Asset Prices and Conditional Moments in Multifactor Non-Affine Models

Robert L. Kimmel
Fisher College of Business
The Ohio State University*

This Version: 4 March 2008

Abstract

In recent research, methods are developed to approximate conditional moments and contingent claims prices in a large class of non-affine continuous-time models. These methods employ series approximations in the time variable, and use change of variable techniques to improve convergence properties. However, at present, these methods have only been developed for scalar diffusion problems. We extend this earlier research to allow accurate approximation of contingent claim prices in a much broader class of multifactor models. Our technique proceeds through several stages. First, a large class of non-affine multifactor pricing problems is shown to be equivalent, after change of variables, to a class of conditional moment problems in a multifactor affine diffusion setting. We then develop a method for approximation of the solution to the transformed problem, such that the approximations often converge uniformly for arbitrary time horizons. Our technique is to embed the path of "true" time in a higher-dimensional space of "artificial" time. The bond (or other contingent claim) price is meaningful only as a function of "true" time; however, extending this function to "artificial" time greatly simplifies the problem of constructing series approximations with good convergence properties. The true bond price function is then the restriction of the extended price function to a curve (representing "true" time) through the higher-dimensional space of "artificial" time. We show through examples, in which the bond price function is known in closed-form, that the approximations are easy to derive and converge very rapidly, in many cases, for arbitrarily long time horizons. We then develop a method for construction of a large class of non-affine multifactor term structure models to which our technique can be applied, resulting in accurate approximations of bond prices and yields (regardless of maturity) even when the true bond price function is not known in closed-form.

*Columbus, OH 43210-1144. Phone: +1 614 292 1875. E-mail: kimmel.42@osu.edu.

1. Introduction

In recent research, Kimmel (2008c) and Kimmel (2008b) develop methods for approximation of contingent claims prices (such as bond prices in a term structure model) and conditional moments. Through the use of non-affine transformations of the time variable, the convergence properties of the approximations can often be improved dramatically relative to approximations attained by a naïve method; examples in Kimmel (2008b) and Kimmel (2008a) show that even when approximations by the naïve method diverge for relatively short time horizons, it is nonetheless often possible to achieve rapid and *uniform* convergence for arbitrarily long time horizons through non-affine time transformation methods. Further applications of this technique appear in Jarrow, Li, Liu, and Wu (2006).

Despite the desirable convergence properties that can result from appropriate use of time transformation methods, a significant limitation of these techniques is that they currently apply only to scalar diffusion problems. It is possible in some cases to break a multiple diffusion problem into several independent scalar problems. For example, in an N-factor term structure model, if state variable processes are independent (under a risk-neutral probability measure), and the state variables enter into the interest rate function in an additively separable way, then the price of a zero-coupon bond can be formulated as the products of N zero-coupon bonds found in N different scalar term structure models. However, we may be interested in more general bond pricing (and other asset pricing) models that do not satisfy these constraints. As we show, in such models, it is often possible to establish convergence of approximations, using the method of time transformations, for arbitrarily long time horizons; however, it is rarely possible, using straightforward extensions of the scalar methods, to show that such convergence is *uniform* in time horizon. Examples in Kimmel (2008b) show the importance of uniform convergence in term structure settings; using models for which bond prices are known in closed-form, he shows that even when approximations converge for all time horizons, if the convergence is not uniform, it can be so slow as to be impossible to apply reasonably in practice.

We therefore develop techniques to allow the extension of time transformation methods to more general multifactor problems. The key to our method is the introduction of multiple, "artificial" time dimensions. Practical problems have solutions that are only meaningful in a single time dimension, and only for positive values; we would like to know, for example, the conditional moment of a process three months into the future, or the price of a zero-coupon bond that matures in five years. Nonetheless, the structure of many multifactor asset pricing or conditional moment problems is such that the solutions are most easily approximated in a setting with multiple time dimensions, even though the extra time dimensions have no physical meaning. The "true" conditional moment or asset price, as a function of ordinary one-dimensional time, is found simply by restricting the solution in multiple-dimensional time to a curve that represents the path of "true" time through the higher-dimensional space. Although perhaps unintuitive, the introduction of multiple time dimensions is very helpful in the development of approximations to solutions of practical asset pricing or conditional moment problems that arise in the familiar setting of one-dimensional "true" time.

The rest of this paper proceeds as follows. First, we review the techniques of Kimmel (2008c) and Kimmel (2008b), as they apply to scalar diffusion problems, and illustrate the difficulties of extending the uniform convergence results to a general multi-dimensional setting. The next section introduces multiple-dimensional "artificial" time, and shows how a large class of asset pricing and conditional moment problems in a multifactor setting can be embedded in a problem formulated with multiple, artificial time dimensions. The solution to the problem in multi-dimensional time can then be found by approximation methods, using the non-affine transformations of the (multiple) time variables, to improve the convergence properties of the approximations. Since true time is simply a curve through the higher-dimensional space of artificial time, the solution to the real-world problem is given by the restriction of the solution to the artificial problem, in multiple-dimensional time, to this curve. The next section explores several examples, in which bond prices are known in closed-form, to illustrate our technique and test its accuracy. We explore a particular family of the affine term structure models of Duffie and Kan (1996) (as further studied and extended by, for example, Dai and Singleton (2000), Duffee (2002), Dai and Singleton (2002) and Cheridito, Filipović, and Kimmel (2007)) and the non-affine models of Ahn, Dittmar, and Gallant (2002). In both cases, we find our technique is easy to apply, and, for reasonable parameter values, yields approximate bond prices and yields that are very accurate for arbitrarily long time horizons. We also examine term structure models in which bond prices are *not* known in closed-form, and show how to apply our technique to find approximations. Numeric analysis suggests that our technique is very accurate for this class of models as well, but the use of closed-form approximations allows much more rapid computation of bond prices, and use of non-affine models that would otherwise be too difficult to estimate. Finally, we conclude, and summarize several possible avenues for future research.

2. Series Solutions with Time Transformations

In this section, we discuss results from the extant literature on series solutions to scalar diffusion problems, and the problems with naïve extension of these methods to multiple diffusion problems.

2.1. Scalar Results

Motivated by the probabilistic problem of finding conditional moments or contingent claims prices, Kimmel (2008c) constructs approximate solutions to second order parabolic partial differential equations. The scalar version of the PDE is:

$$\frac{\partial f}{\partial \Delta}(\Delta, x) = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\Delta, x) - r(x) f(\Delta, x) \tag{2.1}$$

$$f(0, x) = g(x) \tag{2.2}$$

This partial differential equation problem, subject to technical restrictions,[1] is equivalent to a probabilistic problem in which x is the current value of a Brownian motion, and Δ is a time horizon (for conditional moment problems) or maturity (for asset pricing problems).[2] If $r(x) = 0$, then the solution to (2.1) and (2.2) is the expected value Δ units of time into the future of $g(x)$. If $r(x) \neq 0$, but instead specifies the instantaneous

[1] See Levendorskii (2004a) and Levendorskii (2004b) for a recent discussion of the equivalence between the probabilistic and partial differential equation problem.

[2] In either case, note that Δ flows in the reverse direction from calendar time, that is, Δ begins with a positive value, and then decreases with the passage of time to zero.

interest rate as a function of the state variable process, then the solution to (2.1) and (2.2) specifies the price of a contingent claim with final payoff Δ units of time into the future given by $g(x)$ (provided the state variable is a Brownian motion under a risk-neutral probability measure).[3]

The PDE problem specified by (2.1) and (2.2) may appear to be extremely restrictive, since the implied state variable process is always a Brownian motion. However, much more general problems can be expressed in this form after changes of variables. Colton (1979) shows that *any* second order parabolic differential equation with one spatial state variable and time-independent coefficients can be converted to this form by changes of dependent and independent variable;[4] Aït-Sahalia (2002) employs the same change of independent variable (but no change of dependent variable) to simplify the problem of approximating the conditional density function of the state variable process. If a PDE problem is transformed to (2.1) and (2.2) by change of variables, and a solution to the transformed problem can be found, then a solution to the original problem can be found by reversing the change of variables. Therefore, despite its apparent restrictiveness, the problem (2.1) and (2.2) effectively encompasses *all* conditional moment and contingent claim pricing problems motivated by a scalar diffusion.

Kimmel (2008c) refers to (2.1) and (2.2) as the "canonical" form PDE, and characterizes explicitly the final conditions, $g(x)$, such that the PDE solution, $f(\Delta, x)$, is analytic in some neighborhood of $\Delta = 0$ (and therefore can be approximated by a convergent power series) when the $r(x)$ function takes one of two specific forms:[5]

$$r(x) = bx^2 + cx + d \tag{2.3}$$

$$r(x) = \frac{a}{x^2} + d + bx^2 \tag{2.4}$$

Although the final condition $g(x)$ is only meaningful for real x in typical applications, Kimmel (2008c) shows that existence of a convergent series solution (in the time variable) can be established if $g(x)$ has certain properties for all *complex* values of x. Specifically, if $g(x)$ is everywhere analytic and satisfies an exponential quadratic growth condition in *all* directions in the complex plane, then a power series approximation to the solution of (2.1), when $r(x)$ is specified by (2.3), converges within a circle $|\Delta| < s$ for some $s > 0$. The radius of the circle of convergence depends on the rate of growth of $g(x)$ (in all directions on the complex plane) and on the value of b, but does not depend on the values of c or d. For the case where $r(x)$ is given by (2.4), analogous, but somewhat more complicated results are derived, in which the final conditions $g(x)$ yielding analytic solutions $f(\Delta, x)$ are expressed in terms of two everywhere analytic and even functions with exponential quadratic growth conditions in all directions in the complex plane. As in the previous case, the

[3] Solutions to (2.1) and (2.2) are in general not unique unless a growth condition on x is also imposed. However, our focus is on solutions that are analytic in the time variable, and these solutions are unique, and also satisfy the growth condition. Throughout, we therefore do not state the growth condition explicitly.

[4] Colton (1979) studies problems of heat diffusion rather than probabilistic problems. However, his change of variable techniques can still be applied. The probabilistic problem has an equivalent partial differential equation representation, in which there is a first spatial derivative term, and in which the coefficient of the second spatial derivative may not be constant; this PDE can be transformed to one of the form (2.1) and (2.2) by the same methods used in the heat diffusion setting.

[5] The notation here is slightly different than in Kimmel (2008c).

radius of the circle of convergence depends on the rate of growth of the final condition and on the value of the b parameter, but does not depend on the values of a or d. For either (2.3) or (2.4), if the growth rate of $g(x)$ is sufficiently constrained in *all* directions in the complex plane, then a power series approximation converges for all values of Δ. Such convergence is uniform on compact sets of Δ, but is in general not uniform for all Δ.

Kimmel (2008b) shows that it possible in many cases to improve the convergence properties of series approximations, sometimes achieving *uniform* convergence for arbitrarily large time horizons. Specifically, if the value of the b parameter in either (2.3) or (2.4) is negative,[6] then it is possible to construct series approximations that converge uniformly for all $\Delta \in [0, +\infty)$ for many specifications of $g(x)$, by changing the time, state, and dependent variables before deriving the power series approximation. Kimmel (2008b) uses the term structure model of Cox, Ingersoll, and Ross (1985) to illustrate and evaluate the technique, since bond prices in this model are known in closed-form, and finds that series approximations to bond prices using this method are extremely accurate for very long time horizons; he considers bond yields for maturities of up to 10,000 years, and initial interest rates ranging from 0.06% to 600%, and finds the series approximations, even with only a few terms, are very accurate over the entire range of initial interest rates and bond maturities. By contrast, convergence properties of a series approximation to $f(\Delta, x)$ directly in Δ (i. e., without applying the change of variables techniques in Kimmel (2008b)) are accurate for short maturities, but their performance deteriorates dramatically for maturities much beyond a few years, even with a very large number of terms. Kimmel (2008a) considers additional examples, including some for which the solutions are not known in closed-form, and finds the approximation method to be very accurate in these cases as well. Jarrow, Li, Liu, and Wu (2006) use this method to price callable corporate bonds.

For the moment, we focus on the scalar case. We find it convenient to use a different canonical form PDE than that used in Kimmel (2008c):

$$\frac{\partial f}{\partial \Delta}(\Delta, x) = \mu(x)\frac{\partial f}{\partial x}(\Delta, x) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\Delta, x) \tag{2.5}$$

$$f(0, x) = g(x) \tag{2.6}$$

If $\mu(x)$ is a linear function of x:

$$\mu(x) = \kappa(\theta - x) \tag{2.7}$$

then (2.5) and (2.6) are effectively equivalent to (2.1) and (2.2) where $r(x)$ is specified by (2.3). Specifically, if $f(\Delta, x)$ is a solution to (2.5) and (2.6) with $\mu(x)$ given by (2.7), then:

$$f^*(\Delta, x) = e^{-\frac{\kappa}{2}(x-\theta)^2} f(\Delta, x) \tag{2.8}$$

is a solution to (2.1) and (2.2) with $g(x)$ replaced by $g^*(x)$:

$$g^*(x) = e^{-\frac{\kappa}{2}(x-\theta)^2} g(x) \tag{2.9}$$

As described in Kimmel (2008c), it is possible to calculate, through a simple recursive procedure, a power series

[6] It is only necessary that b have a negative real part, but since b is real in typical applications, we focus on this case.

approximation to the solution to the general PDE (2.1) with final condition (2.2); with slight modification, this approach carries over to the general PDE (2.5), with corresponding final condition (2.6). If the solution to (2.5) satisfies certain smoothness conditions (which can be established by showing that the final condition (2.6) satisfies certain smoothness and growth conditions), then its power series converges in some neighborhood of $\Delta = 0$. If $g(x)$ satisfies stronger growth conditions, then the power series converges for all values of Δ.

