JOURNAL OF GEOPHYSICAL RESEARCH: SOLID EARTH, VOL. 118, 2865–2877, doi:10.1002/jgrb.50178, 2013

Earthquake clusters in southern California II: Classification and relation to physical properties of the crust

Ilya Zaliapin[1] and Yehuda Ben-Zion[2]

Received 9 December 2012; revised 20 March 2013; accepted 28 March 2013; published 10 June 2013.

[1] This is a second paper in a study of statistical identification and classification of earthquake clusters using a relocated catalog of 1981–2011 seismicity in southern California and synthetic catalogs produced by the Epidemic Type Aftershock Sequence model. Here we focus on classification of event *families*—statistically significant clusters composed of *foreshocks, mainshocks*, and *aftershocks*—that are detected with the methodology discussed in part I of the study. The families are analyzed using their representation as time oriented tree graphs. The results (1) demonstrate that the clustering associated with the largest earthquakes, $m > 7$, is statistically different from that of small-to-medium earthquakes; (2) establish the existence of two dominant types of small-to-medium magnitude earthquake families—*burst-like* and *swarm-like sequences*—and a variety of intermediate cluster forms obtained as a mixture of the two dominant types; (3) suggest a simple new quantitative measure for identifying the cluster type based on its topological structure; (4) demonstrate systematic spatial variability of the cluster characteristics on a scale of tens of kilometers in relation to heat flow and other properties governing the effective viscosity of a region; and (5) establish correlation between the family topological structure and a dozen of metric properties traditionally considered in the literature (number of aftershocks, duration, spatial properties, *b*-value, parameters of Omori-Utsu and Båth law, etc.). The burst-like clusters likely reflect highly brittle failures in relatively cold regions, while the swarm-like clusters are likely associated with mixed brittle-ductile failures in regions with relatively high temperature and/or fluid content. The results of this and paper I may be used to develop improved region-specific hazard estimates and earthquake forecasts.

Citation: Zaliapin, I., and Y. Ben-Zion (2013), Earthquake clusters in southern California II: Classification and relation to physical properties of the crust, *J. Geophys. Res. Solid Earth*, *118*, 2865–2877, doi:10.1002/jgrb.50178.

1. Introduction

[2] Seismicity patterns are prime examples of complex dynamic phenomena with variations in space, time, and size over very broad ranges of scales [e.g., *Aki*, 1981; *Utsu*, 2002; *Keilis-Borok and Soloviev*, 2003; *Rundle et al.*, 2003; *Sornette*, 2004; *Ben-Zion*, 2008; *Shearer*, 2012]. Despite many years of research, robust methods for characterization of the dynamics of seismicity and basic understanding of the underlying processes are lacking. One fundamental unresolved problem is whether seismicity follows universal laws and reported differences stem only from statistical fluctuations [e.g., *Kagan*, 1994; *Bak*, 1996], or distinct seismic patterns reflect different governing properties of fault zones and crustal domains [e.g., *Mogi*, 1963; *Ben-Zion*, 2008]. This paper is the second in a series aiming to address this basic issue by developing and applying robust methods for detecting and analyzing earthquake clustering.

[3] Our specific goals are to: (1) identify statistically significant earthquake clusters understood in the broadest sense as deviations from a time-stationary space-inhomogeneous marked point process, (2) classify the detected clusters into several main types according to their statistical properties, and (3) relate the detected cluster types to key properties of the deforming region. The problem of cluster identification was addressed by *Zaliapin and Ben-Zion* [2013] referred to below as ZBZ13. This companion paper focuses on classification of the detected clusters and their relations to physical properties of the region.

[4] ZBZ13 applied the earthquake cluster methodology of *Zaliapin et al.* [2008] to identify statistically significant clusters in a relocated catalog of earthquakes in southern California [*Hauksson et al.*, 2012]. The methodology is based on a bimodal distribution of nearest-neighbor

This article is a companion to *Zaliapin and Ben-Zion* [2013] doi:10.1002/jgrb.50179.

Additional supporting information may be found in the online version of this article.

[1]Department of Mathematics and Statistics, University of Nevada, Reno, Nevada, USA.

[2]Department of Earth Sciences, University of Southern California, Los Angeles, California, USA.

Corresponding author: Ilya Zaliapin, Department of Mathematics and Statistics, University of Nevada, Reno, 89557, NV, USA. (zal@unr.edu)

2169-9313/13/10.1002/jgrb.50178

earthquake distances in a combined space-time-magnitude domain. The bimodality was shown to result from dependent space-time seismicity structures (associated primarily with foreshock-mainshock-aftershock sequences), as opposed to possible marginal independent time or spatial inhomogeneities. This intrinsic property of natural seismicity can be used for partitioning an examined catalog into separate *clusters*. The clusters are divided into *singles* that contain just one event, and *families* having multiple events that are subclassified into *foreshocks, mainshocks, and aftershocks* (section 3 of ZBZ13). We have demonstrated in ZBZ13, supporting information sections D, E, using the observed seismicity and the Epidemic Type Aftershock Sequence (ETAS) model that the employed cluster detection method is accurate and robust with respect to (1) the three numerical parameters of the algorithm, (2) minimum reported magnitude, (3) catalog incompleteness, and (4) location errors.

[5] Earthquake families (multi-event clusters) are the main object of this study. We demonstrate that the detected families of seismicity in southern California can be classified into two basic types and a third special class. These correspond to (1) *burst-like sequences* dominated by first-generation offspring of mainshocks and having additional spatiotemporal and internal topological properties consistent with highly brittle behavior; (2) *swarm-like sequences* composed of many generations of events and having an additional set of properties consistent with mixed brittle-ductile behavior; and (3) the most prominent (and least frequent) sequences with mainshocks size $m \geq 7$ that consist likely of combined populations of types (1) and (2).

[6] Traditional studies of earthquake sequences focus on class (3) which are the most conspicuous and most relevant for hazard. The aftershocks of the largest mainshocks are very numerous so their identification can be done effectively with simple window methods [e.g., *Gardner and Knopoff*, 1974]. However, the detection and analysis of the multitudinous ongoing burst- and swarm-like families, each of which contains a smaller number of events, require refined methods. The analysis done in this paper shows that family types (1) and (2) have clear separate identity; these two types provide the basic building blocks for making the largest clusters.

[7] We demonstrate that burst- and swarm-like sequences have distinct preferred geographic locations (section 3) and differ significantly according to a dozen of examined cluster statistics (supporting information section C). Importantly, the family type can be efficiently quantified by a simple scalar measure related to its topological structure—the *average leaf depth* introduced in section 2, which is significantly correlated with a dozen of other cluster statistics (supporting information section C)—including sequence duration, area, magnitude distribution, parameters of Omori-Utsu and Båth laws, and angular surface distribution of family members. The average leaf depth seems to provide the best association of different cluster types with local geographic regions characterized by different levels of heat flow and other properties governing the effective viscosity of the region [*Ben-Zion and Lyakhovsky*, 2006]. The results of this study are consistent with existing knowledge on properties of the largest aftershock sequences and prominent swarms, while providing additional analysis tools and further insight on seismicity patterns.

2. Two Types of Small-to-Medium Magnitude Earthquake Clusters

[8] The employed earthquake catalog and Δ-*analysis* of seismicity that uses data with fixed range of magnitude differences are described in section 2 of ZBZ13. The method of cluster identification is described in section 3 and supporting information sections A, B of ZBZ13. Here we demonstrate the existence of two dominant types of small-to-medium magnitude families; the types are quantified via the topologic structure of the time-oriented trees that represent families. As mentioned above, one type corresponds to *burst-like clustering*, and the other corresponds to *swarm-like clustering*. We start with two visual case studies that help developing intuition and then continue with systematic quantitative analysis of the entire catalog.

2.1. First Case Study: A Simple Example

[9] To illustrate the two dominant types of earthquake families in their clearest and simplest form, we perform the nearest-neighbor analysis of earthquakes with $m \geq 4$ in southern California and consider two families with relatively small number of events (Figure 1). The first family (Figures 1a, 1b, and 1c) occurred in the Coso area and contains five earthquakes; its mainshock has magnitude $m = 5.12$ and coordinates (35.7360 N, 117.7558 W). The second family (Figures 1d, 1e, and 1f) occurred in the San Bernardino region and contains eight earthquakes; its mainshock has magnitude $m = 4.99$ and coordinates (34.3074 N, 116.8437 W). Despite the similarity in the number of events and the mainshock magnitude, the structures of the two families differ dramatically. The Coso family (Figures 1a, 1b, and 1c) has a *linear* topologic structure—each event is the child of the immediately previous one, and its largest event happens in the middle of the sequence. In contrast, the San Bernardino family (Figures 1d, 1e, and 1f) has a *spray-shaped* topologic structure—six out of seven aftershocks in this cluster are children of the mainshock, which happens first in the sequence. In addition, the spatial extent of the Coso family (measured by the area of the convex hull that spans the family events) is somewhat larger than that of the San Bernardino family. Notably, the essential difference between the two families is nicely captured by their *topologic tree graph* (Figures 1c and 1f); such a tree describes the order in which events are connected within the cluster while ignoring all the metric properties (magnitude, location, etc.).

[10] We note that the linear Coso sequence in the above example occurred in an area characterized by relatively high heat flow, while the spray-like San Bernardino sequence occurred in an area with relatively low heat flow. The San Bernardino sequence exhibits an abrupt *brittle* behavior with no foreshock activity and a burst of aftershocks, while the Coso sequence behaves more gradually in a swarm-like fashion, with equally active foreshock and aftershock subsequences. These differences are consistent with theoretical expectations based on a viscoelastic damage rheology model [*Ben-Zion and Lyakhovsky*, 2006]. The spatial distributions of events in the two clusters (Figures 1b and 1e) also appear to have informative differences consistent with other properties of the clusters. These are seen more clearly in the next example.

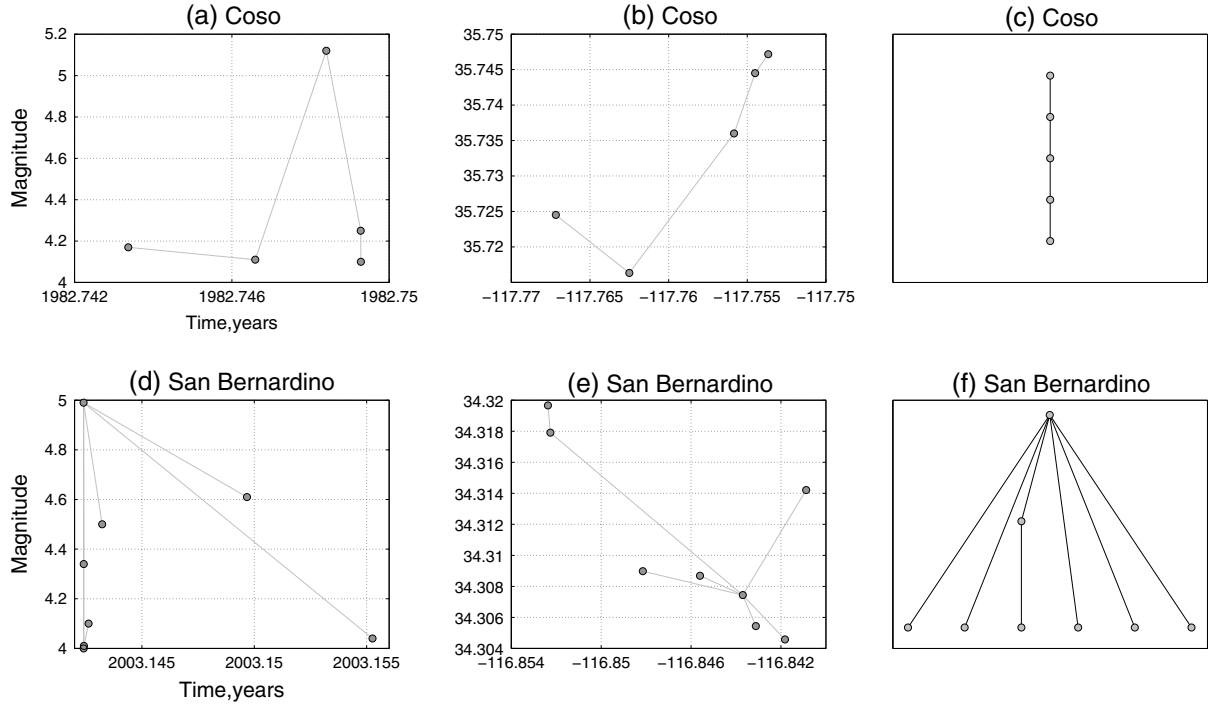

Figure 1. Two types of earthquake families: a simple example. Circles correspond to earthquakes, lines to the parent links. (a,b,c) Family in Coso area. (d,e,f) Family in San Bernardino area. (a,d) Magnitude vs. time. (b,e) Space map. (c,f) Topologic tree.

2.2. Second Case Study: Large Families

[11] The differences between the two dominant family types are visually more prominent for clusters of much larger size. Here we examine two families detected in the nearest-neighbor analysis with magnitude cutoff $m_c = 2$. The first family is located in the Salton trough area (Figures 2a, 2b, and 2c). It has mainshock magnitude $m = 5.75$ and coordinates (33.0875 N, 115.6195 W). The second family is in the San Gabriel region (Figures 2d, 2e, and 2f). It has magnitude $m = 5.51$ and coordinates (34.1380 N, 117.7082 W); it occurred near Claremont, CA.

[12] The Salton trough family (Figures 2a, 2b, and 2c) has 315 events, with 81 foreshocks and 233 aftershocks. Topologically (Figure 2c), the family consists of multiple chains and one dominant burst that includes 136 events (43% of the cluster). The San Gabriel family (Figures 2d, 2e, and 2f) has 400 events, with three foreshocks and 396 aftershocks. Topologically (Figure 2f), this family is mainly composed of a burst that includes 261 events (65%). The spatial extent of the San Gabriel family (Figure 2e) is much smaller (note the different scales in the plots) than that of the Salton trough family (Figure 2b). We also observe that the San Gabriel family has an approximate isotropic shape (reminiscent of explosion) consistent with a highly brittle failure process. In contrast, the Salton trough family is concentrated in a small number of directions (suggesting flow-like failure in specific channels) consistent with partially ductile-like behavior.

[13] Similarly to the previous example, the sequence in the area with higher heat flow (Salton trough) demonstrates more swarm-like behavior (although it may not be a classical swarm), characterized by increased foreshock production, larger spatial extent, and smaller temporal concentration. In contrast, the sequence in the region with lower heat flow (San Gabriel) follows a clear burst-like behavior, characterized by a small number of foreshocks, dominance of first-generation aftershocks, and high concentration of events in space and time. A more detailed analysis of these two families in supporting information section A shows that the cluster structure is stable with respect to the magnitude threshold of the analysis.

2.3. Quantitative Analysis: Average Leaf Depth

[14] We now introduce a quantitative topological approach for identification of the earthquake cluster type. To quantify the differences between the topologic trees illustrated in sections 2.1, 2.2, we use the concept of *vertex depth*—the minimal number d of links that connects a given vertex (earthquake) to the tree root (the first earthquake in the family). A useful scalar measure suitable for the current analysis is the *average leaf depth* $\langle d \rangle$—the vertex depth d averaged over the tree leaves (vertices with no children). This definition is illustrated in Figure 3, where we compute the average leaf depth $\langle d \rangle$ for the relatively small trees from section 2.1, Figures 1c and 1f. The simple intuition behind this definition is that a linearly shaped tree should have much larger depth than a spray-shaped tree with the same number of leaves. Indeed, the tree from the Coso family has $\langle d \rangle = 4$, while the tree from the San Bernardino family has $\langle d \rangle = 1.17$.

