



**HAL**  
open science

# Markov Kernels Models for Natural Exponential Families

Farouk Mselmi

► **To cite this version:**

Farouk Mselmi. Markov Kernels Models for Natural Exponential Families. Statistics [stat]. Université de Bourgogne Franche-Comté, 2023. tel-04069084

**HAL Id: tel-04069084**

**<https://theses.hal.science/tel-04069084>**

Submitted on 14 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Document de synthèse en vue de l'obtention du diplôme

## **D'HABILITATION À DIRIGER DES RECHERCHES**

# **LES MODÈLES DE NOYAUX DE MARKOV POUR FAMILLES EXPONENTIELLES NATURELLES**

**Farouk MSELMI**

Le 03/04/2023

Présenté devant le jury composé de

|                         |            |                                       |
|-------------------------|------------|---------------------------------------|
| Pr. Célestin Kokonendji | Président  | Université de Bourgogne Franche-Comté |
| Pr. Jean-François Dupuy | Rapporteur | Université de Rennes                  |
| Pr. Abderrazek Karoui   | Rapporteur | Université de Carthage                |
| Pr. Yousri Slaoui       | Rapporteur | Université de Poitiers                |
| Pr. Mahdi Louati        | Examineur  | Université de Sfax                    |
| Pr. Angelo Koudou       | Examineur  | Université de Lorraine                |





# Acknowledgements

While it may not be customary to include this page in the preamble of this document, it is my moral obligation and heartfelt appreciation that impels me to begin with it. To neglect expressing my gratitude and recognition to those who have supported and enabled me to complete this work would be a great disservice.

I would like to express my heartfelt gratitude to the jury members of my HDR thesis: Pr. Célestin Kokonendji, Pr. Jean-François Dupuy, Pr. Abderrazek Karoui, Pr. Yousri Slaoui, Pr. Mahdi Louati and Pr. Angelo Koudou who generously gave their time, expertise, and critical feedback. Their insightful comments and suggestions have helped shape my work and have been invaluable in improving its quality.

I am deeply grateful to all those who have contributed to this research project, whether directly or indirectly. From the bottom of my heart, I extend my sincere appreciation to everyone who has offered their assistance, encouragement, and guidance along the way.

I am also grateful to my family and friends who have stood by me through thick and thin and provided me with the support and encouragement I needed to stay the course. Their love, patience, and unwavering support have been a constant source of strength and inspiration to me.



# Contents

|  |           |
|--|-----------|
| <b>Acknowledgements</b>  | <b>i</b>  |
| <b>Introduction</b>  | <b>1</b>  |
| <b>1 Some background</b>   | <b>3</b>  |
| 1.1 Natural exponential families . . . . .                         | 3         |
| 1.2 Markov kernels . . . . .                                       | 5         |
| 1.3 Stable processes . . . . .                                     | 6         |
| <b>2 Univariate Markov kernel for natural exponential family</b>   | <b>9</b>  |
| 2.1 Variance function . . . . .                                    | 10        |
| 2.2 The Lambert class . . . . .                                    | 14        |
| 2.3 The reciprocal class . . . . .                                 | 17        |
| 2.4 Some results on the duality . . . . .                          | 18        |
| 2.5 Generalized linear model . . . . .                             | 24        |
| 2.5.1 Quasi-likelihood function . . . . .                          | 24        |
| 2.5.2 Estimation . . . . .   | 25        |
| 2.5.3 Residual analysis . . . . .                                  | 27        |
| 2.6 Data analyses . . . . .  | 28        |
| 2.6.1 Data 1 . . . . .   | 28        |
| 2.6.2 Data 2 . . . . .   | 30        |
| <b>3 Multivariate Markov kernel for natural exponential family</b> | <b>33</b> |
| 3.1 Variance and generalized variance functions . . . . .          | 33        |

|          |  |           |
|----------|--|-----------|
| 3.2      | Some illustrative examples . . . . .                           | 35        |
| 3.3      | Associated kernel estimations for variance functions . . . . . | 38        |
| 3.4      | Multivariate generalized linear model . . . . .                | 41        |
| 3.5      | Applications to multivariate daily exchange rates . . . . .    | 43        |
| <b>4</b> | <b>Characterizations of stable and related processes</b>       | <b>45</b> |
| 4.1      | Characterization of multivariate stable processes . . . . .    | 45        |
| 4.2      | Characterization of the inverse stable subordinator . . . . .  | 47        |
|          | <b>Perspectives</b>  | <b>49</b> |
|          | <b>Bibliography</b>  | <b>51</b> |
|          | <b>Appendix 1</b>  | <b>59</b> |
|          | <b>Appendix 2</b>  | <b>67</b> |

# List of publications

## In Journal Papers peer-reviewed

- [1a ] S. Charfi, **F. Mselmi**, *Modeling exchange rate volatility: application of GARCH models with a Normal Tempered Stable distribution*. Quantitative Finance and Economics (AIMS Press). Vol 6, No 2, 206-222. 2022.
- [2a ] **F. Mselmi**, *Generalized linear model for subordinated Lévy processes*. Scandinavian Journal of Statistics (Wiley). Vol 49, No 2, 772-801. 2021.
- [3a ] **F. Mselmi**, *Approximation of the quasi-deviance function for the time-changed Lévy processes by the first-exit time of the inverse Gaussian subordinator*. Stat (Wiley), Vol 10, No 1, 1-9. 2021.
- [4a ] M. Louati, A. Masmoudi, **F. Mselmi**, *The normal tempered stable regression model*. Communication in Statistics: Theory & Methods (Taylor & Francis), Vol 49, No 2, 500-512. 2020.
- [5a ] **F. Mselmi**, *Lévy processes time-changed by the first-exit time of the inverse Gaussian subordinator*. Filomat, Vol 32, No 7, 2545-2552. 2018.
- [6a ] **F. Mselmi**, C.C. Kokonendji, M. Louati, A. Masmoudi, *Generalized variance functions for infinitely divisible mixture distributions*. Mediterranean Journal of Mathematics (Springer), Vol 14, No 4:165, 1-19. 2018.



- [7a ] **F. Mselmi**, *Characterization of the inverse stable subordinator*. Statistics and Probability Letters (Elsevier), Vol 140, 37-43. 2018.
- [8a ] **F. Mselmi**, *The stable processes on symmetric matrices*. Arab Journal of Mathematical Sciences (Elsevier), Vol 24, No 1, 63-69. 2018.
- [9a ] M. Louati, A. Masmoudi, **F. Mselmi**, *Characterizations of multivariate stable processes*. Lithuanian Mathematical Journal (Springer), Vol 57, No 1, 59-68. 2017.
- [10a ] M. Louati, A. Masmoudi, **F. Mselmi**, *Multivariate normal  $\alpha$ -stable exponential families*. Mediterranean Journal of Mathematics (Springer), Vol 13, No 3, 1307-1323. 2015.

#### **Book Chapters peer-reviewed**

- [1b ] M. Louati, A. Masmoudi, **F. Mselmi**, *Gamma stopping and drifted stable processes*. Applied Mathematics in Tunisia, Springer Proceedings in Mathematics and Statistics (Springer), 131, 223-232. 2015.

#### **Submitted to Journal Papers peer-reviewed**

- [1s ] **F. Mselmi**, *On the duality for mixture models*. 2022.
- [2s ] **F. Mselmi**, C. Kokonendji, *Multivariate generalized linear models for Markov kernels with (un)known variance functions*. 2022.
- [3s ] **F. Mselmi**, *A note on the parameters estimation of the normal tempered stable model using the tail regression method*. 2022.

#### **In Preparation for Journal Papers peer-reviewed**

- [1p ] **F. Mselmi**, *Natural exponential family for the convolution of two probability measures*. 2022-2023.
- [2p ] F. Fourati, **F. Mselmi**, *The multivariate Markov-Krein transform*. 2022-2023.

# List of Figures

|     |   |    |
|-----|---|----|
| 2.1 | The NEFs classes of some Markov kernels. . . . .  | 16 |
| 3.1 | Plots of link and variance as well as their smoothed functions by the Gaussian kernel for some normal mixture distributions (NIG, NG and NP) with $\lambda \neq 1$ and $\lambda = 1$ . Note: NG means normal gamma or variance gamma. . . . . | 40 |



# List of Tables

|     |   |    |
|-----|---|----|
| 2.1 | Markov kernels and their reciprocals . . . . .  | 18 |
| 2.2 | Summary of some results about some known NEFs. . . . .  | 20 |
| 2.3 | The estimated values of the (N-IG)-GLM. . . . .   | 29 |
| 2.4 | The estimated values of the (N-G)-GLM. . . . .  | 30 |
| 2.5 | The estimated values of the (P-IG)-GLM. . . . .   | 31 |
| 2.6 | The estimated values of the (P-G)-GLM. . . . .  | 32 |
| 3.1 | Mean squared errors for three normal mixture models (NIG, NG and NP) with different values of $\lambda$ and fixed $m \in [-1.5, 1.5]$ and $h = 0.8$ in the univariate standard Gaussian kernel. . . . . | 41 |
| 3.2 | Mean squared errors for three bivariate normal mixture models (NIG, NG and NP) with fixed $\lambda = 1$ and $\mathbf{H} = \mathbf{Diag}_2(h, h)$ in the bivariate standard Gaussian kernel. . . . .     | 41 |
| 3.3 | On regression coefficients of MNIG-MGLM by a multivariate Gaussian kernel for the full dataset. . . . .   | 44 |
| 3.4 | On the nonparametric regression coefficients of MNIG-MGLM by a multivariate Gaussian kernel for the full dataset. . . . .   | 44 |



# Introduction

The research works presented below represent a part of the general framework of the class of natural exponential families. This class has drawn considerable attention of researchers and numerous works have been realized on probability theory and statistical modeling. This gives us the opportunity to develop some theoretical and applied results that study the class of the Markov kernels distributions. In the particular case, this class represents a one of variance-mean mixture models which widely used in the statistical field and, in terms of stochastic processes, it could be considered as a class of subordinated Lévy processes.

My research has focused on studying three interrelated aspects of Markov kernels using the natural exponential family theory, building upon the findings of my Ph.D. thesis. The research axes represent an extension of my earlier work on the normal  $\alpha$ -stable model, as published in [ 4a], [ 9a], and [ 10a]. To be more specific, these research axes can be divided into the following parts:

- Univariate Markov kernel natural exponential family: in this part, we characterize the Markov kernel natural exponential family through its variance function. Besides, we determine some related probabilistic properties such as the reciprocity and the duality. We illustrate the results by some examples. In statistical point of view, we analyze the proposed generalized linear model through the Markov kernel quasi-likelihood function. We apply the results on the daily exchange rate returns of the Tunisian Dinar against

the US Dollar and the damage incidents of ships. These results and other related ones are given in [1a ], [2a ], [3a ], [4a ], [5a ] and [1s ] in the list of publications.

- Multivariate Markov kernel natural exponential family: this part is devoted to the study of the theory of multivariate generalized linear models for Markov kernels with known and also unknown link and variance functions. Theoretical results are first shown for the parametric case. To reduce the assumptions in a full parametric case, we consider both link and variance functions to be unknown but smooth. These functions are then estimated nonparametrically, through associated kernels as smoothers with respect to the appropriate support. Hence, we obtain a three-level approach to this semiparametric model from link function, variance function and the vector of regression coefficients in the linear predictor of the model. Also, we theoretically suggest total deviances and other deduced criteria, as Akaike and Bayesian information criteria. Finally, we apply the results on the multivariate daily exchange rate returns. These results and other related ones are given in [6a ], [10a ] and [2s ] in the list of publications.
- Characterizations of stable and related processes: in this part, we characterize the classes of multivariate stable processes and the inverse stable subordinators through some independence properties. This led us to determine other characterizations related to the Markov kernel models. These results and other related ones are given in [7a ], [8a ], [9a ] and [1b ] in the list of publications.

# Chapter 1

## Some background

In this chapter, we recall some definitions and notations dealing with the Natural Exponential Families (NEFs) and their variance functions. Furthermore, we introduce the classes of Markov kernels and stable processes.

### 1.1 Natural exponential families

The NEF generated by the finite positive measure  $\mu$  on  $\mathbf{R}^d$ , with  $d \geq 1$ , is the set of probability distributions

$$F = F(\mu) = \{P(\theta, \mu)(d\mathbf{x}) = \exp[\langle \theta, \mathbf{x} \rangle - \log L_\mu(\theta)]\mu(d\mathbf{x}); \theta \in \Theta(\mu)\},$$

where  $\langle \cdot, \cdot \rangle$  represents the scalar product and  $L_\mu(\theta) = \int_{\mathbf{R}^d} \exp(\langle \theta, \mathbf{x} \rangle) \mu(d\mathbf{x})$  defines the Laplace transform of  $\mu$  which is finite for  $\theta$  in the domain  $\Theta(\mu)$ . The so-called cumulant function of  $\mu$  is defined by  $K_\mu(\theta) = \log L_\mu(\theta)$ . The map

$$\Theta(\mu) \longrightarrow M_{F(\mu)}; \theta \longmapsto m = K'_\mu(\theta) = \int_{\mathbf{R}^d} \mathbf{x} P(\theta, \mu)(d\mathbf{x})$$

is a diffeomorphism between  $\Theta(\mu)$  and its image  $M_{F(\mu)}$ , called the domain of the means of  $F(\mu)$ . The inverse function of the first derivative  $K'_\mu$  is denoted by  $\psi_\mu$  and is called the link function. The second derivative  $K''_\mu$  defines the covariance operator of  $P(\theta, \mu)$  and is equal to

$$K''_\mu(\theta) = \int_{\mathbf{R}^d} \mathbf{x} \otimes \mathbf{x} P(\theta, \mu)(d\mathbf{x}) - K'_\mu(\theta) \otimes K'_\mu(\theta),$$



where  $\mathbf{x} \otimes \mathbf{x}$  is the matrix  $\mathbf{x}^\top \mathbf{x}$  obtained from the product of the vector  $\mathbf{x}$  and its transpose  $\mathbf{x}^\top$  which is a column vector. The variance function of the NEF  $F(\mu)$  is defined on  $M_{F(\mu)}$  by

$$m \mapsto V_{F(\mu)}(m) = K''_{\mu}(\psi_{\mu}(m)) = [\psi'_{\mu}(m)]^{-1}.$$

Its importance emanates from the fact that it characterizes the family  $F$  within the class of all NEFs. More exactly, if  $F_1$  and  $F_2$  are two NEFs such that the variance functions  $V_{F_1}(m)$  and  $V_{F_2}(m)$  are equal on a nonempty open set of the intersection of the means domains  $M_{F_1} \cap M_{F_2}$ , then  $F_1 = F_2$ . This means that the knowledge of the NEF is obtained by the knowledge of its variance function, and conversely; see, e.g., [Letac \(1992\)](#) for more details. In this context, many classifications of the NEFs according to the form of their variance functions have been established; see, e.g., [Casalis \(1996\)](#) and [Kokonendji and Moypemna Sembona \(2018\)](#) for two classes of multivariate NEFs and references. Thus, the generalized variance function is the determinant of the variance function and it can be successively defined by

$$\det K''_{\mu}(\theta) = \det K''_{\mu}(\psi_{\mu}(m)) = \det V_{F(\mu)}(m).$$

One can refer to [Boubacar Mainassara and Kokonendji \(2014\)](#) and [Mselmi \*et al.\* \(2018\)](#) for some details, and [Kokonendji \*et al.\* \(2020\)](#) for characterization attempts through a Monge-Ampère property. Otherwise, a concise guide to statistical inference of multivariate NEFs can be found in [Bedbur and Kamps \(2021\)](#).

In the univariate case, the notion of reciprocity in NEFs was introduced and characterized by [Letac \(1986\)](#). In fact, let  $\mu$  and  $\nu$  be two positive measures in  $\mathbf{R}$  and  $\tilde{\Theta}(\mu) = \{\theta \in \Theta(\mu) \text{ such that } K'_{\mu}(\theta) > 0\}$ . The pair  $(\mu, \nu)$  is called reciprocal if

- i.*  $\tilde{\Theta}(\mu)$  and  $\tilde{\Theta}(\nu)$  are nonempty.
- ii.* The image of  $\tilde{\Theta}(\mu)$  by  $\theta \mapsto -K_{\mu}(\theta)$  is  $\tilde{\Theta}(\nu)$ , and the image of  $\tilde{\Theta}(\nu)$  by  $\theta \mapsto -K_{\nu}(\theta)$  is  $\tilde{\Theta}(\mu)$ .
- iii.*  $-K_{\nu}(-K_{\mu}(\theta)) = \theta$  for all  $\theta \in \tilde{\Theta}(\mu)$ .

## 1.2 Markov kernels

The concept of reciprocity in NEFs was characterized by variance functions. This characterization is given in the following theorem.

**Theorem 1.1.1** (*Letac and Mora (1990), Theorem 5.2*) *Let  $F(\mu)$  and  $F(\nu)$  be two NEFs in  $\mathbf{R}$ , and denote  $\widetilde{M}_{F(\mu)} = M_{F(\mu)} \cap (0, +\infty)$  and  $\widetilde{M}_{F(\nu)} = M_{F(\nu)} \cap (0, +\infty)$ . Then  $(F(\mu), F(\nu))$  is a reciprocal pair if and only if the following conditions hold:*

- i.  $\widetilde{M}_{F(\mu)}$  and  $\widetilde{M}_{F(\nu)}$  are nonempty.*
- ii.  $m \rightarrow \frac{1}{m}$  is a bijective mapping from  $\widetilde{M}_{F(\mu)}$  onto  $\widetilde{M}_{F(\nu)}$ .*
- iii.  $V_{F(\mu)}(m) = m^3 V_{F(\nu)}(1/m)$  for all  $m \in \widetilde{M}_{F(\mu)}$ .*

## 1.2 Markov kernels

Now, we briefly introduce the Markov kernel [e.g., [Çınlar \(2011, pp. 37-38\)](#)] of interest as a mixture of distributions. Let  $(\mathcal{Y}, Y)$  be a couple of random vectors on  $(0, \infty) \times \mathbf{R}^{d-1}$  ( $d \geq 2$ ) with distribution  $\eta(dy, dy) = Q_{\mathbf{y}}(dy)\mu_1(dy)$  such that  $\mu_1$  represents a mixing measure with  $\text{Supp}(\mu_1) \subset (0, \infty)$  and the kernel  $Q_{\mathbf{y}}$  defines a convolution semigroup of the  $\mathbf{y}^{th}$ -convolution power of  $Q_1$ . The mixture distribution of interest is the marginal distribution of  $\eta(dy, dy)$  with respect the second variable; and, it is denoted by  $\mu(dy) = \int_0^{+\infty} Q_{\mathbf{y}}(dy)\mu_1(dy)$ . For all  $\theta \in \{(\theta_1, \theta_2) \in \mathbf{R} \times \Theta(Q_1); \theta_1 + K_{Q_1}(\theta_2) \in \Theta(\mu_1)\}$ , the cumulant function of  $\eta$  is equal to  $K_{\eta}(\theta) = K_{\mu_1}(\theta_1 + K_{Q_1}(\theta_2))$ . This holds because its Laplace transform can be expressed as follows:

$$\begin{aligned}
 L_{\eta}(\theta) &= \mathbf{E}(e^{\theta_1 \mathcal{Y} + \langle \theta_2, Y \rangle}) = \int_0^{\infty} \int_{\mathbf{R}^{d-1}} \exp[\theta_1 \mathbf{y} + \langle \theta_2, y \rangle] Q_{\mathbf{y}}(dy) \mu_1(dy) \\
 &= \int_0^{\infty} e^{\theta_1 \mathbf{y}} \left( \int_{\mathbf{R}^{d-1}} e^{\langle \theta_2, y \rangle} Q_{\mathbf{y}}(dy) \right) \mu_1(dy) = \int_0^{\infty} e^{\mathbf{y}(\theta_1 + K_{Q_1}(\theta_2))} \mu_1(dy) \\
 &= L_{\mu_1}(\theta_1 + K_{Q_1}(\theta_2)) = \exp[K_{\mu_1}(\theta_1 + K_{Q_1}(\theta_2))].
 \end{aligned} \tag{1.1}$$

By taking  $\theta_1 = 0$  in (1.1), we get the Laplace transform of the mixture distributions

$$L_{\mu}(\theta_2) = L_{\eta}(0, \theta_2) = \exp[K_{\mu_1}(K_{Q_1}(\theta_2))], \tag{1.2}$$

with  $\theta_2 \in \Theta(\mu) := \Theta(Q_1) \cap K_{Q_1}^{-1}(\Theta(\mu_1))$ .

In the context of NEFs, let us recall the notion of cut in NEF introduced by [Barndorff-Nielsen and Koudou \(1995\)](#). In fact, Let  $pro : (0, \infty) \times \mathbb{R}^{d-1} \rightarrow (0, \infty); (\mathbf{y}, y) \mapsto \mathbf{y}$  be a canonical projection on  $(0, \infty)$ . The NEF  $F = F(\eta)$  has a cut on  $(0, \infty)$ , if  $pro(F)$  is also a NEF on  $(0, \infty)$ . The following lemma gives some results about the notion of cut in NEF.

**Lemma 1.2.1** (*Barndorff-Nielsen and Koudou (1995)*) *Let  $pro(F) = \{pro(\eta); \eta \in F\}$ . The following statements are equivalent:*

1. *The family  $pro(F)$  is a NEF on  $(0, \infty)$  ( $F$  has a cut on  $(0, \infty)$ ).*
2. *For all  $\theta \in \Theta(Q_{\mathbf{y}})$ ,  $\mathbf{y} \mapsto K_{Q_{\mathbf{y}}}(\theta)$  is an affine function on  $(0, \infty)$ . There exist maps  $\mathbb{H} : \Theta(Q_{\mathbf{y}}) \rightarrow \mathbf{R}$  and  $\mathbb{K} : \Theta(Q_{\mathbf{y}}) \rightarrow \mathbf{R}$ , such that the conditional cumulant function of the distribution  $Q_{\mathbf{y}}$  is given by  $K_{Q_{\mathbf{y}}}(\theta) = \mathbb{K}(\theta) - \mathbf{y}\mathbb{H}(\theta)$ ,  $\theta \in \Theta(Q_{\mathbf{y}})$ .*

### 1.3 Stable processes

Before defining the class of stable processes, we need first to define a larger class which is the Lévy one.

**Definition 1.3.1** *A stochastic process  $(\mathfrak{L}(t))_{t \geq 0}$  on  $\mathbf{R}^d$  is a Lévy process, if the following conditions are satisfied:*

1.  *$\mathfrak{L}(0) = 0$  almost surely (a.s.).*
2. *For any choice of  $n \geq 1$  and  $0 < t_1 < \dots < t_n$ , the random variables  $\mathfrak{L}(t_1), \mathfrak{L}(t_2) - \mathfrak{L}(t_1), \dots, \mathfrak{L}(t_n) - \mathfrak{L}(t_{n-1})$  are independent.*
3. *The distribution of  $\mathfrak{L}(t+s) - \mathfrak{L}(s) \stackrel{\mathcal{L}}{=} \mathfrak{L}(t)$ , where  $\stackrel{\mathcal{L}}{=}$  denotes the equality in distribution.*
4. *For every  $\omega \in \Omega$ ,  $t \mapsto \mathfrak{L}(t, \omega)$  is càdlàg ( i.e., right-continuous in  $t \geq 0$  and has left-limits in  $t > 0$ .)*

A stable process  $(\mathbb{S}(t))_{t \geq 0}$ , with a parameter  $\alpha \in (0, 2]$ , is a Lévy process distributed by a stable law. In fact, a random variable  $\mathbb{S}$  on  $\mathbf{R}^d$  is called  $\alpha$ -stable if, for each  $n \geq 2$ , there exist a vector  $f_n$  and  $n$  independent copies  $\mathbb{S}_1, \mathbb{S}_2, \dots, \mathbb{S}_n$  of  $\mathbb{S}$ , such that

$$\mathbb{S}_1 + \mathbb{S}_2 + \dots + \mathbb{S}_n \stackrel{\mathcal{L}}{=} n^{1/\alpha} \mathbb{S} + f_n.$$

### 1.3 Stable processes

A Lévy process  $\mathfrak{L} = (\mathfrak{L}(t))_{t \geq 0}$  valued in  $\mathbb{R}^d$  is said strictly stable of parameter  $\alpha \in (0, 2]$  if  $\mathfrak{L}(t) \stackrel{\mathcal{L}}{=} t^{1/\alpha} \mathfrak{L}(1)$ , when  $\alpha \neq 1$ . Furthermore,  $\mathfrak{L} = (\mathfrak{L}(t))_{t \geq 0}$  is said strictly 1-stable if, there exists  $\mathbf{k} \in \mathbf{R}^d$  such that  $\mathfrak{L}(t) \stackrel{\mathcal{L}}{=} t(\mathfrak{L}(1) - \log(t)\mathbf{k})$ .

