# A new construction of CARMA models via iterations of Ornstein-Uhlenbeck processes. 

Argimiro Arratia joint work with Alejandra Cabaña and Enrique M. Cabaña

Universitat Politècnica de Catalunya,
Universitat Autònoma de Barcelona and
Universidad de la República

CRM Workshop Quantitative Finance Barcelona, June 26, 2015

## The Ornstein-Uhlenbeck process

For parameters $\lambda>0$ and $\sigma>0$,

$$
\begin{equation*}
x_{\lambda, \sigma}(t)=\sigma \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} d W(s) \tag{1}
\end{equation*}
$$

where $W$ is standard Wiener process.
(Gaussian, centered, with independent increments and variance
$\left.E(W(t)-W(s))^{2}=|t-s|\right)$
In differential form

$$
\begin{equation*}
d x_{\lambda, \sigma}(t)=-\lambda x_{\lambda, \sigma}(t) d t+\sigma d W(t) \tag{2}
\end{equation*}
$$

G. E. Uhlenbeck and L. S. Ornstein (1930) On the theory of Brownian motion, Phys. Rev. 36.

## OU as continuous time interpolation of AR(1)

Sample the OU process $x_{\lambda, \sigma}$ at equally spaced times $\{i \tau: i=0,1,2, \ldots, n\}, \tau>0$, to get the (discrete time) series

$$
X_{i}=x(i \tau)=\sigma \int_{-\infty}^{i \tau} \mathrm{e}^{-\lambda(i \tau-s)} d W(s)
$$

Proposition: $\left\{X_{i}\right\}$ obeys an $\operatorname{AR}(1)$ model
Show that

$$
X_{i+1}=e^{-\lambda \tau} X_{i}+Z_{i+1}
$$

where $Z_{i+1}=\sigma \int_{i \tau}^{(i+1) \tau} e^{-\lambda((i+1) \tau-s)} d W(s)$, Gaussian innovation.

## $\mathrm{OU}(p)$ : Ornstein-Uhlenbeck processes of order $p$

For complex $\kappa=\lambda+\imath \mu, \lambda>0, \mu \in \mathbf{R}$ define $\mathcal{O} \mathcal{U}_{\kappa}$ operator acting on process $Y(t), t \in \mathbf{R}$, as

$$
\begin{equation*}
\mathcal{O} \mathcal{U}_{\kappa} Y(t)=\int_{-\infty}^{t} \mathrm{e}^{-\kappa(t-s)} d Y(s) \tag{3}
\end{equation*}
$$

For $p \geq 1, \boldsymbol{\kappa}=\left(\kappa_{1}, \ldots, \kappa_{p}\right) \in\left(\mathbb{C}^{+}\right)^{p}$ and $\sigma>0$, we present the
OU process of order $p$

$$
\begin{aligned}
x_{\boldsymbol{\kappa}, \sigma}=\mathcal{O} \mathcal{U}_{\boldsymbol{\kappa}}(\sigma W) & :=\prod_{j=1}^{p} \mathcal{O U}_{\kappa_{j}}(\sigma W) \\
& =\mathcal{O U}_{\kappa_{1}} \mathcal{O U}_{\kappa_{2}} \cdots \mathcal{O U}_{\kappa_{p}}(\sigma W)
\end{aligned}
$$

## Example: $\mathrm{OU}(2)$ as linear combination of $\mathrm{OU}(1)$

Consider $\boldsymbol{\kappa}=\left(\kappa_{1}, \kappa_{2}\right)$, so that

$$
x_{\boldsymbol{\kappa}, \sigma}=\mathcal{O}_{\kappa_{1}} \mathcal{O} \mathcal{U}_{\kappa_{2}}(\sigma W)=\int_{-\infty}^{t} \mathrm{e}^{-\kappa_{1}(t-s)} d \mathcal{O}_{\kappa_{2}}(\sigma W(s))
$$

## Proposition:

When $\kappa_{1} \neq \kappa_{2}$

The above linear expression generalizes to $\mathrm{OU}(p)$.
Everything holds when the underlying noise is Lévy.
From now on $W$ gets substituted by $\Lambda$, a second-order Lévy process.

## Example: $\mathrm{OU}(2)$ as linear combination of $\mathrm{OU}(1)$

Consider $\boldsymbol{\kappa}=\left(\kappa_{1}, \kappa_{2}\right)$, so that

$$
x_{\boldsymbol{\kappa}, \sigma}=\mathcal{O}_{\kappa_{1}} \mathcal{O U}_{\kappa_{2}}(\sigma W)=\int_{-\infty}^{t} \mathrm{e}^{-\kappa_{1}(t-s)} d \mathcal{O U}_{\kappa_{2}}(\sigma W(s))
$$

## Proposition:

When $\kappa_{1} \neq \kappa_{2}$

$$
\mathcal{O U}_{\kappa_{2}} \mathcal{O U}_{\kappa_{1}}=\frac{\kappa_{1}}{\kappa_{1}-\kappa_{2}} \mathcal{O U}_{\kappa_{1}}+\frac{\kappa_{2}}{\kappa_{2}-\kappa_{1}} \mathcal{O U}_{\kappa_{2}}
$$

The above linear expression generalizes to $\mathrm{OU}(p)$.

From now on $W$ gets substituted by $\Lambda$, a second-order Lévy
process.

