Elect. Comm. in Probab. **13** (2008), 113–120

ELECTRONIC COMMUNICATIONS in PROBABILITY

OPTIMISING PREDICTION ERROR AMONG COMPLETELY MONOTONE COVARIANCE SEQUENCES

ROSS MCVINISH[1]
Queensland University of Technology, GPO Box 2434, Brisbane, Q4001, Australia.
email: `r.mcvinish@qut.edu.au`

Submitted September 18, 2007, *accepted in final form* February 8, 2008

AMS 2000 Subject classification: 60G25; 44A60
Keywords: Aggregation; Maximum entropy; Moment space.

Abstract
We provide a characterisation of Gaussian time series which optimise the one-step prediction error subject to the covariance sequence being completely monotone with the first m covariances specified.

1 Introduction

In [9] the aggregation of simple dynamic equations, in particular AR(1) models, was introduced as a means for generating processes displaying long memory. Since this paper, the idea of aggregating has become a popular approach for constructing time series with a flexible covariance structure (see, for instance [1, 12, 2, 19, 15, 17]). Aggregation has also become a useful way to represent certain time series and provides a method for simulation. Examples of time series which can be represented in this way include the fractional ARIMA$(0, d, 0)$ model proposed by Granger and Joyeaux in [10] and Lamperti's transformation of fractional Brownian motion with $H < 1/2$ [5]. Other examples are given in [14]. The problem of representing long memory processes in terms of an aggregation of short memory processes is studied in [7].
Consider the aggregation of independent Gaussian AR(1) time series with positive correlation parameters. The covariance sequence of this time series is represented as

$$\gamma_k = \int_{[0,1]} \rho^{|k|} \sigma(d\rho), \tag{1}$$

for some Borel measure σ on $[0, 1]$. The representation (1) is equivalent to stating that $\{\gamma_k\}_{k=0}^{\infty}$ has the completely monotone property, that is, for all $n, k \geq 0$

$$(-1)^n \Delta^n \gamma_k \geq 0$$

[1]RESEARCH SUPPORTED BY THE ARC CENTRE FOR COMPLEX DYNAMIC SYSTEMS AND CONTROL CEO348165. THE AUTHOR GRATEFULLY ACKNOWLEDGES HELPFUL DISCUSSION WITH K. MENGERSEN.

where Δ is the difference operator $\Delta a_n = a_{n+1} - a_n$ ([18]). In this note we consider the problem of optimising the one-step prediction error variance, denoted by $\psi(\sigma)$, subject to the covariance sequence having representation (1) and satisfying equality constraints

$$\gamma_k = c_k, \quad k = 0, 1, \ldots, m. \tag{2}$$

The measure which maximises the prediction error will also provide the maximum entropy process ([6]) subject to constraints (1) and (2). While the measure which minimises the prediction error will not necessarily be the minimum entropy process, it will still provide a lower bound on the prediction error for Gaussian time series with completely monotone covariance.

For the case where $m = 1$, the problem of maximising $\psi(\sigma)$ subject to (1) and (2) can be solved from the result of Burg [3]. He showed that the time series which maximises prediction error subject to (2) is a Gaussian AR(m) time series. When $m = 1$ this result can be applied so that the measure which maximises the prediction error subject to $\gamma_0 = c_0,\ \gamma_1 = c_1$ is the measure

$$\sigma(A) = c_0 \chi_A(c_1/c_0),$$

where $\chi_A(x)$ is the characteristic function of the set A taking the value 1 if $x \in A$ and zero otherwise. The measure which minimises $\psi(\sigma)$ subject to (1) and (2) was shown in [14] to be

$$\sigma(A) = (c_0 - c_1)\chi_A(0) + c_1 \chi_A(1).$$

It is interesting to note that for $m = 1$ the measure which maximises prediction error minimises the covariance γ_k for $k \geq 2$ and similarly, the measure which minimises prediction error maximises the covariance γ_k for $k \geq 2$. In the terminology of moment spaces [13] these two measures are called the extremising measures of the moment space [11]

The paper is organised as follows. In Section 2 we recall some basic properties of time series with completely monotone covariance sequences and some properties of moment spaces. The main result of this paper, that the extremising measures of the moment space are precisely the measures which maximise and minimise the prediction error, is stated in Section 3. The proof of the result is given first assuming the measure is supported on $[0, \phi]$ with $\phi < 1$ and then for the case of measures supported on $[0, 1]$. The paper concludes with a discussion of how these measures are constructed.

