WKB Approximation for the Sum of Two Correlated Lognormal Random Variables

C.F. Lo

Institute of Theoretical Physics and Department of Physics
The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
cflo@phy.cuhk.edu.hk

Copyright © 2013 C.F. Lo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

In this paper we apply the idea of the WKB method to derive an effective single lognormal approximation for the probability distribution of the sum of two correlated lognormal variables. An approximate probability distribution of the sum is determined in closed form, and illustrative numerical examples are presented to demonstrate the validity and accuracy of the approximate distribution. Our analysis shows that the proposed method is able to provide a simple, efficient and accurate approximation to this probability distribution of the sum of two correlated lognormal variables. We also discuss how this new approach can be straightforwardly extended to study the sum of $N$ lognormals.

Keywords: Lognormal random variables, probability distribution functions, backward Kolmogorov equation, Lie-Trotter splitting approximation, WKB approximation

1 Introduction

The probability distribution of the sum of two correlated lognormal stochastic variables has many important applications in various fields such as telecommunication studies [1-6], financial modelling [7-9], actuarial science [10-12], biosciences [13], physics [14,15], etc. Although the lognormal distribution is well known in the literature [16,17], yet almost nothing is known of the probability distribution of the sum of two correlated lognormal variables. Thus,
much effort has been made to look for good analytical approximations for the desired probability distribution [1-6,8,18-27]. Essentially, these analytical approximations assume a specific distribution that the sum of the two correlated lognormal variables follows, and then use a variety of methods, e.g. moment matching, moment generating function matching, least squares fitting, etc., to identify the parameters for that specific distribution. The most commonly used approximate distribution for the sum is a single lognormal distribution because numerical computations have shown that the probability distribution of the sum is a distribution which bears some resemblance to the lognormal distribution [1]. However, no explicit mathematical justification for the specific distribution has been given. In spite of this shortcoming, these approximations attract considerable attention and have been extended to approximate the sums of $N$ correlated lognormal variables, too.

Recently, by means of the Lie-Trotter operator splitting method (Trotter, 1959), Lo [28] showed that both the sum and difference of two correlated lognormal stochastic processes could be approximated by a shifted lognormal stochastic process, and approximate probability distributions were determined in closed form. Unlike previous studies which treat the sum and difference in a separate manner [1-6,8,18-27], Lo’s method provides a new unified approach to accurately model the dynamics of both the sum and difference of two lognormal variables. In this communication, based upon the Lie-Trotter operator splitting approximation proposed by Lo, we apply the idea of the WKB method [29] to derive an effective single lognormal approximation for the dynamics of the sum of two correlated lognormal variables. An approximate probability distribution of the sum is determined in closed form, and illustrative numerical examples are presented to demonstrate the validity and accuracy of the approximate distribution. In accordance with the analysis, the proposed method is not only able to provide a simple, efficient and accurate approximation to the probability distribution of the sum of two correlated lognormal variables, but also it has a better performance than Lo’s approximation. Furthermore, we discuss how this new approach can be extended to study the sum of $N$ lognormals as well.

2 Effective Single Lognormal Approximation

Given two lognormal stochastic variables $S_1$ and $S_2$ obeying the stochastic differential equations:

$$\frac{dS_i}{S_i} = \sigma_i dZ_i, \quad i = 1, 2 \quad (1)$$

where $\sigma_i^2 = \text{Var}(\ln S_i)$, $dZ_i$ denotes a standard Weiner process associated with $S_i$, and the two Weiner processes are correlated as $dZ_1dZ_2 = \rho dt$, the probability
distribution function of the sum of two correlated lognormal variables, i.e. $P_+ (S+, t; S_{10}, S_{20}, t_0)$, satisfies the backward Kolmogorov equation [30-33]

$$\left\{ \frac{\partial}{\partial t_0} + \hat{L} \right\} P_+ (S+, t; S_{10}, S_{20}, t_0) = 0 \quad (2)$$

for $t > t_0$ where

$$\hat{L} = \frac{1}{2} \sigma_1^2 S_{10}^2 \frac{\partial^2}{\partial S_{10}^2} + \rho \sigma_1 \sigma_2 S_{10} S_{20} \frac{\partial^2}{\partial S_{10} \partial S_{20}} + \frac{1}{2} \sigma_2^2 S_{20} \frac{\partial^2}{\partial S_{20}^2},$$

subject to the boundary condition [31]

$$P_+ (S+, t; S_{10}, S_{20}, t_0 \to t) = \delta (S_{10} + S_{20} - S_+) \quad (3)$$

