

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

Stable numerical Laplace Transform inversion technique without over- and undershoot

Author:

Williams Zanga

Supervisor:

William J. Knottenbelt

Submitted in partial fulfilment of the requirements for the MSc Degree MSc in
Computing Science of Imperial College London

September 2018

Abstract

Numerical inversion of Laplace transform is known to be equivalent to approximating a shifted version of the Dirac impulse function with a linear combination of complex exponentials. From this knowledge, we construct a general framework to approximate that function with concentrated matrix exponential distributions, characterized by low coefficient of variation. That structure generalizes the method proposed by Horváth, Talyigás and Telek; and it guarantees numerical inversions without positive or negative overshoots. Optimization is done for a specific class of inversion methods within that framework, with a semi-deterministic algorithm based upon evolution strategy and gradient descent. This results in approximation errors evolving as $O(1/n^2)$. Finally, we propose an analytical method with error of type $O(1/n)$ to bypass optimization.

Acknowledgments

I would like to thank Prof. William J. Knottenbelt for his support and encouragement throughout this MSc dissertation. I also thank my family for their absolute patience, support and ongoing love and without whom this work would never have been possible.

Contents

1	Introduction	v
2	Background	2
2.1	Definition of Laplace transform	2
2.2	Basic properties	4
2.3	Examples	5
3	Inversion problem	7
3.1	Abate-Whitt framework	8
3.2	Background on probability distributions	9
3.3	Classical methods	12
3.3.1	Euler method	12
3.3.2	Gaver-Stehfest method	12
3.3.3	Talbot method	13
3.4	Horváth-Talyigás-Telek (HTT) method	14
3.4.1	Background on matrix exponential distributions	15
3.4.2	ME based inversion method	16
4	General matrix exponential inversion method	17
4.1	Principle	17
4.2	Linearization of the product form	18
4.3	Semi-frequency classes	19
4.3.1	Polynomials semi-frequency class	20
4.3.2	Monomials semi-frequency class	21
5	Monomials class	22
5.1	Properties of the semi-frequencies	22
5.2	Pseudo-periodicity	26
5.3	Reduced moments	27
5.3.1	Reduced moments theorems	28

5.4	Nodes	30
5.5	Calculation of the weights	32
5.5.1	An integral based method	32
5.5.2	Recursive method	33
5.6	Width of the distribution function	34
5.7	Reinterpretation of the HTT method	35
6	Unimodal method	37
6.1	Weights and nodes	37
6.2	Optimal parameters	38
7	Optimization for the monomials class	40
7.1	Squared coefficient of variation	40
7.2	Time derivatives of the distribution function	40
7.3	Optimization strategy	41
7.3.1	Gradient descent	41
7.3.2	Gradient of the squared coefficient of variation	42
7.3.3	Gradient of the peak of the distribution	43
7.3.4	Hybridization of (1+1)-ES and gradient descent	43
8	Numerical results	45
9	Conclusion	50

Chapter 1

Introduction

Laplace transform is an integral based transform widely applied in various areas of science. Within the (complex-valued) Laplace frequency domain, mathematical operations that are expensive in the (real-valued) time domain can become relatively simple. For example, differentiation involves a simple multiplication operation, integration a simple division operation and convolution of two functions involves finding the product of their respective Laplace transforms.

It is often desirable to return to the time domain through the inversion of a Laplace transform. This is often not possible to do analytically. In this context, numerical Laplace transform inversion is a useful tool; see for example its use in the response time analysis of concurrent systems [1]. There are various techniques for numerical inversion of Laplace transform, e.g. the Euler, Gaver-Stehfest, Talbot methods, but they all run into stability problems of various kinds, especially when inverting discontinuous functions.

In [2], J. Abate and W. Whitt introduce a unified framework to numerically invert Laplace transform. The general idea is to approximate the inverse Laplace transform with a finite linear combination of values of the transform. From a theoretical viewpoint, this framework is strictly equivalent to approximating a shifted version of the Dirac impulse function with a linear combination of complex exponentials. Using this equivalence, Horváth, Talyigás and Telek propose a numerical inversion method without positive or negative overshoots in [3], unlike the methods mentioned earlier. Their method consists in approximating the Dirac impulse function with matrix exponential distributions, characterized by minimal coefficient of variation.

In this thesis, we propose a natural framework to approximate the Dirac impulse function with strictly positive and concentrated matrix exponential distributions, as a generalization of the Horváth-Talyigás-Telek (HTT) method. Thus, inversion methods derived from that framework are free of positive and negative overshoots. Then a class of inversion methods with squared coefficient of variation of type $O(1/n^2)$ is specifically studied. Significant results include the proposition of an explicit and optimal approximation of Dirac impulse function with squared coefficient of variation evolving as $O(1/n)$. Other major result is the reduction of the length of the integral for the calculation of the moments of the approximation of the shifted Dirac impulse. This constitutes the idea behind the reduced moments triangulation (RMT) and the reduced moments convergence (RMC). Furthermore, we propose two methods to compute the coefficients used in the Abate-Whitt framework. The first method takes inspiration in the integral calculation of Fourier series coefficients. The second and faster method is based upon recursion. Finally, we design a simple and efficient algorithm, to minimize the coefficient of variation, based upon a combination of evolution strategy and gradient descent.

Organization

- (i) In Chapter 2, Laplace transform is formally defined, followed by an enumeration of its basic properties which are then illustrated with classical examples.
- (ii) The inversion problem is introduced in Chapter 3, followed by an examination of the Abate-Whitt framework. After that, we review classical inversion methods and present the HTT method.
- (iii) In Chapter 4, we introduce our framework, along with its properties, before defining approximation classes inherited from function spaces.
- (iv) In Chapter 5, the monomials semi-frequencies class is studied in depth. In particular, we formulate and prove the RMT and the RMC. Then, we derive two methods to calculate the coefficients for the Abate-Whitt framework. Finally, we propose an additional measure of concentration and reinterpret the HTT method.
- (v) In Chapter 6, we design an optimal and analytical inversion method for a subclass the monomials semi-frequency class.
- (vi) In chapter 7, we discuss optimization methods to minimize the coefficient of variation for the monomials class. The results are then evaluated in Chapter 8.

Chapter 2

Background

2.1 Definition of Laplace transform

The unilateral or one-sided Laplace transform (LT) is an integral operator which maps real-valued or complex-valued functions, defined on the positive real axis, into complex-valued functions, defined in a region of the complex plane.

Definition 2.1.1. *If the function $f : \mathbb{R}^+ \rightarrow \mathbb{K}$ is real-valued ($\mathbb{K} = \mathbb{R}$) or complex-valued ($\mathbb{K} = \mathbb{C}$), then its Laplace transform is the continuous summation of exponentially weighted values taken on the positive real axis and defined as follows*

$$\mathcal{L}\{f\}(s) = \int_0^{\infty} f(t)e^{-st}dt, \quad s \in \mathbb{C}$$

That definition is valid when the integral is convergent. In practice, it depends on the Laplace domain variable $s \in \mathbb{C}$ and the asymptotic behavior of the function to be transformed as the upper-bound of integration is not finite. Yet, a sufficient condition for convergence of the integral is that s verifies the inequality $\text{Real}(s) \geq \gamma$, where γ is a real number such that f is dominated by a positive multiple of the exponential function $t \mapsto e^{\gamma t}$, i.e

$$\exists A > 0, \quad |f(t)| \leq A \cdot e^{\gamma t} \quad \text{as } t \rightarrow \infty \quad (2.1)$$

If f verifies that condition, then it is said to be of exponential type and in such case, the Laplace transform is well-defined when s belongs to the set $\{z \in \mathbb{C} \mid \text{Real}(z) > \gamma\}$. That portion of the complex plane is commonly referred to as the s -plane or the region of convergence of Laplace transform. It is common to use $\mathcal{L}\{.\}$ to denote the Laplace transform operator and $F(s)$, $f^*(s)$ or $\mathcal{L}\{f(t)\}(s)$ to denote $\mathcal{L}\{f\}(s)$.

Calculation of a Laplace transform

We calculate the expression of the LT for the general power function $p_n : t \mapsto t^n$, where n is a positive integer. In this specific case, the convergence condition reduces to the inequality $\text{Re}(s) > 0$ because $\lim_{x \rightarrow \infty} x^n e^{-x} = 0$. We start by calculating the transform for the special case when $n = 0$; which means that $p_0(t) = 1$ for $t \in \mathbb{R}^+$.

$$\mathcal{L}\{1\}(s) = \int_0^\infty e^{-st} dt = \left[\frac{e^{-st}}{-s} \right]_0^\infty = \frac{1}{s}$$

We carry on by establishing a recursive formula linking $\mathcal{L}\{t^{n+1}\}$ and $\mathcal{L}\{t^n\}$

$$\begin{aligned} \mathcal{L}\{t^{n+1}\}(s) &= \int_0^\infty t^{n+1} \cdot e^{-st} dt \\ &= \left[t^{n+1} \cdot \frac{e^{-st}}{-s} \right]_0^\infty + \frac{n+1}{s} \int_0^\infty t^n \cdot e^{-st} dt \quad (\text{integration by parts}) \\ &= \frac{n+1}{s} \cdot \mathcal{L}\{t^n\}(s) \end{aligned}$$

Using that relation, we obtain the following results

$$\mathcal{L}\{t^n\}(s) = \left(\frac{n}{s}\right) \left(\frac{n-1}{s}\right) \cdots \left(\frac{1}{s}\right) \cdot \mathcal{L}\{t^0\}(s) = \frac{n!}{s^n} \cdot \mathcal{L}\{1\}(s) = \frac{n!}{s^{n+1}}$$

Table of common Laplace transforms

Time domain $f(t)$	Laplace domain $F(s)$	s-plane
$\delta(t)$ (Dirac impulse)	1	\mathbb{C}
$u(t)$ (Heaviside)	$\frac{1}{s}$	$\text{Re}(s) > 0$
$t^n, n \in \mathbb{N}$	$\frac{n!}{s^{n+1}}$	$\text{Re}(s) > 0$
$t^{1/n}, n \in \mathbb{N}$	$\frac{1}{s^{1+1/n}} \Gamma(1 + \frac{1}{n})$	$\text{Re}(s) > 0$
e^{-at}	$\frac{1}{s+a}$	$\text{Re}(s) > -a$
$\sin(\omega t)$	$\frac{\omega}{s^2 + \omega^2}$	$\text{Re}(s) > 0$
$\cos(\omega t)$	$\frac{s}{s^2 + \omega^2}$	$\text{Re}(s) > 0$
$\sinh(\alpha t)$	$\frac{\alpha}{s^2 - \alpha^2}$	$\text{Re}(s) > \alpha $
$\cosh(\alpha t)$	$\frac{s}{s^2 - \alpha^2}$	$\text{Re}(s) > \alpha $
$\ln(t)$	$-\frac{1}{s} (\ln(s) + \gamma)$	$\text{Re}(s) > 0$

Remark 1. In this table, $\Gamma(z)$ is Euler's Gamma function which verifies the relation $\Gamma(z+1) = z \cdot \Gamma(z)$ and γ is the Euler–Mascheroni constant.

Assumptions made in this work

1. f is defined for positive real numbers
2. f is a real-valued function
3. $\mathcal{L}\{f\}$ is well-defined

2.2 Basic properties

We continue with a non-exhaustive enumeration of basic properties of Laplace transform. These properties are often used for the resolution of linear differential equations and the analysis of the time response of dynamical systems.

Proposition 1. (*Linearity*) If a and b are complex numbers and the functions f and g are complex-valued, then the $\mathcal{L}\{.\}$ operator verifies the linearity identity

$$\mathcal{L}\{af + bg\} = a\mathcal{L}\{f\} + b\mathcal{L}\{g\}$$

Proposition 2. (*Time differentiation*) If f is n times differentiable with derivatives of exponential type, then the LT of its n -th derivative verifies the identity

$$\mathcal{L}\{f^{(n)}\}(s) = s^n F(s) - \sum_{k=1}^n s^{n-k} f^{(k-1)}(0^+)$$

Proposition 3. (*Time integration*) If f is continuous on \mathbb{R}^+ , then

$$\mathcal{L}\left\{\int_0^t f(\tau)d\tau\right\}(s) = \frac{1}{s} \cdot F(s)$$

With these properties, any linear differential equation can be turned into an algebraic equation because differentiation (resp. integration) in the time domain becomes multiplication (resp. division) by s in Laplace domain.

Proposition 4. (*Differentiation (resp. integration) in Laplace domain is equivalent to multiplication (resp. division) of the original function by the time domain variable.*

(i) *Frequency differentiation:* $\mathcal{L}\{t^n \cdot f(t)\}(s) = (-1)^n F^{(n)}(s), \quad n \in \mathbb{N}$

(ii) *Frequency integration:* $\mathcal{L}\left\{\frac{f(t)}{t}\right\}(s) = \int_s^\infty F(z) dz$

Proposition 5. *The $\mathcal{L}\{.\}$ operator associates shifts or delays with multiplication by exponential functions and scaling of variables is reversed.*

(i) *Time shift*: $\mathcal{L}\{f(t - a)\}(s) = e^{-as}F(s)$

(ii) *Frequency shift*: $\mathcal{L}\{e^{at} \cdot f(t)\}(s) = F(s - a)$

(iii) *Time scaling*: $\mathcal{L}\{f(at)\}(s) = \frac{1}{a}F\left(\frac{s}{a}\right), \quad a \neq 0$

Proposition 6. (*Multiplication*) $\mathcal{L}\{f \cdot g\}(s) = \frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{\sigma - iT}^{\sigma + iT} F(z)G(s - z) dz$ with $\text{Real}(z) = \sigma$ where σ is in the region of convergence of F .

An interpretation of Proposition 6 is that $\mathcal{L}\{f \cdot g\}$ is a special convolution of F and G , based upon the imaginary part of Laplace domain variable. As a reminder, if u and v are defined on the real axis, then their convolution is defined as the parametrized integral $(u * v)(t) = \int_{-\infty}^{\infty} u(\tau)v(t - \tau) d\tau$.

Proposition 7. (*Convolution*) $\mathcal{L}\{f * g\}(s) = F(s) \cdot G(s)$

Proposition 7 mirrors back Proposition 6 as the LT of a convolution is obtained by multiplying the individual transforms of the functions involved. As a result, the $\mathcal{L}\{\cdot\}$ operator establishes an equivalence between products and convolutions.

Proposition 8. (*Periodic function*) If f is a periodic function of period T , then $\mathcal{L}\{f\}$ has the reduced expression

$$\mathcal{L}\{f\}(s) = \frac{1}{1 - e^{-Ts}} \int_0^T e^{-st} f(t) dt$$

Theorem 1. (*Initial value Theorem*)

$$\lim_{t \rightarrow 0^+} f(t) = \lim_{s \rightarrow \infty} sF(s)$$

Theorem 2. (*Final value Theorem*) If all the poles (zeros of the denominator) of $sF(s)$ have strictly negative real parts, then

$$\lim_{t \rightarrow +\infty} f(t) = \lim_{s \rightarrow 0} sF(s)$$

2.3 Examples

In this section, we illustrate some properties of Laplace transform with classical examples. We begin with the resolution of a linear differential equation, before evaluating Dirichlet's integral.

Resolution of a linear differential equation

The aim of this example is to illustrate Proposition 1 and Proposition 2 with the first order differential equation : $y' + 5y = 0$ and $y(0) = 1$.

$$\begin{aligned}\mathcal{L}\{y' + 5y\} &= \mathcal{L}\{0\} \\ \mathcal{L}\{y'\} + 5\mathcal{L}\{y\} &= 0 && \text{(linearity)} \\ \left\{ s \cdot Y(s) - y(0) \right\} + 5Y(s) &= 0 && \text{(time differentiation)} \\ (s + 5) \cdot Y(s) - 1 &= 0 \\ Y(s) &= \frac{1}{s + 5}\end{aligned}$$

The inverse transform of $Y(s)$ is calculated, either with the time shift property or the table. Which gives the solution

$$y(t) = \mathcal{L}^{-1}\{Y\}(t) = \mathcal{L}^{-1}\left\{\frac{1}{s + 5}\right\}(t) = e^{-5t}$$

This solution is strictly identical to the one that we would obtain with the classical technique involving the characteristic polynomial of the differential equation.

