ISBN: 978-3-95450-219-6 ISSN: 2226-0366 doi: 10.18429/JACoW-LINAC2024-TUPB058

A COMPARISON OF RMS MOMENTS AND STATISTICAL DIVERGENCES AS WAYS TO QUANTIFY THE DIFFERENCE BETWEEN BEAM PHASE SPACE DISTRIBUTIONS

Y. Du*[1,2,3], Z. J. Wang[1,2,3], J. C. Wong[1,2,3], L. W. Liu[1,2,3], C. G. Su[1,2,3], M. Yi[3], L. Y. Gong[1,2], T. Zhang[1,2,3], H. Y. Zhou[1,2,3], B. H. Ma[1,2,3], Y. M. Chu[1,2,3], T. Y. Li[1,2,3], T. L. Wang[1,2,3]

[1]Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]Advanced Energy Science and Technology Guangdong Laboratory, Huizhou, China

Abstract

Accurately assessing the difference between two beam distributions in high-dimensional phase space is crucial for interpreting experimental or simulation results. In this paper, we compare the common method of RMS moments and mismatch factors, and the method of statistical divergences that give the total contribution of differences at all points. We first show that, in the case of commonly used initial distributions, there is a one-to-one correspondence between mismatch factors and statistical divergences. This enables us to show how the values of several popular divergences vary with the mismatch factors, independent of the orientation of the phase space ellipsoid. We utilize these results to propose evaluation standards for these popular divergences, which will help interpret their values in the context of beam phase space distributions.

INTRODUCTION

In the research fields of beam transport line design, beam diagnostics, phase space tomography and so on, accurately assessing the difference between two beam distributions in high-dimensional phase space is crucial for interpreting experimental or simulation results [1–3]. A common approach to quantify these differences is through the use of RMS moments and mismatch factor, which gauges the overlap of RMS boundaries [4]. Another more detailed approach is to use statistical divergences, which give the total contribution of differences at all points, commonly including: Kullback-Leibler divergence, Jensen-Shannon divergence, Total Variation distance, Hellinger distance, collectively referred to as f-divergence [5–9]. This paper compares these two different methods, aiming to find the connection between them.

BASIC TERMS

RMS Moments and Mismatch Factor

Considering a beam distribution centered at the origin of the (x, x') phase space, if we conduct RMS statistics on it, we can obtain the following covariance matrix:

$$\Sigma = \begin{pmatrix} \langle xx \rangle & \langle xx' \rangle \\ \langle x'x \rangle & \langle x'x' \rangle \end{pmatrix}. \tag{1}$$

This is a symmetric matrix constituted by four second-order moments. Each Σ corresponds to a quadratic form RMS phase ellipse equation

$$\mathbf{x}^T \Sigma^{-1} \mathbf{x} = 1, \tag{2}$$

where $\mathbf{x} = (x, x')^T$ is the 2D phase space coordinates of the beam.

The mismatch factor measures the difference between two ellipses in phase space that have the same center and area but do not completely overlap. They both have the form of Eq. (2) and share the same RMS emittance. If one ellipse can exactly enclose the other after being scaled by a certain proportion, then the mismatch factor is defined as the value of this scaling factor minus 1. The commonly used calculation formula is as follows [10]:

$$M = \left[1 + \frac{\Delta + \sqrt{\Delta(\Delta + 4)}}{2}\right]^{1/2} - 1, \tag{3}$$

where

$$\Delta = \Delta\alpha^2 - \Delta\beta\Delta\gamma, \tag{4}$$

with $\Delta\alpha$, $\Delta\beta$ and $\Delta\gamma$ being the differences in Twiss parameters between the two phase ellipses.

f-Divergence

The f-divergence is a class of methods used to measure the difference between two probability distributions, defined as follows (1D):

$$D_f[p(x)\|q(x)] = \int q(x) f\left[\frac{p(x)}{q(x)}\right] \mathrm{d}x; \tag{5}$$

where $p(x)$ and $q(x)$ are the probability density functions of two distributions. $f(\cdot)$ is a convex function and satisfies $f(1) = 0$. Different $f(\cdot)$ correspond to different statistical divergences. This paper selects four popular divergences to study the beam distribution, with corresponding forms and calculation formulas are shown in Table 1.

