Signal Processing with Lévy Information

Lane P. Hughston

Department of Mathematics
Brunel University London
Uxbridge UB8 3PH, United Kingdom

14 May 2015

Barrett Memorial Lectures
University of Tennessee, Knoxville

Collaborators: D. C. Brody (Brunel University London),
E. Mackie (J. P. Morgan, London), and X. Yang (Shell International).
This presentation is based on the following papers:

   “General theory of geometric Lévy models for dynamic asset pricing”

2. Brody, D.C., Hughston, L.P. & Yang, X.
   “Signal processing with Lévy information”

   “Lévy information and the aggregation of risk aversion”
Theory of Noise Type


A real-valued process \( \{\xi_t\}_{t \geq 0} \) on a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) is a Lévy process if: (i) \( \mathbb{P}(\xi_0 = 0) = 1 \), (ii) \( \{\xi_t\} \) has stationary and independent increments, (iii) \( \lim_{t \to s} \mathbb{P}(|\xi_t - \xi_s| > \epsilon) = 0 \), and (iv) \( \{\xi_t\} \) is almost surely càdlàg.

For a Lévy process \( \{\xi_t\} \) to give rise to a class of information processes, we require that it should possess exponential moments.

Let us consider the set defined for some (equivalently for all) \( t > 0 \) by

\[
A = \{ w \in \mathbb{R} : \mathbb{E}^\mathbb{P}[\exp(w \xi_t)] < \infty \}. 
\]

If \( A \) contains points other than \( w = 0 \), then we say that \( \{\xi_t\} \) possesses exponential moments.

We define a function \( \psi : A \to \mathbb{R} \) called the Lévy exponent (or cumulant function), such that for \( \alpha \in A \) we have

\[
\mathbb{E}^\mathbb{P}[\exp(\alpha \xi_t)] = \exp(\psi(\alpha) t). 
\]
If a Lévy process possesses exponential moments, then an exercise shows that:

(i) $\psi(\alpha)$ is convex on $A$,

(ii) the mean and variance of $\xi_t$ are given respectively by $\psi'(0) t$ and $\psi''(0) t$, and

(iii) as a consequence of the convexity of $\psi(\alpha)$, the marginal exponent $\psi'(\alpha)$ possesses a unique inverse $I(y)$ such that $I(\psi'(\alpha)) = \alpha$ for $\alpha \in A$.

The Lévy exponent extends to a function $\psi : A_C \to \mathbb{C}$ where $A_C = \{ w \in \mathbb{C} : \text{Re } w \in A \}$.

It can be shown (Sato 1999, Theorem 25.17) that $\psi(\alpha)$ admits a Lévy-Khintchine representation of the form

$$\psi(\alpha) = p\alpha + \frac{1}{2}q\alpha^2 + \int_{\mathbb{R}\setminus\{0\}} (e^{\alpha z} - 1 - \alpha z 1\{|z| < 1\}) \nu(dz) \tag{3}$$

with the property that (2) holds for all $\alpha \in A_C$.

Here $1\{|\cdot|\}$ is the indicator function, $p \in \mathbb{R}$ and $q \geq 0$ are constants. The so-called Lévy measure $\nu(dz)$ is a positive measure defined on $\mathbb{R}\setminus\{0\}$ satisfying

$$\int_{\mathbb{R}\setminus\{0\}} (1 \wedge z^2) \nu(dz) < \infty. \tag{4}$$
If the Lévy process possesses exponential moments then for $\alpha \in A$ we have

$$\int_{\mathbb{R}\setminus\{0\}} e^{\alpha z} \mathbf{1}\{|z| \geq 1\} \nu(dz) < \infty.$$  \hspace{1cm} (5)

The Lévy measure has the following interpretation: if $B$ is a measurable subset of $\mathbb{R}\setminus\{0\}$, then $\nu(B)$ is the rate at which jumps arrive for which the jump size lies in $B$.

If $\nu(\mathbb{R}\setminus\{0\}) < \infty$ one says that $\{\xi_t\}$ has finite activity.

Thus we can classify a Lévy process abstractly by the specification of its exponent $\psi(\alpha)$.

This means one can speak of a “type” of Lévy noise by reference to the associated exponent.

In this way we obtain a kind of general classification of the different types of “pure noise”.
Families of Noise Types

Suppose we fix a measure $\mathbb{P}_0$ on a measurable space $(\Omega, \mathcal{F})$, and let the process $\{\xi_t\}$ be $\mathbb{P}_0$-Lévy, with exponent $\psi_0(\alpha)$.

There exists a parametric family of probability measures $\{\mathbb{P}_\lambda\}_{\lambda \in A}$ on $(\Omega, \mathcal{F})$ such that for each choice of $\lambda$ the process $\{\xi_t\}$ is $\mathbb{P}_\lambda$-Lévy.


