Review of Discrete and Continuous Processes in Finance
Theory and Applications

Attilio Meucci[1]
attilio_meucci@symmys.com

This version: May 2010
last version available at ssrn.com

Abstract

We review the main processes used to model financial variables. We emphasize the parallel between discrete-time processes, mainly used by econometricians for risk- and portfolio-management, and their continuous-time counterparts, mainly used by mathematicians to price derivatives. We highlight the relationship of such processes with the building blocks of stochastic dynamics and statistical inference, namely the invariants. Figures and practical examples support intuition. Fully documented code illustrating these processes in practice is available at MATLAB Central File Exchange under the author's name.

JEL Classification: C1, G11

Keywords: invariants, random walk, Levy processes, autocorrelation, ARMA, Ornstein-Uhlenbeck, long memory, fractional integration, fractional Brownian motion, volatility clustering, GARCH, stochastic volatility, subordination, real measure, risk-neutral measure, fat tails.

[1] The author is grateful to Gianluca Fusai, Fabio Mercurio and Arne Staal for their helpful feedback

Contents

1 Introduction

Since the financial markets evolve in a random fashion, the natural object to model its evolution are stochastic processes. Stochastic processes are thus essential in risk management, portfolio management and trade optimization to model the evolution of the risk factors that determine the price of a set of securities at a given point in the future. We denote by X_t the value of one such factor X at a generic time t, i.e. the stochastic process for X. This stochastic process is fully described by a multivariate function, the process cumulative distribution function, i.e. the joint probability of its realizations at any set of times

$$F_{t_1,t_2,\dots}(x_1, x_2, \dots) \equiv \mathbb{P}\{X_{t_1} \leq x_1, X_{t_2} \leq x_2, \dots\}. \tag{1}$$

For the above mentioned purposes of risk management and portfolio management , the monitoring times $t_1, t_2, \dots$ are typically chosen on a discrete, equally-spaced, grid $t, t+\Delta, t+2\Delta, \dots$: for instance Δ can be one day, or one week, or one month, etc.

The probability measure "$\mathbb{P}$" is the "real" probability that governs the evolution of the process. In reality, such probability is not known and needs to be estimated: indeed, estimation is the main concern of the real-measure $\mathbb{P}$-world.

$$\mathbb{P} : \textbf{estimate the future} \tag{2}$$

There exists a parallel, much developed, area of application of stochastic processes in finance: derivatives pricing. According to the *fundamental theorem of pricing*, the normalized price $\widetilde{P}_t$ of a security is arbitrage-free only if there exists a fictitious stochastic process describing its future evolution such that its expected value is always constant and equals the current price:

$$\widetilde{P}_t = \mathbb{E}\left\{\widetilde{P}_{t+\tau}\right\}, \quad \tau \geq 0. \tag{3}$$

A process satisfying (3) is called a *martingale*: since a martingale does not reward risk, the fictitious probability of the pricing process is called *risk-neutral* and is typically denoted by "$\mathbb{Q}$". Since some prices are liquid and readily observable, (3) becomes a powerful tool to assign a fair price to a non-liquid security, once a fictitious martingale has been calibrated to other traded securities. In other words, derivatives pricing is a very high-tech interpolation exercise which lives in the risk neutral $\mathbb{Q}$-world

$$\mathbb{Q} : \textbf{interpolate the present} \tag{4}$$

Notice that martingales represent a very restrictive class of processes. A tractable and much more general class of processes are the so-called *semimartingales*, on which a theory of integration can be defined that makes sense for financial modeling. However, for risk and portfolio management, we do not need to restrict our attention to semimartingales: any process that suitably models the behavior of financial times series is a-priori admissible.

To summarize, risk management, portfolio management and trade optimization rely on discrete-time processes that live in the real-world probability measure $\mathbb{P}$. Derivatives pricing relies on continuous-time processes that live in a fictitious risk-neutral probability measure $\mathbb{Q}$.

Interactions between the $\mathbb{P}$-world and the $\mathbb{Q}$-world occur frequently. For instance, the sensitivities of a security to a set of risk factors, the so-called *Greeks*, are computed using $\mathbb{Q}$ models, but then they are applied in the $\mathbb{P}$-world for hedging purposes. Similarly, the evolution through time of the $\mathbb{Q}$ parameters of a given pricing model can be used to predict the $\mathbb{P}$-world distribution of the same prices in the future and generate such statistics as the value at risk.

Here we present an overview of the main processes in finance, highlighting the relationships between the $\mathbb{P}$-world, discrete-time models and their $\mathbb{Q}$-world, continuous-time counterparts.

We start in Part I with the $\mathbb{P}$-world discrete-time processes. In Section 2 we introduce the random walk, which is the cumulative sum of invariants. This process represents the benchmark for buy-side modelling: often considered too simplistic by $\mathbb{Q}$ quants and $\mathbb{P}$ econometricians not involved in risk and portfolio management, the random walk proves to be very hard to outperform when estimation error is accounted for. In Section 3 we relax the hypothesis that the increments of the process be invariant: we introduce autocorrelation, thereby obtaining ARMA processes. In Section 4 we account for the empirical observations that at times autocorrelation, when present, decays very slowly with the number of lags: this behavior is suitably modeled by long memory-processes. In Section 5 we discuss volatility clustering: the scatter of the process, rather than the process itself, displays autocorrelation: GARCH and generalizations thereof capture this feature.

In Part II we discuss the $\mathbb{Q}$-world continuous-time counterparts of the above $\mathbb{P}$-world models. In Section 6 we present Levy processes, the continuous-time version of the random walk. In Section 7 we model autocorrelation with the Ornstein-Uhlenbeck and related processes. In Section 8 we model long memory by means of the fractional Brownian motion. In Section 9 we tackle volatility clustering in two flavors: stochastic volatility and subordination.

We summarize in the table below the processes covered and their key characteristics

	discrete time ($\mathbb{P}$)	continuous time ($\mathbb{Q}$)
base case	random walk	Levy processes
autocorrelation	ARMA	Ornstein-Uhlenbeck
long memory	fractional integration	fractional Brownian motion
volatility clustering	GARCH	stochastic volatility subordination

Fully documented code that illustrates the theory and the empirical aspects of the models in this article is available at MATLAB Central File Exchange under the author's name.

Part I
$\mathbb{P}$: discrete-time processes

We assume that we monitor a financial process at discrete, equally spaced time intervals. Without loss of generality, we can assume this interval to have unit length: therefore, denoting time by t, we assume that $t \in \mathbb{Z}$.

2 Random walk

The random walk with drift is a process that cumulates invariants

$$X_{t+1} = X_t + \epsilon_{t+1}, \tag{5}$$

The *invariants* ϵ_t, are independent and identically distributed (i.i.d.) shocks with *any* distribution. Since the distribution of the invariants is fully general, the random walk has a much richer structure than one would expect.

Two very simple tests to spot invariance in a realized time series are discussed in Meucci (2005).

First, due to the Glivenko-Cantelli theorem, if ϵ_t is an invariant the histogram of its time series represents its probability density function (pdf). Therefore, also any portion of its time series represents the pdf. As a consequence, the histogram of the first half of the time series of ϵ_t and that of the second half should look alike, see the top portion of Figure 1 and the discussion in the examples below.

Second, if ϵ_t is an invariant, then ϵ_t and ϵ_{t+1} are independent and identically distributed. Therefore the mean-covariance ellipsoid of ϵ_t and ϵ_{t+1} must be a circle. A scatter plot of ϵ_t versus ϵ_{t+1} with the plot of the mean-covariance ellipsoid should convey this impression, see the bottom portion of Figure 1 and the discussion in the examples below.

Consider a security that trades at time t at the price P_t and consider the *compounded return*

$$C_t \equiv \ln\left(\frac{P_t}{P_{t-1}}\right). \tag{6}$$

In the equity world the benchmark assumption is that the compounded returns are invariants, i.e. $C_t \equiv \epsilon_t$. This amounts to saying that $X_t \equiv \ln P_t$ evolves according to a random walk:

$$\ln(P_{t+1}) = \ln(P_t) + \epsilon_{t+1}. \tag{7}$$

The plots in Figure 1 refer to this case. Also refer to Figure 3.2 and related discussion in Meucci (2005).

