Rational Approximation of Transfer Functions for Non-Negative EPT Densities

Conor Sexton * **Martine Olivi** ** **Bernard Hanzon** ***

* *School of Mathematical Sciences, University College Cork, Western Rd., Cork, Ireland (e-mail: hcsexton@hotmail.com)*
** *2004 route des Lucioles, BP 93, 06902, Sophia Antipolis, Cedex, France (e-mail: Martine.Olivi@sophia.inria.fr*
*** *School of Mathematical Sciences, University College Cork, Western Rd., Cork, Ireland (e-mail: b.hanzon@ucc.ie)*

Abstract: A stable Exponential-Polynomial-Trigonometric (EPT) probability density function is fitted to a large set of financial data. The class of EPT functions have a strictly proper continuous time rational transform. An isometry is used to derive its discrete time transfer function, also rational. This function can be written in terms of Fourier coefficients which are used as inputs to RARL2. A minimal state space realization is returned to approximate the density function. Non-negativity of the EPT function must be examined which is carried out using the Budan-Fourier technique. A convex optimisation algorithm is then implemented to ensure an optimal non-negative approximation.

Keywords: Probability Density Function, Rational Matrices, Fourier Transforms, Impulse Responses, Lyapunov Function, Convex Optimisation, Function Approximation, Sampled Data, Financial Systems

1 Introduction

The problem of approximating empirical data with a non-negative 2-EPT probability density function is considered here. Hanzon and Sexton [2011] introduce 2-EPT probability density functions and illustrate how they provide an ideal framework for probabilistic and other calculations. Much Financial Mathematics literature assumes that log returns follow a Gaussian distribution which has been shown to provide a poor fit. The Gaussian distribution is often used due to its analytic tractability rather than its goodness of fit.

We demonstrate how to fit an EPT density (i.e. a 2-EPT density defined on $[0, \infty)$ only) to a set of positive data. The dataset to be approximated is the positive daily Dow Jones Industrial Average (DJIA) log returns over 80 years which are shown in Figure (1). The procedure is identical to fit another EPT function on $(-\infty, 0]$ to the negative log returns.

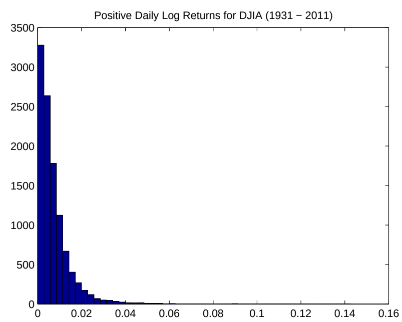

Fig. 1. Daily Log returns for DJIA for (1931 - 2011)

An EPT density can be interpreted as the impulse response of a causal stable system whose Laplace transform is the rational transfer function in continuous time. This transfer function can be viewed as the rational characteristic function of the EPT density. Assuming the data follows an EPT distribution we know from Parsevals Theorem that approximating the transform is equivalent to approximating the density itself. RARL2 is a software package developed in INRIA to approximate rational functions in discrete time. Using an isometry we can transform the continuous time transfer function into the discrete time transfer function which is also rational and can be approximated with RARL2. A description of the capabilities of RARL2 and further references can be found in Olivi [2010].

2 RARL2 Approximation

Firstly the data must be scaled so it fills the interval $[0, 1]$ so all returns are scaled by the largest observed return which in this case is 0.1427. This is necessary to estimate the Fourier coefficients correctly which we will see in Section (3). Figure (2) shows the normalised (i.e. area under curve is one) density to be approximated.

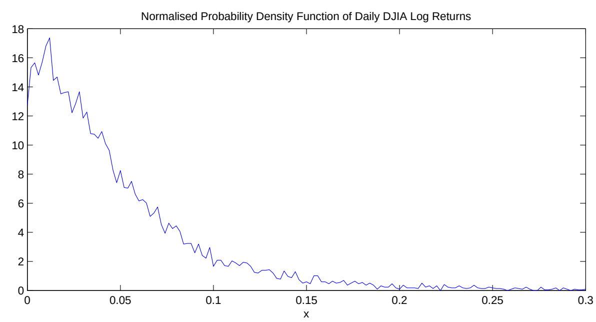

Fig. 2. Normalised Density for observed daily log returns

The observed density will be denoted $h(x)$ where $0 \le x \le 1$ which will be approximated with the EPT probabilty density function $f(x) = \mathbf{c}e^{\mathbf{A}x}\mathbf{b}$. The criterion function to be minimised is

$$\min_{\mathbf{A},\mathbf{b},\mathbf{c}|x\geq 0} ||f(x) - h(x)||_2^2 \quad (1)$$

over the set of minimal triples $(\mathbf{A}, \mathbf{b}, \mathbf{c})$ where $\mathbf{A}$ is stable and of fixed degree.

