Margrabe Revisited

Hans J.H. Tuenter

Abstract We introduce a new representation of the bivariate normal distribution to first give a short derivation of the classic Margrabe exchange-option formula, using elementary integration methods. The second application is a new and simple technique to provide an accurate lower bound for the value of a spread option with a nonzero strike.

1 Introduction

Exchange options were introduced by William Margrabe in a seminal paper [7], published in 1978. This type of option allows the holder to exchange one asset for another at expiration. Such options are ubiquitous in foreign exchange markets, bond markets, stock markets, and commodity markets, among others. In energy markets, in particular, they have found applications in locational spreads, calendar spreads, crack spreads, and spark spreads. (See Clewlow and Strickland [4, pp. 80–81], Geman [5, pp. 287–294], and Pilipovic [8, pp. 361–374].) The survey by Carmona and Durrleman [2] provides a good introduction to the topic.

Margrabe studied European-style exchange options in a Black-Scholes framework, where the rate of return on each asset is given by

$$d\,S_i(t) = S_i(t)\,[r\,dt + \sigma_i\,dW_i(t)]\,, \quad i = 0, 1, \tag{1}$$

with r the risk-free interest rate, σ_i the instantaneous volatilities, $W_i(t)$ Wiener processes, and ρ the correlation coefficient between the increments dW_0 and dW_1. The payoff on the option to exchange S_0 for S_1 at time T, is given by

$$\Big(S_1(T) - S_0(T)\Big)^+\,, \tag{2}$$

where $x^+ = \max\{x, 0\}$.

H.J.H. Tuenter (✉)
Mathematical Finance Program, University of Toronto, Toronto, ON, Canada
e-mail: hans.tuenter@utoronto.ca

R. Aïd et al. (eds.), *Commodities, Energy and Environmental Finance*,
Fields Institute Communications 74, DOI 10.1007/978-1-4939-2733-3_4

Margrabe derived the risk-neutral value of this option as

$$e^{-rT}\,\mathbb{E}\left(S_1(T) - S_0(T)\right)^+ = S_1(0)\Phi(d_+) - S_0(0)\Phi(d_-), \tag{3}$$

where $\mathbb{E}$ denotes the expectation operator, Φ the cumulative density function of the standard normal distribution, and

$$d_\pm = \frac{\ln(S_1(0)/S_0(0))}{\sigma\sqrt{T}} \pm \frac{1}{2}\sigma\sqrt{T} \quad \text{and} \quad \sigma^2 = \sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1.$$

He obtains his formula by deriving a partial differential equation for the price of the option, together with its initial and boundary conditions. He postulates a solution and shows that it is the unique solution by employing a change-of-numéraire approach that transforms the valuation into a Black-Scholes type problem.

The first objective of this article is to provide a brief and simple derivation of the Margrabe formula that is based on a new representation of the bivariate normal distribution. This approach reduces the derivation to an elementary integration, and improves on previous approaches using plain integration, such as the one used by Li et al. [6, Proposition 2]. The second and main objective is to showcase a new lower bound for the value of a spread option with a nonzero strike, similar to the one derived by Carmona and Durrleman [2, § 6.1], but arrived at by a much simpler technique.

2 Bivariate Normal Distribution

In the Black-Scholes framework the logarithms of the asset prices at maturity follow a bivariate normal distribution. So, it seems only natural to first study the expected value of $(X_1 - X_0)^+$, where $\ln X_0$ and $\ln X_1$ are correlated normal variables, without the distraction of the stochastic process that generates them.

We derive this expectation for a particular case that is readily evaluated, and then show that the general case can always be mapped to it. This implies that, within the Black-Scholes framework, the particular case can be interpreted as a canonical formulation for an exchange option.

2.1 The Particular Case

As the particular case, we take

$$\ln X_0 = \mu_0 + aY + bZ \quad \text{and} \quad \ln X_1 = \mu_1 + aY + cZ, \tag{4}$$

where $a \geq 0$, $b > c$, and Y and Z are independent, standard normal variables.

Lemma 1.

$$\mathbb{E}\ (X_1 - X_0)^+ = \mathbb{E}\,X_1\,\Phi\left(z^* - c\right) - \mathbb{E}\,X_0\,\Phi\left(z^* - b\right), \tag{5}$$

where $z^* = (\mu_1 - \mu_0)/(b - c)$.

Proof. Substitute the expressions for X_0 and X_1, separate the factor e^{aY}, and take the expectation over Y, to give

$$\mathbb{E}\ (X_1 - X_0)^+ = e^{\frac{1}{2}a^2}\mathbb{E}\ \left(e^{\mu_1 + cZ} - e^{\mu_0 + bZ}\right)^+. \tag{6}$$

The value for Z, that renders the expression within brackets equal to zero, is given by $z^* = (\mu_1 - \mu_0)/(b - c)$. The expression is positive for values of Z smaller than z^*, and negative for values of Z larger than z^*. Now integrate over Z, and simple algebra, combined with the expectations $\mathbb{E}\,X_1 = e^{\mu_1 + \frac{1}{2}a^2 + \frac{1}{2}c^2}$ and $\mathbb{E}\,X_0 = e^{\mu_0 + \frac{1}{2}a^2 + \frac{1}{2}b^2}$, will show the validity of (5), and proves the lemma. □

Note that the constant z^* also has significance in that $\Phi(z^*)$ is the probability that X_0 is smaller than or equal to X_1, and the corresponding exchange option pays out.

