



Série des Documents de Travail

n° 2019-02

**Noncausal Affine Processes with Applications
to Derivative Pricing**

C.GOURIEROUX¹

Y.LU²

Les documents de travail ne reflètent pas la position du CREST et n'engagent que leurs auteurs.
Working papers do not reflect the position of CREST but only the views of the authors.

¹ CREST; University of Toronto; Toulouse School of Economics. E-mail:
Christian.Gourieroux@ensae.fr

² University of Paris 13

Noncausal Affine Processes with Applications to Derivative Pricing

Christian Gouriéroux⁽¹⁾ and Yang Lu⁽²⁾

January 2019

We gratefully acknowledge financial support of the Chair ACPR "Regulation and Systemic Risk", and the ERC DYSMOIA. We thank A. Monfort for helpful comments.

¹University of Toronto, Toulouse School of Economics and CREST.

²University of Paris 13.

Noncausal Affine Processes with Application to Derivative Pricing
Abstract

Linear factor models, where the factors are affine processes, play a key role in Finance, since they allow for quasi-closed form expressions of the term structure of risks. We introduce the class of noncausal affine linear factor models by considering factors that are affine in reverse time. These models are especially relevant for pricing sequences of speculative bubbles. We show that they feature much more complicated non affine dynamics in calendar time, while still providing (quasi) closed form term structures and derivative pricing formulas. The framework is illustrated with zero-coupon bond and European call option pricing examples.

Keywords : Derivative Pricing, Term Structure, Affine Process, Noncausal Process, Speculative Bubble, Reverse Time.

1 Introduction

Linear factor models, where the factors are affine processes, play a key role in Finance, since they allow for quasi-closed form expressions of the term structures of risks [Duffie et al. (2003)]. These term structures can be computed either under the risk-neutral probability for pricing purpose, or under the historical probability for prediction purpose. They can concern the riskfree interest rates, leading to the affine term structure models (ATSM) [Duffie, Kan (1996)], the corporate rates and default events [see e.g. Gouriéroux, Monfort, Polimenis (2006)], as well as the prices of European derivatives written on a given index.

This paper introduces a new class of models with quasi-closed form expressions of the term structures. These are linear factor models, where the factors are noncausal affine processes, that are affine processes in reverse time, and feature a much more complicated nonlinear dynamics in calendar time. There is a recent growing literature on noncausal processes, which seem to better fit financial and economic time series [see e.g. Lanne, Saikkonen (2013), Gouriéroux, Jasiak (2016), Gouriéroux, Zakoian (2017)]. Indeed noncausal processes are appropriate for modelling sequences of speculative bubbles, that are increasing patterns followed by a burst, and such patterns are frequently encountered in financial series such as commodity prices and cryptocurrencies series [see e.g. Lee, Phillips (2016), Gouriéroux, Jasiak, Monfort (2017)].

The plan of the paper is as follows : in Section 2 we derive the quasi-closed form expression of the joint conditional Laplace transform (LT) of the future and cumulated future values of a noncausal affine process. This derivation is based on a Bayes formula applied to LT's, i.e. the Bartlett's formula. Section 3 considers the noncausal linear autoregressive processes of order 1 and discusses carefully their use for predicting and pricing speculative bubbles. Section 4 applies the methodology to the term structure of interest rates, and Section 5 proposes applications to the pricing of European calls. Section 6 concludes. Proof are gathered in Appendices.

2 Unconditional and conditional LTs

2.1 Properties in reverse time

Let us consider a n -dimensional stationary noncausal affine process. We define this process in discrete time and assume that it is a Markov process in reverse time : thus the conditional distribution of the current value y_t given the future path $\underline{y_{t+1}} = (y_{t+1}, y_{t+2}, \dots)$ is equal to its conditional distribution given the closest future observation y_{t+1} only.

The distribution of the process is characterized by the noncausal transition, which can be defined by the noncausal LT :

$$\mathbb{E}[\exp(uy_t)|y_{t+1}] = \exp[a(u)y_{t+1} + b(u)], \quad (2.1)$$

where $u \in \mathcal{U} \subset \mathbb{C}^n$, is a complex argument such that the expectation of the left hand side of equation (2.1) exists and uy_t (resp. $a(u)y_{t+1}$) denotes the inner product between u and y_t (resp. $a(u)$ and y_{t+1}). The logarithm of this conditional LT is an affine function of y_{t+1} . This is the standard condition for an affine process³ [see Duffie, Filipovic, Schachermayer (2003), Darolles, Gouriéroux, Jasiak (2006)]. The n -dimensional function $a(u)$ summarizes the nonlinear serial dependence and is the log-LT of an infinitely divisible distribution, whereas the scalar function $b(u)$ is the log-LT of an appropriate innovation (see Section 3). The process is stationary if function a satisfies the following stability condition [see Darolles, Gouriéroux, Jasiak (2006), Proposition 6] :

$$\lim_{h \rightarrow \infty} a^{oh}(u) = 0, \quad \forall u \in \mathcal{U}, \quad (2.2)$$

where o denotes the composition of functions.

Under the stationarity, the unconditional LT of y_t defined by :

$$\mathbb{E}[\exp(uy_t)] = \exp c(u), \quad (2.3)$$

is such that :

$$b(u) = c(u) - c[a(u)]. \quad (2.4)$$

³Called Compound Autoregressive (CaR) process in discrete time.

By standard properties of affine processes, we derive the (quasi) closed form expression of the LT of the process conditional on the future (see Appendix 1) :

Proposition 1. *We have :*

$$\mathbb{E}\left\{\exp[uy_t + vy_{t+H} + w(y_{t+1} + \dots + y_{t+H})] | y_{t+H}\right\} = \exp\left\{[v + w + a(u, w, H)]y_{t+H} + b(u, w; H)\right\},$$

where functions a and b satisfy the recursive equations :

$$a(u, w; H) = a[w + a(u, w; H - 1)],$$

$$b(u, w; H) = b[w + a(u, w; H - 1)] + b(u, w; H - 1), H \geq 2,$$

with initial conditions :

$$a(u, w; 1) = a(u), \quad b(u, w; 1) = b(u).$$

2.2 Properties in calendar time

The use of an exponential transform of both y_{T+H} and cumulative values $y_{t+1} + \dots + y_{t+H}$ is needed for pricing interest rates and derivatives [see e.g. Duffie et al. (2003), Gouriéroux et al. (2006)]. The introduction of the additional term uy_t is useful for applying a Bayes formula to the LT (see Lemma 2 below).

We deduce from Proposition 1 the unconditional joint LT.

Corollary 1 : We have :

$$\mathbb{E}\{\exp[uy_t + vy_{t+H} + w(y_{t+1} + \dots + y_{t+H})]\} = \exp\{c[v + w + a(u, w; H)] + b(u, w; H)\},$$

where u, v, w are complex arguments such that the conditional expectation exists.

Proof. By the iterated expectation theorem, we get :

$$\begin{aligned}
& \mathbb{E}\{\exp[uy_t + vy_{t+H} + w(y_{t+1} + \dots + y_{t+H})]\} \\
&= \mathbb{E}\{\mathbb{E}\{\exp[uy_t + vy_{t+H} + w(y_{t+1} + \dots + y_{t+H})]|y_{t+H}\}\} \\
&= \mathbb{E}\{\exp\{[v + w + a(u, w; H)]y_{t+H} + b(u, w; H)\}\}, \text{ by Proposition 1,} \\
&= \exp\{c[v + w + a(u, w; H) + b(u, w; H)]\}.
\end{aligned}$$

Corollary 1 follows. □

Let us now explain how to deduce from this joint unconditional LT the conditional LT of $(y_{t+H}, y_{t+1} + \dots + y_{t+H})$ given the past $\underline{y}_t = (y_t, y_{t-1}, \dots)$. We use the two following lemmas :

Lemma 1. *The stationary Markov process in reverse time is also a stationary Markov process in calendar time.*

This result is proven in Cambanis, Fakhre-Zakeri (1994).

Thus we only have to derive the expression of the conditional LT of $(y_{t+H}, y_{t+1} + \dots + y_{t+H})$ given y_t only. Its expression is a consequence of the next lemma due to Bartlett (1938)⁴.

Lemma 2. *Let y and x be random vectors in \mathbb{R}^n and \mathbb{R}^m with joint LT : $\psi(u, v) = \mathbb{E}[\exp(uy + vx)]$. Then the conditional LT of x given y is :*

$$\mathbb{E}[\exp(vx)|y] = \frac{1}{(2\pi)^n} \frac{\int \psi(iz, v) \exp(-izy) dz}{\ell(y)},$$

where $\ell(y)$ is the marginal (stationary) p.d.f. of y and the integral is over \mathbb{R}^n .

⁴This lemma has been used by Bates (2006) to define a filtering algorithm based on LT for semi-affine state space models.

By applying Lemmas 1 and 2, with $y \equiv y_t, x = (y'_{t+H}, y'_{t+1} + \dots + y'_{t+H})', m = 2n$, we get the following proposition :

Proposition 2. *The causal LT is:*

$$\begin{aligned} & \mathbb{E}\{\exp[vy_{t+H} + w(y_{t+1} + \dots + y_{t+H})]|y_t\} \\ &= \frac{1}{(2\pi)^n \ell(y_t)} \int \exp\{c[v + w + a(iz, w; H)] + b(iz, w; H)\} \exp(-izy_t) dz, \end{aligned} \quad (2.5)$$

where $\ell(y_t)$ is the marginal density of y_t .

Finally, the marginal density at the denominator can be computed from the unconditional LT by Fourier inverse as :

$$\ell(y_t) = \frac{1}{(2\pi)^n} \int \exp[c(iz)] \exp(-izy_t) dz. \quad (2.6)$$

From (2.5)-(2.6), we get a quasi-closed form expression of the causal conditional LT, which can be used to derive quasi-closed form expressions for the term structures of either interest rates, or derivative prices, as seen from the examples below.

Remark 1 : The numerical derivation of the causal LT can be difficult if the dimension n is too large. This difficulty is usually circumvented by assuming that the short interest rate, or the return, is a linear combination of independent factors, $y_t = Ay_t^*$, say, where each of the y^* is either a causal, or a noncausal affine process [see e.g. Singleton (2009), Dai, Singleton (2003) for this practice with causal affine Cox, Ingersoll, Ross processes].

2.3 A typology of noncausal affine processes

The inversion technique to pass from properties in reverse time to properties in calendar time is useful if the noncausal affine process does not admit also a causal affine representation. In this respect, three cases can arise :

- i)* The noncausal affine process is reversible, that is, it admits the same affine representation in both reverse and calendar times.
- ii)* The noncausal affine process admits a causal affine representation, but with different functional measures of dependence, that are different functions $a(\cdot)$.

iii) The noncausal affine process has no causal affine representation.

Let us provide some examples.

i) A stationary Gaussian autoregressive process or a time discretized one dimensional affine diffusion process⁵ as the autoregressive gamma (ARG) process⁶ are reversible.

ii) A stationary Gaussian Vector Autoregressive process is affine in both reverse and calendar time, and is not reversible whenever the autoregressive matrix does not satisfy a symmetry property.

For one-dimensional positive processes a non reversible process with both causal and noncausal affine representations cannot exist as shown in the Proposition below :

Proposition 3. *A univariate nonnegative process, which is weakly ergodic in reverse and calendar time and has both a causal and a noncausal affine representation, if and only if it is reversible.*

Proof. See Appendix 2. □

Thus in general a noncausal affine process is not affine in calendar time, that is, the causal log-LT is non-affine in y_t .

3 Noncausal linear autoregressive processes

The linear autoregressive (AR) processes are a special class of affine processes, corresponding to the case where function a is linear: $a(u) = \rho u$. They can be considered as the discrete time analogue of the non-Gaussian Ornstein-Uhlenbeck introduced by Barndorff-Nielsen and Shephard (2001) in the finance literature. Let us consider a one-dimensional AR process of order 1 written in reverse time.⁷ The dynamics can be defined by the noncausal regression model :

⁵A one dimensional diffusion process is reversible, i.e has the same dynamics in direct and reverse time.

⁶The ARG process [see Gouriéroux and Jasiak (2006)] is the time discretized Cox, Ingersoll, Ross process.