2 Preliminaries

2.1 Wold decomposition

Consider a time series with covariance sequence (1) and such that $\sigma(\{1\}) = 0$. Then the corresponding spectral density function is given by

$$f(\omega; \sigma) = \frac{1}{2\pi} \int_{[0,1)} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho\cos\omega} \sigma(d\rho) = \int_{[0,1)} g(\omega; \rho)\sigma(d\rho),$$

where $g(\omega; \rho)$ is the spectral density of an AR(1) time series with variance 1. This spectral density satisfies the following bounds

$$f(\omega; \sigma) \geq \frac{1}{8\pi}\left(\gamma_0 - \gamma_2\right), \quad \omega \in (-\pi, \pi], \tag{3}$$

and

$$f(\omega;\sigma) \le \frac{\gamma_0}{2\pi\left(1-\cos^2(\omega)\right)}, \quad \omega \in (-\pi/2, \pi/2). \tag{4}$$

From (3) it is seen that if $\sigma(\{1\}) = 0$ then the time series is completely non-deterministic and the prediction error is given by the Kolmogorov formula

$$\psi(\sigma) = 2\pi \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\omega;\sigma) d\omega\right\}. \tag{5}$$

In general, the Wold decomposition of a time series with covariance sequence (1) is simply

$$X_t = Z + Y_t,$$

where Z is a random variable whose variance is $\sigma(\{1\})$ and Y_t has covariance sequence $\gamma_k - \gamma_\infty$. The prediction error is still given by (5) provided the spectral density is understood as that of the completely nondeterministic part.

2.2 Moment spaces

Let $\mathcal{D} \mid c$ be the set of measures on $[0,1]$ satisfying (2). When we need to consider measures on $[0,\phi]$, $\phi \le 1$ we shall write $\mathcal{D}_\phi \mid c$. A necessary and sufficient condition for $\mathcal{D} \mid c$ to consist of at least one measure is that the quadratic forms

$$\sum_{j=0}^{q}\sum_{k=0}^{q} c_{j+k} x_j x_k, \quad \sum_{j=0}^{q-1}\sum_{k=0}^{q-1} \left(c_{j+k+1} - c_{j+k+2}\right) x_j x_k, \tag{6}$$

for $m = 2q$ or

$$\sum_{j=0}^{q}\sum_{k=0}^{q} c_{j+k+1} x_j x_k, \quad \sum_{j=0}^{q}\sum_{k=0}^{q} \left(c_{j+k} - c_{j+k+1}\right) x_j x_k, \tag{7}$$

for $m = 2q+1$, are positive definite or semi-definite (see Theorems 16.1a and 16.1b in [13]). From Theorem 16.2 and Theorem 20.1 of [13] it follows that if (6) or (7), as required, is positive definite then there is more than one measure in $\mathcal{D} \mid c$. In this case we say $\mathcal{D} \mid c$ is non-degenerate. If (6) or (7), as required, is only positive semi-definite then there is only one measure in $\mathcal{D} \mid c$ and we say it is degenerate. These conditions are easily modified for $\mathcal{D}_\phi \mid c$.

2.3 Extremising measures

Let $\mathcal{G}_\phi^{(m)}$ be the set of continuous, bounded functions on $[0,\phi]$ whose first m derivatives are monotonic on $(0,\phi)$ and let $g \in \mathcal{G}_\phi^{(m)}$. Assume that $\mathcal{D}_\phi \mid c$ is non-degenerate. The problem of optimising $\mathbb{E}_\sigma g(\rho)$ over $\mathcal{D}_\phi \mid c$ has been studied by [11]. If $m = 2q$ then the two extremising measures of $\mathcal{D}_\phi \mid c$ are defined by:

$\sigma^*_{1,\phi}$, the measure on $[0,\phi]$ having q atoms in $(0,\phi)$ and an atom at 0.

