



**HAL**  
open science

# Structural interpretation of sparse fault data using graph theory and geological rules Fault data interpretation

Gabriel Godefroy, Guillaume Caumon, Gautier Laurent, François Bonneau

## ► To cite this version:

Gabriel Godefroy, Guillaume Caumon, Gautier Laurent, François Bonneau. Structural interpretation of sparse fault data using graph theory and geological rules Fault data interpretation. *Mathematical Geosciences*, 2019, 51 (8), pp.1091-1107. 10.1007/s11004-019-09800-0 . hal-01879697

**HAL Id: hal-01879697**

**<https://hal.univ-lorraine.fr/hal-01879697>**

Submitted on 24 Sep 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Structural interpretation of sparse fault data using graph theory and geological rules

## Fault data interpretation

G. Godefroy · G. Caumon ·  
G. Laurent · F. Bonneau

Received: . . / . . / . . . . Accepted: . . / . . / . . . .

**Abstract** Structural uncertainties arise partly from the association of sparse fault interpretations made from 2D seismic lines or limited outcrop observations. We propose a graph formalism to describe the problem of associating spatial fault evidence. A combinatorial analysis shows that the number of association scenarios is related to the Bell number and increases exponentially. This makes the complete exploration of the uncertainties computationally highly challenging. We formulate prior geological knowledge as numerical rules to reduce the number of scenarios and to make structural interpretation more repeatable and objective. We use the Bron-Kerbosch graph algorithm to detect the major possible structures. This framework opens the way to a numerically assisted exploration of uncertainties during structural interpretation.

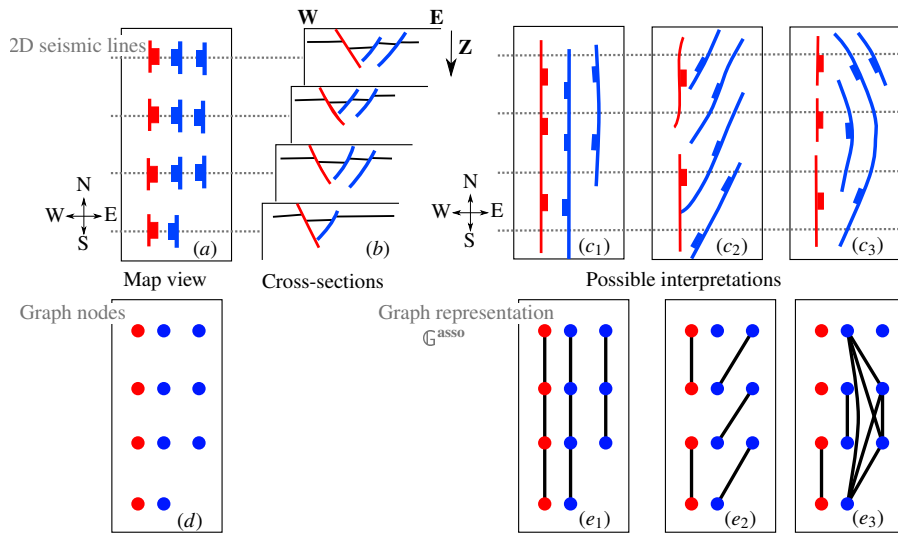
**Keywords:** Structural interpretation, Fault network, Graph processing.

## 1 Introduction

In structural geology and geological mapping, uncertainties arise both from the sparsity and the limited quality of observations [e.g., Thore et al., 2002, Wellmann et al., 2010, Jessell et al., 2010] and from the concepts chosen for their interpretation [e.g., Bond et al., 2007]. Observation gaps between 2D seismic lines, outcrops, boreholes or focal mechanisms are common. Even when 3D seismic data is available, image quality can be poor near faults [Botter et al., 2014, Iacopini et al., 2016, Weinzierl et al., 2016, Botter et al., 2017]. As a result, defining a structural architecture is often tedious and prone to uncertainty. These structural uncertainties impact natural resource exploration

---

G. Godefroy, G. Caumon, G. Laurent and F. Bonneau  
GeoResources, Université de Lorraine / CNRS / CREGU, 2 Rue du Doyen Marcel  
Roubault, 54500 Vandoeuvre-lès-Nancy, France  
E-mail: gabriel.godefroy@univ-lorraine.fr  
Tel: +33 (0)6 68 11 24 42

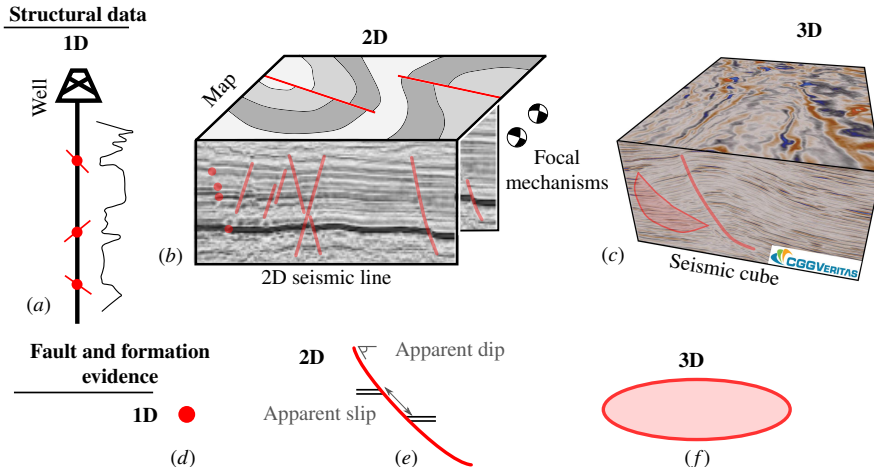


**Fig. 1:** Association of labeled fault evidence (red: east-dipping and blue: west-dipping) interpreted in map view (a) or on 2D seismic lines (b) is an under-constrained problem. Several structural interpretations are possible ( $c_1$ ,  $c_2$  and  $c_3$ ) [Modified from Freeman et al., 1990]. Each interpretation is represented by an association graph ( $e_1$ ,  $e_2$  and  $e_3$ ) whose labeled nodes are the fault evidence (d). The edges link fault evidence belonging to the same fault.

and exploitation [Hollund et al., 2002, Thore et al., 2002, Rivenæs et al., 2005, Richards et al., 2015, Julio et al., 2015a], waste disposal [Mann, 1993, Schneeberger et al., 2017], environmental engineering [Rosenbaum and Culshaw, 2003] and geotechnical works [Zhu et al., 2003].

