Pricing American-style Derivatives under the Heston Model Dynamics: A Fast Fourier Transformation in the Geske–Johnson Scheme

Oleksandr Zhylyevskyy[1]

February 12, 2005

[1]I am greatly indebted to T. Wake Epps for reviewing the draft and to Serguey Khovansky for insightful comments.
Correspondence: Department of Economics, University of Virginia, P.O. Box 400182, Charlottesville, VA 22904-4182. E-mail: oz9a@virginia.edu.

Abstract

Theoretical research on option valuation tends to focus on pricing the plain-vanilla European-style derivatives. Duffie, Pan, and Singleton (2000) have recently developed a general transform method to determine the value of European options for a broad class of the underlying price dynamics. Contrastingly, no universal and analytically attractive approach to pricing of American-style derivatives is yet available. When the underlying price follows simple dynamics, literature suggests using finite difference methods. Simulation methods are often applied in more complicated cases. This paper addresses the valuation of American-style derivatives when the price of an underlying asset follows the Heston model dynamics. The model belongs to the class of stochastic volatility models, which have been proposed in the hope of remedying the strike-price biases of the Black–Scholes formula. Option values are obtained by a variant of the Geske–Johnson scheme (1984), which has been devised in the context of the Black–Scholes model. The scheme exploits the fact that an American option is the limit of a sequence of "Bermudan" derivatives. The latter ones can be priced recursively according to a simple formula, and iterations start from valuing a corresponding European-style security. To implement the recursion, one needs to obtain the expected value of "Bermudan" prices in the joint measure of the state variables of the model. Since the joint density must be, in turn, recovered by inverting the joint characteristic function, an unmodified Geske–Johnson algorithm implies a computationally unfeasible multiple integration. To drastically reduce the cost of numerical integration, I suggest applying a kernel-smoothed bivariate fast Fourier transformation to obtain the density function. Numerical accuracy of the method is assessed by predicting option prices of the S&P 100 index options.

Keywords: American-style option, stochastic volatility model, Geske–Johnson scheme, characteristic function inversion, fast Fourier transform

JEL Code: G13

1. Introduction

Theoretical research on option valuation tends to focus on pricing the plain-vanilla European-style derivatives. Duffie, Pan, and Singleton (2000) showed that such options can be priced by transform methods whenever the state vector (which includes functions of asset prices, unobserved volatilities, etc.) follows a multivariate Gaussian-Poisson affine jump-diffusion. As a result, for a wide class of pricing problems a general solution method has been found.

Contrastingly, no universal and analytically attractive approach is yet available for the American-style derivatives. Still, most traded equity and FX-rate derivatives are the American-style ones. Accurate and efficient pricing of such options is of a significant practical value.

Stochastic volatility models have been proposed in the hope of remedying the strike-price biases in option valuation by the Black–Scholes formula. A model due to Heston (1993) has received considerable attention in the literature. Heston's original method has been modified and simplified by other scholars to deliver a very efficient formula for the European-style puts. To price the American-style derivatives in the two-state-variable setting of the model, authoritative sources (for instance, Wilmott, 2000) strongly suggest using the finite difference (FD) schemes. FD methods, in which the partial differential equation (p.d.e.) in the value function of a derivative security is approximated and solved for the initial option price numerically, are very popular among practitioners and in academia. Applications of FD schemes for the Heston dynamics are available (e.g., Winkler, 2001).

A number of efficient non-FD methods to price the American-style options have been proposed in the context of the Black–Scholes model. A technique due to Broadie and Detemple (1996) is a smoothed binomial scheme. The MacMillan–Barone-Adesi–Whaley approach relies on decomposing the value of an American-style derivative into the value of a corresponding European-style option and early exercise premium. The premium follows the fundamental p.d.e., which can be approximated by the 2^{nd}-order ordinary differential equation that is solved analytically. The Geske–Johnson scheme (1984) exploits the fact that an American-style option is the limit of a sequence of "Bermudan" derivatives. The latter ones can be priced recursively according to a simple formula.

In this paper, I adapt the Geske–Johnson method to the dynamics of the Heston model. As an empirical test of the numerical accuracy of this approach, I consider pricing of the American-style S&P 100 index options (OEX).

The rest of the paper is organized as follows. In Section 2, I state the assumption of the model, which are used in Section 3, to derive the p.d.e. in the value function of a derivative security. Section 4 proceeds at a slow pace from an analytical solution for the joint characteristic function (ch.f.) of log-price and squared volatility

to the specifics of a method of recovering their joint probability density function (p.d.f.): a kernel-smoothed bivariate fast Fourier transformation (FFT). Relevant properties of the proposed kernel are analyzed in the Appendix. Section 5 presents the "Bermudan" recursion formula and outlines the linear Richardson extrapolation scheme. In Section 6, I describe the data, calibrate the parameters, provide illustrations for two selected ch.f.'s and corresponding p.d.f.'s, and, lastly, present the results of option pricing. In Section 7, I conclude.

2. Assumptions

I assume that the following conditions are true. The interest rate is constant and known. The direct costs-of-carry and transaction costs are negligibly small.

The financial market is assumed to admit no arbitrage. At least one asset is traded at a strictly positive price in all states of the world. Then, by the first fundamental theorem of asset pricing, there exists a measure equivalent to the "natural" measure, under which the properly discounted asset prices are martingales. The financial market need not be complete.

I will take for granted that under the equivalent martingale measure, $\hat{\mathbb{P}}$, the evolution of the underlying's price is described by the s.d.e.s:

$$dS_t = (r - \delta)\, S_t dt + \sqrt{v_t} S_t dW_{1t}, \tag{1}$$

$$dv_t = (\alpha - \beta v_t)\, dt + \gamma \sqrt{v_t} dW_{2t}. \tag{2}$$

In equations (1) and (2), symbols have the following meaning. S_t stands for the underlying's price at time t. $r \geq 0$ and $\delta \geq 0$ represent the constant interest rate and (continuous) dividend rate, respectively. v_t is the unobserved state variable and $\sqrt{v_t}$ is referred to as "volatility". Parameters α, β, γ are non-negative. $\{W_{1t}, W_{2t}\}_{t \geq 0}$ are standard Brownian motions on the probability space with filtration mechanism $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \hat{\mathbb{P}}\right)$. The Brownian motion processes are such that $d\langle W_1, W_2 \rangle_t = \rho dt$, where $|\rho| < 1$. In other words, the evolution of the price is governed by two imperfectly correlated sources of risk.

As noted by Chernov and Ghysels (2000), a restriction $\gamma^2 \leq 2\alpha$ must be imposed on the parameters in s.d.e. (2). The restriction guarantees that v_t stays in the open interval $(0, \infty)$ almost surely.[1]

For simplicity, I assume that there exists a money-market fund with a (traded) share worth $M_t = M_0 e^{rt}$, where $M_0 > 0$. M_t may serve are the "numeraire" asset and $\hat{\mathbb{P}}$ may be referenced as the "risk-neutral probability measure".

[1] Chernov and Ghysels restate a result from Cox, Ingersoll, and Ross (1985, p. 391). The latter paper, in turn, refers to Feller (1951).

3. P.D.E.

There are three[2] state variables in the model: observed S_t, unobserved v_t, and observed t. Throughout, t will stand for the current time, T will represent the time of expiration, and $\tau = T - t$ will be referred to as the "time to expiration".

It will be more convenient to operate with a different set of state variables: observed $s_t = \ln S_t$, unobserved v_t, and observed τ. Clearly, for fixed T there is a one-to-one correspondence between the two sets of state variables.

Given equations (1) and (2), by Ito's lemma:

$$ds_t = \left(r - \delta - \frac{v_t}{2}\right) dt + \sqrt{v_t} dW_{1t}, \tag{3}$$

$$dv_t = (\alpha - \beta v_t)\, dt + \gamma \sqrt{v_t} dW_{2t}. \tag{4}$$

It is reasonable to model the value function of a derivative security, D, as a function of the state variables, $D = \mathsf{D}(S_t, v_t, t) = \mathsf{D}(e^{s_t}, v_t, T - (T-t)) = D(s_t, v_t, \tau)$. D is hypothesized to be twice continuously differentiable in (s_t, v_t) and once continuously differentiable in τ.

Let $u \in (t, T]$. By the fundamental theorem, $\hat{E}\left[\frac{D(s_u, v_u, T-u)}{M_u} | \mathcal{F}_t\right] = \frac{D(s_t, v_t, \tau)}{M_t}$. Taking u arbitrarily close to t, $\hat{E}\left[d\frac{D(s_t, v_t, \tau)}{M_t} | \mathcal{F}_t\right] = 0$.

It follows that:

$$\begin{aligned} 0 &= \hat{E}\left[d\frac{D(s_t, v_t, \tau)}{M_t} | \mathcal{F}_t\right] = \hat{E}\left[\frac{dD(s_t, v_t, \tau)}{M_t} - \frac{D(s_t, v_t, \tau)}{M_t^2} dM_t | \mathcal{F}_t\right] = \\ &= M_t^{-1} \hat{E}\left[dD(s_t, v_t, \tau) - D(s_t, v_t, \tau)\, r dt | \mathcal{F}_t\right]. \end{aligned} \tag{5}$$

Expressing $dD(s_t, v_t, \tau)$ by Ito's formula as:

$$\begin{aligned} dD = & \\ &= -D_\tau dt + D_s ds_t + D_v dv_t + \frac{1}{2} D_{ss} d\langle s \rangle_t + \frac{1}{2} D_{vv} d\langle v \rangle_t + D_{sv} d\langle s, v \rangle_t = \\ &= -D_\tau dt + D_s \left[\left(r - \delta - \frac{v_t}{2}\right) dt + \sqrt{v_t} dW_{1t}\right] + D_v \left[(\alpha - \beta v_t)\, dt + \gamma \sqrt{v_t} dW_{2t}\right] + \\ &+ \frac{1}{2} D_{ss} v_t dt + \frac{1}{2} D_{vv} \gamma^2 v_t dt + D_{sv} \rho \gamma v_t dt. \end{aligned}$$

[2] For methodological reasons, I prefer to treat time as a separate state variable. Clearly, it has a deterministic and trivial law of motion.

Taking expectation and simplifying equation (5) :

$$\begin{aligned} 0 &= -D_\tau + D_s\left(r-\delta-\frac{v_t}{2}\right) + D_v\left(\alpha-\beta v_t\right) + \\ &\quad + D_{ss}\frac{v_t}{2} + D_{vv}\gamma^2\frac{v_t}{2} + D_{sv}\rho\gamma v_t - Dr. \end{aligned} \tag{6}$$

Provided that the assumption of continuous differentiability holds and the initial (terminal) conditions are well specified, the value function of a derivative asset can be determined by solving p.d.e. (6). Epps (2004b) shows how to obtain the solution for a special case of a European-style $D : D\left(s_T, v_T, 0\right) = e^{\zeta s_T}$. No closed-form solution is available for an American-style derivative security.

