doi:10.1016/j.ejor.2017.12.039 • www.stefancreemers.be • info@stefancreemers.be

Moments and distribution of the net present value of a serial project

Stefan Creemers

Abstract - **We study the Net Present Value (NPV) of a project with multiple stages that are executed in sequence. A cash flow (positive or negative) may be incurred at the start of each stage, and a payoff is obtained at the end of the project. The duration of a stage is a random variable with a general distribution function. For such projects, we obtain exact, closed-form expressions for the moments of the NPV, and develop a highly accurate closed-form approximation of the NPV distribution itself. In addition, we show that the optimal sequence of stages (that maximizes the expected NPV) can be obtained efficiently, and demonstrate that the problem of finding this optimal sequence is equivalent to the least cost fault detection problem. We also illustrate how our results can be applied to a general project scheduling problem where stages are not necessarily executed in series. Lastly, we prove two limit theorems that allow to approximate the NPV distribution. Our work has direct applications in the fields of project selection, project portfolio management, and project valuation.**

Keywords - **project scheduling, project management, net present value, NPV distribution, least cost fault detection problem**

1 Introduction

We consider a project with multiple stages that are executed in sequence. Each stage of the project has a random duration with general distribution function. At the start of a stage, a deterministic cash flow (positive or negative) may be incurred, and a deterministic payoff is obtained upon completion of the project. Continuous compounding is used to determine the Net Present Value (NPV) of the project (i.e., the sum of the discounted cash flows that are incurred during the project lifetime; the convolution of the NPV distributions of the individual cash flows). We develop exact, closed-form expressions to obtain the moments of the project NPV distribution. In addition, we provide a highly accurate approximation of the NPV distribution itself. The approximation uses a three-parameter lognormal distribution to match the first three moments of the NPV distribution. A lognormal distribution was chosen because: (1) the moment-matching procedure uses closed-form expressions, and (2) we show that the NPV of a cash flow converges to a (reflected) lognormal distribution if the cash flow is not incurred during the early stages of the project. We also show that, if a sufficient number of cash flows are incurred, the project NPV converges to a normal distribution. In addition, we show that the sequence of stages that maximizes the expected NPV (eNPV) over all possible sequences can be found efficiently, and that the problem of finding this optimal sequence is equivalent to the Least Cost Fault Detection Problem (LCFDP). Lastly,

if stages are not executed in sequence, we demonstrate that our approach can still be used to approximate the moments and the distribution of the project NPV. We use examples to illustrate our results, and to show that our approach can easily be implemented.

Our work has direct applications in the fields of project selection, project portfolio management, and project valuation. In these fields, the detailed scheduling of activities is often not considered, and it is assumed that: (1) a project is a sequence of stages with cash flows that are incurred at the start of a stage, and (2) a (uncertain) payoff is obtained upon completion of the project. Such projects are not only prevalent in the real world, but also in the literature. For instance, Huchzermeier and Loch (2001) consider an R&D project that is divided in sequential stages, and develop a dynamic program to determine the expected value of the project. Santiago and Vakili (2005) build on the work of Huchzermeier and Loch, and also consider an R&D project with sequential stages. De Reyck and Leus (2008) discuss the literature on R&D project scheduling, and conclude that most of the literature is limited to sequential R&D stages only. Girotra et al. (2007) determine the eNPV of a drug development project where stages have been defined by a regulator. They also mention that a stage-gate development process is prevalent in most industries. Chao et al. (2014) also investigate the use of a state-gate processes to manage NPD projects. They argue that decisions based on eNPV alone are dangerous, and that risk should be taken into account when making project selection/investment decisions. Often, the risk of a project is modeled using the variance of the NPV (Van Horne 1966). Other measures of risk are the skewness and/or kurtosis of the NPV, and the probability to have a negative NPV. Until now, however, Monte Carlo simulation was the only technique available to obtain higher moments and/or the NPV distribution itself. In this article, we develop a closed-form characterization of the NPV distribution of a project that is a valid alternative to Monte Carlo simulation, and that can be directly applied to evaluate project selection/investment decisions.

On the tactical/strategical level, projects are often seen as a sequence of stages. Operational factors, however, may also result in the serial execution of a project. For instance, a bottleneck resource may force activities to be executed in series. A bottleneck resource has been considered, among others, by Kaviadias and Loch (2003), who study NPD projects that compete for a scarce resource, and that are divided into stages. In addition, some industries are more likely to have a serial project execution due to an abundance of technical precedence relationships (e.g., the construction industry).

In the (more operational) field of project scheduling, our work is related to CPM/PERT in the sense that we also focus on a single sequence of stages, and that we also use normal (lognormal) approximations. The study of CPM/PERT dates back to the work of Kelley and Walker (1959) and Malcolmn et al. (1959), and still continues today (refer to Demeulemeester and Herroelen (2002) and Trietsch and Baker (2012) for an overview of the literature). Whereas CPM/PERT deals with the project completion time, we focus on the NPV. In a recent survey, Wiesemann and Kuhn (2015) not only highlight the importance of NPV over project completion time, but also stress the importance of stochastic project scheduling. In stochastic project scheduling, stage durations and/or cash flows are random variables, and as a result the project NPV is a random variable as well. Although most of the literature deals with minimizing the expected completion time of a project (Herroelen 2005; Ballestín and Leus 2009), some research has already been devoted to maximizing the eNPV of a project (Vanhoucke et al. 2001; Szmerekovsky 2005). Higher moments of the NPV distribution,

and/or the NPV distribution of a project itself, have never been studied before. In general, it is considered to be impossible to efficiently determine the NPV distribution of a project (Wiesemann and Kuhn 2015). In fact, for the completion time of a project, Hagström (1988) has shown that it is $\#P$-complete to determine even a single point of the Cumulative Distribution Function (CDF). Even for serial projects, Kamburowski (1986) has shown that the result of Hagström holds.

The remainder of this article is structured as follows. Section 2 develops exact, closed-form expressions for the moments and the distribution of the NPV of a cash flow that is obtained after a single stage. Multiple stages are considered in Section 3. In Section 3, we also show that the NPV of a single cash flow converges to a (reflected) lognormal distribution if the cash flow is not incurred during the early stages of the project. Section 4 introduces the lognormal approximation that can be used to model the NPV distributions of both individual cash flows as well as projects. In Section 5, we develop exact, closed-form expressions for the moments of the NPV distribution of a multi-stage project with intermediate cash flows. In addition, we also show that the NPV of a project converges to a normal distribution, and assess the accuracy of the lognormal and normal approximations of the project NPV distribution. In Section 6, we show that: (1) the problem of finding the optimal sequence of stages is equivalent to the LCFDP, (2) if stages are not precedence related, a well-known result from the literature on the LCFDP can be used to obtain the optimal sequence in polynomial time, and (3) efficient methods exist to obtain the optimal sequence of stages if they are precedence related. Section 7 illustrates how our results can be used to approximate the moments and the distribution of the NPV of a general project where stages are scheduled using a scheduling policy. Section 8 discusses a number of model extensions, and Section 9 concludes and provides directions for future research.

2 NPV of a cash flow obtained after a single stage

In this section, we investigate the basic case where a cash flow c is incurred after a single stage. Under continuous compounding, the NPV of a cash flow c is given by:

$$v = ce^{-rt}, \tag{1}$$

where r is the discount rate, and t is the time at which cash flow c is incurred. If t is a realization of T, and if T is a random variable with probability function $f(t)$, the eNPV of cash flow c is given by:

$$\mu = \int_0^\infty f(t)ce^{-rt}\mathrm{d}t.$$

Lemma 1. *Consider a cash flow c that is incurred at time T, where T is a random variable with probability function $f(t)$. Given a discount rate r, the eNPV of c is given by:*

$$\mu = cM_T(-r),$$

where $M_T(u)$ is the Moment Generating Function (MGF) of T.

For notational convenience, let $\phi(r) \equiv M_T(-r)$ such that:

$$\mu = cM_T(-r) = c\phi(r). \tag{2}$$

$\phi(r)$ can be interpreted as the eNPV of a cash flow $c = 1$ that is obtained at time T if discount rate r applies. For most distributions, the MGF (and hence $\phi(r)$) is readily available. There are some distributions, however, for which the MGF does not have a closed-form expression (e.g., the Weibull distribution), or for which the MGF is undefined (e.g., the lognormal distribution). For those distributions, $\phi(r)$ has to be approximated. In addition, note that $\phi(r)$ is not always defined for all values of r. For instance, if T is exponentially distributed, its MGF is given by $M_T(u) = \lambda(\lambda - u)^{-1}$. Hence, if $r = -\lambda$, the MGF about $-r$ is undefined, and μ cannot be determined. In practice, however, this is rarely an issue.

We use an example to illustrate Lemma 1. Consider a cash flow $c = 1,000$ that is incurred at time T, where T follows a gamma distribution with shape parameter $k = 5$ and scale parameter $\tau = 1$. The MGF of the gamma distribution is $M_T(u) = (1 - \tau u)^{-k}$. As a result, $\phi(r) = (1 + \tau r)^{-k}$, and the eNPV of cash flow c is $\mu = c\phi(r) = 620.92$ for discount rate $r = 0.1$.

Theorem 1. *Consider a cash flow c that is incurred at time T, where T is a random variable with probability function $f(t)$. Given a discount rate r, the mean, variance, skewness, and kurtosis of the NPV of c are given by:*

$$\begin{aligned}
\mu &= c\phi(r), \\
\sigma^2 &= c^2(\phi(2r) - \phi^2(r)), \\
\gamma &= c^3\left(\phi(3r) - 3\phi(2r)\phi(r) + 2\phi^3(r)\right)\sigma^{-3}, \\
\theta &= \left(\phi(4r) - 4\phi(3r)\phi(r) + 6\phi(2r)\phi^2(r) - 3\phi^4(r)\right)\left(\phi(2r) - \phi^2(r)\right)^{-2}.
\end{aligned} \tag{3}$$

If we revisit the previous example, the moments of the NPV distribution of cash flow c are: $\mu = 620.92$, $\sigma^2 = 16,334$, $\gamma = -0.2347$, and $\theta = 2.7064$ for discount rate $r = 0.1$.

Theorem 2. *Consider a cash flow c that is incurred at time T, where T is a random variable with probability function $f(t)$. Given a discount rate r, the CDF and Probability Density Function (PDF) of the NPV of cash flow c are given by:*

$$\begin{aligned}
G(v) &= 1 - F\left(\ln\left(\frac{c}{v}\right)r^{-1}\right), \\
g(v) &= \frac{f\left(\ln\left(\frac{c}{v}\right)r^{-1}\right)}{|r|v},
\end{aligned}$$

where $F(t)$ is the CDF of T. Note that: (1) if $r > 0$, then v has range $0 \leq v < c$, (2) if $r = 0$, then $v = c$, and (3) if $r < 0$, then v has range $c < v \leq \infty$.

We illustrate Theorem 2 by means of an example. In the example, a cash flow c is incurred at time T, where T follows an exponential distribution with rate parameter λ. For a given discount rate r, the CDF of the NPV of cash flow c is:

$$G(v) = \left(\frac{c}{v}\right)^{-\lambda r^{-1}}.$$

Similar results can be obtained for other probability functions.

3 NPV of a cash flow obtained after multiple stages

In this section, we consider the NPV of a cash flow that is incurred after multiple stages. Below, we use payoff p to demonstrate our results (as payoff p is obtained at the end of the project; after all stages have been completed). Note, however, that the results in this section hold for any cash flow that is incurred during the lifetime of the project.

Lemma 2. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. Each stage $w : w \in \mathbf{N}$ has duration distribution $f_w(t)$ and corresponding factor $\phi_w(r)$ that is obtained using Eq. (2). If the durations of the individual stages are independent, the duration of the project itself has factor:*

$$\phi_{1,n}(r) = \prod_{w \in \mathbf{N}} \phi_w(r).$$

We can combine Theorem 1 with Lemma 2 to determine the moments of the NPV of a cash flow that is incurred after multiple stages. For instance, consider the NPV of a payoff p that is obtained upon completion of a project with three stages. The stages have factors $\phi_1(r)$, $\phi_2(r)$, and $\phi_3(r)$, respectively. The mean and variance of the NPV of payoff p are given by:

$$\begin{aligned}
\mu &= p\phi_1(r)\phi_2(r)\phi_3(r) = p\phi_{1,3}(r), \\
\sigma^2 &= p^2(\phi_1(2r)\phi_2(2r)\phi_3(2r) - \phi_1^2(r)\phi_2^2(r)\phi_3^2(r)) = p^2(\phi_{1,3}(2r) - \phi_{1,3}^2(r)).
\end{aligned}$$

The skewness, kurtosis, and higher-order moments are obtained in the same way.

