

## MULTIDIMENSIONAL FOURIER INVERSION USING IMPORTANCE SAMPLING WITH APPLICATION TO OPTION PRICING

Santanu Dey

School of Technology and Computer Science  
Tata Institute of Fundamental Research  
Mumbai, 400 005, INDIA

Sandeep Juneja

School of Technology and Computer Science  
Tata Institute of Fundamental Research  
Mumbai, 400 005, INDIA

### ABSTRACT

In this paper we present our ongoing effort to use importance sampling to develop unbiased, bounded estimators of densities, distribution functions and expectations of functions of a random vector, when the characteristic function of the (multi-dimensional) random vector is available in analytic or semi-analytic form. This is especially of interest in options pricing as stochastic processes such as affine jump processes and Levy processes are ubiquitous in financial modeling and typically have characteristic functions (of their value at a given time) that are easily evaluated while their density or distribution functions have no readily computable closed form. Typically, for pricing options via Monte Carlo, a discretized version of the underlying SDE is simulated using Euler or a related method and the resultant estimator has a discretization bias. A noteworthy feature of our Monte Carlo approach is that, when applicable, it provides unbiased estimators.

### 1 INTRODUCTION

As is well known, in the options pricing literature, stochastic differential equations (SDE) are used to model the movement of underlying financial instruments such as stock, bond, commodities prices, interest and foreign exchange rates, and default intensities of corporates (see, e.g., [Duffie \(1992\)](#), [Shreve \(2004\)](#) for an introduction). Typically, under the no-arbitrage condition, the price of an option can be expressed as the expectation of its discounted payoff (as a function of the underlying financial instruments) under the risk neutral-measure. This representation allows Monte-Carlo simulation to be used to price such options by taking the average of sample discounted pay-offs generated via simulation. However, often the density or the distribution function of the underlying financial instruments is not available and there are no straightforward ways to generate samples with the correct distribution. Typically, a sample is obtained by discretizing the underlying SDE. This may lead to a discretization bias in generated samples. Substantial literature has come up recently on how to reduce or remove this bias. One strand of literature focusses on developing correction terms in the discretization schemes to reduce the bias (see [Glasserman \(2004\)](#), [Asmussen and Glynn \(2007\)](#) for bias related issues in sampling from SDEs). Another more recent strand focuses on generating exact samples from an SDE (see, e.g., [Beskos and Roberts \(2005\)](#), [Chen \(2010\)](#)). Although, the latter has had success in a limited framework involving single dimensional SDE's.

There is also a significant literature that exploits the fact that although the density or the distribution functions of an SDE evaluated at any given time may not be available, its characteristic function may have a closed or a semi-analytic form (i.e., easily computable form). This is true for a large class of models including those involving affine-jump diffusions and Levy processes (see, e.g., [Kahl and Lord \(2010\)](#), [Borovkov and Novikov \(2002\)](#), [Carr and Madan \(1999\)](#), also see [Abate and Whitt \(1992a\)](#), [Abate and Whitt \(1992b\)](#) for related queueing literature). Then, using Fourier inversion or related Parseval/Plancherel identity, options can be priced by numerically integrating suitably adjusted characteristic function. In this paper, we build upon this literature by noting that often the integral involved may be multi-dimensional. For instance, price of a call option with two separate barriers may have a representation as an expectation

$$E \left[ (X - K)^+ I_{\{Y \leq b_1\}} I_{\{Z \geq b_2\}} \right]$$

where  $X, Y, Z$  are random variables whose joint characteristic function may be known,  $b_1, b_2$  are constants, and  $I_A(\cdot)$  is the indicator function of the set  $A$ . Here, and more generally in examples involving dimensionality of three or more,

the traditional methods for evaluating integrals such as quadrature or fast Fourier transform, may be inefficient. We propose that importance sampling based Monte-Carlo be used instead in such cases, and we develop conditions under which the resulting estimator is bounded (in particular, the variance is finite).

Note that the general idea of using importance sampling to aid in evaluating integrals involving Fourier inversion has wide applicability. For instance, consider a case where the density function of a random vector depends on a parameter set  $\alpha$ . For each  $\alpha$ , the density function may be unavailable but the characteristic function may be known. Then, importance sampling can be used to estimate the density function for any  $\alpha$ . This, in turn may be useful in maximum likelihood estimation of  $\alpha$  from the samples of the random vector.

In Section 2 we discuss how importance sampling may be used to estimate joint probability density functions, options prices using Parseval/Pancheral identity, and multi-dimensional distribution functions. In Section 3 we report numerical results for Heston's Stochastic Volatility model. We end with a brief conclusion in Section 4.