4. Joint Density of (s_T, v_T) under $\hat{\mathbb{P}}$

4.1. Joint Ch.F.

Consider time $u \in (t, T]$. Let the conditional joint ch.f. of (s_T, v_T) be:

$$\Psi\left(u\right) \equiv \hat{E}\left[e^{i(\zeta_1 s_T + \zeta_2 v_T)} | \mathcal{F}_u\right] = \Psi\left(\zeta_1, \zeta_2; s_u, v_u, T-u\right).$$

Above, conditioning on just the three state variables $s_u, v_u, T-u$ vs. the whole information set $\mathcal{F}_u$ is motivated by the Markovian property of Brownian motions.

By the tower property:

$$\Psi\left(t\right) = \hat{E}\left[e^{i(\zeta_1 s_T + \zeta_2 v_T)} | \mathcal{F}_t\right] = \hat{E}\left[\hat{E}\left[e^{i(\zeta_1 s_T + \zeta_2 v_T)} | \mathcal{F}_u\right] | \mathcal{F}_t\right] = \hat{E}\left[\Psi\left(u\right) | \mathcal{F}_t\right].$$

Therefore, taking u arbitrarily close to t, it follows that $\hat{E}\left[d\Psi\left(t\right) | \mathcal{F}_t\right] = 0$.

By Ito's lemma:

$$d\Psi\left(\zeta_1, \zeta_2; s_t, v_t, \tau\right) = -\Psi_\tau dt + \Psi_s ds_t + \Psi_v dv_t + \frac{1}{2}\Psi_{ss} d\left\langle s\right\rangle_t + \frac{1}{2}\Psi_{vv} d\left\langle v\right\rangle_t + \Psi_{sv} d\left\langle s, v\right\rangle_t.$$

Using equations (3) and (4), taking expectation, factoring out dt and simplifying, the conditional joint ch.f. $\Psi\left(\zeta_1, \zeta_2; s_t, v_t, \tau\right)$ solves p.d.e.:

$$0 = -\Psi_\tau + \Psi_s\left(r-\delta-\frac{v_t}{2}\right) + \Psi_v\left(\alpha-\beta v_t\right) + \Psi_{ss}\frac{v_t}{2} + \Psi_{vv}\gamma^2\frac{v_t}{2} + \Psi_{sv}\rho\gamma v_t. \tag{7}$$

Let the trial solution be:

$$\Psi(t) = \Psi(\zeta_1, \zeta_2; s_t, v_t, \tau) = \exp\left[p(\tau; \zeta_1, \zeta_2) + q(\tau; \zeta_1, \zeta_2) v_t + i\zeta_1 s_t\right],$$

where $p(\tau; \zeta_1, \zeta_2)$ and $q(\tau; \zeta_1, \zeta_2)$ are complex-valued functions.

Obviously:

$$\begin{aligned} \Psi_\tau/\Psi &= p_\tau + q_\tau v_t,\ \Psi_s/\Psi = i\zeta_1,\ \Psi_v/\Psi = q, \\ \Psi_{ss}/\Psi &= (i\zeta_1)^2,\ \Psi_{vv}/\Psi = q^2,\ \Psi_{sv}/\Psi = i\zeta_1 q. \end{aligned}$$

Plugging these into (7), factoring out Ψ, and simplifying:

$$0 = \left[-p_\tau + i\zeta_1(r-\delta) + \alpha q\right] + v_t\left[-q_\tau - \frac{1}{2}i\zeta_1 - \beta q + \frac{1}{2}(i\zeta_1)^2 + \frac{1}{2}\gamma^2 q^2 + i\zeta_1\rho\gamma q\right].$$

Since p.d.e. (7) holds for all values of v_t, it must be the case that functions p and q are the solution to the system of ordinary differential equations:

$$\begin{aligned} q_\tau &= \frac{1}{2}\left[(i\zeta_1)^2 - i\zeta_1\right] + \left[i\zeta_1\rho\gamma - \beta\right] q + \frac{1}{2}\gamma^2 q^2, \\ p_\tau &= i\zeta_1(r-\delta) + \alpha q. \end{aligned}$$

These equations are similar to the ones in Epps (2004b). However, the initial conditions are determined by:

$$\Psi(\zeta_1, \zeta_2; s_T, v_T, 0) = \exp\left[i\zeta_2 v_T + i\zeta_1 s_T\right].$$

Therefore, $q(0; \zeta_1, \zeta_2) = i\zeta_2$ and $p(0; \zeta_1, \zeta_2) = 0$.

Solutions were obtained with Maple.[3]

Case 1. $\gamma \neq 0$.

Let:

$$\begin{aligned} A &\equiv A(\zeta_1) = \gamma^2\left(1-\rho^2\right)\zeta_1^2 + \left(\gamma^2 - 2\rho\gamma\beta\right) i\zeta_1 + \beta^2, \\ B &\equiv B(\zeta_1, \zeta_2) = -\frac{\rho\gamma i\zeta_1 - \beta - \sqrt{A(\zeta_1)} + \gamma^2 i\zeta_2}{\rho\gamma i\zeta_1 - \beta + \sqrt{A(\zeta_1)} + \gamma^2 i\zeta_2}. \end{aligned}$$

[3] In the most interesting case, $\gamma \neq 0$, Maple gives solution either in terms of trigonometric and inverse trigonometric functions or in terms of hyperbolic functions with a non-closed-form expression for $p(\tau; \cdot)$. The "trigonometric" solution is perfectly acceptable, but is not convenient to program. I started with the "hyperbolic" solution and derived a closed-form expression for $p(\tau; \cdot)$.

Then:

$$\begin{aligned} q(\tau;\zeta_1,\zeta_2) &= \frac{1}{\gamma^2}\left[\beta-\rho\gamma i\zeta_1-\sqrt{A}\frac{Be^{\tau\sqrt{A}}-1}{Be^{\tau\sqrt{A}}+1}\right]. \\ p(\tau;\zeta_1,\zeta_2) &= \tau\left(r-\delta-\frac{\alpha\rho}{\gamma}\right)i\zeta_1+\frac{\alpha}{\gamma^2}\left[\tau\beta+\tau\sqrt{A}+2\ln\frac{B+1}{Be^{\tau\sqrt{A}}+1}\right]. \end{aligned}$$

Caveat. *This solution is valid for* $(\zeta_1,\zeta_2)' \neq (0,0)'$. *Given the trial solution for* Ψ *and the definition of a ch.f., it must be the case that* $q(\tau;0,0) = p(\tau;0,0) = 0$, $\forall\tau$. *I check that the above expressions for* $q(\tau;\zeta_1,\zeta_2)$ *and* $p(\tau;\zeta_1,\zeta_2)$ *do not contradict the uniform continuity property of* Ψ, *that is, whether* $\lim_{(\zeta_1,\zeta_2)'\to(0,0)'} q(\tau;\zeta_1,\zeta_2) = \lim_{(\zeta_1,\zeta_2)'\to(0,0)'} p(\tau;\zeta_1,\zeta_2) = 0$:

$$\begin{aligned} \lim_{(\zeta_1,\zeta_2)'\to(0,0)'} q(\tau;\zeta_1,\zeta_2) &= \frac{1}{\gamma^2}\left[\beta-\sqrt{\beta^2}\frac{e^{\tau\sqrt{\beta^2}}-\lim 1/B(\zeta_1,\zeta_2)}{e^{\tau\sqrt{\beta^2}}+\lim 1/B(\zeta_1,\zeta_2)}\right] = \\ &= \frac{1}{\gamma^2}[\beta-\beta]=0, \\ \lim_{(\zeta_1,\zeta_2)'\to(0,0)'} p(\tau;\zeta_1,\zeta_2) &= \frac{\alpha}{\gamma^2}\left[\tau\beta+\tau\sqrt{\beta^2}+2\ln\frac{1+\lim 1/B(\zeta_1,\zeta_2)}{e^{\tau\sqrt{\beta^2}}+\lim 1/B(\zeta_1,\zeta_2)}\right] = \\ &= \frac{\alpha}{\gamma^2}\left[\tau\beta+\tau\beta-2\ln e^{\tau\beta}\right]=0, \end{aligned}$$

because $\lim_{(\zeta_1,\zeta_2)'\to(0,0)'} B(\zeta_1,\zeta_2) = -\frac{-\beta-\sqrt{\beta^2}}{-\beta+\sqrt{\beta^2}} = \infty$, *if* $\beta > 0$ *(it is straightforward to extend the proof to the subcase* $\beta = 0$*).*

Case 2. $\gamma = 0$, $\beta > 0$.

$$\begin{aligned} q(\tau;\zeta_1,\zeta_2) &= \frac{1}{2\beta}\left[e^{-\beta\tau}\left(\zeta_1^2+i\zeta_1+2\beta i\zeta_2\right)-\zeta_1^2-i\zeta_1\right], \\ p(\tau;\zeta_1,\zeta_2) &= \tau(r-\delta)i\zeta_1-\frac{\alpha}{2\beta}\left[2\left(e^{-\beta\tau}-1\right)i\zeta_2+\tau i\zeta_1+\tau\zeta_1^2\right]- \\ &\quad -\frac{\alpha\left(e^{-\beta\tau}-1\right)}{2\beta^2}\left[\zeta_1^2+i\zeta_1\right]. \end{aligned}$$

Case 3. $\gamma = \beta = 0$.

$$\begin{aligned} q(\tau;\zeta_1,\zeta_2) &= -\frac{\tau}{2}\left[\zeta_1^2+i\zeta_1\right]+i\zeta_2, \\ p(\tau;\zeta_1,\zeta_2) &= \tau[r-\delta]i\zeta_1-\frac{\alpha\tau}{4}\left[\tau\left(\zeta_1^2+i\zeta_1\right)-4i\zeta_2\right]. \end{aligned}$$

Black–Scholes obtains with $\gamma = \beta = \alpha = 0$ *and* $p(\tau;\zeta_1,\zeta_2) = \tau[r-\delta]i\zeta_1$.

4.2. Mathematical Issues

4.2.1. Multivariate Ch.F. and Inversion Theorem

Let $\mathbf{X} = (X_1, ..., X_p)'$ be a $p \times 1$ random vector with c.d.f. $F_{\mathbf{X}}$ and consider arbitrary $\boldsymbol{\zeta} \in \Re^p$. The ch.f. of $\mathbf{X}$ is the Fourier transform of $F_{\mathbf{X}}$:

$$\Psi_{\mathbf{X}}(\boldsymbol{\zeta}) = \int \cdots \int_{\Re^p} e^{i\boldsymbol{\zeta}'\mathbf{x}} dF_{\mathbf{X}}.$$

Now, suppose that $\mathbf{X}$ has a density function $f_{\mathbf{X}}$. If $\Psi_{\mathbf{X}}$ is Lebesgue integrable ($\Psi_{\mathbf{X}} \in L^1(\Re^p)$), then, by the inversion theorem:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^p} \int \cdots \int_{\Re^p} e^{-i\boldsymbol{\zeta}'\mathbf{x}} \Psi_{\mathbf{X}}(\boldsymbol{\zeta})\, d\boldsymbol{\zeta}.$$

Further extensions are considered in Shephard (1991a) and Shephard (1991b).

