Innovation in Statistics and Probability
2(1), 1–25
DOI:https://doi.org/10.64389/isp.2026.02174
https://sphinxsp.org/journal/index.php/isp/

Research article

Study of a new bivariate trigonometric Gaussian distribution

Julien Samyn[1], **Soan Bailly**[1], **Christophe Chesneau**[1*]

1 Department of Mathematics, LMNO, University of Caen-Normandie, 14032 Caen, France; juliensamyn14@gmail.com, soan.bailly@orange.fr, christophe.chesneau@gmail.com.

* **Correspondence:** christophe.chesneau@gmail.com

ARTICLE INFO

Keywords:
Bivariate Gaussian distribution
Trigonometric distribution
Nonlinear oscillatory modeling
Modified Gaussian distribution
Simulation.

Mathematics Subject Classification:
62E10, 62H05

Important Dates:
Received: 7 March 2026
Revised: 13 April 2026
Accepted: 1 May 2026
Online: 8 May 2026

ABSTRACT

This article introduces a new bivariate distribution derived by extending the standard independent Gaussian framework with trigonometric components. The proposed distribution, known as the bivariate trigonometric Gaussian distribution, enhances the classical Gaussian distribution structure while maintaining key analytical tractability. We outline its key theoretical properties, including the explicit forms of the marginal distributions and several structural characteristics of the joint probability density function. Graphical illustrations highlight the influence of the trigonometric parameters on the shapes of the distributions. Finally, we demonstrate the practical value of this theoretical framework by analyzing a real-world turbine dataset. Our findings demonstrate the empirical superiority of the bivariate trigonometric Gaussian distribution over classical distributions when it comes to fitting complex, nonlinear physical phenomena.

1. Introduction

The analysis of multivariate (probability) distributions is fundamental to statistics and probability theory, as well as their applications in fields such as physics, finance and data science. Bivariate models, in particular, are widely used to describe the dependence and interaction structures of random variables. Gaussian-type distributions remain central among the many available models due to their mathematical tractability and broad applicability. However, classical Gaussian frameworks can be too restrictive for capturing more complex structural features, particularly when nonlinear or oscillatory effects are present. See, for example, [9, 3, 8, 10, 4]. This article contributes to this line of research by introducing a new bivariate trigonometric

Gaussian distribution. This combines the standard independent Gaussian density function with additional trigonometric terms, creating a richer family of probability density functions. This modification enables the distribution to capture structural features that differ from those of the classical Gaussian model, yet it still retains analytical tractability. This new distribution, known as the bivariate trigonometric Gaussian (BTG) distribution, enhances the structure of the classical Gaussian distribution while maintaining key properties.

To bridge the gap between mathematical theory and empirical practice, we propose a comprehensive statistical data analysis pipeline to complement this framework. Specifically, we develop parameter estimation techniques that combine maximum likelihood estimation (MLE) with particle swarm optimization (PSO), as well as an accept-reject simulation algorithm. Furthermore, we establish a robust goodness-of-fit testing methodology based on energy distance via parametric bootstrap. To demonstrate the practical relevance of the BTG distribution, we apply this pipeline to a real-world thermodynamics dataset from a gas turbine. By evaluating variables such as ambient pressure and relative humidity, we demonstrate that our distribution outperforms classical bivariate distributions in capturing complex nonlinear structural dependencies and saturation plateaus.

The rest of the article is organized as follows: Section 2 introduces the BTG distribution and presents its key analytical properties. Section 3 is devoted to studying the marginal and conditional distributions, as well as conditional expectations. Section 4 discusses statistical inference in detail, covering the MLE procedure and the simulation algorithm. Section 5 outlines the applied methodology, including the energy distance test and the model selection framework using information criteria. Section 6 presents the empirical results and performance of our model when applied to the turbine dataset. Finally, Section 7 concludes the article.

2. A new bivariate trigonometric Gaussian distribution

2.1. Definition of the distribution

We begin by introducing a probability density function that extends the standard independent bivariate Gaussian framework by incorporating trigonometric components.

Proposition 2.1. *Let $\alpha, \beta \in \mathbb{R}$. Then the following function is a valid bivariate probability density function:*

$$f(x,y) = K^{-1} e^{-x^2/2-y^2/2} \left(2 + \cos(\alpha x) + \sin(\beta y)\right), \qquad (x,y) \in \mathbb{R}^2,$$

where

$$K = 2\pi \left(2 + e^{-\alpha^2/2}\right).$$

Proof:

Since $\cos(\alpha x) \geq -1$ and $\sin(\beta y) \geq -1$, for any $(x, y) \in \mathbb{R}^2$, we have

$$2 + \cos(\alpha x) + \sin(\beta y) \geq 0,$$

implying that $f(x, y) \geq 0$.

It remains to prove that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1.$$

Set

$$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2/2-y^2/2}\Big(2+\cos(\alpha x)+\sin(\beta y)\Big)dxdy.$$

By linearity, we have

$$I = I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &= 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2/2-y^2/2}dxdy,\\ I_2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2/2-y^2/2}\cos(\alpha x)dxdy,\\ I_3 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2/2-y^2/2}\sin(\beta y)dxdy. \end{aligned}$$

Using

$$\int_{-\infty}^{\infty} e^{-x^2/2}dx = \sqrt{2\pi},$$

we obtain

$$I_1 = 2\left(\sqrt{2\pi}\right)^2 = 4\pi.$$

By the Fubini integral theorem, we have

$$I_2 = \int_{-\infty}^{\infty} e^{-y^2/2}dy\int_{-\infty}^{\infty} e^{-x^2/2}\cos(\alpha x)dx.$$

Using the classical formula

$$\int_{-\infty}^{\infty} e^{-px^2}\cos(qx)dx = \sqrt{\frac{\pi}{p}}e^{-q^2/(4p)},$$

(see, e.g., [6, Formula 3.896.4]) and taking $p = 1/2$ and $q = \alpha$, we obtain

$$\int_{-\infty}^{\infty} e^{-x^2/2}\cos(\alpha x)dx = \sqrt{2\pi}e^{-\alpha^2/2}.$$

Therefore, we estbalish that

$$I_2 = \sqrt{2\pi}e^{-\alpha^2/2}\int_{-\infty}^{\infty} e^{-y^2/2}dy = 2\pi e^{-\alpha^2/2}.$$

Applying again the Fubini integral theorem, we have

$$I_3 = \int_{-\infty}^{\infty} e^{-x^2/2}dx\int_{-\infty}^{\infty} e^{-y^2/2}\sin(\beta y)dy.$$

The inner integrand is odd, hence

$$\int_{-\infty}^{\infty} e^{-y^2/2}\sin(\beta y)dy = 0,$$

so $I_3 = 0$.

Combining the main equations above, we get

$$I = 4\pi + 2\pi e^{-\alpha^2/2} = 2\pi(2 + e^{-\alpha^2/2}) = K.$$

Thus

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = K^{-1}K = 1.$$

This completes the proof of the proposition. □

Definition 2.2. *(BTG distribution) For $\alpha, \beta \in \mathbb{R}$, the BTG distribution is defined by the probability density function*

$$f(x, y) = K^{-1}e^{-x^2/2-y^2/2}\Big(2 + \cos(\alpha x) + \sin(\beta y)\Big), \qquad (x, y) \in \mathbb{R}^2$$

where

$$K = 2\pi(2 + e^{-\alpha^2/2}).$$

We will sometime denote this distribution as BTG(α, β) to specify the parameters involved.