The observation gap is commonly addressed by the construction of geological interpretative sections from available data, followed by deterministic 3D model [Caumon et al., 2009]. This process generally produces one single structural model which does not capture interpretation uncertainty. As geological concepts and knowledge involved in the interpretation are generally not fully described, this make interpretation very subjective and very difficult to reproduce [Bond et al., 2007]. In this paper, we propose instead to focus on the association of spatial fault evidence. This problem typically occurs during the interpretation of 2D seismic lines or limited outcrop information. This is called the *fault correlation problem* by Freeman et al. [1990] or *fault evidence association problem* in this paper (Fig. 1.a-c).

To address this challenge, we propose to formalise the fault evidence association problem using graph theory (Sect. 2). Graph theory has already been proposed to carry the topology of geological structures such as karsts [Collon et al., 2017], fractures networks [Valentini et al., 2007, Sanderson and Nixon, 2015, Santiago et al., 2016], or 3D geological models [Thiele et al., 2016a,b].



**Fig. 2:** Faults are observed from outcrops, wells or seismic data (*a, b, c*). The structural data is interpreted as points, lines and surfaces, hereafter designed as *fault evidence* (*d, e, f*). Apparent fault geometry and displacement are sometimes extracted. [Seismic data courtesy of Rob Butler and David Iacopini through the Virtual Seismic Atlas.]

In this contribution, we suggest using graphs to represent how fault evidence may be associated together. The proposed graph structure offers a simple representation of a structural interpretation. Each association scenario is represented by an *association graph*  $\mathbb{G}^{\text{asso}}$ . Building such graphs is constrained by the prior geological knowledge coming from the literature, the geological context or analog data. In our approach, this prior knowledge is formulated as numerical rules to make structural interpretation more objective; these rules are used to build a graph  $\mathbb{G}_{\varphi}^{\text{all}}$  in which all possible associations for a given fault family (denoted by the index  $\varphi$ ) are considered. This is a way to focus on topological aspects rather than on geometrical details [as suggested in Thomson, 2014].

Therefore, our approach considers how fault evidence may be associated, either in a deterministic ( $\mathbb{G}^{\text{asso}}$ ) or probabilistic sense  $\mathbb{G}_{\varphi}^{\text{all}}$ . Building such graphs rely on numerical geological rules that define how likely and unlikely is that disjoint fault evidence actually belong to the same fault of the same family.

In this frame, we show that a fault surface corresponds to a clique in  $\mathbb{G}_{\varphi}^{\text{all}}$ , the graph of all possible associations (Sect. 2). This observation makes it possible to detect the major possible structures using a maximal clique detection algorithm (Sect. 3). Overall, the knowledge of different possible interpretations has the potential to limit the risk of anchoring to a single initial interpretation [e.g., Bond, 2015].

## 2 A theoretical framework to associate fault evidence using graph theory

### 2.1 Fault evidence and geological knowledge

Faults are typically observed from outcrops, wells or seismic data or through the inversion of focal mechanisms (Fig. 2.a,b,c). The interpretation of these data is summarized by points, lines or surfaces indicating the fault position, hereafter designed as *fault evidence* (Fig. 2.d,e,f). On a cross-section, the local fault geometry (apparent dip and strike) and the apparent fault displacement also provide information that is commonly used by structural geologists to associate sparse fault evidence [e.g., Barnett et al., 1987, Dee et al., 2005, Boulton and Freeman, 2007]. All these observations can be obtained during field work or may be extracted (manually or automatically) from reflection seismic data [e.g., Hale, 2013, Lacaze et al., 2016, Wu and Zhu, 2017].

### 2.2 Mathematical formalism using graphs

We propose to represent the previously described fault evidence as the labeled nodes of a graph (Fig. 1.d). The labels can be used to store information about the fault observations such as the orientation or apparent orientation, the displacement or the size of the observations. For example, the graph nodes in Figure 1 are labeled according to the apparent dip of the interpreted faults (red for east-dipping or blue for west-dipping). The graph edges correspond to the association of fault evidence into the same fault objects so that each fault is a set of connected nodes (Fig. 1.e).

More formally, let  $\mathbb{V}_{obs} = \{\mathbf{v}_{i=1,n}\}$  be a set of  $n$  fault evidences. Our goal is to associate these  $n$  observations into  $k$  ( $k \leq n$ ) fault surfaces. Each possible interpretation can be represented by an association graph  $\mathbb{G}^{asso} = \{\mathbb{V}_{obs}, \mathbb{E}\}$ . In this graph, two nodes (representing fault evidence) are linked by an edge  $\mathbf{e}_{ij} \in \mathbb{E}$  if they are interpreted as belonging to the same fault. The goal of the association problem is to define clusters (or subsets) of observations belonging to the same fault  $\mathbb{F}_f \subseteq \mathbb{V}_{obs}$  where  $f$  is a fault index. In this case, the graph  $\mathbb{G}^{asso}$  must honor the following conditions:

- $\bigcup_f \mathbb{F}_f = \mathbb{V}_{obs}$ : all evidences belong to a fault.
- $\mathbb{F}_f$  is a clique (a subset of nodes that are all mutually adjacent, Fig. 3.b).
- $\mathbb{F}_{f_i} \cap \mathbb{F}_{f_j} = \emptyset$  for all possible faults  $f_i$  and  $f_j$ ,  $f_i \neq f_j$ : an evidence cannot belong to several faults. Indeed, in general branch points and branch lines are seldom observed directly in geoscience data [Yielding, 2016].

### 2.3 Combinatorial analysis of the association problem

The number of ways to partition a set of  $n$  fault evidence into  $k$  non-empty faults is given by the Stirling number of the second kind  $S(n, k)$  [e.g., Graham et al., 1994].  $S(n, k)$  satisfies the relation of recurrence:

$$\forall k \in ]0, n], \quad S(n, k) = S(n-1, k-1) + kS(n-1, k) \quad (1)$$

with initial conditions:

$$S(0, 0) = 1 \quad \text{and} \quad \forall n > 0, S(n, 0) = S(0, n) = 0. \quad (2)$$

The number of partitions of a set of  $n$  fault evidence into any number of faults is given by the Bell number:

$$B_n = \sum_{k=1}^n S(n, k). \quad (3)$$

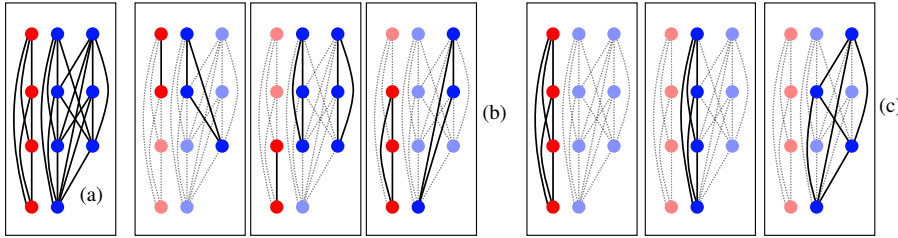
$n$	1	2	3	4	5	6	7	8	9	10	11	12	13
$B_n$	1	2	5	15	52	203	877	4.1e3	2.1e4	1.1e5	6.8e5	4.2e6	2.8e7

**Table 1:** The number of partitions of a set of  $n$  objects is given the Bell number  $B_n$ .

All these  $B_n$  association scenarios can theoretically be listed using the set partition algorithm available in Knuth [2005]. However, the combinatorial complexity (Table 1) makes the computation on current computer hardware impossible for more than thousands of nodes.