- The condition $b > c$ is not really a restriction, as we can switch easily from the case $b < c$, by taking $-Z$ instead of Z in (4), using the fact that the standard normal distribution is symmetric. It was chosen for convenience in the proof of Lemma 1 to give the range of integration for Z as $(-\infty, z^*]$.
- The special case $b = c$ implies that $\ln X_0$ and $\ln X_1$ are the same random variable, except for a difference in their mean. The valuation in this case is simple as $\mathbb{E}\ (X_1 - X_0)^+ = \mathbb{E}\ \left(e^{\mu_1 + \sigma_0 Z} - e^{\mu_0 + \sigma_0 Z}\right)^+ = (e^{\mu_1} - e^{\mu_0})^+\, e^{\frac{1}{2}\sigma_0^2}$, and is equal to $\mathbb{E}\,X_1 - \mathbb{E}\,X_0$, when $\mu_1 > \mu_0$, and zero otherwise. This is not a practical case that one would encounter in the setting of an exchange option. However, it is worth noting that Lemma 1 includes this as a boundary case and thus ensures continuity of solution. Taking the limit of $(\mu_1 - \mu_0)/(b - c)$, as b approaches c from above, gives $z^* = +\infty$ and $z^* = -\infty$, when $\mu_1 > \mu_0$ and $\mu_1 < \mu_0$, respectively, so that (5) gives the correct limit values.

2.2 The General Case

In the general case, we have a bivariate normal distribution where the distributions of the logarithm of X_0 and X_1 are normal with mean μ_0 and μ_1, standard deviation σ_0 and σ_1, and correlation coefficient ρ. We make the very mild assumption that $\sigma_0^2 + \sigma_1^2 \neq 2\rho\sigma_0\sigma_1$. These two variables can be represented in several ways as a linear combination of independent, standard normal variables. The linear combination that is of interest in our setting is the following:

$$\ln X_0 = \mu_0 + \frac{\sigma_0}{\sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1}} \left(\sigma_1\sqrt{1-\rho^2}\,Y + (\sigma_0 - \rho\sigma_1)Z\right) \tag{7}$$

and

$$\ln X_1 = \mu_1 + \frac{\sigma_1}{\sqrt{\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1}} \left(\sigma_0\sqrt{1-\rho^2}\,Y + (\rho\sigma_0 - \sigma_1)Z\right), \tag{8}$$

where Y and Z are independent, standard normal variables. It is easy to verify that this construct gives two normal variables with the required means, standard deviations and correlation coefficient. Now note that the coefficients of Y in (7) and (8) are identical and nonnegative, and that the coefficient of Z in (7) is strictly larger than the coefficient of Z in (8). This implies that (7) and (8) are of the form (4), with

$$a = \frac{\sigma_0\sigma_1}{\sigma}\sqrt{1-\rho^2}, \quad b = \frac{\sigma_0}{\sigma}(\sigma_0 - \rho\sigma_1), \text{ and } c = \frac{\sigma_1}{\sigma}(\rho\sigma_0 - \sigma_1), \tag{9}$$

where $\sigma^2 = \sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1$. We note the following properties: $a^2 + b^2 = \sigma_0^2$, $a^2 + c^2 = \sigma_1^2$, and $b - c = \sigma$.

- The above shows that one can always cast the general bivariate normal distribution into the form (4); thus, they are equivalent and justifies referring to (4) as a canonical formulation. The representation of the bivariate normal distribution, given in (7) and (8), seems to be new. Although, it should be noted that, when $\sigma_0 = \sigma_1 = 1$ (and $\mu_0 = \mu_1 = 0$), it reduces to a well-known form that is used to generate correlated, standard normal variables. (See Tong [9, p. 11].)
- We imposed the condition $\sigma_0^2 + \sigma_1^2 \neq 2\rho\sigma_0\sigma_1$, but this is not restrictive. Equality holds if, and only if, $\rho = 1$ and $\sigma_1 = \sigma_0$. The implication is that $\ln X_0$ and $\ln X_1$ are the same random variable, except for a difference in their mean. This case was dealt with in the last bullet point of Sect. 2.1.

3 Margrabe's Formula

As noted in the introduction, the risk-neutral value of the exchange option within the Black-Scholes framework is given by $e^{-rT}\mathbb{E}\,(S_1(T) - S_0(T))^+$, where $S_0(T)$ and $S_1(T)$ are correlated lognormal variables. This means that we can apply Lemma 1, and, to do this, we take $X_1 = e^{-rT}S_1(T)$ and $X_0 = e^{-rT}S_0(T)$.