2.2. Fundamental property

- **Evenness with respect to x.** Since both the square and cosine functions are even, the probability density function is even in x, i.e.,

  $$f(-x, y) = f(x, y),$$

  for any $(x, y) \in \mathbb{R}^2$.
- **Lack of symmetry in y.** In general, we have

  $$f(x, -y) \neq f(x, y) \qquad \text{and} \qquad f(x, -y) \neq -f(x, y),$$

  so the probability density function is neither even nor odd with respect to y.
- **Asymmetry between coordinates.** The probability density function is not symmetric in (x, y); in general,

  $$f(x, y) \neq f(y, x).$$
- **Gaussian limit case.** When $\alpha = 0$ and $\beta = 0$, since $\cos(0) = 1$, $\sin(0) = 0$ and

  $$K = 2\pi(2 + 1) = 6\pi,$$

  we obtain

  $$f(x, y) = \frac{1}{2\pi}e^{-x^2/2-y^2/2},$$

  which is the standard bivariate Gaussian distribution with independent components.
- **Gaussian domination.** Since $|\cos(\alpha x)| \leq 1$ and $|\sin(\beta y)| \leq 1$, we have

  $$2 + \cos(\alpha x) + \sin(\beta y) \leq 4,$$

  hence

  $$f(x, y) \leq 4K^{-1}e^{-x^2/2-y^2/2}.$$

  This Gaussian upper bound ensures integrability and guarantees the existence of moments of all orders for the BTG distribution.

To provide a visual representation, Figures 1, 2, 3 and 4 display perspective plots of $f(x, y)$ for $(\alpha, \beta) = (1, 1)$, $(\alpha, \beta) = (3, 4)$, $(\alpha, \beta) = (5, 1)$ and $(\alpha, \beta) = (2, 20)$, respectively.

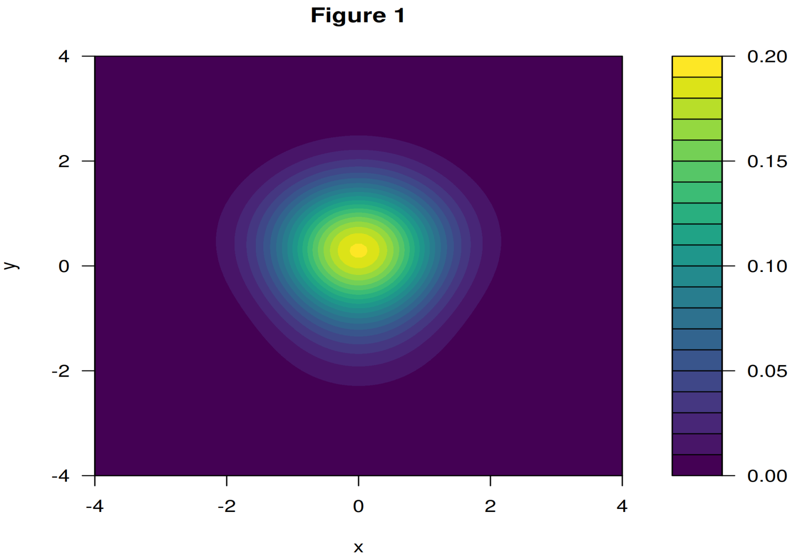

Figure 1. Perspective plot of the probability density function of the BTG distribution for $(\alpha, \beta) = (1, 1)$.

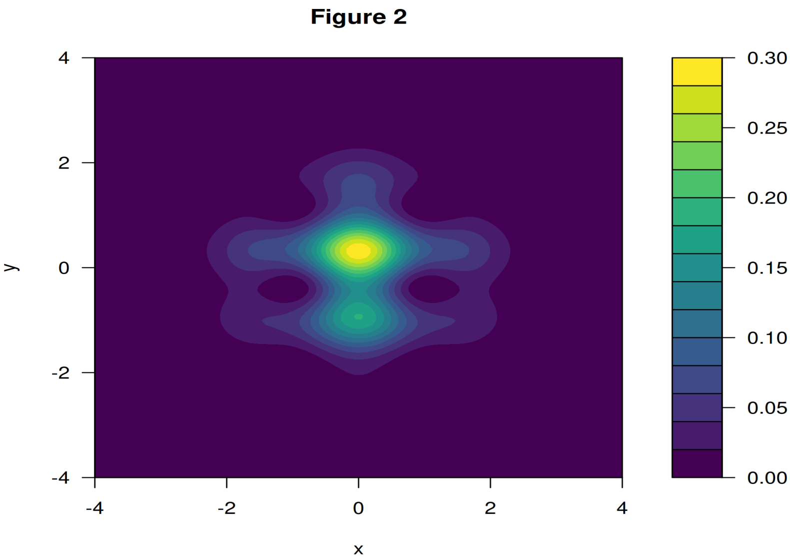

Figure 2. Perspective plot of the probability density function of the BTG distribution for $(\alpha, \beta) = (3, 4)$.

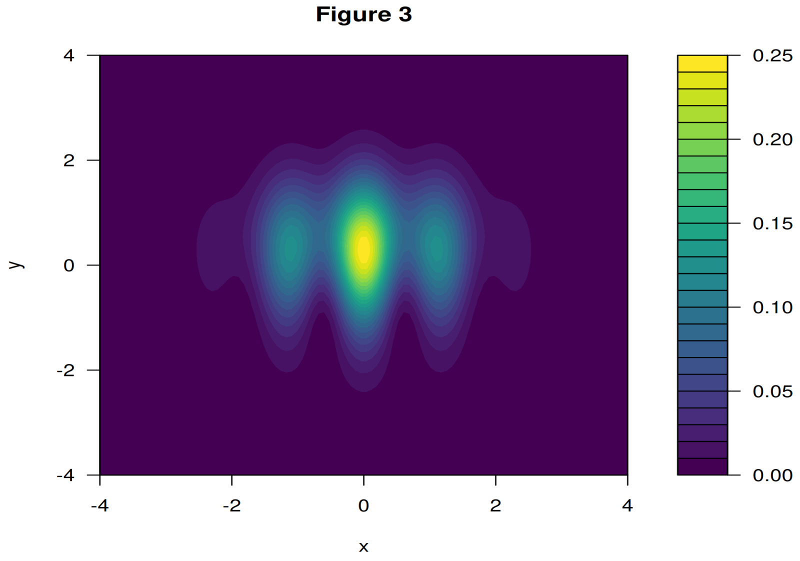

Figure 3. Perspective plot of the probability density function of the BTG distribution for $(\alpha, \beta) = (5, 1)$.

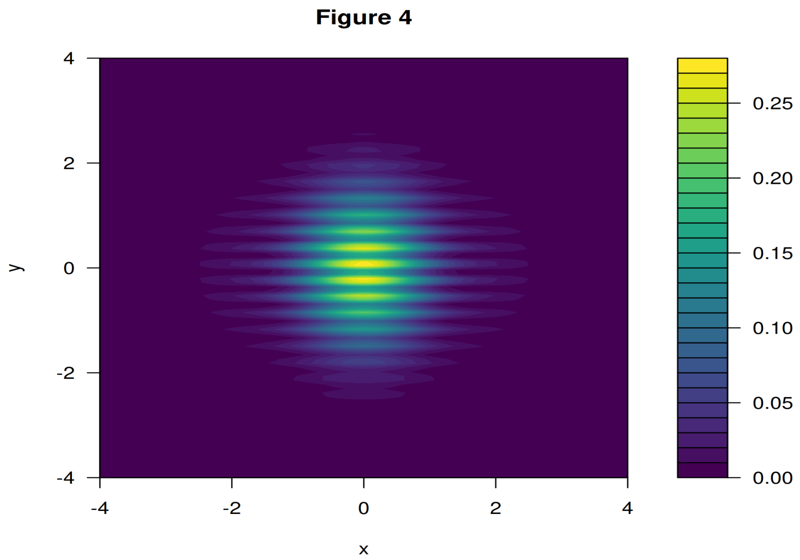

Figure 4. Perspective plot of the probability density function of the BTG distribution for $(\alpha, \beta) = (2, 20)$.