In the univariate case and according to [Nolan \(2008\)](#), page 94, the Laplace transform of the stable process  $\mathbb{S}_t := \mathbb{S}(t)$  is given, for all  $\theta < 0$ , by

$$L_{\mathbb{S}_t}(\theta) = \begin{cases} \exp(-c_\alpha t(-\theta)^\alpha), & \alpha \in (0, 1) \cup (1, 2] \text{ and } c_\alpha > 0 \\ -c(-\theta) \log(-\theta), & \alpha = 1 \text{ and } c > 0. \end{cases}$$

A Lévy process increasing in time is called a subordinator. As example of subordinator, we could cite the stable one with  $\alpha \in (0, 1)$ . The inverse  $\alpha$ -stable subordinator  $(E(t))_{t > 0}$  defined, for any fixed  $t > 0$ , by  $E(t) = \inf\{z > 0 : \mathbb{S}(z) > t\}$  represents the first-exit time of the  $\alpha$ -stable subordinator  $(\mathbb{S}(t))_{t > 0}$ , with  $\alpha \in (0, 1)$ . They are widely used as a time change subordinator (see [Kumar et al. \(2011\)](#) and [Meerschaert et al. \(2011\)](#)). Recently, [Meerschaert and Scheffler \(2008\)](#) have showed that the Laplace transform of the inverse stable subordinator is expressed as follows

$$L_{E(t)}(\theta) = c_\alpha (-\theta)^{\alpha-1} e^{-c_\alpha t(-\theta)^\alpha}, \quad \forall \theta < 0.$$

Moreover, according to [Meerschaert and Scheffler \(2004, 2008\)](#), we have the following relations  $\{E(t) \leq z\} \stackrel{\mathcal{L}}{=} \{\mathbb{S}(z) \geq t\}$  and  $E(t) \stackrel{\mathcal{L}}{=} t^\alpha E(1)$ .



## Chapter 2

# Univariate Markov kernel for natural exponential family

In this chapter, we determine the link and variance functions of the natural exponential family generated by the class of Markov kernels. In this framework, we introduce a class of variance functions that depends on the Lambert function. We call it the Lambert class. The concept of reciprocity in natural exponential families was given in order to obtain an exponential family from another one. In this context, we get the reciprocal class of the natural exponential family generated by the class of Markov kernels. Besides, we study the notion of duality in the class of Markov kernels and we give some illustrative examples. It is well known that the variance function represents an essential element for the determination of the quasi-likelihood and deviance functions. Then, we use the expression of our variance function in order to maintain them. This leads us to analyze the proposed generalized linear model. We illustrate some of our models with applications to the daily exchange rate returns of the Tunisian Dinar against the US Dollar and the damage incidents of ships. The proofs and the auxiliary results of Section 2.4 are given in Appendix 1. Other proofs are given in the work of [Mselmi \(2021a\)](#).

## 2.1 Variance function

In this section, we determine the link function of the class of Markov kernels and we characterize its NEF by the associated variance function. We illustrate our results by the most known examples of Markov kernels in the literature.

In the following theorem, we state the link and variance functions which depend only on a first-order differential equation.

**Theorem 2.1.1** *The link and the variance functions of the class of Markov kernels are given, for all  $m \in M_{F(\mu)} = M_{F(Q_1)}$ , by*

1.  $\psi_\mu(m) = \psi_{Q_1} \left( \frac{m}{g(m)} \right)$
2.  $V_{F(\mu)}(m) = \left( \frac{m}{g(m)} \right)^2 V_{F(\mu_1)}(g(m)) + g(m) V_{F(Q_1)} \left( \frac{m}{g(m)} \right)$

where  $g : M_{F(\mu)} \rightarrow (0, +\infty)$  is a unique continuous function satisfying

$$g'(m) = \frac{mg(m) \left( V_{F(Q_1)} \left( \frac{m}{g(m)} \right) \right)^{-1}}{(g(m))^3 \left( V_{F(\mu_1)}(g(m)) \right)^{-1} + m^2 \left( V_{F(Q_1)} \left( \frac{m}{g(m)} \right) \right)^{-1}}$$

with initial condition  $g(K'_{Q_1}(0)K'_{\mu_1}(0)) = K'_{\mu_1}(0)$ .

Next, we express some examples that illustrate our results. More precisely, we treat the cases of the normal Markov kernel mixed by the inverse Gaussian, the gamma, and the Poisson distributions. Besides, we study the examples of the gamma and Poisson Markov kernels mixed by the inverse Gaussian distribution.

*Example 1: Normal inverse Gaussian distribution*

The normal Markov kernel represents the most useful Markov kernel in scientific research and it has applications in nearly all areas. Note that, in terms of stochastic process, it represents the distribution of a subordinated Brownian motion. In this context, many works have focused on the study of this Markov kernel (see [Ganti et al. \(2009\)](#), [Kim et al. \(2009, 2012\)](#)),

### 2.1 Variance function

Madan and Yor (2008), and Song and Vondraček (2003)). It has been applied on the modeling of stochastic volatility (see Barndorff-Nielsen (1997), Barndorff-Nielsen and Levendorskii (2001), Barndorff-Nielsen and Shephard (2001) and Rydberg (1996)).

In this case, the distributions  $Q_1$  and  $\mu_1$  are normal and inverse Gaussian respectively. According to McCullagh and Nelder (1989) Page 30, we have  $\psi_{Q_1} = id$ ,  $V_{F(Q_1)} = 1$  and  $V_{F(\mu_1)}(m) = m^3$ . Inserting these link and variance functions in Theorem 2.1.1, we obtain  $\psi_\mu(m) = \frac{m}{g(m)}$ ,  $V_{F(\mu)}(m) = g(m) [m^2 + 1]$  and the function  $g$  satisfies the differential equation  $g'(m) = \frac{mg(m)}{1 + m^2}$  with initial condition  $g(K'_{Q_1}(0)K'_{\mu_1}(0)) = K'_{\mu_1}(0)$ . Solving the differential equation and inserting the solution in the expression of  $\psi_\mu$  and  $V_{F(\mu)}$ , we get  $g(m) = c\sqrt{1 + m^2}$ , where  $c = \frac{K'_{\mu_1}(0)}{\sqrt{1 + (K'_{\mu_1}(0)K'_{Q_1}(0))^2}}$ ,  $\psi_\mu(m) = \frac{m}{c\sqrt{1 + m^2}}$  and  $V_{F(\mu)}(m) = c(1 + m^2)^{3/2}$ .

The variance function of the normal inverse Gaussian type was obtained by Kokonendji (1995) in Example 2.5. Note that Louati *et al.* (2015a) have determined it, in the multivariate case, in Remark 3.8.

#### Example 2: Normal gamma distribution

The normal gamma distribution, also known as the variance gamma distribution, is very useful in the pricing options and applied in the modeling of credit risk in structural models (see Fiorani *et al.* (2010), Hirsra and Madan (2003), Madan *et al.* (1998) and Madan and Seneta (1990)). In this case  $g(m) = \frac{1}{2} (c^2 + c\sqrt{2m^2 + c^2})$ , where  $c = \frac{\sqrt{2}K'_{\mu_1}(0)}{\sqrt{(K'_{\mu_1}(0)K'_{Q_1}(0))^2 + 2K'_{\mu_1}(0)}}$ ,  $\psi_\mu(m) = \frac{2m}{c^2 + c\sqrt{2m^2 + c^2}}$ , and  $V_{F(\mu)}(m) = \frac{1}{2} (2m^2 + c^2) + \frac{c}{2} \sqrt{2m^2 + c^2}$ .

#### Example 3: Normal Poisson distribution



In this example, the function  $g(m)$  equal to  $c \exp\left(\frac{1}{2}\mathcal{W}\left(\left(\frac{m}{c}\right)^2\right)\right) = \frac{|m|}{\sqrt{\mathcal{W}\left(\left(\frac{m}{c}\right)^2\right)}}$ , where  $c = K'_{\mu_1}(0) \exp\left(-\frac{1}{2}(K'_{Q_1}(0))^2\right)$  and  $\mathcal{W}$  is the Lambert function. Notice that this special function is the solution of the equation  $u = \mathcal{W}(u)e^{\mathcal{W}(u)}$  and its inverse function is given by  $\mathcal{W}^{-1}(u) = ue^u$ . Recently, this function has been used in many probabilistic research works. In fact, some authors have noted the relationship between the mixture of Poisson distribution and the Lambert function (see [Bondesson and Steutel \(2004\)](#) and [Pakes \(2011\)](#)). Other research works investigated this function in the study of some other probability distributions such as the log-normal and the Kendall-Ressel ones (see [Asmussen et al. \(2014\)](#) and [Vinogradov \(2011\)](#)). The link and the variance functions of this distribution are  $\psi_{\mu}(m) = \frac{m}{c} \exp\left(-\frac{1}{2}\mathcal{W}\left(\left(\frac{m}{c}\right)^2\right)\right)$  and  $V_{F(\mu)}(m) = c \exp\left(\frac{1}{2}\mathcal{W}\left(\left(\frac{m}{c}\right)^2\right)\right) + m^2 \frac{1}{c} \exp\left(-\frac{1}{2}\mathcal{W}\left(\left(\frac{m}{c}\right)^2\right)\right)$ .

*Example 4: Gamma inverse Gaussian distribution*

This distribution was applied in the statistical modeling of the loss functions and the life testing (see [Bhattacharya and Kumar \(1986\)](#) and [Frangos and Karlis \(2004\)](#)). Furthermore, [Gómez-Déniz et al. \(2013\)](#) have determined some of its properties and have given some of its applications. In this example, we obtain the following results  $g(m) = cm \exp\left(\frac{1}{2}\mathcal{W}\left(\frac{1}{(cm)^2}\right)\right)$ , where  $c = (K'_{Q_1}(0))^{-1} \exp\left(-\frac{1}{2(K'_{\mu_1}(0))^2}\right)$ ,  $\psi_{\mu}(m) = c \exp\left(\frac{1}{2}\mathcal{W}\left(\left(\frac{1}{cm}\right)^2\right)\right)$  and  $V_{F(\mu)}(m) = m^2 \left[ \frac{1}{cm} \exp\left(-\frac{1}{2}\mathcal{W}\left(\left(\frac{1}{cm}\right)^2\right)\right) + cm \exp\left(\frac{1}{2}\mathcal{W}\left(\left(\frac{1}{cm}\right)^2\right)\right) \right]$ .

*Example 5: Poisson inverse Gaussian distribution*

The Poisson inverse Gaussian distribution has a lot of attention in the statistical field and it was used in the study of regression models (see [Dean et al. \(1989\)](#), [Holla \(1967\)](#), [Ord and Whitmore \(1986\)](#), [Rigby et al. \(2008\)](#), [Tremblay \(1992\)](#) and [Willmot \(1987\)](#)). In this example, we have  $g(m) = \frac{c}{2} \left( cm + \sqrt{c^2 m^2 + 2} \right)$ , where  $c = \frac{\sqrt{2}K'_{\mu_1}(0)}{\sqrt{2(K'_{\mu_1}(0))^2 K'_{Q_1}(0) + 1}}$ ,

### 2.1 Variance function

$\psi_\mu(m) = \log\left(\frac{2m}{c^2m + c\sqrt{c^2m^2 + 2}}\right)$  and  $V_{F(\mu)}(m) = \frac{m}{2} [c^2m^2 + 2] + cm^2\sqrt{c^2m^2 + 2}$ . Note that [Kokonendji \(1995\)](#) has determined the form of the variance function of the Poisson inverse Gaussian distribution in Example 2.3.

**Remarks 2.1.2** 1. Several classifications of the NEFs according to the form of their variance functions have been given. More exactly, [Morris \(1982\)](#) has determined the real quadratic NEFs that have the variance function of the form of a polynomial with a degree less than or equal to two. Besides, [Letac and Mora \(1990\)](#) have studied the real cubic NEFs that have the variance function which constitutes a polynomial with a degree less than or equal to three. The quadratic and the cubic NEFs are called also the Morris and the Letac-Mora classes. More larger class than the Morris and the Letac-Mora ones was described by [Letac \(1992\)](#) Chapter 5 and called the Babel class. The form of the variance function of this latter class is given by  $P\Delta + Q\sqrt{\Delta}$ , where  $P$ ,  $Q$  and  $\Delta$  are polynomials of degree  $\leq 1$ ,  $\leq 2$ , and  $\leq 2$  respectively. Note that the word "Babel" built by the first letters of the names of the authors of [Bar-Lev et al. \(1991\)](#) and [Letac \(1991\)](#). In the examples given above, we remark that five examples belong to the Babel class. These examples are the normal inverse Gaussian and the normal gamma distributions (N-IG and N-G), the gamma Poisson distribution (G-P), and the Poisson inverse Gaussian and the Poisson gamma distributions (P-IG and P-G). It is worth mentioning that the Babel class is closed by reciprocity, i.e. the reciprocal NEF of  $\mu$  belongs to the Babel class. In fact, let  $\mu$  and  $\nu$  be a reciprocal pair, then  $V_{F(\nu)} = P_1\Delta_1 + Q_1\sqrt{\Delta_1}$ , where  $P_1$ ,  $Q_1$ , and  $\Delta_1$  are polynomials of degree  $\leq 1$ ,  $\leq 2$ , and  $\leq 2$  respectively (see [Kokonendji \(1995\)](#)).

We remark that we define a new class of NEFs that have the variance function depends on the Lambert function. We called it the Lambert class. The normal Poisson distribution (N-P), the gamma inverse Gaussian and the gamma gamma distributions (G-IG and G-G), the inverse Gaussian Poisson distribution (IG-P), and the Poisson Poisson distribution (P-P) belong to this class.

It is important to notice that from the examples, we get a new form of variance function which is the one of the inverse Gaussian gamma distribution (IG-G). This example is considered as a special case that we have not an idea about its class.

2. In some cases of Markov kernels, the explicit solution  $g(m)$  of the first order differential equation, given in Theorem 2.1.1, cannot be found. As examples of these Markov kernels, we may cite the Cauchy (1-stable), the stable (with  $\alpha \in (0, 2) \setminus \{1\}$ ), and the compound Poisson Markov kernels. Therefore, we cannot determine the explicit form of their link and variance functions. However, we may approximate them using the Euler method. In fact, using the first steps  $m^{(0)} = K'_{Q_1}(0)K'_{\mu_1}(0)$  and  $g^{(0)}(m^{(0)}) = K'_{\mu_1}(0)$ , we obtain the

following Euler approximation of the function  $g$

$$g^{(l+1)}(m^{(l+1)}) = g^{(l)}(m^{(l)}) + \delta \frac{m^{(l)} g^{(l)}(m^{(l)}) \left( V_{F(Q_1)} \left( \frac{m^{(l)}}{g^{(l)}(m^{(l)})} \right) \right)^{-1}}{\left( g^{(l)}(m^{(l)}) \right)^3 \left( V_{F(\mu_1)}(g^{(l)}(m^{(l)})) \right)^{-1} + \left( m^{(l)} \right)^2 \left( V_{F(Q_1)} \left( \frac{m^{(l)}}{g^{(l)}(m^{(l)})} \right) \right)^{-1}},$$

where  $l \in \mathbb{N}$  and  $\delta$  is a chosen size value of every step. This leads us to get the approximations of the link and variance functions of the class of Markov kernels. Note that [Mselmi \(2021b\)](#) has approximated the link and variance functions of the class Markov kernels of the distribution of the time-changed Lévy processes by the first-exit time of the inverse Gaussian subordinator by using this method. He investigated these approximations in order to get the one of the quasi-deviance functions.

## 2.2 The Lambert class

From the examples given in Section 2.1, we have noted that all the examples of the gamma Markov kernels and the examples of Markov kernels with the Poisson mixing distribution, except the gamma Poisson one, have variance functions that depend on the Lambert function. Then, in this section, we focus on the study of the variance functions of the classes of the gamma Markov kernels and the Markov kernels with the Poisson mixing distribution. It is important to notice that we will use the Tweedie distribution which covers the normal, gamma, Poisson, compound Poisson, and stable distributions (for more details about Tweedie model the reader can see [Boubacar Mainassara and Kokonendji \(2014\)](#)). Hence, we have the following corollary:

**Corollary 2.2.1** 1. Let  $\mu_1$  be the Poisson distribution and suppose that  $Q_1$  is the Tweedie one (its variance function  $V_{F(Q_1)}(m) = m^p$ , with  $p \in \mathbf{R} \setminus \{2 \cup (0, 1)\}$ ). Then, the link and the variance functions are equal to

$$\psi_\mu(m) = \begin{cases} \frac{1}{1-p} \left( \frac{m}{c} \right)^{1-p} \exp \left( \frac{1-p}{p-2} \mathcal{W} \left( \left( \frac{m}{c} \right)^{2-p} \right) \right), & \text{if } p \in \mathbf{R} \setminus \{1, 2\} \cup (0, 1) \\ \log \left( \mathcal{W} \left( \frac{m}{c} \right) \right), & \text{if } p = 1 \end{cases}$$

and

$$V_{F(\mu)}(m) = m^p c^{1-p} \exp \left( \frac{p-1}{p-2} \mathcal{W} \left( \left( \frac{m}{c} \right)^{2-p} \right) \right) + \frac{m^2}{c} \exp \left( \frac{1}{p-2} \mathcal{W} \left( \left( \frac{m}{c} \right)^{2-p} \right) \right),$$

2.2 The Lambert class

$$\text{where } c = K'_{\mu_1}(0) \exp\left(\frac{\left(K'_{Q_1}(0)\right)^{2-p}}{p-2}\right).$$

2. Let  $Q_1$  be the gamma distribution and suppose that  $\mu_1$  is Tweedie distributed with positive support. Then, the link and the variance functions are given by

$$\psi_{\mu}(m) = c \exp\left(\frac{1}{p-1} \mathcal{W}\left((cm)^{1-p}\right)\right), \quad p \neq 1$$

and

$$V_{F(\mu)}(m) = m^2 \left[ \frac{1}{cm} \exp\left(\frac{1}{1-p} \mathcal{W}\left((cm)^{1-p}\right)\right) + (cm)^{p-2} \exp\left(\frac{p-2}{p-1} \mathcal{W}\left((cm)^{1-p}\right)\right) \right],$$

$$\text{where } c = (K'_{Q_1}(0))^{-1} \exp\left(\frac{\left(K'_{\mu_1}(0)\right)^{1-p}}{1-p}\right).$$

**Remarks 2.2.2** 1. Corollary 2.2.1 (1.) allows us to determine the link and the variance functions of  $\alpha$ -stable Poisson distribution (S-P), with  $\alpha \in (0, 2] \setminus \{1\}$  (for more details about stable distributions and processes see [Louati et al. \(2015b, 2017, 2020\)](#) and [Mselmi \(2018b,c\)](#)). We have the following results

$$\psi_{\mu}(m) = (\alpha - 1) \left(\frac{m}{c}\right)^{\frac{1}{\alpha-1}} \exp\left(-\frac{1}{\alpha} \mathcal{W}\left(\left(\frac{m}{c}\right)^{\frac{\alpha}{\alpha-1}}\right)\right)$$

and

$$V_{F(\mu)}(m) = m^{\frac{\alpha-2}{\alpha-1}} c^{\frac{1}{\alpha-1}} \exp\left(\frac{1}{\alpha} \mathcal{W}\left(\left(\frac{m}{c}\right)^{\frac{\alpha}{\alpha-1}}\right)\right) + \frac{m^2}{c} \exp\left(\frac{1-\alpha}{\alpha} \mathcal{W}\left(\left(\frac{m}{c}\right)^{\frac{\alpha}{\alpha-1}}\right)\right),$$

$$\text{where } c = K'_{\mu_1}(0) \exp\left(\frac{1-\alpha}{\alpha} (K'_{Q_1}(0))^{\frac{\alpha}{\alpha-1}}\right).$$

2. Using Corollary 2.2.1 (2.), we may express the link and variance functions of the gamma  $\alpha$ -stable distribution (G-S), with  $\alpha \in (0, 1)$ . More precisely, we have

$$\psi_{\mu}(m) = c \exp\left((1-\alpha) \mathcal{W}\left((cm)^{\frac{1}{\alpha-1}}\right)\right)$$

and

$$V_{F(\mu)}(m) = m^2 \left[ \frac{1}{cm} \exp\left((\alpha-1) \mathcal{W}\left((cm)^{\frac{1}{\alpha-1}}\right)\right) + (cm)^{\frac{\alpha}{1-\alpha}} \exp\left(\alpha \mathcal{W}\left((cm)^{\frac{1}{\alpha-1}}\right)\right) \right],$$

$$\text{where } c = (K'_{Q_1}(0))^{-1} \exp\left((\alpha-1)(K'_{\mu_1}(0))^{\frac{1}{\alpha-1}}\right).$$

We present the following figure which summarizes all the examples given in this chapter with respect to their classes:

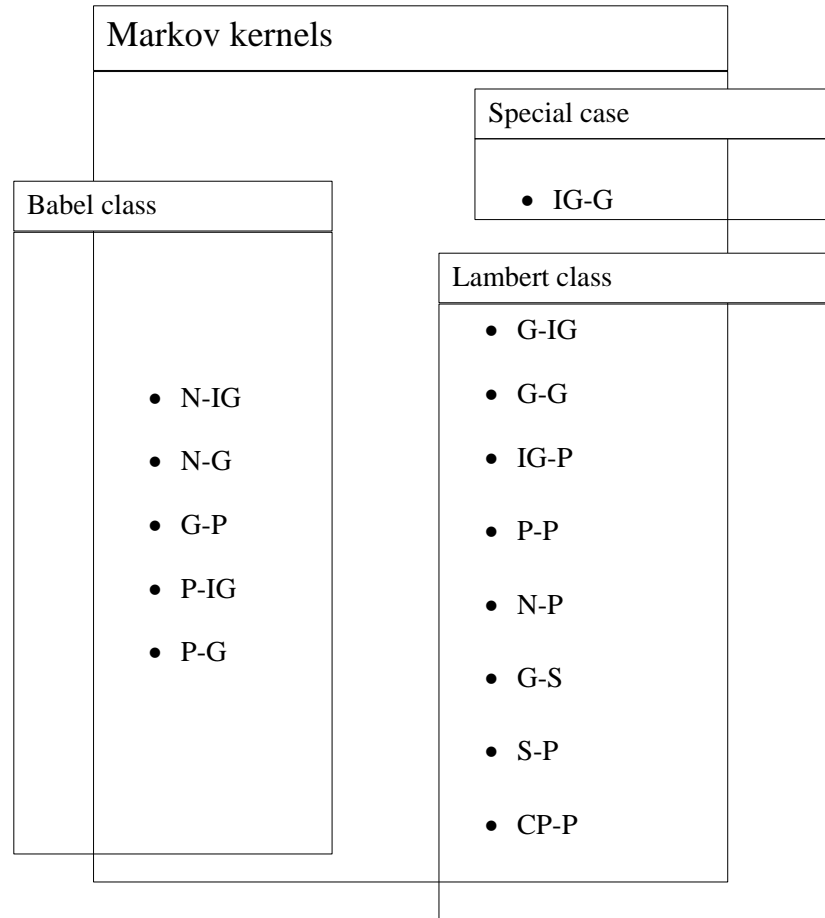


Figure 2.1: The NEFs classes of some Markov kernels.

## 2.3 The reciprocal class

In this part, we focus on the concept of reciprocity in the class of Markov kernels. More precisely, we have determined the form of the link and the variance functions of the reciprocal NEF of the class of the Markov kernels. We denote by  $\mathcal{R}\mu$ ,  $\mathcal{R}Q_1$ , and  $\mathcal{R}\mu_1$  the reciprocal measures of  $\mu$ ,  $Q_1$ , and  $\mu_1$  respectively.