Lemma 3. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. Each stage $w : w \in \mathbf{N}$ has a duration distribution $f_w(t)$ with mean d_w and variance s_w^2. If the durations of the individual stages are independent, the mean and variance of the project duration are given by:*

$$\begin{aligned}
d_{\mathbf{N}} &= \sum_{w \in \mathbf{N}} d_w, \\
s_{\mathbf{N}}^2 &= \sum_{w \in \mathbf{N}} s_w^2.
\end{aligned}$$

If n is sufficiently large, and if no stage dominates the others, the duration of the project will converge to a normal distribution with mean $d_{\mathbf{N}}$ and standard deviation $s_{\mathbf{N}}$.

Lemma 3 is a well-known result in the literature (Malcolm et al. 1959; Van Slyke 1963; Moder and Phillips 1970), and allows to predict the completion time of a project. We will use Lemma 3 to show that the NPV of a payoff p converges to a (reflected) lognormal distribution if n is sufficiently large.

Theorem 3. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. If the durations of the individual stages are independent, and if n is sufficiently large, the NPV of payoff p converges to a (reflected) lognormal distribution $g(v)$ with location parameter $\alpha = \ln(p) - rd_{\mathbf{N}}$ and scale parameter $\beta = rs_{\mathbf{N}}$.*

Exact NPV distribution						
n	1	5	10	25	50	100
μ	666.67	620.92	613.91	609.53	608.04	607.29
σ^2	55,556	16,334	8,654	3,589	1,817	914
γ	-0.566	-0.235	-0.163	-0.101	-0.071	-0.050
θ	2.4000	2.7060	2.8300	2.9252	2.9613	2.9803

$\mathcal{L}_{\mathrm{N}}$ approximation						
n	1	5	10	25	50	100
$\mu_{\mathcal{L}\mathrm{N}}$	687.29	621.89	614.16	609.57	608.05	607.29
$\sigma^2_{\mathcal{L}\mathrm{N}}$	134,164	19,829	9,549	3,734	1,853	923
$\gamma_{\mathcal{L}\mathrm{N}}$	1.7500	0.6909	0.4814	0.3018	0.2128	0.1502
$\theta_{\mathcal{L}\mathrm{N}}$	8.8980	3.8606	3.4148	3.1623	3.0806	3.0401
K–S	0.1587	0.0596	0.0421	0.0266	0.0188	0.0133

Table 1: Accuracy of the $\mathcal{L}_{\mathrm{N}}$ approximation for various number of stages

In order to illustrate Theorem 3, consider a project with n stages that are executed in sequence, and that have i.i.d. exponential durations with rate parameter λ (i.e., the project duration follows an Erlang distribution with parameters n and λ). A payoff p is obtained upon completion of the project. After applying Theorem 2, we obtain the PDF of the NPV of payoff p:

$$g(v) = \frac{\lambda \left(\frac{p}{v}\right)^{-\lambda r^{-1}} \left(\ln\left(\frac{p}{v}\right) \lambda r^{-1}\right)^{n-1}}{|r| v (n-1)!}.$$

The approximate lognormal distribution has location parameter $\alpha = \ln(p) - rn\lambda^{-1}$ and scale parameter $\beta = r\sqrt{n}\lambda^{-1}$, and is denoted by $\mathcal{L}_{\mathrm{N}}$. Given a payoff $p = 1,000$, and a rate parameter $\lambda = 1$, Fig. 1 shows the exact and the approximate PDF of the distribution of the NPV of payoff p for various values of n. The discount rate r is set equal to $0.5n^{-1}$. Table 1 reports the mean, variance, skewness, kurtosis, and Kolmogorov–Smirnov (K–S) test statistic (i.e., the maximum absolute difference in cumulative probability; the maximum absolute difference between $G(v)$ and the CDF of $\mathcal{L}_{\mathrm{N}}$). We observe that, if n is small, Lemma 3 (and hence Theorem 3) does not hold, and the approximation performs poorly. If, on the other hand, n is large, the approximation is fairly accurate, and the NPV of a payoff p may be approximated by a lognormal distribution.

4 A lognormal approximation of the NPV distribution

Theorem 3 only holds for cash flows that are incurred after a sufficient number of stages. Hence, the NPV of a cash flow does not always follow a lognormal distribution. Often, it is impossible to characterize the exact NPV distribution of a cash flow, however, we can use Theorem 1 to obtain its moments. A moment-matching procedure can then be used to define a distribution that approximates the true NPV distribution.

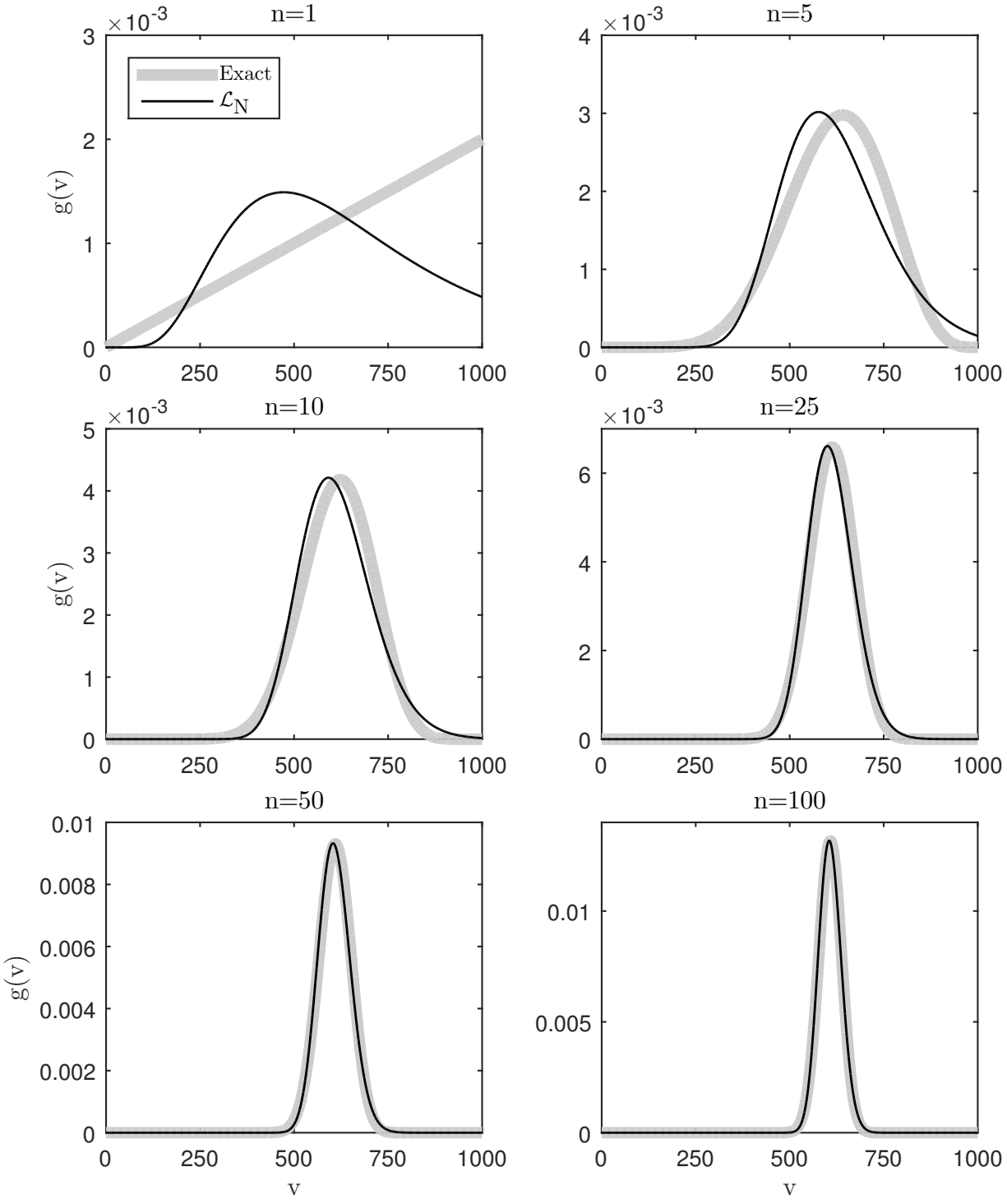

Figure 1: PDF of the exact NPV and the $\mathcal{L}_N$ approximation for various number of stages

Moment-matching procedures can be evaluated along three lines: (1) the number of moments matched, (2) the computational efficiency, and (3) the generality of the solution. Ideally, a moment-matching procedure uses closed-form expressions to match as many moments as possible under general conditions. Most of the literature on moment matching has focussed on the use of phase-type (PH) distributions (Osogami 2005). Using PH distributions, up to three moments can be matched using closed-form expressions (Osogami and Harchol-Balter 2006). In this article, we do not adopt PH distributions, however, we use a lognormal approximation of the NPV distribution of a cash flow c. Not only does the lognormal distribution allow us to develop closed-form expressions to match up to three moments of any real-valued distribution with non-zero skew, it is also a logical choice as the NPV distribution of a cash flow c converges to a (reflected) lognormal distribution if it is incurred after a sufficient number of stages (see also Theorem 3).

In what follows, we define two moment-matching procedures. In a first procedure, we match the first two moments of the NPV distribution. A second procedure matches the first three moments. We use $\mathcal{L}_2$ and $\mathcal{L}_3$ to denote both approximations, respectively.

Proposition 1. *We can approximate the NPV distribution by matching its first two moments using a (reflected) lognormal distribution with scale and location parameters:*

$$\beta = \sqrt{\ln\left(1+\eta^2\right)},$$
$$\alpha = \ln(\mu) - 0.5\beta^2,$$

where μ and $\eta = \sigma^2\mu^{-2}$ are the mean and Squared Coefficient of Variation (SCV) of the NPV distribution, respectively.

In order to match three moments, we use a three-parameter (or bounded) lognormal distribution (Aitchison and Brown 1957) with location, shape, and threshold parameters α, β, and κ, respectively. The threshold parameter can be used to bound the support of the distribution, and can either serve as a lower or as an upper bound (for matching distributions with positive/negative skew, respectively). The mean, variance, skewness, kurtosis, PDF, and CDF of the three-parameter lognormal distribution are given by:

$$\mu_{\mathcal{L}3} = \kappa + \delta e^{\alpha+0.5\beta^2}, \tag{4}$$
$$\sigma^2_{\mathcal{L}3} = \left(e^{\beta^2}-1\right)e^{2\alpha+\beta^2}, \tag{5}$$
$$\gamma_{\mathcal{L}3} = \delta\left(2+e^{\beta^2}\right)\sqrt{e^{\beta^2}-1}, \tag{6}$$
$$\theta_{\mathcal{L}3} = e^{2\beta^2}\left(3+e^{\beta^2}\left(2+e^{\beta^2}\right)\right)-3,$$
$$g_{\mathcal{L}3}(v) = \frac{1}{\delta(v-\kappa)\beta\sqrt{2\pi}}e^{\left(\frac{(\ln(\delta(v-\kappa))-\alpha)^2}{2\beta^2}\right)},$$
$$G_{\mathcal{L}3}(v) = \frac{1}{2}-\frac{\delta}{2}\operatorname{Erf}\left(\frac{\alpha-\ln\left(\delta(v-\kappa)\right)}{\beta\sqrt{2}}\right),$$

where $\delta = -1$ if the distribution has negative skew, and $\delta = 1$ otherwise.