However, this convergence will not be uniform in most cases. Kimmel (2008b) establishes uniform convergence of power series representations of the solution to (2.1) and (2.2) by changes of time, state, and dependent variable. Working with (2.5) and (2.6) instead, no change of dependent variable is needed. The two other changes, of independent and time variables, are still needed however. Kimmel (2008b) implements these changes of variables simultaneously; but we find it advantageous to consider the implications of the two changes separately. The change of independent variable eliminates the first spatial derivative term from the right-hand side of (2.5), at the price of making the coefficient on the second spatial derivative term a function of time. The change of time variable then eliminates the time-dependency introduced by the change of independent variable.

We first use the change of independent variable:[7]

$$y(\Delta, x) = \theta + (x - \theta) e^{-\kappa\Delta} \tag{2.10}$$

We then express the solution in terms of Δ and y rather than Δ and x:

$$f(\Delta, x) = \phi(\Delta, y(\Delta, x)) \tag{2.11}$$

and find that $\phi(\Delta, y)$ satisfies:

$$\frac{\partial \phi}{\partial \Delta}(\Delta, y) = \frac{e^{-2\kappa\Delta}}{2} \frac{\partial^2 \phi}{\partial y^2}(\Delta, y) \tag{2.12}$$

$$\phi(0, y) = g(y) \tag{2.13}$$

The change of the state variable thus eliminates the first spatial derivative term from the PDE, at the price of complicating the second spatial derivative term; it is now difficult or impossible to calculate power series coefficients by a recursive relation. However, the time transformation:

$$\tau(\Delta) = \begin{bmatrix} \dfrac{1 - e^{-2\kappa\Delta}}{2\kappa} & \text{if} & \kappa \neq 0 \\ \Delta & \text{if} & \kappa = 0 \end{bmatrix} \tag{2.14}$$

sets things right again. Expressing the solution now as:

$$f(\Delta, x) = \phi(\Delta, y(\Delta, x)) = h(\tau(\Delta), y(\Delta, x)) \tag{2.15}$$

[7] The transformations of the state and time variables used here are scaled versions of the transformations used in Kimmel (2008b). The reason for using this modified approach is to allow the case with $\kappa = 0$ to be handled in the same way as $\kappa \neq 0$. Kimmel (2008b) considers only the scalar case, in which there is no need to deal with the case of $\kappa = 0$. However, when we consider the multivariate case in later sections, it will become advantageous to be able to treat $\kappa = 0$ as part of the general case.

we find that $h(\tau, y)$ satisfies the PDE:

$$\frac{\partial h}{\partial \tau}(\tau, y) = \frac{1}{2}\frac{\partial^2 h}{\partial y^2}(\tau, y) \tag{2.16}$$

$$h(0, y) = g(y) \tag{2.17}$$

Thus, in the transformed variables τ and y, we have the partial differential equation, with final condition, that is satisfied by conditional moments of a Brownian motion. A power series approximation to $h(\tau, y)$, rather than $f(\Delta, x)$ directly, can now be found, by the same recursive procedure used to find solutions to (2.18) and (2.19). Analyticity of $h(\tau, y)$ in some neighborhood of $\tau = 0$ (which follows immediately from analyticity of $f(\Delta, x)$ in some neighborhood of $\Delta = 0$) establishes convergence of the power series approximation for at least some values of τ. There is, though, a critical practical difference between the two problems. For (2.5) and (2.6), with the linear drift specification of (2.7), convergence of a power series approximation for all values of Δ does not, in general, establish uniform convergence on $\Delta \in [0, +\infty)$. Such uniform convergence may be important, for example, in bond pricing applications, where we might consider bonds with maturities of many years. However, for (2.16) and (2.17), provided $\kappa > 0$, the interval $\Delta \in [0, +\infty)$ maps to $\tau \in [0, 1/(2\kappa))$. Consequently, if the power series approximation to $h(\tau, y)$ converges for all $|\tau| < s$ for any $s > 1/(2\kappa)$, it converges uniformly on $|\tau| \leq 1/(2\kappa)$, and this last set includes $\Delta \in [0, +\infty)$. The presence of the κ term in the drift makes characterization of the region of analyticity of the solution more complicated. However, if κ is positive, is presence can nonetheless be exploited to establish uniform convergence of approximations to solutions to the original PDE problem, by changing the time and state variables, and approximating the solution to the transformed problem.

2.2. Naïve Multivariate Extensions

The series approximation technique with time transformation, as shown by examples in Kimmel (2008b) and Kimmel (2008a), performs well in a scalar setting. However, it does not extend readily to problems motivated by multiple diffusions. To see the problems that arise, we consider a naïve extension of the methods of Kimmel (2008c) and Kimmel (2008b) to this case, and show that in general, it is not possible to construct series approximations that converge uniformly even for some very simple analogous extensions of the scalar case. We use a canonical form PDE of the form:

$$\frac{\partial f}{\partial \Delta}(\Delta, x) = \sum_{i=1}^{N} \mu_i(x)\frac{\partial f}{\partial x_i}(\Delta, x) + \frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2 f}{\partial x_i^2}(\Delta, x) \tag{2.18}$$

$$f(0, x) = g(x) \tag{2.19}$$

where x without a subscript denotes the N-element vector of state variables. In the scalar case, this PDE can be converted to the form of (2.1) (the form used by Kimmel (2008c) and Kimmel (2008b)), but the reverse transformation is, in general, neither easy to find nor unique. The importance of this case is not so much for pricing or conditional moment problems in the form described by these two equations, but for problems that, analogously to the scalar problems of Kimmel (2008a), can be transformed to (2.18) and (2.19) by a change of independent and dependent variables. Many non-linear pricing and conditional moment problems fit into this category. Note, however, that the multivariate case is different than the scalar case, in that there do exist diffusion problems that cannot be transformed to (2.18) and (2.19) by change of variables; in the scalar case, every diffusion problem can be converted to the scalar version of (2.18) and (2.19).

Taking $N = 2$ as an example, we consider the case in which the elements of $\mu_i(x)$ are linear in x:

$$\mu_1(x) = b_{11}x_1 + b_{12}x_2 \tag{2.20}$$

$$\mu_2(x) = b_{21}x_1 + b_{22}x_2 \tag{2.21}$$

We find it convenient to reparameterize the slope coefficients in the $\mu_1(x)$ and $\mu_2(x)$ functions. We first arrange these parameters in a matrix; if that matrix has a spectral representation, it can be written as:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \frac{\begin{bmatrix} \sin\beta & -\sin\alpha \\ -\cos\beta & \cos\alpha \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha \\ \cos\beta & \sin\beta \end{bmatrix}}{\sin(\beta - \alpha)} \tag{2.22}$$

with $\sin(\beta - \alpha) \neq 0$.

A transform of independent variables, analogous to that used in the scalar case, successfully eliminates the spatial derivative terms from the PDE:

$$y_1(\Delta, x_1, x_2) = e^{\lambda_1 \Delta}(x_1 \cos\alpha + x_2 \sin\alpha) \tag{2.23}$$

$$y_2(\Delta, x_1, x_2) = e^{\lambda_2 \Delta}(x_1 \cos\beta + x_2 \sin\beta) \tag{2.24}$$

Expressing the solution as:

$$f(\Delta, x_1, x_2) = \phi(\Delta, y_1(\Delta, x_1, x_2), y_2(\Delta, x_1, x_2)) \tag{2.25}$$

we find that $\phi(\Delta, y_1, y_2)$ satisfies the PDE:

$$\frac{\partial\phi}{\partial\Delta}(\Delta, y_1, y_2) = \begin{bmatrix} \dfrac{e^{2\lambda_1\Delta}}{2}\dfrac{\partial^2\phi}{\partial y_1^2}(\Delta, y_1, y_2) + \dfrac{e^{2\lambda_2\Delta}}{2}\dfrac{\partial^2\phi}{\partial y_2^2}(\Delta, y_1, y_2) \\ +\cos(\beta - \alpha)\, e^{(\lambda_1+\lambda_2)\Delta}\dfrac{\partial^2\phi}{\partial y_1 \partial y_2}(\Delta, y_1, y_2) \end{bmatrix} \tag{2.26}$$

$$\phi(0, y_1, y_2) = g\left(\frac{y_1 \sin\beta - y_2 \sin\alpha}{\sin(\beta - \alpha)}, \frac{y_2 \cos\alpha - y_1 \cos\beta}{\sin(\beta - \alpha)}\right) \tag{2.27}$$

If the left-hand side of (2.22) does not have a spectral representation (this is the case if its two eigenvalues are equal, but at least one of b_{12} or b_{21} is not zero), then it can be expressed instead in the following form:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \lambda - \delta\sin\alpha & \delta(\cos\alpha + 1) \\ \delta(\cos\alpha - 1) & \lambda + \delta\sin\alpha \end{bmatrix} \tag{2.28}$$

In this case, the state variable transformations that eliminate the first spatial derivative terms from the PDE are:

$$y_1(\Delta, x_1, x_2) = e^{\lambda\Delta}\left[x_1(1-\delta\Delta\sin\alpha) + x_2\delta\Delta(\cos\alpha+1)\right] \tag{2.29}$$

$$y_2(\Delta, x_1, x_2) = e^{\lambda\Delta}\left[x_1\delta\Delta(\cos\alpha-1) + x_2(1+\delta\Delta\sin\alpha)\right] \tag{2.30}$$

The transformed PDE is then:

$$\frac{\partial\phi}{\partial\Delta}(\Delta, y_1, y_2) = \frac{e^{2\lambda\Delta}}{2}\begin{bmatrix} \left[1-2\delta\Delta\sin\alpha+2\delta^2\Delta^2(\cos\alpha+1)\right]\dfrac{\partial^2\phi}{\partial y_1^2}(\Delta, y_1, y_2) \\ +\left[4\delta\Delta(\cos\alpha+\delta\Delta\sin\alpha)\right]\dfrac{\partial^2\phi}{\partial y_1\partial y_2}(\Delta, y_1, y_2) \\ +\left[1+2\delta\Delta\sin\alpha-2\delta^2\Delta^2(\cos\alpha-1)\right]\dfrac{\partial^2\phi}{\partial y_2^2}(\Delta, y_1, y_2) \end{bmatrix} \tag{2.31}$$

$$\phi(0, y_1, y_2) = g(y_1, y_2) \tag{2.32}$$

The two PDEs (2.26) and (2.31), together with their respective final conditions (2.27) and (2.32), cover all possible cases in a two-dimensional setting. Furthermore, as in the scalar case, time-dependent change of the state variable eliminates the first spatial derivative terms. However, whichever PDE we end up with, the next step, of eliminating the time-dependence in the coefficients of the second derivative terms, is not straightforward. The same change change of the time variable used in the scalar case can eliminates the time-dependent coefficient of *one* of the second partial derivative terms, but it is impossible to eliminate all three simultaneously, unless extremely strong additional restrictions are imposed. Even in the case of two independent Ornstein-Uhlenbeck processes, it is not possible in general. For example, in (2.22), choose $\alpha = 0$ and $\beta = \pi/2$, so that:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{2.33}$$

Then (2.26) and (2.27) become:

$$\frac{\partial\phi}{\partial\Delta}(\Delta, y_1, y_2) = \frac{e^{2\lambda_1\Delta}}{2}\frac{\partial^2\phi}{\partial y_1^2}(\Delta, y_1, y_2) + \frac{e^{2\lambda_2\Delta}}{2}\frac{\partial^2\phi}{\partial y_2^2}(\Delta, y_1, y_2) \tag{2.34}$$

$$\phi(0, y_1, y_2) = g(y_1, y_2) \tag{2.35}$$

Here, the time transformation:

$$\tau(\Delta) = \begin{bmatrix} \dfrac{e^{2\lambda_1\Delta}-1}{2\lambda_1} & \text{if} & \lambda_1 \neq 0 \\ \Delta & \text{if} & \lambda_1 = 0 \end{bmatrix} \tag{2.36}$$

eliminates the coefficient in front of one of the two second derivative terms in (2.34), and the transformation:

$$\tau(\Delta) = \begin{bmatrix} \dfrac{e^{2\lambda_2\Delta}-1}{2\lambda_2} & \text{if} & \lambda_2 \neq 0 \\ \Delta & \text{if} & \lambda_2 = 0 \end{bmatrix} \tag{2.37}$$

eliminates the other. However, unless $\lambda_1 = \lambda_2$, it is not possible to achieve both goals simultaneously.[8] If $\lambda_1 \neq \lambda_2$ and the final condition can be written in, for example, one of the forms:

$$g(y_1, y_2) = g_1(y_1)\, g_2(y_2) \tag{2.38}$$

$$g(y_1, y_2) = g_1(y_1) + g_2(y_2) \tag{2.39}$$

then the conditional moment problem can be broken into two separate probability problems, and a separate time transformation can be applied to each. However, this kind of separation is not possible in general, and although the time transformation method of Kimmel (2008b) can be applied here, it does not achieve the desired result of uniform convergence on $\Delta \in [0, +\infty)$ in a multifactor setting.

The method we propose in the next section is to embed the half-line corresponding to positive values of time, $\Delta \in [0, +\infty)$, in a higher-dimensional space built from multiple, artificial time dimensions. We then apply a time transformation technique similar to that of Kimmel (2008b), but in multiple dimensions, which allows straightforward derivation of power series approximations to the desired solution in the space of multi-dimensional time. The solution in real time is then simply the restriction of the more general solution to the path of true time, which in general is a curve through the higher-dimensional space of artificial time. Furthermore, this curve of true time through the higher-dimensional space of artificial time is contained within a compact set, provided the eigenvalues of the matrix of slope coefficients in the drift (e. g., either (2.22) or (2.28) in the 2-dimensional case) all have negative real parts. If convergence of a power series approximation in artificial time can be established in a region that contains this compact set, then the desired result of uniform convergence on $\Delta \in [0, +\infty)$ is established.