To solve for $P_+ (S+, t; S_{10}, S_{20}, t_0)$, the backward Kolmogorov equation is first re-written in terms of the new variables $S_{\pm 0} \equiv S_{10} \pm S_{20}$ as

$$\left\{ \frac{\partial}{\partial t_0} + \hat{L}_+ + \hat{L}_R \right\} P_+ (S+, t; S_{+0}, S_{-0}, t_0) = 0 \quad (4)$$

where

$$\hat{L}_+ = \frac{1}{8} \left[ \sigma_+^2 (S_{+0})^2 + 2 (\sigma_1^2 - \sigma_2^2) S_{+0} S_{-0} + \sigma_-^2 (S_{-0})^2 \right] \frac{\partial^2}{\partial S_{+0}^2} \quad (5)$$

$$\hat{L}_R = \frac{1}{4} \left\{ (\sigma_1^2 - \sigma_2^2) [(S_{+0})^2 + (S_{-0})^2] + (\sigma_1^2 + \sigma_2^2) S_{+0} S_{-0} \right\} \frac{\partial^2}{\partial S_{+0} \partial S_{-0}} +$$

$$\frac{1}{8} \left[ \sigma_+^2 (S_{-0})^2 + 2 (\sigma_1^2 - \sigma_2^2) S_{+0} S_{-0} + \sigma_-^2 (S_{+0})^2 \right] \frac{\partial^2}{\partial S_{-0}^2} \quad (6)$$

$$\sigma_\pm = \sqrt{\sigma_1^2 + \sigma_2^2 \pm 2 \rho \sigma_1 \sigma_2}. \quad (7)$$

The corresponding boundary condition now becomes

$$P_+ (S+, t; S_{+0}, S_{-0}, t_0 \to t) = \delta (S_{+0} - S_+) \quad (8)$$

Accordingly, the formal solution of Eq.(5) is given by

$$P_+ (S+, t; S_{+0}, S_{-0}, t_0) = \exp \left\{ (t - t_0) \left( \hat{L}_+ + \hat{L}_R \right) \right\} \delta (S_{+0} - S_+) \quad (9)$$

Since the exponential operator in Eq.(10) is difficult to evaluate, the Lie-Trotter operator splitting method [34] can be applied to approximate the op-
erator by
\[
\hat{O}^{LT}_{+} = \exp\left\{ (t - t_0) \hat{\mathcal{L}}_+ \right\} \exp\left\{ (t - t_0) \hat{\mathcal{L}}_R \right\},
\]
and obtain an approximation to the formal solution \( P_+ (S_+, t; S_{+0}, S_{-0}, t_0) \), namely
\[
P^{LT}_+ (S_+, t; S_{+0}, S_{-0}, t_0) = \hat{O}^{LT}_{+} \delta (S_{+0} - S_+) = \exp\left\{ (t - t_0) \hat{\mathcal{L}}_+ \right\} \delta (S_{+0} - S_+) \quad (11)
\]
where the relation \( \exp\left\{ (t - t_0) \hat{\mathcal{L}}_R \right\} \delta (S_{+0} - S_+) = \delta (S_{+0} - S_+ + \sigma) \) is utilized. It is apparent that the approximate solution \( P^{LT}_+ (S_+, t; S_{+0}, S_{-0}, t_0) \) satisfies the backward Kolmogorov equation
\[
\left\{ \frac{\partial}{\partial t_0} + \frac{1}{2} \sigma_{eff}^2 S^2_{+0} \frac{\partial^2}{\partial S^2_{+0}} \right\} P^{LT}_+ (S_+, t; S_{+0}, S_{-0}, t_0) = 0 \quad (12)
\]
with the boundary condition
\[
P^{LT}_+ (S_+, t; S_{+0}, S_{-0}, t_0 \to t) = \delta (S_{+0} - S_+) , \quad (13)
\]
where
\[
\sigma_{eff}^2 = \frac{1}{4} \left[ \sigma_+^2 + 2 (\sigma_1^2 - \sigma_2^2) \frac{S_{-0}}{S_{+0}} + \sigma_2^2 \left( \frac{S_{-0}}{S_{+0}} \right)^2 \right]
\]
\[= \sigma_1^2 \left( \frac{S_{10}}{S_{+0}} \right)^2 + 2 \rho \sigma_1 \sigma_2 \left( \frac{S_{10}}{S_{+0}} \right) \left( \frac{S_{20}}{S_{+0}} \right) + \sigma_2^2 \left( \frac{S_{20}}{S_{+0}} \right)^2 . \quad (14)
\]
As the volatility \( \sigma_{eff} \) is a function of \( S_{+0} \), the solution \( P^{LT}_+ (S_+, t; S_{+0}, S_{-0}, t_0) \) cannot be a lognormal distribution. Then, here comes the question: “Is it possible to derive an accurate single lognormal approximation to the solution

