Journal of Structural Geology 33 (2011) 1000–1009

Contents lists available at ScienceDirect

Journal of Structural Geology

journal homepage: www.elsevier.com/locate/jsg

# A statistical examination of the Fry method of strain analysis

Yehua Shan [a,b,*], Wenjiao Xiao [c]

[a] Key Laboratory of Marginal Sea Geology, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou City 510640, PR China
[b] Xinjiang Research Centre for Mineral Resources, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, PR China
[c] State Key Laboratory of Lithospheric Evolution, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing City 100029, PR China

ARTICLE INFO

*Article history:*
Received 16 June 2009
Received in revised form
9 February 2011
Accepted 17 February 2011
Available online 24 February 2011

*Keywords:*
The Fry method
Strain analysis
Statistics
Probability density function
Likelihood function
Grid search

ABSTRACT

The statistics of the Fry method of strain analysis are examined in this paper. Point objects in the pre-deformed state are assumed in the Poisson distribution that is truncated at the maximum probability density function (PDF) of a certain point object. The mean log likelihood function (MLLF) is defined as the average sum of the log PDF of each individual point in the deformed state, and is maximized by use of a grid search to solve for unknown parameters, strain and cutoff radius. In order to demonstrate the feasibility of the new strain method, it is applied to artificial sets of point objects generated at the prescribed parameters. Results deliver a very high accuracy of the strain estimate even for artificial sets with many less packed point objects. Increasing the number of point objects in the MLLF tends to give a higher accuracy of strain estimate, particularly for a smaller point number, less than 60 in this case. A deformed conglomerate is analyzed to illustrate the new approach. A probable spurious point in these data is identified using the detection method proposed in this paper. Its removal leads to a more robust estimate of the strain.



## 1. Introduction

Although similar to the classical center-to-center method (Ramsay, 1967) in dealing with the points or centers of deformed objects, the ingenious Fry method (Fry, 1979) is more flexible in taking the distribution of objects into account, and requires no knowledge of the nearest neighbour points. These strengths, as well as the fact that deformed objects exposed at outcrop and/or under the microscope are generally not good strain markers on account of their unknown shapes in the pre-deformed state, make the Fry method applicable as a simple and powerful tool to quantify homogeneous deformation recorded by the distribution of points in rocks of variable kinds (e.g., Hanna and Fry, 1979; González-Casado et al., 1983; Crespi, 1986; Onasch, 1986; Ghaleb and Fry, 1995; Genier and Epard, 2007), and even the spatial distribution of mineralization (Vearncombe and Vearncombe, 1999).

The Fry method was devised to graphically extract strain from object–object separations in deformed rocks (Fry, 1979). For point objects that possess an isotropic anti-clustered distribution and then undergo homogeneous deformation, an elliptical vacancy or void is surrounded by points near the center of the Fry plot, that is an analogue of a finite strain ellipse. Theoretical improvements mainly concentrate on the subjectivity and reproducibility of the strain ellipse estimated by use of the Fry method (Ailleres and Champenois, 1994). The uncertainty in defining the finite strain ellipse near the central vacancy is due to the deviation of real point distributions from Fry's (1979) assumption of an anti-clustered distribution. For objects of strongly variable sizes that appear less anti-clustered, Erslev (1988) developed a normalized Fry method that modified each object–object separation according to the shapes and sizes of the two neighbouring objects. Erslev (1988), Dunne et al. (1990) and McNaught (1994) examined how the way of defining object centers affects the estimated finite strain ellipse from a practical viewpoint. Erslev and Ge (1990) later developed a least-squares method to determine the best-fit strain ellipse from nearest neighbour points in the normalized Fry plot. Fry (1999) claimed that the use of the first or second summed moments in estimating strain from some or all parts of points in the Fry plots is unjustified. Waldron and Wallace (2007) attempted to determine the strain ellipse in the non-normalized or normalized Fry plot by locating the locus of the observed maximum point-density gradient that best fits the calculated point density within a central elliptical domain.

In practice, the distribution of points in the Fry plot is influenced by factors such as the distribution of point objects in the sampled area, the shape and the size of the sampled area, and deformation.

* Corresponding author. Key Laboratory of Marginal Sea Geology, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou City 510640, PR China. Tel.: +86 20 85292403; fax: +86 20 85290130.
*E-mail address:* shanyehua@yahoo.com.cn (Y. Shan).

0191-8141/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsg.2011.02.004

The second factor can be neglected only if a large circular sampled area is chosen. In the absence of the second factor, outside the central vacancy field, points tend to decrease their density outwards from the center, and to be evenly distributed in all directions (Fig. 1c). This is probably the simple reason that most structural geologists have chosen to extract strain only from the points near the center, using some improved Fry methods. Disregarding most points in the Fry plot in this way certainly weakens the claim of obtaining the strain from all separations between objects rather than those lying closest to the center, as Fry (1979,1999) believed. This makes apparent an unclear issue about whether or not all object separations in the Fry plot can be used for strain analysis.

In this communication, we examine some statistical aspects of the Fry method, including the randomness and the truncation of point distribution. The probability density function of distance between the $k$-th nearest neighbour points is then formulated, from which we define the likelihood function. The mean log likelihood function is then maximized using a grid search to solve for strain, which provides us with a new method of determining strain from point distribution. This method is validated and illustrated by applying it to artificial sets and a real set.

The symbols used in this paper are listed and defined in the Appendix.

## 2. Strain analysis using the Fry plot

It is necessary to rehearse the procedure for making the Fry plot (Fry, 1979; Hanna and Fry, 1979) before we make some theoretical investigation of it. Let us take a template made up of the positions of $n$ the centres of deformed objects (Fig. 1b). Move this template without any rotation to a new plot so as to locate one of the points at the origin, and then mark other points from the template onto the plot. This process is repeated until all the points in the template have been used up. The map of all marked points, $n(n-1)$ in number, is called the Fry plot (Fig. 1c).