3. Artificial Time

In this section, we resolve the problem of the previous section by the introduction of multiple time dimensions. Although obviously an abstract formalism, this method often establishes uniform convergence of power series approximations in the multivariate setting. We focus on independent state variable processes; a more general setting is considered in Section 4.

3.1. Extended Time

Continuing with the example of the previous section, we find that, in the two-dimensional case, it is possible to transform the PDE of (2.18) with final condition (2.19) (and the linear drift specification of (2.7)) to either (2.26) or (2.31), with final condition (2.27) or (2.32) respectively, depending on whether the matrix of the b coefficients has an eigenvalue/eigenvector representation. Although the first spatial derivative terms have been eliminated from the PDE, the coefficients of the three second derivative terms are, in general, time-dependent, and, when considered as functions of Δ, linearly independent. Consequently, it is not possible to use a single

[8] If one of λ_1 or λ_2 is an integer multiple of the other, then it is possible to change the time variable so that, in the transformed PDE, the coefficient on one second-order term is constant, and the coefficient on the other second-order term is a polynomial function of the new time variable. If the polynomial is simple enough, series approximation may still be practical.

time transformation, as in Kimmel (2008b), to normalize all of these coefficients simultaneously to constants. However, the situation is considerably improved if we embed the time variable, $\Delta \in [0, +\infty)$, as a curve through a higher-dimensional space consisting of multiple, artificial time dimensions. These additional time dimensions have no physical meaning; only the path of Δ through this space can truly be regarded as "time." Nonetheless, this abstract formalism is very useful in solving real-world conditional moment or asset pricing problems.

Focusing on the example of (2.26) and (2.27) in which $\alpha = 0$ and $\beta = \pi/2$, the PDE (with final condition) is then:

$$\frac{\partial \phi}{\partial \Delta}(\Delta, y_1, y_2) = \frac{e^{2\lambda_1 \Delta}}{2}\frac{\partial^2 \phi}{\partial y_1^2}(\Delta, y_1, y_2) + \frac{e^{2\lambda_2 \Delta}}{2}\frac{\partial^2 \phi}{\partial y_2^2}(\Delta, y_1, y_2) \tag{3.1}$$

$$\phi(0, y_1, y_2) = g(y_1, y_2) \tag{3.2}$$

There is no time transformation that can simultaneously eliminate both the coefficients on the right-hand side of (3.1), except in the restricted special case $\lambda_1 = \lambda_2$. However, we can consider solutions $w(\tau_1, \tau_2, y_1, y_2)$ to the alternate problem:

$$\frac{\partial w}{\partial \tau_{11}}(\tau_{11}, \tau_{22}, y_1, y_2) = \frac{1}{2}\frac{\partial^2 w}{\partial y_1^2}(\tau_{11}, \tau_{22}, y_1, y_2) \tag{3.3}$$

$$\frac{\partial w}{\partial \tau_{22}}(\tau_{11}, \tau_{22}, y_1, y_2) = \frac{1}{2}\frac{\partial^2 w}{\partial y_2^2}(\tau_{11}, \tau_{22}, y_1, y_2) \tag{3.4}$$

$$w(0, 0, y_1, y_2) = g(y_1, y_2) \tag{3.5}$$

If a solution to the alternate problem can be found, then we can construct a solution to the original problem as follows:

$$\phi(\Delta, y_1, y_2) = w(\tau_{11}(\Delta), \tau_{22}(\Delta), y_1, y_2) \tag{3.6}$$

$$\tau_{11}(\Delta) = \begin{bmatrix} \dfrac{e^{2\lambda_1 \Delta} - 1}{2\lambda_1} & \text{if} & \lambda_1 \neq 0 \\ \Delta & \text{if} & \lambda_1 = 0 \end{bmatrix} \tag{3.7}$$

$$\tau_{22}(\Delta) = \begin{bmatrix} \dfrac{e^{2\lambda_2 \Delta} - 1}{2\lambda_2} & \text{if} & \lambda_2 \neq 0 \\ \Delta & \text{if} & \lambda_2 = 0 \end{bmatrix} \tag{3.8}$$

and substituting (3.6) into (3.1) and (3.2). Using the fact that $w(\tau_{11}, \tau_{22}, y_1, y_2)$ is a solution to (3.3), (3.4), and (3.5), it can readily be verified that $\phi(\Delta, y_1, y_2)$ is a solution to the original problem. The practical problem, of finding a contingent claim price or conditional moment in a two-dimensional diffusion, can therefore be found as a restriction of a solution to an alternate problem in an artificial setting with two time dimensions. If $\lambda_1 < 0$ and $\lambda_2 < 0$, then the time transformations of (3.7) and (3.8) are (for non-negative real values of Δ)

monotonically increasing, and have the properties:

$$\tau_{11}(0) = 0 \qquad \tau_{11}(+\infty) = -\frac{1}{2\lambda_1} \tag{3.9}$$

$$\tau_{22}(0) = 0 \qquad \tau_{22}(+\infty) = -\frac{1}{2\lambda_2} \tag{3.10}$$

The set of all points $[\tau_{11}(\Delta), \tau_{22}(\Delta)]$ corresponding to $\Delta \in [0, +\infty)$ then forms a curve contained in a rectangle, beginning at $(0, 0)$ (for $\Delta = 0$) and approaching $(-1/(2\lambda_1), -1(2\lambda_1))$ asymptotically as Δ approaches $+\infty$, as shown in Figure 1.

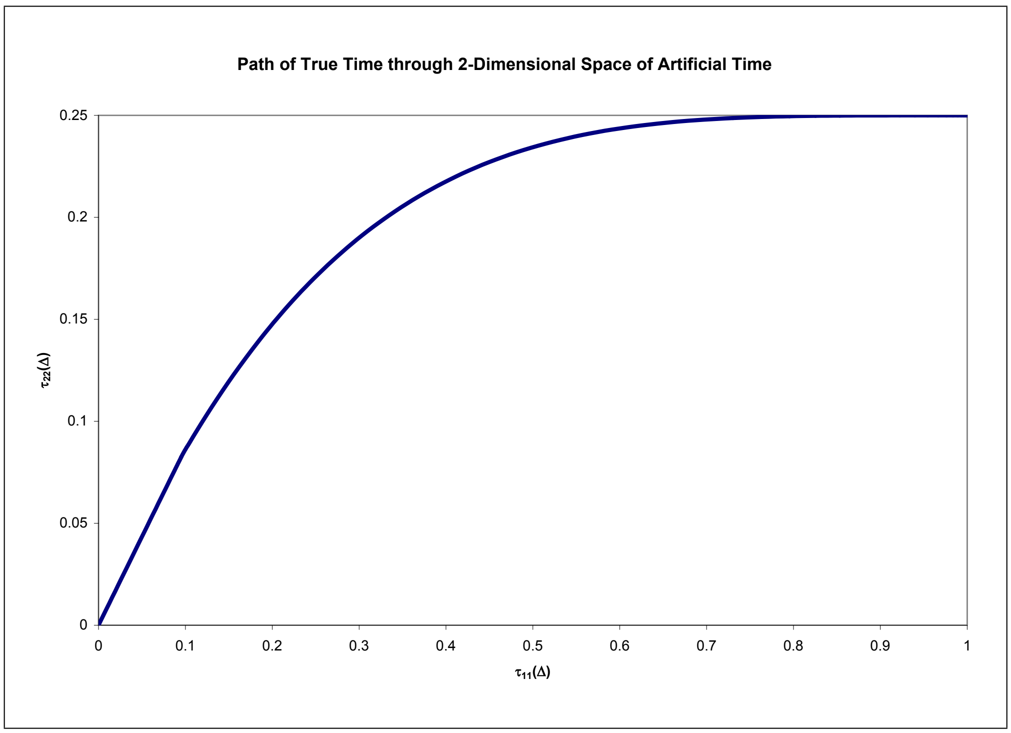

Figure 1: This figure shows the path of true time through a unit square, using the time transformations of (3.7) and (3.8), with $\lambda_1 = -0.5$ and $\lambda_2 = -2.0$. $\Delta = 0$ corresponds to the point $\tau_{11} = \tau_{22} = 0$, and the curve approaches $(\tau_{11}, \tau_{22}) = (1, 0.25)$ as Δ approaches $+\infty$. The joint time transformation therefore maps the infinite interval $\Delta \in [0, +\infty)$ to a subset of a compact set in a two-dimensional space.

If a solution $w(\tau_{11}, \tau_{22}, y_1, y_2)$ to (3.3), (3.4), and (3.5) can be shown to be analytic in τ_{11} and τ_{22} in some neighborhood of $\tau_{11} = \tau_{22} = 0$, then a power series approximation can be found by a simple recursive procedure. We express the solution as:

$$w(\tau_{11}, \tau_{22}, y_1, y_2) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{a_{ij}(y_1, y_2)}{i!j!} \tau_{11}^i \tau_{22}^j \tag{3.11}$$

Substituting this expression into (3.3), (3.4), and (3.5), and, in the first two, gathering terms of like order in

τ_{11} and τ_{22}, we find the coefficients must satisfy:

$$a_{ij}(y_1, y_2) = g(y_1, y_2) \qquad i, j = 0 \tag{3.12}$$

$$a_{ij}(y_1, y_2) = \frac{1}{2}\frac{\partial^2 a_{i-1j}}{\partial y_1^2}(y_1, y_2) \qquad i > 0 \tag{3.13}$$

$$a_{ij}(y_1, y_2) = \frac{1}{2}\frac{\partial^2 a_{ij-1}}{\partial y_2^2}(y_1, y_2) \qquad j > 0 \tag{3.14}$$

It may not be obvious that the above relations are consistent, since for $i, j > 0$, the $a_{ij}(y_1, y_2)$ coefficient is defined in terms of both $a_{i-1j}(y_1, y_2)$ and $a_{ij-1}(y_1, y_2)$. However, the two definitions are consistent, due to the commutativity of the operators on the right-hand sides. The consistency also follows from (3.3) and (3.4); any function that is analytic in the two time variables and that solves both equations must have a power series representation, and the recursions (3.12) through (3.14) are derived from (3.3) and (3.4), respectively.[9] Provided the final condition is infinitely differentiable in both variables, the power series coefficients can be calculated to any desired order, although the same can be said about a power series solution in Δ to the original problem, (3.1) and (3.2). However, the advantage of the artificial formalism of multi-dimensional time is clear when a solution to (3.3), (3.4), and (3.5) can be shown to be analytic for all $|\tau_{11}| < s_{11}$ and $|\tau_{22}| < s_{22}$, for some $s_{11} > -1/(2\lambda_{11})$ and $s_{22} > -1/(2\lambda_{22})$. The two-dimensional power series approximation to $w(\tau_{11}, \tau_{22}, y_1, y_2)$ then converges *uniformly* for $|\tau_{11}| \leq -1/(2\lambda_{11})$ and $|\tau_{22}| \leq -1/(2\lambda_{22})$, and the set of all $\tau_{11}(\Delta)$ and $\tau_{22}(\Delta)$ for $\Delta \in [0, +\infty)$ falls within this range. It is therefore possible, using this method, to construct a power series approximation that converges uniformly for all positive Δ, when it is not possible to construct a uniformly convergent series in a single-dimensional time, transformed or not. The introduction of artificial time dimensions therefore allows us to construct series approximations to problems that arise in the familiar setting of one-dimensional time, but with improved convergence properties.

3.2. Example

A very simple example, in which the conditional moment sought is known in closed-form, serves to illustrate how multi-dimensional time can lead to approximations with improved convergence properties. Of course, since the solution is already known, this case serves only as an illustrative example; furthermore, the two state variables enter into the final condition in an additively separable way, allowing the problem to be split into two scalar problems. However, it is easy to construct examples in which it is not possible to decompose the problem into independent scalar problems in this way. Consider:

$$\frac{\partial f}{\partial \Delta}(\Delta, x_1, x_2) = \begin{bmatrix} \lambda_1 x_1 \dfrac{\partial f}{\partial x_1}(\Delta, x_1, x_2) + \dfrac{1}{2}\dfrac{\partial^2 f}{\partial x_1^2}(\Delta, x_1, x_2) \\ +\lambda_2 x_2 \dfrac{\partial f}{\partial x_2}(\Delta, x_1, x_2) + \dfrac{1}{2}\dfrac{\partial^2 f}{\partial x_2^2}(\Delta, x_1, x_2) \end{bmatrix} \tag{3.15}$$

$$f(0, x_1, x_2) = x_1^4 + x_2^4 \tag{3.16}$$

[9] Note that existence of an analytic solution to (3.3), (3.4), and (3.5) has not yet been proven; this issue is addressed in Section 3.3. However, if such a solution does exist, its power series representation must obey the recursion relations (3.12) through (3.14).

with $\lambda_1, \lambda_2 < 0$. The change of variables:

$$y_1 = x_1 e^{\lambda_1 \Delta} \tag{3.17}$$

$$y_2 = x_2 e^{\lambda_2 \Delta} \tag{3.18}$$

and the definition of $f(\Delta, x_1, x_2) = \phi(\Delta, y_1(\Delta, x_1, x_2), y_2(\Delta, x_1, x_2))$ results in a PDE of the form:

$$\frac{\partial \phi}{\partial \Delta}(\Delta, y_1, y_2) = \frac{e^{2\lambda_1 \Delta}}{2} \frac{\partial^2 \phi}{\partial y_1^2}(\Delta, y_1, y_2) + \frac{e^{2\lambda_2 \Delta}}{2} \frac{\partial^2 \phi}{\partial y_2^2}(\Delta, y_1, y_2) \tag{3.19}$$

$$\phi(0, y_1, y_2) = y_1^4 + y_2^4 \tag{3.20}$$

The solution to this transformed PDE is given by:

$$\phi(\Delta, y_1, y_2) = 3\left[\left(y_1^2 + \frac{e^{2\lambda_1 \Delta} - 1}{2\lambda_1}\right)^2 + \left(y_2^2 + \frac{e^{2\lambda_2 \Delta} - 1}{2\lambda_2}\right)^2\right] - 2y_1^4 - 2y_2^4 \tag{3.21}$$