\footnote{Suppose that one needs to exponentiate an operator \( \hat{C} \) which can be split into two different parts, namely \( \hat{A} \) and \( \hat{B} \). For simplicity, let us assume that \( \hat{C} = \hat{A} + \hat{B} \), where the exponential operator \( \exp\left( \hat{C} \right) \) is difficult to evaluate but \( \exp\left( \hat{A} \right) \) and \( \exp\left( \hat{B} \right) \) are either solvable or easy to deal with. Under such circumstances the exponential operator \( \exp\left( \varepsilon \hat{C} \right) \), with \( \varepsilon \) being a small parameter, can be approximated by the Lie-Trotter splitting formula:
\[
\exp\left( \varepsilon \hat{C} \right) = \exp\left( \varepsilon \hat{A} \right) \exp\left( \varepsilon \hat{B} \right) + O (\varepsilon^2) .
\]
This can be seen as the approximation to the solution at \( t = \varepsilon \) of the equation \( d\hat{Y}/dt = (\hat{A} + \hat{B}) \hat{Y} \) by a composition of the exact solutions of the equations \( d\hat{Y}/dt = \hat{A} \hat{Y} \) and \( d\hat{Y}/dt = \hat{B} \hat{Y} \) at time \( t = \varepsilon \). Details of the Lie-Trotter splitting approximation can be found in [34-39].}
Sum of two correlated lognormals

To answer this question, we may apply the idea of the WKB method which is a powerful tool for obtaining a global approximation to the solution of a linear ordinary differential equation.\(^2\)

**Proposition:**

If \(\sigma_{\text{eff}}\) is a slowly-varying function of \(S_{+0}\), i.e.

\[
\frac{S_{+0}}{\sigma_{\text{eff}}} \left| \frac{\partial \sigma_{\text{eff}}}{\partial S_{+0}} \right| \ll 1 ,
\]

then the solution \(P_{L+}^{LT} (S_{+}, t; S_{+0}, S_{-0}, t_0)\) can be approximated by

\[
P_{\text{eff}}^{LN} (S_{+}, t; S_{+0}, S_{-0}, t_0) = \frac{1}{S_{+}} \sqrt{\frac{2\pi \sigma^2_{\text{eff}} (t - t_0)}} \times \\
\exp \left\{ - \left[ \ln \left( \frac{S_{+}}{S_{+0}} \right) + \frac{1}{2} \sigma^2_{\text{eff}} (t - t_0) \right]^2 \right\}
\]

which resembles the lognormal distribution very closely.

