

DISPLAYING ASYMMETRIC UNCERTAINTIES IN CORRELATION PLOTS

Frédéric GALLIANO

AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France

Created on March 6, 2018
 Updated on October 4, 2019

Contents

1 EQUI-PROBABLE CONTOURS OF SYMMETRIC BIVARIATE DISTRIBUTIONS	1
1.1 Ellipse Equation	1
1.2 Rotation Matrix Formalism	2
2 USING THE BIVARIATE SPLIT-NORMAL DISTRIBUTION	3
2.1 The Univariate Split-Normal Distribution	3
2.2 The Bivariate Case with Correlations	3
2.3 Different Ways of Displaying Uncertainties	5
A USEFUL FORMULAE	7
A.1 Partial Integration of Univariate Normal Distributions	7
A.2 Moments of a Centered, Uncorrelated BSN	7
References	7

1 EQUI-PROBABLE CONTOURS OF SYMMETRIC BIVARIATE DISTRIBUTIONS

1.1 Ellipse Equation

Let's consider the joint distribution of two parameters x and y , with means \bar{x} and \bar{y} and standard deviations σ_x and σ_y . The 1σ ellipse is described by:

$$\vec{X}^T \overleftrightarrow{V}^{-1} \vec{X} = 1, \quad (1)$$

where $\vec{X}^T = (x - \bar{x}, y - \bar{y})$ and the covariance matrix of the parameters is:

$$\overleftrightarrow{V} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}. \quad (2)$$

Since, we have:

$$\overleftrightarrow{V}^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x\sigma_y} \\ -\frac{\rho}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}, \quad (3)$$

the equation of the 1σ ellipse is then:

$$\frac{1}{1 - \rho^2} \times \left(\frac{(x - \bar{x})^2}{\sigma_x^2} + \frac{(y - \bar{y})^2}{\sigma_y^2} - 2\frac{\rho(x - \bar{x})(y - \bar{y})}{\sigma_x\sigma_y} \right) = 1, \quad (4)$$

or:

$$y = \sigma_y \times \left(\rho \frac{x - \bar{x}}{\sigma_x} \pm \sqrt{(1 - \rho^2) \left(1 - \frac{(x - \bar{x})^2}{\sigma_x^2} \right)} \right) + \bar{y}. \quad (5)$$

This is valid for typical symmetric unimodal distributions, such as gaussian or Student's *t*. The difference is that the 1- σ ellipse will correspond to a lower confidence in the Student's *t* case than in the normal case (67 %).

1.2 Rotation Matrix Formalism

Another way to look at this problem is to consider that there exists a rotated vector $\vec{X}' = (x', y')$, where x' and y' are independent (no correlation term). If we note θ the rotation angle, the rotation matrix is:

$$\vec{R} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad (6)$$

with the particular property $\vec{R}^T = \vec{R}^{-1}$. Working with centered variables ($\bar{x} = 0, \bar{y} = 0$), for simplicity, the transformation is expressed: $\vec{X}' = \vec{R} \vec{X}$. In the rotated reference, the covariance matrix, $\vec{\Lambda}$, is diagonal, with standard deviations λ_1 and λ_2 , as well as its inverse:

$$\vec{\Lambda}^{-1} = \begin{pmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_2^2 \end{pmatrix}. \quad (7)$$

The quadratic form can be expressed in both references as:

$$\vec{X}'^T \vec{\Lambda}^{-1} \vec{X}' = (\vec{R} \vec{X})^T \vec{\Lambda}^{-1} (\vec{R} \vec{X}) = \vec{X}^T \vec{\Lambda}^{-1} \vec{X}. \quad (8)$$