Under an Esscher transformation the characteristics of a Lévy process are transformed from one type to another, and thus one can speak of a “family” of Lévy processes interrelated by Esscher transformations.

The relevant change of measure can be specified by use of the process $\{\rho_t^\lambda\}$ defined for $\lambda \in A$ by

$$\rho_t^\lambda := \left. \frac{d\mathbb{P}_\lambda}{d\mathbb{P}_0} \right|_{\mathcal{F}_t} = \exp (\lambda \xi_t - \psi_0(\lambda)t), \quad (6)$$

where $\mathcal{F}_t = \sigma[\{\xi_s\}_{0 \leq s \leq t}]$. 


One can check that \( \{\rho_t^\lambda\} \) is an \( (\{\mathcal{F}_t\}, \mathbb{P}_0) \)-martingale: we find that
\[
\mathbb{E}^{\mathbb{P}_0}_s[\rho_t^\lambda] = \rho_s^\lambda
\]
for \( s \leq t \), where \( \mathbb{E}^{\mathbb{P}_0}_t[\cdot] \) denotes conditional expectation under \( \mathbb{P}_0 \) with respect to \( \mathcal{F}_t \).

One can show that \( \{\xi_t\} \) has \( \mathbb{P}_\lambda \) stationary and independent increments, and that the \( \mathbb{P}_\lambda \) exponent of \( \{\xi_t\} \), defined on the set
\[
A^\lambda_C := \{w \in \mathbb{C} \mid \text{Re } w + \lambda \in A\},
\]
is given by
\[
\psi_\lambda(\alpha) := t^{-1} \ln \mathbb{E}^{\mathbb{P}_\lambda}[\exp(\alpha \xi_t)] = \psi_0(\alpha + \lambda) - \psi_0(\lambda).
\]

In what follows we write \( \mathbb{C}^I = \{w \in \mathbb{C} : \text{Re } w = 0\} \). For any random variable \( Z \) on \( (\Omega, \mathcal{F}, \mathbb{P}) \) we write \( \mathcal{F}^Z = \sigma[Z] \), and we write \( \mathbb{E}^{\mathbb{P}}[\cdot | Z] \) for \( \mathbb{E}^{\mathbb{P}}[\cdot | \mathcal{F}^Z] \).

For processes we use both of the notations \( \{Z_t\} \) and \( \{Z(t)\} \), depending on the context.
Lévy information

With these background remarks in mind, we are in a position to define a Lévy information process.

We confine the discussion to the case of a “simple” message, represented by a single random variable $X$.

In the situation when the noise is Brownian motion, the information admits a linear decomposition into signal and noise.

In the general situation the relation between signal and noise is more subtle, and has something of the character of a fibre space, where one thinks of the points of the base space as representing the different noise types, and the points of the fibres as corresponding to the different information processes that one can construct in association with a given noise type.

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and an Esscher family Lévy exponents $\psi_{\lambda}(\alpha), \alpha \in A_{\lambda}^\lambda$.

We refer to $\psi_0(\alpha)$ as the fiducial exponent.
The intuition here is that the abstract Lévy process of exponent $\psi_0(\alpha)$, which we call the “fiducial” process, represents the noise type of the associated information process.

Thus we can use $\psi_0(\alpha)$, to represent the noise type.

**Definition 1.** By a Lévy information process carrying a message $X$, we mean a random process $\{\xi_t\}$, together with a random variable $X$, such that $\{\xi_t\}$ is conditionally $\psi_X$-Lévy given $\mathcal{F}^X$.

Thus, given $\mathcal{F}^X$ we require $\{\xi_t\}$ to have conditionally independent and stationary increments under $\mathbb{P}$, and to possess a conditional exponent of the form

$$\psi_X(\alpha) := t^{-1} \ln \mathbb{E}^\mathbb{P}[\exp(\alpha \xi_t) \mid \mathcal{F}^X] = \psi_0(\alpha + X) - \psi_0(X)$$

for $\alpha \in \mathbb{C}^I$, where $\psi_0(\alpha)$ is the fiducial exponent of the specified noise type.

It is implicit in the statement of Definition 1 that a certain compatibility condition holds between the message and the noise type.

For any random variable $X$ we define its support $S_X$ to be the smallest closed set $F$ with the property that $\mathbb{P}(X \in F) = 1$.

Then we say that $X$ is compatible with the fiducial exponent $\psi_0(\alpha)$ if $S_X \subset A$. 
Intuitively speaking, this condition ensures that we can use $X$ to make a random Esscher transformation.

We are thus able to state the Lévy noise-filtering problem as follows: given observations of the Lévy information process up to time $t$, what is the best estimate for $X$?

To gain a better understanding of the sense in which the information process $\{\xi_t\}$ actually “carries” the message $X$, it will be useful to investigate its asymptotic behaviour.

We write $I_0(y)$ for the inverse marginal fiducial exponent.

