PROCEEDINGS OF THE ROYAL SOCIETY A

Proc. R. Soc. A (2005) **461**, 459–479
doi:10.1098/rspa.2004.1366
Published online 16 November 2004

Wiener chaos and the Cox–Ingersoll–Ross model

By M. R. Grasselli and T. R. Hurd

Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1, Canada (grasselli@math.mcmaster.ca)

In this paper we recast the Cox–Ingersoll–Ross (CIR) model of interest rates into the chaotic representation recently introduced by Hughston and Rafailidis. Beginning with the 'squared Gaussian representation' of the CIR model, we find a simple expression for the fundamental random variable X_∞. By use of techniques from the theory of infinite-dimensional Gaussian integration, we derive an explicit formula for the nth term of the Wiener chaos expansion of the CIR model, for $n = 0, 1, 2, \dots$. We then derive a new expression for the price of a zero coupon bond which reveals a connection between Gaussian measures and Ricatti differential equations.

Keywords: interest-rate models; Wiener chaos; functional integrals; squared Gaussian models

1. Introduction

In this paper we will study the best-known example of a term structure with positive interest rates, namely the Cox–Ingersoll–Ross (CIR) model (Cox *et al.* 1985), in the context of the 'chaotic approach' to interest-rate dynamics introduced recently by Hughston & Rafailidis (2005) (see also Brody & Hughston 2004). By an interest-rate model we mean the specification of a spot-rate process r_t and of a market price of risk process λ_t both under the 'natural' or physical measure in the economy P. In the chaotic approach, the random nature of the model is assumed to be given by a probability space $(\Omega, \mathcal{F}, P)$ equipped with a Brownian filtration $(\mathcal{F}_t)_{0\leqslant t\leqslant\infty}$. The essence of the approach by Hughston & Rafailidis is the specification of the most general term structure with positive interest rates in terms of a single unconstrained random variable they denote X_∞. They then apply a Wiener chaos expansion to X_∞, and interpret the resulting terms as building blocks for models of increasing complexity.

In one version of the CIR model, r_t is governed by the equation

$$\mathrm{d}r_t = a(b - r_t)\,\mathrm{d}t + c\sqrt{r_t}\,\mathrm{d}\tilde{W}_t, \tag{1.1}$$

for some positive constants a, b, c with $4ab > c^2$, where $\tilde{W}_t$ is a standard one-dimensional P-Brownian motion, and the market price of risk is taken to be proportional to $\sqrt{r}$. By embedding this model inside a general class of squared Gaussian models, we will be led to a natural choice for the basic square-integrable random variable X_∞ associated with it. As we shall then demonstrate, the resulting stochastic process $X_t = E_t[X_\infty]$ admits an explicit chaos expansion, one which in general includes terms of every chaotic order.

Received 14 July 2003
Revised 28 March 2004
Accepted 9 June 2004

A central idea in the present paper is the link between the Wiener chaos expansion and the theory of Gaussian functional integration—an essential tool invented to study the mathematical structure of quantum field theory. In fact, the mathematics underlying our example is a consequence of certain basic results in that theory, and can be found in, for example, Glimm & Jaffe (1981).

The organization of the paper is as follows. In § 2, we review the essential ingredients for the construction of positive-interest-rate models in both Flesaker–Hughston and the state price density approaches, and then compare these approaches with the recently introduced chaotic representation of Hughston & Rafailidis (2005). We end the section by describing the structure of the Wiener–Itô chaos expansion and show how it can be expressed in terms of a certain generating functional acting on the space $L^2(\mathbb{R}^N)$.

In § 3, we describe the squared Gaussian formulation of the CIR model and show how the spot-rate process can be explicitly computed. Based on this representation, we state the form of the random variable X_∞, and give a proof that it lies in L^2. In § 4, we state the exponential quadratic formula, which is the main technical tool in this paper. It is a formula for the generating functional of random variables of the form $X = \mathrm{e}^{-Y}$ for Y lying in a general class of elements in the second chaos space $\mathcal{H}_2$. In § 5, we compute the generating function for the random variables X_t in the CIR model and, as the main result of the paper, derive their chaos expansion. In § 6, we show that the usual CIR bond-pricing formula has a natural derivation within the chaotic framework.

Three appendices focus on the theory of Gaussian functional integration and its relation to the Wiener chaos expansion. Appendix A explores the white-noise calculus. Appendix B states and provides a proof of the generating functional theorem. Appendix C provides a proof of the exponential quadratic formula.

2. Positive interest rates

(a) *State price density and the potential approach*

Rather than focus on the spot-rate process, one can model the system of bond prices directly. Let P_{tT}, $0 \leqslant t \leqslant T$, denote the price at time t for a zero coupon bond which pays one unit of currency at its maturity T. Clearly, $P_{tt} = 1$ for all $0 \leqslant t < \infty$ and, furthermore, positivity of the interest rate is equivalent to having

$$P_{ts} \leqslant P_{tu}, \tag{2.1}$$

for all $0 \leqslant t \leqslant u \leqslant s$.

A general way to model bond prices (Rogers 1997; Rutkowski 1997) is to write

$$P_{tT} = \frac{E_t[V_T]}{V_t} \tag{2.2}$$

for a positive adapted continuous process V_t, called the *state price density*. Positivity of the interest rates is then equivalent to V_t being a supermartingale. In order to match the initial term structure, this supermartingale needs to be chosen so that $E[V_T] = P_{0T}$. If we further impose that $P_{0T} \to 0$ as $T \to \infty$, then V_t satisfies all the properties of what is known in probability theory as a *potential* (namely, a positive supermartingale with expected value going to zero at infinity).

It follows from the Doob–Meyer decomposition that any continuous potential satisfying

$$E\Big(\sup_{0\leqslant t\leqslant\infty} V_t^2\Big) < \infty \tag{2.3}$$

can be written as

$$V_t = E_t[A_\infty] - A_t \tag{2.4}$$

for a unique (up to indistinguishability) adapted continuous increasing process A_t with $E(A_\infty^2) < \infty$. Therefore, the model is completely specified by the process A_t, which can be freely chosen apart from the constraint that

$$E\left[\frac{\partial A_T}{\partial T}\right] = -\frac{\partial P_{0T}}{\partial T}, \quad \text{for a.a. } T. \tag{2.5}$$

(*b*) *Related quantities and absence of arbitrage*

An earlier framework for positive interest rates was introduced by Flesaker & Hughston (1996, eqn (8)), who observed that any arbitrage-free system of zero coupon bond prices has the form

$$P_{tT} = \frac{\int_T^\infty h_s M_{ts}\,\mathrm{d}s}{\int_t^\infty h_s M_{ts}\,\mathrm{d}s}, \quad \text{for } 0 \leqslant t \leqslant T < \infty. \tag{2.6}$$

Here,

$$h_T = -\frac{\partial P_{0T}}{\partial T}$$

is a positive deterministic function obtained from the initial term structure and M_{ts} is a family of strictly positive continuous martingales satisfying $M_{0s} = 1$. Any such system of prices can be put into a potential form by setting

$$V_t = \int_t^\infty h_s M_{ts}\,\mathrm{d}s. \tag{2.7}$$

The converse result is less direct and was first established by Jin & Glasserman (2001, lemma 1).

These equivalent ways of modelling positive interest rates can now be related to other standard financial objects. A particularly straightforward path is to follow Rutkowski (1997, proposition 1): given a strictly positive supermartingale V_t, there exists a unique strictly positive (local) martingale Λ_t such that the process $B_t = \Lambda_t/V_t$ is strictly increasing and $V_0 = \Lambda_0$. We identify B_t with a risk-free money-market account initialized at $B_0 = 1$ and write it as

$$B_t = \exp\left(\int_0^t r_s\,\mathrm{d}s\right), \tag{2.8}$$

for an adapted process $r_s > 0$, the short-rate process.

A sufficient condition for an arbitrage-free bond price structure in the potential approach is to require that the local martingale Λ_t in fact be a martingale, since it can then be used as the density for an equivalent martingale measure. It is an interesting open question in the theory to isolate what general conditions on the potential V_t would suffice for that.

The formulation up to this point is quite general, in the sense that is does not make use of any particular structure of the underlying filtration (other than the usual conditions). Let us now assume that $(\mathcal{F}_t)_{0\leqslant t\leqslant\infty}$ is actually generated by an N-dimensional Brownian motion W_t. The market price of risk then arises as the adapted vector-valued process λ_t such that

$$\mathrm{d}\Lambda_t = -\Lambda_t \lambda_t^\dagger \,\mathrm{d}W_t, \quad \Lambda_0 = 1, \tag{2.9}$$

where we suppress vector indices by adopting a matrix multiplication convention, including '†' for transpose. It is also immediate to see that the state price density process is the solution to

$$\mathrm{d}V_t = -r_t V_t \,\mathrm{d}t - V_t \lambda_t^\dagger \,\mathrm{d}W_t, \quad V_0 = 1, \tag{2.10}$$

so that the specification of the process V_t is enough to produce both the short rate r_t and the market price of risk λ_t.