[15] An important property of the average leaf depth can be established using the following branching model of a cluster. We assume that (1) initially (at $t = 0$), there exists a single element, and (2) each element that exists at integer

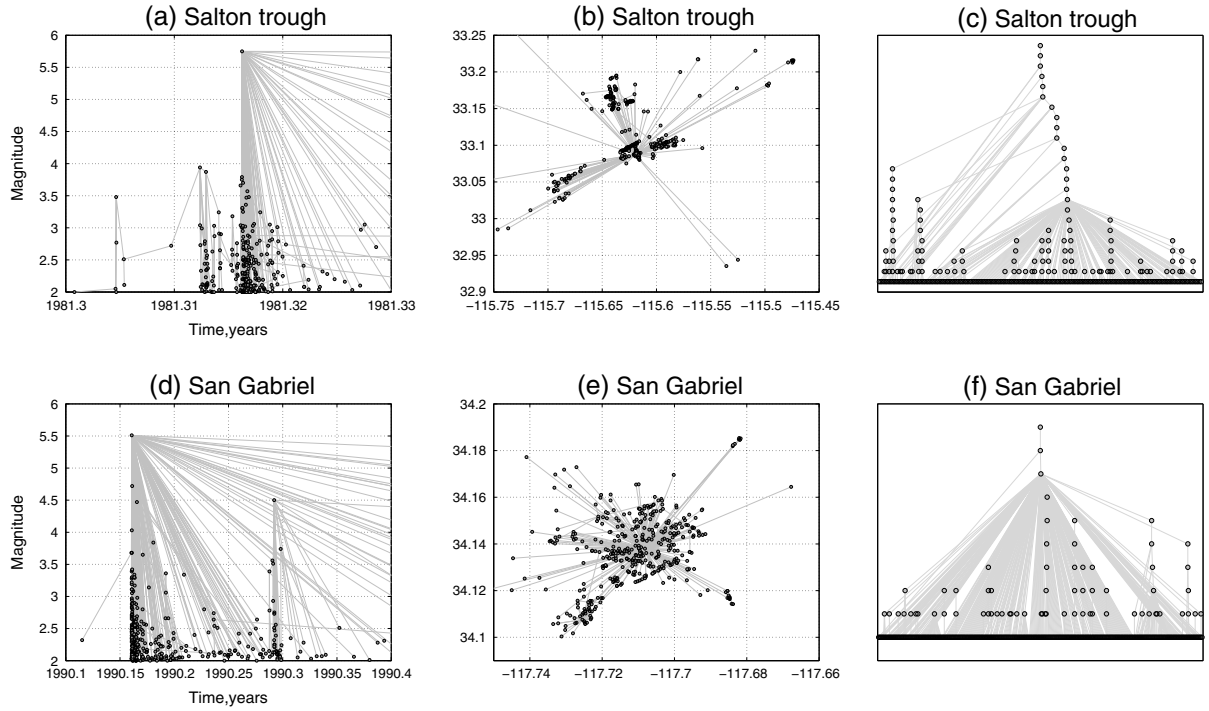

Figure 2. Two types of earthquake families: an example. Circles correspond to earthquakes, lines to parent links. (a,b,c) Family in Salton trough area. (d,e,f) Family in San Gabriel area. (a,d) Magnitude vs. time. (b,e) Space map. (c,f) Topologic tree.

step $t \geq 0$ branches into a random number $Z \geq 0$ of offspring according to a distribution $g(k) = P(Z = k)$, $k \geq 0$ independently of the other elements. This model is known as a Galton-Watson branching process with a single progenitor [*Watson and Galton*, 1875; *Pitman*, 2006]. We focus on finite clusters produced by this model; see *Pitman* [2006] for details. It is readily seen that each leaf at depth k is formed because each of its k ancestors produced some offspring while it produces no offspring. Accordingly, the depth d of a randomly chosen leaf of a finite cluster is given by the random geometric variable $P(d = k) = g(0)(1- g(0))^k$, where $g(0)$ is the probability of having no offspring. The mean leaf depth in this model is hence $(1\text{-}g(0))/g(0)$. The depths of different leaves are dependent, since some of the leaves have multiple common ancestors. This dependence, however, does not affect the average of the leaf depths, which, for a cluster with n leaves of depth d_i, $i = 1,\ldots,n$, is given by

$$E[\langle d \rangle] = E\left[\frac{1}{n}\sum_{i=1}^{n} d_i\right] = \frac{1 - g(0)}{g(0)}. \qquad (1)$$

[16] This suggests, in addition, that the average leaf depth $\langle d \rangle$ should only weakly depend on the observed number of leaves in a cluster. A detailed analysis of the average leaf depth in a branching model will be performed elsewhere.

[17] The ETAS model of seismicity (supporting information section D) creates clusters according to the Galton-Watson process described above. Specifically, each cluster starts with a background event. The "internal" discrete time of a cluster corresponds to consecutive generations of aftershocks (the first event is zero generation, etc.). Each event in a cluster generates a random number of offspring according to a Poisson distribution with parameter depending on the event magnitude. Randomization with respect to the event magnitude (recall that ETAS assigns magnitudes independently according to the Gutenberg-Richter law) leads to an ETAS-specific offspring distribution $g_{\mathrm{ETAS}}(k)$. The exact form of this law is of no importance to our current work, although it can be readily derived. Empirical estimations using the examined catalog of earthquakes in southern California suggest $g_{\mathrm{ETAS}}(0) \approx 0.761 \pm 0.003$. In this estimation, we only considered the offspring (children) from the clustered part of the bimodal distribution of the nearest-neighbor event distances (see ZBZ13, section 3), which corresponds to counting only offspring within individual families. Counting all the offspring lowers the estimation to $g_{\mathrm{ETAS}}(0) \approx 0.663 \pm 0.003$, while counting only the offspring of internal family events (hence excluding singles and first family events from potential parents) gives $g_{\mathrm{ETAS}}(0) \approx 0.713 \pm 0.003$. These estimations provide feasible probabilities of having no offspring for the observed catalog. The average leaf depth that corresponds to $g_{\mathrm{ETAS}}(0) \approx 0.761$ is $\langle d \rangle = 0.31$. The corresponding probability of having a leaf of depth 5 is of the order of 10^{-4}. This implies that such depths are highly unlikely in a branching model, and, hence, the average leaf depth must be much smaller than 5 for the majority of the clusters.

[18] Figure 4 illustrates significant deviations of the average leaf depth distribution in the observed clusters with respect to the expectations from a branching model. Figure 4a shows the average leaf depth $\langle d \rangle$ vs. cluster size (number of events) N for 41,393 clusters in the examined catalog. Most of the clusters (34,836) are singles with size $N = 1$ and depth $\langle d \rangle = 0$ (see Table 1 of ZBZ13), so the informative part of this plot contains much smaller number of points. The first

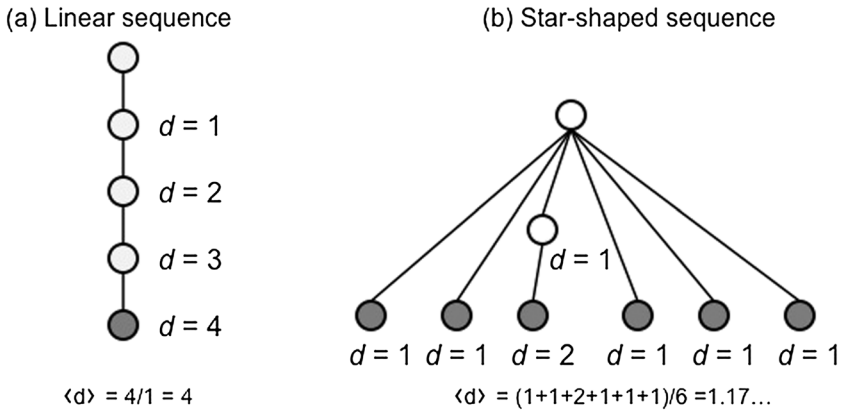

Figure 3. Average leaf depth ‹*d*› of a tree: an illustration. The figure shows how the average leaf depth is computed for a linear tree (panel a) and a spray-shaped tree (panel b). Leaves of the tree are shown by shaded circles, other vertices—by open circles. The vertex depths *d* are indicated in the figure. The examined trees are taken from the case study illustrated in section 2.1, Figures 1c and 1f.

observation is that the three families with the largest mainshocks—Landers M7.3, Hector Mine M7.1, and El Mayor-Cucapah M7.2 (big black circles) —are statistically different from the rest of the clusters (small gray circles). This group of three families corresponds to the most conspicuous aftershock sequences, which have cluster properties different from that of the other 41,390 clusters. The above branching analysis implies that it is highly improbable for a family in a model such as ETAS to have an average leaf depth higher than 5. However, the depths of the three largest families are, respectively, 11.6, 6.0, and 9.2. We emphasize that these families are composed of 14,622, 5689, and 7495 events and have 10,311, 4630, and 5724 leaves, respectively. Accordingly, the relatively high values of the estimated average leaf depths are not artifacts of sample fluctuations. The detailed behavior of the families of the largest events deserves further study.

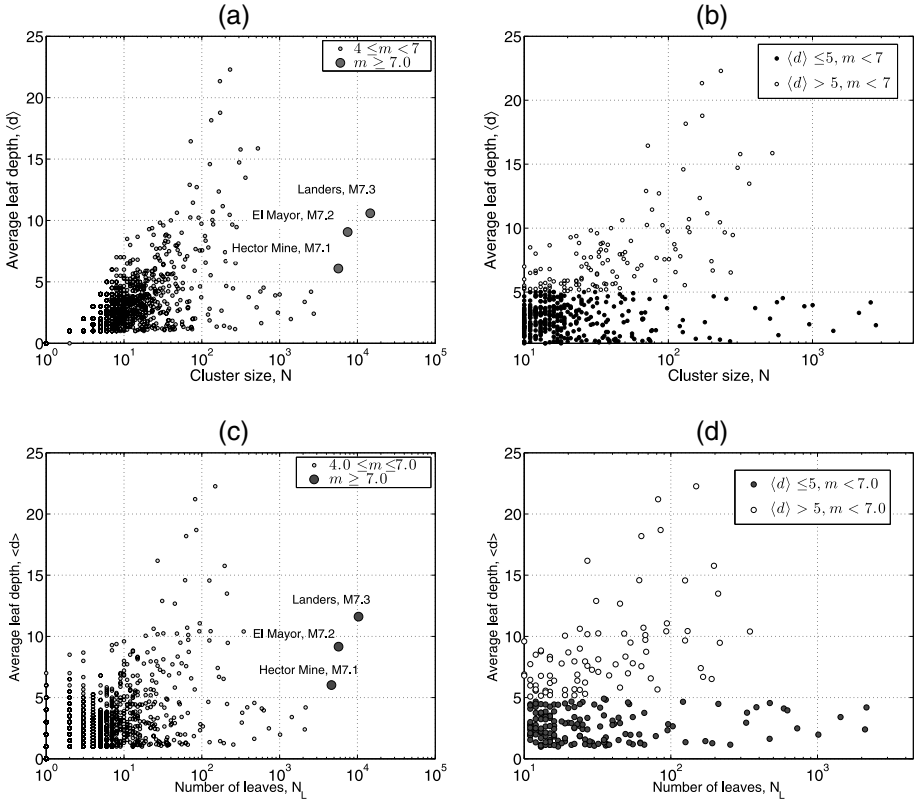

Figure 4. Average leaf depth ‹*d*› vs. cluster size *N* (panels a, b) or number of leaves N_L (panels c, d). (a,c) All 41,393 clusters in the examined catalog. Clusters with mainshock magnitude $m \ge 7$ are shown by large dark circles. (b,d) A subset of 489 clusters with maximal magnitude $m < 7$ and size $N \ge 10$. The bimodal structure is emphasized by the color code explained in the legend.

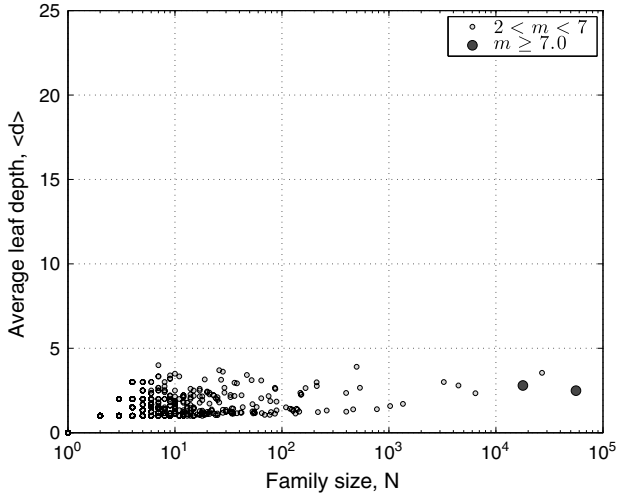

Figure 5. Average leaf depth ‹*d*› vs. cluster size *N* in ETAS model. The axes limits are the same as in Figure 4a. All families have low topological depth ‹*d*› independent of the family size *N*.

[19] We also see that small size clusters always have small average leaf depth, which is a natural consequence of the depth definition (depth ‹*d*› is always smaller than cluster size *N*). If the large-magnitude and small-size clusters are excluded, the remaining clusters exhibit a prominently bimodal behavior suggestive of two basic populations of families with different internal cluster organization. This is illustrated in Figure 4b with a zoom up on 489 families with mainshock magnitude $m < 7$ and family size $N \geq 10$. One population of families (black circles) has small average leaf depth, $\langle d \rangle \leq 5$, which is independent of family size *N* in the range between 10 and 3000. The other population (light circles) has larger average leaf depth, $5 < \langle d \rangle < 25$ that, furthermore, increases with *N*. These observations are strengthened by analysis of Figures 4c and 4d, which show the average leaf depth ‹*d*› vs. the number N_L of leaves in a family, as in the forgoing branching argument. The main difference of this analysis is that, in accordance with a branching model, there is no strong dependence of ‹*d*› on N_L for small leaf numbers.

[20] As mentioned, the existence of leaves with depth over 5 is highly improbable according to a branching model estimated from the data; nevertheless, such leaves are numerous in the observed catalog. We conjecture that clusters with large values of ‹*d*› are related to intrinsic deviations of the observed clustering mechanism from that of a pure branching model (e.g., ETAS); accordingly, there exists an alternative clustering mechanism that is responsible for a substantial population of clusters with large average leaf depth that increases with the cluster size. Despite the fact that the end-members of the two populations are dramatically different, there is no sharp boundary between the two populations, and the smaller is the family size, the harder it is to determine the family type. Nevertheless, as demonstrated below the recognition of two end-member family types may largely facilitate the study of seismicity clusters.

[21] Figure 5 shows for comparison the average leaf depth ‹*d*› as a function of family size *N* for a synthetic catalog of length 146,432 produced by a homogeneous isotropic ETAS model with parameters derived for seismicity in southern California. This is the same synthetic catalog used in ZBZ13, and it is described briefly in supporting information section D. For visual convenience, the axes limits in Figure 5 are the same as in Figure 4a. The most striking qualitative difference is the absence of the mode for which the topological depth increases with the family size; all families in the ETAS model have low topological depth $\langle d \rangle < 5$, independent of the family size. In addition, the largest events with magnitude $m > 7$ do not demonstrate behavior distinct from the rest of the clusters. This is in agreement with the argument based on a branching model and further supports our claim on the existence of an alternative clustering mechanism in the observed seismicity.

[22] Figure 6 shows the distribution of the average leaf depth ‹*d*› in regular analysis (as opposed to Δ-analysis) for families with mainshock magnitude $m \geq 4$ and size $N > 1$; it provides a useful reference for our later analyses that consider families in different percentile groups according to the value of the average leaf depth ‹*d*›. An exponential tail of the distribution (i.e., almost linear behavior of the empirical tail on the semilogarithmic scale) implies that most of the families have small depth ‹*d*› and hence exhibit explosive burst-like behavior, while a smaller number of families have large depth indicative of a swarm-type behavior. This is consistent with previous observations that the number of swarm-like sequences in southern California is relatively small [e.g., *Vidale and Shearer*, 2006; *Shearer*, 2012].