It is straightforward to show that this implies $\mathbb{E}\,X_i = S_i(0)$, $\mu_i = \ln S_i(0) - \frac{1}{2}\sigma_i^2 T$, $b = \frac{\sigma_0}{\sigma}(\sigma_0 - \rho\sigma_1)\sqrt{T}$ and $c = \frac{\sigma_1}{\sigma}(\rho\sigma_0 - \sigma_1)\sqrt{T}$, with $\sigma^2 = \sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1$. Simple substitution of these expressions in (5) gives the Margrabe formula (3).

4 Technical Interlude

In the subsequent analysis, it turns out to be beneficial to distinguish between a few basic case types for the canonical formulation (4). These correspond to whether or not the volatility parameters b and c are each positive or negative. Since we imposed the condition $b > c$, this gives three cases, as listed in Table 1.

The breakdown into the different case types has a straightforward interpretation in the general formulation in terms of a bound for the correlation coefficient. Note that, all things being equal, type II is the one most likely to be encountered, as it covers negative, zero and small positive correlations. We also note that, when one is holding the asset with lower volatility, it corresponds to either type II or III, but never to type I. Conversely, when one is holding the asset with higher volatility, it corresponds to either type I or II, but never to type III.

4.1 *Classification and Roots*

The classification into different case types is not merely an exercise in taxonomy, but plays a key role in the subsequent section on spread options. To ease the notational burden, let us define the function

$$f(z) = e^{\mu_1 + cz} - e^{\mu_0 + bz}. \tag{10}$$

This function already appeared in the proof of Lemma 1, in particular Eq. (6), where we had to find the unique root of $f(z) = 0$, in order to determine the range of integration. The situation is different for the general equation $f(z) = k$, as it may have a unique solution, none or two. The basic shape of $f(z)$ depends only on the signs of b and c, as can be seen more clearly from $f(z + z^*) = e^{\mu_1 + cz^*}\left(e^{cz} - e^{bz}\right)$, where z^* is the unique solution to $f(z) = 0$, that already featured in Lemma 1.

The different shapes are graphed in Fig. 1, from which one can infer the location of the roots of $f(z) = k$.

Table 1 Characteristics of the different case types

Formulation	Type I	Type II	Type III
Canonical	$b > c > 0$	$b > 0 > c$	$0 > b > c$
General	$\rho > \sigma_1/\sigma_0$	$\rho < \min\{\sigma_0/\sigma_1, \sigma_1/\sigma_0\}$	$\rho > \sigma_0/\sigma_1$
$\sigma_0 < \sigma_1$	–	$\rho < \sigma_0/\sigma_1$	$\rho > \sigma_0/\sigma_1$
$\sigma_0 > \sigma_1$	$\rho > \sigma_1/\sigma_0$	$\rho < \sigma_1/\sigma_0$	–

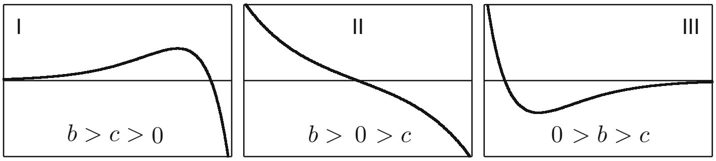

Fig. 1 The shape of $f(z)$ for the different case types

4.2 Case Type Analysis

To simplify the analysis, we will work with the normalized version $f_0(t) = e^{ct} - e^{bt}$, as the relevant properties of $f(z)$ are readily inferred from it. The former has exactly one root at zero, and, under the imposed condition $b > c$, is positive for $t < 0$, and negative for $t > 0$. The solvability of $f_0(t) = k$, depends on the case type. Since the basic shapes of $f_0(t)$ are the same as those of $f(z)$, the graphs in Fig. 1 will be helpful in visualizing the analysis of the different case types.

Type I When $b > c > 0$, then $f_0(t)$ has a unique extremum at

$$t^* = -\frac{\ln(b/c)}{b-c}, \quad \text{with value} \quad f_0^* = \left(\frac{b}{c}\right)^{\frac{b-c}{c}} - \left(\frac{b}{c}\right)^{\frac{b-c}{b}}, \tag{11}$$

and this extremum is positive and a maximum. The function $f_0(t)$ has zero as a left asymptote, is strictly increasing for $t < t^*$, and strictly decreasing for $t > t^*$. This implies that $f_0(t) = k$ has no solution when $k > f_0^*$, two solutions when $0 < k < f_0^*$, and one when $k = f_0^*$ or $k \leq 0$.

Type II When $b > 0 > c$, then $f_0(t)$ is a strictly decreasing function of t, and has no minimum or maximum. This implies that $f_0(t) = k$ has exactly one solution, for every value of k.

Type III When $0 > b > c$, we have the mirror image of case type I, and $f_0(t)$ has a unique extremum at t^* with value f_0^*, as given in (11), with the difference that this extremum is negative and a minimum. The function $f_0(t)$ is strictly decreasing for $t < t^*$, strictly increasing for $t > t^*$, and has zero as a right asymptote. This implies that $f_0(t) = k$ has no solution when $k < f_0^*$, two solutions when $f_0^* < k < 0$, and one when $k = f_0^*$ or $k \geq 0$.