$\mathcal{L}_2$ approximation						
n	1	5	10	25	50	100
$\mu_{\mathcal{L}2}$	666.67	620.92	613.91	609.53	608.04	607.29
$\sigma^2_{\mathcal{L}2}$	55,556	16,334	8,654	3,589	1,817	914
$\gamma_{\mathcal{L}2}$	1.1049	0.6262	0.4581	0.2958	0.2106	0.1495
$\theta_{\mathcal{L}2}$	5.2463	3.7053	3.3754	3.1560	3.0790	3.0397
K–S	0.1357	0.0597	0.0421	0.0266	0.0188	0.0133

$\mathcal{L}_3$ approximation						
n	1	5	10	25	50	100
$\mu_{\mathcal{L}3}$	666.67	620.92	613.91	609.53	608.04	607.29
$\sigma^2_{\mathcal{L}3}$	55,556	16,334	8,654	3,589	1,817	914
$\gamma_{\mathcal{L}3}$	-0.566	-0.235	-0.163	-0.101	-0.071	-0.050
$\theta_{\mathcal{L}3}$	3.5743	3.0981	3.0470	3.0182	3.0090	3.0045
K–S	0.0590	0.0118	0.0059	0.0023	0.0011	0.0006

Table 2: Accuracy of the $\mathcal{L}_2$ and $\mathcal{L}_3$ approximations for various number of stages

Proposition 2. *We can approximate the NPV distribution by matching its first three moments using a bounded lognormal distribution with parameters:*

$$\beta = \sqrt{\ln\left(\frac{2^{1/3}}{\left(2+\gamma^2+\sqrt{4\gamma^2+\gamma^4}\right)^{1/3}} + \frac{\left(2+\gamma^2+\sqrt{4\gamma^2+\gamma^4}\right)^{1/3}}{2^{1/3}} - 1\right)},$$

$$\alpha = 0.5\left(\ln\left(\frac{\sigma^2}{e^{\beta^2}-1}\right) - \beta^2\right),$$

$$\kappa = \mu - \delta e^{\alpha+0.5\beta^2},$$

where μ, σ^2, and γ are the mean, variance, and skewness of the NPV distribution.

In order to illustrate the accuracy of the lognormal approximations, we revisit the last example of Section 3. Fig. 2 shows the exact and the approximate PDF of the NPV distribution for various values of n. Table 2 reports the mean, variance, skewness, kurtosis, and Kolmogorov-Smirnov test statistic. We observe that the $\mathcal{L}_3$ approximation is almost always very accurate, whereas the $\mathcal{L}_2$ approximation has more or less the same accuracy as the $\mathcal{L}_N$ approximation. This latter observation is no surprise. If n is small, neither Lemma 3 nor Theorem 3 hold, and the approximations fail to achieve a good accuracy. In addition, the $\mathcal{L}_2$ and $\mathcal{L}_N$ approximations only take into account the first two moments. As a result, they are always dominated by the $\mathcal{L}_3$ approximation.

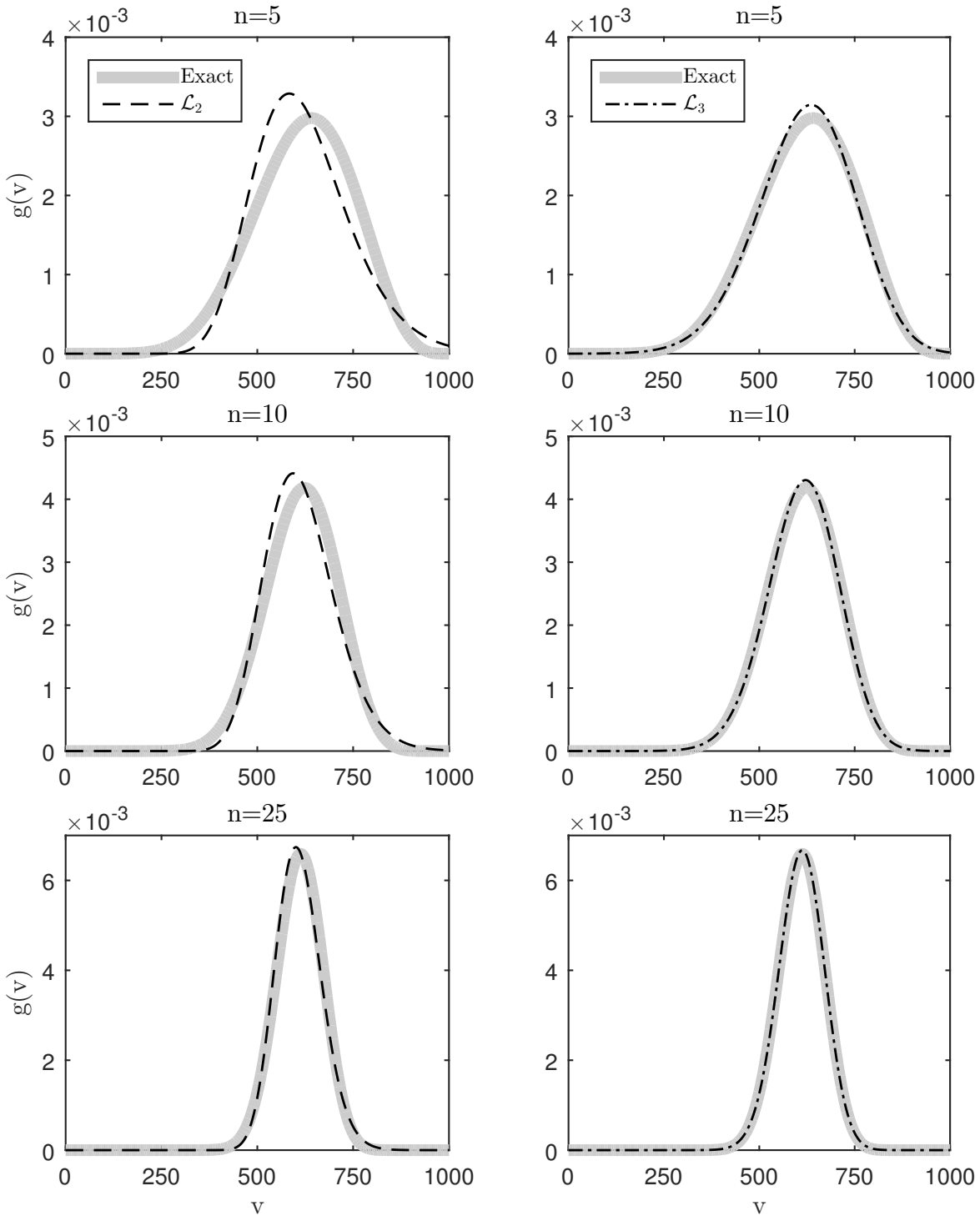

Figure 2: PDF of the exact NPV, the $\mathcal{L}_2$, and the $\mathcal{L}_3$ approximation for various number of stages

5 NPV of a project with multiple stages and intermediate cash flows

In this section, we consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$, and assume that a cash flow c_w is incurred at the start of stage w. A payoff p is obtained upon completion of the project. For notational convenience, we let $c_{n+1} \equiv p$. Let $\mathbf{c} = \{c_1, c_2, \ldots, c_n, c_{n+1}\}$ denote the set of cash flows that are incurred during the lifetime of the project. In addition, define V_w, the random variable that represents the NPV of cash flow c_w, and let $V_{\mathbf{c}} = \sum_{w=1}^{n+1} V_w$ denote the random variable that captures the NPV of the project. Because the NPV of a cash flow c_x depends on the NPV of an earlier cash flow c_w, V_x depends on V_w for all $x, w : 1 \leq w < x \leq n+1$. Hence, $V_{\mathbf{c}}$ is the sum of a number of dependent random variables whose distribution converges to a (reflected) lognormal distribution if their associated cash flow is not incurred during the early stages of the project.

Determining the distribution of $V_{\mathbf{c}}$ is closely related to finding the distribution of the lognormal sum (i.e., the sum of a number of random variables that follow a lognormal distribution). Even though the lognormal sum has received considerable attention in the literature, few exact results are available (Yan et al. 2016). In what follows, we first develop exact, closed-form expressions for the moments of the distribution of $V_{\mathbf{c}}$. We then use the lognormal approximation developed in Section 4 to approximate the NPV distribution and illustrate its accuracy by means of an example. Next, we show that $V_{\mathbf{c}}$ is normally distributed if the number of cash flows is sufficiently large, and propose a new approximation based on the normal distribution. Again, we illustrate the accuracy of this approximation by means of an example.

Theorem 4. *Consider a project with multiple stages $w : w \in \mathbf{N}$, and let $\mathbf{c} = \{c_1, c_2, \ldots, c_n, c_{n+1}\}$ denote the set of cash flows that are incurred at the start of each stage (where $c_{n+1} \equiv p$ is the payoff that is obtained upon project completion). In addition, V_w denotes the random variable that represents the NPV of cash flow c_w, and $V_{\mathbf{c}} = \sum_{w=1}^{n+1} V_w$ is the random variable that captures the NPV of the project. The moments of the distribution of $V_{\mathbf{c}}$ are:*

$$\begin{aligned}
\mu_{\mathbf{c}} &= \sum_{w=1}^{n+1} \mu_w, \\
\sigma_{\mathbf{c}}^2 &= \mathbf{e}\boldsymbol{\Sigma}_{\mathbf{c}}\mathbf{e}, \\
\gamma_{\mathbf{c}} &= (\mathbf{e}\boldsymbol{\Gamma}_{\mathbf{c}}\mathbf{e})\, \sigma_{\mathbf{c}}^{-3}, \\
\theta_{\mathbf{c}} &= (\mathbf{e}\boldsymbol{\Theta}_{\mathbf{c}}\mathbf{e})\, \sigma_{\mathbf{c}}^{-4},
\end{aligned}$$

where $\mathbf{e}$ is a vector of ones, and $\boldsymbol{\Sigma}_{\mathbf{c}}$, $\boldsymbol{\Gamma}_{\mathbf{c}}$, and $\boldsymbol{\Theta}_{\mathbf{c}}$ are the central covariance, coskewness, and cokurtosis matrices, respectively. $\boldsymbol{\Sigma}_{\mathbf{c}}$, $\boldsymbol{\Gamma}_{\mathbf{c}}$, and $\boldsymbol{\Theta}_{\mathbf{c}}$ capture the covariance, coskewness, and cokurtosis of the NPV of the cash flows in $\mathbf{c}$. Table 3 provides a summary of the closed-form expressions that allow to calculate the entries of these cross-moment matrices.

In order to illustrate Theorem 4, we use an example project with 3 stages. In the example, cash outflows are incurred at the start of the project, and at the start of the third stage. Cash inflows, on the other hand, are received at the start of the second stage, and upon completion

Mean μ
$\mu_w = c_w a_1$

Covariance matrix $\mathbf{\Sigma_c}$
$\mathbf{\Sigma_c}(w,w) = \sigma_w^2 = c_w^2(a_2 - a^2)$
$\mathbf{\Sigma_c}(w,x) = c_w c_x b_1 (a_2 - a^2) = c_w^{-1} c_x b_1 \mathbf{\Sigma_c}(w,w)$

Central coskewness matrix $\mathbf{\Gamma_c}$
$\mathbf{\Gamma_c}(w,w,w) = \gamma_w \sigma_w^3 = c_w^3 (a_3 - 3a_2 a_1 + 2a^3)$
$\mathbf{\Gamma_c}(w,w,x) = c_w^{-1} c_x b_1 \mathbf{\Gamma_c}(w,w,w)$
$\mathbf{\Gamma_c}(w,x,x) = c_w c_x^2 (a_3 b_2 - a_2 a_1 (2b^2 + b_2) + 2a^3 b^2)$
$\mathbf{\Gamma_c}(w,x,y) = c_x^{-1} c_y h_1 \mathbf{\Gamma_c}(w,x,x)$

Central cokurtosis matrix $\mathbf{\Theta_c}$
$\mathbf{\Theta_c}(w,w,w,w) = \theta_w \sigma_w^4 = c_w^4 (a_4 - 4a_3 a_1 + 6a_2 a^2 - 3a^4)$
$\mathbf{\Theta_c}(w,w,w,x) = c_w^{-1} c_x b_1 \mathbf{\Theta_c}(w,w,w,w)$
$\mathbf{\Theta_c}(w,w,x,x) = c_w^2 c_x^2 (a_4 b_2 - 2a_3 a_1 (b_2 + b^2) + a_2 a^2 (b_2 + 5b^2) - 3a^4 b^2)$
$\mathbf{\Theta_c}(w,x,x,x) = c_w c_x^3 (a_4 b_3 - a_3 a_1 (b_3 + 3b_2 b_1) + 3a_2 a^2 (b_2 b_1 + b^3) - 3a^4 b^3)$
$\mathbf{\Theta_c}(w,w,x,y) = c_x^{-1} c_y h_1 \mathbf{\Theta_c}(w,w,x,x)$
$\mathbf{\Theta_c}(w,x,x,y) = c_x^{-1} c_y h_1 \mathbf{\Theta_c}(w,x,x,x)$
$\mathbf{\Theta_c}(w,x,y,y) = c_w c_x c_y^2 ((a_4 - a_3 a_1) b_3 h_2 - (h_2 + 2h^2) ((a_3 a_1 - a_2 a^2) b_2 b_1) + (a_2 a^2 - a^4) 3b^3 h^2)$
$\mathbf{\Theta_c}(w,x,y,z) = c_y^{-1} c_z o_1(r) \mathbf{\Theta_c}(w,x,y,y)$