4.2.2. Riemann–Lebesgue Lemma

Suppose that $\mathbf{X}$ has a density function $f_{\mathbf{X}}$. Since the density function is nonnegative, then, $f_{\mathbf{X}} = |f_{\mathbf{X}}|$, and since it integrates to 1 on $\Re^p$, $f_{\mathbf{X}} \in L^1(\Re^p)$. Then, by the multivariate extension of the Riemann–Lebesgue lemma in Rudin (1991, theorem 7.5), its Fourier transform, $\Psi_{\mathbf{X}}$, belongs to $C_0(\Re^p)$, where $C_0(\Re^p)$ is the supremum-normed Banach space of all complex continuous functions on $\Re^p$ that *vanish at infinity*.

A precise statement of this important fact is that for any $\epsilon > 0$ there exists a *compact* subset $\mathbf{K}_\epsilon$ in the domain $\boldsymbol{\vartheta}$ such that $|\Psi_{\mathbf{X}}(\boldsymbol{\zeta})| < \epsilon$ for any $\boldsymbol{\zeta} \in \boldsymbol{\vartheta} \setminus \mathbf{K}_\epsilon$. In particular, $\Psi_{\mathbf{X}} \in C_0(\Re^p)$ implies:

$$\lim_{\text{any } \zeta \to \infty} \Psi_{\mathbf{X}}(\boldsymbol{\zeta}) = 0 + i \cdot 0, \tag{8}$$

as a sequence of vectors with at least one exploding component cannot be contained in any compact set.

Moreover, $\|\Psi_{\mathbf{X}}\|_\infty \leq \|f_{\mathbf{X}}\|_1$, where $\|g\|_r = \left\{(2\pi)^{-\frac{p}{2}} \int \cdots \int_{\Re^p} |g|^r \, d\mathbf{x}\right\}^{\frac{1}{r}}$, $1 \leq r < \infty$ (and for $r = \infty$ it is the essential supremum of $|g| \times (2\pi)^{-\frac{p}{2}}$).[4] $\|f_{\mathbf{X}}\|_1$ is a finite number, $(2\pi)^{-\frac{p}{2}}$. Then, $\Psi_{\mathbf{X}} \in L^\infty(\Re^p)$, which, unfortunately, does *not* imply that $\Psi_{\mathbf{X}} \in L^1(\Re^p)$.[5]

4.3. Joint P.D.F.

It is beyond the scope of this paper to establish restrictions on the underlying parameter vector, $\mathbf{c} = (r, \delta, s_t, v_t, \tau, \alpha, \beta, \gamma, \rho)'$, that guarantee absolute continuity of the random vector (s_T, v_T), that is, existence of the density *per se*. My best guess is that a rigorous proof would appeal to conditions under which Ito processes are continuous semimartingales and, certainly, to the condition under which $\{v_l\}_{t \leq l \leq T} \geq 0$,

[4] Rudin integrates with respect to the *normalized* Lebesgue measure, m_p, to preserve the symmetry of the Fourier forward and inverse transformations.

[5] Relationship $L^r(\Re^p) \subset L^s(\Re^p)$ when $0 < s \leq r \leq \infty$ does not hold for m_p, since $m_p(\Re^p) = +\infty$.

$\gamma^2 \leq 2\alpha$. Likewise, I do not undertake a research into the issue of which $\mathbf{c}$ makes $\Psi_{\mathbf{X}} \in L^1(\Re^p)$. To proceed, I make an assumption that the density exists and the ch.f. is Lebesgue integrable.

As noted above, given the closed-form expressions for $p(\tau;\zeta_1,\zeta_2)$ and $q(\tau;\zeta_1,\zeta_2)$, the analytical solution for the conditional joint ch.f. is $\Psi(\zeta_1,\zeta_2;s_t,v_t,\tau) = \exp[p(\tau;\zeta_1,\zeta_2) + q(\tau;\zeta_1,\zeta_2)v_t + i\zeta_1 s_t]$.

Fourier transforming Ψ, the joint density of (s_T, v_T) is:

$$f(s_T,v_T;s_t,v_t,\tau) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2;s_t,v_t,\tau)\,d\zeta_1 d\zeta_2. \tag{9}$$

For the sake of completeness, note that the *argument vector* of the joint ch.f. is $\boldsymbol{\zeta} = (\zeta_1,\zeta_2)'$ and the argument vector of the joint p.d.f. is $\mathbf{x} = (s_T, v_T)$. Both functions have the same *parameter vector* $\mathbf{c}$ (defined above).

4.4. Numerical Integration

In what follows, the p.d.f. and ch.f. are denoted as $f(s_T,v_T)$ and $\Psi(\zeta_1,\zeta_2)$. $e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\cdot\Psi(\zeta_1,\zeta_2)$ will be referred to as "the integrand function".

The primary interest in implementing inversion (9) is whether there exists a compact subset such that $\left|\mathrm{Re}\left[e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right]\right|$ is negligible on its complement. By specifying a rectangle that would encompass this compact subset, (9) can be approximated by a definite Riemann integral.

Establishing that $\mathrm{Re}\left[e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right]$ vanishes at infinity is fairly easy, provided that $f(s_T,v_T)$ exists. Apparently:

$$\begin{aligned}\left|\mathrm{Re}\left[e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right]\right| &\equiv \sqrt{\mathrm{Re}\left[e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right]^2} \\ \leq \sqrt{\mathrm{Re}\left[e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right]^2 + \mathrm{Im}\left[e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right]^2} &\equiv \\ \equiv \left|e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right| = \left|e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\right|\cdot\left|\Psi(\zeta_1,\zeta_2)\right| &= \left|\Psi(\zeta_1,\zeta_2)\right|.\end{aligned} \tag{10}$$

Since $\Psi(\zeta_1,\zeta_2)$ vanishes by the Riemann–Lebesgue lemma, $\mathrm{Re}[\cdot]$ vanishes as well.

Therefore, if a_1, b_1, a_2, b_2 are "large" in absolute value: $a_1 < 0, b_1 > 0, a_2 < 0, b_2 > 0$, then, a valid approximation is:

$$f(s_T,v_T) \cong \frac{1}{(2\pi)^2}\int_{a_2}^{b_2}\int_{a_1}^{b_1}\mathrm{Re}\left[e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\right]d\zeta_1 d\zeta_2.$$

The integrand function may approach zero at different rates in each direction. Given the expressions for A, B, p, and q, it is very likely that the integrand goes to zero faster in ζ_1 direction. This has direct implications for numerical integration: the cutoff points may be chosen such that: $|a_1| < |a_2|$ and $|b_1| < |b_2|$.

A choice of some particular cutoff points is case dependent.

The integrand function may exhibit periodicity, as it can always be rewritten in terms of sines and cosines. Implications of this fact are explored in detail later.

4.5. FFT

Numerical integration of complicated bivariate functions is computationally demanding. Fortunately, (9) is a Fourier integral. It may be possible to apply very efficient FFT algorithms.

A symmetric Fourier transform pair of bivariate functions is:

$$\begin{aligned} f(x,y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(k_x,k_y)\, e^{-i\cdot 2\pi(xk_x+yk_y)} dk_x dk_y, \\ F(k_x,k_y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{i\cdot 2\pi(k_x x+k_y y)} dx dy. \end{aligned}$$

FFT algorithms are applied to a discrete version of the above relationship.

Returning to equation (9), the goal is to approximate $(2\pi)^2 f(s_T, v_T)$ as $\int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\, d\zeta_1 d\zeta_2$, where the cutoff points are sufficiently large in absolute value. Define:

$$\Delta_1 = \frac{b_1-a_1}{N_1},\ \Delta_2 = \frac{b_2-a_2}{N_2},$$

$$\zeta_{j_1} = a_1 + j_1\Delta_1,\ \zeta_{j_2} = a_2 + j_2\Delta_2,\ \Psi_{j_1,j_2} = \Psi\left(\zeta_{j_1},\zeta_{j_2}\right),\ j_1 = 0,...,N_1,\ j_2 = 0,...,N_2.$$

Then:

$$\int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i(\zeta_1 s_T+\zeta_2 v_T)}\Psi(\zeta_1,\zeta_2)\, d\zeta_1 d\zeta_2 \cong \Delta_1\Delta_2 \sum_{j_2=0}^{N_2-1}\sum_{j_1=0}^{N_1-1} e^{-i\left(s_T\zeta_{j_1}+v_T\zeta_{j_2}\right)}\Psi_{j_1,j_2}.$$

Now, *define*:

$$s_{T,k_1} = \frac{2\pi k_1}{N_1\Delta_1} = \frac{2\pi k_1}{b_1-a_1},\ v_{T,k_2} = \frac{2\pi k_2}{N_2\Delta_2} = \frac{2\pi k_2}{b_2-a_2},$$

where k_1 and k_2 are on the same grid as j_1 and j_2, respectively.

Then:

$$\begin{aligned} (2\pi)^2 f(s_{T,k_1}, v_{T,k_2}) &\cong \Delta_1\Delta_2 \sum_{j_2=0}^{N_2-1}\sum_{j_1=0}^{N_1-1} e^{-i\left(s_{T,k_1}\zeta_{j_1}+v_{T,k_2}\zeta_{j_2}\right)}\Psi_{j_1,j_2} = \\ &= \Delta_1\Delta_2 e^{-i\left(s_{T,k_1}a_1+v_{T,k_2}a_2\right)} \sum_{j_2=0}^{N_2-1}\sum_{j_1=0}^{N_1-1} e^{-i\cdot 2\pi\left(k_1\frac{j_1}{N_1}+k_2\frac{j_2}{N_2}\right)}\Psi_{j_1,j_2}. \end{aligned}$$

The double summation term is *by definition* a discrete Fourier transform, for which FFT algorithms are

available.

As noted by Press et al. (1992, p. 578), this approximation is apt to be imprecise. The sources of the innacuracy are the possible error of truncation (if the function is not negligibly small at the cutoff boundary), and the periodicity of the integrand function.

The first criticism does not apply as the property of vanishing at infinity has been established for the integrand function. The second criticism does apply, however. Numerical integration of oscillatory functions is a painstaking and often unsuccessful endeavor.

One possible way to mitigate the problem is to do kernel smoothing. I extend Press et al. solution to the bivariate case. Interpolate $\Psi(\zeta_1, \zeta_2)$ as follows:

$$\begin{aligned} \Psi(\zeta_1, \zeta_2) \quad \cong \quad & \sum_{j_2=0}^{N_2} \sum_{j_1=0}^{N_1} \Psi_{j_1,j_2} K\left(\frac{\zeta_1 - \zeta_{j_1}}{\Delta_1}, \frac{\zeta_2 - \zeta_{j_2}}{\Delta_2}\right) + \\ & \sum_{j_1,j_2 \in \{\text{endpoints}\}} \Psi_{j_1,j_2} K_{j_1,j_2}\left(\frac{\zeta_1 - \zeta_{j_1}}{\Delta_1}, \frac{\zeta_2 - \zeta_{j_2}}{\Delta_2}\right), \end{aligned} \tag{11}$$

where $K\left(\frac{\zeta_1-\zeta_{j_1}}{\Delta_1}, \frac{\zeta_2-\zeta_{j_2}}{\Delta_2}\right)$ is the kernel function and $K_{j_1,j_2}\left(\frac{\zeta_1-\zeta_{j_1}}{\Delta_1}, \frac{\zeta_2-\zeta_{j_2}}{\Delta_2}\right)$ is the difference between the true kernel function at endpoints and $K(\cdot,\cdot)$.

