$, any change in the schedule would require a relaxation in these bounds. This contradicts with the condition that a feasible solution exists with r_o relaxations on the bounds. Furthermore, any partial investment schedule for periods after $t^s_{o'}$ would not violate feasibility, since there is no nonanticipativity requirements after period $t^s_{o'}$. Hence, it is possible to improve this feasible solution by reoptimizing the allocations in each scenario s for periods after $t^s_{o'}$. This will lead to an improvement of $\sum_s z^s_c$ in the objective value, implying that $\Delta Z \leq \sum_{i=i^s_{(1)}}^{i^s_{(r_o)}} z^s_i - \sum_s z^s_c$. It follows from Proposition 2 that the bound can be expressed similarly by replacing r_o with r^U_o. Since the algorithm performs N/S iterations to obtain a feasible solution, the total difference is the sum over all iterations, and the result follows. □

Calculation of the above bound requires the solution of small optimization problems for each scenario. These problems include only a subset of the projects in the portfolio, and contain periods after $t^s_{o'}$. Noting that these small problems can be solved significantly fast, the difficulty of bound calculations is only dependent on the number of scenarios considered.

Using the bounding schemes discussed, a branch and bound algorithm with branching on the nonanticipativity constraints that are not satisfied in the solution of the Lagrangian dual can be implemented to close the duality gap. In the case of the project portfolio management problem, the nonanticipativity constraints are on the continuous variables x_{it}^k. Hence a branching rule could use the average investment in the scenario solutions of the dual problem, or the most frequent occurrence of x_{it}^k values to branch on. However, the branch and bound scheme is usually computationally efficient only for very small scale problems. On the other hand, duality gaps are not significantly high for the approximate solutions produced by the feasible dual conversion algorithm for larger models as noted in Tables 3 and 4. Thus, in most instances, it will suffice to obtain approximate solutions through the feasible dual conversion algorithm, and use them as the solutions to the SAA problems. In parallel with this analysis, computational studies in Section 6 have been implemented without the branch and bound step for efficiency purposes.

5.3 Solution Algorithm Overview

The overall procedure to solve the project portfolio optimization problem is summarized below, which is also shown in Figure 4.

Algorithm 2 (Solution Algorithm for $MPPM$ and $2PPM$)**.** *The general solution algorithm can be summarized as follows:*

Step 1. Obtain N samples from the set of scenarios, and form the SAA problem with these scenarios.

Step 2. Perform Lagrangian relaxation on the SAA problem, decomposing the problem into individual scenario subproblems.

Step 3. Use subgradient algorithm with the proposed step size measure to obtain an upper bound for the SAA problem.

3a. If computationally feasible, solve the LP relaxation of (4)-(23)*, and set the corresponding dual values as the initial Lagrangian multipliers. Use a rounding heuristic to obtain an initial lowerbound on the problem, i.e if $\beta_{it}^k \geq 0.5$ and $\beta_{it'}^k \geq 0.5$ for all $t' > t$ in the LP relaxation solution, then set $\dot{\beta}_{it}^k = 1$, else set $\dot{\beta}_{it}^k = 0$. Then use the feasible dual conversion algorithm.*

3b. At each iteration j of the algorithm, determine a lowerbound for the scenario subproblems by calculating $\dot{L}_l(\mathbf{x}^l, \lambda^{j+1}, \mu^{j+1})$, and selecting the minimum.

3c. Based on an improvement threshold for the dual solution or at every f_o iterations, apply the feasible dual conversion algorithm, to obtain a lowerbound for the SAA problem, as well as for the scenario subproblems.

3d. Use the best lowerbounds for the scenario subproblems as the starting solution for the subproblems at iteration $j+1$.

4. Calculate the duality gap upon convergence of the subgradient algorithm. If the gap is less than or equal to ϵ, go to step 5. Else, if computationally feasible, use branch and bound to close the duality gap, by branching on the nonanticipativity conditions.