$a_i = \phi_{1,w-1}(ir)$	$b_i = \phi_{w,x-1}(ir)$	$h_i = \phi_{x,y-1}(ir)$	$o_i = \phi_{y,z-1}(ir)$
$a^i = \phi^i_{1,w-1}(r)$	$b^i = \phi^i_{w,x-1}(r)$	$h^i = \phi^i_{x,y-1}(r)$	

Table 3: Summary of closed-form expressions that allow to calculate the moments of the NPV distribution of a project

w	c_w	$f_w(t)$	k_w	τ_w	d_w	s_w^2
1	-300	gamma	1.5	1.0	1.5	1.5
2	250	gamma	2.5	1.0	2.5	2.5
3	-750	gamma	0.5	1.0	0.5	0.5

p	1000
r	0.05

Table 4: Data of the example project with three stages

	Exact	Simulation	$\mathcal{L}_2$	$\mathcal{L}_3$	$\mathcal{L}_3$ without cross moments
μ	168.21	168.21	168.21	168.21	168.21
σ^2	1,533	1,533	1,533	1,533	10,276
γ	-1.035	-1.035	0.1006	-1.035	-2.620
θ	4.7421	4.7420	3.0180	4.9631	17.269
$G(0)$	NA	0.0105	0.0008	0.0105	0.1018
K–S	NA	NA	0.0734	0.0055	0.1018

Table 5: Accuracy of the $\mathcal{L}_2$ and $\mathcal{L}_3$ approximations of the NPV distribution of a project with intermediate cash flows

of the project. Each stage w has a duration that follows a gamma distribution with shape and scale parameters k_w and τ_w, respectively. The gamma distribution was chosen because it is a general distribution (e.g., the exponential, Erlang, and chi-squared distributions are all special cases of the gamma distribution) that has many practical applications. We assume a discount rate $r = 0.05$. The data of the example project are summarized in Table 4. Fig. 3 shows the $\mathcal{L}_2$ and $\mathcal{L}_3$ approximations, as well as the simulated PDF of the project NPV (note that we have to resort to simulation as it is no longer an easy task to determine the exact NPV distribution). It is clear that, in this example, the $\mathcal{L}_2$ approximation performs very poorly. The $\mathcal{L}_3$ approximation, however, is once more very accurate. Table 5 reports on the moments of the NPV distribution, the probability to have a negative project NPV, and the Kolmogorov–Smirnov test statistic. The exact moments have been obtained using Theorem 4. We observe that the simulation (with 1 billion replications) almost perfectly matches the exact moments, which supports the claim that the simulated PDF is close to the true PDF of the project NPV. As was also shown by Fig. 3, the $\mathcal{L}_3$ approximation yields excellent accuracy. If, however, cross moments are ignored (i.e., if we assume that the NPVs of the cash flows are independent), the accuracy is abysmal. This is also reflected in the probability to have a negative project NPV. Only the $\mathcal{L}_3$ approximation is able to provide an accurate estimate.

Theorem 5. *Consider a project with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. At the start of each stage $w : w \in \mathbf{N}$, a cash flow c_w is incurred, and a payoff $p \equiv c_{n+1}$ is obtained upon completion of the project. Let V_w denote the random*

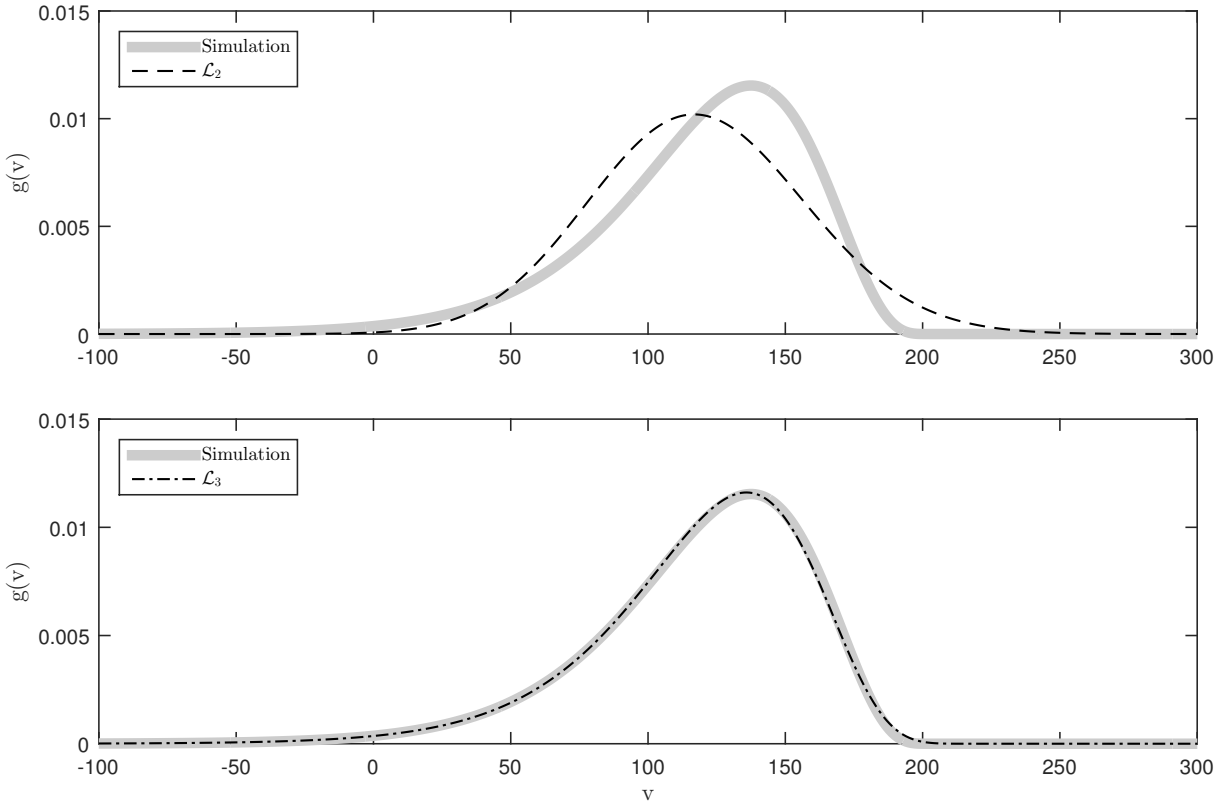

Figure 3: PDF of the simulated NPV, and the $\mathcal{L}_2$ and $\mathcal{L}_3$ approximations for a project with intermediate cash flows

variable that represents the NPV of cash flow c_w, and let $V_{\mathbf{c}} = \sum_{w=1}^{n+1} V_w$ denote the random variable that captures the NPV of the project. If $r > 0$, and if $s_w^2 > 0$ for all $w \in \mathbf{N}$, the project NPV converges to a normal distribution, with mean $\mu_{\mathbf{c}}$ and variance $\sigma_{\mathbf{c}}^2$, as the number of stages increases.

Note that Theorem 5 also applies in a more general context where stages are not necessarily executed in sequence. In fact, Theorem 5 holds as long as a sufficient number of cash flows are incurred during the lifetime of a project.

We use an example to illustrate Theorem 5. The example project has n stages with gamma-distributed durations with shape and scale parameters k_i and τ_i, respectively. Cash outflows are incurred at the start of odd stages. Cash inflows, on the other hand, are obtained at the start of even stages, and upon completion of the project. The discount rate r equals $0.1n^{-1}$. Table 6 summarizes the data of the example project. Fig. 4 shows the simulated and the approximate PDF of the distribution of the project NPV. Next to the lognormal $\mathcal{L}_3$ approximation, we now also include a normal approximation that has mean $\mu_{\mathbf{c}}$ and variance $\sigma_{\mathbf{c}}^2$, and that is denoted by $\mathcal{N}$. We observe that, as n increases, the project NPV converges to a normal distribution, and the accuracy of the $\mathcal{N}$ approximation improves. Even so, the $\mathcal{L}_3$ approximation still performs better due to the extra moment matched. These findings are confirmed by Table 7 that reports on the moments of the NPV distribution, and on the Kolmogorov–Smirnov test statistic. For reference, we have also included the CPU time required to run the simulation (with 1 billion replications) and to calculate the moments using the closed-form expressions provided in Table 3. Both the simulation as well as the exact approach were implemented in Visual Studio C++. Although the simulation model yields good accuracy (also for a lower number of replications), it can hardly compete with an exact, closed-form approach that requires less than a second of CPU time when 100 stages are considered. In addition, most of the computation time is spent on calculating the cokurtosis matrix. Our approach, however, only requires that the first three moments are specified (i.e., there is no need to determine the kurtosis). If only three moments are calculated, the required CPU time drops to 0.046 seconds (for $n = 100$). Even if, for very large n, the computation of the coskewness matrix becomes too time consuming, it suffices to calculate only the first two moments as the $\mathcal{N}$ approximation becomes more accurate as n increases (i.e., for large n, it is no longer necessary to calculate the coskewness matrix).

6 Optimal sequence of stages

The problem of finding the optimal sequence (that maximizes the eNPV over all possible sequences) can be seen as a special case of the stochastic NPV maximization problem (SNPV). The SNPV tries to maximize the eNPV of a project with n stages that do not necessarily have to be scheduled in series. A solution to the SNPV is a policy that schedules stages such that the eNPV of the project (i.e., the expected sum of the discounted cash flows that are incurred during the lifetime of the project) is maximized. The SNPV has been considered by, among others, Sobel et al. (2009), Creemers et al. (2010), and Wiesemann et al. (2010). For a review of the literature on the SNPV, refer to Wiesemann and Kuhn (2015).

Another related problem is the LCFDP; a variant of the Sequential Testing Problem (STP) where n precedence-related tests have to be scheduled such that the expected cost

$$c_w = \begin{cases} 250 & \text{if } w \text{ is even} \\ -250 & \text{if } w \text{ is odd} \end{cases}$$

$$k_w = \begin{cases} 0.5 & \text{if } w \in \{1, 6, 11, \ldots\} \\ 1.0 & \text{if } w \in \{2, 7, 12, \ldots\} \\ 1.5 & \text{if } w \in \{3, 8, 13, \ldots\} \\ 2.0 & \text{if } w \in \{4, 9, 14, \ldots\} \\ 2.5 & \text{if } w \in \{5, 10, 15, \ldots\} \end{cases}$$

$$f_w(t) = \begin{cases} \text{gamma} & \text{if } w \in \{1, 6, 11, \ldots\} \\ \text{exponential} & \text{if } w \in \{2, 7, 12, \ldots\} \\ \text{gamma} & \text{if } w \in \{3, 8, 13, \ldots\} \\ \text{Erlang} & \text{if } w \in \{4, 9, 14, \ldots\} \\ \text{gamma} & \text{if } w \in \{5, 10, 15, \ldots\} \end{cases}$$

$$\tau_w = \begin{cases} 2.0 & \text{if } w \text{ is even} \\ 1.0 & \text{if } w \text{ is odd} \end{cases}$$

$$p = 1000$$
$$r = 0.1n^{-1}$$

Table 6: Data of the example project with n stages and intermediate cash flows

	$n = 10$			$n = 30$			$n = 100$		
	Sim	$\mathcal{N}$	$\mathcal{L}_3$	Sim	$\mathcal{N}$	$\mathcal{L}_3$	Sim	$\mathcal{N}$	$\mathcal{L}_3$
μ	783.04	783.04	783.04	782.16	782.16	782.16	781.86	781.86	781.86
σ^2	2,584	2,584	2,584	875	875	875	264	264	264
γ	-0.361	0.0	-0.361	-0.211	0.0	-0.211	-0.117	0.0	-0.116
θ	3.1159	3.0	3.1162	3.0415	3.0	3.0402	3.0769	3.0	3.0122
K–S	NA	0.0297	0.0032	NA	0.0155	0.0001	NA	0.0080	0.0003
CPU (s)	2,250	0.000		11,279	0.015		90,955	0.967	

Table 7: Accuracy of the $\mathcal{N}$ and $\mathcal{L}_3$ approximations for the NPV of a project with intermediate cash flows and n stages