Since the cutoff points are sufficiently large in absolute value, Ψ_{j_1,j_2} for $j_1, j_2 \in \{\text{endpoints}\}$ is negligibly small. So, the second summation in (11) may be ignored.

Now, apply $\int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i(\zeta_1 s_T + \zeta_2 v_T)} d\zeta_1 d\zeta_2$ to both sides of (11):

$$\int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i(\zeta_1 s_T + \zeta_2 v_T)} \Psi(\zeta_1, \zeta_2)\, d\zeta_1 d\zeta_2 \cong$$

$$\int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i(\zeta_1 s_T + \zeta_2 v_T)} \sum_{j_2=0}^{N_2} \sum_{j_1=0}^{N_1} \Psi_{j_1,j_2} K\left(\frac{\zeta_1 - \zeta_{j_1}}{\Delta_1}, \frac{\zeta_2 - \zeta_{j_2}}{\Delta_2}\right) d\zeta_1 d\zeta_2 =$$

$$= \sum_{j_2=0}^{N_2} \sum_{j_1=0}^{N_1} \left[\int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i(\zeta_1 s_T + \zeta_2 v_T)} K\left(\frac{\zeta_1 - \zeta_{j_1}}{\Delta_1}, \frac{\zeta_2 - \zeta_{j_2}}{\Delta_2}\right) d\zeta_1 d\zeta_2\right] \Psi_{j_1,j_2}.$$

Change variables as $\frac{\zeta_1-\zeta_{j_1}}{\Delta_1} = x$, $\frac{\zeta_2-\zeta_{j_2}}{\Delta_2} = y$:

$$(2\pi)^2 f(s_T, v_T) \cong \Delta_1 \Delta_2 W(\theta_{s_T}, \theta_{v_T}) \sum_{j_2=0}^{N_2} \sum_{j_1=0}^{N_1} e^{-i\left(s_T \zeta_{j_1} + v_T \zeta_{j_2}\right)} \Psi_{j_1,j_2},$$

where $W(\theta_{s_T}, \theta_{v_T}) = \left[\int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i\left(\theta_{s_T} x + \theta_{v_T} y\right)} K(x,y)\, dxdy\right]$, $\theta_{s_T} = \Delta_1 s_T$, $\theta_{v_T} = \Delta_2 v_T$. An important fact

to notice is that the weighting function $W\left(\theta_{s_T},\theta_{v_T}\right)$ must be real and nonnegative.

Then, at $s_{T,k_1}=\frac{2\pi k_1}{N_1\Delta_1}$, $v_{T,k_2}=\frac{2\pi k_2}{N_2\Delta_2}$:

$$\left(2\pi\right)^2 f\left(s_{T,k_1},v_{T,k_2}\right)\cong\Delta_1\Delta_2 W\left(\theta_{s_{T,k_1}},\theta_{v_{T,k_2}}\right)\sum_{j_2=0}^{N_2}\sum_{j_1=0}^{N_1}e^{-i\left(s_{T,k_1}\zeta_{j_1}+v_{T,k_2}\zeta_{j_2}\right)}\Psi_{j_1,j_2}\cong$$

$$\cong W\left(\theta_{s_{T,k_1}},\theta_{v_{T,k_2}}\right)\Delta_1\Delta_2\sum_{j_2=0}^{N_2-1}\sum_{j_1=0}^{N_1-1}e^{-i\left(s_{T,k_1}\zeta_{j_1}+v_{T,k_2}\zeta_{j_2}\right)}\Psi_{j_1,j_2}=$$

$$=W\left(\theta_{s_{T,k_1}},\theta_{v_{T,k_2}}\right)\Delta_1\Delta_2 e^{-i\left(s_{T,k_1}a_1+v_{T,k_2}a_2\right)}\sum_{j_2=0}^{N_2-1}\sum_{j_1=0}^{N_1-1}e^{-i\cdot 2\pi\left(k_1\frac{j_1}{N_1}+k_2\frac{j_2}{N_2}\right)}\Psi_{j_1,j_2}.$$

It remains to propose a kernel function and derive $W\left(\theta_{s_T},\theta_{v_T}\right)$.

My suggestion is to use a simple kernel function of the form:

$$K\left(x,y\right)=\begin{cases}\frac{\left(1-|x|\right)^2\left(1-|y|\right)^2}{x^2y^2+x^2\left(1-|y|\right)^2+\left(1-|x|\right)^2y^2+\left(1-|x|\right)^2\left(1-|y|\right)^2}, & \text{if both } |x|\leq 1 \text{ and } |y|\leq 1\ .\\ 0, & \text{elsewhere.}\end{cases}$$

Properties of $K\left(x,y\right)$ are discussed in the Appendix. It turns out that the weighting function does not have a closed-form expression. However, $W\left(\theta_{s_T},\theta_{v_T}\right)$ can be approximated with arbitrary precision.

5. Geske–Johnson Scheme with Richardson Extrapolation

Consider a sequence of "Bermudan"-style derivative securities, $\left\{D_n\left(s_t,v_t,T-t\right)\right\}_{n=1}^{\infty}$, where each D_n can be exercised just at times $t_j=t+\frac{j(T-t)}{n}$, $j=1,...,n$, prior to expiration at T. An American-style option is the limit of the sequence. D_1 is the value of a corresponding European-style derivative.

If $EX\left(s_{t'},v_{t'},T-t'\right)$ is the exercise value of the security at t', then, D_n's obey the following recursion:

$$D_n\left(s_t,v_t,T-t\right)= \tag{12}$$

$$=e^{-r(t_1-t)}\hat{E}\left[\max\left\{EX\left(s_{t_1},v_{t_1},T-t_1\right),D_{n-1}\left(s_{t_1},v_{t_1},T-t_1\right)\right\}\right].$$

For a put, $EX\left(s_{t'},v_{t'},T-t'\right)=\left(X-e^{s_{t'}}\right)^+$ and for a call, $EX\left(s_{t'},v_{t'},T-t'\right)=\left(e^{s_{t'}}-X\right)^+$. X denotes the strike price. Equation (12) is similar to (7.7) in Epps (2004a).

Next, define:

$$h_n=\frac{T-t}{n},\ n=1,2,...$$

Derivative security D_n can be represented as a polynomial of infinite order in h_n :

$$D_n = a_0 + a_1 h_n + a_2 h_n^2 + \cdots .$$

Clearly, $D_\infty = a_0$ and its linear approximation is:

$$\begin{array}{rcl} D_1 & \cong & a_0 + a_1 h_1 \\ D_2 & \cong & a_0 + a_1 h_2 \end{array} \Longrightarrow D_\infty \cong \frac{h_2 D_1 - h_1 D_2}{h_2 - h_1} = 2D_2 - D_1. \tag{13}$$

As noted, D_1 can be obtained as the value of the corresponding European-style option. D_2 can be calculated from recursion (12). This task is sufficiently difficult *per se*, so, quadratic or higher order approximations for D_∞ are not feasible.

6. Empirical Application

6.1. Data

The tools developed in the preceding sections are used to predict prices of S&P 100 index options.

The S&P U.S. 100, a subset of the S&P 500, is comprised of 100 leading U.S. stocks, which together represent almost 45 percent of the market capitalization of the U.S. equity market. As of June 2004, the 5 largest companies included were: General Electric, Exxon Mobil Corp., Microsoft Corp., Pfizer Inc., and Citigroup Inc. The S&P 100 index was originally developed by the CBOE and later transfered to Standard & Poor's for management.

The CBOE offers three distinct options on S&P 100. Two of these have relevance to this paper. The most popular one is the American-style S&P 100 index option, commonly known by its ticker OEX. As a rule, on each trading day, OEX puts and calls have high trading volumes and open interest for a wide range of strikes.

Less popular, but still actively traded at CBOE is the European-style S&P 100 index option, ticker XEO. Both OEX and XEO are cash-settled, and apart from the difference in the exercise style, share same characteristics (exercise dates, minimal strike intervals, minimum ticks, etc.).

I collected closing CBOE prices of both OEX and XEO on 7 consecutive trading days: June 30^{th} – July 2^{nd}, July 6^{th} – July 9^{th}. Only options with positive trading volume and open interest are used in estimations. July 9^{th} data is set aside for out-of-the-sample predictions.

As a proxy for the risk-free interest rate I use the T-bill rate. Rates for different times to maturity are obtained by interpolation: linear OLS fit to the whole set of T-bill quotes reported in the Wall Street Journal

(R^2 is above 90 percent). These interpolated rates range from 1 percent to 1.5 percent (annual).

Unlike S&P 500, S&P 100 index futures are not traded at organized exchanges. Therefore, an independent source of data on the dividend rate is not available. Standard & Poor's routinely collects data related to the performance of its indices and publishes dividend rates once a month. However, these rates are the *ex-post* ones, which are apt to diverge from the daily market's assessment of the future dividend streams. Therefore, I employ the data on the European-style S&P 100 index options to infer dividend rates from the European put-call parity relationship, $C_t - P_t = S_t e^{-\delta(T-t)} - e^{-r(T-t)}X$. To obtain δ, I always take XEO calls and puts with high trading volume and open interest at a strike price closest to the underlying's price. Interpolated dividend rates range from 1 percent to 4 percent (annual).

Since XEO options are used to calibrate the parameters only, I converted XEO calls into puts according to the parity relationship. A summary of the option data is presented in Tables 1 and 2.[6]

6.2. Parameter Calibration

In principle, there are several ways to obtain the parameters of the model. A direct and analytically appealing approach is to use the underlying's price data only (together with "known" r and δ) and estimate the parameters by maximum likelihood.

However, for several reasons, I do not dare to use MLE. First, *even if* v_t were observable for every t corresponding to the sample time-series data points, sample likelihood may be computationally prohibitive to evaluate just for *one* trial set of parameters. In present context, the sample likelihood function is the joint density of a *series* of random vectors $(s_t, v_t)'$. A well-known technique from the time-series analysis is to represent this density as a product of conditional densities. Each of these would have the form specified by equation (9), as it is sufficient to condition on the most recent previous values of the state variables. Therefore, in analytical terms, sample likelihood presents no difficulty. Nevertheless, numerical integration in (9) is computationally demanding and must be performed as many times as there are data points in the sample. Further parameter search would require this whole task to be repeated iteratively, until the maximum of the sample likelihood is achieved.

Second, in reality the state variable v_t is not observed. Still, it is a necessary component in the option pricing formula. The combined task of simultaneously estimating v_t series and maximizing the sample likelihood is, indeed, daunting.

An approach to recovering v_t series has been suggested by Chernov and Ghysels (2000). Their filtering method is based on the reprojection procedure introduced by Gallant and Tauchen (1998).[7] Chernov

[6] S&P U.S. 100 index closing prices were: 553.87 (June 30), 549.01, 547.17, 543.33, 544.25, 540.21, and 542.63 (July 9).

[7] "Reprojecting partially observable systems with application to interest rate diffusions," *Journal of American Statistical Association* 93, 10–24.

and Ghysels claim that other existing filtration schemes (extensions of Kalman filter) cannot accommodate derivative security market information.

Unmodified MLE is, clearly, not feasible. To my best knowledge, attempts are being undertaken elsewhere to reduce the computational burden of the Fourier inversion (e.g., by approximating the integrand function with an easy-to-integrate function, by using quasi-Monte-Carlo integration schemes, etc.).