The probability density function can be written as a standard Gaussian term multiplied by a periodic modulation:

$$f(x, y) = K^{-1} e^{-x^2/2 - y^2/2} (2 + \cos(\alpha x) + \sin(\beta y)).$$

The Gaussian part $e^{-x^2/2 - y^2/2}$ gives the usual bell-shaped structure centered at the origin. The cosine and sine terms introduce oscillations.

The parameter α appears only in the term $\cos(\alpha x)$. Therefore, it affects the probability density function only in the x-direction. When $|\alpha|$ increases, the oscillations along the x-axis become more frequent. The amplitude does not change, only the frequency.

Similarly, the parameter β appears only in $\sin(\beta y)$. It controls oscillations in the y-direction. Larger values of $|\beta|$ produce faster vertical oscillations.

Since there is no interaction term such as xy, the effects of α and β are independent. The oscillations are horizontal for α and vertical for β, while the overall shape remains that of the standard Gaussian distribution..

3. Marginal probability density functions

The following proposition provides the explicit forms of the marginal distributions of the BTG distribution.

Proposition 3.1. *Let (X, Y) be a random vector following the BTG distribution. The marginal distributions of X and Y admit the following probability density functions.*

- *The marginal probability density function of X is given by*

$$g(x) = L\,(2 + \cos(\alpha x))\, e^{-x^2/2}, \qquad x \in \mathbb{R}.$$

- *The marginal probability density function of Y is given by*

$$h(y) = L\left(2 + e^{-\alpha^2/2} + \sin(\beta y)\right) e^{-y^2/2}, \qquad y \in \mathbb{R},$$

where

$$L = \sqrt{2\pi} K^{-1} = \frac{1}{\sqrt{2\pi}\left(2 + e^{-\alpha^2/2}\right)}.$$

Proof:

We recall that (X, Y) has the following probability density function:

$$f(x, y) = K^{-1} e^{-x^2/2 - y^2/2} (2 + \cos(\alpha x) + \sin(\beta y)).$$

- Marginal probability density function of X.
  By definition, we have

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy.$$

  Hence,

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} K^{-1} e^{-x^2/2 - y^2/2} (2 + \cos(\alpha x) + \sin(\beta y)) dy \\ &= K^{-1} e^{-x^2/2} (2 + \cos(\alpha x)) \int_{-\infty}^{\infty} e^{-y^2/2} dy + K^{-1} e^{-x^2/2} \int_{-\infty}^{\infty} \sin(\beta y) e^{-y^2/2} dy. \end{aligned}$$

Since

$$\int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{2\pi}$$

and

$$\int_{-\infty}^{\infty} \sin(\beta y) e^{-y^2/2} dy = 0,$$

we obtain

$$g(x) = K^{-1} \sqrt{2\pi}\,(2 + \cos(\alpha x))\, e^{-x^2/2}.$$

Setting $L = \sqrt{2\pi}K^{-1}$ gives the announced expression for $g(x)$.

- Marginal probability density function of Y.
  Similarly, we have

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

Thus, we get

$$\begin{aligned} h(y) &= \int_{-\infty}^{\infty} K^{-1} e^{-x^2/2 - y^2/2} (2 + \cos(\alpha x) + \sin(\beta y)) dx \\ &= K^{-1} e^{-y^2/2} \left[2 \int_{-\infty}^{\infty} e^{-x^2/2} dx + \int_{-\infty}^{\infty} e^{-x^2/2} \cos(\alpha x) dx + \sin(\beta y) \int_{-\infty}^{\infty} e^{-x^2/2} dx \right]. \end{aligned}$$

Using

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

and

$$\int_{-\infty}^{\infty} e^{-x^2/2} \cos(\alpha x) dx = \sqrt{2\pi} e^{-\alpha^2/2},$$

we obtain

$$h(y) = K^{-1} \sqrt{2\pi} \left(2 + e^{-\alpha^2/2} + \sin(\beta y)\right) e^{-y^2/2}.$$

Again, substituting $L = \sqrt{2\pi}K^{-1}$ yields the desired expression for $h(y)$.

This completes the proof of the proposition. □

Graphical representation

Before proceeding to further analytical developments, Figure 5 illustrates the behavior of $g(x)$ for $\alpha \in \{0, 0.5, 1.5, 3\}$. Figure 6, Figure 7 and Figure 8 display the corresponding marginal probability density function $h(y)$ for fixed α and $\beta \in \{0, 0.5, 1.5, 3\}$. Finally, Figure 9 also presents $h(y)$ for β fixed at 2 and $\alpha \in \{0, 1, 3, 6\}$.

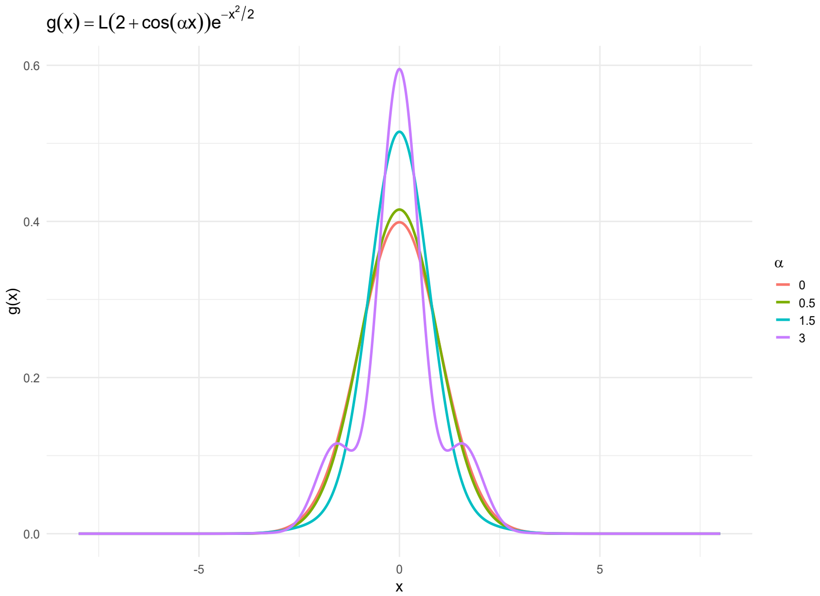

Figure 5. Standard plots of the marginal probability density function with respect to x for $\alpha = 0, \alpha = 0.5, \alpha = 1.5$ and $\alpha = 3$

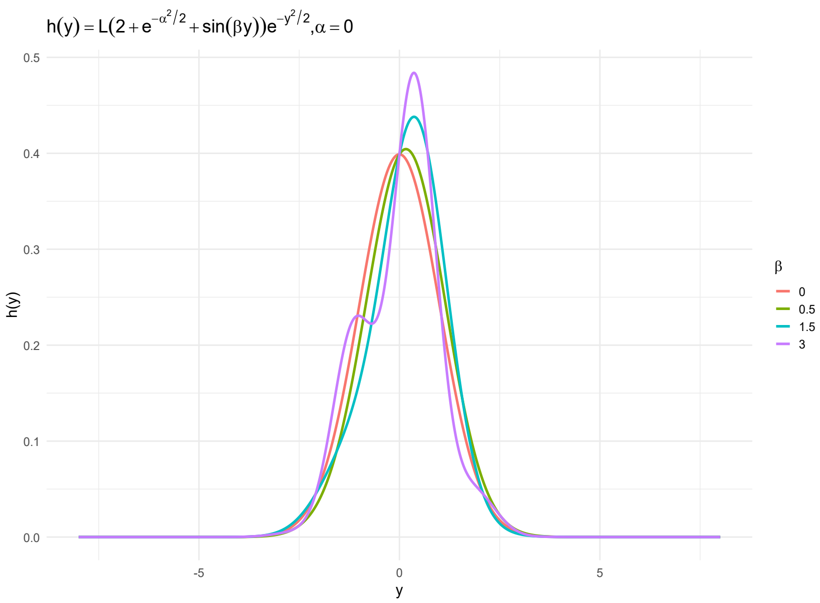

Figure 6. Standard plots of the marginal probability density function with respect to y for $\beta = 0, \beta = 0.5, \beta = 1.5$ and $\beta = 3$, α fixed to 0

