stats

MDPI

Communication

Goodness of Chi-Square for Linearly Parameterized Fitting

George Livadiotis

Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA; glivadiotis@princeton.edu

Abstract

The paper shows an alternative perspective of the reduced chi-square as a measure of the goodness of fitting methods. The reduced chi-square is given by the ratio of the fitting over the propagation errors, that is, a universal relationship that holds for any linearity, but not for a nonlinearly parameterized fitting model. We begin by providing the proof for the traditional examples of one-parametric fitting of a constant and the bi-parametric fitting of a linear model, and then, for the general case of any linearly multi-parameterized model. We also show that this characterization is not generally true for nonlinearly parameterized fitting. Finally, we demonstrate these theoretical developments with an application in real data from the plasma protons in the heliosphere.

Keywords: fitting methods; errors; chi-square; space science

1. Introduction

The paper revisits a well-known concept—the reduced chi-square—and introduces and formalizes a novel interpretation of the ratio between fitting and propagation errors. This supports the usage of the reduced chi-square and the criterion of its deviation from one as a measure of the goodness of fitting.

The chi-square sums the squares of the residuals between a given dataset and a model, and is suitable for fitting this dataset [1]. Given the dataset of N observations, $\{y_i\}_{i=1}^N$, the respective modeled values $\{f_i\}_{i=1}^N$ are determined as $f_i \equiv f(x_i)$ from the statistical model $f(x)$, which is expected to best describe the data. Often, the statistics of observational data are unknown, thus a general fitting method is employed to adjust the respective modeled values to data by minimizing their total deviations $\sum_{i=1}^N \Phi(y_i - f_i)$ defined through a norm Φ. For the Euclidean norm, the sum of the square deviations is used, given by $\sum_{i=1}^N (y_i - f_i)^2$.

A mono-parametric fitting model, $f(x; p)$, depends on one single parameter p, e.g., the simple case of fitting a constant, $f(x; p) = p$. A bi-parametric fitting model $f(x; p_1, p_2)$ depends on two independent parameters p_1 and p_2, e.g., the famous least-square methods of a linear fit, $f(x; p_1, p_2) = p_1 + p_2 \cdot x$. In general, we consider a multi-parametric fitting model $f(x; \{p_k\}_{k=1}^M)$ that depends on M independent parameters, $\{p_k\}_{k=1}^M = p_1, p_2, \ldots, p_M$. In all of the cases, the fitting method involves finding the optimal values of the parameters that minimize the sum of the squares $\sum_{i=1}^N (y_i - f_i)^2$ [2,3]. The chi-square is a similar sum expression to be minimized but also includes the data uncertainties.

Observational data are typically accompanied by the respective uncertainties, $\{y_i \pm \sigma_{yi}\}_{i=1}^N$, which are used for normalizing the deviation of data with the fitted model,

check for updates

Academic Editor: Wei Zhu

Received: 21 September 2025
Revised: 25 November 2025
Accepted: 28 November 2025
Published: 1 December 2025

Citation: Livadiotis, G. Goodness of Chi-Square for Linearly Parameterized Fitting. *Stats* **2025**, *8*, 113. https://doi.org/10.3390/stats8040113

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

$\{|y_i - f_i|/\sigma_{yi}\}_{i=1}^{N}$. These ratios are summed up to construct the chi-square, using the Euclidean norm L_2, that is, by adding the squares,

$$\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{y_i - f_i}{\sigma_{yi}} \right)^2. \tag{1}$$

There are four straightforward generalizations of the chi-square. These are described and formulated as follows:

(1) When a non-Euclidean norm, L_q, is being used to sum the vector components $\{|y_i - f_i|/\sigma_{yi}\}_{i=1}^{N}$, then, we minimized the chi-q (instead of the chi-square), defined by [4–6]

$$\chi^q \equiv \sum_{i=1}^{N} \left| \frac{y_i - f_i}{\sigma_{yi}} \right|^q. \tag{2}$$

The optimization involves finding, aside from the optimal value of the fitting parameters, the optimal value of q that characterizes the non-Euclidean norm, L_q.

(2) When the observations of the independent variable x have significant uncertainties $\{x_i \pm \sigma_{xi}\}_{i=1}^{N}$, then, the optimization involves minimizing the chi-square defined by [7].

$$\chi^2 \equiv \sum_{i=1}^{N} \frac{(y_i - f_i)^2}{{\sigma_{yi}}^2 + (\partial f_i/\partial x_i)^2 {\sigma_{xi}}^2}. \tag{3}$$

(3) When the vector components are characterized by correlations, then, a non-diagonal covariance matrix ${\sigma^2}_{ij}$ exists, whose inverse matrix is involved in the construction of the chi-square [8], i.e.,

$$\chi^2 \equiv \sum_{i=1}^{N} \sum_{j=1}^{N} (y_i - f_i)(\sigma^2)_{ij}^{-1} (y_j - f_j). \tag{4}$$

(4) The relation between degrees of freedom and information is significant in statistics and information theory. Degrees of freedom represent the number of independent values in a dataset that can vary while satisfying statistical constraints. Therefore, this number indicates the amount of information available to estimate parameters. Furthermore, when a significant variability characterizes the uncertainties of observations, then, a potential statistical bias exists in their use as statistical weights (expressed as inverse variance). This bias is reduced through the effective statistical degrees of freedom N_{e} (replacing the number of data points, N), which is as follows [9]:

$$N_{\mathrm{e}} \equiv \left(\sum_{i=1}^{N} {\sigma_{yi}}^{-2} \right)^2 / \sum_{i=1}^{N} {\sigma_{yi}}^{-4}. \tag{5}$$

This is reduced to the number of observations, N, when all the uncertainties are equal, $\{\sigma_{yi} = \sigma_y\}_{i=1}^{N}$. Then, the effective independent degrees of freedom are given by $N_{\mathrm{e}} - M$, instead of $N - M$, where M is the number of the independent parameters to be optimized by the fitting. The notation of the effective independent degrees of freedom is used for formulating the reduced chi-square.

All the above four scenarios are well-studied and used in statistical analyses (e.g., see the cited references). In this paper, we focus on the standard form of (1), in order to demonstrate the statistical meaning of the chi-square in characterizing the goodness of fitting.

The statistical model is expressed by a function, f, of the independent variable x and of a number of independent parameters $\{p_k\}_{k=1}^{M}$, in order to describe the observations of $\{y_i \pm \sigma_{yi}\}_{i=1}^{N}$ over $\{x_i\}_{i=1}^{N}$, so that $y_i = f\left(x_i; \{p_k\}_{k=1}^{M}\right) + e_i$, where e_i is the error term that follows a probability distribution (typically assumed to be a normal distribution). A linearly parametric model is referred to as a linear relationship of the model's function with the involved parameter(s), and not with the variable x; e.g., the model function $f(x;p) = x^2 + p \cdot x^3$ is linearly parameterized, while the function $f(x;p) = p + p^2 \cdot x$ is nonlinear.

In general, the linearly parameterized model function can be written as

$$f\left(x; \{p_k\}_{k=1}^{M}\right) = g_0(x) + \sum_{k=1}^{M} p_k \cdot g_k(x) \tag{6}$$

which involves arbitrary functions of the variable x, $\{g_k(x)\}_{k=0}^{M}$.

Minimization of the constructed chi-square,

$$\chi^2\left(\{p_k\}_{k=1}^{M}\right) \equiv \sum_{i=1}^{N} \left[\frac{y_i - f\left(x_i; \{p_k\}_{k=1}^{M}\right)}{\sigma_{yi}}\right]^2, \tag{7}$$

involves finding the optimal values of the fitting parameters, $\{p_k = {p_k}^*\}_{k=1}^{M}$, so that

$$\partial\chi^2\left(\{{p_k}^*\}_{k=1}^{M}\right)/\partial p_j = 0, \text{ for all } j: \ 1, \ \ldots, \ M. \tag{8}$$

The minimum value of the chi-square, $\chi^2\left(\{{p_k}^*\}_{k=1}^{M}\right)$, equals the sum of the residuals and can provide a measure of the goodness of the fitting. This is precisely given by the minimum reduced chi-square, that is, the chi-square over the independent degrees of freedom $N - M$,

$$\chi^2_{\text{red}} \equiv \frac{1}{N-M} \cdot \chi^2_{\text{min}}. \tag{9}$$

How can a chi-square be used to characterize the goodness of a fitting method? The first thought is that the smaller the minimum chi-square value, the better the fitting. Ideally, the fitting is perfect when the sum of the residuals turns into zero for some value of the fitting parameters, $\{p_k = {p_k}^*\}_{k=1}^{M}$; when this occurs, all the vector components of the data-model deviation (fitting residuals) are zero, $\{\,|y_i - f_i^*| = 0\}_{i=1}^{N}$; (note that we used the abbreviation $f_i^* \equiv f(x_i; \{{p_k}^*\}_{k=1}^{M})$). Such a fitting would have been meaningful, if the observation uncertainties were negligible, $\{\,\sigma_{yi} \sim 0\}_{i=1}^{N}$. On the contrary, when these uncertainties are finite and not negligible, such an adaptation of a perfect fitting becomes meaningless; indeed, what is the point of having residuals between observations and the model tending to zero $|y_i - f_i^*| \sim 0$, if the uncertainties of these observations are significantly larger than these deviations, i.e., $\sigma_{yi} >> |y_i - f_i^*|$?