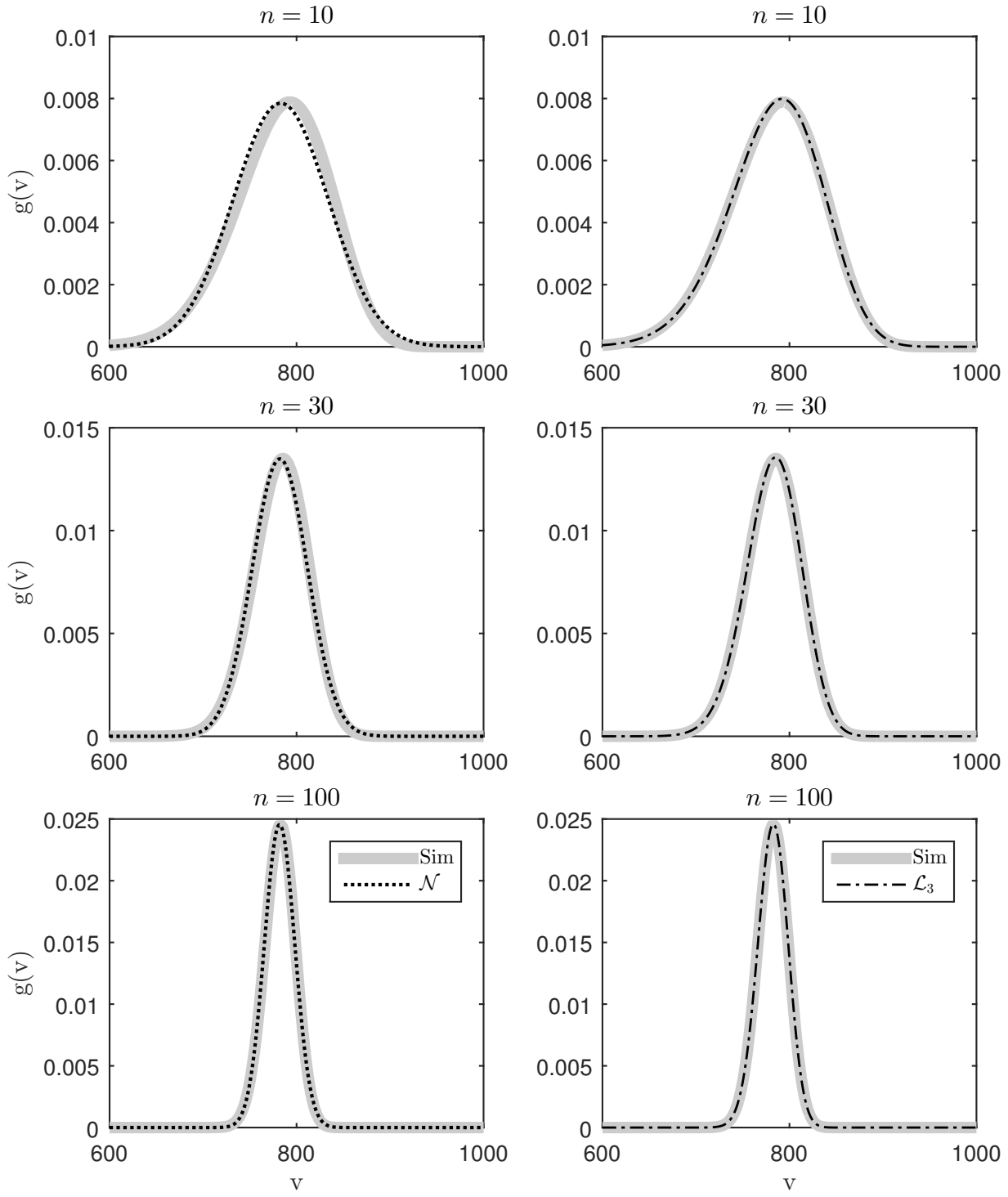

Figure 4: PDF of the simulated NPV, and the $\mathcal{N}$ and $\mathcal{L}_3$ approximations for various number of stages

of the diagnosis of a system is minimized. Each test $w : w \in \mathbf{N} = \{1, \ldots, n\}$ has a known cost c_w and a failure probability q_w. In this article, we consider the setting where a single test results in the failure of the system (i.e., we study so-called n-out-of-n or serial systems). For such a setting, it can be shown that there exists a full order sequence of tests in $\mathbf{N}$ that is globally optimal. The LCFDP is related to the R&D project scheduling problem studied in De Reyck and Leus (2008), who show that their problem is NP-hard. It follows that the LCFDP is also NP-hard if tests are precedence-related (Wei et al., 2013). The LCFDP arises in many practical contexts, such as the inspection of containers arriving at a port (Madigan et al., 2011) and the identification of toxic chemicals (Gowtham et al., 2012). A literature review on the STP in general, and on the LCFDP in particular, may be found with Ünlüyurt (2004), Wei et al. (2013), and Coolen et al. (2014).

We define the serial SNPV as the problem to find the optimal sequence of stages that maximizes the eNPV over all possible sequences. For serial projects, the serial SNPV is equivalent to the SNPV. In what follows, we show that: (1) the LCFDP is equivalent to the serial SNPV, (2) a well-known result from the literature on the LCFDP may be used to obtain the optimal solution to the serial SNPV if stages are not precedence related, and (3) methods for solving the SNPV can also be used to solve the LCFDP. In addition, we perform a computational experiment that shows that the state-of-the-art procedure for solving the SNPV (a more general problem where stages are allowed to be executed in parallel) outperforms the state-of-the-art procedure for solving the LCFDP.

6.1 Equivalence of the serial SNPV and the LCFDP

Let $\mathbf{s} = \{s_1, \ldots, s_n\}$ denote a sequence of n stages, where s_w is the stage at position w in the sequence. As shown in Section 5, the eNPV of a sequence $\mathbf{s}$ is given by:

$$c_{s_1} + \sum_{w=2}^{n+1} \phi_{1,(w-1)}(r) c_{s_w},$$

where $\phi_{1,w}$ is the discount factor for a sequence of stages $\{s_1, \ldots, s_w\}$. The objective of the serial SNPV is to find a sequence that maximizes:

$$c_{s_1} + \left(\sum_{x=2}^{n} c_{s_x} \prod_{w=1}^{x-1} \phi_{s_w}(r) \right) + \left(c_{n+1} \prod_{w=1}^{n} \phi_w(r) \right),$$

where the latter term is a constant that does not depend on the sequence of stages (i.e., the latter term may be ignored when making sequencing decisions), and hence the objective reduces to:

$$\max_{\mathbf{s}} c_{s_1} + \left(\sum_{x=2}^{n} c_{s_x} \prod_{w=1}^{x-1} \phi_{s_w}(r) \right). \tag{7}$$

The objective of the LCFDP, on the other hand, is to find a sequence of tests that minimizes the cost of the sequential diagnosis of a system, and is given by:

$$\max_{\mathbf{s}} c_{s_1} + \left(\sum_{x=2}^{n} c_{s_x} \prod_{w=1}^{x-1} p_{s_w} \right), \tag{8}$$

w	c_w	λ_w	ϕ_w	$c_w(1-\phi_w)^{-1}$	s_w
1	-10	1/2	0.833	-60	5
2	10	1/4	0.714	35	2
3	-15	1/6	0.625	-40	3
4	20	1/8	0.556	45	1
5	-36	1/20	0.250	-48	4

p	100
r	0.1

Table 8: Data of the example project

where $p_w = 1 - q_w$ is the success probability of test w. Eq. (7) and Eq. (8) are equivalent if $\phi_w(r) \equiv p_w$ for all $w : w \in \mathbf{N}$. We conclude that the LCFDP is equivalent to the serial SNPV, which in turn is a special case of the SNPV.

6.2 Optimal sequence

In the absence of precedence relationships, Boothroyd (1960) has shown that the optimal solution to the LCFDP is a sequence that arranges tests in (increasing) order of their ratio of cost over failure probability. Therefore, for the LCFDP, the optimal sequence can be determined in polynomial time, and if:

$$\frac{c_{s_1}}{q_{s_1}} \leq \frac{c_{s_2}}{q_{s_2}} \leq \ldots \leq \frac{c_{s_n}}{q_{s_n}},$$

then $\mathbf{s} = \{s_1, s_2, \ldots, s_n\}$ is optimal.

The above result can also be used to determine the optimal sequence that maximizes the eNPV of a project where stages are not precedence related. More precisely, in the absence of precedence relationships, sequence $\mathbf{s} = \{s_1, s_2, \ldots, s_n\}$ is optimal if:

$$\frac{c_{s_1}}{1-\phi_{s_1}(r)} \leq \frac{c_{s_2}}{1-\phi_{s_2}(r)} \leq \ldots \leq \frac{c_{s_n}}{1-\phi_{s_n}(r)}.$$

To illustrate this finding, we use an example project with 5 stages that have exponentially-distributed durations with rate parameter $\lambda_w : w \in \mathbf{N} = \{1, 2, 3, 4, 5\}$. The data of the example project are summarized in Table 8. Note that ϕ_w is the moment-generating function of f_w about $-r$, and is given by:

$$\phi_w(r) = \frac{\lambda_w}{\lambda_w + r} \tag{9}$$

for an exponentially-distributed duration with rate parameter λ_w. The optimal sequence executes stages 4, 2, 3, 5, and 1 in series, and yields an eNPV of 15.22.

6.3 Solving the LCFDP

We solve the LCFDP using the state-of-the-art procedure of Creemers (2017) that was designed to solve the SNPV. Creemers (2017) assumes that stage durations are exponentially distributed, and uses a Continuous-Time Markov Chain (CTMC) to model the state space. A backward Stochastic Dynamic Program (SDP) is used to obtain the globally optimal policy that maximizes the eNPV of a project (note that a solution to the SNPV is a scheduling policy rather than a sequence). In contrast to most of the literature on the scheduling of Markovian PERT networks (i.e., PERT networks where stages have exponentially-distributed durations), Creemers (2017) does not use Uniformly Directed Cuts (UDCs) to structure the state space, nor does he represent the state of the system using sets of idle, ongoing, and finished stages (see e.g., Creemers et al., 2010). Instead, Creemers (2017) uses arrays to store states that are defined only by the set of finished stages. The cardinality of a state (i.e., the number of finished stages) determines the array in which the state is stored (there is one array for each number of finished stages). Because states with cardinality $(i + 1)$ are only accessible from states with cardinality i, at most two arrays need to be stored in memory (i.e., after calculating all value functions of states with cardinality i, states with cardinality $(i + 1)$ are no longer needed, and they can be removed from memory). Together with a stricter definition of the state space (by only using the set of finished stages), this more efficient structuring of the state space results in a significant reduction of memory and computational requirements (when compared to other methods that solve the SNPV).

In order to solve an instance of the LCFDP by means of a procedure for solving the SNPV, tests first need to be "transformed" into stages. As explained in Section 6.1, the serial SNPV and the LCFDP are equivalent if $\phi_w(r) \equiv p_w$ for all $w : w \in \mathbf{N}$. In the procedure of Creemers (2017), stages are assumed to have exponentially-distributed durations, and therefore, the discount factor of a stage w is given by Eq. (9). As such, a test w with cost c_w and failure probability q_w can be transformed into a stage w with cost c_w and rate parameter:

$$\lambda_w = \frac{q_w r}{1 - q_w}.$$

After transforming all tests into stages, the procedure of Creemers (2017) can be used to solve an instance of the LCFDP. Note that, in order to make sure that stages are executed in a sequence, we impose a resource constraint (i.e., each stage requires one unit of a renewable resource that has unit availability).

We compare the performance of the above approach with the state-of-the-art procedure of Wei et al. (2013). Wei et al. (2013) propose both a Branch-and-Bound (B&B) as well as an SDP procedure to solve the k-out-of-n STP (i.e., at least k out of n components should be functional, otherwise the system is down). The SDP procedure significantly outperforms the B&B, and in what follows, we will compare its performance with that of the procedure of Creemers (2017). Note that, if we let $k = n$, the k-out-of-n STP corresponds to the LCFDP as defined by Boothroyd (1960). Similar to the procedures of Creemers et al. (2010) and Coolen et al. (2014), The SDP procedure of Wei et al. (2013) uses UDCs to structure the state space. Once the states of a UDC are no longer required, the UDC is discarded, and the memory is freed.