As another standard example, denote by $Z_t^{(E)}$ the price at time t of a zero-coupon bond that matures at time E. As discussed in Meucci (2005), the raw

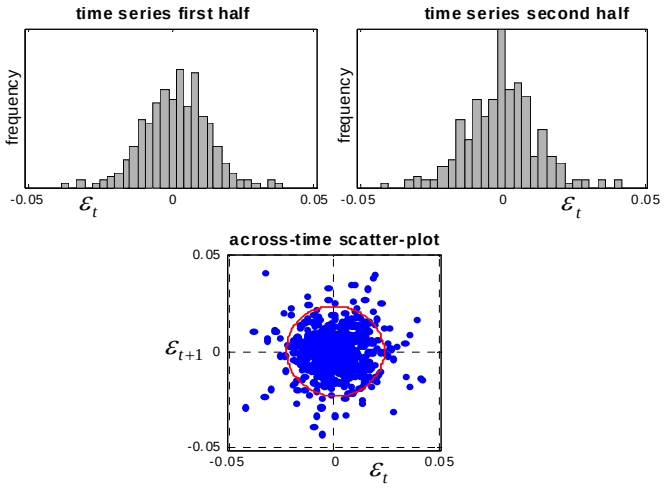

Figure 1: Invariance check

bond price cannot give rise to an invariant, because the approaching maturity dates breaks the time translation invariance of the analysis. Therefore, one introduces the *yield to maturity*, which is the annualized compounded return over the life of the bond

$$Y_t^{(\upsilon)} \equiv -\frac{1}{\upsilon} \ln\left(Z_t^{(t+\upsilon)}\right). \tag{8}$$

The time series of the yield *to* maturity, as opposed to the yield *of* maturity, is invariant under time translations and therefore we can hope to use it to generate invariants. Notice however that this time series does not correspond to the price of one single bond throughout, but rather to the price of a different bond at each time. In the fixed-income world the benchmark assumption is that the changes in yield to maturity are invariants

$$Y_{t+1}^{(\upsilon)} = Y_t^{(\upsilon)} + \epsilon_{t+1}^{(\upsilon)}, \tag{9}$$

which is clearly in the format (5). One can check this assumption with the simple test described in Figure 1. Also refer to Figure 3.5 and related discussion in Meucci (2005).

2.1 Continuous invariants

As a base case, the invariants can have a *normal distribution*

$$\epsilon_t \sim \mathrm{N}\left(\mu, \sigma^2\right), \tag{10}$$

where μ is the expectation and σ^2 is the variance. Independent normal distributions are closed under the sum. With slight abuse of notation we write:

$$\mathrm{N}\left(\mu_1, \sigma_1^2\right) + \mathrm{N}\left(\mu_2, \sigma_2^2\right) \stackrel{d}{=} \mathrm{N}\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right). \tag{11}$$

Since the sum of independent normal variables is normal, from normal invariants follows a normally distributed random walk, i.e. a Brownian motion, see Section 6.1.

Different continuous distributions for the invariants give rise to different distributions for the random walk (5). For instance, one can model the invariants using stable distributions, Student t distributions, skewed distributions, etc.

2.2 Discrete invariants

The invariants can also take on a set $\mathbf{a} \equiv \{a_0, \ldots, a_K\}$ of discrete values with respective probabilities $\mathbf{p} \equiv \{p_0, \ldots, p_K\}$ that sum to one. This is the *generalized Bernoulli distribution*:

$$\epsilon_t \sim \mathrm{Be}\left(\mathbf{p}; \mathbf{a}\right). \tag{12}$$

If the invariant has a generalized Bernoulli distribution, the ensuing random walk also has a generalized Bernoulli distribution with different parameters at each time step.

When the number of values in $\mathbf{a}$ is large, a more parsimonious parameterization becomes necessary. An important case is the *Poisson distribution*:

$$\epsilon_t \sim \mathrm{Po}\left(\lambda; \Delta\right). \tag{13}$$

This represents a special case of (12) where the variable takes values on an infinite, equally spaced grid $a_k \equiv k\Delta$, also denoted as $\Delta\mathbb{N}$, with probabilities

$$p_k \equiv \mathbb{P}\left\{\epsilon_t = k\Delta\right\} \equiv \frac{\lambda^k e^{-\lambda}}{k!}. \tag{14}$$

It follows that λ represents both the expectation and the variance of the invariant (13), as one can prove using the interchangeability of summation and derivation in the power series expansion of the exponential function.

Independent Poisson distributions are closed under the sum. With slight abuse of notation we write:

$$\mathrm{Po}\left(\lambda_1; \Delta\right) + \mathrm{Po}\left(\lambda_2; \Delta\right) \stackrel{d}{=} \mathrm{Po}\left(\lambda_1 + \lambda_2; \Delta\right). \tag{15}$$

Therefore the random walk (5) ensuing from a Poisson invariant is Poisson distributed on the same grid.

2.3 Generalized representations

Since any function of invariants is an invariant, we can create invariants by aggregation.

2.3.1 Stochastic volatility

A flexible formulation in this direction is the *stochastic volatility* formulation:

$$\epsilon_t \overset{d}{=} \mu_t + \sigma_t Z_t. \tag{16}$$

In this expression μ_t is a fully general invariant, σ_t is a positive invariant and Z_t is a general invariant with unit scatter and null location: these invariants represent respectively the location, the scatter and the shape of ϵ_t. specification.

For instance, for specific choices of the building blocks we recover the *Student t distribution*:

$$\left.\begin{array}{l} \mu_t \equiv \mu \\ \nu/\sigma_t^2 \sim \chi^2_\nu \\ Z_t \sim \mathrm{N}\left(0, \sigma^2\right) \end{array}\right\} \Rightarrow \epsilon_t \sim \mathrm{St}\left(\nu, \mu, \sigma^2\right). \tag{17}$$

We stress that the term stochastic volatility is used mostly when σ_t in (16) is not an invariant, see Section 5.

2.3.2 Mixture models

Another flexible aggregate representation is the *mixture model*

$$\epsilon_t \overset{d}{=} (1 - B_t)\, Y_t + B_t Z_t \tag{18}$$

where Y_t and Z_t are generic invariants and B_t are invariants with a standard Bernoulli distribution (12) with values in $\{0, 1\}$: the ensuing model for ϵ_t is a mixture of Y_t and Z_t respectively, with a given probability: it is useful, for instance, to model the *regime switch* between a regular market environment Y_t and a market crash model Z_t. More general mixture models select among S states

$$\epsilon_t \overset{d}{=} \sum_{s=1}^{S} B_t^{(s)} Y_t^{(s)}, \tag{19}$$

where $\mathbf{B}_t$ is a multiple-selection process with values in the canonical basis of $\mathbb{R}^S$.

2.3.3 Stochastic volatility as mixture models

Stochastic volatility models can be seen as special cases of multi-mixture models with a continuum of components. Indeed

Consider a generic stochastic volatility model (16), which we report here

$$\epsilon_t \overset{d}{=} \mu_t + \sigma_t Z_t. \tag{20}$$

Denote by $f_{\mu,\sigma}$ the joint pdf of the location and scatter invariants μ_t and σ_t respectively. Define the following multi-mixture model

$$\epsilon_t \equiv \int\int B_t^{(\mu,\sigma)} Y_t^{(\mu,\sigma)} d\mu d\sigma, \tag{21}$$

where

$$Y_t^{(\mu,\sigma)} \stackrel{d}{=} \mu + \sigma Z_t \tag{22}$$

and $B_t^{(\mu,\sigma)}$ is a continuous multiple-selection process with values in the indicators of infinitesimal intervals of $\mathbb{R} \times \mathbb{R}^+$ such that

$$\mathbb{P}\left\{B_t^{(\mu,\sigma)} = \mathbb{I}_{[d\mu \times d\sigma]}\right\} = f_{\mu,\sigma}(\mu,\sigma)\, d\mu d\sigma. \tag{23}$$

Then the multi-mixture model (21) is distributed as the stochastic volatility model (20).

From (17) the Student t distribution is a special instance of the stochastic volatility representation (16) and thus it is also a multi-mixture distribution with a continuum of components: Z_t in (22) is standard normal, and $f_{\mu,\sigma} \equiv f_\mu f_\sigma$, where f_μ is a Dirac delta and f_σ is the density of the square root of the inverse chi-square, suitably rescaled.

2.4 Heavy tails

Heavy-tailed distributions are those whose pdf decreases away from the location of the distribution slower than exponentially. Suitable models for heavy tails can be found among the continuous distributions, the discrete distributions, or the generalized representations discussed above.

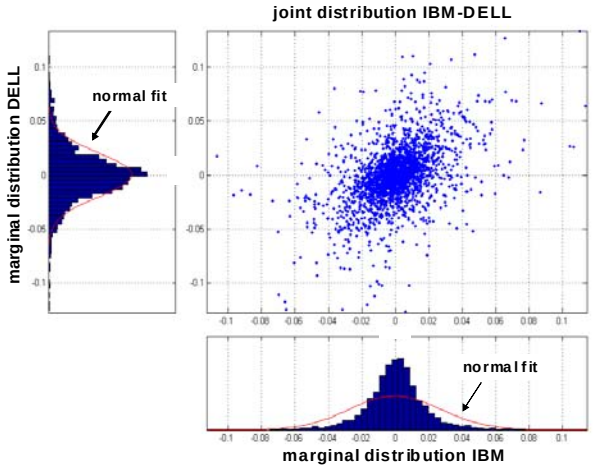

Figure 2: Heavy tails in empirical distribution of compounded stock returns

Stable distributions are continuous models popular for their tractability, see e.g. Rachev (2003) and Meucci (2005): if the invariants ϵ_t are stable, the random walk X_t has the same distribution as ϵ_t, modulo rescaling.