5. Repeat Steps 1-4 M times. Each solution is a candidate solution for the true problem.

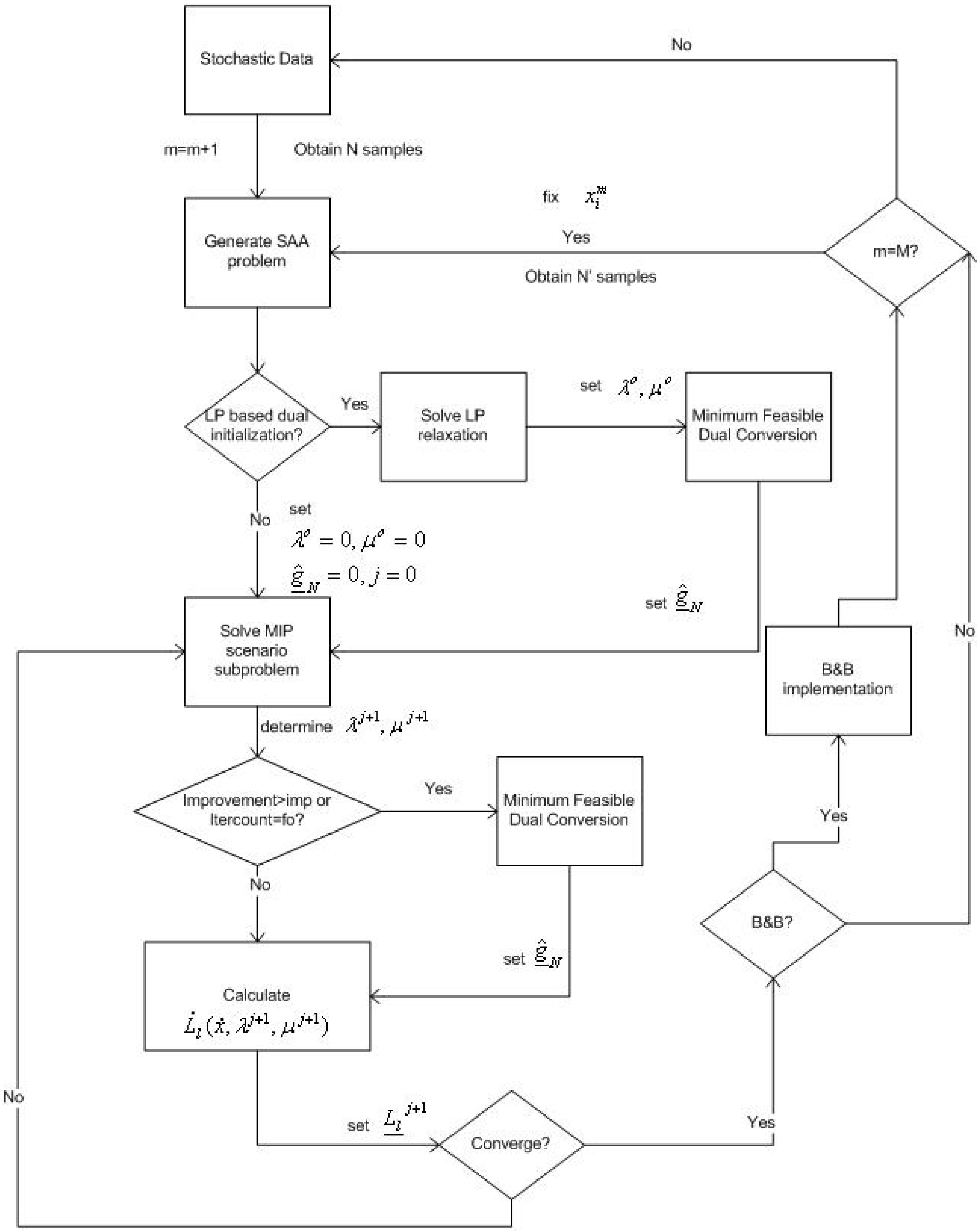

Figure 4: Solution algorithm for $MPPM$ and $2PPM$

6. For some or all of the candidate solutions, perform N' replications by fixing the values of the first stage variables according to the solution, and repeating steps $1-4$ with these fixed values to estimate the objective value of the candidate solutions.

7. Select a solution as the best solution using an appropriate criterion.

For the lower bounding procedure, both the LP relaxation based and dual solution based heuristics can be applied and the maximum objective value can be selected as the better lowerbound. Computational studies have shown that the LP relaxation based heuristic can often produce good solutions.