## 2 FOURIER INVERSION AND SIMULATION

The essential idea of using importance sampling in conjunction with Fourier inversion is straightforward. We first discuss it for the simple case of estimating the joint probability density function of a random vector. We then note that using Parseval/Pancheral identity, option prices may often be expressed as integrals involving Fourier transforms, and importance sampling may be effectively used to estimate these. We end this section with an algorithm that uses importance sampling to estimate multi-dimensional distribution functions. We also discuss few examples where these ideas are applied.

Some notation: Consider a  $p$  dimensional random vector  $\mathbf{X} = (X_1, \dots, X_p)$ . The (joint) characteristic function of  $\mathbf{X}$  is defined as

$$\phi_{\mathbf{X}}(\mathbf{t}) = \phi_{\mathbf{X}}(t_1, t_2, \dots, t_p) = E[e^{i(t_1 X_1 + t_2 X_2 + \dots + t_p X_p)}] = E[e^{i\mathbf{t}^T \mathbf{X}}]$$

where  $t = \sqrt{-1}$  is the complex root of unity.

### 2.1 Estimating joint PDF

Let  $f$  denote the joint density function of  $\mathbf{X}$  (assuming this exists). Then,

$$\phi_{\mathbf{X}}(\mathbf{t}) = \int_{\mathbb{R}^p} e^{i\mathbf{t}^T \mathbf{x}} f(\mathbf{x}) d\mathbf{x}.$$

Furthermore, under the assumption that  $|\phi_{\mathbf{X}}(\mathbf{t})|$  is integrable, we have by the inversion formula:

$$f(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^p \int_{\mathbb{R}^p} \phi_{\mathbf{X}}(\mathbf{t}) e^{-i\mathbf{t}^T \mathbf{x}} d\mathbf{t}. \quad (1)$$

When  $p = 1$ , it can be seen (Lemma XV 4, Feller (1971)) that if  $f$  has integrable derivatives  $f', f'', \dots, f^{(n)}$ , then  $\phi_{\mathbf{X}}(t)$  is  $o(|t|^{-n})$ . In the general  $p \geq 1$  case, we assume that there exists  $\delta > 1$  such that

$$\phi_{\mathbf{X}}(\mathbf{t}) = o\left(\left(\prod_{i=1}^p |t_i|\right)^{-\delta}\right). \quad (2)$$

This may be interpreted as  $\phi_{\mathbf{X}}(\mathbf{t}) (\prod_{i=1}^p |t_i|)^{\delta}$  converges to zero as  $\max_i t_i \rightarrow \infty$ .

Well known, Hormander Theorem and related conditions guarantee smoothness of the joint density function of  $\mathbf{X}$  (see, e.g., Komatsu and Takeuchi (2001), M.Bismut (1981)), from which (2) may follow.

Then, by setting the importance sampling density function as

$$g(\mathbf{t}) = \frac{c_{\varepsilon}}{\prod_{i=1}^p (1 + |t_i|)^{\delta - \varepsilon}}$$

for  $\varepsilon \in [0, \delta - 1)$  and appropriate normalization constant  $c_{\varepsilon}$ , we have

$$f(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^p \int_{\mathbb{R}^p} \frac{\phi_{\mathbf{X}}(\mathbf{t}) e^{-i\mathbf{t}^T \mathbf{x}}}{g(\mathbf{t})} g(\mathbf{t}) d\mathbf{t}. \quad (3)$$

This then provides a straightforward method for estimating the joint density function: Generate i.i.d. samples  $(\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_N)$  from the density  $g$ . The sample average

$$\left(\frac{1}{2\pi}\right)^p \frac{1}{N} \sum_{i=1}^N \frac{\phi_{\mathbf{X}}(\mathbf{T}_i) e^{-i\mathbf{T}_i^T \mathbf{x}}}{g(\mathbf{T}_i)}$$

is an unbiased estimator of  $f(\mathbf{x})$  for each  $\mathbf{x}$ . Since each sample  $\frac{\phi_{\mathbf{X}}(\mathbf{T}) e^{-i\mathbf{T}^T \mathbf{x}}}{g(\mathbf{T})}$  is bounded by  $\sup_t \frac{\prod_{i=1}^p (1+|t_i|)^{\delta-\varepsilon} \phi_{\mathbf{X}}(\mathbf{t})}{c_\varepsilon}$ , which in turn, being a supremum of a continuous function that can be made arbitrarily small outside a compact set (by assumption (2)), is finite. The resultant estimator has all moments finite and the variance converges to zero as  $N \rightarrow \infty$ . In practice, one may ignore the imaginary component from the generated samples because its mean equals zero. In fact, this may be an effective control variate.