⁷The extension to Vector Autoregressive (VAR) process is straightforward and omitted.

$$y_t = \rho y_{t+1} + \varepsilon_t, \quad (3.1)$$

with $|\rho| < 1$ and a sequence of i.i.d. variables $\varepsilon_t, t = 1, \dots$

When $|\rho| < 1$, the recursive equation (3.1) has a strictly stationary solution, if $\mathbb{E}|\varepsilon_t|^s < \infty$, for a power $s > 0$ [Gourieroux, Zakoian (2017)]. In particular the ε_t 's may have fat tails.

3.1 The noncausal LT

The noncausal transition of this process is characterized by its LT :

$$\mathbb{E}[\exp(uy_t)|y_{t+1}] = \exp[u\rho y_{t+1} + b(u)], \quad (3.2)$$

where $b(u) = \log \mathbb{E}[\exp(u\varepsilon_t)]$ is the log-LT of the noncausal innovation ε_t . This is a noncausal affine process, where $a(u) = \rho u$ is linear in argument u and $b(u)$ characterizes the distribution of ε_t .

The corresponding noncausal conditional LT in Proposition 1 can be derived explicitly (see Appendix 3). We have :

$$\begin{aligned} & \mathbb{E}\{\exp[uy_t + vy_{t+H} + w(y_{t+1} + \dots + y_{t+H})]|y_{t+H}\} \\ &= \exp\{[v + w + a(u, w; H)]y_{t+H} + b(u, w; H)\}, \end{aligned} \quad (3.3)$$

where functions a and b are given by :

$$\begin{cases} v + w + a(u, w; H) = u\rho^H + v + w\frac{1 - \rho^H}{1 - \rho}, \\ b(u, w; H) = b(u) + b(u\rho + w\frac{1 - \rho}{1 - \rho}) + \dots + b(u\rho^{H-1} + w\frac{1 - \rho^{H-1}}{1 - \rho}). \end{cases} \quad (3.4)$$

Thus the causal LTs depend on the scalar parameter ρ and function b , or equivalently on ρ and function c by applying Proposition 2.

3.2 Examples

The general formulas of Sections 2.2 and 3.1 can be more explicit for special choices of the noncausal dynamics of the process, within the family of noncausal linear AR models. This dynamics can either be characterized by

scalar ρ and function b , which can be the log-LT of any arbitrary distribution, or alternatively it can be characterized by scalar ρ and function c . In order for this specification to be self-consistent, the stationary distribution should belong to the family of self-decomposable distributions [see e.g. Feller (1968), Barndorff-Nielsen and Shephard (2001), Carr et al. (2007)], which is an important sub-family of infinitely divisible distributions. Roughly speaking, the distribution of a variable Y is self-decomposable if and only if for any $\rho \in]0, 1[$, there exists another variable ϵ independent of Y such that Y and $\rho Y + \epsilon$ have the same distribution. Thus, by definition, if process (y_t) is stationary, then its stationary distribution is necessarily self-decomposable. Conversely, any self-decomposable distribution can be chosen as the stationary distribution of process (y_t) .

In the following let us first look at two (Cauchy and Lévy) examples with alpha-stable noncausal innovation. For these cases the stationary distribution belongs to the same parametric family as the error distribution. Then we introduce additional examples (exponential, gamma, and Laplace) in which we specify directly the stationary distribution. We see that this latter approach can lead to simple expressions for causal conditional LT.

3.2.1 Noncausal AR(1) Cauchy process

Let us assume that the noncausal innovations follow a Cauchy distribution. This distribution is continuous with probability density function (p.d.f.) : $f(\varepsilon) = \frac{1}{\pi} \frac{1}{1 + \varepsilon^2}$, and characteristic function⁸ : $\mathbb{E}[\exp(iz\varepsilon)] = \exp b(iz) = \exp(-|z|)$. The associated noncausal AR(1) process (y_t) is stationary and the stationary distribution of process (y_t) is a re-scaled Cauchy distribution, with characteristic function :

$$\mathbb{E}[\exp(izY)] = \exp c(iz) = \exp\left(-\frac{|z|}{1 - |\rho|}\right), \quad (3.5)$$

and p.d.f. :

$$\ell(y) = \frac{1 - |\rho|}{\pi} \frac{1}{1 + (1 - |\rho|)^2 y^2}.$$

⁸For the Cauchy distribution, the real LT is not defined, but the characteristic function of a distribution is always defined.

The joint characteristic function of $y_t, y_{t+H}, y_{t+1} + \dots + y_{t+H}$ is obtained for imaginary arguments $u = iz, v = i\nu_0, \omega = i\nu_1$, say, where z, ν_0, ν_1 are real numbers. We get :

$$\begin{aligned} & \mathbb{E} \left\{ \exp[izy_t + i\nu_0 y_{t+H} + i\nu_1(y_{t+1} + \dots + y_{t+H})] \right\} \\ &= \exp \left\{ - \left| z\rho^H + \nu_0 + \nu_1 \frac{1 - \rho^H}{1 - \rho} \right| \frac{1}{1 - |\rho|} - \sum_{h=0}^{H-1} \left| z\rho^h + \nu_1 \frac{1 - \rho^h}{1 - \rho} \right| \right\}. \end{aligned}$$

By applying Proposition 2, we deduce the causal characteristic function :

Corollary 2 : The causal characteristic function of a noncausal AR(1) Cauchy process is :

$$\begin{aligned} & \mathbb{E} \{ \exp[i\nu_0 y_{t+H} + i\nu_1(y_{t+1} + \dots + y_{t+H})] | y_t \} \tag{3.6} \\ &= \frac{1 + (1 - |\rho|)^2 y^2}{2(1 - |\rho|)} \int \exp \left\{ - \left| z\rho^H + \nu_0 + \nu_1 \frac{1 - \rho^H}{1 - \rho} \right| \frac{1}{1 - |\rho|} - \sum_{h=0}^{H-1} \left| z\rho^h + \nu_1 \frac{1 - \rho^h}{1 - \rho} \right| - izy_t \right\} dz. \end{aligned}$$

The computations are greatly simplified when we focus on the causal prediction of y_{t+H} , that is, when $\nu_1 = 0$. For instance in (3.6), the sum $\sum_{h=0}^{H-1} \left| z\rho^h + \nu_1 \frac{1 - \rho^h}{1 - \rho} \right|$ equals $|z| \frac{1 - |\rho|^H}{1 - |\rho|}$.

The prediction formulas can also be derived directly by noting that :

$$\begin{aligned} y_t &= \rho^H y_{t+H} + \rho^{H-1} \varepsilon_{t+H-1} + \dots + \varepsilon_t \\ &\equiv \rho^H y_{t+H} + u_t, \end{aligned} \tag{3.7}$$

where the variables y_{t+H} and u_t are independent, Cauchy distributed with scales $\eta = \frac{1}{1 - |\rho|}$ and $\sigma = \frac{1 - |\rho|^H}{1 - |\rho|}$, respectively. Equation (3.7) is a linear regression with both Cauchy error u_t and regressor y_{t+H} . The conditional distribution of y_{t+H} given y_t is easily derived by the Bayes formula (see Appendix 4).

Proposition 4. For the noncausal AR(1) Cauchy, the conditional causal density at horizon H is :

$$\begin{aligned}
f(y_{t+H}|y_t) &= \frac{1}{\pi} \frac{[1 + |\rho|(1 - |\rho|^H)](1 - |\rho|)}{(1 - |\rho|^H)} \\
&\{1 + (y_t \frac{1 - |\rho|}{1 + |\rho|(1 - |\rho|^H)})^2\}^{-1} \{1 + (y_t - \rho^H y_{t+H})^2 (1 - |\rho|^2)^2\}^{-1} \{1 + y_{t+H}^2 \frac{(1 - |\rho|)^2}{(1 - |\rho|^H)^2}\}^{-1}.
\end{aligned} \tag{3.8}$$

The noncausal AR(1) Cauchy process is not integrable. It has no marginal mean, no mean conditional to the future. But the expression of the causal transition (3.8) shows that the causal conditional moments exist up to power 3. It has been shown [Gourieroux, Zakoian (2017)], that its conditional causal mean is $E(y_{t+H}|y_t) = y_t$, that is the martingale condition. Thus, if (y_t) is a value process and the riskfree rate is zero⁹, formula (3.8) applied to the risk-neutral distribution provides the term structure of state price densities at date t . As expected this process is not affine in calendar time. Indeed the conditional variance in calendar time [Gourieroux, Zakoian (2017), Proposition 3.5] :

$$V(y_{t+1}|y_t) = \left(\frac{1}{|\rho|} - 1\right)y_t^2 + \frac{1}{|\rho|(1 - |\rho|)},$$

is not affine¹⁰ in y_t .

3.2.2 Noncausal AR(1) Lévy process

Let us now assume that the noncausal innovation ε_t follows a Lévy distribution with zero drift and scale σ . Its density is :

$$f(\varepsilon) = \sqrt{\frac{\sigma}{2\pi}} \frac{1}{\varepsilon^{3/2}} \exp\left(-\frac{\sigma}{2\varepsilon}\right) \mathbb{1}_{\varepsilon>0}. \tag{3.9}$$

Variable ε_t is positive and its real LT is :

$$\mathbb{E}[\exp(-u\varepsilon)] = \exp(-\sqrt{2\sigma u}), \quad \forall u \geq 0.$$

Let us assume that $\rho > 0$, then process (y_t) is also positive and it is shown in Appendix 5 that:

⁹Otherwise the reasoning has to be applied to the discounted prices.

¹⁰A causal affine process is such that all its existing conditional cumulants are affine in y_t , in particular the conditional variance.

- The stationary distribution of this process is still Lévy with zero drift and scale $\sigma/[1 - \sqrt{\rho}]^2$.
- The joint LT is:

$$\mathbb{E}[e^{-uy_t - vy_{t+1}}] = \exp\left(-\frac{\sqrt{2\sigma(\rho u + v)}}{1 - \sqrt{\rho}} - \sqrt{2\sigma u}\right).$$

- The transition density in calendar time is:

$$\ell(y_{t+1}|y_t) = \sqrt{\frac{\sigma}{2\pi}} \exp\left\{-\frac{\sigma}{2(y_t - \rho y_{t+1})} - \frac{\sigma}{2(1 - \sqrt{\rho})^2} \left(\frac{1}{y_{t+1}} - \frac{1}{y_t}\right)\right\} \frac{y_t^{3/2} \mathbb{1}_{y_t > \rho y_{t+1}}}{y_{t+1}^{3/2} (y_t - \rho y_{t+1})^{3/2}}. \quad (3.10)$$

Since the joint LT of (y_t, y_{t+1}) is not symmetric in y_t, y_{t+1} , the process is not reversible. Then, by Proposition 3, the process has no causal affine representation.

Since the causal conditional distribution has a bounded support $[0, y_t/\rho]$, all its moments are finite, on the contrary to the noncausal conditional integer moments, which are all infinite. In particular, it has been shown by Cambanis, Fakhre-Zakeri (1994) that the first order moment is:

$$\mathbb{E}[y_{t+1}|y_t] = \frac{1}{\sqrt{\rho}} y_t.$$

Thus process (y_t) is a stationary sub-martingale. Finally, the conditional higher order moments do not seem to have explicit expressions, but we can check that the conditional distribution is not symmetric. Indeed, simple calculation shows that the conditional density defined in (3.10) satisfies $\ell(y|y_t) = \ell(\frac{y_t}{\rho} - y|y_t)$ for any $y \in [0, \frac{y_t}{\rho}]$, if and only if $\rho = 1/4$. If $\rho > 1/4$ (resp. $\rho < 1/4$), the mean $\frac{1}{\sqrt{\rho}} y_t$ is larger (resp. smaller) than the middle of the interval $\frac{y_t}{2\rho}$, in which case the conditional distribution is skewed to the left (resp. right) regardless of the value of y_t .