The fit, once optimized, is considered good when the respective residuals are neither larger, $|y_i - f_i^*| > \sigma_{yi}$ (underfitting), nor smaller, $|y_i - f_i^*| < \sigma_{yi}$ (overfitting), than the observations uncertainties, but rather, when they have near values, $|y_i - f_i^*| \sim \sigma_{yi}$. Ideally, the observation uncertainty σ_{yi} would be equal to the minimum value of each data-model deviation, $|y_i - f_i^*|$, instead of being equal to zero, in order for the fitting to be considered meaningful. In this case, the summation on all the points to derive the best chi-square value simply provides the number of points, $(\chi^2)^* \sim N$, while the average of all the vector components $\{\,|y_i - f_i|/\sigma_{yi}\}_{i=1}^{N}$ provides unity, $\frac{1}{N}(\chi^2)^* \sim 1$. The average chi-square approximates its reduced value, whose optimal value is ~1.

In a more detailed derivation, the reduced chi-square is exactly involved in this interpretation. In particular, let the observation's uncertainties to be equal, $\{\sigma_{yi} = \sigma_y\}_{i=1}^{N}$; then, the chi-square is $\chi^2 = {\sigma_y}^{-2} \cdot \sum_{i=1}^{N} (y_i - f_i)^2$ and its reduced value is $\chi^2_{\text{red}} \equiv \frac{1}{N-M} \cdot \chi^2 = {\sigma_y}^{-2} \cdot \frac{1}{N-M} \cdot \sum_{i=1}^{N} (y_i - f_i)^2$. The averaged squared residuals, that is, the variance between the data $\{y_i\}_{i=1}^{N}$ and model $\{f_i\}_{i=1}^{N}$ values, are determined by $\overline{|y-f|}^2 = \frac{1}{N-M} \cdot \sum_{i=1}^{N} (y_i - f_i)^2$; hence,

$$\overline{|y-f|} = \sqrt{\frac{1}{N-M} \cdot \sum_{i=1}^{N} (y_i - f_i)^2} = \sqrt{\chi^2_{\text{red}}} \cdot \sigma_y. \tag{10}$$

Therefore, the standard deviation $\overline{|y-f|}$ is minimized to its smallest meaningful value, that is, the standard deviation of the observation values, σ_y, only when χ^2_{red} is equal to unity, i.e., $\overline{|y-f^*|} \sim \sigma_y \Leftrightarrow \chi^2_{\text{red}} \sim 1$. This observation is fundamental, interwoven with the definition, rather than a simple property, of the chi-square.

The concept and interpretation of the chi-square is interwoven with two types of errors characterizing the fitting parameters. The first, the fitting error, comes from the regression, caused by the nonzero deviation between the data and modeled values, $\overline{|y-f|}$. The second, the propagation error, comes from the existence of the observation data uncertainties, that is, the nonzero value of σ_y. Equation (10) provides insights that the chi-square expresses the ratio between the two types of errors, the fitting over the propagation error. We demonstrate this relation between the errors in Figure 1.

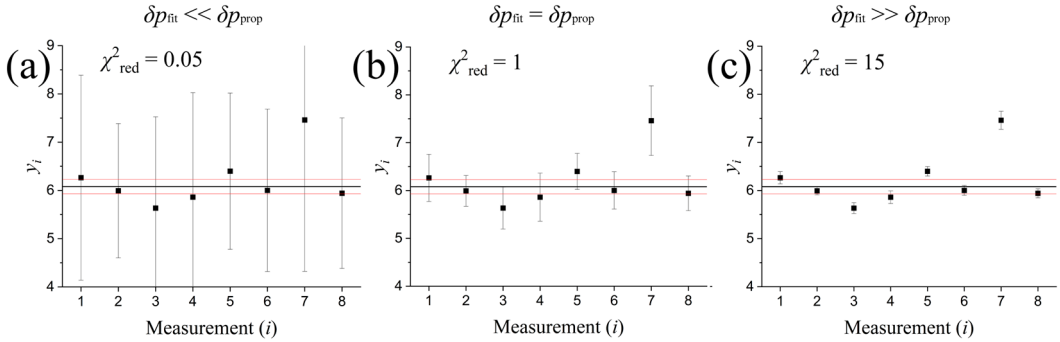


Figure 1. A constant statistical model fits eight data points with different errors: (**a**) when the errors are too large (overestimation), the reduced chi-square is less than 1, the fitted line that represents the average of data points (black line) does pass through all the data points or their error lines, but the variation in the fitted model (red lines) is small compared to the size of errors, leading to a meaningless rate of good fitting. (**c**) When the errors are too small (underestimation), the calculated reduced chi-square is more than 1, and the fitted line does not pass through all the data points or their error lines, leading to an obvious bad fitting. (**b**) In the case where the errors are similar to the deviations of the data points from the model, the reduced chi-square is about 1, and the fitting is characterized as good.

In particular, Figure 1 shows the fitting of eight data values with a constant fitting model (a special case of a linear model is shown in detail in Section 2.1; see: [10–12]). This fitting is performed to test the statistical hypothesis that the data values are well represented by a constant. The characterization of the fitting depends on the accompanying observational uncertainties. Too large uncertainties (Figure 1a) lead to small value of the reduced chi-square ($\chi^2_{\text{red}} < 1$); if the fitting model is the right choice, then, this result indicates that there is an error overestimation characterizing the observational data. On the other hand, too small uncertainties (Figure 1c) lead to a large value of the reduced chi-square ($\chi^2_{\text{red}} > 1$); again, if the fitting model is the right choice, then, this outcome

indicates that there is an error underestimation of the observational data. The optimal fitting goodness would be for observation uncertainties similar to the data variation (Figure 1b), meaning that the two types of errors are about the same, and the reduced chi-square is about one ($\chi^2_{\text{red}} \sim 1$).

The statistical behavior of chi-squared values derived from model residuals has been thoroughly understood for over a century, thanks to the foundational work of R.A. Fisher [13,14]. Fisher not only performed the necessary algebraic derivations but also rigorously established the associated probability distributions. Building on this classical framework, the current study introduces a fresh algebraic and intuitive perspective on reduced chi-squared statistics.

The purpose of this paper is to show an alternative perspective to the statistical meaning of the reduced chi-square as a measure of the goodness of fitting methods, and finally, to demonstrate it with several examples and applications. This expression is given by the ratio between the two types of errors, the fitting over the propagation error, which holds for any linearly parameterized fitting. We show this characterization of chi-square, as follows. We first work with the traditional examples of one-parametric fitting of a constant and the bi-parametric fitting of a linear model; then, we provide the proof for the general case of any linearly multi-parameterized model. Nevertheless, we present a counterexample of this characterization of chi-square, showing that it is not generally true for nonlinearly parameterized fitting. Furthermore, we focus on the nature of the two types of errors and show their independent origin. We also examine how the observation uncertainties affect the formulations of the fitting and propagation errors and the reduced chi-square. The developments of this paper in regard to the chi-square formulation and characterization are applied in the case of plasma protons of the heliosheath; we find that the acceptable characterization of the fitting goodness is given by one order of magnitude of the reduced chi-square around its optimal unity value, $0.1 < \chi^2_{\text{red}} < 10$.

The paper is organized as follows: In Section 2, we show both the cases of linearly and nonlinearly parametric fitting models. We start with a linearly mono-parametric model, then work with bi-parametric model, and end up with the general case of a multi-parametric fitting model of an arbitrary number of independent parameters. In the same section, we show that the reduced chi-square cannot be used for nonlinear parametric fitting models. We show this separately for nonlinearly mono-parameterized and multi-parameterized models. In Section 3, we provide an example application from the plasma protons in the heliosphere. Finally, in Section 4, we summarize the conclusions and findings of this paper.

2. Methods

2.1. Linearly Parametrical Fitting

2.1.1. Linearly Mono-Parametric Fitting

The linearly mono-parametric model function can be generally written as

$$f(x;p) = g_0(x) + p \cdot g_1(x), \tag{11}$$

where the variable functions $g_0(x)$ and $g_1(x)$ may be linear or nonlinear. Here, we use the simplest case of a constant function, $f(x;p) = p$. Minimization of the chi-square,

$$\chi^2(p) = \sum_{i=1}^{N} \left(\frac{y_i - p}{\sigma_{yi}} \right)^2, \tag{12}$$

involves finding the optimal values of the fitting parameters, $p = p^*$, so that

$$p^* = \overline{y} = \sum_{i=1}^{N} \frac{y_i}{{\sigma_{yi}}^2} / \sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2}. \tag{13}$$

In order to derive the fitting error, the chi-square is expanded near its minimum,

$$\chi^2(p) \cong A_0 + A_2 \cdot (p - p^*)^2 + O(p - p^*)^3. \tag{14}$$

The square of the fitting error is proportional to the (reduced) chi-square and inversely proportional to the curvature coefficient A_2 [5], i.e.,

$$\delta {p^*}_{\text{fit}} = \sqrt{\frac{1}{N-1} A_0 \cdot A_2^{-1}}, \tag{15}$$

where $A_0 = \chi^2(p^*)$ is the minimum chi-square and $\frac{1}{N-1} A_0 = \chi^2_{\text{red}}$ is the reduced chi-square. For more details, see [5]. In our case, the expansion of the chi-square, given by Equation (14), is

$$\chi^2(p) = \sum_{i=1}^{N} \left[\frac{(y_i - p^*) - (p - p^*)}{\sigma_{yi}} \right]^2 = \chi^2(p^*) + \left(\sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \right) \cdot (p - p^*)^2. \tag{16}$$

Hence, the curvature coefficient equals

$$A_2 = \sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2}, \tag{17}$$

thus, the fitting error equals

$$\delta {p^*}_{\text{fit}} = \sqrt{\chi^2_{\text{red}}} / \sqrt{\sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2}}. \tag{18}$$

The origin of the fitting error comes from the residuals, i.e., the deviations between the data and model, $\{ y_i - f_i \}_{i=1}^{N}$. The role of the observational uncertainties of the data, $\{\sigma_{yi}\}_{i=1}^{N}$, in the formulation of this deviation, is simply to express the statistical weights. They do not contribute to the measures of these deviations. For example, if all the observational uncertainties were equal, either big or small, then, they would have always zero impact on the value of the fitting error. Indeed, if $\{ \sigma_{yi} = \sigma \}_{i=1}^{N}$, then,

$$\delta {p^*}_{\text{fit}} = \sqrt{\frac{1}{N(N-1)} \cdot \textstyle\sum_{i=1}^{N} (y_i - p^*)^2}, \tag{19}$$

that is, the fitting error becomes independent of the uncertainty σ.