**Theorem 2.3.1** 1. *The variance function of the reciprocal NEF of the class of Markov kernels is given, for all  $\tilde{m} \in \widetilde{M}_{F(\mathcal{R}\mu)}$ , by*

$$V_{F(\mathcal{R}\mu)}(\tilde{m}) = \left( \frac{\tilde{m}}{g_1(\tilde{m})} \right)^2 V_{F(\mathcal{R}Q_1)}(g_1(\tilde{m})) + g_1(\tilde{m}) V_{F(\mathcal{R}\mu_1)} \left( \frac{\tilde{m}}{g_1(\tilde{m})} \right), \text{ where } g_1(\tilde{m}) = \tilde{m} g \left( \frac{1}{\tilde{m}} \right).$$

2. *The function  $g_1$  satisfies the following differential equation*

$$g_1'(\tilde{m}) = \frac{\tilde{m} g_1(\tilde{m}) \left( V_{F(\mathcal{R}\mu_1)} \left( \frac{\tilde{m}}{g_1(\tilde{m})} \right) \right)^{-1}}{(g_1(\tilde{m}))^3 \left( V_{F(\mathcal{R}Q_1)}(g_1(\tilde{m})) \right)^{-1} + \tilde{m}^2 \left( V_{F(\mathcal{R}\mu_1)} \left( \frac{\tilde{m}}{g_1(\tilde{m})} \right) \right)^{-1}},$$

*with initial condition  $g_1(K'_{\mathcal{R}Q_1}(0)K'_{\mathcal{R}\mu_1}(0)) = K'_{\mathcal{R}Q_1}(0)$ .*

3. *The link function of the reciprocal NEF of the class of Markov kernels is given, for all  $\tilde{m} \in \widetilde{M}_{F(\mathcal{R}\mu)}$ , by*

$$\psi_{\mathcal{R}\mu}(\tilde{m}) = \psi_{\mathcal{R}\mu_1} \left( \frac{\tilde{m}}{g_1(\tilde{m})} \right).$$

4. *If  $\mathcal{R}\mu_1$  defines a convolution semigroup and  $\mathcal{R}Q_1$  has a positive support, then the distribution that generates the reciprocal NEF of the class of Markov kernels is given by*

$$\mathcal{R}\mu(dy) = \int_0^{+\infty} \mathcal{R}\mu_{1y}(dy) \mathcal{R}Q_1(dy).$$

**Remarks 2.3.2** 1. *From the distribution of the Markov kernel, we deduce that the reciprocal of the kernel  $Q_1$  is considered as a mixing distribution and the reciprocal of the mixing distribution  $\mu_1$  is taken as a kernel. Using this and [Letac and Mora \(1990\)](#) page 23, we obtain the reciprocal distributions of some examples of the class of Markov kernels given in [Section 2.1](#). These examples are summarized in the following table:*

*From [Table 2.1](#), we deduce that we have three self-reciprocal examples which are the N-IG, G-P, and the P-G distributions.*

|                |      |      |      |      |     |     |      |     |     |
|----------------|------|------|------|------|-----|-----|------|-----|-----|
| Markov kernel  | N-IG | N-G  | N-P  | G-IG | G-G | G-P | P-IG | P-G | P-P |
| Its reciprocal | N-IG | P-IG | G-IG | N-P  | P-P | G-P | N-G  | P-G | G-G |

Table 2.1: Markov kernels and their reciprocals

2. In 1990, [Letac and Mora \(1990\)](#), page 25, have noted that the reciprocal of an  $\alpha$ -stable distribution  $S_\alpha$ , with  $1 < \alpha < 2$ , is also an  $\alpha_1$ -stable distribution  $S_{\alpha_1}$ , with  $\alpha_1 = \frac{1}{\alpha} \in \left(\frac{1}{2}, 1\right)$ . Besides, they indicated, in page 24, that for  $0 < \alpha < 1/2$ , there is no reciprocal for the  $\alpha$ -stable distribution. Using these results and [Theorem 2.3.1](#), we deduce that the reciprocal of the gamma  $\alpha_1$ -stable distribution  $G-S_{\alpha_1}$ , with  $1/2 < \alpha_1 < 1$ , is the  $\alpha$ -stable Poisson distribution  $S_\alpha$ -P, with  $1 < \alpha < 2$ .
3. The distribution of the inverse Gaussian Markov kernel could not have a reciprocal one. In fact, the reciprocal of its kernel, which is the normal distribution, could not be considered as a mixing distribution since its support equal to  $\mathbf{R}$ . We announce the same remark for the compound Poisson Poisson distribution CP-P.

## 2.4 Some results on the duality

Recently, [Letac \(2021\)](#) has defined and studied the notion of duality for NEFs which could be investigated, in the case of existence, in the large deviation principle. In this context, let us use some notations, given by [Letac \(2021\)](#), in order to simplify the description of duality between two natural exponential families. Let us, first, define the bilateral Laplace transform  $\mathbb{B}_\mu(s) = L_\mu(-s)$ , for all  $s \in S(\mu) = -\Theta(\mu)$ . We modify, also, the description of some classical objects associated to the NEF  $F = F(\mu)$  as follows

$$\ell_\mu(s) = K_\mu(-s), \quad m = -\ell'_\mu(s) = K'_\mu(-s), \quad s = \pi_\mu(m) = -\psi_\mu(m), \quad \ell''_\mu(s) = K''_\mu(-s) \quad (2.1)$$

and  $V_{F(\mu)}(m) = -(\pi'_\mu(m))^{-1}$ . Now, we could define the duality of NEFs by the following way: if there exists a positive measure  $\mu^*$  such that

$$-\ell'_{\mu^*}(-\ell'_\mu(s)) = s, \quad (2.2)$$

then we say that  $\mu^*$  is the dual measure of  $\mu$ . In case of existence, the dual measure is investigated in the large deviation principle. In fact, let  $Y_1, \dots, Y_n$  be  $n$ -independent random

#### 2.4 Some results on the duality

variables with the same distribution  $\mathcal{P}$  in  $TF(\mu)$  with mean  $m_0$  ( $TF(\mu)$  is the union of all the translations of the NEF  $F(\mu)$  and is equal to  $\{P(-s, \mu) * \delta_{m_0}; s \in S(\mu), m_0 \in \mathbf{R}\}$ ). Then

$$-\lim_{n \rightarrow +\infty} \frac{1}{n} \log \left( \mathbf{P} \left( \frac{1}{n} \sum_{i=1}^n Y_i > m \right) \right) = h(m_0, m) = \int_{m_0}^m \frac{m-t}{V_{F(\mathcal{P}^*)}(t)} dt.$$

Note that, the function  $h(m_0, m)$  represents the rate function. For  $m_0 < m$  two points of the mean domain  $M_{F(\mu)}$ , if the dual measure  $\mathcal{P}^*$  of  $\mathcal{P}$  exists, then

$$\log(B_{\mathcal{P}_{m_0}^*}(m)) = h(m_0, m), \text{ where } \mathcal{P}_{m_0}^* = e^{\ell_{\mathcal{P}^*}(m_0) - m_0 \ell'_{\mathcal{P}^*}(m_0)} \mathcal{P}^* * \delta_{\ell_{\mathcal{P}^*}(m_0)}.$$

This gives the link between duality and large deviation principle (For more details, the reader can see [Letac \(2021\)](#), Subsection 3.5). We express the cumulant, the link and variance functions of the dual measure  $\mu^*$  in function of the ones of the kernel and mixing distributions in the following proposition:

**Proposition 2.4.1** *Suppose that the dual measure, denoted by  $\mu^*$ , of  $\mu$  exists. Then, for all  $s \in S(\mu)$  and  $m \in M_{F(\mu)}$ ,  $\ell_{\mu^*}(m) = \int \psi_{Q_1} \left( \frac{m}{g(m)} \right) dm$ ,  $\psi_{\mu^*}(s) = -K'_{Q_1}(-s)K'_{\mu_1}(K_{Q_1}(-s))$  and  $V_{F(\mu^*)}(s) = (K''_{Q_1}(-s)K'_{\mu_1}(K_{Q_1}(-s)) + (K'_{Q_1}(-s))^2 K''_{\mu_1}(K_{Q_1}(-s)))^{-1}$ .*

#### Some existing dual measures

We illustrate the results of Proposition 2.4.1 by some examples. We have used the expression of  $\psi_{\mu}(m) = \psi_{Q_1} \left( \frac{m}{g(m)} \right)$ , given in the examples of the work of [Mselmi \(2021a\)](#), in order to obtain the function  $\ell_{\mu^*}(m)$ . Before stating the expression of the link and variance functions of  $F(\mu^*)$ , we need first to prove the existence of dual measure through  $\ell_{\mu^*}$ . We give the following table in which summarizes some known results about the most useful NEFs in the literature. These results will be used in the examples below.



| NEF              | $\psi_\mu$                         | $V_{F(\mu)}$                    | $M_F$          | $K_\mu$   | $\Theta(\mu)$  |
|------------------|------------------------------------|---------------------------------|----------------|---|----------------|
| Gaussian         | $m$                                | 1                               | $\mathbf{R}$   | $\theta^2/2$  | $\mathbf{R}$   |
| Inverse Gaussian | $-1/2m^2$                          | $m^3$                           | $(0, +\infty)$ | $-2\sqrt{-\theta}$  | $(-\infty, 0)$ |
| Poisson          | $\log(m)$                          | $m$                             | $(0, +\infty)$ | $\exp(\theta)$  | $\mathbf{R}$   |
| Gamma            | $-1/m$                             | $m^2$                           | $(0, +\infty)$ | $-\log(-\theta)$  | $(-\infty, 0)$ |
| Compound Poisson | $\frac{m^{1-p}}{1-p}$              | $m^p$                           | $(0, +\infty)$ | $\frac{[(1-p)\theta]^{(p-2)/(p-1)}}{(2-p)}$                             | $(-\infty, 0)$ |
| Positive stable  | $(\alpha-1)m^{\frac{1}{\alpha-1}}$ | $m^{\frac{\alpha-2}{\alpha-1}}$ | $(0, +\infty)$ | $\frac{\alpha-1}{\alpha} \left[ \frac{\theta}{\alpha-1} \right]^\alpha$ | $(-\infty, 0)$ |

Table 2.2: Summary of some results about some known NEFs.

Note that for the compound Poisson distribution, one has  $p \in (1, 2)$  and for the positive stable distribution, the parameter  $\alpha$  in the set  $(0, 1)$ . Before expressing the examples, we need first to state the following remarks:

- Remarks 2.4.2** 1. In the examples of the gamma and inverse Gaussian Markov kernels, [Mselmi \(2021a\)](#) has considered the canonical link functions  $\frac{1}{m}$  for the gamma distribution and  $\frac{1}{m^2}$  for the inverse Gaussian one. However, in this section, we add some minor changes in the expression of these link functions. In fact, by using [Table 2.2](#), the link functions of the mixture of the gamma and inverse Gaussian kernels are, respectively, given by  $\psi_\mu(m) = -\frac{g(m)}{m}$  and  $\psi_\mu(m) = -\frac{1}{2} \left( \frac{g(m)}{m} \right)^2$ .
2. In some examples, we use a drifted cumulant function  $K_\mu(\theta + \theta^*) - K_\mu(\theta^*)$ , where  $\theta^* \in \Theta(\mu)$ , instead of the canonical one as indicated in [Table 2.2](#).
3. We recall the definition of Bernstein function which will be used for proving the existence of the dual measure  $\mu^*$ . A function  $f : (0, +\infty) \rightarrow \mathbf{R}$  is a Bernstein function if  $f$  is of class  $C^\infty$ ,  $f(\vartheta) \geq 0$ , for  $\vartheta > 0$ , and  $(-1)^{n-1} f^{(n)}(\vartheta) \geq 0$ , for all  $n \in \mathbf{N}^*$  and  $\vartheta > 0$ . Besides, A function  $f : (0, +\infty) \rightarrow \mathbf{R}$  is a Bernstein function if, and only if, it admits the representation  $f(\vartheta) = a + b\vartheta + \int_0^{+\infty} (1 - e^{-\vartheta u}) \zeta(du)$ , where  $a, b > 0$  are non-negative constants and  $\zeta$  is a measure on  $(0, +\infty)$  such that  $\int_0^{+\infty} (1 \wedge u) \zeta(du) < +\infty$ ,

#### 2.4 Some results on the duality

called the Lévy measure. From this, we deduce that if  $f : (0, +\infty) \longrightarrow [0, +\infty)$  is a Bernstein function then,  $e^{-f}$  represents the Laplace transform of a unique positive measure (for more details see [Schilling et al. \(2010\)](#)).

*Example 1: The negative Binomial distribution (the Poisson gamma distribution: P-G)*

The negative Binomial distribution represents a mixture between a Poisson kernel and gamma mixing distribution. In this case,  $K_{Q_1}(\theta) = \exp(\theta)$  and  $K_{\mu_1}(\theta_0) = -\log(1 - \theta_0)$  for all  $\theta \in \mathbf{R}$  and  $\theta_0 \in (-\infty, 1)$ . For all  $s \in (0, +\infty)$ , one has the following cumulant function  $K_{\mu_1}(K_{Q_1}(-s)) = -\log(1 - e^{-s})$ . Recall that, in this case,  $\psi_{Q_1}\left(\frac{m}{g(m)}\right) = \log\left(\frac{m}{m+1}\right)$  on  $(0, +\infty)$  (see [Mselmi \(2021a\)](#), supplementary material, example 3). Thus, we obtain the function  $\ell_{\mu^*}(m) = m \log(m) - (m+1) \log(m+1)$ . Note that  $-\ell_{\mu^*}(m)$  represents a Bernstein function (see [Schilling et al. \(2010\)](#), example 31, p. 230) and then  $\exp(\ell_{\mu^*}(m))$  represents the Laplace of a positive probability measure. This prove the existence of the dual measure  $\mu^*$ . We remind that this example was treated by [Letac \(2021\)](#) in Subsection 5.1. For all  $s \in (0, +\infty)$ , the link and variance functions of  $F(\mu^*)$  are given by  $\psi_{\mu^*}(s) = \frac{e^{-s}}{e^{-s} - 1}$  and  $V_{F(\mu^*)}(s) = e^s(e^{-s} - 1)^2$ .

*Example 2: The gamma Poisson distribution: G-P*

The gamma Poisson distribution represents a mixture between a gamma kernel and Poisson mixing distribution. Recall that, in this case,  $K_{Q_1}(\theta) = -\log(-\theta)$  and  $K_{\mu_1}(\theta_0) = \exp(\theta_0)$ , for all  $\theta \in (-\infty, 0)$  and  $\theta_0 \in \mathbf{R}$ . The cumulant function of the gamma Poisson distribution is equal, for all  $s \in (0, +\infty)$ , to  $K_{\mu_1}(K_{Q_1}(-s)) = \frac{1}{s}$ . For all  $m \in (0, +\infty)$ , one has  $\psi_{\mu}(m) = -1/\sqrt{m}$  (see [Mselmi \(2021a\)](#) supplementary material, example 2 and Remark 2.4.2(1.)). It follows that  $\ell_{\mu^*}(m) = -2\sqrt{m}$ . This represents the cumulant function of the canonical inverse Gaussian distribution, and then, we deduce that the dual measure of the gamma Poisson distribution represents the inverse Gaussian one. Its associated link and variance functions are given by  $\psi_{\mu^*}(s) = \frac{-1}{s^2}$  and  $V_{F(\mu^*)}(s) = \frac{s^3}{2}$ .

In the following Proposition, we express some non trivial examples of dual measures. Some of them depend on the Lambert function  $\mathcal{W}$ .

**Proposition 2.4.3** 1. The gamma gamma distribution (G-G) has a dual measure. Furthermore, we have, for all  $m \in (0, +\infty)$  and  $s \in (1, +\infty)$ ,

$$-\ell_{\mu^*}(m) = \log \left( \frac{1}{\mathcal{W}\left(\frac{1}{m}\right)} \right) + \frac{1}{\mathcal{W}\left(\frac{1}{m}\right)} + C, \text{ where } C \in \mathbf{R}, \psi_{\mu^*}(s) = \frac{-1}{s \log(s)}$$

$$\text{and } V_{F(\mu^*)}(s) = (s \log(s))^2 (1 + \log(s))^{-1}.$$

2. The inverse Gaussian gamma distribution (IG-G) has a dual measure. Besides, for all  $m \in (0, +\infty)$  and  $s \in (0, +\infty)$ , one has

$$-\ell_{\mu^*}(m) = \frac{1}{4}m + \frac{1}{2} \log(m) + \frac{1}{4} \sqrt{m+4} \left( \sqrt{m} + \frac{4 \sinh^{-1}(\sqrt{m}/2)}{\sqrt{m+4}} \right) + C, \text{ where } C \in \mathbf{R} \text{ and}$$

$$\sinh \text{ is the hyperbolic sine functions. } \psi_{\mu^*}(s) = \frac{-1}{\sqrt{s}(1+2\sqrt{s})} \text{ and } V_{F(\mu^*)}(s) = \frac{2s^{3/2}(2\sqrt{s}+1)^2}{4\sqrt{s}+1}.$$

3. The inverse Gaussian Poisson distribution (IG-P) has a dual measure. For all  $m \in (0, +\infty)$  and  $s \in (0, +\infty)$ , we obtain

$$-\ell_{\mu^*}(m) = \frac{1}{2} \exp \left( -\mathcal{W} \left( \frac{1}{m} \right) \right) \left( \mathcal{W} \left( \frac{1}{m} \right) + 2 \right) + C, \text{ where } C \in (0, +\infty),$$

$$\psi_{\mu^*}(s) = \frac{-1}{\sqrt{s}} \exp(-2\sqrt{s}) \text{ and } V_{F(\mu^*)}(s) = \frac{2s^{3/2} \exp(2\sqrt{s})}{1+2\sqrt{s}}.$$

4. The compound Poisson Poisson distribution (CP-P) has a dual measure. For all  $m \in (0, +\infty)$  and  $s \in (0, +\infty)$ , one gets  $-\ell_{\mu^*}(m) = \frac{(\mathcal{W}(m^{2-p}) + p - 1)}{p - 1} \exp \left( \frac{1}{2-p} \mathcal{W}(m^{2-p}) \right) - 1$ ,

$$\psi_{\mu^*}(s) = -\frac{\exp \left( \frac{1}{2-p} [(p-1)s]^{\frac{p-2}{p-1}} \right)}{[(p-1)s]^{\frac{1}{p-1}}} \text{ and } V_{F(\mu^*)}(s) = \frac{\exp \left( \frac{1}{p-2} [(p-1)s]^{\frac{p-2}{p-1}} \right)}{\left( [(p-1)s]^{\frac{p}{1-p}} + [(p-1)s]^{\frac{2}{1-p}} \right)}.$$

5. For  $\alpha \in (0, 1/2)$ , the stable Poisson distribution (S-P) has a dual measure. For all  $m \in (0, +\infty)$  and  $s \in (0, +\infty)$ , we obtain

$$-\ell_{\mu^*}(m) = (1-\alpha) \left( \mathcal{W}(m^{\frac{\alpha}{1-\alpha}}) + \frac{1}{1-\alpha} \right) \exp \left( \frac{\alpha-1}{\alpha} \mathcal{W}(m^{\frac{\alpha}{1-\alpha}}) \right) + C, \text{ where } C \in (0, +\infty),$$

$$\psi_{\mu^*}(s) = -\frac{\exp \left( \frac{\alpha-1}{\alpha} \left[ \frac{s}{1-\alpha} \right]^\alpha \right)}{\left[ \frac{s}{1-\alpha} \right]^{1-\alpha}} \text{ and } V_{F(\mu^*)}(s) = \frac{\exp \left( \frac{1-\alpha}{\alpha} \left[ \frac{s}{1-\alpha} \right]^\alpha \right)}{\left( \left[ \frac{s}{1-\alpha} \right]^{\alpha-2} + \left[ \frac{s}{1-\alpha} \right]^{2(\alpha-1)} \right)}.$$

#### 2.4 Some results on the duality

Next, we focus on the notion of infinite divisibility of some examples given in Proposition 2.4.3.

**Corollary 2.4.4** *The dual measures of the gamma gamma (G-G), the Inverse Gaussian Poisson (IG-P), the compound Poisson Poisson (CP-P) and the stable Poisson (S-P with  $\alpha \in (0, 1/2)$ ) distributions are infinitely divisible.*

#### Non existence of some dual measures

We check the non existence of the dual measures by two ways. For the first one, we refer to Letac (2021), Subsection 3.6, who has studied the non existence of duality of some NEFs by checking the Taylor series expansion of  $\exp(\ell_{\mu^*}(m))$ . In fact, if some coefficients the Taylor series expansion are negative (for even order), we deduce that the dual measure does not exists. For the second way, we investigate the Bernstein function theory. We consider the following examples of Markov kernels NEFs that have not dual:

*Example 1: The normal inverse Gaussian distribution: N-IG*

The kernel  $Q_{x_1}$  here represents the centered normal distribution  $N(0, x_1)$  with cumulant function  $K_{Q_1}(\theta) = \frac{\theta^2}{2}$  on  $\mathbf{R}$ . Its mixing distribution represents the inverse Gaussian one,  $IG(1, 1)$ , and it has the following cumulant function  $K_{\mu_1}(\theta_0) = 1 - \sqrt{1 - 2\theta_0}$  for  $\theta_0 \in \left(-\infty, \frac{1}{2}\right)$ . Then, we obtain the cumulant function of the normal inverse Gaussian distribution which is given, for all  $s \in (-1, 1)$ , by  $K_{\mu_1}(K_{Q_1}(-s)) = 1 - \sqrt{1 - s^2}$ . For all  $m \in \mathbf{R}$ ,  $\psi_{\mu}(m) = \frac{m}{\sqrt{1 + m^2}}$  and then, we deduce that  $\ell_{\mu^*}(m) = \sqrt{1 + m^2} - 1$ . In this case,  $\exp(\ell_{\mu^*}(m)) = 1 + \frac{m^2}{2} + o(m^5)$ . It follows that  $\int_{\mathbf{R}} x^4 \mu^*(dx) = 0$ . This prove the non existence of the dual measure  $\mu^*$ . Note that Letac (2021) has studied this example, in Subsection 3.6, without mentioning that it represents the normal inverse Gaussian model.

*Example 2: The variance gamma distribution (the normal gamma distribution: N-G)*

In this example the kernel distribution represents the normal one as in the case of the normal inverse Gaussian model. However, the mixing distribution represents the gamma one,  $G(1, 1)$ , and it has the following cumulant function  $K_{\mu_1}(\theta_0) = -\log(1 - \theta_0)$  for  $\theta_0 \in$

$(-\infty, 1)$ . Thus, we have the following cumulant function  $K_{\mu_1}(K_{Q_1}(-s)) = -\log\left(1 - \frac{s^2}{2}\right)$ , for all  $s \in (-\sqrt{2}, \sqrt{2})$ . For all  $m \in \mathbf{R}$ , one has the link function  $\psi_{\mu}(m) = \frac{2m}{1 + \sqrt{2m^2 + 1}}$  and then  $\ell_{\mu^*}(m) = \sqrt{2m^2 + 1} - \log(\sqrt{2m^2 + 1} + 1) + \log(2) - 1$ . Thus, we have the following Taylor series  $\exp(\ell_{\mu^*}(m)) = 1 + \frac{m^2}{2} + o(m^5)$ . As in (1.)  $\int_{\mathbf{R}} x^4 \mu^*(dx) = 0$  and then, the dual of variance gamma distribution does not exist.

*Example 3: Neyman type A distribution (the Poisson Poisson distribution: P-P)*

The mixture of two Poisson distributions called in the literature the Neyman type A model. It has, for all  $s \in \mathbf{R}$ ,  $K_{\mu_1}(K_{Q_1}(-s)) = \exp(\exp(-s))$ . According to [Mselmi \(2021a\)](#), supplementary material, example 4, we have, for all  $m \in (0, +\infty)$ ,

$$-\ell'_{\mu^*}(m) = -\psi_{Q_1}\left(\frac{m}{g(m)}\right) = -\log(\mathcal{W}(m)).$$

This function does not keep a nonnegative sign and then its primitive does not represent a Bernstein function (see Chapter 3 of the book of [Schilling et al. \(2010\)](#)). This implies that the Neyman type A distribution has not a dual measure.