6 Computational Results for $MPPM$ and $2PPM$

Computational tests for the developed solution procedures were conducted on two sets of project portfolio data under different algorithmic configurations. The data sets consist of five and ten technology projects and are represented as 5T and 10T in the results tables. The stochastic data for the ten project instance is shown in Table 1. The probability distributions for the uncertain parameters, i.e. required investment levels, initial return estimates and realized return levels, were assumed to be discrete with low and high levels. Corresponding probabilities for each case are also listed in Table 1. The probability distributions for the uncertain parameters, i.e. required investment levels, initial return estimates and realized return levels, were assumed to be discrete with low and high levels. Although the dependence of the probability distributions of return estimates and realizations are modeled to reflect a gradual resolution of uncertainty, all other stochastic parameters are assumed to be independent. Joint return effects are defined according to the description in Section 2. Several implementations with varying sample sizes and number of replications are displayed in Tables 3 and 4. The number preceding the letter S in the table notation represents the number of samples, while the number preceding the letter R is the number of replications.

Computations were performed on a PC with an Intel Core 2 Duo 2.0 GHz processor and 2GB of internal memory, using ILOG CPLEX Version 10.0. Although the computational studies were conducted on a single computer, the proposed solution procedure can easily be parallelized by solving the scenario subproblems on multiple machines to improve the solution times significantly.

The first two columns after the problem size information in Tables 3 and 4 display the time in seconds per replication of the SAA implementation and the expected value estimation for a given solution, respectively. The next column is the average duality gap, which is an average of the gap over all replications. The adjusted optimality gap estimate is given in the last column, and is calculated according to (44), based on the best solution obtained using the developed procedure. The sample size N' to estimate the corresponding objective value of a candidate solution was selected as 100 and 50 for 5T and 10T implementations. As it is shown in these results tables, the calculations of the objective values when the first stage decisions are fixed can be performed significantly faster then the solution of the SAA problem. Table 2 displays the first stage solutions for all tested configurations of the SAA algorithm. In most cases, different configurations return the same solution, based on the methodology used to select the best solution among the candidate solutions. Furthermore, for these instances, the two stage and multi-stage solutions are not significantly different than each other.

Overall, the computational results show that the developed procedure is effective and efficient in solving the project portfolio optimization problem, which is a difficult multistage stochastic program with endogenous uncertainty. Even without the implementation of a branch and bound procedure to close the duality gap, obtained lower bounds are very close to the Lagrangian upper bounds. As expected, the duality gap is less in instances with small sample sizes, while the optimality gap estimate is very low for large sample sizes. For the latter case, the variances are much lower and convergence of $\bar{v}_N^M$ and $\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)$ occur significantly faster, in the expense of slower computation times. The selection of the best solution out of several SAA solutions was done in two steps. In the first step, candidate solutions were identified based on the frequency of occurrences in the SAA solutions. Then the expected returns were estimated for these candidate solutions as described above, and the solution with the highest expected return estimate was selected. In Figure 5, we show the different levels of variance and convergence in this process on the 5T instances for both $MPPM$ and $2PPM$. The horizontal line in each plot represents the value of the estimate $\bar{v}_N^M$ for the corresponding algorithmic configuration. The effects of large sample sizes are evident in these plots, as it can be seen that convergence to the corresponding objective value is much faster in these cases. In addition, when compared with the two-stage model, convergence is better in the multi-stage case, mainly due the flexibility in a multistage model in rebalancing the portfolio in later stages. Hence, the results for different scenarios do not vary significantly.

7 Conclusions and Future Work

Project portfolio optimization problem has not been studied at the detailed level considered in this study before. It was also noted that the problem has a unique structure with endogenous uncertainty of the stochastic parameters, and development of a solution methodology would also contribute to the general class of such problems. We have presented a detailed and comprehensive description of the problem, the solution characteristics, and an efficient solution approach that can be used to solve this large-scale problem.

Implementation of the proposed models in project portfolio selection by organizations will lead to significant increases in returns, as all relevant inputs and uncertainty are captured in the models, as opposed to existing project portfolio selection tools. The developed methodology is in the process of being used by the Federal Aviation Administration (FAA) in determining resource allocations to a portfolio of aviation modernization technologies. A significant contribution of the developed models is that they include a common but less studied characteristic of endogenous uncertainty. Problems of this type are usually difficult to model, since the nonanticipativity conditions require comparisons of scenario pairs. We present a compact decomposable structure which can be exploited by methods that are commonly used in the solution of classical stochastic programming problems. It must be noted that even if the endogenous uncertainty were to be ignored, the resulting problem would be a multistage stochastic integer program with several stages for which no general solution procedures are available. Hence, to handle the difficulty, an effective lower bounding algorithm and performance bounds have been developed as a part of the overall solution procedure. The algorithm has been tested with promising results, and it is believed that such a procedure can be implemented in several other similar problems. Additional extensions of the study are possible in several areas. Integration of risk is an important part of the technology portfolio selection, since most practical decisions are made while considering risks associated with the investment decisions. This can be analyzed through the introduction of other objective functions capturing risk, such as value-at-risk models. One other extension includes capturing the effects of dependencies in probability distributions on the investment decisions.