[23] Supporting information section B discusses an alternative approach to quantifying the family type, using the average leaf depth normalized by the family size *N*. That approach gives the same qualitative results as the one considered in the main text.

3. Spatial Variations of Cluster Properties

[24] Here we show that the essential cluster statistics are characterized by substantial spatial variability, which cannot be attributed to spurious random fluctuations, within South-

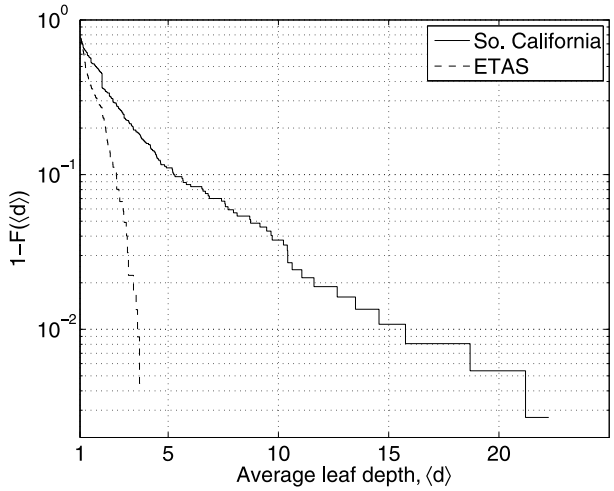

Figure 6. Distribution of the average leaf depth ‹*d*› for regular families with mainshock magnitude $m \geq 4$ and size $N > 1$ in the observed catalog (solid line) and ETAS model (dashed line). $F(x)$ denotes the cumulative distribution function.

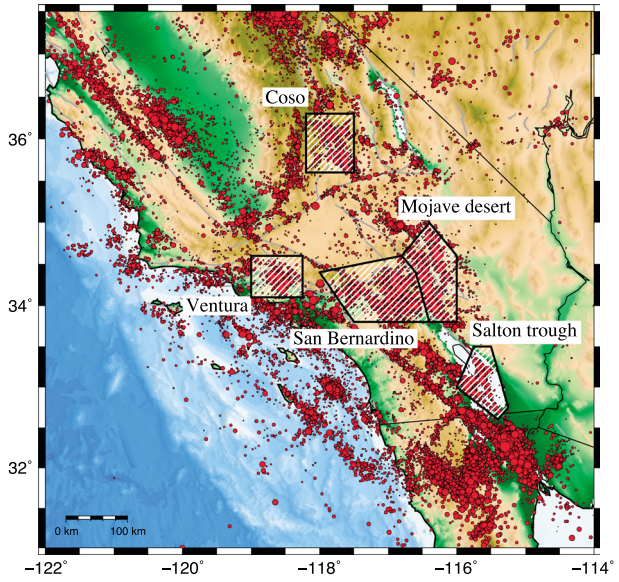

Figure 7. Seismicity of southern California and five special study regions: Coso, Ventura, San Bernardino/San Gabriel, Mojave Desert, and Salton trough. Earthquakes with $m \geq 2$ are shown by circles whose radius is proportional to the magnitude. Main faults are shown by gray lines.

ern California and that the observed variability is closely connected to the characterization of a region in terms of its effective viscosity.

3.1. Special Study Regions

[25] The main cluster analysis in this study is done uniformly over the entire examined space-time domain, without preselection of special fault zones or study areas. However, in order to emphasize the coupling between properties of seismicity and local physical characteristics of the crust, in section 3.3, we pay special attention to five regions that have both high level of seismic activity and well-documented physical characteristics—the Ventura basin, San Gabriel and San Bernardino basins, Mojave desert, Coso, and Salton trough. These five regions were examined previously by *Yang and Ben-Zion* [2009] and are shown in Figure 7. They contain 44% of all the earthquakes in the examined catalog and 26% of all the mainshocks. Figure 8 displays the averaged heat flow (panel a) and averaged hypocentral depth (panel b) for the families with $N \geq 10$ and mainshock magnitude $m \geq 4$; the results provide partial justification for selecting the above five regions. Two of the regions—Salton trough and Coso—are characterized by geothermal activity associated with high heat flow and fluid content, along with relatively thin seismogenic zone manifested by relatively shallow seismicity. These regions, referred to for brevity as "hot", can be characterized mechanically as having decreased level of effective viscosity. Three of the regions—Mojave, San Bernardino, and Ventura—are characterized by little-to-no geothermal activity, decreased heat flow, and relatively thick seismogenic zone with higher seismicity depth. These regions, referred to for brevity as "cold", can be characterized as having increased level of effective viscosity. To further emphasize the prominent differences between the cold and hot regions, the subsequent Figure 10d shows the comparative boxplots of the heat flow within the cold and hot regions.

3.2. Analysis for Entire Southern California

[26] In this section, we perform analysis of spatial distribution of various family statistics for families with mainshock magnitude $m \geq 4.0$ and size $N \geq 10$. The magnitude threshold is established so that the mainshock magnitudes are at least two units above the lowest reported magnitude ($m_{\text{cut}} = 2$). The size threshold is necessary for a meaningful computation of family statistics—small families by construction may only have restricted values for many statistics, e.g. average leaf depth, branching index, etc. The following analysis is mostly visual; quantitative analysis is done below in section 3.3.

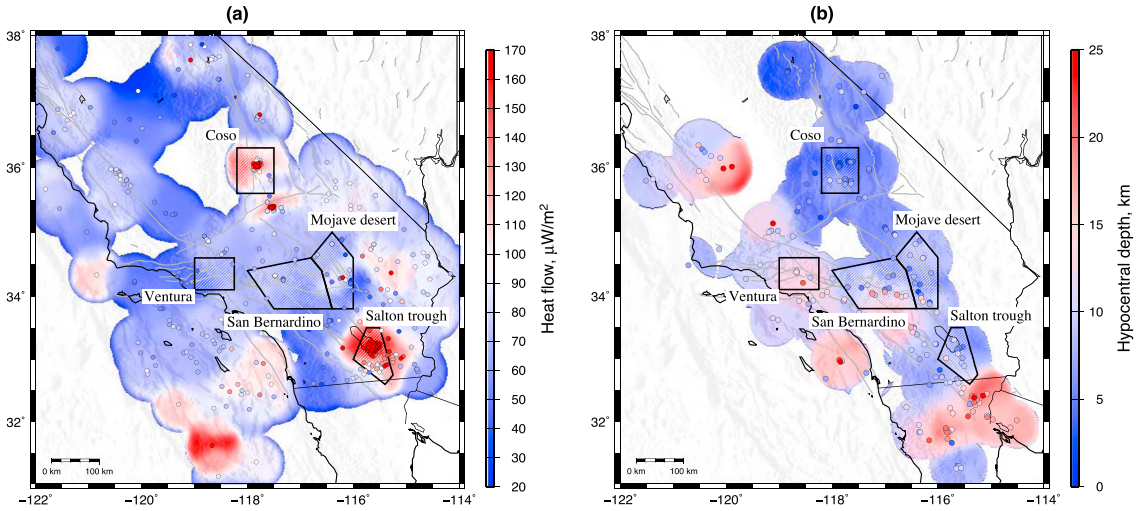

Figure 8. (a) Heat flow in southern California. Colored circles denote measurements at individual wells; color-code represents smoothed heat flow values; data from http:/smu.edu/geothermal [*Blackwell and Richards*, 2004]. (b) Hypocentral depth for the families with mainshock magnitude $m \geq 4$ and $N \geq 10$. In both panels, gray lines depict the major faults; black lines outline five special study regions.

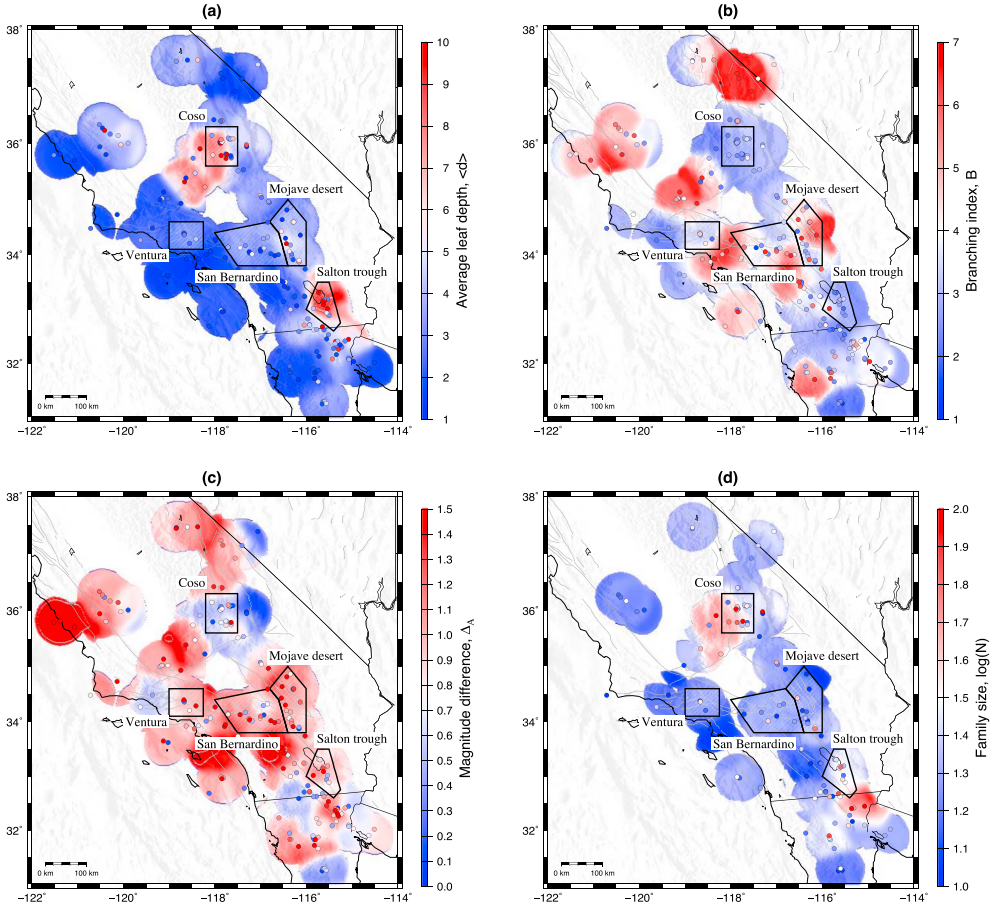

Figure 9. Maps of family statistics. Individual family statistics are shown by colored circles; average values are shown by color code. The analysis is done for all families with mainshock magnitude $m \geq 4$ and size $N \geq 10$. Gray lines show the major faults; shaded polynomials depict special study regions. (a) Average leaf depth ‹*d*›; (b) Family branching number, *B*; (c) Magnitude difference between mainshock and the largest aftershock, Δ_A and (d) Family size *N* on logarithmic scale. Panels (a–c) refer to 195 regular families, panel (d) to 112 Δ-families.

[27] Figure 9 shows the spatially averaged values for the four family statistics. The spatial averaging is done for the five spatially nearest families, or for all families within a circle of 50 km radius, whichever is smaller. Spatial locations that have no data points (family mainshock) within a 50 km circle are not analyzed. The values of family statistics are assigned to the epicenter of the family mainshock. The boundaries of the special study regions are shown in each panel. Next, we discuss the results for each of the examined statistics.

[28] The *Average leaf depth* ‹*d*› (Figure 9a) exhibits a notable spatial variability, with larger values of ‹*d*› > 5 being concentrated in two regions: Salton trough and Coso including its southwestern vicinity. The rest of southern California is characterized by smaller values of ‹*d*›; the lowest values of the topological depth are seen in the Ventura basin and its vicinity. Figure 10a shows the comparative boxplots of the average leaf depth ‹*d*› within the cold (Ventura, San Gabriel/San Bernardino, Mojave) and hot (Coso, Salton trough) regions described above (section 3.1, Figure 7). It further confirms the visual impression from Figure 9a by demonstrating that the distribution of ‹*d*› within the hot regions is shifted toward high values, compared to the distribution of ‹*d*› within the cold regions. The boxplots also show that the observed difference has statistical rather than deterministic nature, for instance low values of ‹*d*› can be seen in both cold and hot regions.

[29] The *Family branching number B* (Figure 9b) is defined as the average number of offspring over all earthquakes in the family that have at least one offspring (so the terminal events with zero offspring are not counted). The family branching number has several notable spikes (red regions in the figure), related to several families with particularly large value of *B*. Notably, the Salton trough and Coso regions are characterized by consistently small values of the branching index, $B < 3$ on average. Comparing Figures 9a and 9b, one may notice that the topological depth ‹*d*› and the branching index *B* seem to be anticorrelated. This impression is confirmed by formal analysis in supporting information section C. Figure 10b shows the comparative boxplots of the branching index *B* in the cold and hot regions.

[30] The *Magnitude difference* Δ_A between the mainshock and the largest aftershock (Figure 9c) exhibits spatial variability reminiscent of that seen for the branching

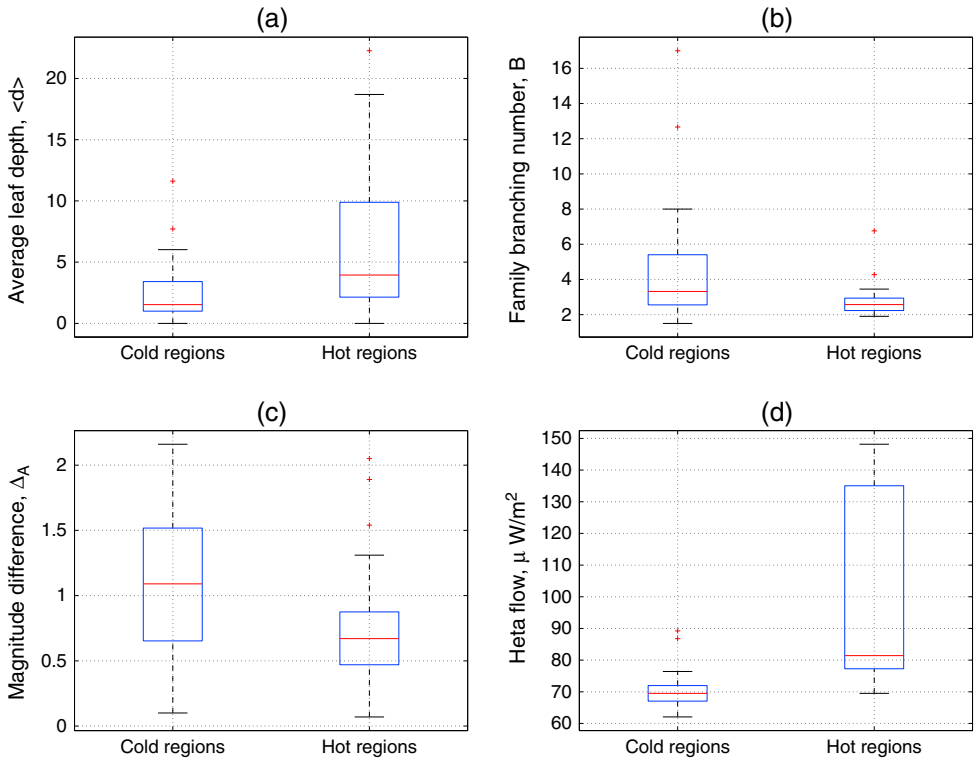

Figure 10. Family statistics in cold vs. hot special study regions. Boxplots of (a) average leaf depth, ‹*d*›; (b) family branching number, *B*; (c) magnitude difference between the mainshock and the largest aftershock, Δ_A; and (d) heat flow. Significance of the observed differences is assessed in Table 1.

number *B* (Figure 9b). The difference shows variability within the range 0.3 – 1.6, with higher values, $\Delta_A > 1$, in Mojave desert and San Gabriel regions and lower values, $\Delta_A < 1$, in Coso and Salton trough regions. Figure 10c shows the comparative boxplots of the magnitude difference Δ_A in cold and hot regions.