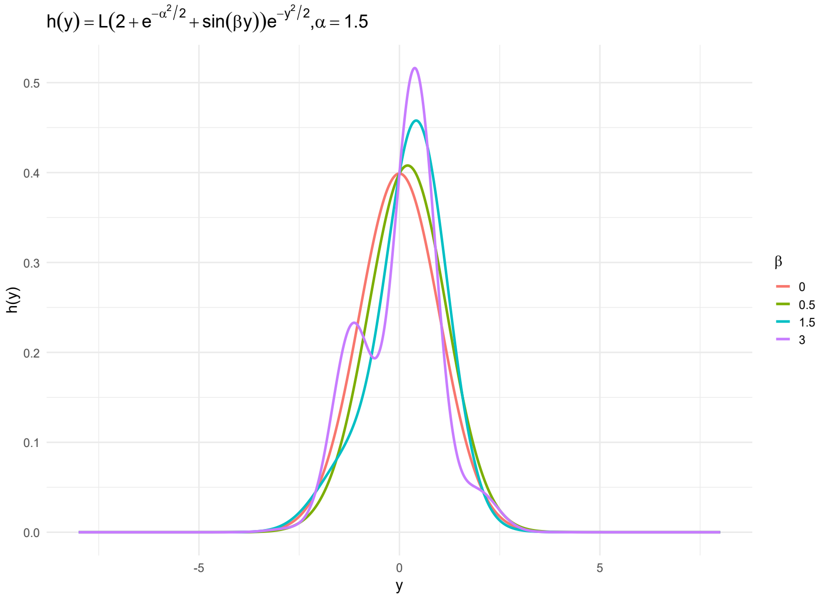

Figure 7. Standard plots of the marginal probability density function with respect to y for $\beta = 0, \beta = 0.5, \beta = 1.5$ and $\beta = 3$, α fixed to 1.5

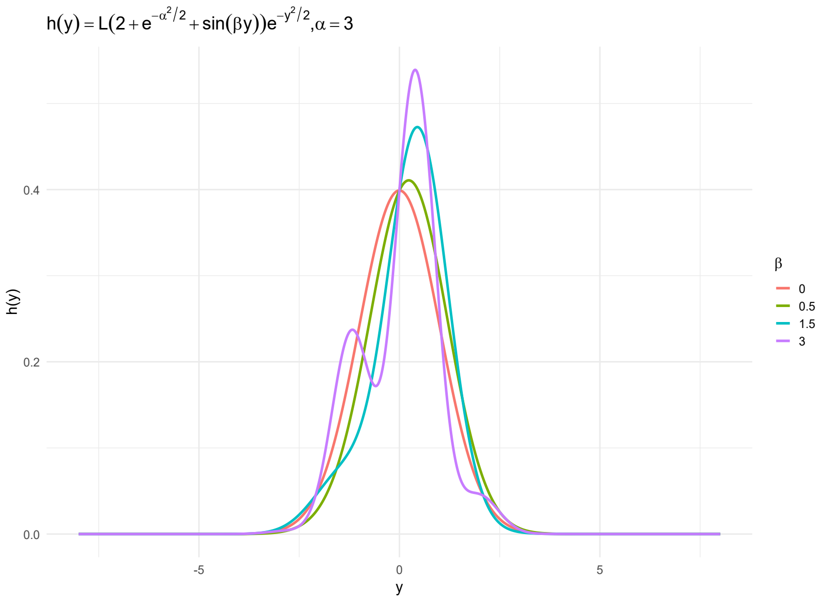

Figure 8. Standard plots of the marginal probability density function with respect to y for $\beta = 0, \beta = 0.5, \beta = 1.5$ and $\beta = 3$, α fixed to 3

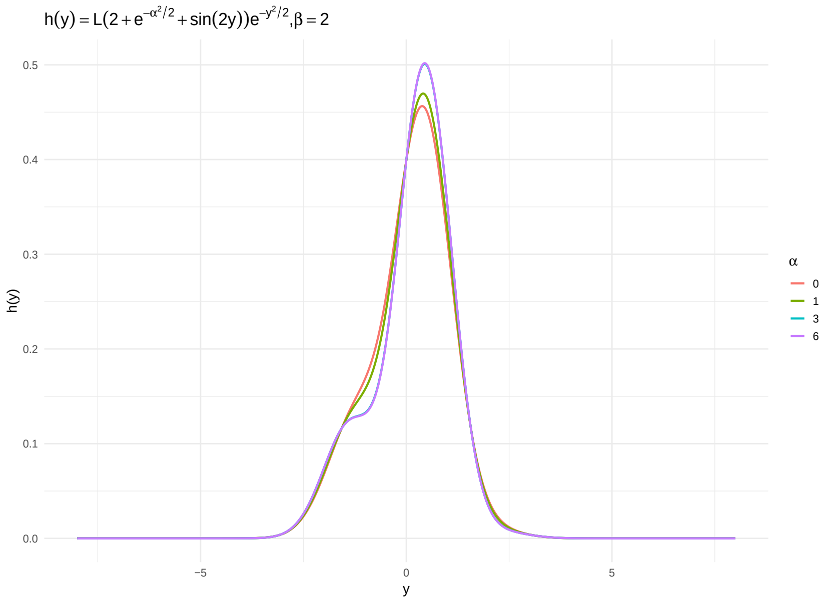

Figure 9. Standard plots of the marginal probability density function with respect to y for $\alpha = 0, \alpha = 1, \alpha = 3$ and $\alpha = 6, \beta$ fixed to 2

From the graphical representations, one observes that the parameter α has only a very limited impact on the marginal density in y.

Indeed, although $h(y)$ depends on α, this dependence appears only through the normalizing constant

$$\frac{1}{2 + e^{-\alpha^2/2}}.$$

Since $e^{-\alpha^2/2} \in]0, 1]$, the quantity $2 + e^{-\alpha^2/2}$ varies between 2 and 3, so that

$$\frac{1}{2 + e^{-\alpha^2/2}} \in \left[\frac{1}{3}, \frac{1}{2}\right).$$

Consequently, changing α modifies only the amplitude of the oscillatory term $\sin(\beta y)$, and this modification remains moderate. The Gaussian envelope $e^{-y^2/2}$, which dominates the overall shape of the curve, is completely unaffected by α.

This explains why, on the graphs, the curves corresponding to different values of α are very close to each other: the global bell-shaped structure remains essentially unchanged, and only small vertical adjustments in the oscillation amplitude can be observed.

3.1. Conditional densities

The proposition below determines the conditional distributions associated with the BTG distribution.