## 2.5 Generalized linear model

Generalized Linear Model (GLM) was introduced by [Nelder and Wedderburn \(1972\)](#) and [Wedderburn \(1974\)](#) in order to study the statistical inference for the models with normal and non-normal data. In fact, it extends the concept of linear regression and takes into account that the linearity of the model cannot be obtained in a variety of practical situations.

Then, in this section, we focus on the study of the GLM of the distribution of the class of Markov kernels. In fact, we introduce the method of estimation and give the residual analysis of the GLM.

### 2.5.1 Quasi-likelihood function

The variance function is used in the determination of the quasi-likelihood estimator and is applied in the analysis of GLMs (see [McCullagh and Nelder \(1989\)](#), [Nelder and Wedderburn](#)

(1972) and Wedderburn (1974)). The quasi-likelihood function is given by

$$\mathcal{Q}(y, m) = \int_y^m \frac{y - u}{V_{F(\mu)}(u)} du. \quad (2.3)$$

Note that the properties of the quasi-likelihood function are similar to the log-likelihood one. That is why it can be useful when the likelihood function cannot be obtained.

In the case where we could not determine the explicit quasi-likelihood function, we may approximate it using Monte Carlo's method. In fact, the quasi-likelihood function is given by

$$\mathcal{Q}(y, m) = \int_y^m \frac{y - u}{V_{F(\mu)}(u)} du = (m - y) \int_{\mathbf{R}} \frac{y - u}{V_{F(\mu)}(u)} \frac{1}{m - y} \mathbf{1}_{[y, m]}(u) du = (m - y) \mathbf{E} \left( \frac{y - U}{V_{F(\mu)}(U)} \right),$$

where  $U$  is a continuous uniform random variable on  $[y, m]$ . This leads us to get the following approximation

$$\mathcal{Q}(y, m) \approx \frac{(m - y)}{N} \sum_{l=1}^N \frac{y - u_l}{V_{F(\mu)}(u_l)},$$

where  $N \in \mathbb{N} \setminus \{0\}$  is large enough and  $u_1, \dots, u_N$  are  $N$  simulated elements from the continuous uniform random variable  $U$ .

Using the quasi-likelihood function given in (2.3), we obtain the deviance function which measures the discrepancy between observations and its expected value

$$\mathcal{D}(y, m) = -2\mathcal{Q}(y, m). \quad (2.4)$$

### 2.5.2 Estimation

Let  $n \in \mathbb{N} \setminus \{0\}$  and  $\mathbf{Y}^\top = (Y_1, \dots, Y_n)$  be a random vector with i.i.d. components generated by the distribution  $\mu$ . Recall that the vector  $\mathbf{Y}^\top$  represents the transpose of  $\mathbf{Y}$ . Suppose that, for  $i \in \{1, \dots, n\}$ , the response random variable  $Y_i$  has a mean  $m_i$  and covariates  $x_{ij}$  for  $j \in \{1, \dots, k\}$  with  $k \in \mathbb{N} \setminus \{0\}$  and  $x_{i1} = 1$ . In the particular case, when the response data are normal, we obtain a linear regression model with mean  $m_i$  assumed to be a linear function of the covariates  $x_{ij}$ . Since the GLM provides the relationship between the linear predictor and the mean of the distribution function using the link function, then we have, for  $i \in \{1, \dots, n\}$ , the following relation

$$\psi_\mu(m_i) = \psi_\mu(\mathbf{E}(Y_i | \mathbf{X}_i)) = \mathbf{X}_i^\top \beta,$$

where  $\mathbf{X}_i^\top = (x_{i1}, \dots, x_{ik})$  and  $\beta = (\beta_1, \dots, \beta_k)^\top$  is the unknown vector of regression coefficients. Assume that, the random variable  $Y_i$  has the variance  $V(Y_i) = \sigma^2 V_{F(\mu)}(m_i)$ , where  $i \in \{1, \dots, n\}$  and  $\sigma^2$  is a known parameter or is an additional one to be estimated from the data. In what follows, we denote by  $\mathbf{m}$  the transpose of the vector  $(m_1, \dots, m_n)$  and we write as  $\mathbf{Y}$  or  $\mathbf{y}$  without subscripts. According to [McCullagh and Nelder \(1989\)](#) page 327, the estimation of the unknown vector of regression coefficients is obtained by the differentiating the quasi-likelihood function  $\mathcal{Q}(\mathbf{y}, \mathbf{m}) = \frac{1}{\sigma^2} \sum_{i=1}^n \mathcal{Q}(y_i, m_i)$ . Then, in order to calculate  $\hat{\beta}$ , we solve the quasi-score equation corresponding to

$$\mathcal{U}(\beta) = \frac{1}{\sigma^2} \sum_{i=1}^n D_{ir} V_{F(\mu)}^{-1}(m_i)(y_i - m_i) = 0,$$

where  $r \in \{1, \dots, k\}$  and  $D_{ir} = \frac{\partial m_i}{\partial \beta_r}$ . Using the fact that  $D_{ir} = \frac{\partial m_i}{\partial \psi_\mu(m_i)} \frac{\partial \psi_\mu(m_i)}{\partial \beta_r} = V_{F(\mu)}(m_i) x_{ir}$ , we obtain the following quasi-score equation

$$\mathcal{U}(\beta) = \frac{1}{\sigma^2} \sum_{i=1}^n x_{ir}(y_i - m_i) = \frac{1}{\sigma^2} \sum_{i=1}^n x_{ir}(y_i - K'_\mu(\mathbf{X}_i^\top \beta)) = 0. \quad (2.5)$$

Note that the quasi-likelihood estimating equations coincide with the ordinary likelihood ones. In general, the equation (2.5) is a nonlinear equation and is solved iteratively by Fisher's scoring method

$$\hat{\beta}^{new} = \hat{\beta}^{old} + \left( D^\top V_{F(\mu)}^{-1}(\mathbf{m}) D \right)^{-1} D^\top V_{F(\mu)}^{-1}(\mathbf{m})(\mathbf{y} - \mathbf{m}) = \hat{\beta}^{old} + \left( \mathbf{X} V_{F(\mu)}(\mathbf{m}) \mathbf{X}^\top \right)^{-1} \mathbf{X}(\mathbf{y} - \mathbf{m}),$$

where  $V_{F(\mu)}(\mathbf{m}) = \text{diag}(V_{F(\mu)}(m_1), \dots, V_{F(\mu)}(m_n))$ ,  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  and  $D = (D_{ir})_{1 \leq i \leq n, 1 \leq r \leq k} = V_{F(\mu)}(\mathbf{m}) \mathbf{X}^\top$ . The asymptotic variance of  $\hat{\beta}$  is estimated by

$$V(\hat{\beta}) \approx \sigma^2 \left( \mathbf{X} V_{F(\mu)}(\mathbf{m}) \mathbf{X}^\top \right)^{-1}. \quad (2.6)$$

It is important to notice that this matrix plays the same role as the inverse matrix of the Fisher information for the ordinary likelihood functions. In the case where the parameter  $\sigma^2$  is unknown, it could be estimated by the Pearson estimator

$$\hat{\sigma}^2 = \frac{1}{n-k} \sum_{i=1}^n \frac{(y_i - \hat{m}_i)^2}{V_{F(\mu)}(\hat{m}_i)} = \frac{1}{n-k} \sum_{i=1}^n \frac{\left( y_i - K'_\mu(\mathbf{X}_i^\top \hat{\beta}) \right)^2}{V_{F(\mu)} \left( K'_\mu(\mathbf{X}_i^\top \hat{\beta}) \right)}. \quad (2.7)$$

## 2.5 Generalized linear model

The confidence interval of the coefficients  $\beta_j$ ,  $j \in \{1, \dots, k\}$  was given by Jørgensen (2013) for the cases when  $\sigma^2$  is known and unknown. More precisely, for  $\sigma^2$  known the confidence interval for  $\beta_j$  is given by  $\hat{\beta}_j \pm se(\hat{\beta}_j)z_{1-\frac{\xi}{2}}$ , where  $se(\hat{\beta}_j)$  is calculated as the square-root of the  $j^{th}$  diagonal element of the matrix  $V(\hat{\beta})$  and  $z_{1-\frac{\xi}{2}}$  is the  $1 - \frac{\xi}{2}$  standard normal quantile. When  $\sigma^2$  is unknown, we replace  $\sigma^2$  by its estimated value in (2.6) and we take  $t_{1-\frac{\xi}{2}}(n-k)$  instead of  $z_{1-\frac{\xi}{2}}$  in (2.7), where  $t_{1-\frac{\xi}{2}}(n-k)$  is the  $1 - \frac{\xi}{2}$  quantile of the Student's t distribution with  $n - k$  degrees of freedom.

### 2.5.3 Residual analysis

The residual analysis is used to check the adequacy of fit of a model with respect to the choice of the link and variance functions. In this context, McCullagh and Nelder (1989) have introduced three forms of residuals that are the Pearson, the Anscombe, and the deviance. The classical proposal is the Pearson residual which is defined by  $r_i^P = \frac{y_i - \hat{m}_i}{\sqrt{V_{F(\mu)}(\hat{m}_i)}}$ . Unfortunately, this residual is inadequate for the study of non-normal data (see McCullagh and Nelder (1989), page 38). For this reason, the use of the Anscombe and the deviance residuals deal better with GLMs. The Anscombe residual is defined using a function  $A(\bullet) = \int_{-\infty}^{\bullet} (V_{F(\mu)}(u))^{-\frac{1}{3}} du$  and has the following form

$$r_i^A = \frac{A(y_i) - A(\hat{m}_i)}{A'(\hat{m}_i)\sqrt{V_{F(\mu)}(\hat{m}_i)}}.$$

For the examples given as above, we get an explicit form for the Anscombe residual of the of the N-IG and the G-P distributions. In fact, the Anscombe residual of the N-IG distribution is given by  $r_i^A = \frac{asinh(y_i) - asinh(\hat{m}_i)}{\sqrt{c}(1 + \hat{m}_i^2)^{\frac{1}{4}}}$ . Furthermore, the Anscombe residual of the G-P distribution is equal to  $r_i^A = \sqrt{2c}\hat{m}_i^{-\frac{1}{4}} (\sqrt{y_i} - \sqrt{\hat{m}_i})$ . Despite we cannot explicit the Anscombe residual of the majority of examples, we can calculate it by using the Monte-Carlo's approximation. In fact,

$$A(y_i) - A(\hat{m}_i) = \int_{\hat{m}_i}^{y_i} (V_{F(\mu)}(u))^{-\frac{1}{3}} du = (y_i - \hat{m}_i) \mathbf{E} \left( (V_{F(\mu)}(U))^{-\frac{1}{3}} \right) \approx \frac{y_i - \hat{m}_i}{N} \sum_{l=1}^N (V_{F(\mu)}(u_l))^{-\frac{1}{3}},$$



where  $N \in \mathbb{N} \setminus \{0\}$  is large enough and  $u_1, \dots, u_N$  are  $N$  simulated elements from a continuous uniform random variable  $U$  on  $[\hat{m}_i, y_i]$ . This implies that the Anscombe residual has the following approximation

$$r_i^A \approx \frac{(y_i - \hat{m}_i) \sum_{l=1}^N (V_{F(\mu)}(u_l))^{-\frac{1}{3}}}{N (V_{F(\mu)}(\hat{m}_i))^{\frac{1}{6}}}.$$

In 1986, [Pierce and Schafer \(1986\)](#) have studied and compared all types of residuals of GLMs. They have concluded that the most appropriate residual to GLMs is the deviance one. It is defined as follows

$$r_i^D = \text{sign}(y_i - \hat{m}_i) \sqrt{\mathcal{D}(y_i, \hat{m}_i)}, \quad (2.8)$$

where  $\mathcal{D}(y_i, \hat{m}_i)$  is the deviance function given in (2.4). Recall that [Pierce and Schafer \(1986\)](#) have noted that the deviance residual has the following bias  $-\frac{\mathbf{E}((Y_i - \hat{m}_i)^3)}{6\sigma^2 (V_{F(\mu)}(\hat{m}_i))^{\frac{3}{2}}}$  which must be subtracted from (2.8). In the numerical study of some examples, [McCullagh and Nelder \(1989\)](#) and [Pierce and Schafer \(1986\)](#) have indicated that the Anscombe and deviance residuals are very similar, in spite of the difference of their functionals.

## 2.6 Data analyses

In this part of our study, we focus on the applications of the GLMs for discrete and continuous data. We analyze this set of data by the GLMs of the N-IG, N-G, P-IG, and P-G distributions.

### 2.6.1 Data 1

This data represents the daily exchange rate returns of the Tunisian Dinar against the US Dollar. It spans over the period from 4 January 2016 to 30 December 2016. The series of exchange rate data are collected from the Datastream database with size  $n = 250$ . The daily exchange rate returns are calculated in terms of:

$$y_i = 100 \times (\log(P_i) - \log(P_{i-1})),$$

2.6 Data analyses

where  $y_i$  denotes the daily exchange rate return at time  $i$ ,  $P_i$  represents the nominal exchange rate at time  $i$  and  $P_{i-1}$  corresponds to the nominal exchange rate at time  $i - 1$ . We focus on the study of this data through the autoregressive model of order 1 (AR(1)).

In this example, we choose to study a nonlinear AR model with errors N-IG and N-G distributions. Then, we have, respectively, the representations of the (N-IG)-AR(1) and the (N-G)-AR(1) models

$$\psi_{\mu}^{N-IG}(m_i) = \frac{m_i}{\sqrt{1 + m_i^2}} = \beta_0 + \beta_1 y_{i-1} \text{ and } \psi_{\mu}^{N-G}(m_i) = \frac{2m_i}{1 + \sqrt{2m_i^2 + 1}} = \beta_0 + \beta_1 y_{i-1}.$$

Notice that, in this application, the constant  $c$  is fixed and is equal to 1 and the number of coefficients  $k = 2$ . Tables 2.3 and 2.4 give us the estimated values of  $\beta_0$  and  $\beta_1$  of the (N-IG)-AR(1) and the (N-G)-AR(1) models, their variances and their 95% confidence intervals. In order to study the adequacy of the models, we add the estimated variance  $\hat{\sigma}^2$  introduced in

(2.7), the deviance  $\mathcal{D}(\mathbf{y}, \mathbf{m}) = \sum_{i=1}^n \mathcal{D}(y_i, \hat{m}_i)$ , the deviance residual  $r^D = \sum_{i=1}^n r_i^D$ , the unbiased

deviance residual  $r^{D*} = \sum_{i=1}^n r_i^D + \frac{\mathbf{E} \left( (Y_i - \hat{m}_i)^3 \right)}{6\hat{\sigma}^2 (V_{F(\mu)}(\hat{m}_i))^{\frac{3}{2}}}$ , the Anscombe residual  $r^A = \sum_{i=1}^n r_i^A$ , the

Akaike information criterion  $AIC = \mathcal{D}(\mathbf{y}, \mathbf{m}) + 2k$  and the Bayesian information criterion  $BIC = \mathcal{D}(\mathbf{y}, \mathbf{m}) + \log(n)k$ .

| Variable                              | Estimate | Variance | 95% confidence interval |
|---------------------------------------|----------|----------|-------------------------|
| $\beta_0$                             | -0.0239  | 0.0002   | (-0.0486,0.0009)        |
| $\beta_1$                             | -0.1673  | 0.0039   | (-0.2891,-0.0455)       |
| $\hat{\sigma}^2$                      | 0.0396   | -        | -                       |
| $\mathcal{D}(\mathbf{y}, \mathbf{m})$ | 9.2982   | -        | -                       |
| $r^D$                                 | 0.2243   | -        | -                       |
| $r^{D*}$                              | -3.8005  | -        | -                       |
| $r^A$                                 | 0.2467   | -        | -                       |
| $AIC$                                 | 13.2982  | -        | -                       |
| $BIC$                                 | 20.3411  | -        | -                       |

Table 2.3: The estimated values of the (N-IG)-GLM.

| Variable                              | Estimate             | Variance             | 95% confidence interval |
|---------------------------------------|----------------------|----------------------|-------------------------|
| $\beta_0$                             | 0.4261               | $2.5 \times 10^{-5}$ | (0.4162,0.4359)         |
| $\beta_1$                             | 2.3788               | 0.0001               | (2.3577,2.4000)         |
| $\hat{\sigma}^2$                      | 0.2856               | -                    | -                       |
| $\mathcal{D}(\mathbf{y}, \mathbf{m})$ | 142.6275             | -                    | -                       |
| $r^D$                                 | -99.9491             | -                    | -                       |
| $r^{D*}$                              | $6.6296 \times 10^5$ | -                    | -                       |
| $r^A$                                 | -96.0790             | -                    | -                       |
| $AIC$                                 | 146.6275             | -                    | -                       |
| $BIC$                                 | 153.6705             | -                    | -                       |

Table 2.4: The estimated values of the (N-G)-GLM.

From Tables 2.3 and 2.4, we deduce that the Anscombe residuals  $r^A$  are closer to zero than the unbiased deviance residuals  $r^{D*}$ . We notice that all the residuals of the (N-IG)-GLM are better than the ones of the (N-G)-GLM. Besides, the AIC and BIC information criteria of the (N-IG)-GLM are less than the ones of the (N-G)-GLM. From this, we deduce that the modeling by the (N-IG)-GLM is adequate to the data more than the modeling by the (N-G)-GLM.

### 2.6.2 Data 2

This data is obtained from the book of McCullagh and Nelder (1989), page 204. It represents discrete data in which the outputs are the numbers of damage incidents of ships and the observed inputs are the type of ship, the year of construction, the period of operation, and the months of service. McCullagh and Nelder (1989) have analyzed this data by using the GLM of Poisson distribution. That is why we study the same data by the GLMs of the P-IG and P-G distributions. Then, we have, respectively, the following representations of the P-IG and the P-G GLMs

$$\psi_{\mu}^{P-IG}(m_i) = \log \left( \frac{2m_i}{m_i + \sqrt{m_i^2 + 2}} \right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}$$

2.6 Data analyses

and

$$\psi_{\mu}^{P-G}(m_i) = \log\left(\frac{m_i}{m_i + 1}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}.$$

In these GLMs, we fix the constant  $c = 1$ . As in the first application, we estimate  $\beta_0$  and  $\beta_1$ , their variances, their 95% confidence intervals, the parameter  $\hat{\sigma}^2$ , the deviance  $\mathcal{D}(\mathbf{y}, \mathbf{m})$ , the deviance residual  $r^D$ , the unbiased deviance residual  $r^{D*}$ , the Anscombe residual  $r^A$ , the Akaike information criterion  $AIC$  and the Bayesian information criterion  $BIC$ .

| Variable                              | Estimate              | Variance                | 95% confidence interval                     |
|---------------------------------------|-----------------------|-------------------------|---|
| $\beta_0$                             | -0.0296               | $0.8817 \times 10^{-4}$ | (-0.0480,-0.0112)                           |
| $\beta_1$                             | -0.0021               | $0.014 \times 10^{-4}$  | (-0.0044,0.0002)                            |
| $\beta_2$                             | 0.0032                | $0.0107 \times 10^{-4}$ | (0.0012,0.0052)                             |
| $\beta_3$                             | 0.0087                | $0.0653 \times 10^{-4}$ | (0.0037,0.0137)                             |
| $\beta_4$                             | $4.88 \times 10^{-7}$ | $1.85 \times 10^{-14}$  | $(2.2 \times 10^{-7}, 7.55 \times 10^{-7})$ |
| $\hat{\sigma}^2$                      | 0.1167                | -                       | -   |
| $\mathcal{D}(\mathbf{y}, \mathbf{m})$ | 16.3748               | -                       | -   |
| $r^D$                                 | -10.9034              | -                       | -   |
| $r^{D*}$                              | -0.1630               | -                       | -   |
| $r^A$                                 | -6.5225               | -                       | -   |
| $AIC$                                 | 26.3748               | -                       | -   |
| $BIC$                                 | 34.0066               | -                       | -   |

Table 2.5: The estimated values of the (P-IG)-GLM.

Tables 2.5 and 2.6 give us unbiased deviance residuals  $r^{D*}$  more close to zero than the Anscombe ones  $r^A$ . Furthermore, we note that all the residuals of the (P-IG)-GLM are better than the ones of the (P-G)-GLM. For the information criteria, we remark that the AIC and BIC criteria of the (P-IG)-GLM are less than the ones of the (P-G)-GLM. Then, we deduce that the modeling by the (P-IG)-GLM is adequate to the data more than the modeling by the (P-G)-GLM.

| Variable                              | Estimate              | Variance               | 95% confidence interval                      |
|---------------------------------------|-----------------------|------------------------|--|
| $\beta_0$                             | -0.4170               | 0.0229                 | (-0.7139,-0.1201)                            |
| $\beta_1$                             | -0.0245               | 0.0004                 | (-0.0651,0.0160)                             |
| $\beta_2$                             | 0.0427                | 0.0005                 | (-0.0003,0.0858)                             |
| $\beta_3$                             | 0.1076                | 0.0026                 | (0.0082,0.2069)                              |
| $\beta_4$                             | $6.77 \times 10^{-6}$ | $3.67 \times 10^{-12}$ | $(3.02 \times 10^{-6}, 1.05 \times 10^{-5})$ |
| $\hat{\sigma}^2$                      | 1.2022                | -                      | -  |
| $\mathcal{D}(\mathbf{y}, \mathbf{m})$ | 42.002                | -                      | -  |
| $r^D$                                 | -13.0242              | -                      | -  |
| $r^{D*}$                              | -8.9989               | -                      | -  |
| $r^A$                                 | -12.8316              | -                      | -  |
| $AIC$                                 | 52.0020               | -                      | -  |
| $BIC$                                 | 59.6338               | -                      | -  |

Table 2.6: The estimated values of the (P-G)-GLM.

## Chapter 3

# Multivariate Markov kernel for natural exponential family

In this chapter, we characterize the NEF generated by the distribution of multivariate Markov kernels through its associated variance function. Additional results as the link and generalized variance functions have been established. Some examples related to the multivariate normal and Poisson Markov kernels have been given. Since all the results related essentially to a nonlinear first order differential equation which has a complicated structure, then we consider a nonparametric method for estimating the associated link and variance functions. We check the performance of our estimated results through the mean squared error. We propose a MGLM for the class of Markov kernels by investigating the parametric and nonparametric results. Besides, we suggest a new multivariate total deviance function which was used for introducing the multivariate Akaike and Bayesian information criteria. We apply our parametric and nonparametric works to the multivariate daily exchange rates against US dollar. Proofs and auxiliary results are given in Appendix 2.

### 3.1 Variance and generalized variance functions

Here, we present some theoretical properties around variance functions of Markov kernels such as link, variance and also generalized variance functions.

**Theorem 3.1.1** *The link function of the NEF  $F(\mu)$  is given, for all  $m \in M_{F(\mu)} = M_{F(Q_1)}$ , by*

$$\psi_\mu(m) = \psi_{Q_1} \left( \frac{m}{g(m)} \right), \quad (3.1)$$

where  $g : \mathbf{R}^{d-1} \rightarrow (0, \infty)$  is a unique continuously differentiable function satisfying

$$g'(m) = \frac{g(m) [V_{F(Q_1)}(m/g(m))]^{-1} m^\top}{[g(m)]^3 [V_{F(\mu_1)}(g(m))]^{-1} + m [V_{F(Q_1)}(m/g(m))]^{-1} m^\top} \quad (3.2)$$

with initial condition  $g(K'_{Q_1}(0)K'_{\mu_1}(0)) = K'_{\mu_1}(0)$ .