[31] The *Family size N* (Figure 9d) demonstrates systematic spatial fluctuations reminiscent of those for the average leaf depth in panel (a). The analysis is done for Δ-families, whose size is independent of the mainshock magnitude (see ZBZ13, Fig. 15). In particular, the family size is relatively large, $\log N > 1.5$, in the Salton trough and Coso area, and relatively small, $\log N < 1.5$ in Mojave desert, San Bernardino, and Ventura basin.

[32] Figure 10 uses statistical boxplots to further summarize the differences of selected statistics of families with mainshock magnitude $m \geq 4$ (panels a–c) and heat flow (panel d) in cold and hot regions. Each panel of the figure shows two boxplots for a particular examined statistic, one for the families in hot regions (right) and for the families in cold regions (left). Recall that a boxplot is a graphical summary of a sample: the box corresponds to the range between the first (Q1) and third (Q3) quartiles, the sample median is represented by the central horizontal line, the side bars extend to the 1.5 times the interquartile range Q1–Q3. The outliers (points outside of the sidebars) are shown individually. The figure clearly depicts the statistical differences between the cold and hot regions.

[33] The results of Figures 8–10 suggest that (1) seismicity clusters have different structures in different regions, (2) the character of clustering may change on the scale of tens of kilometers, and (3) the cluster properties vary systematically and are related to the effective viscosity of the crust. In the next section, we systematically explore this observation using the hypothesis testing approach.

3.3. Analysis for the Five Special Study Regions

[34] To show that the spatial variability of cluster statistics illustrated in section 3.2 is statistically significant, and hence is caused by physical differences of the deforming media rather than spurious statistical fluctuations, we perform here comparative analysis of various cluster statistics in the cold and hot special study regions defined in section 3.1 and Figure 7. Specifically, we use the Analysis of Variance (ANOVA) approach [*Freedman*, 2005] to test the null hypothesis that the average value of a particular family statistic *S* is the same within the combined cold and hot regions; the ANOVA approach is briefly reviewed in supporting information E. The results for 18 cluster statistics are summarized in Table 1. The table lists the examined statistic *S*; indicator of whether the test uses Δ-analysis (Δ = Y) or regular analysis (Δ = N) of clusters; the number of examined families in cold and hot regions; the values of pooled (using all available data) average in hot and cold regions; the value of the *F*-statistic; the corresponding *P*-value that is the probability that a random variable with *F*-distribution will be larger than the observed value of the *F*-statistic; and the final decision at the 5% level. Recall that under the null hypothesis and the standard assumptions of the ANOVA model, the *F*-statistic has *F*-distribution with $(1, n-1)$ degrees of freedom, where *n* is the total number of examined families. The ANOVA test is known to be fairly robust to mild deviations from its main

Table 1. ANOVA Test of the Null Hypothesis H_0: *The Average Value of a Family Statistic S Is the Same in Cold and Hot Regions* for Families With Mainshock Magnitude $m \geq 4$

#	Statistic, S	Δ	No. of Families		Pooled Average		F	P	Decision at 5% Level
			Cold	Hot	Cold	Hot			
1	No. of aftershocks, N_A	Y	66	37	7.36	20.89	12.89	5×10^{-4}	Reject
2	No. of foreshocks, N_F	Y	66	37	0.77	6.14	27.09	1×10^{-6}	Reject
3	Relative aftershock moment, M_A/M_M	Y	66	37	0.13	0.40	9.37	3×10^{-3}	Reject
4	Relative foreshock moment, M_F/M_M	Y	66	37	0.01	0.13	24.56	3×10^{-6}	Reject
5	Aftershock duration, $\log_{10}(D_A$ [years])	Y	55	34	−1.35	−0.92	2.69	0.10	Do not reject
6	Foreshock duration, $\log_{10}(D_F$ [years])	Y	21	26	−3.15	−2.32	7.22	2×10^{-2}	Reject
7	Aftershock distance, $\log_{10}(r_A$ [km])	Y	55	34	−0.05	−0.03	0.04	0.83	Do not reject
8	Foreshock distance, $\log_{10}(r_F$ [km])	Y	21	26	−0.67	−0.18	13.84	5×10^{-3}	Reject
9	Average leaf depth, $\langle d\rangle$	N	66	37	2.33	6.04	24.63	3×10^{-6}	Reject
10	Maximal leaf depth, d_{max}	N	66	37	4.70	10.43	15.96	1×10^{-4}	Reject
11	Mainshock depth, d_{main}	N	66	37	0.83	4.38	25.25	2×10^{-6}	Reject
12	Branching, B	N	60	35	3.79	2.74	4.97	2×10^{-2}	Reject
13	Aftershock magnitude difference, Δ_A	N	60	35	1.22	0.83	10.52	2×10^{-3}	Reject
14	Foreshock magnitude difference, Δ_F	N	27	27	1.69	0.93	17.72	1×10^{-4}	Reject
15	Circular homogeneity, $p_{KS} \geq 0.01$	N	66	37	0.39	0.27	1.59	0.21	Do not reject
16	Depth, z [km]	-	66	37	9.87	6.57	20.16	2×10^{-5}	Reject
17	Heat flow, [μW/m^2]	-	66	37	70.2	100.2	64.04	2×10^{-12}	Reject
18	Magnitude, m	-	66	37	4.55	4.50	0.16	0.69	Do not reject

assumptions, so the main conclusions of this test will remain valid for the examined data. In some cases, we consider transformed values of the cluster statistics to better satisfy the test assumptions.

[35] The analysis demonstrates that both aftershock and foreshock activity is significantly stronger in the hot regions. Specifically, the number of Δ-aftershocks (line 1 of Table 1) and Δ-foreshocks (line 2), and the relative aftershock (foreshock) moment defined as the ratio of the total aftershock (foreshock) moment $M_{A(F)}$ to the mainshock moment M_M (lines 3–4), are significantly larger within the hot regions. For instance, there is an average of 20.89 Δ-aftershocks per family in the hot regions vs. 7.36 in the cold regions; the total moment of aftershocks in the hot regions equals 40% of the mainshock moment vs. 13% in the cold regions. In addition, the Δ-foreshock duration (line 6) and average distance from a Δ-foreshock to the mainshock (line 8)—both considered on a logarithmic scale to better satisfy the assumptions of the ANOVA model—are significantly larger in the hot regions. The same statistics for Δ-aftershocks (lines 5 and 7) do not show significant differences, although the average values of the statistics are also larger within the hot regions.

[36] Next, the Bath's law parameter—the magnitude difference Δ_A between the mainshock and the largest aftershock (line 13) —is significantly smaller in the hot regions ($\Delta_A = 0.83$) than in the cold regions ($\Delta_A = 1.22$). A similar significant difference is seen for the foreshock magnitude difference Δ_F (line 14).

[37] Finally, the topological structure of the families differs significantly between the hot and cold regions. Specifically, the average leaf depth $\langle d\rangle$ (line 9) discussed in sections 2.3, 3.2, the maximal leaf depth d_{max} (line 10), the mainshock depth (the number of generations from the first foreshock to the mainshock) d_{main} (line 11) are all significantly larger within the hot regions. On the other hand, the family branching index B (line 12) is significantly lower within the hot regions. The negative association between the branching index and the topological depth is very intuitive, since under a fixed family size, those two parameters should be roughly reciprocal.

[38] Table 1 also analyzes the heat flow (line 17) and epicentral depth distribution (line 16) within the hot and cold regions, showing very significant differences for both these statistics.

[39] The analysis summarized in Table 1 is restricted to the families with mainshock magnitude $m \geq 4.0$. To show that this does not affect our overall conclusions, we perform ANOVA testing of a particular hypothesis H_0: *The average value of depth $\langle d\rangle$ is the same in hot and cold regions* (line 9 of Table 1) for a range of family selections; the results are summarized in Table 2. We see that the main conclusion—rejection of the null hypothesis—remains

Table 2. ANOVA Test of the Null Hypothesis H_0: *The Average Value of Depth $\langle d\rangle$ Is the Same in Hot and Cold Regions*

#	Family Thresholds	No. of Families			F	P-value	Decision at 5% Level
		Total	Hot	Cold			
1	$2 \leq m < 6$, $L \geq 10$	192	107	85	6.6	0.01	Reject
2	$m \geq 2$, $L \geq 10$	197	108	89	5.84	0.02	Reject
3	$3 \leq m < 6$, $L \geq 10$	157	89	68	5.92	0.02	Reject
4	$4 \leq m < 6$, $L \geq 10$	61	28	33	19.82	0.000039	Reject
5	$m \geq 4$, $L \geq 10$	66	29	37	17.22	0.0001	Reject
6	$4 \leq m < 6$, $L \geq 30$	37	22	15	11.74	0.002	Reject
7	$4 \leq m < 6$, $L \geq 50$	23	15	8	7.04	0.01	Reject

unchanged under a broad range of selection criteria, hence supporting the stability of our results.

4. Discussion

[40] This study attempts to provide a quantitative method for classification of the earthquake clusters detected in ZBZ13 in a detailed catalog of 1980–2011 earthquakes in southern California [*Hauksson et al*., 2012]. The classification is done with a variety of traditional and new measures. A simple and very informative characteristic for classification of clusters is the topological average leaf depth $\langle d \rangle$ of section 2. It provides a sharp spatial localization of clusters with distinct topological structures (Figures 9a and 10a) and is statistically correlated with a dozen of cluster statistics traditionally considered in aftershock studies (supporting information C, Table C1).

[41] The results are consistent generally with nonuniversal region-dependent seismicity patterns manifested by different seismicity cluster types and related statistical laws. Our comprehensive analysis indicates that the structure of earthquake clusters (and probably other patterns of seismicity) is closely related to physical characteristics of a region and varies systematically on a scale of tens of kilometers within southern California. This dependence can guide using complementary data sources (heat flow, depth of seismicity, etc.) for improving seismic hazards assessment and forecasting strategies.

[42] Our findings can be explained by the existence of two dominant underlying failure mechanisms, *brittle* and *ductile*, whose interplay creates the observed variety of earthquake patterns [e.g., *Ben-Zion and Lyakhovsky*, 2006]. Regions with relatively cold crystalline rocks (and nonelevated fluid content) have highly brittle rapid failure process (small number of foreshocks) leading to topologically shallow clusters. The stress transfer from mainshocks in such regions triggers aftershocks with high spatiotemporal localization and broad angular distribution owing to the existence of brittle hypocenter sites in all directions. In contrast, regions with elevated temperature and/or fluid content (or relatively soft sediments) have some relaxation via ductile deformation. This leads to a more gradual failure process (larger number of foreshocks) and topologically deep clusters. Such mixed brittle-ductile failures are expected to be dominant in partially weakened spatial channels, so they have narrow angular distribution and may trigger additional failure events to large distances.

[43] It is natural to assume that the two dominant cluster types should have preferred spatial locations related to physical properties of the crust (e.g., heat flow, fluids, soft sediments). This is confirmed by the detailed spatial analysis of section 3 (Figures 9 and 10; Tables 1 and 2), which illustrates the spatial variability of family statistics (Figure 9) and demonstrates that the differences in the values of essential family statistics, traditional and new, within the hot and cold regions (Figure 10; Tables 1 and 2) cannot be explained by random fluctuations and hence should be due to different physical mechanisms.

[44] The results of this study are consistent with previous findings by *Yang and Ben-Zion* [2009] and *Enescu et al*. [2009] that the aftershock productivity in a region, expressed via parameters of the Omori-Utsu law and ETAS model, is positively correlated with the heat flow. The original contributions of this work include (1) systematic and stable cluster detection approach, (2) absence of relying on particular parameterized forms of clustering (e.g., Omori-Utsu law, ETAS model), (3) uniform analysis of small-to-intermediate magnitude clusters, (4) uniform spatial analysis of the entire southern California (not only special study regions), (5) large number of examined complementary cluster statistics, and (6) analysis of the newly available high-quality relocated catalog of *Hauksson et al*. [2012]. Notably, the self-adapted methodology developed in this study can be applied to other regions without re-fitting numerical parameters. The agreement among the main conclusions based on the complementary approaches of *Yang and Ben-Zion* [2009], *Enescu et al*. [2009] and this study suggests that the detected patterns represent physical effects rather than statistical artifacts.

[45] Notably, the three largest aftershock sequences—M7.3 Landers, M7.2 El Mayor–Cucapah, and M7.1 Hector Mine—behave statistically different from the remaining seismicity (Figure 4a). All other seismic clusters, associated with small-to-medium magnitude mainshocks, are shown to exhibit a bimodal $\langle d \rangle$ pattern that facilitates their classification into two basic types: *burst-like clusters* reflecting highly brittle failures and *swarm-like clusters* associated with mixed brittle-ductile failures [e.g., *Ben-Zion and Lyakhovsky*, 2006]. The average leaf depth $\langle d \rangle$ of section 2 provides an effective way of quantifying the cluster type. The largest aftershock sequences with mainshock magnitudes $m > 7$ in the examined catalog appear to have a mixture of the two basic cluster types.

[46] Although there is no sharp boundary between the cluster types, the distribution of $\langle d \rangle$ is a mixture (Figure 6) of two exponential distributions with distinct slopes and change point at $\langle d \rangle \approx 5$, which is consistent with our rough division of the clusters into burst-like ($\langle d \rangle < 5$) and swarm-like ($\langle d \rangle > 5$). A clear separation into two basic cluster types, swarm-like sequences with linear topology and burst-like sequences with spray-like topology, can only be done for families of sufficient large size, say $N \geq 100$ (and mainshock magnitude $m < 7$). As illustrated in Figure 4, families of small size may exhibit various cluster forms, from purely linear ($\langle d \rangle = N-1$) to purely spray-shaped ($\langle d \rangle = 1$) with no clear boundary. This dependence on family size is natural and would likely affect any measure related to the cluster type, as it is objectively hard if not impossible to decide what type a small cluster (with say three events) belongs to.

[47] The cluster type measure suggested here is purely topological—it describes how events are connected to each other within a cluster and ignores all metric properties of earthquakes (magnitude, depth, etc.). However, the topology of the detected clusters is shown to be closely connected to their metric properties. We establish strong coupling between the (topological) cluster type measured by the depth $\langle d \rangle$ and each of 12 examined cluster statistics (supporting information C, Table C1), most of which are traditionally considered in earthquake aftershock studies.

[48] The majority (up to 90%) of detected clusters in southern California belong to *burst-like sequences* characterized by small average leaf depth $\langle d \rangle$ (Figures 6, 9a, and 10a). These sequences develop in a wide range of directions (Figs. C1 and C10) within relatively small space-time regions (Figs. C4, C5, C7, and C8), they are mostly composed

of the first-generation aftershocks (Fig. C9) of a prominently larger mainshock (Fig. C6), have relatively small total number of foreshocks and aftershocks (Figs. C2 and C3a), small proportion of foreshocks (Fig. C3b), fast decay rate with time from the mainshock (Fig. C5), and *b*-value of mainshocks close to unity (Fig. C1). In contrast, *swarm-like clusters* are characterized by large average leaf depth $\langle d \rangle$ (Figures 6, 9a, and 10a). These sequences develop in strongly anisotropic fashion propagating along effectively one-dimensional channels (Figs. C1 and C10), their spatiotemporal extents are relatively large (Figs. C4, C5, C7, and C8), they form sequences of multiple-generation events with each parent having a small number of offspring of comparable magnitude (Figs. C9 and C6), they have relatively large number of foreshocks and aftershocks (Figs. C2 and C3a), large proportion of foreshocks (Fig. C3b), relatively slow decay rate with time from the mainshock (Fig. C5), and *b*-value for mainshocks significantly less that unity (Fig. C1).

[49] While most of the detected clusters are burst-like, the size of *burst-like* clusters is generally smaller than that of *swarm-like* clusters. In the southern California earthquake catalog of *Hauksson et al*. [2012], the events in the *burst-like* (defined here as $\langle d \rangle < 5$) and *swarm-like* (defined here as $\langle d \rangle \geq 5$) clusters account, respectively, for 47% and 53% of the data; these computations do not include singles, which account for 31% of the catalog.