In essence, the Fry plot is a realization of $n(n-1)$ points around a fixed point (origin) in a way peculiar to it of repeating the same random point pattern within the same domain but at differing locations. This peculiarity never violates the assumption of a statistically homogeneous or Poisson distribution of the point objects in the sampled area (Fry, 1979). The density of marked points in the Fry plot depends upon the cumulative occurrences of

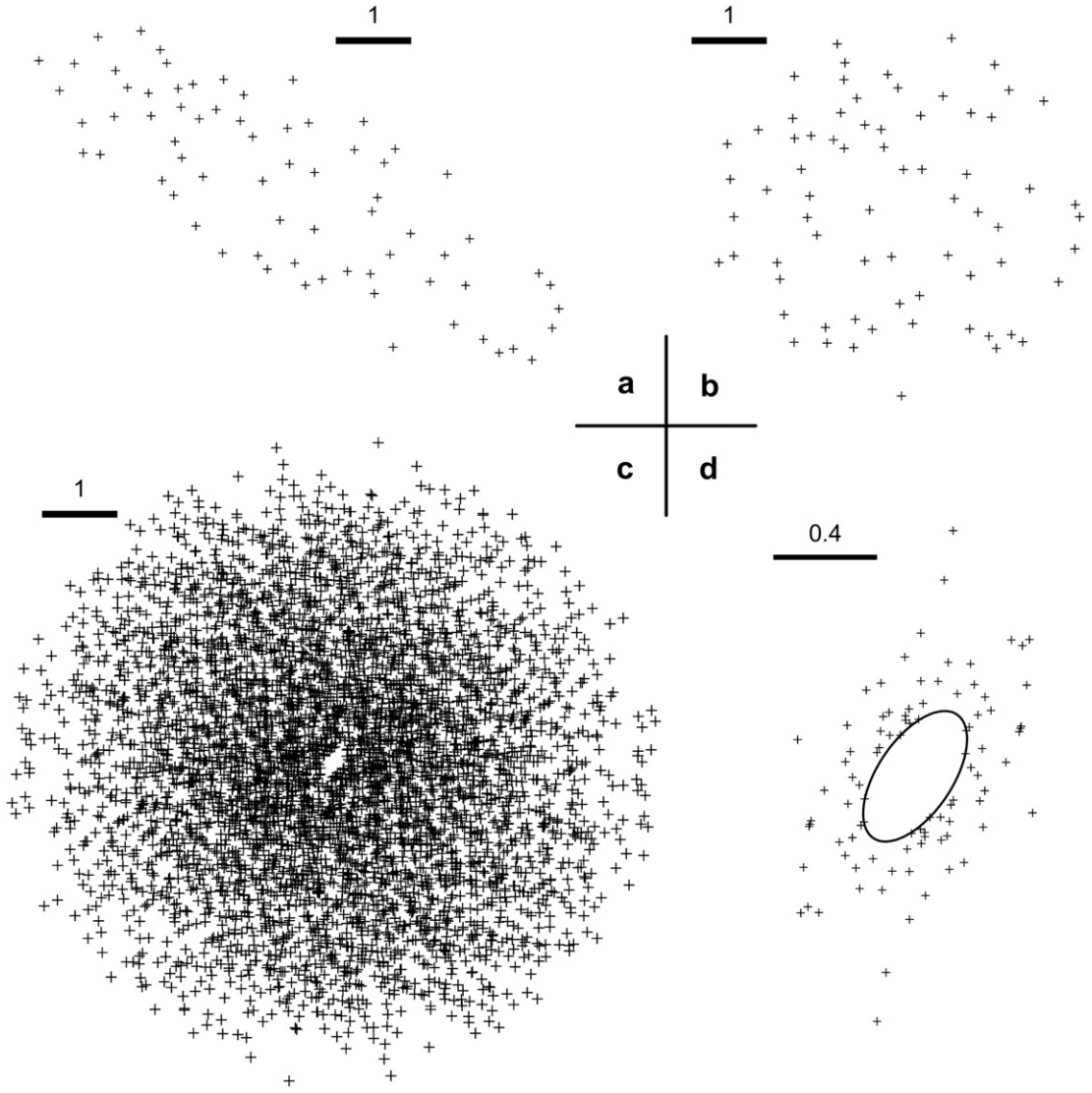


**Fig. 1.** An artificial set of 70 points in the pre-deformed state (a) and in the deformed state (b). Fry plot of the set in the deformed state (c), and the best fit elliptical void in the center of the plot (d). All points are marked by "plus". See Section 4.1 for information on data generation. The best fit elliptical void, as well as the first nearest neighbour points in the pre-deformed state is determined by using the proposed method in the case of one outcome ($k = 1$).

each repetition. A little away from the central vacancy, the tendency of outward decrease in density appears to tell us no more information on strain than the distribution of points in the area.

In fact, the determination of strain requires the assumption of a truncated Poisson distribution of points in the sampled area (Fry, 1979) of which the first or second or other near-neighbour points tend to define the boundary of an elliptical void near the center of the plot which can be used as a passive strain marker. The more anti-clustered the points are, the better defined elliptical void we get. However, point data available from rocks are generally not strictly anti-clustered. As a remedy, a variety of empirical methods have been developed, including the normalized Fry method (Erslev, 1988) and its derivatives such as the enhanced normalized Fry method (Erslev and Ge, 1990), the refined Fry method (Ailleres and Champenois, 1994) and the continuous function method (Waldron and Wallace, 2007). These methods try to enhance in different ways the definition of the central elliptical void in the Fry plot. Apart from strain, the Fry plot contains other information on point distribution, the pattern of points in the study area, for instance.

## 3. Statistics in the Fry method

The basic assumption of the Fry method of strain analysis (Fry, 1979) is a truncated Poisson distribution of point objects in the sampled area. In the method, randomness, truncation and homogeneity are therefore three requirements of the point distribution in the undeformed state. The first two requirements will be discussed in detail below. More advanced consideration of spatial point processes is needed to address the last requirement (see Lindsey, 1997; Daley and Vere-Jones, 2008; and others), and this is beyond the scope of this simple communication. For simplicity, the spatial homogeneity of point objects of interest is assumed.