### 2.2 Options pricing using Parseval/Pancheral's Formula

If  $\mathbf{X}$  is an  $\mathbb{R}^p$  valued random variable with density function  $f_{\mathbf{X}}$  and characteristic function  $\phi_{\mathbf{X}}$  then by Parseval/Pancheral identity, under technical conditions we have

$$E[H(\mathbf{X})] = \int_{\mathbb{R}^p} H(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) dx_1 dx_2 \dots dx_p = \left(\frac{1}{2\pi}\right)^p \int_{\mathbb{R}^p} h(\mathbf{t}) \phi_{\mathbf{X}}(-\mathbf{t}) dt_1 dt_2 \dots dt_p \tag{4}$$

where  $h(\mathbf{t}) = \int_{\mathbb{R}^p} e^{i\mathbf{t}^T \mathbf{x}} H(\mathbf{x}) dx_1 dx_2 \dots dx_p$  is the Fourier transform of  $H(\mathbf{x})$  (see (Feller 1971), (Rudin 1986)). This is useful as an option price often takes the form  $\int_{\mathbb{R}^p} H(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) dx_1 \dots dx_p$  where  $H(\mathbf{x})$  denotes the discounted payoff as a function of asset value  $\mathbf{x}$  and  $f_{\mathbf{X}}$  denotes the pdf of these asset values under the risk neutral measure. As mentioned earlier, often we may not have explicit knowledge about the density function  $f_{\mathbf{X}}$  but instead the characteristic function  $\phi_{\mathbf{X}}$  may be readily available in an analytic or semi-analytic form. In such a case, the option price can be evaluated using the right most integral in (4), provided the Fourier transform  $h(\mathbf{t})$  for the function  $H(\mathbf{x})$  is defined and well behaved. See, e.g., Borovkov and Novikov (2002), Carr and Madan (1999), Kahl and Lord (2010) for further details on when the latter condition holds and relaxations based on damping technique as well as numerous applications of this representation. The following example illustrates one scenario where Fourier transform  $h(\mathbf{t})$  for a popular cost function  $H(\mathbf{x})$  is known.

**Example 1.** Consider the example of pricing a European call option on a single asset. While methods such as quadrature or Fast Fourier Transform for evaluating integrals may be preferred over Monte Carlo in the single dimension setting, this example is useful to simply illustrate the proposed method. The option price equals  $E(e^X - K)^+$  where  $X$  denotes the logarithm of the asset price at maturity. Let  $f_X$  denote its density function.

As in Borovkov and Novikov (2002), write:

$$(e^x - K)^+ = (e^x - e^b)^+ = e^x \left(1 - e^{-(x-b)}\right)^+ = e^{(1+a)x} e^{-ax} \left(1 - e^{-(x-b)}\right)^+ = e^{(1+a)x} G(x),$$

where  $b = \ln K$  and  $G(x) = e^{-ax} \left(1 - e^{-(x-b)}\right)^+$ . It is easily seen that fourier transform of  $G(x)$  is given by

$$\hat{G}(t) = \int_{-\infty}^{\infty} e^{itx} G(x) dx = \frac{e^{b(t-a)}}{(a-it)(a+1-it)}.$$

Also, under the assumption  $E[e^{(1+a)X}] < \infty$ , we have

$$\int_{-\infty}^{\infty} e^{itx} e^{(1+a)x} f_X(x) dx = \phi_X(t - 1 - ia)$$

so that

$$E[(e^X - K)^+] = \int_{-\infty}^{\infty} e^{(1+a)x} G(x) f_X(x) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_X(-t - 1 - ia) \hat{G}(t) dt \tag{5}$$

Note that  $|\hat{G}(t)| = O(|t|^{-2})$ , so that when  $X$  has a Gaussian distribution, we have  $\phi_X(t) \hat{G}(t) = O(|t|^{-\delta})$  for any  $\delta > 0$ .

Next we consider the case of estimating the RHS in (4). The idea is same as the previous case. We write this integral as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left[ \frac{h(\mathbf{t})\phi_{\mathbf{X}}(-\mathbf{t})}{g(\mathbf{t})} g(\mathbf{t}) \right] dt_1 dt_2 \dots dt_p = E_g \left[ \frac{h(\mathbf{T})\phi_{\mathbf{X}}(-\mathbf{T})}{g(\mathbf{T})} \right]$$

Where  $g$  is a density function on  $\mathbb{R}^p$ , and  $\mathbf{T}$  is a random vector with this density. This density may be chosen so that the random variable  $\frac{h(\mathbf{T})\phi_{\mathbf{X}}(-\mathbf{T})}{g(\mathbf{T})}$  is bounded a.s. Again, we assume that,

$$h(\mathbf{t})\phi_{\mathbf{X}}(-\mathbf{t}) = O\left(\left(\prod_{i=1}^p |t_i|\right)^{-\delta}\right) \tag{6}$$

for some  $\delta > 1$ .  
Then by selecting

$$g(\mathbf{t}) = \frac{c_\varepsilon}{\prod_{i=1}^p (1 + |t_i|)^{\delta-\varepsilon}}$$

for an appropriate constant  $c_\varepsilon$  for  $\varepsilon \in [0, \delta - 1)$ , it is easy to see that  $\frac{h(\mathbf{T})\phi_{\mathbf{X}}(-\mathbf{T})}{g(\mathbf{T})}$  is bounded a.s., so that the simulation results in an estimator with finite variance that decreases to zero as the number of generated samples increases to infinity.