**Proposition 1.** Let $\{\xi_t\}$ be a Lévy information process with fiducial exponent $\psi_0(\alpha)$ and message $X$. Then for every $\epsilon > 0$ we have

$$\lim_{t \to \infty} \mathbb{P}[|I_0(t^{-1}\xi_t) - X| \geq \epsilon] = 0.$$  \hspace{1cm} (11)

It follows that the information process does indeed carry information about the message, and in the long run “reveals” it.

The intuition here is that as more information is gained we improve our estimate of $X$ to the point that $X$ eventually becomes known with near certainty.
Properties of Lévy information

It will be useful if we present a construction that ensures the existence of Lévy information processes.

First we select a noise type by specification of a fiducial exponent.

Next we introduce a probability space \((\Omega, \mathcal{F}, \mathbb{P}_0)\) that supports the existence of a \(\mathbb{P}_0\)-Lévy process \(\{\xi_t\}\) with the given fiducial exponent, together with an independent random variable \(X\) that is compatible with it.

Write \(\{\mathcal{F}_t\}\) for the filtration generated by \(\{\xi_t\}\), and \(\{\mathcal{G}_t\}\) for the filtration generated by \(\{\xi_t\}\) and \(X\) jointly: \(\mathcal{G}_t = \sigma[\{\xi_t\}_{0 \leq s \leq t}, X]\).

Let \(\psi_0(\alpha)\) be the fiducial exponent.

One can check that the process \(\{\rho^X_t\}\) defined by

\[
\rho^X_t = \exp \left( X \xi_t - \psi_0(X) t \right)
\]

is a \(\{\mathcal{G}_t\}, \mathbb{P}_0\)-martingale.
We are thus able to introduce a change of measure \( \mathbb{P}_0 \to \mathbb{P}_X \) on \((\Omega, \mathcal{F}, \mathbb{P}_0)\) by setting
\[
\frac{d\mathbb{P}_X}{d\mathbb{P}_0} \bigg|_{\mathcal{G}_t} = \rho^X_t. \tag{13}
\]

It should be evident that \( \{\xi_t\} \) is conditionally \( \mathbb{P}_X \)-Lévy given \( \mathcal{F}^X \), since for fixed \( X \) the measure change is an Esscher transformation.

In particular, a calculation shows that the conditional exponent of \( \xi_t \) under \( \mathbb{P}_X \) is given by
\[
t^{-1} \ln \mathbb{E}^{\mathbb{P}_X} \left[ \exp(\alpha \xi_t) \mid \mathcal{F}^X \right] = \psi_0(\alpha + X) - \psi_0(X) \tag{14}
\]
for \( \alpha \in \mathbb{C}^I \), which shows that the conditions of Definition 1 are satisfied, allowing us to conclude the following:

**Proposition 2.** The \( \mathbb{P}_0 \)-Lévy process \( \{\xi_t\} \) is a Lévy information process under \( \mathbb{P}_X \) with noise type \( \psi_0(\alpha) \) and message \( X \).

In fact, the converse also holds: if we are given a Lévy information process, then by a change of measure we obtain a Lévy process and an independent “message” variable.
Going forward, we adopt the convention that $\mathbb{P}$ always denotes the “physical” measure in relation to which an information process with message $X$ is defined, and that $\mathbb{P}_0$ denotes the transformed measure with respect to which the information process and the message decouple.

The idea is to work out properties of information processes by referring the calculations back to $\mathbb{P}_0$.

We consider as an example the problem of working out the $\mathcal{F}_t$-conditional expectation under $\mathbb{P}$ of a $\mathcal{G}_t$-measureable integrable random variable $Z$.

The $\mathbb{P}$ expectation can be written in terms of $\mathbb{P}_0$ expectations, and is given by a “generalised Bayes formula” (Kallianpur & Striebel 1968) of the form

$$E^{\mathbb{P}}[Z | \mathcal{F}_t] = \frac{E^{\mathbb{P}_0}[\rho^X_t Z | \mathcal{F}_t]}{E^{\mathbb{P}_0}[\rho^X_t | \mathcal{F}_t]}.$$  \hfill (15)

This formula can be used to obtain the $\mathcal{F}_t$-conditional distribution of $X$.

We have the following.
Proposition 3. Let \( \{\xi_t\} \) be a Lévy information process under \( \mathbb{P} \) with noise type \( \psi_0(\alpha) \), and let the a priori distribution of the associated message \( X \) be \( \pi(dx) \). Then the \( \mathcal{F}_t \)-conditional a posteriori distribution of \( X \) is

\[
\pi_t(dx) = \frac{\exp(x\xi_t - \psi_0(x)t)}{\int \exp(x\xi_t - \psi_0(x)t) \pi(dx)} \pi(dx).
\] (16)

In particular, when \( X \) is a continuous random variable with a density function \( p(x) \) one can write \( \pi_t(dx) = p_t(x)dx \), where \( p_t(x) \) is the conditional density.