4.3 Generalization of Lemma 1

In this section, we consider the expected value of $(X_1 - X_0 - K)^+$, where K is a constant. Using the canonical formulation (4), one can write

$$\mathbb{E}\,(X_1 - X_0 - K)^+ = \mathbb{E}\,\left(e^{aY} f(Z) - K\right)^+, \tag{12}$$

where Y and Z are standard normal and independent random variables, and $f(z)$ is as defined in (10). For $K = 0$, the expectation in (12) factorizes and results in (5). For general values of K, a simple and closed-form expression is not known. However, as noted by several researchers, expressions that provide good lower bounds do exist. Carmona and Durrleman [2, § 6.1], in the context of spread options, provide such an expression, although it involves solving a not-so-simple equation with quite a few trigonometric functions. An important feature of their lower bound is that it has a structure that is like that of the Margrabe formula, but with the addition of a term that is linear in K. It turns out that one can use our canonical representation to derive similar lower bounds with far less effort.

Lemma 2. *When the equation*

$$f(z) = Ke^{-\frac{1}{2}a^2}, \tag{13}$$

where K is nonnegative, has one or more solutions and z^* is the largest of them, then

$$\mathbb{E}\,(X_1 - X_0 - K)^+ \geq \mathbb{E}\,X_1\,\Phi(z^* - c) - \mathbb{E}\,X_0\,\Phi(z^* - b) - K\Phi(z^*). \tag{14}$$

Proof. We use representation (12). By conditioning on the value of Z, and then taking the expectation over Y, one can derive the lower bound

$$\mathbb{E}\,\left(e^{aY} f(Z) - K\right)^+ \geq \mathbb{E}\,\left(e^{\frac{1}{2}a^2} f(Z) - K\right)^+. \tag{15}$$

This follows easily from the fact that the function x^+ is convex and an application of Jensen's inequality. The problem is now reduced to that of determining the roots of (13). These roots provide the range of integration that allows the lower bound to be evaluated. From Sect. 4.2, we know that we have either one, none or two roots. When we are dealing with case type II, then Eq. (13) has a unique solution z^*, and the same applies when we are dealing with case type III, as K is nonnegative. This means that we can evaluate the right-hand side of (15) as

$$\int_{-\infty}^{z^*} \left(e^{\frac{1}{2}a^2} f(z) - K\right) d\Phi(z) \;=\; \mathbb{E}\,X_1\,\Phi(z^* - c) - \mathbb{E}\,X_0\,\Phi(z^* - b) - K\Phi(z^*).$$

When we are dealing with case type I, then Eq. (13) has roots $z_2^* \leq z_1^*$, as per condition of the lemma, and the right-hand side of (15) evaluates to

$$\int_{z_2^*}^{z_1^*} \left(e^{\frac{1}{2}a^2} f(z) - K\right) d\Phi(z) > \int_{-\infty}^{z_1^*} \left(e^{\frac{1}{2}a^2} f(z) - K\right) d\Phi(z), \tag{16}$$

with the latter integral evaluating to the same expression as for case types II and III, only with z_1^* instead of z^*. This shows, when (13) has a solution, that the largest root provides the lower bound (14) and proves the Lemma. □

- The condition in Lemma 2 that the constant K is nonnegative is not restrictive. Using the fact that $x^+ = x + (-x)^+$, gives the parity formula $\mathbb{E}\,(X_1 - X_0 - K)^+ = \mathbb{E}\,X_1 - \mathbb{E}\,X_0 - K + \mathbb{E}\,(X_0 - X_1 + K)^+$, so that one can easily convert from a negative value of K to a positive one.
- The condition in Lemma 2 that Eq. (13), where K is nonnegative, has one or more solutions is easily verified. Case types II and III always have exactly one solution. Case type I has one or two solutions if, and only if, $\max f(z) \geq Ke^{-\frac{1}{2}a^2}$. Using the analysis in Sect. 4.2, one can show that this is equivalent to

$$\frac{\mu_1 b - \mu_0 c}{b - c} + \ln f_0^* \geq \ln K - \frac{1}{2}a^2, \tag{17}$$

  with f_0^* as defined in (11).
- The convexity-based lower bound (14) has a format that is similar to the lower bound derived by Carmona and Durrleman [2, Eqn. 6.3]. However, the arguments of and inputs to their lower bound are much more complex and their computation more involved. The computational effort to compute our lower bound is limited to a simple root-finding procedure that can be implemented efficiently by a binary search or a Newton-Raphson method, and has no convergence issues.
- The bound (15) holds with equality when either $K = 0$ or $a = 0$. This implies that the convexity bound (14) is exact for the case $K = 0$, and thus that Lemma 2 is indeed a generalization of Lemma 1. When a is zero, this corresponds to X_0 and X_1 being perfectly correlated and $\rho = \pm 1$.