### 2.3 Estimating Joint Distribution Function

Let  $F(\cdot)$  be the distribution function of  $\mathbf{X}$ . That is,

$$F(x_1, x_2, \dots, x_p) := F_{\mathbf{X}}(x_1, x_2, \dots, x_p) := P[X_1 \leq x_1, X_2 \leq x_2, \dots, X_p \leq x_p]$$

We will drop the subscripts from  $F$  when there is no possible confusion about the random variables involved. Recall that the joint characteristic function of  $\mathbf{X}$  is defined as

$$\phi_{\mathbf{X}}(\mathbf{t}) = \phi_{\mathbf{X}}(t_1, t_2, \dots, t_p) = E[e^{i(t_1 X_1 + t_2 X_2 + \dots + t_p X_p)}] = E[e^{i\mathbf{t}^T \mathbf{X}}].$$

For  $1 \leq i_1 < i_2 < \dots < i_k \leq p$ , denote the joint df of  $(X_{i_1}, X_{i_2}, \dots, X_{i_k})$  by

$$F_{X_{i_1}, X_{i_2}, \dots, X_{i_k}}(x_{i_1}, x_{i_2}, \dots, x_{i_k}) := P[X_{i_1} \leq x_{i_1}, X_{i_2} \leq x_{i_2}, \dots, X_{i_k} \leq x_{i_k}].$$

Likewise, denote the joint characteristic function by

$$\phi_{X_{i_1}, X_{i_2}, \dots, X_{i_k}}(t_{i_1}, t_{i_2}, \dots, t_{i_k}) := E[e^{i(t_{i_1} X_{i_1} + t_{i_2} X_{i_2} + \dots + t_{i_k} X_{i_k})}]$$

Note that

$$F_{X_{i_1}, X_{i_2}, \dots, X_{i_k}}(x_{i_1}, x_{i_2}, \dots, x_{i_k}) = \lim_{x_j \rightarrow \infty, j \notin \{i_1, i_2, \dots, i_k\}} F_{\mathbf{X}}(x_1, x_2, \dots, x_p)$$

and

$$\phi_{X_{i_1}, X_{i_2}, \dots, X_{i_k}}(t_{i_1}, t_{i_2}, \dots, t_{i_k}) = \phi_{\mathbf{X}}(0, \dots, t_{i_1}, 0, \dots, 0, t_{i_2}, 0, \dots, 0, t_{i_k}, 0, \dots, 0)$$

Define,

$$u_0(\mathbf{x}) = u_0(x_1, x_2, \dots, x_p) := 1$$

and for  $1 \leq k \leq p$

$$u_k(\mathbf{x}) = u_k(x_1, x_2, \dots, x_p) := \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq p} F_{X_{i_1}, X_{i_2}, \dots, X_{i_k}}(x_{i_1}, x_{i_2}, \dots, x_{i_k})$$

The following relation is well known (see, e.g., (Gurland 2001), (Shephard 1991a), (Shephard 1991b)):

$$\sum_{k=0}^p (-1)^k 2^{p-k} u_{p-k}(\mathbf{x}) = \left(\frac{-1}{\pi}\right)^p \int_0^\infty \int_0^\infty \dots \int_0^\infty \Delta_{t_1} \Delta_{t_2} \dots \Delta_{t_p} \left[ \frac{\phi_{\mathbf{X}}(\mathbf{t}) e^{-i\mathbf{t}^T \mathbf{x}}}{i t_1 i t_2 \dots i t_p} \right] dt_1 dt_2 \dots dt_p \tag{7}$$

where

$$\Delta_t h(t) = h(t) + h(-t).$$

Note that by applying  $\frac{\partial^p}{\partial x_1 \partial x_2 \dots \partial x_p}$  to both sides of equation (7) we get the inversion formula for density as:

$$f(x_1, x_2, \dots, x_p) = \left(\frac{1}{2\pi}\right)^p \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \phi_{\mathbf{X}}(\mathbf{t}) e^{-i\mathbf{t}'\mathbf{x}} dt_1 dt_2 \dots dt_p, \tag{8}$$

which holds whenever  $|\phi_{\mathbf{X}}(\mathbf{t})|$  is integrable.