It has already been remarked by Flesaker & Hughston (1996) that in the Brownian filtration with finite time horizon any positive-interest-rate model in their formulation corresponds to a model in the Heath–Jarrow–Morton (HJM) family with positive instantaneous forward rates f_{tT}. The converse result that any interest-rate model in HJM form with positive instantaneous forward rates can be written in the Flesaker–Hughston form was also obtained by Jin & Glasserman (2001, theorem 5). In order to prove this result they found a rather technical necessary and sufficient condition for positivity in terms of the volatility structure of the HJM form, confirming that the HJM formulation is not the most natural one to investigate positive interest rates.

(c) *The chaotic approach*

We have seen in the potential approach that the fundamental ingredient to model the random behaviour of the interest rates is the increasing process A_t in the decomposition $V_t = E_t[A_\infty] - A_t$, whereas in the Flesaker–Hughston construction the corresponding role is played by the martingales M_{ts}.

In Hughston & Rafailidis (2005), an elegant construction of general positive-interest-rate models based on a Brownian filtration using simpler fundamentals was introduced. If we assume that V_t is a potential for which the process Λ_t is a true martingale, then integrating (2.10) on the interval (t, T), taking conditional expectations at time t and the limit $T \to \infty$, one finds that

$$V_t = E_t\left[\int_t^\infty r_s V_s \,\mathrm{d}s\right]. \tag{2.11}$$

In particular,

$$E\left[\int_0^\infty r_s V_s \,\mathrm{d}s\right] < \infty. \tag{2.12}$$

Now let σ_t be a vector-valued process such that

$$\sigma_t^\dagger \sigma_t = r_t V_t. \tag{2.13}$$

Due to (2.12), we can define the random variable

$$X_\infty = \int_0^\infty \sigma_s \,\mathrm{d}W_s, \tag{2.14}$$

and it follows from the Itô isometry that

$$V_t = E_t[X_\infty^2] - E_t[X_\infty]^2, \tag{2.15}$$

which is called the conditional variance representation of the state price density V_t. To obtain the connection between this representation and the Flesaker–Hughston framework, observe that a direct comparison between (2.11) and (2.7) gives that

$$h_s M_{ts} = E_t[\sigma_s^\dagger \sigma_s]. \tag{2.16}$$

Similarly, by comparing the conditional variance representation (2.15) with the decomposition (2.4), we see that

$$E_t[X_\infty^2] - X_t^2 = E_t[A_\infty] - A_t,$$

where $X_t = E_t[X_\infty]$. It follows from the uniqueness of the Doob–Meyer decomposition that

$$A_t = [X, X]_t,$$

that is, the quadratic variation of the process X_t.

Conversely, given a zero-mean random variable $X_\infty \in L^2(\Omega, \mathcal{F}, P)$, the representation (2.15) defines a potential V_t, which can then be used as a state price density to obtain a system of bond prices. The issue of absence of arbitrage can then be addressed in terms of necessary and sufficient conditions on X_∞, and is, by and large, an open question at this point. The construction in Hughston & Rafailidis (2005) flows in the opposite direction, in the sense that the authors first enumerate a series of axioms to be satisfied by an arbitrage-free interest-rate model and then obtain a square-integrable random variable X_∞ corresponding to it.

(*d*) *Wiener chaos*

As Hughston & Rafailidis (2005) also observed, the L^2 condition on X_∞ is necessary and sufficient for X_∞ to have the type of orthogonal decomposition known as a Wiener chaos expansion (Itô 1951; Nualart 1995; Wiener 1938). They interpret the different orders of this decomposition as basic building blocks for models of increasing complexity.

Let W_t be an N-dimensional Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, P)$. We introduce a compact notation

$$\tau = (s, \mu) \in \Delta \doteq \mathbb{R}_+ \times \{1, \ldots, N\}$$

and express integrals as

$$\begin{aligned} \int_\Delta f(\tau)\,\mathrm{d}\tau &\doteq \sum_\mu \int_0^\infty f(s,\mu)\,\mathrm{d}s, \\ \int_\Delta f(\tau)\,\mathrm{d}W_\tau &\doteq \sum_\mu \int_0^\infty f(s,\mu)\,\mathrm{d}W_s^\mu. \end{aligned} \tag{2.17}$$

For each $n \geqslant 0$, let

$$H_n(x) = (-1)^n \mathrm{e}^{x^2/2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2/2} \tag{2.18}$$

be the nth Hermite polynomial. For $h \in L^2(\Delta)$, let $\|h\|^2 = \int_\Delta h(\tau)^2 \, d\tau$ and define the Gaussian random variable

$$W(h) := \int_\Delta h(\tau) \, dW_\tau.$$

The spaces

$$\mathcal{H}_n \doteq \mathrm{span}\{H_n(W(h)) \mid h \in L^2(\Delta)\}, \quad n \geqslant 1,$$
$$\mathcal{H}_0 \doteq \mathbb{C},$$

form an orthogonal decomposition of the space $L^2(\Omega, \mathcal{F}_\infty, P)$ of square-integrable random variables:

$$L^2(\Omega, \mathcal{F}_\infty, P) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n.$$

Each $\mathcal{H}_n$ can be identified with $L^2(\Delta_n)$ via the isometries

$$J_n : L^2(\Delta_n) \to \mathcal{H}_n$$

given by

$$f_n \mapsto J_n(f_n) = \int_{\Delta_n} f_n(\tau_1, \ldots, \tau_n) \, dW_{\tau_1} \cdots dW_{\tau_n}, \tag{2.19}$$

where

$$\Delta_n \doteq \{(\tau_1, \ldots, \tau_n) \mid \tau_i = (s_i, \mu_i) \in \Delta, \; 0 \leqslant s_1 \leqslant s_2 \leqslant \cdots \leqslant s_n < \infty\}.$$

With these ingredients, one is then led to the result that any $X \in L^2(\Omega, \mathcal{F}_\infty, P)$ can be represented as a *Wiener chaos expansion*

$$X = \sum_{n=0}^{\infty} J_n(f_n), \tag{2.20}$$

where the deterministic functions $f_n \in L^2(\Delta_n)$ are uniquely determined by the random variable X.

A special example arises by noting that, for $h \in L^2(\Delta)$,

$$n! J_n(h^{\otimes n}) = \|h\|^n H_n\left(\frac{W(h)}{\|h\|}\right) \tag{2.21}$$

and, furthermore,

$$\exp\left[W(h) - \frac{1}{2} \int h(\tau)^2 \, d\tau\right] = \sum_{n=0}^{\infty} \frac{\|h\|^n}{n!} H_n\left(\frac{W(h)}{\|h\|}\right). \tag{2.22}$$

In the notation of quantum field theory (see Appendix A), this example defines the Wick-ordered exponential and Wick powers

$$:\exp[W(h)]: \doteq \exp\left[W(h) - \frac{1}{2} \int h(\tau)^2 \, d\tau\right],$$
$$:W(h)^n: \doteq n! J_n(h^{\otimes n}). \tag{2.23}$$

Generating functionals provide one systematic approach to developing explicit formulae for the terms of the chaos expansion in specific examples.

Theorem 2.1. *For any random variable $X \in L^2(\Omega, \mathcal{F}_\infty, P)$, the generating functional $Z_X(h) : L^2(\Delta) \to \mathbb{C}$ defined by*

$$Z_X(h) \doteq E\left[X \exp\left[W(h) - \frac{1}{2}\int h(\tau)^2 \,\mathrm{d}\tau\right]\right] \tag{2.24}$$

is an entire analytic functional of $h \in L^2(\Delta)$ and hence has an absolutely convergent expansion

$$Z_X(h) = \sum_{n\geqslant 0} F_X^{(n)}(h), \tag{2.25}$$

where

$$F_X^{(n)}(h) = \int_{\Delta_n} f_X^{(n)}(\tau_1, \ldots, \tau_n) h(\tau_1)\cdots h(\tau_n)\,\mathrm{d}\tau_1\cdots \mathrm{d}\tau_n. \tag{2.26}$$

The nth Fréchet derivative of Z_X at $h = 0$, $f_X^{(n)}(\tau_1, \ldots, \tau_n)$, lies in $L^2(\Delta)$. Finally, the Wiener–Itô chaos expansion of X is

$$X = \sum_{n\geqslant 0}\int_{\Delta_n} f_X^{(n)}(\tau_1, \ldots, \tau_n)\,\mathrm{d}W_{\tau_1}\cdots \mathrm{d}W_{\tau_n}. \tag{2.27}$$

Proof. See Appendix B. ■

3. Squared Gaussian models

A number of authors (Jamshidian 1995; Maghsoodi 1996; Rogers 1997) have observed that the CIR model (Cox *et al.* 1985) with an integer constraint

$$N \doteq \frac{4ab}{c^2} \in \mathbb{N}_+ \setminus \{0, 1\}$$

lies in the class of so-called squared Gaussian models. By introducing an $\mathbb{R}^N$-valued Ornstein–Uhlenbeck process R_t, governed by the stochastic differential equation

$$\mathrm{d}R_t = -\tfrac{1}{2}aR_t\,\mathrm{d}t + \tfrac{1}{2}c\,\mathrm{d}W_t, \tag{3.1}$$

where W_t is N-dimensional Brownian motion, the Itô formula together with Lévy's criterion for Brownian motion shows that the square $r_t = R_t^\dagger R_t$ satisfies (1.1), where

$$\tilde{W}_t = \int_0^t (R_s^\dagger R_s)^{-1/2} R_s^\dagger \mathrm{d}W_s$$

is itself a one-dimensional Brownian motion.