## Example: $\mathrm{OU}(2)$ as linear combination of $\mathrm{OU}(1)$

Consider $\boldsymbol{\kappa}=\left(\kappa_{1}, \kappa_{2}\right)$, so that

$$
x_{\kappa, \sigma}=\mathcal{O}_{\kappa_{1}} \mathcal{O U}_{\kappa_{2}}(\sigma W)=\int_{-\infty}^{t} \mathrm{e}^{-\kappa_{1}(t-s)} d \mathcal{O}_{\kappa_{2}}(\sigma W(s))
$$

## Proposition:

When $\kappa_{1} \neq \kappa_{2}$

$$
\mathcal{O U}_{\kappa_{2}} \mathcal{O U}_{\kappa_{1}}=\frac{\kappa_{1}}{\kappa_{1}-\kappa_{2}} \mathcal{O U}_{\kappa_{1}}+\frac{\kappa_{2}}{\kappa_{2}-\kappa_{1}} \mathcal{O U}_{\kappa_{2}}
$$

The above linear expression generalizes to $\mathrm{OU}(p)$.
Everything holds when the underlying noise is Lévy.
From now on $W$ gets substituted by $\Lambda$, a second-order Lévy process.

A Lévy process $\Lambda(t)$ is a càdlàg function, with independent and stationary increments, that vanishes in $t=0$.
As a consequence, $\Lambda(t)$ is, for each $t$, a random variable with an infinitely divisible law.
The characteristic function of $\Lambda(t)$ is $\mathbf{E e}^{i u \Lambda(t)}=\left(\mathbf{E e}^{i u \Lambda(1)}\right)^{t}$, and is usually written as $\mathbf{E e}^{i u \Lambda(1)}=\mathrm{e}^{\psi_{\Lambda}(i u)}$.

The function $\psi_{\Lambda}$ is called characteristic exponent and has the form
$\psi_{\Lambda}(i u)=a i u-\frac{\sigma^{2}}{2} u^{2}+\int_{|x|<1}\left(\mathrm{e}^{i u x}-1-i u x\right) d v(x)+\int_{|x| \geq 1}\left(\mathrm{e}^{i u x}-1\right) d v(x)$
where $v(\{0\})=0, \int_{|x|<1} x^{2} d v(x)<\infty, \int_{|x| \geq 1} d v(x)<\infty$.

Wiener process $W$ satisfies these properties, and, moreover, is the unique continuous Lévy process.

The compound Poisson process with rate $\lambda$ and i.i.d. jumps $Y_{j}$ with $\mathbf{E} Y_{j}=0, \operatorname{Var}\left(Y_{j}\right)=\eta<$ $\infty$ is also a Lévy process.



## $\mathrm{OU}(p)$ as a superposition of $\mathrm{OU}(1)$

The Ornstein-Uhlenbeck process with parameters $\boldsymbol{\kappa}=\left(\kappa_{1}, \ldots, \kappa_{p}\right)$ and $\sigma$,

$$
x_{\kappa, \sigma}=\prod_{j=1}^{p} \mathcal{O U}_{\kappa_{j}}(\sigma \Lambda)
$$

can be written as a linear combination of $p$ processes of order 1 :

1) When the components of $\kappa$ are pairwise different:

The coefficients are:

$$
K_{j}(\boldsymbol{\kappa})=\frac{1}{\prod_{\kappa_{l} \neq \kappa_{j}}\left(1-\kappa_{l} / \kappa_{j}\right)}
$$

## $\mathrm{OU}(p)$ as a superposition of $\mathrm{OU}(1)$

The Ornstein-Uhlenbeck process with parameters $\boldsymbol{\kappa}=\left(\kappa_{1}, \ldots, \kappa_{p}\right)$ and $\sigma$,

$$
x_{\kappa, \sigma}=\prod_{j=1}^{p} \mathcal{O U}_{\kappa_{j}}(\sigma \Lambda)
$$

can be written as a linear combination of $p$ processes of order 1 :

1) When the components of $\boldsymbol{\kappa}$ are pairwise different:

$$
x_{\boldsymbol{\kappa}, \boldsymbol{\sigma}}=\sum_{j=1}^{p} K_{j}(\boldsymbol{\kappa}) \xi_{\kappa_{j}}, \quad \xi_{\kappa_{j}}(t)=\int_{-\infty}^{t} \mathrm{e}^{-\kappa_{j}(t-s)} d(\sigma \Lambda(s))
$$

The coefficients are:

$$
K_{j}(\boldsymbol{\kappa})=\frac{1}{\prod_{\kappa_{l} \neq \kappa_{j}}\left(1-\kappa_{l} / \kappa_{j}\right)}
$$

2) When $\boldsymbol{\kappa}$ has components $\kappa_{h}$ repeated $p_{h}$ times $(h=1,2, \ldots, q$, $\sum_{h=1}^{q} p_{h}=p$ ) the linear combination is:

$$
x_{\boldsymbol{\kappa}, \boldsymbol{\sigma}}=\sum_{h=1}^{q} K_{h}(\boldsymbol{\kappa}) \sum_{j=0}^{p_{h}-1}\binom{p_{h}-1}{j} \xi_{\kappa_{h}}^{(j)}
$$

where

$$
\xi_{\kappa_{h}}^{(j)}(t)=\int_{-\infty}^{t} \mathrm{e}^{-\kappa_{h}(t-s)} \frac{\left(-\kappa_{h}(t-s)\right)^{j}}{j!} d(\sigma \Lambda(s))
$$