### 2.3.1 Monte Carlo estimation

We now discuss one way to get an unbiased estimator  $\hat{F}_N(x_1, x_2, \dots, x_p)$  of a multivariate df  $F(x_1, x_2, \dots, x_p)$  at a given point  $(x_1, x_2, \dots, x_p)$ . We exploit the multivariate inversion formula (7). Importance sampling is used to estimate the various integrals that appear in the proposed algorithm. Its implementation is described later. In the proposed algorithm, to estimate the df of dimension  $p$  we need to estimate all the lower dimensional marginal dfs. we do this iteratively in  $p$  steps. We spell out the first few steps to add clarity to the algorithm:

**Step 1:** Note from (7) (with  $p$  set to 1) that

$$2F_{X_i}(x_i) - 1 = \left(\frac{-1}{\pi}\right) \int_0^{\infty} \Delta_t \left[ \frac{\phi_{X_i}(t) e^{-tx_i}}{t} \right] dt.$$

By using importance sampling to estimate the integral on the RHS, unbiased estimates of  $F_{X_i}(x_i)$  for  $i = 1, 2, \dots, p$  are obtained.

**Step 2:** Again, from (7) (with  $p$  set to 2):

$$4F_{X_i, X_j}(x_i, x_j) - 2[F_{X_i}(x_i) + F_{X_j}(x_j)] + 1 = \left(\frac{-1}{\pi}\right)^2 \int_0^{\infty} \int_0^{\infty} \Delta_{t_1} \Delta_{t_2} \left[ \frac{\phi_{X_i, X_j}(t_1, t_2) e^{-t_1 x_i - t_2 x_j}}{t_1 t_2} \right] dt_1 dt_2.$$

Using importance sampling to get a samples of integral in the RHS and combining this with the estimates of  $F_{X_i}(x_i)$  in Step 1, we get unbiased estimates of  $F_{X_i, X_j}(x_i, x_j)$  for  $1 \leq i < j \leq p$ .

**Step  $k$  ( $k = 3, \dots, p$ ):** From (7), we have

$$\begin{aligned} & 2^k F_{X_{i_1}, \dots, X_{i_k}}(x_{i_1}, \dots, x_{i_k}) + \sum_{j=1}^k (-1)^j 2^{k-j} u_{k-j}(x_{i_1}, \dots, x_{i_k}) \\ &= \left(\frac{-1}{\pi}\right)^k \int_{\mathbb{R}^{+k}} \Delta_{t_1} \dots \Delta_{t_k} \left[ \frac{\phi_{X_{i_1}, \dots, X_{i_k}}(t_1, \dots, t_k) e^{-\sum_{j=1}^k t_j x_{i_j}}}{t_1 \dots t_k} \right] dt_1 \dots dt_k. \end{aligned}$$

Using the earlier generated estimates of  $u_{k-j}(x_{i_1}, \dots, x_{i_k})$  for  $1 \leq j \leq k-1$ , and sampling to estimate the integral on the RHS, we get an unbiased estimator of  $F_{X_{i_1}, \dots, X_{i_k}}(x_{i_1}, \dots, x_{i_k})$  for  $1 \leq i_1 < i_2 < \dots < i_k \leq p$ .

Note that at each step in the above algorithm, we need to estimate the integral of the form

$$\int_{\mathbb{R}^{+k}} \Delta_{t_1} \dots \Delta_{t_k} \left[ \frac{\phi_{\mathbf{X}}(\mathbf{t}) e^{-i\mathbf{t}'\mathbf{x}}}{t_1 t_2 \dots t_k} \right] dt_1 \dots dt_k = \int_{\mathbb{R}^{+k}} \left[ \frac{\psi(\mathbf{x}, \mathbf{t})}{t_1 t_2 \dots t_k} \right] dt_1 \dots dt_k, \tag{9}$$

where the bold vectors above are now of dimension  $k$  and

$$\psi(\mathbf{x}, \mathbf{t}) = \sum_{\varepsilon_i \in \{1, -1\}} (\varepsilon_1 \dots \varepsilon_k) \phi_{\mathbf{X}}(\varepsilon_1 t_1, \dots, \varepsilon_k t_k) e^{-i(\varepsilon_1 t_1 x_1 + \dots + \varepsilon_k t_k x_k)}.$$

Alternatively,

$$\psi(\mathbf{x}, \mathbf{t}) = 2^k E \left( \left( \prod_{i=1}^k B_i \right) \phi_{\mathbf{X}}(B_1 t_1, \dots, B_k t_k) e^{-t(B_1 t_1 x_1 + \dots + B_k t_k x_k)} \right),$$

where each  $B_i$  is an independent Bernoulli random variable taking values  $-1$  or  $1$  with equal probability.