We focus on a general family of interest-rate models which includes this example, the so-called extended CIR model, and more. Note that we always work in the physical measure and thus to specify the term-structure model one needs to determine the market price of risk vector λ_t as well as the spot-rate process r_t.

Definition 3.1. A pair (r_t, λ_t) of $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ processes is called an N-dimensional squared Gaussian model of interest rates ($N \geqslant 2$) if there is an $\mathbb{R}^N$-valued Ornstein–Uhlenbeck process such that $r_t = R_t^\dagger R_t$ and $\lambda_t = \bar{\lambda}(t)R_t$. R_t satisfies

$$\mathrm{d}R_t = \alpha(t)(\bar{R}(t) - R_t)\,\mathrm{d}t + \gamma(t)\,\mathrm{d}W_t, \quad R|_{t=0} = R_0, \tag{3.2}$$

where the symmetric matrices α, γ, $\bar{\lambda}$ and the vector $\bar{R}$ are deterministic Lipschitz functions on $\mathbb{R}_+$, and W is standard N-dimensional Brownian motion. In addition we impose boundedness conditions that there is some constant $M > 0$ such that

$$\bar{\lambda}^2(t) \leqslant MI, \qquad \alpha(t) \geqslant M^{-1}I, \qquad \alpha(t) + \gamma(t)\bar{\lambda}(t) \geqslant M^{-1}I, \qquad \gamma^2(t) \geqslant M^{-1}I,$$

for all t.

The exact solution of (3.2) is easily seen to be

$$R_t = \tilde{R}(t) + \int_0^t K(t, t_1)(\gamma\, \mathrm{d}W)_{t_1}, \tag{3.3}$$

where

$$\tilde{R}(t) = K(t, 0)R_0 + \int_0^t K(t, t_1)\alpha(t_1)\bar{R}(t_1)\, \mathrm{d}t_1 \tag{3.4}$$

and $K(t, s), t \geqslant s$ is the matrix-valued solution of

$$\begin{cases} \mathrm{d}K(t, s)/\mathrm{d}t = -\alpha(t)K(t, s), & 0 \leqslant s \leqslant t, \\ K(t, t) = I, & 0 \leqslant t, \end{cases} \tag{3.5}$$

which generates the Ornstein–Uhlenbeck semigroup.

In accordance with (2.10), we define the state price density process to be

$$V_t = \exp\left[-\int_0^t (R_s^\dagger(I + \tfrac{1}{2}\bar{\lambda}(s)^2)R_s\, \mathrm{d}s + R_s^\dagger\bar{\lambda}(s)\, \mathrm{d}W_s)\right]. \tag{3.6}$$

We thus have a natural candidate for the random variable X_∞:

$$X_\infty = \int_0^\infty \sigma_t^\dagger\, \mathrm{d}W_t, \tag{3.7}$$

where the $\mathbb{R}^N$-valued process

$$\sigma_t \doteq \exp\left[-\int_0^t (R_s^\dagger(\tfrac{1}{2}I + \tfrac{1}{4}\bar{\lambda}(s)^2)R_s\, \mathrm{d}s + \tfrac{1}{2}R_s^\dagger\bar{\lambda}(s)\, \mathrm{d}W_s)\right]R_t \tag{3.8}$$

is the natural solution of $\sigma_t^\dagger\sigma_t = r_tV_t$.

Before proceeding to analyse X_∞, we show that squared Gaussian models give rise to a state price density V_t satisfying the conditions of the previous section, and consequently $E[X_\infty^2] = 1$.

Proposition 3.2. *For squared Gaussian models, the process*

$$\Lambda_t = \exp\left(-\int_0^t \lambda_s^\dagger\, \mathrm{d}W_s - \frac{1}{2}\int_0^t \lambda_s^\dagger\lambda_s\, \mathrm{d}s\right) \tag{3.9}$$

is a martingale for $0 \leqslant t \leqslant T$, and the state price density V_t defined in (3.6) is a potential.

Proof. To show that the positive local martingale Λ_t is a martingale, it suffices to show that $E[\Lambda_T]=\Lambda_0$. The proof, which we sketch, follows from a similar proof in Cheridito *et al.* (2003). We introduce a new process

$$\mathrm{d}\hat{R}_t=\alpha(t)\bar{R}(t)\,\mathrm{d}t-[\alpha(t)+\gamma(t)\bar{\lambda}(t)]\hat{R}_t\,\mathrm{d}t+\gamma(t)\,\mathrm{d}W_t$$

and two sequences of stopping times for $n\in\mathbb{N}$:

$$\tau_n=\inf\left\{t:\int_0^t R_s^{\dagger}\bar{\lambda}^2(s)R_s\,\mathrm{d}s\geqslant n\right\}\wedge T,$$
$$\hat{\tau}_n=\inf\left\{t:\int_0^t \hat{R}_s^{\dagger}\bar{\lambda}^2(s)\hat{R}_s\,\mathrm{d}s\geqslant n\right\}\wedge T.$$

Under the conditions of definition 3.1, we have that both

$$\lim_{n\to\infty}P\{\tau_n=T\}=1\quad\text{and}\quad\lim_{n\to\infty}P\{\hat{\tau}_n=T\}=1.$$

Moreover, the processes $\lambda_t^{(n)}=\lambda_t\mathbb{1}_{\{t<\tau_n\}}$ satisfy the Novikov condition, so that from the Girsanov theorem we obtain that their stochastic exponentials $\Lambda_t^{(n)}$ are densities of equivalent measures $Q^{(n)}\sim P$. Therefore,

$$\begin{aligned}E[\Lambda_T]&=\lim_{n\to\infty}E[\Lambda_T\mathbb{1}_{\{\tau_n=T\}}]=\lim_{n\to\infty}E[\Lambda_T^{(n)}\mathbb{1}_{\{\tau_n=T\}}]\\&=\lim_{n\to\infty}E^{Q^{(n)}}[\mathbb{1}_{\{\tau_n=T\}}]=\lim_{n\to\infty}E[\mathbb{1}_{\{\hat{\tau}_n=T\}}]=1,\end{aligned}$$

where we have used the fact that the distribution of $R_{t\wedge\tau_n}$ under $Q^{(n)}$ is the same as the distribution of $\hat{R}_{t\wedge\hat{\tau}_n}$ under P.

To prove that the supermartingale V_t defined in (3.6) is a potential, let $0<\epsilon<M$ and write $V_T=\mathrm{e}^{-Y_1-Y_2}$, where

$$Y_1=\int_0^T R_t^{\dagger}(I-\tfrac{1}{2}\epsilon\bar{\lambda}(t)^2)R_t\,\mathrm{d}t$$

and

$$Y_2=\frac{1}{1+\epsilon}\int_0^T[\tfrac{1}{2}R_t^{\dagger}(I+\epsilon)^2\bar{\lambda}(t)^2R_t\,\mathrm{d}t+(I+\epsilon)R_t^{\dagger}\bar{\lambda}(t)\,\mathrm{d}W_t].$$

By the Hölder inequality,

$$E[V_T]\leqslant(E[\mathrm{e}^{-(1+1/\epsilon)Y_1}])^{\epsilon/(1+\epsilon)}(E[\mathrm{e}^{-(1+\epsilon)Y_2}])^{1/(1+\epsilon)},$$

with the second factor less than or equal to 1, since $\mathrm{e}^{-(1+\epsilon)Y_2}$ is a positive local martingale. Now Y_1 is a positive random variable for which a direct computation using the lower bound on γ shows

$$\operatorname{mean}(Y_1)=C_1NT(1+\mathcal{O}(T)),\tag{3.10}$$
$$\operatorname{var}(Y_1)=C_2NT(1+\mathcal{O}(T)),\tag{3.11}$$

for positive constants C_1, C_2. An easy application of Chebyshev's inequality,

$$\operatorname{Prob}(Y_1\leqslant\tfrac{1}{2}C_1NT)\leqslant\mathcal{O}\left(\frac{1}{NT}\right),\tag{3.12}$$

then implies that $\lim_{T\to\infty}E[V_T]=0$. ∎

4. Exponentiated second chaos

The chaos expansion we seek for the CIR model will be derived from a closed formula for expectations of e^{-Y} for elements

$$Y = A + \int_{\Delta} B(\tau_1)\, \mathrm{d}W_{\tau_1} + \int_{\Delta_2} C(\tau_1, \tau_2)\, \mathrm{d}W_{\tau_1}\, \mathrm{d}W_{\tau_2} \tag{4.1}$$

in a certain subset

$$\mathcal{C}^+ \subset \mathcal{H}_{\leqslant 2} \doteq \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2.$$

In the integrals above, recall that compact notation using τs carries a summation over vector indices as well as integration over time. The formula we present is well known in the theory of Gaussian functional integration (Glimm & Jaffe 1981, ch. 9). In probability theory, this result gives the Laplace transform of a general class of quadratic functionals of Brownian motion. Many special cases of this result have been studied in probability theory (see, for example, Yor (1992, ch. 2) and the references contained therein).