The actual measure of the observational uncertainties comes with the propagation error. This is a type of error different from the fitting error; it generally differs in its nature, origin, and value. The propagation error originates from the function that expresses the optimal value in terms of the observation data, $p^*(\{y_i\}_{i=1}^{N})$. The variation in this function leads to the propagation error,

$$\delta {p^*}_{\text{prop}} = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial p^*}{\partial y_i} \right)^2 {\sigma_{yi}}^2}. \tag{20}$$

In our case, where $p^*(\{y_i\}_{i=1}^{N}) = \overline{y}$, the propagation error is specifically equal to

$$\delta p^*{}_{\text{prop}} = 1/\sqrt{\sum_{i=1}^{N} \frac{1}{\sigma_{yi}{}^2}}. \tag{21}$$

We note that the square of the propagation error equals the inverse curvature coefficient,

$$\delta p^*{}^2_{\text{fit}} = A_2^{-1}, \tag{22}$$

where we conclude that the chi-square is related to the ratio of the two errors, i.e.,

$$\sqrt{\chi^2_{\text{red}}} = \delta p^*{}_{\text{fit}} / \delta p^*{}_{\text{prop}}. \tag{23}$$

The fitting error is given by the formulation of the standard error, i.e.,

$$\delta \overline{y}^2 = \frac{1}{N-1} \sum_{i=1}^{N} \frac{(y_i - \overline{y})^2}{\sigma_{yi}{}^2} / \sum_{i=1}^{N} \frac{1}{\sigma_{yi}{}^2}. \tag{24}$$

Thus, its connection to the reduced chi-square is

$$\chi^2_{\text{red}} = \frac{1}{N-1} \sum_{i=1}^{N} \frac{(y_i - \overline{y})^2}{\sigma_{yi}{}^2} = \left(\sum_{i=1}^{N} \frac{1}{\sigma_{yi}{}^2} \right) \cdot \delta \overline{y}^2, \tag{25}$$

or

$$\delta p^*{}_{\text{fit}} = \delta \overline{y} = \sqrt{\chi^2_{\text{red}}} / \sqrt{\textstyle\sum_{i=1}^{N} \sigma_{yi}{}^{-2}},\ \delta p^*{}_{\text{prop}} = 1/\sqrt{\textstyle\sum_{i=1}^{N} \sigma_{yi}{}^{-2}}. \tag{26}$$

An example of this fitting is shown in Figure 1. Large deviations between propagation and fitting errors lead to a reduced chi-square significantly different than one, corresponding to low fitting goodness.

2.1.2. Linearly Bi-Parametric Fitting

The linearly bi-parametric model function can be written as

$$f(x; p_1, p_2) = g_0(x) + p_1 \cdot g_1(x) + p_2 \cdot g_2(x). \tag{27}$$

As an example, we use the simplest case of a linear function, $f(x; p_1, p_2) = p_1 + p_2 \cdot x$, with the fitting parameters of intercept p_1 and slope p_2. Minimization of the chi-square,

$$\chi^2(p_1, p_2) = \sum_{i=1}^{N} \left(\frac{y_i - p_1 - p_2 \cdot x_i}{\sigma_{yi}} \right)^2, \tag{28}$$

leads to the optimal fitting parameters:

$$p_1^* = \overline{y} - p_2^* \cdot \overline{x},\ p_2^* = \sigma_{xy}^2 / \sigma_{xx}^2. \tag{29}$$

The minimum value of the chi-square equals

$$\chi^2(p_1^*, p_2^*) = \sum_{i=1}^{N} \left[\frac{(y_i - \overline{y}) - p_2^* \cdot (x_i - \overline{x})}{\sigma_{yi}} \right]^2 = \sigma_{xx}^{-2} \cdot (\sigma_{xx}^2 \sigma_{yy}^2 - \sigma_{xy}^4) \cdot \sum_{i=1}^{N} \frac{1}{\sigma_{yi}{}^2}, \tag{30}$$

where we find

$$\begin{aligned}\sigma_{yy}^2 &= \sum_{i=1}^{N} \frac{(y_i-\overline{y})^2}{\sigma_{yi}^2} / \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2},\\ \sigma_{xy}^2 &= \sum_{i=1}^{N} \frac{(x_i-\overline{x})(y_i-\overline{y})}{\sigma_{yi}^2} / \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2},\\ \sigma_{xx}^2 &= \sum_{i=1}^{N} \frac{(x_i-\overline{x})^2}{\sigma_{yi}^2} / \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2},\end{aligned} \tag{31}$$

and

$$\overline{x} = \sum_{i=1}^{N} \frac{x_i}{\sigma_{yi}^2} / \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2},\ \overline{y} = \sum_{i=1}^{N} \frac{y_i}{\sigma_{yi}^2} / \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2}. \tag{32}$$

Next, we calculate the two types of errors. First, we find the propagation error of the optimal slope, which is,

$$\delta p_{2\text{prop}}^* = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial p_2^*}{\partial y_i}\right)^2 \sigma_{yi}^2} = \sigma_{xx}^{-2} \sqrt{\sum_{i=1}^{N} \left(\frac{\partial \sigma_{xy}^2}{\partial y_i}\right)^2 \sigma_{yi}^2}. \tag{33}$$

We have

$$\frac{\partial \sigma_{xy}^2}{\partial y_j} \sigma_{yi} = \frac{(x_i-\overline{x})}{\sigma_{yi}} / \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2}, \tag{34}$$

or

$$\sum_{j=1}^{N} \left(\frac{\partial \sigma_{xy}^2}{\partial y_i}\right)^2 \sigma_{yi}^2 = \sum_{i=1}^{N} \frac{(x_i-\overline{x})^2}{\sigma_{yi}^2} / \left(\sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2}\right)^2 = \sigma_{xx}^2 / \left(\sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2}\right). \tag{35}$$

Hence, the propagation error of the optimal slope is

$$\delta p_{2\text{prop}}^* = \sigma_{xx}^{-1} \cdot \left(\sum_{i=1}^{N} \sigma_{yi}^{-2}\right)^{-1/2}. \tag{36}$$

The propagation error of the optimal intercept is

$$\delta p_{1\text{prop}}^* = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial p_1^*}{\partial y_i}\right)^2 \sigma_{yi}^2}, \tag{37}$$

where $p_1^* = \overline{y} - p_2^* \cdot \overline{x}$; thus, we obtain

$$\frac{\partial p_1^*}{\partial y_j} = \frac{\partial \overline{y}}{\partial y_j} - \frac{\partial p_2^*}{\partial y_j} \cdot \overline{x} = \left[1 - (x_j - \overline{x}) \cdot \overline{x} \sigma_{xx}^{-2}\right] \sigma_{yj}^{-2} / \sum_{i=1}^{N} \sigma_{yi}^{-2}, \tag{38}$$

or

$$\left(\frac{\partial p_1^*}{\partial y_j}\right)^2 \sigma_{yj}^{-2} = \left[1 - (x_j - \overline{x}) \cdot \overline{x} \sigma_{xx}^{-2}\right]^2 \sigma_{yj}^{-2} / \left(\sum_{i=1}^{N} \sigma_{yi}^{-2}\right)^2, \tag{39}$$

namely,

$$\sum_{i=1}^{N} \left(\frac{\partial p_1^*}{\partial y_i}\right)^2 \sigma_{yi}^2 = (1 + \overline{x}^2 \sigma_{xx}^{-2}) / \sum_{i=1}^{N} \sigma_{yi}^{-2} = \overline{x^2} \sigma_{xx}^{-2} / \sum_{i=1}^{N} \sigma_{yi}^{-2}. \tag{40}$$

Hence, we find that the propagation error of the optimal intercept is

$$\delta p_{1\text{prop}}^* = \sqrt{\overline{x^2}} \cdot \sigma_{xx}^{-1} \cdot \left(\sum_{i=1}^{N} \sigma_{yi}^{-2}\right)^{-1/2}. \tag{41}$$