We use the instances of Wei et al. (2013) to compare the performance of both SDP procedures. Wei et al. (2013) use RanGen (Demeulemeester et al., 2003) to generate three

	Creemers (2017)			Wei et al. (2013)		
n	OS = 0.8	OS = 0.6	OS = 0.4	OS = 0.8	OS = 0.6	OS = 0.4
10	10	10	10	10	10	10
20	10	10	10	10	10	10
30	10	10	10	10	10	10
40	10	10	10	10	10	10
50	10	10	10	10	10	10
60	10	10	10	10	10	10
70	10	10	10	10	10	9
80	10	10	10	10	10	0
90	10	10	0	10	10	0
100	10	10	0	10	10	0
110	10	10	0	10	7	0
120	10	9	0	10	0	0

Table 9: Number of instances solved (out of 10) by the procedures of Creemers (2017) and Wei et al. (2013)

data sets that each contain 10 instances for each value of $n : n \in \{10, 20, \ldots, 120\}$ and for each value of OS : OS $\in \{0.4, 0.6, 0.8\}$ (where OS is the Order Strength; a measure of the density of the project network). Each data set has different failure probabilities. Because failure probabilities do not impact the computational performance of the SDP procedures, we select the data set that has the lowest failure probabilities (i.e., failure probabilities are drawn from a uniform distribution with a minimum of 0.8 and a maximum of 1). Both procedures are tested on an Intel 3.3 GHz desktop computer with 16GB of RAM.

Table 9 reports on the number of instances solved by each approach, and shows that the procedure of Creemers (2017) outperforms the procedure of Wei et al. (2013). This can be explained by the more efficient memory-management techniques adopted by Creemers (2017). When comparing average computation times (in seconds) on instances that could be solved by Wei et al., however, Table 10 shows that the procedure of Creemers (2017) is somewhat slower (26.5% on average). Since the procedure of Creemers (2017) was designed to solve the SNPV (a more general problem where stages are allowed to be executed in parallel), this does not come as a surprise. In addition, it is clear that memory requirements rather than CPU times are the bottleneck for the problem at hand. Even for larger problems, Table 11 shows that the procedure of Creemers (2017) is able to solve instances within a reasonable time frame. We conclude that the procedure of Creemers (2017) is the current state-of-the-art for solving the LCFDP/serial SNPV.

7 NPV of a general project with multiple stages that are scheduled using a scheduling policy

In this section, we illustrate how to determine the NPV of a general project with multiple stages that are scheduled using a scheduling policy. To this end, we use an example project

	Creemers (2017)			Wei et al. (2013)		
n	OS = 0.8	OS = 0.6	OS = 0.4	OS = 0.8	OS = 0.6	OS = 0.4
10	0	0	0	0	0	0
20	0	0	0	0	0	0
30	0	0	0.01	0	0	0.01
40	0	0.01	0.17	0	0.01	0.13
50	0	0.04	1.99	0	0.02	1.44
60	0.01	0.19	25.4	0	0.11	19.3
70	0.01	0.94	178	0.01	0.58	156
80	0.03	4.00	–	0.01	2.40	–
90	0.05	15.0	–	0.02	9.48	–
100	0.11	77.1	–	0.05	45	–
110	0.24	223	–	0.10	151	–
120	0.56	–	–	0.24	–	–

Table 10: Comparison of average computation time (in seconds) for the instances that could be solved by Wei et al. (2013)

n	OS = 0.8	OS = 0.6	OS = 0.4
10	0	0	0
20	0	0	0
30	0	0	0.01
40	0	0.01	0.17
50	0	0.04	1.99
60	0.01	0.19	25.4
70	0.01	0.94	205
80	0.03	4.00	2,013
90	0.05	15.0	–
100	0.11	77.1	–
110	0.24	323	–
120	0.56	1,009	–

Table 11: Average computation time (in seconds) for the procedure of Creemers (2017)

w	c_w	d_w	λ_w	Predecessor
1	-50	1	1.0	$\emptyset$
2	-20	2	0.5	$\emptyset$
3	-10	2	0.5	$\{2\}$

p	200
r	0.1

Table 12: Data of the example project with general structure

with three stages. The first stage can be executed together with any of the other two stages. The third stage, however, can only start after the second stage has finished. At the start of each stage w, a cash flow c_w is incurred. Upon completion of the project, a payoff $p = 200$ is obtained. The discount rate $r = 0.1$. Each stage w has an exponentially-distributed duration with rate parameter λ_w. The data of the example project are summarized in Table 12.

There are several policies that allow to schedule the stages during the execution of the project. We discuss two: the Early-Start (ES) policy and the optimal policy. The ES policy first starts stages 1 and 2, and upon completion of stage 2, stage 3 is started (note that stage 1 can still be ongoing at that time). The optimal policy maximizes the eNPV of the project, and starts stages 1 and 3 only upon completion of stage 2. In what follows, we first discuss how to determine the NPV of the ES policy.

In the ES policy, cash flows c_1 and c_2 are incurred at the start of the project (i.e., no discounting is required). Cash flow c_3 is incurred upon completion of stage 2 (i.e., when stage 3 starts). Hence, the NPV of cash flow c_3 can be determined using factor $\phi_2(r) = \lambda_2(\lambda_2+r)^{-1}$. The NPV of payoff p, on the other hand, is only obtained if both stage 1 and the series of stage 2 and 3 are finished. In other words, the distribution of time until we obtain payoff p is the distribution of the maximum of an exponential distribution with rate parameter λ_1 and an Erlang distribution with two phases that both have rate parameter $\lambda_2 = \lambda_3$. The corresponding factor can be determined as follows:

$$\phi_p(r) = \left(\frac{\lambda_2}{\lambda_2 + r}\right)^2 - r\frac{\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^2}{\lambda_1 + r}.$$

The eNPV of the ES policy can then be computed using Theorem 4, and amounts to 58.78. Note that Theorem 4 can always be used to calculate the exact eNPV of any project/policy as the eNPV does not require to calculate cross moments. For higher moments of the NPV distribution, however, cross moments may need to be approximated if the sequence of cash flows is not fixed (i.e., if the sequence of cash flows is probabilistic). In the ES policy, for instance, the NPV of payoff p can be impacted by stage 1 (if it takes long enough) as well as by stages 2 and 3. In other words, there are two possible sequences that impact the NPV of payoff p. Theorem 4 assumes that there is only a single sequence, and as a result, we have to approximate the cross moments between payoff p and the cash flows of the preceding stages. In our example, we can follow three different approaches:

- We can assume that stage 1 was finished before stage 2. In this case, the higher

	Exact	Simulation	Stage 1 finished	Stage 1 ongoing	Weighted approach
μ	58.780	58.780	58.780	58.780	58.780
σ^2	971.08	971.10	967.25	969.50	968.00
γ	-0.580	-0.581	-0.579	-0.580	-0.580
θ	2.8549	2.8552	2.8361	2.8695	2.8457

Table 13: Accuracy of different approaches to determine the NPV of the ES policy

moments of the NPV of payoff p are determined solely by stages 2 and 3 (i.e., the sequence consists of stages 2 and 3).

- We can assume that stage 1 was still ongoing after completion of stage 2. In this case, the higher moments of the NPV of payoff p are determined by stage 2 and the maximum of stages 1 and 3 (i.e., the sequence consists of stage 2 and the maximum of stages 1 and 3).
- We can adopt a weighted approach where both of the previous approaches are weighed depending on their probability of occurrence.

The moments corresponding to each of the above approaches are given in Table 13. Table 13 also reports the exact and simulated moments. We observe that the error is relatively small, however, further research is required to assess the accuracy for larger projects.

Next, we observe the optimal policy. In the optimal policy, there is only one possible sequence (stage 2 is followed by stages 1 and 3 that are executed in parallel), and as result, the moments of the project NPV can be determined in an exact manner using Theorem 4. Whereas cash flow c_2 is incurred at the start of the project, cash flows c_1 and c_3 are incurred upon completion of stage 2 (i.e., factor $\phi_2(r)$ applies). Payoff p is obtained after both stage 1 and stage 3 are finished (i.e., after the maximum of two exponential durations). The corresponding factor is given by:

$$\phi_p(r) = \frac{\lambda_1\lambda_3(\lambda_1+\lambda_3+r)}{(\lambda_1+r)(\lambda_3+r)(\lambda_1+\lambda_3+r)}.$$

The moments of the NPV distribution of the optimal policy are $\mu = 64.154$, $\sigma^2 = 698.43$, $\gamma = -0.5342$, and $\theta = 2.9649$.

We also assess the accuracy of the $\mathcal{L}_3$ approximation. Table 14 summarizes the results, and Fig. 5 shows the simulated and the approximate PDF of the NPV distribution of the ES and the optimal policy. We conclude that the $\mathcal{L}_3$ approximation is once more very accurate.

8 Model extensions

In this section, we discuss two model extensions. A first extension allows stages (and hence projects) to fail. Stage/project failure is common in the literature on R&D projects (Sommer 2004; De Reyck and Leus 2008; Creemers et al. 2015), and can easily be incorporated in our

	ES policy		Optimal policy	
	Simulation	$\mathcal{L}_3$	Simulation	$\mathcal{L}_3$
μ	58.780	58.780	64.154	64.154
σ^2	971.10	971.08	698.45	698.43
γ	-0.581	-0.580	-0.534	-0.534
θ	2.8552	3.6048	2.9654	3.5116
K–S	NA	0.0267	NA	0.0190

Table 14: Accuracy of the $\mathcal{L}_3$ approximations to model the NPV of the ES policy and the optimal policy

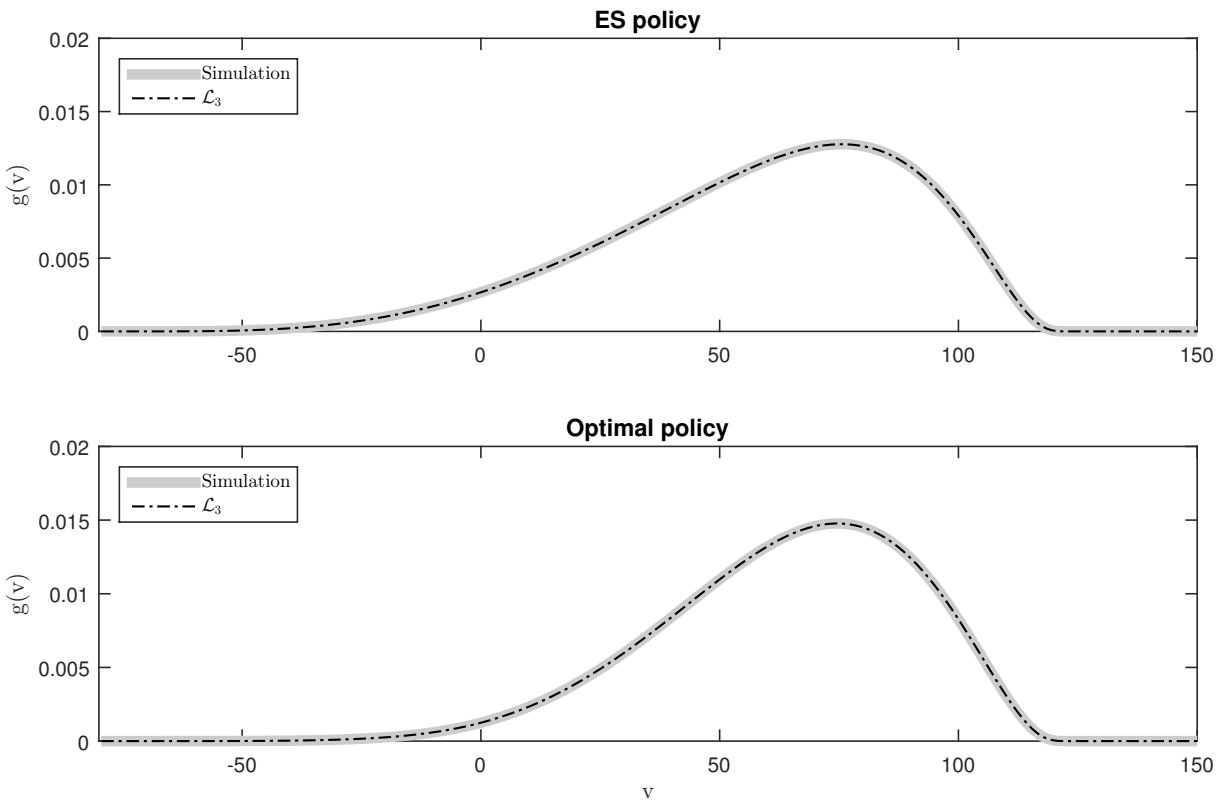

Figure 5: PDF of the simulated NPV and the $\mathcal{L}_3$ approximation of the ES policy and the optimal policy

approach. We need only to redefine factor $\phi_w(r)$:

$$\phi_w(r) = p_w \phi_w^*(r),$$

where p_w is the probability of success of stage w, and $\phi_w^*(r)$ is the factor given by Eq. (2) (i.e., the factor that does not take into account stage/project failure).