**Proof:**

First of all, it is not difficult to show that

\[
\frac{S_{+0}}{\sigma_{\text{eff}}} \left| \frac{\partial \sigma_{\text{eff}}}{\partial S_{+0}} \right| = \frac{S_{+0}}{2\sigma^2_{\text{eff}}} \left| \frac{\partial \sigma^2_{\text{eff}}}{\partial S_{+0}} \right| \\
= \left\{ \frac{(\sigma^2_{1} - \sigma^2_{2}) + \sigma^2_{-} [S_{-0}/S_{+0}]}{\sigma^2_{+} + 2 (\sigma^2_{1} - \sigma^2_{2}) [S_{-0}/S_{+0}] + \sigma^2_{-} [S_{-0}/S_{+0}]^2} \right\} \\
\times \frac{S_{-0}}{S_{+0}} \ll 1
\]

\(^2\)The WKB method provides approximate solutions of differential equations of the form

\[
\frac{d^2 y (x)}{dx^2} + k (x)^2 y (x) = 0 ,
\]

provided that \(k (x)\) is slowly varying, i.e.

\[
\left| \frac{1}{k (x)} \frac{dk (x)}{dx} \right| \ll 1 .
\]

The completed approximate solution is given by

\[
y (x) \approx \frac{1}{\sqrt{k (x)}} \exp \left\{ \pm i \int k (x) dx \right\} .
\]

It is obvious that the approximate solution will be reduced to the usual plane-wave solution if \(k (x)\) is replaced by a constant. Details of the method can be found in [29,40,41].
provided \( \frac{\sigma_1^2 - \sigma_2^2}{\sigma_+^2} \leq 1 \) (or equivalently, \( \rho \geq -\sigma_2/\sigma_1 \)) and \( |S_{-0}/S_{+0}| \ll 1 \). Similarly, we can also show that

\[
\frac{S_{+0}^2}{4\sigma_{eff}^2} \frac{\partial^2 \sigma_{eff}^2}{\partial S_{+0}^2} = \left\{ \frac{(\sigma_1^2 - \sigma_2^2) + \frac{3}{2} \sigma_-^2 [S_{-0}/S_{+0}]}{\sigma_+^2 + 2(\sigma_1^2 - \sigma_2^2) [S_{-0}/S_{+0}] + \sigma_+^2 [S_{-0}/S_{+0}]^2} \right\} \times \frac{S_{-0}}{S_{+0}} \ll 1 .
\]

(18)

Then, substituting \( P_{eff}^{LN} (S_+, t; S_{+0}, t_0) \) into the left-hand side (L.H.S.) of Eq.(13), we obtain, after simplification,

\[
L.H.S. = -\frac{S_{+0}}{2(t-t_0)} \left[ \ln \left( \frac{S_+}{S_{+0}} \right) + \frac{1}{2} \sigma_{eff}^2 (t-t_0) \right] \left( \frac{\partial \sigma_{eff}^2}{\partial S_{+0}} \right) \frac{\partial P_{eff}^{LN}}{\partial (\sigma_{eff}^2)} - \frac{1}{2} \sigma_{eff}^2 S_{+0} \left[ \left( \frac{\partial^2 \sigma_{eff}^2}{\partial S_{+0}^2} \right) \frac{\partial P_{eff}^{LN}}{\partial (\sigma_{eff}^2)} + \left( \frac{\partial \sigma_{eff}^2}{\partial S_{+0}} \right)^2 \frac{\partial^2 P_{eff}^{LN}}{\partial (\sigma_{eff}^2)^2} \right] .
\]

(19)

where

\[
\frac{\partial P_{eff}^{LN}}{\partial (\sigma_{eff}^2)} = \frac{1}{2\sigma_{eff}^2} \left\{ \frac{1}{\sigma_{eff}^2 (t-t_0)} \left[ \ln \left( \frac{S_+}{S_{+0}} \right) + \frac{1}{2} \sigma_{eff}^2 (t-t_0) \right]^2 - \left[ \ln \left( \frac{S_+}{S_{+0}} \right) + \frac{1}{2} \sigma_{eff}^2 (t-t_0) \right] - 1 \right\} P_{eff}^{LN}
\]