Recently, [Mselmi et al. \(2018, Theorems 3.4 and 4.2\)](#) have determined the form of the variance and generalized variance functions of the NEF generated by the class of variance-mean mixture of two infinitely divisible distributions. Nevertheless, these functions have not an explicit form and they depend on the unknown link function. Note that the variance and generalized variance functions of the Markov kernel have the same form of the ones given by [Mselmi et al. \(2018\)](#). In fact, they have the following expressions

**Proposition 3.1.2** *One has:*

$$\begin{aligned} V_{F(\mu)}(m) &= \frac{\|K'_{Q_1}(\psi_\mu(m))\|^2}{\|m\|^2} V_{F(\mu_1)} \left( \frac{\|m\|}{\|K'_{Q_1}(\psi_\mu(m))\|} \right) m \otimes m \\ &\quad + \frac{\|m\|}{\|K'_{Q_1}(\psi_\mu(m))\|} V_{F(Q_1)} \left( \frac{m \|K'_{Q_1}(\psi_\mu(m))\|}{\|m\|} \right) \end{aligned}$$

and

$$\det V_{F(\mu)}(m) = \frac{[\varphi(m)]^{d-1} \det V_{F(Q_1)}(\phi(m))}{1 - [\varphi(m)]^{-2} V_{F(\mu_1)}(\varphi(m)) m [V_{F(\mu)}(m)]^{-1} m^\top},$$

with

$$\phi(m) = \frac{m \|K'_{Q_1}(\psi_\mu(m))\|}{\|m\|}, \quad \varphi(m) = \frac{\|m\|}{\|K'_{Q_1}(\psi_\mu(m))\|}$$

and

$$\begin{aligned} [V_{F(\mu)}(m)]^{-1} &= \frac{V_{F(\mu_1)}(\varphi(m)) [V_{F(Q_1)}(\phi(m))]^{-1} m \otimes m [V_{F(Q_1)}(\phi(m))]^{-1}}{\varphi(m) \left\{ (\varphi(m))^3 + V_{F(\mu_1)}(\varphi(m)) m [V_{F(Q_1)}(\phi(m))]^{-1} m^\top \right\}} \\ &\quad + [\varphi(m) V_{F(Q_1)}(\phi(m))]^{-1}. \end{aligned}$$

### 3.2 Some illustrative examples

Inserting the link function  $\psi_\mu$  given in Theorem 3.1.1 in the expression of the (generalized) variance functions of Proposition 3.1.2, we directly obtain the following corollary.

**Corollary 3.1.3** *From the NEF  $F(\mu)$  of Proposition 3.1.2 and Theorem 3.1.1, one has:*

$$V_{F(\mu)}(m) = g(m)V_{F(Q_1)}\left(\frac{m}{g(m)}\right) + \frac{1}{g(m)^2}V_{F(\mu_1)}(g(m))m \otimes m$$

and

$$\det V_{F(\mu)}(m) = \frac{[g(m)]^{d-1} \det V_{F(Q_1)}(m/g(m))}{1 - [g(m)]^{-2}V_{F(\mu_1)}(g(m)) m [V_{F(\mu)}(m)]^{-1} m^\top},$$

with

$$\begin{aligned} [V_{F(\mu)}(m)]^{-1} &= -\frac{V_{F(\mu_1)}(g(m)) \left[V_{F(Q_1)}\left(\frac{m}{g(m)}\right)\right]^{-1} m \otimes m \left[V_{F(Q_1)}\left(\frac{m}{g(m)}\right)\right]^{-1}}{g(m) \left\{g(m)^3 + V_{F(\mu_1)}(g(m)) m \left[V_{F(Q_1)}\left(\frac{m}{g(m)}\right)\right]^{-1} m^\top\right\}} \\ &\quad + \left[g(m)V_{F(Q_1)}\left(\frac{m}{g(m)}\right)\right]^{-1}. \end{aligned}$$

## 3.2 Some illustrative examples

We now present two kinds of examples related to the multivariate normal and also of Poisson Markov kernels that illustrate our previous results.

*Example 1: Multivariate normal Markov kernel*

The conditional distribution  $Q_{\mathbf{y}}$  here represents the  $(d-1)$ -variate centered normal distributed  $N(0, \mathbf{y}\mathbf{I}_{d-1})$ . Its unit variance function is given by  $V_{F(Q_1)}(m) = \mathbf{I}_{d-1}$  on  $\mathbf{R}^{d-1}$  with  $K_{Q_1}(\theta) = (\theta_1^2 + \dots + \theta_{d-1}^2)/2$  on  $\mathbf{R}^{d-1}$ .

### 1. Multivariate normal inverse Gaussian distribution (MNIG)

For instance, we refer to [Boubacar Mainassara and Kokonendji \(2014, Example 2.1\)](#) and reference therein with  $K_{\mu_1}(\theta_0) = -\sqrt{-2\theta_0}$  on  $(-\infty, 0]$ . For  $c = \frac{K'_{\mu_1}(0)}{\sqrt{1 + (K'_{\mu_1}(0))^2 \|K'_{Q_1}(0)\|^2}}$ , one has the following results  $g(m) = c\sqrt{1 + \|m\|^2}$ ,  $\psi_\mu(m) = m/(c\sqrt{1 + \|m\|^2})$ ,  $V_{F(\mu)}(m) =$



$$c\sqrt{1 + \|m\|^2} (\mathbf{I}_{d-1} + m \otimes m), [V_{F(\mu)}(m)]^{-1} = \frac{\mathbf{I}_{d-1}}{c\sqrt{1 + \|m\|^2}} - \frac{m \otimes m}{c(1 + \|m\|^2)^{3/2}} \text{ and}$$

$$\det V_{F(\mu)}(m) = \frac{\left(c\sqrt{1 + \|m\|^2}\right)^{d-1}}{1 - c\sqrt{1 + \|m\|^2} m [V_{F(\mu)}(m)]^{-1} m^\top}.$$

2. *Multivariate normal gamma distribution (MNG)*

Also, following [Boubacar Maïnassara and Kokonendji \(2014, Example 2.2\)](#) one has  $K_{\mu_1}(\theta_0) = -\log(-\theta_0)$  on  $(-\infty, 0)$ . Let  $c = \frac{\sqrt{2}K'_{\mu_1}(0)}{\sqrt{(K'_{\mu_1}(0))^2 \|K'_{Q_1}(0)\|^2 + 2K'_{\mu_1}(0)}}$ . Then,

we have the following results  $g(m) = (c^2 + c\sqrt{2\|m\|^2 + c^2})/2$ ,  $\psi_\mu(m) = 2m/(c^2 + c\sqrt{2\|m\|^2 + c^2})$ ,  $V_{F(\mu)}(m) = m \otimes m + 2^{-1} (c^2 + c\sqrt{2\|m\|^2 + c^2}) \mathbf{I}_{d-1}$ ,

$$[V_{F(\mu)}(m)]^{-1} = \frac{2 \mathbf{I}_{d-1}}{(c^2 + c\sqrt{2\|m\|^2 + c^2})} - \frac{4 m \otimes m}{(c^2 + c\sqrt{2\|m\|^2 + c^2}) \left[ (c^2 + c\sqrt{2\|m\|^2 + c^2}) + \|m\|^2 \right]}$$

$$\text{and } \det V_{F(\mu)}(m) = \frac{\left\{ (c^2 + c\sqrt{2\|m\|^2 + c^2}) / 2 \right\}^{d-1}}{1 - m [V_{F(\mu)}(m)]^{-1} m^\top}.$$

*Example 2: Multivariate Poisson Markov kernel*

Here, we consider  $Q_{\mathbf{y}}$  as the  $(d-1)$ -variate uncorrelated Poisson distribution with parameter  $\mathbf{y}\mathbf{I}_{d-1}$ . Its unit variance function is equal to  $V_{F(Q_1)}(m) = \mathbf{Diag}_{d-1}(m) = \mathbf{Diag}_{d-1}(m_1, \dots, m_{d-1})$ , the  $(d-1)$ -diagonal matrix of the vector  $m = (m_1, \dots, m_{d-1}) \in (0, \infty)^{d-1}$  with  $K_{Q_1}(\theta) = \exp \theta_1 + \dots + \exp \theta_{d-1}$  on  $\mathbf{R}^{d-1}$ .

1. *Multivariate Poisson inverse Gaussian distribution (MPIG)*

From [Mselmi et al. \(2018, Section 6.1\)](#), the MPIG is the particular case of the multivariate Poisson  $\alpha$ -stable with  $\alpha = 1/2$ ; see also [Jørgensen and Kokonendji \(2016, Section 3.5\)](#) for the univariate case. With  $K_{\mu_1}(\theta_0) = -\sqrt{-2\theta_0}$  on  $(-\infty, 0]$ , we consider the constant  $c = \sqrt{2}K'_{\mu_1}(0) / \sqrt{2(K'_{\mu_1}(0))^2 \langle K'_{Q_1}(0) \mathbf{I}_{d-1} \rangle + 1}$ . One gets  $g(m) = c \left( c \langle m, \mathbf{I}_{d-1} \rangle + \sqrt{c^2 \langle m, \mathbf{I}_{d-1} \rangle^2 + 2} \right) / 2$ ,

$$\psi_\mu(m) = \left( \log \frac{2m_1}{c \left( c \langle m, \mathbf{I}_{d-1} \rangle + \sqrt{c^2 \langle m, \mathbf{I}_{d-1} \rangle^2 + 2} \right)}, \dots, \log \frac{2m_{d-1}}{c \left( c \langle m, \mathbf{I}_{d-1} \rangle + \sqrt{c^2 \langle m, \mathbf{I}_{d-1} \rangle^2 + 2} \right)} \right),$$

### 3.2 Some illustrative examples

$$V_{F(\mu)}(m) = \mathbf{Diag}_{d-1}(m) + \frac{c}{2} \left( c \langle m, \mathbf{I}_{d-1} \rangle + \sqrt{c^2 \langle m, \mathbf{I}_{d-1} \rangle^2 + 2} \right) m \otimes m,$$

$$\begin{aligned} [V_{F(\mu)}(m)]^{-1} &= \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) \\ &\quad - \frac{c \left( c \langle m, \mathbf{I}_{d-1} \rangle + \sqrt{c^2 \langle m, \mathbf{I}_{d-1} \rangle^2 + 2} \right)}{2 + c \left( c \langle m, \mathbf{I}_{d-1} \rangle + \sqrt{c^2 \langle m, \mathbf{I}_{d-1} \rangle^2 + 2} \right) m \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) m^\top} \times \\ &\quad \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) m \otimes m \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) \end{aligned}$$

$$\text{and } \det V_{F(\mu)}(m) = \frac{2}{2 - c \left( c \langle m, \mathbf{I}_{d-1} \rangle + \sqrt{c^2 \langle m, \mathbf{I}_{d-1} \rangle^2 + 2} \right) m [V_{F(\mu)}(m)]^{-1} m^\top} \prod_{i=1}^{d-1} m_i.$$

#### 2. Multivariate negative binomial distribution (MNB)

We here consider a classical mixture of the multivariate Poisson  $Q_{\mathbf{y}}$  by a gamma generated by  $\mu_1$  with  $K_{\mu_1}(\theta_0) = -\log(-\theta_0)$  on  $(-\infty, 0)$ . For  $c = K'_{\mu_1}(0)/[1 + K'_{\mu_1}(0)\langle K'_{Q_1}(0), \mathbf{I}_{d-1} \rangle]$ , one has  $g(m) = c(1 + \langle m, \mathbf{I}_{d-1} \rangle)$ ,  $V_{F(\mu)}(m) = \mathbf{Diag}_{d-1}(m_1, \dots, m_{d-1}) + m \otimes m$ , with  $\psi_{\mu}(m) = \left( \log \frac{m_1}{c(\langle m, \mathbf{I}_{d-1} \rangle + 1)}, \dots, \log \frac{m_{d-1}}{c(\langle m, \mathbf{I}_{d-1} \rangle + 1)} \right)$ ,

$$\begin{aligned} [V_{F(\mu)}(m)]^{-1} &= - \frac{\mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) m \otimes m \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right)}{1 + m \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) m^\top} \\ &\quad + \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) \end{aligned}$$

$$\text{and } \det V_{F(\mu)}(m) = \frac{1}{1 - m [V_{F(\mu)}(m)]^{-1} m^\top} \prod_{i=1}^{d-1} m_i.$$

#### 3. Multivariate Neyman type A distribution (MNTA)

According to the univariate Neyman type A model, e.g. [Jørgensen and Kokonendji \(2016, Example 2.1\)](#), the MNTA is the mixture of the multivariate Poisson  $Q_{\mathbf{y}}$  by the univariate Poisson generated by  $\mu_1$  with  $K_{\mu_1}(\theta_0) = \exp \theta_0$  on  $\mathbf{R}$ . Let  $c = K'_{\mu_1}(0) \exp(-\langle K'_{Q_1}(0), \mathbf{I}_{d-1} \rangle)$ , then  $g(m) = c \exp \{ \mathcal{W}(\langle m, \mathbf{I}_{d-1} \rangle / c) \}$ ,

$$V_{F(\mu)}(m) = \mathbf{Diag}_{d-1}(m) + c^{-1} \exp \{ \mathcal{W}(\langle m, \mathbf{I}_{d-1} \rangle / c) \} m \otimes m,$$

$$\psi_\mu(m) = \left( \log \left\{ m_1 c^{-1} \exp \left[ -\mathcal{W} \left( \frac{\langle m, \mathbf{I}_{d-1} \rangle}{c} \right) \right] \right\}, \dots, \log \left\{ m_{d-1} c^{-1} \exp \left[ -\mathcal{W} \left( \frac{\langle m, \mathbf{I}_{d-1} \rangle}{c} \right) \right] \right\} \right),$$

$$\begin{aligned} [V_{F(\mu)}(m)]^{-1} &= -\frac{\mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) m \otimes m \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right)}{c \exp \left[ \mathcal{W} \left( \frac{\langle m, \mathbf{I}_{d-1} \rangle}{c} \right) \right] + m \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) m^\top} \\ &\quad + \mathbf{Diag}_{d-1} \left( \frac{1}{m_1}, \dots, \frac{1}{m_{d-1}} \right) \end{aligned}$$

and  $\det V_{F(\mu)}(m) = \frac{1}{1 - c^{-1} \exp[-\mathcal{W}(\langle m, \mathbf{I}_{d-1} \rangle / c)] m [V_{F(\mu)}(m)]^{-1} m^\top} \prod_{i=1}^{d-1} m_i$ , where  $\mathcal{W}$  is the Lambert function.

### 3.3 Associated kernel estimations for variance functions

In this section, we nonparametrically estimate the function  $g$  introduced in Theorem 3.1.1. This leads us to smooth the link and variance functions. In the next results, we express two nonparametric estimations through an associated kernel  $\mathcal{K}_{\mathbf{m}, \mathbf{H}}(\cdot)$  of both functions  $\psi_\mu(\mathbf{m})$  and  $V_{F(\mu)}(\mathbf{m})$ , for a given bandwidths matrix  $\mathbf{H} = (h_{ij})_{1 \leq i, j \leq d-1}$ ; see, e.g., Kokonendji and Somé (2018, 2021). According to the support of  $\mathbf{m}$ , we will consider two kinds of continuous associated kernels:

$$\mathcal{K}_{\mathbf{m}, \mathbf{H}}(\cdot) = |\mathbf{H}|^{-1/2} \mathcal{K} \left[ \mathbf{H}^{-1/2} (\mathbf{m} - \cdot) \right] \quad \text{and} \quad \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\cdot) = \prod_{j=1}^{d-1} \mathcal{K}_{m_j, h_{jj}}(\cdot),$$

where  $\mathcal{K}$  is a  $d$ -variate classical (symmetric) kernel (e.g., standard normal),  $|\mathbf{H}|$  is the determinant of the matrix  $\mathbf{H}$ ,  $\mathbf{m} = (m_1, \dots, m_{d-1})$  and  $\mathcal{K}_{m_j, h_{jj}}$ ,  $j = 1, \dots, d-1$ , belong either to the same family or to different families of univariate (classical or purely as gamma) associated kernels; see also Somé *et al.* (2022).

**Proposition 3.3.1** *Let  $f : (0, \infty) \rightarrow (0, \infty)$  be a bijective continuous function such that  $\lambda := \int_{\mathbb{R}^{d-1}} f \circ g(\mathbf{m}) d\mathbf{m} \in (0, \infty)$ . Then, the associated kernel estimators of the link and variance functions  $\psi_\mu(\mathbf{m})$  and  $V_{F(\mu)}(\mathbf{m})$  are given, respectively, by*

$$\widehat{\psi}_\mu(\mathbf{m}) = \psi_{Q_1} \left( \frac{\mathbf{m}}{f^{-1}(\lambda n^{-1} \sum_{i=1}^n \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\mathbf{m}_i))} \right)$$

### 3.3 Associated kernel estimations for variance functions

and

$$\begin{aligned} \widehat{V}_{F(\mu)}(\mathbf{m}) &= \left\{ f^{-1}(\lambda n^{-1} \sum_{i=1}^n \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\mathbf{m}_i)) \right\}^{-2} V_{F(\mu_1)} \left( f^{-1} \left( \frac{\lambda}{n} \sum_{i=1}^n \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\mathbf{m}_i) \right) \right) \mathbf{m} \otimes \mathbf{m} \\ &\quad + f^{-1} \left( \frac{\lambda}{n} \sum_{i=1}^n \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\mathbf{m}_i) \right) V_{F(Q_1)} \left( \frac{\mathbf{m}}{f^{-1}(\lambda n^{-1} \sum_{i=1}^n \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\mathbf{m}_i))} \right), \end{aligned}$$

where  $\mathbf{m} = (m_1, \dots, m_{d-1})$  and  $\mathbf{m}_i = (m_{i1}, \dots, m_{i(d-1)})$  for  $i = 1, \dots, n$  ( $n \in \mathbf{N}$ ) are  $(d-1)$ -vectors.

**Remarks 3.3.2** 1. Unfortunately, the estimation of the parameter  $\lambda$  was not possible in this work. For this reason and in order to compare our numerical results, we have computed  $\lambda \in (0, \infty)$  for some known examples and we have compared the results of our estimation for the real value of  $\lambda$  and  $\lambda = 1$ . We have noted from Table 3.1 and Figure 3.1 that the value of  $\lambda$  has not an important impact on the estimation of the link and variance functions.

2. The choice of the function  $f$  is crucial in this nonparametric part. In practice, we propose an appropriated function  $f(\cdot)$ , like  $\exp(\cdot)$  but other examples of functions, that satisfies the conditions of Proposition 3.3.1, can be used.

Figure 3.1 provides some plots of link and variance functions as well as their smoothed versions using the Gaussian kernel for some normal mixture distributions (NIG, NG and NP) with  $\lambda \neq 1$  and  $\lambda = 1$ . They generally have similar behaviours.

Multivariate Markov kernel for natural exponential family

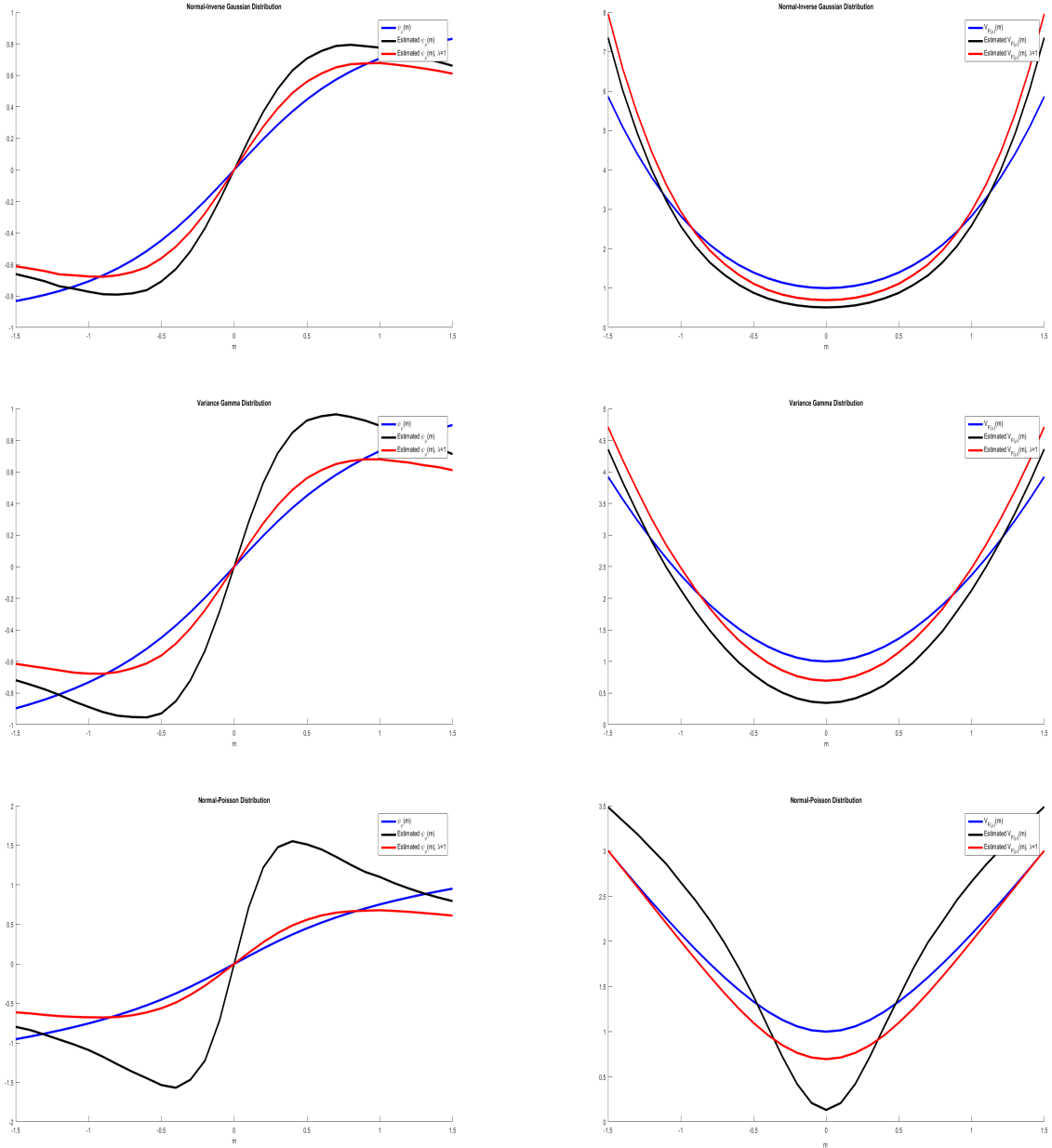


Figure 3.1: Plots of link and variance as well as their smoothed functions by the Gaussian kernel for some normal mixture distributions (NIG, NG and NP) with  $\lambda \neq 1$  and  $\lambda = 1$ . Note: NG means normal gamma or variance gamma.

### 3.4 Multivariate generalized linear model

Tables 3.1 and 3.2 present numerical results of mean squared errors for three normal mixture models (NIG, NG and NP) with different values of parameters in univariate and bivariate, respectively, setups.

| Model | Link function | Variance function | $\lambda$ | $h$ | $m$           |
|-------|---------------|-------------------|-----------|-----|---------------|
| NIG   | 0.0278/0.0120 | 0.3705/0.5759     | 1.2038/1  | 0.8 | $[-1.5, 1.5]$ |
| NG    | 0.0881/0.0176 | 0.2116/0.1197     | 1.4208/1  | 0.8 | $[-1.5, 1.5]$ |
| NP    | 0.4974/0.0237 | 0.2711/0.0348     | 1.7563/1  | 0.8 | $[-1.5, 1.5]$ |

Table 3.1: Mean squared errors for three normal mixture models (NIG, NG and NP) with different values of  $\lambda$  and fixed  $m \in [-1.5, 1.5]$  and  $h = 0.8$  in the univariate standard Gaussian kernel.

| Model | Link function | Variance function | $\lambda$ | $h$ | $m$             |
|-------|---------------|-------------------|-----------|-----|-----------------|
| NIG   | 0.0156        | 0.9366            | 1         | 0.3 | $[-1.5, 1.5]^2$ |
| NG    | 0.0217        | 0.3704            | 1         | 0.5 | $[-1.5, 1.5]^2$ |
| NP    | 0.0171        | 0.0686            | 1         | 0.3 | $[-1.5, 1.5]^2$ |

Table 3.2: Mean squared errors for three bivariate normal mixture models (NIG, NG and NP) with fixed  $\lambda = 1$  and  $\mathbf{H} = \mathbf{Diag}_2(h, h)$  in the bivariate standard Gaussian kernel.