It is straightforward to establish by use of a variational argument that the best estimate for the message \( X \) conditional on the information \( \mathcal{F}_t \) is given by

\[
\hat{X}_t := \mathbb{E}_t^{\mathbb{P}}[X | \mathcal{F}_t] = \int x \pi_t(dx).
\] (17)

By “best estimate” we mean the \( \mathcal{F}_t \)-measurable random variable \( \hat{X}_t \) that minimises the quadratic error \( \mathbb{E}_t^{\mathbb{P}}[(X - \hat{X}_t)^2 | \mathcal{F}_t] \).
It will be observed that at any given time $t$ the best estimate can be expressed as a function of $\xi_t$ and $t$, and does not involve values of the information process at times earlier than $t$.

That this should be the case can be seen as a consequence of the following:

**Proposition 4.** The Lévy information process $\{\xi_t\}$ has the Markov property under $\mathbb{P}$.

It should be apparent that simulation of the dynamics of the filter is readily approachable on account of this property.
Examples of Lévy information processes

In a number of situations one can construct explicit examples of information processes, categorised by noise type.

The Brownian and Poisson constructions, which are familiar in other contexts, can be seen as belonging to a unified scheme that brings out their differences and similarities.

We then proceed to construct information processes of the gamma, the variance gamma, and the negative binomomial type.

It is interesting to take note of the diverse nature of noise, and to observe the many different ways in which messages can be conveyed in a noisy environment.

Example 1: Brownian information. On a probability space \((\Omega, \mathcal{F}, \mathbb{P})\), let \(\{B_t\}\) be a Brownian motion, let \(X\) be an independent random variable, and set

\[
\xi_t = X t + B_t.
\]  

(18)

The random process \(\{\xi_t\}\), which we call the Brownian information process, is \(\mathcal{F}^X\)-conditionally Lévy, with conditional exponent \(\psi_X(\alpha) = X \alpha + \frac{1}{2} \alpha^2\).
The fiducial exponent is $\psi_0(\alpha) = \frac{1}{2} \alpha^2$, and the associated fiducial process or “noise type” is standard Brownian motion.

In the case of Brownian information, there is a linear separation of the process into signal and noise.

This model, considered by Wonham (1965), is perhaps the simplest continuous-time generalisation of the example originally described by Wiener (1948).

The message is given by the value of $X$, but $X$ can only be observed indirectly, through $\{\xi_t\}$.

The observations of $X$ are obscured by the noise represented by the Brownian motion $\{B_t\}$.

Since the signal term grows linearly in time, whereas $|B_t| \sim \sqrt{t}$, it is intuitively plausible that observations of $\{\xi_t\}$ will asymptotically reveal the value of $X$, and a direct calculation using properties of the normal distribution function confirms that $t^{-1}\xi_t$ converges $X$, which is consistent with the fact that $\psi'_0(\alpha) = \alpha$ and $I_0(y) = y$ in the standard Brownian case.
We conclude our discussion of Brownian information with the following remark. In problems involving prediction and valuation, it is not uncommon that the message is revealed after the passage of a finite amount of time.

This is often the case in applications to finance, where the message takes the form of a random cash flow at some future date, or, more generally, a random factor that affects such a cash flow.

There are numerous examples coming from the physical sciences, economics and operations research where the goal of an agent is to form a view concerning the outcome of a future event by monitoring the flow of information relating to it.

One way of modelling such situations in the present context is by use of a time change.

If \( \{ \xi_t \} \) is a Lévy information process with message \( X \) and a specified fiducial exponent, then a generalisation of Proposition 1 shows that the process \( \{ \xi_{tT} \} \) defined over the time interval \( 0 \leq t < T \) by

\[
\xi_{tT} = \frac{T - t}{T} \xi \left( \frac{tT}{T - t} \right)
\]  

reveals the value of \( X \) in the limit as \( t \to T \).
One can check for $0 \leq s \leq t < T$ that

$$\text{Cov} \left[ \xi_{sT}, \xi_{tT} \mid \mathcal{F}^X \right] = \frac{s(T - t)}{T} \psi_0''(X).$$

(20)

In the case where $\{\xi_t\}$ is a Brownian information process represented as above in the form $\xi_t = X_t + B_t$, the time-changed process takes the form

$$\xi_{tT} = X_t + \beta_{tT},$$

(21)

where $\{\beta_{tT}\}$ is a Brownian bridge over the interval $[0, T]$.


It seems reasonable to conjecture that time-changed Lévy information processes of the more general type proposed above may be similarly applicable.

**Example 2: Poisson information.** Consider a situation in which an agent observes a series of events taking place at a random rate, and the agent wishes to determine this unknown rate as best as possible since its value conveys an important piece of information.
One can model the information flow in this example by a modulated Poisson process for which the jump rate is itself an independent random variable.

Such a scenario arises in many real-world situations, and has been investigated in the literature (Segall & Kailath 1975, Segall et al. 1975, Brémaud 1981, Di Masi & Runggaldier 1983, Kailath & Poor 1998).