### 2.4 Geological interpretation rules to reduce the number of possible scenarios

We propose to focus on the most likely scenarios with user-defined numerical geological rules. Geological rules are essential to structural interpretation, as



**Fig. 3:** (a) Graph of all possible associations between labeled fault evidence. (b) Example of cliques (a clique is a subset of fault evidence that are all mutually adjacent) extracted from the graph shown in (a). (c) Maximal cliques are cliques that cannot be enlarged.

they tell about how likely (and unlikely) it is that disjoint fault evidence actually belong to the same fault. These rules completely drive the geological features that can emerge from the available data. Geological rules are already used to test and validate structural models [e.g., Freeman et al., 1990, 2010] and in the frame of Bayesian approach [e.g., de la Varga and Wellmann, 2016]. We propose to integrate this geological knowledge during structural interpretation to limit the time-consuming loop between structural interpretation, modeling and validation [Caumon et al., 2013].

Within a fault network, faults are generally organized into families that have similar geometrical parameters [e.g., Cherpeau et al., 2010, Henza et al., 2011, Nixon et al., 2011]. Because a structural interpretation depends on the considered tectonic context, we define numerical rules indicating to which fault family each observation may belong to (*family rules*). We also define rules to identify associations in a given fault family (*association rules*).

A *family rule*  $R_\varphi^{\text{fam}}(\mathbf{v}_i)$  quantifies the likelihood that a fault evidence  $\mathbf{v}_i$  belongs to a given family ( $\varphi$ ). Such a rule returns a number between 0 ( $\mathbf{v}_i$  cannot belong to the fault family  $\varphi$ ) and 1 (if it is highly likely that  $\mathbf{v}_i$  belongs to the fault family  $\varphi$ ). Family rules are defined from both the available structural data and the prior regional information such as outcropping analogs or regional tectonic data [as in Cherpeau and Caumon, 2015, Aydin and Caers, 2017]. In practice, such rules can use the apparent orientation of an interpreted fault, or the structural style of a fault (i.e. normal or reverse) stored as node labels.

*Association rules* define the likelihood that two observations of the same family belong to the same fault. An association rule  $R_\varphi^{\text{assoc}}(\mathbf{v}_i \leftrightarrow \mathbf{v}_j)$  returns a number between 0 ( $\mathbf{v}_i$  and  $\mathbf{v}_j$  cannot belong to the same fault of the family  $\varphi$ ) and 1 (if it is likely that the two observations belong to the same fault). We present two basic examples below, but more elaborate rules have been proposed using the distance, the orientation [Cherpeau and Caumon, 2015], the throw gradient [e.g., Barnett et al., 1987, Freeman et al., 1990, Cherpeau and Caumon, 2015], or the estimated strain [Freeman et al., 2010] between two observations.

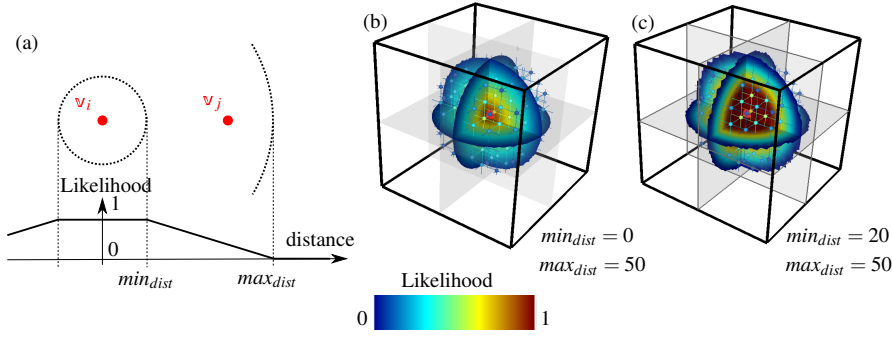
These geological rules are used to compute a graph of all possible associations  $\mathbb{G}_\varphi^{\text{all}}$  for each family  $\varphi$ . Each edge  $e^\varphi(\mathbf{v}_i, \mathbf{v}_j)$  (linking two observations  $\mathbf{v}_i$  and  $\mathbf{v}_j$ ) of the graph carries an association likelihood for each family:

$$L_\varphi^{\text{all}}(\mathbf{v}_i \leftrightarrow \mathbf{v}_j) = R_\varphi^{\text{fam}}(\mathbf{v}_i) R_\varphi^{\text{fam}}(\mathbf{v}_j) R_\varphi^{\text{assoc}}(\mathbf{v}_i \leftrightarrow \mathbf{v}_j). \quad (4)$$

In this framework, the edges where  $L_\varphi^{\text{all}}(\mathbf{v}_i \leftrightarrow \mathbf{v}_j) = 0$  are removed from the graph  $\mathbb{G}_\varphi^{\text{all}}$  of all the possible associations for each family  $\varphi$ , so that the corresponding associations will not be considered later on.

#### *Examples of association rules $R_\varphi^{\text{assoc}}$*

A simple likelihood that two observations  $\mathbf{v}_i$  and  $\mathbf{v}_j$  (whose barycenters are separated by a distance  $d_{ij}$ ) belong to the same fault is given using the rule



**Fig. 4:** Association likelihood computed on points and surfaces from a uniform distance rule. Only the points with a positive likelihood to be associated with the red central point are shown. (a) Linear decrease of the association likelihood with the distance between  $\mathbf{v}_i$  and  $\mathbf{v}_j$ . (b,c) Examples in 3D view.