### 3.1. Probability density function

Let us consider a two-dimensional Poisson process that has a number of $n$ outcomes or points, $\boldsymbol{X_i}$ ($i$ = 1, 2, …, $n$) within a bounded region $\boldsymbol{W}$ of space. For any circular subregion $\boldsymbol{V}$ ($V \in W$) with a radius of $r$, the probability density function (PDF) has the following expression,

$$P(k,\lambda,r) = \frac{e^{-\lambda\pi r^2}\left(\lambda\pi r^2\right)^k}{k!} \tag{1}$$

where $k$ is the number of points within the subregion $\boldsymbol{V}$, $\lambda$ is the mean density of points, and $k!$ is the factorial of integer $k$. By setting the differential PDF with $r$ to zero and solving this equation, the PDF reaches the maximum when:

$$\lambda\pi r^2 = k \tag{2}$$

By randomness, we mean that each individual outcome has the same possibility of occurring throughout the subregion, and therefore is independent of other outcomes. This independence of any two outcomes becomes invalid, particularly for those nearest neighbour outcome pairs, when a truncated Poisson process is taken into account. As required in the Fry method, the truncation is isotropic in the pre-deformed state, thus giving rise to an empty circular area for each individual outcome, within which any other outcome is prohibited (Fig. 2a). Accordingly, the PDF of the truncated Poisson distribution has the following expression,

$$P(k,\lambda,r,r_0) = \begin{cases} a(k,\lambda,r_0)P(k,\lambda,r) & r \geq r_0 \\ 0 & r < r_0 \end{cases} \tag{3}$$

$$a(k,\lambda,r_0) = \frac{\int_0^{+\infty} P(k,\lambda,r)2\pi r\mathrm{d}r}{\int_{r_0}^{+\infty} P(k,\lambda,r)2\pi r\mathrm{d}r} \tag{4}$$

where $r_0$ denotes the radius of the cutoff circular area, or cutoff radius.

As explicitly required by the truncation, each sample point becomes a circular point object with the same radius. The more

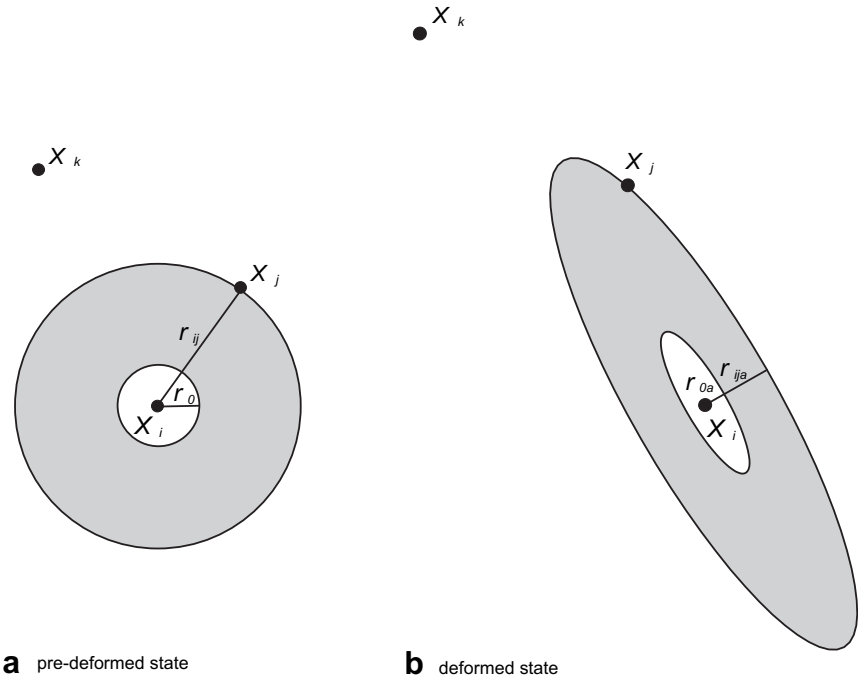


**Fig. 2.** Both the empty void and the subregion at point $X_i$ have a circular shape in the pre-deformed state (a) and an elliptical one in the deformed state (b), as required in the Fry method. This anisotropy is recorded by the distance between the $k$-th nearest neighbour points ($k$ = 1, 2, 3, …), which in turn makes it possible to extract strain from them in some way.

well-packed these point objects are, the more well-defined elliptical void in the central Fry plot we have. In the case of well-packed point objects, it is thus justifiable to consider the occurrence of the truncation at the maximum value of the PDF in formula (1) for $k = 1$, for simplicity.

$$\hat{\lambda}\pi r_0^2 = 1 \tag{5}$$

where $\hat{\lambda}$ is the intensity of well-packed point objects. In most cases,$\lambda$ does not equal to $\hat{\lambda}$. Their difference reflects how the point objects are packed.

However, other truncated Poisson distributions may be employed to describe the packed points. They are beyond the scope of this paper.

### 3.2. Strain analysis

In the Fry method, the isotropically truncated distribution corresponds to the pre-deformed state where marked points in the Fry plot tend to define a circular void near the center. When homogeneous deformation takes place in the region of interest, the isotropic truncation is changed to the anisotropic truncation that permits the use of the elliptical central void as a passive strain marker (Fig. 2). For the sake of strain analysis, the study point objects are assumed well packed. Their PDF in the deformed state and the PDF in the pre-deformed state are mutually related in the following expressions:

$$P(k, \hat{\lambda}, r_{ia}, r_{0a}, \theta, Rs) = P(k, \hat{\lambda}, r_i, r_0) \tag{6}$$

$$r_i = r_{ia}\sqrt{Rs} \tag{7}$$

$$r_0 = r_{0a}\sqrt{Rs} \tag{8}$$

where $r_{0a}$ is the minor axial length of the elliptical void near the center, $r_{ia}$ the minor axial length of the elliptical subregion of interest, $\theta$ the direction of the major strain axis, and $Rs$ strain ratio. $\hat{\lambda}$ is calculated from $r_{0a}$, according to Eqs. (5) and (8).

For the purpose of strain analysis, the unknown, independent parameters in formula (6) are $r_{0a}$, $\theta$ and $Rs$. They are estimated below in a general way by maximization of the likelihood function. For example, in the case of formula (4), the likelihood function $PL(k, \hat{\lambda}, r_{0a}, \theta, Rs)$ is formulated as:

$$PL(k, \hat{\lambda}, r_{0a}, \theta, Rs) = \prod_{i=1}^{n} P(k, \hat{\lambda}, r_{ia}, r_{0a}, \theta, Rs) \tag{9}$$

In practice, instead of formula (9), the mean log likelihood function is adopted that permits the comparison in value between any two PDFs having different formulas,

$$\frac{\log(PL(k, \hat{\lambda}, r_{0a}, \theta, Rs))}{n} = \frac{1}{n}\sum_{i=1}^{n} \log(P(k, \hat{\lambda}, r_{ia}, r_{0a}, \theta, Rs)) \tag{10}$$

By virtue of the complexity of formula (10), a grid search is used in this paper to solve the maximization of the equation for the best estimators of $r_{0a}$, $\theta$ and $Rs$. This simple algorithm is time consuming, but competent in searching for the global solution at a certain prescribed resolution. Much about grid search will be described in the next section.