$\sigma^*_{2,\phi}$, the measure on $[0,\phi]$ having q atoms in $(0,\phi)$ and an atom at ϕ.

If $m = 2q+1$ then the two extremising measures of $\mathcal{D}_\phi \mid c$ are defined by:

$\sigma^*_{1,\phi}$, the measure on $[0,\phi]$ having $q+1$ atoms in $(0,\phi)$.

$\sigma^*_{2,\phi}$, the measure on $[0,\phi]$ having q atoms in $(0,\phi)$, an atom at 0 and another atom at ϕ.

Note that these conditions define unique measures (see Theorem 5 of [11]). In [11] it was demonstrated that one of $\sigma^*_{1,\phi}, \sigma^*_{2,\phi}$ will maximise $\mathbb{E}_\sigma g(\rho)$ while the other measure will minimise the expectation. Which of these measures acts to minimise and which acts to maximise depends on the function g.

3 Main result

We first consider the case of measures on $[0,\phi]$ with $\phi < 1$. The following result confirms the observation made in the introduction that in order to maximise (minimise) the prediction error we should minimise (maximise) the covariance sequence.

Lemma 1. *Let $\gamma, \tilde\gamma$ be the covariance sequences associated with measures $\sigma, \tilde\sigma \in \mathcal{D}_\phi \mid c, \phi < 1$. If $\gamma_k \geq \tilde\gamma_k$ for all $k \geq m$ then $\psi(\sigma) \leq \psi(\tilde\sigma)$.*

Proof. To show that $\psi(\sigma) \leq \psi(\tilde\sigma)$ it is sufficient to show that

$$\int_{-\pi}^{\pi} \log \frac{f(\omega;\sigma)}{f(\omega;\tilde\sigma)} d\omega \leq 0. \tag{8}$$

From the inequality $x - 1 - \log x \geq 0$, for all $x > 0$,

$$\int_{-\pi}^{\pi} \log \frac{f(\omega;\sigma)}{f(\omega;\tilde\sigma)} d\omega \leq \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} (\gamma_k - \tilde\gamma_k) \cos(\omega k) f^{-1}(\omega;\tilde\sigma) d\omega. \tag{9}$$

The interchange of summation and integration is permitted since the covariance sequences are in ℓ_1 and $f^{-1}(\omega;\tilde\sigma)$ is bounded. Applying Jensen's inequality

$$\begin{aligned}\int_{-\pi}^{\pi} \cos(\omega k) f^{-1}(\omega;\tilde\sigma) d\omega &\leq \int_{-\pi}^{\pi} \cos(k\omega) \int_{[0,\phi]} g^{-1}(\omega;\rho)\tilde\sigma(d\rho) d\omega \\ &\leq \int_{[0,\phi]} \int_{-\pi}^{\pi} \cos(k\omega) g^{-1}(\omega;\rho) d\omega \tilde\sigma(d\rho) \\ &\leq 0, \quad \text{for } k \neq 0, \end{aligned} \tag{10}$$

where the final inequality follows from the fact that the inverse autocovariance of an AR(1) time series with positive correlation parameter is non-positive for all $k \neq 0$ [4]. The result now follows from (8)-(10).

Lemma 2. *The measure which maximises $\mathbb{E}_\sigma \rho^j,\ j > m$ over $\mathcal{D}_\phi \mid c$ is $\sigma^*_{2,\phi}$. Conversely, the measure which minimises $\mathbb{E}_\sigma \rho^j,\ j > m$ over $\mathcal{D}_\phi \mid c$ is $\sigma^*_{1,\phi}$.*