The Segall-Kailath scheme can be seen to emerge rather naturally as an example of our general model for Lévy information.

As in the Brownian case, one can construct the relevant information process directly. The setup is as follows.

On a probability space \((\Omega, \mathcal{F}, P)\), let \(\{N(t)\}_{t \geq 0}\) be a standard Poisson process with jump rate \(m > 0\), let \(X\) be an independent random variable, and set

\[
\xi_t = N(e^X t).
\]  

Thus \(\{\xi_t\}\) is a time-changed Poisson process, and the effect of the signal is to randomly modulate the rate at which the process jumps.
It is evident that \( \{\xi_t\} \) is \( \mathcal{F}^X \)-conditionally Lévy and satisfies the conditions of Definition 1.

In particular, we have

\[
\mathbb{E} \left[ \exp(\alpha N(e^X t)) \mid \mathcal{F}^X \right] = \exp(me^X(e^\alpha - 1) t), \tag{23}
\]

and for fixed \( X \) one obtains a Poisson process with rate \( me^X \).

It follows that (22) is an information process.

The fiducial exponent is \( \psi_0(\alpha) = m(e^\alpha - 1) \).

A calculation using (10) shows that \( \psi_X(\alpha) = me^X(e^\alpha - 1) \).

The relation between signal and noise in the case of Poisson information is rather subtle.

The noise is associated with the random fluctuations of the inter-arrival times of the jumps, whereas the message determines the average rate at which the jumps occur.
In the case of Poisson information the conditional distribution of $X$ is

$$\pi_t(dx) = \frac{\exp(x\xi_t - m(e^x - 1)t)}{\int \exp(x\xi_t - m(e^x - 1)t) \pi(dx)} \pi(dx).$$

(24)

Example 3: Gamma information.

Let $m$ and $\kappa$ be positive numbers. By a gamma process with rate $m$ and scale $\kappa$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we mean a Lévy process $\{\gamma_t\}_{t \geq 0}$ with exponent

$$t^{-1} \ln \mathbb{E}\mathbb{P} [\exp(\alpha \gamma_t)] = -m \ln (1 - \kappa \alpha)$$

for $\alpha \in A_C = \{w \in \mathbb{C} \mid \text{Re } w < \kappa^{-1}\}$.

The probability density for $\gamma_t$ is

$$\mathbb{P}(\gamma_t \in dx) = \mathbb{1}\{x > 0\} \kappa^{-mt} x^{mt-1} \exp(-x/\kappa) \frac{d\chi}{\Gamma[mt]},$$

(26)

where $\Gamma[a]$ is the gamma function.

A short calculation making use of the functional equation $\Gamma[a + 1] = a\Gamma[a]$ shows that $\mathbb{E}\mathbb{P}[\gamma_t] = m\kappa t$ and $\text{Var}\mathbb{P}[\gamma_t] = m\kappa^2 t$.

Clearly, the mean and variance determine the rate and scale.
The Lévy measure in this example is given by

\[ \nu(dz) = 1\{z > 0\} m z^{-1} \exp(-\kappa z) \, dz. \] (27)

One can check that \( \nu(\mathbb{R}\setminus\{0\}) = \infty \) and thus that the gamma process has infinite activity.

If \( \kappa = 1 \) we say that \( \{\gamma_t\} \) is a *standard* gamma process with rate \( m \), and in that case one finds that \( \{\kappa \gamma_t\} \) is a scaled gamma process with rate \( m \) and scale \( \kappa \).

Now let \( \{\xi_t\} \) be a standard gamma process with rate \( m \) on a probability space \((\Omega, \mathcal{F}, \mathbb{P}_0)\), and let \( \lambda \in \mathbb{R} \) satisfy \( \lambda < 1 \).

Then the process \( \{\rho_t^\lambda\} \) defined by

\[ \rho_t^\lambda = (1 - \lambda)^m t e^{\lambda \gamma_t} \] (28)

is an \((\{\mathcal{F}_t\}, \mathbb{P}_0)\)-martingale.

If we let \( \{\rho_t^\lambda\} \) act as a change of measure density for the transformation \( \mathbb{P}_0 \to \mathbb{P}_\lambda \), then we find that \( \{\gamma_t\} \) is a *scaled* gamma process under \( \mathbb{P}_\lambda \), with rate \( m \) and scale \( 1/(1 - \lambda) \). Thus the effect of an Esscher transformation on a gamma process is to alter its scale.

With these facts in mind, one can establish the following result.
Proposition 5. Let \( \{\gamma_t\} \) be a standard gamma process with rate \( m \) on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\), and let the independent random variable \( X \) satisfy \( X < 1 \) almost surely. Then the process \( \{\xi_t\} \) defined by

\[
\xi_t = \frac{1}{1 - X} \gamma_t
\]  

(29)

is a Lévy information process with message \( X \) and gamma noise, with fiducial exponent \( \psi_0(\alpha) = -m \ln(1 - \alpha) \) for \( \alpha \in \{w \in \mathbb{C} | \Re w < 1\} \).