If in (4.1) we define $C(\tau_1, \tau_2) = C(\tau_2, \tau_1)$ when $\tau_1 > \tau_2$, then C is the kernel of a symmetric integral operator on $L^2(\Delta)$:

$$[Cf](\tau) = \int_0^{\infty} C(\tau, \tau_1) f(\tau_1)\, \mathrm{d}\tau_1. \tag{4.2}$$

Recall that Hilbert–Schmidt operators on $L^2(\Delta)$ are finite norm operators under the norm

$$\|C\|_{\mathrm{HS}}^2 = \int_{\Delta^2} C(\tau_1, \tau_2)^2\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2.$$

We say that $Y \in \mathcal{H}_{\leqslant 2}$ is in $\mathcal{C}^+$ if C is the kernel of a symmetric Hilbert–Schmidt operator on $L^2(\Delta)$ such that $(1 + C)$ has positive spectrum.

Proposition 4.1. *Let $Y \in \mathcal{C}^+$. Then*

$$E[\mathrm{e}^{-Y}] = [\det{}_2(1 + C)]^{-1/2} \exp\left[-A + \frac{1}{2} \int_{\Delta_2} B(\tau_1)(1 + C)^{-1}(\tau_1, \tau_2) B(\tau_2)\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2\right]. \tag{4.3}$$

Remark 4.2. The Carleman–Fredholm determinant is defined as the extension of the formula

$$\det{}_2(1 + C) = \det(1 + C) \exp[-\operatorname{tr}(C)] \tag{4.4}$$

from finite-rank operators to bounded Hilbert–Schmidt operators; the operator kernel $(1 + C)^{-1}(\tau_1, \tau_2)$ is also the natural extension from the finite-rank case.

Proof. See Appendix C. ■

Using this proposition, it is possible to deduce the chaos expansion of the random variables e^{-Y}, $Y \in \mathcal{C}^+$, a result known in quantum field theory as Wick's theorem.

Corollary 4.3. *If*

$$Y = \int_{\Delta_2} C(\tau_1, \tau_2)\, \mathrm{d}W_{\tau_1}\, \mathrm{d}W_{\tau_2} \in \mathcal{C}^+,$$

then the random variable $X = \mathrm{e}^{-Y}$ *has Wiener chaos coefficient functions*

$$f_n(\tau_1, \ldots, \tau_n) = \begin{cases} K \sum_{G \in \mathcal{G}_n} \prod_{g \in G} [C(1+C)^{-1}](\tau_{g_1}, \tau_{g_2}), & n \text{ even}, \\ 0, & n \text{ odd}, \end{cases} \tag{4.5}$$

where $K = [\det_2(1+C)]^{-1/2}$ *and, for* n *even,* $\mathcal{G}_n$ *is the set of Feynman graphs on the* n *marked points* $\{\tau_1, \ldots, \tau_n\}$. *Each Feynman graph* G *is a disjoint union of unordered pairs* $g = (\tau_{g_1}, \tau_{g_2})$ *with* $\bigcup_{g \in G} g = \{\tau_1, \ldots, \tau_n\}$.

Proof. The generating functional for $X = \mathrm{e}^{-Y}$ is

$$\begin{aligned} Z_X(h) &= E\left[X \exp\left(\int h(\tau)\,\mathrm{d}W_\tau - \frac{1}{2}\int h(\tau)^2\,\mathrm{d}\tau\right)\right] \\ &= E\left[\exp\left(\int h(\tau)\,\mathrm{d}W_\tau - \frac{1}{2}\int h(\tau)^2\,\mathrm{d}\tau - \int_{\Delta_2} C(\tau_1, \tau_2)\,\mathrm{d}W_{\tau_1}\,\mathrm{d}W_{\tau_2}\right)\right], \end{aligned}$$

so we can use proposition 4.1 with $A = \frac{1}{2}\int h(\tau)^2\,\mathrm{d}\tau$ and $B(\tau) = -h(\tau)$, which yields

$$\begin{aligned} Z_X(h) &= \det_2(1+C)^{-1/2} \\ &\quad \times \exp\left[-\frac{1}{2}\int_{\Delta^2} h^\dagger(\tau_1)[\delta(\tau_1, \tau_2) - (1+C)^{-1}(\tau_1, \tau_2)]h(\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right]. \end{aligned} \tag{4.6}$$

Using the last part of theorem 2.1, the result comes by evaluating the nth Fréchet derivative at $h = 0$, or equivalently by expanding the exponential and symmetrizing over the points $\tau_1, \ldots, \tau_n$ in the $\frac{1}{2}n$th term. ■

5. The chaotic expansion for squared Gaussian models

We now derive the chaos expansion for the squared Gaussian model defined by (3.2). In view of (3.7) it will be enough to find the chaos expansion for σ_T^μ, $T < \infty$. We start by finding its generating functional $Z_{\sigma_T^\mu}$. For $h, k \in L^2(\Delta)$, define the auxiliary functional $Z(h, k) = E[\mathrm{e}^{-Y_T}]$ with

$$\begin{aligned} Y_T = \int_0^T R_t^\dagger(\tfrac{1}{2}I + \tfrac{1}{4}\bar{\lambda}^2)R_t\,\mathrm{d}t + \frac{1}{2}\int_0^T R_t^\dagger\bar{\lambda}(t)\,\mathrm{d}W_t - \int_0^T h^\dagger(t)\,\mathrm{d}W_t \\ + \frac{1}{2}\int_0^T h^\dagger(t)h(t)\,\mathrm{d}t - \int_0^T k^\dagger(t)R_t\,\mathrm{d}t. \end{aligned} \tag{5.1}$$

Proposition 5.1. *$Z(h, k)$ is an entire analytic functional on* $L^2(\Delta) \times L^2(\Delta)$. *Moreover,*

$$\lim_{t \to T^-} \left.\frac{\delta Z(h, k)}{\delta k^\mu(t)}\right|_{k=0} = Z_{\sigma_T^\mu}(h), \tag{5.2}$$

where $Z_{\sigma_T^\mu}(h)$ *is defined by (2.24) with* $X = \sigma_T^\mu$, $\mu = 1, \ldots, N$.

Proof. Analyticity in (h, k) follows by repeating the argument given in Appendix B. By the definition of Fréchet differentiation and continuity of the $t \to T^-$ limit, (5.2) follows. ■