The autocovariances of $x_{\kappa, \sigma}$ are

$$
\begin{gathered}
\gamma_{\boldsymbol{\kappa}, \sigma}(t)=\sum_{h^{\prime}=1}^{q} \sum_{i^{\prime}=0}^{p_{h^{\prime}}-1} \sum_{h^{\prime \prime}=1}^{q} \sum_{i^{\prime \prime}=0}^{p_{h^{\prime \prime}}-1} K_{h^{\prime}}(\boldsymbol{\kappa}) \bar{K}_{h^{\prime \prime}}(\boldsymbol{\kappa})\binom{p_{h^{\prime}}-1}{i^{\prime}}\binom{p_{h^{\prime \prime}}-1}{i^{\prime \prime}} \gamma_{\kappa_{h^{\prime}}, \kappa_{h^{\prime \prime}}, \sigma}^{\left(i^{\prime}, i^{\prime \prime}\right)}(t) \\
\gamma_{\kappa_{1}, \kappa_{2}, \sigma}^{\left(i_{1}, i_{2}\right)}(t)=\mathbf{E} \xi_{\kappa_{1}}^{\left(i_{1}\right)}(t) \overline{\xi_{\kappa_{2}}^{\left(i_{2}\right)}(0)} \\
=\sigma^{2}\left(-\kappa_{1}\right)^{i_{1}}\left(-\bar{\kappa}_{2}\right)^{i_{2}} \int_{-\infty}^{0} \mathrm{e}^{-\kappa_{1}(t-s)} \frac{(t-s)^{i_{1}}}{i_{1}!} \mathrm{e}^{-\bar{\kappa}_{2}(-s)} \frac{(-s)^{i_{2}}}{i_{2}!} d s
\end{gathered}
$$

and when the components of $\boldsymbol{\kappa}$ are pairwise different, the covariances can be written as

$$
\gamma_{\boldsymbol{\kappa}, \sigma}(t)=\sum_{h^{\prime}=1}^{p} \sum_{h^{\prime \prime}=1}^{p} K_{h^{\prime}}(\boldsymbol{\kappa}) \bar{K}_{h^{\prime \prime}}(\boldsymbol{\kappa}) \gamma_{\kappa_{h^{\prime}}, \kappa_{h^{\prime \prime}}, \sigma}^{(0,0)}(t) .
$$

## A state space representation of the $\mathrm{OU}(p)$ process

The decomposition of the $\mathrm{OU}(p)$ process

$$
x_{\boldsymbol{\kappa}, \sigma}(t)
$$

as a linear combination of simpler processes of order 1, leads to an expression of the process by means of a state space model.

State space modelling provides us with

- a unified approach for computing the likelihood of $x_{\kappa, \sigma}(t)$ through a Kalman filter.
- a tool to show that the covariances of $x_{\kappa, \sigma}(t)$ coincide with those of an ARMA $(p, p-1)$ whose coefficients can be computed from $\boldsymbol{\kappa}$.

In order to ease notation, we consider that the components of $\boldsymbol{\kappa}$ are all different.
The decomposition of $x_{\boldsymbol{\kappa}, \sigma}(t)=\sum_{j=1}^{p} K_{j} \xi_{\kappa_{j}}(t)$ as a linear combination of the $\mathrm{OU}(1)$ processes
$\xi_{\kappa_{j}}(t)=\int_{-\infty}^{t} \mathrm{e}^{-\kappa_{j}(t-s)} d(\sigma \Lambda(s))=\mathrm{e}^{-\kappa_{j}} \xi_{\kappa_{j}}(t-1)+\int_{t-1}^{t} \mathrm{e}^{-\kappa_{j}(t-s)} d(\sigma \Lambda(s))$
with innovations $\boldsymbol{\eta}_{\boldsymbol{\kappa}}$ with components

$$
\eta_{\kappa_{j}}(t)=\int_{t-1}^{t} \mathrm{e}^{-\kappa_{j}(t-s)} d \Lambda(s)
$$

provides a representation of the $\mathrm{OU}(p)$ process in the space of states $\xi_{\kappa}=\left(\xi_{\kappa_{1}}, \ldots, \xi_{\kappa_{p}}\right)^{\mathrm{tr}}$.

The transitions in the state space are

$$
\boldsymbol{\xi}_{\boldsymbol{\kappa}}(t)=\operatorname{diag}\left(\mathrm{e}^{-\kappa_{1}}, \ldots, \mathrm{e}^{-\kappa_{p}}\right) \boldsymbol{\xi}_{\boldsymbol{\kappa}}(t-1)+\boldsymbol{\eta}_{\boldsymbol{\kappa}}(t),
$$

and

$$
\boldsymbol{x}(t)=\boldsymbol{K}^{\operatorname{tr}}(\boldsymbol{\kappa}) \boldsymbol{\xi}(t)
$$

The assumption $\mathbf{E} \Lambda(1)^{2}=1$ implies that the innovations have variance

$$
\begin{gathered}
\operatorname{Var}\left(\boldsymbol{\eta}_{\boldsymbol{\kappa}, \boldsymbol{\tau}}(t)\right)=\left(\left(v_{j, l}\right)\right) \\
v_{j, l}=\mathbf{E} \int_{t-1}^{t} \mathrm{e}^{-\left(\kappa_{j}+\bar{\kappa}_{l}\right)(t-s)} d s=\frac{1-\mathrm{e}^{-\left(\kappa_{j}+\bar{\kappa}_{l}\right)}}{\kappa_{j}+\bar{\kappa}_{l}}
\end{gathered} .
$$

A state space representation and its implications on the covariances of the OU process in the general case are slightly more complicated.