Thus, to estimate (9) via importance sampling we re-express it as

$$\int_{\mathbb{R}^{+k}} \left[ \frac{\psi_{\mathbf{X}}(\mathbf{t})}{(t_1 \dots t_k) g(\mathbf{t})} \right] g(\mathbf{t}) dt_1 \dots dt_k = E_g \left[ \frac{\psi_{\mathbf{X}}(\mathbf{T})}{(T_1 T_2 \dots T_k) g(\mathbf{T})} \right]$$

Assuming (2) as before, if we set

$$g(\mathbf{t}) = \frac{d_\varepsilon}{\prod_{i=1}^k (1 + t_i)^{\delta + 1 - \varepsilon}}$$

for appropriate constant  $d_\varepsilon$ ,  $0 \leq \varepsilon < \delta$ , then the unbiased sample for (9) equals

$$\frac{2^k (\prod_{i=1}^k B_i) \phi_{\mathbf{X}}(B_1 T_1, \dots, B_k T_k) e^{-t(B_1 T_1 x_1 + \dots + B_k T_k x_k)}}{(T_1 T_2 \dots T_k) g(\mathbf{T})}$$

where  $(T_1, T_2, \dots, T_k)$  are generated using  $g$  and the  $B_i$ 's are independently generated as discussed earlier. It can be easily seen that under mild conditions, this output is bounded and hence the simulation estimator involving average of such samples has finite variance that converges to zero as the generated samples increase to infinity. Also, one can generate one set of  $(B_i : i \leq p)$  and  $(T_i : i \leq p)$  and use them to estimate each term of the form (9) to get an unbiased sample for  $F(x_1, x_2, \dots, x_p)$ .

Example 2 below illustrates a case where finding the options price can be reduced to evaluating a multi-dimensional distribution function using above Fourier inversion ideas.

**Example 2.** Suppose we want to price a derivative security whose pay-off is given by

$$V(X_1, X_2, X_3) = (X_1 - K_1)^+ I_{\{X_2 \leq K_2\}} I_{\{X_3 \geq K_3\}}.$$

Under fundamental theorem of asset pricing this reduces to evaluating the expectation (under the risk neutral measure):

$$\begin{aligned} E[V(X_1, X_2, X_3)] &= E[(X_1 - K_1)^+ I_{\{X_2 \leq K_2\}} I_{\{X_3 \geq K_3\}}] \\ &= E[X_1 I_{\{X_2 \leq K_2\}} I_{\{X_3 \geq K_3\}}] - K_1 P[X_1 \geq K_1, X_2 \leq K_2, X_3 \geq K_3] \\ &= EX_1 P^*[X_1 \geq K_1, X_2 \leq K_2, X_3 \geq K_3] - K_1 P[X_1 \geq K_1, X_2 \leq K_2, X_3 \geq K_3] \end{aligned}$$

where  $P^*$  is a measure defined by  $P^*(A) = \frac{EX_1 I(A)}{EX_1}$  for all measurable  $A$ . In many settings in finance, the density or the distribution function of  $(X_1, X_2, X_3)$  is not readily available, however, their joint characteristic function is known. One popular example corresponds to the case where  $(X_1, X_2, X_3)$  are a solution at a particular time to an affine process SDE. Specifically, a  $p$ -dimensional Ito process  $\mathbf{X}(t) = (X_1(t), \dots, X_p(t))$  satisfying SDE

$$d\mathbf{X}(t) = \mu(\mathbf{X}(t))dt + \sigma(\mathbf{X}(t))d\mathbf{W}(t),$$

is called affine if the  $\mathcal{F}_t$ -conditional characteristic function of  $\mathbf{X}(T)$  is exponential affine in  $\mathbf{X}(t)$ , for all  $t \leq T$ . That is there exists functions  $A : (0, \infty) \rightarrow \mathbb{C}$  and  $\mathbf{B} : (0, \infty) \rightarrow \mathbb{C}^p$  such that

$$E[e^{\mathbf{u}^T \mathbf{X}(T)} | \mathcal{F}_t] = e^{A(T-t, \mathbf{u}) + \mathbf{B}(T-t, \mathbf{u})^T \mathbf{X}(t)}$$

for all  $\mathbf{u} \in \mathbb{R}^p$  and  $t \leq T$ .

Assuming that  $A$  and  $\mathbf{B}$  are sufficiently smooth standard argument using Feynman-Kac formula implies that  $A$  and  $\mathbf{B}$  satisfies two coupled ordinary differential equations (referred to as generalized Riccati equations), subject to two boundary conditions  $A(0, \mathbf{u}) = 0$ ,  $\mathbf{B}(0, \mathbf{u}) = \mathbf{u}$  for all  $\mathbf{u} \in \mathbb{R}^p$  (for details see Duffie (1992)). These two equations can be solved for  $A$  and  $\mathbf{B}$  in closed form (e.g., Heston's model) or can be approximated easily using standard numerical techniques like Runge-Kutta method. Furthermore, the process  $\mathbf{X}(t)$  remains affine under the probability measure  $P^*$ .

So the option price involves estimating multidimensional distribution function where the characteristic function is easily computed.

### 3 Numerical example

In this section we apply our method in a simple setting of Heston's stochastic volatility model. This has the advantage that the results are available from literature for comparison.