We want to use proposition 4.1 in order to compute $Z(h,k)$. Substitution of (3.3) into the first term of (5.1) leads to

$$\begin{aligned}\int_0^T R_t^\dagger(\tfrac{1}{2}I+\tfrac{1}{4}\bar\lambda(t)^2)R_t\,\mathrm{d}t \\ &= \int_0^T \tilde R^\dagger(t)(\tfrac{1}{2}I+\tfrac{1}{4}\bar\lambda(t)^2)\tilde R(t)\,\mathrm{d}t \\ &\quad + \int_0^T\left[\int_0^T \tilde R^\dagger(s)(I+\tfrac{1}{2}\bar\lambda(t)^2)K_T(s,t)\,\mathrm{d}s\right]\gamma(t)\,\mathrm{d}W_t \\ &\quad + \int_{\Delta_2}\gamma(t_1)\left[\int_0^T K_T^\dagger(t_1,s)(I+\tfrac{1}{2}\bar\lambda(t)^2)K_T(s,t_2)\,\mathrm{d}s\right]\gamma(t_2)\,\mathrm{d}W_{t_1}\,\mathrm{d}W_{t_2} \\ &\quad + \int_0^T \mathrm{tr}\left\{\gamma(t)\left[\int_0^T K_T^\dagger(t,s)(\tfrac{1}{2}I+\tfrac{1}{4}\bar\lambda(t)^2)K_T(s,t)\,\mathrm{d}s\right]\gamma(t)\right\}\mathrm{d}t,\end{aligned}$$

where we define $K_T(t_1,t_2)=\mathbb{1}(t_1\leqslant T)K(t_1,t_2)$. For the second and the last terms of (5.1) we have

$$\begin{aligned}\frac{1}{2}\int_0^T R_t^\dagger\bar\lambda(t)\,\mathrm{d}W_t &= \frac{1}{2}\int_0^T \tilde R^\dagger(t)\bar\lambda(t)\,\mathrm{d}W_t+\frac{1}{2}\int_0^T \mathrm{tr}\left[\int_0^T \gamma(s)K_T^\dagger(s,t)\bar\lambda(t)\,\mathrm{d}s\right]\mathrm{d}t \\ &\quad +\frac{1}{2}\int_{\Delta_2}(\gamma(t_1)K_T^\dagger(t_1,t_2)\bar\lambda(t)+\bar\lambda(t)K_T(t_1,t_2)\gamma(t_2))\,\mathrm{d}W_{t_1}\,\mathrm{d}W_{t_2},\end{aligned}$$

$$\int_0^T k^\dagger(t)R_t\,\mathrm{d}t=\int_0^T k^\dagger(t)\tilde R(t)\,\mathrm{d}t+\int_0^T\left(\int_0^T k^\dagger(s)K_T(s,t)\,\mathrm{d}s\right)\gamma(t)\,\mathrm{d}W_t.$$

Thus the exponent Y_T appearing in (5.1) has the form of (4.1) with

$$\begin{aligned}A_T &= \int_0^T\left[\tilde R^\dagger(t)(\tfrac{1}{2}I+\tfrac{1}{4}\bar\lambda(t)^2)\tilde R(t)+\tfrac{1}{2}h^\dagger(t)h(t)-k^\dagger(t)\tilde R(t)\right]\mathrm{d}t \\ &\quad +\int_0^T \mathrm{tr}\left\{\gamma(t)\left[\int_0^T K_T^\dagger(t,s)(\tfrac{1}{2}I+\tfrac{1}{4}\bar\lambda(t)^2)K_T(s,t)\,\mathrm{d}s\right]\gamma(t)\right\}\mathrm{d}t \\ &\quad +\frac{1}{2}\int_0^T \mathrm{tr}\left[\int_0^T \gamma(s)K_T^\dagger(s,t)\bar\lambda(t)\,\mathrm{d}s\right]\mathrm{d}t, \\ B_T(t) &= -h(t)-\gamma(t)\int_0^T K_T^\dagger(t,s)k(s)\,\mathrm{d}s+\tfrac{1}{2}\bar\lambda(t)\tilde R(t) \\ &\quad +\gamma(t)\int_0^T K_T^\dagger(t,s)(I+\tfrac{1}{2}\bar\lambda(t)^2)\tilde R(s)\,\mathrm{d}s, \\ C_T(t_1,t_2) &= \gamma(t_1)\left[\int_0^T K_T^\dagger(t_1,s)(I+\tfrac{1}{2}\bar\lambda(t)^2)K_T(s,t_2)\,\mathrm{d}s\right]\gamma(t_2) \\ &\quad +\tfrac{1}{2}[\gamma(t_1)K_T^\dagger(t_1,t_2)\bar\lambda(t)+\bar\lambda(t)K_T(t_1,t_2)\gamma(t_2)].\end{aligned}$$

It is clear that the operator C_T has Hilbert–Schmidt norm $\|C_T\|^2_{\mathrm{HS}}=\mathcal{O}(T)$. Moreover, if we denote by $\gamma K_T^\dagger(1+\frac{1}{2}\bar\lambda^2)K_T\gamma$, $\gamma K_T^\dagger\bar\lambda$ and $\bar\lambda K_T\gamma$ the operators whose kernels appear in the expression above, then C_T can be written as

$$C_T = \gamma K_T^\dagger K_T \gamma + \tfrac{1}{2}(\gamma K_T^\dagger \bar{\lambda} + 1)(\bar{\lambda} K_T \gamma + 1) - \tfrac{1}{2}, \tag{5.3}$$

from which we see that $(1 + C_T)$ is positive. Therefore, we can use proposition 4.1 for $E[\mathrm{e}^{-Y_T}]$, leading to a general formula for the generating functional $Z(h, k)$:

$$\begin{aligned} Z(h,k) &\\ &= \det{}_2(1 + C_T)^{-1/2} \exp(-\tfrac{1}{2} \operatorname{tr} C_T) \\ &\quad \times \exp\left\{ -\int_0^T \left[\tilde{R}^\dagger(t)(\tfrac{1}{2} I + \tfrac{1}{4}\bar{\lambda}(t)^2)\tilde{R}(t) + \tfrac{1}{2} h^\dagger(t) h(t) - k^\dagger(t)\tilde{R}(t) \right] \mathrm{d}t \right\} \\ &\quad \times \exp\left\{ \frac{1}{2} \int_{\Delta_2} \left[h^\dagger(t_1) + \int_0^T k^\dagger(s) K_T(s, t_1)\gamma(t_1)\, \mathrm{d}s - \tfrac{1}{2}\tilde{R}^\dagger(t_1)\bar{\lambda}(t) \right.\right. \\ &\qquad\qquad \left. - \int_0^T \tilde{R}^\dagger(s)(I + \tfrac{1}{2}\bar{\lambda}(t)^2) K_T(s, t_1)\gamma(t_1)\, \mathrm{d}s \right] (1 + C_T)^{-1}(t_1, t_2) \\ &\qquad\qquad \times \left[h(t_2) + \gamma(t_2) \int_0^T K_T^\dagger(t_2, s) k_s\, \mathrm{d}s - \tfrac{1}{2}\bar{\lambda}(t)\tilde{R}(t_2) \right. \\ &\qquad\qquad\qquad \left.\left. - \gamma(t_2) \int_0^T K_T^\dagger(t_2, s)(I + \tfrac{1}{2}\bar{\lambda}(t)^2)\tilde{R}(s)\, \mathrm{d}s \right] \mathrm{d}t_1\, \mathrm{d}t_2 \right\}. \end{aligned} \tag{5.4}$$

Differentiation once with respect to k then yields

$$\begin{aligned} Z_{\sigma_T}(h) = M_T &\exp\left\{ -\int_0^T \left[\tilde{R}^\dagger(t)(\tfrac{1}{2} I + \tfrac{1}{4}\bar{\lambda}(t)^2)\tilde{R}(t) + \tfrac{1}{2} h^\dagger(t) h(t) \right] \mathrm{d}t \right\} \\ &\times \{ -\tilde{R} + K_T \gamma (1 + C_T)^{-1} [h - \tfrac{1}{2}\bar{\lambda}(t)\tilde{R} - \gamma K_T^\dagger (I + \tfrac{1}{2}\bar{\lambda}(t)^2)\tilde{R}] \}(T) \\ &\times \exp\left\{ \frac{1}{2} \int_{\Delta_2} [h^\dagger - \tfrac{1}{2}\tilde{R}^\dagger \bar{\lambda}(t) - \tilde{R}^\dagger (I + \tfrac{1}{2}\bar{\lambda}(t)^2) K_T \gamma](t_1)(1 + C_T)^{-1}(t_1, t_2) \right. \\ &\qquad \left. \times [h - \tfrac{1}{2}\bar{\lambda}(t)\tilde{R} - \gamma K_T^\dagger (I + \tfrac{1}{2}\bar{\lambda}(t)^2)\tilde{R}](t_2)\, \mathrm{d}t_1\, \mathrm{d}t_2 \right\}, \end{aligned} \tag{5.5}$$

where

$$M_T = \mathrm{e}^{-(1/2)\operatorname{tr} C_T} (\det{}_2(1 + C_T))^{-1/2} = (\det(1 + C_T))^{-1/2}. \tag{5.6}$$

By comparing (5.4) and (5.5), the reader can observe our use of an operator notation, which suppresses some time integrals; for example, in the very last term,

$$[\gamma K_T^\dagger (I + \tfrac{1}{2}\bar{\lambda}^2)\tilde{R}](t) \doteq \gamma(t) \int_0^T K_T^\dagger(t, s)(I + \tfrac{1}{2}\bar{\lambda}^2)\tilde{R}(s)\, \mathrm{d}s,$$

and similarly for other terms.