If one were to use the time transformation of Kimmel (2008b) (or any other time transformation, for that matter), it would quickly be discovered that for any transform that maps $\Delta \in [0, +\infty)$ to some bounded set, the solution is *not* analytic in the transformed time variable at the point corresponding to $\Delta = +\infty$, unless either λ_1/λ_2 or λ_2/λ_1 is an integer. Apart from these special cases, a power series approximation in the transformed variable may converge for all $\Delta \in [0, +\infty)$, but will not do so uniformly. However, using the multi-dimensional time transform of (3.7) and (3.8), the solution is:

$$\phi(\Delta, y_1, y_2) = 3\left[\left(y_1^2 + \tau_{11}\right)^2 + \left(y_2^2 + \tau_{22}\right)^2\right] - 2y_1^4 - 2y_2^4 \tag{3.22}$$

which is everywhere analytic in both τ_{11} and τ_{22}; a power series expansion to the solution can easily be found by recursive calculation, and converges for all τ_{11} and τ_{22}. Furthermore, this power series converges *uniformly* on a rectangle in (τ_{11}, τ_{22}) with opposite corners given by $(0, 0)$ and $(-1/(2\lambda_1), -1/(2\lambda_2))$. Since the path of $[\tau_{11}(\Delta), \tau_{22}(\Delta)]$ for $\Delta \in [0, +\infty)$ falls within this rectangle, with $\tau_{11}(0) = \tau_{22}(0) = 0$, $\tau_{11}(+\infty) = -1/(2\lambda_1)$, and $\tau_{22}(+\infty) = -1/(2\lambda_2)$, convergence on $\Delta \in [0, +\infty)$ is uniform.

3.3. Existence of Solutions

So far, we have focused on transforming a pricing or conditional moment problem based on an two-dimensional diffusion into a problem of the form:

$$\frac{\partial w}{\partial \tau_{11}}(\tau, y) = \frac{1}{2} \frac{\partial^2 w}{\partial y_1^2}(\tau, y) \tag{3.23}$$

$$\frac{\partial w}{\partial \tau_{22}}(\tau, y) = \frac{1}{2} \frac{\partial^2 w}{\partial y_2^2}(\tau, y) \tag{3.24}$$

$$w(0, y_1, y_2) = g(y_1, y_2) \tag{3.25}$$

where τ and y without subscripts denote vectors. If $\lambda_1 < 0$ and $\lambda_2 < 0$ (whether the parameterization of (2.22) or (2.28) applies), then the set of all times τ_{11} and τ_{22} corresponding to $\Delta \in [0, +\infty)$ lies within a bounded rectangle. If a power series approximation to the solution converges in a region including this rectangle, it then converges uniformly for all positive real Δ. We have paid little attention, however, to the problem of determining whether a solution to the transformed problem exists, and, if so, whether it is analytic in a region including the required rectangle. The following result addresses this problem in a general N-dimensional setting, but requires the implied state variable processes to be independent. A discussion of the more general case, with dependent state variable processes, is deferred until Section 4.

Theorem 1 *Let y be a vector of N state variables, and let $g(y)$ be an everywhere analytic function satisfying the growth condition:*

$$|g(y)| \leq c e^{\sum_{i=1}^{N} \frac{\|y_i\|_i^2}{2}} \tag{3.26}$$

for all complex vectors y, where $c > 0$, and $\|y\|_i$ is a norm (over the reals) for each $1 \leq i \leq N$. Then there exists a solution $w(\tau, y)$ to the system of PDEs with final condition:

$$\frac{\partial w}{\partial \tau_{ii}}(\tau, y) = \frac{1}{2}\frac{\partial^2 w}{\partial y_i^2}(\tau, y) \qquad 1 \leq i \leq N \tag{3.27}$$

$$w(0, y) = g(y) \tag{3.28}$$

that is defined and analytic for all complex vectors y and all complex vectors τ such that $\left\|\sqrt{\tau_{ii}}\right\|_i \leq 1$ for all $1 \leq i \leq N$.

Proof: See Appendix.

In particular, if in Theorem 1, we have $\left\|\sqrt{-1/(2\lambda_i)}\right\|_i < 1$ for each $1 \leq i \leq N$, then a power series approximation to $w(\tau, y)$ converges on a rectangle that includes all future values of "true" time, provided the λ_i are all negative. Furthermore, the power series converges uniformly for these times.

We can state a result in terms of the original multiple-dimensional Ornstein-Uhlenbeck process, rather than the transformed system of PDEs considered in Theorem 1 above.

Corollary 1 *Let x be a vector of N state variables, and let $g(x)$ be an everywhere analytic function satisfying the growth condition:*

$$|g(x)| \leq c e^{\sum_{i=1}^{N} \frac{\|x_i\|_i^2}{2}} \tag{3.29}$$

for all complex vectors y, where $c > 0$, and $\|x\|_i$ is a norm (over the reals) for each $1 \leq i \leq N$. Then there exists a solution $f(\Delta, x)$ to the PDE with final condition:

$$\frac{\partial f}{\partial \Delta}(\Delta, x) = \sum_{i=1}^{N}\left[\mu_i(x)\frac{\partial f}{\partial x_i}(\Delta, x)\right] + \frac{1}{2}\sum_{i=1}^{N}\frac{\partial f}{\partial x_i^2}(\Delta, x) \tag{3.30}$$

$$f(0, x) = g(x) \tag{3.31}$$

where $\mu_i(x) = \lambda_i x_i$ for all $1 \leq i \leq N$, that is defined and analytic for all complex vectors x and all complex

Δ *such that* $\left\| \sqrt{\tau_{ii}(\Delta)} \right\|_i < 1$ *for all* $1 \le i \le N$*, where:*

$$\tau_{ii}(\Delta) = \begin{bmatrix} \dfrac{e^{2\lambda_i \Delta} - 1}{2\lambda_i} & \quad \textit{if} & \quad \lambda_i \neq 0 \\ \Delta & \quad \textit{if} & \quad \lambda_i = 0 \end{bmatrix} \tag{3.32}$$

Furthermore, the solution can be expressed as:

$$f(\Delta, x) = w(\tau(\Delta), y(\Delta, x)) \tag{3.33}$$

with

$$y_i(\Delta, x) = x_i e^{\lambda_i \Delta} \tag{3.34}$$

where $w(\tau, y)$ *is defined and analytic for all* y *and all* $\left\| \sqrt{\tau_i} \right\|_i < 1$*, and solves:*

$$\frac{\partial w}{\partial \tau_i} = \frac{1}{2} \frac{\partial^2 w}{\partial y_i^2} \qquad 1 \le i \le N \tag{3.35}$$

$$w(0, y) = g(y) \tag{3.36}$$

Proof: See Appendix.

The above result provides a powerful tool for establishing *uniform* convergence of power series approximations to conditional moments of independent Ornstein-Uhlenbeck processes. However, this result also achieves the same result for the large class of non-affine conditional moment and contingent claim pricing problems that can be transformed, by change of dependent variable, to conditional moment problems in a multiple-dimensional independent Ornstein-Uhlenbeck process setting. Although this class of problems is very large, it is possible to do better still. In the next section, however, we present two different methods for extending the results to a general multiple-dimensional Ornstein-Uhlenbeck setting, in which the state variable processes need not be independent.

4. Dependent Processes

The transformation of time by (3.7) and (3.8) does not suffice to simply the PDE problem in the general two-dimensional setting. For example, if we choose $\alpha = 0$ (as before), but make no restriction on β (other than $\sin\beta \neq 0$), then:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ (\lambda_2 - \lambda_1)\cos\beta & \lambda_2 \sin\beta \end{bmatrix} \tag{4.1}$$

Then (2.26) and (2.27) simplify to:

$$\frac{\partial \phi}{\partial \Delta}(\Delta, y_1, y_2) = \frac{e^{2\lambda_1 \Delta}}{2}\frac{\partial^2 \phi}{\partial y_1^2}(\Delta, y_1, y_2) + \cos\beta e^{(\lambda_1+\lambda_2)\Delta}\frac{\partial^2 \phi}{\partial y_1 \partial y_2}(\Delta, y_1, y_2) + \frac{e^{2\lambda_2 \Delta}}{2}\frac{\partial^2 \phi}{\partial y_2^2}(\Delta, y_1, y_2) \tag{4.2}$$

$$\phi(0, y_1, y_2) = g\left(y_1, \frac{y_2 - y_1 \cos\beta}{\sin\beta}\right) \tag{4.3}$$

However, unless $\cos\beta = 0$ (in which case (4.2) and (4.3) reduce to (3.1) and (3.2)), the time transformations (3.7) and (3.8) no longer suffice, because however the cross-derivative term is allocated when the PDE is broken into two, it still has a time-dependent coefficient. We now discuss two distinct methods of solving this problem, each with their own advantages and disadvantages. One method retains the simplicity of the recursion equations used in the case of independent processes, but in general introduces $N(N+1)/2$ artificial time dimensions for a PDE with N state variables. The other method uses only N artificial time dimensions for an N state variable system, but at the price of a more complicated recursion equation.

4.1. Additional Time Variables

Introduction of a third time dimension is one way to extend the method used for independent processes to the case of two dependent processes:

$$\tau_{12}(\Delta) = \begin{bmatrix} \dfrac{e^{(\lambda_1+\lambda_2)\Delta}}{\lambda_1+\lambda_2} & \text{if} & \lambda_1+\lambda_2 \neq 0 \\ \Delta & \text{if} & \lambda_1+\lambda_2 = 0 \end{bmatrix} \tag{4.4}$$

The map of true time corresponding to $\Delta \in [0, +\infty)$ through this three-dimensional artificial time construct is a curve in $(\tau_{11}, \tau_{22}, \tau_{12})$, beginning at $(0,0,0)$, and ending at $(-1/(2\lambda_1), -1/(2\lambda_2), -1/(\lambda_1+\lambda_2))$. If a solution to the problem:

$$\frac{\partial w}{\partial \tau_{11}}(\tau_{11}, \tau_{22}, \tau_{12}, y_1, y_2) = \frac{1}{2}\frac{\partial^2 w}{\partial y_1^2}(\tau_{11}, \tau_{22}, \tau_{12}, y_1, y_2) \tag{4.5}$$

$$\frac{\partial w}{\partial \tau_{22}}(\tau_{11}, \tau_{22}, \tau_{12}, y_1, y_2) = \frac{1}{2}\frac{\partial^2 w}{\partial y_2^2}(\tau_{11}, \tau_{22}, \tau_{12}, y_1, y_2) \tag{4.6}$$

$$\frac{\partial w}{\partial \tau_{12}}(\tau_{11}, \tau_{22}, \tau_{12}, y_1, y_2) = \cos\beta \frac{\partial^2 w}{\partial y_1 \partial y_2}(\tau_{11}, \tau_{22}, \tau_{12}, y_1, y_2) \tag{4.7}$$

$$w(0,0,0, y_1, y_2) = g\left(y_1, \frac{y_2 - y_1\cos\beta}{\sin\beta}\right) \tag{4.8}$$

can be found, then $\phi(\Delta, y_1, y_2) = w(\tau_{11}(\Delta), \tau_{22}(\Delta), \tau_{12}(\Delta), y_1, y_2)$ is a solution to (4.2) and (4.3), as can readily be verified by substituting in this proposed solution. More generally, any N-dimensional version of this problem can be converted to a problem with $N(N+1)/2$ artificial time variables, and provided the eigenvalues of the b matrix all have negative real parts, then all values of $\Delta \in [0, +\infty)$ are contained within a bounded $N(N+1)/2$-dimensional rectangle. If the solution to the transformed problem can then be shown to be analytic for all $|\tau_{ij}| < s_{ij}$ where $s_{ij} > 0$ are chosen to include such a rectangle, then the power series approximation to that solution converges *uniformly* in a region that includes $\Delta \in [0, +\infty)$. In a term structure setting, for example, it would then be possible to approximate bond prices and yields uniformly accurately, even for very long maturities.

It is not so difficult to characterize solutions to the transformed problem (4.5) through (4.8) that are analytic in the three time variables in some unspecified neighborhood of the origin, but explicitly characterizing the region of analyticity can be tricky. The following result is describes sufficient conditions for existence of a solution that is everywhere analytic in all τ_{ij} variables:

Theorem 2 *Let y be a vector of N state variables, let $g(y)$ be an everywhere analytic function, and for each $k > 0$, let there be some $c_k > 0$ such that $g(y)$ satisfies the growth condition:*

$$|g(y)| \leq c_k e^{\sum_{i=1}^{N} \frac{|y_i|^2}{2k}} \tag{4.9}$$

for all complex vectors y. Let d_{ij}, $1 \leq i \leq j \leq N$, be any set of numbers. Then there exists a solution $w(\tau, y)$ to the system of PDEs with final condition:

$$\frac{\partial w}{\partial \tau_{ij}}(\tau, y) = d_{ij} \frac{\partial^2 w}{\partial y_i \partial y_j}(\tau, y) \qquad 1 \leq i \leq j \leq N \tag{4.10}$$

$$w(0, y) = g(y) \tag{4.11}$$

that is defined and analytic for all complex vectors y and all complex vectors τ.