The two sides are:

$$\vec{X}'^T \vec{\Lambda}^{-1} \vec{X}' = \left(\frac{\cos^2 \theta}{\lambda_1^2} + \frac{\sin^2 \theta}{\lambda_2^2} \right) x^2 + \left(\frac{\sin^2 \theta}{\lambda_1^2} + \frac{\cos^2 \theta}{\lambda_2^2} \right) y^2 - 2 \cos \theta \sin \theta \left(\frac{1}{\lambda_1^2} - \frac{1}{\lambda_2^2} \right) xy \quad (9)$$

$$\vec{X}^T \vec{\Lambda}^{-1} \vec{X} = \frac{1}{1 - \rho^2} \times \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - 2 \frac{\rho xy}{\sigma_x \sigma_y} \right). \quad (10)$$

Since these equations have to be true $\forall x, y$, we have the following system that we can solve to derive the unknown θ , λ_1 and λ_2 as a function of the parameters ρ , σ_x and σ_y :

$$\frac{\cos^2 \theta}{\lambda_1^2} + \frac{\sin^2 \theta}{\lambda_2^2} = \frac{1}{1 - \rho^2} \frac{1}{\sigma_x^2} \quad (11)$$

$$\frac{\sin^2 \theta}{\lambda_1^2} + \frac{\cos^2 \theta}{\lambda_2^2} = \frac{1}{1 - \rho^2} \frac{1}{\sigma_y^2} \quad (12)$$

$$\cos \theta \sin \theta \left(\frac{1}{\lambda_1^2} - \frac{1}{\lambda_2^2} \right) = \frac{\rho}{1 - \rho^2} \frac{1}{\sigma_x \sigma_y}. \quad (13)$$

This system can be solved, posing $\alpha = \cos^2 \theta$ and:

$$\beta = \left(\frac{\rho \sigma_x \sigma_y}{\sigma_x^2 - \sigma_y^2} \right)^2, \quad (14)$$

we get:

$$\alpha = \begin{cases} \frac{1}{2} \left(1 + \frac{1}{\sqrt{4\beta + 1}} \right) & \text{if } \sigma_x > \sigma_y \\ \frac{1}{2} \left(1 - \frac{1}{\sqrt{4\beta + 1}} \right) & \text{if } \sigma_x < \sigma_y \\ 1 & \text{if } \sigma_x = \sigma_y. \end{cases} \quad (15)$$

The rotation angle is then simply: $\theta = \text{sgn}\rho \times \arccos \alpha$. The standard deviations in the new reference are:

$$\frac{1}{\lambda_1^2} = \frac{1}{1 - \rho^2} \frac{1}{1 - 2\alpha} \left(-\frac{\alpha}{\sigma_x^2} + \frac{1 - \alpha}{\sigma_y^2} \right) \quad (16)$$

$$\frac{1}{\lambda_2^2} = \frac{1}{1 - \rho^2} \frac{1}{1 - 2\alpha} \left(\frac{1 - \alpha}{\sigma_x^2} - \frac{\alpha}{\sigma_y^2} \right). \quad (17)$$

2 USING THE BIVARIATE SPLIT-NORMAL DISTRIBUTION

2.1 The Univariate Split-Normal Distribution

The split-normal distribution (Villani & Larsson, 2006) is a good choice when there is a non negligible skewness in the noise. It is defined as:

$$p(x) = \begin{cases} A \times \exp \left[-\frac{1}{2} \left(\frac{x - x_0}{\lambda} \right)^2 \right] & \text{if } x \leq x_0 \\ A \times \exp \left[-\frac{1}{2} \left(\frac{x - x_0}{\lambda \tau} \right)^2 \right] & \text{if } x > x_0, \end{cases} \quad (18)$$

where:

$$A = \sqrt{\frac{2}{\pi}} \frac{1}{\lambda(1 + \tau)}, \quad (19)$$

and x_0 is a position parameter, λ a scale parameter, and τ a shape parameter. These parameters are linked to the mean μ , standard deviation σ and skewness γ_1 through the following set of equations:

$$b = \left(1 - \frac{2}{\pi} \right) (\tau - 1)^2 + \tau \quad (20)$$

$$\mu = x_0 + \sqrt{\frac{2}{\pi}} \lambda (\tau - 1) \quad (21)$$

$$\sigma = \sqrt{b} \lambda^2 \quad (22)$$

$$\gamma_1 = b^{-3/2} \sqrt{\frac{2}{\pi}} (\tau - 1) \times \left[\left(\frac{4}{\pi} - 1 \right) (\tau - 1)^2 + \tau \right]. \quad (23)$$

There is no simple inversion. We therefore solve τ numerically from Eq. (23), and then inverse Eqs. (21)-(22).