Proposition 3.2. *Let (X, Y) be a random vector following the BTG distribution. Then the conditional distributions of X and Y are defined by the following probability density functions.*

- The probability density function of Y given $X = x$ with $x \in \mathbb{R}$ is given by

$$k(y \mid x) = \phi(y)\frac{2 + \cos(\alpha x) + \sin(\beta y)}{2 + \cos(\alpha x)},$$

where

$$\phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

- The probability density function of X given $Y = y$ with $y \in \mathbb{R}$ is given by

$$l(x \mid y) = \phi(x) \frac{2 + \cos(\alpha x) + \sin(\beta y)}{2 + e^{-\alpha^2/2} + \sin(\beta y)}.$$

Proof:

- By definition, the conditional probability density function of Y given $X = x$ is given by

$$k(y \mid x) = \frac{f(x, y)}{g(x)}.$$

Using the previous expressions, we get

$$\begin{aligned} k(y \mid x) &= \frac{K^{-1} e^{-x^2/2 - y^2/2}(2 + \cos(\alpha x) + \sin(\beta y))}{\frac{\sqrt{2\pi}}{K}(2 + \cos(\alpha x)) e^{-x^2/2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \frac{2 + \cos(\alpha x) + \sin(\beta y)}{2 + \cos(\alpha x)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \left(1 + \frac{\sin(\beta y)}{2 + \cos(\alpha x)}\right) \\ &= \phi(y) \left(1 + \frac{\sin(\beta y)}{2 + \cos(\alpha x)}\right). \end{aligned}$$

- Similarly, the conditional probability density function of X given $Y = y$ is given by

$$l(x \mid y) = \frac{f(x, y)}{h(y)}.$$

We obtain

$$\begin{aligned} l(x \mid y) &= \frac{K^{-1} e^{-x^2/2 - y^2/2}(2 + \cos(\alpha x) + \sin(\beta y))}{\frac{\sqrt{2\pi}}{K}(2 + e^{-\alpha^2/2} + \sin(\beta y)) e^{-y^2/2}} \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{2 + \cos(\alpha x) + \sin(\beta y)}{2 + e^{-\alpha^2/2} + \sin(\beta y)} \\ &= \phi(x) \frac{2 + \cos(\alpha x) + \sin(\beta y)}{2 + e^{-\alpha^2/2} + \sin(\beta y)}. \end{aligned}$$

This completes the proof of the proposition. □

3.2. Conditional expectation

The following proposition determines the conditional expectation associated with the BTG distribution.

Proposition 3.3. *Let* (X, Y) *be a random vector following the BTG distribution. Then we have*

$$\mathbb{E}(X \mid Y = y) = 0, \qquad \mathbb{E}(Y \mid X = x) = \frac{\beta e^{-\beta^2/2}}{2 + \cos(\alpha x)}.$$

Proof:

- By definition, the conditional expectation of Y given $X = x$ is given by

$$\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y k(y \mid x) dy.$$

Recall that the conditional density of Y given $X = x$ is

$$k(y \mid x) = \phi(y)\left(1 + \frac{\sin(\beta y)}{2 + \cos(\alpha x)}\right), \qquad \phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

Substituting the expression of $k(y \mid x)$, we obtain

$$\begin{aligned}\mathbb{E}(Y \mid X = x) &= \int_{-\infty}^{\infty} y\phi(y)\left(1 + \frac{\sin(\beta y)}{2 + \cos(\alpha x)}\right) dy \\ &= \int_{-\infty}^{\infty} y\phi(y) dy + \frac{1}{2 + \cos(\alpha x)} \int_{-\infty}^{\infty} y\phi(y) \sin(\beta y) dy.\end{aligned}$$

Since the standard Gaussian distribution is centered, we have

$$\int_{-\infty}^{\infty} y\phi(y) dy = 0.$$

It remains to compute

$$I(\beta) = \int_{-\infty}^{\infty} y\phi(y) \sin(\beta y) dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-y^2/2} \sin(\beta y) dy.$$

Using

$$\frac{d}{dy}\left(e^{-y^2/2}\right) = -y e^{-y^2/2},$$

we write

$$I(\beta) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sin(\beta y) \frac{d}{dy}\left(e^{-y^2/2}\right) dy.$$

Integrating by parts with

$$u = \sin(\beta y), \qquad dv = \frac{d}{dy}\left(e^{-y^2/2}\right) dy,$$

so that

$$du = \beta \cos(\beta y) dy, \qquad v = e^{-y^2/2},$$

we obtain

$$I(\beta) = -\frac{1}{\sqrt{2\pi}} \left[\sin(\beta y) e^{-y^2/2}\right]_{-\infty}^{\infty} + \frac{\beta}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} \cos(\beta y) dy.$$

Since $e^{-y^2/2} \to 0$ as $y \to \pm\infty$, the boundary term vanishes. Hence, we have

$$I(\beta) = \frac{\beta}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} \cos(\beta y) dy.$$

Using the classical Gaussian oscillatory integral

$$\int_{-\infty}^{\infty} e^{-py^2} \cos(qy) dy = \sqrt{\frac{\pi}{p}} e^{-q^2/(4p)}, \qquad p > 0, q \in \mathbb{R},$$

(see, e.g., [6, Formula 3.896.4]) and taking $p = 1/2$ and $q = \beta$, we obtain

$$\int_{-\infty}^{\infty} e^{-y^2/2} \cos(\beta y) dy = \sqrt{2\pi} e^{-\beta^2/2}.$$

Therefore, we get

$$I(\beta) = \frac{\beta}{\sqrt{2\pi}} \sqrt{2\pi} e^{-\beta^2/2} = \beta e^{-\beta^2/2}.$$

Finally, we obtain

$$\mathbb{E}(Y \mid X = x) = \frac{\beta e^{-\beta^2/2}}{2 + \cos(\alpha x)}.$$

- Similarly, the conditional expectation of X given $Y = y$ is given by

$$\mathbb{E}(X \mid Y = y) = \int_{-\infty}^{\infty} x l(x \mid y) dx.$$

Recall that the conditional density of X given $Y = y$ is

$$k(x \mid y) = \phi(x) \frac{2 + \cos(\alpha x) + \sin(\beta y)}{2 + e^{-\alpha^2/2} + \sin(\beta y)}, \qquad \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Substituting the expression of $l(x \mid y)$, we obtain

$$\mathbb{E}(X \mid Y = y) = \frac{1}{2 + e^{-\alpha^2/2} + \sin(\beta y)} \int_{-\infty}^{\infty} x\phi(x)(2 + \cos(\alpha x) + \sin(\beta y)) dx.$$

Expanding the integral, we get

$$\mathbb{E}(X \mid Y = y) = \frac{1}{2 + e^{-\alpha^2/2} + \sin(\beta y)}\Big[2\int_{-\infty}^{\infty} x\phi(x)dx + \int_{-\infty}^{\infty} x\phi(x)\cos(\alpha x)dx + \sin(\beta y)\int_{-\infty}^{\infty} x\phi(x)dx\Big].$$

Since the standard Gaussian distribution is centered, we have

$$\int_{-\infty}^{\infty} x\phi(x)dx = 0.$$

Moreover, the function $x\phi(x)\cos(\alpha x)$ is odd, because x is odd while both $\phi(x)$ and $\cos(\alpha x)$ are even. Hence, we have

$$\int_{-\infty}^{\infty} x\phi(x)\cos(\alpha x)dx = 0.$$

Therefore, we obtain

$$\mathbb{E}(X \mid Y = y) = 0.$$

This completes the proof of the proposition. □

3.3. *First-order moments*

The proposition below establishes the first-order moments of the distribution.

Proposition 3.4. *Let (X, Y) be a random vector following the BTG distribution. Then we have*

$$\mathbb{E}(X) = 0, \qquad \mathbb{E}(Y) = \frac{be^{-b^2/2}}{2 + e^{-a^2/2}}, \qquad \mathbb{E}(XY) = 0.$$

Proof:

We recall that (X, Y) has the following probability density function:

$$f(x, y) = K^{-1}e^{-x^2/2-y^2/2}(2 + \cos(\alpha x) + \sin(\beta y)).$$

We compute each moment separately.