Evaluation of the Dirichlet integral

We now evaluate the integral of the cardinal sine function $\text{sinc} : t \mapsto \frac{\sin(t)}{t}$ on the positive real axis i.e

$$\int_0^\infty \frac{\sin(t)}{t} dt$$

The first step consists in applying Proposition 4 to $f : t \mapsto \sin(t)$ as follows

$$\mathcal{L}\left\{\frac{f(t)}{t}\right\}(s) = \int_s^\infty F(z) dz = \int_s^\infty \frac{1}{z^2 + 1} dz = \frac{\pi}{2} - \arctan(s)$$

An alternative expression for the leftmost term is obtained with the original definition of Laplace transform

$$\mathcal{L}\left\{\frac{f(t)}{t}\right\}(s) = \mathcal{L}\left\{\frac{\sin(t)}{t}\right\}(s) = \int_0^\infty \frac{\sin(t)}{t} e^{-st} dt$$

Finally, we take the limit of these expressions as the Laplace domain variable approaches zero; which gives the value of the Dirichlet integral

$$\int_0^\infty \frac{\sin(t)}{t} dt = \frac{\pi}{2}$$

Chapter 3

Inversion problem

As mentioned in Section 2.2, Laplace transform facilitates the analysis of dynamical systems; specially with Proposition 1, Proposition 2, Theorem 1 and Theorem 2. However, some transforms must be inverted in order to recover all the information contained in the time domain signal. For instance, the resolution of the first order differential equation in Section 2.1 required to inverse a Laplace transform. That inversion problem can be formally stated as follows

Problem 1. *Given a function $s \mapsto F(s)$ in the Laplace domain, evaluate the inverse Laplace transform $t \mapsto \mathcal{L}^{-1}\{F\}(t)$ for any real argument $t \geq 0$.*

Except for the special cases where F corresponds to the transform of a known function, there is not general closed form expression for the inverse transform. As discussed in [4, Chapter 3], various inversion techniques rely on expansions or approximation of F with power series or orthogonal polynomials for which the inverse Laplace transforms are easier to determine. Although no explicit expression for the inverse transform is known, there exists two analytical inversion formulas to calculate $\mathcal{L}^{-1}\{F\}$. The first one uses a complex integral known as Bromwich integral, Fourier-Mellin integral or Mellin's inverse formula and the second one relies on the calculation of a limit known as Post-Widder formula.

Theorem 3. (*Bromwich integral*)

$$f(t) = \frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{\gamma - iT}^{\gamma + iT} F(s) e^{st} ds$$

γ is a real number such that the contour path of integration is in the region of convergence of F .

Theorem 4. (*Post-Widder formula*)

$$f(t) = \lim_{n \rightarrow \infty} \frac{(-1)^n}{n!} \left(\frac{n}{t}\right)^{n+1} F^{(n)}\left(\frac{n}{t}\right)$$

Formal proofs for these inversion formulas can be found in [4, Section 2.2, Section 2.3]. Even though these formulas constitute systematic ways to calculate the inverse LT of a function, their effective use often requires heavy calculations which can rarely be performed by hand on a reasonable amount of time. For that reason, practical applications of these techniques are numerical. However, the complexity of hand-made calculations is turned into numerical precision and stability problems.

3.1 Abate-Whitt framework

Multiple approaches used to estimate the inverse LT were examined by J. Abate, W. Whitt et al. They notably introduced a unified framework to construct numerical LT inversion methods in [2]. As discussed in Section 3.3, classical inversion methods (Euler, Gaver-Stehfest, Talbot) can be rewritten in this framework.

Definition 3.1.1. (*Abate-Whitt framework*) *The inverse transform $\mathcal{L}^{-1}\{F\}$ or f is approximated by a finite linear combination of values of F as follows*

$$f(t) \approx f_n(t) = \frac{1}{t} \sum_{k=1}^n \eta_k F\left(\frac{\beta_k}{t}\right), \quad t \geq 0$$

The nodes β_k and the weights η_k are complex numbers which depend neither on the transform F nor on the time argument t but only on the order of approximation n .

The independence of the weights and nodes, *or simply Abate-Whitt coefficients*, from both the time argument and the transform is a key requirement in order for this framework to be applicable to multiple functions at various time points. If the inverse transform is real-valued, then it is approximated by the real part of that sum, i.e

$$\mathbf{Real}\{f(t)\} \approx \mathbf{Real}\{f_n(t)\} = \frac{1}{t} \sum_{k=1}^n \mathbf{Real}\left\{\eta_k F\left(\frac{\beta_k}{t}\right)\right\} \quad (3.1)$$

Proposition 9. (*Integral interpretation*) *When the nodes verify $\mathbf{Real}\{\beta_k\} \geq 0$, the Abate-Whitt framework is equivalent to approximating of the shifted-scaled Dirac impulse $\delta(x/t - 1)$ with a finite linear combination of exponential functions i.e*

$$\delta(x/t - 1) \approx \delta_n(x/t - 1) = \frac{1}{t} \sum_{k=1}^n \eta_k e^{-\beta_k \frac{x}{t}}$$

Remark 2. The weights η_k and the nodes β_k are identical to those in Definition 3.1.1

Remark 3. To simplify the notations, the approximation of the shifted-scaled Dirac impulse $\delta_n(x/t - 1)$ is also noted $\delta_n^t(x)$

Proof.

$$\begin{aligned}
f(t) \approx f_n(t) &= \frac{1}{t} \sum_{k=1}^n \eta_k F\left(\frac{\beta_k}{t}\right) \\
&= \frac{1}{t} \sum_{k=1}^n \eta_k \left\{ \int_0^\infty f(x) \cdot e^{-\frac{\beta_k}{t}x} dx \right\} \\
&= \int_0^\infty f(x) \left\{ \frac{1}{t} \sum_{k=1}^n \eta_k e^{-\frac{\beta_k}{t}x} \right\} dx \\
&= \int_0^\infty f(x) \cdot \delta_n^t(x) dx
\end{aligned}$$

□

The inversion is perfect when $x \mapsto \delta_n^t(x)$ is exactly the Dirac impulse function about time t . But as stated in [3, Section 3], the accuracy of this approximation depends on the order n and the Abate-Whitt coefficients. Generally, the exactness of the approximation gets better when the order increases. Finally, we can remark that $\delta_n^t(x)$ verifies a scaling relation which is similar to that of Proposition 5

$$\delta_n^t(x) = \frac{1}{t} \cdot \delta_n^1\left(\frac{x}{t}\right) \quad (3.2)$$

and which leads to a simplified interpretation of the Abate-Whitt framework.

Theorem 5. The Abate-Whitt framework is equivalent to approximating of the shifted Dirac impulse about the point $t=1$ with a finite linear combination of exponentials i.e

$$\delta_n^1(x) = \sum_{k=1}^n \eta_k e^{-\beta_k x}$$

3.2 Background on probability distributions

Originally defined in [5, p.58], Dirac impulse function (equally known as Dirac distribution, delta distribution or δ -distribution) represents the space density of a particle as a function which is null everywhere except for the argument zero and with integral over the real numbers equal to one. In principle, that distribution is similar to probability density functions (pdf).

Definition 3.2.1. In probability theory, a function f defined in a given set I is a probability density function if it satisfies the conditions

$$(i) \quad \forall t \in I \quad 0 \leq f(t) \leq 1$$

$$(ii) \quad \int_{t \in I} f(t) dt = 1$$

That proximity between Dirac distribution and pdf is legitimized by the fact that $x \mapsto \delta(x)$ is the limit of various density functions. For instance, the uniform distribution $\mathcal{U}([0; 1/n])$ converges toward Dirac distribution as n grows larger. Similarly, the normal distribution $\mathcal{N}(0, \sigma^2)$ converges toward Dirac distribution as σ gets smaller.

Dispersion metrics

We now move on to discuss some measures used in Statistics and Probability Theory in order to study a data set or the behavior of a random variable. These metrics are averaged quantities over all the distribution and called moments.

Definition 3.2.2. The n -th order moment about the point $c \in I_x$ for a random variable X with a probability density function f_x is defined as the following integral

$$E[(X - c)^n] = \int_{t \in I_x} (t - c)^n f_x(t) dt$$

The moments can be given a physical interpretation when f_x represents the repartition of mass within a body. In particular, the 0^{th} order moment represents the total mass, the 1^{st} order moment represents the center of mass and the 2^{nd} order moment represents the moment of inertia around an axis. The 1^{st} order moment $E[X]$ is known as the average or mean value of the distribution. In the special cases when $c = E[X]$, the n -th order moment is called *central moment* and serves as a dispersion measure. The 2^{nd} order central moment, called *variance*, measures how far a set of random numbers are from their average value. The 3^{rd} order central moment, known as the *skewness*, measures how evenly spread is a set of random numbers around their average value. Finally, the 4^{th} order central moment, referred to as the *kurtosis*, measures the sharpness of the distribution in the vicinity of the average value.

Definition 3.2.3. The variance of a random variable is defined as

$$var(X) = E[(X - \mu)^2] = \int_{t \in I_x} (t - \mu)^2 f_x(t) dt$$

where $\mu = E[X]$ is the average value of the distribution. The standard deviation σ is then defined as the square root of the variance i.e $\sigma = \sqrt{\text{var}(X)}$

Definition 3.2.4. The squared coefficient of variation $\text{scv}(X)$ or $\text{scv}(f_x)$ of a random variable X is defined as the variance normalized by the average value

$$\text{scv}(X) = E \left[\left(\frac{X - \mu}{\mu} \right)^2 \right]$$

Definition 3.2.5. Pearson's moment coefficient of skewness $\gamma(X)$ or $\gamma(f_x)$ is defined as the 3rd central moment normalized by the standard deviation

$$\gamma(X) = E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right]$$

Definition 3.2.6. The kurtosis $\text{Kurt}(X)$ or $\text{Kurt}(f_x)$ is defined as the 4th central moment normalized by the standard deviation

$$\text{Kurt}[X] = E \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right]$$

The normalization process corresponds to nondimensionalization; such that those measures can be compared for two random variables with different units. Moving forward, we can mention another commonly used dispersion metric, the differential entropy.

Definition 3.2.7. The differential entropy $h(X)$ or $h(f_x)$ is defined as follows

$$h(X) = - \int_{t \in I_X} f_x(t) \log f_x(t) dt.$$

This quantity measures equipartition in the distribution. On one side, the uniform distribution is known to be the probability distribution with maximum entropy; which is due to the fact that its pdf is constant on I_X . On the other side, Dirac distribution is an example of distributions that minimize the differential entropy; which evaluates to $-\infty$ in this case. This result is explained by the concentration of all the information at a single point; to such extent that uncertainty about the state of the random variable vanishes. A standard dispersion coefficient based upon the differential entropy and similar to the squared coefficient of variation is discussed in [6].

3.3 Classical methods

3.3.1 Euler method

The Euler method is based upon the transformation of the Bromwich integral into a Fourier transform, thereafter approximated by a series via trapezoidal discretization. The Euler summation is then applied to accelerate the convergence of that series. The complete procedure produces an approximation of the inverted transform as follows

$$f(t) \approx \frac{e^{A/2}}{2t} \cdot \mathbf{Real} \left\{ F \left(\frac{A}{2t} \right) \right\} + \frac{e^{A/2}}{2t} \cdot \sum_{k=1}^n (-1)^k \mathbf{Real} \left\{ F \left(\frac{A + 2ik\pi}{2t} \right) \right\}$$

The derivation of the above formula is fully developed in [7] and the general expressions for the Abate-Whitt coefficients are given in [2, p.16]. We recall them here

(Nodes and weights for odd n)

$$\begin{aligned} \beta_k &= \frac{n-1}{6} \cdot \ln(10) + i\pi(k-1) & 1 \leq k \leq n \\ \eta_k &= (-1)^k \cdot 10^{(n-1)/6} \cdot \xi_k & 1 \leq k \leq n \end{aligned}$$

$$\begin{aligned} \text{where} \quad \xi_1 &= \frac{1}{2} \\ \xi_n &= 2^{-(n-1)/2} \\ \xi_k &= 1 & 1 \leq k \leq (n+1)/2 \\ \xi_{n-k} &= \xi_{n-k+1} + \xi_n \cdot \binom{(n-1)/2}{2} & 1 \leq k \leq (n-1)/2 \end{aligned}$$

3.3.2 Gaver-Stehfest method

The original Gaver method [8] relies on a sampling of the transform about the real axis and an approximation of the exponential function in the Bromwich integral by a rational function. Cauchy integral formula then gives the inverse transform in the Abate-Whitt framework. Stehfest proposed an acceleration of that method with Salzer's accelerating scheme for infinite series. The Abate-Whitt coefficients are rewritten here

(Nodes and weights for even n)

$$\beta_k = k \cdot \ln(2), \quad 1 \leq k \leq n$$

$$\eta_k = \ln(2) \cdot (-1)^{n/2+k} \sum_{j=\lfloor (k+1)/2 \rfloor}^{\min(k, n/2)} \frac{j^{n/2+1}}{(n/2)!} \binom{n/2}{j} \binom{2j}{j} \binom{j}{k-j} \quad 1 \leq k \leq n$$

3.3.3 Talbot method

The Talbot method is based upon a deformation of the Bromwich contour into an open contour on the negative real axis side. The expressions for the Abate-Whitt coefficients are given in [2, p.17] and recalled immediately

(Nodes and weights for all n)

$$\begin{aligned} \beta_1 &= \frac{2n}{5} \\ \beta_k &= \beta_1 \theta_k \left(\cot(\theta_k) + i \right), \quad 2 \leq k \leq n \\ \eta_1 &= \frac{1}{5} e^{\beta_1} \\ \eta_k &= \frac{2}{5} \left[1 + i \theta_k \left(1 + \cot(\theta_k)^2 \right) - i \cot(\theta_k) \right] e^{\beta_k}, \quad 1 \leq k \leq n \\ \theta_k &= \frac{(k-1)\pi}{n} \end{aligned}$$

Numerical approximations of the shifted Dirac impulse function

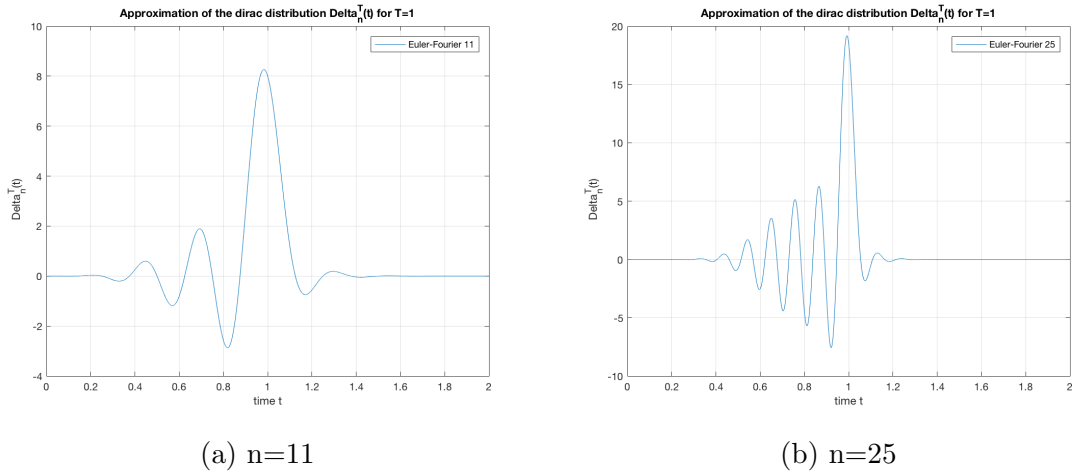
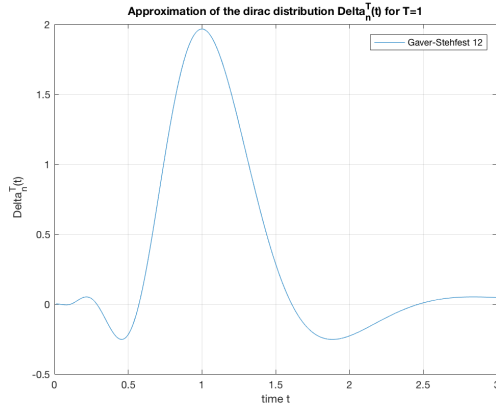
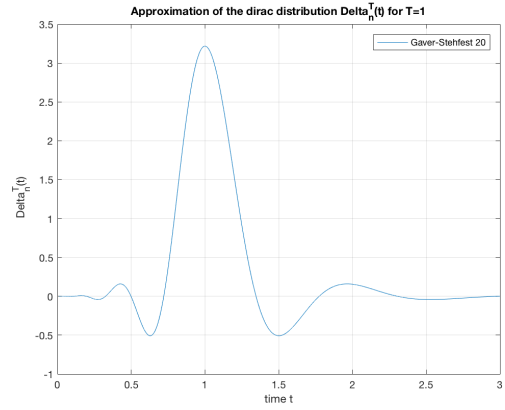


Figure 3.1: Approximation of the shifted Dirac impulse with the Euler method



(a) $n=12$



(b) $n=20$

Figure 3.2: Approximation of the shifted Dirac impulse with the Gaver-Stehfest method

Figures (3.1) and (3.2) illustrate the approximation of the shifted Dirac impulse with Euler and Gaver-Stehfest methods. Those approximations take negative values while Dirac impulse is strictly positive. As a consequence, positive and negative overshoots can be expected on the inverse Laplace transforms which use those methods. Furthermore, the factorials and the binomial coefficients in the expression of the weights can cause numerical instability for high orders. This problem is avoidable for the Euler method by using the inversion formula in Section 3.3.1. For the Gaver-Stehfest method, stability can be enhanced with the $\log \Gamma$ or GammaLn function which verifies

$$\forall x \in \mathbb{R} \quad x! = \Gamma(x+1) = e^{\text{GammaLn}(x+1)}$$

An ultimate solution to stabilize these methods consists in numerically increasing the floating-point precision in the calculations of the factorials and the binomial coefficients; which can be done with a multi-precision software.