$R_{dist}$  parameterized by two parameters  $min_{dist}$  and  $max_{dist}$  (Fig. 4):

$$R_{dist}(\mathbf{v}_i \leftrightarrow \mathbf{v}_j) = \begin{cases} 1 & \text{if } d_{ij} \leq min_{dist} \\ \frac{d_{ij} - min_{dist}}{max_{dist} - min_{dist}} & \text{if } min_{dist} \leq d_{ij} \leq max_{dist} \\ 0 & \text{if } d_{ij} \geq max_{dist} \end{cases} \quad (5)$$

As another example, an orientation rule  $R_{orientation}$  gives a likelihood that two observations  $\mathbf{v}_i$  and  $\mathbf{v}_j$  belong to the same fault plane whose orientation is defined by the unit normal vector  $\mathbf{n}_\varphi$ . The normalized vector between the barycenters of  $\mathbf{v}_i$  and  $\mathbf{v}_j$  is noted  $\mathbf{l}_{ij}$ . This rule is parameterized by the unit normal vector to the mean fault plane  $\mathbf{n}_\varphi$  and a tolerance angle  $\alpha_{tol}$  (Fig. 5). The likelihood is computed using:

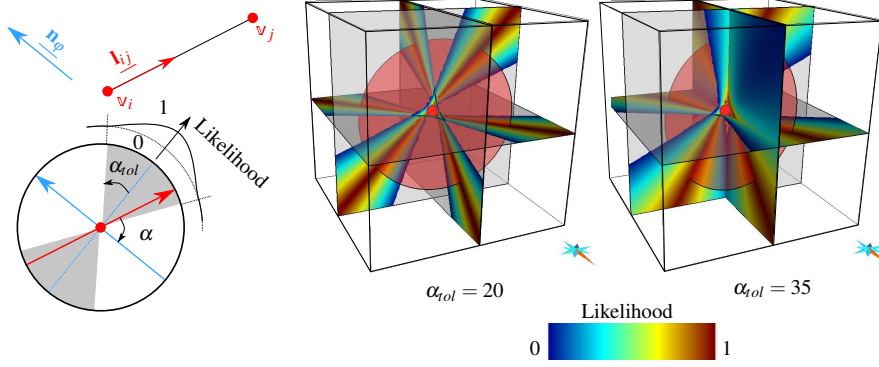
$$R_{orientation}(\mathbf{v}_i \leftrightarrow \mathbf{v}_j) = \max \left( 1 - \frac{|\mathbf{n}_\varphi \cdot \mathbf{l}_{ij}|}{\sin(\alpha_{tol})}, 0 \right). \quad (6)$$

If each evidence is composed of several points (for example, the observation is a fault trace), it is convenient to define  $R_{orientation}$  using the minimum value computed from each couple of points.

A wide range of alternative association rules can be defined, for example with different angular tolerances in the dip and strike directions, using Fisher-Bingham angular distributions [Kent, 1982, Carmichael and Ailleres, 2016], accounting for a maximum throw gradient, or accounting for the scaling laws describing fault length distributions [e.g., Bonnet et al., 2001, Torabi and Berg, 2011].

An interpretation rule  $R_{mix}$  can also be defined by combining several simple rules. For example, one may want to account for the distance separating the observations and for orientation, using two dedicated rules  $R_{dist}^{fam}$  and  $R_{orient}^{fam}$ .





**Fig. 5:** Association likelihood computed on points and surfaces from an orientation rule. Sectors of a horizontal and two orthogonal vertical planes a non-zero likelihood to be associated with the red central point are shown.  $\mathbf{n}_\varphi$  is the normal vector to a surface (in red in b,c) defining the mean orientation;  $\alpha_{tol}$  is a tolerance angle. (a) Decrease of the likelihood computed using an orientation rule. (b,c) Examples in 3D view.

This can be done by considering that both rules should be respected altogether:

$$R_{mix} = R_{dist}^{fam} R_{orient}^{fam},$$

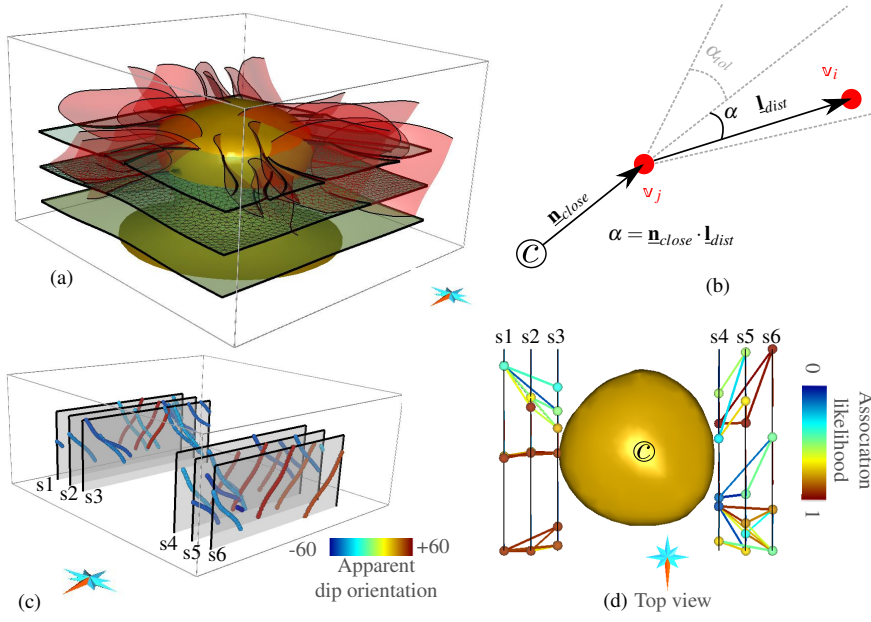
or that either rules apply:

$$R_{mix} = \frac{1}{2} (R_{dist}^{fam} + R_{orient}^{fam}).$$

## 2.5 Definition of geological rules according to the context

There is room for using a large set of geological rules during structural interpretation. In the previous works of Cherpeau and Caumon [2015] and Aydin and Caers [2017], the rules used during stochastic fault modeling are related to the three-dimensional modeling algorithm. Defining new interpretation rules would probably call for intrusive changes in their algorithms. For example, in both works, the fault orientation is supposed constant for a given fault family. This hypothesis is not respected in the case of normal faulting related to active salt diapirism where faults form a radial pattern [Davison et al., 2000, Yin et al., 2009]. In our framework, new interpretation rules can be added according to the context, for example to account for the non-stationarity of the orientations of faults that are radial around a salt diapir (Fig. 6). For this, we propose a new rule  $R_{radial}$  :

$$R_{radial}(\mathbf{v}_i \leftrightarrow \mathbf{v}_j) = \max \left( 0, \frac{\mathbf{n}_{close} \cdot \mathbf{l}_{dist} - \cos(\alpha_{tol})}{1 - \cos(\alpha_{tol})} \right) \quad (7)$$