[50] The most conspicuous sequences of mainshocks with $m > 7$ have intermediate properties likely produced by a mixture of both cluster types. These events are somewhat larger than the size associated with rupturing the entire crust ($m \approx 6.2$ in southern California) and are likely to have increased radiation efficiency than the smaller events contained within the crust. We note that estimating the radiation efficiency from seismograms involves various assumptions and there is a debate on whether large events have indeed higher radiation efficiency than small ones (see, e.g., section 2 of *Ben-Zion* [2008] and references therein).

[51] Recall that our results suggest that the burst-like topologically shallow sequences are characterized by *low* intensity (and, hence, low total number) of aftershocks, *short* duration, and *small* area, in comparison to the swarm-like topologically deep sequences with similar mainshock magnitude. This may seem counterintuitive, as it challenges the traditional perception that aftershock sequences are the most prominent form of earthquake clustering that is responsible for most inhomogeneities in observed catalogs. We note that most traditional aftershock studies were made with the most prominent sequences that follow large mainshocks. In the examined catalog, this would include the M7.3 Landers, M7.2 El Mayor-Cucapah, M7.1 Hector Mine, etc. Such sequences indeed contain the majority of the observed earthquakes and present the most visible deviation from stationarity/homogeneity in visual inspection of the catalog. For example, in our analysis, the M7.3 Landers cluster has size $N = 14{,}622$, which is 13% of the entire catalog, and the three largest events of magnitude $m \geq 7$ have 27,806 earthquakes in their respective families, which is 25% of the entire catalog. Our results suggest that such large events belong to a special class with intermediate clustering properties, likely resulting from triggering a mixture of brittle and ductile-brittle cluster types. The comparative properties of the burst-like vs. swarm-like sequences reported in this study are based exclusively on the small-to-medium magnitude clusters. While there is no sharp boundary between the "largest" and "medium"-magnitude events, the $m \geq 7$ events are clear outliers in the analysis of Figure 4. The six mainshocks with $m \geq 6.2$ in the examined catalog behave in a similar fashion, exhibiting an intermediate-type clustering. The analysis of supporting information B on scaling of the normalized topological depth δ with cluster size N shows (Fig. B1) that all $m \geq 6.2$ events belong to the same (right) mode of the bimodal distribution.

[52] It is worth noting that the largest earthquakes are *not* outliers in aftershock production when considered within the Δ-analysis. For example, the M7.3 Landers mainshock has only 11 aftershocks with magnitude above 5.3 (= 7.3 − 2), which is merely a 79% percentile of the distribution of the sizes of Δ-clusters; the M7.2 El Mayor mainshock has six Δ-aftershocks, and the M7.1 Hector Mine has only two Δ-aftershocks. In contrast, the largest number of Δ-aftershocks ($N = 200$) is observed in a swarm-like cluster with mainshock magnitude $m = 4.11$. The exact values of various estimated parameters (e.g., of the Omori-Utsu, Båth, Gutenberg-Richter laws, etc.) may be revised in future studies with more accurate catalogs and additional analysis methods. Nevertheless, we conjecture that our main results on the existence of several dominant cluster types with different topological structure correlated with numerous metric characteristics, which have preferred spatial locations associated with different physical properties and explain the essential properties of foreshock-mainshock-aftershock sequences in different regions, will remain valid.

[53] **Acknowledgments.** We are grateful to Rick Schoenberg for sharing with us the R-code for simulating the ETAS model and to Yevgeniy Kovchegov for helpful discussions. The manuscript benefited from constructive comments of three anonymous referees. The study was supported by the Southern California Earthquake Center (based on NSF Cooperative Agreement EAR-0529922 and USGS Cooperative Agreement 07HQAC0008) and the National Science Foundation (grants EAR-0908903 and DMS-0934871). The SCEC contribution number for this paper is 1719.

References

Aki, K. (1981), A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction: An International Review, Maurice Ewing Set*., vol. 4, edited by D. W. Simpson and P. G. Richards, pp. 566–574, AGU, Washington, D.C.

Bak, P. (1996), *How Nature Works: The Science of Self-Organized Criticality*, Copernicus, New York, pp. 212 .

Ben-Zion, Y. (2008), Collective Behavior of Earthquakes and Faults: Continuum-Discrete Transitions, Evolutionary Changes and Corresponding Dynamic Regimes, *Rev. Geophys.*, *46*, RG4006, doi:10.1029/2008RG000260.

Ben-Zion, Y., and V. Lyakhovsky (2006), Analysis of Aftershocks in a Lithospheric Model with Seismogenic Zone Governed by Damage Rheology, *Geophys. J. Int.*, *165*, 197–210, doi:10.1111/j.1365-246X.2006.02878.X.

Blackwell, D. D., and M. Richards (2004), *Geothermal Map of North America*, American Assoc. Petroleum Geologist (AAPG), 1 sheet, scale 1:6,500,000.

Enescu, B., S. Hainzl, and Y. Ben-Zion (2009), Correlations of Seismicity Patterns in Southern California with Surface Heat Flow Data, *Bull. Seismol. Soc. Am.*, *99*, 3114–3123, doi:10.1785/0120080038.

Freedman, D. A. (2005), *Statistical Models: Theory and Practice*, pp. 456, Cambridge University Press, Cambridge, U.K.

Gardner, J. K., and L. Knopoff (1974), Is the sequence of earthquakes in Southern California, with aftershocks removed, Possionian?, *Bull. Seismol. Soc. Am.*, *64*(5), 1363–1367.

Hauksson, E., W. Yang, and P. M. Shearer (2012), Waveform Relocated Earthquake Catalog for Southern California (1981 to June 2011), *Bull. Seismol. Soc. Am.*, *102*(5), 2239–2244, doi:10.1785/0120120010.

Kagan, Y. Y. (1994), Observational evidence for earthquakes as a nonlinear dynamic process, *Physica D*, *77*, 160–192.

Keilis-Borok, V. I., and A. A. Soloviev (Eds) (2003), *Nonlinear Dynamics of the Lithosphere and Earthquake Prediction*, Springer-Verlag, Berlin-Heidelberg, pp. 337.

Mogi, K. (1963), Some discussions on aftershocks, foreshocks, and earthquake swarms: The fracture of a semi-infinite body caused by an inner stress origin and its relation to the earthquake phenomenon, *Bull. Earthquake Res. Inst. Univ. Tokyo*, *41*, 615–658.

Pitman, J. (2006), Combinatorial stochastic processes, in *Lecture Notes in Mathematics*, vol. 1875, p. 270, Springer-Verlag, Berlin, Heidelberg.

Rundle, J. B., D. L. Turcotte, R. Shcherbakov, W. Klein, and C. Sammis (2003), Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems, *Rev. Geophys.*, *41*(4), 1019, doi:10.1029/2003RG000135.

Shearer, P. M. (2012) Self-similar earthquake triggering, Bath's law, and foreshock/aftershock magnitudes: Simulations, theory, and results for southern California. *J. Geophys. Res.*, *117*, B06310.

Sornette, D. (2004), Critical phenomena in natural sciences, in *Chaos Fractals, Self-Organization and Disorder: Concepts and Tools*, p. 528, Springer, Heidelberg.

Utsu, T. (2002), Statistical features of seismology, in *International Handbook of Earthquake and Engineering Seismology, Part A*, edited by W. HK Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger, 719–732, Academic Press, London, U.K.

Vidale, J. E., and P. M. Shearer (2006), A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations and aseismic slip as drivers, *J. Geophys. Res.*, *111*, B05312, doi:10.1029/2005JB004034.

Watson, H. W., and F. Galton (1875), On the Probability of the Extinction of Families, *J. Anthrop. Inst. Great Britain*, *4*, 138–144.

Yang, W., and Y. Ben-Zion (2009), Observational analysis of correlations between aftershock productivities and regional conditions in the context of a damage rheology model, *Geophys. J. Int.*, *177*, 481–490, doi:10.1111/j.1365-246X.2009.04145.X.

Zaliapin, I., and Y. Ben-Zion (2013), Earthquake clusters in southern California I: Identification and stability, *J. Geophys. Res. Solid Earth*, *118*, doi:10.1002/jgrb.50179.

Zaliapin, I., A. Gabrielov, H. Wong, and V. Keilis-Borok (2008), Clustering analysis of seismicity and aftershock identification, *Phys. Rev. Lett.*, *101*, 018501, doi:10.1103/PhysRevLett.101.018501.

Auxiliary material for

Earthquake clusters in southern California II: Classification and relation to physical properties of the crust

Ilya Zaliapin[1] and Yehuda Ben-Zion[2]

[1]Department of Mathematics and Statistics, University of Nevada, Reno, 89557 (zal@unr.edu)
[2] Department of Earth Sciences, University of Southern California, Los Angeles, 90089-0740 (benzion@usc.edu)

J. Geophys. Res.

Section A. *Stability of family structure with respect to the magnitude cutoff*

The goal of this section is to (i) further illustrate the dominant family types and (ii) explore stability of the family type with respect to the magnitude threshold of the analysis. We consider here two families and study how they transform depending on the minimal magnitude of the analysis. The first family is located in the Salton trough area (Fig. A1). The largest event in this family has mainshock magnitude 5.75 and coordinates (33.0875N, 115.6195W). The second family spans the San Gabriel valley and mountains (Fig. A2). The mainshock of this family has magnitude 5.51 and coordinates (34.1380N, 117.7082W); it occurred right off the San Gabriel mountains near Claremont, CA.

When the nearest-neighbor analysis is done for earthquakes with $m \geq 4.0$, the Salton trough sequence consists of a single event (Fig. A1a,b,c), while the San Gabriel family (Fig. A2a,b,c) combines 6 events in a spray-like configuration with a single foreshock.

When the magnitude cutoff in the nearest-neighbor analysis is lowered to 3.0, the number of events in both the families increases. The Salton trough family (Fig. A1d,e,f) now has 31 events, including 12 foreshocks and 18 aftershocks. The topology of the family (Fig. A1f) combines a chain of 10 events and a burst of 15 events. The San Gabriel family (Fig. A2d,e,f) now contains 34 events; and still has a single foreshock. It must be noted that the foreshock that was present in the $m \geq 4.0$ analysis no longer belongs to the cluster; such reshuffling of the nearest-neighbor cluster structure may occur even in conventional Euclidean spaces. The number of events that change their clusters under changing the magnitude cutoff of the analysis is however very small. The events in the Salton trough family are organized in a prominently spray-like shape (Fig. A2e,f), with 26 out of 34 earthquakes being direct aftershocks of the largest event. The spatial extent of the San Gabriel family (Fig. A2e) is smaller than that of the Salton trough family (Fig. A1e).

Finally, we decrease the magnitude cutoff to 2.0. The Salton trough family (Fig. A1g,h,i) now has 315 events, with 81 foreshocks and 233 aftershocks. Topologically (Fig. A1i), the family consists of multiple chains and a dominant burst that includes 136 events (43%). The San Gabriel family (Fig. A2g,h,i) has 400 events, with 3 foreshocks and 396 aftershocks. Topologically (Fig. A2i), the family is mainly comprised of a burst that includes 261 events (65%). It is now clearly seen that the spatial extend of the San Gabriel family (Fig. A2h) is much smaller than that of the Salton trough family (Fig. A1h). We also note that the San Gabriel family has roughly isotropic shape (reminiscent of explosion) whereas the Salton trough family is concentrated in a small number of directions (suggesting flow in specific channels). To conclude, our results suggest that the cluster structure is stable with respect to the magnitude threshold of the nearest-neighbor analysis.

Section B. *Normalized tree depth*

The main text of the paper analyzes the average leaf depth $\langle d \rangle$. An alternative approach to treat the bimodal distribution of the average leaf depth is related to the depth scaling with family size N. Note that a linear chain of size N has depth $\langle d \rangle = N-1 \sim N$; a perfect spray-shaped tree of size N (a tree with $N-1$ leaves directly attached to the root) has depth $1 \sim N^0$. Here the sign "~" stands for "scales as when N increases". It is hence natural to expect for the observed trees to behave like $\langle d \rangle \sim N^{\kappa}$, with $0 < \kappa < 1$. Figure B1a shows the average leaf depth $\langle d \rangle$ as a function of the family size N for the regular families obtained in the nearest-neighbor analysis for $m \geq 2$ earthquakes. The figure only shows 452 families with the mainshock magnitude 4.0 or above. A notable observation is the existence of two principal modes of the expected increase of the depth $\langle d \rangle$ with family size N; they are depicted by two lines $\langle d \rangle \propto N^{0.5}$. One of the modes (located to the left) corresponds to the much higher tree depths for the same family size. To quantify the mode separation, we introduce the normalized tree depth $\delta = \langle d \rangle \times N^{-0.5}$, which balances the effect of depth increase with the family size. Figure B1b further illustrates the modes of the depth-size dependence, using different levels of the normalized depth δ. Figure B2 shows three examples of trees with different values of the normalized depth δ.

We note that the tree structure is affected by the magnitude of the events in the family. In particular, large-magnitude events tend to attract more offspring, according to the nearest-neighbor distance of Eq. (1) that exponentially decreases with the magnitude of the parent. Figure B3 shows the normalized depth δ as a function of the family mainshock magnitude m for the 51 regular families of size $N \geq 100$. The values of the normalized depth span the range $0.05 < \delta < 2$. Notably, there exists a *transition in the family formation* process between medium-magnitude and large-magnitude events with the transition range $4.7 < m < 6.2$. Namely, *all* $m > 6.2$ mainshocks form prominently spray-shaped clusters with very small tree depth, $\delta < 0.11$; such clusters would be commonly referred to as *aftershock sequences*. We note that large-magnitude events typically break the entire seismogenic zone and reach the free surface; this may be related to the topologic structure of the respective families. *All* medium-magnitude mainshocks, $m < 4.7$, form families with high normalized depth, $\delta > 0.5$; such clusters would be commonly referred to as *swarms*. Finally, the mainshocks in the transition range $4.7 < m < 6.2$ may form clusters with a wide variety of normalized depths, $0.05 < \delta < 2.0$, which includes linear (for $\delta > 0.5$) and spray-shaped (for $\delta < 0.2$) families as well as all intermediate types ($0.2 < \delta < 0.5$).

Section C. *Cluster statistics vs. average leaf depth*

We have noticed already in Sect. 3 that some family statistics considered in our analysis are related to each other. Say, it is natural to expect the topological depth $\langle d \rangle$ to be negatively associated with the family branching index B. We demonstrate in this section that numerous statistical properties of the nearest-neighbor families are indeed strongly coupled with the average leaf depth $\langle d \rangle$. Such coupling in many cases is a natural consequence of conditional family construction, and hence presents purely statistical rather than physical effect. Nevertheless, systematic exploration of these dependencies seems necessary for better understanding of the earthquake family structure. We choose the topological depth $\langle d \rangle$ as the governing parameter since it exhibits the strongest association with the regional properties among the examined family statistics not exclusively related to foreshocks (see Fig. 7, Table 1).

The frequency-magnitude distribution for cluster mainshocks in two groups with different ranges of family depth ‹*d*› is shown in Fig. C1. The mainshock distribution within the topologically shallow families (solid line) is reminiscent of that for the entire mainshock population [cf. ZBZ13, Fig. 10] and can be closely approximated by the exponential Gutenberg-Richter law with *b*-value (slope of the line) $b = 1$. The mainshocks of the topologically deep families (dashed line) also follow an exponential distribution, although with significantly lower *b*-value $b \approx 0.6$. Accordingly, the proportion of large-magnitude mainshocks is higher within topologically deep, swarm-like families. Recall that the *b*-value can be interpreted as $b = 0.5d_f$ with d_f being the fractal dimension of epicenters [*Aki*, 1981]. This implies that the epicenters of the burst-like families with small topological depth occupy statistically the entire surface ($d_f \approx 2$), while those of the swarm-like families occur within essentially one-dimensional channels ($d_f \approx 1.2$). This property is explicitly confirmed below in Fig. C10. Another interesting observation is that while the number of swarm-like low-magnitude clusters is much smaller than the number of burst-like low-magnitude clusters; the number of large-magnitude clusters is comparable for both cluster types.