We wrote a MS Fortran 5.0 program for finding the best strain estimate by maximising the mean log likelihood function by a grid search strategy. The code is available to interested readers on request.

### 3.3. Randomness

The pattern of $n$ outcomes in the region $\boldsymbol{W}$ has a direct impact on how the nearest neighbour points define the elliptical void in the center of the Fry plot, as mentioned above. The spatial randomness of these outcomes may be evaluated by using either first-order measures (e.g. Clark and Evans, 1954) or second-order measures (e.g. Ripley, 1976).

However, in this paper, the degree of packing of a given set of point objects is qualitatively described by intensity ratio, the ratio of $\hat{\lambda}$ to $\lambda$. When the intensity ratio approaches 1, the point objects tend to be well packed.

In the frame of well packing, the randomness of point objects is tested using the Kolmogorov–Smirnov (K–S) test statistic (see Gibbons, 1985),

$$D(n) = \max|F_{cal}(k, \hat{\lambda}, r_{ia}, r_{0a}, \theta, Rs) - F_{obs}(k, \hat{\lambda}, r_{ia}, r_{0a}, \theta, Rs)| \tag{11}$$

$$F_{cal}(k, \hat{\lambda}, r_{ia}, r_{0a}, \theta, Rs) = \frac{\int_{r_{0a}\sqrt{Rs}}^{r_{ia}\sqrt{Rs}} P(k, \hat{\lambda}, r)2\pi r dr}{\int_{r_{0a}\sqrt{Rs}}^{+\infty} P(k, \hat{\lambda}, r)2\pi r dr} \tag{12}$$

where $F_{cal}$ and $F_{obs}$ are the calculated and the observed cumulative distribution frequencies of the $k$-th nearest neighbour in the retro-deformed state at a distance of $r_{ia}\sqrt{Rs}$. With a confidence $\alpha$, the study point objects are random if the calculated statistic is no more than$D_\alpha(n)$. Otherwise, they are either clustered or anti-clustered.

## 4. Tests using artificial examples

In this section, artificial sets of random points generated by the prescribed parameters are taken as examples to demonstrate the feasibility of the new method, and to examine what factors the accuracy of the strain estimate depends upon.

### 4.1. Data generation

The procedure for generating sets of point objects at random is summarized below.

1. Input the values of the prescribed parameters such as the radius of the cutoff circular area ($r_0$), the direction of the major strain axis ($\theta$), and strain ratio ($Rs$). They are 0.02, 60° and 3.0, respectively, in this case.
2. Generate a random point within a 10 × 10 square region, and then look for its minimum nearest neighbour distance to other saved points. Save it if the distance is no less than $r_0$. Otherwise, repeat this step.
3. Displace the new saved point using the prescribed strain.
4. Accept the new saved point if its distance to the center of the square in the deformed state is no more than 5; otherwise, return to step 2. For tens of saved points, calculate strain from them using the method proposed in this paper. (The circular shape of the re-sampled area in the deformed state ensures the isotropy of data sampling, in hope of the directional spread of the nearest neighbour points in the Fry plot, as mentioned in the introduction.)

### 4.2. Grid search

As state previously, an exhaustive grid search scheme is adopted in this paper to look for the best estimates of cutoff radius ($r_0$), the direction of the major strain axis ($\theta$) and the strain ratio ($Rs$) through maximizing the mean log likelihood function. In the parameter space, these parameters are discretized in the following way:

1. $\theta$ is taken as an integer degree of 0–179°.
2. $Rs$ is taken as a real of 1.0–20.0 with a spacing of 0.1. The upper limit of $Rs$ may be set to a larger or smaller value, depending upon the extent of strain ratio. 20.0 is sufficiently large in the cases of this paper.
3. Once both $\theta$ and $Rs$ are obtained, they are used to retro-deform the point set, and then calculate the minimum distance between the nearest neighbour points. An open interval of 0 to the distance is then evenly meshed into 50 nodes that we search for the best estimate of $r_0$.

### 4.3. Results

Shown in Figs. 1 and 3 are results from applying the proposed strain method for $k = 1$, 2 and 3, respectively, to a large number of point object sets generated in the way described above.

In Fig. 3d–e, the accuracy of strain estimate has a great fluctuation at the start, and then strongly tends to increase with the number of points in the set. The estimated strain matches the prescribed strain very well, when point number is no less than 60. Meanwhile, the estimate of cutoff radius appears to decrease monotonically, and approaches the prescribed cutoff radius when the number of points becomes greater than 60 (Fig. 3f). With the increase of outcomes ($k$), there exists a weak tendency of increasing the accuracy of strain estimate and cutoff radius, particularly for a point number of less than 60.

As a whole, 60 is a turning point for the maximised mean log likelihood function (MMLLF) (Fig. 3a). Above this number, the measure displays a stronger tendency to increase. Apparently, over this number there must exist a sufficient portion of well-packed point objects that play a significant role in guaranteeing the great accuracy of strain estimate. This is also reflected by the increase of the intensity ratio with point number (Fig. 3b). All the values of the intensity ratio are larger than 1.0, implying that the point sets should be not well packed. It is thus fortunate that we can obtain a high accuracy of strain estimate from some less packed point objects.

In contrast, the K–S test statistic has a great fluctuation, without any tendency of increasing or decreasing (Fig. 3c). Occasionally, it becomes smaller than $D_{0.10}$, indicating the randomness of points within the confidence level.

## 5. Application

A real example of clast supported conglomerate (Fig. 4) was chosen in this section. Pebbles in the rock are intensely deformed and aligned approximately east – west, with an axial ratio of 2.5–16.0 or more. Although a difference in viscosity between the matrix and the pebbles would cause a departure of the estimated strain from the real value, this effect is thought to decrease with the amount of the matrix (Gay, 1968a,b). For this reason, we consider only the upper part that contains a majority of more closely packed pebbles. The geometrical centers of these pebbles (Fig. 5a) are recognized by eye as point objects.

Applying our proposed method to the set gives rise to results shown in Figs. 5e and 6. In the estimated strain, $Rs = 16.2$ and $\theta = 170°$. Both the intensity ratio and the K–S test statistic have a relatively large value, 2.01 and 0.24, indicating that the point objects in the set are less packed and less random. In this instance, the strain ratio must be over-estimated for it is approximately equivalent to the axial ratio of the most elongated pebbles. A plausible reason for this over-estimation is the existence of some spurious point object(s) due to the under packed state of the set.