3.4 Horváth-Talyigás-Telek (HTT) method

In [3], Horváth, Talyigás and Telek propose a method to numerically invert Laplace Transform without positive or negative overshoots. That method approximates the shifted Dirac impulse function by means of concentrated matrix exponential (ME) distributions which can be expressed as linear combination of complex exponentials.

3.4.1 Background on matrix exponential distributions

Definition 3.4.1. *In probability theory, matrix exponential distributions have probability density functions of the form*

$$f_{ME}(t) = -\boldsymbol{\alpha} A e^{tA} \mathbf{1}, \quad t \geq 0$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{1 \times n}$, $A \in \mathbb{R}^{n \times n}$ and $\mathbf{1} \in \mathbb{R}^{n \times 1}$ is a column vector with one as coefficients.

A random variable with pdf of that form is said to be $ME(\boldsymbol{\alpha}, A)$ -distributed; and as discussed in e.g [9, Section 2.1], a given ME distribution can be represented by multiple pairs, e.g $(\boldsymbol{\alpha}_1, A_1)$ and $(\boldsymbol{\alpha}_2, A_2)$ where A_1 and A_2 have different size.

Definition 3.4.2. *The class $ME(n)$ contains matrix exponential distributions which have a representation of order at most n .*

Definition 3.4.3. *If a random variable X is $ME(\boldsymbol{\alpha}, A)$ -distributed, such that $\boldsymbol{\alpha}$ and A satisfy the following assumptions*

1. $\alpha_i \geq 0$
2. $A_{i,i} \leq 0$
3. $A_{i,j} \geq 0$ for $i \neq j$
4. *All the components of the vector $A\mathbf{1}$ are negative*

then X is said to be phase type (PH) distributed or $PH(\boldsymbol{\alpha}, A)$ -distributed.

Proposition 10. *The probability density function of $ME(\boldsymbol{\alpha}, A)$ distributions have the following Jordan decomposition*

$$f_{ME}(t) = \sum_{i=1}^k \sum_{j=0}^{mult(\lambda_i)-1} c_{i,j} t^j e^{\lambda_i t}$$

where $\lambda_1, \dots, \lambda_k$ with $k \leq n$ are the eigenvalues of A , $mult(\lambda_i)$ is the multiplicity of λ_i and $c_{i,j}$ are complex numbers.

If any eigenvalue of A has multiplicity equal to one, then the Jordan decomposition transforms f_{ME} into a linear combination of complex exponential function; which means that concentrated matrix exponential distributions can be used for numerical inversion of Laplace transform. The most commonly used measure for concentration is the squared coefficient of variation. The minimum *scv* is known analytically for phase-type distributions with the following result, proven by Aldous and Shepp in [10].

Theorem 6.

$$\min_{X \in PH(N)} scv(X) = \frac{1}{N}$$

and the equality is obtained for the Erlang distribution of parameter (N, λ) with $\lambda > 0$.

3.4.2 ME based inversion method

Back to the inversion problem, Horváth, Talyigás and Telek [3] propose to approximate the shifted Dirac impulse with matrix exponential distributions as follows

$$\delta_{ME}(t) = c \cdot e^{-\lambda t} \prod_{i=1}^{(N-1)/2} \cos^2(\omega t - \phi_i) = \eta_1 e^{-\beta_1 t} + 2 \sum_{k=2}^n \text{Real} \left\{ \eta_k e^{-\beta_k t} \right\} \quad (3.3)$$

where N is an integer and $c, \lambda, \omega, \phi_1, \phi_2, \dots, \phi_{(n-1)/2}$ are real numbers, optimized to minimize the squared coefficient of variation. The optimization is done with a standard heuristic search algorithm [11, Section 2.8]. Further details about that optimization process and a list of optimized coefficients are given in [12, Section 3.3] and [13, Section 3]. In this method, the nodes β_k for $k = 1, \dots, n$ share the same real part λ but have

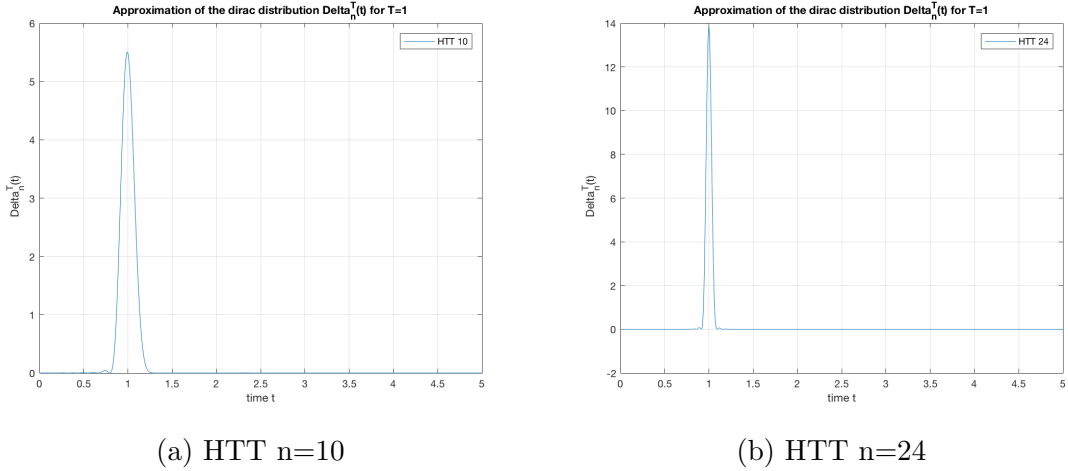


Figure 3.3: Approximation of Dirac distribution with the HHT method

different imaginary parts, forming an arithmetic sequence. Figure (3.3) depicts the resulting distribution and unlike those of the Euler and the Gaver-Stehfest methods, this function only takes positive values; which prevents positive or negative overshoots in the inverted transform. Further numerical optimization carried in [13, Section 3.4] suggests that the convergence of the squared coefficient of variation is quadratic and follows the asymptotic law

$$scv(X) \sim \frac{2}{N^2}$$

Chapter 4

General matrix exponential inversion method

4.1 Principle

We now introduce a general framework to approximate the shifted Dirac impulse about the abscissa $t = 1$ with matrix exponential distributions; as a generalization of the Horváth-Talyigás-Telek method. Our work is focused on matrix exponential distributions expressed as products of trigonometric functions.

Definition 4.1.1. (π -form) *The product form or π -form of the distribution function is defined as follows*

$$\delta_N(t) = e^{-\lambda t} \prod_{k=1}^N \cos^{2\alpha}(w_k t + \phi_k)$$

where N and α are positive integers; the decay coefficient λ and the modal frequencies w_1, w_2, \dots, w_N are positive real numbers and the modal phases $\phi_1, \phi_2, \dots, \phi_N$ are real numbers. The terms $\cos(w_k t + \phi_k)$ in the product are referred to as the π -modes.

The π -form can be interpreted as a product-wise superposition of trigonometric functions. The global exponent 2α is intended to make the resulting function positive, as for the HTT method. We carry on by noticing that this product can be transformed into a sum of other trigonometric functions. To illustrate that property, we examine a classic trigonometric formula which relates a product and a sum of cosine functions

$$2 \cdot \cos(x) \cdot \cos(y) = \cos(x + y) + \cos(x - y) \quad (4.1)$$

Such transformation remains feasible when the product includes powers of trigonometric functions $\cos^2(x), \cos^3(x), \dots, \cos^n(x)$ $n \in \mathbb{N}$ since we can apply Equation (4.1) repeatedly to reconstruct the individual powers before considering the global product.

4.2 Linearization of the product form

A consequence of Equation (4.1) is that the π -form can be rewritten as a sum of trigonometric functions. We choose to express this alternative form for δ_N as a linear combination of complex exponentials in order to remain within the boundaries of the Abate-Whitt framework. The linearization process is summarized immediately.

Theorem 7. (*Euler's identity*) Let $x \in \mathbb{R}$ and i such that $i^2 = -1$

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}$$

Theorem 8. (*Binomial theorem*) Let $n \geq 0$ be a integer and $(x, y) \in \mathbb{C}^2$.

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

Step 1 : Each π -mode is transformed into a sum of conjugated complex exponentials via Euler's identity

$$\delta_N(t) = e^{-\lambda t} \prod_{p=1}^N \left(\frac{e^{i(w_p t + \phi_p)} + e^{-i(w_p t + \phi_p)}}{2} \right)^{2\alpha} \quad (4.2)$$

Step 2 : The Binomial Theorem is then applied to each π -mode

$$\delta_N(t) = \frac{e^{-\lambda t}}{4^{\alpha N}} \prod_{p=1}^N \left[\sum_{k_p=0}^{2\alpha} \binom{2\alpha}{k_p} e^{2i(k_p - \alpha)(w_p t + \phi_p)} \right] \quad (4.3)$$

Step 3 : The product is expanded and the binomial coefficients are separated from the complex exponentials

$$\delta_N(t) = \frac{e^{-\lambda t}}{4^{\alpha N}} \sum_{k_1, k_2, \dots, k_N=0}^{2\alpha} \left[\prod_{p=1}^N \binom{2\alpha}{k_p} \right] \left[\prod_{p=1}^N e^{2i(k_p - \alpha)(w_p t + \phi_p)} \right] \quad (4.4)$$

Step 4 : The product of complex exponentials is transformed into a single complex exponential, then the modal frequencies and the modal phases are separated

$$\delta_N(t) = \frac{e^{-\lambda t}}{4^{\alpha N}} \sum_{k_1, k_2, \dots, k_N=0}^{2\alpha} \left[\prod_{p=1}^N \binom{2\alpha}{k_p} \right] e^{2i \left[t \sum_{p=1}^N (k_p - \alpha) w_p + \sum_{p=1}^N (k_p - \alpha) \phi_p \right]} \quad (4.5)$$

Equation (4.5) depicts the expected linear combination of complex exponentials that we now simplify by introducing the following quantities

- (a) The *combinatorial set*: $\mathcal{K}_\alpha = \{0, 1, \dots, 2\alpha\}^N$
- (b) The *combinatorial vector*: $\mathbf{k} = (k_1, k_2, \dots, k_N)^\top \in \mathcal{K}_\alpha$
- (c) The *combinatorial factor*: $\begin{bmatrix} 2\alpha \\ \mathbf{k} \end{bmatrix} = \frac{1}{4^{\alpha N}} \prod_{p=1}^N \binom{2\alpha}{k_p}$
- (d) The *semi-frequency*: $W(\mathbf{k}) = \sum_{p=1}^N (k_p - \alpha)w_p$
- (e) The *semi-phase*: $\Phi(\mathbf{k}) = \sum_{p=1}^N (k_p - \alpha)\phi_p$

The linearized π -form is then given a more compact expression that we call: sum form.

Definition 4.2.1. (σ -form I) *The first sum form or σ -form I of the distribution function is defined as follows*

$$\delta_N(t) = e^{-\lambda t} \sum_{\mathbf{k} \in \mathcal{K}_\alpha} \begin{bmatrix} 2\alpha \\ \mathbf{k} \end{bmatrix} e^{2i(t \cdot W(\mathbf{k}) + \Phi(\mathbf{k}))}$$

The derivation of the σ -form I proves that the π -form can be rewritten in a form which is compatible the Abate-Whitt framework and constitutes the starting point for further analysis. We can remark that the nodes are entirely determined by the decay coefficient and the semi-frequencies, while the weights are determined by the combinatorial factors and the semi-phases.

4.3 Semi-frequency classes

In virtue of Proposition 10, semi-frequencies constitute the imaginary parts of the eigenvalues of the matrix associated to these ME-distributions. Then as indicated by their definition, the semi-frequencies are completely determined by the modal frequencies w_1, w_2, \dots, w_N . Theoretically, there are infinitely many expressions for the modal frequencies; therefore, we consider the general case when the terms w_p for $p = 1, \dots, N$ are generated by a function that we call: modal frequency generator.

Definition 4.3.1. *The modal frequency generator is defined as follows*

$$\begin{aligned} \psi_0: \mathbb{N} &\rightarrow \mathbb{R} \\ p &\mapsto w_p \end{aligned}$$

With that function defined, we now introduce a class which extends the concept of semi-frequencies $W(\mathbf{k}) = \sum_{p=1}^N (k_p - \alpha)w_p$ and constitutes a starting point for the analysis of the properties that they inherit from modal frequencies.

Definition 4.3.2. *Semi-frequency class (SFC)*

$$\Omega^{(N,\alpha)} = \left\{ \sum_{p=1}^N (k_p - \alpha)\psi_0(p) \mid \psi_0 \in \mathcal{F}(\mathbb{N}, \mathbb{R}), \mathbf{k} \in \{0, 1, \dots, 2\alpha\}^N \right\}$$

4.3.1 Polynomials semi-frequency class

We can further observe that standtard functions can be approximated by polynomials; either with Taylor's theorem for differentiable functions or with Lagrange polynomial interpolation theorem.

Theorem 9. (*Taylor's theorem*) *Let $n \geq 1$ be an integer and let the function $f: \mathbb{R} \rightarrow \mathbb{R}$ be n times differentiable at the point $a \in \mathbb{R}$, then there exists a function $h_n: \mathbb{R} \rightarrow \mathbb{R}$ such that*

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k + h_n(x) \cdot (x - a)^n$$

Theorem 10. (*Lagrange polynomial interpolation theorem*)

Let $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ be data points such that no two x_j are identical. Then the polynomial

$$\mathcal{P}(x) = \sum_{k=1}^n y_k \left[\prod_{j=1, j \neq k}^n \frac{x - x_j}{x_k - x_j} \right]$$

is the unique polynomial of degree n that satisfies $\mathcal{P}(x_k) = y_k$ for $k = 1, 2, \dots, n$

Taylor's theorem is essential for MacLaurin Series and power series. It is particularly efficient for functions which have simple expressions for their derivatives; such as exponential functions $\{D^k e^{\alpha x} = \alpha^k e^{\alpha x}\}$, power functions $\{D^k x^n = \frac{n!}{(n-k)!} x^{n-k}\}$ or trigonometric functions $\{D^k \cos(x) = \cos(x + \frac{k\pi}{2})\}$. However, not every function is differentiable or has simple derivatives; so the Lagrange polynomial interpolation theorem provides an alternative to Taylor's theorem since it only requires values of the function. Although applicable for any function, the interpolation can diverge when not enough values of the function are given. These theorems give rise to the introduction of the subclass of SFC in which the modal frequency generator is a polynomial.