(20)

and

\[
\frac{\partial^2 P_{eff}^{LN}}{\partial (\sigma_{eff}^2)^2} = \frac{1}{2 (\sigma_{eff}^2)^2} \left\{ \frac{3}{2} + 3 \ln \left( \frac{S_+}{S_{+0}} \right) + \sigma_{eff}^2 (t-t_0) + \left( \frac{1}{2} - \frac{3}{\sigma_{eff}^2 (t-t_0)} \right) \left[ \ln \left( \frac{S_+}{S_{+0}} \right) + \frac{1}{2} \sigma_{eff}^2 (t-t_0) \right]^2 - \frac{1}{\sigma_{eff}^2 (t-t_0)} \left[ \ln \left( \frac{S_+}{S_{+0}} \right) + \frac{1}{2} \sigma_{eff}^2 (t-t_0) \right]^3 + \frac{1}{2 (\sigma_{eff}^2)^2 (t-t_0)^2} \left[ \ln \left( \frac{S_+}{S_{+0}} \right) + \frac{1}{2} \sigma_{eff}^2 (t-t_0) \right]^4 \right\} P_{eff}^{LN}
\]

(21)

Since \( \sigma_{eff} \) is a slowly-varying function of \( S_{+0} \) as shown in Eq.(18) and Eq.(19), and

\[
\frac{[\sigma_1 S_{10} - \sigma_2 S_{20}]}{S_{+0}} < \sigma_{eff} < \frac{\sigma_1 S_{10} + \sigma_2 S_{20}}{S_{+0}},
\]

(22)

it can be inferred that \( L.H.S. \approx 0 \) in Eq.(20) and \( P_{eff}^{LN} (S_+, t; S_{+0}, S_{-0}, t_0) \) can be a good approximate solution of Eq.(13). (Q.E.D.)
As a consequence, we have succeeded in deriving an effective single lognormal approximation $P_{\text{eff}}^\text{LN}(S_+, t; S_{+0}, S_{-0}, t_0)$ to the probability distribution of the sum of two correlated lognormal variables, namely $P_{+}(S_+, t; S_{+0}, S_{-0}, t_0)$. The essence of the proposed approximation can be summarised as follows:

1. We first assume that the sum $S_+$ of the two correlated lognormal variables is governed by the lognormal process:

$$dS_+ = \sigma S_+ dZ_+$$

for some constant parameter $\sigma$. Then the probability distribution function of the sum $S_+$ is given by the lognormal distribution

$$f(S_+, t; S_{+0}, t_0) = \frac{1}{S_+ \sqrt{2\pi \sigma^2 (t - t_0)}} \exp \left\{-\frac{[\ln (S_+/S_{+0}) + \frac{1}{2} \sigma^2 (t - t_0)]^2}{2\sigma^2 (t - t_0)}\right\}$$

for $t > t_0$.

2. Next, we replace the constant $\sigma$ of the lognormal distribution $f(S_+, t; S_{+0}, t_0)$ by $\sigma_{\text{eff}}$ which is a function of the values of the two correlated lognormal variables $S_1$ and $S_2$ at $t = t_0$. The resultant function $P_{\text{eff}}^\text{LN}(S_+, t; S_{+0}, S_{-0}, t_0)$ gives a good approximation of the probability distribution of the sum of the two correlated lognormal variables, provided that $\sigma_{\text{eff}}$ is a slowly-varying function of $S_+$.

3 **Illustrative Numerical Results**

In Figure 1 we plot the approximate closed-form probability distribution function of the sum $S_+$ given in Eq.(17) for different values of the input parameters. We start with $S_{10} = 110$, $S_{20} = 100$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.15$ in Figure 1(a). Then, in order to examine the effect of $S_{20}$, we decrease its value to 70 in Figure 1(b) and to 40 in Figure 1(c). In Figures 1(d)-(f) we repeat the same investigation with a new set of values for $\sigma_1$ and $\sigma_2$, namely $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$. Without loss of generality, we set $t - t_0 = 1$ for simplicity. Both the (numerically) exact results obtained by numerical integrations and the approximate probability distribution function of Lo’s approximation [28] are also included for comparison. It is clear that the proposed approximation can provide accurate estimates for the exact values and have a better performance than Lo’s approximation. Moreover, to have a clearer picture of the accuracy, we plot the corresponding errors of the estimation in Figure 2. We can easily see that major discrepancies appear around the peak of the probability distribution function and that the estimation deteriorates as the
correlation parameter $\rho$ decreases from 0.5 to $-0.5$. It is also observed that the errors increase with the ratio $S_{-0}/S_{+0}$ as expected but they seem to be less sensitive to the changes in $\sigma_1$ and $\sigma_2$.