Spurious points, if they exist in the set, need be picked out before we can make meaningful estimation of strain from other points. For this purpose, we develop a detection method, similar to leave-one-out cross validation (Evans and Rosenthal, 2000), whose procedure is summarized below.

1) For a set of $n$ point objects, delete each individual point in a sequence and calculate the MMLLF for the rest, or $n-1$ points by use of the grid search described in the previous section.
2) Delete from the set the point corresponding to the minimum value of the MMLLF, and return to step 1 until the number of deleted points reaches the prescribed value, 10 in this case.
3) Look for the spurious point(s) at the swift change of the minimum value of the MMLLF with the number of deleted points.

Figs. 5f and 6 show the results from applying the above procedure to the set. The best solution of strain has a direction of 167° for the major strain axis, and a strain ratio of 7.3. All strain estimates vary very slightly, in contrast with strain estimated from the original data set. This makes us believe that there exists only one spurious point in the set. Hence, we accept with confidence strain estimated from any of the above sets excluding this spurious point, although the MMLLF reaches the maximum at a deleted data number of 10. In addition, the central void in Fry plot appears more apparent in the set excluding a number of ten deleted data (Fig. 5d) than in the original set (Fig. 5c), because the MMLLF is about two orders of magnitude larger for the former than for the later (Fig. 6a).

## 6. Discussion

Some theoretical and practical aspects about this method will be discussed below.

### 6.1. Advantages and disadvantages

The proposed strain method highlights a new approach for estimating of strain from point distributions not by visual appreciation but by numerical solution. We believe that it provides a more objective and more robust measure of strain than pre-existing improvements on the Fry method, many of which are based upon visualization. It is however a pity that the numerical treatment does not need the Fry plot as a working platform. This would certainly disappoint those structural geologists preferring manual manipulation and visual appreciation, although the numerical estimates are superimposed on a visual plot (Figs. 1d and 5e–f). In addition, the grid search adopted in this paper is very time-consuming, and should be replaced by other numerical algorithms much more efficient in search of the best solution of strain.

### 6.2. Use of all the points in the Fry plot

Can we use all the points in the Fry plot for estimation of strain? This is raised in the Introduction section, since most pre-existing improvements on the Fry method (e.g. Erslev, 1988; Erslev and Ge,

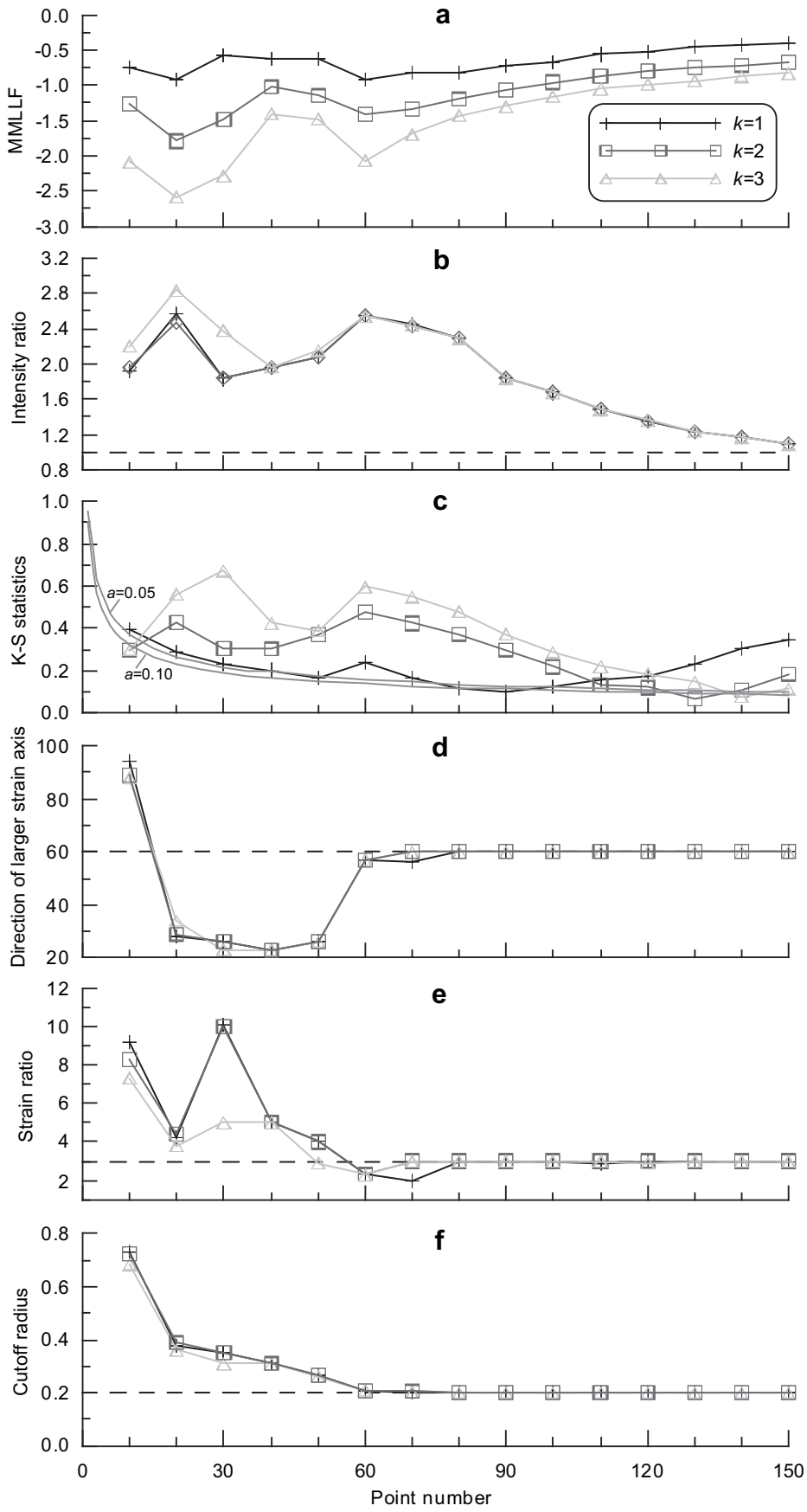


**Fig. 3.** The maximized mean log likelihood function (MMLLF) (a), the intensity ratio (b), the K–S test statistic (c), the estimated direction in degree of the major strain axis (d), the estimated strain ratio (e), and the estimated cutoff radius (f) by applying the proposed strain method for $k = 1$, 2 and 3, respectively, to a series of artificial sets. Values of $D_{0.10}(n)$ and $D_{0.05}(n)$ are shown in gray lines in (c). Dashed lines shown in (d)–(f) represent prescribed values. See the text for more information.