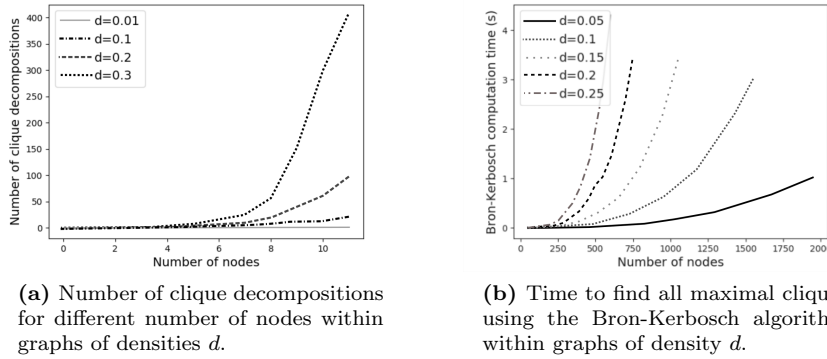
**Fig. 6:** (a) Synthetic structural model with radial faults around a salt diapir. (c) Synthetic fault evidence extracted along six seismic lines. (d) Graph of the possible associations, computed with the rule  $R_{radial}$  illustrated on (b) and Eq. 7.

where  $\mathbf{n}_{close}$  and  $\mathbf{l}_{dist}$  are the unit vectors respectively from the center of the diapir to the closest evidence and from this observation to the second one (Fig. 6.b).  $\alpha_{tol}$  is a tolerance angle.

### 3 Bron-Kerbosch algorithm to detect the major structures

The graph  $\mathbb{G}_\varphi^{all}$  proposes an abstract view of all possible faults for a given family and offers a way to screen an existing interpretation with a set of numerical geological rules. It can also be used to quickly screen which pairwise associations are impossible or highly unlikely. Nonetheless, the number of interpretation scenarios respecting these rules (i.e. the number of decompositions of  $\mathbb{G}_\varphi^{all}$  into non-overlapping cliques) is generally too large to list all possible scenarios, even in the case of small and sparse graphs (Fig. 7a).

Alternatively, one may want to generate some possible structural scenarios by processing this graph. For this, we propose a parsimonious strategy which detects all the possible major structures for each fault family using  $\mathbb{G}_\varphi^{all}$ . These major structures are represented by *maximal cliques* in  $\mathbb{G}_\varphi^{all}$ . A clique is maximal when no other node can be added while preserving its complete connectivity (Fig. 3.c). The detected structures can be helpful to provide several structural interpretations. To sample the uncertainties arising from fault



(a) Number of clique decompositions for different number of nodes within graphs of densities  $d$ .

(b) Time to find all maximal cliques using the Bron-Kerbosch algorithm within graphs of density  $d$ .

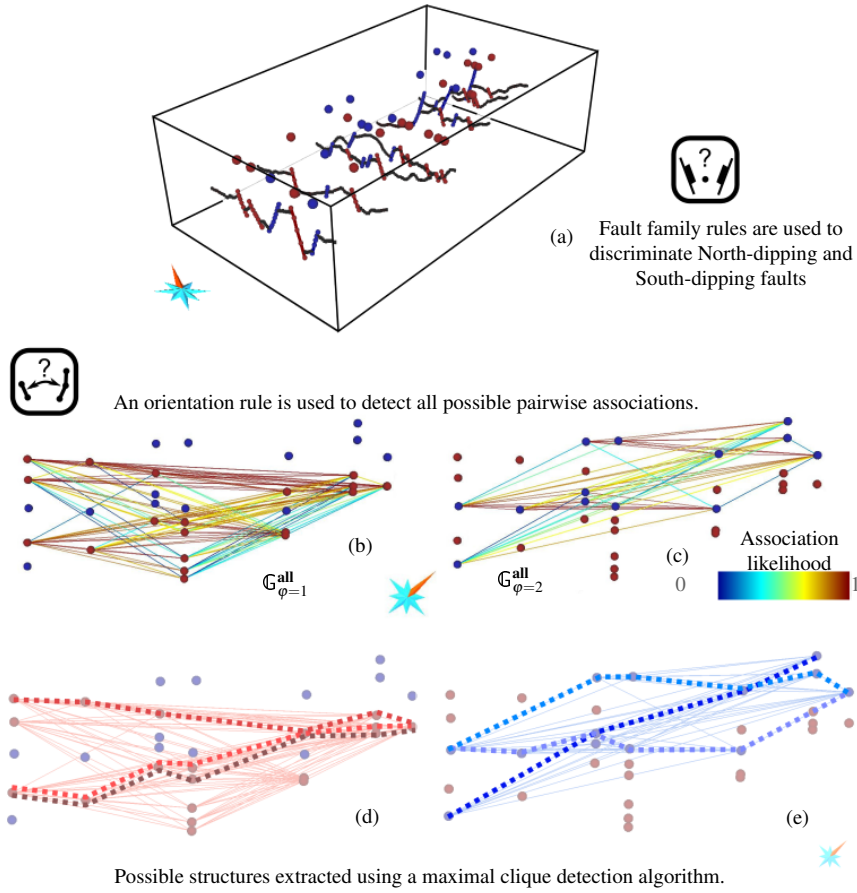
**Fig. 7:** Numerical experiments illustrating the combinatorial dimension of graph decomposition problems.  $d$  is the graph density. Computations run with *Networkx*, a Python package [Hagberg et al., 2008] on a personal laptop with a processor *Intel(R) Core(TM) i7-4510U* clocked at 2.00 GHz and 16 GB of Random-Access Memory.

segmentation, these main structures can then be divided into several fault segments, eventually separated by relay zones. Such a downscaling process can rely on the geometry of the fault surfaces [Julio et al., 2015a,b] or on kinematic criteria [Manighetti et al., 2009, 2015].

We use the Bron-Kerbosch [Bron and Kerbosch, 1973] algorithm to find all maximal cliques. Other maximal clique detection algorithms have theoretical better running times, but the Bron-Kerbosch algorithm is frequently reported as being more efficient in practice for sparse graphs [Tomita et al., 2006, Cazals and Karande, 2008]. The running time is exponentially related to the number of nodes and to the density of the graph ( $d = \frac{\text{number of edges}}{\text{maximum number of edges}}$ ). It remains acceptable for sparse graphs with less than 1000 nodes (Fig. 7b).