5 Spread Options

In many practical settings it is not really the intent to exchange one asset for another, but to lock in a price differential. In these settings, one has to overcome a fixed cost or pay a price to exercise (or strike) the option. In the set-up of Sect. 2, this means that we are considering the expected value of $(X_1 - X_0 - K)^+$, where K denotes the strike price. The exchange option can thus be seen as a spread option with a zero strike.

5.1 Validation

To validate Lemma 2 and get a sense of its applicability and usefulness in deriving a lower bound for the value of an exchange option, we use the test case from Carmona and Durrleman [3, Table 1], and compare against their results. They take

$S_0(0) = 100, q_0 = 2\,\%, \sigma_0 = 15\,\%$, and $S_1(0) = 110, q_1 = 3\,\%, \sigma_1 = 10\,\%$, where the q_i represent the continuous dividend yields. The time to maturity is taken as $T = 1$ year and the risk-free rate as $r = 5\,\%$. For the strike and the correlation coefficient, all combinations of $K \in \{-20, -10, 0, 5, 15, 25\}$ and $\rho \in \{-1, -0.5, 0, 0.3, 0.8, 1\}$ are considered.

Although Margrabe's original formulation did not include dividends (he only considered capital gains), the introduction of the continuous dividend yields q_i does not fundamentally change things. The spot prices are now assumed to follow the stochastic differential equation

$$d\,S_i(t) = S_i(t)\left[(r - q_i)\,dt + \sigma_i\,dW_i(t)\right], \quad i = 0, 1. \tag{18}$$

The risk-neutral value of the spread option is given by $e^{-rT}\mathbb{E}\,(S_1(T) - S_0(T) - K)^+$. This means that we can apply Lemma 2, with $X_1 = e^{-rT}S_1(T)$ and $X_0 = e^{-rT}S_0(T)$, and a strike price of $e^{-rT}K$. It is straightforward to show that $\mathbb{E}\,X_i = e^{-q_iT}S_i(0)$, $\mu_i = \ln S_i(0) - (q_i + \frac{1}{2}\sigma_i^2)T$, $a = \frac{\sigma_0\sigma_1}{\sigma}\sqrt{1-\rho^2}\sqrt{T}$, $b = \frac{\sigma_0}{\sigma}(\sigma_0 - \rho\sigma_1)\sqrt{T}$, and $c = \frac{\sigma_1}{\sigma}(\rho\sigma_0 - \sigma_1)\sqrt{T}$, with $\sigma^2 = \sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1$. Substitution of these expressions in (14) gives the following lower bound for the value of this spread option:

$$e^{-q_1T}S_1(0)\Phi(z^* - c) - e^{-q_0T}S_0(0)\Phi(z^* - b) - e^{-rT}K\Phi(z^*), \tag{19}$$

where z^* solves $f(z) = e^{-rT}Ke^{-\frac{1}{2}a^2}$, and $f(z)$ is as defined in (10).

The numerical values of the three volatility parameters a, b and c, under each of the six correlation coefficients considered, are given in Table 2. The results of the lower bound (19) for the value of the spread option are listed in Table 3, together with the lower bound, as derived by the Carmona-Durrleman method [3, p. 24].

As can be seen, the two methods give virtually the same results, with the Carmona-Durrleman bound being a little bit better for large, positive values of the strike.

5.2 *Relevance of the Lower Bounds*

Using a Monte Carlo simulation with 100,000 trials, Carmona and Durrleman [3, p. 24] showed that these lower bounds are extremely close to the true value. This finding is corroborated in a study by van der Hoek and Korolkiewicz [10] who use a different valuation technique. Their approach is based on a perturbation

Table 2 Volatility parameters for the Carmona-Durrleman test case under the canonical formulation (4)

ρ	−1	−0.5	0	0.3	0.8	1
a	0	0.05960	0.08321	0.09334	0.09762	0
b	0.15	0.13765	0.12481	0.11742	0.11389	0.15
c	−0.1	−0.08030	−0.05547	−0.03588	0.02169	0.1

Table 3 Comparison of the lower bounds

K	ρ					
	-1	-0.5	0	0.3	0.8	1
-20	29.656	28.994	28.381	28.070	27.770	27.754
	29.656	28.990	28.373	28.062	27.769	27.754
-10	21.868	20.904	19.888	19.270	18.381	18.244
	21.869	20.903	19.885	19.265	18.379	18.244
0	15.133	13.917	12.523	11.561	9.632	8.821
	15.133	13.918	12.524	11.562	9.633	8.821
5	12.244	10.956	9.445	8.367	5.967	4.454
	12.244	10.956	9.444	8.365	5.963	4.454
15	7.521	6.242	4.744	3.679	1.342	0.049
	7.522	6.236	4.729	3.657	1.303	0.049
25	4.201	3.129	1.961	1.219	0.103	0
	4.201	3.115	1.930	1.178	0.076	0

For every strike value, the first row gives the bound by the Carmona-Durrleman method, and the second row the bound (19) by our method

expansion of the solution to the differential equation that the price of the spread-option satisfies, and they use this expansion to derive analytic formulae for second-order approximations. They employ the same test case and validate their bounds with a recombining binomial tree model. Bjerksund and Stensland [1] study a modified version of the Kirk approximation, and verify their results with a quasi-Monte Carlo method using a two-dimensional Halton sequence with 100,000 pairs, combined with a variance reduction technique. They also use the test case from Carmona and Durrleman, and show that the lower bounds are extremely accurate. Their methodology appears simpler and as accurate as the Carmona-Durrleman approach. The conclusion in each of these studies is the same: these lower bounds provide extremely accurate approximations.