3.2.3 Noncausal AR(1) exponential process

Let us consider a noncausal AR(1) process with :

$$a(u) = \rho u, \quad b(u) = \log\left(\frac{1 - \rho u}{1 - u}\right).$$

where parameter ρ is such that $0 \leq \rho < 1$. Thus process (y_t) is again positive. The causal analogue of this process has been introduced by Gaver, Lewis (1980). Its marginal distribution is standard exponential, i.e. $c(u) = -\log(1 - u)$, whereas its conditional distribution given the future is, up to a translation of ρy_{t+1} , the mixture of a point mass at zero and an exponential distribution with weights ρ and $1 - \rho$, respectively. Indeed we have :

$$\exp b(u) = \rho + (1 - \rho) \frac{1}{1 - u}.$$

As an illustration, Figure 1 displays a simulated trajectory of a noncausal exponential process. The parameter ρ is set to $\rho = 0.2, 0.8$. These simulated trajectories show a sequence of explosive bubbles, that are increasing pattern, followed by a burst.

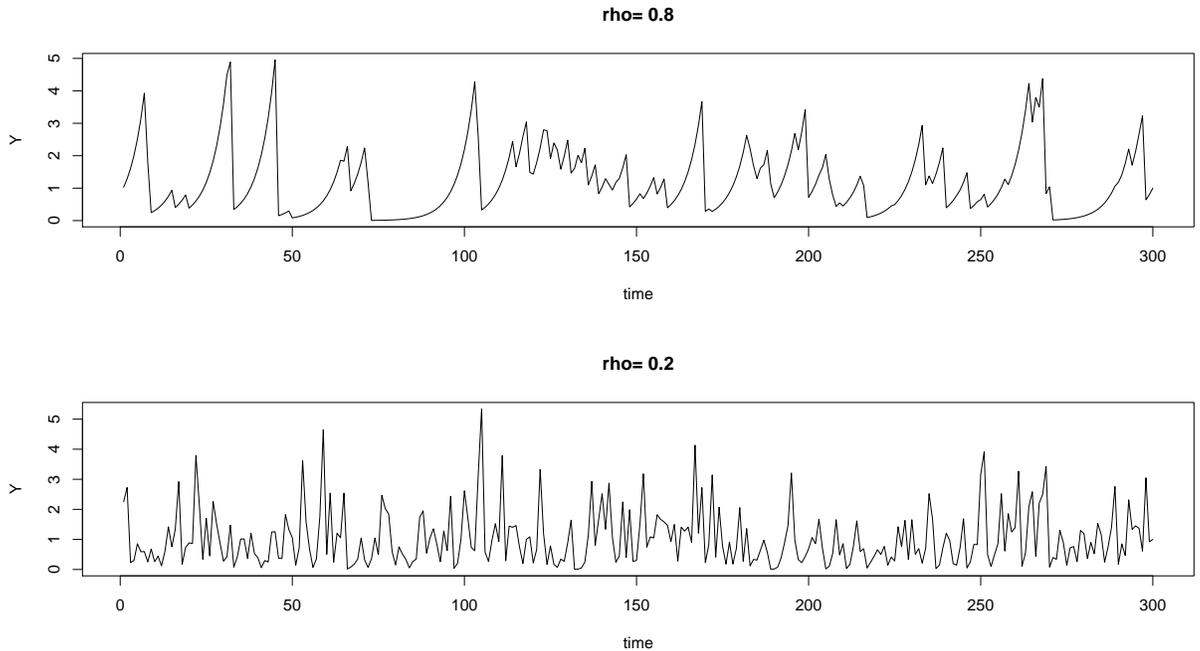


Figure 1: Trajectory of a noncausal gamma autoregressive process for two values of ρ , $\rho = 0.2, 0.8$.

The conditional distribution of y_{t+1} given y_t is easily derived either by the Bayes formula, or by applying residual calculus. More precisely we have

(see Appendix 6) :

$$\mathbb{E}[\exp(vy_{t+1})|y_t] = \frac{1-\rho}{1-\rho-v} + \frac{-v}{1-\rho-v} \exp\left[-\frac{y_t(1-v)}{\rho} + y_t\right]. \quad (3.11)$$

Thus the process is not affine in calendar time. The conditional causal transition is the mixture of a point mass at y_t/ρ :

$$\mathbb{P}[y_{t+1} = \frac{y_t}{\rho} | y_t] = \rho \exp\left[-\frac{(1-\rho)y_t}{\rho}\right],$$

and a continuous component with density :

$$f(y_{t+1}|y_t) = (1-\rho) \exp[-(1-\rho)y_{t+1}] \mathbb{1}_{y_{t+1} < \frac{y_t}{\rho}},$$

which is the exponential distribution $\gamma(1, 1-\rho)$, right censored at $y_{t+1} = \frac{y_t}{\rho}$.

Equation (3.11) can be used to derive the successive conditional cumulants (see Appendix 6.4). Figure 2 displays the conditional skewness and excess kurtosis as functions of y_t , for $\rho = \frac{1}{2}$.

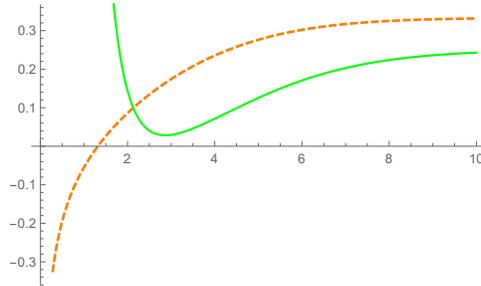


Figure 2: Conditional skewness (dashed line) and excess kurtosis (full line) as functions of y_t .

The conditional excess kurtosis is always positive, whereas the conditional skewness is negative for small values of y_t , but positive for large values of x .

Finally, the same residue method can be used to compute the LT of the cumulative factor $\mathbb{E}[e^{w(y_{t+1}+y_{t+2}+\dots+y_{t+H})}|y_t]$, for any horizon $H \in \mathbb{N}$ (see Appendix A.6.3).

The noncausal AR(1) exponential process can be extended to a noncausal gamma mixture process, whose stationary distribution is gamma with an arbitrary, real shape parameter $\alpha > 0$. This extension is discussed in Section 4.2.

3.2.4 Noncausal AR(1) Laplace process

Let us assume that process (y_t) has a symmetric Laplace stationary distribution with density $\ell(y_t) = \frac{1}{2\sigma} e^{-|y_t|/\sigma}$ and LT:

$$\mathbb{E}[e^{uy_t}] = \frac{1}{1 - \sigma^2 u^2}, \quad \forall |u| < 1/\sigma,$$

where parameter $\sigma > 0$. This distribution has a variance $2\sigma^2$; thus, when y_t is an asset return, σ can be interpreted as a volatility parameter. This LT is analytic in the complex region $\{u \in \mathbb{C}, |u| < 1/\sigma\}$ and thus the above equation is valid even for any complex number u satisfying $|u| < 1/\sigma$. In particular, the marginal moments of y_t exists at any order. The Laplace distribution has been recently proposed as a model for forecasting asset returns by Taylor (2017), and is a special case of the Variance gamma (VG) distribution [see e.g. Madan, Carr and Chang (1998) for use of the VG distribution for option pricing].¹¹ This distribution is self-decomposable [see e.g. Madan, Carr and Chang (1998)], and is therefore compatible with a noncausal linear AR model. Under the Laplace assumption of the stationary distribution, the distribution of the noncausal innovation (ϵ_t) has the LT:

$$\mathbb{E}[e^{u\epsilon_t}] = \frac{1 - \rho^2 \sigma^2 u^2}{1 - \sigma^2 u^2} = \rho^2 + (1 - \rho^2) \frac{1}{1 - \sigma^2 u^2}. \quad (3.12)$$

In other words (ϵ_t) has a mixture distribution with a point mass at zero with weight ρ^2 , and a continuous Laplace-distributed component, with weight $1 - \rho^2$. In Appendix 7 we prove that the causal conditional LT is:

$$\begin{aligned} \mathbb{E}[e^{vy_{t+1}} | y_t] &= \frac{(-1)^{\mathbb{1}_{y_t > 0}} 2\pi i \sigma}{\pi \exp(-|y_t|/\sigma)} \left\{ \frac{1 - \rho^2}{2i\sigma[1 - (\sigma v - \rho)^2]} e^{y_t/\sigma} \mathbb{1}_{y_t < 0} + \frac{1 - \rho^2}{-2i\sigma[1 - (\sigma v + \rho)^2]} e^{-y_t/\sigma} \mathbb{1}_{y_t \geq 0} \right. \\ &\quad + \frac{1 - (1 - \sigma v)^2}{[1 - \frac{(1 - \sigma v)^2}{\rho^2}](-2\rho i \sigma)} e^{-\frac{1/\sigma - v}{\rho} y_t} \mathbb{1}_{y_t \operatorname{Im}(-\frac{i}{\rho}(1-v)) < 0} \\ &\quad \left. + \frac{1 - (1 + \sigma v)^2}{[1 - \frac{(1 + \sigma v)^2}{\rho^2}](2\rho i \sigma)} e^{-\frac{-1/\sigma - v}{\rho} y_t} \mathbb{1}_{y_t \operatorname{Im}(-\frac{i}{\rho}(-1-v)) < 0} \right\}, \quad (3.13) \end{aligned}$$

for any complex number v such that $|v| < 1/\sigma$. This formula can be used to compute the conditional cumulants (see e.g. Appendix A.6.4. for the

¹¹The simplest, symmetric variance gamma distribution has a LT $\mathbb{E}[e^{uy_t}] = \frac{1}{(1 - \sigma^2 u^2)^\delta}$. When $\delta = 1$ we get the symmetric Laplace distribution.

noncausal exponential process). This is done using Mathematica and Figure 3 plots the conditional skewness and conditional excess kurtosis, for $\rho = 0.2$ and $\sigma = 1$

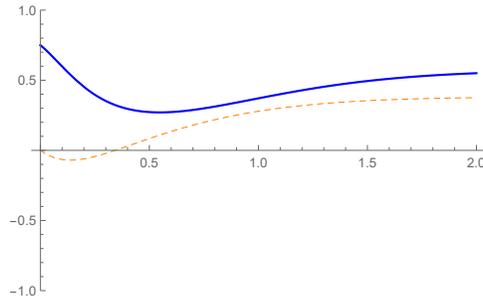


Figure 3: Conditional skewness (dashed line) and conditional excess kurtosis (full line) as a function of y_t .

Again, we see that the conditional skewness changes the sign when y_t increases, whereas the excess kurtosis is always positive. We also remark that when $y_t = 0$, the conditional skewness is zero. This is expected since the Laplace distribution is symmetric with respect to origin.

4 Application to the term structure of interest rate

This section shows how noncausal affine processes can be used to price risk-free bond prices. Let us denote by $B(t, h)$ the price at date t of the zero-coupon bond with time-to-maturity (term) h and by $r_t(h) = -\frac{1}{h} \log B(t, h)$ the associated geometric rate. Under the absence of arbitrage opportunity, we have :

$$B(t, h) = \mathbb{E}_t[\exp(-r_t - \dots - r_{t+h-1})], \quad (4.1)$$

where $r_t = r_t(1)$ is the short term interest rate and the conditional expectation is computed under the risk-neutral probability.

It is usual to assume that the short term interest rate is an affine combination of underlying factors :

$$r_t = w^* y_t + w_0, \text{ say.} \quad (4.2)$$

Then Proposition 2 can be directly applied to derive the term structure of interest rates, when the factors y satisfy a noncausal affine dynamic model under the risk-neutral probability.

The main difference between this noncausal affine term structure model and a standard affine term structure model [Duffie et al. (2003)] is the following one : In both causal and noncausal affine specifications the short term interest rate is an affine function of underlying factors. However, in the standard affine term structure model, the rates $r(t, h)$ are also affine function of these factors. This is no longer the case for noncausal affine models, where they can depend nonlinearly on the current factor values y_t :

$$r(t, h) \equiv g(y_t, h), \quad \text{say.} \quad (4.3)$$

Since function g is deterministic, it is possible to deduce the underlying factors y_t from the rates $r(t, h)$ where h varying, by “inverting” function g . In the standard causal affine framework, this inverse is affine and the components of y_t are interpreted as rates of mimicking coupon bonds. When the factors are noncausal affine, each component of y_t is a nonlinear function of the interest rates, and this standard interpretation is no longer valid. For the same reason shocks on the factors have effect in the term structure of interest rates, which can no longer be interpreted in terms of level, slope and curvature effects.