Table 1: Data for the ten project test instance of stochastic project portfolio optimization problem

Attributes / Projects	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Fixed activity cost (mil.$)	0.2	0.1	0.3	0.2	0.2	0.05	0.1	0.2	0.05	0.3
Min. req. inv.(mil.$) / probability	2 / 0.35	3 / 0.3	4 / 0.5	2 / 0.4	1 / 0.5	1 / 0.6	5 / 0.5	1 / 0.5	2 / 0.55	1 / 0.3
Max. req. inv.(mil.$) / probability	4 / 0.65	5 / 0.7	6 / 0.5	6 / 0.6	3 / 0.5	3 / 0.4	7 / 0.5	1 / 0.5	4 / 0.45	3 / 0.7
Implementation time (yrs)	5	2	3	3	2	3	1	4	2	2
Min. initial return est. (mil.$) / prob.	1.5 / 0.6	1 / 0.4	1.5 / 0.14	0 / 0.225	0 / 0.5	2 / 0.25	3 / 0.83	2 / 0.67	1.5 / 0.5	1 / 0.5
Max. initial return est. (mil.$) / prob.	4.5 / 0.4	3.5 / 0.6	4.5 / 0.86	3.5 / 0.775	1 / 0.5	5 / 0.75	7 / 0.17	3 / 0.33	4/ 0.5	4.5 / 0.5
Prob. of min. return after low initial real.	*0.8*	*0.8*	*0.8*	*0.6*	*0.5*	*0.7*	*0.5*	*0.4*	*0.7*	*0.6*
Prob. of max. return after low initial real.	*0.2*	*0.2*	*0.2*	*0.4*	*0.5*	*0.3*	*0.5*	*0.6*	*0.3*	*0.4*
Prob. of min. return after high initial real.	*0.3*	*0.3*	*0.3*	*0.2*	*0.5*	*0.3*	*0.2*	*0.1*	*0.5*	*0.4*
Prob. of max. return after high initial real.	*0.7*	*0.7*	*0.7*	*0.8*	*0.5*	*0.7*	*0.8*	*0.9*	*0.5*	*0.6*
Dependent	-	C	B	E	D	-	J	-	-	G
Joint effect after min-min return (mil.$)	-	-0.5	-0.5	0	0	-	0.5	-	-	0.5
Joint effect after min-max return (mil.$)	-	-1.5	-2	0	0	-	-0.5	-	-	0
Joint effect after max-max return (mil.$)	-	-3	-3	-0.5	-0.5	-	-3	-	-	-3

Table 2: First period solutions for different configurations of the SAA algorithm for the ten project test instance of stochastic project portfolio optimization problem

Project / SAA Configuration	**10T5S100R-MPPM**	**10T10S50R-MPPM**	**10T25S20R-MPPM**	**10T50S10R-MPPM**	**10T5S100R-2PPM**	**10T10S50R-2PPM**	**10T25S20R-2PPM**	**10T50S10R-2PPM**
A								
B								
C								
D								
E								
F	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05
G	0.75				0.75	0.75		0.75
H	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
I							0.75	
J		0.75	0.75	0.75				
First Period Budget	3	3	3	3	3	3	3	3
First Period Investment	3	3	3	3	3	3	3	3

Table 3: Computational Results for $MPPM$ - * and ** indicate that the best solutions were the same