In the case of a three-dimensional seismic acquisition, dealing with more than 1000 points during fault interpretation is common. However, the interpreter usually knows that several of these points explain the same fault, for example by interpreting both on vertical and time seismic slices. These points can then be considered as one node in the graph representation, thus reducing the combinatorial size of the problem.

Figure 8 shows an example of such detection on data interpreted along 2D seismic lines located offshore Morocco. Two families of normal faults have been interpreted by Walter [2016] (in red and in blue, Fig. 8.a) regarding the apparent dip orientation of the faults. For each family, the graph of all possible associations ( $\mathbb{G}_\varphi^{\text{all}}$ ) is built using the orientation rule (Eq. 6 and Fig. 8.b,c). Some of the possible largest faults extracted using the Bron-Kerbosch algorithm are shown Figure 8.d,e.



**Fig. 8:** (a) Structural fault evidence interpreted from seismic data offshore Morocco. The corresponding graph nodes are colored using a family rule relying on dip orientation. (b)  $\mathbb{G}_{\varphi}^{\text{all}}$  for the South-dipping, and for the (c) North-dipping fault families using an orientation rule defined from a close onshore analog. (d, e) Some of the maximal structures detected using the Bron-Kerbosch algorithm.

#### 4 Discussion and perspectives

This contribution presents a lightweight graph-based framework to associate the interpreted fault evidence during structural interpretation. The methodology is parameterized by user-defined geological rules to discriminate likely and unlikely association scenarios. We argue that using numerical interpretation rules makes geological interpretation easier to reproduce and less subjective. These rules can also be used to analyze existing expert-based interpretations.

We have shown that the combinatorial complexity of the association problem makes the complete exploration of structural interpretation very difficult, uncertainties unless very strong rules can be found. The major possible struc-

tures can be detected using the Bron-Kerbosch algorithm, a maximal clique detection method. These major structures can also be seen as important alternative scenarios while trying to generate several fault network models to assess structural uncertainties. We discuss below some important aspects of this work and new perspectives. Nonetheless, the proposed graph abstraction makes it possible to define and test various geometric and kinematic rules to guide structural interpretation in a wide range of contexts. For example, the relationship and the scaling laws between the fault length and the throw gradient could be considered [e.g., Gillespie et al., 1992, Torabi and Berg, 2011].

In this paper, we did not consider the uncertainty associated to the fault evidence. In the case where a particular evidence is uncertain, its probability of existence can easily be integrated using a "null-family" we regroup erroneous interpretation. In the case of a geometric uncertainty about a fault evidence [e.g., Thore et al., 2002, Wellmann et al., 2010] the specific rules should be designed, reflecting for instance the position or orientation uncertainty.

#### 4.1 Geological rules considering multi-points costs and conditional likelihoods

At this point, the interpretation rules only consider pairwise associations of observations and do not consider the multi-point likelihood of associating several data at once. Such a multi-point rule could consider the throw distribution on the fault surface [as in Freeman et al., 1990] and could further reduce the number of possible scenarios. The association likelihoods should also be revised during the interpretation to capture that crossing faults at large-scale are unlikely [Schneeberger et al., 2017]. The conditional likelihood of associating two evidence, knowing that two other observations belong to the same fault, has not been considered yet.

#### 4.2 From graph representations to 3D structural models

In this paper, we focused on the association of sparse fault evidence. In order to head for a more comprehensive description of the data interpretation topology, it is necessary to handle the interactions between faults. A second oriented graph could be used to represent the branch lines.

Once associations of fault evidence have been defined, fault geometries should be modeled. Implicit or explicit fault surfaces can be interpolated in order to obtain a fault network in which horizon surfaces are built. One perspective is to partly automate this process in order to propagate uncertainties to structural mapping, waste disposal, seismic hazard assessment, and mining or petroleum production forecasts.

### 4.3 Stochastic fault interpretation

Even though the use of appropriate rules significantly increases the chance to find the correct association, it is unlikely that a single interpretation is correct when working with sparse data. Therefore, we strongly advocate stochastic interpretation to test scenarios and to assess structural uncertainties.

We suggest that the presented data structure can be used in the early stage of stochastic modeling, prior to building three-dimensional structural models, to assess uncertainties arising from fault segmentation using geometric [Julio et al., 2015a,b] or kinematic [Manighetti et al., 2009, 2015] criteria. However, the graph clustering methods proposed in the literature [e.g., Schaeffer, 2007] are deterministic and provide only one graph decomposition. A preliminary stochastic method generating several decompositions of non-oriented weighted graphs into non-overlapping cliques is proposed by Godefroy [2018].

It is important to mention that, in this paper, we focused on generating a minimal set of faults, in the sense that all faults provided by our method are supported by at least one fault evidence. Point processes could be used in complement to generate faults not directly supported by observations [e.g., Holden et al., 2003, Cherpeau et al., 2010, Davy et al., 2013, Bonneau et al., 2016, Aydin and Caers, 2017].

**Acknowledgements** This work was performed in the frame of the RING project at Université de Lorraine. We thank for their support the industrial and academic sponsors of the RING-GOCAD Consortium managed by ASGA (<http://ring.georessources.univ-lorraine.fr>). Software corresponding to this paper is available to sponsors in the RING software package FaultMod2. We acknowledge Paradigm for the SKUA-GOCAD software and API.

### References

- O. Aydin and J. K. Caers. Quantifying structural uncertainty on fault networks using a marked point process within a Bayesian framework. *Tectonophysics*, 2017. doi: 10.1016/j.tecto.2017.04.027.
- J. A. Barnett, J. Mortimer, J. H. Rippon, J. J. Walsh, and J. Watterson. Displacement geometry in the volume containing a single normal fault. *AAPG Bulletin*, 71(8):925–937, 1987.
- C. Bond, A. Gibbs, Z. Shipton, and S. Jones. What do you think this is? “Conceptual uncertainty” in geoscience interpretation. *GSA today*, 17(11): 4, 2007. doi: 10.1130/GSAT01711A.1.
- C. E. Bond. Uncertainty in structural interpretation: Lessons to be learnt. *Journal of Structural Geology*, 74:185–200, 2015. doi: 10.1016/j.jsg.2015.03.003.
- F. Bonneau, G. Caumon, and P. Renard. Impact of a stochastic sequential initiation of fractures on the spatial correlations and connectivity of discrete fracture networks. *Journal of Geophysical Research: Solid Earth*, 121(8): 5641–5658, 2016. doi: 10.1002/2015JB012451.