Proof: See Appendix.

The above result establishes that, provided the final condition is analytic and grows at less than an exponential quadratic rate in *all* directions in the complex plane, the solution is everywhere analytic in the time variables. Recalling that the problem is motivated by a related problem in a multivariate Ornstein-Uhlenbeck process, we note that if the positive real line in $\Delta \in [0, +\infty)$ maps to bounded intervals in the τ_{ij} (e. g., if the eigenvalues of the b matrix all have negative real parts), then a power series approximation in the τ_{ij} to the solution $w(\tau, y)$ converges *uniformly* for all $\Delta \in [0, +\infty)$.

The above result can be useful in establish uniform convergence of approximations in some cases, but is too restrictive for other cases. In particular, it is useful to consider cases where the solution exists and is analytic for some region in the τ_{ij}, but not everywhere. If $\Delta \in [0, +\infty)$ maps to a bounded interval in the τ_{ij}, then existence and analyticity of a solution in a hypersphere containing that interval is sufficient to establish uniform convergence on $\Delta \in [0, +\infty)$. The following result establishes the sufficient existence and analyticity properties with a weaker growth condition:

Theorem 3 *Let y be a vector of N state variables, let $g(y)$ be an everywhere analytic function, and let there be some $c > 0$ such that $g(y)$ satisfies the growth condition:*

$$|g(y)| \leq c e^{\sum_{i=1}^{N} \frac{|y_i|^2}{2}} \tag{4.12}$$

for all complex vectors y. Let $T_{ij} > 0$ and d_{ij}, $1 \leq i \leq j \leq N$, be numbers such that:

$$\det \left\{ I - \begin{bmatrix} d_{11}\tau_{11} & \cdots & d_{1N}\tau_{1N} \\ \vdots & \ddots & \vdots \\ d_{N1}\tau_{N1} & \cdots & d_{NN}\tau_{NN} \end{bmatrix} \right\} \neq 0 \tag{4.13}$$

for all $|\tau_{ij}| < T_{ij}$, $1 \leq i \leq j \leq N$, where I denotes the $N \times N$ identity matrix. Then there exists a solution $w(\tau, y)$ to the system of PDEs with final condition:

$$\frac{\partial w}{\partial \tau_{ij}}(\tau, y) = d_{ij} \frac{\partial^2 w}{\partial y_i \partial y_j}(\tau, y) \qquad 1 \leq i \leq j \leq N \tag{4.14}$$

$$w(0, y) = g(y) \tag{4.15}$$

that is defined and analytic for all complex vectors y and all values of $|\tau_{ij}| < T_{ij}$, $1 \leq i \leq j \leq N$.

Proof: See Appendix.

The above result establishes existence and analyticity of a solution within circles $|\tau_{ij}| < T_{ij}$ in each of the τ_{ij}. In the motivating example (providing the λ_i are negative), $\Delta \in [0, +\infty)$ maps to a subset of $|\tau_{ij}| < -1/(\lambda_i + \lambda_j)$ for each $1 \leq i \leq j \leq N$. Therefore, an appropriate choice is $T_{ij} = -\lambda_i - \lambda_j + \epsilon$ where $\epsilon > 0$ is an arbitrarily small positive number. If the solution $w(\tau, y)$ can be shown to be analytic for $|\tau_{ij}| < T_{ij}$ with this choice of the T_{ij}, then the power series approximation will converge for values of τ_{ij} corresponding to $\Delta \in [0, +\infty)$.

4.2. Alternate Parameterization of Time

The approach discussed in the previous section has the advantage of the retaining the simple recursion equation. However, the number of time variables required grows quadratically with the number of state variables. Furthermore, the growth restriction on the PDE final condition needed to establish analyticity within a given region is sometimes difficult to characterize. It is possible to use an alternate parameterization of the time variables, such that one time variable is required for each state variable, and such that the growth restriction on the final condition is more precisely characterized than in the previous section. These advantages come at the price of a more complicated recursion equation.

We first note that the three time variables (in the case of a two-dimensional process) of the previous section can be written as:

$$\tau_{11}(\Delta) = \lambda_1 s_1^2(\Delta) + 2s_1(\Delta) \tag{4.16}$$

$$\tau_{12}(\Delta) = \frac{\lambda_1 \lambda_2 s_1(\Delta) s_2(\Delta) + \lambda_1 s_1(\Delta) + \lambda_2 s_2(\Delta)}{\lambda_1 + \lambda_2} \tag{4.17}$$

$$\tau_{22}(\Delta) = \lambda_2 s_2^2(\Delta) + 2s_2(\Delta) \tag{4.18}$$

where:

$$s_1(\Delta) = \begin{bmatrix} \dfrac{e^{\lambda_1 \Delta} - 1}{\lambda_1} & \text{if} & \lambda_1 \neq 0 \\ \Delta & \text{if} & \lambda_1 = 0 \end{bmatrix} \tag{4.19}$$

$$s_2(\Delta) = \begin{bmatrix} \dfrac{e^{\lambda_2 \Delta} - 1}{\lambda_2} & \text{if} & \lambda_2 \neq 0 \\ \Delta & \text{if} & \lambda_2 = 0 \end{bmatrix} \tag{4.20}$$

Note that, if λ_1 and λ_2 are both negative, then $s_1(\Delta)$ and $s_2(\Delta)$ take values in $[0, -1/\lambda_1)$ and $[0, -1/\lambda_2)$, respectively, for $\Delta \in [0, +\infty)$. If a solution to a conditional moment or asset pricing problem is analytic in τ_{11}, τ_{12}, and τ_{22} at the point specified by s_1 and s_2 (treating the τ variables as functions of the s variables), then the solution is analytic in s_1 and s_2 at that point. When the solution is expressed as a function of s_1 and s_2:

$$\phi(\Delta, y_1, y_2) = \psi(s_1(\Delta), s_2(\Delta), y_1, y_2) \tag{4.21}$$

then the general PDE (4.5) can then be written as:

$$\begin{bmatrix} (\lambda_1 s_1 + 1) \dfrac{\partial \psi}{\partial s_1}(s_1, s_2, y_1, y_2) \\ + (\lambda_2 s_2 + 1) \dfrac{\partial \phi}{\partial s_2}(s_1, s_2, y_1, y_2) \end{bmatrix} = \begin{bmatrix} \dfrac{1}{2}(\lambda_1 s_1 + 1)^2 \dfrac{\partial^2 \psi}{\partial y_1^2}(s_1, s_2, y_1, y_2) \\ + \dfrac{1}{2}(\lambda_2 s_2 + 1)^2 \dfrac{\partial^2 \psi}{\partial y_2^2}(s_1, s_2, y_1, y_2) \\ + \cos\beta (\lambda_1 s_1 + 1)(\lambda_2 s_2 + 1) \dfrac{\partial^2 \psi}{\partial y_1 \partial y_2}(s_1, s_2, y_1, y_2) \end{bmatrix} \tag{4.22}$$

As in Section 4.1, we would like to split this equation into two, each with a derivative with respect to one of the time variables (s_1 or s_2) on the left-hand side, and only terms with second-order derivatives with respect to the state variables on the right-hand side. It is not immediately obvious, though, how to allocate the second cross derivative between the two equations. But $\psi(s_1, s_2, y_1, y_2)$ can also be expressed in terms of the solution $w(\tau_{11}, \tau_{12}, \tau_{22}, y_1, y_2)$ derived in Section 4.1:

$$\psi(s_1, s_2, y_1, y_2) = w(tau_{11}(s_1), tau_{12}(s_1, s_2), tau_{22}(s_2), y_1, y_2) \tag{4.23}$$

Differentiating both sides with respect to s_1 and s_2, we find:

$$\frac{\partial \psi}{\partial s_1}(s_1, s_2, y_1, y_2) = 2(\lambda_1 s_1 + 1)\frac{\partial w}{\partial \tau_{11}} + \frac{\lambda_1}{\lambda_1 + \lambda_2}(\lambda_2 s_2 + 1)\frac{\partial w}{\partial \tau_{12}} \tag{4.24}$$

$$\frac{\partial \psi}{\partial s_2}(s_1, s_2, y_1, y_2) = 2(\lambda_2 s_2 + 1)\frac{\partial w}{\partial \tau_{22}} + \frac{\lambda_2}{\lambda_1 + \lambda_2}(\lambda_1 s_1 + 1)\frac{\partial w}{\partial \tau_{12}} \tag{4.25}$$

Replacing the derivatives of ϕ with respect to the τ in (4.24) and (4.25) by the spatial derivatives on the right-hand side of (4.5), (4.6), and (4.7), and noting that the spatial derivatives of $w(\tau_{11}, \tau_{12}, \tau_{22}, y_1, y_2)$ are all equal to the corresponding spatial derivatives of $\psi(s_1, s_2, y_1, y_2)$, we derive the following system of PDEs:

$$\frac{\partial \psi}{\partial s_1}(s_1, s_2, y_1, y_2) = (\lambda_1 s_1 + 1)\frac{\partial^2 \psi}{\partial y_1^2}(s_1, s_2, y_1, y_2) + \cos\beta \frac{\lambda_1}{\lambda_1 + \lambda_2}(\lambda_2 s_2 + 1)\frac{\partial^2 \psi}{\partial y_1 \partial y_2}(s_1, s_2, y_1, y_2) \quad (4.26)$$

$$\frac{\partial \psi}{\partial s_2}(s_1, s_2, y_1, y_2) = (\lambda_2 s_2 + 1)\frac{\partial^2 \psi}{\partial y_2^2}(s_1, s_2, y_1, y_2) + \cos\beta \frac{\lambda_2}{\lambda_1 + \lambda_2}(\lambda_1 s_1 + 1)\frac{\partial^2 \psi}{\partial y_1 \partial y_2}(s_1, s_2, y_1, y_2) \quad (4.27)$$

$$\psi(0, 0, y_1, y_2) = g(y_1, y_2) \quad (4.28)$$

It is clear (by substititon of the $\psi(s_1, s_2, y_1, y_2)$ constructed here) that a solution to (4.26), (4.27), and (4.28) is also a solution to (2.26). The reverse is less obvious (and in general, not true). However, existence of the solutions established by Theorems 2 and 3 establishes by the construction shown here a solution to (4.26), (4.27), and (4.28). These latter equations also form the basis for recursive calculation of a power series in s_1 and s_2 instead of τ_{11}, τ_{12}, and τ_{22}. Of course, this alternate approach can be extended to higher-dimensional problems. Note that the operators, applied to $\phi(s_1, s_2, y_1, y_2)$, on the right-hand sides of (4.26), (4.27) do not, in general, commute. Nonetheless, because the solution is constructed from $w(\tau_{11}, \tau_{12}, \tau_{22}, y_1, y_2)$, which is proven to exist under the assumptions of Theorems 2 and 3, and which is also analytic in s_1 and s_2, the solution must satisfy each of the PDEs above, and using the power series representation of the solution, we can use these two PDEs to calculate coefficients recursively. For all terms in the power series that contain both s_1 and s_2 raised to some non-zero power, the coefficient can be calculated from previously calculated coefficients using either of the two PDEs. However, the definitions are necessarily consistent.

Throughout the remainder, we consider only the technique described in Section 4.1. However, we note that an alternate approach with N time variables instead of $N(N+1)/2$ is possible, at the price of a more complicated recursion equation.

5. A General Family of Non-Affine Term Structure Models

Solutions to the general multiple-dimensional PDE (2.18), with final condition (2.19), in which the elements of $\mu(x)$ are linear in x, correspond to conditional moments of a multifactor Ornstein-Uhlenbeck process. However, we have already noted that the applicability of our results is much wider, since many problems can be converted the problem of finding a conditional moment in such a process, by change of dependent variable. This fact makes our results particularly useful in constructing non-affine multifactor term structure models in which bond prices can be approximated very accurately. Recall the general problem:

$$\frac{\partial f}{\partial \Delta}(\Delta, x) = \sum_{i=1}^{N} \mu_i(x)\frac{\partial f}{\partial x_i}(\Delta, x) + \frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2 f}{\partial x_i^2}(\Delta, x) \quad (5.1)$$

$$f(0, x) = g(x) \quad (5.2)$$

where $\mu_i(x)$ is an affine function of x for each $1 \leq i \leq N$. We have derived conditions on $g(x)$ which establish existence and analyticity (in Δ) of the solution $f(\Delta, x)$, and joint restrictions on $g(x)$ and $\mu_i(x)$ to allow derivation of approximations to the solution that converge *uniformly* for all $\Delta \in [0, +\infty)$.

We can also reverse the change of dependent variables. If we define:

$$P(\Delta, x) = \frac{f(\Delta, x)}{g(x)} \tag{5.3}$$

then it should be noted that $P(0, x) = 1$, which is the final condition for a zero-coupon bond price. We can also plug this definition of $P(\Delta, x)$ into (5.1), and derive a PDE satisfied by $P(\Delta, x)$; in general, the coefficients of the first derivatives in the PDE satisfied by $P(\Delta, x)$ will not be affine, and the coefficient on $P(\Delta, x)$ itself (which is absent in the PDE (5.1) prior to transformation) will also not be affine. It follows that *every* specification of $g(x)$ that satisfies the conditions of Theorem 3, effectively specifies a multifactor term structure model. Furthermore, except for very specific choices of $g(x)$, this term structure model will not be affine. Despite its non-affinity, however, we are able to approximate bond prices using our technique.