Definition 4.3.3. *Polynomials semi-frequency class (P-SFC)*

$$\Theta_r^{(N,\alpha)} = \left\{ \sum_{p=1}^N (k_p - \alpha) \psi_0(p) \mid \psi_0(X) \in \mathbb{R}_r[X], \mathbf{k} \in \{0, 1, \dots, 2\alpha\}^N \right\}$$

In this definition for the polynomials semi-frequency class (*or simply polynomials class*), α , N and the class index r are positive integers. Furthermore, $\mathbb{R}_r[X]$ stands for the set of polynomials of degree at most r and with real coefficients; which implies that $\psi_0(x) = a_r x^r + a_{r-1} x^{r-1} + \dots + a_1 x + a_0$ where $(a_0, a_1, \dots, a_r) \in \mathbb{R}^{1 \times (r+1)}$. This class is by construction a subclass of $\Omega^{(N,\alpha)}$ since $\mathbb{R}_r[X]$ is a subset of $\mathcal{F}(N, \mathbb{R})$. Moreover, the sets of polynomials $\mathbb{R}_r[X]$ form an increasing family of sets because any polynomial of degree r can be seen as polynomial of degree $r+1$ with zero as the coefficient of X^{r+1} . The polynomials semi-frequency class inherits that property, which can be summarized with the following inclusions

$$\Theta_0^{(N,\alpha)} \subset \Theta_1^{(N,\alpha)} \subset \Theta_2^{(N,\alpha)} \subset \dots \subset \Theta_\infty^{(N,\alpha)}$$

4.3.2 Monomials semi-frequency class

The polynomials semi-frequency class can be further reduced by observing that any polynomial is a linear combination of monomials. As a result, we can define the subclass of $\Theta_r^{(N,\alpha)}$ in which the modal frequency generator is a monomial.

Definition 4.3.4. *Monomials semi-frequency class (M-SFC)*

$$\theta_r^{(N,\alpha)} = \left\{ \omega \sum_{p=1}^N (k_p - \alpha) p^r \mid \omega \in \mathbb{R}, \mathbf{k} \in \{0, 1, \dots, 2\alpha\}^N \right\}$$

Again by construction, the monomials semi-frequency class (*or simply monomials class*) verifies the inclusion $\theta_r^{(N,\alpha)} \subset \Theta_r^{(N,\alpha)}$ and constitutes the main area of investigation in the next chapters. It is already remarkable that the HTT method is in this particular class when $r = 0$. A complete reinterpretation of that method is done in Section 5.7.

Chapter 5

Monomials class

This chapter focuses on the properties of distribution functions in the monomials semi-frequency class. First, we discuss main properties of that family of matrix exponential distributions. Then, we propose a general method to calculate the Abate-Whitt coefficients. Finally, we give a new interpretation of the HTT method. In the general case of the monomials class, the π -form has the expression

$$\delta_N(t) = ce^{-\lambda t} \prod_{k=1}^N \cos^{2\alpha}(\omega k^r t + \phi_k) \quad (5.1)$$

where N and α are positive integers, $r \in \mathbb{N}$ is the class index, $\lambda \in \mathbb{R}^+$ is the decay coefficient, $\omega \in \mathbb{R}^{+*}$ is the main frequency and $(\phi_1, \phi_2, \dots, \phi_N) \in \mathbb{R}^N$ are the modal phases and $c \in \mathbb{R}^{+*}$ is a normalization constant.

5.1 Properties of the semi-frequencies

In order to study their properties, we introduce a function which generates the semi-frequencies in the complex exponentials of the σ -form corresponding to the π -form from Equation (5.1). That function is defined on $\mathcal{K}_\alpha = \{0, 1, \dots, 2\alpha\}^N$ and referred to as the semi-frequency generator.

Definition 5.1.1. *The semi-frequency generator is defined as follows*

$$\begin{aligned} \psi_1 : \{0, 1, \dots, 2\alpha\}^N &\rightarrow \mathbb{R} \\ \mathbf{k} &\mapsto \omega \sum_{p=1}^N (k_p - \alpha) p^r \end{aligned}$$

Remark 4. *From now on, r, α and ω are considered to be fixed parameters.*

Proposition 11. *The semi-frequency generator can be rewritten as the inner product*

$$\psi_1(\mathbf{k}) = \omega \langle \mathbf{k} - \alpha \mathbf{1}, \mathbf{p}^r \rangle$$

where \mathbf{k} is a vector in the combinational set \mathcal{K}_α , $\mathbf{p}^r = (1, 2^r, \dots, N^r)^\top$ and $\mathbf{1}$ are vectors in $\mathbb{R}^{N \times 1}$, and $\langle \cdot, \cdot \rangle$ is the canonical inner product of $\mathbb{R}^{N \times 1}$.

This result is straightforward and is due to linearity of sums. A particular consequence of that alternative form is that $\psi_1(\alpha \mathbf{1}) = 0$. Such identity means that $\alpha \mathbf{1}$ is mapped into the zero semi-frequency, which accounts for a constant term in the σ -form. That constant corresponds the mean value of δ_N as shown in Section 5.5.1.

Proposition 12. *(Anti-reflection) The semi-frequency generator verifies the identity*

$$\psi_1(2\alpha \mathbf{1} - \mathbf{k}) = -\psi_1(\mathbf{k})$$

Proof. It is a direct consequence of properties of the inner product

$$\begin{aligned} \psi_1(2\alpha \mathbf{1} - \mathbf{k}) &= \omega \langle 2\alpha \mathbf{1} - \mathbf{k} - \alpha \mathbf{1}, \mathbf{p}^r \rangle \\ &= \omega \langle \alpha \mathbf{1} - \mathbf{k}, \mathbf{p}^r \rangle \\ &= -\omega \langle \mathbf{k} - \alpha \mathbf{1}, \mathbf{p}^r \rangle \\ &= -\psi_1(\mathbf{k}) \end{aligned}$$

□

The anti-reflection implies that any positive semi-frequency has a negative counterpart; which is due to the expansion of the π -modes into complex exponentials via Euler's identity. Furthermore, symmetries in Pascal's triangle give the following result

$$\begin{bmatrix} 2\alpha \\ 2\alpha \mathbf{1} - \mathbf{k} \end{bmatrix} = \frac{1}{4^{\alpha N}} \prod_{p=1}^N \binom{2\alpha}{2\alpha - k_p} = \frac{1}{4^{\alpha N}} \prod_{p=1}^N \binom{2\alpha}{k_p} = \begin{bmatrix} 2\alpha \\ \mathbf{k} \end{bmatrix}$$

So Proposition 11 and the symmetry of combinatorial factors guarantee that the resulting distribution function is real-valued because conjugate complex exponentials are grouped together to reform a cosine function.

Proposition 13. *(Jensen's equality) If β is a real number on the interval $[0;1]$, \mathbf{u} and \mathbf{v} two vectors in \mathcal{K}_α , then*

$$\psi_1(\beta \mathbf{u} + (1 - \beta) \mathbf{v}) = \beta \cdot \psi_1(\mathbf{u}) + (1 - \beta) \cdot \psi_1(\mathbf{v})$$

Proof. The demonstration is based upon the linearity of the inner product $\langle \cdot, \cdot \rangle$ and the trivial vector equality $\alpha \mathbf{1} = \beta \alpha \mathbf{1} + (1 - \beta) \alpha \mathbf{1}$.

$$\begin{aligned}
\psi_1(\beta \mathbf{u} + (1 - \beta) \mathbf{v}) &= \omega \langle \beta \mathbf{u} + (1 - \beta) \mathbf{v} - \alpha \mathbf{1}, \mathbf{p}^r \rangle \\
&= \omega \langle \beta(\mathbf{u} - \alpha \mathbf{1}) + (1 - \beta)(\mathbf{v} - \alpha \mathbf{1}), \mathbf{p}^r \rangle \\
&= \omega \beta \langle \mathbf{u} - \alpha \mathbf{1}, \mathbf{p}^r \rangle + \omega(1 - \beta) \langle \mathbf{v} - \alpha \mathbf{1}, \mathbf{p}^r \rangle \\
&= \beta \cdot \psi_1(\mathbf{u}) + (1 - \beta) \cdot \psi_1(\mathbf{v})
\end{aligned}$$

□

Jensen's equality implies that any semi-frequency belongs to a segment whose endpoints are the minimum and the maximum value of ψ_1 . Accordingly, they can be calculated via linear combinations of those extrema. Then in order to determine the minimum and the maximum of ψ_1 , the contribution of the components of the input vectors must be examined; which leads to the next proposition.

Proposition 14. *The semi-frequency generator admits a minimum (resp. a maximum) in \mathcal{K}_α for the input vector $\mathbf{0}$ (resp. $2\alpha \mathbf{1}$).*

$$(i) \min\{\psi_1\} = \psi_1(\mathbf{0}) = -\omega \alpha \langle \mathbf{1}, \mathbf{p}^r \rangle$$

$$(ii) \max\{\psi_1\} = \psi_1(2\alpha \mathbf{1}) = -\psi_1(\mathbf{0}) = \omega \alpha \langle \mathbf{1}, \mathbf{p}^r \rangle$$

Proof. The minimum value of ψ_1 is obtained when all the individual contributions $k_p - \alpha$ are minimal. That situation occurs when $\mathbf{k} = \mathbf{0}$. In such case, ψ_1 has the value $\psi_1(\mathbf{0}) = \omega \langle \mathbf{0} - \alpha \mathbf{1}, \mathbf{p}^r \rangle = -\omega \alpha \langle \mathbf{1}, \mathbf{p}^r \rangle$. Then Proposition 12 *i.e anti-reflection* guarantees that the maximum is $\psi_1(2\alpha \mathbf{1} - \mathbf{0}) = -\psi_1(\mathbf{0})$ □

Proposition 15. *The semi-frequencies are all multiples of ω in the closed interval defined by the extrema of the semi-frequency generator *i.e**

$$\forall \mathbf{k} \in \mathcal{K}_\alpha, \exists m \in \mathbb{N} \mid m\omega \in [\psi_1(\mathbf{0}); -\psi_1(\mathbf{0})] \text{ and } \psi_1(\mathbf{k}) = m\omega$$

Proof. The Jensen equality implies that all semi-frequencies are in an closed interval, bounded by the extrema of ψ_1 . In virtue of Proposition 14, that interval is $[\psi_1(\mathbf{0}); -\psi_1(\mathbf{0})]$. Furthermore, the sum in Definition 5.1.1 involves only subtractions and multiplications of integers; therefore, any output of ψ_1 is a multiple of ω . □

We continue with the development of a potential parametrization of the semi-frequencies with the real number β from Proposition 13. The idea is to seek for a general form for a vector \mathbf{k} which generates a given semi-frequency $m\omega$. First of all, we use Jensen's equality to parametrize \mathbf{k} in the form $\mathbf{k}_\beta = 2\beta\alpha\mathbf{1} + (1 - \beta)\mathbf{0}$; which leads to the semi-frequency $\psi_1(\mathbf{k}_\beta) = (1 - 2\beta)\psi_1(\mathbf{0})$. Then, we apply Proposition 5.1 to write $\psi_1(\mathbf{k}_\beta) = m\omega$ for some integer $m \in [0; \alpha\langle\mathbf{1}, \mathbf{p}^r\rangle]$. From these two expressions for the semi-frequency, we can deduce the equality

$$2\beta = 1 - \frac{m\omega}{\psi_1(\mathbf{0})}$$

The original input vector is then given by the relation $\mathbf{k}_\beta = \alpha\left(1 + \frac{m}{\alpha\langle\mathbf{1}, \mathbf{p}^r\rangle}\right)\mathbf{1}$. Although this parametrization gives a general form for the input vectors, it is only valid for $m = 0$ and $m = \pm\alpha\langle\mathbf{1}, \mathbf{p}^r\rangle$. This is explained by the fact that those vectors are not in $\{0, 1, \dots, 2\alpha\}^N$. However, if we ignore that infringement and consider the special case $\alpha = 1$, then the N -th root of the combinatorial factor becomes

$$\left[\begin{matrix} 2 \\ \mathbf{k} \end{matrix}\right]^{1/N} = \frac{1}{4} \left[\begin{matrix} 2 \\ (1 + \frac{m}{\langle\mathbf{1}, \mathbf{p}^r\rangle})\mathbf{1} \end{matrix}\right]^{1/N} = \frac{1/2}{\left(1 + \frac{m}{\langle\mathbf{1}, \mathbf{p}^r\rangle}\right)! \left(1 - \frac{m}{\langle\mathbf{1}, \mathbf{p}^r\rangle}\right)!}$$

If $x = \frac{m}{\langle\mathbf{1}, \mathbf{p}^r\rangle}$ and the factorials are replaced with their definitions with the gamma function; followed by the application of Euler's reflective formula, then the expression for the N -th root of the combinatorial factor can be simplified as follows

$$\left[\begin{matrix} 2 \\ \mathbf{k} \end{matrix}\right]^{1/N} = \frac{1/2}{(1+x)x(1-x)\Gamma(x)\Gamma(1-x)} = \frac{1/2}{(1+x)(1-x)} \cdot \frac{\sin(x\pi)}{x\pi} \quad (5.2)$$

Theorem 11. (*Euler's reflective formula*)

$$\Gamma(z) \cdot \Gamma(1-z) = \frac{\pi}{\sin(z\pi)} \quad z \notin \mathbb{Z}$$

The singularities of Equation (5.2) correspond to the cases where \mathbf{k}_β is the single vector which generates $m\omega$ i.e for $m = 0$ and $m = \pm\alpha\langle\mathbf{1}, \mathbf{p}^r\rangle$. In practice, that expression is not applicable because ψ_1 is not an injective function. Besides, finding all the antecedents of a given semi-frequency is equivalent to solving a non trivial combinatorial problem, see Section 5.5.

5.2 Pseudo-periodicity

As mentioned in Proposition 15, all semi-frequencies are multiples of the main frequency. Subsequently, they share a common period that we now determine. To start with, we can notice that any integer $m \in [0; \alpha \langle \mathbf{1}, \mathbf{p}^r \rangle]$ is associated to the semi-frequencies $\pm m\omega$. The corresponding period is given by the relation

$$T_m = \frac{2\pi}{2m\omega} = \frac{\pi}{m\omega}$$

We can inject this expression in the π -modes to determine the optimal value of m

$$\cos^{2\alpha}(\omega k^r(t + T_m) + \phi_k) = \cos^{2\alpha}(\omega k^r t + \omega k^r T_m + \phi_k) = \cos^{2\alpha}\left(\omega k^r t + \pi \frac{k^r}{m} + \phi_k\right)$$

Since the trigonometric function $x \mapsto \cos^2(x)$ has period π , the optimal value of m is such that k^r/m is an integer for $k = 1, \dots, N$; which is only possible for $m = 1$.

Definition 5.2.1. (*Pseudo-period*) The π -modes share a common period, referred to as the pseudo-period and defined as

$$T = \frac{\pi}{\omega}$$

Proposition 16. (*Auto-similarity*) If $t \geq 0$ is a real number then the distribution function verifies the relation

$$\delta_N(t + T) = e^{-\lambda T} \cdot \delta_N(t)$$

Proof. To prove this result, we apply the property of the pseudo-period to calculate the image of the input $t + T$ with the distribution function.

$$\begin{aligned} \delta_N(t + T) &= e^{-\lambda(t+T)} \prod_{k=1}^N \cos^{2\alpha}(\omega k^r(t + T) + \phi_k) \\ &= e^{-\lambda(t+T)} \prod_{k=1}^N \cos^{2\alpha}(\omega k^r t + \phi_k) \\ &= e^{-\lambda T} \cdot \delta_N(t) \end{aligned}$$

□

The auto-similarity indicates that the oscillations of the π -modes are periodically damped with a logarithmic decrement λT . So in order to concentrate the distribution

at the point $t = 1$, that decrement ought to be large enough such that oscillations after the first period are negligible. For instance, λ can be chosen such that in the second period, the amplitudes are 10^3 or 10^6 smaller than in the first one.

5.3 Reduced moments

As discussed in the previous section, the auto-similarity reduces the analysis of the distribution function to the first period when the decay coefficient λ is large enough. Intuitively, we can expect the integral in the definition of the moments to be reduced to the first period only. This conjecture is formally proven with Theorem 13.