Figure C2 shows the average number of aftershocks and foreshocks per family, in regular analysis for events with small-to-intermediate mainshock magnitudes $2 \le m \le 6$, as a function of family mainshock magnitude *m*. The analysis is done separately for deep families (‹*d*› > 5, diamonds, dashed line) and shallow families (‹*d*› ≤ 5, circles, solid line). The average number of foreshocks and aftershocks is larger in deep families. This is true for aftershocks in families with mainshock magnitude $m \le 5$, and for foreshocks in families with mainshocks magnitude $m \le 6$. At the same time, the large magnitude families seem to have fore/aftershock productivity that is independent of the tree depth. The figure does not show mainshock magnitudes above 6, which add to the variability of the plot without changing the above conclusions. The average fore/aftershock number for intermediate mainshock magnitudes can be approximated by an exponential law

$$N = K_N \times 10^{\beta m}. \qquad \text{(C1)}$$

The value of the productivity index for aftershocks in deep (‹*d*› > 5) and shallow (‹*d*› ≤ 5) families is $\beta \approx 0.7$, $\beta \approx 0.9$ respectively; these estimates are done within the magnitude range [2.5-5] and may be different if larger mainshocks are considered. While it is harder to estimate the productivity index for foreshocks due to large fluctuations of the foreshock number, it is safe to say that the index value is close to $\beta = 0.5$ for the mainshock magnitude range [2.5-5].

In part I of this study it was shown [ZBZ13, Fig. 14] that the average number of aftershocks per family N_A, ignoring the family depth, scales with the mainshock magnitude *m* as $N_a \propto 10^{\beta m}$, $\beta \approx 1$; this result is consistent with the other studies that report the productivity index α of about unity [e.g., *Helmstetter et al*., 2005]. The depth-independent index $\beta \approx 1$ may seem inconsistent with the depth-dependent indices $\beta \approx 0.7$ and $\beta \approx 0.9$, which are both significantly less that unity. This effect is explained by the depth-dependent mainshock distribution illustrated in Fig. C1. Namely, the proportion of topologically deep low-magnitude clusters is small; hence, the total number of aftershocks for low-magnitude clusters is about the same as the number of aftershocks for shallow low-magnitude clusters. At the same time, the proportion of topologically deep large-magnitude clusters is much larger; hence, the total number of aftershocks for large-magnitude clusters is the sum of that number in both deep and shallow

clusters. This effect leads to increase of the scaling exponent in the entire population compared to the subpopulations of deep and shallow clusters.

In part I it was shown [ZBZ13, Fig. 15] that the aftershock and foreshock productivity in Δ-analysis is independent of the family mainshock magnitude. This motivates examination of the average number of Δ-foreshocks and Δ-aftershocks per family grouped by the family depth. The results are shown in Fig. C3a, where all families are divided into 5 equal percentile groups according to the increasing value of the average leaf depth ‹*d*›. The number of foreshocks and aftershocks clearly increases with the topological depth. At the same time, the number of foreshocks is always less than the number of aftershocks. Figure C3b shows the proportion of Δ-foreshocks in the families with size $N \geq 10$, according to the average leaf depth ‹*d*›. The proportion of foreshocks increases with the depth from almost 0 for shallow families to above 0.25 for the deepest ones.

Next, we focus on the temporal intensity of events within a family around the mainshock. Figure C4 shows the estimated earthquake intensity, in events per day per cluster, in regular clusters with mainshock magnitude $m \geq 4$ within 30 days from a mainshock. The analysis is done separately for shallow clusters (‹*d*› ≤ 5, solid line, circles) and deep clusters (‹*d*› > 5, dashed line, diamonds). This analysis includes clusters with no foreshocks and/or aftershocks. The intensity of events decreases away from the mainshock, in agreement with the depth-independent results [ZBZ13, Fig. 16]. The intensity of topologically deep clusters is order of magnitude higher than that of shallow ones (consistent with the results of Fig. C3), independently of the time away from mainshock. Moreover, the intensity of events decays faster within 10 days from the mainshock (for both foreshocks and aftershocks) in shallow clusters. This visual impression is confirmed by the analysis of Fig. C5 below. We note also that the foreshock intensity for shallow clusters is always below 0.1 event/day/cluster, and it decreases to 0.01 events/day/cluster 10 days away from a mainshock. This explains the observation that only 27% of the shallow clusters (‹*d*› ≤ 5, $m \geq 4$) have foreshocks; while among the deep clusters (‹*d*› > 5, $m \geq 4$) 95% have foreshocks.

Figure C5 presents more focused results on event intensity within 10 days from mainshocks for families with at least one Δ-aftershock or Δ-foreshock; the earthquake intensity is measured in events per day per family. The aftershock decay in the examined cases closely follows the Omori-Utsu law [*Omori*, 1894; *Utsu et al.*, 1995]:

$$\Lambda = K_\Lambda \times (t + c)^{-p}. \qquad \text{(C2)}$$

The decay rate is higher for topologically shallow families ($p \approx 0.85$) than for deep ones ($p \approx 0.65$). The foreshock decay shows much more scattered results due to smaller number of events, but it can also be coarsely approximated by a power law. Due to sampling problems, we are not trying to estimate the exact foreshock decay rates, although it is clear that the overall decay in shallow families is faster than in deep ones. For visual convenience we show in Fig. C5b two lines that correspond to power law decay with rates of 0.65 and 1.1. The results are consistent with the depth-independent intensity decay illustrated in Fig. 17 of ZBZ13. The results confirm the suggestion in Fig. C4 that the intensities of events in topologically shallow sequences tend to decay faster as time from mainshock (in both directions) increases.

Figure C6 illustrates results of regular analysis of the magnitude difference Δ_m between the mainshock and the largest foreshock (diamonds, dashed line) and aftershock (circles, solid

line). The magnitude difference for both event types tends to be smaller for deeper families; the effect although is much stronger for aftershocks than for foreshocks. Notably, the depth-dependent magnitude difference for the foreshocks is always statistically indistinguishable from the depth-independent average of Δ_m = 1.2 [see ZBZ13, Fig. 18]. In contrast, the depth-dependent magnitude differences for aftershocks do deviate significantly from the depth-independent average Δ_m = 1.1 for very shallow and very deep families.

The duration of foreshocks and aftershocks in Δ-analysis is illustrated in Fig. C7; both foreshock and aftershocks sequences are longer for deep families. The duration of Δ-foreshocks is order of magnitude smaller than that of Δ-aftershocks, independently of the family depth. The distribution of the area for aftershock sequences according to the family depth is shown in Fig. C8. The area tends to increase with increasing topological depth, in agreement with the example results shown in Figs. 1 and 2 (see also Figs. A1, A2). The dependence of area on the family depth is more scattered than the other characteristics examined in this study; this prevents a robust analysis of the foreshock area.

Next, we examine the immediate child productivity by analyzing the average number B of children per parent (Fig. C9). In graph-theoretical terminology, this is known as the *branching number*; in seismological context it is usually called *the number of first-generation offspring*. Specifically, we (i) consider *every* parent event within each family, (ii) focus on the first-generation offspring only, and (iii) do not consider events with no children, so the minimal number is B = 1. The branching numbers are averaged within each family. Fig. C9a shows that the shallow families have a prominently higher average B. This observation is further illustrated in Fig. C9b that displays the distribution of B for shallow ($\langle d \rangle \le 3$) and deep ($\langle d \rangle > 3$) families. As shown, the distribution of B for deep families has exponential tail $1-F(B) = C_B \times 10^{-\alpha B}$ with $\alpha \approx 0.3$.

Finally, we analyze the directional dependency of events in families of different types. Specifically, consider the empirical distribution $F(\theta)$ of the surface angle θ between the epicenters of family mainshock and the other events. The angle, in degrees, is counted counterclockwise assuming that East corresponds to $\theta = 0$. For each family we perform a one sample Kolmogorov-Smirnov test [*Conover*, 1971] that compares $F(\theta)$ to the uniform distribution on the interval [0, 360]. The proportion U of families with at least 5 events and mainshock $m \ge 4$ that pass this test at level 0.01 (we call such families *isotropic*) for different average leaf depths is shown in Fig. C10. The proportion of isotropic families decreases as the tree depth increases. In other words, burst-like sequences develop in spatially isotropic fashion reflected in uniform circular event distribution, while deeper swarm-like sequences propagate along preferred channels in particular directions. The existence of preferred propagation channels may also explain the observation of Fig. C8 that the area of aftershock sequences increases with the family depth (and related results in Figs. 1 and 2); a failure cascade along specific (presumably weaker) directions can extend larger distance from the mainshock compared to the isotropic failures characterizing the burst-like shallow sequences.

Section D. *ETAS model: specification and parameters*

The ETAS model is specified in terms of the conditional intensity $m(t,\mathbf{f},m|H_t)$ of a process $Z_t = \{t_i,\mathbf{f}_i,m_i\}$ given its history $H_t = (\{t_i,\mathbf{f}_i,m_i\} : t_i < t)$ up to time t. Here t_i represents earthquake occurrence times, $\mathbf{f}_i$ their coordinates (e.g., epicenter, hypocenter, or centroid) and m_i the magnitudes [*Daley and Vere-Jones*, 2002]. The statistical analysis and inference for Z_t are done using the conditional likelihood

$$\log L_t = \sum_{t_i<t} \log \mu(t_i, \mathbf{f}_i, m_i \mid H_t) - \int_0^t \int_M \int_F \mu(t, \mathbf{f}, m \mid H_t)\, dt\, dm\, d\mathbf{f}\,, \tag{D1}$$

where M and F denote the magnitude range and spatial domain of events, respectively. We assume furthermore that the magnitudes of events are independent and drawn from the Gutenberg-Richter (exponential) distribution with a constant b-value. This reduces conditional intensity to the following special form, which allows various particular parameterizations [*Ogata*, 1998, 1999]:

$$\mu(t, \mathbf{f} | \mathrm{H}_t) = \mu_0(t, \mathbf{f}) + \sum_{i:t_i<t} g(t - t_i, \mathbf{f} - \mathbf{f}_i, m_i)\,.$$

We use in this study a homogeneous background intensity $\mu_0 = \mu$ and the following parameterization for the response function g suggested by *Ogata* [1998, Eq. (2.3)]:

$$g(t, x, y, m) = \frac{K}{(t+c)^p} \frac{\exp\left(\alpha(m - m_0)\right)}{\left(x^2 + y^2 + d\right)^q}\,. \tag{D2}$$

Here m_0 is the lowest considered magnitude, and (x,y) are Cartesian coordinates of the epicenters. The model is specified by 8 scalar parameters $\theta = \{\mu, b, K, c, p, \alpha, d, q\}$. In this study, we generate synthetic ETAS catalogs using parameters consistent with those reported in the literature [e.g., *Wang et al*., 2010; *Chu et al*., 2011; *Marzocchi and Zhuang*, 2011]: $\mu = 0.003$ (km^2 year)$^{-1}$, $b = \alpha = 1$, $K = 0.007$ (km^2 year)$^{-1}$, $c = 0.00001$ year, $p = 1.17$, $q = 1.7$, $d = 30$ km^2; the simulations are done within a region of 500×500 km during 15 years. The catalog consists of 146,432 earthquakes.

Section E. *Analysis of Variance: Review*

The *one-way* ANOVA test (*Freedman*, 2005) compares the means of several groups of observations by examining the variance within the groups relative to the variance between the groups. Formally, consider samples X_{ij}, where index $i = 1,\ldots,G$ counts different groups and index $j = 1,\ldots,N_i$ counts observations within group i; and let $N = N_1 + \ldots + N_G$. Let $\bar{X}_i$ denote the sample average for the group i and μ_i denote the population mean for the same group. The ANOVA tests the null hypothesis H_0: $\mu_1 = \ldots = \mu_G$ vs. the alternate hypothesis that at least two groups have different means. The test statistic is computed as

$$F = \frac{SSG}{SSE} \frac{N-1}{G-1},$$

where *SSG* is the *group sum of squares* and *SSE* is the *error sum of squares*:

$$SSG = \sum_{i=1}^{G} N_i \left(\bar{X}_i - \bar{X}\right)^2;\; SSE = \sum_{i=1}^{G} \sum_{j=1}^{N_i} \left(X_{ij} - \bar{X}_i\right)^2.$$

The intuition behind the test is that if all groups have the same mean, then SSG/(G-1) ≈ SSE/(N-1) and the test statistic F should be close to unity; while if the groups have different means, then SSG/(G-1) > SSE/(N-1) and the values of F will increase. Namely, if (i) the observations are normally distributed and (ii) the variances of all the groups are the same, then the test statistic F has F-distribution with (G-1) and (N-1) degrees of freedom (*Freedman*, 2005). The ANOVA test is reasonably robust with respect to the violation of both the above assumptions and it is known to have large power with respect to numerous alternative hypotheses. When applying the ANOVA test, we always transform the variables to make the samples approximately Normally distributed.

Table C1: Earthquake cluster statistics related to the earthquake family type

Statistic	Burst-like family	Swarm-like family	Figure #	
			Δ-analysis*	Regular analysis
Average leaf depth, ‹*d*›	Low	High	1-4	
b-value for mainshocks	High	Low	C1	
Ave. no. of aftershocks per family, N_a	Low	High	C3	C2a
Ave. no. of foreshocks per family, N_f	Low	High	C3	C2b
Intensity of aftershocks, Λ_a	Low	High	C4, C5a	
Intensity of foreshocks, Λ_f	Low	High	C4, C5b	
Magnitude difference between mainshock and largest aftershock, Δ_m	High	Low		C6
Magnitude difference between mainshock and largest foreshock, Δ_m	High	Low		C6
Duration of aftershocks, D_a	Low	High	C7a	
Duration of foreshocks, D_f	Low	High	C7b	
Area of aftershocks, A_a	Low	High	C8	
Branching index, *B*	High	Low		C9
Angular surface isotropy, *U*	High	Low		C10

* Defined in Sect. 2 of ZBZ13

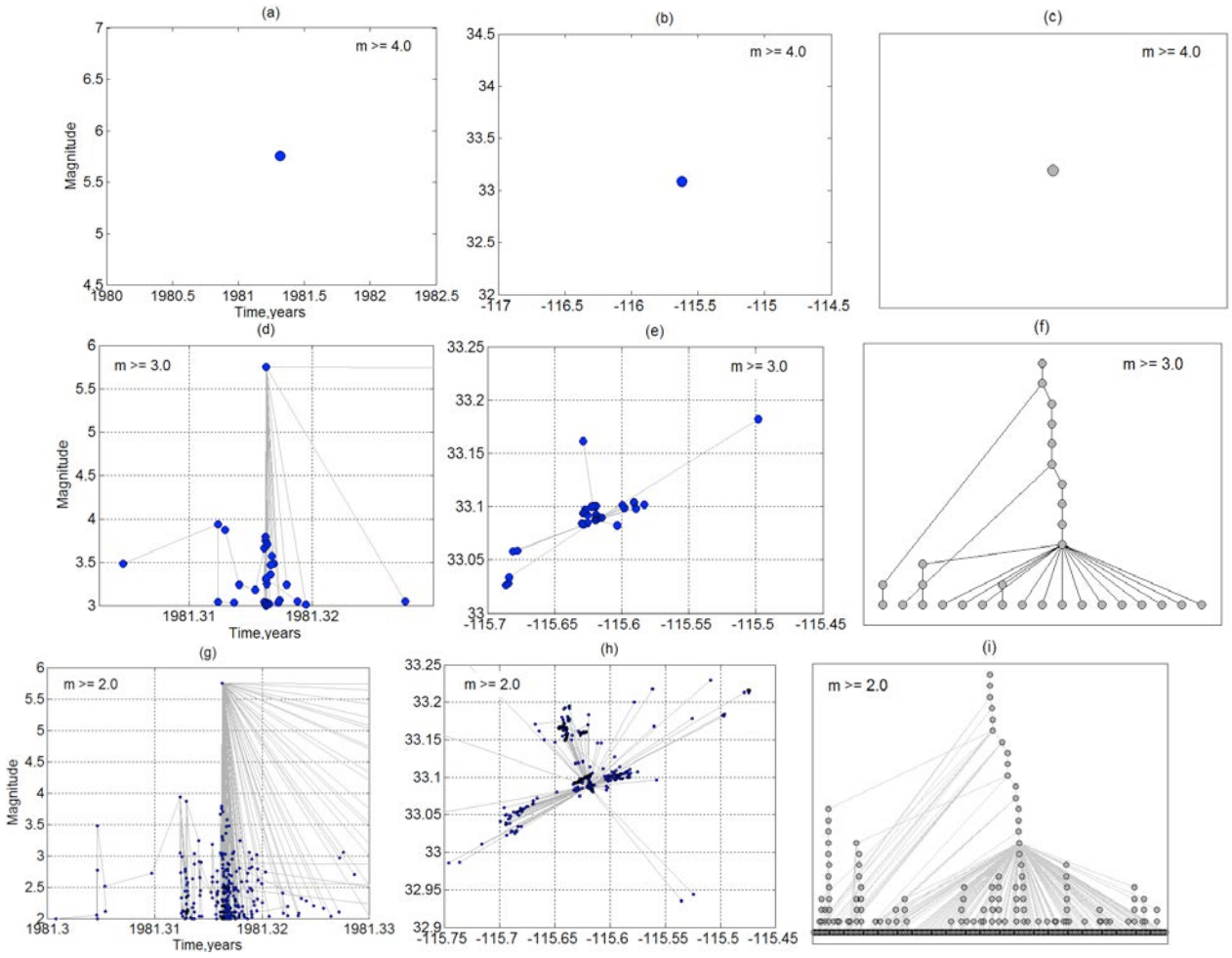

Figure A1: Cluster in Salton trough area. Circles correspond to earthquakes, lines to parent links. Figure shows results for different magnitude thresholds of the nearest-neighbor analysis: (a,b,c) $m \geq 4.0$, (d,e,f) $m \geq 3.0$, (g,h,i) $m \geq 2.0$. (a,d,g) Magnitude as a function of time. (b,e,h) Space map. (c,f,i) Topologic tree.