Next, we derive the respective fitting errors. For this, we expand the chi-square at its minimum value,

$$\chi^2(p_1, p_2) \cong A_0 + \sum_{m=1}^{2}\sum_{n=1}^{2}(A_2)_{mn}\cdot(p_m - p_m^*)(p_n - p_n^*) + O[(p_m - p_m^*)^a(p_n - p_n^*)^b]. \quad (42)$$

ignoring higher terms for $a + b \geq 3$. Then, the chi-square minimum is

$$A_0 = \chi^2(p_1^*, p_2^*),\ \frac{1}{N-2}A_0 = \chi^2_{\text{red}}, \quad (43)$$

while the curvature coefficient A_2 is a matrix equal to half the Hessian matrix H [14],

$$(A_2)_{m\,n} = \frac{1}{2}H_{mn}, \quad (44)$$

defined by

$$H_{mn} \equiv \frac{\partial^2\chi^2(p_1^*, p_2^*)}{\partial p_m \partial p_n}. \quad (45)$$

The curvature coefficient matrix is

$$(A_2) = \left(\sum_{i=1}^{N}{\sigma_{yi}}^{-2}\right)\cdot\begin{pmatrix}1 & \overline{x}\\ \overline{x} & \overline{x^2}\end{pmatrix} \text{or } (A_2)^{-1} = {\sigma_{xx}}^{-2}\left(\sum_{i=1}^{N}{\sigma_{yi}}^{-2}\right)^{-1}\cdot\begin{pmatrix}\overline{x^2} & -\overline{x}\\ -\overline{x} & 1\end{pmatrix}. \quad (46)$$

Thus,

$$(A_2)^{-1}_{11} = \overline{x^2}\cdot{\sigma_{xx}}^{-2}\cdot\left(\sum_{i=1}^{N}{\sigma_{yi}}^{-2}\right)^{-1} \text{and } (A_2)^{-1}_{22} = {\sigma_{xx}}^{-2}\cdot\left(\sum_{i=1}^{N}{\sigma_{yi}}^{-2}\right)^{-1}. \quad (47)$$

Therefore, the fitting errors are given by

$$\delta p^*_{m\text{fit}} = \sqrt{\frac{1}{N-2}A_0\cdot(A_2)^{-1}_{mm}}, \quad (48)$$

or

$$\delta p^*_{1\text{fit}} = \sigma_{xx}^{-2}\cdot\sqrt{\frac{1}{N-2}\overline{x^2}(\sigma^2_{xx}\sigma^2_{yy} - \sigma^4_{xy})},\ \delta p^*_{2\text{fit}} = \sigma_{xx}^{-2}\cdot\sqrt{\frac{1}{N-2}(\sigma^2_{xx}\sigma^2_{yy} - \sigma^4_{xy})}. \quad (49)$$

We observe that the respective propagation errors are

$$\delta p^*_{1\text{prop}} = \sqrt{(A_2)^{-1}_{11}} \text{and } \delta p^*_{2\text{prop}} = \sqrt{(A_2)^{-1}_{22}}. \quad (50)$$

Finally, the ratio of the fitting over the propagation errors, for either the slope or the intercept, provides the chi-square,

$$\delta p^*_{1\text{prop}} = \sqrt{(A_2)^{-1}_{11}} \text{and } \delta p^*_{2\text{prop}} = \sqrt{(A_2)^{-1}_{22}}. \quad (51)$$

which can be given more explicitly through the following formulation:

$$\chi^2_{\text{red}} = \frac{1}{N-2}\cdot{\sigma_{xx}}^{-2}\cdot(\sigma^2_{xx}\sigma^2_{yy} - \sigma^4_{xy})\cdot\sum_{i=1}^{N}{\sigma_{yi}}^{-2}. \quad (52)$$

2.1.3. Linearly Multi-Parametric Fitting

Lastly, in this section, we examine the most general linearly multi-parametric model function, i.e.,

$$f\left(x;\{p_k\}_{k=1}^{M}\right)=\sum_{k=1}^{M}p_k\cdot g_k(x). \tag{53}$$

that involves the arbitrary nonlinear functions of x, $\{g_k(x)\}_{k=1}^{M}$. (Note that in this compact version of Equation (53), we have included the former function $g_0(x)$ in the sum with $p_0 = 1$.)

The chi-square is now formulated by

$$\chi^2\left(\{p_k\}_{k=1}^{M}\right)=\sum_{i=1}^{N}\left[\frac{y_i-\sum\limits_{k=1}^{M}p_k\cdot g_k(x_i)}{\sigma_{yi}}\right]^2. \tag{54}$$

The expansion around its minimum is

$$\chi^2\left(\{p_k\}_{k=1}^{M}\right)\cong A_0+\sum_{m=1}^{M}\sum_{n=1}^{M}(A_2)_{mn}\cdot(p_m-p_m^*)(p_n-p_n^*)+O[(p_m-p_m^*)^a(p_n-p_n^*)^b], \tag{55}$$

where the minimum chi-square value is given by

$$A_0=\chi^2\left(\{p_k^*\}_{k=1}^{M}\right),\ \frac{1}{N-M}A_0=\chi^2_{\text{red}}, \tag{56}$$

while the curvature coefficient A_2 is a matrix equal to half the Hessian matrix H,

$$(A_2)_{mn}=\frac{1}{2}H_{mn},\ H_{mn}\equiv\frac{\partial^2\chi^2\left(\{p_k^*\}_{k=1}^{M}\right)}{\partial p_m\partial p_n}. \tag{57}$$

The minimization of the chi-square, that is,

$$\frac{\partial\chi^2\left(\{p_k\}_{k=1}^{M}\right)}{\partial p_m}=-2\sum_{i=1}^{N}\left[\frac{y_i-\sum\limits_{k=1}^{M}p_k\cdot g_k(x_i)}{{\sigma_{yi}}^2}\right]\cdot g_m(x_i)=0, \tag{58}$$

leads to the optimal values of the parameters, which are implicitly given by

$$\sum_{i=1}^{N}\frac{y_i\,g_m(x_i)}{{\sigma_{yi}}^2}=\sum_{k=1}^{M}p_k^*\cdot\sum_{i=1}^{N}\frac{g_k(x_i)g_m(x_i)}{{\sigma_{yi}}^2}. \tag{59}$$

Note that the argument $\sum_{i=1}^{N}{\sigma_{yi}}^{-2}g_a(x_i)g_b(x_i)$ is a matrix (two free indices, a and b), which is connected with Hessian matrix. Indeed, the Hessian or curvature matrix is derived from differentiating Equation (58) once more, that is,

$$(A_2)_{mn}=\frac{1}{2}H_{mn}=\frac{1}{2}\frac{\partial^2\chi^2\left(\{p_k\}_{k=1}^{M}\right)}{\partial p_m\partial p_n}=\sum_{i=1}^{N}\frac{g_m(x_i)g_n(x_i)}{{\sigma_{yi}}^2}. \tag{60}$$

Hence, we observe that Equation (59) can be written in terms of the curvature matrix, i.e.,

$$\sum_{i=1}^{N}\frac{y_i\,g_m(x_i)}{{\sigma_{yi}}^2}=\sum_{k=1}^{M}p_k^*\cdot(A_2)_{km}. \tag{61}$$

Some calculations lead to

$$\sum_{i=1}^{N}\left[\frac{y_i}{\sigma_{yi}^{\,2}}\sum_{m=1}^{M} g_m(x_i)(A_2)^{-1}_{m\,n}\right] = \sum_{k=1}^{M} p_k^* \cdot \sum_{m=1}^{M}(A_2)_{km}(A_2)^{-1}_{m\,n} = \sum_{k=1}^{M} p_k^*\delta_{kn} = p_n^*, \tag{62}$$

and

$$\begin{aligned}\left(\tfrac{\partial p_n^*}{\partial y_j}\right)^2\sigma_{yj}^{\,2} &= \sigma_{yj}^{\,2}\left[\sum_{m=1}^{M} g_m(x_j)(A_2)^{-1}_{m\,n}\right]^2 \\ &= \sigma_{yj}^{\,2}\sum_{m=1}^{M}\sum_{m'=1}^{M} g_m(x_j)g_{m'}(x_j)(A_2)^{-1}_{m\,n}(A_2)^{-1}_{m'\,n}\,,\end{aligned} \tag{63}$$

thus,

$$\begin{aligned}&\sum_{j=1}^{N}\left(\frac{\partial p_n^*}{\partial y_j}\right)^2\sigma_{yj}^{\,2} = \sum_{m=1}^{M}\sum_{m'=1}^{M}\left[\sum_{j=1}^{N}\frac{g_m(x_j)g_{m'}(x_j)}{\sigma_{yj}^{\,2}}\right](A_2)^{-1}_{m\,n}(A_2)^{-1}_{m'\,n} \\ &= \sum_{m=1}^{M}\sum_{m'=1}^{M}(A_2)_{mm'}(A_2)^{-1}_{m\,n}(A_2)^{-1}_{m'\,n} = \sum_{m=1}^{M}\delta_{mn}(A_2)^{-1}_{mn} = (A_2)^{-1}_{nn}\,,\end{aligned} \tag{64}$$

leading to

$$\delta p^*_{n\mathrm{prop}} = \sqrt{(A_2)^{-1}_{nn}}. \tag{65}$$

Given the standard equation of fitting error,

$$\delta p^*_{n\mathrm{fit}} = \sqrt{\chi^2_{\mathrm{red}}\cdot(A_2)^{-1}_{nn}}, \tag{66}$$

with

$$\chi^2_{\mathrm{red}} = \frac{1}{N-M}A_0, \tag{67}$$

we finally conclude that

$$\delta p^*_{n\mathrm{fit}}/\delta p^*_{n\mathrm{prop}} = \sqrt{\chi^2_{\mathrm{red}}}, \text{ for all } n:\ 1,\ \ldots,\ M. \tag{68}$$

Namely, the ratio of the fitting over the propagation errors is fixed for any of the M fitting parameters and provides (the square root of) the chi-square.