Instance	app. #of rows	app. #of columns	sec/rep	sec/rep (soln.)	avg. duality gap (%)	$\bar{v}_N^M$	$\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)$	$\hat{\sigma}^2_{\bar{v}_N^M}$	$\hat{\sigma}^2_{\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)}$	adj.opt.gap est.
5T5S100R-MPPM	4,500	3,000	572	5.69	1.189	82.009	81.583	0.889	0.672	2.875
5T10S50R-MPPM	11,000	5,500	780	6.85	1.170	80.695	81.745	0.956	0.398	1.231 *
5T25S20R-MPPM	45,000	13,500	1398	23.09	2.438	80.373	81.117	0.731	0.217	1.164 *
5T50S10R-MPPM	146,000	27,000	3552	77.52	4.408	78.583	80.421	0.889	0.073	0.084 *
10T5S100R-MPPM	8,500	6,000	1706	84.32	1.020	228.596	230.736	5.851	2.090	3.383
10T10S50R-MPPM	22,000	11,000	1860	122.34	1.280	229.444	232.035	2.836	5.580	3.095 **
10T25S20R-MPPM	88,000	27,000	3480	204.10	1.870	229.033	232.314	2.034	1.754	0.534 **
10T50S10R-MPPM	290,000	54,000	3552	652.45	2.438	225.321	230.022	9.226	1.365	1.678 **

Table 4: Computational Results for $2PPM$ - * and ** indicate that the best solutions were the same

Instance	app. #of rows	app. #of columns	sec/rep	sec/rep (soln.)	avg. duality gap (%)	$\bar{v}_N^M$	$\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)$	$\hat{\sigma}^2_{\bar{v}_N^M}$	$\hat{\sigma}^2_{\hat{g}_{N'}(\hat{\mathbf{x}}_N^m)}$	adj.opt.gap est.
5T5S100R-2PPM	2,500	1,500	58	2.28	0.769	82.230	83.210	0.833	0.774	1.505 *
5T10S50R-2PPM	4,500	3,000	69	4.50	1.753	80.808	81.701	0.741	0.346	1.150 *
5T25S20R-2PPM	11,500	7,000	168	12.01	2.770	79.900	81.356	0.617	0.185	0.299 *
5T50S10R-2PPM	23,000	14,000	204	28.05	3.060	79.394	81.650	1.582	0.104	0.289 *
10T5S100R-2PPM	4,500	3,000	795	33.27	0.586	232.231	232.282	3.613	3.152	5.047 **
10T10S50R-2PPM	8,500	6,000	992	39.60	0.894	232.679	232.793	2.458	2.483	4.243 **
10T25S20R-2PPM	21,500	14,000	1446	48.12	1.709	228.650	230.204	1.747	2.564	2.516
10T50S10R-2PPM	43,000	28,000	2040	66.70	2.625	228.610	231.917	1.268	1.804	0.128 **