Definition 5.3.1. (*Complete moments*) The n -th order complete moment of δ_N is defined as follows

$$m_n = \int_0^\infty t^n \cdot \delta_N(t) dt$$

Definition 5.3.2. (*Reduced moments*) If T is the pseudo period then the n -th order restricted moment of δ_N is defined as follows

$$\mu_n = \int_0^T t^n \cdot \delta_N(t) dt$$

Proposition 17. The Laplace transform of δ_N verifies the relation

$$\Delta_N(s) = \frac{1}{1 - e^{-(s+\lambda)T}} \int_0^T e^{-st} \cdot \delta_N(t) dt, \quad \text{Real}(s) > -\lambda$$

Proof. The Laplace transform of δ_N can be directly obtained with Property 8 and the frequency shift property. Nonetheless, we outline the elements of proof below.

$$\begin{aligned} \Delta_N(s) &= \int_0^\infty e^{-st} \cdot \delta_N(t) dt \\ &= \sum_{k=0}^\infty \int_{kT}^{(k+1)T} e^{-st} \cdot \delta_N(t) dt && \text{(Chasles relation)} \\ &= \sum_{k=0}^\infty \int_{kT}^{(k+1)T} e^{-s(u+kT)} \cdot \delta_N(u+kT) du && \text{(Substitution } t=u+kT) \\ &= \sum_{k=0}^\infty e^{-k(s+\lambda)T} \int_0^T e^{-su} \delta_N(u) du && \text{(Auto-similarity)} \\ &= \frac{1}{1 - e^{-(s+\lambda)T}} \int_0^T e^{-su} \delta_N(u) du && \text{(Geometric series)} \end{aligned}$$

□

5.3.1 Reduced moments theorems

Theorem 12. *Reduced Moments Triangulation (RMT)*

If n is a positive integer, then the n -th order complete moment of δ_N is a linear combination of reduced moments of orders at most n . More precisely,

$$\forall n \in \mathbb{N}, m_n = \sum_{k=0}^n a_{n,k} \cdot \mu_k \quad \text{with } a_{n,k} = (-1)^{n-k} \binom{n}{k} \frac{d^{n-k}}{ds^{n-k}} \left\{ \frac{1}{1 - e^{-(s+\lambda)T}} \right\}_{s=0}$$

Proof. To prove this result, we use the Laplace transform as a generator of the *complete moments* of a random variable. Then we apply the General Leibniz rule, which expresses the derivatives of a product of two functions as linear combination of products of their individual derivatives.

$$\begin{aligned} m_n &= (-1)^n \frac{d^n}{ds^n} \left\{ \Delta_N(s) \right\}_{s=0} \\ &= (-1)^n \frac{d^n}{ds^n} \left\{ \frac{1}{1 - e^{-(s+\lambda)T}} \int_0^T e^{-st} \cdot \delta_N(t) dt \right\}_{s=0} \\ &= (-1)^n \sum_{k=0}^n \binom{n}{k} \frac{d^{n-k}}{ds^{n-k}} \left\{ \frac{1}{1 - e^{-(s+\lambda)T}} \right\}_{s=0} \cdot \frac{d^k}{ds^k} \left\{ \int_0^T e^{-st} \cdot \delta_N(t) dt \right\}_{s=0} \end{aligned}$$

The rightmost derivatives simply as follows

$$\frac{d^k}{ds^k} \left\{ \int_0^T e^{-st} \cdot \delta_N(t) dt \right\}_{s=0} = \int_0^T (-t)^k \cdot \delta_N(t) dt = (-1)^k \cdot \mu_k$$

□

The RMT transforms the improper integral in the definition of *complete moments* into a sum of finite integrals. However, such reduction is compensated by the fact that the coefficients $a_{n,k}$ for $k = 0, \dots, n$ are derivatives of a composition of functions. Yet without explicit differentiation, we can note that $a_{n,n}$ is independent of n as follows

$$a_{n,n} = \frac{1}{1 - e^{-\lambda T}}$$

The remaining coefficients can be calculated recursively in the order $a_{n,n-1}, a_{n,n-2}, \dots, a_{n,0}$.

Corollary 1. If n is a positive integer, $\mathbf{m} = (m_0, m_1, \dots, m_n)^\top$ and $\boldsymbol{\mu} = (\mu_0, \mu_1, \dots, \mu_n)^\top$, then exists an $n \times n$ lower triangular matrix $\mathcal{T}_{m\mu}$ such that

$$\mathbf{m} = \mathcal{T}_{m\mu} \cdot \boldsymbol{\mu}$$

Proof. In virtue of the RMT, the coefficient of $T_{m\mu}$ are defined as follows

$$\mathcal{T}_{m\mu}^{i,j} = \begin{cases} a_{i,j} & \text{if } j \leq i. \\ 0 & \text{otherwise.} \end{cases}$$

□

As mentioned earlier, the information contained in the distribution ought to be concentrated into the first period when the decay coefficient increases. Without further calculations, it is straightforward that $a_{n,n}$ approaches 1 for $\lambda \rightarrow \infty$. Therefore, to show that the *complete moments* converge toward the *reduced moments*, we have to prove that any coefficient outside the principal diagonal of $\mathcal{T}_{m\mu}$ approaches zero when λ gets larger. To that end, we can apply Faà di Bruno's formula; which generalizes the chain rule to higher derivatives for a composition of functions.

Theorem 13. (*Faà di Bruno's formula*)

$$(f \circ g)^{(n)}(x) = \sum \frac{n!}{m_1! 1!^{m_1} m_2! 2!^{m_2} \dots m_n! n!^{m_n}} f^{(m_1+\dots+m_n)}(g(x)) \prod_{j=1}^n (g^{(j)}(x))^{m_j},$$

Where m_1, m_2, \dots, m_n are positive numbers such that $1m_1 + 2m_2 + 3m_3 + \dots + nm_n = n$

In Faà di Bruno's formula, f and g are considered to be differentiable enough such that their k -th derivative for $0 \leq k \leq n$ are well-defined. Then, the n -th derivative of $f \circ g$ is expressed in terms of sums of products of derivatives of f and g .

Theorem 14. *Reduced Moments Convergence (RMC)*

The complete moments converge toward the reduced moments as the decay coefficient λ grows larger. Which is summarized with the following points

- (i) $\lim_{\lambda \rightarrow \infty} a_{n,n} = 1$
- (ii) $\lim_{\lambda \rightarrow \infty} a_{n,p} = 0$ for $0 \leq p < n$
- (iii) $\forall n \in \mathbb{N}, \lim_{\lambda \rightarrow \infty} m_n = \mu_n$

Proof. The first results (i) has already been discussed. For the other cases, the functions of interest for Faà di Bruno's formula are $f : x \mapsto \frac{1}{1-x}$ and $g : x \mapsto e^{-Tx}$. Those functions are both infinitely differentiable and their k -th derivatives are defined by the equalities

$$f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}}$$

$$g^{(k)}(x) = (-T)^k e^{-Tx}$$

The application of Faà di Bruno's formula gives the following result

$$(f \circ g)^{(n)}(x) = \sum \frac{n!}{m_1! 1!^{m_1} m_2! 2!^{m_2} \dots m_n! n!^{m_n}} \cdot \left[\frac{(m_1 + \dots + m_n)!}{(1 - e^{-Tx})^{m_1 + \dots + m_n + 1}} \right] (-T)^n \cdot e^{-x \cdot T(m_1 + \dots + m_n)}$$

Given their definition with the RMT, the terms outside the principal diagonal of $\mathcal{T}_{m\mu}$ correspond to $n \neq 0$ in Faà Di Bruno's formula. The exponential term enforces the convergence to zero as x grows larger, hence (ii). The last point (iii) is simply the applications of (i) and (ii) to the RMT. \square

We continue with an illustration of the RMT and the RMC on a practical example. To that end, we have calculated the first three moments of δ_N . If $a = e^{-\lambda T}$, $b = 1 - a$ and $c = \frac{a}{b}$ then we can write the matrix equality

$$\begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix} = \frac{1}{b} \begin{bmatrix} 1 & 0 & 0 \\ Tc & 1 & 0 \\ T^2(c + 2c^2) & Tc & 1 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix}$$

It important to notice that this equality is homogeneous from a physical viewpoint. In fact, if the period T has the dimension of t , then the product $T\mu_0$ has the same dimension as μ_1 . Which entails that m_1 is homogeneous to time, as we could expect. The same analysis is applicable to the other moments. Finally, it is straightforward that $a \rightarrow 0$, $b \rightarrow 1$ and $c \rightarrow 0$ when $\lambda \rightarrow \infty$. Therefore, we obtain the expected result.

$$\lim_{\lambda \rightarrow \infty} \begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix}$$

With the RMC, the complexity of computations involving the moments of the distribution can be tremendously reduced if λ is large enough. In practice, it is the case because all the information about the distribution must concentrated into the first period, precisely at the point $t = 1$.

5.4 Nodes

We continue the analysis of δ_N with a simplification of the σ -form I, in virtue of the properties of the semi-frequency generator. The first elements to consider is anti-reflection and symmetry of combinatorial factors which guarantee that δ_N is real-valued. In fact, any positive semi-frequency has a negative counterpart that can be

grouped together with Euler's identity, as mentioned in Section 5.1.

Proposition 18. (*Order of approximation*) *The order of approximation for the Dirac impulse function corresponds to the number of pairs of positive-negative semi-frequencies. The general expression for the order is given by the following relation*

$$\text{Order}(r, \alpha, N) = 1 + \alpha \langle \mathbf{1}, \mathbf{p}^r \rangle$$

Proof. This result is a consequence of the fact that any semi-frequency is a multiple of ω in the interval $[\psi_1(\mathbf{0}); -\psi_1(\mathbf{0})]$. And as mentioned before, pairs of negative-positive semi-frequencies are grouped together to form a cosine function, in virtue of Euler's identity. The extra term "1" accounts for the pair of semi-frequencies $(-0, +0)$. \square

Knowing the exact order of approximation of the shifted Dirac impulse, we can move forward by proposing a new expression for the σ -form, explicitly in the Abate-Whitt framework.

Definition 5.4.1. (σ -**form II**) *The sum form has the alternative form*

$$\delta_N(t) = e^{-\lambda t} \sum_{k=0}^{\alpha \langle \mathbf{1}, \mathbf{p}^r \rangle} \left(\eta_k e^{2i\omega k t} + \bar{\eta}_k e^{-2i\omega k t} \right)$$

where η_k for $k = 1, \dots, \alpha \langle \mathbf{1}, \mathbf{p}^r \rangle$ are complex numbers which account for the combinatorial factors and the semi-phases; $\bar{\eta}_k$ are their complex conjugate numbers.

Proposition 19. (σ -**form III**) *The sum form has the third alternative form*

$$\delta_N(t) = 2e^{-\lambda t} \cdot \text{Real} \left(\sum_{k=0}^{\alpha \langle \mathbf{1}, \mathbf{p}^r \rangle} \eta_k e^{2i\omega k t} \right)$$

The σ -form III is obtained with the application of the identity $z + \bar{z} = 2 \cdot \text{real}(z)$. The factor "2" in the arguments of the exponentials originates from the fact that any frequency is twice a given semi-frequency. Subsequently, we can deduce a close form expression for the nodes in the Abate-Whitt framework.

Proposition 20. (**Nodes**) *The node of δ_N have the same real part and imaginary parts forming an arithmetic sequence with common difference factor 2ω .*

$$\beta_k = \lambda - 2i\omega k$$

This result was proven recursively in [13, Appendix], for the case $r = 0$.

5.5 Calculation of the weights

We now focus on the remaining unknowns, the weights η_k . As mentioned in Section 5.1, their calculation requires to solve a combinatorial problem because the nodes are defined as follows

$$\eta_p = \sum_{\mathbf{k} \in \mathcal{K}_\alpha | \psi_1(\mathbf{k}) = p\omega} \frac{1}{4^{\alpha N}} \begin{bmatrix} 2\alpha \\ \mathbf{k} \end{bmatrix} e^{2i\Phi(\mathbf{k})}$$

It is difficult and impractical to apply this definition for large values of N . That is the reason why we propose two methods to bypass that shortcoming.

5.5.1 An integral based method

A first method is based upon the orthogonality of complex exponentials such that the weights are calculated in the same way as the coefficients of a Fourier series.

Definition 5.5.1. (*Projection integral*) If k and p are positive integers on the interval $[0; \alpha\langle \mathbf{1}, \mathbf{p}^r \rangle]$ then the projection integral is defined as

$$\mathcal{I}(k, p) = \int_0^T \left(\eta_k e^{2i\omega kt} + \bar{\eta}_k e^{-2i\omega kt} \right) e^{-2i\omega pt} dt \quad T = \frac{\pi}{\omega}$$

Proposition 21. The integral projector $\mathcal{I}(\cdot, \cdot)$ is derived from an hermitian inner product and verifies the following relations

- (i) $\mathcal{I}(0, 0) = (\eta_0 + \bar{\eta}_0) \cdot T$
- (ii) $\mathcal{I}(p, p) = \eta_p \cdot T$
- (iii) $\mathcal{I}(k, p) = 0$ if $k \neq p$

Proof. Let us assume that $k \neq p$, then the projection integral becomes

$$\begin{aligned} \mathcal{I}(k, p) &= \int_0^T \left(\eta_k e^{2i\omega kt} + \bar{\eta}_k e^{-2i\omega kt} \right) e^{-2i\omega pt} dt \\ &= \eta_k \cdot \left[\frac{e^{2i\omega(k-p)t}}{2i\omega(k-p)} \right]_0^T + \bar{\eta}_k \cdot \left[\frac{e^{-2i\omega(k+p)t}}{-2i\omega(k+p)} \right]_0^T \\ &= \eta_k \cdot \frac{e^{2i(k-p)\pi} - 1}{2i\omega(k-p)} + \bar{\eta}_k \cdot \frac{e^{-2i(k+p)\pi} - 1}{-2i\omega(k+p)} \\ &= 0 \end{aligned}$$

Hence the result (iii). The result (i) is trivial since the exponential terms vanish under the integral sign and the result (ii) follows the same principle as (iii). \square

Proposition 22. (Weights) *The weights of δ_N in the Abate-Whitt framework are given by the relations*

$$\eta_0 = \bar{\eta}_0 = \frac{1}{2T} \int_0^T \left(\prod_{k=1}^N \cos^{2\alpha}(\omega k^r t + \phi_k) \right) dt$$

$$\eta_p = \frac{1}{T} \int_0^T \left(e^{-2i\omega p t} \prod_{k=1}^N \cos^{2\alpha}(\omega k^r t + \phi_k) \right) dt$$

Proof. The first step is to equate the π -form and the σ -form II.

$$e^{-\lambda t} \prod_{k=1}^N \cos^{2\alpha}(\omega k^r t + \phi_k) = e^{-\lambda t} \sum_{k=0}^{\alpha\langle \mathbf{1}, \mathbf{p}^r \rangle} \left(\eta_k e^{2i\omega k t} + \bar{\eta}_k e^{-2i\omega k t} \right)$$

Then one replaces $e^{-\lambda t}$ by $e^{-2i\omega p t}$, integrates over the first period and gets

$$\int_0^T \left(e^{-2i\omega p t} \prod_{k=1}^N \cos^{2\alpha}(\omega k^r t + \phi_k) \right) dt = \sum_{k=0}^{\alpha\langle \mathbf{1}, \mathbf{p}^r \rangle} \mathcal{I}(k, p)$$

Using the previous results on the values of $\mathcal{I}(k, p)$, we can obtain the expression of each individual weight. We choose the zero-frequency weights η_0 and $\bar{\eta}_0$ to be real numbers. In fact, their imaginary parts cancel each other, leaving only the real part. \square

5.5.2 Recursive method

Despite being simple to implement, the integral based method requires heavy calculations when N grows larger. The precision required is ever increasing as the complex exponentials have high oscillation frequencies. So a more stable method is needed in order to reduce the cost of computations and to preserve accuracy in the results. We summarize, a recursive method for $\alpha = 1$ in the next 4 steps.