Proof. The second statement is an immediate implication of the first. The result for all $\phi \leq 1$ will follow immediately if it can be shown to be true for $\phi = 1$. Since σ^*_2 always has an atom at 1 it follows that

$$\lim_{j\to\infty} \mathbb{E}_{\sigma^*_2} \rho^j = \alpha > 0.$$

On the other hand σ^*_1 only has atoms in $[0,1)$ and so

$$\lim_{j\to\infty} \mathbb{E}_{\sigma^*_1} \rho^j = 0.$$

Therefore, there exists a $J \in \mathbb{N}$ such that for all $j \geq J$

$$\mathbb{E}_{\sigma_1^*}\rho^j < \mathbb{E}_{\sigma_2^*}\rho^j. \tag{11}$$

Now assume that $\mathbb{E}_\sigma \rho^j$ is not maximised by σ_2^* for some $j > m$ and let k be the largest integer such that

$$\mathbb{E}_{\sigma_1^*}\rho^k > \mathbb{E}_{\sigma_2^*}\rho^k. \tag{12}$$

From (11) and (12) there exists an $\omega_j \in (0, 1)$ such that

$$\mathbb{E}_{\sigma_1^*}\left(\omega_j\rho^k + (1-\omega_j)\rho^j\right) = \mathbb{E}_{\sigma_2^*}\left(\omega_j\rho^k + (1-\omega_j)\rho^j\right).$$

As $\omega_j\rho^k + (1-\omega)\rho^j \in \mathcal{G}^{(m)}$ we can apply Theorem 7 of [11] so that

$$\mathbb{E}_\sigma\left(\omega_j\rho^k + (1-\omega_j)\rho^j\right) = C_j \tag{13}$$

for all $\sigma \in \mathcal{D} \mid c$. Now take any $\sigma \in \mathcal{D} \mid c$ which does not have an atom at 1. As $\lim_{j\to\infty}\omega_j = \omega > 0$ we can apply (13) to conclude that

$$\lim_{j\to\infty}\frac{C_j}{\omega_j} = \mathbb{E}_\sigma\rho^k,$$

and all higher order moments can be determined from (13). Therefore, for any two measures $\sigma, \sigma' \in \mathcal{D} \mid c$ which do not have atoms at 1, $\mathbb{E}_\sigma\rho^j = \mathbb{E}_{\sigma'}\rho^j, j \geq k$. It follows that there is only one measure in $\mathcal{D} \mid c$ which does not have an atom at 1 ([8]) and this is σ_1^*. If we can show that there exists another measure in $\mathcal{D} \mid c$ which does not have an atom at 1 then a contradiction will occur and the result will be established. Take $\phi = 1 - \epsilon$ with $\epsilon > 0$ and such that $\sigma_1^*([\phi, 1]) = 0$. Define $\tilde{c}_k = \phi^{-k}c_k, k = 0, \ldots, m$. As $\mathcal{D} \mid c$ is non-degenerate and (6) and (7) are continuous, we may take ϵ sufficiently small so that $\mathcal{D} \mid \tilde{c}$ is also non-degenerate. Let $\tilde{\sigma}_2^*$ be the extremising measure in $\mathcal{D} \mid \tilde{c}$ which has an atom at 1. The measure $\tilde{\sigma}_2^*(\phi^{-1}d\rho)$ is in $\mathcal{D} \mid c$ and does not have an atom at 1. By construction, this measure is different to σ_1^* and hence there is more than one measure in $\mathcal{D} \mid c$ which does not have an atom at 1. Having arrived at a contradiction, it follows that there is no k for which (12) holds and the result now follows.

The following theorem is a direct result of lemmas 1 and 2.

Theorem 3. *Let $\phi < 1$. The measure in $\mathcal{D}_\phi \mid c$ which maximises prediction error is $\sigma_{1,\phi}^*$. The measure in $\mathcal{D}_\phi \mid c$ which minimises the prediction error is $\sigma_{2,\phi}^*$.*

The space $\mathcal{D} \mid c$ introduces a problem since the covariance sequence will not necessarily be in ℓ_1. To overcome this problem we first show that $\psi(\sigma)$ is continuous with respect to weak convergence of σ. We then consider the limit of the extremising measures $\sigma_{1,\phi}^*,\ \sigma_{2,\phi}^*$ as $\phi \to 1$.