- E. Bonnet, O. Bour, N. E. Odling, P. Davy, I. Main, P. Cowie, and B. Berkowitz. Scaling of fracture systems in geological media. Reviews of geophysics, 39(3):347–383, 2001. doi: 10.1029/1999RG000074.
- C. Botter, N. Cardozo, S. Hardy, I. Lecomte, and A. Escalona. From mechanical modeling to seismic imaging of faults: A synthetic workflow to study the impact of faults on seismic. Marine and Petroleum Geology, 57:187–207, 2014. doi: 10.1016/j.marpetgeo.2014.05.013.
- C. Botter, N. Cardozo, D. Qu, J. Tveranger, and D. Kolyukhin. Seismic characterization of fault facies models. Interpretation, 5(4):1–53, 2017. doi: 10.1190/INT-2016-0226.1.
- P. Boulton and B. Freeman. Using faults as an aid to horizon mapping in areas of poor 2d seismic: Otway basin example. MESA Journal, 46, 2007.
- C. Bron and J. Kerbosch. Algorithm 457: finding all cliques of an undirected graph. Communications of the ACM, 16(9):575–577, 1973. doi: 10.1145/362342.362367.
- T. Carmichael and L. Ailleres. Method and analysis for the upscaling of structural data. Journal of Structural Geology, 83:121–133, 2016. doi: 10.1016/j.jsg.2015.09.002.
- G. Caumon, P. Collon-Drouaillet, C. L. C. De Veslud, S. Viseur, and J. Sausse. Surface-based 3D modeling of geological structures. Mathematical Geosciences, 41(8):927–945, 2009. doi: 10.1007/s11004-009-9244-2.
- G. Caumon, G. Laurent, J. Pellerin, N. Cherpeau, F. Lallier, R. Merland, and F. Bonneau. Current bottlenecks in geomodeling workflows and ways forward. Closing the Gap: Advances in Applied Geomodeling for Hydrocarbon Reservoirs. CSPG Memoir, 20:43–52, 2013.
- F. Cazals and C. Karande. A note on the problem of reporting maximal cliques. Theoretical Computer Science, 407(1-3):564–568, 2008. doi: 10.1016/j.tcs.2008.05.010.
- N. Cherpeau and G. Caumon. Stochastic structural modelling in sparse data situations. Petroleum Geoscience, 21(4):233–247, 2015. doi: 10.1144/petgeo2013-030.
- N. Cherpeau, G. Caumon, and B. Lévy. Stochastic simulations of fault networks in 3d structural modeling. Comptes Rendus Géoscience, 342(9):687–694, 2010. doi: 10.1016/j.crte.2010.04.008.
- P. Collon, D. Bernasconi, C. Vuilleumier, and P. Renard. Statistical metrics for the characterization of karst network geometry and topology. Geomorphology, 283:122–142, 2017. doi: 10.1016/j.geomorph.2017.01.034.
- I. Davison, I. Alsop, P. Birch, C. Elders, N. Evans, H. Nicholson, P. Rorison, D. Wade, J. Woodward, and M. Young. Geometry and late-stage structural evolution of central graben salt diapirs, north sea. Marine and Petroleum Geology, 17(4):499–522, 2000. doi: 10.1016/S0264-8172(99)00068-9.
- P. Davy, R. Le Goc, and C. Darcel. A model of fracture nucleation, growth and arrest, and consequences for fracture density and scaling. Journal of Geophysical Research: Solid Earth, 118(4):1393–1407, 2013. doi: 10.1002/jgrb.50120.

- 
- M. de la Varga and J. F. Wellmann. Structural geologic modeling as an inference problem: A bayesian perspective. *Interpretation*, 4(3):SM1–SM16, 2016. doi: 10.1190/INT-2015-0188.1.
- S. Dee, B. Freeman, G. Yielding, A. Roberts, and P. Bretan. Best practice in structural geological analysis. *First Break*, 23(4), 2005.
- B. Freeman, G. Yielding, and M. Badley. Fault correlation during seismic interpretation. *First Break*, 8(3):87–95, 1990. doi: 10.3997/1365-2397.1990006.
- B. Freeman, P. J. Boulton, G. Yielding, and S. Menpes. Using empirical geological rules to reduce structural uncertainty in seismic interpretation of faults. *Journal of Structural Geology*, 32(11):1668–1676, 2010. doi: 10.1016/j.jsg.2009.11.001.
- P. Gillespie, J. t. Walsh, and J. Watterson. Limitations of dimension and displacement data from single faults and the consequences for data analysis and interpretation. *Journal of Structural Geology*, 14(10):1157–1172, 1992. doi: 10.1016/0191-8141(92)90067-7.
- G. Godefroy. *Modélisation cinématique et stochastique des failles pour la prise en compte des incertitudes structurales*. PhD thesis, Université de Lorraine, 2018.
- R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete mathematics: a foundation for computer science*. Pearson Education India, 1994.
- A. A. Hagberg, D. A. Schult, and P. J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conferences (SciPy 2008)*, Gaël Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), volume 2008, pages 11–15, 2008.
- D. Hale. Methods to compute fault images, extract fault surfaces, and estimate fault throws from 3D seismic images. *Geophysics*, 78(2):O33–O43, 2013. doi: 10.1190/geo2012-0331.1.
- A. A. Henza, M. O. Withjack, and R. W. Schlische. How do the properties of a pre-existing normal-fault population influence fault development during a subsequent phase of extension? *Journal of Structural Geology*, 33(9):1312–1324, 2011. doi: 10.1016/j.jsg.2011.06.010.
- L. Holden, P. Mostad, B. F. Nielsen, J. Gjerde, C. Townsend, and S. Ottesen. Stochastic structural modeling. *Mathematical Geology*, 35(8):899–914, 2003.
- K. Hollund, P. Mostad, B. F. Nielsen, L. Holden, J. Gjerde, M. G. Contursi, A. J. McCann, C. Townsend, and E. Sverdrup. Havana - a fault modeling tool. *Norwegian Petroleum Society Special Publications*, 11:157–171, 2002.
- D. Iacopini, R. Butler, S. Purves, N. McArdle, and N. De Freslon. Exploring the seismic expression of fault zones in 3D seismic volumes. *Journal of Structural Geology*, 89:54–73, 2016. doi: 10.1016/j.jsg.2016.05.005.
- M. W. Jessell, L. Ailleres, and E. A. De Kemp. Towards an integrated inversion of geoscientific data: What price of geology? *Tectonophysics*, 490(3):294–306, 2010. doi: 10.1016/j.tecto.2010.05.020.
- C. Julio, G. Caumon, and M. Ford. Impact of the en echelon fault connectivity on reservoir flow simulations. *Interpretation*, 3(4):SAC23–SAC34, 2015a.