I take a different route by literally *calibrating*, rather than estimating, the parameters. This is done by minimizing the sum of squared differences of actual XEO put prices and their corresponding predicted values on June 30^{th} through July 8^{th} data (134 options). As the parameters are specific to the price evolution of the underlying asset, there is no need to use OEX data at this stage.

Since volatility is changing from day to day, the objective function, SSQR, is treated as dependent on ten parameters: $v_{t,01}, v_{t,02}, v_{t,03}, v_{t,04}, v_{t,05}, v_{t,06}, \alpha, \beta, \gamma, \rho$. Parameters $\alpha, \beta, \gamma, \rho$ are common across all options. Parameters $v_{t,\#\#}$ are "day-specific", that is, $v_{t,01}$ is v_t at market's closingon June 30^{th},..., $v_{t,06}$ is v_t at market's closing on July 8^{th}.

Previous research has shown that the objective function is difficult to minimize as local extrema abound. Therefore, I use a relatively powerful, but resource consuming simulated annealing algorithm. The objective function is set to a "penalty" value once the trial parameter set violates $\gamma^2 \leq 2\alpha$. The algorithm allows to customize the lower and upper bounds for all parameters. For ρ, the bounds are set to theoretical -1 and 1. For the remaining nine parameters, the bounds are practically unrestricted.

Optimal parameter values are presented in Table 3. A few facts are worth noting. First, in the optimum, the restriction $\gamma^2 \leq 2\alpha$ is *non-binding*. This important result guarantees that $v_t \in (0, \infty)$ almost surely. Second, estimated $v_{t,\#\#}$'s are of the same order of magnitude as the long-run value of v, $\frac{\alpha}{\beta}$. Third, calibrated $\rho < 0$. Loosely speaking, this conforms to the empirical observation that the variance of log-returns is inversely related to the initial price level. So, I conclude that the calibrated parameters make sense.

To make out-of-the-sample predictions, $v_{t,07}$ (at market's closing time on July 9^{th}) is set to: $v_{t,07} = v_{t,06} + (\alpha - \beta v_{t\,,06})\,\Delta_t \cong 0.015848$, where $\Delta_t = \frac{1}{366}$.

6.3. Ch.F.

The shape of the integrand function, $e^{-i(\zeta_1 s_T + \zeta_2 v_T)} \cdot \Psi(\zeta_1, \zeta_2)$, has important implications for FFT. As a rough illustration, consider Figures 1 and 2. In these Figures, I plot the real part of the integrand function with parameters as of June 30. In view of result (10), for simplicity, I look at a special case of $s_T = v_T = 0$, that is, when the integrand is identically the ch.f. Also, I restrict attention to just 2 values of τ: a "small" one, which corresponds to $\frac{1}{2}$ of the time span (in years) between June 30 and 3^{rd} Friday of July, and a "large" one, which stands for $\frac{1}{2}$ of the time till 3^{rd} Friday of December.

$\mathrm{Re}\,[\Psi]$ is an even function with damped oscillations. Interestingly, for small τ, oscillations appear to be more persistent and the function is still *non*-negligibly small at $\zeta_1 = \pm 300$ and $\zeta_2 \in (-300, 300)$. For large τ, the integrand has a pronounced peak at $(0, 0)$ and oscillations quickly fade out. The function approaches zero at different rates in each direction. As predicted, the rate of convergence is slower in ζ_2 direction.

Therefore, in FFT, the cutoff points may be chosen (and in Fortran programs *are* chosen) such that: $|a_1| < |a_2|$ and $|b_1| < |b_2|$. Moreover, for smaller τ it is desirable to have relatively large cutoffs (in absolute value). For larger τ, $|a_1|$, $|a_2|$, $|b_1|$, $|b_2|$ may be much smaller, but Δ_1 and Δ_2 must be as tiny as possible to "capture" the peak.

It should be evident that numerical integration of the integrand function is unlikely to succeed. My experiments with different numerical integration routines reveal the following. Fast 5-degree polynomial methods produce very innacurate results and large negative values for many choices of s_T and v_T. Quadrature methods (Gauss-Kronrod with adaptive integration) usually fail to reach desired precision when the cutoff points are large in absolute value, even if the number of abscissae is high.[8] Romberg integration routines are reasonably accurate, but accuracy comes at a cost of very long time spans to compute integrals with desired precision.

6.4. P.D.F.

Numerical integration of $e^{-i(\zeta_1 s_T + \zeta_2 v_T)} \cdot \Psi(\zeta_1, \zeta_2)$ to recover the p.d.f. of $(s_T, v_T)'$ under $\hat{\mathbb{P}}$ is impractical. Accurate evaluation of the integral for *just one* point $(s_T, v_T)'$ takes a long time, and such evaluations must be done (iteratively) for all points of a reasonably fine grid of $(s_T, v_T)'$ to implement recursion (12).

Contrastingly, FFT allows to recover $f(s_T, v_T)$ on the whole grid quickly in *one* step. While being economical in terms of computational time, FFT places high demand on computer memory.

It is desirable to have relatively large (in absolute value) cutoff points and small step sizes, Δ_1 and Δ_2, in $(\zeta_1, \zeta_2)'$ grid. Since $\Delta_1 = \frac{b_1 - a_1}{N_1}$, $\Delta_2 = \frac{b_2 - a_2}{N_2}$, the dimensions of the Fourier coefficient matrix, $2N_1$ and N_2, are necessarily large.[9] Worse, FFT algorithms require N_1 and N_2 to be powers of 2. This leaves little leeway in choosing the step sizes Δ_1 and Δ_2, once the cutoffs are set to values beyond which the integrand is negligibly small. Note that the step sizes in $(s_T, v_T)'$ grid are $\frac{2\pi}{b_1 - a_1}$ and $\frac{2\pi}{b_2 - a_2}$, respectively; bounds of this grid are: $s_T^{\min} = v_T^{\min} = 0$, and $s_T^{\max} = \frac{2\pi(N_1 - 1)}{b_1 - a_1}, v_T^{\max} = \frac{2\pi(N_2 - 1)}{b_2 - a_2}$.

Picking a finer grid of $(\zeta_1, \zeta_2)'$ implies at least doubling the size of the Fourier coefficient matrix. Modern 32-bit computers theoretically cannot address arrays with more than $2^{32} - 1$ elements. In practice, Fortran

[8] A standard recommendation for oscillatory functions is to set this number to the maximum allowed by the algorithm.

[9] The Fourier coefficients, Ψ_{j_1, j_2}, are complex numbers. The bivariate discrete complex FFT routine I use to program the transformation stores an $N_1 \times N_2$ array of complex numbers as a $2N_1 \times N_2$ array of real numbers: real parts are in odd-numbered rows, imaginary parts are in even-numbered rows.

compilers presently fail to allocate arrays that require memory in excess of 2GB. It can be shown that with double precision number storage, the largest Fourier coefficient matrix one can use is of size $2N_1 \times N_2$: $N_1 \cdot N_2 = 2^{26}$.

FFT routines were compiled and run on the Birch Linux cluster at UVA. Each node of the cluster is a double-Pentium IV 2.4GHz system with 2GB RAM. FFT program requires memory storage of slightly above 1GB, therefore, there is no need to use virtual memory and the transformation itself takes approximately 1 minute. Extra 2 minutes are needed to form a 2^{26}-element Fourier coefficient matrix and to write out the results: one round of FFT takes slightly more than 3 minutes in total. For comparison, one round of FFT takes 6–8 minutes on the Aspen Linux cluster and over 25 minutes on the Unixlab cluster.

Each $(2k_1 + 1, k_2 + 1)$-indexed element of the inverse Fourier coefficient matrix[10] (normalized and smoothed by $(2\pi)^{-2} \Delta_1 \Delta_2 W\left(\theta_{s_{T,k_1}}, \theta_{v_{T,k_2}}\right)$) is $f\left(s_{T,k_1}, v_{T,k_2}\right)$, the value of the p.d.f. corresponding to one particular point in $(s_T, v_T)'$ grid.

There is no need to save and use the whole 2^{26}-element inverse coefficient matrix. The p.d.f. converges to 0 long before s_T and v_T attain their maximum grid values and long before s_T goes all the way down to the minimum grid value. For practical purposes of option pricing, it is more than sufficient to extract the part of the matrix that covers: 5-times up and down movement of the underlying's price $S_T = e^{s_T}$ from S_t and 4-times up and down movement of $\sqrt{v_T}$ from $\sqrt{v_t}$. Resulting grids of $(s_T, v_T)'$ on average contain $800 \cdot 120$ elements. This extracted matrix needs to be processed to zero out very small negative values at some levels of s_T far from s_t, which infrequently occur because $(\zeta_1, \zeta_2)'$ grid is bounded and imperfectly fine. Such negative values are always small in absolute value (on average, $< -10^{-6}$) and negligibly small if compared to the values of the p.d.f. around the peak (on order of $+10^2$). I also verified that the imaginary parts of the inverse Fourier coefficients were close to zero (the imaginary part of the integrand function must integrate out to zero by the inversion theorem).

As an illustration, consider Figures 3 and 4. In these Figures, I plot $f\left(s_T, v_T\right)$ with parameters as of June 30 for two different τ's: a "small" one, which corresponds to $\frac{1}{2}$ of the time span (in years) between June 30 and 3$^{\text{rd}}$ Friday of July, and a "large" one, which stands for $\frac{1}{2}$ of the time till 3$^{\text{rd}}$ Friday of December. These p.d.f.'s are, in fact, used in pricing "Bermudan" options expiring in July and December that offer an opportunity of an early exercise half-way to expiration. The p.d.f. is a unimodal function with a peak at $(s_T, v_T) = (s_t, v_t)$. It has a long "tail" in v_T direction. It is also evident that for smaller τ, the function is "concentrated" around the peak; for larger τ, it is more "diffuse". This result has an intuitive explanation: one should be more uncertain about relatively distant future.

[10] Note that $k_1 = 0, ..., N_1 - 1$, $k_2 = 0, ..., N_2 - 1$ and recall that real parts of the coefficients are stored in odd-numbered rows and imaginary parts are in even-numbered rows.

6.5. Pricing OEX Options

Once the joint p.d.f.'s are obtained, it is straightforward to program pricing of OEX options by Richardson extrapolation (13). The two corresponding options are: D_1, the value of a European-style derivative security, and D_2, the value of a "Bermudan" security with one intermittent exercise date at t_1 (half-way to expiration). In turn, D_2 is calculated according to (12), where the integral, $\hat{E}\left[\max\left\{\cdot\right\}\right]$, is approximated in the simplest non-adaptive way on the equispaced and fixed grid of (s, v).

On a Birch Linux cluster node, pricing of one option takes approximately $1\frac{1}{2}$ minutes (including the time spent on reading-in a p.d.f. matrix from a corresponding saved file).

Consider Table 4, where I report signed pricing errors of all traded OEX options that expire in September. Pricing error is the difference between the actual CBOE last sale value and its corresponding predicted value.