2.2. Nonlinearly Parametric Fitting

2.2.1. Nonlinearly Mono-Parameterized Fitting

We start with the mono-parametric model function $f(x;p)$, where the chi-square is given by

$$\chi^2(p) = \sum_{i=1}^{N}\left(\frac{y_i - f(x_i;p)}{\sigma_{yi}}\right)^2. \tag{69}$$

Minimization of the chi-square involves finding the optimal value $p = p^*$ that solves the equation

$$\sum_{i=1}^{N}\frac{y_i - f(x_i;p^*)}{\sigma_{yi}^{\,2}}\cdot\frac{\partial f(x_i;p^*)}{\partial p^*} = 0. \tag{70}$$

Applying the partial derivative $\partial/\partial y_j$ to Equation (70), we have

$$\begin{aligned}&\sum_{i=1}^{N}\frac{1}{\sigma_{yi}^{\,2}}\cdot\left\{\left[\delta_{ij} - \frac{\partial p^*}{\partial y_j}\cdot\frac{\partial f(x_i;p^*)}{\partial p^*}\right]\cdot\frac{\partial f(x_i;p^*)}{\partial p^*} + [y_i - f(x_i;p^*)]\cdot\frac{\partial p^*}{\partial y_j}\cdot\frac{\partial^2 f(x_i;p^*)}{\partial p^{*2}}\right\} = 0 \\ &\frac{1}{\sigma_{yj}^{\,2}}\cdot\frac{\partial f(x_j;p^*)}{\partial p^*} = \frac{\partial p^*}{\partial y_j}\cdot\sum_{i=1}^{N}\frac{1}{\sigma_{yi}^{\,2}}\cdot\left\{\left[\frac{\partial f(x_i;p^*)}{\partial p^*}\right]^2 - [y_i - f(x_i;p^*)]\cdot\frac{\partial^2 f(x_i;p^*)}{\partial p^{*2}}\right\},\end{aligned} \tag{71}$$

Thus, we obtain

$$\delta p^*_{\text{prop}} = \sqrt{\sum_{j=1}^{N} \left(\frac{\partial p^*}{\partial y_j} \right)^2 \sigma_{yj}^2} = \frac{\sqrt{\sum_{j=1}^{N} \frac{1}{\sigma_{yj}^2} \cdot \left[\frac{\partial f(x_j; p^*)}{\partial p^*} \right]^2}}{\sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2} \cdot \left\{ \left[\frac{\partial f(x_i; p^*)}{\partial p^*} \right]^2 - [y_i - f(x_i; p^*)] \cdot \frac{\partial^2 f(x_i; p^*)}{\partial p^{*2}} \right\}}. \quad (72)$$

The respective curvature coefficient is given by

$$A_2 = \frac{1}{2} \frac{\partial^2 \chi^2(p^*)}{\partial p^{*2}} = \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2} \left\{ \left[\frac{\partial f(x_i; p^*)}{\partial p^*} \right]^2 - [y_i - f(x_i; p^*)] \frac{\partial^2 f(x_i; p^*)}{\partial p^{*2}} \right\}, \quad (73)$$

which can be included in the expression of propagation error, as follows:

$$\delta p^*_{\text{prop}} = \sqrt{(A_2)^{-1} \cdot \left[1 + (A_2)^{-1} \sum_{i=1}^{N} \frac{y_i - f(x_i; p^*)}{\sigma_{yi}^2} \cdot \frac{\partial^2 f(x_i; p^*)}{\partial p^{*2}} \right]}, \quad (74)$$

or

$$\delta p^*_{\text{prop}} = \sqrt{(A_2)^{-1} + (A_2)^{-2} \sum_{i=1}^{N} \frac{y_i - f(x_i; p^*)}{\sigma_{yi}^2} \cdot \frac{\partial^2 f(x_i; p^*)}{\partial p^{*2}}}. \quad (75)$$

Then, we find that the relationship of the two errors, propagation and fitting, involves the reduced chi-square of the nonlinearly parameterized fitting, $\chi^2_{\text{red,NL}}$; however, the relationship is not universal:

$$\delta p^{*\,2}_{\text{prop}} = \delta p^{*\,2}_{\text{fit}} / \chi^2_{\text{red,NL}} + (A_2)^{-2} \sum_{i=1}^{N} \frac{y_i - f(x_i; p^*)}{\sigma_{yi}^2} \cdot \frac{\partial^2 f(x_i; p^*)}{\partial p^{*2}} \neq \delta p^{*\,2}_{\text{fit}} / \chi^2_{\text{red,NL}}. \quad (76)$$

Therefore, the chi-square cannot be expressed through the ratio of these errors; thus, there is no universal value of the best chi-square.

2.2.2. Nonlinearly Multi-Parameterized Fitting

The counterexample of the mono-parametric case is enough to show that the chi-square is a meaningless tool for characterizing the goodness of a nonlinearly parameterized fit. Nevertheless, we extend this example to the case of a multi-parametric model function $f\left(x; \{p_k\}_{k=1}^{M}\right)$.

Minimization of the chi-square,

$$\delta p^*_{\text{prop}} = \sqrt{(A_2)^{-1} + (A_2)^{-2} \sum_{i=1}^{N} \frac{y_i - f(x_i; p^*)}{\sigma_{yi}^2} \cdot \frac{\partial^2 f(x_i; p^*)}{\partial p^{*2}}}, \quad (77)$$

involves finding the optimal parameter values, $\{p_k = {p_k}^*\}_{k=1}^{M}$, that solve the $m = 1, \ldots, M$, canonical equations:

$$\left. \frac{\partial \chi^2\left(\{p_k\}_{k=1}^{M}\right)}{\partial p_m} \right|_{\{p_k = {p_k}^*\}_{k=1}^{M}} = -2 \sum_{i=1}^{N} \frac{y_i - f(x_i; \{{p_k}^*\}_{k=1}^{M})}{\sigma_{yi}^2} \cdot \frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^{M})}{\partial p_m^*} = 0. \quad (78)$$

Applying the derivative, $\partial/\partial p_n$, once more on the chi-square gives the curvature matrix

$$(A_2)_{mn} = \sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot \left\{ \begin{array}{l} \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*} \cdot \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_n}^*} \\ -\left[y_i - f(x_i; \{{p_k}^*\}_{k=1}^M)\right] \cdot \dfrac{\partial^2 f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^* \partial {p_n}^*} \end{array} \right\}, \tag{79}$$

while the derivative $\partial/\partial y_j$ on the optimal parameter values, as shown in Equation (78), gives

$$\sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot \left\{ \begin{array}{l} \left[\delta_{ij} - \sum\limits_{k=1}^{M} \dfrac{\partial {p_k}^*}{\partial y_j} \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_k}^*}\right] \cdot \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*} \\ +\left[y_i - f(x_i; \{{p_k}^*\}_{k=1}^M)\right] \cdot \sum\limits_{k=1}^{M} \dfrac{\partial {p_k}^*}{\partial y_j} \dfrac{\partial^2 f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_k}^* \partial {p_m}^*} \end{array} \right\} = 0, \tag{80}$$

or

$$\begin{aligned} \frac{1}{{\sigma_{yj}}^2} \cdot \frac{\partial f(x_j; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*} &= \sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot \left\{ \begin{array}{l} \left[\sum\limits_{l=1}^{M} \dfrac{\partial {p_l}^*}{\partial y_j} \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_l}^*}\right] \cdot \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*} \\ -\left[y_i - f(x_i; \{{p_k}^*\}_{k=1}^M)\right] \cdot \sum\limits_{l=1}^{M} \dfrac{\partial {p_l}^*}{\partial y_j} \dfrac{\partial^2 f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_l}^* \partial {p_m}^*} \end{array} \right\} \\ &= \sum_{l=1}^{M} \frac{\partial {p_l}^*}{\partial y_j} \cdot \sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot \left\{ \begin{array}{l} \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_l}^*} \cdot \dfrac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*} \\ -\left[y_i - f(x_i; \{{p_k}^*\}_{k=1}^M)\right] \cdot \dfrac{\partial^2 f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_l}^* \partial {p_m}^*} \end{array} \right\} = \sum_{l=1}^{M} \frac{\partial {p_l}^*}{\partial y_j} \cdot (A_2)_{lm}, \end{aligned} \tag{81}$$

where the curvature matrix is given by Equation (79). Then, inversing Equation (81), we obtain

$$\frac{\partial {p_n}^*}{\partial y_j} = \frac{1}{{\sigma_{yj}}^2} \cdot \sum_{m=1}^{M} \left[\frac{\partial f(x_j; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*}\right] \cdot (A_2)^{-1}_{mn}. \tag{82}$$

Then, we derive the propagation error, as follows:

$$\delta {p_{n\text{prop}}^*}^2 = \sum_{i=1}^{N} \left(\frac{\partial {p_n}^*}{\partial y_i}\right)^2 {\sigma_{yi}}^2 = \sum_{i=1}^{N} \left\{ \frac{1}{{\sigma_{yi}}^2} \cdot \left\{ \sum_{m=1}^{M} \left[\frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*}\right] \cdot (A_2)^{-1}_{mn} \right\}^2 \right\}. \tag{83}$$