- By definition, the expectation of X is given by

$$\mathbb{E}(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)dxdy.$$

Using the Fubini integral theorem and separating the terms, we obtain

$$\mathbb{E}(X) = K^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\,(2 + \cos(ax) + \sin(by))\,e^{-x^2/2-y^2/2}dxdy.$$

Each term contains

$$\int_{-\infty}^{\infty} xe^{-x^2/2}dx = 0$$

or

$$\int_{-\infty}^{\infty} x\cos(\alpha x)e^{-x^2/2}dx = 0$$

since the integrand is an odd function. Therefore,

$$\mathbb{E}(X) = 0.$$

- Similarly, the expectation of X is given by

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)dxdy.$$

Again, we separate the terms. The parts involving

$$ye^{-y^2/2}$$

are odd and integrate to zero. The only non-vanishing contribution comes from

$$K\int_{-\infty}^{\infty} y\sin(by)e^{-y^2/2}dy\int_{-\infty}^{\infty} e^{-x^2/2}dx.$$

Using the classical Gaussian integral identity

$$\int_{-\infty}^{\infty} y\sin(by)e^{-y^2/2}dy = b\sqrt{2\pi}e^{-b^2/2},$$

and

$$\int_{-\infty}^{\infty} e^{-x^2/2}dx = \sqrt{2\pi},$$

we obtain

$$\mathbb{E}(Y) = K^{-1}(2\pi)be^{-b^2/2}.$$

Substituting the value of K gives

$$\mathbb{E}(Y) = \frac{be^{-b^2/2}}{2 + e^{-a^2/2}}.$$

- By definition, the expectation of XY is given by

$$\mathbb{E}(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)dxdy.$$

After expanding, each term contains at least one factor of the form

$$\int_{-\infty}^{\infty} xe^{-x^2/2}dx$$

or

$$\int_{-\infty}^{\infty} ye^{-y^2/2}dy,$$

which vanish by symmetry.
Hence, we have

$$\mathbb{E}(XY) = 0.$$

This completes the proof of the proposition. □

4. Estimation

4.1. Maximum likelihood estimation

Let n be a positive integer and let $(x_1, y_1), \ldots, (x_n, y_n)$ be an independent and identically distributed sample from a random vector (X, Y) following the BTG distribution with parameters α and β. The associated likelihood function is

$$L(\alpha,\beta) = \prod_{i=1}^{n} f(x_i, y_i; \alpha,\beta),$$

where

$$f(x, y; \alpha,\beta) = K(\alpha)^{-1} e^{-x^2/2-y^2/2}\Big(2 + \cos(\alpha x) + \sin(\beta y)\Big), \qquad (x, y) \in \mathbb{R}^2,$$

and

$$K(\alpha) = 2\pi(2 + e^{-\alpha^2/2}).$$

Maximizing $L(\alpha,\beta)$ is equivalent to maximizing the log-likelihood function

$$\ell(\alpha,\beta) = \log L(\alpha,\beta) = \sum_{i=1}^{n} \log f(x_i, y_i; \alpha,\beta).$$

By substituting the expression of f, we obtain

$$\begin{aligned}\ell(\alpha,\beta) &= \sum_{i=1}^{n}\left[-\log K(\alpha) - \frac{x_i^2 + y_i^2}{2} + \log\Big(2 + \cos(\alpha x_i) + \sin(\beta y_i)\Big)\right]\\ &= -n\log K(\alpha) - \frac{1}{2}\sum_{i=1}^{n}(x_i^2 + y_i^2) + \sum_{i=1}^{n}\log\Big(2 + \cos(\alpha x_i) + \sin(\beta y_i)\Big).\end{aligned}$$

The term $(1/2)\sum_{i=1}^{n}(x_i^2+y_i^2)$ does not depend on (α,β) and can be dropped for optimization purposes. Hence, the maximum likelihood estimator (MLE) of (α,β) satisfies

$$(\hat{\alpha},\hat{\beta}) = \arg\max_{(\alpha,\beta)\in\mathbb{R}^2}\left\{-n\log K(\alpha) + \sum_{i=1}^{n}\log\Big(2 + \cos(\alpha x_i) + \sin(\beta y_i)\Big)\right\}.$$

Score equations. The first-order conditions are obtained by differentiating $\ell(\alpha,\beta)$. With respect to α, we get

$$\frac{\partial\ell}{\partial\alpha}(\alpha,\beta) = \sum_{i=1}^{n}\frac{-x_i\sin(\alpha x_i)}{2 + \cos(\alpha x_i) + \sin(\beta y_i)} - n\frac{K'(\alpha)}{K(\alpha)},$$

where

$$K'(\alpha) = 2\pi \times \frac{d}{d\alpha}\Big(e^{-\alpha^2/2}\Big) = -2\pi\alpha e^{-\alpha^2/2}, \qquad \frac{K'(\alpha)}{K(\alpha)} = \frac{-\alpha e^{-\alpha^2/2}}{2 + e^{-\alpha^2/2}}.$$

With respect to β, we obtain

$$\frac{\partial\ell}{\partial\beta}(\alpha,\beta) = \sum_{i=1}^{n}\frac{y_i\cos(\beta y_i)}{2 + \cos(\alpha x_i) + \sin(\beta y_i)}.$$

The MLE $(\hat{\alpha}, \hat{\beta})$ is therefore a solution (when it exists) of the nonlinear system

$$\frac{\partial \ell}{\partial \alpha}(\alpha, \beta) = 0, \qquad \frac{\partial \ell}{\partial \beta}(\alpha, \beta) = 0.$$

These likelihood equations do not admit a closed-form solution in general. The presence of the trigonometric terms $\cos(\alpha x_i)$ and $\sin(\beta y_i)$ inside logarithms and denominators leads to highly nonlinear score functions. Moreover, due to the periodic nature of sine and cosine, the log-likelihood function is typically non-concave and may exhibit multiple local maxima. As a consequence, $(\hat{\alpha}, \hat{\beta})$ cannot be obtained analytically and must be computed numerically using standard optimization procedures (see, e.g., [2]).

In practice, we estimate (α, β) by maximizing $\ell(\alpha, \beta)$ using a numerical optimization procedure. Because the objective function may be multimodal, gradient-based methods can be sensitive to initialization. We therefore rely on a heuristic/global optimization strategy (e.g., particle swarm optimization, simulated annealing, or multi-start local optimization) to obtain a reliable approximation of the global maximizer. The resulting maximizer is taken as the maximum likelihood estimate of (α, β).

4.2. Simulation procedure

Since the BTG distribution does not admit a closed-form sampling method, we rely on the accept-reject algorithm with the standard bivariate independante Gaussian distribution as proposoal. The procedure can be summarized in the following four steps:

Step 1. Recall the probability density function of the BTG distribution with parameters α and β is given by

$$f(x, y) = K^{-1} e^{-x^2/2 - y^2/2} \Big(2 + \cos(\alpha x) + \sin(\beta y)\Big), \qquad (x, y) \in \mathbb{R}^2,$$

where

$$K = 2\pi \left(2 + e^{-\alpha^2/2}\right).$$

Step 2. We choose the standard bivariate independent probability density function defined by:

$$\phi(x, y) = \frac{1}{2\pi} e^{-x^2/2 - y^2/2}, (x, y) \in \mathbb{R}^2$$

as the proposal distribution. Notice that, for any $(x, y) \in \mathbb{R}^2$, we have

$$K f(x, y) = 2\pi \phi(x, y) \Big(2 + \cos(\alpha x) + \sin(\beta y)\Big) \leq 8\pi \phi(x, y),$$

that is

$$f(x, y) \leq M \phi(x, y),$$

where

$$M = K^{-1} 8\pi.$$

Step 3. To generate a random value (x, y) from the BTG distribution, the process is as follows:

1. Generate a candidate (x^*, y^*) from the standard bivariate independent Gaussian distribution.
2. Generate a value u from the uniform distribution over $(0, 1)$.
3. Accept (x^*, y^*) if

$$u < \frac{f(x^*, y^*)}{M \phi(x^*, y^*)} = \frac{2 + \cos(\alpha x^*) + \sin(\beta y^*)}{4}$$

Otherwise, reject and repeat.