Clearly, a perfect fit can never be attained. Still, it is remarkable that with several layers of approximation, pricing errors look more or less reasonable. For September-expiring options, instances when the absolute pricing error is in excess of 2.00 are rare and most predicted option values do not deviate from CBOE last sale values by more than 1.00. The method seems to underprice far out-of-the-money puts and deep-in-the-money calls. Apart from that, it is hard to find a consistent trend in pricing errors across strikes. Out-of-the-sample pricing errors (July 9^{th}) are not particularly different from their in-the-sample counterparts (June 30^{th} through July 8^{th}).

To assess how pricing accuracy varies with the time to expiration, examine Table 5, where I report root MSE's for almost all traded OEX options that mature in 2004, by trading day and expiration month. There is no clear indication that pricing errors decrease on average for options with longer maturities. On the contrary, with two exceptions (July 7^{th} and July 8^{th}), predictions for December-expiring derivatives are relatively coarse. Probably, one should expect the precision of the linear Richarson extrapolation to degrade as the time to maturity of an American-style security increases. Overall, $\sqrt{MSE}$'s are of sensible magnitudes.

7. Conclusion

In this paper, I consider valuation of the American-style derivative securities when the price of an underlying asset follows the dynamics of the Heston model. As a feasible alternative to popular FD techniques, I employ a version of the Geske–Johnson scheme with linear Richardson extrapolation. The method requires knowledge of the joint p.d.f. of the future log-price, s, and squared volatility, v, to price the "Bermudan" option with one intermittent early exercise date half-way to expiration.

The joint p.d.f. can be recovered by inverting the corresponding joint ch.f. Unfortunately, the pronouncedly oscillatory nature of the integrand function makes accurate bivariate numerical integration pro-

hibitively costly in terms of computational time. Instead, I suggest that the joint ch.f. be inverted using the FFT algorithm, and a kernel-smoothing scheme be used to mitigate the problem of precision loss. A several orders of magnitude reduction in the time of the inversion, however, comes at a cost of dramatically increasing demand for computer memory. A fast FT algorithm that would produce satisfactory results will not be "fast" on an average modern PC. It is more efficient to obtain p.d.f.'s separately on a server with large RAM and use the results to price options on a, possibly, less powerful machine.

To assess the numerical accuracy of the method, I apply it to price all S&P U.S. 100 index options ("OEX") that expire in 2004 and are quoted at market's closing on 7 consecutive trading days, starting with June 30^{th}. Parameters are calibrated on "XEO" options and July 9^{th} data is set aside for out-of-the-sample predictions. Pricing errors overall look reasonable and out-of-the-sample errors are not fundamentally different from the in-sample ones. Still, the method tends to underprice out-of-the-money puts and predictions are relatively coarse for options with "long" ($\frac{1}{2}$ of the year) time to expiration. The latter result is, most likely, an artifact of the degrading accuracy of the *linear* Richardson extrapolation.

References

[1] Barone-Adesi, Giovanni, Robert E. Whaley. 1987. "Efficient Analytic Approximation of American Option Values," *Journal of Finance*, Vol. 42, No. 2 (June), 301–320

[2] Bates, David S. 1996. "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies,* Vol. 9, No.1 (Spring), 69–108

[3] Billingsley, Patrick. 1995. *Probability and Measure*. 3rd ed. John Wiley & Sons Ltd.

[4] Broadie, Mark, Jerome Detemple. 1996. "American Option Valuation: New Bounds, Approximations, and a Comparison of Existing Methods," *Review of Financial Studies,* Vol. 9, No. 4 (Winter), 1211–1250

[5] Chernov, Mikhail, Eric Ghysels. 2000. "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*. Vol. 56, 407–458

[6] Chicago Board of Options Exchange. 2004a. *European-style S&P 100 Index Options: Product Specifications*. http://www.cboe.com/OptProd/indexopts/xeo_spec.asp

[7] Chicago Board of Options Exchange. 2004b. *OEX S&P 100 Index Options: Product Specifications*. http://www.cboe.com/OptProd/indexopts/oex_spec.asp

[8] Cox, John C., Jonathan E. Ingersoll, Stephen A. Ross. 1985. "A Theory of the Term Structure of Interest Rates," *Econometrica*, Vol. 53, No. 2 (March), 385–408

[9] Duffie Darrell, Jun Pan, and Kenneth Singleton. 2000. "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*. Vol. 68, No. 6 (November), 1343–1376

[10] Epps, Thomas W. 2004a. *Econ 834: Derivative Securities*. UVA

[11] Epps, Thomas W. 2004b. Option Pricing Under Stochastic Volatility with Jumps. UVA, mimeo

[12] Feller, William. 1951. "Two Singular Diffusion Problems," *Annals of Mathematics*, Vol. 54, No. 1 (July), 173–182

[13] Geske R., H. Johnson. 1984. "The American Put Option Valued Analytically," *Journal of Finance*, Vol. 39, 1511–1524

[14] Heston, Steven L. 1993. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, Vol. 6, No. 2, 327–343

[15] Hirsa, Ali, Dilip B. Madan. 2003. "Pricing American Options Under Variance Gamma," *Journal of Computational Finance,* Vol. 7, No. 2 (Winter)

[16] Longstaff, Francis A., Eduardo S. Schwartz. 2001. "Valuing American Options by Simulation: A Simple Least-Squares Approach," *Review of Financial Studies,* Vol. 14, No. 1 (Spring), 113–147

[17] Press, William H., Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. 1992. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. 2nd ed. Vol. 1. Cambridge University Press

[18] Rudin, Walter. 1991. *Functional Analysis*. 2nd ed. McGraw-Hill Book Co.

[19] Rudin, Walter. 1987. *Real and Complex Analysis*. 3rd ed. McGraw-Hill Book Co.

[20] Shephard, Neil G. 1991a. "From Characteristic Function to Distribution Function: A Simple Framework for the Theory," *Econometric Theory,* 7, 519–529

[21] Shephard, Neil G. 1991b. "Numerical Integration Rules for Multivariate Inversions," *Journal of Statistical Computation and Simulation,* 39, 37–46

[22] Standard & Poor's. 2004. *Press Release*. July 1st. http://www2.standardandpoors.com/spf/pdf/index/tr.pdf

[23] Standard & Poor's. 2003. *S&P 100*. December 31st. http://www.standardandpoors.com/indices

[24] *The Wall Street Journal*. Tuesday, July 6, 2004. Vol. CCXLIV No. 3, p. C12

[25] Wilmott, Paul. 2000. *Paul Wilmott on Quantitative Finance*. 2-volume set. John Wiley & Sons Ltd.

[26] Winkler, Gunter. 2001. *Analytische und numerische Untersuchung des Modells von Heston zur Optionspreisberechnung*. Chemnitz, Fakultät für Mathematik

8. Appendix

8.A. Kernel Function

Let the kernel function be:

$$K(x,y) = \begin{cases} \frac{(1-|x|)^2(1-|y|)^2}{x^2y^2+x^2(1-|y|)^2+(1-|x|)^2y^2+(1-|x|)^2(1-|y|)^2}, & \text{if both } |x| \leq 1 \text{ and } |y| \leq 1 \ . \\ 0, & \text{elsewhere.} \end{cases}$$

$K(x,y)$ has the following desirable properties. First, it is symmetric and nonnegative.

Second, it reaches a max of 1 if both $x = 0$ and $y = 0$. Since $x = \frac{\zeta_1 - \zeta_{j_1}}{\Delta_1}$, $y = \frac{\zeta_2 - \zeta_{j_2}}{\Delta_2}$ this implies a weight of 1 in case (ζ_1, ζ_2) coincides with one interpolation point.

Third, since the grid of (ζ_1, ζ_2) is equispaced, $K(x,y)$ will assign positive weights to at most 4 interpolation points closest to (ζ_1, ζ_2).

Fourth, it can be verified that:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(x,y)\,dxdy = \int_{-1}^{1}\int_{-1}^{1} K(x,y)\,dxdy = 1,$$

which establishes that $K(x,y)$ *is* a valid kernel function. See Figure 5 for a graphical representation.

8.B. Weighting Function

By construction, $W(\theta_{s_T}, \theta_{v_T}) = \int_{a_2}^{b_2}\int_{a_1}^{b_1} e^{-i\left(\theta_{s_T}x+\theta_{v_T}y\right)}K(x,y)\,dxdy$. With the discussed choice of the kernel function, this specializes as:

$$W(\theta_{s_T}, \theta_{v_T}) = \int_{-1}^{1}\int_{-1}^{1} e^{-i\left(\theta_{s_T}x+\theta_{v_T}y\right)}K(x,y)\,dxdy =$$

$$= \int_{-1}^{1}\int_{-1}^{1} \cos(\theta_{s_T}x + \theta_{v_T}y)K(x,y)\,dxdy - i\int_{-1}^{1}\int_{-1}^{1} \sin(\theta_{s_T}x + \theta_{v_T}y)K(x,y)\,dxdy =$$

$$= \int_{-1}^{1}\int_{-1}^{1} \cos(\theta_{s_T}x + \theta_{v_T}y)K(x,y)\,dxdy.$$

$W(\theta_{s_T}, \theta_{v_T})$ may be evaluated in a number of ways. A direct approach is to approximate it by numerical integration.

A better way is to exploit the properties of the cosine function. Taylor expanding $\cos(\theta_{s_T}x + \theta_{v_T}y)$ and

interchanging summation and integration:

$$W\left(\theta_{s_T},\theta_{v_T}\right)=\sum_{n=0}^{\infty}\int_{-1}^{1}\int_{-1}^{1}\frac{(-1)^n\left[\theta_{s_T}x+\theta_{v_T}y\right]^{2n}}{(2n)!}K\left(x,y\right)dxdy.$$

Next, in the grid, $\theta_{s_{T,k_1}}=\Delta_1 s_{T,k_1}=\frac{2\pi k_1}{N_1}$, $\theta_{v_{T,k_2}}=\Delta_2 v_{T,k_2}=\frac{2\pi k_2}{N_2}$, then, on the rectangle of integration:

$$\left|\theta_{s_{T,k_1}}x+\theta_{v_{T,k_2}}y\right|\leq\theta_{s_{T,k_1}}\left|x\right|+\theta_{v_{T,k_2}}\left|y\right|\leq\theta_{s_{T,k_1}}+\theta_{v_{T,k_2}}<4\pi.$$

So, $\left|\int_{-1}^{1}\int_{-1}^{1}\left[\theta_{s_T}x+\theta_{v_T}y\right]^{2n}K\left(x,y\right)dxdy\right|\leq\int_{-1}^{1}\int_{-1}^{1}\left|\theta_{s_T}x+\theta_{v_T}y\right|^{2n}K\left(x,y\right)dxdy<\left(4\pi\right)^{2n}$. $(2n)!$ swamps $\left(4\pi\right)^{2n}$ at a fast rate. Of course, there will be additional attenuation due to the kernel function itself.

Therefore:

$$W\left(\theta_{s_T},\theta_{v_T}\right)\cong 1+\sum_{n=1}^{N}\frac{(-1)^n}{(2n)!}\int_{-1}^{1}\int_{-1}^{1}\left[\theta_{s_T}x+\theta_{v_T}y\right]^{2n}K\left(x,y\right)dxdy,$$

even for a moderate choice of N. Still, truncation at a very low N, may result in a negative value of W.