Another popular parametric model for heavy tails is the Student t distribution (17), which for $\nu > 2$ has finite variance. For $\nu \equiv 1$ the Student t distribution is stable and it is known as *Cauchy distribution*.

In Figure 2 we display the empirical distribution of the daily compounded returns (6) of two technology stocks. It is apparent that a normal fit is unsuitable to describe the heavy-body, heavy tail empirical behavior of these returns, refer to MATLAB Central File Exchange under the author's name for the code.

The heavy tails of a random walk tend to disappear with aggregation. More precisely, consider the aggregate invariant

$$\widetilde{\epsilon}_{t,\tau} \equiv \epsilon_t + \epsilon_{t-1} \cdots + \epsilon_{t-\tau+1}. \tag{24}$$

If the one-step dynamics is the random walk (5), the τ-step dynamics is also a random walk:

$$X_{t+\tau} = X_t + \widetilde{\epsilon}_{t+\tau,\tau}, \tag{25}$$

However, the *central limit theorem* prescribes that, if the invariant ϵ_t has finite variance, the aggregate invariant $\widetilde{\epsilon}_{t,\tau}$ becomes normal for large enough an aggregation size.

In particular It follows that stable distributions must have infinite variance, or else they would violate the central limit theorem.

Therefore, random walks for large enough an aggregation size behave like discrete samples from the Brownian motion, see Section 6.1.

From (24) it follows that the variance of the steps of a random walk is a linear function of the time interval of the step. This is the *square root rule* propagation of risk: risk, defined as the standard deviation of the step, increases as the square root of the time interval:

$$\mathrm{Sd}\left\{\widetilde{\epsilon}_{t,\tau}\right\} \propto \sqrt{\tau}. \tag{26}$$

For stable distributions, where the variance is not defined, one can use other definitions of risk, such as the inter-quantile range. In particular, for the Cauchy distribution one can check that the propagation law of risk does not grow as the square root of the time interval, but instead it grows linearly with the time interval.

3 ARMA processes

The steps of a random walk display no dependence across time, because they are the invariants. This feature makes random walks not stationary: if X_t evolves

as in (5), its distribution never stabilizes. To model stationarity we consider the one-lag *autoregressive process*:

$$X_{t+1} = aX_t + \epsilon_{t+1}, \tag{27}$$

where ϵ_t are the invariants. The limit $a \equiv 1$ is the random walk (5). If $|a| < 1$ the distribution of X_t eventually stabilizes and thus the process becomes stationary. The autocorrelation of this process reads:

$$\text{Cor}\{X_t, X_{t-\tau}\} = e^{(\ln a)\tau}, \tag{28}$$

see Appendix A.1.

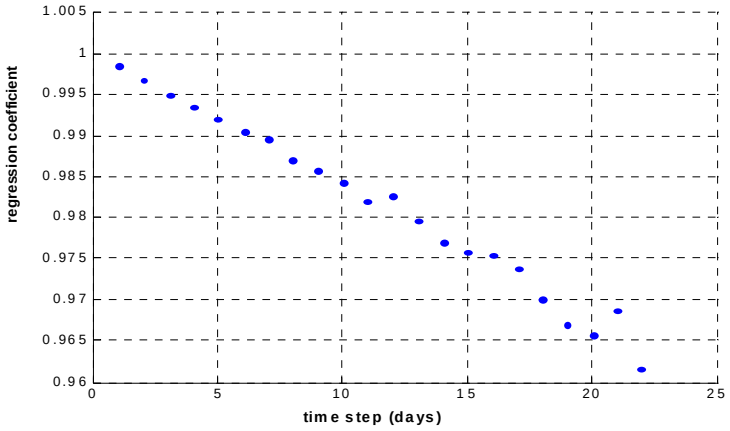

Figure 3: AR(1) fit of five-year swap rate

For instance, interest rates cannot evolve as random walks at any scale. If (9) held true exactly then for any aggregation size τ the following would hold true

$$Y^{(\upsilon)}_{t+\tau} = Y^{(\upsilon)}_t + \tilde{\epsilon}^{(\upsilon)}_{t+\tau,\tau}, \tag{29}$$

where the aggregate invariant $\tilde{\epsilon}^{(\upsilon)}_{t+\tau}$ is defined as in (24). However, rates cannot diffuse indefinitely. Therefore, for aggregation size τ of the order of a month or larger mean-reverting effects must become apparent. We see this in Figure 3 where we plot the fitted value of a as a function of the aggregation size for the time series of the five-year par swap rate.

We can generalize (27) by adding more lags of the process and of the invariant. The so-called autoregressive moving average process, or ARMA(p, q) process is defined as

$$\prod_{j=1}^{p} (1 - a_j \mathcal{L})\, X_t = D_t + \prod_{j=1}^{q} (1 - b_j \mathcal{L})\, \epsilon_t, \tag{30}$$

where $\mathcal{L}$ denotes the lag operator $\mathcal{L}X_t \equiv X_{t-1}$, and D_t is a deterministic component that we have added in such a way that we can assume the location of ϵ_t to be zero. It can be proved that the process X_t stabilizes in the long run, and therefore it is stationary, only if all the a's in (30) lie inside the unit circle, see Hamilton (1994). The fast, exponential decay (28) of the autocorrelation is common to all ARMA(p, q) processes with finite p and q.

It is always possible to switch from an ARMA(p, q), to an ARMA$(0, \infty)$ or an ARMA$(\infty, 0)$ representation. Indeed, the following identity holds

$$(1 - \gamma\mathcal{L})^{-1} \equiv \sum_{k=0}^{\infty} (\gamma\mathcal{L})^k \,, \tag{31}$$

which can be checked by applying the operator $(1 - \gamma\mathcal{L})$ to the right hand side. Therefore, we can remove iteratively all the lags from the MA portion of an ARMA(p, q) process, making it ARMA$(\infty, 0)$, as long as $q < \infty$. Similarly, we can remove all the lags from the AR portion of an ARMA(p, q) process, making it ARMA$(0, \infty)$, as long as $p < \infty$.

Although the converse is not true, the ARMA$(0, \infty)$ is extremely important because of *Wold's theorem*, which states that any stationary process can be expressed as ARMA$(0, \infty)$. The infinite parameters of such a process are impossible to estimate, but a parsimonious ARMA(p, q) should in most cases provide a good approximation.

4 Long memory

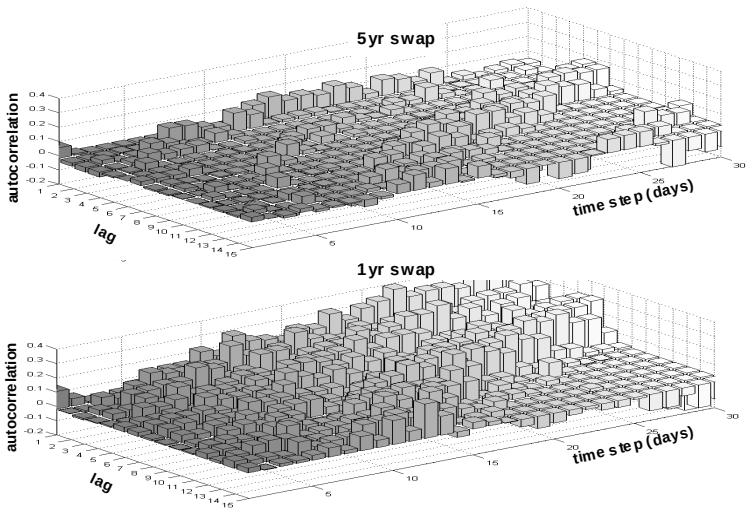

Figure 4: Autocorrelation decay properties of swap rate changes

Long memory is a phenomenon whereby the autocorrelation decays slower than exponentially, see e.g. Beran (1994).

For instance, in Figure 4 we display the autocorrelation decay of the five-year and the one year swap rates changes. The five year autocorrelation quickly decays to zero. This is consistent with (29) and comments thereafter. On the other hand, the one-year autocorrelation consistently displays a small, yet non-zero pattern. This pattern is persistent, or else the non-linear aggregation properties of the sample variance displayed in Figure 5 would not be justified, refer to MATLAB Central File Exchange under the author's name for the code.