5. Application

Classical univariate goodness-of-fit procedures (e.g., the Kolmogorov-Smirnov test) are not directly applicable to assessing whether an empirical dataset is compatible with the proposed BTG model. Standard extensions to the multivariate setting are particularly challenging and usually require access to the cumulative distribution function in a tractable form. We therefore adopt a two-sample goodness-of-fit strategy based on the energy distance [12], which is well-suited to multivariate data.

5.1. Energy statistic (energy distance)

The energy distance provides a global measure of discrepancy between two distributions. In our setting, we use a two-sample procedure to compare:

1. the observed sample $\{(x_i, y_i)\}_{i=1}^{n}$,
2. a synthetic sample of the same size n generated under the fitted BTG model using the accept–reject sampler and the parameter estimates $\hat{\alpha}$ and $\hat{\beta}$ obtained by PSO.

The `energy` package in R [11] (function `eqdist.e`) yields the observed test statistic T_{obs}, quantifying the distributional discrepancy between the two samples.

Adjustment via parametric bootstrap. Because the synthetic sample is generated using the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$, the null hypothesis is composite and the usual asymptotic calibration of the test statistic is no longer valid. To account for parameter uncertainty, we implement a parametric bootstrap [5] that replicates the full pipeline (estimation + testing):

- **Observed statistic:** Compute T_{obs} by comparing the data to a synthetic sample generated under $(\hat{\alpha}, \hat{\beta})$.
- **Bootstrap iterations** ($B = 199$)**:** For each $b = 1, \ldots, B$, generate a bootstrap sample from the BTG model with parameter $(\hat{\alpha}, \hat{\beta})$, re-estimate the parameters $(\hat{\alpha}, \hat{\beta})_b$ on this bootstrap sample via PSO, then compute a bootstrap statistic $T^{(b)}_{\text{boot}}$ using the same two-sample energy procedure.
- **Bootstrap p-value:** The p-value is estimated as

$$p = \frac{1 + \sum_{b=1}^{B} \mathbf{1}\left\{T^{(b)}_{\text{boot}} \geq T_{\text{obs}}\right\}}{B + 1}.$$

Hypothesis testing framework. Formally, the goodness-of-fit procedure is conducted under the composite null hypothesis

$$\mathcal{H}_0 : \text{ there is } (\alpha, \beta) \text{ such that } (X, Y) \text{ follows the BTG}(\alpha, \beta) \text{ distribution},$$

against the general alternative

$$\mathcal{H}_1 : (X, Y) \text{ does not follow any BTG distribution.}$$

Since the parameters α and β are unknown, they are first estimated from the data, yielding $(\hat{\alpha}, \hat{\beta})$. The parametric bootstrap then approximates the null distribution of the test statistic under $\mathcal{H}_0$ while accounting for parameter estimation. In each bootstrap replication, both the data-generating step and the re-estimation step are reproduced, thereby mimicking the full inferential procedure.

Interpretation of the p-value. The resulting bootstrap p-value quantifies the probability, under the fitted BTG model, of observing a discrepancy (as measured by the energy statistic) at least as large as the one computed from the real data.

A large p-value indicates that the observed structural discrepancy is compatible with sampling variability under the BTG model and therefore provides no evidence against $\mathcal{H}_0$. Conversely, a small p-value suggests that the empirical distribution exhibits features that are unlikely to arise from any BTG distribution, thus leading to rejection of $\mathcal{H}_0$.

Implementation in R. This entire logic, coupling PSO optimization for re-fitting, accept-reject simulation, and energy distance, was coded as follows:

Listing 1. Goodness-of-fit test by energy distance and parametric bootstrap

```
gof_energy_bootstrap_btg <- function(
  X, lim_inf = c(0, 0), lim_sup = c(1, 0.5),
  B = 199, M = 1,
  # PSO hyperparameters
  S = 350, e = 1e-4, N = 500, prop = 0.2,) {
  stopifnot(is.matrix(X), nrow(X) == 2, length(lim_inf) == 2,
  length(lim_sup) == 2)
  stopifnot(B >= 1, M >= 1, S >= 2, N >= 10)

  n <- ncol(X)

  #------------------------------------------------------------
  # 1) Fit on observed data via PSO: theta_hat = (alpha_hat, beta_hat)
  #------------------------------------------------------------
  theta_hat <- estimate_btg_pso(
    X = X,
    lim_inf = lim_inf,
    lim_sup = lim_sup,
    S = S, e = e, N = N, prop = prop )
  alpha_hat <- theta_hat[1]
  beta_hat  <- theta_hat[2]

  #------------------------------------------------------------
  # 2) Observed statistic T_obs (optional averaging over M replicates)
  #------------------------------------------------------------
  T_obs <- mean(replicate(M, {
    Y0 <- rbtg(n, alpha = alpha_hat, beta = beta_hat)
    energy_stat_2sample(X, Y0)
  }))
```

```
#--------------------------------------------------------------
# 3) Parametric bootstrap with re-estimation (sequential)
#--------------------------------------------------------------
theta_boot <- matrix(NA_real_, nrow = B, ncol = 2)
colnames(theta_boot) <- c("alpha", "beta")
T_boot <- numeric(B)

for (b in seq_len(B)) {
  # (i) Bootstrap dataset under the fitted model
  Xb <- rbtg(n, alpha = alpha_hat, beta = beta_hat)

  # (ii) Re-estimate parameters on Xb via PSO
  theta_b <- estimate_btg_pso(
    X = Xb,
    lim_inf = lim_inf,
    lim_sup = lim_sup,
    S = S, e = e, N = N, prop = prop,
    omega = omega, phi_p = phi_p, phi_g = phi_g
  )
  theta_boot[b, ] <- theta_b

  # (iii) Bootstrap statistic (same procedure as for T_obs)
  alpha_b <- theta_b[1]
  beta_b  <- theta_b[2]

  T_boot[b] <- mean(replicate(M, {
    Yb <- rbtg(n, alpha = alpha_b, beta = beta_b)
    energy_stat_2sample(Xb, Yb)
  }))
}

# Bootstrap p-value with +1 correction
p_value <- (1 + sum(T_boot >= T_obs)) / (B + 1)

list(
  theta_hat  = theta_hat,   # c(alpha_hat, beta_hat)
  T_obs      = T_obs,
  p_value    = p_value,)
}
```

5.2. Model selection and information criteria

In order to validate the practical utility of the BTG distribution, it is necessary to prove that it models real data with greater precision than existing classical distributions. We therefore set up a model selection procedure comparing our distribution against 9 other standard bivariate distributions.

Directly comparing the maximum Log-Likelihood ($\log \hat{L}$) of these models would be statistically biased. Indeed, a model with many parameters (such as the bivariate Student's t-distribution, which has 6) will mechanically always fit the data better than a simple, parsimonious model (like our BTG distribution, which relies on only two parameters, α and β).

To overcome this overfitting problem, we use information criteria that impose a trade-off between the goodness-of-fit and the complexity of the model:

1. The Akaike information criterion (AIC). The AIC estimates the relative information loss when a model is used to represent the real data-generating process. It is defined by

$$\text{AIC} = 2k - 2\log(\hat{L}),$$

where k represents the number of estimated parameters of the model and $\hat{L}$ the maximum value of the likelihood function. The best model is the one that minimizes this value. The $2k$ term acts as a strict penalty against adding unnecessary parameters.

2. The Bayesian information criterion (BIC). The BIC relies on a similar approach but penalizes complexity much more heavily when the sample size n is large. It is defined by

$$\text{BIC} = k\log(n) - 2\log(\hat{L}).$$

In our study on Turbine dataset (where n is often greater than several hundreds), the BIC is an extremely strict referee that will favor parsimony.