## 3.4 Multivariate generalized linear model

In this section, we focus on the study of the multivariate generalized linear model (MGLM) which represents an extension of the GLM; e.g., McCullagh and Nelder (1989). In fact, we have the same concept in order to determine the coefficients and some related new properties.

Let  $k, n \in \mathbb{N} \setminus \{0\}$  and  $\mathbf{Y}_1 = (Y_{11}, \dots, Y_{1k})^\top, \dots, \mathbf{Y}_n = (Y_{n1}, \dots, Y_{nk})^\top$  be  $n$  random vector with i.i.d. components generated by the mixture distribution  $\mu$  and let  $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)^\top$  be  $nk \times 1$  response vector. Suppose that, for  $i \in \{1, \dots, n\}$ , the response random vector  $\mathbf{Y}_i$  has a mean vector  $\mathbf{m}_i = (m_{i1}, \dots, m_{ik})^\top$ . Consider  $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)^\top$  be  $nk \times pk$  covariates matrix, where  $\mathbf{X}_i = ([x_{i}]_{jl})_{1 \leq j \leq k, 1 \leq l \leq pk}$  is a  $k \times pk$  matrix with  $p \in \mathbb{N} \setminus \{0\}$ . The MGLM provides the following relationship between the linear predictor and the mean vector of the

distribution function by using its link function for all  $i \in \{1, \dots, n\}$

$$\psi_\mu(\mathbf{m}_i) = \psi_\mu(\mathbf{E}(\mathbf{Y}_i|\mathbf{X}_i)) = \mathbf{X}_i\boldsymbol{\beta},$$

where  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_k)^\top$  is the  $pk \times 1$ -unknown vector of regression coefficients such that  $\boldsymbol{\beta}_j = (\beta_{j1}, \dots, \beta_{jp})^\top$  for all  $j \in \{1, \dots, k\}$ . The estimation of the unknown matrix of regression coefficients is obtained by solving the following quasi-score equation

$$\mathcal{U}(\boldsymbol{\beta}) = \sum_{i=1}^n D_i^\top V_{F(\mu)}^{-1}(\mathbf{m}_i)(\mathbf{Y}_i - \mathbf{m}_i) = 0, \quad (3.3)$$

where  $D_i = \nabla_{\boldsymbol{\beta}}(\mathbf{m}_i)$  is a  $k \times pk$  matrix and  $\nabla_{\boldsymbol{\beta}}$  is the gradient operator with respect to  $\boldsymbol{\beta}$ . Note that

$$D_i = \nabla_{\boldsymbol{\beta}}(\mathbf{m}_i) = \nabla_{\psi_\mu(\mathbf{m}_i)}(\mathbf{m}_i)\nabla_{\boldsymbol{\beta}}(\psi_\mu(\mathbf{m}_i)) = V_{F(\mu)}(\mathbf{m}_i)\mathbf{X}_i.$$

The equation (3.3) is a nonlinear equation and is solved iteratively by the Fisher scoring method

$$\widehat{\boldsymbol{\beta}}^{new} = \widehat{\boldsymbol{\beta}}^{old} - S_{\boldsymbol{\beta}}^{-1}\mathcal{U}(\boldsymbol{\beta}),$$

with  $S_{\boldsymbol{\beta}} = -\sum_{i=1}^n D_i^\top V_{F(\mu)}^{-1}(\mathbf{m}_i)D_i$ . We approximate the covariance matrix of  $\widehat{\boldsymbol{\beta}}$  as follows:

$Cov(\widehat{\boldsymbol{\beta}}) \simeq \left( \sum_{i=1}^n \mathbf{X}_i^\top V_{F(\mu)}(\mathbf{m}_i)\mathbf{X}_i \right)^{-1}$ . The confidence interval of the coefficients  $\beta_{j,\ell}$ , for  $j \in \{1, \dots, k\}$  and  $\ell \in \{1, \dots, p\}$ , is given by  $\widehat{\beta}_{j,\ell} \pm z_{1-\xi/2}\sqrt{V(\widehat{\beta}_{j,\ell})}$ , where  $z_{1-\xi/2}$  is the  $1 - \xi/2$  standard normal quantile.

In order to check the adequacy of fit of a model with respect to the choice of the link and variance functions, we introduce the new quasi-deviance function, for all  $i \in \{1, \dots, n\}$ , as follows:

$$\begin{aligned} \mathcal{D}(\mathbf{Y}_i, \widehat{\mathbf{m}}_i) &= \int_{\widehat{\mathbf{m}}_i}^{\mathbf{Y}_i} V_{F(\mu)}^{-1}(\mathbf{u}_i)(\mathbf{Y}_i - \mathbf{u}_i)d\mathbf{u}_i = \int_{\widehat{m}_{i1}}^{Y_{i1}} \dots \int_{\widehat{m}_{ik}}^{Y_{ik}} V_{F(\mu)}^{-1}(\mathbf{u}_i)(\mathbf{Y}_i - \mathbf{u}_i)du_{i1} \dots du_{ik} \\ &= (\mathcal{D}_1(\mathbf{Y}_i, \widehat{\mathbf{m}}_i), \dots, \mathcal{D}_k(\mathbf{Y}_i, \widehat{\mathbf{m}}_i))^\top, \end{aligned}$$

where  $\mathbf{u}_i = (u_{i1}, \dots, u_{ik})^\top$ ,  $\widehat{\mathbf{m}}_i = K'_\mu(\mathbf{X}_i\widehat{\boldsymbol{\beta}})$  and  $V_{F(\mu)}^{-1}$  is given as in Corollary 3.1.3. We therefore define the total deviance and the residual deviance, respectively, in the following

### 3.5 Applications to multivariate daily exchange rates

sense:  $\mathcal{TD}(\mathbf{Y}_i, \hat{\mathbf{m}}_i) = \sum_{j=1}^k \mathcal{D}_j(\mathbf{Y}_i, \hat{\mathbf{m}}_i)$  and  $r_i^D = \langle \text{sign}(\mathbf{Y}_i - \hat{\mathbf{m}}_i), \mathcal{D}(\mathbf{Y}_i, \hat{\mathbf{m}}_i) \rangle$ , with

$\text{sign}(\mathbf{Y}_i - \hat{\mathbf{m}}_i) = (\text{sign}(Y_{i1} - \hat{m}_{i1}), \dots, \text{sign}(Y_{ik} - \hat{m}_{ik}))^\top$ . This implies that the total residual deviance  $r^D = \sum_{i=1}^n r_i^D$ , the Akaike information criterion  $AIC = \mathcal{TD}(\mathbf{Y}, \hat{\mathbf{m}}) + 2kp$  and the

Bayesian information criterion  $BIC = \mathcal{TD}(\mathbf{Y}, \hat{\mathbf{m}}) + kp \log(kn)$ , with  $\mathcal{TD}(\mathbf{Y}, \hat{\mathbf{m}}) = \sum_{i=1}^n \mathcal{TD}(\mathbf{Y}_i, \hat{\mathbf{m}}_i)$ .

### 3.5 Applications to multivariate daily exchange rates

We here analyze some sets of data in finance by the (M)GLMs for multivariate mixture models with (un)known link and variance functions.

In this general example, we consider the 7-variate daily exchange rate returns of the British pound (GBP), Canadian dollar (CAD), Chinese yuan (CNY), EURO (EUR), Moroccan dirham (MAD), Tunisian dinar (TND) and Turkish lira (TRL) against the US Dollar (USD).

In fact, let  $Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4}, Y_{i5}, Y_{i6}$  and  $Y_{i7}$  be, respectively, the exchange rate returns of GBP/USD, CAD/USD, CNY/USD, EUR/USD, MAD/USD, TND/USD and TRL/USD on the day  $i \in \{1, 2, \dots, n = 4366\}$ . Note that the parameter  $k$  is here equal to 7. Thus, the vector  $\mathbf{Y}_i$  which is defined by  $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{i7})^\top$  has a mean vector  $\mathbf{m}_i = (m_{i1}, \dots, m_{i7})^\top$ .

The covariates matrix  $\mathbf{X}_i$  can be written as:

$$\mathbf{X}_i = \begin{pmatrix} 1 & Y_{(i-1)1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & Y_{(i-1)2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & Y_{(i-1)3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & Y_{(i-1)4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & Y_{(i-1)5} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & Y_{(i-1)6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & Y_{(i-1)7} \end{pmatrix}.$$

Hence, we define the following multivariate nonlinear autoregressive model MAR(1) with (un)known errors MNIG distributions:  $\psi_\mu(\mathbf{m}_i) = \psi_\mu(\mathbf{E}(\mathbf{Y}_i | \mathbf{X}_i)) = \mathbf{X}_i \boldsymbol{\beta}$ , where  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_7)^\top$  is the  $14 \times 1$ -unknown vector of regression coefficients such that  $\boldsymbol{\beta}_j = (\beta_{j1}, \beta_{j2})^\top$  for all  $j \in \{1, \dots, 7\}$ . Here we also have  $p = 2$ . The results of estimations under known (parametric) and unknown (nonparametric) link and variance functions are given in Tables 3.3 and 3.4.



| Parametric   |                        |                       |                         |
|--------------|------------------------|-----------------------|-------------------------|
| Variable     | Estimate               | Variance              | 95% confidence interval |
| $\beta_{11}$ | $-3.02 \times 10^{-3}$ | $2.29 \times 10^{-4}$ | (-0.0293, 0.0300)       |
| $\beta_{12}$ | $2.10 \times 10^{-2}$  | $3.64 \times 10^{-3}$ | (-0.1001, 0.1364)       |
| $\beta_{21}$ | $2.57 \times 10^{-4}$  | $2.29 \times 10^{-4}$ | (-0.0269, 0.0324)       |
| $\beta_{22}$ | $-6.23 \times 10^{-3}$ | $3.66 \times 10^{-3}$ | (-0.1206, 0.1166)       |
| $\beta_{31}$ | $1.18 \times 10^{-3}$  | $2.30 \times 10^{-4}$ | (-0.0262, 0.0332)       |
| $\beta_{32}$ | $-1.65 \times 10^{-2}$ | $1.20 \times 10^{-1}$ | (-0.6900, 0.6708)       |
| $\beta_{41}$ | $-2.07 \times 10^{-3}$ | $2.29 \times 10^{-4}$ | (-0.0286, 0.0307)       |
| $\beta_{42}$ | $-1.75 \times 10^{-5}$ | $2.97 \times 10^{-3}$ | (-0.1042, 0.1094)       |
| $\beta_{51}$ | $-2.11 \times 10^{-3}$ | $2.29 \times 10^{-4}$ | (-0.0286, 0.0307)       |
| $\beta_{52}$ | $-4.88 \times 10^{-2}$ | $4.56 \times 10^{-3}$ | (-0.1659, 0.0990)       |
| $\beta_{61}$ | $-9.74 \times 10^{-3}$ | $2.29 \times 10^{-4}$ | (-0.0343, 0.0250)       |
| $\beta_{62}$ | $-8.54 \times 10^{-2}$ | $5.00 \times 10^{-3}$ | (-0.1993, 0.0781)       |
| $\beta_{71}$ | $-2.72 \times 10^{-2}$ | $2.29 \times 10^{-4}$ | (-0.0472, 0.0122)       |
| $\beta_{72}$ | $-1.13 \times 10^{-2}$ | $9.90 \times 10^{-4}$ | (-0.0675, 0.0560)       |

Table 3.3: On regression coefficients of MNIG-MGLM by a multivariate Gaussian kernel for the full dataset.

| Nonparametric |                        |                       |                         |
|---------------|------------------------|-----------------------|-------------------------|
| Variable      | Estimate               | Variance              | 95% confidence interval |
| $\beta_{11}$  | $2.38 \times 10^{-3}$  | $1.24 \times 10^{-3}$ | (-0.0115, 0.0162)       |
| $\beta_{12}$  | $1.64 \times 10^{-2}$  | $1.97 \times 10^{-3}$ | (-0.0136, 0.0463)       |
| $\beta_{21}$  | $4.30 \times 10^{-3}$  | $1.24 \times 10^{-3}$ | (-0.0089, 0.0175)       |
| $\beta_{22}$  | $5.10 \times 10^{-4}$  | $1.98 \times 10^{-2}$ | (-0.0182, 0.0192)       |
| $\beta_{31}$  | $4.84 \times 10^{-3}$  | $1.24 \times 10^{-3}$ | (-0.0097, 0.0193)       |
| $\beta_{32}$  | $-5.51 \times 10^{-3}$ | $6.51 \times 10^{-1}$ | (-0.3245, 0.3135)       |
| $\beta_{41}$  | $2.94 \times 10^{-3}$  | $1.24 \times 10^{-3}$ | (-0.0112, 0.0171)       |
| $\beta_{42}$  | $4.14 \times 10^{-3}$  | $1.60 \times 10^{-3}$ | (-0.0374, 0.0456)       |
| $\beta_{51}$  | $2.92 \times 10^{-3}$  | $1.24 \times 10^{-3}$ | (-0.0113, 0.0171)       |
| $\beta_{52}$  | $-2.43 \times 10^{-2}$ | $2.47 \times 10^{-2}$ | (-0.0537, 0.0051)       |
| $\beta_{61}$  | $-1.54 \times 10^{-3}$ | $1.24 \times 10^{-3}$ | (-0.0158, 0.0128)       |
| $\beta_{62}$  | $-4.57 \times 10^{-2}$ | $2.71 \times 10^{-2}$ | (-0.0979, 0.0068)       |
| $\beta_{71}$  | $-1.17 \times 10^{-2}$ | $1.24 \times 10^{-3}$ | (-0.0215, -0.0019)      |
| $\beta_{72}$  | $-2.44 \times 10^{-3}$ | $5.36 \times 10^{-3}$ | (-0.0116, 0.0067)       |

Table 3.4: On the nonparametric regression coefficients of MNIG-MGLM by a multivariate Gaussian kernel for the full dataset.

## Chapter 4

# Characterizations of stable and related processes

This chapter deals with some characterizations of the classes of multivariate stable processes and the inverse stable subordinators using some independence properties with a positive random variable. Moreover, we establish some characterizations of some multivariate Lévy processes, related to the class of Markov kernels, based on the notion of cut in a natural exponential family. This allows us to draw some related characteristic properties. Proofs and auxiliary results are given in [Louati \*et al.\* \(2017\)](#) and [Mselmi \(2018b\)](#).

### 4.1 Characterization of multivariate stable processes

This section is devoted to a characterization of the multivariate stable processes. We study the cases  $\alpha \neq 1$  and  $\alpha = 1$  separately. In the following theorem, we prove that a necessary and sufficient condition for a Lévy process to be  $\alpha$ -stable is that it satisfies some independence properties with a positive random variable  $T$ .

**Theorem 4.1.1** 1. *Let  $\alpha \in (0, 1) \cup (1, 2]$ ,  $\mathfrak{L}$  a multivariate Lévy process independent of  $T > 0$ , where the closed support of  $T$  is  $[0, +\infty)$ . Then  $\mathfrak{L}$  is strictly stable of parameter  $\alpha$  if and only if  $T^{-1/\alpha}\mathfrak{L}(T)$  and  $T$  are independent.*

2. *Let  $\mathfrak{L}$  be a multivariate Lévy process independent of  $T > 0$ , where the closed support of  $T$  is  $[0, +\infty)$ . Then  $\mathfrak{L}$  is strictly 1-stable if and only if there exists  $k \in \mathbf{R}^d \setminus \{0\}$  such that  $T^{-1}\mathfrak{L}(T) + \log(T)k$  and  $T$  are independent.*

Now, we consider the joint distribution  $(\eta_t(dy, dy) = \mu_{1,t}(dy)Q_{\mathbf{y},t}(dy))_{t>0}$  on  $(0, +\infty) \times \mathbf{R}^{d-1}$  with the two following properties:

- The cut property: if  $(\mathcal{Y}, Y)$  has a distribution belonging to the natural exponential family generated by  $\eta_t(dy, dy)$ , then the distribution of  $\mathcal{Y}$  belongs to the natural exponential family generated by  $\mu_{1,t}$ .
- For all  $\mathbf{y} > 0$ ,  $(Q_{\mathbf{y},t}(dy))_{t>0}$  is a convolution semigroup. We consider a Lévy process  $Z_{\mathbf{y}}(t)$  governed by this convolution semigroup. We denote by  $F_t = F(\eta_t)$  the NEF generated by  $\eta_t$ .

**Theorem 4.1.2** *Let  $(\mathcal{Y}, Z_{\mathbf{y}})$  and  $T$  be independent with  $Pr(T > t) = e^{-t}$ , and let  $p > 1$  be fixed. Then the two following statements are equivalent:*

1.  $T^{-p}\mathcal{Y}(T)$  and  $\mathcal{Y}(T)$  are independent and  $Z_{\mathbf{y}}(T)/\sqrt{T}$  and  $T$  are independent.
2. there exist  $\kappa > 0$  and two symmetric positive matrices  $B_0$  and  $B_1$  such that, for all  $(\mathbf{a}, \mathbf{b}) \in \{(\mathbf{a}, \mathbf{b}) \in \mathbf{R} \times \mathbf{R}^{d-1}; \mathbf{a} + \frac{1}{2}\langle \mathbf{b}, B_1 \mathbf{b} \rangle > -1/\kappa\}$

$$\int_{(0,+\infty) \times \mathbf{R}^{d-1}} e^{-\mathbf{a}\mathbf{y} + i\langle \mathbf{b}, \mathbf{y} \rangle} \eta_t(dy, dy) = \exp t \left( 1 - \left( 1 + \kappa \left( \mathbf{a} + \frac{1}{2}\langle \mathbf{b}, B_1 \mathbf{b} \rangle \right) \right)^{\frac{1}{p}} - \frac{1}{2}\langle \mathbf{b}, B_0 \mathbf{b} \rangle \right).$$

**Remark 4.1.3** *Note that, under the same conditions of Theorem 4.1.2 and by taking  $\mathbf{a} = 0$ , we get the Fourier transform of the normal tempered stable process  $(Y(t))_{t \geq 0}$ , that's*

$$\mathfrak{F}_{Y(t)}(\mathbf{b}) = e^{t \left( 1 - \left( 1 + \frac{\kappa}{2}\langle \mathbf{b}, B_1 \mathbf{b} \rangle \right)^{\frac{1}{p}} - \frac{1}{2}\langle \mathbf{b}, B_0 \mathbf{b} \rangle \right)}.$$

Next, we prove that the conditional distribution of  $Z_{\mathbf{y}}(T)$  is multivariate Laplace distributed. For this purpose, we recall some facts concerning this distribution. Let  $\Gamma$  be a  $(d-1)$ -symmetric definite positive matrix. Recall that the multivariate Laplace distribution on  $\mathbf{R}^{d-1}$  (for more details, we can see [Kotz et al. \(2001\)](#), p. 235), with a variance-covariance parameter  $\Gamma$ , is

$$f_{\Gamma}(z) = \frac{1}{(2\pi)^{\frac{d-1}{2}} \det(\Gamma)^{\frac{1}{2}}} \frac{2\mathfrak{K}_{\frac{d-3}{2}} \left( \sqrt{2\langle z, \Gamma^{-1}z \rangle} \right)}{\left( \sqrt{\frac{1}{2}\langle z, \Gamma^{-1}z \rangle} \right)^{\frac{d-3}{2}}}, \quad (4.1)$$

where  $\mathfrak{K}_{\chi}$  denotes the modified Bessel function of the third kind with order  $\chi$  (see [Kotz et al. \(2001\)](#), p. 315). For all  $\theta \in \mathbf{R}^{d-1}$ , this distribution is characterized by its Fourier transform

$$\mathfrak{F}_{\Gamma}(\theta) = \frac{1}{1 + \frac{1}{2}\langle \theta, \Gamma \theta \rangle}. \quad (4.2)$$

#### 4.2 Characterization of the inverse stable subordinator

(see [Kotz et al. \(2001\)](#), p. 235).

**Theorem 4.1.4** *Let  $Z_{\mathbf{y}}$  and  $T$  be independent with  $\Pr(T > t) = e^{-t}$ , and suppose that  $Z_{\mathbf{y}}(T)/\sqrt{T}$  and  $T$  are independent. Then  $Z_{\mathbf{y}}(T)$  is multivariate Laplace distributed a variance-covariance parameter  $B_0 + \mathbf{y}B_1$ , where  $B_0$  and  $B_1$  are  $(d - 1)$ -symmetric positive matrices.*

## 4.2 Characterization of the inverse stable subordinator

In this section, we characterize the inverse stable subordinator  $(E(t))_{t>0}$  by an independence condition with a positive random variable  $T$ . Furthermore, we extend this subordinator to a bivariate stochastic process  $((E_1(t), E_2(t)))_{t>0}$  and we characterize this process by its Laplace transform.

**Theorem 4.2.1** *Let  $\alpha \in (0, 1)$ ,  $(\mathfrak{L}(t))_{t>0}$  be a Lévy process independent of a random variable  $T$  with closed support  $[0, +\infty)$ , and  $E(t) = \inf\{z > 0 : \mathfrak{L}(z) > t\}$  be the first-exit time of  $\mathfrak{L}(t)$ . Then  $(E(t))_{t>0}$  is an inverse stable subordinator with a stability index  $\alpha$  if and only if,  $T^{-\alpha}E(T)$  and  $T$  are independent.*

Note that the Laplace transform of  $E(T)$  is given, for all  $\theta < 0$ , by

$$\mathbb{E}(e^{\theta E(T)}) = \int_0^{+\infty} \mathbb{E}(e^{\theta E(t)})f_T(dt) = c_\alpha(-\theta)^{\alpha-1} \int_0^{+\infty} e^{-tc_\alpha(-\theta)^\alpha} f_T(dt) = c_\alpha(-\theta)^{\alpha-1} L_T(-c_\alpha(-\theta)^\alpha),$$

where  $f_T$  is the density of  $T$ . In the case where  $T$  has the standard exponential distribution, we deduce that the Laplace transform of  $E(T)$  is equal to

$$\mathbb{E}(e^{\theta E(T)}) = \frac{1}{(-\theta)} \frac{c_\alpha(-\theta)^\alpha}{1 + c_\alpha(-\theta)^\alpha}.$$

This Laplace transform represents the double Laplace transform of the inverse stable subordinator given by [Alrawashdeh et al. \(2017\)](#) and [Meerschaert and Straka \(2013\)](#) and it has been used in order to solve some diffusion equations such as the tempered fractional telegraph equation.

Consider now a bivariate stochastic process  $((E_1(t), E_2(t)))_{t>0}$  defined on  $(0, +\infty) \times \mathbf{R}$  governed by a distribution belonging to the natural exponential family generated by  $\eta_t(dx_1, dx_2) = \mu_{1,t}(dx_1)Q_{x_1,t}(dx_2)$ . In what follows, we assume that the conditional distribution  $Q_{x_1,t}(dx_2)$  generate the continuous time Lévy process denoted by  $(\mathcal{A}_{x_1}(t))_{t>0} = (E_2(t)|E_1(t) = x_1)_{t>0}$ .

Besides, we suppose that, for all  $t > 0$ ,  $F_t$  has a cut on  $(0, +\infty)$  and that  $(E_1(t))_{t>0}$  is the first-exit time of a Lévy process.

**Theorem 4.2.2** *Let  $\alpha \in (0, 1)$ ,  $\tau \in (0, 1) \cup (1, 2]$  and  $(E_1(t))_{t>0}$ ,  $(\mathcal{A}_{x_1}(t))_{t>0}$  be independent of a random variable  $T$  with closed support  $[0, +\infty)$ . Then the following two statements are equivalent:*

1.  $T^{-\alpha}E_1(T)$  and  $T$  are independent and  $T^{-\frac{1}{\tau}}\mathcal{A}_{x_1}(T)$  and  $T$  are independent.
2. The Laplace transform of  $(E_1(t), E_2(t))$  is given, for all  $\theta_1, \theta_2 \leq 0$ , by

$$L_{\eta_t}(\theta_1, \theta_2) = c_\alpha(t\varepsilon_\tau(-\theta_2)^\tau - \theta_1)^{\alpha-1} \exp(-t[\delta_\tau(-\theta_2)^\tau + c_\alpha(t\varepsilon_\tau(-\theta_2)^\tau - \theta_1)^\alpha]),$$

where  $c_\alpha$ ,  $\varepsilon_\tau$  and  $\delta_\tau$  are three strictly positive constants.