By plugging (5.3) into (5.1) and gathering like derivatives in $P(\Delta, x)$, we can state explicitly the PDE satisfied by bond prices in a non-affine model constructed as described above:

$$\frac{\partial P}{\partial \Delta}(\Delta, x) = \sum_{i=1}^{N} \frac{\partial P}{\partial x_i}(\Delta, x)\left[\mu_i(x) + \sum_{j=1}^{N} \frac{\frac{\partial g}{\partial x_j}(x)}{g(x)}\right] + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2 P}{\partial x_i \partial x_j}(\Delta, x) \tag{5.4}$$

$$+ P(\Delta, x)\left[\mu_i \frac{\frac{\partial g}{\partial x_i}}{g(x)} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\frac{\partial^2 g}{\partial x_i \partial x_j}(x)}{g(x)}\right] \tag{5.5}$$

Any choice of $g(x)$ and affine $\mu_i(x)$ such that the transformed system (after the change of dependent variables $f(\Delta, x) = g(x) P(\Delta, x)$) satisfies the conditions of Theorem 3 specifies a term structure model in which bond prices can be approximated using our technique, and the approximations converge *uniformly* for all maturities $\Delta \in [0, +\infty)$.

6. Examples

In this section, we apply the approximation methods developed in the preceding sections to the problem of bond pricing in several multifactor term structure models, and examine the accuracy of the results. For two of the models, bond prices are already known in closed form, so the approximations serve as illustrative examples only. For the third model, bond prices are not known in closed-form, and we compare the accuracy of the approximations found by our method to bond prices calculated through numeric methods. In all three cases, we find that our method often produces highly accurate approximations for arbitrarily long maturities.

6.1. Affine Yield Models

In this section, we use our technique to approximate bond prices in an $A_0(2)$ affine yield model. Since bond prices are known in closed-form for this model, it serves as an illustrative example by which we can evaluate the accuracy of the approximations. The parameter values we use are from Cheridito, Filipović, and Kimmel (2007).

The state variable process in the $A_0(2)$ model (under risk-neutral probabilities) is given by:

$$d\begin{bmatrix} X_t^{(1)} \\ X_t^{(2)} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\begin{bmatrix} X_t^{(1)} \\ X_t^{(2)} \end{bmatrix} dt + d\begin{bmatrix} W_t^{(1)} \\ W_t^{(2)} \end{bmatrix} \tag{6.1}$$

with the interest rate process given by:

$$r_t = d_0 + d_1 X_t^{(1)} + d_2 X_t^{(2)} \tag{6.2}$$

In other words, the risk-neutral state process is a multivariate Ornstein-Uhlenbeck process, and the instantaneous interest rate is an affine function of the state process. The choice of instantaneously uncorrelated state variables, each with normalized variance, is not restrictive, since an affine change of the state vector can convert a general two-dimensional Ornstein-Uhlenbeck process into this form; see Dai and Singleton (2000). Bond prices $P(\Delta, x_1, x_2)$ then satisfy a pricing PDE with final condition:

$$\frac{\partial P}{\partial \Delta}(\Delta, x_1, x_2) = \begin{bmatrix} (b_{11}x_1 + b_{12}x_2)\dfrac{\partial P}{\partial x_1}(\Delta, x_1, x_2) + (b_{21}x_1 + b_{22}x_2)\dfrac{\partial P}{\partial x_2}(\Delta, x_1, x_2) \\ +\dfrac{1}{2}\dfrac{\partial^2 P}{\partial x_1^2}(\Delta, x_1, x_2) + \dfrac{1}{2}\dfrac{\partial^2 P}{\partial x_2^2}(\Delta, x_1, x_2) - (d_0 + d_1 x_1 + d_2 x_2) P(\Delta, x_1, x_2) \end{bmatrix} \tag{6.3}$$

$$P(0, x_1, x_2) = 1 \tag{6.4}$$

The general PDE effectively states that the expected instantaneous rate of return on the bond, under risk-neutral probabilities, is equal to the instantaneous risk-free rate. The final condition simply specifies that the bond price must reach face value at maturity. The following change of dependent variable:

$$P(\Delta, x_1, x_2) = \exp\left(\gamma_1 x_1 + \gamma_2 x_2 - \left(d_0 - \frac{\gamma_1^2 + \gamma_2^2}{2}\right)\Delta\right) f(\Delta, x_1, x_2) \tag{6.5}$$

$$\gamma_1 = \frac{b_{22}d_1 - b_{21}d_2}{b_{11}b_{22} - b_{12}b_{21}} \tag{6.6}$$

$$\gamma_2 = \frac{b_{11}d_2 - b_{12}d_1}{b_{11}b_{22} - b_{12}b_{21}} \tag{6.7}$$

changes the PDE to the canonical form:

$$\frac{\partial f}{\partial \Delta}(\Delta, x_1, x_2) = \begin{bmatrix} (\gamma_1 + b_{11}x_1 + b_{12}x_2)\dfrac{\partial f}{\partial x_1}(\Delta, x_1, x_2) + (\gamma_2 + b_{21}x_1 + b_{22}x_2)\dfrac{\partial f}{\partial x_2}(\Delta, x_1, x_2) \\ +\dfrac{1}{2}\dfrac{\partial^2 f}{\partial x_1^2}(\Delta, x_1, x_2) + \dfrac{1}{2}\dfrac{\partial^2 f}{\partial x_2^2}(\Delta, x_1, x_2) \end{bmatrix} \tag{6.8}$$

$$f(0, x_1, x_2) = \exp(-\gamma_1 x_1 - \gamma_2 x_2) \tag{6.9}$$

The original bond pricing problem was one of finding the conditional expectation of the final payoff (equal to one) discounted back to the present at the risk-free rate of return, where the instantaneous interest rate is an affine function of a set of variables that follow an Ornstein-Uhlenbeck process. By change of variables, this problem has effectively been converted into the problem of finding the conditional expectation of an exponential function of the state vector, where the state variables follow a *different* state variable process.

The next change of variables eliminates the first spatial derivatives from the PDE:

$$f(\Delta, x_1, x_2) = \phi(\Delta, y_1(\Delta, x_1, x_2), y_2(\Delta, x_1, x_2)) \tag{6.10}$$

$$y_1(\Delta, x_1, x_2) = (c_{10} + c_{11}x_1 + c_{12}x_2)\, e^{\lambda_1 \Delta} \tag{6.11}$$

$$y_2(\Delta, x_1, x_2) = (c_{20} + c_{21}x_1 + c_{22}x_2)\, e^{\lambda_2 \Delta} \tag{6.12}$$

where:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \tag{6.13}$$

$$c_{10} = \frac{c_{11}b_{22}\gamma_1 - c_{11}b_{12}\gamma_2 - c_{12}b_{21}\gamma_1 + c_{12}b_{11}\gamma_2}{b_{11}b_{22} - b_{12}b_{21}} \tag{6.14}$$

$$c_{20} = \frac{c_{22}b_{11}\gamma_2 - c_{22}b_{21}\gamma_1 - c_{21}b_{12}\gamma_2 + c_{21}b_{22}\gamma_1}{b_{11}b_{22} - b_{12}b_{21}} \tag{6.15}$$

Substituting the right-hand side of (6.10) into (6.8) in place of $f(\Delta, x_1, x_2)$, we find that $\phi(\Delta, y_1, y_2)$ must satisfy the following PDE with final condition:

$$\frac{\partial \phi}{\partial \Delta}(\Delta, z_1, z_2) = \begin{bmatrix} \frac{c_{11}^2 + c_{12}^2}{2} e^{2\lambda_1 \Delta} \frac{\partial^2 \phi}{\partial z_1^2}(\Delta, z_1, z_2) \\ + (c_{11}c_{21} + c_{12}c_{22})\, e^{(\lambda_1 + \lambda_2)\Delta} \frac{\partial^2 \phi}{\partial z_1 \partial z_2}(\Delta, z_1, z_2) \\ + \frac{c_{21}^2 + c_{22}^2}{2} e^{2\lambda_2 \Delta} \frac{\partial^2 \phi}{\partial z_2^2}(\Delta, z_1, z_2) \end{bmatrix} \tag{6.16}$$

$$\phi(0, z_1, z_2) = \exp\left(-\gamma_1 \frac{c_{12}c_{20} - c_{10}c_{22} + c_{22}z_1 - c_{12}z_2}{c_{11}c_{22} - c_{12}c_{21}} - \gamma_2 \frac{c_{11}c_{20} - c_{10}c_{21} + c_{21}z_1 - c_{11}z_2}{c_{11}c_{22} - c_{12}c_{21}} \right) \tag{6.17}$$

As in Section 4, the change of independent variables has transformed the canonical form PDE into a different PDE with no first-order spatial derivatives, but time-dependent coefficients on the second-order spatial derivatives. Continuing as in Section 4, we now introduce three time variables:

$$\phi(\Delta, z_1, z_2) = w(\tau_{11}(\Delta), \tau_{12}(\Delta), \tau_{22}(\Delta), z_1, z_2) \tag{6.18}$$

$$\tau_{11}(\Delta) = \begin{bmatrix} \frac{e^{2\lambda_1 \Delta} - 1}{2\lambda_1} & \text{if} & \lambda_1 \neq 0 \\ \Delta & \text{if} & \lambda_1 = 0 \end{bmatrix} \tag{6.19}$$

$$\tau_{12}(\Delta) = \begin{bmatrix} \frac{e^{(\lambda_1 + \lambda_2)\Delta} - 1}{\lambda_1 + \lambda_2} & \text{if} & \lambda_1 + \lambda_2 \neq 0 \\ \Delta & \text{if} & \lambda_1 + \lambda_2 = 0 \end{bmatrix} \tag{6.20}$$

$$\tau_{22}(\Delta) = \begin{bmatrix} \frac{e^{2\lambda_2 \Delta} - 1}{2\lambda_2} & \text{if} & \lambda_2 \neq 0 \\ \Delta & \text{if} & \lambda_2 = 0 \end{bmatrix} \tag{6.21}$$

These three time variables effectively restore the time-homogeneity in the coefficients of (6.16), which can now be restated as the following system of PDEs:

$$\frac{\partial w}{\partial \tau_{11}}\left(\tau_{11},\tau_{12},\tau_{22},z_1,z_2\right)=\frac{c_{11}^2+c_{12}^2}{2}\frac{\partial^2\phi}{\partial z_1^2}\left(\Delta,z_1,z_2\right) \tag{6.22}$$

$$\frac{\partial w}{\partial \tau_{12}}\left(\tau_{11},\tau_{12},\tau_{22},z_1,z_2\right)=\left(c_{11}c_{21}+c_{12}c_{22}\right)\frac{\partial^2\phi}{\partial z_1\partial z_2}\left(\Delta,z_1,z_2\right) \tag{6.23}$$

$$\frac{\partial w}{\partial \tau_{22}}\left(\tau_{11},\tau_{12},\tau_{22},z_1,z_2\right)=\frac{c_{21}^2+c_{22}^2}{2}\frac{\partial^2\phi}{\partial z_2^2}\left(\Delta,z_1,z_2\right) \tag{6.24}$$

$$w\left(0,0,0,z_1,z_2\right)=\exp\begin{pmatrix}-\gamma_1\dfrac{c_{12}c_{20}-c_{10}c_{22}+c_{22}z_1-c_{12}z_2}{c_{11}c_{22}-c_{12}c_{21}}\\ -\gamma_2\dfrac{c_{11}c_{20}-c_{10}c_{21}+c_{21}z_1-c_{11}z_2}{c_{11}c_{22}-c_{12}c_{21}}\end{pmatrix} \tag{6.25}$$

Existence of a solution to the system of PDEs implies existence of a solution to the original PDE. Furthermore, the system satisfies the conditions of Theorem 2, so there is a solution that is everywhere analytic in τ_{11}, τ_{12}, and τ_{22}, and which therefore has a convergent power series representation.

We use the estimated (risk-neutral) parameters of Cheridito, Filipović, and Kimmel (2007) to examine the accuracy of the approximations. Specifically, we have:

$$b_{11}=-0.1607 \qquad b_{12}=-1.3462 \tag{6.26}$$

$$b_{21}=-0.1741 \qquad b_{22}=-1.9840 \tag{6.27}$$

$$d_0=0.0562 \qquad d_1=0.0194 \tag{6.28}$$

$$d_2=0.0178 \tag{6.29}$$

For the initial values of the state variables, we choose $x_1=0.3$ and $x_2=-0.2$, although it should be noted that, for this model, the relative accuracy of approximations by our method does not depend on the initial values of the state variables. It should also be noted that with these parameters, the risk-neutral state variable process is stationary, and bond prices tend to a limit. However, there is a slowly mean-reverting linear combination of the state variables; specifically, one of the eigenvalues λ_1 or λ_2 has a value of approximately -0.0401. Because of this slowly mean-reverting component, more terms are required for the level of accuracy obtained in the examples in Kimmel (2008b) and Kimmel (2008a). However, it is still possible to construct approximations with a reasonable number of terms that have very high level of accuracy.

Figure 2 compares exact zero-coupon bond prices in this model to power series approximations *directly* in Δ, that is, without using our time transformation technique. The approximations include seven terms, that is, terms up to order of the sixth power of Δ. As shown, the approximations perform well for very short maturities, but become extremely inaccurate for maturities much beyond one year. At a maturity of one year, the relative pricing error (the difference between the approximate price and the true price, divided by the true price), is -0.00015, which corresponds to a yield error of about 1.5 basis points. However, at a maturity of two years, the relative pricing error increases to -0.0159, corresponding to a yield error of about 80 basis points. At a maturity of three years, the seven-term approximations are essentially worthless, as the relative pricing error increases to -0.236, corresponding to a pricing error of 900 basis points. Adding more terms to the approximation improves the accuracy, but for very long maturities, the number of terms required is very large, and makes direct approximation of bond prices in Δ completely impractical.