Suppose that under risk neutral measure the stock price and its volatility process follow:

$$dS(t) = \mu S(t)dt + \sqrt{V(t)}S(t)dW(t) \tag{10}$$

$$dV(t) = \kappa(\theta - V(t))dt + \omega\sqrt{V(t)}dZ(t) \tag{11}$$

$$dZ(t)dW(t) = \rho dt \tag{12}$$

The parameters values as well as their interpretations are specified in Table (1). This is not an affine process, but we can convert it to one by substituting  $X(t) = \ln S(t)$ . In what follows we assume that  $\mu = 0$  and write:

$$dW(t) = dW_1(t) \text{ and } dZ(t) = \rho dW_1(t) + \sqrt{1 - \rho^2}dW_2(t)$$

The system of SDEs under these transformations is:

$$dX(t) = -\frac{1}{2}V(t)dt + \sqrt{V(t)}dW_1(t) \tag{13}$$

$$dV(t) = \kappa(\theta - V(t))dt + \omega\rho\sqrt{V(t)}dW_1(t) + \omega\sqrt{1 - \rho^2}\sqrt{V(t)}dW_2(t) \tag{14}$$

The resulting coupled Ricatti equation can be solved in closed form for an expression of the characteristic function  $\varphi(u)$  of  $X(T)$  (see, e.g., [Kahl and Lord \(2010\)](#)). We get

$$\varphi(u) = e^{C(T,u)+D(T,u)V_0+tu\ln F},$$

where

$$C(T, u) = \frac{\kappa\theta}{\omega^2} \left[ (\kappa - \rho\omega u + d(u))T - 2\ln \frac{c(u)e^{d(u)T} - 1}{c(u) - 1} \right],$$

$$D(T, u) = \frac{\kappa - \rho\omega u + d(u)}{\omega^2} \times \frac{e^{d(u)T} - 1}{c(u)e^{d(u)T} - 1},$$

with

$$c(u) = \frac{\rho\omega u - \kappa - d(u)}{\rho\omega u - \kappa + d(u)} \text{ and } d(u) = \sqrt{(\rho\omega u - \kappa)^2 + \omega^2(u^2 + u)}.$$

See [Kahl and Jackel \(2005\)](#), for technical issues concerning the continuity of the complex logarithm embedded in the expression of  $D$ . They also implement numerical integration algorithm of Gauss and Lobatto to evaluate the price of a European call option. Along the lines of Example (2), using Fourier inversion, it is well known that this price is given by

$$C = \frac{1}{2}(S_0 - K) + \frac{1}{\pi} \left[ S_0 \int_0^\infty f_1(u)du - K \int_0^\infty f_2(u)du \right], \tag{15}$$

where

$$f_1(u) := \Re \left( \frac{e^{-iu\ln K} \varphi(u - i)}{iuS_0} \right) \text{ and } f_2(u) := \Re \left( \frac{e^{-iu\ln K} \varphi(u)}{iu} \right).$$

We perform our experiment in the same set up as [Kahl and Jackel \(2005\)](#). Specifically, we use their method to evaluate the characteristic function, and consider the same parameter values in our numerical experiment (as shown in Table 1). For these parameters using Gauss Lobatto algorithm for numerical integration [Kahl and Jackel \(2005\)](#) report the price of the call option to equal 0.0495212.

We use importance sampling to estimate the two integrals appearing in (15). Our importance sampling density is :

Table 1: Model specifications.

Symbol	for	Value
$S_0$	Initial stock price	1.0
$V_0$	initial volatility	0.16
$K$	Strike price	2.0
$T$	Maturity	10.0
$\omega$	volatility of volatility	2.0
$\kappa$	Mean reversion rate	1.0
$\theta$	Long term mean volatility	0.16
$\rho$	Correlation	-0.8

$$g(t) = \frac{\varepsilon}{(1+t)^{\varepsilon+1}}$$

for  $\varepsilon = 2, 3, 4$ . We can estimate the two integrals separately, and can in principle use different importance sampling density to estimate them. But empirically we observe considerable variance reduction by using the same IS density and same samples to estimate the two integrals (suggesting positive correlation in the two outputs). For instance, when we estimated the two integrals using independent samples,  $N = 100,000$ , with same density ( $\varepsilon = 3$ ), our price point estimate equalled 0.0467 with 95% confidence interval (0.0420, 0.0514) (confidence width= 0.0094). When this experiment was done using common random numbers for the two integrals, our point estimate equalled 0.0496 with 95% and confidence interval (0.0480, 0.0511) (confidence width= 0.0031). So over nine times variance reduction through using common random numbers. The results in Tables 2, 3 and 4 are using common random numbers. We note a reasonable rate of convergence to the true value. In this example  $\varepsilon = 2$  appears to have lower variance that appears to increase with increasing  $\varepsilon$ .