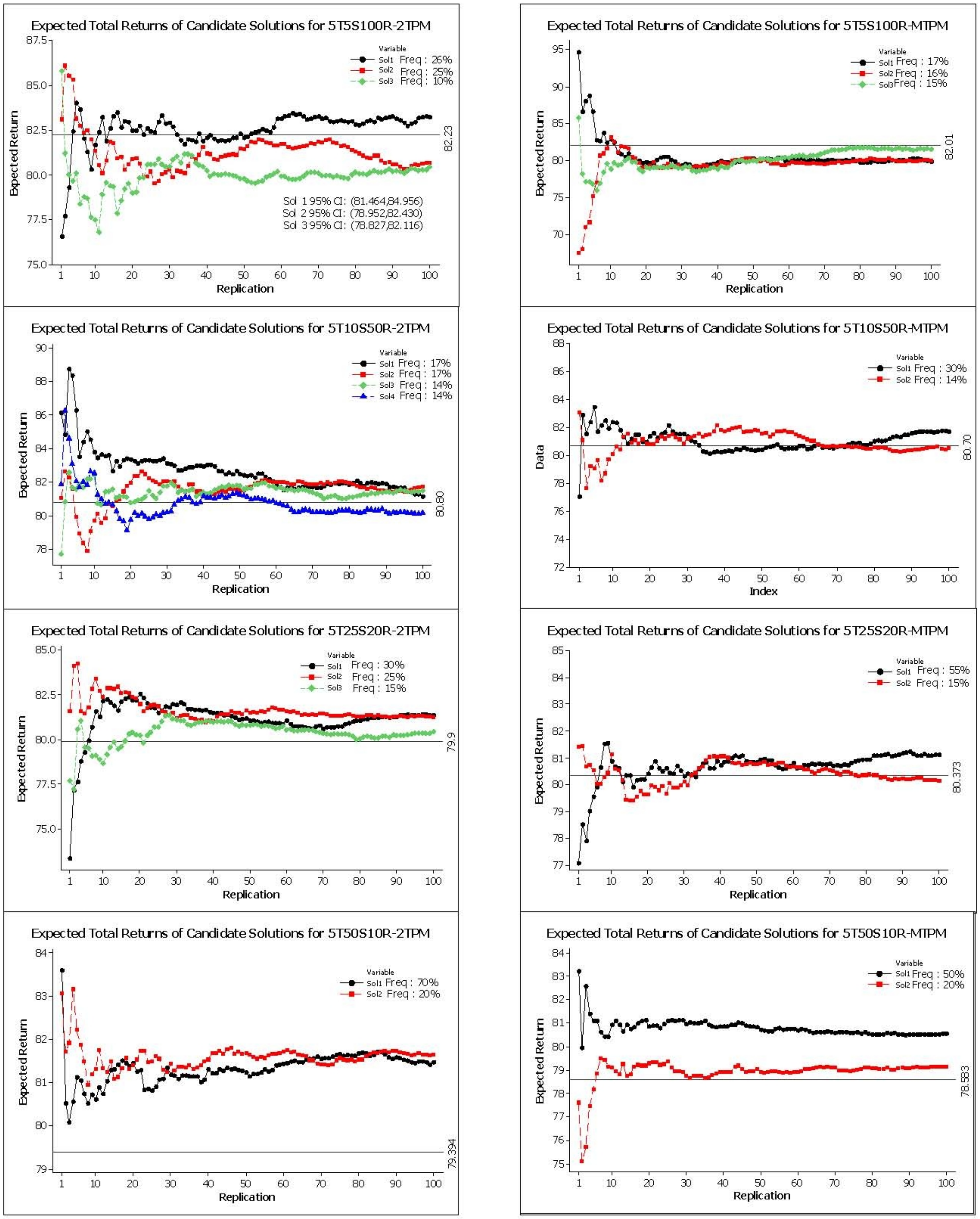

Figure 5: Estimation of expected value of the objective function for candidate solutions using samples sizes of $N' = 100$

References

April, J., F. Glover, J.P. Kelly. 2003. Optfolio-a simulation optimization system for project portfolio planning. *Proceedings of the 2003 Winter Simulation Conference*. New Orleans, LA.

Bardhan, I.R., S. Dagehi, R. Sougstad. 2004. A real options approach for prioritization of a portfolio of information technology projects: a case study of a utility company. *Proceedings of the 37th Hawaii International Conference on System Sciences*.

Bardhan, I.R., R. J. Kauffman, S. Naranpanawe. 2006. Optimizing an it project portfolio with time-wise interdependencies. *Proceedings of the 39th Hawaii International Conference on System Sciences*.

Birge, J.R., F. Louveaux. 1997. *Introduction to Stochastic Programming*. Springer, New York.

Campbell, J.A. 2001. Identifying an optimal start-up date for information technologyprojects: a real options approach. *Proceedings of Portland International Conference on Management of Engineering and Technology* **1** 37.

Chan, T., J.A. Nickerson, H. Owan. 2007. Strategic management of r&d pipelines with cospecialized investments and technology markets. *Management Science* **53**(4) 667–682.

Cooper, R.G., S.J. Edgett, E.J. Kleinschmidt. 2001. *Portfolio Management for New Products*. Perseus Publishing, New York.

Dickinson, M.W., A.C. Thornton, S. Graves. 2001. Technology portfolio management: optimizing interdependent projects over multiple time periods. *IEEE Transactions on Engineering Management* **48**(4) 518–527.

Elfes, A., W.P. Lincoln, G. Rodriguez, C.R. Weisbin, J.A. Wertz. 2005. Risk-based technology portfolio optimization for early space mission design. *IEEE Transactions on Aerospace and Electronic Systems* **42**(1) 22–36.

Ghasemzadeh, F., N. Archer, P. Iyogun. 1999. A zero-one model for project portfolio selection and scheduling. *The Journal of the Operational Research Society* **50** 745–755.

Goel, V., I.E. Grossmann. 2004a. A lagrangian duality based branch and bound for solving linear stochastic programs with decision dependent uncertainty. Tech. rep., Carnegie Mellon University.

Goel, V., I.E. Grossmann. 2004b. A stochastic programming approach to planning of offshore gas field developments. *Computers and Chemical Engineering* **28**(8) 1409–1429.