The continuous time transfer function can be found via the transform

$$\mathbb{E}[e^{-sx}] = F(s) = \int_0^\infty \mathbf{c}e^{\mathbf{A}x}\mathbf{b}e^{-sx}dx = \mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \quad (2)$$

which is a strictly proper rational function. $F(is)$ is the characteristic function of f. In continuous time the $\mathbf{A}$ matrix must be stable with all eigenvalues located in the open left half plane since we are using minimal realizations.

Using the normalised density $h(x)$, which is observable at $N+1$ points, we can find its transform H.

$$H(s) = \sum_{n=0}^{N-1} h(x_n)\Big(-\frac{1}{s}e^{-sx_{n+1}} + \frac{1}{s}e^{-sx_n}\Big) \quad (3)$$

where $x_0 = 0$, $x_N = 1$ and $x_{n+1} - x_n = \delta x$ for all n where $\delta x = \frac{1}{N}$.

To transform the continuous time rational function into discrete time we use a well known isometry. In discrete time the stable $\mathbf{A}$ matrix will have all its eigenvalues located inside the unit disk. The map from continuous to discrete time is

$$F(s) \mapsto \tilde{F}(z) = \frac{\sqrt{2}}{z-1}F(\frac{z+1}{z-1}) \quad (4)$$

achieved using the mobius transform

$$s \mapsto z = \frac{s+1}{s-1} \quad (5)$$

These formulae allow us to derive a discrete time realization, $(\tilde{\mathbf{A}},\ \tilde{\mathbf{b}},\ \tilde{\mathbf{c}})$, of the continuous time realization $(\mathbf{A},\ \mathbf{b},\ \mathbf{c})$. The state space formulae of the transformation are symmetric and given by

$$\begin{aligned}\tilde{\mathbf{A}} &= -(\mathbf{I}-\mathbf{A})^{-1}(\mathbf{I}+\mathbf{A}) \quad (6)\\ \tilde{\mathbf{b}} &= \sqrt{2}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}\\ \tilde{\mathbf{c}} &= \mathbf{c}\end{aligned}$$

Hence the discrete time rational function is given by

$$\tilde{F}(z) = \tilde{\mathbf{c}}(z\mathbf{I} - \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{b}} \quad (7)$$

Equations (4) and (5) can be used to transform H into its discrete time counterpart $\tilde{H}$. RARL2 requires discrete time pointwise samples on the unit circle as inputs in the form $\tilde{H}(e^{i\theta})$ where $\theta \in [-\pi, \pi]$. Hence we let

$$i\omega = \frac{e^{i\theta}+1}{e^{i\theta}-1} \quad (8)$$

giving us

$$\begin{aligned}\tilde{H}(e^{i\theta}) &= \frac{\sqrt{2}}{e^{i\theta}-1}H(i\omega) \quad (9)\\ &= \frac{\sqrt{2}}{e^{i\theta}-1}\sum_{n=0}^{N-1} h(x_n)\Big(-\frac{1}{i\omega}e^{-i\omega x_{n+1}} + \frac{1}{i\omega}e^{-i\omega x_n}\Big)\end{aligned}$$

It is clear that problems are encountered for $\theta = 0$ as the denominator $e^{i\theta} - 1 = 0$. For this reason θ must be sufficiently far away from zero. θ is then sampled on the intervals $\theta \in [-\pi, -\frac{\pi}{M}]$ and $\theta \in [\frac{\pi}{M}, \pi]$ where $M > 0$ is chosen to be sufficiently small.

RARL2 seeks to minimise the L_2 norm between

$$\min_{\tilde{\mathbf{A}},\ \tilde{\mathbf{b}},\ \tilde{\mathbf{c}}} ||\tilde{H} - \tilde{F}||_2^2 \quad (10)$$

which is equivalent to Eq. (1).