$$
\begin{align*}
\boldsymbol{\xi}(t) & =A \boldsymbol{\xi}(t-1)+\boldsymbol{\eta}(t)  \tag{4}\\
x(t) & =\boldsymbol{K}^{\operatorname{tr}} \boldsymbol{\xi}(t) \tag{5}
\end{align*}
$$

When $\kappa_{1}, \ldots, \kappa_{q}$ are all different, $p_{1}, \ldots, p_{q}$ are positive integers, $\sum_{h=1}^{q} p_{h}=p$ and $\boldsymbol{\kappa}$ is a $p$-vector with $p_{h}$ repeated components equal to $\kappa_{h}$, the $\mathrm{OU}(p)$ process $x_{\kappa}$ is a linear function of the state space vector

$$
\left(\xi_{\kappa_{1}}^{(0)}, \xi_{\kappa_{1}}^{(1)}, \ldots, \xi_{\kappa_{1}}^{\left(p_{1}-1\right)}, \ldots, \xi_{\kappa_{q}}^{(0)}, \xi_{\kappa_{q}}^{(1)}, \ldots, \xi_{\kappa_{q}}^{\left(p_{q}-1\right)}\right)
$$

and the transition equation is no longer expressed by a diagonal matrix.

## $\operatorname{OU}(p)$ as a CARMA $(p, p-1)$

Eq. (4) shows that $\boldsymbol{\xi}(t)$ is a $p$-dimensional VARMA $(1,0)$ process Eq. (5) expresses $x(t)$ as a linear transformation of $\boldsymbol{\xi}(t)$ by the $(1 \times p)$ matrix $F=\boldsymbol{K}^{\mathrm{tr}}$. Use Theorem below (with $d=p$, $\tilde{p}=1, \tilde{q}=0, m=1$ )

## Theorem (Lütkepohl (2005) Cor. 11.1.2)

Let $y_{t}$ be a d-dimensional, stable, invertible $\operatorname{VARMA}(\tilde{p}, \tilde{q})$ process and let $F$ be an $(m \times d)$ matrix of rank $m$. Then the process $z_{t}=F y_{t}$ has a VARMA( $(\check{p}, \check{q})$ representation with $\check{p} \leq(d-m+1) \tilde{p}$ and $\check{q} \leq(d-m) \tilde{p}+\tilde{q}$.

## $\operatorname{OU}(p)$ as a CARMA $(p, p-1)$

We conclude that $(x(t): t=0,1, \ldots, n)$ is an $\operatorname{ARMA}(\check{p}, \check{q})$ process with $\check{p} \leq p$ and $\check{q} \leq p-1$ :

$$
x(i)=\sum_{j=1}^{p} \phi_{j} x(i-j)+\sum_{l=0}^{p-1} \theta_{l} \epsilon_{i-l}
$$

where $\epsilon$ is a white noise with variance one and
the parameters $\boldsymbol{\phi}=\left(\phi_{1}, \ldots, \phi_{p}\right)^{\operatorname{tr}}, \boldsymbol{\theta}=\left(\theta_{0}, \ldots, \theta_{p-1}\right)^{\operatorname{tr}}$ of the ARMA $(p, p-1)$ process are functions of the parameters $\boldsymbol{\kappa}$ and $\sigma$ of the $\mathrm{OU}(p)$.

Hence,
Given $\boldsymbol{\kappa}$ we know exactly to which $\operatorname{ARMA}(p, p-1)$ corresponds the $\mathrm{OU}_{\boldsymbol{\kappa}}$

## and $\boldsymbol{\theta}$ as functions of $\boldsymbol{\kappa}$

Apply the AR operator $\prod_{j=1}^{p}\left(1-\mathrm{e}^{-\kappa_{j}} B\right)$ to $x_{\boldsymbol{\kappa}}$ and obtain

$$
\prod_{j=1}^{p}\left(1-\mathrm{e}^{-\kappa_{j}} B\right) x_{\kappa}(t)=\sum_{j=1}^{p} K_{j} G_{j}(B) \eta_{\kappa_{j}}(t)=: \zeta(t)
$$

with $G_{j}(z)=\prod_{l \neq j}\left(1-\mathrm{e}^{-\kappa_{l}} z\right):=1-\sum_{l=1}^{p-1} g_{j, l} z^{l}$.
This process has the same second-order moments as the $\operatorname{ARMA}(p, p-1)^{1}$
$\prod_{j=1}^{p}\left(1-\mathrm{e}^{-\kappa_{j}} B\right) x_{\kappa}(t)=\sum_{j=0}^{p-1} \theta_{j} \epsilon(t-j)=: \zeta^{\prime}(t) \quad(\epsilon$ is a white noise $)$
when the covariances $c_{j}=\mathbf{E} \zeta(t) \bar{\zeta}(t-j)$ and $c_{j}^{\prime}=\mathbf{E} \zeta^{\prime}(t) \bar{\zeta}^{\prime}(t-j)$ coincide.

[^0]The covariances $c_{j}^{\prime}$ and $c_{j}$ are given respectively by the generating functions

$$
\sum_{l=-p+1}^{p-1} c_{l}^{\prime} z^{l}=\left(\sum_{h=0}^{p-1} \theta_{h} z^{h}\right)\left(\sum_{k=0}^{p-1} \bar{\theta}_{k} z^{-h}\right)
$$

and

$$
\sum_{l=-p+1}^{p-1} c_{l} z^{l}=\sum_{j=1}^{p} \sum_{l=1}^{p} K_{j} \bar{K}_{l} G_{j}(z) \bar{G}_{l}(1 / z) v_{j, l}=: J(z)
$$

Since $J(z)$ can be computed once $\boldsymbol{\kappa}$ is known, the coefficients $\boldsymbol{\theta}=$ $\left(\theta_{0}, \theta_{1}, \ldots \theta_{p-1}\right)$ are obtained by identifying the coefficients of the polynomials $z^{p-1} J(z)$ and $z^{p-1}\left(\sum_{h=0}^{p-1} \theta_{h} z^{h}\right)\left(\sum_{k=0}^{p-1} \bar{\theta}_{k} z^{-h}\right)$.