Step 1 : For a given integer N , the π -forms of δ_{N+1} and δ_N verify the relation

$$\delta_{N+1}(t) = \cos^2 \left(\omega(N+1)^r t + \phi_{N+1} \right) \cdot \delta_N(t)$$

Step 2 : The π -form of δ_N is expanded into the σ -form II

$$\delta_N(t) = e^{-\lambda t} \sum_{k=0}^{\alpha\langle \mathbf{1}, \mathbf{p}^r \rangle} \left(\eta_k e^{2i\omega k t} + \bar{\eta}_k e^{-2i\omega k t} \right)$$

Step 3 : The cosine term is transformed into its complex exponential form. For clarity, we use ϕ to designate ϕ_{N+1}

$$\cos^2\left(\omega(N+1)^r t + \phi\right) = \frac{1}{4} \left(e^{2i(\omega(N+1)^r t + \phi)} + 2 + e^{-2i(\omega(N+1)^r t + \phi)} \right)$$

Step 4 : The product of these forms yields

$$\begin{aligned} \delta_{N+1}(t) &= \frac{1}{2} \cdot \delta_N(t) \\ &+ \frac{e^{-\lambda t}}{4} \cdot e^{2i(\omega(N+1)^r t + \phi)} \cdot \sum_{k=0}^{\alpha\langle \mathbf{1}, \mathbf{p}^r \rangle} \left(\eta_k e^{2i\omega k t} + \bar{\eta}_k e^{-2i\omega k t} \right) \\ &+ \frac{e^{-\lambda t}}{4} \cdot e^{-2i(\omega(N+1)^r t + \phi)} \cdot \sum_{k=0}^{\alpha\langle \mathbf{1}, \mathbf{p}^r \rangle} \left(\eta_k e^{2i\omega k t} + \bar{\eta}_k e^{-2i\omega k t} \right) \end{aligned}$$

The weights of δ_{N+1} can then be explicitly calculated from those of δ_N . In application, only the terms with positive frequencies are relevant since the coefficients for negative frequencies can be obtained with complex conjugation. In the light of this observation, only the products $\eta_k \cdot e^{2i\phi}$ for $(N+1)^r + k \geq 0$ and $\bar{\eta}_k \cdot e^{2i\phi}$ for $(N+1)^r - k \geq 0$ are relevant in the first sum; whereas the products $\eta_k \cdot e^{-2i\phi}$ for $k - (N+1)^r \geq 0$ are of interest in the second sum.

5.6 Width of the distribution function

The approximation of the shifted Dirac impulse function about the point $t = 1$ is imperfect since the distribution function has a non negligible width, resulting from the superposition of the π -modes as shown in Figure (5.1). In this section, we determine an upper bound for that additional measure for concentration by analyzing the π -form.

Definition 5.6.1. *The width ϵ_N of the distribution function is defined the greatest difference between two consecutive zeros of δ_N in the vicinity of the point $t = 1$.*

Proposition 23. *The width of the distribution function verifies the inequalities*

$$0 < \epsilon_N \leq \frac{\pi}{\omega N^r}$$

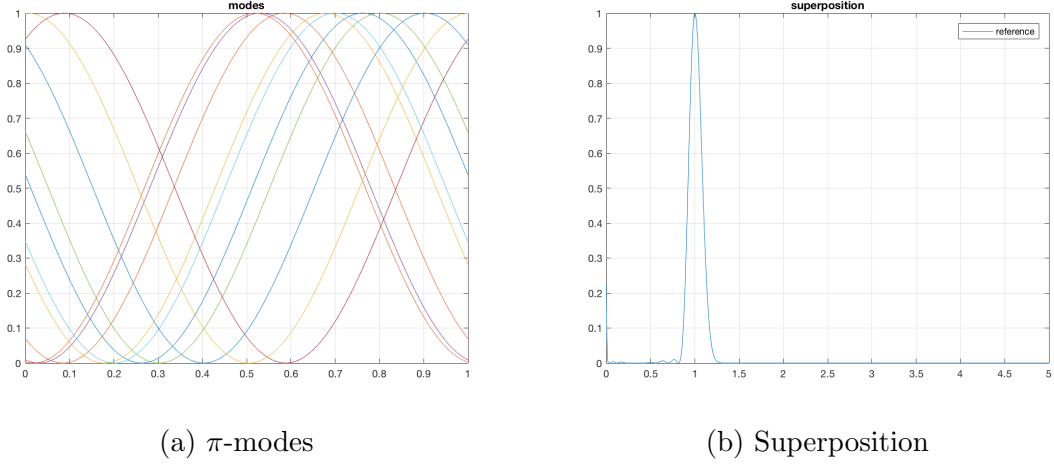


Figure 5.1: Superposition of π -modes in a normalized scale $k = 1, \dots, 10$ for $r = 0$, $\omega = 2.7233$. The phases are ϕ_k are optimized to reduce scv

Proof. The mode with the highest frequency in the π -form is $\cos(\omega \cdot N^r t + \phi_n)$. This mode has the smallest period, thus also has the most zeros within a period. These zeros can be parametrized as follows

$$z_p = \frac{1}{\omega \cdot N^r} \left(\left(p + \frac{1}{2}\right)\pi - \phi_N \right), \quad p \in \mathbb{N}$$

The relation between two consecutive zeros is given as the arithmetic sequence

$$z_{p+1} = z_p + \frac{\pi}{\omega \cdot N^r}$$

Hence the expression for the upper-bound of the width. □

We can note that the upper-bound of the width ignores the modal phases ϕ_k because they only serve to shift the position of the zeros of the π -modes on the real axis. So an optimal parametrization of those phases leads to a reduction of the width and a significant increase in the concentration at $t = 1$, i.e $\delta_N(1)$ is maximized.

5.7 Reinterpretation of the HTT method

We now reinterpret the HTT method within this framework. The original distribution function in Equation (3.3) can be rewritten is the standard π -form

$$\delta_{ME}(t) = c \cdot e^{-\lambda t} \prod_{k=1}^M \cos^2(\omega t + \phi_k), \quad M = (N - 1)/2 \quad (5.3)$$

from which we deduce that the HTT method uses matrix exponential distributions in the monomials class $\theta_r^{(M,1)}$ with index $r = 0$, general exponent $\alpha = 1$ and $M = \frac{N-1}{2}$. The order of approximation of the shifted Dirac impulse function is given by the relations

$$Order(r, \alpha, M) = 1 + \alpha \langle \mathbf{1}, \mathbf{p}^r \rangle = 1 + \langle \mathbf{1}, \mathbf{1} \rangle = 1 + M = \frac{N+1}{2}$$

This result for the order is strictly identical to the order which is given in [3, Section 4.4] and the σ -form III simplifies to

$$\delta_M(t) = 2e^{-\lambda t} \sum_{k=0}^{\frac{N-1}{2}} \mathbf{Real}(\eta_k e^{2i\omega_k t})$$

The main frequency ω and the modal phases ϕ_1, \dots, ϕ_M are obtained via optimization in order to minimize the squared coefficient of variation.

Chapter 6

Unimodal method

In this chapter, we propose an inversion method without positive or negative overshoots and with explicit expressions for the Abate-Whitt coefficients. The shifted Dirac distribution is approximated as follows

Definition 6.0.1. (*unimodal π -form*)

$$\delta_N(t) = ce^{-\lambda t} \cdot \cos^{2N}(\omega t + \phi)$$

The coefficient α in the general formula for the π -form from Chapter 2 is not relevant anymore, since it only changes the overall power from $2N$ to $2\alpha N$.

6.1 Weights and nodes

The π -form can be expanded into the σ -form directly with Euler's identity and the Binomial theorem. Which leads to the following propositions

Proposition 24. (*unimodal σ -form*)

$$\delta_N(t) = \frac{c}{4^N} \binom{2N}{N} e^{-\lambda t} + \frac{c}{4^N} \sum_{k=0}^{N-1} \binom{2N}{k} \left\{ e^{2i(N-k)(\omega t + \phi)} + e^{-2i(N-k)(\omega t + \phi)} \right\} e^{-\lambda t}$$

Proposition 25. (*order*) The order of approximation of the Dirac impulse function in the unimodal class is linear and given by the expression $\text{Order}(N) = 1 + N$

Proposition 26. (*weights and nodes*) The σ -form for the unimodal distribution function has the alternative form

$$\delta_N(t) = 2 \cdot \text{Real} \left(\sum_{k=0}^N \eta_k e^{-\beta_k t} \right)$$

where the nodes are given by the expressions

$$\begin{aligned}\beta_k &= \lambda - 2i\omega(N - k) & 0 \leq k \leq N \\ \eta_k &= \frac{c}{4^N} \binom{2N}{k} e^{2i(N-k)\phi} & 0 \leq k \leq N - 1 \\ \eta_N &= \frac{1}{2} \cdot \frac{c}{4^N} \binom{2N}{N}\end{aligned}$$

Remark 5. To improve numerical stability, the weights may be calculated via the recursive method proposed in section 5.5

Remark 6. The normalization coefficient c correspond to the inverse of the 0^{th} order moment of density function i.e $1/c = m_0 = \int_0^\infty \delta_N(t)dt = \sum_{k=0}^N \frac{\eta_k}{\beta_k}$

6.2 Optimal parameters

We continue by deriving the optimal expressions for the main frequency, the decay and the phase. The ideas is to look for δ_N as a sine function in order to take advantage of the results $\lim_{x \rightarrow 0} \sin(x) = 0$, $\sin(\pi/2) = 1$ and $\sin(\pi) = 0$. A preliminary parametrization for the unimodal π -form is as follows

$$\delta_N(t) = ce^{-\lambda t} \cdot \cos^{2N}(\omega(t-1) + \xi), \quad \omega = \frac{\pi}{2}$$

The phase coefficient then becomes $\phi = -\pi/2 + \xi$ and we can then notice that if $\xi = 0$ then $\delta_N(0) = 0$, $\delta_N(1) = ce^{-\lambda}$ and $\delta_N(2) = 0$. In practice, the decay coefficient shifts the argument of the maximum to a value $t \neq 1$. To offset this shift, we use the extra term in the phase coefficient to cancel the derivative of the density function at $t = 1$,

$$\frac{d\delta_N}{dt} = \left(-\lambda - 2N\omega \tan(\omega(t-1) + \xi) \right) \delta_N(t)$$

which yields $\xi = -\arctan\left(\frac{\lambda}{2N\omega}\right)$. Therefore, the optimal phase coefficient is

$$\phi = -\frac{\pi}{2} - \arctan\left(\frac{\lambda}{2N\omega}\right)$$

The decay coefficient λ is chosen such that the amplitudes in the second period are 10^r times smaller than those in the first period. This is equivalent to solving the equation

$\delta_N(T) = 10^{-r} \cdot \delta_N(0)$ or $e^{-\lambda T} = 10^{-r}$; which yields

$$\lambda = \frac{r\omega \log(10)}{\pi}, \quad r \in \mathbb{N}^*$$

Figures 6.1a to 6.1b depict the resulting distribution function.

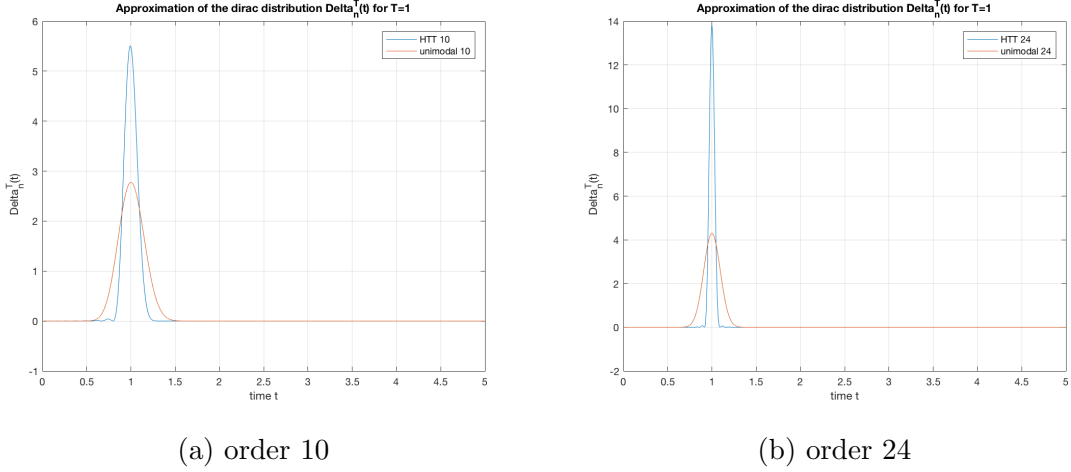


Figure 6.1: Comparison of the shifted Dirac impulse function with the HTT method and the unimodal method

The convergence of the unimodal distribution toward the shifted-scaled Dirac impulse function is slower than that of the HTT method. This behavior accounts for the use of a single mode, which has a constant width given by the expression $\epsilon_N = \frac{\pi}{\omega} = 2$. The width is a constant as there is exactly one zero per period. The squared coefficient of variation has been computed numerically up to $N=4000$ and follows the asymptotic law

$$scv(X) \sim \frac{0.2}{N}$$

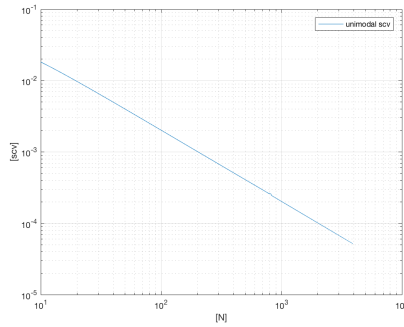


Figure 6.2: Squared coefficient of variation as a function of the order in loglog scale

Chapter 7

Optimization for the monomials class

As mentioned in Section 3.4, the parameters of δ_N must be optimized in order to reduce the squared coefficient of variation. Those parameters are represented by $\mathbf{x} = [\lambda, \omega, \phi_1, \dots, \phi_N] \in \mathbb{R}^{N+2}$, and the distribution function can be written more explicitly as $\delta_N(t, \mathbf{x})$. Therefore, any quantity derived from it is also a function of \mathbf{x} .

7.1 Squared coefficient of variation

We move on to deriving another expression for scv as function of \mathbf{x} . To that end, we first recall that δ_N is a matrix exponential distribution; thus must be normalized. That normalization is equivalent to dividing all the moments by $m_0(\mathbf{x})$; which give the new expression for the n -th order moment,

$$\tilde{m}_n(\mathbf{x}) = \frac{m_n(\mathbf{x})}{m_0(\mathbf{x})}$$

from which we can derive the new expression for the squared coefficient of variation

$$scv(\mathbf{x}) = \frac{m_0(\mathbf{x})m_2(\mathbf{x})}{m_1(\mathbf{x})^2} - 1$$

7.2 Time derivatives of the distribution function

We now determine conditions of the derivative of δ_N at $t = 1$, by analogy with the standard normal distribution $\mathcal{N}(0, 1)$ of density function $g : t \mapsto \frac{1}{\sqrt{2\pi}}e^{-t^2}$. This function, also called *gaussian*, verifies that $g^{(4k)}$ has global maximum $g^{(4k)}(0)$ and $g^{(4k+2)}$ has global minimum $g^{(4k+2)}(0)$ for $k \in \mathbb{N}$. These results can be proven by induction and are illustrated in Figure (7.1) for $k = 0$.

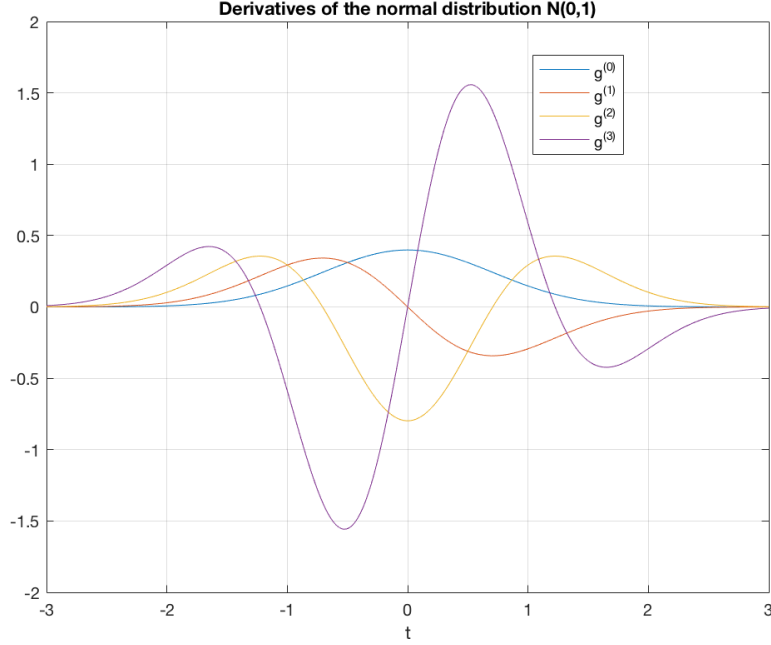


Figure 7.1: Derivative of a gaussian

Similarly, the parameters of δ_N have to be optimized such that $\delta_N^{(4k)}$ has global maximum $\delta_N^{(4k)}(1, \mathbf{x})$ and $\delta_N^{(4k+2)}$ has global minimum $\delta_N^{(4k+2)}(1, \mathbf{x})$. For simplicity and as a follow up to the argument given for the reduction of the width of the density function in Section 5.6, we can seek maximize $\delta_N(1, \mathbf{x})$.