3 RARL2 Approximation from Fourier Coefficients

It is also possible to perform rational approximation using RARL2 by providing the Fourier Coefficients of the discrete time rational function as inputs. It is shown here how the Fourier coefficients may be calculated directly from the data. Let $\tilde{H}(z)$ be the discrete time rational function and note that it can be written as

$$\begin{aligned}\tilde{H}(z) &= \sum_{n=-\infty}^{\infty} c_n z^n \quad (11)\\ \tilde{H}(e^{it}) &= \sum_{n=-\infty}^{\infty} c_n e^{int}\end{aligned}$$

where c_n are the Fourier coefficients given by

$$\begin{aligned}c_n &= \frac{1}{2\pi}\int_0^{2\pi} \tilde{H}(e^{it})e^{-int}dt \quad (12)\\ &= < \tilde{H},\ e^{int} >_{L^2(T)}\end{aligned}$$

The problem can be transformed into continuous time using Eq. (4) and the Mobius transforms from Eq. (5)

$$H(s) = \frac{\sqrt{2}}{s-1}\tilde{H}\Big(\frac{s+1}{s-1}\Big)$$

The continuous time transfer function in terms of Fourier Coefficients is

$$H(s) = \frac{\sqrt{2}}{s-1}\sum_{n=-\infty}^{\infty} c_n\Big(\frac{s+1}{s-1}\Big)^n \quad (13)$$

Letting

$$\begin{aligned}G_n(s) &= \frac{\sqrt{2}}{s-1}\Big(\frac{s+1}{s-1}\Big)^n \quad (14)\\ &= \sqrt{2}\frac{(s+1)^n}{(s-1)^{n+1}}\end{aligned}$$

The Fourier coefficients can then be found as solutions to

$$\begin{aligned}c_n &= < H, G_n >_{L^2(\mathbb{R})} \quad (15)\\ &= < h, g_n >_{L^2(\mathbb{R})}\end{aligned}$$

using Plancherals Theorem where h is the observed normalised density. g_n, the Fourier inverse of G_n, must be computed. The general solution is

$$\begin{aligned}g_n(t) &= \mathcal{F}^{-1}(G_n(t)) \quad (16)\\ &= \sqrt{2}e^{-t}P_n(t)\end{aligned}$$

For $n \geq 0$ we can derive $g_n(t)$, shown here without the $\sqrt{2}$ constant for clarity

$$
\begin{aligned}
n = 0 \qquad & \frac{1}{s-1} \mapsto -e^t u(-t) && (17) \\
n = 1 \qquad & \frac{s+1}{(s-1)^2} = \frac{1}{s-1} + \frac{2}{(s-1)^2} \mapsto -e^t u(-t)(1+2t) \\
n = 2 \qquad & \frac{((s-1)+2)^2}{(s-1)^3} = \frac{(s-1)^2}{(s-1)^3} + \frac{4(s-1)}{(s-1)^3} + \frac{4}{(s-1)^3} \\
& \mapsto -e^t u(-t)(1+4t+2t^2)
\end{aligned}
$$

where $u(x) = \mathbb{I}_{x>0}$. Thus for $t \leq 0$ only terms $n \geq 0$ are required

$$
g_n = \mathcal{F}^{-1}(G_n) = -e^t \sqrt{2} u(-t) \Big(\sum_{j=0}^{n} \binom{n}{j} \frac{(2t)^j}{j!} \Big) \qquad (18)
$$

Similarly for $n < 0$ we let $n' = -n - 1$ and note that

$$
\begin{aligned}
\frac{(s-1)^{-n-1}}{(s+1)^{-n}} &= \frac{(s-1)^{n'}}{(s+1)^{n'+1}} = \frac{(s+1-2)^{n'}}{(s+1)^{n'+1}} \qquad (19) \\
&= \sum_{j=0}^{n'} \binom{n'}{j} \frac{(s+1)^{n'-j}(-2)^j}{(s+1)^{n'+1}} \\
&= \sum_{j=0}^{n'} \binom{n'}{j} \frac{(-2)^j}{(s+1)^{j+1}}
\end{aligned}
$$

so $G_{n'}$ is given by

$$
G_{n'}(s) \;=\; \sqrt{2} \sum_{j=0}^{n'} \binom{n'}{j} \frac{(-2)^j}{(s+1)^{j+1}} \qquad (20)
$$

The Fourier inverse of which is

$$
\begin{aligned}
g_n = \mathcal{F}^{-1}(G_n) &= \sqrt{2} e^{-t} u(t) \Big(\sum_{j=0}^{n'} \binom{n'}{j} \frac{(-2t)^j}{j!} \Big) \qquad (21) \\
&= \sqrt{2} e^{-t} u(t) \Big(\sum_{j=0}^{-n-1} \binom{-n-1}{j} \frac{(-2t)^j}{j!} \Big)
\end{aligned}
$$

Clearly for $t > 0$ the Fourier coefficients with $n \leq 0$ are considered.