Proof. It is evident that \( \{\xi_t\} \) is \( \mathcal{F}^X \)-conditionally a scaled gamma process. As a consequence of (25) we have

\[
t^{-1} \ln \mathbb{E}^\mathbb{P} \left[ \exp(\alpha \xi_t) | X \right] = t^{-1} \ln \mathbb{E}^\mathbb{P} \left[ \exp \left( \frac{\alpha}{1 - X} \gamma_t \right) | X \right] = -m \ln \left( 1 - \frac{\alpha}{1 - X} \right)
\]

for \( \alpha \in \mathbb{C}^I \). Then we note that

\[
-m \ln \left( 1 - \frac{\alpha}{1 - X} \right) = -m \ln (1 - (X + \alpha)) + m \ln (1 - X),
\]

(30)

from which it follows that the \( \mathcal{F}^X \)-conditional \( \mathbb{P} \) exponent of \( \{\xi_t\} \) is \( \psi_0(X + \alpha) - \psi_0(X) \).

The gamma filter arises as follows.
An agent observes a process of accumulation: typically there are many small increments, but now and then there are large increments.

The unknown rate at which the process is growing on average is an important figure that the agent wishes to determine as accurately as possible.

The accumulation process can be modelled by gamma information, and the associated filter can be utilised to estimate the growth rate.

It has long been recognised that the gamma process is useful in characterising phenomena such as the water level of a dam or the totality of the claims made in a large portfolio of insurance contracts (Gani 1957, Kendall 1957, Gani & Pyke 1960). Use of the gamma information process and related bridge processes, with applications in finance and insurance, is pursued in Brody *et al.* (2008b), Hoyle (2010), and Hoyle *et al.* (2011).

For the conditional distribution we deduce that

$$
\pi_t(dx) = \frac{(1 - x)^{mt} \exp(x\xi_t)}{\int_{-\infty}^{1} (1 - x)^{mt} \exp(x\xi_t) \pi(dx)} \pi(dx), \tag{31}
$$

and this gives us the optimal filter for the case of gamma information.
**Example 4: Variance-gamma information.** The so-called variance-gamma or VG process (Madan & Seneta 1990, Madan & Milne 1991, Madan *et al.* 1998) was introduced in the theory of finance and has been extensively investigated.

The relevant definitions and conventions are as follows.

By a VG process with drift $\mu \in \mathbb{R}$, volatility $\sigma \geq 0$, and rate $m > 0$, we mean a Lévy process with exponent

$$
\psi(\alpha) = -m \ln \left(1 - \frac{\mu}{m} \alpha - \frac{\sigma^2}{2m} \alpha^2\right). \tag{32}
$$

The VG process admits representations in terms of simpler Lévy processes.

Let $\{\gamma_t\}$ be a standard gamma process on $(\Omega, \mathcal{F}, \mathbb{P})$, with rate $m$, and let $\{B_t\}$ be a standard Brownian motion, independent of $\{\gamma_t\}$.

We call the scaled process $\{\Gamma_t\}$ defined by $\Gamma_t = m^{-1} \gamma_t$ a gamma subordinator with rate $m$. Note that $\Gamma_t$ has dimensions of time and that $\mathbb{E}^\mathbb{P}[\Gamma_t] = t$.

A calculation shows that the Lévy process $\{V_t\}$ defined by

$$
V_t = \mu \Gamma_t + \sigma B_{\Gamma_t} \tag{33}
$$

has the exponent (32).
If $\mu = 0$ and $\sigma = 1$, we say that $\{V_t\}$ is a “standard” VG process, with rate parameter $m$.

If $\mu \neq 0$, we say that $\{V_t\}$ is a “drifted” VG process.

One can always choose units of time such that $m = 1$, but for applications it is better to choose conventional units of time (e.g., seconds for physical applications, years for economic applications), and treat $m$ as a model parameter.

In the limiting case $\sigma \to 0$ we obtain a gamma process with rate parameter $m$ and scale parameter $\mu/m$.

In the limiting case $m \to \infty$ we obtain a Brownian motion with drift $\mu$ and volatility $\sigma$.

An important alternative representation of the VG process results if we let $\{\gamma^1_t\}$ and $\{\gamma^2_t\}$ be a pair of independent standard gamma processes on $(\Omega, \mathcal{F}, \mathbb{P})$, each with rate $m$, and set

$$V_t = \kappa_1 \gamma^1_t - \kappa_2 \gamma^2_t,$$

where $\kappa_1$ and $\kappa_2$ are nonnegative constants.
A calculation shows that the exponent is of the form (32).