Wold's theorem states that any stationary process can be represented as an ARMA process. However, a parsimonious, finite-lag ARMA specification gives rise to an exponential decay in the autocorrelation and thus it does not yield long memory. On the other hand, a generic infinite-lag ARMA specification does not make sense from an estimation perspective because of the excessive number of parameters to estimate. Therefore, in order to model long-memory, it becomes necessary to impose a parametric structure on an infinite-lag ARMA process.

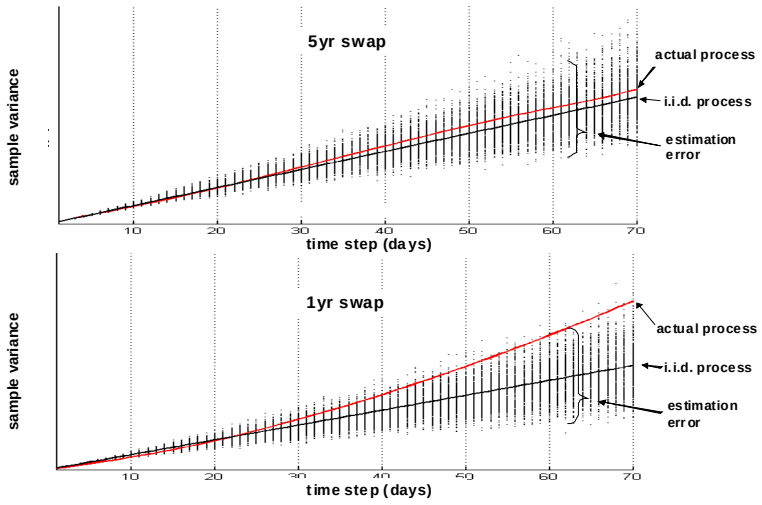

Figure 5: Variance aggregation properties of swap rate changes

This can be achieved by generalizing the random walk (5) into a fractionally integrated process. First we twist the invariants as follows

$$\widetilde{\epsilon}_t = (1-\mathcal{L})^{-d}\,\epsilon_t, \tag{32}$$

where $0 \leq d < 1/2$; then we define

$$X_{t+1} = X_t + \widetilde{\epsilon}_{t+1}, \tag{33}$$

If $d \equiv 0$ in (32) we obtain the standard random walk. If $0 < d < 1/2$ we obtain a highly structured integrated ARMA$(0, \infty)$ series. Indeed, the following identity holds

$$(1 - \mathcal{L})^{-d} = \sum_{k=0}^{\infty} \frac{\Gamma(k+d)}{\Gamma(k+1)\,\Gamma(d)} \mathcal{L}^k. \tag{34}$$

In particular, the autocorrelation of the shocks decays according to a power law

$$\text{Cor}\{\widetilde{\epsilon}_t, \widetilde{\epsilon}_{t-\tau}\} \approx \frac{\Gamma(1-d)}{\Gamma(d)} \tau^{2d-1}, \tag{35}$$

see Baillie (1996). Therefore the fractional process (32) suitably describes the long memory phenomenon. Notice that the autocorrelation decay is reflected in a non-linear, power-law increase of the variance of $X_{t+\tau} - X_t$ as a function of τ.

For instance, the non-linear aggregation properties of the sample variance of the five-year swap rate changes displayed in Figure 5 are consistent with a value $d \approx 0$, whereas for the one-year swap rate changes we have $d \approx 0.1$, refer to MATLAB Central File Exchange under the author's name.

Further generalization of the fractional process (32)-(33) include adding several fractional lags to both the process and the invariant, similarly to (30).

5 Volatility clustering

Volatility clustering is the phenomenon whereby the scatter of a financial variable displays evident autocorrelation, whereas the variable itself does not. This is the case among others for the daily returns of most stocks, see e.g. Ghoulmie, Cont, and Nadal (2005) and Cont (2005).

For instance, in the left portion of Figure 6 we display the scatter plot of the daily returns of the IBM stock versus their lagged values one day before. The location-dispersion ellipsoid is a circle and thus the log-returns are not autocorrelated.

On the other hand, in the right portion of Figure 6 we display the scatter plot of the squared daily returns of the IBM stock versus their lagged values one day before. The location-dispersion ellipsoid now is an ellipse tilted across the positive quadrant: the square log-returns, which proxy volatility, are positively autocorrelated, refer to MATLAB Central File Exchange under the author's name.

The generalized representation (16) of the invariants suggests to model such volatility clustering in terms of stochastic volatility:

$$\epsilon_t \overset{d}{=} \mu_t + \sigma_t Z_t. \tag{36}$$

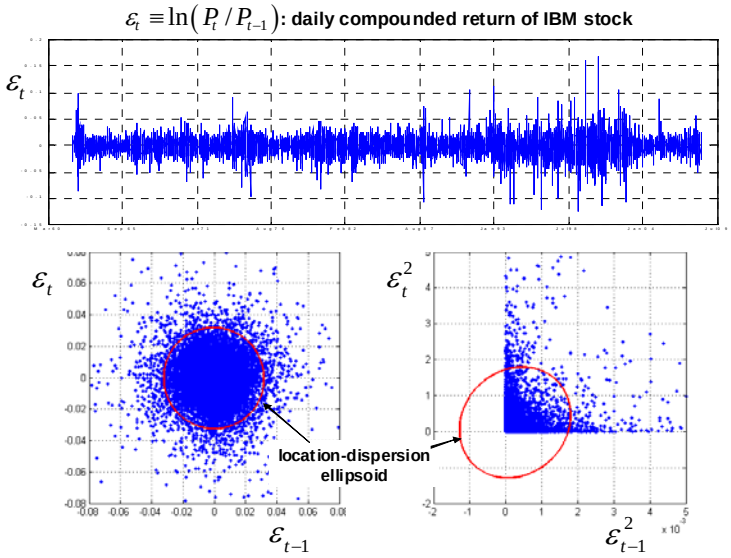

Figure 6: Autocorrelation properties of stocks log-returns

In this expression, the location invariant is typically constant $\mu_t \equiv \mu$ and Z_t is a generic standardized shape invariant. However, unlike in (16), the scatter term σ_t displays a *non-invariant* dynamics. The dynamics of σ_t can be specified in terms of an ARMA(p, q) process, a long memory process, or any other model that is not i.i.d. across time.

In particular, if the process for σ_t^2 is ARMA(p, q), with the *same* set of invariants Z_t as in (36), then the process ϵ_t, which is no longer an invariant, is called a generalized autoregressive conditional heteroskedastic, or GARCH(p, q), process. Among the most popular and parsimonious such specifications is the GARCH$(1, 1)$ model:

$$\sigma_t^2 = \sigma^2 + a\sigma_{t-1}^2 + bZ_{t-1}^2, \tag{37}$$

which is stationary if $0 < a + b < 1$.

In Figure 11 we display a sample path from the GARCH process, along with a sample path of the volatility (66) that generates it, refer to MATLAB Central File Exchange under the author's name.

If the process for σ_t^2 is driven by sources of randomness other than the invariants Z_t that appear in (36) then we have the most common type of stochastic volatility models, where the variance is not conditionally deterministic.

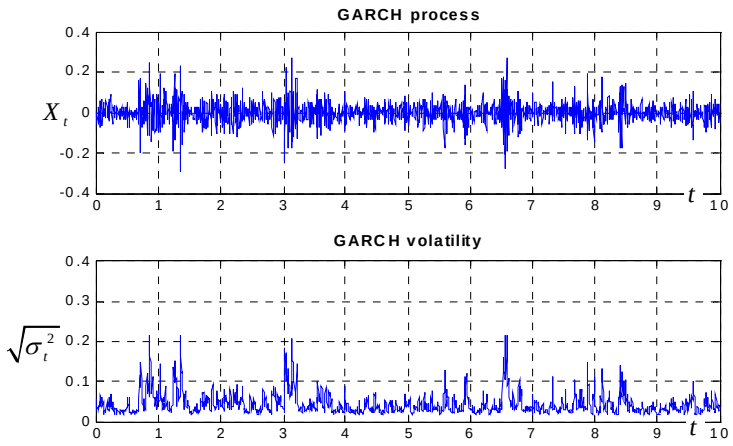

Figure 7: GARCH model for volatility clustering

Part II
$\mathbb{Q}$: continuous-time processes

Here we discuss stochastic processes in continuous time: therefore we assume $t \in \mathbb{R}$. In order to support intuition, each section mirrors its discrete-time counterpart in Part I.

6 Levy processes

A Levy process X_t is the continuous-time generalization of the random walk with drift. Here we highlights the main features; for more on this subject refer to Schoutens (2003) and Cont and Tankov (2008).