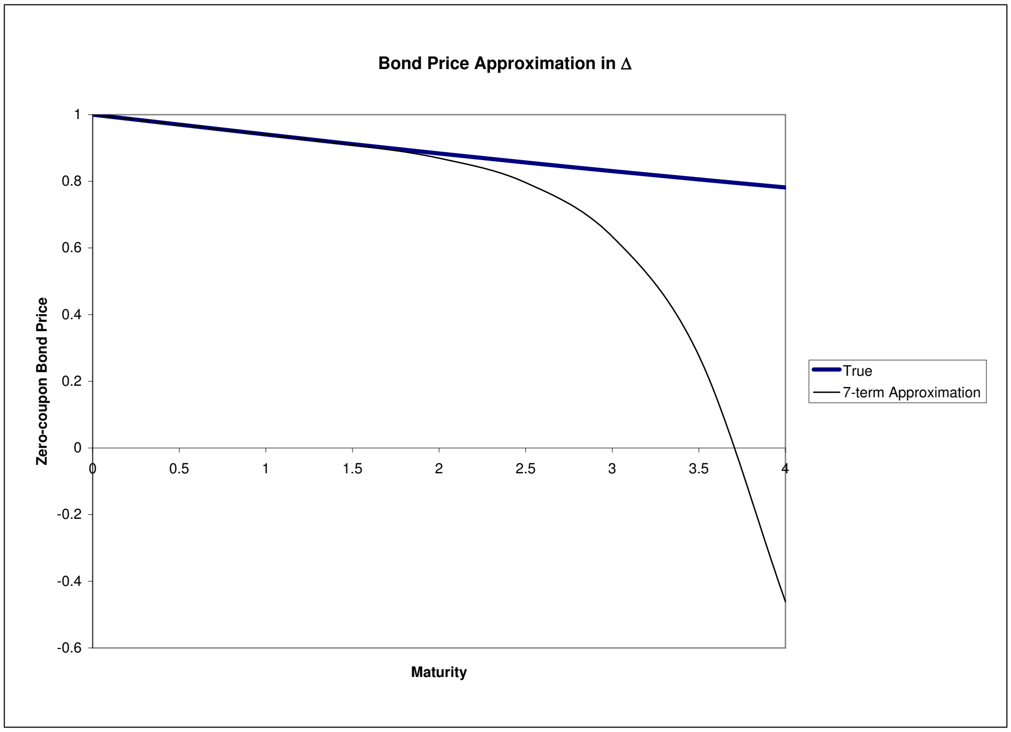

Figure 2: This figure shows exact and approximate bond prices in the $A_0\,(2)$ model. The approximations are obtained using a seven-term power series in Δ (including terms of order up to the sixth power of Δ), without using the method of time transformations. As shown, the approximations before well for short maturities, but begin to diverge badly from the true bond prices for maturities around two years and longer. Although power series approximations in Δ converge for all maturities in this model, the convergence is not uniform in Δ, is very slow for long maturities, and requires a very large number of terms to be accurate for long maturities.

The situation is very different using our the time transformation methods. Figure 3 shows exact zero-coupon bond prices and power series approximations using time transformation methods, also with seven terms (that is, including terms of order up to the sixth power of τ_{11}, τ_{12}, and τ_{22}). As shown, the approximations are very accurate (in the graph, the differences are smaller than the thickness of the line) for maturities of up to twenty years. (Although not shown, the approximations are accurate for arbitrarily large maturities, the pricing error reaching a limit with increasing maturity.) At a maturity of 20 years, the relative pricing error is -0.00038, corresponding to a yield error of less than two-tenths of a basis point.

It should be noted that seven terms were used in the approximations based on our time transformation method because of the slowly mean-reverting component in the state variable process. If the b_{11} parameter were doubled (which increases the speed of mean-reversion on the slowly mean-reverting component), the approximations using the time transformation method result in yield errors of less than two-tenths of a basis point for all maturities, using just two-term approximations (i. e., including terms of up to order one in τ_{11}, τ_{12}, and τ_{22}). More terms are needed when there is a slowly mean-reverting component, but it should be

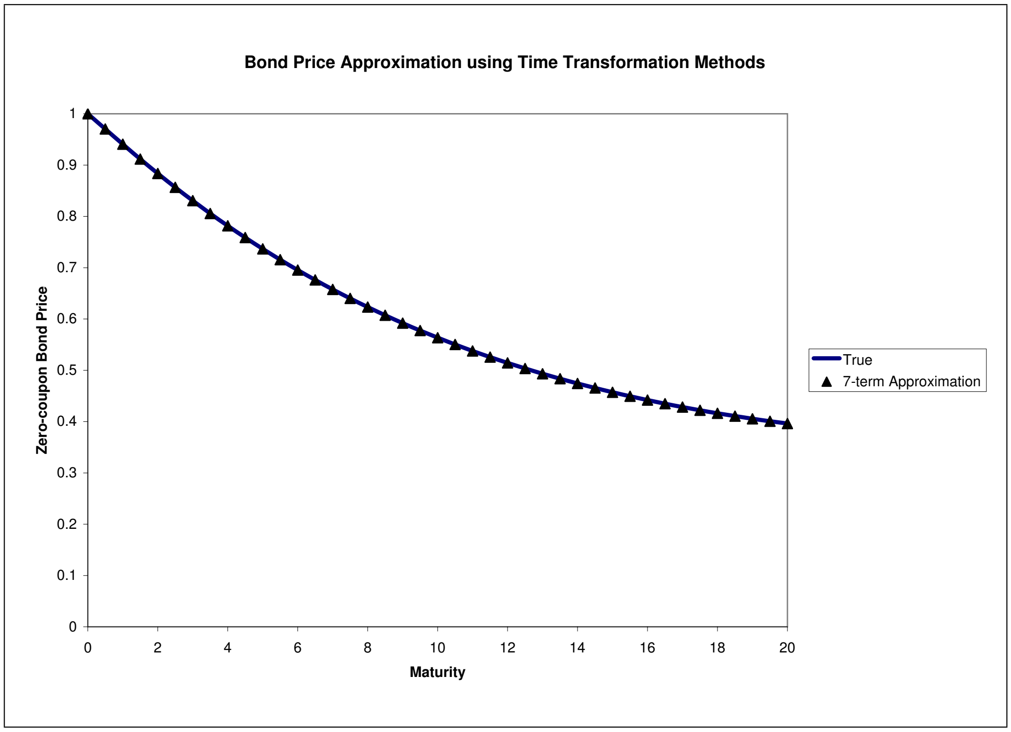

Figure 3: This figure shows exact and approximate bond prices in the $A_0(2)$ model. The approximations are obtained using a seven-term power series in τ_{11}, τ_{12}, and τ_{22} (including terms of order up to the sixth power), using the method of time transformations. As shown, the approximations are very accurate for all maturities.

noted that, for example, in the sixth order approximations used to construct the graph, not *all* terms need to be included for the given level of accuracy. As the slowly mean-reverting component is associated with only one eigenvalue of the (transformed) instantaneous variance/covariance matrix of the state variables, only higher-order terms in the time variables associated with this component need be included. For example, if we assign $\lambda_1 = -0.0041$, then the approximations are nearly as accurate if only sixth-order terms in τ_{11}, third-order terms in τ_{12}, and first-order terms in τ_{22} are included.

7. Conclusion

In this paper, we have developed techniques for the approximation of solutions to a large class of non-affine multifactor conditional moment and contingent claims pricing problems. Following Kimmel (2008c) and Kimmel (2008b), who consider only scalar diffusion problems, we transform the problem of finding the conditional moment or contingent claims price in the non-affine model into the problem of finding a *different* conditional moment in a multifactor affine model. As in the scalar case, power series approximations to the transformed problem can be found by a simple recursive relation, and can be shown to converge if the final condition in the conditional moment problem satisfies appropriate smoothness and growth conditions. However, the *uniform* convergence results established by Kimmel (2008b) in the scalar case do not carry over easily to the multifactor case. To derive analogous results in a multifactor setting, we introduce the abstract formalism of multi-dimensional time, and embed the problem from the setting of real, single-dimensional time, in this abstract setting. We further derive approximation and convergence results, including results of uniform convergence are arbitrarily long (true) time horizons, and recover the solution to the real-world problem by restricting the solution to the abstract problem to a curve representing the path of true time through the abstract multi-dimensional time setting. We also present a method for construction of a large family of non-affine term structure models for which bond prices can be approximated with uniform (in maturity) accuracy. Finally, using examples, we show that our technique is easy to apply and typically highly accurate.

Several possible avenues for future research remain open. Although the class of non-affine problems covered by our technique is broad, it could possibly be larger still. In particular, in the scalar setting, Kimmel (2008c) and Kimmel (2008b) explore problems that, after change of dependent variable, transform to problems based on either an Ornstein-Uhlenbeck process or a Feller square-root process. We consider only the multifactor extension of the former; it is possible that our technique could be extended to the latter as well. Care must be taken in derivation of the power series in a multi-dimensional time setting, because the right-hand side operators corresponding to multifactor Feller square-root processes do not necessarily commute. However, these difficulties can likely be overcome. Far more formidable, but also far more rewarding if it can be achieved, is the possibility of approximating solutions to *any* conditional moment or contingent claim pricing problem in a general multifactor setting. Ability to solve the general problem will almost certainly require consideration of non-analytic (at time of zero) solutions, but if the nature of the non-analyticity can be characterized accurately enough, perhaps it can be achieved. These problems are left for future work.

References

Ahn, D.-H., R. F. Dittmar, and A. R. Gallant (2002): "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, 15, 243–288.

Aït-Sahalia, Y. (2002): "Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach," *Econometrica*, 70, 223–262.

Cheridito, P., D. Filipović, and R. L. Kimmel (2007): "Market Price of Risk Specifications for Affine Models: Theory and Evidence," *Journal of Financial Economics*, 83, 123–170.

Colton, D. (1979): "The Approximation of Solutions to the Backwards Heat Equation in a Nonhomogeneous Medium," *Journal of Mathematical Analysis and Applications*, 72, 418–429.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–408.

Dai, Q., and K. J. Singleton (2000): "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943–1978.

——— (2002): "Expectations puzzles, time-varying risk premia, and affine models of the term structure," *Journal of Financial Economics*, 63, 415–441.

Duffee, G. R. (2002): "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, 405–443.

Duffie, D., and R. Kan (1996): "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 379–406.

Jarrow, R., H. Li, S. Liu, and C. Wu (2006): "Reduced-Form Valuation of Callable Corporate Bonds: Theory and Evidence," working paper.

Kimmel, R. (2008a): "Bond Pricing in Non-Affine Term Structure Models," Ohio State University working paper.

——— (2008b): "Changing Times: Accurate Solutions to Pricing and Conditional Moment Problems in Non-Affine Continuous-Time Models," Ohio State University working paper.

——— (2008c): "Complex Times: Asset Pricing and Conditional Moments under Non-Affine Diffusions," Ohio State University working paper.

Levendorskii, S. (2004a): "Consistency Conditions for Affine Term Structure Models," *Stochastic Processes and Their Applications*, 109, 225–261.

——— (2004b): "Consistency Conditions for Affine Term Structure Models II: Option Pricing under Diffusions with Embedded Jumps," University of Texas working paper.

A. Appendix

A.1. Proof of Theorem 1

We show that the solution is given by:

$$f(\tau, y) = \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} g\left(y + T^{\frac{1}{2}} u\right) du \tag{A.1}$$

where $T^{\frac{1}{2}}$ is the diagonal matrix:

$$T^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\tau_{11}} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \sqrt{\tau_{ii}} & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \sqrt{\tau_{NN}} \end{bmatrix} \tag{A.2}$$

where, for each diagonal element, either square root (for $\tau_{ii} \neq 0$) is chosen. As will soon be demonstrated, the value of the integral does not depend on which square roots are chosen. It must be shown that the integral converges, and that the solution is analytic in both τ and y.

Using the bound on $g(y)$, we can establish the existence of the integral:

$$\begin{aligned} |f(\tau, y)| &\leq \left| \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} g\left(y + T^{\frac{1}{2}} u\right) du \right| \\ &\leq \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} \left| g\left(y + T^{\frac{1}{2}} u\right) \right| du \\ &\leq \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} c e^{\sum_{i=1}^N \frac{\|y_i + u_i \sqrt{\tau_{ii}}\|_i^2}{2}} du \\ &\leq \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} c e^{\sum_{i=1}^N \frac{u_i^2\left(\|\sqrt{\tau_{ii}}\|_i^2 - 1\right) + 2|u_i| \|y_i\| \|\sqrt{\tau_{ii}}\| + \|y_i\|_i^2}{2}} du \end{aligned} \tag{A.3}$$

If $\left\|\sqrt{\tau_{ii}}\right\| < 1$ for all $1 \leq i \leq N$, the coefficients of the u_i^2 terms in the exponent of the integrand all have negative real parts. The integral therefore converges. To show that it is analytic in y, we note that the integrand is analytic in y and continuous in u. The integrand is also uniformly bounded, and the integral therefore uniformly convergent, on compact sets of y. The integral therefore inherits the analyticity of the integrand in y.

Analyticity in τ is somewhat more subtle. Choosing some $i \leq i \leq N$, we introduce:

$$T_i^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\tau_{11}} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & -\sqrt{\tau_{ii}} & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \sqrt{\tau_{NN}} \end{bmatrix} \tag{A.4}$$

In other words, $T_i^{\frac{1}{2}}$ is the same as $T^{\frac{1}{2}}$, except that the ith diagonal element is negated. Then:

$$\begin{aligned} f(\Delta, y) &= \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} g\left(y + T^{\frac{1}{2}} u\right) du \\ &= \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} \frac{g\left(y + T^{\frac{1}{2}} u\right) + g\left(y - T_i^{\frac{1}{2}} u\right)}{2} du \\ &+ \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} \frac{g\left(y + T^{\frac{1}{2}} u\right) - g\left(y - T_i^{\frac{1}{2}} u\right)}{2} du \end{aligned} \tag{A.5}$$

In the second integral of the last expression, the integrand is odd in u_i. This integral therefore evaluates to zero. In the first integral in the last expression, the integrand is even in $\sqrt{\tau_{ii}}$, and therefore analytic in τ_{ii}; this follows from the power series expansion of $g(y)$. Since the same argument can be applied for each $1 \leq i \leq N$, $f(\tau, y)$ is analytic in all τ_{ii}, $1 \leq i \leq N$.