The link between discrete ARMA processes and stationary processes with continuous time has been of interest for many years and has been studied, among others, by

Doob, J.L. (1944) The elementary Gaussian Processes, Ann. Math. Statist. 25
Durbin, J. (1961) Efficient fitting of linear models for continuous stationary time series from discrete data
Bergstrom (1984) Handbook of Econometrics, (1990) Continuous Time Econometric Modelling, Oxford U Press
and there is a recent upsurge of interest in continuous-time models, because they can be used in presence of irregularly spaced data, and in non Gaussian processes mainly due to the fact that jumps play an important role in realistic modelling in finance and other fields of application.

One approach is via the stochastic volatility model ${ }^{2}$, in which the volatility process $V$ and the log asset price $G$ satisfy (apart from a deterministic rescaling of time)

$$
\begin{aligned}
d V(t) & =-\lambda V(t)+d \Lambda(t) \\
d G(t) & =(\gamma+\beta V(t)) d t+\sqrt{V(t)} d W(t)+\rho d \Lambda(t)
\end{aligned}
$$

where $\lambda>0, \Lambda$ is a non-decreasing Lévy process and $W$ is a standard Brownian motion independent of $\Lambda$.
The volatility $V$ is a Lévy-driven Ornstein-Uhlenbeck process, or a continuous- time autoregression of order 1: CAR(1).
The autocorrelations of $V$ decay exponentially, hence they constitute a very restrictive family.

[^1]In order to include a wider family of covariances, econometric or physical models apply frequently linear combinations (superpositions) of OU processes driven by either uncorrelated or correlated noise

$$
\sum_{j=1}^{p} a_{j} \int_{-\infty}^{t} \mathrm{e}^{-\kappa_{j}(t-s)} d \Lambda_{j}(s)
$$

Eliazar, I, and Klafter, J. (2009), From Ornstein-Uhlenbeck dynamics to long-memory processes and fractional Brownian motion, Physical Review E, 79
Barndorff- Nielsen, O. and Shephard, N. (2001), op.cit. JRSS, 63.
or models that replace the finite linear combination by a continuous version

$$
\int_{s=-\infty}^{t} \int_{\Re(\kappa)>0} \mathrm{e}^{-\kappa(t-s)} d \Lambda(s, \kappa)
$$

Bergstrom, A. R. (1984), Continuous time stochastic models and issues of aggregation over time, in Handbook of Econometrics, Volume ll, Edited by Z. Griliches and M. D. Intriligator,

Brockwell, P.J. (2004), Representations of continuous time ARMA models, Jr. Appl. Probab, 41.
Brockwell, P.J. (2009), Levy-driven continuous time ARMA processes, in Handbook of financial Time Series.

Chambers, M. J. and Thornton, M.A. (2012), Discrete time representation of continuous time ARMA processes

Brockwell proposes to define CARMA processes via a state space representation of the formal equation

$$
a(D) Y(t)=\sigma b(D) D \Lambda(t)
$$

where $\sigma>0$ is a scale parameter, $D$ denotes differentiation w.r.t $t$, $\Lambda$ is a second-order Lèvy process, $a(z)=z^{p}+a_{1} z^{p-1}+\ldots+a_{p}$ is a polynomial of order $p$ and $b(z)=b_{0}+b_{1}+\ldots+b_{q} z^{q}$ a polynomial of order $q \leq p-1$.

He proves that, under reasonable conditions on the eigenvalues of the matrix of the space equations, the process $Y(t)$ can be written as

$$
Y(t)=\frac{\sigma}{2 \pi} \iint \exp (i(t-u) \lambda) \frac{b(i \lambda)}{a(i \lambda)} d \lambda d \Lambda(u)
$$

This representation allows the computation of moments, and to identify the resulting Gaussian CARMA with the class of stationary Gaussian processes with rational spectral density.

When the zeroes of the AR polynomial are all different, he obtains a representation of the CARMA as a sum of Lèvy-driven OrnsteinUhlenbeck processes.

He proposes to estimate the CARMA parameters by adjusting an $\operatorname{ARMA}(p, q), q<p$ to regularly spaced data and then obtain the parameters of the CARMA whose values at the observation times have the same distribution of the fitted ARMA.

- Is this always possible?

Even in the Gaussian case, he shows that not all $\operatorname{ARMA}(p, q)$ with $q<p$ are embeddable.

We have proposed a different construction of continuous versions of ARMA processes that solves at least in part, these inconveniences.

Our heuristics is completely different, though some of the features of the resulting process are the same.