A Levy process is by definition a process such that its increments over any time interval are invariants, i.e. i.i.d. random variables. Therefore, any invariant discussed in Section 2 gives rise to a Levy process, as long as the distribution of the invariant is *infinitely divisible*. This technical condition is important because it ensures that each invariant can be represented in turns as the sum of invariants. More precisely, for an invariant is infinitely divisible if for any integer K the following holds

$$\begin{aligned} \epsilon_t \equiv X_t - X_{t-1} &= \left(X_t - X_{t-\frac{1}{K}}\right) + \cdots + \left(X_{t-1+\frac{1}{K}} - X_{t-1}\right) \\ &\stackrel{d}{=} Z_1 + \cdots + Z_K, \end{aligned} \tag{38}$$

where the Z_k's are i.i.d. This requirements is important to guarantee that we

can consider time intervals of any size in the definition of the invariant.

Levy processes are fundamentally of two kinds: Brownian diffusion and Poisson jumps. As it turns out, any Levy process can be generated with these fundamental bricks.

6.1 Diffusion

In order to construct a stochastic diffusion, we need its building blocks, i.e. the increments over any time interval, to be infinitely divisible and to span a continuum of potential values. In order for the model to be tractable, we need to be able to parametrize the distribution of these increments. Therefore, it is natural to turn to continuous stable distributions, and in particular to the one stable distribution that displays finite variance: the normal distribution

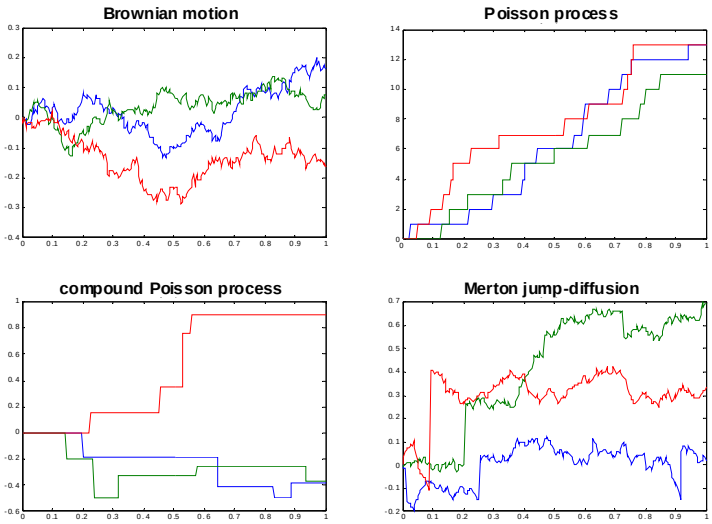

Figure 8: Sample paths from Levy processes

The *Brownian motion* with drift B_t^{μ,σ^2} is a continuous, diffusive process whose invariants are normally distributed with expectation and variance proportional to the time lag between the increments:

$$\epsilon_{t,\tau} \equiv B_t^{\mu,\sigma^2} - B_{t-\tau}^{\mu,\sigma^2} \sim \mathrm{N}\left(\mu\tau, \sigma^2\tau\right). \tag{39}$$

This formulation generalizes (10). Since the sum of independent normal variables is normal, the increments of the Brownian motion are infinitely divisible as in (38), where each term Z_k is normally distributed. Therefore, the Brownian motion is properly defined and it is normally distributed at any time.

The benchmark process in the $\mathbb{Q}$ -measure quantitative finance is the Black-Scholes-Merton *geometric Brownian motion*, which is used to model, among many others financial variables, the price P_t of stocks. This is obtained by modeling the variable $X_t \equiv \ln P_t$ as a Brownian motion.

$$\ln P_t \overset{d}{\equiv} B_t^{\mu,\sigma^2}. \tag{40}$$

Notice that the random walk (7) is much more general, in that the invariants, i.e. the compounded returns, need not be normally distributed. In Figure 8 we plot a few paths from the Brownian motion, refer to MATLAB Central File Exchange under the author's name for the code.

6.2 Jumps

In order to construct a stochastic jump, we need its building blocks, i.e. the increments over any time interval, to be infinitely divisible and to span a discrete set of potential values. In order for the model to be tractable, we need to be able to parametrize the distribution of these increments. From (15) it follows that the natural choice in this case is the Poisson distribution.

The *Poisson process* $P_t^{\Delta,\lambda}$ jumps by positive multiples of a base (positive or negative) step Δ according to a Poisson distribution:

$$\epsilon_{t,\tau} \equiv P_t^{\Delta,\lambda} - P_{t-\tau}^{\Delta,\lambda} \sim \mathrm{Po}\left(\lambda\tau; \Delta\right). \tag{41}$$

Since from (15) the sum of independent Poisson variables has a Poisson distribution, the increments of the Poisson process are infinitely divisible as in (38), where each term Z_k is Poisson distributed. Therefore, the Poisson process is properly defined and it is Poisson distributed at any non-integer time.

In 8 we plot a few paths from the standard Poisson process, which takes values on the integer grid, refer to MATLAB Central File Exchange under the author's name for the code.

As the time step decreases so does the probability of a jump. Indeed from (41) and (14) we obtain

$$\begin{aligned}
\mathbb{P}\left\{\epsilon_{t,\tau} = 0\right\} &\approx 1 - \lambda\tau & (42)\\
\mathbb{P}\left\{\epsilon_{t,\tau} = \Delta\right\} &\approx \lambda\tau & (43)\\
\mathbb{P}\left\{\epsilon_{t,\tau} > \Delta\right\} &\approx 0. & (44)
\end{aligned}$$

Since $\mathrm{E}\left\{\epsilon_{t,\tau}\right\} = \lambda\tau$, the parameter λ plays the role of the *intensity* of the jumps.

6.3 Generalized representations

The generic Levy process is a sum of the above processes. From (11) and (39), the weighted sum of independent Brownian motions is a Brownian motion:

$$\sum_{k=1}^{K} B_{t,k}^{\mu_k,\sigma_k^2} \stackrel{d}{=} B_t^{\mu,\sigma^2}, \tag{45}$$

where $\mu \equiv \sum_{k=1}^{K} \mu_k$ and $\sigma^2 \equiv \sum_{k=1}^{K} \sigma_k^2$. Also, from (15) and (41), the sum of independent Poisson processes with the same base step Δ is a Poisson process with that base step:

$$\sum_{k=1}^{K} P_t^{\Delta,\lambda_k} \stackrel{d}{=} P_t^{\Delta,\lambda}, \tag{46}$$

where $\lambda \equiv \sum_{k=1}^{K} \lambda_k$. Therefore we can construct more general Levy processes from our building blocks by considering a continuum of terms in the sum as follows

$$X_t = B_t^{\mu,\sigma^2} + \int_{-\infty}^{+\infty} P_t^{\Delta,\lambda(\Delta)} d\Delta, \tag{47}$$

where the intensity $\lambda(\Delta)$ determines the relative weight of the discrete Poisson jumps on the grid $\Delta\mathbb{N}$.

As it turns out, this representation is exhaustive: any Levy process, stemming from slicing as in (38) an arbitrarily distributed infinitely divisible invariant, can be represented as (47)[2]. As we show in Appendix A.3, from (47) follows the *Levy-Khintchine* representation of Levy processes in terms of their characteristic function:

$$\ln\left(\mathrm{E}\left\{e^{i\omega X_t}\right\}\right) = i\omega\mu t - \frac{1}{2}\sigma^2\omega^2 t + t\int_{-\infty}^{+\infty}\left(e^{i\omega\Delta} - 1\right)\lambda(\Delta)\, d\Delta. \tag{48}$$

An alternative representation of the Levy processes follows from a result by Monroe (1978), according to which Levy processes can be written as

$$X_t \stackrel{d}{=} B_{T_t}^{\mu,\sigma^2}, \tag{49}$$

where B_t is a Brownian motion and T_t is another Levy process that never decreases, called *subordinator*. In practice, T_t is a *stochastic time* that indicates the activity of the financial markets. The specific case $T_t \equiv t$ recovers the Brownian motion.

6.4 Notable examples

In addition to (geometric) Brownian motion and Poisson processes, another important class of Levy processes are the *α-stable processes* by Mandelbrot (1963),

[2] This is a slightly abridged version that avoids the problem of infinite activity near the zero

see also Samorodnitsky and Taqqu (1994), Rachev (2003), whose distribution is closed under the sum. However, in order not to contradict the *central limit theorem* these processes have infinite variance.