Table 2:  $\varepsilon = 2$ 

N	price	95% confidence interval
1000	0.0523	(0.0469, 0.0578)
10000	0.0496	(0.0479, 0.0514)
100000	0.0494	(0.0485, 0.0508)

Table 3:  $\varepsilon = 3$ 

N	price	95% confidence interval
1000	0.0488	(0.0338, 0.0637)
10000	0.0521	(0.0472, 0.0569)
100000	0.0496	(0.0480, 0.0511)

#### 4 Conclusions

In this paper we reported some observations from our ongoing research on using importance sampling to estimate multi-dimensional probability density functions, distribution functions and related measures that can be obtained by inverting Fourier transforms, whenever the Fourier transform, or equivalently, the characteristic function are available in a closed or easily computable form. This may be a useful approach as in many applications, particularly in financial options pricing, Fourier transforms of underlying functions or random variables are readily available while their distributions may not have a closed form representation. The proposed methodology may then become a viable alternative especially when multiple dimensions are involved.

#### REFERENCES

- Abate, J., and W. Whitt. 1992a. The Fourier series method for inverting transforms of probability distribution. *Queueing Systems Theory and Applications* 10:5–88.
- Abate, J., and W. Whitt. 1992b. Numerical inversion of probability generating functions. *Operation Research Letters* 12:245–251.



Table 4:  $\varepsilon = 4$ 

N	price	95% confidence interval
1000	0.0401	(0.00939,0.0709)
10000	0.0533	(0.0428,0.064)
100000	0.0509	(0.0476,0.0542)
1000000	0.0491	(0.0481,0.0502)
10000000	0.0494	(0.0490,0.0497)

- Asmussen, S., and P. Glynn. 2007. *Stochastic simulation: Algorithms and analysis*. New York: Springer.
- Beskos, A., and G. Roberts. 2005. Exact simulation of diffusions. *Ann. Appl. Probab.* 15 (4): 2422–2444.
- Borovkov, K., and A. Novikov. 2002. On a new approach to calculating expectations for option pricing. *Journal of Applied Probability.* 39 (4): 889–895.
- Carr, P., and D. Madan. 1999. Option valuation using fast Fourier transform. *Journal of Computational Finance* 2:61–73.
- Chen, N. 2010. Exact simulation of brownian motion driven stochastic differential equations. *Preprint*.
- Duffie, D. 1992. *Dynamic asset pricing theory*. Princeton NJ: Princeton University Press.
- Feller, W. 1971. *An introduction to probability theory and its applications*, 2. New York: John Wiley and Sons Inc.
- Glasserman, P. 2004. *Monte Carlo methods in financial engineering*. New York: Springer.
- Gurland, J. 2001. Inversion formula for the distribution of ratios. *Annals of Mathematical Statistics*.
- Kahl, C., and P. Jackel. Sept 2005. Not-so-complex logarithms in Heston model. *Wilmott*:94–103.
- Kahl, C., and R. Lord. 2010. Fourier inversion method in finance. *Forthcoming in Handbook of Computational Finance*.
- Komatsu, T., and A. Takeuchi. 2001. Simplified probabilistic approach to the Hormander theorem. *Osaka J. Math.* 38:681–691.
- M. Bismut, J. 1981. Martingales, Malliavin calculus and hypoellipticity under general Hormander’s conditions. *Zeit. Fur Wahr.* 56:469–505.
- Rudin, W. 1986. *Real and complex analysis, 3rd edition*. International Series in Pure and Applied Mathematics: Tata McGrawhill.
- Shephard, N. G. 1991a. Characteristic function to distribution functions: A simple framework for the theory. *Econometric Theory* 7 (4): 519–529.
- Shephard, N. G. 1991b. Numerical integration rules for multivariate inversions. *Journal of Statistical Computation and Simulation* 39 (1): 37–46.
- Shreve, S. E. 2004. *Stochastic calculus for finance II: Continuous-time models*. New York: Springer.

#### AUTHOR BIOGRAPHIES

**SANTANU DEY** is a PhD Student in the School of Technology and Computer Sciences at the Tata Institute of Fundamental Research, Mumbai, India. He received a Bachelors of Statistics from Indian Statistical Institute (ISI)-Kolkata, India, a Master of Statistics from Indian Statistical Institute (ISI)-Kolkata, India. His research focuses on application of stochastic processes and simulation to computational finance. His email address is [dsantanu@tcs.tifr.res.in](mailto:dsantanu@tcs.tifr.res.in).

**SANDEEP JUNEJA** is an Associate Professor School of Technology and Computer Science at the Tata Institute of Fundamental Research. His Ph.D. was from Department of Operations Research at Stanford University (1994). His research interests include applied probability, computational finance, simulation and rare-event analysis. He serves on the editorial board on Mathematics of Operations Research. His email address is [jjuneja@tifr.res.in](mailto:jjuneja@tifr.res.in), and his web page is [www.tcs.tifr.res.in/~sandeepj](http://www.tcs.tifr.res.in/~sandeepj).