After some calculations, we have

$$\begin{aligned} &\left\{ \sum_{m=1}^{M} \left[\frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*}\right] \cdot (A_2)^{-1}_{mn} \right\}^2 = \\ &\sum_{m=1}^{M} \sum_{m'=1}^{M} \frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*} \frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_{m'}}^*} \cdot (A_2)^{-1}_{mn} (A_2)^{-1}_{m'n}, \end{aligned} \tag{84}$$

and

$$\begin{aligned} &\sum_{i=1}^{N} \left\{ \frac{1}{{\sigma_{yi}}^2} \cdot \left\{ \sum_{m=1}^{M} \left[\frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*}\right] \cdot [(A_2)^{-1}_{mn}] \right\}^2 \right\} = \\ &\sum_{m=1}^{M} \sum_{m'=1}^{M} \left[\sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot \frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_m}^*} \frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^M)}{\partial {p_{m'}}^*}\right] \cdot (A_2)^{-1}_{mn} (A_2)^{-1}_{m'n}. \end{aligned} \tag{85}$$

Hence,

$$\delta {p_{n\text{prop}}^*}^2 = \sum_{i=1}^{N} \left(\frac{\partial {p_n}^*}{\partial y_i} \right)^2 {\sigma_{yi}}^2 = \sum_{m=1}^{M} \sum_{m'=1}^{M} \left[\sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot \frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^{M})}{\partial {p_m}^*} \frac{\partial f(x_i; \{{p_k}^*\}_{k=1}^{M})}{\partial {p_{m'}}^*} \right] \cdot (A_2)_{mn}^{-1} (A_2)_{m'n}^{-1}, \tag{86}$$

or

$$\delta {p_{n\text{prop}}^*}^2 = \sum_{m=1}^{M} \sum_{m'=1}^{M} \left\{ \begin{array}{l} (A_2)_{mm'} + \\ \sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot [y_i - f(x_i; \{{p_k}^*\}_{k=1}^{M})] \cdot \frac{\partial^2 f(x_i; \{{p_k}^*\}_{k=1}^{M})}{\partial {p_m}^* \partial {p_{m'}}^*} \end{array} \right\} \cdot (A_2)_{mn}^{-1} (A_2)_{m'n}^{-1}$$
$$= (A_2)_{nn}^{-1} + \sum_{m=1}^{M} \sum_{m'=1}^{M} \left\{ \sum_{i=1}^{N} \frac{1}{{\sigma_{yi}}^2} \cdot [y_i - f(x_i; \{{p_k}^*\}_{k=1}^{M})] \cdot \frac{\partial^2 f(x_i; \{{p_k}^*\}_{k=1}^{M})}{\partial {p_m}^* \partial {p_{m'}}^*} \right\} (A_2)_{mn}^{-1} (A_2)_{m'n}^{-1}. \tag{87}$$

Therefore,

$$\delta {p_{n\text{prop}}^*}^2 = \delta {p_{n\text{fit}}^*}^2 / \chi^2_{\text{red,NL}} + \sum_{m=1}^{M} \sum_{m'=1}^{M} (A_2)_{mn}^{-1} (A_2)_{m'n}^{-1} \sum_{i=1}^{N} \frac{y_i - f(x_i; \{{p_k}^*\}_{k=1}^{M})}{{\sigma_{yi}}^2} \cdot \frac{\partial^2 f(x_i; \{{p_k}^*\}_{k=1}^{M})}{\partial {p_m}^* \partial {p_{m'}}^*}, \tag{88}$$

or

$$\delta {p_{n\text{prop}}^*}^2 \neq \delta {p_{n\text{fit}}^*}^2 / \chi^2_{\text{red,NL}}. \tag{89}$$

Namely, in any of the general cases of a fitting with multi-parametric model, the chi-square is not given by the ratio of the respective errors (squares); neither has a universal value, because the difference $\delta {p_{n\text{prop}}^*}^2 - \delta {p_{n\text{fit}}^*}^2 / \chi^2_{\text{red,NL}}$ differs from case to case depending on the particular fitting model.

3. Discussion: Application in the Outer Heliosphere

We use the reduced chi-square to characterize the goodness of the fitting involved in the method that determines the thermodynamic parameters of the inner heliosheath, that is, the distant outer shelf of our heliosphere. In addition, we use these datasets to compare the two measures related to the estimated chi-square, the reduced chi-square, and the respective *p*-value.

The Interstellar Boundary Explorer (IBEX) mission observes energetic neutral atom (ENA) emissions, constructing images of the whole sky every six months [15]. The persistence of these maps and its origins has been studied in detail [16,17]. Using these datasets, Livadiotis et al. [10] derived the sky maps of the temperature T, density n, and other thermodynamic quantities that characterize the plasma protons in the inner heliosheath (see also: [18–20]). These authors also showed the negative correlation between the values of temperature and density, which originates from the thermodynamic processes of the plasma.

The sky maps of temperature and density constitute 1800 pixels, each with angular dimensions of $6^0 \times 6^0$, while each map's grid in longitude and latitude is separated in 60×30 pixels. The method used by [11] derives eight (statistically independent) measurements of the logarithms of the temperature and density of each pixel. This is based on fitting kappa distributions on the plasma protons, that is, the distribution function of particle velocities linked to the thermodynamics of space plasma particle populations [21–30]. Once these eight values (temperature and density) are well-fitted by a constant, then, the respective pixel is assigned by this thermodynamic parameter. Figure 2 shows two examples of these eight measurements of temperature and density values. As is shown in Section 2, the

fitting parameter is a constant whose constant is the weighted mean. Both types of errors are plotted together with their weighted mean.

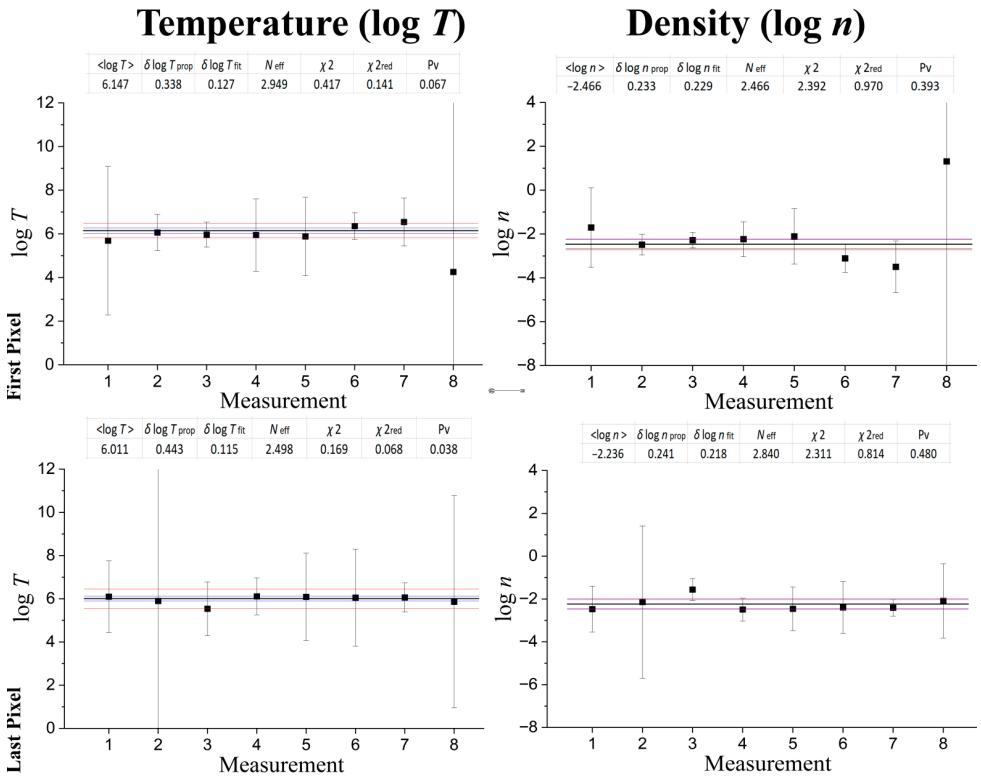


Figure 2. Two pixels' examples of eight measurements of temperature and density values that characterize the sky map of the plasma protons in the inner heliosheath. The weighted average (black) is plotted together with the ± deviation caused by fitting errors (blue) and propagation errors (red). (Tables show statistical details for each fit.).

The two errors were calculated as follows:

$$\delta p^*_{\text{prop}} \cong \sum_{i=1}^{N} \left[p^*(y_i + \sigma_{yi}) - p^*(y_i) \right]^2, \tag{90}$$

and

$$\delta p^*_{\text{prop}} \cong \varepsilon^{-1} \cdot \sqrt{\sum_{i=1}^{N} \left[p^*(y_i + \varepsilon \cdot \sigma_{yi}) - p^*(y_i) \right]^2} \tag{91}$$

for some small value of ε (as this tends to zero, the value of the resultant errors becomes independent of ε).