THE RELATIONSHIP BETWEEN MISMATCH FACTOR AND F-DIVERGENCE

Theorem 1 *In 2D phase space, there is a one-to-one correspondence between the f-divergence and the mismatch factor for two beam distributions with elliptical symmetry.*

* duyu@impcas.ac.cn

32nd Linear Accelerator Conference (LINAC2024),Chicago, IL, USA JACoW Publishing
ISBN: 978-3-95450-219-6 ISSN: 2226-0366 doi: 10.18429/JACoW-LINAC2024-TUPB058

Table 1: Four Forms of f-divergence; $t = \frac{p(x)}{q(x)}$

Name	$f(t)$	$D_f[p(x)\|q(x)]$
Kullback-Leibler	$t\ln t$	$\int p(x)\ln\left[\frac{p(x)}{q(x)}\right]\mathrm{d}x$
Jensen-Shannon	$\frac{1}{2}\left[\ln\left(\frac{2}{t+1}\right)^{t+1} + t\ln t\right]$	$\frac{1}{2}\int\left\{q(x)\ln\left[\frac{2q(x)}{p(x)+q(x)}\right] + p(x)\ln\left[\frac{2p(x)}{p(x)+q(x)}\right]\right\}\mathrm{d}x$
Total Variation	$\frac{1}{2}\lvert t-1\rvert$	$\frac{1}{2}\int\lvert p(x)-q(x)\rvert\mathrm{d}x$
Squared Hellinger	$(\sqrt{t}-1)^2$	$\int\left[\sqrt{p(x)}-\sqrt{q(x)}\right]^2\mathrm{d}x$

We now prove Theorem 1. In the (x, x') phase space, the f-divergence between the two beam distributions is:

$$D_f[\rho_1(\mathbf{x})\|\rho_2(\mathbf{x})] = \iint \rho_2(\mathbf{x}) f\left[\frac{\rho_1(\mathbf{x})}{\rho_2(\mathbf{x})}\right]\mathrm{d}x\mathrm{d}x', \quad (6)$$

where $\rho_1(\mathbf{x})$ and $\rho_2(\mathbf{x})$ are the probability density functions of the selected matched and mismatched beam distributions. Their RMS phase ellipses are shown in Fig. 1a.

Assuming there is a distribution with the same mismatch factor as the matched beam but with different second-order moments, with its probability density function denoted as $\tilde{\rho}_2(\mathbf{x})$, then the f-divergence between them is given by:

$$\widetilde{D}_f[\rho_1(\mathbf{x})\|\tilde{\rho}_2(\mathbf{x})] = \iint \tilde{\rho}_2(\mathbf{x}) f\left[\frac{\rho_1(\mathbf{x})}{\tilde{\rho}_2(\mathbf{x})}\right]\mathrm{d}x\mathrm{d}x'. \quad (7)$$

Their RMS phase ellipses are shown in Fig. 1c.

To facilitate the observation of the relationship between the mismatch factor and the f-divergence, we first perform the following transformation on the phase space:

$$\mathbf{m} = \mathbf{P}^{-1}\mathbf{x}, \quad (8)$$

where $\mathbf{m} = (m, m')^T$ is the transformed new phase space coordinate, and

$$\mathbf{P} = \begin{pmatrix} \sqrt{\beta} & 0 \\ -\frac{\alpha}{\sqrt{\beta}} & \frac{1}{\sqrt{\beta}} \end{pmatrix}, \quad (9)$$