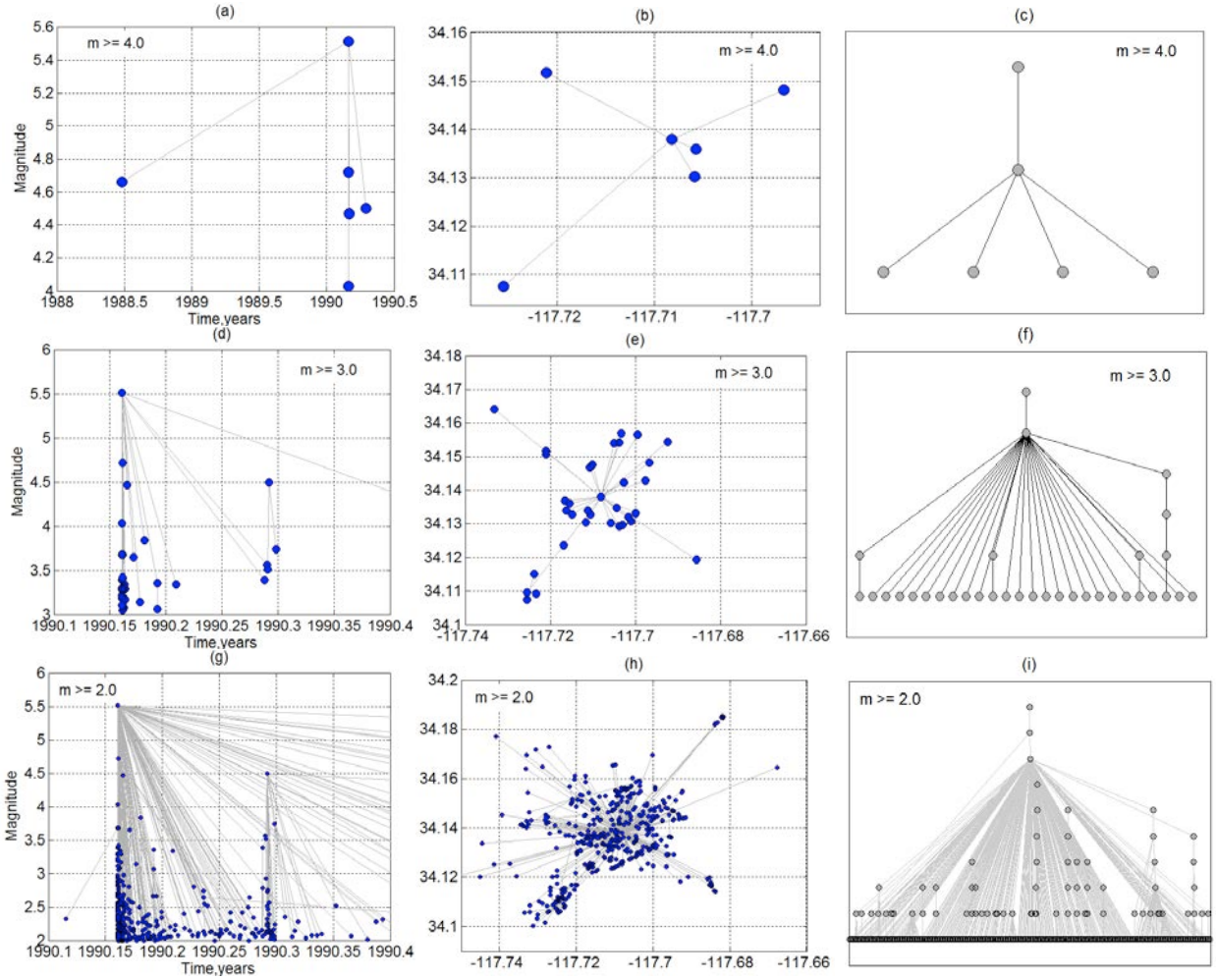

Figure A2: Family in San Gabriel area. The other notations are the same as in Fig. A1.

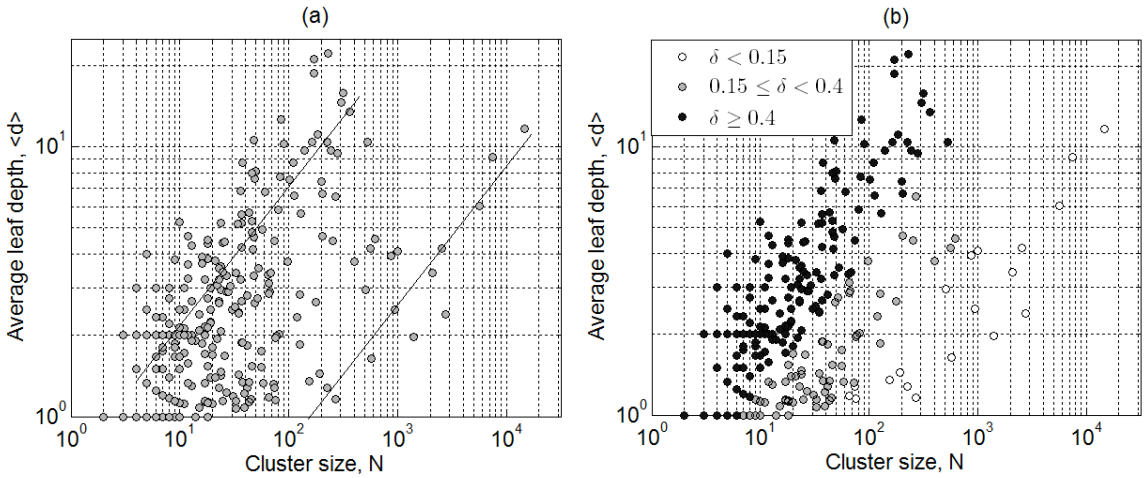

Figure B1: Two types of nearest-neighbor families. The figure shows the average leaf depth ‹*d*› as a function of the family size *N* for 452 regular families with maximal magnitude $m \geq 4$. The nearest-neighbor analysis is done for $m \geq 2.0$. Panel (a) depicts the two modes by lines $\langle d \rangle \propto N^{0.5}$: one of the modes is characterized by much larger average leaf depth for the same family size. Panel (b) further illustrates the two modes by using different colors for families with different normalized depth $\delta = \langle d \rangle \times N^{-0.5}$, as described in the legend.

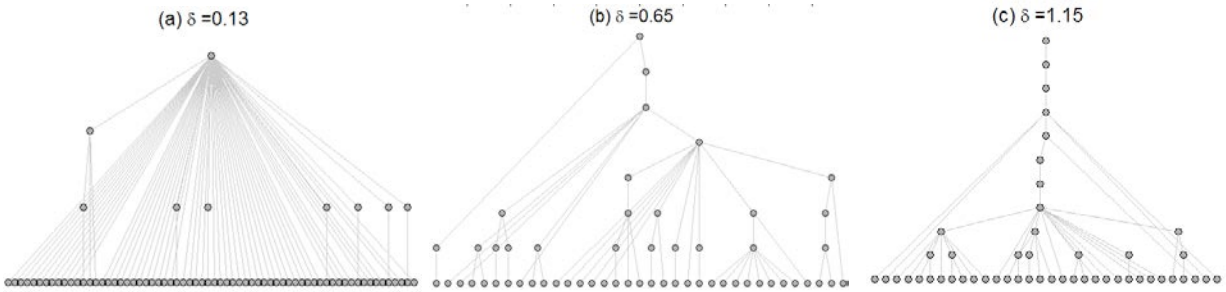

Figure B2: Examples of trees with different values of the normalized tree depth δ. All trees correspond to the earthquake families observed in southern California.

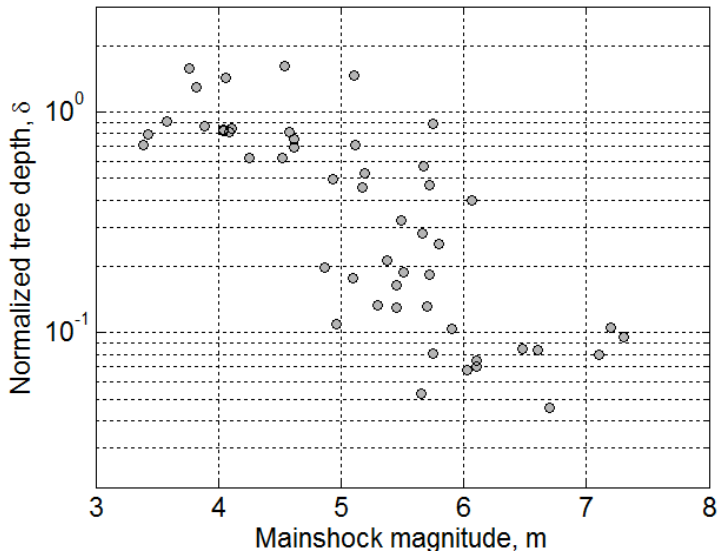

Figure B3: Two types of nearest-neighbor families. The figure shows the normalized tree depth δ as a function of the family mainshock magnitude m for 51 regular families with size $N \geq 100$. There exists a transition in the family formation process: all mainshocks with $m < 4.7$ correspond to large-depth trees (swarm-like families), $\delta > 0.5$; all mainshocks with $m > 6.2$ correspond to small-depth trees (burst-like families), $\delta < 0.11$; the mainshocks in the transition range $4.7 < m < 6.2$ may form families of various types, with $0.05 < \delta < 2$.

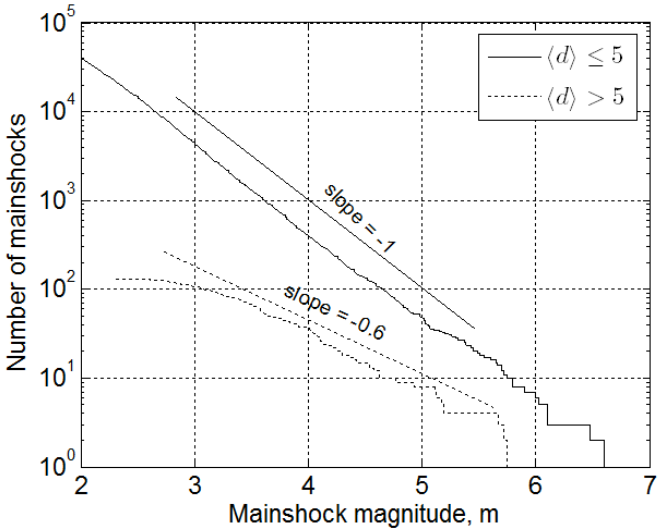

Figure C1: Number of mainshocks with magnitude equal or above m. The analysis is done separately for clusters with average leaf depth $\langle d \rangle \leq 5$ (solid line) and $\langle d \rangle > 5$ (dashed line). Deep, swarm-like families have significantly larger proportion of high-magnitude mainshocks.

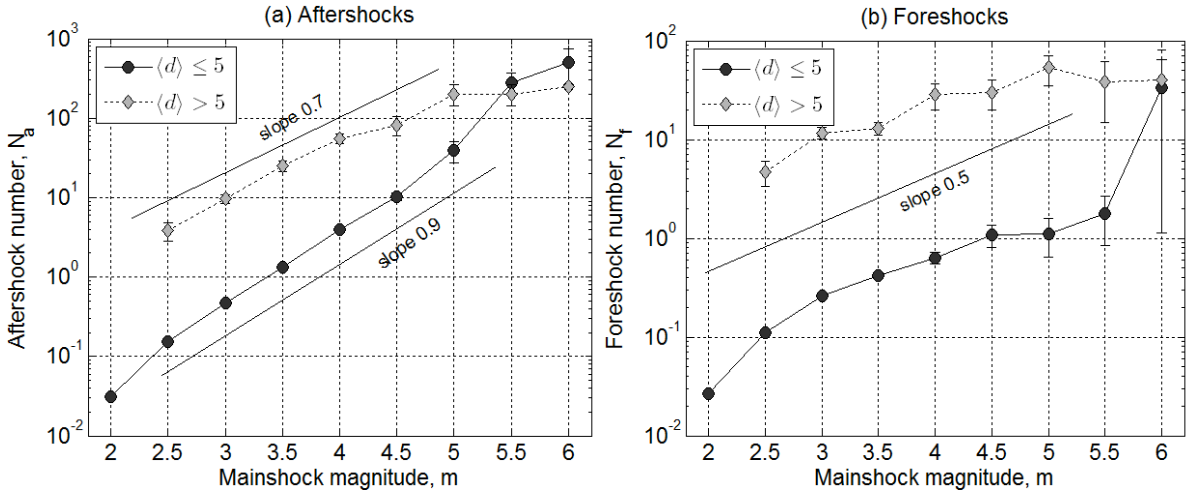

Figure C2: Average number of aftershocks (panel a) and foreshocks (panel b) per family for different mainshock magnitude. The analysis is done separately for shallow families, ‹*d*› ≤ 5, (solid line, circles) and deep families, ‹*d*› > 5, (dashed line, diamonds). The productivity is significantly larger in deep families.

