



Information Research – Vol. 31 No. iConf (2026)

The duality of connectivity: how the connectivity of collaborative networks shapes disruptive innovation

Yingqun Li, Chuhui Zhang, Zhiwei Hu, and Lei Pei

DOI: <https://doi.org/10.47989/ir31iConf64132>

Abstract

Introduction. This study first reveals the nonlinear impact mechanism of algebraic connectivity in scientific collaboration networks on disruptive innovation, offering a novel perspective for understanding the boundary effects of knowledge flow and the emergence of innovation.

Method. Using academic papers in the field of computer science as empirical samples, we first quantify the degree of disruptive innovation in scientific papers using CD-index. Then, we employ the Louvain algorithm to identify the specific weighted collaboration network embedded by scientists and calculate its algebraic connectivity. A two-way fixed effects model is employed, with the quadratic term of algebraic connectivity introduced to test the nonlinear impact mechanism.

Results. A significant ‘inverted U-shaped’ relationship is found between the algebraic connectivity of scientific collaboration networks and disruptive innovation, with a inflection point at 0.174. Moderate levels of connectivity facilitate the integration of global knowledge, thereby promoting disruptive innovation. Both excessively high and low levels of connectivity suppress breakthroughs in disruptive innovation. Moderating effect reveals that higher knowledge density accelerates the marginal benefits of connectivity, allowing the network to form an innovation advantage at relatively low connectivity levels.

Conclusion. The emergence of disruptive innovation requires an emphasis on seeking structural balance between knowledge integration and cognitive diversity.

Introduction

Disruptive innovation has long been regarded as a key driver of social progress and transformation (Kuhn, 1996; Funk & Owen-Smith, 2016). However, recent studies indicate that the rate of disruptive innovation generation is gradually slowing down, even showing a marked decline (Park et al., 2023; Wu et al., 2019). The structural characteristics of scientific collaboration networks have consistently served as a critical analytical pathway for understanding innovation ecosystems, knowledge emergence, and research performance (Ahuja, 2000; Gilsing et al., 2008; Gonzalez-Brambila et al., 2013). Traditional research has often focused on the micro-level perspective, examining the structural positions of scientists within local networks (e.g., centrality, structural holes), and inferring how their information acquisition and resource control capabilities impact scientific output (Burt, 2004; Gonzalez-Brambila et al., 2013; Lazega & Burt, 1995). However, disruptive innovations often arise from the deep integration and recombination of knowledge across fields (Janssen & Frenken, 2019; Xu et al., 2024). The generation of such innovations highly relies on the emergent intelligence of groups, rather than simply aggregating individual capabilities (Ibrahimi et al., 2023). Therefore, a purely micro-level perspective is insufficient for fully uncovering the dynamics of systemic innovation, necessitating a shift towards a whole-network level analysis.

However, the proliferation of internet technologies and the rise of open science movements have facilitated greater mobility and autonomy in scientific collaboration models (Lu et al., 2024), while also exacerbating the fragmentation of scientific collaboration networks (Fares et al., 2021; Perc, 2010). The proliferation of ad hoc small collaborative teams formed to address specific problems (Driskell et al., 2024; Mayo, 2022) has contributed to a significant increase in network modularity and a persistent decrease in long-range connections. This structural evolution may hinder cross-group knowledge exchange and integration, thus impeding the emergence of disruptive innovations. In this context, algebraic connectivity, through spectral graph theory, integrates the topological information of the entire network, directly mapping the impact of network fragmentation on global knowledge flow (Di Lorenzo & Barbarossa, 2014; Zanetti et al., 2013). Unlike local metrics based on shortest paths (e.g., network efficiency), algebraic connectivity reflects the overall connectivity strength of the network via the second smallest eigenvalue of the Laplacian matrix: a lower value indicates that the network is more easily decomposed into isolated components, restricting the breadth and depth of knowledge flow; conversely, a higher value indicates stronger global connectivity, facilitating smoother knowledge exchange and flow (Di Lorenzo & Barbarossa, 2014), thereby creating favorable conditions for disruptive innovation.