2.2 The Bivariate Case with Correlations

The bivariate split-normal distribution (BSN) is the generalization of Sect. 2.1 in 2D. It is paramaterized by its position vector, $\vec{X}_0 = (x_0, y_0)$, its scale vector, $\vec{\Lambda} = (\lambda_x, \lambda_y)$, its shape vector, $\vec{T} = (\tau_x, \tau_y)$, and its rotation angle, $-\pi/2 < \theta < \pi/2$. We use its $1-\sigma$ contour to introduce the concept of *Skewed Uncertainty Ellipse* (SUE; Fig. 1). Expressing the density in the centered, rotated, reference frame (Fig. 1):

$$\vec{X}' = \begin{pmatrix} x'(x, y) \\ y'(x, y) \end{pmatrix} = \vec{R}^T (\vec{X} - \vec{X}_0) = \begin{pmatrix} (x - x_0) \cos \theta + (y - y_0) \sin \theta \\ -(x - x_0) \sin \theta + (y - y_0) \cos \theta \end{pmatrix}, \quad (24)$$

the density can be expressed as:

$$p(x, y) = \frac{2}{\pi \lambda_x \lambda_y (1 + \tau_x)(1 + \tau_y)} \times \begin{cases} \exp \left[-\frac{1}{2} \left(\frac{x'(x, y)}{\lambda_x} \right)^2 - \frac{1}{2} \left(\frac{y'(x, y)}{\lambda_y} \right)^2 \right] & \text{if } x'(x, y) < 0 \text{ and } y'(x, y) < 0 \\ \exp \left[-\frac{1}{2} \left(\frac{x'(x, y)}{\lambda_x \tau_x} \right)^2 - \frac{1}{2} \left(\frac{y'(x, y)}{\lambda_y} \right)^2 \right] & \text{if } x'(x, y) \geq 0 \text{ and } y'(x, y) < 0 \\ \exp \left[-\frac{1}{2} \left(\frac{x'(x, y)}{\lambda_x} \right)^2 - \frac{1}{2} \left(\frac{y'(x, y)}{\lambda_y \tau_y} \right)^2 \right] & \text{if } x'(x, y) < 0 \text{ and } y'(x, y) \geq 0 \\ \exp \left[-\frac{1}{2} \left(\frac{x'(x, y)}{\lambda_x \tau_x} \right)^2 - \frac{1}{2} \left(\frac{y'(x, y)}{\lambda_y \tau_y} \right)^2 \right] & \text{if } x'(x, y) \geq 0 \text{ and } y'(x, y) \geq 0. \end{cases} \quad (25)$$

We can estimate the moments of \vec{X} , as a linear combination of moments of \vec{X}' , using the formulae of Appendix A.2 and noting that:

$$x = x' \cos \theta - y' \sin \theta + x_0 \quad (26)$$

$$y = x' \sin \theta + y' \cos \theta + y_0. \quad (27)$$

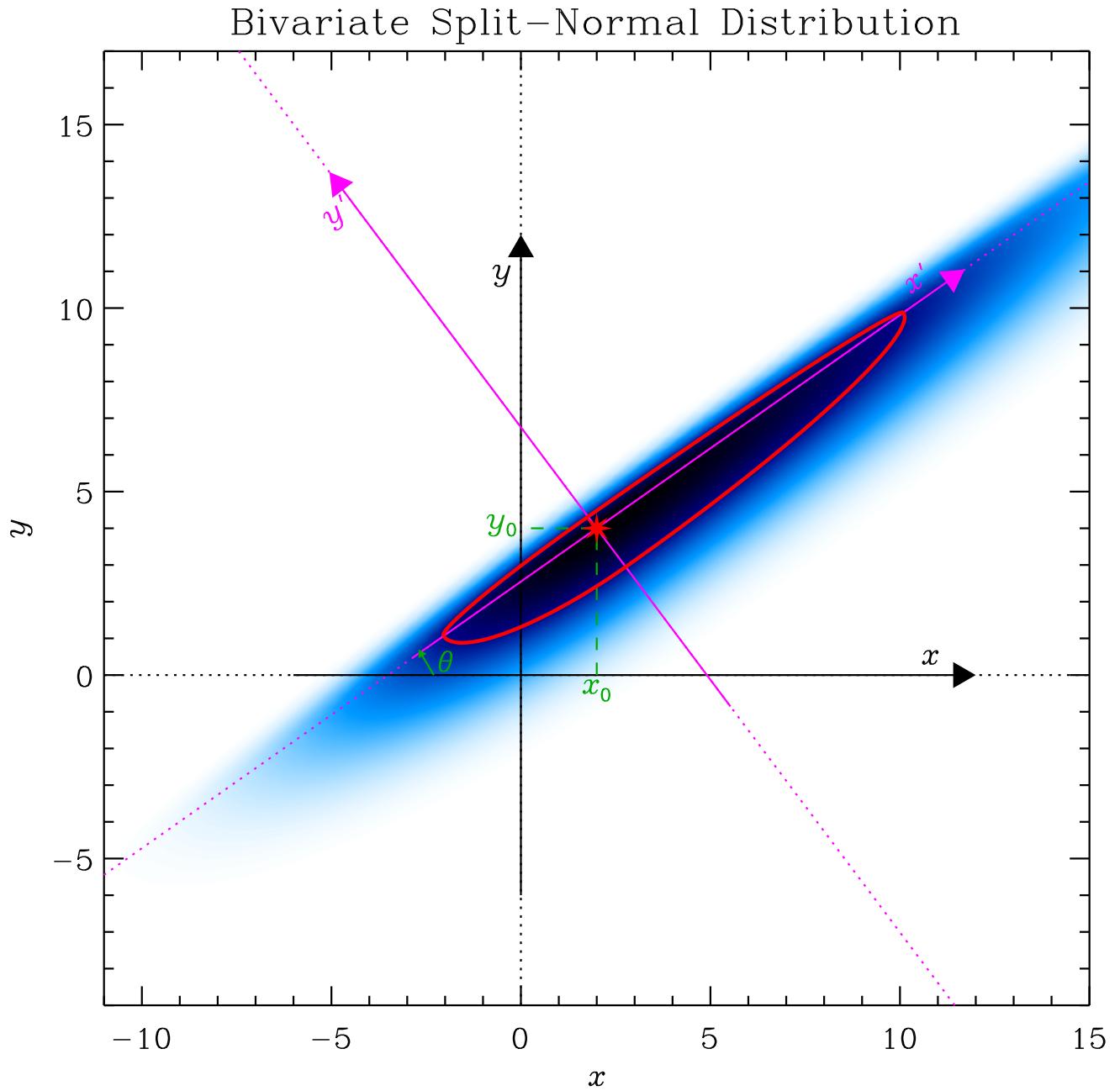


Figure 1: *Notation conventions.* This figure shows a BSN with $\vec{X}_0 = (2, 4)$, $\vec{\Lambda} = (5, 1.3)$, $\vec{T} = (2, 0.3)$ and $\theta = \pi/5$. The PDF is displayed with a blue density. The red line represents the asymmetric 1- σ *Skewed Uncertainty Ellipse* (SUE).