Closed-form expressions for the integrals are straightforward to derive and are not presented here to save space. Some experimentation reveals that very accurate results for W are obtained for $N=8$. This choice also guarantees strict positivity of W for all $\theta_{s_T},\theta_{v_T}\in[0,2\pi]$.

Table 1: OEX Option Data: Summary

	June 30	July 1	July 2	July 6	July 7	July 8	July 9
included puts, #	57	56	48	56	51	55	56
included calls, #	38	41	39	32	34	35	36
put (July expiring) strike range	440-570	450-580	440-570	450-600	480-600	480-600	460-565
call (July expiring) strike range	500-590	500-595	510-590	520-580	520-580	525-580	525-590

Table 2: XEO Option Data: Summary

	June 30	July 1	July 2	July 6	July 7	July 8	July 9
included options, #	21	12	27	31	23	20	23
inferred by parity, #	10	2	9	12	10	8	9
option (July expiring) strike range	510-580	520-575	520-570	510-585	525-565	520-555	520-565

Table 3: Calibrated Parameters

Parameter	Value	Parameter	Value
$v_{t,01}$	0.011392	α	0.353948
$v_{t,02}$	0.010889	β	9.561292
$v_{t,03}$	0.008932	γ	0.763721
$v_{t,04}$	0.016582	ρ	-0.692404
$v_{t,05}$	0.012725		
$v_{t,06}$	0.015280		

Table 4: Pricing Errors: September Expiring OEX Options

	June 30	July 1	July 2	July 6	July 7	July 8	July 9
S_t	553.87	549.01	547.17	543.33	544.25	540.21	542.63
X, type							
400, put	...	...	...	...	0.116577	...	0.170837
420, put	...	...	0.467125	...	...	...	...
440, put	0.272164	0.365445	0.337645	...	...	...	...
450, put	...	...	0.578615	...	0.476928	0.448424	...
460, put	0.477434	0.636137	...	0.555188	0.561046	0.419271	...
480, put	0.233909	0.738647	...	...	0.561046	...	0.926725
490, put	...	...	0.795218	...	...	...	...
500, put	...	0.852520	0.544044	0.434873	0.339729	0.603371	0.735753
510, put	0.694260	0.274149	...	...	0.101801	0.229474	0.419345
520, put	-0.503378	-0.343453	0.302446	-0.048642	-0.521728	-0.126598	-0.174747
530, put	-0.591617	...	...	0.175511	-0.740207	-0.090279	0.525561
540, put	-2.429237	-0.154515	-1.027132	-1.014578	-1.725362	...	-0.131715
550, put	-2.538264	0.520023	-0.366053	-0.103850	-0.295665	-1.805092	...
560, put	-1.522215	...	...	-1.752594	...	0.002115	-1.306124
580, put	...	-1.011618	...	0.074978	...	...	...
500, call	...	2.438584	...	...	...	...	...
520, call	...	...	...	...	...	3.296368	...
530, call	...	...	...	1.810402	3.058499	...	...
540, call	1.619276	...	1.490850	1.081886	2.824853	0.942793	1.105416
550, call	1.902790	1.470218	0.824900	...	...	0.172171	0.736377
560, call	0.590900	-0.086137	0.518619	0.387300	1.983925	-0.040926	0.415623
570, call	0.673686	-0.082919	-0.062988	0.107370	0.934013	0.116647	0.351140
580, call	0.035720	-0.205257	0.191112	...	0.720421	0.001702	0.070248
590, call	0.028256	-0.161321	-0.078389	...	...	0.109251	0.073845
600, call	-0.202037	-0.190903	...	...	...	...	-0.075595

Table 5: RMSE's: OEX Options, by Expiration Month and Trading Day

	June 30	July 1	July 2	July 6	July 7	July 8	July 9
July	0.563594	0.729401	0.555327	0.640847	0.529694	1.097807	0.271871
August	1.316759	0.408102	0.490070	0.367468	0.686766	0.419591	0.311062
September	1.203947	0.849578	0.659790	0.878324	1.394570	1.029171	0.617068
October	0.442581	1.059348	0.551232	0.387529	1.425867	0.633056	0.563490
December	2.024635	2.129737	1.723453	1.578420	1.229747	0.984973	1.707936