Another important subclass are the *compound Poisson processes*, which generalize Poisson processes by allowing jumps to take on random values, instead of values in a fixed grid $\Delta\mathbb{N}$. The compound Poisson processes can be expressed as

$$X_t = B_t^{\mu,\sigma^2} + \sum_{n=1}^{P_t^\lambda} Z_n, \tag{50}$$

where P_t^λ is a Poisson process with values in the unit-step grid $\mathbb{N}$ and Z_n are arbitrary i.i.d. variables. As we show in Appendix A.4, the parameterization (47) for jump-diffusion models (50) is given by

$$\lambda(\Delta) \equiv \lambda f_Z(\Delta), \tag{51}$$

were f_Z is the pdf of Z_n.

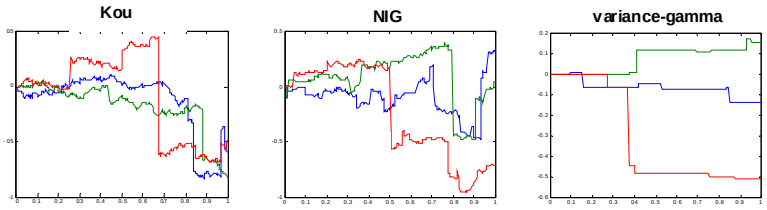

Figure 9: Sample paths from Levy processes

A notable example in the class of compound Poisson processes is the the-jump diffusion by Merton (1976), where Z_n is normally distributed. We plot in Figure 8 a few paths sampled from this process, refer to MATLAB Central File Exchange under the author's name.

Another tractable jump-diffusion process is the double-exponential by Kou (2002). Other notable parametric Levy processes include the Normal-Inverse-Gamma by Barndorff-Nielsen (1998), the variance-gamma process by Madan and Milne (1991) and the CGMY by Carr, Geman, Madan, and Yor (2003), see Figure 9.

7 Autocorrelated processes

The continuous-time version of the AR(1) process (27) is the *Ornstein-Uhlenbeck process*. Its dynamics is defined by a stochastic differential equation

$$dX_t = -\theta(X_t - \mu)\,dt + dL_t, \tag{52}$$

where μ and $\sigma > 0$ are constants and L is a Levy process. This process is defined for $\theta \geq 0$ and is stationary if $\theta > 0$. This dynamics is easy to interpret: a change in the process is due to a deterministic component, which tends to pull the process back towards the equilibrium asymptote μ at an exponential rate, and a random shock dL_t that perturbs this exponential convergence.

Of particular importance is the case where the Levy driver is a Brownian motion:

$$dX_t = -\theta\left(X_t - \mu\right)dt + \sigma dB_t. \tag{53}$$

As we show in Appendix A.1 this dynamics can be integrated explicitly:

$$X_t \overset{d}{=} \left(1 - e^{-\theta\tau}\right)\mu + e^{-\theta\tau}X_{t-\tau} + \eta_{t,\tau}, \tag{54}$$

where $\eta_{t,\tau}$ are normally distributed invariants

$$\eta_{t,\tau} \sim \mathrm{N}\left(0, \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta\tau}\right)\right). \tag{55}$$

Then the long-term autocorrelation reads:

$$\mathrm{Cor}\left\{X_t, X_{t+\tau}\right\} = e^{-\theta\tau}. \tag{56}$$

A comparison of (27) with (54) yields the correspondence between discrete and continuous time parameters.

Notice that for small time steps τ the process (53) behaves approximately as a Brownian motion. Indeed, a Taylor expansion of the terms in (54)-(55) yields

$$X_t \approx X_{t-\tau} + \epsilon_{t,\tau}, \tag{57}$$

where $\epsilon_{t,\tau}$ is in the form of the normal invariant in (39). On the other hand, as the step τ increases, the distribution of the process stabilizes to its stationary behavior

$$X_t \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{2\theta}\right). \tag{58}$$

The AR(1) process for the five year swap rate in Figure 3 can be seen as the discrete-time sampling of an Ornstein-Uhlenbeck process, which in the context of interest rate modeling is known as the *Vasicek model*, see Vasicek (1977).

Another useful autocorrelated process is obtained by applying Ito's rule to the square $Y_t \equiv X_t^2$ of the process(53) for $m \equiv 0$. As we show in Appendix A.2, this yields the *CIR process* by Cox, Ingersoll, and Ross (1985):

$$dY_t = -\widetilde{\theta}\left(Y_t - \widetilde{m}\right)dt + \widetilde{\sigma}\sqrt{Y_t}dB_t, \tag{59}$$

where the parameters $\widetilde{\theta}$, $\widetilde{m}$ and $\widetilde{\sigma}$ are simple functions of the parameters in (53). This process, is useful to model the evolution of positive random variables.

For instance, the CIR dynamics provides an alternative to the Vasicek model for interest rates. Furthermore, the CIR process can model stochastic volatility, see Section 9.1.

The same way as by adding lags to the one-lag autoregression (27) we obtain the ARMA processes (30), so it is possible to add differentials to the Ornstein-Uhlenbeck process (52): the result are the so-called *continuous autoregressive moving average*, or *CARMA*, processes.

8 Long memory

As discussed in Section 4, some financial variables display long memory: the empirical autocorrelation displays a decay slower than the exponential pattern (56) prescribed by the Ornstein-Uhlenbeck or, in discrete time, by finite-lags ARMA processes.

The continuous-time version of the fractional process with normal invariants is the *fractional Brownian motion*. Similarly to (32)-(34), this process is defined as a structured average of past invariants[3]

$$X_t \equiv \mu t + \sigma \int_{-\infty}^{t} \frac{(t-s)^d}{\Gamma(d+1)} dB_s, \tag{60}$$

where $0 \leq d < 1/2$. This increments of this process

$$X_t = X_{t-\tau} + \widetilde{\epsilon}_{t,\tau}, \tag{61}$$

are normally distributed:

$$\widetilde{\epsilon}_{t,\tau} \sim \mathrm{N}\left(\mu\tau, \sigma^2\tau^{2H}\right), \tag{62}$$

where

$$H \equiv d + \frac{1}{2} \tag{63}$$

is the *Hurst coefficient*.

For $d \equiv 0$ we recover the regular Brownian motion, where the increments are invariants as in (39). If $0 < d < 1/2$, the increments are identically distributed and normal, but they are not independent, and therefore they do not represent invariants.

[3] The integral (60) is not defined in the neighborhood of $-\infty$: this is to be interpreted as a definition of the increments

$$X_t - X_s \equiv \int_{-\infty}^{t} \frac{(t-s)^d}{\Gamma(d+1)} dB_s - \int_{-\infty}^{s} \frac{(t-s)^d}{\Gamma(d+1)} dB,$$

where the limit to $-\infty$ is taken after the subtraction

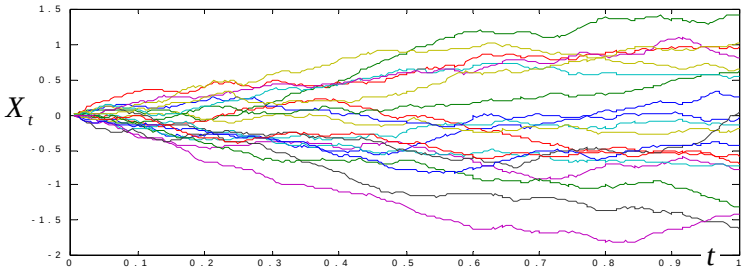

Figure 10: Sample paths from a fractional Brownian motion

Figure 10 displays the persistent pattern of a few sample paths of fractional Brownian motion with Hurst coefficient $H = 0.8$, refer to MATLAB Central File Exchange under the author's name for the code. For a financial application of this process, refer to the one-year swap rate in Figure 5.

Unlike for the Ornstein-Uhlenbeck process (56), the differences of the fractional Brownian motion display a power law decay:

$$\text{Cor}\left\{\Delta X_t, \Delta X_{t+\tau}\right\} \approx d\,(2d+1)\,\tau^{2d-1}, \tag{64}$$

see Dieker (2004), Comte and Renault (1996), Baillie (1996).

9 Volatility clustering

To model volatility clustering in continuous time, we can proceed in two ways.

9.1 Stochastic volatility

To generate volatility clustering in discrete time we modified the stochastic volatility representation of the invariants (16) by imposing autocorrelation on the scatter as in (36).

Similarly, in continuous time first we represent the dynamics of the cumulative invariants, i.e. the generic Levy process (47) as follows:

$$X_t \equiv \mu t + \sigma Z_t. \tag{65}$$

In this expression the first term is the location of X_t, which must grow linearly; Z_t is a zero-location Levy process whose scatter at time $t \equiv 1$ is one; and σ is a positive constant. Then we modify the (trivial) Levy process for σ into a stationary process σ_t that displays autocorrelation: in this situation X_t is no longer a Levy process and volatility clustering follows.