6 Improving the Lower Bounds

To obtain better lower bounds for $\mathbb{E}\,(X_1 - X_0 - K)^+ = \mathbb{E}\,\left(e^{aY}f(Z) - K\right)^+$, we can partition the range for Y, from the one interval, covering the totality of the real line, into n intervals: $I_i = [y_i, y_{i+1})$, $i = 0, \ldots n-1$, with $y_0 = -\infty$ and $y_n = +\infty$. This gives

$$\mathbb{E}\,\left(e^{aY}f(Z) - K\right)^+ = \sum_{i=0}^{n-1} \mathbb{E}\left[\left(e^{aY}f(Z) - K\right)^+ \mid Y \in I_i\right] \text{Prob}\,(Y \in I_i) \;\geq$$

$$\sum_{i=0}^{n-1} \mathbb{E}\left(\left(\Phi(y_{i+1} - a) - \Phi(y_i - a)\right) e^{\frac{1}{2}a^2} f(Z) - \left(\Phi(y_{i+1}) - \Phi(y_i)\right) K\right)^+,$$

where the lower bound follows from applying Jensen's inequality to the conditional expectation. For this lower bound, similar to Lemma 2, let z_i^* be the largest value for Z that equates the argument of the expectation in the ith summand to zero. This expectation is then evaluated by integrating Z over the interval $(-\infty, z_i^*]$. To simplify the notation, we define $p_i = \text{Prob}\,(Y \in I_i)$ and $\tilde{p}_i = \text{Prob}\,(Y + a \in I_i)$, and derive the following generalization of the lower bound (14):

$$\sum_{i=0}^{n-1} \tilde{p}_i \left[\mathbb{E}\, X_1 \Phi(z_i^* - c) - \mathbb{E}\, X_0 \Phi(z_i^* - b) - K\Phi(z_i^*) p_i/\tilde{p}_i \right], \tag{20}$$

where z_i^* is the largest of the values that solves $f(z) = Ke^{-\frac{1}{2}a^2} p_i/\tilde{p}_i$.

This leaves us with a multitude of ways to choose the partition. An appealing choice is the construction where the intervals $[y_i - a, y_{i+1} - a)$ are equally probable, so that $\tilde{p}_i = 1/n$. This implies $y_i = \Phi^{-1}(i/n) + a$, and gives the lower bound as

$$\mathbb{E}\, \left(e^{aY} f(Z) - K\right)^+ \geq \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\, \left(e^{\frac{1}{2}a^2} f(Z) - np_i K\right)^+ \tag{21}$$

$$\geq \frac{1}{n} \sum_{i=0}^{n-1} \left[\mathbb{E}\, X_1\, \Phi(z_i^* - c) - \mathbb{E}\, X_0\, \Phi(z_i^* - b) - np_i K\Phi(z_i^*) \right]. \tag{22}$$

We note that these bounds might also give some insight into the structure of the spread option formula. As the number of partitions goes to infinity, the limit of (21) will converge to the true value of the spread option. For case type II, this evaluates to $\frac{1}{n} \sum_{i=0}^{n-1} \int_{-\infty}^{z_i^*} e^{\frac{1}{2}a^2} f(z)\, dz - \sum_{i=0}^{n-1} p_i K\Phi(z_i^*)$. This structure implies that there is a ζ_n, such that the first expression is the same as the integral $\int_{-\infty}^{\zeta_n} e^{\frac{1}{2}a^2} f(z)\, dz$, and that there is a ξ_n, such that the second expression is the same as $K\Phi(\xi_n)$. These aggregations imply that, for case type II, the value of $\mathbb{E}\,(X_1 - X_0 - K)^+$ can be given by a formula of the type:

$$\mathbb{E}\, X_1\, \Phi(d_1) - \mathbb{E}\, X_0\, \Phi(d_0) - K\Phi(d_2),$$

where $d_1 - d_0 = b - c = \sqrt{\sigma_0^2 - 2\rho\sigma_0\sigma_1 + \sigma_1^2}$, which has a pleasant resemblance to the Black-Scholes formula.

6.1 Test Results

To measure the effect of increasing the number of partitions, we implemented lower bound (22) and again used the Carmona-Durrleman test case. The results for a selected number of partitions are given in Table 4. We note that our benchmark values agree perfectly with those given by Bjerksund and Stensland [1, Table 1].