A simple test to examine the performance of the Fourier Approximation is that the density should be recovered via

$$
h(t) \;=\; \sum_{n=-\infty}^{0} c_n g_n(t) \qquad (22)
$$

The RARL2 software then converts these Fourier coefficients into a discrete time minimal realization $(\tilde{\mathbf{A}}, \; \tilde{\mathbf{b}}, \; \tilde{\mathbf{c}})$ of degree n. This realization is then used as an initial starting approximation for RARL2 which then uses a gradient search algorithm to locate the minimal realization of degree n subject to the criterion in Eq. (10). The Nyquist plot in Figure (2) shows the performance of the rational approximation.

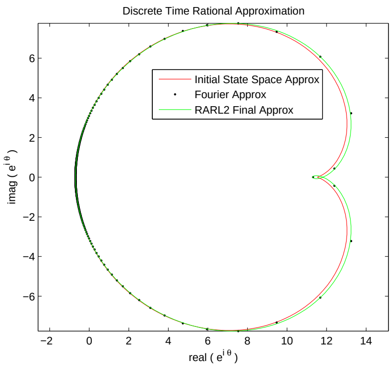

Fig. 3. Rational Approximation of order 3.

From inspection of the above plot the final RARL2 approximation appears to be better than the initial approximation which was obtained directly from converting the Fourier coefficients into a minimal realization. We will now examine how these approximations compare when the realizations are transformed into continuous time.

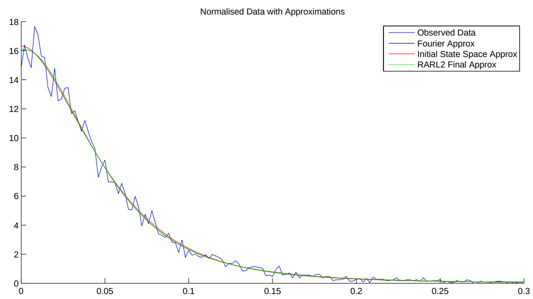

Fig. 4. Normalised Density for observed daily log returns with Fourier Approximation, Initial State Space Approximation of order 3 and Final RARL2 Approximation of Order 3

The Fourier approximation in Figure 2 is found from Eq. (22) with 170 Fourier coefficients. It is difficult to distinguish between Fourier approximation (black) and the final RARL2 approximation (green). The square error for the raw Fourier approach in Eq. (22) is 5.61% compared to the final RARL2 approximation of order 3 which has a square error of 5.59% also. The improvement provided by the RARL2 algorithm is clear when consider the initial state space approximation (red) which had a square error of 5.87%.

4 Non-Negative EPT Approximations

We now consider the problem of ensuring the approximating EPT densities are non-negative on the entire half real line $[0, \infty]$. This is an essential requirement in terms of probability density functions where the probabilities should not be negative. The Budan Fourier algorithm of Hanzon and Holland [2010] can be used to test for non-negativity of an EPT function on a finite interval $[0, T]$.

Again $h(x)$ is the true function we are approximating with the EPT density $f(x) = \mathbf{c}e^{\mathbf{A}x}\mathbf{b}$, the minimization problem we wish to solve is

$$
\min_{\{\mathbf{A},\mathbf{b},\mathbf{c}\} | f(x) \geq 0} ||h(x) - f(x)||_2^2 \qquad (23)
$$

Despite the fact that the true function $h(x)$ may be non-negative, for all $x \geq 0$, it is still possible that the approximating density $f(x)$ is negative on certain intervals. This has been observed in practice and is also present when we use the Fourier approximation in Eq. (22). The problem is also encountered in model reduction when approximating a high order non-negative system with a lower order system.

As we proceed to examining potential solutions to this problem we elucidate to the importance of the Perron-Frobenius type result which states *that if* $f(x) = \boldsymbol{c}e^{\boldsymbol{A}x}\boldsymbol{b}$ *is non-negative* $\forall\ x \geq 0$ *then* $\lambda_M = max_{\lambda \in \sigma(A)}\mathbb{R}(\lambda)$ *is an element of the spectrum of A*.This puts an obvious rather difficult constraint on the eigenvalues of the approximating **A** matrix in that it must contain a dominant real pole.