Note that in the case where  $\tau = 2$ , the process  $((E_1(t), E_2(t)))_{t>0}$  is related to the notion of normal variance-mean mixture introduced by [Barndorff-Nielsen \(1997\)](#).

Now, we focus our attention to the case where  $\tau = 1$  which is slightly different of the case where  $\tau \in (0, 1) \cup (1, 2]$ . In this case, the process  $(\mathcal{A}_{x_1}(t))_{t>0}$  is a 1-stable process.

**Theorem 4.2.3** *Let  $\alpha \in (0, 1)$  and  $(E_1(t))_{t>0}$ ,  $(\mathcal{A}_{x_1}(t))_{t>0}$  be independent of a random variable  $T$  with closed support  $[0, +\infty)$ . Then the following two statements are equivalent:*

1. There exists  $k \in \mathbf{R} \setminus \{0\}$  such that  $T^{-1}\mathcal{A}_{x_1}(T) + \log(T)k$  and  $T$  are independent and  $T^{-\alpha}E_1(T)$  and  $T$  are independent.
2. The Laplace transform of  $(E_1(t), E_2(t))$  is given, for all  $\theta_1, \theta_2 \leq 0$ , by

$$L_{\eta_t}(\theta_1, \theta_2) = c_\alpha(t\lambda_0(-\theta_2) \log(-\theta_2) - \theta_1)^{\alpha-1} e^{(-t[\lambda_1(-\theta_2) \log(-\theta_2) + c_\alpha(t\lambda_0(-\theta_2) \log(-\theta_2) - \theta_1)^\alpha])},$$

where  $c_\alpha$ ,  $\lambda_0$  and  $\lambda_1$  are three strictly positive constants.

# Perspectives

We end this habilitation thesis by the following short-term perspectives:

## 1. Normal tempered stable Bayesian network:

Bayesian network is a probabilistic graphical model that analyzes a set of conditional independence relationships between a set of variables and consists of a directed acyclic graph. Let us consider  $d \in \mathbf{N} \setminus \{0\}$  and  $X = (X_1, X_2, \dots, X_d)$  a set of variables (commonly called nodes) connected with a set of directed edges and a set of conditional probability distributions. The distribution of  $X$  is defined through the chain rule  $\mathbf{P}(X_1, X_2, \dots, X_d) = \prod_{i=1}^d \mathbf{P}(X_i | Pa(X_i))$ , where  $Pa(X_i)$  denotes the parents of the node  $X_i$ . In this field, I will suppose that the distribution of the random variables  $X_i | Pa(X_i)$ , for  $i \in \{1, \dots, d\}$ , is normal tempered stable. Note that, in this work, I will use the techniques and results of GLM given in the work of Mselmi (2021a), and Mselmi and Kokonendji (2022).

## 2. Generalized probit model

A probit model represents a classification model that uses the cumulative distribution function of the normal distribution  $\Phi_{N(0,1)}$ , and is defined as follows:

$$\mathbf{P}(Y = 1|X) = \mathbf{P}(Y^* > 0) = \Phi_{N(0,1)}(X^T \beta),$$

where  $Y$  is a binary response variable,  $X$  is a vector of regressors,  $\beta$  is an unknown vector of regression coefficients to be estimated and  $Y^*$  is an auxiliary random variable

called the latent variable. The latent variable has the following relation  $Y^* = X^T\beta + \epsilon$ , where  $\epsilon \sim N(0, 1)$ , and it is positive if  $Y = 1$ . In this framework, I will change the error of the expression of the latent variable by the following way:  $Y^* = g(X^T\beta) + \epsilon$ , where  $\epsilon$  is normal tempered stable distributed. Note that, the function  $g$  represents the link function of the GLM of the normal tempered model (see Mselmi (2021a) and Mselmi and Kokonendji (2022)). Also, it is important to notice that the choice of the normal tempered stable distribution comes from the fact that it represents a generalization of the normal one. The random variable  $Y$  still a binary one and the change of the model will be on the cumulative distribution function, *i.e.* I will use the one of the normal tempered stable distribution  $\Phi_{NTS}(g(X^T\beta))$ . This cumulative distribution function will be estimated through the Monte-Carlo method.

### 3. R and Matlab packages

In this part, I will develop some packages for GLMs of the mixture distributions of interest by using R and Matlab. Some examples have been done in the work of Mselmi (2021a) and Mselmi and Kokonendji (2022). In fact, I will take the following steps:

- Develop the GLM packages for some univariate examples.
- Develop the GLM packages for some multivariate examples.
- Develop the nonparametric GLM packages.
- Develop a package for the normal tempered stable Bayesian network.
- Develop a package for the Generalized probit model.

# Bibliography

- Alrawashdeh, M.S., Kelly, J.F., Meerschaert, M.M. and Scheffler, H.P. (2017). Applications of inverse tempered stable subordinators. *Comp. Math. App.*, 73(6), 892–905.
- Asmussen, S., Jensen, J.L. and Rojas-Nandayapa, L. (2014). On the Laplace transform of the lognormal distribution. *Meth. Comp. App. Prob.*, 18(2), 441–458.
- Bar-Lev, S.K., Bshouty, D. and Enis, P. (1991). Variance functions with meromorphic means. *Ann. Prob.*, 19(3), 1349–1366.
- Barndorff-Nielsen, O.E. (1997). Normal inverse Gaussian distributions and stochastic volatility modeling. *Scand. J. Stat.*, 24, 1–13.
- Barndorff-Nielsen, O.E. and Koudou, A.E. (1995). Cuts in natural exponential families. *Theo. Prob. Appl.* 40, 361–372.
- Barndorff-Nielsen, O.E. and Levendorskii, S.Z. (2001). Feller processes of normal inverse Gaussian type. *Quan. Fin.*, 1(3), 318–331.
- Barndorff-Nielsen, O.E. and Shephard, N. (2001). Normal modified stable processes. *Theo. Prob. Math. Stat.*, 65, 1–19.
- Bedbur, S. and Kamps, U. (2021). *Multivariate Exponential Families: A Concise Guide to Statistical Inference*. Switzerland: Springer.
- Bhattacharya, S.K. and Kumar, S. (1986). E-IG model in life testing. *Cal. Stat. Asso. Bull.*, 35, 85–90.



- Bondesson, L. and Steutel, F. (2004). A class of infinitely divisible distributions connected to branching processes and random walks. *J. Math. Anal. Appl.*, 295(1), 134–143.
- Boubacar Maïnassara, Y. and Kokonendji, C.C. (2014). On normal stable Tweedie models and power-generalized variance functions of only one component. *Test*, 23, 585–606.
- Casalis, M. (1996). The  $(2d + 4)$  simple quadratic natural exponential families on  $\mathbf{R}^d$ . *Ann. Stat.*, 24, 1828–1854.
- Charfi, S. and Mselmi, F. (2022). Modeling exchange rate volatility: application of GARCH models with a Normal Tempered Stable distribution. *Quan. Fin. Eco.*, 6(2), 206–222.
- Çınlar, E. (2011). *Probability and Stochastics*. New York: Springer-Verlag.
- Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actu. Sci. Ind.*, 736, 5–23.
- Dean, C., Lawless, J.F. and Willmot, G.E. (1989). A mixed Poisson-inverse Gaussian regression model. *Can. J. Stat.*, 17, 171–181.
- Fiorani, F., Luciano, E. and Semeraro, P. (2010). Single and joint default in a structural model with purely discontinuous asset prices. *Quan. Fin.*, 10, 249–263.
- Frangos, N. and Karlis, D. (2004). Modelling losses using an exponential-inverse Gaussian distribution. *Insur. Math. Eco.*, 35, 53–67.
- Ganti, V., Singh, A., Passalacqua, P. and Foufoula-Georgiou, E. (2009). Subordinated Brownian motion model for sediment transport. *Phys. Rev. E*, 80, 011111.
- Gómez-Déniz, E., Calderín-Ojeda, E. and Sarabia, J.M. (2013). Gamma-generalized inverse Gaussian class of distributions with applications. *Comm. Stat. Theo. Meth.*, 42(6), 919–933
- Hirsa, A. and Madan, D.B. (2003). Pricing american options under variance gamma. *J. Comput. Fin.*, 7(2), 63–80.

BIBLIOGRAPHY

- Holla, M.S. (1967). On a Poisson-inverse Gaussian distribution. *Metrika*, 11, 115–121.
- Jørgensen, B. (2013). Generalized linear models. In A. H. El-Shaarawi, & W. W. Piegorsch (Eds.), *Encyclopedia of Environmetrics (2 ed., pp. 1152–1159)*. Chichester: Wiley.
- Jørgensen, B. and Kokonendji, C.C. (2016). Discrete dispersion models and their Tweedie asymptotics. *AStA Adv. Stat. Anal.*, 100, 43–78.
- Kalugin, G., Jeffrey, D., Corless, R. and Borwein, P.B. (2012). Stieltjes and other integral representations for functions of Lambert W. *Int. Trans. Spec. Funct.*, 23(8), 581–593.
- Kim, P., Song, R. and Vondraček, Z. (2009). Boundary Harnack principle for subordinate Brownian motions. *Stoch. Proce. Appl.*, 119(5), 1601–1631.
- Kim, P., Song, R. and Vondraček, Z. (2012). Potential theory of subordinate Brownian motions revisited. *Stochastic Analysis and Applications to Finance. Essays in honour of Jia-an Yan, T. Zhang, X. Zhou (eds), pp. 243-290*. World Sci. Publ., Hackensack, NJ.
- Kokonendji, C.C. (1995). Sur les familles exponentielles naturelles de grand-Babel. *Ann. Facul. Sci. Toul.*, IV,4, 763–799.
- Kokonendji, C.C. and Moypemna Sembona, C.C. (2018). Characterization and classification of multiple stable Tweedie models. *Lith. Math. J.*, 58, 441–456.
- Kokonendji, C.C., Moypemna Sembona, C.C. and Nisa, K. (2020). A complete characterization of multivariate normal stable-Tweedie models through a Monge-Ampère property. *Act. Math. Sin. Eng. Ser.*, 36, 1232–1244.
- Kokonendji, C.C. and Somé, S.M. (2018). On multivariate associated kernels to estimate general density functions. *J. Korean Statist. Soc.*, 47, 112–126.
- Kokonendji, C.C. and Somé, S.M. (2021). Bayesian bandwidths in semiparametric modelling for nonnegative orthant data with diagnostics. *Stats*, 4, 162–183.

- Kotz, S., Kozubowski, T.J. and Podgorski, K. (2001). *The Laplace Distribution and generalizations. A revisit with applications to communications, economics, engineering, and finance*, Birkhauser, Boston.
- Kumar, A., Nane, E. and Vellaisamy, P. (2011). Time-changed Poisson processes. *Stat. Prob. Lett.*, 81, 1899–1910.
- Letac, G. (1986). La réciprocité des familles exponentielles naturelles sur  $\mathbf{R}$ . *C. R. Acad. Sci., Paris Sér. I.*, 303, 61–64.
- Letac, G. (1991). The classification of natural exponential families by their variances functions. *Proceeding of the 48th session of the International Statistical Institute*, Vol. LIV, Book 3
- Letac, G. (1992). *Lecture on natural exponential families and their variance functions*. Monografias de Mathematica 50. IMPA, Rio de Janeiro.
- Letac, G. (2021). Duality for real and multivariate exponential families. *J. Mult. Anal.*, 188, 104811.
- Letac, G. and Mora, M. (1990). Natural real exponential families with cubic variance function. *Ann. Stat.*, 18, 1–37.
- Louati, M., Masmoudi, A. and Mselmi, F. (2015a). Multivariate normal  $\alpha$ -stable exponential families. *Medit. J. Math.*, 13, 1307–1323.
- Louati, M., Masmoudi, A. and Mselmi, F. (2015b). *Gamma stopping and drifted stable processes*. Applied Mathematics in Tunisia, Springer Proceedings in Mathematics and Statistics (Springer), 131, 223–232.
- Louati, M., Masmoudi, A. and Mselmi, F. (2017). Characterizations of multivariate stable processes. *Lith. Math. J.*, 57(1), 59–68.
- Louati, M., Masmoudi, A. and Mselmi, F. (2020). The normal tempered stable regression model. *Comm. Stat. Theo. Meth.*, 49(2), 500–512.

BIBLIOGRAPHY

- Madan, D.B., Carr, P. and Chang, E. (1998). The variance gamma process and option pricing. *Euro. Fin. Rev.*, 2, 79–105.
- Madan, D.B. and Seneta, E. (1990). The Variance Gamma (V.G.) model for share market returns. *J. Bus.*, 63(4), 511–524.
- Madan, D.B. and Yor, M. (2008). Representing the CGMY and Meixner processes as time changed Brownian motions. *J. Comput. Fin.*, 12(1), 27–47.
- McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*. New York: Chapman and Hall.
- Meerschaert, M.M., Nane, E. and Vellaisamy, P. (2011). The fractional Poisson process and the inverse stable subordinator. *Elect. J. Prob.*, 16, 1600–1620.
- Meerschaert, M.M. and Scheffler, H.P. (2004). Limit theorems for continuous time random walks with infinite mean waiting times. *J. Appl. Prob.*, 41(3), 623–638.
- Meerschaert, M.M. and Scheffler, H.P. (2008). Triangular array limits for continuous time random walks. *Stoch. Proc. Appl.*, 118, 1606–1633.
- Meerschaert, M.M. and Straka, P. (2013). Inverse Stable Subordinators. *Math. Model. Nat. Phen.*, 8(2), 1–16.
- Morris, C.N. (1982). Natural exponential families with quadratic variance functions. *Ann. Stat.*, 10, 65–80.
- Mselmi, F. (2018a). Lévy processes time-changed by the first-exit time of the inverse Gaussian subordinator. *Filomat*, 32(7), 2545–2552.
- Mselmi, F. (2018b). Characterization of the inverse stable subordinator. *Stat. Prob. Lett.*, 140, 37–43.
- Mselmi, F. (2018c). The stable processes on symmetric matrices. *Ara. J. Math. Sci.*, 24(1), 63–69.

- Mselmi, F. (2021a). Generalized linear model for subordinated Lévy processes. *Scand. J. Stat.*, 49(2), 772–801.
- Mselmi, F. (2021b). Approximation of the quasi-deviance function for the time-changed Lévy processes by the first-exit time of the inverse Gaussian subordinator. *Stat*, 10(1), 1–9.
- Mselmi, F., Kokonendji, C.C., Louati, M. and Masmoudi, A. (2018). Generalized variance functions for infinitely divisible mixture distributions. *Medit. J. Math.*, 14(4):165, 1–19.
- Nelder, J. and Wedderburn, R.W.M. (1972). Generalized linear models. *J. Roy. Stat. Soc. Series A*, 135(3), 370–384.
- Nolan, J.P. (2008). *Stable distributions: Models for heavy tailed data*. Birkhäuser, Boston.
- Ord, J.K. and Whitmore, G.A. (1986). The Poisson-inverse Gaussian distribution as a model for species abundance. *Comm. Stat. Theo. Meth.*, 15, 853–871.
- Pakes, A.G. (2011). Lambert’s W, infinite divisibility and Poisson mixtures. *J. Math. Anal. Appl.*, 378(2), 480–492.
- Pierce, D.A. and Schafer, D.W. (1986). Residuals in generalized linear models. *J. Amer. Stat. Asso.*, 81, 977–986.
- Rigby, R.A., Stasinopoulos, D.M. and Akantziliotou, C. (2008). A framework for modelling overdispersed count data, including the Poisson-shifted generalized inverse Gaussian distribution. *Comp. Stat. Data Anal.*, 53, 381–393.
- Rydberg, T. (1996). The normal inverse Gaussian Lévy Process: simulation and approximation. *Comm. Stat. Stoch. Mod.*, 13, 887–910.
- Schilling, R.L., Song, R. and Vondraček, Z. (2010). *Bernstein Functions: Theory and Applications*. De Gruyter, Berlin.

BIBLIOGRAPHY

- Somé, M.S., Kokonendji, C.C., Mamode Khan, N.A., Adjabi, S. and Beddek, S. (2022). Multiple combined gamma kernel estimations for nonnegative data with Bayesian adaptive bandwidths. *Comput. Stat.*, Submitted. <https://arxiv.org/abs/2202.09314>.
- Song, R. and Vondraček, Z. (2003). Potential theory of subordinate killed Brownian motion in a domain. *Prob. Theo. Rel. Fiel.*, 125, 578–592.
- Tremblay, L. (1992). Using the Poisson inverse Gaussian in bonus-malus systems. *Ast. Bull.*, 22, 97–106.
- Vinogradov, V. (2011). On Kendall-Ressel and related distributions. *Stat. Prob. Lett.*, 81, 1493–1501.
- Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, 61(3), 439–447.
- Willmot, G.E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scand. Actu. J.*, 113–127.
- Zhang, W. and Ge, S.S. (2006). A global implicit function theorem without initial point and its applications to control of non-affine systems of high dimensions. *J. Math. Anal. Appl.*, 313, 251–261.



# Appendix 1

## Proofs and auxiliary results for Section 2.4

We refer to the book of [Schilling et al. \(2010\)](#) for introducing the notion of Bernstein function theory. Let us, first, give some definitions related to this theory.

### Definitions 4.2.4 (Bernstein function theory)

1. A function  $f : (0, +\infty) \rightarrow \mathbf{R}$  is a completely monotone function if  $f$  is of class  $C^\infty$  and  $(-1)^n f^{(n)}(\vartheta) \geq 0$ , for all  $n \in \mathbf{N}$  and  $\vartheta > 0$ . The family of all completely monotone functions will be denoted by  $\mathcal{CM}$ .
2. A (non-negative) Stieltjes function is a function  $f : (0, +\infty) \rightarrow [0, +\infty)$  which can be written in the form  $f(\vartheta) = \frac{a}{\vartheta} + b + \int_0^{+\infty} \frac{1}{\vartheta + u} \varsigma(du)$ , where  $a, b > 0$  are non-negative constants and  $\varsigma$  is a measure on  $(0, +\infty)$  such that  $\int_0^{+\infty} \frac{1}{1+u} \varsigma(du) < +\infty$ . We denote the family of all Stieltjes functions by  $\mathcal{S}$ .
3. A function  $f : (0, +\infty) \rightarrow \mathbf{R}$  is a Bernstein function if  $f$  is of class  $C^\infty$ ,  $f(\vartheta) \geq 0$ , for  $\vartheta > 0$ , and  $(-1)^{n-1} f^{(n)}(\vartheta) \geq 0$ , for all  $n \in \mathbf{N}^*$  and  $\vartheta > 0$ . The family of all Bernstein functions will be denoted by  $\mathcal{BF}$ . Besides, A function  $f : (0, +\infty) \rightarrow \mathbf{R}$  is a Bernstein function if, and only if, it admits the representation  $f(\vartheta) = a + b\vartheta + \int_0^{+\infty} (1 - e^{-\vartheta u}) \zeta(du)$ , where  $a, b > 0$  are non-negative constants and  $\zeta$  is a measure on  $(0, +\infty)$  such that  $\int_0^{+\infty} (1 \wedge u) \zeta(du) < +\infty$ , called the Lévy measure.
4. A function  $f : (0, +\infty) \rightarrow \mathbf{R}$  is an extended Bernstein function if  $f$  is of class  $C^\infty$  and  $(-1)^{n-1} f^{(n)}(\vartheta) \geq 0$ , for all  $n \in \mathbf{N}^*$  and  $\vartheta > 0$ . The family of all extended Bernstein functions will be denoted by  $\mathcal{EBF}$ .
5. A Bernstein function  $f$  is said to be a complete Bernstein function if its Lévy measure  $\eta$  has a completely monotone density  $\varrho(u)$  with respect to Lebesgue measure. The family of all complete Bernstein functions will be denoted by  $\mathcal{CBF}$ .



6. A  $C^\infty$  function  $f : (0, +\infty) \rightarrow (0, +\infty)$  is said to be logarithmically completely monotone if  $-(\log(f))' \in \mathcal{CM}$ .

Now, we express, in the following lemma, some known properties that give the relationship between the classes  $\mathcal{S}$ ,  $\mathcal{CM}$ ,  $\mathcal{CBF}$ ,  $\mathcal{BF}$  and  $\mathcal{EBF}$ .

**Lemma 4.2.5** 1.  $\mathcal{S} \subset \mathcal{CM}$  and  $\mathcal{CBF} \subset \mathcal{BF} \subset \mathcal{EBF}$ .

2. If  $f : (0, +\infty) \rightarrow \mathbf{R}$  is a completely monotone function, then it represents the Laplace transform of a unique positive measure  $\mu$  (see [Schilling et al. \(2010\)](#), Theorem 1.4, page 3).
3.  $f : (0, +\infty) \rightarrow [0, +\infty)$  is a Bernstein function if, and only if,  $e^{-f}$  is a completely monotone function (see [Schilling et al. \(2010\)](#), Theorem 3.6, page 19).
4. A non-negative  $C^\infty$ -function  $f : (0, +\infty) \rightarrow \mathbf{R}$  is a Bernstein function if, and only if,  $f'$  is a completely monotone function.
5. If  $f$  an extended Bernstein function, then  $f'$  is a completely monotone function.
6. If  $f \in \mathcal{S} \setminus \{0\}$ , then  $\frac{1}{f} \in \mathcal{BF}$  (see [Kalugin et al. \(2012\)](#), Section 3).
7.  $f \in \mathcal{CBF} \iff \frac{1}{f} \in \mathcal{S} \setminus \{0\}$  (see [Kalugin et al. \(2012\)](#), Section 4).

### Proof of Proposition 2.4.1

According to (2.1), we have  $m = -\ell'_\mu(s) = K'_\mu(-s)$ . It follows that,  $s = -\psi_\mu(m)$ . This and (2.2) imply  $-\ell'_{\mu^*}(m) = -\ell'_{\mu^*}(-\ell'_\mu(s)) = s = -\psi_\mu(m) = -\psi_{Q_1}\left(\frac{m}{g(m)}\right)$ . Integrating this, we obtain  $\ell_{\mu^*}(m) = \int \psi_{Q_1}\left(\frac{m}{g(m)}\right) dm$ . Using the fact that  $K'_{\mu^*}(-m) = -\ell'_{\mu^*}(m) = s$ , we obtain  $\psi_{\mu^*}(s) = -m = -K'_\mu(-s) = -K'_{Q_1}(-s)K'_{\mu_1}(K_{Q_1}(-s))$ . Differentiating this link function, we obtain the expression of the variance function as stated in Proposition 2.4.1. ■

### Proof of Proposition 2.4.3

1. According to [Mselmi \(2021a\)](#), supplementary material, example 1 and Remark 2.4.2(1.), one has  $\psi_\mu(m) = -\frac{1}{m} \left( \mathcal{W}\left(\frac{1}{m}\right) \right)^{-1}$  (Note that according to [Kalugin et al. \(2012\)](#) Theorem 2.6(c),  $-\psi_\mu(m) \in \mathcal{S}$ ). It follows that

$$-\ell_{\mu^*}(m) = \log \left( \frac{1}{\mathcal{W}\left(\frac{1}{m}\right)} \right) + \frac{1}{\mathcal{W}\left(\frac{1}{m}\right)} + C,$$

where  $C \in \mathbf{R}$ . From Lemma 4.2.5 and Theorem 2.6(b) of Kalugin *et al.* (2012), we deduce that  $\mathcal{W}\left(\frac{1}{m}\right) \in \mathcal{S}$  and then  $\frac{1}{\mathcal{W}\left(\frac{1}{m}\right)} \in \mathcal{BF}$ . So, our problem, now, is to show that  $\log\left(\frac{1}{\mathcal{W}\left(\frac{1}{m}\right)}\right) \in \mathcal{EBF}$ . According to Kalugin *et al.* (2012), Section 4, one has  $\mathcal{W}(m) \in \mathcal{CBF}$ . Using equation (7.3) of Schilling *et al.* (2010), page 66, we obtain  $m\mathcal{W}\left(\frac{1}{m}\right) \in \mathcal{CBF}$ . Since the function  $m \mapsto m$  is a trivial  $\mathcal{CBF}$  (see Remark 7.8, Chapter 7 of Schilling *et al.* (2010)), then  $m + m\mathcal{W}\left(\frac{1}{m}\right) \in \mathcal{CBF}$ . Using Lemma 4.2.5, one gets  $\frac{1}{m + m\mathcal{W}\left(\frac{1}{m}\right)} \in \mathcal{S} \subset \mathcal{CM}$ . From this, we deduce that

$$-\left(\log\left(\mathcal{W}\left(\frac{1}{m}\right)\right)\right)' = -\frac{\left(\mathcal{W}\left(\frac{1}{m}\right)\right)'}{\mathcal{W}\left(\frac{1}{m}\right)} = \frac{\mathcal{W}'\left(\frac{1}{m}\right)}{m^2\mathcal{W}\left(\frac{1}{m}\right)} = \frac{1}{m + m\mathcal{W}\left(\frac{1}{m}\right)} \in \mathcal{CM}.$$

It follows that  $\mathcal{W}\left(\frac{1}{m}\right)$  is logarithmically completely monotone. According to Schilling *et al.* (2010), Theorem 5.9, page 38, we obtain  $\log\left(\frac{1}{\mathcal{W}\left(\frac{1}{m}\right)}\right) \in \mathcal{EBF}$ . Using this and the fact that  $\frac{1}{\mathcal{W}\left(\frac{1}{m}\right)} \in \mathcal{BF}$ , one has  $-\ell(m) \in \mathcal{EBF}$ . Thus, we get the existence of the dual measure of the gamma-gamma distribution. For all  $s \in (1, +\infty)$ , one has  $K_{\mu_1}(K_{Q_1}(-s)) = -\log(\log(s))$ . From the first and second derivatives of this function, we obtain  $\psi_{\mu^*}$  and  $V_{F(\mu^*)}$ .