A.2. Proof of Corollary 1

A solution to (3.35) and (3.36) exists by Theorem 1. Then $f(\Delta, x)$ is well-defined for all x and all Δ such that $\left\|\sqrt{\tau_{ii}}(\Delta)\right\|_i < 1$, $1 \leq i \leq N$. Furthermore, $\tau(\Delta)$ is everywhere analytic in Δ, and $y(\Delta, x)$ is everywhere analytic in Δ and x. Since $w(\tau, y)$ is analytic in y and τ for $\left\|\sqrt{\tau_{ii}}\right\|_i < 1$, the function $f(\Delta, x)$ is therefore an analytic function of x and Δ for all values such that $\left\|\sqrt{\tau_{ii}(\Delta)}\right\|_i < 1$. The partial derivatives of $f(\Delta, x)$ with respect to Δ and x are given by:

$$\frac{\partial f}{\partial \Delta}(\Delta, x) = \sum_{i=0}^{N} \left[e^{2\lambda_i \Delta} \frac{\partial w}{\partial \tau_{ii}}(\tau, y) + \lambda_i x_i e^{\lambda_i \Delta} \frac{\partial w}{\partial y_i}(\tau, y) \right] \tag{A.6}$$

$$\frac{\partial f}{\partial x_i}(\Delta, x) = e^{\lambda_i \Delta} \frac{\partial w}{\partial y_i}(\tau, y) \tag{A.7}$$

$$\frac{\partial^2 f}{\partial x_i^2}(\Delta, x) = e^{2\lambda_i \Delta} \frac{\partial^2 w}{\partial y_i^2}(\tau, y) \tag{A.8}$$

Substituting of these derivatives into (3.30), one finds immediately that $f(\Delta, x)$ is a solution. Furthermore, $f(\Delta, x)$ satisfies the final condition:

$$f(0, x) = w(0, y(0, x)) = g(x) \tag{A.9}$$

A.3. Proof of Theorem 2

We show that the solution is given by:

$$f(\tau, y) = \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} g(y + Tu)\, du \tag{A.10}$$

where T is any matrix such that $TT^T = \Omega$, with the elements of the $N \times N$ matrix Ω given by $\Omega_{ij} = d_{ij}\tau_{ij}$, $i \leq i, j \leq N$. We first show that the integral converges, does not depend on the choice of T, and is an analytic function of the elements of Ω.

First, we note that, for any y and T in compact sets, there exist d_0, d_1, and d_2 such that:

$$\left|(y+Tu)^*(y+Tu)\right| \leq d_0 + d_1\,|u| + d_2 u^T u \tag{A.11}$$

But then there exists a bound of the form:

$$|g(y+Tu)| \leq d_k e^{\frac{(y+Tu)^*(y+Tu)}{2k}} \leq d_k e^{\frac{d_0+d_1|u|+d_2u^Tu}{2k}} \tag{A.12}$$

for any arbitrary $k > 0$. The integrand of (A.10) then satisfies the bound:

$$\left|\frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^Tu}{2}} g(y+Tu)\right| \leq \frac{d_k}{(2\pi)^{\frac{N}{2}}} e^{\frac{d_0+d_1|u|+(d_2-k)u^Tu}{2k}} \tag{A.13}$$

Since we can choose $k > d_2$, the integrand is uniformly bounded, and the integral converges. Furthermore, since the integrand is analytic in y and T, and continuous in u, the integral inherits the analyticity of y and T. Since every combination of y and T is in some compact set, the integral is everywhere in analytic in y and T.

So far, we have established the existence of the proposed solution for all values of y and T, and shown that it is analytic. It remains to show that the proposed solution is uniquely defined (that is, does not depend on the particular choice of T as long as $TT^T = \Omega$), and that is solve the PDE with final condition. We do so by representing the function $g(\bullet)$ as a power series, and performing the integration term-by-term. We therefore first consider the special case in which $g(z)$ has the form:

$$g(z) = g_{i_1\cdots i_N}(z) \equiv \prod_{j=1}^{N} z_j^{i_j} \tag{A.14}$$

with $f_{i_1\cdots i_N}(\tau, y)$ denoting the corresponding PDE solution. Since $g_{0\cdots 0}(z) = 1$:

$$f_{0\cdots 0}(\tau, y) = \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^Tu}{2}} du = 1 \tag{A.15}$$

One verifies immediately that $f_{0\cdots 0}(\tau, y)$ is a solution to the general PDE satisfying $f_{0\cdots 0}(0, y) = 1$; it is also trivially an analytic function of the elements of τ. We demonstrate the same properties for general $f_{i_1\cdots i_N}(\tau, y)$ by induction. We take $i_N \neq 0$; if any i_j are non-zero, we can reorder the indices so that i_N is not zero. Then using integration by parts:

$$\begin{aligned}
f_{i_1\cdots i_N+1}(\tau, y) &= \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^Tu}{2}} g_{i_1\cdots i_N+1}(y+Tu)\,du \\
&= \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^Tu}{2}} (y_N + T_N u)\, g_{i_1\cdots i_N}(y+Tu)\,du \\
&= y_N f_{i_1\cdots i_N}(\tau, y) + \sum_{k=1}^{N} T_{Nk} \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^Tu}{2}} \frac{\partial g_{i_1\cdots i_N}}{\partial u_k}(y+Tu)\,du \\
&= y_N f_{i_1\cdots i_N}(\tau, y) + \sum_{k=1}^{N}\sum_{l=1}^{N} T_{Nk}T_{lk} i_l \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^Tu}{2}} i_k g_{i_1\cdots i_l-1\cdots i_N}(y+Tu)\,du \\
&= y_N f_{i_1\cdots i_N}(\tau, y) + \sum_{l=1}^{N} T_N T_l^T i_l f_{i_1\cdots i_l-1\cdots i_N}(\tau, y) \\
&= y_N f_{i_1\cdots i_N}(\tau, y) + \sum_{l=1}^{N} i_l \tau_{Nl} f_{i_1\cdots i_l-1\cdots i_N}(\tau, y)
\end{aligned} \tag{A.16}$$

Thus, by induction, it is clear that all $f_{i_1\cdots i_N}(\tau, y)$ are well-defined (that is, the integral converges and does not depend on the particular choice of τ), and are everywhere analytic functions of the elements of τ and y. We show that the $f_{i_1\cdots i_N}(\tau, y)$ solve the general PDE with final condition as well. To this end, we define the operators:

$$L_{ij} = -\frac{\partial}{\partial \tau_{ij}} + d_{ij}\frac{\partial^2}{\partial y_i \partial y_j} \tag{A.17}$$

A function $f(\tau, y)$ solves the system of PDEs if $L_{ij}f(\tau, y) = 0$ for all $1 \le i, j \le N$. If $i \ne N$ and $j \ne N$, applying L_{ij} to both sides of (A.16), we have:

$$L_{ij}f_{i_1\cdots i_N+1}(\tau, y) = y_N L_{ij} f_{i_1\cdots i_N}(\tau, y) + \sum_{l=1}^{N} i_l \tau_{Nl} L_{ij} f_{i_1\cdots i_l-1\cdots i_N}(\tau, y) = 0 \tag{A.18}$$

If $i = j = N$, then:

$$\begin{aligned} L_{NN}f_{i_1\cdots i_N+1}(\tau, y) =& 2\frac{\partial f_{i_1\cdots i_N}}{\partial y_N}(\tau, y) \\ &+ y_N L_{NN} f_{i_1\cdots i_N}(\tau, y) \\ &- i_N f_{i_1\cdots i_N-1}(\tau, y) \\ &+ \sum_{l=1}^{N} i_l \tau_{Nl} L_{NN} f_{i_1\cdots i_l-1\cdots i_N}(\tau, y) = 0 \end{aligned} \tag{A.19}$$

If $i = N$ and $j \ne N$, then:

$$\begin{aligned} L_{Nj}f_{i_1\cdots i_N+1}(\tau, y) =& \frac{\partial f_{i_1\cdots i_N}}{\partial y_j}(\tau, y) \\ &+ y_N L_{ij} f_{i_1\cdots i_N}(\tau, y) \\ &- i_j f_{i_1\cdots i_j-1\cdots i_N}(\tau, y) \\ &+ \sum_{l=1}^{N} i_l \tau_{Nl} L_{Nj} f_{i_1\cdots i_l-1\cdots i_N}(\tau, y) = 0 \end{aligned} \tag{A.20}$$

The case of $i \ne N$ and $j = N$ is exactly analogous. It follows by induction that $f_{i_1\cdots i_N}(\tau, y)$ solves the system of PDEs. Since for $\tau = 0$, (A.16) simplifies to:

$$f_{i_1\cdots i_N+1}(0, y) = y_N f_{i_1\cdots i_N}(0, y) \tag{A.21}$$

it also follows by induction that $f_{i_1\cdots i_N+1}(0, y) = g_{i_1\cdots i_N+1}(y)$.

The theorem is now proven for the restricted special case when $g(z)$ takes the form $g_{i_1\cdots i_N}(z)$. But recall that the integrand in the definition of the general solution $f(\tau, y)$ is uniformly convergent on compact sets of τ and y. Therefore, if the general $g(z)$ is represented by the power series:

$$g(z) = \sum_{i_1=0}^{\infty} \cdots \sum_{i_N=0}^{\infty} a_{i_1\cdots i_N} g_{i_1\cdots i_N}(z) \tag{A.22}$$

then the integral of (A.10) can be evaluated term-by-term, so that the solution is:

$$f(\tau, y) = \sum_{i_1=0}^{\infty} \cdots \sum_{i_N=0}^{\infty} a_{i_1\cdots i_N} f_{i_1\cdots i_N}(\tau, y) \tag{A.23}$$

This power series is uniformly convergent on compact sets of y and τ, so the sum inherits the properties of the individual terms; it is an analytic function of τ and y. Furthermore, since the system of PDEs is linear, $f(\tau, y)$ is a solution to the system, with $f(0, y) = g(y)$.

A.4. Proof of Theorem 3

We show that the solution is given by:

$$f(\tau, y) = \int_0^{+\infty} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{r^2}{2}} \int_{u^T u = r^2} g(y + Tu)\, dS_N(r) dr \tag{A.24}$$

where T is any matrix such that:

$$TT^T = \begin{bmatrix} \tau_{11} & \tau_{12} & \cdots & \cdots & \tau_{1N} \\ \tau_{12} & \ddots & & & \vdots \\ \vdots & & \tau_{ii} & & \vdots \\ \vdots & & & \ddots & \tau_{N-1N} \\ \tau_{1N} & \cdots & \cdots & \tau_{N-1N} & \tau_{NN} \end{bmatrix} \tag{A.25}$$

where $dS_N(r)$ is the surface measure of a sphere of radius r. We first show that the inner integral converges, does not depend on the choice of T, and is an analytic function of the elements of TT^T.

Evaluating the two integrals in the order indicated is important. It might be tempting to apply Fubini's theorem and write a single integral over $\mathbb{R}^N$:

$$f(\tau, y) = \int_{\mathbb{R}^N} \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{u^T u}{2}} g(y + Tu)\, du \tag{A.26}$$

However, convergence of the integral can then depend on the choice of T. For example, if we take $N = 2$ and:

$$g(y) = e^{\frac{y^T y}{2}} \tag{A.27}$$

$$T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{A.28}$$

We then have $f(\tau, y) = e^{\frac{y^T y}{2}}$ when all elements of τ are zero. However, if we choose:

$$T = \begin{bmatrix} 2 & 2\imath \\ 0 & 0 \end{bmatrix} \tag{A.29}$$

then the single integral expression for $f(\tau, y)$ does not even converge, let alone have the same value as with the previous choice of T.

First, we note that $g(y)$ is a continuous function, so that, for any choice of T and r, $g(y + Tu)$ is bounded on $u^T u = r^2$. The inner integral therefore converges. Consider the power series representation of $g(z)$ around $z = y$:

$$g(z) = \sum_{i_1=1}^{+\infty} \cdots \sum_{i_N=1}^{+\infty} a_{i_1 \cdots i_N} \prod_{j=1}^{N} (z_j - y_j)^{i_j} \tag{A.30}$$

Since $g(z)$ is everywhere analytic, this series converges uniformly on any compact set of z. The inner integral may therefore be applied term by term. Consider a single term. If the i_j are all zero, then the integral evaluates to the surface area of an N-dimensional hypersphere with radius r. If at least one i_j is non-zero (by reordering of the indices, we take it to be i_N), we can apply integration by parts:

$$\int_{u^T u=r^2} a_{i_1 \cdots i_N} \prod_{j=1}^{N} (T_j u)^{i_j} dS_N(r) = a_{i_1 \cdots i_N} \sum_{k=1}^{N} \int_{u^T u=r^2} \left[(T_N u)^{i_j - 1} \prod_{j=1}^{N-1} (T_j u)^{i_j} \right] T_{Nk} u_k dS_N(r) \qquad \text{(A.31)}$$

Thus, by induction, every term in the power series representation integrates to a function that is analytic in the elements of TT^T. The sum of these terms is therefore also analytic.