- doi: 10.1190/INT-2015-0060.1.
- C. Julio, G. Caumon, and M. Ford. Sampling the uncertainty associated with segmented normal fault interpretation using a stochastic downscaling method. *Tectonophysics*, 639:56–67, 2015b. doi: 10.1016/j.tecto.2014.11.013.
- J. T. Kent. The Fisher-Bingham distribution on the sphere. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 71–80, 1982.
- D. E. Knuth. *The Art of Computer Programming, Volume 4: Generating all Combinations and Partitions, Fascicle 3*, 2005.
- S. Lacaze, B. Luquet, T. Valding, and F. Pauget. Enhanced fault imaging from seismic and geological model. In *SEG Technical Program Expanded Abstracts 2016*, pages 1971–1975. 2016.
- I. Manighetti, D. Zigone, M. Campillo, and F. Cotton. Self-similarity of the largest-scale segmentation of the faults: Implications for earthquake behavior. *Earth and Planetary Science Letters*, 288(3):370–381, 2009. doi: 10.1016/j.epsl.2009.09.040.
- I. Manighetti, C. Caulet, L. Barros, C. Perrin, F. Cappa, and Y. Gaudemer. Generic along-strike segmentation of Afar normal faults, East Africa: Implications on fault growth and stress heterogeneity on seismogenic fault planes. *Geochemistry, Geophysics, Geosystems*, 16(2):443–467, 2015.
- C. J. Mann. Uncertainty in geology. In *Computers in geology—25 years of progress*, pages 241–254. Oxford University Press, Inc., 1993.
- C. W. Nixon, D. J. Sanderson, and J. M. Bull. Deformation within a strike-slip fault network at Westward Ho!, Devon UK: Domino vs conjugate faulting. *Journal of Structural Geology*, 33(5):833–843, 2011.
- F. L. Richards, N. J. Richardson, C. E. Bond, and M. Cowgill. Interpretational variability of structural traps: implications for exploration risk and volume uncertainty. *Geological Society, London, Special Publications*, 421(1):7–27, 2015.
- J. C. Rivenæs, C. Otterlei, E. Zachariassen, C. Dart, and J. Sjøholm. A 3D stochastic model integrating depth, fault and property uncertainty for planning robust wells, Njord Field, offshore Norway. *Petroleum Geoscience*, 11(1):57–65, 2005. doi: 10.1144/1354-079303-612.
- M. S. Rosenbaum and M. G. Culshaw. Communicating the risks arising from geohazards. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 166(2):261–270, 2003.
- D. J. Sanderson and C. W. Nixon. The use of topology in fracture network characterization. *Journal of Structural Geology*, 72:55–66, 2015.
- E. Santiago, J. X. Velasco-Hernández, and M. Romero-Salcedo. A descriptive study of fracture networks in rocks using complex network metrics. *Computers & Geosciences*, 88:97–114, 2016. doi: 10.1016/j.cageo.2015.12.021.
- S. E. Schaeffer. Graph clustering. *Computer science review*, 1(1):27–64, 2007. doi: 10.1016/j.cosrev.2007.05.001.
- R. Schneeberger, M. de La Varga, D. Egli, A. Berger, F. Kober, F. Wellmann, and M. Herwegh. Methods and uncertainty-estimations of 3D structural modelling in crystalline rocks: A case study. *Solid Earth Discuss.*, 2017.

- 
- S. T. Thiele, M. W. Jessell, M. Lindsay, V. Ogarko, J. F. Wellmann, and E. Pakyuz-Charrier. The topology of geology 1: Topological analysis. Journal of Structural Geology, 2016a. doi: 10.1016/j.jsg.2016.08.009.
- S. T. Thiele, M. W. Jessell, M. Lindsay, V. Ogarko, J. F. Wellmann, and E. Pakyuz-Charrier. The topology of geology 2: Topological analysis. Journal of Structural Geology, 2016b. doi: 10.1016/j.jsg.2016.08.010.
- B. Thomson. Topology vs. topography: Sometimes less is more. In Second EAGE Integrated Reservoir Modelling Conference, 2014. doi: 10.3997/2214-4609.20147461.
- P. Thore, A. Shtuka, M. Lecour, T. Ait-Ettajer, and R. Cognot. Structural uncertainties: Determination, management, and applications. Geophysics, 67(3):840–852, 2002. doi: 10.1190/1.1484528.
- E. Tomita, A. Tanaka, and H. Takahashi. The worst-case time complexity for generating all maximal cliques and computational experiments. Theoretical Computer Science, 363(1):28–42, 2006.
- A. Torabi and S. S. Berg. Scaling of fault attributes: A review. Marine and Petroleum Geology, 28(8):1444–1460, 2011. doi: 10.1016/j.marpetgeo.2011.04.003.
- L. Valentini, D. Perugini, and G. Poli. The "small-world" topology of rock fracture networks. Physica A: Statistical Mechanics and its Applications, 377(1):323–328, 2007.
- B. Walter. Réservoirs de socle en contexte extensif : Genèse, géométries et circulations de fluides. Exemples du rift intracontinental du Lac Albert (Ouganda) et de la marge proximale d’Ifni (Maroc). PhD thesis, Université de Lorraine, 2016.
- W. Weinzierl, L. Schulte, and V. Aarre. Volumetric of fault perturbation in the first Fresnel zone. Interpretation, 4(4):T419–T426, 2016. doi: 10.1190/INT-2015-0192.1.
- J. F. Wellmann, F. G. Horowitz, E. Schill, and K. Regenauer-Lieb. Towards incorporating uncertainty of structural data in 3D geological inversion. Tectonophysics, 490(3):141–151, 2010. doi: 10.1016/j.tecto.2010.04.022.
- X. Wu and Z. Zhu. Methods to enhance seismic faults and construct fault surfaces. Computers & Geosciences, 107:37–48, 2017.
- G. Yielding. The geometry of branch lines. Geological Society, London, Special Publications, 439:SP439–1, 2016. doi: 10.1144/SP439.1.
- H. Yin, J. Zhang, L. Meng, Y. Liu, and S. Xu. Discrete element modeling of the faulting in the sedimentary cover above an active salt diapir. Journal of Structural Geology, 31(9):989–995, 2009. doi: 10.1016/j.jsg.2008.10.007.
- S. Zhu, R. Hack, K. Turner, and M. Hale. How far will uncertainty of the subsurface limit the sustainability planning of the subsurface? In Proc. Sustainable Development & Management of the Subsurface (SDMS) conference, pages 5–7, 2003.

---

**latexcount:** 2821 words in text + 77 in headers +391 in captions