**Fig. 4.** A clast supported conglomerate from Gäddede, Sweden showing pebbles with elongation in an approximately horizontal direction. A number of 51 pebbles only in the upper part were taken into account for their dense occurrence. (Photo courtesy of R.J. Lisle.)

1990; Ailleres and Champenois, 1994) concentrate on the nearest neighbour points, just a small minority of points in the plot. Our answer to this question is optimistic. In our method, it is fairly easy to use all or a vast majority of points in the plot to estimate strain from them through increasing the number of outcomes in the PDF. The increase tends to enhance the accuracy of strain estimate, as discussed previously. However, as a drawback about doing this, the edge effect becomes more apparent when the more subregions of the outcomes are not confined to the sampled region. How it affects the accuracy of strain estimate is beyond the scope of this paper.

### 6.3. Spurious points

The less densely the packed point objects are, the greater the probability of the existence of spurious points. The presence of spurious points would affect the accuracy of strain estimate to

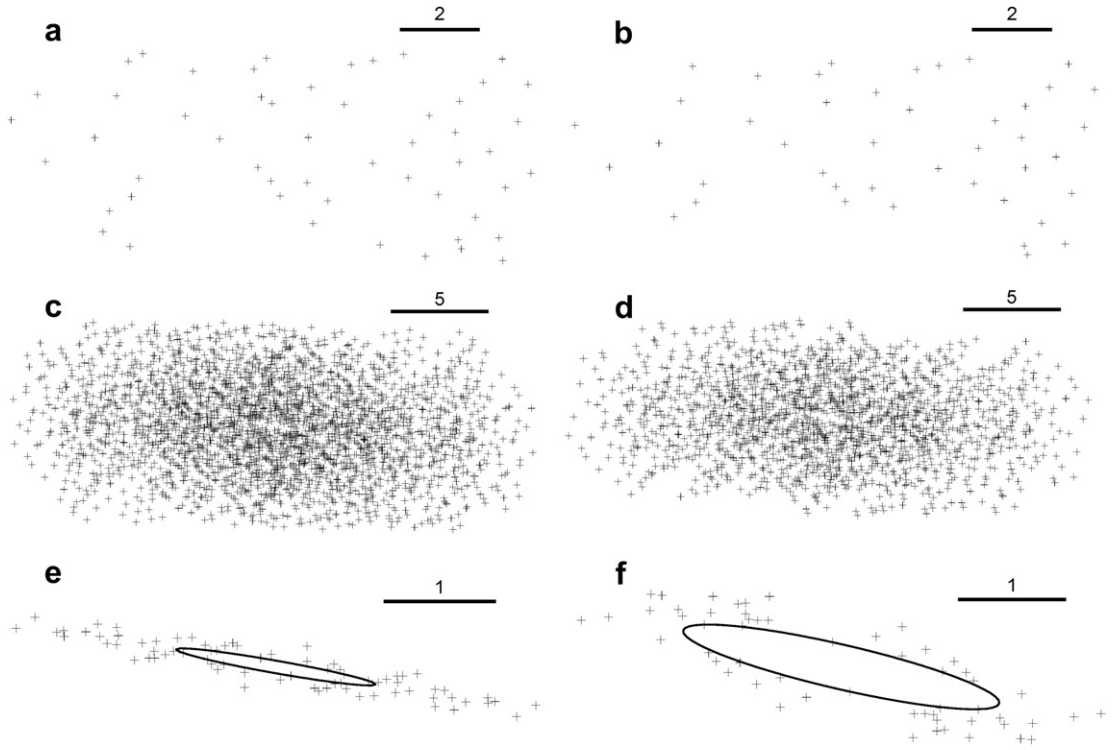


**Fig. 5.** For the original set with and without ten deleted data (a and b), their Fry plots (c and d) and the best fit elliptical voids in the center and the first nearest neighbour points in the pre-deformed state (e and f). All points are marked by “plus”. The deleted points are picked out using the detection method developed in this paper.