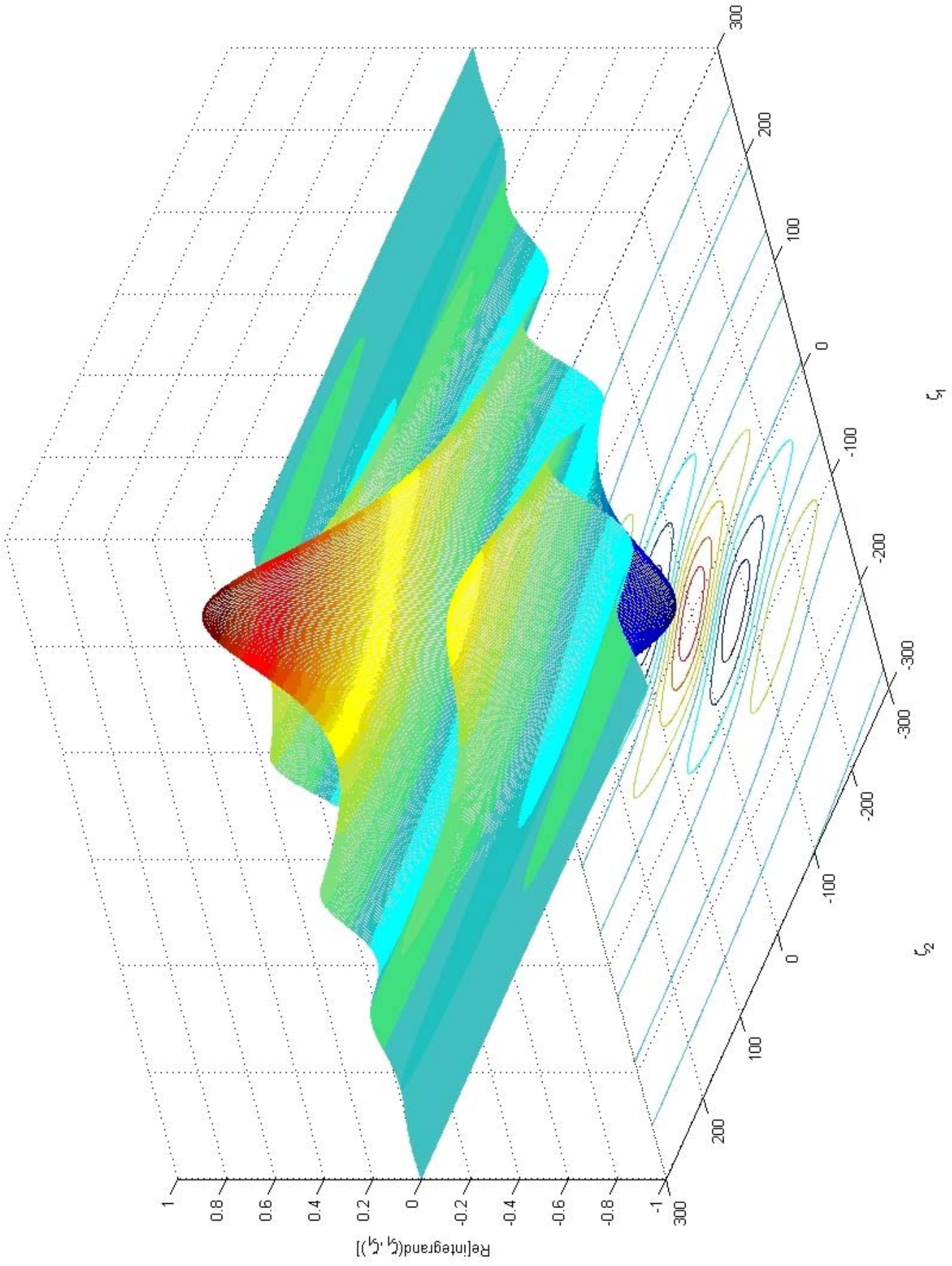

Figure 1: Re [Integrand]: "small" τ

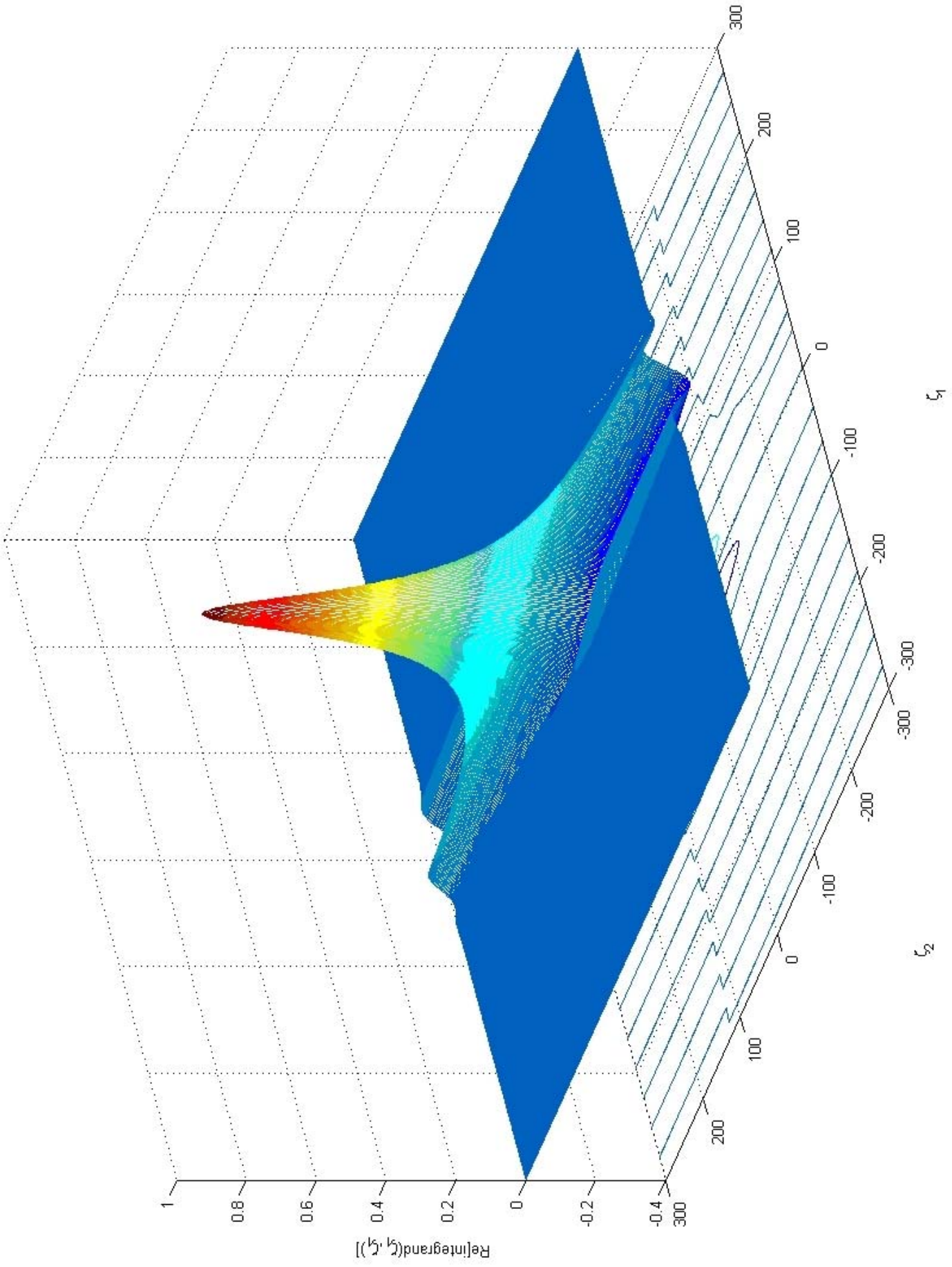

Figure 2: Re [Integrand]: "large" τ

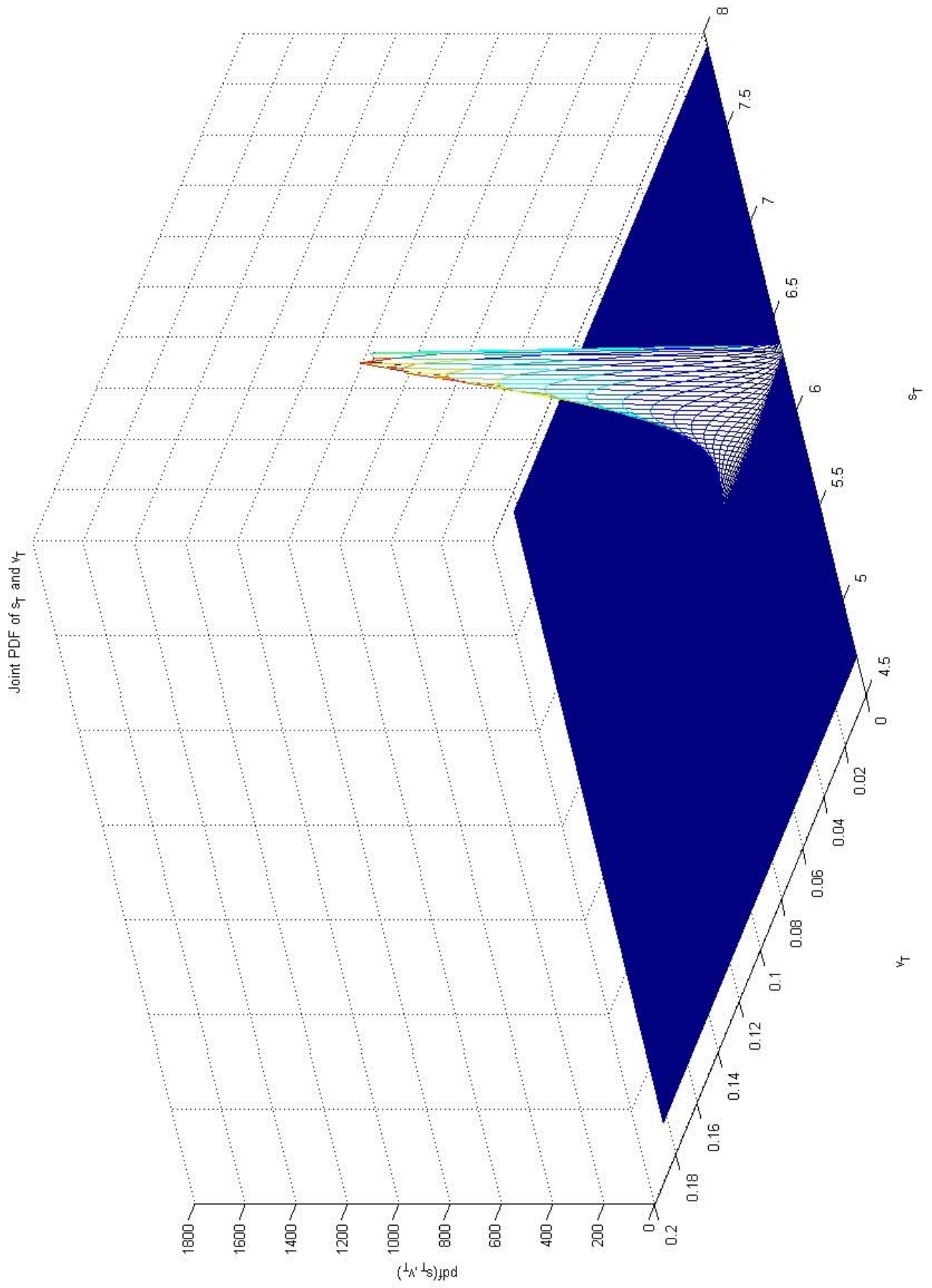

Figure 3: P.D.F. as of June 30: "small" τ

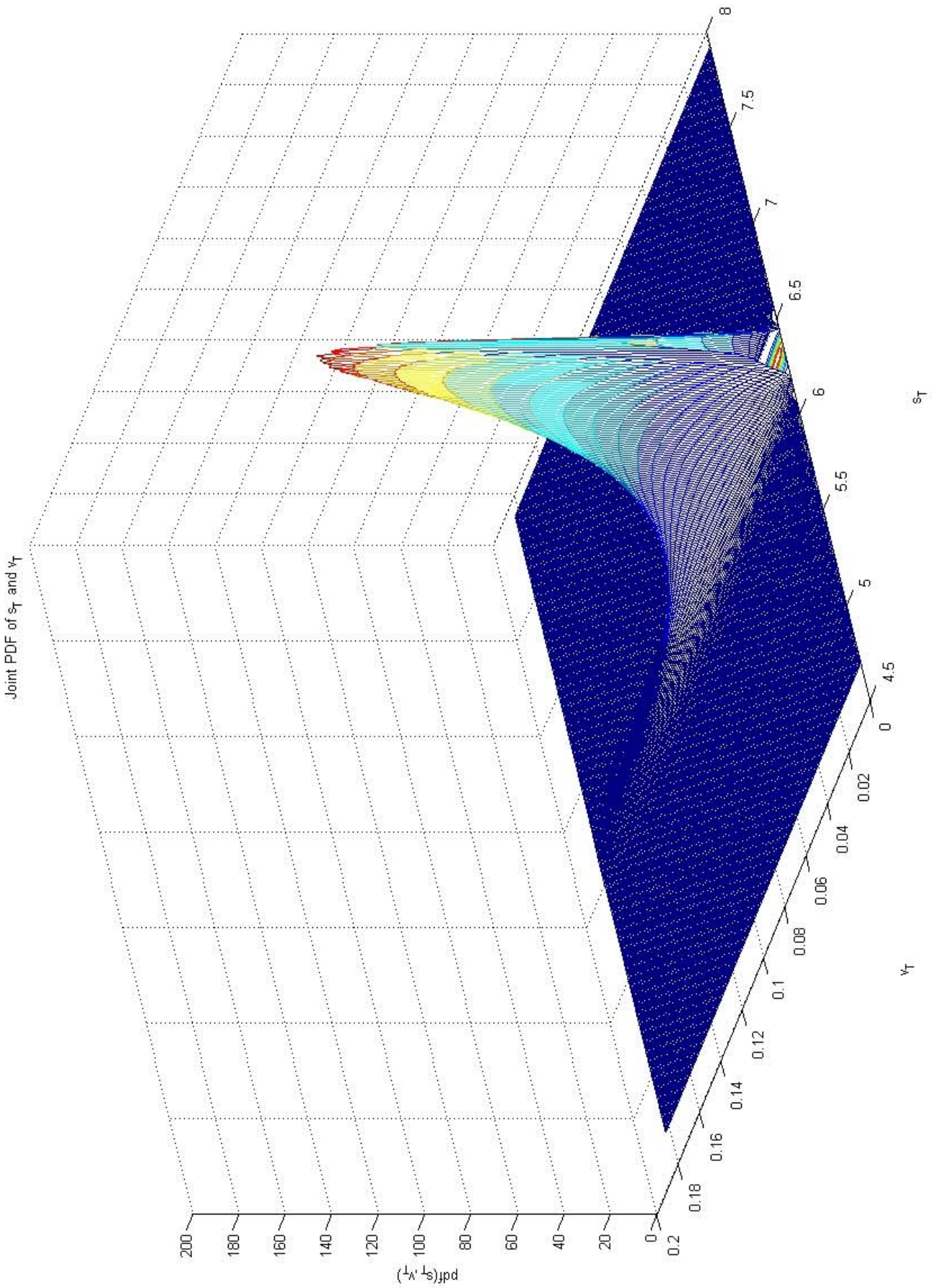

Figure 4: P.D.F. as of June 30: "large" τ

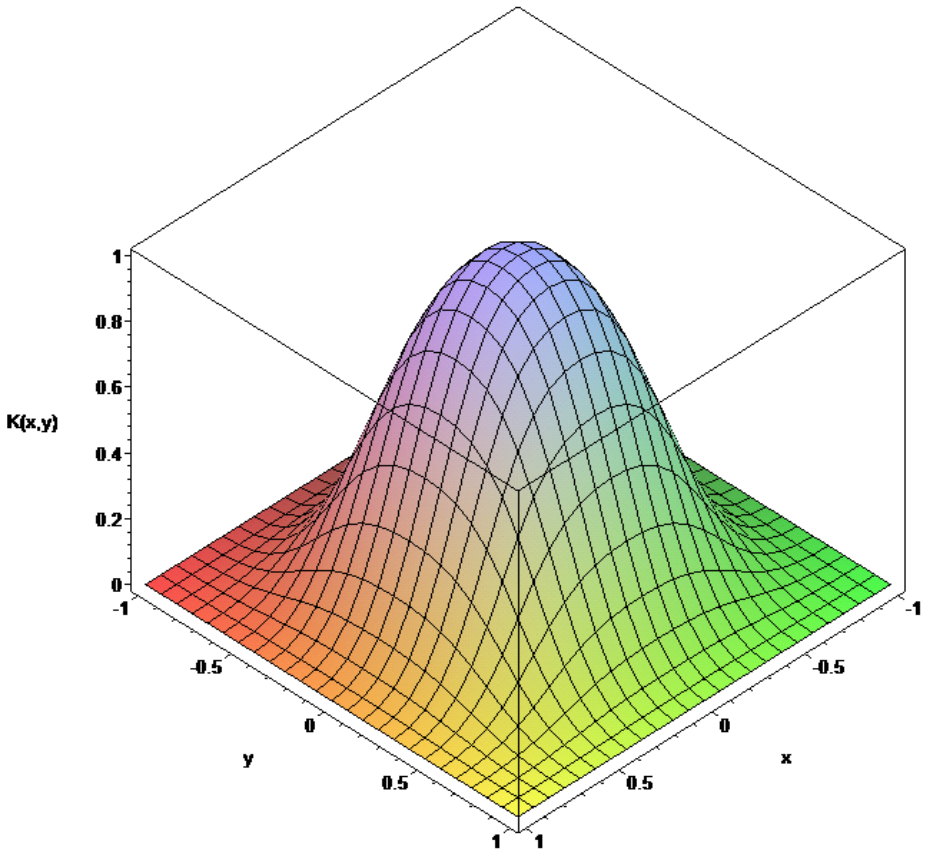

Figure 5: Kernel Function