Though $\gamma(t)$ depends continuously on $\boldsymbol{\kappa}$, the same does not happen with each term in the expression for the covariance, because of the lack of boundedness of the coefficients of the linear combination when two different values of the components of $\boldsymbol{\kappa}$ approach each other. Recall that the autocovariances of $x_{\boldsymbol{\kappa}, \sigma}$ are

$$
\gamma_{\boldsymbol{\kappa}, \sigma}(t)=\sum_{h^{\prime}=1}^{q} \sum_{i^{\prime}=0}^{p_{h^{\prime}}-1} \sum_{h^{\prime \prime}=1}^{q} \sum_{i^{\prime \prime}=0}^{p_{h^{\prime \prime}}-1} K_{h^{\prime}}(\boldsymbol{\kappa}) \bar{K}_{h^{\prime \prime}}(\boldsymbol{\kappa})\binom{p_{h^{\prime}}-1}{i^{\prime}}\left(\begin{array}{c}
p_{h^{\prime \prime}-1}^{\prime \prime}-1
\end{array}\right) \gamma_{i_{h^{\prime}}, \kappa_{h^{\prime \prime}}, \sigma}^{\left(i^{\prime}, i^{\prime \prime}\right)}(t)
$$

with $K_{j}(\boldsymbol{\kappa})=\frac{1}{\prod_{\kappa_{l} \neq \kappa_{j}}\left(1-\kappa_{l} / \kappa_{j}\right)}$
Since we wish to consider real processes $x$ and the process itself and its covariance $\gamma(t)$ depend only of the unordered set of the components of $\boldsymbol{\kappa}$, we shall reparameterize the process.

With the notation $K_{j, i}=\frac{1}{\left(-\kappa_{j}\right)^{i} \prod_{l \neq j}\left(1-\kappa_{l} / \kappa_{j}\right)}$
(in particular, $K_{j, 0}$ is the same as $K_{j}$ )

- The processes $x_{i}(t)=\sum_{j=1}^{p} K_{j, i} \xi_{j}(t)$ and the coefficients $\phi=\left(\phi_{1}, \ldots, \phi_{p}\right)$ of the polynomial (in $\left.z\right)$

$$
\prod_{j=1}^{p}\left(1+\kappa_{j} z\right)=1-\sum_{j=1}^{p} \phi_{j} z^{j}
$$

satisfy

$$
\sum_{i=1}^{p} \phi_{i} x_{i}(t)=x_{\boldsymbol{\kappa}, \sigma}(t)
$$

Therefore, the new parameter $\phi=\left(\phi_{1}, \ldots, \phi_{p}\right) \in \mathbf{R}^{p}$ shall be adopted.

## ML estimation of the parameters of $\mathrm{OU}(p)$ in the Gaussian case

From the observations $\{\mu+x(i): i=0,1, \ldots, n\}$, obtain the likelihood $L$ of the vector $x=(x(1)), \ldots, x(n))$ :

$$
\log L(\boldsymbol{x} ; \boldsymbol{\phi}, \sigma)=-\frac{n}{2} \log (2 \pi)-\frac{1}{2} \log \left(\operatorname{det}(V(\boldsymbol{\phi}, \sigma))-\frac{1}{2} \boldsymbol{x}^{\operatorname{tr}}(V(\boldsymbol{\phi}, \sigma))^{-1} \boldsymbol{x}\right.
$$

with $V(\phi, \sigma)$ equal to the $n \times n$ matrix with components

$$
V_{h, i}=\gamma(|h-i|)
$$

that reduce to $\gamma(0)$ at the diagonal, $\gamma(1)$ at the $1^{\text {st }}$ sub and super diagonals, ...
Obtain via numerical optimisation the MLE $\hat{\boldsymbol{\phi}}$ of $\phi$ and $\hat{\sigma}^{2}$ of $\sigma^{2}$.
The estimations $\hat{\kappa}$ follow by solving $\prod_{j=1}^{p}\left(1+\hat{\kappa}_{j} z\right)=1-\sum_{j=1}^{p} \hat{\phi}_{j} z^{j}$.

## Matching correlations estimation

From the closed formula for the covariance $\gamma$ and the relationship between $\boldsymbol{\kappa}$ and $\boldsymbol{\phi}$, we have a mapping $\left(\boldsymbol{\phi}, \sigma^{2}\right) \mapsto \gamma(t)$, for each $t$. Since $\boldsymbol{\rho}^{(T)}:=(\rho(1), \ldots, \rho(T))^{\operatorname{tr}}=(\gamma(1), \ldots, \gamma(T))^{\operatorname{tr}} / \gamma(0)$ does not depend on $\sigma^{2}$, these equations determine a map

$$
\mathcal{C}:(\phi, T) \mapsto \boldsymbol{\rho}^{(T)}=\mathcal{C}(\phi, T)
$$

for each $T$.
After choosing a value of $T$ and obtaining an estimate $\boldsymbol{\rho}_{e}^{(T)}$ of $\boldsymbol{\rho}^{(T)}$ based on $x$, we propose as a first estimate of $\boldsymbol{\phi}$, the vector $\check{\boldsymbol{\phi}}_{T}$ such that all the components of the corresponding $\boldsymbol{\kappa}$ have positive real parts, and such that the euclidean norm $\left\|\boldsymbol{\rho}_{e}^{(T)}-\mathcal{C}\left(\check{\boldsymbol{\phi}}_{T}, T\right)\right\|$ reaches its minimum, that is, a procedure that resembles the method of moments.

## The Gaussian case: examples

When $\Lambda$ is a Wiener process, the OU process of order $p$ belongs to a subclass with $p+1$ parameters of the classical family of the $2 p$-parameters Gaussian $\operatorname{ARMA}(p, p-1)$

$$
x_{t}=\phi_{1} x_{t-1}+\cdots+\phi_{p} x_{t-p}+\theta_{0} \epsilon_{t}+\theta_{1} \epsilon_{t-1}+\cdots+\theta_{p-1} \epsilon_{t-p+1}
$$

where $\phi_{1}, \ldots, \phi_{p}$ and $\theta_{0}, \ldots, \theta_{q}$ are parameters and $\epsilon_{t}$ is a Gaussian noise with variance 1 .