A possible solution to this problem is to assume the presence of a unique dominant real pole, $\lambda_M < 0$, in the system. The dominant pole and its coefficient, $\mu e^{\lambda_M x}$, can be estimated from $h(x)$. One approach to estimating these parameters is to denote $(N - K)$ points on the normalised density $h(x_i)$ and let $\hat{\mu}$, $\hat{\lambda}_M$ be the estimated parameters for the dominant real pole and coefficient. The function to be minimised is then

$$\min_{\hat{\mu},\hat{\lambda}_M} \sum_{i=N-K}^{N} (h(x_i) - \hat{\mu}e^{\hat{\lambda}_M x_i})^2 \tag{24}$$

K must be chosen sufficiently large to capture tail behaviour only. This minimization problem could be solved by fitting a regression line, $L(\hat{\mu}, \hat{\lambda}_M)$ to the log of the criterion i.e.

$$L(\hat{\mu}, \hat{\lambda}_M) = log(h(x_i)) = log(\hat{\mu}) - \hat{\lambda}_M x_i \tag{25}$$

Once the dominant pole and coefficient have been identified, they can be fixed and excluded from the approximation. $h(x)$ and $f(x)$ are transformed as follows

$$\hat{h}(x) = h(x) - \hat{\mu}e^{\hat{\lambda}_M x} \tag{26}$$
$$\hat{f}(x) = f(x) - \hat{\mu}e^{\hat{\lambda}_M x} = \hat{\mathbf{c}}e^{\hat{\mathbf{A}}x}\hat{\mathbf{b}}$$

The minimization problem is now

$$\min_{\{\hat{\mathbf{A}},\hat{\mathbf{b}},\hat{\mathbf{c}}\}|\hat{f}(x)\geq -\hat{\mu}e^{\hat{\lambda}_M x}\}} ||\hat{h}(x) - \hat{\mathbf{c}}e^{\hat{\mathbf{A}}x}\hat{\mathbf{b}}||_2^2 \tag{27}$$

To ensure $\hat{\lambda}_M$ remains the dominant pole the eigenvalues of $\hat{\mathbf{A}}$ must be constrained such that $\mathbb{R}(\sigma(\hat{\mathbf{A}})) < \hat{\lambda}_M$ where $\sigma(\hat{\mathbf{A}})$ indicates the spectrum of $\hat{\mathbf{A}}$.

RARL2 ensures the stability of $\hat{\mathbf{A}}$ such that $\mathbb{R}(\sigma(\hat{\mathbf{A}})) < 0$ but to implement the above constraint would require more work. If $\hat{h}(x)$ and $\hat{f}(x)$ are scaled by the factor $e^{-\hat{\lambda}_M}$ the minimisation problem is

$$\min_{\{\hat{\mathbf{A}},\hat{\mathbf{b}},\hat{\mathbf{c}}\}|\hat{\mathbf{c}}e^{(\hat{\mathbf{A}}-\mathbf{I}\hat{\lambda}_M)x}\hat{\mathbf{b}}\geq -\mu} ||e^{-\hat{\lambda}_M x}\hat{h}(x) - \hat{\mathbf{c}}e^{(\hat{\mathbf{A}}-\mathbf{I}\hat{\lambda}_M)x}\hat{\mathbf{b}}||_2^2 \tag{28}$$

reducing the eigenvalue constraint to $\mathbb{R}(\sigma(\hat{\mathbf{A}} - \mathbf{I}\hat{\lambda}_M)) < 0$ which is implemented by default. However this does imply that the L_2 minimisation is weighted with greater emphasis placed on the tail behaviour since $\hat{\lambda}_M < 0$.

The RARL2 algorithm does not take the non-negativity constraint into consideration. The RARL2 approximation is to minimize Eq. (29)

$$\min_{\{\hat{\mathbf{A}},\hat{\mathbf{c}}\}} ||e^{-\hat{\lambda}_M x}\hat{h}(x) - e^{-\hat{\lambda}_M x}\hat{f}(x)||_2^2 \tag{29}$$

The pair $(\hat{\mathbf{A}}, \hat{\mathbf{c}})$ ranges over the set of balanced observable pairs as described in Hanzon et al [2006] and implementing the gradient algorithms in Hanzon et al [2006] with such parametrizations allow us to solve for an optimal $\hat{\mathbf{b}}$.