These moments are:

$$\langle x \rangle = \sqrt{\frac{2}{\pi}} [\lambda_x(\tau_x - 1) \cos \theta - \lambda_y(\tau_y - 1) \sin \theta] + x_0 \quad (28)$$

$$\langle y \rangle = \sqrt{\frac{2}{\pi}} [\lambda_x(\tau_x - 1) \sin \theta + \lambda_y(\tau_y - 1) \cos \theta] + y_0 \quad (29)$$

$$\sigma_x^2 \equiv \langle (x - \langle x \rangle)^2 \rangle = \lambda_x^2 B(\tau_x) \cos^2 \theta + \lambda_y^2 B(\tau_y) \sin^2 \theta \quad (30)$$

$$\sigma_y^2 \equiv \langle (y - \langle y \rangle)^2 \rangle = \lambda_x^2 B(\tau_x) \sin^2 \theta + \lambda_y^2 B(\tau_y) \cos^2 \theta \quad (31)$$

$$\rho \sigma_x \sigma_y \equiv \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle = (\lambda_x^2 B(\tau_x) - \lambda_y^2 B(\tau_y)) \cos \theta \sin \theta \quad (32)$$

$$\gamma_x \sigma_x^3 \equiv \langle (x - \langle x \rangle)^3 \rangle = \sqrt{\frac{2}{\pi}} (\lambda_x^3 C(\tau_x) \cos^3 \theta - \lambda_y^3 C(\tau_y) \sin^3 \theta) \quad (33)$$

$$\gamma_y \sigma_y^3 \equiv \langle (y - \langle y \rangle)^3 \rangle = \sqrt{\frac{2}{\pi}} (\lambda_x^3 C(\tau_x) \sin^3 \theta + \lambda_y^3 C(\tau_y) \cos^3 \theta), \quad (34)$$

with:

$$B(\tau) = \left(1 - \frac{2}{\pi}\right) (\tau - 1)^2 + \tau \quad (35)$$

$$C(\tau) = \left[\left(\frac{4}{\pi} - 1\right) \tau^2 + \left(3 - \frac{8}{\pi}\right) \tau + \frac{4}{\pi} - 1 \right] (\tau - 1) \quad (36)$$

We want to be able to derive the distribution's parameters ($x_0, y_0, \lambda_x, \lambda_y, \tau_x, \tau_y, \theta$) from the estimated moments ($\langle x \rangle, \langle y \rangle, \sigma_x, \sigma_y, \gamma_x, \gamma_y, \rho$). We therefore need to solve the system of Eqs. (28) and (34). To do that, we first solve θ :

$$\theta = \frac{1}{2} \arctan \left(\frac{2\rho \sigma_x \sigma_y}{\sigma_x^2 - \sigma_y^2} \right). \quad (37)$$

We then solve numerically for τ_x and τ_y , independently, from the two equations:

$$\frac{C(\tau_x)}{B(\tau_x)^{3/2}} = \sqrt{\frac{\pi}{2}} \frac{\gamma_x \sigma_x^3 \cos^3 \theta + \gamma_y \sigma_y^3 \sin^3 \theta}{\cos^6 \theta + \sin^6 \theta} \left(\frac{\cos^2 \theta - \sin^2 \theta}{\sigma_x^2 \cos^2 \theta - \sigma_y^2 \sin^2 \theta} \right)^{3/2} \quad (38)$$

$$\frac{C(\tau_y)}{B(\tau_y)^{3/2}} = \sqrt{\frac{\pi}{2}} \frac{\gamma_y \sigma_y^3 \cos^3 \theta - \gamma_x \sigma_x^3 \sin^3 \theta}{\cos^6 \theta + \sin^6 \theta} \left(\frac{\cos^2 \theta - \sin^2 \theta}{\sigma_y^2 \cos^2 \theta - \sigma_x^2 \sin^2 \theta} \right)^{3/2}. \quad (39)$$