The most popular such model is the *Heston model*,see Heston (1993): the normalized Levy process Z_t in (65) is a Brownian motion; the scatter parameter follows a CIR process(59) shocked by a different Brownian motion

$$d\sigma_t^2 = -\kappa\left(\sigma_t^2 - \overline{\sigma}^2\right)dt + \lambda\sqrt{\sigma_t^2}dB_t, \tag{66}$$

where

$$2\kappa\overline{\sigma}^2 > \lambda^2; \tag{67}$$

the copula between B_t and Z_t is normal at all times with constant correlation ρ; and $\rho < 0$ due to the *leverage effect*: lower returns tend to correspond to higher volatility. The Heston model is heavily used by $\mathbb{Q}$-finance quants to price options. Indeed, the original Black-Scholes-Merton model (40) is inconsistent with the observed smiles, skews and smirks in the implied volatility profile.

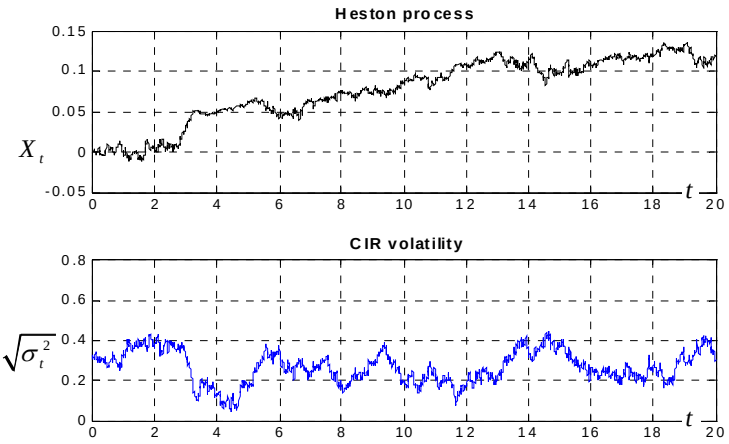

Figure 11: Heston stochastic volatility process

In Figure 11 we display a sample path from the Heston stochastic volatility process, along with a sample path of the volatility (66) that generates it, refer to MATLAB Central File Exchange under the author's name for the code

9.2 Subordination

An alternative way to model volatility clustering in continuous time is inspired by the self-similarity of the Brownian motion:

$$B_{\sigma^2 t} \stackrel{d}{=} \sigma B_t. \tag{68}$$

Consider the subordination (49) of a Levy processes:

$$X_t \stackrel{d}{=} B^{\mu,\sigma^2}_{T_t}. \tag{69}$$

If the subordinator, i.e. the stochastic time T_t increases faster than linearly, than the volatility increases: periods of large volatility correspond to periods where the pace of the market increases. In order to generate volatility clustering, i.e. autocorrelation in volatility, we simply relax the assumption that the pace of the market T_t is a Levy process.

Since T_t must be increasing, it can be defined as the integral of a process $\dot{T}_s$ which is positive at all times

$$T_t \equiv \int_0^t \dot{T}_s ds. \tag{70}$$

A natural choice for $\dot{T}_s$ is the CIR process (59) and a natural choice for the equilibrium value for $\dot{T}_s$ is one:

$$d\dot{T}_t = -\widetilde{\theta}\left(\dot{T}_t - 1\right) dt + \widetilde{\sigma}\sqrt{\dot{T}_t} dZ_t. \tag{71}$$

In order to further increase the flexibility of the subordination approach to stochastic volatility, one can subordinate a generic Levy process instead of the Brownian motion in (69).

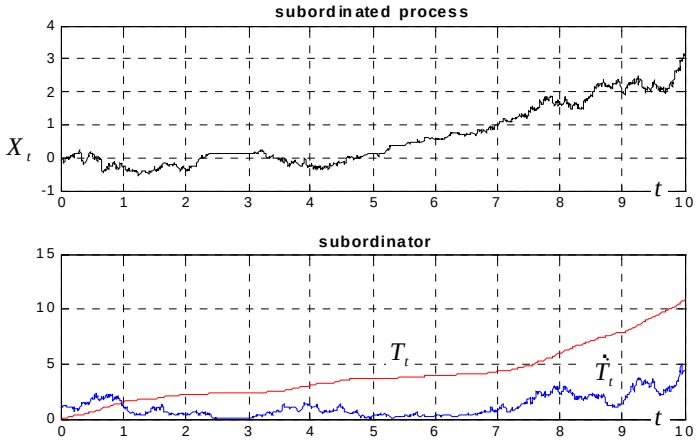

Figure 12: CIR-driven subordinated Brownian motion

In Figure 12 we display a sample path from the subordinated Brownian motion, along with a sample path of the subordinator (70) that generates it, as well as the CIR process that generates the subordinator, refer to MATLAB Central File Exchange under the author's name for the code. Notice that periods of higher volatility correspond to period where the stochastic time elapses faster

References

Baillie, R.T., 1996, Long memory processes and fractional integration in econometrics, *Journal of Econometrics* 73, 5–59.

Barndorff-Nielsen, O.E., 1998, Processes of normal inverse Gaussian type, *Finance and Stochastics* 2, 41–68.

Beran, J., 1994, *Statistics for Long-Memory Processes* (Chapman and Hall).

Carr, P., H. Geman, D. H. Madan, and M. Yor, 2003, Stochastic volatility for Levy processes, *Mathematical Finance* 13, 345–382.

Comte, F., and E. Renault, 1996, Long memory continuous time models, *Journal of Econometrics* 73, 101–149.

Cont, R., 2005, Volatility clustering in financial markets: Empirical facts and agent-based models, in A. Kirman, and G. Teyssiere, ed.: *Long Memory in Economics*. Springer.

———, and P. Tankov, 2008, *Financial Modelling with Jump Processes* (Chapman and Hall/CRC) 2nd edn.

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

Dieker, T., 2004, Simulation of fractional Brownian motion, *Master's Thesis*.

Ghoulmie, F., R. Cont, and J.P. Nadal, 2005, Heterogeneity and feedback in an agent-based market model, *Journal of Physics: Condensed Matter* 17, S1259–S1268.

Hamilton, J. D., 1994, *Time Series Analysis* (Princeton University Press).

Heston, S. L., 1993, Closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies* 6, 327–343.

Kou, S. G., 2002, A jump-diffusion model for option pricing, *Management Science* 48, 1086–1101.

Madan, D., and F. Milne, 1991, Option pricing with VG martingale components, *Mathematical Finance* 1, 39–56.

Mandelbrot, B., 1963, The variation of certain speculative prices, *Journal of Business* 36, 394–419.

Merton, R. C., 1976, Option pricing when underlying stocks are discontinuous, *Journal of Financial Economics* 3, 125–144.

Meucci, A., 2005, *Risk and Asset Allocation* (Springer).

Monroe, I., 1978, Processes that can be imbedded in Brownian motion, *The Annals of Probability* 6, 42–56.

Rachev, S. T., 2003, *Handbook of Heavy Tailed Distributions in Finance* (Elsevier/North-Holland).

Samorodnitsky, G., and M. S. Taqqu, 1994, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* (Chapman and Hall).

Schoutens, W., 2003, *Levy Processes in Finance* (Wiley.).

Vasicek, O., 1977, An equilibrium characterisation of the term structure, *Journal of Financial Economics* 5, 177–188.

Vicente, R., C. M. de Toledo, Leite V. B. P., and N. Caticha, 2005, Underlying dynamics of typical fluctuations of an emerging market price index: The Heston model from minutes to months, *arXiv:physics/0506101v1 [physics.soc-ph]*.

Index

A Appendix

In this appendix we present proofs, results and details that can be skipped at first reading.