The parameters $\boldsymbol{\kappa}, \sigma$ determine the Gaussian likelihood of $\mathcal{O} \mathcal{U}_{\kappa, \sigma} w$, and are estimated by the values $\hat{\boldsymbol{\kappa}}$ and $\hat{\sigma}$ that maximize that likelihood.

The matching correlations estimators can be used as the starting point of an optimization procedure leading to compute the ML estimators.

We have simulated the sample paths for the Wiener-driven $\mathrm{OU}(p)$ for different values of the parameters.


Covariances - $\mathrm{p}=3$


Series A (Box, Jenkins \& Reinsel) consists of 197 lectures of concentration in a certain chemical process, taken every 2 hours.

The original series


The parameters $\phi$ and $\theta$ can be obtained from $\kappa$ and $\sigma$ using the expressions for the covariances of both processes.
For Series A, the parameters of the $\mathrm{OU}(3)$ fitted by maximum likelihood are
$\hat{\boldsymbol{\kappa}}=(0.8293,0.0018+0.0330 i, 0.0018-0.0330 i) \quad$ and $\hat{\sigma}=0.4401$
The corresponding $\operatorname{ARMA}(3,2)$ is
$\left(1-2.4316 B+1.8670 B^{2}-0.4348 B^{3}\right) x=0.4401\left(1-1.9675 B+0.9685 B^{2}\right) \epsilon$

On the other hand, the $\operatorname{ARMA}(3,2)$ fitted by maximum likelihood is
$\left(1-0.7945 B-0.3145 B^{2}+0.1553 B^{3}\right) x=0.3101\left(1-0.4269 B-0.2959 B^{2}\right) \epsilon$.


ARMA (AIC 110.46) and OU(3) (AIC 109.9) fitted by maximum likelihood

## Estimating the shape of the Lévy noise

The estimation of the parameters of $\psi_{\Lambda}$, two real numbers and a measure (the so called Lévy-Khinchin triplet), is difficult and requires a large amount of information.

Jongbloed, van der Meulen and van der Vaart (Bernoulli 11(5), 2005, 759-791)
Valdivieso, Schoutens, Tuerlinks (Stat. Infer. Stoch. Process, 2009, 12: 1-18) have proposed nonparametric estimation for the Lévy noise driving an OrnsteinUhlenbeck process.

A simpler setting is to assume that the admissible exponents belong to a parametric class $\Psi=\left\{\psi_{\theta}: \theta \in \Theta\right\}$ where $\Theta \subset \mathbf{R}^{d}$, and obtain the value of $\theta$ for which a chosen quadratic distance between the exponential of $\psi_{\theta}(i u)$ and the empirical characteristic function of the residuals is minimum.

Let us denote $\psi_{\Lambda}(i u)$ the characteristic exponent of the Lévy process $\Lambda$,

$$
\psi_{\Lambda}(i u)=\log \mathbf{E} \mathrm{e}^{i u \Lambda(1)}
$$

The innovation in each component $\xi_{j}$ is $\eta_{j}(t)=\int_{t-1}^{t} \mathrm{e}^{-\kappa_{j}(t-s)} d \Lambda(s)$, so that the innovation of $x_{\kappa, \sigma}$ is

$$
\eta(t)=\int_{t-1}^{t} g(t-s) d \Lambda(s) \quad \text { where } \quad g(t)=\sum_{j=1}^{p} K_{j} \mathrm{e}^{-\kappa_{j} t}
$$

Hence,

$$
\eta \sim \int_{0}^{1} g(1-s) d \Lambda(s) \sim \int_{0}^{1} g(s) d \Lambda(s)
$$

and its characteristic exponent is therefore

$$
\psi_{\eta}=\log \mathbf{E} \mathrm{e}^{i u \eta}=\log \mathbf{E} \mathrm{e}^{i u \int_{0}^{1} g(s) d \Lambda(s)}=\int_{0}^{1} \psi_{\Lambda}(i u g(s)) d s
$$

A simple example: estimation of a noise sum of a Poisson process plus a Gaussian term
Let us assume that the noise is given by

$$
\Lambda(t)=\sigma W(t)+c(N(t)-\lambda t)
$$

where $W$ is a standard Wiener process and $N$ is a Poisson process with intensity $\lambda$. The family of possible noises depends on the three parameters $(\sigma, \lambda, c)$. In this case, the characteristic exponent has a simple form:

$$
\psi_{\Lambda(1)}(i u)=-\frac{\sigma^{2} u^{2}}{2}+\lambda\left(\mathrm{e}^{i u c}-i u c-1\right)
$$

hence

$$
\psi_{\boldsymbol{\eta}}(i u)=\int_{0}^{1}\left(-\frac{\sigma^{2} u^{2} g^{2}(s)}{2}+\lambda\left(\mathrm{e}^{i u g(s) c}-i u g(s) c-1\right)\right) d s
$$

With $g_{h}=\int_{0}^{1} g^{h}(s) d s$,

$$
\psi_{\boldsymbol{\eta}}(i u)=-\frac{\sigma^{2} u^{2} g_{2}}{2}+\lambda\left(-\frac{u^{2} g_{2} c^{2}}{2}-i \frac{u^{3} g_{3} c^{3}}{6}+\frac{u^{4} g_{4} c^{4}}{24}+\ldots\right)
$$