We derive the remaining parameters, using the following equations:

$$\lambda_x = \sqrt{\frac{1}{B(\tau_x)} \frac{\sigma_x^2 \cos^2 \theta - \sigma_y^2 \sin^2 \theta}{\cos^2 \theta - \sin^2 \theta}} \quad (40)$$

$$\lambda_y = \sqrt{\frac{1}{B(\tau_y)} \frac{\sigma_y^2 \cos^2 \theta - \sigma_x^2 \sin^2 \theta}{\cos^2 \theta - \sin^2 \theta}} \quad (41)$$

$$x_0 = \langle x \rangle - \sqrt{\frac{2}{\pi}} [\lambda_x(\tau_x - 1) \cos \theta - \lambda_y(\tau_y - 1) \sin \theta] \quad (42)$$

$$y_0 = \langle y \rangle - \sqrt{\frac{2}{\pi}} [\lambda_x(\tau_x - 1) \sin \theta + \lambda_y(\tau_y - 1) \cos \theta] \quad (43)$$

2.3 Different Ways of Displaying Uncertainties

Fig. 2 compares the most used ways to display a measurement and its uncertainty from the estimated moments of its PDF. Obviously, the SUEs give the best representation of the uncertainties, provided that one can reasonably estimate the skewnesses.

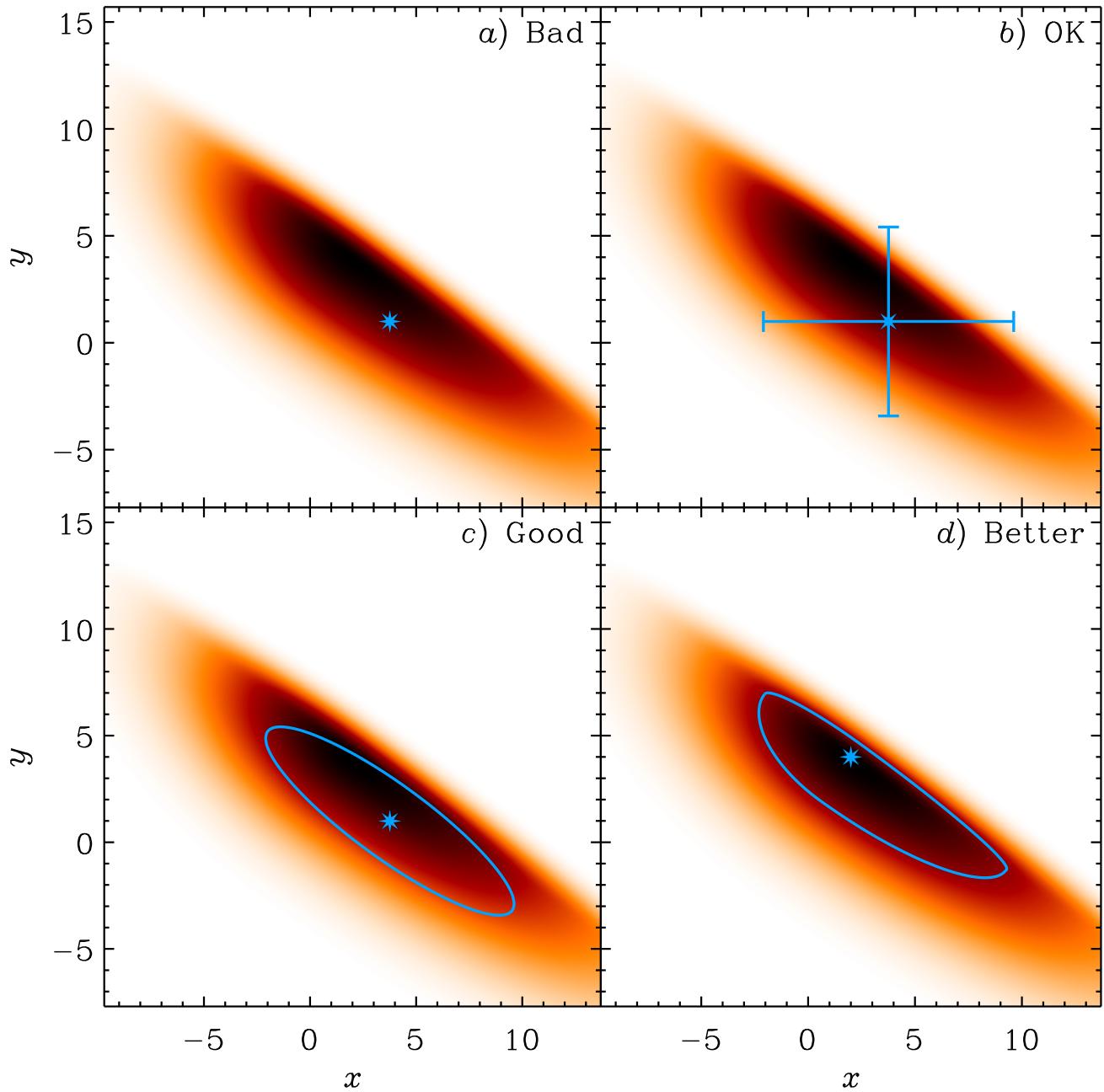


Figure 2: *Different ways of displaying uncertainties.* The PDF (red density) is identical in the four panels. Measurements are displayed in blue. Panel a shows $\langle x \rangle$ and $\langle y \rangle$. Panel b shows the typical $\pm\sigma$ error bars. Panel c shows standard ellipses. Panel d shows the *Skewed Uncertainty Ellipse* (SUE).

A USEFUL FORMULAE

A.1 Partial Integration of Univariate Normal Distributions

$$\int_{-\infty}^0 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \sqrt{\frac{\pi}{2}}\sigma \quad \text{and} \quad \int_0^\infty \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \sqrt{\frac{\pi}{2}}\sigma \quad (44)$$