For negative values of the strike K, the choice of $n = 1$, corresponding to the lower bound (14) from Lemma 2, already gives excellent results. This could already have been surmised from Table 3, but the test results show that the approximations

Table 4 The effect of the number of partitions n on the accuracy of lower bound (22)

		ρ					
K	n	-1	-0.5	0	0.3	0.8	1
−20	1	29.656138	28.989825	28.372967	28.062162	27.768572	27.753786
	2		28.992999	28.378177	28.067242	27.769537	
	5		28.994296	28.380299	28.069303	27.769931	
	10		28.994605	28.380802	28.069789	27.770024	
	20		28.994726	28.380997	28.069977	27.770061	
	50		28.994782	28.381088	28.070064	27.770078	
	100		28.994797	28.381112	28.070086	27.770082	
	1,000		28.994808	28.381129	28.070102	27.770086	
-10	1	21.868637	20.902992	19.884984	19.265409	18.378517	18.243872
	2		20.904239	19.887451	19.268381	18.380153	
	5		20.904750	19.888463	19.269599	18.380818	
	10		20.904873	19.888705	19.269891	18.380976	
	20		20.904921	19.888801	19.270005	18.381036	
	50		20.904943	19.888846	19.270059	18.381065	
	100		20.904949	19.888858	19.270073	18.381072	
	1,000		20.904954	19.888866	19.270083	18.381077	
0	1	15.133217	13.917957	12.523665	11.561761	9.632542	8.821249
5	1	12.244123	10.955506	9.443681	8.364984	*5.962979*	4.454214
	2		10.955956	9.444728	8.366512	5.965534	
	5		10.956141	9.445161	8.367146	5.966597	
	10		10.956185	9.445266	8.367300	5.966857	
	20		10.956203	9.445307	8.367361	5.966961	
	50		10.956211	9.445327	8.367390	5.967011	
	100		10.956213	9.445332	8.367398	5.967025	
	1,000		10.956215	9.445336	8.367404	5.967035	
15	1	7.521812	*6.235713*	*4.729195*	***3.657248***	***1.302566***	0.048825
	2		6.239841	*4.738887*	*3.671555*	***1.328113***	
	5		6.241535	4.742874	*3.677442*	*1.338485*	
	10		6.241941	4.743833	3.678856	*1.340930*	
	20		6.242101	4.744211	3.679413	1.341874	
	50		6.242176	4.744390	3.679677	1.342310	
	100		6.242196	4.744438	3.679748	1.342423	
	1,000		6.242210	4.744472	3.679798	1.342500	

(continued)

Table 4 (continued)

K	*n*	ρ					
		-1	-0.5	0	0.3	0.8	1
25	1	4.201368	*3.114953*	***1.929799***	***1.177618***	***0.076431***	0.000000
	2		*3.124543*	***1.950431***	***1.204798***	***0.094212***	
	5		3.128465	*1.958832*	*1.215786*	***0.101371***	
	10		3.129400	*1.960822*	*1.218362*	***0.103045***	
	20		3.129765	1.961594	*1.219351*	*0.103687*	
	50		3.129936	1.961954	1.219806	*0.103983*	
	100		3.129981	1.962048	1.219923	*0.104059*	
	1,000		3.130013	1.962113	1.220002	0.104112	

For the cases with $K = 0$ or $\rho = \pm 1$, our method is exact and attained for $n = 1$, so these numbers are not replicated further. For all other cases, we have taken the value for $n = 1{,}000$ as the "true" and benchmark value. The numerical results that deviate more than 0.05 % from their benchmark are typeset in italics, and those that deviate more than 0.5 % are typeset in bold-italics

are all within 0.05 % of their benchmark. For positive values of K, the results are still very good, but decrease in accuracy with increasing value of the strike K, and increasing value of the correlation coefficient ρ. However, the choice of $n = 5$ partitions does give an approximation within 0.5 % of the benchmark, for all cases, except for those with the highest correlation coefficient of 0.8.

7 Discussion

In this section we add a few comments and observations that would have obstructed the flow of the discussion had they been incorporated into the main section.

Why Are the Lower Bounds So Accurate? One naturally wonders why the analytical lower bounds for the price of the spread option are so accurate. All the approaches that rely on using convexity arguments and replacing a random variable by its expectation reduce the problem from a two-dimensional to a one-dimensional one. The approximation being so accurate must mean that the problem is, in some sense, close to a one-dimensional problem. The traditional formulation does not show this, but the canonical formulation provides some insight. The convexity argument uses the approximation $\mathbb{E}\left[e^{aY}f(Z) - K\right]^+ \approx \mathbb{E}\left[\left(\mathbb{E}\, e^{aY}\right)f(Z) - K\right]^+$, so that, the lower the variability of e^{ay}, the better the approximation is likely to be. For the Carmona-Durrleman test case, the numerical value of a is close to zero, and Table 5 shows that e^{aY} is close to one, as measured by its expectation and standard deviation.