We repeat the above fitting procedure for eight neighboring pixels in the sky (taken from the annual maps of year 2010; see details in [10]). Figure 3a shows these neighboring pixels in the sky maps of temperature and density. Then, in Figure 3b we plot the derived temperature, density, and respective uncertainties (corresponding to the larger among the two error types). Finally, in Figure 3c we show their mutual relationship, by plotting the density against the temperature (logarithmic values). This is described by a negative correlation relationship, originated by a polytropic process with sub-isothermal polytropic index γ. A polytropic process is described by a power-law relation between temperature and density (i.e., $T \propto n^{\gamma-1}$), or between thermal pressure and density (i.e., $p \propto n^{\gamma}$), along a streamline of the flow. In particular, we find a slope $\nu = 1/(\gamma - 1)$ (that is, the exponent in $n \propto T^{\nu}$) around $\nu \sim -2$, but it is characterized with better statistics when outlining the last point, leading to $\nu \sim -1$ (also, see: [10,30]).

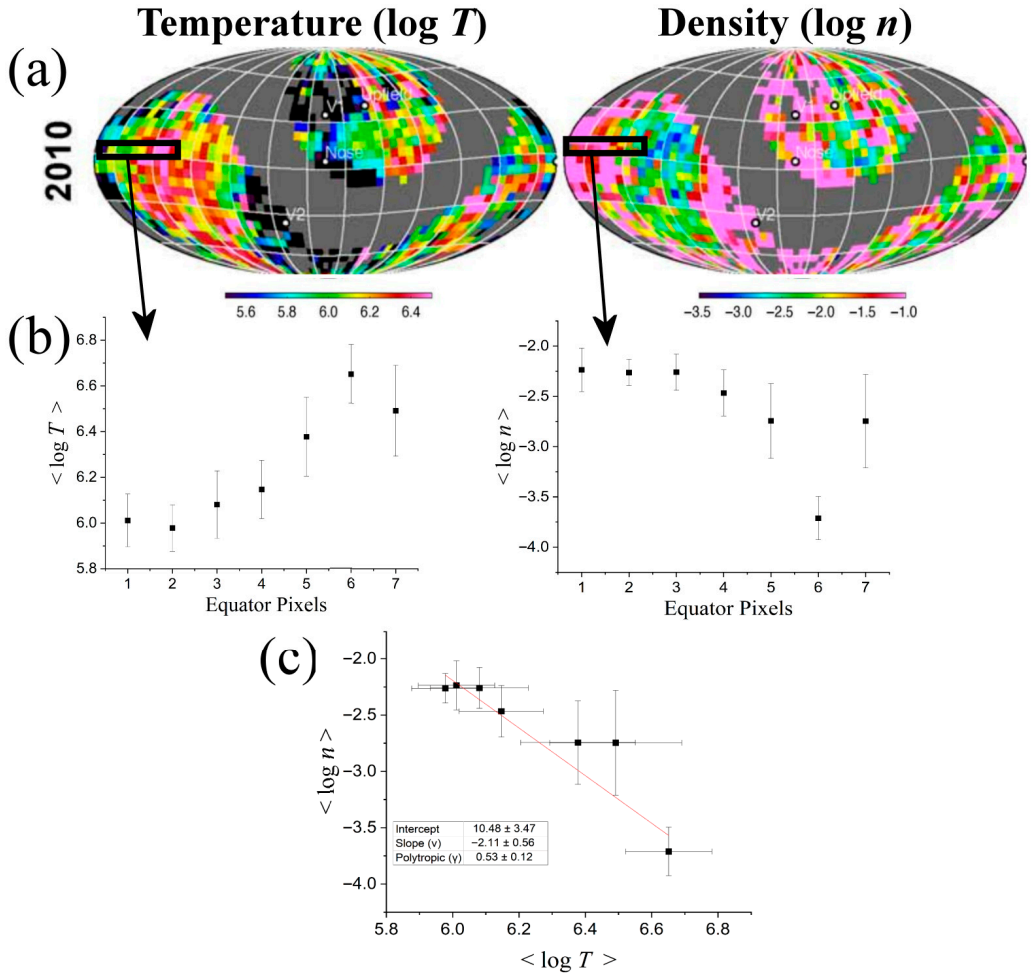


Figure 3. (**a**) The eight examined neighboring pixels in the sky maps of temperature and density. (**b**) The results of weighted mean and the uncertainty (largest between fitting and propagation errors). (**c**) Co-plot of density vs. temperature shows its negative correlation. (Temperature and density are represented with logarithms).

Using the p-value method for evaluating the goodness of fitting [13], we compare the estimated chi-square value, $\chi^2_{\rm est}$, with any possible chi-square value larger than the estimated one, $\chi^2_{\rm est} \leq \chi^2 < \infty$, by deriving this exact probability, $P(\chi^2_{\rm est} \leq \chi^2 < \infty)$; this determined the p-value. This is determined by integrating the chi-square distribution, $P(\chi^2; D) = \frac{2^{-\frac{D}{2}}}{\Gamma(\frac{D}{2})} e^{-\frac{1}{2}\chi^2} (\chi^2)^{\frac{D}{2}-1}$, that is the distribution of all the possible χ^2 values (parameterized by the degrees of freedom D). Therefore, the likelihood of having a chi-square value, χ^2, larger than the estimated one, $\chi^2_{\rm est}$, is given in terms of the complementary cumulative chi-square distribution. The respective probability, the p-value, equals $P(\chi^2_{\rm est} \leq \chi^2 < \infty) = \int_{\chi^2_{\rm est}}^{\infty} P(\chi^2; D)\, d\chi^2$.

The larger the p-value, the better the goodness of the fitting. A p-value larger than 0.5 corresponds to $\chi^2_{\rm est} < D$ or $\chi^2_{\rm red} < 1$. Larger p-values, up to $p = 1$, correspond to smaller chi-squares, down to $\chi^2_{\rm red} \sim 0$. Thus, an increasing p-value above the threshold of 0.5 cannot lead to a better fitting. Rather, it leads to a worse fit, similar to a decreasing $\chi^2_{\rm red} < 1$. On the other hand, a p-value smaller than the significance level of ~0.05 is typically rejected. Using these limits of the p-value, we find the respective limits of the reduced chi-square.

In Figure 4, we demonstrate the relationship between the reduced chi-square $\chi^2_{\rm red}$ and the respective p-value $\rm P_v$. We plot 16 pair values of ($\chi^2_{\rm red}$ and $\rm P_v$), one pair for each of the 16 fitting measures performed (temperature and density fits, both for eight pixels). The deduced relationship between the two measures can be used to deduct the extreme

acceptable values of the reduced chi-square; indeed, the best reduced chi-square value that characterizes the goodness of a fitting is one, but at what extreme should a lower goodness be accepted? The correspondence with the acceptable value of 0.05 of the *p*-value will provide an answer to this issue.

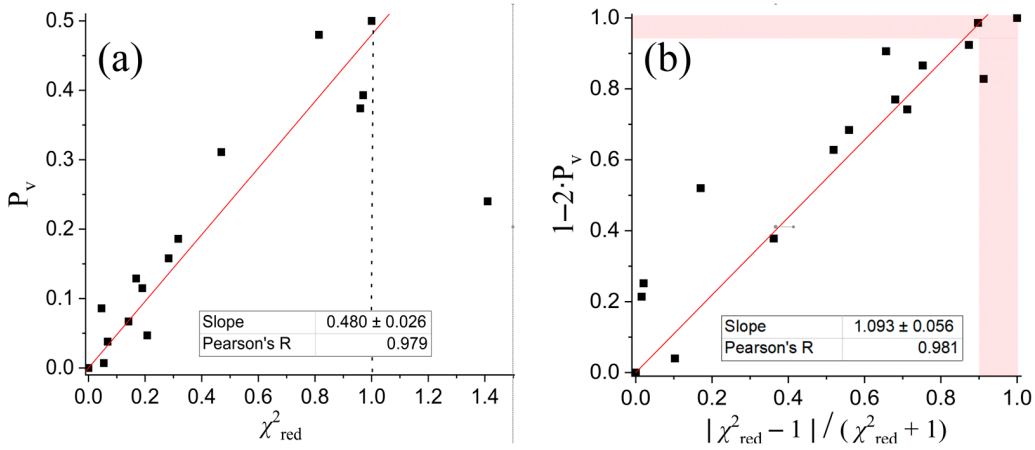


Figure 4. (**a**) Co-plots of the *p*-values and the respective reduced chi-square values, showing their relationship. (**b**) Similar to (a) but with modified functions ranging in [0, 1]; (shaded area corresponds to rejected fits because of low goodness).

In Figure 4a we plot all (χ^2_{red} and P_v) pair values and observe the linear behavior between the two measures. The plot shows the well-described linearity between the low chi-square values $\chi^2_{red} \leq 1$ and the corresponding *p*-values. Nevertheless, in order to also catch the relationship for all the points, including those with $\chi^2_{red} > 1$, in Figure 4b we plot the values of $1–2\cdot P_v$ as a function of the modified chi-square values, that is, $|\chi^2_{red} - 1|/(\chi^2_{red} + 1)$, which surely ranges in the interval [0, 1] (similarly for both axes). We find that the excluded *p*-values ($P_v < 0.05$) correspond to the reduced chi-square values $\chi^2_{red} < 0.1$ and $\chi^2_{red} > 10$; (of course, this is not a universal value, but corresponds to the examined datasets).

4. Conclusions

The paper showed the relationship between the two types of errors that characterize the optimal fitting parameters, the fitting and propagation errors. While the fitting error estimates the uncertainty originated from the variation in the observational data around the model, the propagation error comes from the raw uncertainties of the observational data. The (square) ratio of the two errors determines the value of the reduced chi-square.