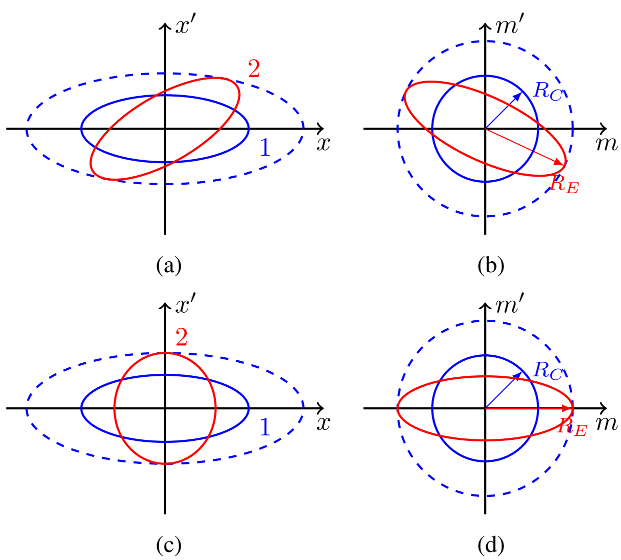


Figure 1: RMS phase ellipses before and after coordinate transformation (Blue: Matched; Red: Mismatched).

with α, β and γ being the Twiss parameters of the matched beam's RMS phase ellipse.

After this transformation, one of the RMS phase ellipses becomes a circle, while the other remains an ellipse, as shown in Figs. 1b and 1d. The relationship between the volume elements before and after the coordinate transformation is:

$$\mathrm{d}x\mathrm{d}x' = |J_1|\mathrm{d}m\mathrm{d}m', \quad (10)$$

where $|J_1|$ is the Jacobian determinant of the coordinate transformation. From Eq. (8) and Eq. (9), we obtain:

$$|J_1| = \left|\frac{\partial(x, x')}{\partial(m, m')}\right| = \begin{vmatrix} \sqrt{\beta} & 0 \\ -\frac{\alpha}{\sqrt{\beta}} & \frac{1}{\sqrt{\beta}} \end{vmatrix} = 1. \quad (11)$$

Thus, in the (m, m')phase space, Eq. (6) and (7) become:

$$D_f[\rho_1(\mathbf{Pm})\|\rho_2(\mathbf{Pm})] = \iint \rho_2(\mathbf{Pm}) f\left[\frac{\rho_1(\mathbf{Pm})}{\rho_2(\mathbf{Pm})}\right]\mathrm{d}m\mathrm{d}m' \quad (12)$$

$$\widetilde{D}_f[\rho_1(\mathbf{Pm})\|\tilde{\rho}_2(\mathbf{Pm})] = \iint \tilde{\rho}_2(\mathbf{Pm}) f\left[\frac{\rho_1(\mathbf{Pm})}{\tilde{\rho}_2(\mathbf{Pm})}\right]\mathrm{d}m\mathrm{d}m'. \quad (13)$$

Obviously, in the (m, m') phase space, Fig. 1b and Fig. 1d have the same mismatch factor. If the beam distributions are all elliptically symmetric, then the distributions in these two cases differ by only a rotation angle. Next, apply a rotation transformation to the (m, m') phase space depicted in Fig. 1b, assuming that after rotating by θ, the two beam distributions coincide with those in Fig. 1d:

$$\mathbf{u} = \mathbf{R}^{-1}\mathbf{m}, \quad (14)$$

where

$$\mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad (15)$$

with $\mathbf{u} = (u, u')^T$ being the new phase space coordinates after the rotation transformation. This results in:

$$\begin{cases} \rho_1(\mathbf{PRu}) = \rho_1(\mathbf{Pm}) \\ \rho_2(\mathbf{PRu}) = \tilde{\rho}_2(\mathbf{Pm}) \end{cases}. \quad (16)$$

At this point, the Jacobian determinant of the coordinate transformation is:

$$|J_2| = \left|\frac{\partial(m, m')}{\partial(u, u')}\right| = \begin{vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{vmatrix} = 1. \quad (17)$$

Thus, in the (u, u') phase space, Eq. (12) becomes:

$$D_f[\rho_1(\mathbf{PRu})\|\rho_2(\mathbf{PRu})] = \iint \rho_2(\mathbf{PRu}) f\left[\frac{\rho_1(\mathbf{PRu})}{\rho_2(\mathbf{PRu})}\right]\mathrm{d}u\mathrm{d}u'. \quad (18)$$