These formulae simplify considerably if the function $\tilde{R}$ vanishes, which is true in the simple CIR model of (3.1) when $r_0 = 0$. In this case we have $\alpha(t) = \frac{1}{2} a I$ and $\gamma(t) = \frac{1}{2} c I$, so that $K_T(s, t) = \mathrm{e}^{-a(s-t)/2} \mathbb{1}(t \leqslant s \leqslant T)$ and

$$C_T(t_1, t_2) = \frac{c^2}{4a}(I + \tfrac{1}{2}\bar{\lambda}^2)[\mathrm{e}^{-(a/2)|t_1 - t_2|} - \mathrm{e}^{(a/2)(t_1 + t_2 - 2T)}] + \tfrac{1}{2} c \bar{\lambda} \mathrm{e}^{-(a/2)|t_1 - t_2|}. \tag{5.7}$$

Moreover, the previous expression for $Z_{\sigma_T}(h)$ reduces to

$$Z_{\sigma_T}(h) = M_T[K_T\gamma(1+C_T)^{-1}h](T) \times \exp\left[-\frac{1}{2}\int_0^T h^\dagger(t)h(t)\,\mathrm{d}t + \frac{1}{2}\int_{\Delta_2} h^\dagger(t_1)(1+C_T)^{-1}(t_1,t_2)h(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2\right].$$

We can then easily evaluate the nth Fréchet derivative of Z_{σ_T} at $h=0$ as in the proof of corollary 4.3 and determine the following partly explicit form for the nth term of the chaos expansion.

Theorem 5.2. *The nth term of the chaos expansion of σ_T for the CIR model with initial condition $r_0 = 0$ is zero for n even. For n odd, the kernel of the expansion is the function $f_T^{(n)}(\cdot) : \Delta_n \to \mathbb{R}$:*

$$f_T(t_1,\dots,t_n) = M_T \sum_{G\in\mathcal{G}_n^*} \prod_{g\in G} L(g), \tag{5.8}$$

where

$$L(g) = \begin{cases} [C_T(1+C_T)^{-1}](t_{g_1},t_{g_2}), & T\notin g,\\ (K_T\gamma(1+C_T)^{-1})(T,t_{g_2}), & T\in g. \end{cases} \tag{5.9}$$

Here, $\mathcal{G}_n^*$ is the set of Feynman graphs, each Feynman graph G being a partition of $\{t_1,\dots,t_n,T\}$ into pairs $g=(t_{g_1},t_{g_2})$.

The chaos expansion for X_∞ itself is exactly the same, except that the variable T is treated as an additional Itô integration variable. The explicit expansion up to fourth order is

$$\begin{aligned} X_\infty = {} & \int_{\Delta_2} M_T[K_T\gamma(1+C_T)^{-1}](T,t_1)\,\mathrm{d}W_{t_1}\,\mathrm{d}W_T \\ & + \int_{\Delta_4} M_T[K_T\gamma(1+C_T)^{-1}](T,t_3)[C_T(1+C_T)^{-1}](t_1,t_2)\,\mathrm{d}W_{t_1}\,\mathrm{d}W_{t_2}\,\mathrm{d}W_{t_3}\,\mathrm{d}W_T \\ & + \int_{\Delta_4} M_T[K_T\gamma(1+C_T)^{-1}](T,t_2)[C_T(1+C_T)^{-1}](t_1,t_3)\,\mathrm{d}W_{t_1}\,\mathrm{d}W_{t_2}\,\mathrm{d}W_{t_3}\,\mathrm{d}W_T \\ & + \int_{\Delta_4} M_T[K_T\gamma(1+C_T)^{-1}](T,t_1)[C_T(1+C_T)^{-1}](t_2,t_3)\,\mathrm{d}W_{t_1}\,\mathrm{d}W_{t_2}\,\mathrm{d}W_{t_3}\,\mathrm{d}W_T \\ & + \cdots. \end{aligned} \tag{5.10}$$

6. Bond-pricing formula

In this section we give a derivation of the price of a zero coupon bond in the CIR model. Recall from § 2 that these are given by

$$P_{tT} = E_t[V_t^{-1}V_T]. \tag{6.1}$$

To keep things 'as simple as possible, but not any simpler', we take $\bar{\lambda}=0$ so $V_t = \exp[-\int_0^t r_s\,\mathrm{d}s]$, or, in terms of the squared Gaussian formulation,

$$V_t = \exp\left[-\int_0^t R_s^\dagger R_s\,\mathrm{d}s\right]. \tag{6.2}$$

As we have seen in the previous section, for $t \leqslant s \leqslant T$,

$$R_s^\mu = K_T(s,t)R_t^\mu + \tfrac{1}{2}c\int_t^s K_T(s,s_1)\,\mathrm{d}W_{s_1}^\mu, \tag{6.3}$$

hence

$$-\log[V_t^{-1}V_T] = \sum_\mu \int_t^T (R_s^\mu)^2\,\mathrm{d}s$$

can be written as

$$\sum_\mu\left[\frac{4}{c^2}(R_t^\mu)^2 C_T(t,t) + \frac{4R_t^\mu}{c}\int_t^T C_T(t,s)\,\mathrm{d}W_s^\mu + 2\int_t^T\int_t^{s_2} C_T(s_1,s_2)\,\mathrm{d}W_{s_1}^\mu\,\mathrm{d}W_{s_2}^\mu\right] + N\int_t^T C_T(s,s)\,\mathrm{d}s, \tag{6.4}$$

where $C_T(s_1,s_2)$ is given by (5.7) with $\bar{\lambda}=0$.

Taking the conditional expectation of $V_t^{-1}V_T$ by use of proposition 4.1 leads to the desired formula:

$$P_{tT} = [\det(1+2C_T)]^{-N/2}\prod_\mu \exp\left[-\frac{4}{c^2}(R_t^\mu)^2(C_T(1+2C_T)^{-1})(t,t)\right]. \tag{6.5}$$

Thus P_{tT} has the exponential affine form $\exp[-\beta(t,T)r_t - \alpha(t,T)]$ with

$$\left.\begin{aligned}\beta(t,T) &= \frac{4}{c^2}[C_T(1+2C_T)^{-1}](t,t),\\ \alpha(t,T) &= \tfrac{1}{2}N\log[\det(1+2C_T)].\end{aligned}\right\} \tag{6.6}$$

The known formula has the same form, with

$$\left.\begin{aligned}\beta(t,T) &= \frac{2(\mathrm{e}^{\rho(T-t)}-1)}{(\rho+a)(\mathrm{e}^{\rho(T-t)}-1)+2\rho}, \quad \rho^2 = a^2+2c^2,\\ \alpha(t,T) &= -\frac{2ab}{\rho^2}\log\left[\frac{2\rho\mathrm{e}^{(a+\rho)(T-t)/2}}{(\rho+a)(\mathrm{e}^{\rho(T-t)}-1)+2\rho}\right],\end{aligned}\right\} \tag{6.7}$$

which can be derived as solutions of the pair of Ricatti ordinary differential equations:

$$\left.\begin{aligned}\frac{\partial\beta}{\partial t} &= \tfrac{1}{2}c^2\beta^2 + a\beta - 1,\\ \frac{\partial\alpha}{\partial t} &= -ab\beta.\end{aligned}\right\} \tag{6.8}$$

One can demonstrate using power-series expansions that (6.6) do in fact solve the Ricatti equations and hence agree with the usual formula. This example points to the rather subtle general relationship between kernels such as $(1+2C_T)^{-1}$ and solutions of Ricatti equations deserving of further study.

7. Discussion

We have shown how the CIR model, at least in integer dimensions, can be viewed within the chaos framework of Hughston & Rafailidis (2005) as arising from a somewhat special random variable X_∞. This random variable can be understood as derived from exponentiated second chaos random variables e^{-Y}, $Y \in \mathcal{C}^+$. Such exponentiated $\mathcal{C}^+$ variables form a rich and natural family that is likely to include many more candidates for applicable interest-rate models. Although their analytic properties are complicated, there do exist approximation schemes, which can, in principle, be the basis for numerical methods.

On the theoretical side, this family is distinguished by its natural invariance properties. Most notably, as will be investigated elsewhere, it is invariant under conditional $\mathcal{F}_t$-expectations: $\log E_t[\mathrm{e}^{-Y}] \in \mathcal{C}^+$ whenever $Y \in \mathcal{C}^+$. Note in particular that this implies that these can be used as the Radon–Nikodym derivatives of measure changes which generalize the Girsanov transform, and which can greatly enrich the tools applicable in finance.

The results of this paper rest on the special properties of squared Gaussian processes, which in turn work within the filtration $\mathcal{F}^W$ of N-dimensional Brownian motion W. However, the CIR process r_t itself is adapted to the much smaller filtration $\mathcal{F}^{\tilde{W}}$ of one-dimensional Brownian motion $\tilde{W}$. It is an important unsolved problem to construct an analytical chaos expansion for the CIR model adapted to the natural filtration $\mathcal{F}^{\tilde{W}}$. Since the expansion we present certainly does not have this property, we do not expect a simple relation between the two constructions.