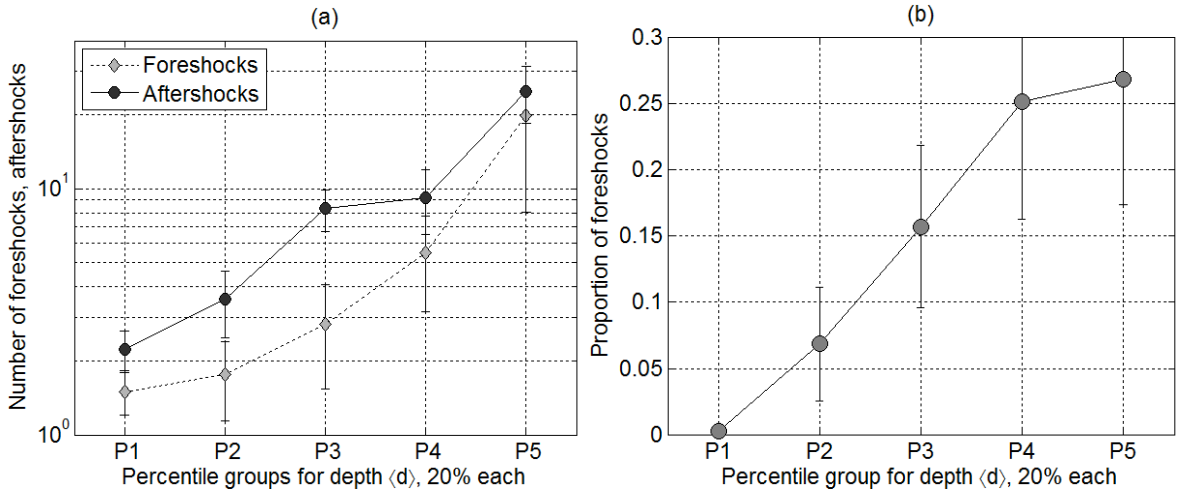

Figure C3: (a) Average number of Δ-foreshocks (diamonds, dashed line) and Δ-aftershocks (circles, solid line) for families with different average leaf depth ‹*d*›. Each group corresponds to 20% of the families in Δ-analysis, according to the increasing ‹*d*›-values. Only families with at least one fore/aftershock are examined. Each foreshock group contains 25 or 26 families; each aftershock group contains 27 or 28 families. The error bars correspond to a 95% confidence interval for the mean. (b) Proportion of foreshocks in Δ-families with size $N \geq 10$. Each group contains 22 or 23 families.

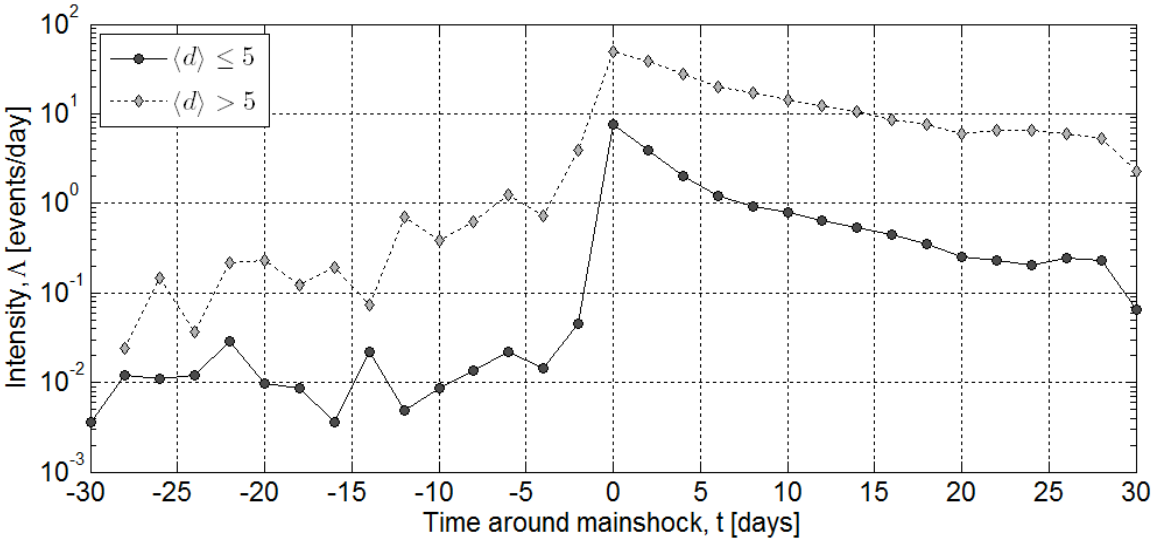

Figure C4: Intensity Λ of events around a mainshock in events per day per cluster; regular analysis, clusters with mainshocks $m \geq 4$. The analysis is done separately for clusters with ‹d› ≤ 5 (solid line, circles) and ‹d› > 5 (dashed line, diamonds).

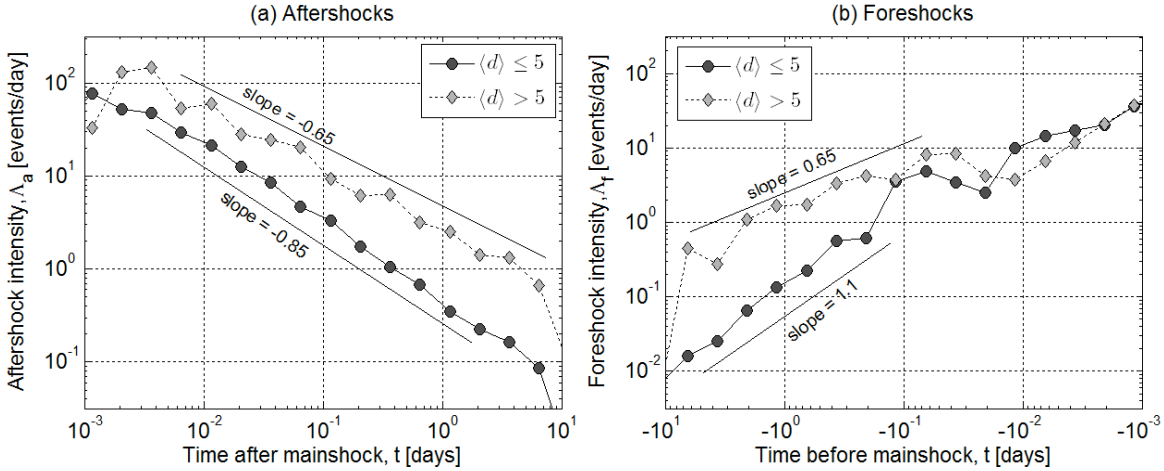

Figure C5: Intensity of aftershocks (panel a) and foreshocks (panel b) in events per day per family for families with mainshock magnitude $m \geq 4$ and at least one aftershock (panel a) or foreshock (panel b) in Δ-analysis. The analysis is done separately for families with ‹d› ≤ 5 (solid line, circles) and ‹d› > 5 (dashed line, diamonds). The event decay away from the mainshock is more rapid in topologically shallow families.

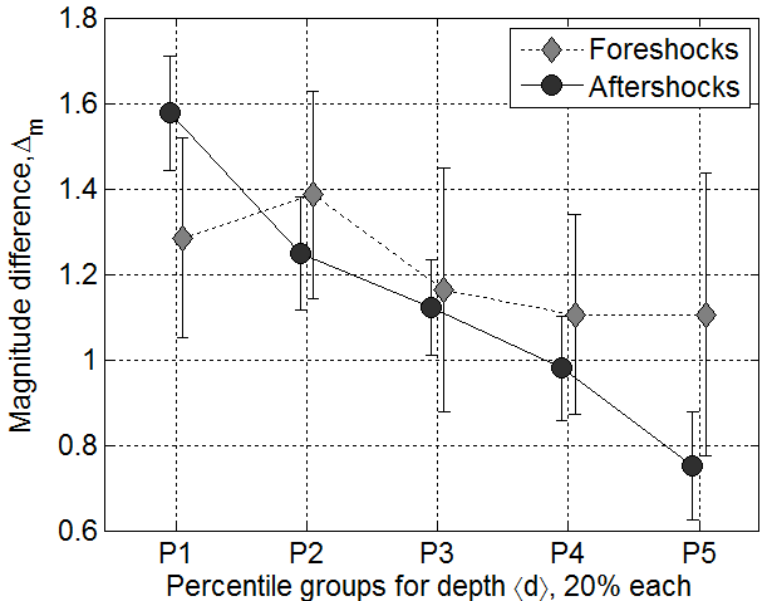

Figure C6: Magnitude difference Δ_m between the mainshock and the largest foreshock (dashed line, diamonds) and aftershock (solid line, circles) in regular analysis. The figure shows the average value of the magnitude difference for different ranges of the average leaf depth ‹*d*› in regular analysis. Each depth group corresponds to 20% of families with at least one fore/aftershock. Each aftershock group contains 67 or 68 events; each foreshock group contains 25 or 26 events. The error bars correspond to a 95% confidence interval for the mean.

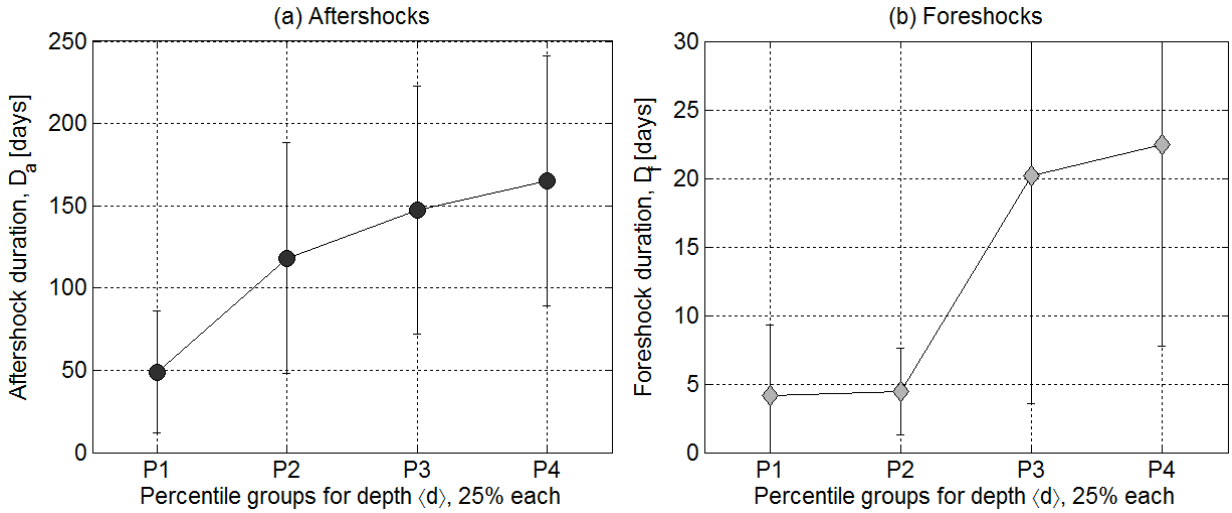

Figure C7: Duration of foreshock and aftershock sequences. The figure shows the average value of duration for different ranges of the average leaf depth ‹*d*› in Δ-analysis. Each depth group corresponds to 25% of families with at least one fore/aftershock. (a) Aftershocks, each group contains 84 or 85 sequences. (b) Foreshocks, each group contains 32 sequences. The error bars correspond to a 95% confidence interval for the mean.

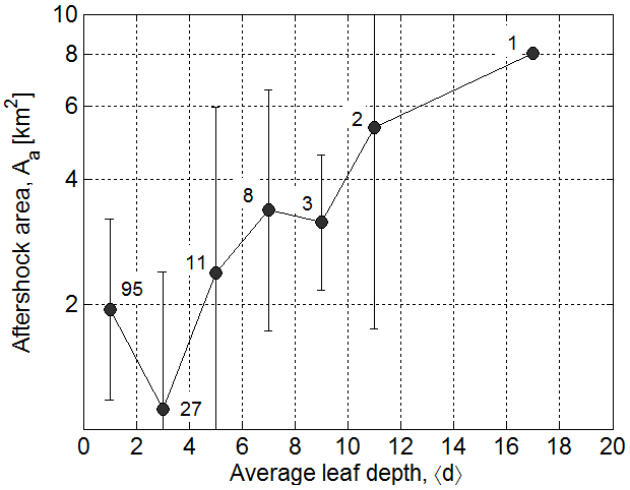

Figure C8: Area A occupied by aftershocks. The analysis only considers Δ-families with at least 5 aftershocks within 5 parent fault rupture lengths from the mainshock. The area is averaged over all families within different ranges of the average leaf depth, each range has length 2. The number of families within each range is indicated in figure.

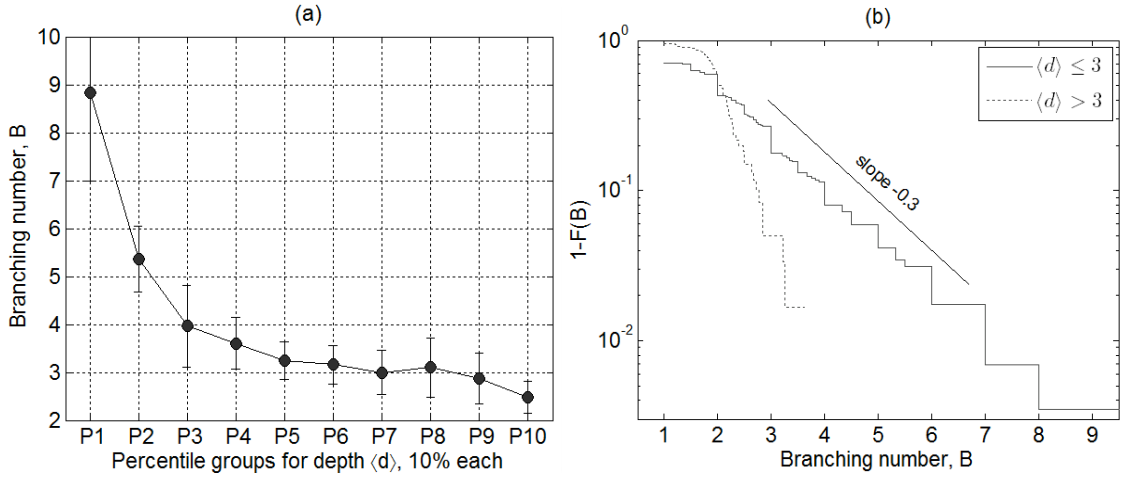

Figure C9: Branching number B. (a) The average value of the branching number for different ranges of the average leaf depth ‹d› in regular analysis. Each depth group corresponds to 10% of families with mainshock magnitude $m \geq 4$ and size $N \geq 10$. Each group contains 19 or 20 families. The error bars correspond to a 95% confidence interval for the mean. (b) The tail of the distribution of the branching number B for families with ‹d› ≤ 3 (solid line) and ‹d› > 3 (dashed line). Branching is larger for shallow families.

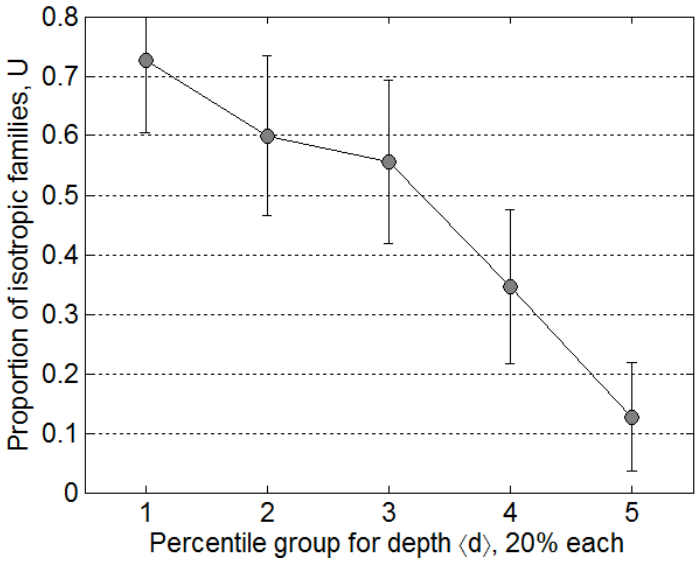

Figure C10: Circular spatial isotropy of family events. The figure shows the proportion of families with circularly uniform distribution of events relative to the mainshock, according to the Kolmogorov-Smirnov test at level 0.01 (see the text for details). Regular families with at least 5 events and mainshock magnitude $m \geq 4$ are considered. The results are averaged within families with different values of the average leaf depth, each point corresponds to 20% of examines families; each group contains 54 or 55 families.