The convex set $\hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ can then be defined for a given pair $(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ such that if $\hat{\mathbf{b}} \in \hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ then non-negativity is guaranteed by $f(x) > 0$ or

$$\hat{f}(x) = \hat{\mathbf{c}}e^{\hat{\mathbf{A}}x}\hat{\mathbf{b}} \geq -\hat{\mu}e^{\hat{\lambda}_M x} \qquad \forall\ x \geq 0 \tag{30}$$

Using the convex optimisation algorithm, illustrated below in Section (5), it is possible to find the constrained optimal $\hat{\mathbf{b}}_C^* \in \hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ s.t. $\hat{f}(x) = \hat{\mathbf{c}}e^{\hat{\mathbf{A}}x}\hat{\mathbf{b}}_C^* \geq -\hat{\mu}e^{\hat{\lambda}_M x}\ \forall\ x \geq 0$.

The final triple can then be represented in state space form by

$$\left(\begin{array}{cc|c} \hat{\lambda}_M & 0 & 1 \\ 0 & \hat{\mathbf{A}} & \hat{\mathbf{b}}_C^* \\ \hline \hat{\mu} & \hat{\mathbf{c}} & 0 \end{array}\right)$$

5 Convex Optimisation with Budan Fourier Algorithm

Implementing RARL2 with the criterion function in Eq. (28) yields the pair $(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ for which the optimal $\hat{\mathbf{b}}^*$ can be derived. If $\hat{\mathbf{b}}^* \notin \hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ then the resultant function $f(x)$ will have negative values, equivalent to $\hat{\mathbf{c}}e^{\hat{\mathbf{A}}x}\hat{\mathbf{b}}^* < -\hat{\mu}e^{\hat{\lambda}_M x}$, for some $x \geq 0$. The constrained optimal $\hat{\mathbf{b}} \in \hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ can be seen to be the $\hat{\mathbf{b}}$ which minimises the L_2 distance

$$\min_{\hat{\mathbf{b}} \in \hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})} ||\hat{\mathbf{b}} - \hat{\mathbf{b}}^*||_2^2 \tag{31}$$

Section (6) derives a T_0 such that if the EPT function is non-negative on $[0, T_0]$ then non-negativity is assured on $[0, \infty)$. This is possible due to the presence of the dominant real pole and a restriction on the norm of $\hat{\mathbf{b}}$. Using the Budan-Fourier algorithm non-negativity can then be examined on the whole positive half line.

An initial $\hat{\mathbf{b}}_0 \in \hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ is always available since the zero vector would trivially satisfy the non-negativity constraint. With the Budan-Fourier algorithm it is possible to find $\hat{\mathbf{B}}_0$ on the boundary of $\hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})$ where $\hat{\mathbf{B}}_0$ is a convex combination of $\hat{\mathbf{b}}_0$ and $\hat{\mathbf{b}}^*$. Setting

$$\hat{\mathbf{B}}_0 = \alpha\hat{\mathbf{b}}_0 + (1 - \alpha)\hat{\mathbf{b}}^* \tag{32}$$

α can be chosen, $0 < \alpha < 1$, such that

$$\min_{\alpha|(\alpha\hat{\mathbf{b}}_0+(1-\alpha)\hat{\mathbf{b}}^*)\in\hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})} \left|\left|(\alpha\hat{\mathbf{b}}_0 + (1-\alpha)\hat{\mathbf{b}}^*) - \hat{\mathbf{b}}^*\right|\right|_2^2 = \min_{\alpha|\hat{\mathbf{B}}_0(\alpha)\in\hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})} \left|\left|\hat{\mathbf{B}}_0 - \hat{\mathbf{b}}^*\right|\right|_2^2 \tag{33}$$

Using a convex optimisation algorithm $\hat{\mathbf{b}}_i$ is updated to $\hat{\mathbf{b}}_{i+1}$ decreasing the criterion

$$\min_{\alpha|\hat{\mathbf{B}}_{i+1}\in\hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})} \left|\left|\hat{\mathbf{B}}_{i+1}(\alpha) - \hat{\mathbf{b}}^*\right|\right|_2^2 < \min_{\alpha|\hat{\mathbf{B}}_i(\alpha)\in\hat{\mathcal{B}}(\hat{\mathbf{A}},\ \hat{\mathbf{c}})} \left|\left|\hat{\mathbf{B}}_i - \hat{\mathbf{b}}^*\right|\right|_2^2$$

This step is repeated for $i > 0$ until no further improvement is possible meaning the global minimum has been located at $\hat{\mathbf{B}}^*$ and the algorithm is illustrated in Figure (5).