Our application of the chaos expansion to squared Gaussian models also illustrates a deep connection between methods developed for quantum field theory and the methods of Malliavin calculus. Many of the very rich analytic properties of this example reflect well-known techniques widely used in mathematical physics.

Appendix A. White-noise calculus

Here we describe the white-noise calculus, which can be regarded as a reformulation of the calculus of Wiener measure (Øksendal 1996) into concepts familiar to practitioners in quantum field theory such as Gaussian functional integration (Glimm & Jaffe 1981)

Let $\mathcal{S}$ be the Schwartz space of smooth functions on $\mathbb{R}_+$ of rapid decrease, and $\mathcal{S}'$ its topological dual, the space of tempered distributions on $\mathbb{R}_+$. If $\phi \in \mathcal{S}'$ and $f \in \mathcal{S}$, we write

$$\phi(f) \doteq \langle \phi, f \rangle$$

for their canonical pairing. We also use the formal notation

$$\phi(f) = \int_{\mathbb{R}_+} f(s)\phi_s \,\mathrm{d}s$$

to represent the 'smearing' of the distribution ϕ over the test function f. It acquires a rigorous meaning, however, in the cases where ϕ is itself a function on $\mathbb{R}_+$ for which the pointwise product $\phi(s)f(s)$ is integrable for all $f \in \mathcal{S}$.

Now define the functional

$$S\{f\} \doteq \mathrm{e}^{-\|f\|^2/2} = \mathrm{e}^{-\langle f,f\rangle_{L^2}/2}, \tag{A 1}$$

where $\langle\cdot,\cdot\rangle_{L^2}$ denotes the real-valued inner product in $L^2(\mathbb{R}_+)$. From its properties, it follows from the Bochner–Minlos theorem that there exists a unique Borel probability measure μ on $\mathcal{S}'$ such that $S\{f\}$ corresponds to a *moment-generating functional*, that is, for all $f\in\mathcal{S}$,

$$\int_{\mathcal{S}'} \mathrm{e}^{\mathrm{i}\phi(f)}\,\mathrm{d}\mu(\phi) = S\{f\} = \mathrm{e}^{-\|f\|^2/2}. \tag{A 2}$$

The measure space $(\mathcal{S}',\mathcal{B},\mu)$, where $\mathcal{B}$ is the Borel σ-algebra of $\mathcal{S}'$, is called the *white-noise probability space*.

In Euclidean quantum field theory, a given Borel measure μ on $\mathcal{S}'$ characterizes the family of 'fields' $\phi\in\mathcal{S}'$ through the properties of the random variables $\phi(f):\mathcal{S}'\to\mathbb{R}$ obtained for each $f\in\mathcal{S}$. One tries to construct measures μ so that the generating functional $S\{f\}$ satisfies the so-called Osterwalder–Schrader axioms, in order to guarantee that the fields ϕ have certain required physical properties. The family of Euclidean free fields is obtained when

$$S_C\{f\} \doteq \mathrm{e}^{-\langle f,Cf\rangle/2} = \int_{\mathcal{S}'} \mathrm{e}^{\mathrm{i}\phi(f)}\,\mathrm{d}\mu_C(\phi),$$

where C is the integral kernel of a positive, continuous, non-degenerate Euclidean covariant bilinear form C on $\mathcal{S}\times\mathcal{S}$. We see that the special case $S\{f\}=\mathrm{e}^{-\|f\|^2/2}$ is obtained when $C(s,t)=\delta(t-s)$, called the 'ultralocal' covariance. For each $\phi\in\mathcal{S}'$, the random variables $\{\phi(f):f\in\mathcal{S}\}\subset L^2(\mathcal{S}',\mathcal{B},\mu)$ form a Gaussian family with mean zero and covariances

$$E_\mu[\phi(f)\phi(g)] = \int_{\mathbb{R}_+\times\mathbb{R}_+} f(s)g(t)C(s,t)\,\mathrm{d}s\,\mathrm{d}t = \langle f,g\rangle. \tag{A 3}$$

The theory of martingales makes its appearance in white-noise calculus through the concept of Wick-ordered random variables. We begin by defining the *Wick-ordered* exponential for any $f\in\mathcal{S}$ to be

$$:\!\mathrm{e}^{\phi(f)}\!: = \mathrm{e}^{\phi(f)-\|f\|^2/2}. \tag{A 4}$$

Then we have the following proposition.

Proposition A 1. *For any* $f\in\mathcal{S}$,

$$:\!\mathrm{e}^{\phi(f)}\!: = \sum_{n\geqslant 0}\frac{\|f\|^n}{n!}H_n\left(\frac{\phi(f)}{\|f\|}\right), \tag{A 5}$$

where H_n *is the* n*th Hermite polynomial. Moreover, for any* $f,g\in\mathcal{S}$,

$$E_\mu\left[H_n\left(\frac{\phi(f)}{\|f\|}\right)H_m\left(\frac{\phi(g)}{\|g\|}\right)\right] = \delta_{nm}(n!)\left(\frac{\langle f,g\rangle}{\|f\|\,\|g\|}\right)^n. \tag{A 6}$$

Proof. For any $a,b\in\mathbb{R}$ we have the absolutely convergent expansions

$$\begin{aligned}\mathrm{e}^{a-b^2/2} &= \mathrm{e}^{-(b-a/b)^2/2+a^2/(2b^2)}\\ &= \sum_{n\geqslant 0}\frac{b^n}{n!}(-1)^n\mathrm{e}^{a^2/(2b^2)}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(\mathrm{e}^{-x^2/2})|_{x=a/b}\\ &= \sum_{n\geqslant 0}\frac{b^n}{n!}H_n(a/b),\end{aligned} \tag{A 7}$$

where the last line makes use of the defining property of the Hermite polynomials. Using this with $a = \phi(f)$, $b = \|f\|$ gives (A 5). The orthogonality relation (A 6) follows by expanding the identity

$$E_\mu[{:}\mathrm{e}^{\phi(f)}{:}\,{:}\mathrm{e}^{\phi(g)}{:}] = \mathrm{e}^{\langle f,g\rangle} \tag{A 8}$$

in powers of f, g and comparing with the expansion derived from (A 5). ■

From this we define the *Wick-ordered* monomials as the random variables

$$:\phi(f)^n: = \|f\|^n H_n\left(\frac{\phi(f)}{\|f\|}\right), \tag{A 9}$$

so that linearity and convergent power series imply that

$$:\mathrm{e}^{\phi(f)}: = \sum_{n\geqslant 0} \frac{:\phi(f)^n:}{n!}.$$

Formally, we express the Wick-ordered monomials as

$$:\phi(f)^n: = \int_{\mathbb{R}_+^n} f(s_1)\cdots f(s_n)\,{:}\phi_{s_1}\cdots\phi_{s_n}{:}\,\mathrm{d}s_1\cdots\mathrm{d}s_n,$$

and from the orthogonalization (A 6) we can define the *Wick products*:

$$:\phi(f_1)\cdots\phi(f_n): = \int_{\mathbb{R}_+^n} f_1(s_1)\cdots f_n(s_n)\,{:}\phi_{s_1}\cdots\phi_{s_n}{:}\,\mathrm{d}s_1\cdots\mathrm{d}s_n. \tag{A 10}$$

The space $\mathcal{H}_n$ is defined to be the span of $\{:\phi(f)^n: \mid f \in \mathcal{S}\}$ and consists of precisely the random variables

$$(n!)^{-1}\int_{\mathbb{R}_+^n} \tilde{f}(s_1,\ldots,s_n)\,{:}\phi_{s_1}\cdots\phi_{s_n}{:}\,\mathrm{d}s_1\cdots\mathrm{d}s_n,$$

where $\tilde{f}$ lies in $\tilde{T}_n$, the L^2 completion of the space of *symmetric* functions in $\mathcal{S}^{\otimes n}$. These subspaces form an orthogonal decomposition of $L^2(\mathcal{S}', \mathcal{B}, \mu)$. In quantum field theory this is known as the Fock-space decomposition of the Hilbert space of quantum states into n particle sectors for $n \geqslant 0$.