7.3 Optimization strategy

In order to find the optimal parameters we need to solve an optimization problem. In practice, if $f : \mathbb{R}^{N+2} \rightarrow \mathbb{R}$ is an objective function, then we have to find $\mathbf{x}^* \in \mathbb{R}^{N+2}$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^{N+2}$.

7.3.1 Gradient descent

When f is differentiable, we can apply the gradient descent algorithm; which produces successive candidates \mathbf{x}_k for $k \in \mathbb{N}^*$ such that $f(\mathbf{x}_0) > f(\mathbf{x}_1) > \dots > f(\mathbf{x}_n)$. The initial guess \mathbf{x}_0 is arbitrarily close to an optimum which may not be the global optimum if f is not convex. In order to update the candidates, this algorithm uses the gradient of f to find the direction of steepest descent. For that reason, it is categorized as a first-order optimization algorithm.

The procedure is summarized in Algorithm 1. In the version presented, the step size γ is a constant but this coefficient can be updated at each iteration in order to

avoid overshoots or divergence from the optimal solution.

Algorithm 1

```

1: procedure GRADIENTDESCENT( $n, \epsilon, \gamma, f, \mathbf{x}_0$ )
2:    $\mathbf{x} = \mathbf{x}_0$ 
3:   for  $i \leftarrow 1, n$  do
4:     if  $\|\nabla f(\mathbf{x})\| < \epsilon$  then
5:       return  $\mathbf{x}$  ▷ Optimal solution found
6:     else
7:        $\mathbf{x} = \mathbf{x} - \gamma \nabla f(\mathbf{x})$ 
8:     end if
9:   end for
10:  return  $\mathbf{x}$  ▷ Maximum iteration reached
11: end procedure

```

7.3.2 Gradient of the squared coefficient of variation

In order to apply the gradient descent to scv , we have to calculate its partial derivatives with respect to the parameters. For the general case, if x_i is a component of \mathbf{x} , then by logarithmic differentiation we have

$$\frac{\partial scv}{\partial x_i} = \left(\frac{1}{m_0} \frac{\partial m_0}{\partial x_i} - \frac{2}{m_1} \frac{\partial m_1}{\partial x_i} + \frac{1}{m_2} \frac{\partial m_2}{\partial x_i} \right) scv$$

$$\frac{\partial m_p}{\partial x_i} = \int_0^\infty t^p \cdot \frac{\partial \delta_N}{\partial x_i} \cdot dt = \int_0^\infty t^p \cdot \frac{\partial \log\{\delta_N\}}{\partial x_i} \cdot \delta_N \cdot dt$$

The logarithm of δ_N has the expression $\log\{\delta_N\} = -\lambda t + 2 \sum_{k=1}^N \log\{\cos(\omega k^r t + \phi_k)\}$ from which we can calculate the partial derivatives with respect to the parameters

$$\frac{\partial \log\{\delta_N\}}{\partial \lambda} = -t \tag{7.1}$$

$$\frac{\partial \log\{\delta_N\}}{\partial \phi_k} = -2 \tan(\omega k^r t + \phi_k) \tag{7.2}$$

$$\frac{\partial \log\{\delta_N\}}{\partial \omega} = -2t \sum_{k=1}^N k^r \tan(\omega k^r t + \phi_k) \tag{7.3}$$

Remark 7. Given that $t \mapsto \cos^2(t)$ has period π , the modal phases can be restricted to the interval $[0; \pi[$; which is done with the modulo function.

Remark 8. For the first guess \mathbf{x}_0 , we can choose $\lambda = 1$ and randomly select $\omega \in [0, 2\pi[$ and $\phi_k \in [0, \pi[$ for $k = 1, \dots, N$.

The gradient descent algorithm applied to the squared coefficient of variation generally finds a satisfying set of parameters. However, it requires heavy calculations because any partial derivative of a moment involves an integral. Although they can be reduced with Theorem 14 by choosing a large enough value for λ , this method remains expensive for large N .

7.3.3 Gradient of the peak of the distribution

An alternative method to reduce the cost of computation is to optimize \mathbf{x} in order to minimize $-\delta_N(1, \mathbf{x})$. That quantity has the less expensive partial derivatives

$$\frac{\partial \{ -\delta_N(1, \mathbf{x}) \}}{\partial \lambda} = \delta_N(1, \mathbf{x}) \quad (7.4)$$

$$\frac{\partial \{ -\delta_N(1, \mathbf{x}) \}}{\partial \phi_k} = 2 \tan(\omega k^r + \phi_k) \cdot \delta_N(1, \mathbf{x}) \quad (7.5)$$

$$\frac{\partial \{ -\delta_N(1, \mathbf{x}) \}}{\partial \omega} = 2 \sum_{k=1}^N k^r \tan(\omega k^r + \phi_k) \cdot \delta_N(1, \mathbf{x}) \quad (7.6)$$

This second method is a cheaper and faster than that with the coefficient of variation as it does not involve any integral. As mentioned earlier, we only use $\delta_N(1, \mathbf{x})$ but higher time-derivatives of $\delta_N(t, \mathbf{x})$ can also be used. A strategy for the first guess is as follows : if $\mathbf{u} = [\lambda^*, \omega^*, \phi_1^*, \dots, \phi_N^*] \in \mathbb{R}^{N+2}$ is the optimal solution for N , then the first guess for $N + 1$ is as follows $\mathbf{x}_0 = [1, \omega^*, \phi_1^*, \dots, \phi_N^*, \xi]$ with $\xi \sim \pi \cdot \mathcal{N}(0, 1)$.

7.3.4 Hybridization of (1+1)-ES and gradient descent

We now examine a hybrid optimization algorithm which mixes gradient descent and standard evolution strategy [11, Section 2.8]. The idea is to replace the generation of candidates via perturbations via a normal distribution sampling by the gradient descent based upon $-\delta_N(1, \mathbf{x})$. Those candidates are then evaluated with the squared coefficient of variation and the one with lowest *scv* is selected as the best solution. To ease the notations, we redefine the objective functions as $f_0(\mathbf{x}) = \text{scv}(\mathbf{x})$ and $f_1(\mathbf{x}) = -\delta_N(1, \mathbf{x})$. It p is the maximum number of candidates, ϵ the targeted precision, n the maximum number of iterations per candidate, γ the step size and \mathbf{u} the optimal solution for δ_{N-1} , then we can define Algorithm 2 and Algorithm 3 as follows

Algorithm 2

```
1: procedure INITIALGUESSGENERATOR( $\mathbf{u}$ )  
2:   Assert  $\mathbf{u} = [\lambda^*, \omega^*, \phi_1^*, \dots, \phi_{N-1}^*] \in \mathbb{R}^{N+1}$  ▷ General form for  $\mathbf{u}$   
3:   Sample  $\xi \sim \pi \cdot \mathcal{N}(0, 1)$   
4:    $\mathbf{x} = [1, \omega^*, \phi_1^*, \dots, \phi_{N-1}^*, \xi]$   
5:   return  $\mathbf{x}$   
6: end procedure
```

Algorithm 3

```
1: procedure ES-GD( $p, n, \epsilon, \gamma, f_0, f_1, \mathbf{u}$ )  
2:   Initialize  $\mathbf{x}^* = \mathbf{0} \in \mathbb{R}^{N+2}$   
3:   Initialize  $\mu^* = \infty$   
4:   for  $i \leftarrow 1, p$  do  
5:      $\mathbf{x}_0 = \text{InitialGuessGenerator}(\mathbf{u})$   
6:      $\mathbf{c} = \text{GradientDescent}(n, \epsilon, \gamma, f_1, \mathbf{x}_0)$   
7:      $\mu_c = f_0(\mathbf{c})$   
8:     if  $\mu_c < \mu^*$  then  
9:        $\mathbf{x}^* = \mathbf{c}$  ▷ Better candidate found  
10:       $\mu^* = \mu_c$   
11:     end if  
12:   end for  
13:   return  $\mathbf{x}^*$  ▷ Best candidate returned  
14: end procedure
```

For the general case, that algorithm produces better solutions than single gradient applied to $f_1(\mathbf{x})$. That is explained by the fact that it takes advantage of the increase in speed with the replacement of $\nabla f_0(\mathbf{x})$ by $\nabla f_1(\mathbf{x})$ and the natural selection enforced by $f_0(\mathbf{x})$.

Remark 9. *The differential entropy is can also serve as a second selection function, along with the squared coefficient of variation. But practical computation of the entropy requires special attention because the π -modes cancel at several time points; which can introduce numerical instability with $\log\{\delta_N\}$.*

Chapter 8

Numerical results

The hybrid optimization method discussed in the previous chapter has been applied to the monomials class of indices $r = 0$ and $r = 1$ when $\alpha = 1$. In this section, we present the global performances obtained with these methods in terms of concentration of the resulting distributions and accuracy on the inverted transforms.

Coefficient of variation: numerical results confirm the asymptotic law unveil in [13]. As shown in Figure 8.1, the squared coefficient of variation decrease linearly in a logarithmic scale as the order increases. A general conjecture for that asymptotic

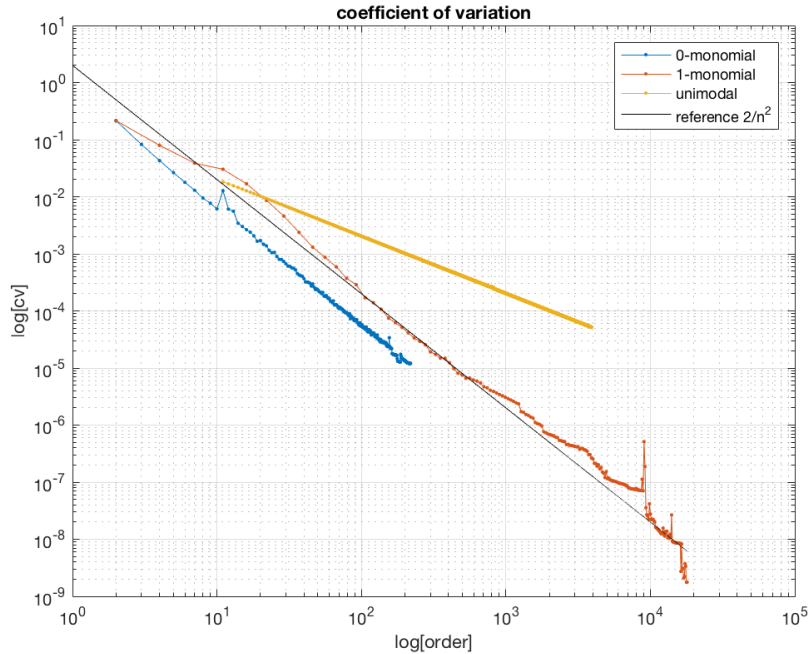


Figure 8.1: Squared coefficient of variation for the monomials class ($r = 0$ and $r = 1$) and the unimodal class

behavior in the monomials class can be summarized as follows

$$scv\left\{\theta_r^{(N,1)}\right\} \sim \frac{a_r}{n^2}$$

From numerical optimizations carried in this work, we have, $a_0 \approx 1.9901$ and $a_1 \approx 2$.

Differential entropy : we have calculated the differential entropy over the first period. Figure 8.2 depicts the expected behavior i.e a decreasing entropy as the order gets larger because all the information about the distribution is getting concentrated at the time point $t = 1$.

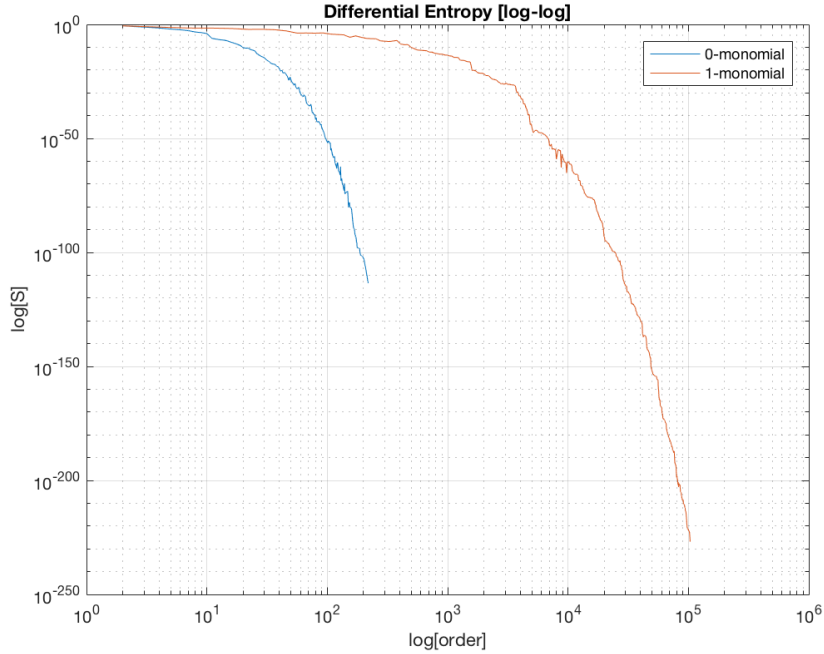
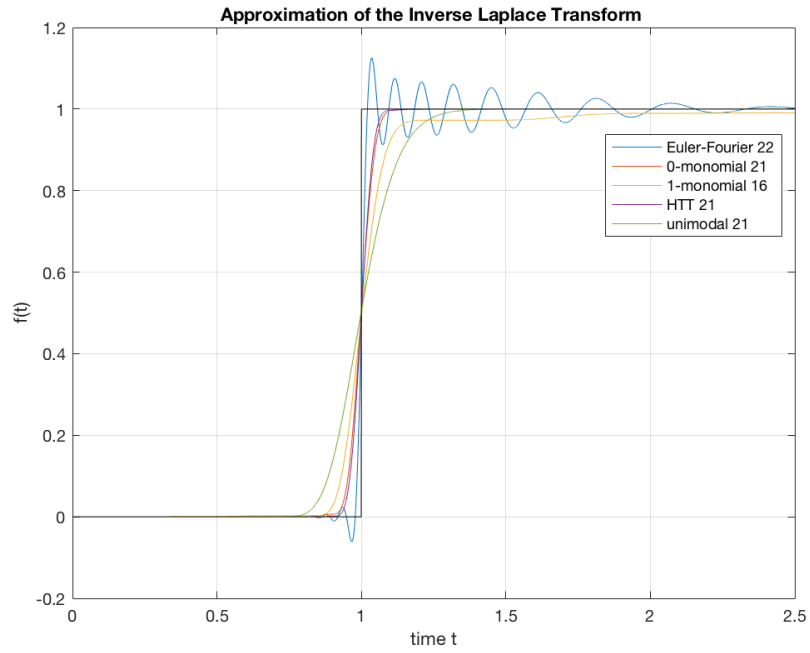


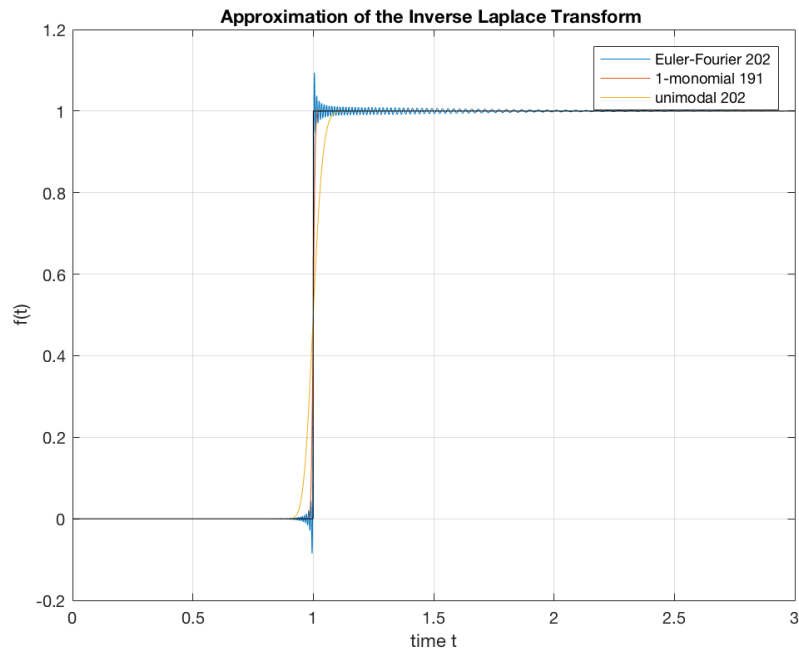
Figure 8.2: Entropy for the monomials class with indices $r = 0$ and $r = 1$

Overall, these two measures tend to confirm that the optimization determines optimal parameters for the approximation of Dirac impulse function. However, we have yet to find out if the functions used numerical optimizations exhibit convex behaviors that the solution that we get are global optimums.