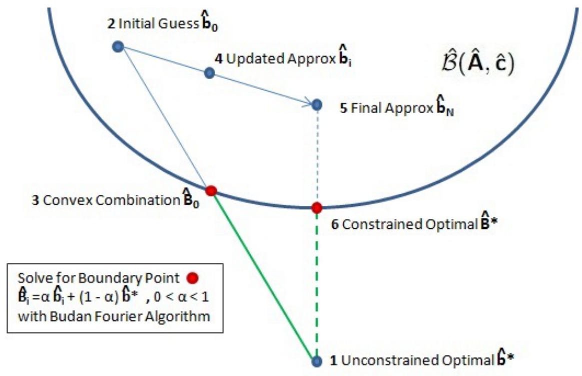

Fig. 5. Convex Optimisation Algorithm for Non-Negative EPT Density Function using Budan Fourier Technique

6 Finite Interval to Test for Non-Negativity in presence of Dominant Pole

Given a minimal realization with strictly dominant real pole $\lambda < 0$

$$\left(\begin{array}{cc|c} \lambda & 0 & 1 \\ 0 & \mathbf{A} & \mathbf{b} \\ \hline \mu & \mathbf{c} & 0 \end{array}\right) \tag{34}$$

The triple $(\mathbf{A}, \mathbf{b}^*, \mathbf{c})$ is the optimal triple satisfying the criterion in Eq. (35).

$$\min_{\mathbf{A},\mathbf{b},\mathbf{c}} ||h(x) - (\mathbf{c}e^{\mathbf{A}x}\mathbf{b} + \mu e^{\lambda x})||_2^2 \tag{35}$$

The non-empty set $\mathcal{B}_1(\mathbf{A}, \mathbf{c})$ is defined for a given pair $(\mathbf{A}, \mathbf{c})$ s.t.

$$\mathcal{B}_1(\mathbf{A}, \mathbf{c}) \;=\; \left\{ \mathbf{b} \;\middle|\; ||\mathbf{b}||_{\mathbf{Q}}^2 \;\leq\; 2||\mathbf{b}^*||_{\mathbf{Q}}^2 \right\} \tag{36}$$

Proposition1

For any minimal realization as described in Eq. (34) with strictly dominant real pole, λ and $\mathbf{b} \in \mathcal{B}_1(\mathbf{A}, \mathbf{c})$ we can construct a $T_0 > 0$ such that if $\mathbf{c}e^{\mathbf{A}x}\mathbf{b} \geq -\mu e^{\lambda x}$ for all $x \in [0, T_0]$ then $\mathbf{c}e^{\mathbf{A}x}\mathbf{b} \geq -\mu e^{\lambda x}$ for all $x \geq 0$.

Proof

The dominant pole constraint implies $\mathbb{R}(\sigma(\mathbf{A})) < \lambda$. Scaling the EPT density above by a factor $e^{-\lambda x}$ transforms the condition to be considered to

$$\mathbf{c}e^{(\mathbf{A}-\mathbf{I}\lambda)x}\mathbf{b}_C \geq -\mu \qquad , \qquad \forall\, x \;\geq\; 0 \tag{37}$$

$$\mathbf{c}e^{\tilde{\mathbf{A}}x}\mathbf{b}_C \geq -\mu \qquad , \qquad \forall\, x \;\geq\; 0$$

where $\tilde{\mathbf{A}} = (\mathbf{A} - \mathbf{I}\lambda)$ such that $\mathbb{R}(\sigma(\tilde{\mathbf{A}})) < 0$.
For any $\mathbf{b}_C \in \mathcal{B}_1(\mathbf{A}, \mathbf{c})$ we have

$$||\mathbf{c}e^{\tilde{\mathbf{A}}x}\mathbf{b}_C||_2^2 \;\leq\; 2||\mathbf{c}e^{\tilde{\mathbf{A}}x}\mathbf{b}^*||_2^2 \tag{38}$$

$\mathbf{Q}$ is defined as the positive definite observability grammian

$$\mathbf{Q} \;=\; \int_0^\infty e^{\tilde{\mathbf{A}}^T y}\, \mathbf{c}^T\, \mathbf{c}\, e^{\tilde{\mathbf{A}}y} dy \tag{39}$$

and L_2 norm is then given by

$$\mathbf{b}_C^T\, \mathbf{Q}\, \mathbf{b}_C \;=\; ||\mathbf{c}\, e^{\tilde{\mathbf{A}}x}\, \mathbf{b}_C||_2^2 \tag{40}$$