4.1 An illustration with noncausal exponential process

Let us apply the above procedure by assuming that the factor (y_t) is univariate and follows the noncausal AR(1) exponential model with $\rho = 0.8$. For illustration purpose we assume equation (4.2) with $w^* = 1$ and $w_0 = 0$ such that $r_t = y_t$. Then we use the residue method to compute the bond price $B(t, H)$ and the corresponding yield $r(t, H)$ for $H = 1, \dots, 60$, and different values of y_t . These term structures are illustrated in Figure 4. We get different patterns. When y_t is small, the valuation takes into account the occurrence of a future bubble and we get a hump shape term structure. When y_t is large, we are still on the explosive part of the bubble and the term structure is decreasing. Moreover, the larger y_t , the larger the rate of decrease.

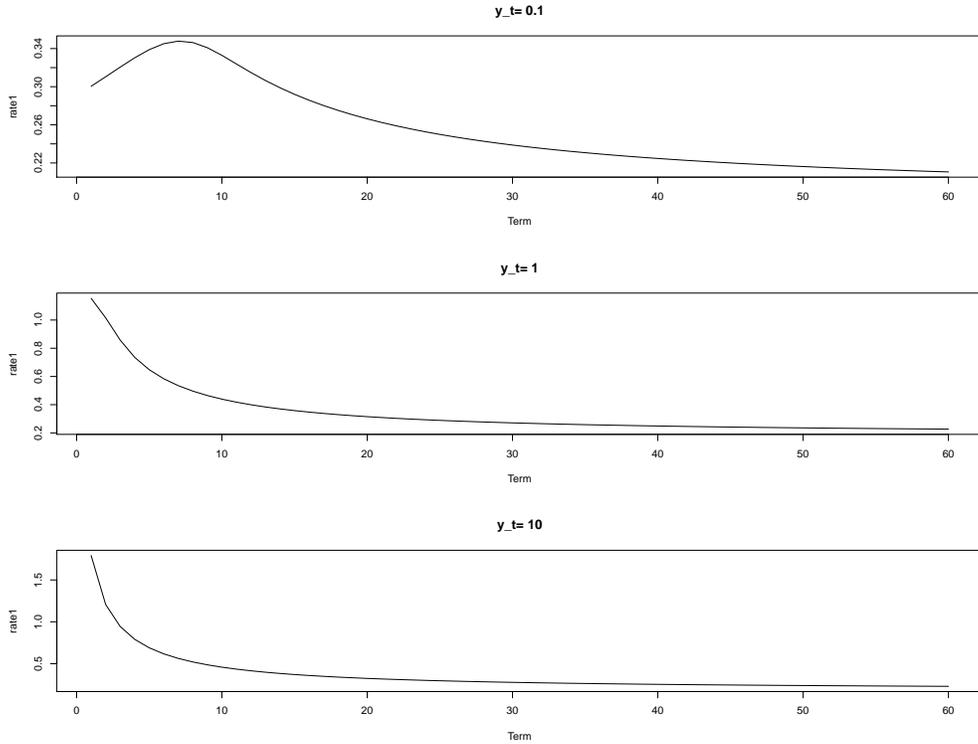


Figure 4: Term structure of interest rates $r(t, H)$: upper panel: $y_t = 0.1$, middle panel: $y_t = 1$, lower panel: $y_t = 10$. Parameter ρ is set to $\rho = 0.8$.

In Figure 5 we plot the value of the rate $r(t, H)$ against the short rate $r_t = r(t, 1)$ at two different horizons $H = 2, 8$. Since the factor is non affine in calendar time, the rates with term $H = 2, 8$ are nonlinear functions of the short term rate. More precisely, we prove in Appendix 6.3 that the bond price $B(t, H)$ is a linear combination of $H + 1$ exponential functions of the short rate $r_t = y_t$. Thus the yield $r(t, H)$ cannot be linear.

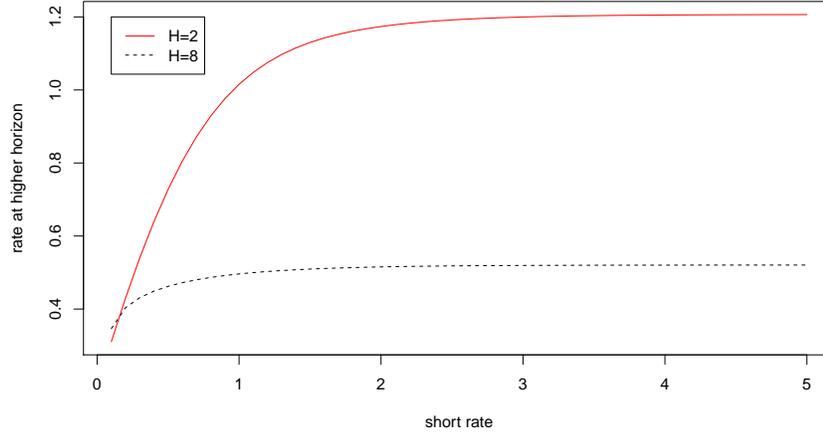


Figure 5: Relationship between the rates $r(t, H)$, with $H = 2$ and $H = 8$ and the short rate $r_t = r(t, 1)$.

These term structures can be compared with the term structures obtained under the reverse time model, that is for a causal AR(1) exponential factor with the same parameter $\rho = 0.8$. More precisely, we can define

$$r_2(t, H) = -\frac{1}{H} \log B_2(t, H) = -\frac{1}{H} \log \mathbb{E}[e^{-(y_{t+1} + \dots + y_{t+H})} | y_t]$$

Since the reverse time dynamics is affine, the term structure is easily computed and is plotted in Figure 6.

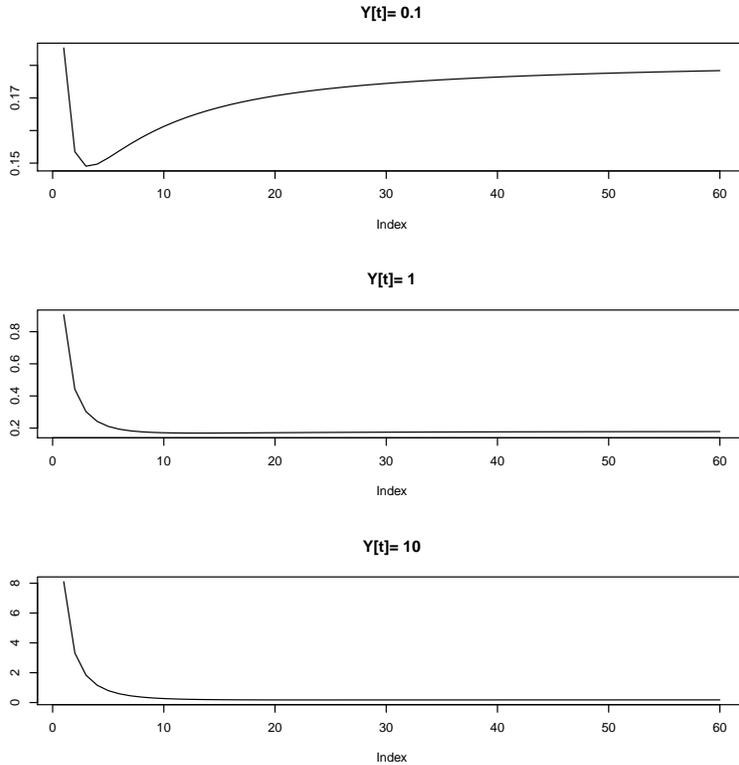


Figure 6: (Reverse time) term structure of interest rates for different values of y_t and a causal AR(1) exponential factor.

4.2 Numerical approximation for noncausal AR gamma mixture process

In general the term structure is derived by numerical approaches. To illustrate the accuracy of numerical approximations we extend the above non-causal exponential process to a process with gamma stationary distribution. More precisely we assume that positive process (y_t) has the noncausal AR(1) representation:

$$y_t = \rho y_{t+1} + \epsilon_t,$$

where the distribution of i.i.d., positive error ϵ_t is characterized by its LT $\mathbb{E}[e^{uy_t}] = \left(\frac{1-\rho u}{1-u}\right)^\alpha$. The causal analogue of this process has been introduced by Bernier (1970) to analyze the daily river flows. Bernier shows that ϵ_t has

a compound Poisson representation and then can be easily simulated. For this process, the joint LT is:

$$\phi(u, v) := \mathbb{E}[e^{vy_t + uy_{t+1}}] = \frac{1}{(1 - \rho v - u)^\alpha} \frac{(1 - \rho v)^\alpha}{(1 - v)^\alpha}.$$

Hence, by Bartlett's formula,

$$\mathbb{E}[e^{uy_{t+1}} | y_t] = \frac{1}{2\pi\ell(y_t)} \int_{-\infty}^{\infty} \frac{1}{(1 - \rho iz - u)^\alpha} \frac{(1 - \rho iz)^\alpha}{(1 - iz)^\alpha} e^{-izy_t} dz, \quad (4.4)$$

where $\ell(y_t)$ is the density of the gamma distribution with shape parameter α and unitary scale parameter. When α is integer valued, we can apply the same residue theorem as in the previous subsection; when α is non-integer, the above integral is approximated numerically, for instance by Riemann sums. To assess the quality of the numerical approximation, we consider the special case where $\alpha = 1$ that corresponds to the mixture exponential model in Section 3.2.3. We report in Table 1 the Riemann approximated values¹² of (4.4) against their exact values given in Section 3.2.3., for different values of y , where we take $u = -1$.

Value of y_t	0.25	0.5	1	2
Exact value of the conditional LT	0.7394	0.5603	0.3526	0.2082
Approximated value	0.7392	0.5602	0.3527	0.2080
Relative error	$-2 * 10^{-4}$	$-3 * 10^{-5}$	$2 * 10^{-4}$	$-3 * 10^{-5}$

Table 1: Exact and approximated values of the conditional LT for $u = -1$ and different values of y_t . The relative error is defined as the approximated value divided by the corresponding exact value, minus 1.

The accuracy of the Bartlett numerical inversion is extremely high and the approach can also be used to compute (4.4), when α is non integer. Similarly, the conditional LT $\mathbb{E}[e^{-y_{t+1} - y_{t+2} - \dots - y_{t+H}} | y_t]$ can also be computed numerically. The details are omitted.

Finally, in the above two examples we have considered term structure models with only one noncausal factor. In case where the dimension of the

¹²We have subdivided the interval $[-10^4, 10^4]$ into intervals of size 0.1 and used the Riemann sum on this subdivision to approximate this one-dimensional integral. The CPU required for conducting this numerical integral is less than 0.05 second in R with a standard PC.

noncausal factor is larger than 1, faster alternative methods can also be employed [such as Fast Fourier Transform, see e.g. Carr and Madan (1999)].

5 Applications to option pricing

In this section we study the pricing of European call options in two models, with or without stochastic volatility. In Section 5.1 we assume that the return follows a noncausal Laplace AR process, whereas in Section 5.2 we consider a stochastic volatility model in which the volatility process follows a noncausal exponential AR process.