2. According to Mselmi (2021a), supplementary material, example 5 and Remark 2.4.2(1.), one has  $-\psi_{\mu}(m) = \frac{1}{8m} [\sqrt{m} + \sqrt{m+4}]^2 = \frac{1}{4} + \frac{1}{2m} + \frac{1}{4}\sqrt{\frac{4+m}{m}}$ . From Remark 7.8 of Schilling *et al.* (2010) page 65, the function  $\sqrt{\frac{m}{4+m}} \in \mathcal{CBF}$  (the functions  $\frac{\vartheta}{\vartheta+t}$  and  $\vartheta^a$  are complete Bernstein functions for  $\vartheta, t > 0$  and  $0 \leq a \leq 1$ ). It follows that  $\sqrt{\frac{4+m}{m}} \in \mathcal{S} \subset \mathcal{CM}$ . Lemma 4.2.5 implies that

$$\int \sqrt{\frac{4+m}{m}} dm = \sqrt{m+4} \left( \sqrt{m} + \frac{4 \sinh^{-1}(\sqrt{m}/2)}{\sqrt{m+4}} \right) \in \mathcal{BF}, \quad (4.3)$$

where  $\sinh^{-1}$  is the inverse hyperbolic sine function. Besides, One has  $\frac{1}{m}$  is logarithmically completely monotone function  $\left(\frac{1}{m} \in \mathcal{S} \subset \mathcal{CM}\right)$ . According to Schilling *et al.* (2010), Theorem 5.9, page 38, we obtain

$$\log(m) \in \mathcal{EBF}. \tag{4.4}$$

Using the fact that the function  $m \mapsto m$  is a trivial Bernstein function and equations (4.3) and (4.4), one gets

$$-\ell_{\mu^*}(m) = -\int \psi_{\mu}(m)dm = \frac{1}{4}m + \frac{1}{2}\log(m) + \frac{1}{4}\sqrt{m+4} \left( \sqrt{m} + \frac{4 \sinh^{-1}(\sqrt{m}/2)}{\sqrt{m+4}} \right) + C \in \mathcal{EBF}.$$

Thus  $\mu^*$  has a dual measure. For the link and variance functions, we consider  $K_{Q_1}(\theta) = -2\sqrt{-\theta}$  and  $K_{\mu_1}(\theta_0) = -\log(1 - \theta_0)$  for all  $\theta \in (-\infty, 0)$  and  $\theta_0 \in (-\infty, 1)$ . For  $s \in (0, +\infty)$ , one has  $K_{\mu_1}(K_{Q_1}(-s)) = -\log(1 + 2\sqrt{s})$ . From the first and second derivatives of this function, we obtain  $\psi_{\mu^*}$  and  $V_{F(\mu^*)}$ .

3. According to Mselmi (2021a), supplementary material, example 6 and Remark 2.4.2(1.), one has  $-\psi_{\mu}(m) = \frac{1}{2} \left( \mathcal{W} \left( \frac{1}{m} \right) \right)^2$ . From Theorem 3.4 (b) of Kalugin *et al.* (2012), we obtain  $-\psi_{\mu}(m) \in \mathcal{CM}$ . Integrating  $-\psi_{\mu}(m)$ , one gets

$$-\ell_{\mu^*}(m) = \frac{1}{2} \exp \left( -\mathcal{W} \left( \frac{1}{m} \right) \right) \left( \mathcal{W} \left( \frac{1}{m} \right) + 2 \right) + C,$$

where  $C \in \mathbf{R}$ . Since  $\lim_{m \rightarrow +\infty} -\ell_{\mu^*}(m) = +\infty$ , then, by using the definition of limits, we deduce that for  $\mathfrak{B} \gg 1$  large enough, there exists  $\mathfrak{A} \gg 0$  large enough such that

$$\mathfrak{B} = \frac{1}{2} \exp \left( -\mathcal{W} \left( \frac{1}{\mathfrak{A}} \right) \right) \left( \mathcal{W} \left( \frac{1}{\mathfrak{A}} \right) + 2 \right) + C \approx 1 + C.$$

It follows that  $C \approx \mathfrak{B} - 1 > 0$   $\left( \lim_{m \rightarrow +\infty} \mathcal{W} \left( \frac{1}{m} \right) = 0 \right)$ . Thus  $C > 0$  and then  $-\ell_{\mu^*}(m) \geq 0$ . This together with the fact that  $-\psi_{\mu}(m) \in \mathcal{CM}$  imply  $-\ell_{\mu^*}(m) \in \mathcal{BF}$  (see Lemma 4.2.5). Thus,  $\mu^*$  has a dual measure. In this case  $K_{\mu_1}(K_{Q_1}(-s)) = \exp(-2\sqrt{s})$ . From the first and second derivatives of this function, we obtain  $\psi_{\mu^*}$  and  $V_{F(\mu^*)}$ .

4. The compound Poisson distribution is characterized by its variance function which equal to  $m^p$  with  $p \in (1, 2)$ . According to Mselmi (2021a), Corollary 1.(1), one has

$$-\psi_{\mu}(m) = \frac{1}{p-1} m^{1-p} \exp \left( \frac{1-p}{p-2} \mathcal{W}(m^{2-p}) \right) = \frac{1}{p-1} \left( \mathcal{W}(m^{2-p}) \right)^{\frac{1-p}{2-p}}.$$

From Theorem 3.4 (c) of [Kalugin et al. \(2012\)](#), we obtain  $-\psi_\mu(m) \in \mathcal{CM}$ . Integrating  $-\psi_\mu(m)$ , we obtain

$$-\ell_{\mu^*}(m) = \frac{1}{p-1} \mathcal{W}(m^{2-p}) \exp\left(\frac{1}{2-p} \mathcal{W}(m^{2-p})\right) + \exp\left(\frac{1}{2-p} \mathcal{W}(m^{2-p})\right) + C,$$

where  $C \in \mathbf{R}$ . Since  $-\ell_{\mu^*}(0) = 1 + C = 0$ , then  $C = -1$ . Using the fact that  $\exp\left(\frac{1}{2-p} \mathcal{W}(m^{2-p})\right) \geq 1$ , we conclude that  $-\ell_{\mu^*}(m) \geq 0$ . This together with the fact that  $-\psi_\mu(m) \in \mathcal{CM}$  imply  $-\ell_{\mu^*}(m) \in \mathcal{BF}$  (see Lemma 4.2.5). Hence,  $\mu^*$  has a dual measure. For all  $s \in (0, +\infty)$ , one has  $K_{\mu_1}(K_{Q_1}(-s)) = \exp\left(\frac{1}{2-p} [(p-1)s]^{\frac{p-2}{p-1}}\right)$ . From the first and second derivatives of this function, we obtain  $\psi_{\mu^*}$  and  $V_{F(\mu^*)}$ .

5. The stable distribution is characterized by its variance function which equal to  $m^{\frac{\alpha-2}{\alpha-1}}$  with  $\alpha \in (0, 2) \setminus \{1\}$ . According to [Mselmi \(2021a\)](#), Remark 2.(1), one has

$$-\psi_\mu(m) = (1-\alpha)m^{\frac{1}{\alpha-1}} \exp\left(-\frac{1}{\alpha} \mathcal{W}\left(m^{\frac{\alpha}{\alpha-1}}\right)\right) = (1-\alpha) \mathcal{W}\left(m^{\frac{\alpha}{\alpha-1}}\right)^{\frac{1}{\alpha}}.$$

For  $\alpha \in (0, 1/2)$  and from Theorem 3.4 (b) of [Kalugin et al. \(2012\)](#), we obtain  $-\psi_\mu(m) \in \mathcal{CM}$ . Hence

$$-\ell_{\mu^*}(m) = (1-\alpha) \left( \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) + \frac{1}{1-\alpha} \right) \exp\left(\frac{\alpha-1}{\alpha} \mathcal{W}(m^{\frac{\alpha}{\alpha-1}})\right) + C, \text{ where } C \in \mathbf{R}.$$

As in (3.), we deduce for  $\mathfrak{B} \gg 1$  large enough, there exists  $\mathfrak{A} \gg 0$  large enough such that

$$\mathfrak{B} = (1-\alpha) \left( \mathcal{W}(\mathfrak{A}^{\frac{\alpha}{\alpha-1}}) + \frac{1}{1-\alpha} \right) \exp\left(\frac{\alpha-1}{\alpha} \mathcal{W}(\mathfrak{A}^{\frac{\alpha}{\alpha-1}})\right) + C \approx 1 + C.$$

It follows that  $C \approx \mathfrak{B} - 1 > 0$ . Thus  $C > 0$  and then  $-\ell_{\mu^*}(m) \geq 0$ . This together with the fact that  $-\psi_\mu(m) \in \mathcal{CM}$  imply  $-\ell_{\mu^*}(m) \in \mathcal{BF}$  (see Lemma 4.2.5). Hence,  $\mu^*$  has a dual measure. For all  $s \in (0, +\infty)$ , one has  $K_{\mu_1}(K_{Q_1}(-s)) = \exp\left(\frac{\alpha-1}{\alpha} \left[\frac{s}{1-\alpha}\right]^\alpha\right)$ . From the first and second derivatives of this function, we obtain  $\psi_{\mu^*}$  and  $V_{F(\mu^*)}$ . ■

**Lemma 4.2.6** (Proposition 3 of [Letac \(2021\)](#)) Let  $\mu \in \mathcal{M}$  and assume that a dual measure  $\mu^*$  does exist. Then  $\mu^*$  is infinitely divisible if, and only if,  $m \mapsto \frac{1}{V_{F(\mu)}(m)}$  is the Laplace transform of some measure.

Now, we are able to express the following corollary which study the infinite divisibility property of some examples.

**Proof of Corollary 2.4.4**

1. The gamma gamma case: According to [Mselmi \(2021a\)](#), supplementary material, example 1, we deduce that  $V_{F(\mu)}(m) = m^2 \left[ 1 + \mathcal{W}\left(\frac{1}{m}\right) \right]$ . Then,  $\frac{1}{V_{F(\mu)}(m)} = \frac{1}{m^2 \left[ 1 + \mathcal{W}\left(\frac{1}{m}\right) \right]}$ . Since  $\frac{1}{m \left[ 1 + \mathcal{W}\left(\frac{1}{m}\right) \right]} \in \mathcal{S}$  and  $m \mapsto \frac{1}{m}$  is a trivial Stieltjes function, thus  $\frac{1}{V_{F(\mu)}(m)}$  is the product of two Stieltjes functions. By using Corollary 1.6 of [Schilling et al. \(2010\)](#), page 5, we deduce that  $\frac{1}{V_{F(\mu)}(m)} \in \mathcal{CM}$  (recall that  $\mathcal{S} \subset \mathcal{CM}$ ). Lemma 4.2.6 implies that the dual measure of the gamma gamma distribution is infinitely divisible.
2. The inverse Gaussian Poisson case: According to [Mselmi \(2021a\)](#), supplementary material, example 6, we deduce that  $V_{F(\mu)}(m) = \frac{m}{\mathcal{W}\left(\frac{1}{m}\right)} \left[ 1 + \frac{1}{\mathcal{W}\left(\frac{1}{m}\right)} \right] = \frac{m}{\left(\mathcal{W}\left(\frac{1}{m}\right)\right)^2} \left[ 1 + \mathcal{W}\left(\frac{1}{m}\right) \right]$ . Then,  $\frac{1}{V_{F(\mu)}(m)} = \frac{\left(\mathcal{W}\left(\frac{1}{m}\right)\right)^2}{m \left[ 1 + \mathcal{W}\left(\frac{1}{m}\right) \right]}$ . Since the function  $\frac{1}{m \left[ 1 + \mathcal{W}\left(\frac{1}{m}\right) \right]} \in \mathcal{S} \subset \mathcal{CM}$  and  $\left(\mathcal{W}\left(\frac{1}{m}\right)\right)^2 \in \mathcal{CM}$  (see Theorem 3.4(b) of [Kalugin et al. \(2012\)](#)), thus  $\frac{1}{V_{F(\mu)}(m)} \in \mathcal{CM}$  (see Corollary 1.6 of [Schilling et al. \(2010\)](#), page 5). This and Lemma 4.2.6 imply that the dual measure of the inverse Gaussian Poisson distribution is infinitely divisible.
3. The compound Poisson Poisson case: According to [Mselmi \(2021a\)](#), corollary 1, we deduce that

$$V_{F(\mu)}(m) = m^p e^{\frac{p-1}{p-2} \mathcal{W}(m^{2-p})} + m^2 e^{\frac{1}{p-2} \mathcal{W}(m^{2-p})} = m \left(\mathcal{W}(m^{2-p})\right)^{-\frac{p-1}{p-2}} \left[ 1 + \mathcal{W}(m^{2-p}) \right].$$

Note that the right side of the latter equation was obtained from  $e^{\mathcal{W}(m^{2-p})} = \frac{m^{2-p}}{\mathcal{W}(m^{2-p})}$ .

Then,  $\frac{1}{V_{F(\mu)}(m)} = \frac{\left(\mathcal{W}(m^{2-p})\right)^{\frac{p-1}{p-2}}}{m \left[ 1 + \mathcal{W}(m^{2-p}) \right]}$ . From [Kalugin et al. \(2012\)](#) Theorem 3.4(c), we

deduce that  $\frac{1}{V_{F(\mu)}(m)} \in \mathcal{CM}$  (in fact  $-1 < p-2 < 0$  and  $\frac{p-1}{p-2} < 0$ ). This and Lemma 4.2.6 imply that the dual measure of the compound Poisson distribution is infinitely divisible.

4. The stable-Poisson case ( $\alpha \in (0, 1/2)$ ): According to Mselmi (2021a), remark 2(1), we deduce that

$$V_{F(\mu)}(m) = m^{\frac{\alpha-2}{\alpha-1}} e^{\frac{1}{\alpha}} \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) + m^2 e^{\frac{1-\alpha}{\alpha}} \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) = m \left( \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \right)^{-\frac{1}{\alpha}} [1 + \mathcal{W}(m^{\frac{\alpha}{\alpha-1}})].$$

Then,  $\frac{1}{V_{F(\mu)}(m)} = \frac{\left( \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \right)^{\frac{1}{\alpha}}}{m \left[ 1 + \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \right]}$ . From Kalugin *et al.* (2012) Theorem 3.4(b),

we deduce that  $\left( \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \right)^{\frac{1}{\alpha}} \in \mathcal{CM}$  (in fact  $-1 < \frac{\alpha}{\alpha-1} < 0$  and  $\frac{1}{\alpha} > 0$ ). So, our main problem, now, is to show that  $m^{-1} \left[ 1 + \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \right]^{-1} \in \mathcal{CM}$ . According to Kalugin *et al.* (2012), Section 4, one has  $\mathcal{W}(m) \in \mathcal{CBF}$ . Furthermore,  $m^{-\frac{\alpha}{\alpha-1}} \in \mathcal{CBF}$  (see Remark 7.8 of Schilling *et al.* (2010) page 65). It follows that the composite function  $\mathcal{W}(m^{-\frac{\alpha}{\alpha-1}}) \in \mathcal{CBF}$  (see Corollary 7.9 of Schilling *et al.* (2010) page 66). Using equation (7.3) of Schilling *et al.* (2010), page 66, we obtain  $m\mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \in \mathcal{CBF}$ . Since the function  $m \mapsto m$  is a trivial  $\mathcal{CBF}$  (see Remark 7.8, Chapter 7 of Schilling *et al.* (2010)), then  $m + m\mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \in \mathcal{CBF}$ . Using Lemma 4.2.5, one gets  $\frac{1}{m + m\mathcal{W}(m^{\frac{\alpha}{\alpha-1}})} \in \mathcal{S} \subset \mathcal{CM}$ .

From this and the fact that  $\left( \mathcal{W}(m^{\frac{\alpha}{\alpha-1}}) \right)^{\frac{1}{\alpha}} \in \mathcal{CM}$ , we deduce that  $\frac{1}{V_{F(\mu)}(m)} \in \mathcal{CM}$  (see Corollary 1.6 of Schilling *et al.* (2010), page 5). This and Lemma 4.2.6 imply that the dual measure of the stable Poisson distribution (with  $\alpha \in (0, 1/2)$ ) is infinitely divisible. ■



# Appendix 2

## Proof of Theorem 3.1.1

Using (1.1), we get

$$\begin{aligned} K'_\eta(\theta_1, \theta_2) &= \left( \frac{\partial K_\eta}{\partial \theta_1}(\theta_1, \theta_2), \frac{\partial K_\eta}{\partial \theta_2}(\theta_1, \theta_2) \right) \\ &= (K'_{\mu_1}(\theta_1 + K_{Q_1}(\theta_2)), K'_{Q_1}(\theta_2)K'_{\mu_1}(\theta_1 + K_{Q_1}(\theta_2))) =: (\mathbf{m}, m). \end{aligned} \quad (4.5)$$

Using the fact that  $K'_{\mu_1}(\theta_1 + K_{Q_1}(\theta_2)) = \mathbf{m} > 0$  and (4.5), we deduce, for all  $m \in M_{F(\mu)} = M_{F(Q_1)}$ , that

$$\theta_2 = \psi_{Q_1}\left(\frac{m}{\mathbf{m}}\right) \text{ and } \theta_1 = \psi_{\mu_1}(\mathbf{m}) - K_{Q_1}\left(\psi_{Q_1}\left(\frac{m}{\mathbf{m}}\right)\right). \quad (4.6)$$

By taking  $\theta_1 = 0$  and using (1.2), we conclude that

$$\frac{\partial K_\eta}{\partial \theta_1}(\theta_1, \theta_2) \Big|_{\theta_1=0} = K'_{\mu_1}(K_{Q_1}(\theta_2)) = \mathbf{m}(0, \theta_2) \quad (4.7)$$

and

$$K'_\mu(\theta_2) = \frac{\partial K_\eta}{\partial \theta_2}(\theta_1, \theta_2) \Big|_{\theta_1=0} = K'_{Q_1}(\theta_2)K'_{\mu_1}(K_{Q_1}(\theta_2)) = m(0, \theta_2). \quad (4.8)$$

From the last equality, we obtain  $\theta_2 = \psi_\mu(m(0, \theta_2))$ . This together with (4.6) imply

$$\theta_2 = \psi_\mu(m(0, \theta_2)) = \psi_{Q_1}\left(\frac{m(0, \theta_2)}{\mathbf{m}(0, \theta_2)}\right) \text{ and } 0 = \psi_{\mu_1}(\mathbf{m}(0, \theta_2)) - K_{Q_1}\left(\psi_{Q_1}\left(\frac{m(0, \theta_2)}{\mathbf{m}(0, \theta_2)}\right)\right). \quad (4.9)$$

Without loss of generality and in what follows, we denote by  $\mathbf{m} := \mathbf{m}(0, \theta_2)$  and  $m := m(0, \theta_2)$ .

Let  $\mathbf{f} : \mathbf{R}^{d-1} \times (0, \infty) \rightarrow (0, \infty); (\mathbf{z}, \mathbf{t}) \mapsto \mathbf{f}(\mathbf{z}, \mathbf{t}) = \psi_{\mu_1}(\mathbf{t}) - K_{Q_1}(\psi_{Q_1}(\mathbf{z}/\mathbf{t}))$ . The function  $\mathbf{f}$  is continuously differentiable and from the second equation of (4.9), we obtain  $\mathbf{f}(m, \mathbf{m}) = 0$ . The partial derivative of  $\mathbf{f}$  with respect to  $\mathbf{t}$  is given by

$$\frac{\partial \mathbf{f}}{\partial \mathbf{t}}(m, \mathbf{m}) = [V_{F(\mu_1)}(\mathbf{m})]^{-1} + \frac{1}{\mathbf{m}^3} m \left[ V_{F(Q_1)}\left(\frac{m}{\mathbf{m}}\right) \right]^{-1} m^\top.$$



Since  $\mathbf{m} > 0$ ,  $V_{F(\mu_1)}(\mathbf{m}) > 0$  and  $V_{F(Q_1)}(m/\mathbf{m})$  is a symmetric definite positive matrix, then we deduce that there exists  $\delta > 0$  such that  $\frac{\partial \mathbf{f}}{\partial \mathbf{t}}(m, \mathbf{m}) > \delta$ . The global implicit function theorem given by [Zhang and Ge \(2006\)](#) implies that there exists a unique continuous function  $g : \mathbf{R}^{d-1} \rightarrow (0, \infty)$  such that  $\mathbf{f}(m, g(m)) = 0$ . Differentiating  $\mathbf{f}$  with respect to  $m$ , we obtain

$$\mathbf{f}'(m, g(m)) = \frac{\partial \mathbf{f}}{\partial \mathbf{z}}(m, g(m)) + \frac{\partial \mathbf{f}}{\partial \mathbf{t}}(m, g(m)) \circ g'(m) = \mathbf{0}_{\mathbf{R}^{d-1}}.$$

It follows that

$$\frac{-1}{[g(m)]^2} \left[ V_{F(Q_1)} \left( \frac{m}{g(m)} \right) \right]^{-1} m^\top + \left\{ [V_{F(\mu_1)}(g(m))]^{-1} + \frac{1}{[g(m)]^3} m \left[ V_{F(Q_1)} \left( \frac{m}{g(m)} \right) \right]^{-1} m^\top \right\} g'(m) = \mathbf{0}_{\mathbf{R}^{d-1}}.$$

This involves that  $g$  satisfies the differential equation [\(3.2\)](#). Inserting  $g$  in the first equality of [\(4.9\)](#), we conclude that Equation [\(3.1\)](#) holds. Setting  $\theta_2 = 0$  in [\(4.7\)](#) and [\(4.8\)](#), we get the initial condition  $g(K'_{Q_1}(0)K'_{\mu_1}(0)) = K'_{\mu_1}(0)$  of [\(3.2\)](#). ■

### Proof of Proposition [3.3.1](#)

Since  $\lambda \in (0, \infty)$  then the function  $G(\mathbf{m}) := f \circ g(\mathbf{m})/\lambda$  represents a multivariate probability density function. Using a multivariate associated kernel estimation method, we obtain  $\widehat{G}(\mathbf{m}) = \frac{1}{n} \sum_{i=1}^n \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\mathbf{m}_i)$ . It follows that

$$\widehat{g}(\mathbf{m}) = f^{-1} \left( \frac{\lambda}{n} \sum_{i=1}^n \mathcal{K}_{\mathbf{m}, \mathbf{H}}(\mathbf{m}_i) \right).$$

Inserting the estimator of  $g$  in the expressions of the link and variance functions, given in [Theorem 3.1.1](#) and [Corollary 3.1.3](#), we obtain the required results. ■