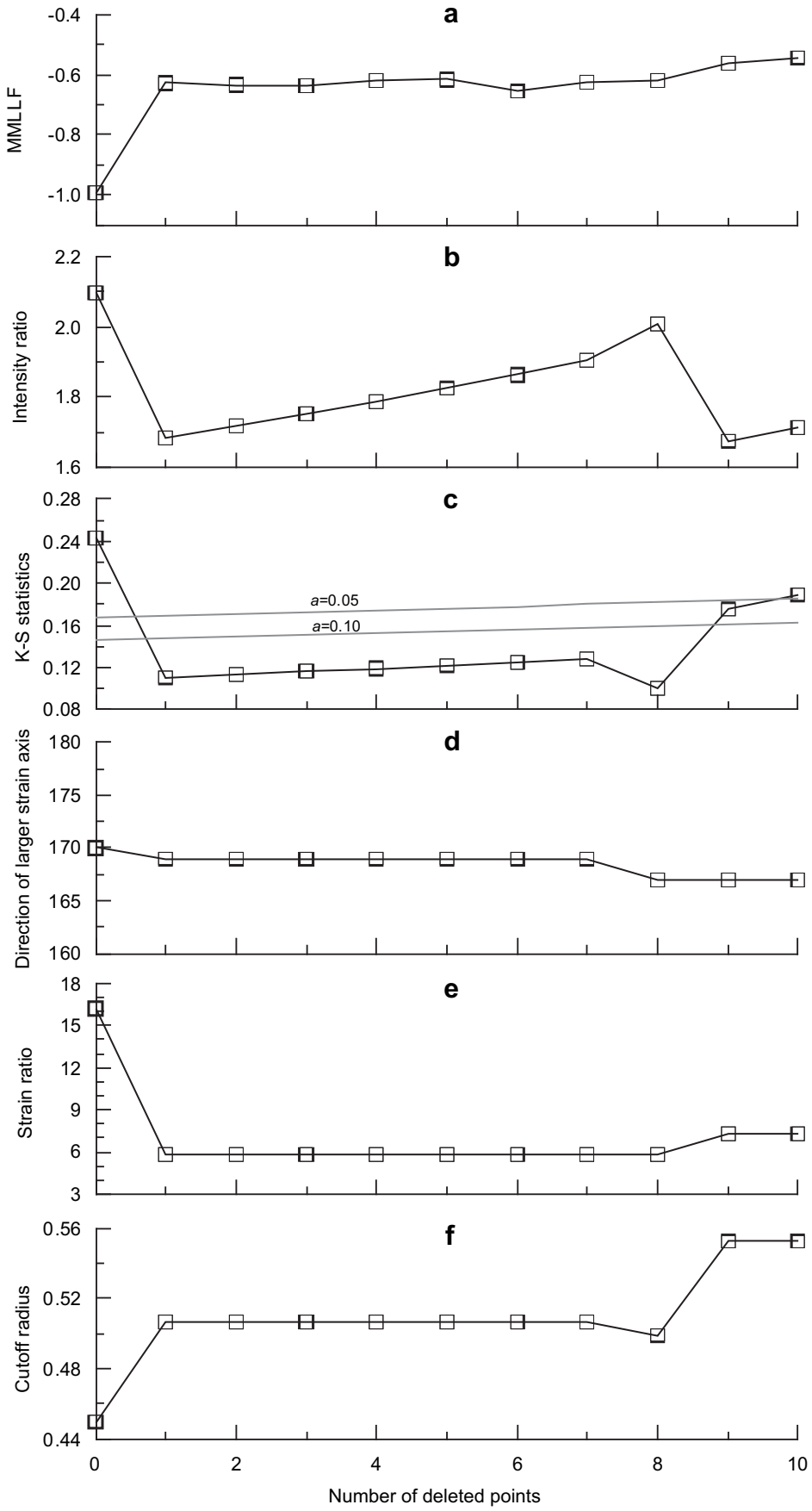


**Fig. 6.** The maximized mean log likelihood function (MMLLF) (a), the intensity ratio (b), the K–S test statistic (c), estimated direction in degree of the major strain axis (d), the estimated strain ratio (e), and the estimated cutoff radius (f) by applying the proposed strain method for $k = 1$ to the original set with 0–10 deleted points, respectively. Values of $D_{0.10}(n)$ and $D_{0.05}(n)$ are shown in gray lines in (c).

a greater or lesser degree. As far as we are concerned, there is no good way to recognize spurious point(s), if present. The detection method used in this paper is inspired by the idea of leave-one-out cross validation. It assumes that, for the point set excluding the spurious point(s) the MMLLF should reach the maximum, and should have a very slight variation, if at all. Otherwise, it is beyond the capacity of the method to pick out spurious point(s).

### 6.4. Point object distribution

As mentioned above, point objects in the pre-deformed state are assumed to obey the Poisson distribution that is truncated at the maximum PDF of one outcome ($k = 1$). The nearest neighbour outcomes thus become mutually dependent if their separation is smaller than twice the cutoff diameter, or independent if it is larger. Strictly speaking, such truncation is better explained by two-dimensional sequential random packing (SRP) than by the truncated Poisson process. Packing is a common phenomenon not only in real life, but also in scientific disciplines such as physics, chemistry and biology. At present we have an analytical solution to one-dimensional packing problems, also called the car-parking problem, for one or two types of the object size (e.g. Rényi, 1958; Mannion, 1964; Rabinovitch et al., 1999; Rawal and Rodgers, 2005), but, unfortunately, solving two-dimensional SRP problems requires numerical methods (Akeda and Hori, 1976; Torquato et al., 2000).

## 7. Conclusions

Statistical aspects of the Fry method of strain analysis have been examined in this paper, including two of the fundamental requirements, truncation and randomness. Point objects in the pre-deformed state are assumed in the Poisson distribution that is truncated at the maximum value of the probability density function (PDF) of one outcome. Under this assumption, the point objects appear well packed, and provide a good strain marker for the Fry method. Multiplying the PDF at each individual outcome in the deformed state gives the likelihood function. The mean log of the likelihood function (MLLF) is maximized using the grid search to look for the best solution of the unknown parameters, such as the direction of the major strain axis, strain ratio and cutoff radius. For point objects in the pre-deformed state, their packing degree is described by the intensity ratio, and their randomness with respect to the well-packed point objects is tested using the Kolmogorov–Smirnov test statistic.

This new method was applied not only to a series of artificial sets of point objects generated at the prescribed parameters, such as strain, cutoff radius, and the number of randomly packing points at each set, but to an example of a deformed conglomerate as well. Results demonstrate the feasibility of the method. The estimated strain has a very high accuracy even for artificial sets of less packed points. The accuracy tends to increase with the number of outcomes in the PDF, particularly for a point number of less than 60.

In the real example, the excessively small value of the maximized MLLF was attributed to the presence of one spurious point in the set. The spurious point was picked out using the detection method developed in this paper, and we had a more robust strain estimated from the rest points.

Finally, some theoretical and practical aspects of this novel method were discussed, including the advantages and disadvantages of the method, the possibility of using all the points in the Fry plot for strain estimation, the recognition of spurious points, and the distribution pattern of point objects.

## Acknowledgement

This work is funded by National Basic Research Program of China 973 (Grant 2007CB411307), Hundred Talent Program of Chinese Academy of Sciences (KZCX0543081001), and National Natural Science Foundation of China (Grants 40672144 and 40725009). Prof. R.J. Lisle is thanked for the permission of using the photo in this paper, and for his kindly checking the draft of this paper twice. The first author had discussed the theory of the Fry method many times with him and Dr. Norman Fry, from which this paper benefited significantly. This paper was reviewed by J. Waldron and T. Blenkinsop who made value suggestions and numerous modifications to it. This is contribution No. IS- 1306 from GIGCAS.