5.1 A model with noncausal return process

Let us now assume that the daily geometric return of a stock (y_t) is a non-causal AR(1) Laplace process. Thus, given information up to time t , the stock price of the next period is $P_{t+1} = P_t \exp(y_{t+1})$. For expository purpose we assume, without loss of generality that $P_t = 1$. Let us denote by $C(K)$ the price of a European call option with horizon 1 and strike $K > 0$. It is shown in Heston (1993) that under the assumption of a constant short rate $r_t = r$, we have¹³:

$$C(K) = S_t \Pi_{1t} - K e^{-r} \Pi_{2t}, \quad (4.5)$$

where

$$\Pi_{1t} = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left(\frac{e^{-iu \log(K)} \Phi_t(u-i)}{iu \Phi_t(-i)} \right) du, \quad (4.6)$$

$$\Pi_{2t} = \mathbb{P}[P_{t+1} > K] = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left(\frac{e^{-iu \log(K)} \Phi_t(u)}{iu} \right) du, \quad (4.7)$$

$$S_t = \Phi_t(-i), \quad (4.8)$$

$$\text{and } \Phi_t(u) = \mathbb{E}[e^{iu \log(P_{t+1})} | P_t] = \mathbb{E}[e^{iuy_{t+1}} | y_t]. \quad (4.9)$$

The integrals in equations (4.6) and (4.7) are easily approximated by Riemann summation and Figure 7 and 8 display the price of the option as a

¹³Inversion of the Fourier transform is not the only possible approach to compute option prices. For instance, Heston and Rossi (2017) propose an orthogonal polynomial approach, which they show to be computationally cheaper than Fourier inversion since it only involves the computation of some conditional moments, which do not depend on the choice of strike K . For expository purpose, this approach is not pursued in this paper.

function of the strike K , ranging from 0.9 to 1.1. The parameter ρ of the noncausal Laplace process is set to $\rho = 0.2$ (dashed line) and $\rho = 0$, respectively. The price depends on the volatility parameter σ , the correlation parameter ρ , and the current return y_t . For $\rho = 0$, the model is a special case of the variance gamma [see e.g. Madan et al. (1998)], in which the conditional distribution of the return given the stochastic volatility is log-normal and the stochastic volatility follows an exponential distribution. In this latter case the price of the European call does not depend on y_t .

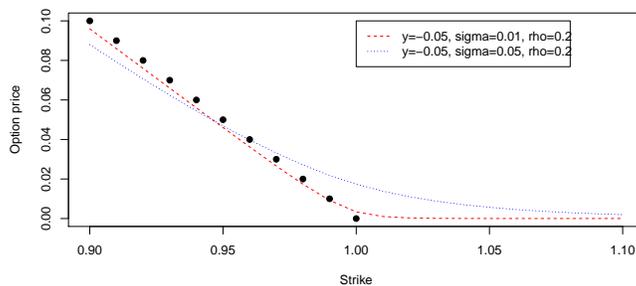


Figure 7: European call price as a function of strike K , for different values of σ and $y_t = -0.05$ in a noncausal Laplace model with $\rho = 0.2$ (dashed line) and $\rho = 0$ (full line).

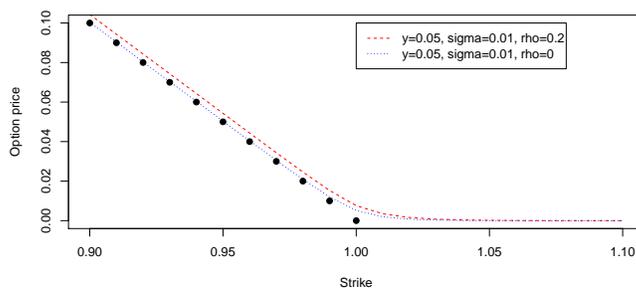


Figure 8: European call price as a function of strike K , for different values of σ and $y_t = 0.05$ in a noncausal Laplace model with $\rho = 0.2$ (dashed line) and $\rho = 0$ (full line).

In both figures we have also plotted K against $S_t - K = 1 - K$ in bullets. We can see that in Figure 7, when the strike is small, the option prices are smaller than $1 - K$. This feature cannot be reproduced by the Black-Scholes model¹⁴ and is due to the serial dependence of the return process, which leads to a downward conditional drift for y_{t+1} .¹⁵ Thus in this case, it is not possible to quote these options using their Black-Scholes implied volatility. On the other hand, in Figure 8, when past return y_t is positive, all the call option prices are above their Black-Scholes limiting value $1 - K$. In Figure 9 we compute the corresponding implied volatilities as a function of K :

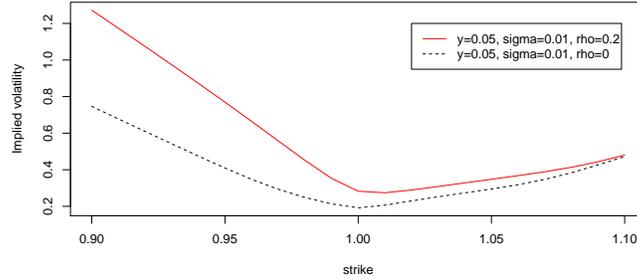


Figure 9: Implied volatility of the call options of Figure 8. We can see that the serial correlation induces asymmetry for the implied volatility, that is the volatility smirk.

The same residual analysis can be conducted to obtain the option prices at higher horizons.

5.2 Noncausal stochastic volatility model

In the analysis of Section 5.1 we have considered the return process to be directly noncausal linear autoregressive. Alternatively, we can consider a noncausal stochastic volatility model, as an alternative to the Heston model.

¹⁴Indeed, in the Black-Scholes model, the call option price is increasing in the volatility parameter. In the limiting case where the volatility is zero, the future return is known at date t and is equal to $S_t = 1$, in this case the call option's price is equal to $1 - K$.

¹⁵We refer the reader to Lo and Wang (1995) for a more detailed analysis of option pricing in the presence of return autocorrelation.

More precisely, let us assume that the asset return x_t satisfies:

$$x_t = \sqrt{y_t}u_t,$$

where (u_t) is a sequence of independent standard normal variables and (y_t) is the positive stochastic variance process. By iterative expectation, the causal characteristic function of X_{t+1} given Y_t is:

$$\mathbb{E}[e^{iu_{t+1}}|y_t] = \mathbb{E}[e^{-\frac{u^2}{2}y_{t+1}}|y_t].$$

Thus the option price is easily computed, whenever the causal LT of process (y_t) is easily computable. Let us now compare the option prices under two models:

- In the first one, process (y_t) follows the noncausal AR(1) exponential process with scale parameter¹⁶ σ^2 . Such a model can be regarded as the noncausal, time-discretized version of the SV model of Barndorff-Nielsen and Shephard (2001).
- In the second model, (y_t) follows the ARG process, which is a time-reversible affine process.¹⁷ This model is the discrete-time analogue of the Heston SV model [see Heston (1993)]. The ARG process has an autocorrelation coefficient βc and gamma stationary distribution with scale parameter $\frac{c}{1-\beta c}$ and shape parameter δ .

To facilitate the comparison, we set $\delta = 1$, $\beta c = \rho = 0.2$, and $c = (1 - \beta c)\sigma^2$, with $\sigma = 0.1$, so that (y_t) has the same stationary distribution and autocorrelation coefficient under the two models. Figure 10 compares the

¹⁶In other words, process (y_t/σ^2) follows the standard noncausal AR(1) exponential process introduced in Section 3.3. By equation (3.1) and a change of variable, we get the causal LT of process (y_t) :

$$\mathbb{E}[e^{uy_{t+1}}|y_t] = \frac{1 - \rho}{1 - \rho - v\sigma^2} + \frac{-v\sigma^2}{1 - \rho - v\sigma^2} e^{-\frac{y_t}{\sigma^2}[\frac{1-v\sigma^2}{\rho}-1]}, \quad \forall u.$$

¹⁷A process (y_t) is ARG if given y_t , variable y_{t+1} conditionally follows gamma distribution with scale parameter c and shape parameter $\delta + Z_t$, where Z_t conditionally follows Poisson distribution with parameter βy_t . The conditional LT of the ARG process is $\mathbb{E}[e^{-uy_{t+1}}|y_t] = \frac{1}{(1+cu)^\delta} e^{-y_t \frac{\beta cu}{1+cu}}$, for any u , thus option prices can also be computed up to some Fourier inversion.

price of the European call options under the two models, for $y_t = 0.5\sigma^2$ and $y = 5\sigma^2$, respectively. The current price of the asset is set to $P_t = 1$ and the maturity equal to 1.

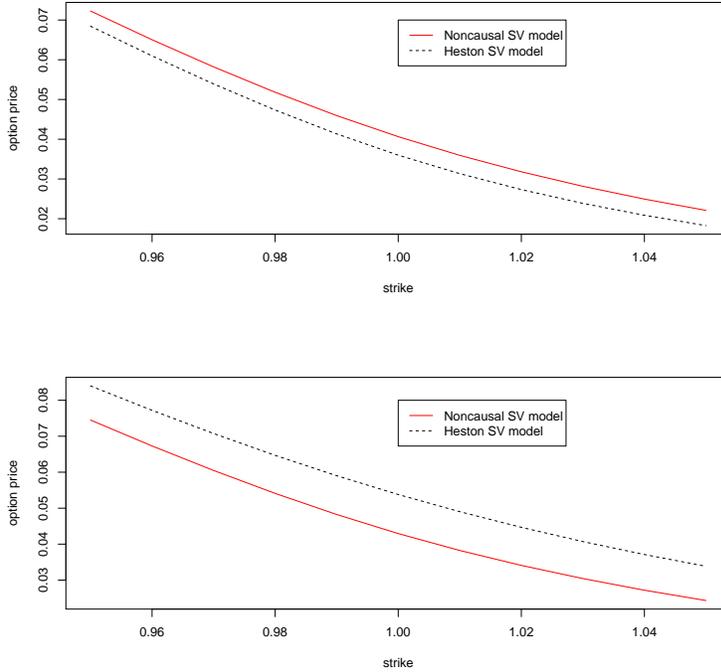


Figure 10: European call price as a function of strike $K \in [0.95, 1.05]$ under the noncausal SV model (full line) and the Heston model (dashed line), when $y_t = 0.5\sigma^2$ (upper panel) and $y = 5\sigma^2$ (lower panel).

When the current volatility $\sqrt{y_t}$ is below its mean value σ (in the upper panel), call options are under the noncausal SV model more expensive than under the Heston SV model when the strike is close to 1, that is when the option is nearly at the money. On the other hand, when the current volatility is high (in the lower panel), call options are cheaper when evaluated using the noncausal SV model. This difference is expected, since in the first case, the noncausal model takes into account a high probability of future volatility bubble, whereas in the second case, since the volatility is already high enough, the noncausal model predicts a much higher probability for the volatility bubble to collapse.

6 Concluding Remarks

We have introduced new parametric families of processes with quasi-closed form expressions for derivative prices. These processes are affine in reverse time, but their causal dynamics is not affine. This causal nonlinearity creates a new type of nonlinear dependence between the derivative prices and the underlying factors. The feature is illustrated by considering noncausal positive AR(1) process for short term interest rate and the associated term structure patterns. We also considered the pricing of European call options based on noncausal AR(1) Laplace process and stochastic volatility models with noncausal AR(1) exponential process, and have compared them with the Variance Gamma, as well as the Heston model, respectively. Such non-causal processes are especially relevant for pricing sequences of speculative bubbles frequently encountered for commodity prices and cryptocurrencies series.

REFERENCES

- Barndorff-Nielsen, O.E. and N., Shephard, (2001) : "Non-Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics", *Journal of the Royal Statistical Society: Series B*, 63(2), 167-241.
- Bartlett, M. (1938) : "The Characteristic Function of a Conditional Statistic", *Journal of the London Mathematical Society*, 13, 63-67.
- Bates, D. (2006) : "Maximum Likelihood Estimation of Latent Affine Processes", *Review of Financial Studies*, 19, 909-965.
- Bernier, J. (1970) : "Inventaire des modeles de processus stochastiques applicables a la description des debits journaliers des rivieres", *International Statistical Review*, 38(1), 49-104.
- Bernstein, S. (1928) : "Sur les fonctions absolument monotones". *Acta Mathematica*, 52: 1-66.
- Cambanis, S., and I., Fakhre-Zakeri (1994) : "On Prediction of Heavy Tailed Autoregressive Sequences : Forward Versus Reversed Time", *Theor. Probab. Appl.*, 39, 217-233.
- Carr, P. and D., Madan, (1999) : "Option Valuation Using the Fast Fourier Transform", *Journal of Computational Finance*, 2(4),61-73.
- Carr, Peter, H., Geman, D., Madan, and M., Yor (2007): "Self-Decomposability and Option Pricing." *Mathematical Finance* 17(1), 31-57.
- Dai, Q., and K., Singleton (2003) : "Term Structure Dynamics in Theory and Reality", *Review of Financial Studies*, 16, 631-378.
- Darolles, S., C., Gouriéroux, and J., Jasiak (2006) : "Structural LT and Compound Autoregressive Models", *Journal of Time Series Analysis*, 27, 477-503.
- Duffie, D., Filipovic, D., and W., Schachermayer (2003) : "Affine Processes and Applications in Finance", *Annals of Applied Probability*, 13, 984-1053.