3. AIC differences (Δ_i) and Akaike weights (w_i). Since the absolute value of the AIC has no intrinsic meaning, we evaluate the models by calculating the AIC difference between each model i and the best model of the tournament (the one with the lowest score, AIC_{min}). It is defined by

$$\Delta_i = \text{AIC}_i - \text{AIC}_{min}.$$

As a simple rule, a model with $\Delta_i > 10$ has no statistical support and can be discarded.

To make these results more intuitive, we transform these differences into Akaike weights [1]. They normalize the scores to provide a probability, ranging from 0 to 1 (or 0% to 100%), that model i is the best among the set of tested models. It is defined by

$$w_i = \frac{e^{-\Delta_i/2}}{\sum_{j=1}^{M} e^{-\Delta_j/2}}.$$

where M is the total number of competing models (here $M = 10$). This probability w_i will serve as our primary indicator to identify the most suitable model for each studied pair of variables.

6. Results

6.1. Study on the ”Turbine” dataset

To assess the empirical relevance of our model, we applied it to the publicly available Turbine dataset [7]. This dataset contains five quantitative variables describing the operational conditions and power output of a turbine-based energy production system:

- **AT**: Ambient Temperature,
- **V**: Exhaust Vacuum,
- **AP**: Ambient Pressure,
- **RH**: Relative Humidity,
- **PE**: Power Output.

These variables capture both environmental conditions (AT, AP, RH), mechanical system behavior (V), and the resulting energy production (PE). Prior to statistical inference, all variables were centered and scaled in order to ensure numerical stability and comparability across dimensions.

6.2. Analysis approach and variable filtering

Given the multivariate structure of the dataset, we implemented a systematic ”funnel” exploratory strategy to identify the bivariate configurations for which the BTG model provides a statistically acceptable fit.

1. **Combinatorial exploration.** The dataset contains five quantitative variables. We generated all possible pairwise combinations among them, resulting in $\binom{5}{2} = 10$ distinct bivariate configurations.
2. **Goodness-of-Fit filtering.** Each pair was subjected to the parametric bootstrap Energy distance test. We retained only the pairs for which the null hypothesis of compatibility with a BTG distribution was not rejected at the 5% significance level (p-value > 0.05). In addition, we required at least one of the estimated parameters $\hat{\alpha}$ or $\hat{\beta}$ to be meaningfully different from zero. This condition excludes trivial cases where the model would collapse to a standard bivariate Gaussian structure.
3. **Validation via information criteria.** The retained pair(s) were subsequently evaluated using AIC, BIC, and Akaike weights, in order to compare the BTG distribution with nine classical reference bivariate models.

6.3. Results: validation and model selection

1. Validation via the Energy test. The Energy distance test (via parametric bootstrap) allowed us to identify the pairs for which the empirical distribution does not reject our trigonometric model. Table 1 illustrates these results for the retained pair.

Variable Crossing (Turbine)	Estimated parameters $(\hat{\alpha}, \hat{\beta})$	T_{obs}	P-value
AP and RH	(0.097, 0.028)	3.21	0.109

Table 1. Goodness-of-fit test results (Energy) on a valid pair from the Turbine dataset.

2. Competition and model selection. For this candidate pair, we then put our distribution in direct competition with 9 other classical bivariate distributions. The evaluation relies on the Akaike information criterion (i.e., AIC $= 2k - 2\log\hat{L}$) and the Bayesian information criterion (i.e., BIC $= k\log(n) - 2\log\hat{L}$), which penalize model complexity. For our BTG distribution, the number of estimated parameters is now $k = 2$.

Table 2 details the ranking for the pair Ambient Pressure (AP) and Relative Humidity (RH).

Model	k	LogLik	AIC	BIC	ΔAIC	Akaike weight (%)
BTG	2	-26980.9	53965.8	53980.1	0.0	100
Bivariate Gaussian ($\mu = 0$)	3	-27104.1	54214.3	54235.8	248.5	0
Bivariate Gaussian	5	-27104.1	54218.3	54254.1	252.5	0
Bivariate Student t	6	-27119.7	54251.4	54294.4	285.6	0
Standard Gaussian (Baseline)	0	-27151.8	54303.6	54303.6	337.8	0
Independent Gaussians	4	-27151.8	54311.6	54340.3	345.8	0
Independent Students	6	-27154.3	54320.6	54363.6	354.8	0
Independent Logistics	4	-27427.9	54863.8	54892.4	897.9	0
Independent Laplaces	4	-28299.4	56606.8	56635.4	2641.0	0
Independent Cauchys	4	-30917.1	61842.1	61870.8	7876.3	0

Table 2. Model comparison based on log-likelihood, AIC, BIC and Akaike weights.

Physical interpretation. The success of our mathematical model on this specific pair can be explained by the meteorological and thermodynamic constraints of the turbine environment. Ambient pressure (AP) and relative humidity (RH) do not follow a simple linear correlation. Instead, their relationship is bounded by saturation thresholds (dew point) and discrete operational ranges induced by the turbine's intake systems. The independent trigonometric components (α and β) of the BTG distribution allow it to perfectly capture these nonlinear saturation plateaus and exclusion zones, whereas classical Gaussian models fail by assuming infinite, continuous, and elliptical dispersion.

7. Conclusion

In this article, we introduced a new trigonometric extension of the Gaussian distribution. The model relies on two independent parameters, α and β, and provides a mathematically rigorous framework whose inference can be fully carried out computationally. The statistical tools implemented in R, PSO optimization, parametric bootstrap based on the energy distance, and model selection using Akaike weights, enabled an effective confrontation of the theoretical model with empirical data.

The application to the Turbine dataset demonstrates the practical relevance of the approach. Its oscillatory components provide a unique ability to capture features often observed in physical systems, such as non-linear thresholds, saturation plateaus, and complex structural dependencies, where standard Gaussian models tend to be inadequate.

Conflict of interest

The authors declare that there are no conflicts of interest regarding this work.

Acknowledgments

The authors gratefully acknowledge the constructive comments and suggestions provided by the anonymous reviewers, which helped improve the quality of this article.

References

1.Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer.

2.Casella, G., Robert, C. P., and Wells, M. T. (2004). Generalized accept-reject sampling schemes. *Lecture Notes-Monograph Series*.

3.Chesneau, C. (2023). On two-dimensional functions with an integral equal to zero over a rectangle: Application to a modified gaussian distribution. *Asian Journal of Mathematics and Applications*, 5:1–23.

4.Chesneau, C. (2025). Theory on a new bivariate trigonometric gaussian distribution. *Innovation in Statistics and Probability*, 1(2):1–17.

5.Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. CRC Press.

6.Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products*. Academic Press, 7 edition.

7.Ishank (2024). Wind turbines data. `https://www.kaggle.com/datasets/ishank2005/wind-turbines-data-csv`. Kaggle dataset.

8.Jwo, D.-J., Cho, T.-S., and Biswal, A. (2023). Geometric insights into the multivariate gaussian distribution and its entropy and mutual information. *Entropy*, 25(8):1177.

9.Mahmoudi, E. and Mahmoodian, H. (2017). A new bivariate distribution obtained by compounding the bivariate normal and geometric distributions. *Journal of Statistical Theory and Applications*, 16:198–208.

10.Mathai, A. M., Provost, S. B., and Haubold, H. J. (2022). The multivariate gaussian and related distributions. In *Multivariate Statistical Analysis in the Real and Complex Domains*, chapter 3. Springer, Cham.

11.R Core Team (2023). R: A language and environment for statistical computing.

12.Székely, G. J. and Rizzo, M. L. (2013). Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference*, 143(8):1249–1272.