Lemma 4. *Let σ_j be a sequence of measures converging weakly to σ. Then $\psi(\sigma_j) \to \psi(\sigma)$.*

Proof. (See Proposition 4.1 (ii) of [14]) We need to show that for any $\epsilon > 0$ there exists a $J \in \mathbb{N}$ such that for all $j \geq J$

$$\left|\int_{-\pi}^{\pi}\log\left[\frac{f(\omega;\sigma)}{f(\omega;\sigma_j)}\right]d\omega\ \right| < \epsilon$$

Take $\delta > 0$, as a family of functions of ρ on $[0, 1]$, parameterized by ω, $g(\omega; \rho)$ is bounded and uniformly continuous for all $|\omega| > \delta$. Applying Theorem 3.2 of [16] we have $f(\omega;\sigma_j) \to f(\omega;\sigma)$

uniformly on $|\omega| > \delta$. As $f(\omega;\sigma)$ is bounded away from zero we may apply the dominated convergence theorem to obtain

$$\lim_{j\to\infty}\int_\delta^\pi \log\left[\frac{f(\omega;\sigma)}{f(\omega;\sigma_j)}\right]d\omega = 0.$$

Applying the bounds (3) and (4) we can show that the integral

$$\int_0^\delta \log\left[\frac{f(\omega;\sigma)}{f(\omega;\sigma_m)}\right]d\omega$$

can be made arbitrarily small by taking δ sufficiently small. The result follows.

Lemma 5. *Let $V(\sigma)$ be a weak neighbourhood around $\sigma \in \mathcal{D} \mid c$. Then there exists a measure $\tilde{\sigma} \in V(\sigma)$ such that $\tilde{\sigma} \in \mathcal{D}_\phi \mid c$ for some $\phi < 1$.*

Proof. It follows from the definition of the Prohorov metric that for $\phi < 1$ but sufficiently large the measure $\sigma(\phi^{-1}d\rho)$ is in $V(\sigma)$.

Lemma 6. *The sequence of extremising measures $\{\sigma^*_{1,\phi}\}$ and $\{\sigma^*_{2,\phi}\}$ converge weakly to σ^*_1 and σ^*_2, respectively, as $\phi \to 1$.*

Proof. For any $\phi' < \phi \leq 1$ it is obvious that $\mathcal{D}_{\phi'} \mid c \subseteq \mathcal{D}_\phi \mid c$. From Lemma 2 and Theorem 7 of [11] it follows that for any $j \in \mathbb{N}$, $\mathbb{E}_{\sigma^*_{1,\phi}}\rho^j$ is bounded and decreasing as $\phi \to 1$. Therefore, the moments converge for all $j \in \mathbb{N}$ and so the measures must converge weakly. That the limiting measure is σ^*_1 follows from Lemma 5 since if $\tilde{\sigma} \in \mathcal{D}_\phi \mid c,\ \phi < 1$ is in a weak neighbourhood of σ^*_1 then

$$\mathbb{E}_{\tilde{\sigma}}\rho^j \geq \mathbb{E}_{\sigma^*_{1,\phi}}\rho^j \geq \mathbb{E}_{\sigma^*_1}\rho^j,$$

and $\sigma^*_{1,\phi}$ must also be in a weak neighbourhood of σ^*_1. A similar argument applies for $\{\sigma^*_{2,\phi}\}$.

Theorem 7. *The measure in $\mathcal{D} \mid c$ which maximises the prediction error is σ^*_1. The measure in $\mathcal{D} \mid c$ which minimises the prediction error is σ^*_2.*

Proof. Let σ^* be the measure in $\mathcal{D} \mid c$ which maximises $\psi(\sigma)$. From Lemma 4 and Lemma 5 it follows that for any $\epsilon > 0$ we can find a $\phi < 1$ and a measure $\tilde{\sigma}_\phi \in \mathcal{D}_\phi \mid c$ such that $\psi(\sigma^*) - \epsilon \leq \psi(\tilde{\sigma}_\phi)$. From Theorem 3 $\psi(\sigma^*) - \epsilon \leq \psi(\sigma^*_{1,\phi})$ and applying Lemma 4 and Lemma 6 it follows that $\psi(\sigma^*) - \epsilon \leq \psi(\sigma^*_1)$. As this holds for all $\epsilon > 0$ we can conclude that $\psi(\sigma^*) = \psi(\sigma^*_1)$ and hence σ^*_1 is the measure which maximises prediction error. A similar argument holds for minimising the prediction error.