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

Feller, W. (1968): "An Introduction to Probability Theory and its Applications", Vol. 2. John Wiley & Sons.

Gaver, D., and P., Lewis (1980) : "First-Order Autoregressive Gamma Sequences and Point Processes", *Advances in Applied Probability*, 12, 727-745.

Gourieroux, C., and J., Jasiak (2006) : "Autoregressive gamma processes", *Journal of Forecasting*, 25(2), 129-152.

Gourieroux, C., and J., Jasiak (2016) : "Filtering, Prediction and Simulation Methods for Noncausal Processes", *Journal of Time Series Analysis*, 37, 405-430.

Gourieroux, C., Jasiak, J., and A., Monfort (2017) : "Stationary Bubble Equilibria in Rational Expectation Models", CREST DP.

Gourieroux, C., Monfort, A., and V., Polimenis (2006) : "Affine Models for Credit Risk Analysis", *Journal of Financial Econometrics*, 4, 494-536.

Gourieroux, C., and J.M., Zakoian (2017) : "Local Explosion Modelling by Noncausal Process", *Journal of the Royal Statistical Society: Series B*, 79, 737-756.

Heston, S. (1993) : "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies*, 6(2), 327-343.

Heston, S., and A., Rossi (2017) : "A Spanning Series Approach to Options", *Review of Asset Pricing Studies*, 7(1), 2-42.

Lanne, M., and P., Saikkonen (2013) : "Noncausal Vector Autoregression", *Econometric Theory*, 29, 447-481.

Lee, J., and P., Phillips (2016) : "Asset Pricing with Financial Bubble Risk", *Journal of Empirical Finance*, 38, 590-622.

Lewis, A. (1998) : "Applications of Eigenfunction Expansions in Continuous-Time Finance." *Mathematical Finance* 8(4), 349-383.

Linetsky, V. (2006) : "Pricing Equity Derivatives Subject to Bankruptcy", *Mathematical Finance*, 16(2), 255-282.

Lo, A., and J., Wang (1995) : "Implementing Option Pricing Models when Asset Returns Are Predictable", *Journal of Finance*, 50(1), 87-129.

Madan, D., Carr, P., and E., Chang (1998) : "The Variance Gamma Process and Option Pricing," *Review of Finance*, 2(1), 79-105.

Pelsser, A. (2000) : "Pricing Double Barrier Options Using LTs", *Finance and Stochastics*, 4(1), 95-104.

Singleton, K. (2009) : "Empirical Dynamic Asset Pricing and Econometric Assessment", Princeton Univ. Press.

Taylor, J. W. (2017) : "Forecasting Value at Risk and Expected Shortfall Using a Semiparametric Approach Based on the Asymmetric Laplace Distribution", forthcoming *Journal of Business and Economic Statistics*, 1-13.

APPENDICES

Appendix 1: Derivation of the recursive equations

i) By the iterated expectation theorem we have :

$$\begin{aligned}
& \mathbb{E}\{\exp[uy_t + w(y_{t+1} + \dots + y_{t+H-1})] | y_{t+H}\} \\
= & \mathbb{E}\{\mathbb{E}\{\exp[uy_t + w(y_{t+1} + \dots + y_{t+H-1})] | \underline{y_{t+H-1}}\} | y_{t+H}\} \\
= & \mathbb{E}\{\exp(wy_{t+H-1}) \mathbb{E}\{\exp[uy_t + w(y_{t+1} + \dots + y_{t+H-2})] | y_{t+H-1}\} | y_{t+H}\} \\
= & \mathbb{E}\{\exp(wy_{t+H-1} + a(u, w; H-1)y_{t+H-1} + b(u, w; H-1)) | y_{t+H}\} \\
= & \exp[b(u, w; H-1)] \mathbb{E}\{\exp[w + a(u, w; H-1)]y_{t+H-1} | y_{t+H}\} \\
= & \exp[b(u, w; H-1)] \exp[a[w + a(u, w; H-1)]y_{t+H} + b[w + a(u, w; H-1)]].
\end{aligned}$$

The result follows by comparing with the other expression of this conditional LT. It is valid for $H \geq 3$.

ii) Let us now consider the case $H = 2$. We get :

$$\begin{aligned}
& \mathbb{E}\{\exp(uy_t + wy_{t+1}) | y_{t+2}\} \\
= & \mathbb{E}\{\mathbb{E}\{\exp(uy_t + wy_{t+1}) | y_{t+1}\} | y_{t+2}\} \\
= & \mathbb{E}\{\exp[wy_{t+1} + a(u)y_{t+1} + b(u)] | y_{t+2}\} \\
= & \exp\{a[w + a(u)]y_{t+2} + b(u) + b[w + a(u)]\}.
\end{aligned}$$

Therefore, we have :

$$\begin{aligned}
a(u, w; 2) &= a[w + a(u)], \\
b(u, w; 2) &= b[w + a(u)] + b(u).
\end{aligned}$$

The expressions above correspond to the recursive equations applied for $H \geq 2$, with initial conditions :

$$a(u, w; 1) = a(u), \quad b(u, w; 1) = b(u).$$

Appendix 2: Proof of Proposition 3

Since the distribution of a nonnegative variable is characterized by its real LT, we consider this function defined by :

$$\mathbb{E}[\exp(-uy_{t+1})|y_t] = \exp[-a(u)y_t - b(u)], \quad u \geq 0, \quad (\text{a.1})$$

where we change the notations of functions a, b for expository purpose. Let us also assume that y_{t+1} can take any value in $(0, \infty)$. By considering the limiting cases $y_{t+1} = 0$ and $y_{t+1} = \infty$, we deduce that both $\exp[-a(u)]$ and $\exp[-b(u)]$ are also real LTs of nonnegative variable. The same for the real LT of the stationary distribution of y_t defined by :

$$\mathbb{E}[\exp(-uy_t)] = \exp[-c(u)], \quad \forall u \geq 0. \quad (\text{a.2})$$

We will use several Lemmas.

Lemma 3 (Bernstein Theorem, see Bernstein 1928). *The log real LT of a nonnegative variable is infinitely right differentiable outside zero. Moreover these right derivatives have alternating signs.*

Lemma 3 can be applied to functions a, b, c .

Lemma 4. *If the nonnegative affine process (y_t) is weakly ergodic, then function $a(\cdot)$ is right differentiable at zero.*

Proof. By Bernstein Theorem the derivative $a'(u)$ exists for any $u > 0$ and is increasing when u decreases to zero. Thus we can define $\lim_{u \rightarrow 0^+} a'(u)$. If this limit is finite, function $a(\cdot)$ is differentiable at zero.

Let us now show that this limit cannot be infinite. If the limit were infinite, for any arbitrary large number $M > 0$, it would be possible to find $\varepsilon > 0$, such that :

$$\frac{a(u)}{u} > M, \quad \text{for any } u < \varepsilon. \quad (\text{a.3})$$

This condition contradicts the property that $a^{oh}(u) \rightarrow 0$, for any $u > 0$, when h goes to infinity. Indeed for given ε, u , there would exist h_0 such that, for any $h > h_0$, we would have $a^{oh}(u) \in (0, \varepsilon)$. Then by equation (a.3) we have $\frac{a^{o(h)}(u)}{a^{o(h-1)}(u)} > M$ for any $h > h_0$. By iterating this inequality, we get : $a^{oh}(u) > M^{h-h_0} a^{oh_0}(u)$. Then the sequence $a^{oh}(u)$ cannot stay in $(0, \varepsilon)$ for large h . This contradicts the fact that $\lim_{u \rightarrow 0^+} a'(u) = \infty, \forall u > 0$. \square

Let us now assume that the stationary process satisfying (a.1) has also an affine representation in reverse time, that is,

$$\mathbb{E}[\exp(-vy_t)|y_{t+1}] = \exp[-a^*(v)y_{t+1} - b^*(v)], \forall v \geq 0. \quad (\text{a.4})$$

The functions b and b^* can be written in terms of (the same) function c as :

$$b(u) = c(u) - c[a(u)], \quad (\text{a.5})$$

$$b^*(u) = c(u) - c[a^*(u)]. \quad (\text{a.6})$$

To prove Proposition 3, we have to show that $a^* = a$. The proof is similar to the characterization of reversibility in Darolles et al. (2006). We consider the joint LT of (y_{t+1}, y_t) . On the one hand we have :

$$\begin{aligned} \mathbb{E}[\exp(-uy_{t+1} - vy_t)] &= \mathbb{E}\left\{ \exp(-vy_t) \mathbb{E}[\exp(-uy_{t+1})|y_t] \right\} \\ &= \exp[-b(u)] \mathbb{E}\{\exp[-(a(u) + v)y_t]\} \\ &= \exp\{-b(u) - c[a(u) + v]\}; \end{aligned}$$

on the other hand :

$$\begin{aligned}
\mathbb{E}[\exp(-uy_{t+1} - vy_t)] &= \mathbb{E}\{\exp(-uy_{t+1})\mathbb{E}[\exp(-vy_t)|y_{t+1}]\} \\
&= \exp[-b^*(v)]\mathbb{E}\{\exp(-[a^*(v) + u]y_{t+1})\} \\
&= \exp\{-b^*(v) - c[a^*(v) + u]\}.
\end{aligned}$$

Therefore we get :

$$b(u) + c[a(u) + v] = b^*(v) + c[a^*(v) + u], \quad \forall u, v > 0, \quad (\text{a.7})$$

or equivalently :

$$c[a(u) + v] - c[a(u)] = \{c[a^*(v) + u] - c(u)\} + b^*(v), \quad \forall u, v > 0. \quad (\text{a.8})$$

Lemma 5. *b^* , b and c are right differentiable at zero.*

Proof. Let us consider equation (a.8). Its left hand side (LHS) is right differentiable at $v = 0$, since c is right differentiable outside zero by Bernstein theorem. Moreover, the first term on the RHS can be written as:

$$\frac{c[a^*(v) + u] - c(u)}{v} = \frac{c[a^*(v) + u] - c(u)}{a^*(v)} \frac{a^*(v)}{v}.$$

It tends to $c'(u)a^{*'}(0)$ when v tends to zero, by Lemma 4. By equation (a.8) we deduce that b^* is also right differentiable at zero. A similar argument can be applied to function b .

As for the right differentiability at zero of c , by iterating (a.5), we get :

$$c(u) = b(u) + b[a(u)] + b[a^{o2}(u)] + \dots$$

By the composition differentiation rule, we conclude that function c is also right differentiable at zero, with $c'(0) = \frac{b'(0)}{1 - a'(0)}$. □

Then equality (a.8) can be rewritten as :

$$c[a(u) + v] - c[a(u)] = \{c[a^*(v) + u] - c(u)\} + c(v) - c[a^*(v)]. \quad (\text{a.9})$$

Then differentiating both sides w.r.t. v at $v = 0^+$ leads to :

$$c'[a(u)] - c'(0) = (a^*)'(0)[c'(u) - c'(0)]. \quad (\text{a.10})$$

By rewriting equality (a.9) as :

$$\{c[a(u) + v] - c(v)\} - c[a(u)] = \{c[a^*(v) + u] - c[a^*(v)]\} - c(u),$$

and differentiating both sides w.r.t. u at $u = 0^+$, we get :

$$c'[a^*(v)] - c'(0) = a'(0)[c'(v) - c'(0)]. \quad (\text{a.11})$$

From equations (a.10)-(a.11), we see that $a^* = a$, if and only if $(a^*)'(0) = a'(0)$.

Let us assume for instance that $a'(0) > (a^*)'(0)$. By Bernstein theorem, function c' is decreasing and $c'(u) - c'(0) < 0$. Thus :

$$a'(0)[c'(u) - c'(0)] < (a^*)'(0)[c'(u) - c'(0)], \quad \forall u > 0,$$

or equivalently by (a.10)-(a.11) :

$$c'[a(u)] > c'[a^*(u)], \quad \forall u > 0, \quad \text{or} \quad a(u) < a^*(u), \quad \forall u > 0.$$

This implies $a'(0) = \lim_{u \rightarrow 0^+} \frac{a(u)}{u} \leq \lim_{u \rightarrow 0^+} \frac{a^*(u)}{u} = (a^*)'(0)$. This contradicts the assumption that $a'(0) > (a^*)'(0)$.