## Appendix A. List of symbols and their definitions

| Symbols | Definitions | Explanations |
|---|---|---|
| $\boldsymbol{W}$ | The sampled region in two dimension | |
| $\boldsymbol{X_i}$ | The occurrence of each point object in $\boldsymbol{W}$ | $i = 1, 2, \ldots, n$ |
| $\boldsymbol{V}$ | A circular subregion of $\boldsymbol{W}$ | $V \in W$ |
| $n$ | The number of outcomes or points in $\boldsymbol{W}$ | See (9)–(11) |
| $P()$ | The probability density function | See Eqs. (1), (3), (6) (9) and (10) |
| $r$ | The radius of $\boldsymbol{V}$ | See Eqs. (1)–(4) |
| $r_0$ | The radius of the empty circular area in the pre-deformed state, or cutoff radius | See Eqs. (3)–(6) |
| $r_{0a}$ | The length of the minor strain axis of the cutoff elliptical area in the deformed state | $r_0 = r_{0a}\sqrt{Rs}$; see Eqs. (6) and (8)–(12) |
| $r_{ia}$ | The minor axial lengths of the elliptical subregion centered at $\boldsymbol{X_i}$ in the deformed state | See Eqs. (6) and (8)–(12) |
| $\lambda$ | The intensity of the Poisson process | See Eqs. (1)–(4) |
| $\hat{\lambda}$ | The density of well-packed point objects | See Eqs. (5)–(6) and (8)–(12) |
| $k!$ | The factorial of integer $k$ | See Eq. (1) |
| $k$ | The number of the nearest neighbour outcomes at the proximity to a certain outcome | See Eqs. (1)–(4), (6) and (9)–(12) |
| $\theta$ | The direction of the major strain axis | See Eqs. (6) and (9)–(12) |
| $Rs$ | Strain ratio | See Eqs. (6)–(12) |
| $D()$ | The Kolmogorov–Smirnov test statistic | See Eq. (11) |
| $F_{cal}()$ | The calculated cumulative distribution frequency of the $k$-th nearest neighbour at a certain distance | See Eqs. (11)–(12) |
| $F_{obs}()$ | The observed cumulative distribution frequency of the $k$-th nearest neighbour at a certain distance | See Eqs. (11)–(12) |

## References

Ailleres, L., Champenois, M., 1994. Refinements to the Fry method (1979) using image processing. Journal of Structural Geology 8, 799–808.
Akeda, Y., Hori, M., 1976. On the random sequential packing in two and three dimensions. Biometrika 63, 361–366.
Clark, P.J., Evans, F.C., 1954. Distance to nearest neighbour as a measure of spatial relationships in populations. Ecology 35, 445–453.
Crespi, J.M., 1986. Some guidelines for the practical application of Fry's method of strain analysis. Journal of Structural Geology 16, 1327–1330.
Daley, D.J., Vere-Jones, D., 2008. An Introduction to the Theory of Point Processes. In: General Theory and Structure, vol. II. Springer-Verlag, New York, p. 573.
Dunne, W.M., Onasch, C.M., Williams, R.T., 1990. The problem of strain-marker centers and the Fry method. Journal of Structural Geology 12, 933–1990.
Erslev, E., 1988. Normalized center-to-center strain analysis of packed aggregates. Journal of Structural Geology 10, 201–209.
Erslev, E., Ge, H., 1990. Least-squares center-to-center and mean object ellipse fabric analysis. Journal of Structural Geology 12, 1047–1059.

Evans, M.J., Rosenthal, J.S., 2000. Probability and Statistics: The Science of Uncertainty. W.H. Freeman & Company, New York, p. 685.
Fry, N., 1979. Random point distributions and strain measurement in rocks. Tectonophysics 60, 89–105.
Fry, N., 1999. Fry plots: warming about summed moments. Journal of Structural Geology 21, 129–134.
Gay, N.C., 1968a. Pure shear and simple shear deformation of inhomogeneous viscous fluids. 1. Theory. Tectonophysics 5, 211–234.
Gay, N.C., 1968b. Pure shear and simple shear deformation of inhomogeneous viscous fluids. 2. The determination of the total finite strain in a rock from objects such as deformed pebbles. Tectonophysics 5, 295–302.
Genier, F., Epard, J.-L., 2007. The Fry method applied to an augen orthogeneiss: problems and results. Journal of Structural Geology 29, 209–224.
Ghaleb, A.M., Fry, N., 1995. CSTRAIN: a FORTRAN 77 program to study Fry's plots in two-dimensional simulated models. Computers & Geosciences 21, 825–831.
Gibbons, J.D., 1985. Nonparametric Methods for Quantitative Analysis, Second Edition. American Sciences Press, New York, p. 481.
González-Casado, J.M., Jimenez-Berrocoso, A., Carcia-Cuevas, C., Elorza, J., 1983. Strain determinations using inoceramid shells as strain makers: a comparison of calcite strain gauge technique and the Fry method. Journal of Structural Geology 25, 1773–1778.
Hanna, S.S., Fry, N., 1979. A comparison of methods of strain determination in rocks from southwest Dyfed (Pembrokeshire) and adjacent areas. Journal of Structural Geology 2, 155–162.
Lindsey, J.K., 1997. Applying Genealized Linear Model. Springer-Verlag, New York, p. 265.
Mannion, D., 1964. Random space-filling in one dimension. Publications on the Mathematical Institute of the Hungarian Academy of Sciences 9, 143–153.
McNaught, M.A., 1994. Modifying the normalized Fry method for aggregates of non-elliptical grains. Journal of Structural Geology 16, 493–503.
Onasch, C.M., 1986. Ability of the Fry method to characterize pressure-solution deformation. Tectonophysics 122, 187–193.
Rabinovitch, A., Bahat, D., Melamed, Z., 1999. A note on joint spacing. Rock Mechanics & Rock Engineering 32, 71–75.
Ramsay, J.G., 1967. Folding and Fracturing of Rocks. McGraw-Hill Book Co., New York, p. 568.
Rawal, S., Rodgers, G.J., 2005. Modelling the gap size distribution of parked cars. Physica A 346, 621–630.
Ripley, B.D., 1976. The second-order analysis of stationary point processes. Journal of Applied Probability 13, 255–266.
Rényi, A., 1958. On a one-dimensional space-filling problem. Publications on the Mathematical Institute of the Hungarian Academy of Sciences 3, 109–127.
Torquato, S., Truskett, T.M., Debenedetti, P.G., 2000. Is random close packing of spheres well defined? Physical Review Letters 84, 2064–2067.
Vearncombe, J., Vearncombe, S., 1999. The spatial distribution of mineralization: application of Fry analysis. Economic Geology 94, 475–486.
Waldron, J.W.F., Wallace, K.D., 2007. Objective fitting of ellipses in the center-to-center (Fry) method of strain analysis. Journal of Structural Geology 29, 1430–1444.