Nevertheless, excessively high algebraic connectivity may also exert a suppressive effect on disruptive innovation. While high network connectivity enhances knowledge dissemination efficiency, it can also lead the network into predicaments of cognitive homogenization and information redundancy (Manso & Pourbabaee, 2022), undermining scientists' motivation and ability to explore non-redundant, long-range knowledge combinations (Ren & He, 2023). Furthermore, highly connected structures are often accompanied by path dependence, making scientists more inclined towards incremental innovation within mainstream paradigms rather than challenging existing knowledge frameworks (Stephen et al., 2015; Stolarczyk et al., 2017). Therefore, a nonlinear, inverted U-shaped relationship likely exists between the connectivity of scientific collaboration networks and disruptive innovation: moderate network connectivity promotes global knowledge integration, whereas excessive connectivity reinforces conservative tendencies, ultimately inhibiting the generation of disruptive breakthroughs. This duality raises a critical research question: *What is the functional form of the relationship between the algebraic connectivity of scientific collaboration networks and disruptive innovation?* We hypothesize that this relationship follows an inverted U-shaped curve: moderate network connectivity promotes global knowledge integration and disruptive innovation, whereas excessive connectivity reinforces conservative tendencies and ultimately inhibits disruptive breakthroughs.

This study uses academic papers from the field of computer science as empirical samples. The choice of this field is due to its dual characteristics: first, breakthroughs in computer science heavily rely on the recombination and cross-domain integration of heterogeneous knowledge (Asplund et al., 2020; Gamidullaeva et al., 2021); second, under the influence of the ‘open-source culture’ and ‘technological sharing’ movements, this field is experiencing significant fragmentation of its scientific collaboration networks, making it an ideal setting for studying the dynamic relationship between network connectivity structure and innovation. We quantify the degree of disruptive innovation in scientific papers using the CD-index and precisely locate the specific collaborative networks in which paper authors are embedded through the Louvain algorithm. To more accurately reflect the flow of knowledge in collaboration, network edge weights are determined based on the frequency of cooperation between scientists, and this information diffusion characteristic is captured using the weighted adjacency matrix in algebraic connectivity. To control for the time dynamics of disruptive innovation and individual heterogeneity of scientists, we construct a two-way fixed-effects model and introduce the quadratic term of algebraic connectivity to examine the nonlinear impact mechanism of scientific collaboration network connectivity on disruptive innovation.

The results show that there is a significant inverted U-shaped relationship between the connectivity of scientific collaboration networks and the degree of disruptive innovation: as connectivity increases, the level of disruptive innovation first rises and then declines. This suggests that moderate connectivity in scientific collaboration networks promotes disruptive innovation, while excessive connectivity inhibits it. The moderating effect reveals that in collaboration networks with higher knowledge density, the promoting effect of connectivity on disruptive innovation reaches its turning point earlier (from 0.105 to 0.079). This indicates that high knowledge density reduces the ‘connectivity threshold’ required to achieve optimal innovation outcomes, enabling scientific collaboration networks to form innovation advantages with relatively lower connectivity levels. This study breaks the limitations of the traditional micro-level perspective and provides the first empirical evidence of the nonlinear relationship between algebraic connectivity in scientific collaboration networks and disruptive innovation, thus offering a network topology-based approach for promoting innovation in research management and policy formulation.

Literature review

Disruptive innovation

Disruptive innovation plays a central role in driving fundamental progress in both science and society (Kuhn, 1996). To systematically understand the types and mechanisms of innovation, existing studies draw upon the theoretical foundations of scientific and technological change (Arthur, 2007; Tushman & Anderson, 1986), distinguishing between disruptive and integrative innovation (Funk & Owen-Smith, 2017). Disruptive research introduces new elements of knowledge, technological paths, or practical models, challenging or even reconstructing existing theoretical and methodological frameworks (Cao et al., 2022; Staudt et al., 2018). It overcomes the inertia that maintains the status quo, eventually leading to a shift in research paradigms and gaining widespread recognition within the academic community (Cole, 1970). In contrast, integrative research focuses on solving problems by utilizing existing theories and methods, strengthening connections between different disciplines or fields, and further consolidating the mainstream scientific paradigm (Funk & Owen-Smith, 2017).