$$\int_{-\infty}^0 x \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = -\sigma^2 \quad \text{and} \quad \int_0^\infty x \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \sigma^2 \quad (45)$$

$$\int_{-\infty}^0 x^2 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \sqrt{\frac{\pi}{2}}\sigma^3 \quad \text{and} \quad \int_0^\infty x^2 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \sqrt{\frac{\pi}{2}}\sigma^3 \quad (46)$$

$$\int_{-\infty}^0 x^3 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = -2\sigma^4 \quad \text{and} \quad \int_0^\infty x^3 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = 2\sigma^4. \quad (47)$$

A.2 Moments of a Centered, Uncorrelated BSN

If $p(x', y')$ is the BSN distribution of Eq. (25), with $x_0 = 0$, $y_0 = 0$ and $\theta = 0$, then:

$$\int_{-\infty}^\infty p(x', y') dx' dy' = 1 \quad (48)$$

$$\langle x' \rangle \equiv \int_{-\infty}^\infty x' \times p(x', y') dx' dy' = \sqrt{\frac{2}{\pi}}\lambda_x(\tau_x - 1) \quad (49)$$

$$\langle x'^2 \rangle \equiv \int_{-\infty}^\infty x'^2 \times p(x', y') dx' dy' = \lambda_x^2(\tau_x^2 - \tau_x + 1) \quad (50)$$

$$\langle x'^3 \rangle \equiv \int_{-\infty}^\infty x'^3 \times p(x', y') dx' dy' = 2\sqrt{\frac{2}{\pi}}\lambda_x^3(\tau_x^3 - \tau_x^2 + \tau_x - 1) \quad (51)$$

$$\langle x'y' \rangle \equiv \int_{-\infty}^\infty x'y' \times p(x', y') dx' dy' = \frac{2}{\pi}\lambda_x\lambda_y(\tau_x - 1)(\tau_y - 1) \quad (52)$$

$$\langle x'^2y' \rangle \equiv \int_{-\infty}^\infty x'^2y' \times p(x', y') dx' dy' = \sqrt{\frac{2}{\pi}}\lambda_x^2\lambda_y(\tau_x^2 - \tau_x + 1)(\tau_y - 1). \quad (53)$$

References

Villani, M. & Larsson, R. 2006, *Communications in Statistics-Theory and Methods*, 35, 1123