Large deviations between propagation and fitting errors lead to a reduced chi-square significantly different than one, corresponding to low fitting goodness. In fact, large uncertainties lead to small values of the reduced chi-square ($\chi^2_{red} < 1$), indicating a possible error overestimation; small uncertainties lead to a large value of the reduced chi-square ($\chi^2_{red} > 1$), indicating that there is an error underestimation of the observation data. The optimal fitting goodness would be for observation uncertainties similar to the data variation, namely, the two types of errors are about the same, and the reduced chi-square is about one ($\chi^2_{red} \sim 1$).

The relationship between the two types of errors and the chi-square was shown for mono-parametric, bi-parametric, and in general, multi-parametric statistical models. In the multi-parametric case, it is interesting that the ratio is universal, namely, independent of the respective parameter; namely, $\delta p^*_{n\text{fit}}/\delta p^*_{n\text{prop}} = \sqrt{\chi^2_{red}}$, for any n: 1, …, M, as shown in Equation (68).

In addition to this theoretical development, we presented a counterexample of this characterization of chi-square, showing that it is not generally true for a nonlinearly parameterized fitting. Some alternative approaches for assessing nonlinear fits involve the bootstrap inference [31], the Bayesian frameworks [32,33], and the technique of averaging the parameter's optimal values derived from solving by pairs [11].

Finally, the method of chi-square characterization as a measure of the fitting goodness is applied on the datasets of temperature and density of the plasma protons in the inner heliosheath, the outer shelf of our heliosphere. A previously applied method derived eight measurements of the logarithms of temperature and density for each pixel of the sky map [10]. Once these eight values (temperature and density) are well-fitted by a constant, then, the respective pixel is assigned by this thermodynamic parameter. We have determined the two types of errors of this analysis; the respective reduced chi-square values were compared with the respective p-values for completeness.

Funding: This research was funded by the interstellar Boundary Explorer (IBEX) mission (80NSSC20K0719), which is a part of NASA's Explorer Program.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: No new data were created by this paper.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Cochran, W.G. The Chi-square Test of Goodness of Fit. *Ann. Math. Stat.* **1952**, *23*, 315–345. [CrossRef]
2. McCullagh, P. What is statistical model? *Ann. Stat.* **2002**, *30*, 1225–1310. [CrossRef]
3. Adèr, H.J. Modelling. In *Advising on Research Methods: A Consultant's Companion*; Johannes van Kessel Publishing: Huizen, The Netherlands, 2008; Chapter 12; pp. 271–304.
4. Burden, R.L.; Faires, J.D. *Numerical Analysis*; PWS Publishing Company: Boston, MA, USA, 1993; pp. 437–438.
5. Livadiotis, G. Approach to general fitting methods and their sensitivity. *Physica A* **2007**, *375*, 518. [CrossRef]
6. Livadiotis, G. Chi-p distribution: Characterization of the goodness of the fitting using L p norms. *Physica A* **2007**, *1*, 4. [CrossRef]
7. Frisch, P.C.; Bzowski, M.; Livadiotis, G.; McComas, D.J.; Moebius, E.; Mueller, H.-R.; Pryor, W.R.; Schwadron, N.A.; Sokół, J.M.; Vallerga, J.V.; et al. Decades-long changes of the interstellar wind through our solar system. *Science* **2013**, *341*, 1080. [CrossRef] [PubMed]
8. Shih, J.H.; Fay, M.P. Pearson's Chi-square Test and Rank Correlation Inferences for Clustered Data. *Biometrics* **2017**, *73*, 822–834. [CrossRef] [PubMed]
9. Bevington, P.R. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw-Hill: New York, NY, USA, 1969; p. 336.
10. Livadiotis, G.; McComas, D.J.; Funsten, H.O.; Schwadron, N.A.; Szalay, J.R.; Zirnstein, E. Thermodynamics of the Inner Heliosheath. *Astrophys. J. Suppl. Ser.* **2022**, *262*, 53. [CrossRef]
11. Livadiotis, G.; Cummings, A.T.; Cuesta, M.E.; Bandyopadhyay, R.; Farooki, H.A.; Khoo, L.Y.; McComas, D.J.; Rankin, J.S.; Sharma, T.; Shen, M.M.; et al. Kappa-tail Technique: Modeling and Application to Solar Energetic Particles Observed by Parker Solar Probe. *Astrophys. J. Suppl. Ser.* **2024**, *973*, 6. [CrossRef]
12. Cuesta, M.E.; Cummings, A.T.; Livadiotis, G.; McComas, D.J.; Cohen, C.M.S.; Khoo, L.Y.; Sharma, T.; Shen, M.M.; Bandyopadhyay, R.; Rankin, J.S.; et al. Observations of Kappa Distributions in Solar Energetic Protons and Derived Thermodynamic Properties. *Astrophys. J. Suppl. Ser.* **2024**, *973*, 76. [CrossRef]
13. Fisher, R.A. On the Interpretation of χ^2 from Contingency Tables and the Calculation of p. *J. R. Stat. Soc.* **1922**, *85*, 87–94. [CrossRef]
14. Melissinos, A.C. *Experiments in Modern Physics*; Academic: London, UK, 1966; p. 438.
15. McComas, D.J.; Allegrini, F.; Bochsler, P.; Bzowski, M.; Christian, E.R.; Crew, G.B.; DeMajistre, R.; Fahr, H.; Fichtner, H.; Frisch, P.C.; et al. Global Observations of the Interstellar Interaction from the Interstellar Boundary Explorer (IBEX). *Science* **2009**, *326*, 959. [CrossRef]
16. Sarlis, N.V.; Livadiotis, G.; McComas, D.J.; Alimaganbetov, M.; Schwadron, N.A.; Fairchild, K. Persistent Behavior in Energetic Neutral Atom Time Series from IBEX. *Astrophys. J. Suppl. Ser.* **2024**, *976*, 45. [CrossRef]
17. Sarlis, N.V.; Livadiotis, G.; McComas, D.J.; Alimaganbetov, M.; Schwadron, N.A. Origins of Persistence in Energetic Neutral Atom Time Series from IBEX. *Astrophys. J. Suppl. Ser.* **2025**, *986*, 131. [CrossRef]

18. Livadiotis, G.; McComas, D.J.; Dayeh, M.A.; Funsten, H.O.; Schwadron, N.A. First Sky Map of the Inner Heliosheath Temperature Using IBEX Spectra. *Astrophys. J. Suppl. Ser.* **2011**, *734*, 1. [CrossRef]
19. Livadiotis, G.; McComas, D.J.; Randol, B.M.; Funsten, H.O.; Möbius, E.S.; Schwadron, N.A.; Dayeh, M.A.; Zank, G.P.; Frisch, P.C. Pick-up Ion Distributions and Their Influence on Energetic Neutral Atom Spectral Curvature. *Astrophys. J. Suppl. Ser.* **2012**, *751*, 64. [CrossRef]
20. Livadiotis, G.; McComas, D.J.; Schwadron, N.A.; Funsten, H.O.; Fuselier, S.A. Pressure of the Proton Plasma in the Inner Heliosheath. *Astrophys. J. Suppl. Ser.* **2013**, *762*, 134. [CrossRef]
21. Livadiotis, G.; McComas, D.J. Beyond kappa distributions: Exploiting Tsallis statistical mechanics in space plasmas. *J. Geophys. Res.* **2009**, *114*, A11105. [CrossRef]
22. Livadiotis, G. "Lagrangian Temperature": Derivation and Physical Meaning for Systems Described by Kappa Distributions. *Entropy* **2014**, *16*, 4290–4308. [CrossRef]
23. Livadiotis, G. *Kappa Distributions: Theory and Applications in Plasmas*; Elsevier: Amsterdam, The Netherlands, 2017; 738p.
24. Nicolaou, G.; Livadiotis, G.; Owen, C.J.; Verscharen, D.; Wicks, R.T. Determining the kappa distributions of space plasmas from observations in a limited energy range. *Astrophys. J.* **2018**, *864*, 3. [CrossRef]
25. Livadiotis, G. Thermodynamic origin of kappa distributions. *Europhys. Lett.* **2018**, *122*, 50001. [CrossRef]
26. Nicolaou, G.; Livadiotis, G. Long-term correlations of polytropic indices with kappa distributions in solar wind plasma near 1 au. *Astrophys. J.* **2019**, *884*, 52. [CrossRef]
27. Livadiotis, G.; McComas, D.J. Physical Correlations Lead to Kappa Distributions. *Astrophys. J.* **2022**, *940*, 83. [CrossRef]
28. Livadiotis, G.; McComas, D.J. Entropy defect in thermodynamics. *Sci. Rep.* **2023**, *13*, 9033. [CrossRef] [PubMed]
29. Livadiotis, G.; McComas, D.J. Thermodynamic Definitions of Temperature and Kappa and Introduction of the Entropy Defect. *Entropy* **2021**, *23*, 1683. [CrossRef]
30. Livadiotis, G. Superposition of Polytropes in the Inner Heliosheath. *Astrophys. J. Suppl. Ser.* **2016**, *223*, 13. [CrossRef]
31. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*; Chapman & Hall/CRC: Boca Raton, FL, USA, 1993.
32. Carlin, B.P.; Louis, T.A. *Bayesian Methods for Data Analysis*, 3rd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2008.
33. McElreath, R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2020.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.