By letting $\lambda_{\min} = \min\{\lambda \in \sigma(\mathbf{Q})\}$ such that

$$\mathbf{b}_C^T\, \mathbf{Q}\, \mathbf{b}_C \;\geq\; \mathbf{b}_C^T\, \mathbf{b}_C\, \lambda_{\min} \;=\; ||\mathbf{b}_C||^2\lambda_{\min} \tag{41}$$

it can be seen that

$$||\mathbf{b}_C||^2\lambda_{\min} \;\leq\; \mathbf{b}_C^T\mathbf{Q}\, \mathbf{b}_C \;\leq\; 2R^2 \tag{42}$$

where $R^2 = \mathbf{b}^{*T}\mathbf{Q}\;\mathbf{b}^*$. Hence there is an obvious bound on $\mathbf{b}_C$

$$||\mathbf{b}_C||^2 \;\leq\; \frac{2R^2}{\lambda_{\min}} \tag{43}$$

As the pair $(\tilde{\mathbf{A}}, \mathbf{c})$ is observable and $\tilde{\mathbf{A}}$ an asymptotically stable matrix it must hold

$$\lim_{x\to\infty} \;\mathbf{c}\, e^{\tilde{\mathbf{A}}x} \;\to\; 0 \tag{44}$$

implying that $\forall\; \epsilon > 0 \;\exists\; T_0 > 0$ s.t. $\forall\; x > T_0$, $||\mathbf{c}e^{\tilde{\mathbf{A}}x}||^2 < \epsilon$.
To find such a T_0 we define the Lyapunov function

$$V(x) \;=\; \mathbf{c}\, e^{\tilde{\mathbf{A}}x}\, \mathbf{M}\, e^{\tilde{\mathbf{A}}^T x}\, \mathbf{c}^T \tag{45}$$

where $\mathbf{M}$ is the positive definite solution to the Lypanunov equation

$$\tilde{\mathbf{A}}\,\mathbf{M} \;+\; \mathbf{M}\,\tilde{\mathbf{A}}^T \;=\; -\mathbf{I} \tag{46}$$

$V(x)$ can be seen to be monotonically decreasing by examining its derivative

$$V'(x) \;=\; \mathbf{c}\, e^{\tilde{\mathbf{A}}x}\, (\tilde{\mathbf{A}}\,\mathbf{M} \;+\; \mathbf{M}\,\tilde{\mathbf{A}}^T)\, e^{\tilde{\mathbf{A}}^T x}\, \mathbf{c}^T \;<\; 0 \tag{47}$$

$\forall\; x > 0$. Letting $\tilde{\lambda}_{\min} = \min\{\lambda \in \sigma(\mathbf{M})\}$ it is clear

$$\tilde{\lambda}_{\min}||\mathbf{c}e^{\tilde{\mathbf{A}}x}||^2 \;\leq\; \mathbf{c}\, e^{\tilde{\mathbf{A}}x}\, \mathbf{M}\, e^{\tilde{\mathbf{A}}^T x}\, \mathbf{c}^T \tag{48}$$

$$||\mathbf{c}e^{\tilde{\mathbf{A}}x}||^2 \;\leq\; \frac{V(x)}{\tilde{\lambda}_{\min}}$$

We can then solve for T_0 such that

$$V(T_0) \;=\; \epsilon\tilde{\lambda}_{\min} \tag{49}$$

where

$$\epsilon \;=\; \frac{\mu}{\frac{2R^2}{\lambda_{\min}}} \;=\; \frac{\mu\lambda_{\min}}{2R^2} \tag{50}$$

proving Proposition 1.

Acknowledgements

The authors would like thank Fabien Seyfert for his invaluable contribution.

References

Hanzon, B., Sexton, C., *State Space Calculations for two-sided EPT Densities with Financial Modelling Applications.* Forthcoming. Draft version available at “www.edgeworth.biz”, 2011

Hanzon, B., Holland, F. *Non-Negativity of Exponential Polynomial Trigonometric Functions - A Budan Fourier Sequence Approach.* BFS Toronto Poster 429, June 22 - 26, 2010. Online version available at “www.edgeworth.biz”

Olivi, M. *Parametrization of Rational Lossless Matrices with Applications to Linear System Theory* HDR Thesis, 2010

Hanzon, B., Olivi, M., Peeters, R.L.M, *Balanced Realizations of Discrete-Time Stable All-Pass Systems and the Tangential Schur Algorithm* Linear Algebra and its Applications, 2006