Combining with Eq. (16), it can be seen that this is consistent with the result obtained by directly calculating the

32nd Linear Accelerator Conference (LINAC2024),Chicago, IL, USA JACoW Publishing
ISBN: 978-3-95450-219-6 ISSN: 2226-0366 doi: 10.18429/JACoW-LINAC2024-TUPB058

f-divergence between the two beam distributions shown in Fig. 1d:

$$D_f[\rho_1(\mathbf{PRu})\|\rho_2(\mathbf{PRu})] = \widetilde{D}_f[\rho_1(\mathbf{Pm})\|\tilde{\rho}_2(\mathbf{Pm})]. \quad (19)$$

That is:

$$D_f[\rho_1(\mathbf{x})\|\rho_2(\mathbf{x})] = \widetilde{D}_f[\rho_1(\mathbf{x})\|\tilde{\rho}_2(\mathbf{x})]. \quad (20)$$

This indicates that for two different 2D beam distributions with elliptical symmetry, the f-divergence between them depends solely on the mismatch factor, and is independent of the orientation of the ellipsoids in phase space. Q.E.D. We can easily extend this theorem to $2n$-D phase space ($n = 1, 2, \cdots$), yielding the following theorem:

Theorem 2 *In the* $(x_1, x_1', x_2, x_2', \cdots, x_i, x_i', \cdots, x_n, x_n')$ *phase space (2n-D), given two uncoupled beam distributions with elliptical symmetry, the f-divergence between them has a one-to-one correspondence with the mismatch factor in the 2D subspaces represented by the elements of the set* $\{(x_i, x_i') \mid i = 1, 2, \cdots, n\}$.

SIMULATION RESULTS

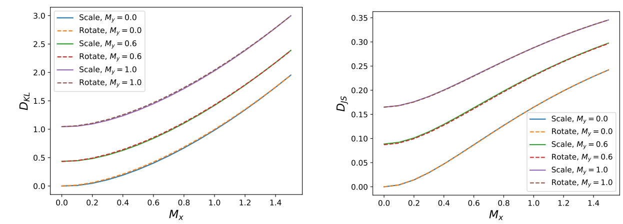
The Relationship Between 4D f-divergence and Transverse Mismatch Factors

We simulated the relationship between the f-divergence and the transverse mismatch dfactors in a 4D phase space for ideal beam distributions without $x - y$ coupling and with elliptical symmetry. First, we selected a standard distribution, then performed rotational and scaling transformations on it, ensuring that the RMS emittance of the projected distributions in both the (x, x') and (y, y') phase spaces remained unchanged. Subsequently, we calculated a series of mismatch factors M_x and M_y generated in a 2D subspace using these two different methods, as well as the 4D f-divergence between the transformed distribution and the selected standard distribution. The simulation results are shown in Fig. 2.

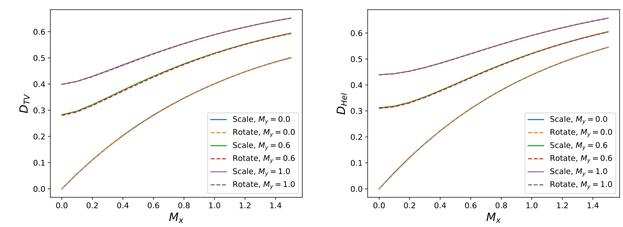
It illustrates how several popular divergences based on the Gaussian distribution vary with the mismatch factor. Additionally, it can be observed that the plots of the relationship between the mismatch factor and the f-divergence under the two methods coincide (the solid and dashed lines overlap in the figure). This indicates that for beams without $x - y$ coupling, in the 4D phase space, as long as the two transverse mismatch factors are fixed, there is a unique f-divergence that corresponds to them, which is consistent with the description of Theorem 2.