For new test cases, with larger values of a, one would hazard a guess that the lower bounds are likely to be less accurate. It would be interesting to compare the

Table 5 Characteristics of the volatility parameter a

ρ	-1	-0.5	0	0.3	0.8	1
a	0	0.05960	0.08321	0.09334	0.09762	0
$\mathbb{E}\, e^{aY}$	1	1.00178	1.00347	1.00437	1.00478	1
SDev e^{aY}	0	0.09763	0.08364	0.09395	0.09832	0

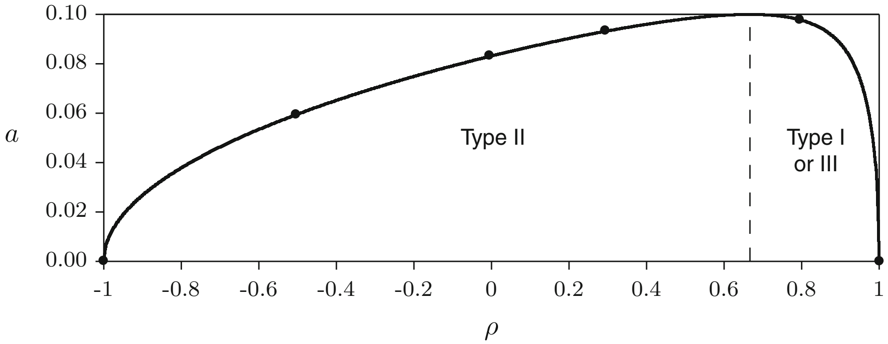

Fig. 2 The effect of the correlation on the volatility parameter a for the Carmona-Durrleman test case. The maximum occurs at $\rho = 2/3$ with value 0.1. The *bullets* correspond to the values for $\rho \in \{-1, -0.5, 0, 0.3, 0.8, 1\}$

results from the Carmona-Durrleman and the Bjerksund-Stensland approaches to ours for a wider range of parameters.

The Effect of Correlation When the correlation is perfect, that is, $\rho = \pm 1$, the value of a is zero, and our lower bound gives the true value of the spread option. Since larger values of a imply a larger standard deviation for e^{aY}, the behavior of a and its maximum, as a function of the correlation coefficient, is of interest. It is easier to look at a^2 and differentiate this with respect to ρ:

$$\frac{\partial a^2}{\partial \rho} = \frac{2\sigma_0^2\sigma_1^2(\sigma_1\rho - \sigma_0)(\rho\sigma_0 - \sigma_1)}{(\sigma_0^2 + \sigma_1^2 - 2\rho\sigma_0\sigma_1)^2}.$$

This derivative is zero for either $\rho = \sigma_0/\sigma_1$ or $\rho = \sigma_1/\sigma_0$. As ρ lies within the interval $[-1, 1]$, it is not difficult to show that a^2 is strictly increasing for $\rho \leq r$, where $r = \min\{\sigma_0/\sigma_1, \sigma_1/\sigma_0\}$, and strictly decreasing for $\rho \geq r$. The unique maximum of a^2 is at $\rho = r$ with value $\min\{\sigma_0^2, \sigma_1^2\}$, so that $a \leq \min\{\sigma_0, \sigma_1\}$. Note that the upper bound also follows from the fact that $a^2 + b^2 = \sigma_0^2$ and $a^2 + c^2 = \sigma_1^2$. For the Carmona-Durrleman test case, the value of a, as a function of ρ, is graphed in Fig. 2. Typically, for smaller values of a, the lower bound (14) is more accurate.

Acknowledgements The simplified proof of the Margrabe formula, using elementary integration techniques, was developed in the context of a course on energy commodities that the author teaches in the Mathematical Finance Program at the University of Toronto. I thank the anonymous referee for comments and suggestions that helped to improve the presentation, and my wife Marguerite Martindale for a professional line edit.

References

1. Bjerksund, P., Stensland, G.: Closed form spread option valuation. Quant. Finance **14**(10), 1785–1794 (2014)
2. Carmona, R., Durrleman, V.: Pricing and hedging spread options. SIAM Rev. **45**(4), 627–685 (2003)
3. Carmona, R., Durrleman, V.: Pricing and hedging spread options in a log-normal model. Technical Report, Department of Operations Research and Financial Engineering. Princeton University, Princeton, NJ, 16 March (2003)
4. Clewlow, L., Strickland, C.: Energy Derivatives: Pricing and Risk Management. Lacima Publications, London (2000)
5. Geman, H.: Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy. Wiley, London (2005)
6. Li, M., Deng, S.-J., Zhou, J.: Closed-form approximations for spread option prices and Greeks. J. Deriv. **15**(3), 58–80 (Spring 2008)
7. Margrabe, W.: The value of an option to exchange one asset for another. J. Finance **33**(1), 177–186 (1978)
8. Pilipovic, D.: Energy Risk: Valuing and Managing Energy Derivatives, 2nd edn. McGraw-Hill, New York (2007)
9. Tong, Y.L.: The Multivariate Normal Distribution. Springer Series in Statistics. Springer, New York (1990)
10. van der Hoek, J., Korolkiewicz, M.W.: New analytic approximations for pricing spread options. In: Cohen, S.N., Madan, D., Siu, T.K.,Yang, H. (eds.) Stochastic Processes, Finance and Control: A Festschrift in Honor of Robert J. Elliott. Advances in Statistics, Probability and Actuarial Science, vol. 1, pp. 259–284. World Scientific, Singapore (2012)