The relationship between Wiener measure and the white-noise measure is to identify $\phi(f) \doteq \int f(s)\phi_s\,\mathrm{d}s$ with $W(f) \doteq \int f(s)\,\mathrm{d}W_s$ for all $f \in \mathcal{S} \subset L^2(\mathbb{R}_+)$. One is then led to the formal relation $\phi_s = \mathrm{d}W_s/\mathrm{d}s$, that is 'white noise' is the 'derivative' of Brownian motion. This identification extends to all orders in the chaos expansion via

$$(n!)^{-1}\int_{\mathbb{R}_+^n} \tilde{f}(s_1,\ldots,s_n)\,{:}\phi_{s_1}\cdots\phi_{s_n}{:}\,\mathrm{d}s_1\cdots\mathrm{d}s_n = \int_{\Delta_n} f(s_1,\ldots,s_n)\,\mathrm{d}W_{s_1}\cdots\mathrm{d}W_{s_n}, \tag{A 11}$$

where the bijection $\tilde{f} \leftrightarrow f$ between $\tilde{T}_n$ and $L^2(\Delta_n)$ is the restriction map and its inverse. Finally, this leads to the identification $L^2(\mathcal{S}', \mathcal{B}, \mu) \equiv L^2(\Omega, \mathcal{F}_\infty, P)$.

One useful consequence of Wick products and the chaos expansion is that conditional expectations can be handled systematically.

Proposition A 2. *Let $f \in L^2(\mathbb{R}_+)$ and $t \in [0, \infty)$. Then*

$$E[:\mathrm{e}^{\phi(f)}: \mid \mathcal{F}_t] = :\mathrm{e}^{\phi(f_{(t)})}:,$$
$$E[:(\phi(f))^n: \mid \mathcal{F}_t] = :(\phi(f_{(t)}))^n:, \tag{A 12}$$

where $[f_{(t)}](s) = \mathbb{1}(s \leqslant t) f(s)$. In other words, the processes $:\mathrm{e}^{\phi(f_{(t)})}:$, $:(\phi(f_{(t)}))^n:$ are martingales.

Appendix B. Generating functional for chaos coefficients

We derive theorem 2.1 in the notation of the white-noise calculus. To begin, we recall the definition of analyticity for a function between complex Banach spaces.

Definition B 1. Let $f : B \to C$, where B, C are complex Banach spaces. Then f is analytic at a point $x \in B$ if

(i) for any finite set $\{x_1, \ldots, x_n\} \subset B$ of points the function $F : \mathbb{C}^n \to \mathbb{C}$

$$F(\zeta_1, \ldots, \zeta_n) = f\left(x + \sum_i \zeta_i x_i\right) \tag{B 1}$$

is analytic at $0 \in \mathbb{C}^n$;

(ii) f is continuous at x.

Proof of theorem 2.1. One can check directly that the map $h \mapsto :\mathrm{e}^{\phi(h)}:$ is entire analytic between $L^2_{\mathbb{C}}(\mathbb{R}_+)$ and $L^2_{\mathbb{C}}(\mathcal{S}', \mathcal{B}, \mu)$. Similarly, $Z_X(h)$ is analytic for any $h \in L^2_{\mathbb{C}}(\mathbb{R}_+)$.

Finally, to verify (2.27), it is enough to take the expectation of the equation multiplied by $:\phi(g)^n:$ for arbitrary $g \in L^2(\Delta), n \geqslant 0$:

$$\begin{aligned} E[X{:}\phi(g)^n{:}] &= \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} Z_X(\lambda g)|_{\lambda=0} \\ &= n! \int_{\Delta_n} f_X^{(n)}(s_1, \ldots, s_n) g(s_1) \cdots g(s_n)\, \mathrm{d}s_1 \cdots \mathrm{d}s_n, \end{aligned} \tag{B 2}$$

whereas, by (A 6),

$$\begin{aligned} &E\left[\sum_{m \geqslant 0} (m!)^{-1} \int_{\mathbb{R}_+^m} f_X^{(m)}(s_1, \ldots, s_m) {:}\phi(s_1) \cdots \phi(s_m){:} \mathrm{d}s_1 \cdots \mathrm{d}s_m {:}\phi(g)^n{:}\right] \\ &\qquad = n! \int_{\mathbb{R}_+^m} g(s_1) \cdots g(s_n) f_X^{(n)}(s_1, \ldots, s_n)\, \mathrm{d}s_1 \cdots \mathrm{d}s_n. \end{aligned} \tag{B 3}$$

■

Appendix C. The exponential quadratic formula

Here we prove the formula (4.3) in the context of one-dimensional white-noise calculus. The proof extends easily to the multidimensional case.

Proof of proposition 4.1.

Step 1. First we note that $Y \in \mathcal{C}^+$ can be approximated in the Hilbert space $\mathcal{H}_{\leqslant 2}$ by random variables of the form

$$Y = A + \sum_{i=1}^{M} [b_i \phi(g_i) + c_i {:}\phi(g_i)^2{:}/2], \tag{C 1}$$

with $\{g_i\}$ a finite orthonormal set in $L^2(\Delta)$ and numbers $c_i > -1, b_i$. Then $X_i = \phi(g_i)$ form a collection of independent $N(0,1)$ random variables. For Y of this type, there is a factorization

$$E[\mathrm{e}^{-Y}] = \mathrm{e}^{-A}(2\pi)^{-M/2} \prod_i \int_{\mathbb{R}} \exp[-b_i x - (1+c_i)x^2/2 + c_i/2]\,\mathrm{d}x \tag{C 2}$$

into one-dimensional Gaussian integrals. Each integral gives the factor

$$(2\pi)^{1/2}(1+c_i)^{-1/2} \exp \tfrac{1}{2}[c_i + b_i^2(1+c_i)^{-1}], \tag{C 3}$$

leading to a formula for $E[\mathrm{e}^{-Y}]$ which agrees with (4.1) for Y of this form.

Step 2. Since the formula is true for Y in a dense subset of $\mathcal{C}^+$, it is now enough to prove that the map $Y \mapsto E[\mathrm{e}^{-Y}]$ is continuous in C provided the kernel C satisfies the stated conditions. By the definition of the Carleman–Fredholm determinant,

$$\det{}_2(1+C) = \exp(\mathrm{tr}[\log(1+C) - C]). \tag{C 4}$$

Since $\log(1+x) - x = \mathcal{O}(x^2)$ for $x \to 0$, we see that $\det_2(1+C)$ is well defined and continuous for a Hilbert–Schmidt operator $C > -1$. Therefore, the entire right-hand side of (4.1) is continuous in C for $Y \in \mathcal{C}^+$. ■

This research was supported by the Natural Sciences and Engineering Research Council of Canada and MITACS—Mathematics of Information Technology and Complex Systems, Canada.

We thank L. P. Hughston, D. Brody and L. C. G. Rogers for discussions and suggestions. We are grateful for the comments by the participants of the Financial and Insurance Mathematics talks at ETH Zurich, May 2003, the Probability seminars at the University of Cambridge, June 2003, and the Financial Mathematics and Applied Probability seminars at King's College London, July 2003, where this work was presented.

References

Brody, D. & Hughston, L. P. 2004 Chaos and coherence: a new framework for interest rate modelling. *Proc. R. Soc. Lond.* A **460**, 85–110.

Cheridito, P., Filipovic, D. & Kimmel, R. 2003 Market price of risk specifications for affine models: theory and evidence. Working paper available at http://www. princeton.edu/~rkimmel/research.htm.

Cox, J. C., Ingersoll Jr, J. E. & Ross, S. A. 1985 A theory of the term structure of interest rates. *Econometrica* **53**(2), 385–407.

Flesaker, B. & Hughston, L. P. 1996 Positive interest. *Risk Magazine* **9**(1), 46–49.

Glimm, J. & Jaffe, A. 1981 *Quantum physics: a functional integral point of view*. Springer.

Hughston, L. P. & Rafailidis, A. 2005 A chaotic approach to interest rate modelling. *Fin. Stochast.* **9**, 43–65.

Itô, K. 1951 Multiple Wiener integral. *J. Math. Soc. Jpn* **3**, 157–169.

Jamshidian, F. 1995 A simple class of square-root interest rate models. *Appl. Math. Fin.* **2**(1), 61–72.

Jin, Y. & Glasserman, P. 2001 Equilibrium positive interest rates: a unified view. *Rev. Fin. Stud.* **14**(1), 187–214.

Maghsoodi, Y. 1996 Solution of the extended CIR term structure and bond option valuation. *Math. Fin.* **6**, 89–109.

Nualart, D. 1995 *The Malliavin calculus and related topics.* Probability and its Applications. Springer.

Øksendal, B. 1996 An introduction to Malliavin calculus with applications to economics. Unpublished lecture notes available at http://www.nhh.no/for/dp/1996/index.htm.

Rogers, L. C. G. 1997 The potential approach to the term structure of interest rates and foreign exchange rates. *Math. Fin.* **7**(2), 157–176.

Rutkowski, M. 1997 A note on the Flesaker–Hughston model of the term structure of interest rates. *Appl. Math. Fin.* **4**, 151–163.

Wiener, N. 1938 The homogeneous chaos. *Am. J. Math.* **60**, 897–936.

Yor, M. 1992 *Some aspects of Brownian motion.* Lectures in Mathematics, ETH Zurich. Basel: Birkhäuser.