A.1 The Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck process is defined by the dynamics

$$dX_t = -\theta\left(X_t - \mu\right)dt + \sigma dB_t. \tag{72}$$

We introduce the integrator

$$Y_t \equiv e^{\theta t}\left(X_t - \mu\right). \tag{73}$$

Then from Ito's lemma

$$\begin{aligned} dY_t &= \theta e^{\theta t}\left(X_t - \mu\right)dt + e^{\theta t}\left(-\theta\left(X_t - \mu\right)dt + \sigma dB_t\right) \\ &= \sigma e^{\theta t} dB_t \end{aligned} \tag{74}$$

Integrating we obtain

$$Y_t = Y_0 + \sigma\int_0^t e^{\theta s} dB_s \tag{75}$$

which from (73) becomes

$$X_t = \left(1 - e^{-\theta t}\right)\mu + e^{-\theta t}X_0 + \sigma\int_0^t e^{-\theta(t-s)} dB_s. \tag{76}$$

Therefore

$$X_t|x_0 \sim \mathrm{N}\left(\mu_t|x_0, \sigma_t^2|x_0\right), \tag{77}$$

where

$$\mu_t|x_0 \equiv \left(1 - e^{-\theta t}\right)\mu + e^{-\theta t}x_0. \tag{78}$$

To compute $\sigma_t^2|x_0$ we use Ito's isometry

$$\begin{aligned} \sigma_t^2|x_0 &= \sigma^2\int_0^t e^{-2\theta(t-s)} ds = \sigma^2\int_0^t e^{-2\theta u} du \\ &= \sigma^2\left.\frac{e^{-2\theta u}}{-2\theta}\right|_0^t = \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta t}\right). \end{aligned} \tag{79}$$

Similarly for the autocovariance

$$\begin{aligned}
\sigma^2_{t,\tau}|x_0 &\equiv \mathrm{Cov}\left\{X_t, X_{t+\tau}|x_0\right\} \\
&= \mathrm{E}\left\{\left(\sigma e^{-\theta t}\int_0^t e^{\theta s}dB_s\right)\left(\sigma e^{-\theta(t+\tau)}\int_0^{t+\tau} e^{\theta s}dB_s\right)\right\} \\
&= \sigma^2 e^{-2\theta t}e^{-\theta\tau}\int_0^t e^{2\theta s}ds \\
&= \sigma^2 e^{-2\theta t}e^{-\theta\tau}\left.\frac{e^{2\theta u}}{2\theta}\right|_0^t \\
&= \frac{\sigma^2}{2\theta}e^{-\theta\tau}\left(1-e^{-2\theta t}\right)
\end{aligned} \tag{80}$$

Therefore the autocorrelation reads

$$\begin{aligned}
\mathrm{Cor}\left\{X_t, X_{t+\tau}|x_0\right\} &= \frac{\sigma^2_{t,\tau}|x_0}{\sqrt{\sigma^2_t|x_0}\sqrt{\sigma^2_{t-\tau}|x_0}} \\
&= e^{-\theta\tau}\sqrt{\frac{1-e^{-2\theta t}}{1-e^{-2\theta(t+\tau)}}},
\end{aligned} \tag{81}$$

which in the limit $t\to\infty$ becomes the unconditional autocorrelation

$$\mathrm{Cor}\left\{X_t, X_{t+\tau}\right\} = e^{-\theta\tau}. \tag{82}$$

A.2 Relation between CIR and OU processes

Consider the OU process with null unconditional expectation

$$dX_t = -\theta X_t dt + \sigma dB_t. \tag{83}$$

Using Ito's rule

$$d\left(X_t^2\right) = 2X_t dX_t + \left(dX_t\right)^2 \tag{84}$$

we obtain

$$\begin{aligned}
dY_t &= 2\sqrt{Y_t}\left(-\theta\sqrt{Y_t}dt + \sigma dB_t\right) + \sigma^2 dt \\
&= \left(-2\theta Y_t + \sigma^2\right)dt + 2\sqrt{Y_t}\sigma dB_t,
\end{aligned} \tag{85}$$

which is a specific instance of the general CIR process (59).

More in general, following Vicente, de Toledo, P., and Caticha (2005), we consider several independent OU processes

$$dX_{j,t} = -\theta X_{j,t}dt + \sigma dB_{j,t}, \tag{86}$$

whose square evolves as follows:

$$\begin{aligned}
d\left(X_{j,t}\right)^2 &= 2X_{j,t}\left(-\theta X_{j,t}dt + \sigma dB_{j,t}\right) + \sigma^2 dt \\
&= -2\theta X^2_{j,t}dt + 2\sigma X_{j,t}dB_{j,t} + \sigma^2 dt.
\end{aligned} \tag{87}$$

Then $v_t \equiv \sum_{j=1}^{J} X_{j,t}^2$ evolves as a CIR process

$$dv_t = \left(J\sigma^2 - 2\theta v_t\right) dt + 2\sigma\sqrt{v_t} dB_t, \tag{88}$$

where in the last step we used the isometry

$$\sum_{j=1}^{J} X_{j,t} dB_{j,t} \overset{d}{=} \sqrt{\sum_{j=1}^{J} X_{j,t}^2} dB_t. \tag{89}$$

A.3 The Levy-Khintchine representation of a Levy process

First we compute the characteristic function of the Poisson process (41), which follows from (14):

$$\begin{aligned} \phi_{\epsilon_t}(\omega) &\equiv \mathrm{E}\left\{e^{i\omega\epsilon_t}\right\} = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-(\lambda t)} e^{i\omega k\Delta} \\ &= e^{-(\lambda t)} \sum_{k=0}^{\infty} \frac{\left(\lambda t e^{i\omega\Delta}\right)^k}{k!} = e^{-\lambda t} e^{\lambda t e^{i\omega\Delta}} \\ &= e^{\lambda t\left(e^{i\omega\Delta}-1\right)}. \end{aligned} \tag{90}$$

Now the generic representation of a Levy process (47), which we report here:

$$X_t = B_t^{\mu,\sigma^2} + \int_{-\infty}^{+\infty} P_t^{\Delta,\lambda(\Delta)} d\Delta. \tag{91}$$

Its characteristic function reads

$$\mathrm{E}\left\{e^{i\omega X_t}\right\} = \mathrm{E}\left\{e^{i\omega\left(\mu t + \sigma B_t + \int_{-\infty}^{+\infty} P_t^{\Delta,\lambda(\Delta)} d\Delta\right)}\right\} \tag{92}$$

Approximating the integral with sums we obtain

$$\begin{aligned} \ln\left(\mathrm{E}\left\{e^{i\omega X_t}\right\}\right) &= i\omega\mu t - \frac{1}{2}\sigma^2 t\omega^2 + \ln\left(\mathrm{E}\left\{e^{i\omega\left(\sum_k P_t^{\Delta_k,\lambda_k}\right)}\right\}\right) \\ &= i\omega\mu t - \frac{1}{2}\sigma^2 t\omega^2 + \ln\prod_k \mathrm{E}\left\{e^{i\omega P_t^{\Delta_k,\lambda_k}}\right\} \\ &= i\omega\mu t - \frac{1}{2}\sigma^2 t\omega^2 + \sum_k \ln e^{\lambda_k t\left(e^{i\omega\Delta_k}-1\right)} \\ &= i\omega\mu t - \frac{1}{2}\sigma^2 t\omega^2 + t\sum_k \lambda_k\left(e^{i\omega\Delta_k}-1\right) \\ &= i\omega\mu t - \frac{1}{2}\sigma^2 t\omega^2 + t\int_{-\infty}^{+\infty} \lambda(\Delta)\left(e^{i\omega\Delta}-1\right) d\Delta \end{aligned} \tag{93}$$

Reverting from the sum to the integral notation we obtain

$$\ln\left(\mathrm{E}\left\{e^{i\omega X_t}\right\}\right) = \left(i\omega\mu - \frac{1}{2}\sigma^2\omega^2 + \int_{-\infty}^{+\infty} \lambda\left(\Delta\right)\left(e^{i\omega\Delta} - 1\right) d\Delta\right) t, \tag{94}$$

where the intensity $\lambda\left(\Delta\right)$ determines the relative weight of the discrete Poisson jumps on the grid $\Delta\mathbb{N}$.

A.4 Representation of compound Poisson process

Consider the general representation (50) of a Poisson process. We do not need to worry about the diffusive portion, so we only consider the jump portion

$$X_t = \sum_{n=1}^{P_t^\lambda} Z_n. \tag{95}$$

The characteristic function reads

$$\begin{aligned} \mathrm{E}\left\{e^{i\omega X_t}\right\} &= \mathrm{E}_P\left\{\mathrm{E}_Z\left\{e^{\sum_{n=1}^{P_t^\lambda} i\omega Z_n}\right\}\right\} \\ &= \mathrm{E}_P\left\{\phi_Z\left(\omega\right)^{P_t^\lambda}\right\} = \mathrm{E}_P\left\{e^{i\gamma P_t^\lambda}\right\}\Big|_{\phi_Z(\omega)=e^{i\gamma}} \\ &= e^{t\lambda(\phi_Z(\omega)-1)}, \end{aligned} \tag{96}$$

where in the last step we used (90).

Suppose now that in (94) we set $\mu \equiv \sigma^2 \equiv 0$ and $\lambda\left(\Delta\right) \equiv \lambda f_Z\left(\Delta\right)$. Then (94) becomes

$$\begin{aligned} \ln\left(\mathrm{E}\left\{e^{i\omega X_t}\right\}\right) &= t\lambda\left(\int_{-\infty}^{+\infty} f_Z\left(\Delta\right) e^{i\omega\Delta} d\Delta - \int_{-\infty}^{+\infty} f_Z\left(\Delta\right) d\Delta\right) \\ &= t\lambda\left(\phi_Z\left(\Delta\right) - 1\right), \end{aligned} \tag{97}$$

which is (96).