In particular, we have

$$\psi(\alpha) = -m \ln \left( 1 - (\kappa_1 - \kappa_2) \alpha - \kappa_1 \kappa_2 \alpha^2 \right),$$

(35)

where $\mu = m(\kappa_1 - \kappa_2)$ and $\sigma^2 = 2m\kappa_1 \kappa_2$, or equivalently

$$\kappa_1 = \frac{1}{2m} \left( \mu + \sqrt{\mu^2 + 2m\sigma^2} \right) \quad \text{and} \quad \kappa_2 = \frac{1}{2m} \left( -\mu + \sqrt{\mu^2 + 2m\sigma^2} \right),$$

(36)

where $\alpha \in \{ w \in \mathbb{C} : -1/\kappa_2 < \Re w < 1/\kappa_1 \}$.

Now let $\{\xi_t\}$ be a standard VG process on a probability space $(\Omega, \mathcal{F}, \mathbb{P}_0)$, with exponent $\psi_0(\alpha) = -m \ln(1 - (2m)^{-1} \alpha^2)$ for $\alpha \in \{ w \in \mathbb{C} : |\Re w| < \sqrt{2m} \}$.

Under the transformed measure $\mathbb{P}_\lambda$ defined by the change-of-measure martingale (6), one finds that $\{\xi_t\}$ is a drifted VG process, with

$$\mu = \lambda \left( 1 - \frac{1}{2m} \lambda^2 \right)^{-1} \quad \text{and} \quad \sigma = \left( 1 - \frac{1}{2m} \lambda^2 \right)^{-\frac{1}{2}},$$

(37)

for $|\lambda| < \sqrt{2m}$.

Thus in the case of the VG process an Esscher transformation affects both the drift and the volatility.
Note that for large $m$ the effect on the volatility is insignificant, whereas the effect on the drift reduces to that of an ordinary Girsanov transformation.

With these facts in hand, we are now in a position to construct the VG information process.

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a number $m > 0$.

**Proposition 6.** Let $\{\Gamma_t\}$ be a standard gamma subordinator with rate $m$, let $\{B_t\}$ be an independent standard Brownian motion, and let the independent random variable $X$ satisfy $|X| < \sqrt{2m}$ almost surely. Then the process $\{\xi_t\}$ defined by

$$\xi_t = X \left(1 - \frac{1}{2m} X^2\right)^{-1} \Gamma_t + \left(1 - \frac{1}{2m} X^2\right)^{-\frac{1}{2}} B(\Gamma_t)$$

is a Lévy information process with message $X$ and VG noise, with fiducial exponent

$$\psi_0(\alpha) = -m \ln \left(1 - \frac{1}{2m} \alpha^2\right)$$

for $\alpha \in \{w \in \mathbb{C} : \text{Re} w < \sqrt{2m}\}$. 
An alternative representation for the VG information process can be established if one randomly rescales the gamma subordinator appearing in the time-changed Brownian motion.

The result is as follows.

**Proposition 7.** Let \( \{ \Gamma_t \} \) be a gamma subordinator with rate \( m \), let \( \{ B_t \} \) be an independent standard Brownian motion, and let the independent random variable \( X \) satisfy \( |X| < \sqrt{2m} \) almost surely. Write \( \{ \Gamma_t^X \} \) for the subordinator defined by

\[
\Gamma_t^X = \left( 1 - \frac{1}{2m} X^2 \right)^{-1} \Gamma_t .
\]  

Then the process \( \{ \xi_t \} \) defined by \( \xi_t = X \Gamma_t^X + B(\Gamma_t^X) \) is a VG information process with message \( X \).

A further representation of the VG information process arises as a consequence of the representation of the VG process as the asymmetric difference between two independent standard gamma processes.

In particular, we have the following:
**Proposition 8.** Let \( \{\gamma^1_t\} \) and \( \{\gamma^2_t\} \) be independent standard gamma processes, each with rate \( m \), and let the independent random variable \( X \) satisfy \( |X| < \sqrt{2m} \) almost surely. Then the process \( \{\xi_t\} \) defined by

\[
\xi_t = \frac{1}{\sqrt{2m} - X} \gamma^1_t - \frac{1}{\sqrt{2m} + X} \gamma^2_t
\]

is a VG information process with message \( X \).

**Example 5: Negative-binomial information.** By a negative binomial process with rate parameter \( m \) and probability parameter \( q \), where \( m > 0 \) and \( 0 < q < 1 \), we mean a Lévy process with exponent

\[
\psi_0(\alpha) = m \ln \left( \frac{1 - q}{1 - q e^\alpha} \right), \quad \alpha \in \{w \in \mathbb{C} \mid \text{Re } w < -\ln q\}. \tag{42}
\]

There are two representations for the negative binomial process. The first is a compound Poisson process for which the jump size \( J \in \mathbb{N} \) has a logarithmic distribution

\[
P_0(J = n) = -\frac{1}{\ln(1 - q)} \frac{1}{n} q^n, \tag{43}
\]

and the intensity determining the timing of the jumps is \(-m \ln(1 - q)\).
One finds that the characteristic function of \( J \) is

\[
\phi_0(\alpha) := \mathbb{E}^{P_0}[\exp(\alpha J)] = \frac{\ln(1 - q e^{\alpha})}{\ln(1 - q)}
\]  

(44)

for \( \alpha \in \{w \in \mathbb{C} \mid \text{Re } w < -\ln q\} \).