Goel, V., I.E. Grossmann. 2006. A class of stochastic programs with decision dependent uncertainty. *Mathematical Programming* **108**(2) 355–394. doi:http://dx.doi.org/10.1007/s10107-006-0715-7.

Goldner, D.R., S.S. Borener. 2006. Evaluating ngats research priorities at jpdo. *Proceedings of the 6th AIAA Aviation Technology, Integration and Operations Conference*. Wichita, KS.

Gustafsson, J., A. Salo. 2005. Contingent portfolio programming for the management of risky projects. *Operations Research* **53**(6) 946–956.

Held, H., D.L. Woodruff. 2005. Heuristics for multistage interdiction of stochastic networks. *Journal of Heuristics* **11**(5-6) 483–500. doi:http://dx.doi.org/10.1007/s10732-005-3122-y.

Hiriart-Urruty, J-B., C. Lemarechal. 1993. *Convex Analysis and Minimization Algorithms I: Fundamentals*. Springer.

Jonsbraten, T.W., R.J. Wets, D.L. Woodruff. 1998. A class of stochastic programs with decision dependent random elements. *Annals of Operations Research* **82** 83–106.

Kleywegt, A.J., A. Shapiro, T.H. De-Mello. 2002. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* **12**(2) 479–502. doi: http://dx.doi.org/10.1137/S1052623499363220.

Lee, J., D. Lee, J. Ahn. 2001. Real options approach for optimal investment decision timing. *Proceedings of Portland International Conference on Management of Engineering and Technology* **1** 121.

Lincoln, W., S. Merida, V. Adumitroaie, C.R. Weisbin. 2006. Capability investment strategy to enable jpl future space missions. ASCE, 100–100. URL `http://link.aip.org/link/?ASC/188/100/1`.

Linton, J. D., S.T. Walsh, J. Morabito. 2002. Analysis, ranking and selection of r&d projects in a portfolio. *R&D Management* **32**(2) 139–148.

Louveaux, F.V., R. Schultz. 2003. *Handbooks in Operations Research and Management Science*, vol. 10, chap. 4. Elsevier Science, 213–266.

Luenberger, D. G. 1998. *Investment Science*. Oxford Univ. Press, New York.

Markowitz, H.M. 1952. Portfolio selection. *Journal of Finance* **6** 71–91.

Norkin, V.I., Y.M. Ermoliev, A. Ruszczynski. 1998. On optimal allocation of indivisibles under uncertainty. *Operations Research* **46**(3) 381–395.

Reyck, B. De, Y. Cockayne, M. Lockett, S.R. Calderini, M. Moura, A. Sloper. 2005. The impact of project portfolio management on information technology projects. *International Journal of Project Management* **23** 524–537.

Sallie, S.; Tapper S., F.; Surujhlal. 2002. Application of r&d portfolio analysis at a de beers business unit. *Proceedings of the IEEE International Engineering Management Conference* **2** 556–561.

Shapiro, A. 2003. Inference of statistical bounds for multistage stochastic programming problems. *Mathematical Methods of Operations Research* **58** 57–68.

Sharpe, W.F. 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* .

Tarhan, B., I.E. Grossman. 2006. A multistage stochastic programming approach with strategies for uncertainty reduction in the synthesis of process networks with uncertain yields. Tech. rep., Carnegie Mellon University.

Thompson, R.W. 2006. System architecture modeling for technology portfolio management using atlas. *Proceedings of the 2006 IEEE Aerospace Conference*. Big Sky, MT.

Tobin, J. 1958. Liquidity preference as behavior towards risk. *Review of Economic Studies* **67** 65–86.

Tralli, D.M. 2004. Valuation of technology development using a workflow approach to compund real options. *Proceedings of IEEE Aerospace Conference*. Big Sky, MT.

Utturwar, A., S. Rallabhandi, D. DeLaurentis, D. Mavris. 2002. A bi-level optimization approach for teechnology selection. *Proceedings of the 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*. Atlanta, GA.

Viswanath, K., S. Peeta, S.F. Salman. 2004. Investing in the links of a stochastic network to minimize expected shortest path length. Tech. rep., Purdue University.

Wu, L., C. Ong. 2007. Management of information technology investment: A framework based on a real options and meanvariance theory perspective. *Technovation* doi:10.1016/j.technovation.2007.05.011.