To conclude we briefly describe how the extremising measures can be constructed from the given covariances $c_0, \ldots, c_m$. First define the polynomials $\underline{\Delta}_n(t)$ and $\overline{\Delta}_n(t)$ for $n = 2k$ by the determinants,

$$\underline{\Delta}_n(t) = \begin{vmatrix} c_0 & c_1 & \cdots & c_{k-1} & 1 \\ c_1 & c_2 & \cdots & c_k & t \\ \vdots & & & & \vdots \\ c_k & c_{k+1} & \cdots & c_{2k-1} & t^k \end{vmatrix}, \quad \overline{\Delta}_n(t) = \begin{vmatrix} c_1 - c_2 & \cdots & c_{k-1} - c_k & 1 \\ c_2 - c_3 & \cdots & c_k - c_{k+1} & t \\ \vdots & & \vdots & \vdots \\ c_k - c_{k+1} & \cdots & c_{2k-2} - c_{2k-1} & t^{k-1} \end{vmatrix},$$

and for $n = 2k + 1$ by the determinants,

$$\underline{\Delta}_n(t) = \begin{vmatrix} c_1 & c_2 & \cdots & c_k & 1 \\ c_2 & c_3 & \cdots & c_k & t \\ \vdots & & & & \vdots \\ c_{k+1} & c_{k+2} & \cdots & c_{2k} & t^k \end{vmatrix}, \quad \overline{\Delta}_n(t) = \begin{vmatrix} c_0 - c_1 & \cdots & c_{k-1} - c_k & 1 \\ c_1 - c_2 & \cdots & c_k - c_{k+1} & t \\ \vdots & & \vdots & \vdots \\ c_k - c_{k+1} & \cdots & c_{2k-1} - c_{2k} & t^k \end{vmatrix}.$$

Define the polynomials $\overline{P}_n(t)$ and $\underline{P}_n(t)$ by

$$\underline{P}_n(t) = \begin{cases} [\underline{\Delta}_n(t)]^2, & \text{if } n \text{ is even,} \\ t\,[\underline{\Delta}_n(t)]^2, & \text{if } n \text{ is odd;} \end{cases}$$

$$\overline{P}_n(t) = \begin{cases} t(1-t)\,\left[\overline{\Delta}_n(t)\right]^2, & \text{if } n \text{ is even,} \\ (1-t)\,\left[\overline{\Delta}_n(t)\right]^2, & \text{if } n \text{ is odd;} \end{cases}$$

From Theorem 20.2 of [13], the location of the atoms in the measures σ_1^* and σ_2^* are given by the distinct roots of the polynomials $\underline{P}_{m+1}(t)$ and $\overline{P}_{m+1}(t)$, respectively. The mass associated with each atom in the measure can then be determined from the linear system

$$w_1\rho_1^k + \ldots + w_r\rho_r^k = c_k, \quad k = 0, 1, \ldots, m,$$

where $\rho_1, \ldots, \rho_r$ are the locations of the atoms (see section 2.3 for the number of atoms in the extremising measures). This system will be uniquely solvable provided $\mathcal{D} \mid c$ contains at least one measure.