Then we propose to estimate the parameters by equating the coefficients of $u^{2}, u^{3}, u^{4}$ in $\psi_{\boldsymbol{\eta}}(i u)$ with the corresponding ones in the logarithm of the empirical characteristic function of the residuals. Assuming that the mean of the residuals $r_{1}, r_{2}, \ldots, r_{n}$ is zero, their empirical characteristic function is

$$
\frac{1}{n} \sum_{h=1}^{n} \mathrm{e}^{i u r_{h}}=1-\frac{1}{2} u^{2} R_{2}-\frac{1}{6} i u^{3} R_{3}+\frac{1}{24} u^{4} R_{4}+\ldots
$$

where $R_{m}=\frac{1}{n} \sum_{h=1}^{n} r_{h}^{m}$.

Then the logarithm has the expansion $\log \frac{1}{n} \sum_{h=1}^{n} \mathrm{e}^{i u r_{h}}=-\frac{1}{2} u^{2} R_{2}-\frac{1}{6} i u^{3} R_{3}+\frac{1}{24} u^{4} R_{4}-\frac{1}{8} u^{4} R_{2}^{2}+\ldots$
Consequently, the estimation equations are

$$
\begin{aligned}
\left(\sigma^{2}+\lambda c^{2}\right) g_{2} & =R_{2} \\
\lambda c^{3} g_{3} & =R_{3} \\
\lambda c^{4} g_{4} & =R_{4}-3 R_{2}^{2}
\end{aligned}
$$

from which the estimators follow:

$$
\begin{gathered}
\tilde{c}=\frac{R_{4}-3 R_{2}^{2}}{R_{3}} \frac{g_{3}}{g_{4}} \quad \tilde{\lambda}=\frac{R_{3}^{4}}{\left(R_{4}-3 R_{2}^{2}\right)^{3}} \frac{g_{4}^{3}}{g_{3}^{4}} \\
\tilde{\sigma}^{2}=\frac{R^{2}}{g_{2}}-\frac{R_{3}^{2}}{\left(R_{4}-3 R_{2}^{2}\right)} \frac{g_{4}}{g_{3}^{2}} .
\end{gathered}
$$

Next figures show the empirical c.d.f. of 90 estimators of the parameters obtained from simulated series of 200 terms. The residuals were obtained by applying a Kalman filter to the space state formulation, starting from the actual vaue of $\boldsymbol{\kappa}$ used at the simulation (red), that in practical situations is unknown, and from the estimators obtained by matching correlations (green) and by maximum likelihood (blue).




Estimation of the parameters of the noise from 90 replications of
$\{x \boldsymbol{\kappa}(t): t=0,1, \ldots, 200\}, \boldsymbol{\kappa}=(0.01 \pm 0.1 i, 0.2)$, driven by $\Lambda(t)=0.1 W(t)+N_{0.3}(t)-0.3 t .{ }^{3}$
The estimators are not sharp at all, but the ones obtained by the same procedure applied directly on the unfiltered noise $\Lambda$ (dashed lines) are equally rough. Larger series (of size 10000 and 1000000) produce sharper estimates, also shown in the figures by dotted lines.

[^2]

Estimation of the parameters of the noise from 90 replications of $\left\{x_{\boldsymbol{\kappa}}(t): t=0,1,2, \ldots, 200\right\}, \boldsymbol{\kappa}=(0.0018+0.033 i, 0.0018-0.033 i, 0.083)$, driven by $\Lambda(t)=W(t)+N_{1}(t)-t .{ }^{4}$
$\kappa$ used at the simulation (red), estimated by matching correlations (green) and by maximum likelihood (blue)

[^3]
## Conclusions

- We have proposed a family of continuous time stationary processes, based on $p$ iterations of the linear operator that maps a Wiener process onto an Ornstein-Uhlenbeck process, or more generally, wrt Levy processes.
- These operators have some nice properties, such as being commutative, and their $p$-compositions decompose as a linear combination of simple operators of the same kind.
- An $\mathrm{OU}(p)$ process depends on $p+1$ parameters that can be easily estimated by either maximum likelihood (ML) or matching correlations (MC) procedures. Matching correlation estimators provide a fair estimation of the covariances of the data, even if the model is not well specified.


## Conclusions

- When sampled on equally spaced instants, the $\mathrm{OU}(p)$ family can be written as a discrete time state space model; i.e., a VARMA model in a space of dimension $p$. As a consequence, the families of $\mathrm{OU}(p)$ models are a parsimonious subfamily of the $\operatorname{ARMA}(p, p-1)$ processes in the Gaussian case.
- Furthermore, the coefficients of the ARMA can be deduced from those of the corresponding $\mathrm{OU}(p)$.
- We have shown examples for which the ML-estimated OU model is able to capture a long term dependence that the ML-estimated ARMA model does not show. This leads to recommend the inclusion of OU models as candidates to represent stationary series to the users interested in such kind of dependence.


[^0]:    ${ }^{1}$ When $\Lambda$ is a Wiener process, it is in fact an $\operatorname{ARMA}(p, p-1)$

[^1]:    ${ }^{2}$ Barndorff- Nielsen, O. and Shephard, N. (2001), Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. JRSS, 63.

[^2]:    ${ }^{3}$ Normality is rejected in $100 \%$ of all cases.

[^3]:    ${ }^{4}$ Normality is rejected in $30 \%, 36 \%$ and $36 \%$ of the cases.