4 Conclusion

Based upon the Lie-Trotter operator splitting approximation proposed by Lo [28], we have applied the idea of the WKB method to derive an effective single lognormal approximation for the dynamics of the sum of two correlated lognormal stochastic variables. An approximate probability distribution of the sum is determined in closed form, and illustrative numerical examples are presented to demonstrate the validity and accuracy of these approximate distributions. The analysis shows that the proposed method provides a simple, efficient and accurate approximation to the probability distribution of the sum of two correlated lognormal variables.

Moreover, this new approach can be straightforwardly extended to study the sum of $N$ lognormals by simply re-define the $\sigma_{eff}^2$ in Eq.(15) as

$$\sigma_{eff}^2 = \sum_{i,j=1}^{N} \rho_{ij} \sigma_i \sigma_j \left( \frac{S_{i0}}{S_{+0}} \right) \left( \frac{S_{j0}}{S_{+0}} \right)$$

where $\rho_{ii} = 1$ and $\rho_{ij} = \rho_{ji}$. With this revised $\sigma_{eff}$, the probability distribution of the sum of $N$ lognormals can be accurately approximated by the distribution function $P_{eff}^{LN}(S_+; t, S_{+0}, S_{-0}, t_0)$ in Eq.(17). The proof can be outlined as follows:

1. As in the case of two lognormals, we first define $N$ new stochastic variables in terms of the $N$ lognormals, one of which represents the sum $S_+$.

2. Then we write down the backward Kolmogorov equation for the sum $S_+$ as in Section II.

3. Next, applying the Lie-Trotter splitting approximation, we can derive an approximate backward Kolmogorov equation for the sum $S_+$. This backward equation is the same as Eq.(13) except that the revised $\sigma_{eff}$ is used instead.

4. Finally, we apply the idea of the WKB method to show that the distribution function $P_{eff}^{LN}(S_+; t, S_{+0}, S_{-0}, t_0)$ with the revised $\sigma_{eff}$ provides a good approximation to the probability distribution of the sum of $N$ lognormals.
Figure 1: Probability density vs. $S_1 + S_2$. The solid lines denote the distributions of the effective single lognormal approximation, the dotted lines show the results of Lo’s approximation [28], and the dash lines represent the (numerically) exact results. (a) $S_{10} = 110$, $S_{20} = 100$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.15$; (b) $S_{10} = 110$, $S_{20} = 70$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.15$; (c) $S_{10} = 110$, $S_{20} = 40$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.15$; (d) $S_{10} = 110$, $S_{20} = 100$, $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$; (e) $S_{10} = 110$, $S_{20} = 70$, $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$; (f) $S_{10} = 110$, $S_{20} = 40$, $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$. 
Figure 2: Error vs. $S_1 + S_2$. The solid lines denote the errors of the effective single lognormal approximation and the dotted lines show the errors of Lo’s approximation [28]. (a) $S_{10} = 110$, $S_{20} = 100$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.15$; (b) $S_{10} = 110$, $S_{20} = 70$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.15$; (c) $S_{10} = 110$, $S_{20} = 40$, $\sigma_1 = 0.25$ and $\sigma_2 = 0.15$; (d) $S_{10} = 110$, $S_{20} = 100$, $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$; (e) $S_{10} = 110$, $S_{20} = 70$, $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$; (f) $S_{10} = 110$, $S_{20} = 40$, $\sigma_1 = 0.5$ and $\sigma_2 = 0.3$. 

![Figure 2(a)](image1)

![Figure 2(b)](image2)

![Figure 2(c)](image3)

![Figure 2(d)](image4)

![Figure 2(e)](image5)

![Figure 2(f)](image6)
References


Received: September 16, 2013