Then if we set

\[
n_t = \sum_{k=1}^{\infty} \mathbb{1}\{k \leq N_t\} J_k,
\]

(45)

where \( \{N_t\} \) is a Poisson process with rate \(-m \ln(1 - q)\), and \( \{J_k\}_{k \in \mathbb{N}} \) denotes a collection of independent identical copies of \( J \), representing the jumps, a calculation shows that

\[
\mathbb{P}_0(n_t = k) = \frac{\Gamma(k + mt)}{\Gamma(mt)\Gamma(k + 1)} q^k (1 - q)^{mt},
\]

(46)

and that the resulting exponent is given by (42).

The second representation of the negative binomial process makes use of the method of subordination.

We take a Poisson process with rate \( \Lambda = mq/(1 - q) \), and time-change it using a gamma subordinator \( \{\Gamma_t\} \) with rate parameter \( m \).
The moment generating function thus obtained, in agreement with (42), is

\[
\mathbb{E}^\mathbb{P}_0 \left[ \exp (\alpha N(\Gamma_t)) \right] = \mathbb{E}^\mathbb{P}_0 \left[ \exp (\Lambda(e^\alpha - 1) \Gamma_t) \right] = \left( \frac{1 - q}{1 - qe^\alpha} \right)^{mt}. \tag{47}
\]

With these results in mind, we fix a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and find the following:

**Proposition 9.** Let \(\{\Gamma_t\}\) be a gamma subordinator with rate \(m\), let \(\{N_t\}\) be an independent Poisson process with rate \(m\), let the independent random variable \(X\) satisfy \(X < -\ln q\) almost surely, and set

\[
\Gamma^X_t = \left( \frac{qe^X}{1 - qe^X} \right) \Gamma_t. \tag{48}
\]

Then the process \(\{\xi_t\}\) defined by

\[
\xi_t = N(\Gamma^X_t) \tag{49}
\]

is a Lévy information process with message \(X\) and negative binomial noise, with fiducial exponent \((42)\).
There is also a representation for negative binomial information based on the compound Poisson process.

This can be obtained by an application of Proposition ??, which shows how the Lévy measure transforms under a random Esscher transformation.

In the case of a negative binomial process with parameters $m$ and $q$, the Lévy measure is given by

$$
\nu(dz) = m \sum_{n=1}^{\infty} \frac{1}{n} q^n \delta_n(dz),
$$

where $\delta_n(dz)$ denotes the Dirac measure with unit mass at the point $z = n$.

The Lévy measure is finite in this case, and we have $\nu(\mathbb{R}) = -m \ln(1 - q)$, which is the overall rate at which the compound Poisson process jumps.

If one normalises the Lévy measure with the overall jump rate, one obtains the probability measure (43) for the jump size.

With these facts in mind, we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and specify the constants $m$ and $q$, where $m > 1$ and $0 < q < 1$. Then as a consequence of Proposition 6 we have the following:
Proposition 10. Let the random variable $X$ satisfy $X < -\ln q$ almost surely, let the random variable $J^X$ have the conditional distribution

$$\mathbb{P}(J^X = n \mid X) = -\frac{1}{\ln(1 - qe^X)} \frac{1}{n} (qe^X)^n,$$

(51)

let $\{J^X_k\}_{k \in \mathbb{N}}$ be a collection of conditionally independent identical copies of $J^X$, and let $\{N_t\}$ be an independent Poisson process with rate $m$. Then the process $\{\xi_t\}$ defined by

$$\xi_t = \sum_{k=1}^{\infty} \mathbb{1}\{k \leq N(-\ln(1 - qe^X)t)\} J^X_k$$

(52)

is a Lévy information process with message $X$ and negative binomial noise, with fiducial exponent (42).
Conclusions

We wrap up this study of Lévy information with the following remarks.

Applications to physical (Brody & Hughston 2006) and economic (Brody et al. 2008a) time series bring out the following points:

(i) Signal processing techniques have far-reaching applications to the identification, characterization and categorization of phenomena, both in the natural and in the social sciences.

(ii) Beyond the conventional remits of prediction, filtering, and smoothing there is a fourth and important new domain of applicability: the description of phenomena, both in physical and social sciences.

In particular, there is scope for the modeling of asset price processes.

In some financial models the “message” is the state variable that determines a cash flow, and the associated information process generates the filtration representing market information.

In another class of financial models the hidden variable represents the excess rate of return per unit of risk demanded by investors in exchange for investment in a risky asset.
References