A second model extension allows for different discount rates to be applied during different stages of the project. This extension requires a redefinition of factor $\phi_{w,x}(r)$:

$$\phi_{w,x}(\mathbf{r}) = \prod_{y=w}^{x} \phi_y(r_y),$$

where r_y is the discount rate that applies for stage y, and $\mathbf{r} = \{w, w+1, \ldots, x\}$ is the vector of discount rates that apply to stages $y : w \leq y \leq x$.

9 Conclusions

In this article, we considered projects with multiple stages $w : w \in \mathbf{N} = \{1, 2, \ldots, n\}$ that are executed in sequence. Each stage $w : w \in \mathbf{N}$ has a random duration T equipped with probability function $f(t)$. A cash flow c_w (positive or negative) may be incurred upon the start of stage w, and a payoff p is obtained at the completion of the project. We use continuous compounding and a discount rate r to determine the NPV of a project.

Our main contributions can be summarized as follows: (1) we obtain exact, closed-form expressions for the moments of the NPV of a project, (2) we develop a highly accurate closed-form approximation of the distribution of the project NPV, (3) we show that the NPV of a single cash flow converges to a (reflected) lognormal distribution if the cash flow is not incurred during the early stages of the project, (4) we show that the NPV of a project converges to a normal distribution if a sufficient number of cash flows are incurred during the lifetime of the project, (5) we show that the problem of finding the optimal sequence of stages is equivalent to the LCFDP, (6) we show that the optimal sequence of stages can be obtained in polynomial time if stages are not precedence related, (7) we perform a computational experiment to identify the state-of-the-art procedure to determine the optimal sequence of stages if they are precedence related, and (8) we illustrate how our approach can be used to determine the moments and the NPV of a general project where stages are scheduled using a scheduling policy.

Our work can be directly applied in the fields of project selection, project portfolio management, and project valuation. In these fields, a project is often seen as a sequence of stages with intermediate cash flows (including a payoff that is obtained upon the successful completion of the project). Project selection/investment decisions can be made based on the eNPV and the risk of a project. The risk of a project/an investment is often modeled using the variance of the NPV. Other measures of risk include the skewness and/or kurtosis of the NPV, and the probability to have a negative NPV. Until now, Monte Carlo simulation was the only tool available to obtain estimates for these measures. Our work, however, offers a valid alternative to Monte Carlo simulation, and allows to obtain a highly accurate approximation of the NPV distribution, and an exact characterization of its moments.

In the (more operational) field of project scheduling, our work is related to CPM/PERT in the sense that we also focus on a single sequence of stages, and also use normal (log-normal) approximations. As a result, the limitations of our work are similar to those of CPM/PERT. Therefore, future research should further investigate methods that have been used to generalize CPM/PERT, and that may also be applied here (e.g., network transformations/reductions and bounding procedures). In addition, the scheduling of stages of a general project is also an important direction for future research. Determining the optimal release dates to maximize the eNPV of a project is also a topic worthy of study. Other directions for extending our results are the inclusion of time-dependent cash flows and interdependent stage durations.

Appendix. Proofs

Proof of Lemma 1. The proof follows from the definition of the MGF:

$$\begin{aligned} M_T(u) &= \int_0^\infty f(t)e^{ut}\mathrm{d}t, \quad \text{if } T \text{ is continuous,} \\ &= \sum_{t=0}^\infty f(t)e^{ut}, \quad \text{if } T \text{ is discrete.} \end{aligned}$$

□

Proof of Theorem 1. Let V denote the random variable that represents the NPV of a cash flow c that is incurred at time T. The MGF of V is:

$$M_V(u) = \sum_{i=0}^\infty \frac{u^i m_i}{i!},$$

where m_i is the ith raw moment of the NPV distribution:

$$\begin{aligned} m_i &= \int_0^\infty f(t)(e^{-rt})^i\mathrm{d}t = \phi(ir), \quad \text{if } T \text{ is continuous,} \\ &= \sum_{t=0}^\infty f(t)(e^{-rt})^i = \phi(ir), \quad \text{if } T \text{ is discrete.} \end{aligned}$$

Using these raw moments, we can obtain the mean, variance, skewness, kurtosis, and even higher-order moments of the NPV of cash flow c. □

Proof of Theorem 2. If we solve Eq. (1) for t, we obtain:

$$t_v = \ln\left(\frac{c}{v}\right) r^{-1}.$$

where t_v is the time at which cash flow c needs to be incurred in order to obtain NPV v for a given discount rate r. As a result, $F(t_v)$ not only represents the probability to have a time $t \leq t_v$, but it also represents the probability to have an NPV $\geq v$. □

Proof of Lemma 2. Factor $\phi_{1,n}(r)$ can be obtained as follows:

$$\begin{aligned}\phi_{1,n}(r) &= \int_0^\infty \cdots \int_0^\infty f_1(t_1)e^{-rt_1} \cdots f_n(t_n)e^{-rt_n} \mathrm{d}t_1 \cdots \mathrm{d}t_n, \\ &= \phi_1(r) \int_0^\infty \cdots \int_0^\infty f_2(t_2)e^{-rt_2} \cdots f_n(t_n)e^{-rt_n} \mathrm{d}t_2 \cdots \mathrm{d}t_n, \\ &\cdots \\ &= \prod_{w \in \mathbf{N}} \phi_w(r).\end{aligned}$$

In general, let $\phi_{w,x}(r)$ denote the factor for stages w to x, where $x \geq w$:

$$\phi_{w,x}(r) = \prod_{y=w}^{x} \phi_y(r).$$

□

Proof of Lemma 3. The proof follows from the Central Limit Theorem (CLT). □

Proof of Theorem 3. The proof is a direct application of Theorem 2 and Lemma 3. If n is sufficiently large, the duration of the project is normally distributed, and if $F(t)$ is a normal distribution function, $G(v)$ can be expressed as follows:

$$G(v) = \frac{1}{2} + \frac{1}{2} \operatorname{Erf}\left(\frac{\ln(v) - (\ln(p) - rd_{\mathbf{N}})}{\sqrt{2} rs_{\mathbf{N}}}\right).$$

When substituting $\ln(p) - rd_{\mathbf{N}}$ by α and $rs_{\mathbf{N}}$ by β, we get:

$$G(v) = \frac{1}{2} + \frac{1}{2} \operatorname{Erf}\left(\frac{\ln(v) - \alpha}{\sqrt{2}\beta}\right),$$

which is the CDF of the lognormal distribution with location parameter $\alpha = \ln(p) - rd_{\mathbf{N}}$ and scale parameter $\beta = rs_{\mathbf{N}}$. Note that, if p is negative, the NPV distribution of p converges to a reflected lognormal distribution. □

Proof of Proposition 1. Consider a (reflected) lognormal distribution with location and scale parameters α and β, respectively. The mean and SCV of that distribution are given by:

$$\mu_{\mathcal{L}2} = e^{\alpha + 0.5\beta^2}, \tag{10}$$

$$\eta^2_{\mathcal{L}2} = e^{(2\alpha + \beta^2)} \left(e^{\beta^2} - 1\right) e^{-2(\alpha + 0.5\beta^2)}. \tag{11}$$

The unique solution for β can easily be obtained by solving Eq. (11):

$$\beta = \sqrt{\ln(1 + \eta^2_{\mathcal{L}2})}.$$

Note that, as long as $\sigma^2_{\mathcal{L}2} > 0$, then $\eta^2_{\mathcal{L}2} > 0$, and therefore $\beta > 0$. Given β, the unique solution for α can easily be obtained by solving Eq. (10):

$$\alpha = \ln(\mu_{\mathcal{L}2}) - 0.5\beta^2.$$

Note that, if μ is negative, a reflected lognormal distribution can be used to match the NPV distribution. We conclude that α and β are the unique solution to Eqs. (10–11), and that they are well defined for all $\mu_{\mathcal{L}2}, \sigma^2_{\mathcal{L}2} \in \mathbb{R}$, and for $\sigma^2_{\mathcal{L}2} > 0$. □

Proof of Proposition 2. First, we obtain the unique solution for β from Eq. (6). We have:

$$\gamma_{\mathcal{L}3} = \delta(2 + e^{\beta^2})\sqrt{e^{\beta^2} - 1},$$
$$\gamma^2_{\mathcal{L}3} = e^{3\beta^2} + 3e^{2\beta^2} - 4.$$

If we let $q = e^{\beta^2}$ and $l = 4 + \gamma^2_{\mathcal{L}3}$, we obtain the following cubic equation:

$$q^3 + 3q^2 - l = 0.$$

The discriminant of $q^3 + 3q^2 - l$ is $\Delta = -27(l-4)l$, and is always negative if $\gamma_{\mathcal{L}3} \neq 0$. For cubic equations, if $\Delta < 0$, the equation has one unique real root and two non-real complex conjugate roots. The unique real root of $q^3 + q^2 - l$ is:

$$q = \frac{2^{1/3}}{\left(-2 + l + \sqrt{-4l + l^2}\right)^{1/3}} + \frac{\left(-2 + l + \sqrt{-4l + l^2}\right)^{1/3}}{2^{1/3}} - 1.$$

After substituting $q = e^{\beta^2}$ and $l = 4 + S^2_{\mathcal{L}3}$, we obtain the unique, real solution for β:

$$\beta = \sqrt{\ln\left(\frac{2^{1/3}}{\left(2 + \gamma^2_{\mathcal{L}3} + \sqrt{4\gamma^2_{\mathcal{L}3} + \gamma^4_{\mathcal{L}3}}\right)^{1/3}} + \frac{\left(2 + \gamma^2_{\mathcal{L}3} + \sqrt{4\gamma^2_{\mathcal{L}3} + \gamma^4_{\mathcal{L}3}}\right)^{1/3}}{2^{1/3}} - 1\right)}.$$

Given β, the unique solution for α can easily be obtained by solving Eq. (5):

$$\alpha = 0.5\left(\ln\left(\frac{\sigma^2_{\mathcal{L}3}}{e^{\beta^2} - 1}\right) - \beta^2\right).$$

Given α and β, the unique solution for κ can easily be obtained by solving Eq. (4):

$$\kappa = \mu_{\mathcal{L}3} - \delta e^{\alpha + 0.5\beta^2}.$$

We conclude that α, β, and κ are the unique solution to Eqs. (4–6), and that they are well defined for all $\mu_{\mathcal{L}3}, \sigma^2_{\mathcal{L}3}, \gamma_{\mathcal{L}3} \in \mathbb{R}$, $\sigma^2_{\mathcal{L}3} > 0$, and for $\gamma_{\mathcal{L}3} \neq 0$. □

Proof of Theorem 4. The covariance between the NPV of cash flow c_x and the NPV of an earlier cash flow c_w is given by:

$$\mathbf{\Sigma_c}(w,x) = \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(c_w e^{-r\left(\sum_{y=1}^{w-1} t_y\right)} - \mu_w \right) \left(c_x e^{-r\left(\sum_{y=1}^{x-1} t_y\right)} - \mu_x \right) \mathrm{d}t_1 ... \mathrm{d}t_{x-1}.$$

Which can be rewritten as a sum of 4 parts:

$$\begin{aligned}
\mathbf{\Sigma_c}(w,x) = & \\
& \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(c_w e^{-r\left(\sum_{y=1}^{w-1} t_y\right)} c_x e^{-r\left(\sum_{y=1}^{x-1} t_y\right)} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1} \\
& - \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(\mu_x c_w e^{-r\left(\sum_{y=1}^{w-1} t_y\right)} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1} \\
& - \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(\mu_w c_x e^{-r\left(\sum_{y=1}^{x-1} t_y\right)} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1} \\
& + \int_0^\infty \cdots \int_0^\infty \prod_{y=1}^{x-1} f_y(t_y) \left(\mu_w \mu_x \right) \mathrm{d}t_1 \cdots \mathrm{d}t_{x-1}.
\end{aligned}$$

After application of Lemma 2, we get:

$$\begin{aligned}
& \mathbf{\Sigma_c}(w,x) = \\
& c_w c_x \phi_{1,w-1}(2r) \phi_{w,x-1}(r) \\
& - \mu_x c_w \phi_{1,w-1}(r) \\
& - \mu_w c_x \phi_{1,w-1}(r) \phi_{w,x-1}(r) \\
& + \mu_w \mu_x.
\end{aligned}$$

From Theorem 1, we have that $\mu_w = c_w \phi_{1,w-1}(r)$ and $\mu_x = c_x \phi_{1,x-1}(r)$, and therefore:

$$\begin{aligned}
& \mathbf{\Sigma_c}(w,x) = \\
& c_w c_x \phi_{1,w-1}(2r) \phi_{w,x-1}(r) \\
& - c_w c_x \phi_{1,w-1}(r) \phi_{1,w-1}(r) \phi_{w,x-1}(r) \\
& - c_w c_x \phi_{1,w-1}(r) \phi_{1,w-1}(r) \phi_{w,x-1}(r) \\
& + c_w c_x \phi_{1,w-1}(r) \phi_{1,w-1}(r) \phi_{w,x-1}(r).
\end{aligned}$$

Which, finally, can be simplified to:

$$\mathbf{\Sigma_c}(w,x) = c_w c_x \phi_{w,x-1}(r) \left(\phi_{1,w-1}(2r) - \phi_{1,w-1}^2(r) \right). \tag{12}$$

The same approach can be used to determine the coskewness, the cokurtosis, and even the higher-order cross moments of the NPV of the cash flows that are incurred during the lifetime of the project. □

Proof of Theorem 5. Let $(V) = \{V_1, V_2, \ldots, V_n, V_{n+1}\}$ denote the non-stationary sequence of dependent random variables $V_w : 1 \leq w \leq n+1$. For such a sequence, Bradley and Tone (2015) have shown that a CLT holds if:

- the sequence is strongly mixing,
- the sequence has a maximum correlation that is strictly smaller than 1 for some V_w and V_{w+1} in (V),
- the Lindeberg condition holds.