In other words we have shown that $a'(0) = (a^*)'(0)$, that is $a^*(u) = a(u), \forall u > 0$. Thus by (a.10)-(a.11), we deduce the reversibility of process (y_t) .

Appendix 3: Noncausal Conditional LT of a Noncausal AR(1) Process.

We deduce by recursion from (3.1) that :

$$\begin{aligned}
y_{t+H-1} &= \rho y_{t+H} + \varepsilon_{t+H-1}, \\
y_{t+H-2} &= \rho^2 y_{t+H} + \rho \varepsilon_{t+H-1} + \varepsilon_{t+H-2}, \\
&\vdots \\
y_{t+1} &= \rho^{H-1} y_{t+H} + \rho^{H-2} \varepsilon_{t+H-1} + \rho^{H-3} \varepsilon_{t+H-2} + \dots + \varepsilon_{t+1}, \\
y_t &= \rho^H y_{t+H} + \rho^{H-1} \varepsilon_{t+H-1} + \rho^{H-2} \varepsilon_{t+H-2} + \dots + \rho \varepsilon_{t+1} + \varepsilon_t.
\end{aligned}$$

Therefore we get :

$$\begin{aligned}
& uy_t + vy_{t+H} + w(y_{t+1} + \dots + y_{t+H}) \\
&= y_{t+H}[u\rho^H + v + w(1 + \dots + \rho^{H-1})] \\
&+ \varepsilon_{t+H-1}[u\rho^{H-1} + w(1 + \dots + \rho^{H-2})] \\
&\vdots \\
&+ \varepsilon_{t+1}[u\rho + w] + \varepsilon_t u \\
&= y_{t+H}[u\rho^H + v + w\frac{1 - \rho^H}{1 - \rho}] \\
&+ \varepsilon_{t+H-1}[u\rho^{H-1} + w\frac{1 - \rho^{H-1}}{1 - \rho}] \\
&\vdots \\
&+ \varepsilon_{t+1}[u\rho + w\frac{1 - \rho}{1 - \rho}] + \varepsilon_t u.
\end{aligned}$$

The expression of the noncausal LT in (3.3)-(3.4) follows by the serial independence between the ε_t 's.

Appendix 4: Noncausal Cauchy AR process

Let us denote $C(\sigma)$ the Cauchy distribution with scale parameter $\sigma, \sigma > 0$. A variable X follows the $C(\sigma)$ distribution if and only if X/σ follows the standard Cauchy. Its characteristic function is : $E \exp(iuX) = \exp(-\sigma|u|)$ and its p.d.f is :

$$f(x; \sigma) = \frac{1}{\sigma\pi} \frac{1}{1 + (x/\sigma)^2}.$$

Let us consider the linear regression model :

$$Y = aX + u, \tag{a.12}$$

where u and X are independent, u following $C(\eta)$ and X following $C(\sigma)$.

Lemma 6. *Under the assumptions of regression model (a.12):*

i) Y follows $C(\eta + |a|\sigma)$;

ii) *The conditional p.d.f. of X given Y is :*

$$f(x|y) = \frac{1}{\pi} \frac{\eta + |a|\sigma}{\eta\sigma} \frac{1 + [y/(\eta + |a|\sigma)]^2}{[1 + (y - ax)^2][1 + (x/\sigma)^2]}.$$

Proof. i) The marginal distribution of Y is directly obtained by computing its characteristic function.

ii) The conditional density is derived by the Bayes formulas :

$$\begin{aligned} f(x|y) &= \frac{f(y|x)f(x)}{f(y)} \\ &= \left[\frac{1}{\pi\eta} \frac{1}{1 + (y - ax)^2/\eta^2} \frac{1}{\pi\sigma} \frac{1}{1 + (x/\sigma)^2} \right] / \left[\frac{1}{\pi(\eta + |a|\sigma)} \frac{1}{1 + [y/(\eta + |a|\sigma)]^2} \right] \\ &= \frac{1}{\pi} \frac{\eta + |a|\sigma}{\eta\sigma} \frac{1 + [y/(\eta + |a|\sigma)]^2}{[1 + (y - ax)^2/\eta^2][1 + (x/\sigma)^2]}. \end{aligned}$$

□

Appendix 5: Noncausal Lévy AR Process

A. 5.1: Marginal and joint distribution

From the noncausal moving average representation of the process :

$$y_t = \varepsilon_t + \rho\varepsilon_{t+1} + \dots + \rho^h\varepsilon_{t+h}, \dots, \quad 0 \leq \rho < 1,$$

we deduce the real LT as :

$$\begin{aligned} \mathbb{E}[\exp(-uy_t)] &= \prod_{h=0}^{\infty} \mathbb{E}[\exp(-u\rho^h\varepsilon_{t+h})] \\ &= \prod_{h=0}^{\infty} \exp\left(-\sqrt{2\sigma\rho^h}u\right) \\ &= \exp\left[-\sqrt{2\sigma}u \sum_{h=0}^{\infty} \rho^{h/2}\right] \\ &= \exp\left[-\sqrt{2\frac{\sigma u}{(1-\sqrt{\rho})^2}}\right]. \end{aligned}$$

This is a Lévy distribution with zero drift and scale $\frac{\sigma}{(1-\sqrt{\rho})^2}$.

As a consequence, the joint LT is:

$$\begin{aligned} \mathbb{E}[e^{-uy_t - vy_{t+1}}] &= \mathbb{E}[e^{-(\rho u + v)y_{t+1}}] e^{-\sqrt{2\sigma}u} \\ &= e^{-\frac{\sqrt{2\sigma(\rho u + v)}}{1-\sqrt{\rho}} - \sqrt{2\sigma}u}. \end{aligned}$$

A. 5.2: Transition density in calendar time

It is derived by the Bayes formula :

$$\ell(y_{t+1}|y_t) = \ell(y_t|y_{t+1}) \frac{\ell(y_{t+1})}{\ell(y_t)},$$

where the transition distribution in reverse time is drift ρy_{t+1} and scale c , and the marginal distribution is Lévy with zero drift and scale $\sigma/[(1-\sqrt{\rho})^2]$.

We get :

$$\ell(y_{t+1}|y_t) = \sqrt{\frac{\sigma}{2\pi}} \exp\left\{-\frac{\sigma}{2(y_t - \rho y_{t+1})} - \frac{\sigma}{2(1-\sqrt{\rho})^2} \left(\frac{1}{y_{t+1}} - \frac{1}{y_t}\right)\right\} \frac{y_t^{3/2} \mathbb{1}_{y_t > \rho y_{t+1}}}{y_{t+1}^{3/2} (y_t - \rho y_{t+1})^{3/2}}.$$

Appendix 6: Noncausal exponential process

The result can be derived either by applying the Bayes formula from the joint density of (y_t, y_{t+1}) , or by the Bayes formula applied to the LTs. We deduce the two approaches below.

A.6.1: Joint density

The joint density is with respect to the sum of two Lebesgue measures, that are a one-dimensional Lebesgue measure λ_{1t} on the line $y_t - \rho y_{t+1} = 0$, and a two dimensional Lebesgue measure λ_{2t} on the set $y_t - \rho y_{t+1} > 0$. This joint density is first derived by taking the product of the conditional density of y_t given y_{t+1} by the marginal density of y_{t+1} . Thus the joint distribution $d\mu_t$ is :

$$d\mu_t = \rho \exp(-y_{t+1})d\lambda_{1t} + (1 - \rho) \exp[-(y_t - \rho y_{t+1}) - y_{t+1}]d\lambda_{2t}.$$

This can be rewritten as :

$$\begin{aligned} d\mu_t &= \rho \exp(-y_t/\rho)d\lambda_{1t} + (1 - \rho) \exp[-(y_t - \rho y_{t+1}) - y_{t+1}]d\lambda_{2t} \\ &= \exp(-y_t) \left\{ \rho \exp[-y_t(\frac{1}{\rho} - 1)]d\lambda_{1t} + (1 - \rho) \exp[-y_{t+1}(1 - \rho)]d\lambda_{2t} \right\}. \end{aligned}$$

Thus the conditional distribution of y_{t+1} given y_t has a point mass with weight $\rho \exp[-y_t(\frac{1 - \rho}{\rho})]$ at $y_{t+1} = \rho y_t$, and a continuous component on $[0, y_t/\rho]$ with density $(1 - \rho) \exp[-(1 - \rho)y_{t+1}]$. The result follows.

A.6.2: Causal LT

Alternatively we can follow the approach in Section 2.2. The joint LT of (y_t, y_{t+1}) is :

$$\begin{aligned}
\psi(u, v) &= \mathbb{E}[\exp(uy_t + vy_{t+1})] \\
&= \frac{1 - \rho u}{1 - u} \mathbb{E}\{\exp[(\rho u + v)y_{t+1}]\} \\
&= \frac{1}{1 - \rho u - v} \frac{1 - \rho u}{1 - u}.
\end{aligned}$$

By the inverse LT formula in Lemma 2, we have :

$$\begin{aligned}
&\mathbb{E}[\exp(vy_{t+1})|y_t] \\
&= \frac{1}{2\pi\ell(y_t)} \int_{-\infty}^{+\infty} \psi(iz, v) \exp(-izy_t) dz \\
&= \frac{1}{2\pi \exp(-y_t)} \int_{-\infty}^{+\infty} \frac{1}{1 - \rho iz - v} \frac{1 - \rho iz}{1 - iz} \exp(-izy_t) dz. \tag{a.13}
\end{aligned}$$

Let us now consider in the complex plane the contour C_a that goes along the real line from $-a$ to a and then counterclockwise along a semi-circle centered at 0 from a to $-a$, where real number a is sufficiently large, i.e. $a > \frac{1-v}{\rho}$.

Since, on the semi-circle, the modulus of e^{-izy_t} is upper bounded by 1, the integral on the semi-circle goes to zero when a goes to infinity and hence the integral on the contours C_a in formula (a.2) is the limit for large a of the integral $\int_{C_a} \frac{1}{1 - \rho iz - v} \frac{1 - \rho iz}{1 - iz} \exp(-izy_t) dz$. By residue theorem¹⁸, this integral is equal to :

$$-2\pi i \sum_{k=1}^2 \text{Res}(a_k),$$

where the negative sign comes from the fact that contour C_a is negatively oriented, $\text{Res}(a_k)$ denote the residuals computed at the pole $z_1 = -i$ and $z_2 = -\frac{1-v}{\rho}i$. Then the residuals are :

$$\text{Res}(-i) = -\frac{1 - \rho}{(1 - \rho - v)i} \exp(-y_t),$$

$$\text{and } \text{Res}\left(-\frac{1-v}{\rho}i\right) = -\frac{-v}{i(1 - \rho - v)} \exp\left[-\frac{1-v}{\rho}y_t\right].$$

¹⁸The residue theorem has been previously applied in the finance literature by, e.g. Lewis (1998), Pelsser (2000) and Linetsky (2006).