Assessment Standard for 4D f-Divergence

We utilized the one-to-one relationship between the 4D f-divergence and the two transverse mismatch factors to plot a heatmap of the f-divergence variation with mismatch factor under ideal distribution conditions, which serves as the evaluation standard. Figure 3 presents an evaluation heatmap based on Gaussian distribution. With this reference in place, when using f-divergence to calculate the differences between non-ideal beam distributions, we can roughly infer the degree of difference represented by the f-divergence values from the existing heatmap.



(a) Kullback-Leibler Divergence (b) Jensen-Shannon Divergence



(c) Total Variation Distance (d) Hellinger Distance

Figure 2: The relationship between 4D f-divergence and transverse mismatch factor (Gaussian).

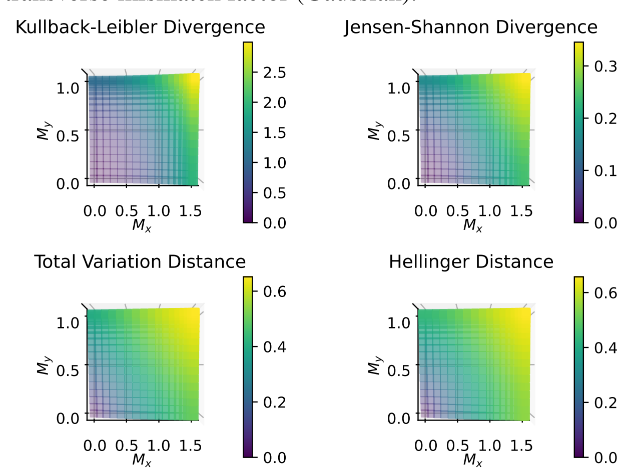


Figure 3: 4D f-divergence evaluation heatmap (Gaussian).

CONCLUSION

In this paper, we demonstrated that the f-divergence between beams with elliptical symmetry in 2D phase space had a one-to-one correspondence with the mismatch factor, and we extended this theorem to $2n$-D phase space. Subsequently, we further corroborated this theorem through simulations. Finally, based on this correspondence, we proposed evaluation standards for f-divergence, which allowed for an intuitive interpretation of the meaning of f-divergence values when assessing differences between non-ideal beam distributions.

ACKNOWLEDGEMENTS

The authors would like to thank Austin Hoover from SNS for many fruitful discussions. This work is supported by the CiADS project.

ISBN: 978-3-95450-219-6 ISSN: 2226-0366 doi: 10.18429/JACoW-LINAC2024-TUPB058

REFERENCES

[1] Y. Yang *et al.*, "A low energy beam transport system for proton beam", *Review of Scientific Instruments*, vol. 84, no. 3, pp. 83–89, 2013.

[2] P. Forck, "Lecture notes on beam instrumentation and diagnostics", Joint University Accelerator School, 2011.

[3] J. C. Wong *et al.*, "4d transverse phase space tomography of an operational hydrogen ion beam via noninvasive 2d measurements using laser wires", *Physical Review. Accelerators and Beams*, vol. 25, no. 4, 2022.

[4] K. R. Crandall, "Trace 3-d documentation", Los Alamos National Laboratory, 1987. doi:10.2172/6290515

[5] I. Csiszar, "On topological properties of f-divergence", *Studia Scientiarum Mathematicarum Hungarica*, vol. 2, 1967.

[6] D.I. Belov *et al.*, "Distributions of the Kullback–Leibler divergence with applications", *British Journal of Mathematical & Statistical Psychology*, vol. 64, no. 2, pp. 291–309, 2011.

[7] F. Nielsen, "On a Variational Definition for the Jensen-Shannon Symmetrization of Distances based on the Information Radius", *Entropy*, vol 23, p. 464, 2021. doi:10.3390/e23040464

[8] J. K. Chung *et al.*, "Measures of distance between probability distributions", *Journal of Mathematical Analysis and Applications*, vol. 138, no. 1, pp. 280–92, 1989.

[9] D. A. Cieslak *et al.*, "Hellinger distance decision trees are robust and skew-insensitive", *Data Mining Knowledge Discovery*, vol. 24, no. 1, pp. 136–158, 2012.

[10] T. Wangler, *RF Linear Accelerators*, Wiley, 2008. doi:10.1002/9783527623426