As an example, consider the case of $m = 2$. The polynomial determining the locations of the atoms of σ_1^* is

$$\underline{P}_3(t) = t(c_1t - c_2)^2.$$

Therefore, σ_1^* has one atom at 0 and another at $\rho = c_2/c_1$. The moment constraints determine the measure to be

$$\sigma_1^*(A) = \frac{c_2 - c_1^2}{c_2}\chi_A(0) + \frac{c_1^2}{c_2}\chi_A\left(\frac{c_2}{c_1}\right).$$

Similarly, the polynomial determining the location of the atoms of σ_2^* is

$$\overline{P}_3(t) = (1-t)((1-c_1)t - (c_1 - c_2))^2.$$

The measure σ_2^* is then

$$\sigma_2^*(A) = \frac{(1-c_1)^2}{1 - 2c_1 + c_2}\chi_A\left(\frac{c_1 - c_2}{1 - c_1}\right) + \frac{c_2 - c_1^2}{1 - 2c_1 + c_2}\chi_A(1).$$

References

[1] Barndorff-Neilsen, O.E., Jensen, J.L., and Sorensen, M. (1995) “Some stationary processes in discrete and continuous time,” *Advances in Applied Probability*, 30, 989-1007. MR1671092

[2] Barndorff-Neilsen, O.E., and Shephard, N. (2001) “Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial econometrics (with discussion),” *Journal of the Royal Statistical Society, Series B*, 63, 167-241. MR1841412

[3] Burg, J.P. (1975) *Maximum entropy spectral analysis*, Ph.D. Thesis, Department of Geophysics, Stanford University.

[4] Chatfield, C. (1979) “Inverse autocorrelations,” *Journal of the Royal Statistical Society, Series A*, 142, 363-377. MR0547168

[5] Cheridito, P., Kawaguchi, H., and Maejima, M. (2003) "Fractional Ornstein-Uhlenbeck processes," *Electronic Journal of Probability*, 8(3), 1-14. MR1961165

[6] Cover, T.M., and Thomas, J.A. (1991) *Elements of information theory*, John Wiley and Sons, New York. MR1122806 MR1122806

[7] Dacunha-Castelle, D., and Fermín, L. (2006) "Disaggregation of long memory processes on $\mathcal{C}^\infty$ class," *Electronic Communications in Probability*, 11, 35-44. MR2219344

[8] Feller, W. (1939) "Completely monotone functions and sequences," *Duke Mathematical Journal*, 5, 661-674. MR0000315

[9] Granger, C.W.J. (1980) "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227-238. MR0597259

[10] Granger, C.W.J., and Joyeux, R. (1980) "An introduction to long-memory time series models and fractional differencing," *Journal of Time Series Analysis*, 1, 15-29. MR0605572

[11] Harris, B. (1959) "Determining bounds on integrals with applications to cataloging problems," *Annals of Mathematical Statistics*, 30, 521-548. MR0102876

[12] Igloi, E., and Terdik, G. (1999) "Long-range dependence through Gamma-mixed Ornstein-Uhlenbeck processes," *Electronic Journal of Probability*, 4(16), 1-33. MR1713649

[13] Kalrin, S., and Shapley, L.S. (1953) "The geometry of moment spaces," *Memoirs of the American Mathematical Society*, No. 12. MR0059329

[14] McVinish, R. (2008) "On the structure and estimation of reflection positive processes," *Journal of Applied Probability*, 45(1), *to appear.*

[15] Oppenheim, G., and Viano, M.-C. (2004) "Aggregation of random parameters Ornstein-Uhlenbeck or AR processes: some convergence results," *Journal of Time Series Analysis*, 25, 335-350. MR2062677

[16] Ranga Rao, R. (1962) "Relations between weak and uniform convergence of measures with applications," *Annals of Mathematical Statistics*, 32, 659-680. MR0137809

[17] Roberts, G.O., Papaspiliopoulos, O., and Dellaportas, P. (2004) "Bayesian inference for non-Gaussian Ornstein-Uhlenbeck stochastic volatility processes," *Journal of the Royal Statistical Society, Series B*, 66, 369-393. MR2062382

[18] Widder, D.V (1946) *The Laplace Transform*, Princeton University Press. MR0005923

[19] Zaffaroni, P. (2004) "Contemporaneous aggregation of linear dynamic models in large economies," *Journal of Econometrics*, 120, 75-102. MR2047781