Inversion of Laplace transform of the unit step function



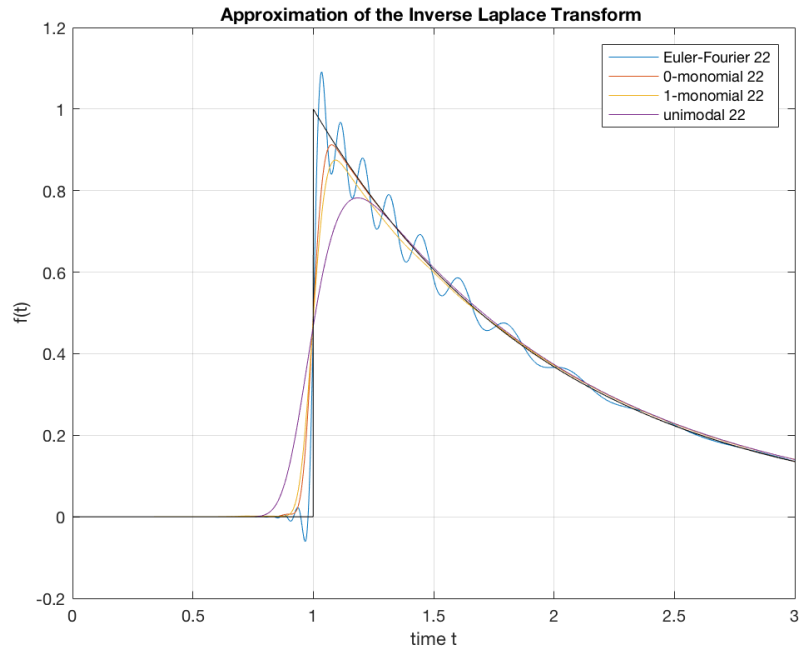
(a) $n=21$



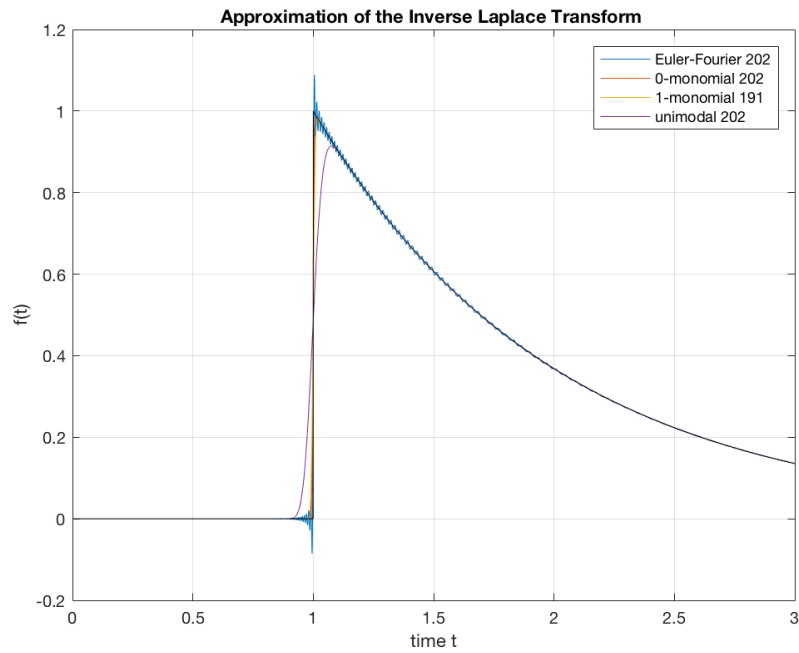
(b) $n=201$

Figure 8.3: Inversion of the step function $f(t) = u(t)$ with the optimized monomials methods, the Euler method and the unimodal method for orders $n = 21$ and $n = 201$.

Inversion of Laplace transform of a delayed exponential function



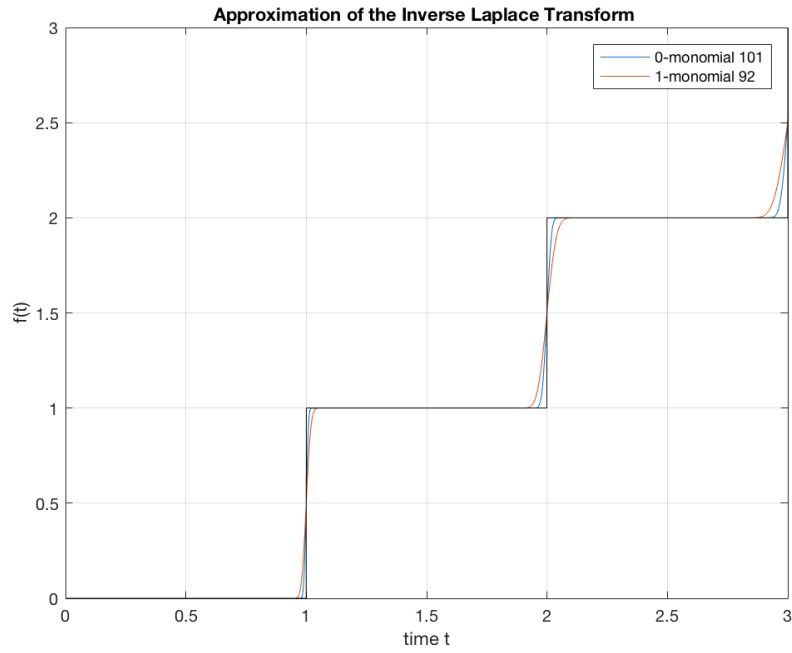
(a) $n=21$



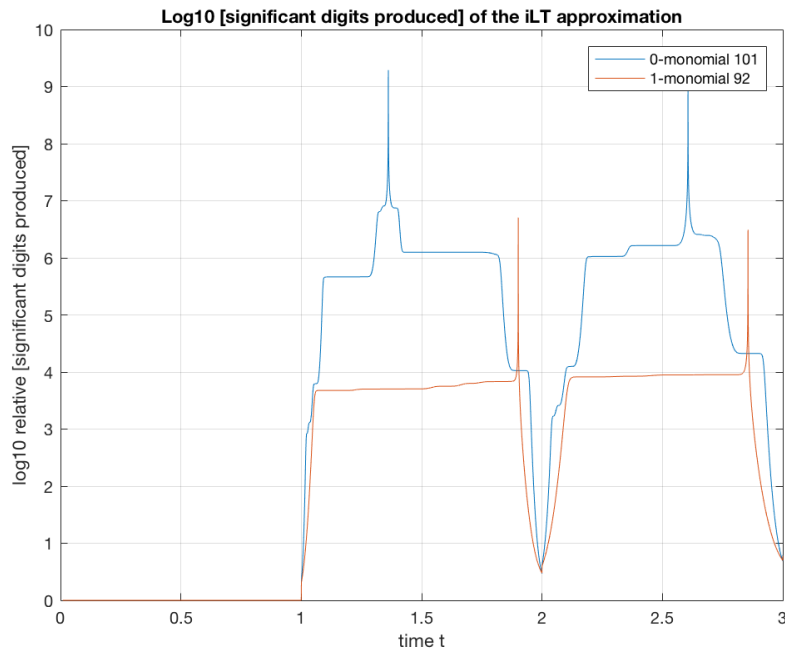
(b) $n=201$

Figure 8.4: Inversion of the delayed step function $f(t) = u(t-1) \cdot e^{-t}$ with the optimized monomials methods, the Euler method and the unimodal method for orders $n = 21$ and $n = 201$.

Inversion of Laplace transform of the floor function



(a) $n=100$



(b) number of significant digits

Figure 8.5: Inversion of the floor function $f(t) = \lfloor t \rfloor$ with the optimized monomials methods, order $n = 100$.

Chapter 9

Conclusion

Summary

In this dissertation, we have constructed a general framework to approximate the shifted-scaled Dirac distribution in order to numerically inverse Laplace transform without positive or negative overshoots. This framework has been constructed around matrix exponential distributions with low coefficient of variation, as a generalization of the inversion method proposed by Horváth, Talyigás and Telek. We have notably studied the specific class of monomials semi-frequencies. The results of this study include the design of two algorithms to calculate the weights and nodes of the distribution functions in the Abate-Whitt framework; the reduction of the complexity in the calculation of the moments of the distribution with the reduced moments triangulation (RMT) and the reduced moments convergence (RMC). Essentially, the initial integral of the positive real axis has been reduced to an finite integral, over one period of the distribution function. Numerical optimization suggests an asymptotic squared coefficient of variation evolving as $O(1/n^2)$. Finally, we have proposed an explicit method to approximate the shifted Dirac impulse function, with an asymptotic squared coefficient of variation behaving like $O(1/n)$, as an attempt to avoid numerical optimization.

Future work

Possible extensions of this work can include in depth examinations of the following areas:

1. The coefficients of variation for the monomials and unimodals classes have been computed numerically but a formal proof for their asymptotic behaviors is still missing.
2. There is no guarantee that the parameters obtained from optimization are the

best candidates. This uncertainty could be cleared by determining if the coefficient of variations is a convex function with respect to these parameters for the density functions.

3. The semi-frequency class of polynomials could be studied by using the fact that any polynomial is a linear combination of monomials; thus that class may inherit some properties of the monomials class.
4. The framework that we have proposed for the approximation of the shifted Dirac distribution can be further extended with the following broader formulation

$$\delta_N(t) = e^{-\lambda t} \prod_{k=1}^N \left(a_k + e^{-\lambda_k t} \cos(w_k t + \phi_k) \right)^{2\alpha_k}$$

Bibliography

- [1] P. G. Harrison and W. J. Knottenbelt, “Passage time distributions in large markov chains,” in *Proceedings of the International Conference on Measurements and Modeling of Computer Systems, SIGMETRICS 2002, June 15-19, 2002, Marina Del Rey, California, USA*, 2002.
- [2] J. Abate and W. Whitt, “A unified framework for numerically inverting laplace transforms,” *INFORMS J. on Computing*, vol. 18, pp. 408–421, Jan. 2006.
- [3] I. Horváth, Z. Talyigás, and M. Telek, “An optimal inverse laplace transform method without positive and negative overshoot - an integral based interpretation,” 2017.
- [4] A. M. Cohen, *Numerical Methods for Laplace Transform Inversion*. Cardiff University: Springer Publishing Company, Incorporated, 1st ed., 2007. <https://www.springer.com/gb/book/9780387282619>.
- [5] P. A. M. Dirac, *The Principles of Quantum Mechanics*. International Series of Monographs on Physics, pub-CLARENDON, fourth ed., 1958.
- [6] L. Kostal, P. Lánský, and O. Pokora, “Measures of statistical dispersion based on shannon and fisher information concepts,” *Inf. Sci.*, vol. 235, pp. 214–223, 2013.
- [7] J. Abate and W. Whitt, “Numerical inversion of laplace transforms of probability distributions,” *INFORMS Journal on Computing*, vol. 7, no. 1, pp. 36–43, 1995.
- [8] D. P. G. Jr., “Observing stochastic processes, and approximate transform inversion,” *Operations Research*, vol. 14, no. 3, pp. 444–459, 1966.
- [9] M. W. Fackrell, *Characterization of Matrix-exponential Distributions*. PhD thesis, University of Adelaide, 11 2003.
- [10] D. Aldous and L. Shepp, “The least variable phase type distribution is erlang,” 1987.

- [11] O. Kramer, *Machine Learning for Evolution Strategies*. Springer International Publishing Switzerland, 2016.
- [12] M. T. P. Buchholz, A. Horváth, “Stochastic petri nets with low variation matrix exponentially distributed firing time,” 2011.
- [13] I. Horváth, O. Sáfár, M. Telek, and B. Zámbo, “Concentrated matrix exponential distributions,” 2016.

Appendix

Ethics summary

As mentioned in Chapter 1, Laplace transform is first and foremost a mathematical tool which can find applications in various fields of science; especially in Statistics, Probability Theory and study of Dynamical Systems. According to the Oxford dictionary, science is an "intellectual and practical activity encompassing the systematic study of the structure and behavior of the physical and natural world through observation and experiment"; which means that science is intrinsically neutral. And as a mathematical tool, the Laplace transform is also neutral and has no inherent ties to biology, leaving beings, personal data, environment, legislation, etc. However, an individual who studies one or multiple sciences may have intentions which lead him to use science for specific purposes. This point is often a concern for theories with potentially controversial interpretations or questionable applications rather than a concern for tools. So overall, Laplace transform and Laplace transform inversion is neutral as for ethics.

Ethics checklist

	Yes	No
Section 1: HUMAN EMBRYOS/FOETUSES		
Does your project involve Human Embryonic Stem Cells?		X
Does your project involve the use of human embryos?		X
Does your project involve the use of human foetal tissues / cells?		X
Section 2: HUMANS		
Does your project involve human participants?		X
Section 3: HUMAN CELLS / TISSUES		

Does your project involve human cells or tissues? (Other than from “Human Embryos/Foetuses” i.e. Section 1)?		X
Section 4: PROTECTION OF PERSONAL DATA		
Does your project involve personal data collection and/or processing?		X
Does it involve the collection and/or processing of sensitive personal data (e.g. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)?		X
Does it involve processing of genetic information?		X
Does it involve tracking or observation of participants? It should be noted that this issue is not limited to surveillance or localization data. It also applies to Wan data such as IP address, MACs, cookies etc.		X
Does your project involve further processing of previously collected personal data (secondary use)? For example Does your project involve merging existing data sets?		X
Section 5: ANIMALS		
Does your project involve animals?		X
Section 6: DEVELOPING COUNTRIES		
Does your project involve developing countries?		X
If your project involves low and/or lower-middle income countries, are any benefit-sharing actions planned?		X
Could the situation in the country put the individuals taking part in the project at risk?		X
Section 7: ENVIRONMENTAL PROTECTION AND SAFETY		
Does your project involve the use of elements that may cause harm to the environment, animals or plants?		X
Does your project deal with endangered fauna and/or flora /protected areas?		X
Does your project involve the use of elements that may cause harm to humans, including project staff?		X

Does your project involve other harmful materials or equipment, e.g. high-powered laser systems?		X
Section 8: DUAL USE		
Does your project have the potential for military applications?		X
Does your project have an exclusive civilian application focus?		X
Will your project use or produce goods or information that will require export licenses in accordance with legislation on dual use items?		X
Does your project affect current standards in military ethics – e.g., global ban on weapons of mass destruction, issues of proportionality, discrimination of combatants and accountability in drone and autonomous robotics developments, incendiary or laser weapons?		X
Section 9: MISUSE		
Does your project have the potential for malevolent/criminal/terrorist abuse?		X
Does your project involve information on/or the use of biological-, chemical-, nuclear/radiological-security sensitive materials and explosives, and means of their delivery?		X
Does your project involve the development of technologies or the creation of information that could have severe negative impacts on human rights standards (e.g. privacy, stigmatization, discrimination), if misapplied?		X
Does your project have the potential for terrorist or criminal abuse e.g. infrastructural vulnerability studies, cybersecurity related project?		X
Section 10: LEGAL ISSUES		
Will your project use or produce software for which there are copyright licensing implications?		X
Will your project use or produce goods or information for which there are data protection, or other legal implications?		X
Section 11: OTHER ETHICS ISSUES		
Are there any other ethics issues that should be taken into consideration?		X