Thus we deduce :

$$\mathbb{E}[\exp(vy_{t+1})|y_t] = \frac{1-\rho}{1-\rho-v} + \frac{-v}{1-\rho-v} \exp[-y_t(\frac{1-v}{\rho} - 1)].$$

A.6.3: Causal LT of the cumulated process

The aim of this subsection is to compute the causal conditional LT of the cumulated process: $\mathbb{E}[e^{w(y_{t+1}+y_{t+2}+\dots+y_{t+H})}|y_t]$. We first derive the joint LT:

$$\begin{aligned} \phi(u, w) &:= \mathbb{E}[e^{uy_t+w(y_{t+1}+y_{t+2}+\dots+y_{t+H})}] \\ &= \mathbb{E}\left[\mathbb{E}[e^{uy_t+w(y_{t+1}+y_{t+2}+\dots+y_{t+H})}|y_{t+H}]\right] \\ &= \mathbb{E}\left[e^{(u\rho^H+w\frac{1-\rho^H}{1-\rho})y_{t+H}}\right] \frac{1-\rho u}{1-u} \frac{1-\rho(\rho u+w)}{1-(\rho u+w)} \frac{1-\rho[\rho^2 u+w(1+\rho)]}{1-[\rho^2 u+w(1+\rho)]} \dots \frac{1-\rho[\rho^{H-1}u+w\frac{1-\rho^{H-1}}{1-\rho}]}{1-[\rho^{H-1}u+w\frac{1-\rho^{H-1}}{1-\rho}]} \\ &= \frac{1}{1-(u\rho^H+w\frac{1-\rho^H}{1-\rho})} \frac{1-\rho u}{1-u} \frac{1-\rho(\rho u+w)}{1-(\rho u+w)} \frac{1-\rho[\rho^2 u+w\frac{1-\rho^2}{1-\rho}]}{1-[\rho^2 u+w\frac{1-\rho^2}{1-\rho}]} \dots \frac{1-\rho[\rho^{H-1}u+w\frac{1-\rho^{H-1}}{1-\rho}]}{1-[\rho^{H-1}u+w\frac{1-\rho^{H-1}}{1-\rho}]} \end{aligned} \quad (\text{a.14})$$

where we have used the formulas derived in Appendix 3. By Lemma 2 we get:

$$\mathbb{E}[e^{w(y_{t+1}+y_{t+2}+\dots+y_{t+H})}|y_t] = \frac{1}{2i\pi e^{-y_t}} \int_{-\infty}^{\infty} \phi(iz, w) e^{-izy_t} d(iz). \quad (\text{a.15})$$

For fixed w and y_t , function $z \mapsto \phi(iz, w) e^{-izy_t}$ has $H+1$ simple poles, that are $z_1 = -i$, $z_2 = -\frac{1-w}{\rho}i$, $z_3 = -\frac{1-w\frac{1-\rho^2}{1-\rho}}{\rho^2}i, \dots, z_H = -\frac{1-w\frac{1-\rho^{H-1}}{1-\rho}}{\rho^{H-1}}i$, and $z_{H+1} = \frac{1-w\frac{1-\rho^H}{1-\rho}}{\rho^H}i$. By the same residue analysis as in Appendix A.6.2:

$$\mathbb{E}[e^{w(y_{t+1}+y_{t+2}+\dots+y_{t+H})}|y_t] = \frac{1}{e^{-y_t}} \sum_{k=1}^{H+1} \text{Res}_t(z_k),$$

where the residues $Res_t(z_k)$ are given by:

$$\begin{aligned}
Res_t(z_1) &:= \frac{[1 - \rho][1 - \rho(\rho + w)][1 - \rho(\rho^2 + w\frac{1-\rho^2}{1-\rho})] \cdots [1 - \rho(\rho^{H-1} + w\frac{1-\rho^{H-1}}{1-\rho})]}{(1 - \rho - w)(1 - \rho^2 - w\frac{1-\rho^2}{1-\rho})(1 - \rho^3 - w\frac{1-\rho^3}{1-\rho}) \cdots (1 - \rho^{H-1} - w\frac{1-\rho^{H-1}}{1-\rho})} e^{-y_t}, \\
Res_t(z_2) &:= \frac{w[1 - \rho][1 - \rho(\rho + w)] \cdots [1 - \rho(\rho^{H-2} + w\frac{1-\rho^{H-2}}{1-\rho})]}{(1 - \frac{1-w}{\rho})(1 - \rho - w)(1 - \rho^2 - w\frac{1-\rho^2}{1-\rho}) \cdots (1 - \rho^{H-2} - w\frac{1-\rho^{H-2}}{1-\rho})} e^{-\frac{1-w}{\rho}y_t}, \\
Res_t(z_3) &:= \frac{[1 - \frac{1-w(1+\rho)}{\rho}]w[1 - \rho][1 - \rho(\rho + w)] \cdots [1 - \rho(\rho^{H-3} + w\frac{1-\rho^{H-3}}{1-\rho})]}{(1 - \frac{1-w(1+\rho)}{\rho^2})(1 - \frac{1-w}{\rho})(1 - \rho - w) \cdots (1 - \rho^{H-3} - w\frac{1-\rho^{H-3}}{1-\rho})} e^{-\frac{1-w(1+\rho)}{\rho^2}y_t} \\
&\dots\dots
\end{aligned}$$

When y_t is a short term rate, and the distribution is the risk-neutral distribution, the price of the zero-coupon bond with term H is obtained by taking $w = -1$. The expressions of the residues show that this bond price is a combination of exponential functions of the short term rate.

A.6.4: Conditional moments

We use the identity:

$$\begin{aligned}
\log \mathbb{E}[\exp(vy_{t+1})|y_t] &= \mathbb{E}[y_{t+1}|y_t]v + \frac{1}{2}\mathbb{V}[y_{t+1}|y_t]v^2 + \frac{1}{6}\mathbb{E}[(y_{t+1} - \mathbb{E}[y_{t+1}|y_t])^3|y_t]v^3 \\
&\quad + \frac{1}{24}\left[\mathbb{E}[(y_{t+1} - \mathbb{E}[y_{t+1}|y_t])^4] - 3(\mathbb{V}[y_{t+1}|y_t])^2\right] + \dots,
\end{aligned}$$

where terms of order higher than 4 are omitted. By expanding the cumulant generating function $\log \mathbb{E}[\exp(vy_{t+1})|y_t]$, we get:

$$\begin{aligned}
\log \mathbb{E}[\exp(vy_{t+1})|y_t] &= \log \frac{1}{1 - \frac{v}{1-\rho}} \left[1 - \frac{v}{1-\rho} e^{\frac{v}{\rho}y_t} e^{-y_t \frac{1-\rho}{\rho}} \right] \\
&= -\log\left(1 - \frac{v}{1-\rho}\right) + \log \left[1 - e^{-y_t \frac{1-\rho}{\rho}} \frac{v}{1-\rho} \left(1 + \frac{vy_t}{\rho} + \frac{v^2 y_t^2}{2\rho^2} + \frac{v^3 y_t^3}{6\rho^3} \right) \right] + o(v^4) \\
&= \left[\frac{v}{1-\rho} + \frac{v^2}{2(1-\rho)^2} + \frac{v^3}{3(1-\rho)^3} + \frac{v^4}{4(1-\rho)^4} \right] - e^{-y_t \frac{1-\rho}{\rho}} \frac{v}{1-\rho} \left[1 + \frac{vy_t}{\rho} + \frac{v^2 y_t^2}{2\rho^2} + \frac{v^3 y_t^3}{6\rho^3} \right] \\
&\quad - \frac{e^{-2y_t \frac{1-\rho}{\rho}} v^2}{2(1-\rho)^2} \left[1 + \frac{2vy_t}{\rho} + \frac{v^2 y_t^2}{\rho^2} \right] - \frac{e^{-3y_t \frac{1-\rho}{\rho}} v^3}{3(1-\rho)^3} \left[1 + 3\frac{vy_t}{\rho} \right] - \frac{e^{-4y_t \frac{1-\rho}{\rho}} v^4}{4(1-\rho)^4} + o(v^4).
\end{aligned}$$

By coefficient matching we deduce that:

$$\begin{aligned}
\mathbb{E}[y_{t+1}|y_t] &= \frac{1}{1-\rho} \left[1 - \exp\left[-y_t \frac{1-\rho}{\rho}\right] \right] < y_t/\rho, \\
\frac{1}{2}\mathbb{V}[y_{t+1}|y_t] &= \frac{1}{2(1-\rho)^2} - e^{-y_t \frac{1-\rho}{\rho}} \frac{y_t}{\rho(1-\rho)} - \frac{e^{-2y_t \frac{1-\rho}{\rho}}}{2(1-\rho)^2} \geq 0, \\
\frac{1}{6}\mathbb{E}[(y_{t+1} - \mathbb{E}[y_{t+1}|y_t])^3|y_t] &= \frac{1}{3(1-\rho)^3} - \frac{e^{-y_t \frac{1-\rho}{\rho}} y_t^2}{2(1-\rho)\rho^2} - \frac{e^{-2y_t \frac{1-\rho}{\rho}} y_t}{(1-\rho)^2 \rho} - \frac{e^{-3y_t \frac{1-\rho}{\rho}}}{3(1-\rho)^3}, \\
\frac{1}{24} \left[\mathbb{E}[(y_{t+1} - \mathbb{E}[y_{t+1}|y_t])^4 - 3(\mathbb{V}[y_{t+1}|y_t])^2] \right] &= \frac{1}{4(1-\rho)^4} - \frac{e^{-y_t \frac{1-\rho}{\rho}} y_t^3}{6(1-\rho)\rho^3} - \frac{e^{-2y_t \frac{1-\rho}{\rho}} y_t^2}{2(1-\rho)^2 \rho^2} - \frac{e^{-3y_t \frac{1-\rho}{\rho}} y_t}{(1-\rho)^3 \rho} \\
&\quad - \frac{e^{-4y_t \frac{1-\rho}{\rho}}}{4(1-\rho)^4}
\end{aligned}$$

Appendix 7: Noncausal Laplace process

Let us now compute the causal LT. By Bartlett's formula we have:

$$\begin{aligned}
\mathbb{E}[e^{vy_{t+1}}|y_t] &= \frac{1}{2\pi\ell(y_t)} \int_{-\infty}^{+\infty} \psi(iz, v) \exp(-izy_t) dz \\
&= \frac{\sigma}{\pi \exp(-|y_t|/\sigma)} \int_{-\infty}^{+\infty} \frac{1 + \rho^2 \sigma^2 z^2}{(1 + \sigma^2 z^2)[1 - \sigma^2(v + i\rho z)^2]} \exp(-izy_t) dz,
\end{aligned} \tag{a.16}$$

where $\phi(\cdot, \cdot)$ is the joint LT of (y_t, y_{t+1}) . The closed form expression of integral (a.16) depends on the sign of y_t . If y_t is positive (resp. negative), then we can consider the contours C_a going along the real line from $-a$ to a and then along the semi-circle centered at 0 from a to $-a$, clockwise (resp. counterclockwise). When z is on this semi-circle, the modulus of e^{-izy_t} is bounded by 1. Thus for large a , the integral on the contours C_a converges to the integral on \mathbb{R} . On the other hand, the integrand in integral (a.16) has simple poles, which are: $z_1 = -i/\sigma, z_2 = i/\sigma, z_3 = -\frac{i}{\rho}(1/\sigma - v)$ and $z_4 = -\frac{i}{\rho}(-1/\sigma - v)$. Among these four poles,

- z_1 is within C_a for all large a , if and only if $y_t > 0$;
- z_2 is within C_a for all large a , if and only if $y_t < 0$;

- z_3 is within C_a for all large a , if and only if $y_t > 0$ and $Im(z_3) < 0$, or $Im(z_3) > 0$ and $y_t < 0$, that is, if and only if $y_t Im(z_3) < 0$, where $Im(\cdot)$ denotes the imaginary part of a complex number
- z_4 is within C_a for all large a , if and only if $y_t Im(z_4) < 0$.

$$\begin{aligned} \mathbb{E}[e^{vy_{t+1}} | y_t] &= \frac{(-1)^{\mathbb{1}_{y_t > 0}} 2\pi i \sigma}{\pi \exp(-|y_t|/\sigma)} \left\{ \frac{1 - \rho^2}{2i\sigma[1 - (\sigma v - \rho)^2]} e^{y_t/\sigma} \mathbb{1}_{y_t < 0} + \frac{1 - \rho^2}{-2i\sigma[1 - (\sigma v + \rho)^2]} e^{-y_t/\sigma} \mathbb{1}_{y_t \geq 0} \right. \\ &\quad + \frac{1 - (1 - \sigma v)^2}{[1 - \frac{(1 - \sigma v)^2}{\rho^2}](-2\rho i \sigma)} e^{-\frac{1/\sigma - v}{\rho} y_t} \mathbb{1}_{y_t Im(-\frac{i}{\rho}(1-v)) < 0} \\ &\quad \left. + \frac{1 - (1 + \sigma v)^2}{[1 - \frac{(1 + \sigma v)^2}{\rho^2}](2\rho i \sigma)} e^{-\frac{-1/\sigma - v}{\rho} y_t} \mathbb{1}_{y_t Im(-\frac{i}{\rho}(-1-v)) < 0} \right\}. \end{aligned}$$