Several mixing conditions have been defined in the literature (for an overview, refer to Bradley (2005)). In this proof, we will show that sequence (V) is ρ-mixing (which automatically implies that (V) is strongly mixing). A sequence is said to be ρ-mixing if the maximum correlation between two random variables $V_w, V_x \in (V)$ tends to zero for some w and x that are "far apart". We use Eq. (12) to obtain the expression for the correlation between two random variables V_w and V_x:

$$\begin{aligned}\mathrm{Corr}(w,x) &= \frac{\phi_{w,x-1}(r)\left(\phi_{1,w-1}(2r) - \phi^2_{1,w-1}(r)\right)}{\sqrt{\phi_{1,w-1}(2r) - \phi^2_{1,w-1}(r)}\sqrt{\phi_{1,x-1}(2r) - \phi^2_{1,x-1}(r)}},\\ &= \phi_{w,x-1}(r)\sqrt{\frac{\phi_{1,w-1}(2r) - \phi^2_{1,w-1}(r)}{\phi_{1,x-1}(2r) - \phi^2_{1,x-1}(r)}}.\end{aligned}$$

It is easy to verify that $\mathrm{Corr}(w,x) \to 0$ if $\phi_{w,x-1}(r) \to 0$, or if $\phi_{1,w-1}(2r) = \phi^2_{1,w-1}(r)$. If $c_w > 0$, and if at least one stage $z : 1 \leq z < w$ has $s^2_z > 0$, then $\sigma_w > 0$, and it follows from Eq. (3) that $\phi_{1,w-1}(2r) > \phi^2_{1,w-1}(r)$. Therefore, we say that $\mathrm{Corr}(w,x) \to 0$ if and only if $\phi_{w,x-1}(r) \to 0$. From Lemma 2 we know that $\phi_{w,x-1}(r) = \prod_{y=w}^{x-1} \phi_y(r)$. In addition, if $r > 0$, and if $s^2_y > 0$, then $\phi_y(r) < 1$, and $\phi_{w,x-1}(r) \to 0$ if $s^2_y > 0$ for sufficient $y : w \leq y < x$, and for $x - w \to \infty$. In other words, if $r > 0$, and if $s^2_y > 0$ for sufficient $y \in \mathbf{N}$, then sequence (V) is ρ-mixing as $n \to \infty$.

In order to show that sequence (V) satisfies the second condition, we observe the correlation between random variables V_w and V_{w+1}:

$$\begin{aligned}\mathrm{Corr}(w,w+1) &= \phi_w(r)\sqrt{\frac{\phi_{1,w-1}(2r) - \phi^2_{1,w-1}(r)}{\phi_{1,w}(2r) - \phi^2_{1,w}(r)}},\\ &= \frac{\phi_{1,w-1}(2r)\phi^2_w(r) - \phi^2_{1,w}(r)}{\phi_{1,w-1}(2r)\phi_w(2r) - \phi^2_{1,w}(r)}.\end{aligned}$$

A perfect correlation is achieved if $\phi_{1,w-1}(2r) \to 0$, or if $\phi_w(2r) = \phi^2_w(r)$. If $s^2_w > 0$, then $\phi_w(2r) > \phi^2_w(r)$, and as a result, $\mathrm{Corr}(w,w+1) \to 1$ if and only if $\phi_{1,w-1}(2r) \to 0$. If $w \to \infty$, $\phi_{1,w-1}(2r) \to 0$, however, because $\phi_{1,w-1}(2r) > \phi^2_{1,w-1}(r)$, $\phi^2_{1,w-1}(r)$ goes to zero even faster. In addition, $\phi_w(2r) > \phi^2_w(r)$, and therefore, the maximum correlation between random variables V_w and V_{w+1} is always strictly smaller than 1.

To complete the proof, we still need to show that the Lindeberg condition holds. Instead of verifying the Lindeberg condition itself, we show that sequence (V) satisfies the more strict

Lyapunov condition. The Lyapunov condition requires that all random variables $V_w \in (V)$ have finite mean, variance, and at least one finite higher-order moment (Greene 2003). In our case, the ith moment of V_x is finite if $\phi_{1,x-1}(ir)$ is finite; if the MGF about $-ir$ is defined for all duration distributions $f_w(t) : 1 \leq w < x$ (see also Lemma 1). In general, the MGF is defined for most duration distributions, and for most values of r. Therefore, we conclude that, in general, the Lyapunov condition (and hence the Lindeberg condition) holds. □

References

[1] Aitchison, J., & Brown, J.A.C. (1957). *The Lognormal Distribution, with Special Reference to Its Use in Economics.* New York: Cambridge University Press.

[2] Ballestín, F., & Leus, R. (2009). Resource-constrained project scheduling for timely project completion with stochastic activity durations. *Production and Operations Management*, 18(4), 459–474.

[3] Boothroyd, H. (1960). Least-cost testing sequence. *Operational Research Society*, 11(3), 137–138.

[4] Bradley, R.C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys*, 2(2005), 107–144.

[5] Bradley, R.C., & Cristina, T. (2015). A central limit theorem for non-stationary strongly mixing random fields. *Journal of Theoretical Probability*, 2015, 1–20.

[6] Chao, R.O., Lichtendahl Jr., K.C., & Grushka-Cockayne, Y. (2005). Incentives in a stage-gate process. *Production and Operations Management*, 23(8), 1286–1298.

[7] Coolen, K., Wei, W., Talla Nobibon, F., & Leus, R. (2014). Scheduling modular projects on a bottleneck resource. *Journal of Scheduling*, 17(1), 67-85.

[8] Creemers, S., Leus, R., & Lambrecht, M. (2010). Scheduling Markovian PERT networks to maximize the net present value. *Operations Research Letters*, 38(1), 51-56.

[9] Creemers, S., Leus, R., & De Reyck, B. (2015). Project planning with alternative technologies in uncertain environments. *European Journal of Operational Research*, 242(2), 465–476.

[10] Creemers, S. (2017). Maximizing the expected net present value of a project with phase-type distributed activity durations: an efficient globally optimal solution procedure, *European Journal of Operational Research*, forthcoming.

[11] Demeulemeester, E., & Herroelen, W. (2002). *Project Scheduling: A Research Handbook.* Dordrecht: Kluwer Academic Publishers.

[12] Demeulemeester, E., Vanhoucke, M., & Herroelen, W. (2003). A random network generator for activity-on-the-node networks. *Journal of Scheduling*, 6(1), 17-38.

[13] De Reyck, B., Degraeve, Z., & Vandenborre, R. (2008). Project options valuation with net present value and decision tree analysis. *European Journal of Operational Research*, 184(1), 341–355.

[14] De Reyck, B., & Leus, R. (2008). R&D-project scheduling when actities may fail. *IIE Transactions*, 40(4), 367-384.

[15] Girotra, K., Terwiesch, C., & Ulrich, K.T. (2007). Valuing R&D projects in a portfolio: evidence from the pharmaceutical industry. *Management Science*, 53(9), 1452–1466.

[16] Greene, W.H. (2003). *Econometric analysis.* New Jersey: Prentice Hall.

[17] Gowtham, B., Bhavnani, S.K., & Scott, C. (2011). group-based active query selection for rapid diagnosis in time-critical situations. *IEEE Transactions on Information Theory*, 58(1), 459–478.

[18] Hagström, J.N. (1988). Computational complexity of PERT problems. *Networks*, 18(2), 139-147.

[19] Herroelen, W. (2005). Project scheduling: theory and practice. *Production and Operations Management*, 14(4), 413–432.

[20] Huchzermeier, A., & Loch, C.H. (2001). Project management under risk: using the real options approach to evaluate flexibility in R&D. *Management Science*, 47(1), 85–101.

[21] Kamburowski, J. (1986). On the computational complexity of the shortest route and maximum flow problems in stochastic networks, *Foundations of Control Engineering* 11(4), 167–175.

[22] Kavadias, S., & Loch, C.H. (2003). Optimal project sequencing with recourse at a scarce resource. *Production and Operations Management*, 12(4), 433–444.

[23] Kelley, J.E. Jr., & Walker, M.R. (1959). Critical-path planning and scheduling. *Proceedings of the Eastern Joint Computer Conference, December 13, Boston*, 160-173.

[24] Madigan, D., Mittal, D., & Roberts, F. (2011). Efficient sequential decision-making algorithms for container inspection operations. *Naval Research Logistics*, 58(7), 637–654.

[25] Malcolm, D.G., Rosenboom, J.H., Clark, C.E., & Fazar, W. (1959). Application of a technique for research and development program evaluation. *Operations Research*, 7(5), 646–669.

[26] Moder, J.J., & Phillips, C.R. (1970). *Project Management with CPM and PERT*, 2nd ed. New York: Van Nostrand Reinhold Company.

[27] Osogami, T. (2005). *Analysis of multiserver systems via dimensionality reduction of Markov chains*, PhD thesis. Pittsburgh: Carnegie Mellon University.

[28] Osogami, T., & Harchol-Balter, M. (2006). Closed form solutions for mapping general distributions to quasi-minimal PH distributions. *Performance Evaluation*, 63(6), 524-552.

[29] Santiago, L.P., & Vakili, P. (2005). On the value of flexibility in R&D projects. *Management Science*, 51(8), 1206–1218.

[30] Sobel, M.J., Szmerekovsky, J.G., & Tilson, V. (2009). Scheduling projects with stochastic activity duration to maximize expected net present value. *European Journal of Operational Research*, 198(1), 697-705.

[31] Sommer, S.C. (2004). Selectionism and learning in projects with complexity and unforseeable uncertainty. *Management Science*, 50(10), 1334–1347.

[32] Szmerekovsky, J.G. (2005). The impact of contractor behavior on the client's payment-scheduling problem. *Management Science*, 51(4), 629–640.

[33] Trietsch, D., & Baker, K.R. (2012). PERT 21: fitting PERT/CPM for use in the 21st century. *International Journal of Project Management*, 30(4), 490–502.

[34] Ünlüyurt, T. (2004). Sequential testing of complex systems: a review. *Discrete Applied Mathematics*, 142(1–3), 189–205.

[35] Van Horne, J. (1966). Capital-budgeting decisions involving combinations of risky investments. *Management Science*, 3(2), B84–B92.

[36] Van Slyke, R.M. (1963). Monte Carlo methods and the PERT problem. *Operations Research*, 11(5), 839–860.

[37] Vanhoucke, M., Demeulemeester, E., & Herroelen, W. (2001). On maximizing the net present value of a project under renewable resource constraints. *Management Science*, 47(8), 1113–1121.

[38] Wei, W., Coolen, K., & Leus, R. (2013). Sequential testing policies for complex systems under precedence constraints. *Expert Systems Applications*, 40(2), 611-620.

[39] Wiesemann, W., Kuhn, D., & Rustem, B. (2010). Maximizing the net present value of a project under uncertainty. *European Journal of Operational Research*, 202(2), 356-367.

[40] Wiesemann, W., & Kuhn, D. (2015). The stochastic time-constrained net present value problem. In C. Schwindt, & J. Zimmermann (Eds.), *Handbook on Project Management and Scheduling Vol. 2* (pp. 753–780). New York: Springer.

[41] Yan, X., Xu, J., Zhu, Y., Wang, J., Yang, Y., & Wang, C.X. (2016). Downlink average rate and SINR distribution in cellular networks. *IEEE Transactions on Communications*, 64(2), 847–862.