A substantial body of research indicates that disruptive innovations often emerge from atypical combinations of knowledge or interdisciplinary concept integration (Uzzi et al., 2013). This process breaks through inherent cognitive boundaries within specific fields (Cao et al., 2022). For instance, Wang et al. (2017) found that highly disruptive papers typically build upon distant knowledge recombinations or non-mainstream ideas. Additionally, institutional environments and resource

allocation mechanisms play an important moderating role in the generation of disruptive innovation. For example, funding structures that support high-risk, high-reward projects are more likely to foster disruptive outcomes (Jacob & Hellström, 2023; Savona et al., 2025). In terms of social mechanisms, the structural characteristics of scientific collaboration networks are considered key factors influencing innovation output. Weak ties can bridge different information communities, facilitating the dissemination and validation of novel ideas (Granovetter, 1973). Structural attributes of collaboration networks, such as centrality and structural holes, have been widely shown to enhance conventional innovation performance (Burt, 2004; Gonzalez-Brambila et al., 2013; Lazega & Burt, 1995). While previous studies have explored the association between the macro characteristics of collaboration networks and disruptive innovation, there is still a lack of systematic theoretical and empirical tests on how the connectivity of scientific collaboration networks specifically influences the mechanisms underlying disruptive innovation.

Connectivity in scientific collaboration networks

Scientific collaboration networks, as key infrastructure for knowledge production, have their connectivity profoundly influencing the efficiency of knowledge flow and the resilience of the system (Crespo et al., 2013; Dallmann et al., 2024; Zhang et al., 2023). Compared to path-dependent metrics (such as average path length or clustering coefficient), algebraic connectivity, based on spectral graph theory, quantifies the global connectivity of a network by calculating the second smallest eigenvalue of the Laplacian matrix (Buldú et al., 2019; Freitas et al., 2022). This metric captures not only the structural robustness of a network against fragmentation but also the synergistic efficiency of information propagation throughout the entire network (Chan & Akoglu, 2016; Zheng et al., 2022). A higher algebraic connectivity value signifies more efficient and stable knowledge dissemination, making it particularly relevant for processes of collective cognitive integration (Di Lorenzo & Barbarossa, 2014; Zanetti et al., 2013). Theoretically, it holds greater potential for supporting the heterogeneous knowledge fusion necessary for disruptive innovation.

However, excessively high global connectivity may lead to cognitive homogenization and information redundancy (Manso & Pourbabaee, 2022). While strong connectivity promotes rapid knowledge diffusion, it may also accelerate the entrenchment of mainstream paradigms (Stephen et al., 2015), inducing groupthink and path dependence, thereby inhibiting support for non-consensus, high-risk explorations (Stolarczyk et al., 2017). In such an environment, researchers are more likely to engage in incremental research rather than challenge existing paradigms (Ren & He, 2023). Thus, the relationship between algebraic connectivity and disruptive innovation may not be monotonic, but rather exhibit an inverted U-shape: moderate connectivity provides the structural foundation for cross-domain knowledge recombination, while excessive connectivity may hinder disruptive innovation by weakening the necessary structural diversity and cognitive heterogeneity.

Methodology

Data

This study is based on the Web of Science dataset, selecting papers from the ‘computer science, interdisciplinary applications’ category as research samples. This category focuses on the integration and application of computer science in interdisciplinary contexts, covering 170 high-quality journals and involving several cutting-edge interdisciplinary fields, such as machine intelligence (ISSN: 2522-5839), medical imaging (ISSN: 1361-8415), smart agriculture (ISSN: 2214-3173), and industrial Internet of Things (ISSN: 0166-3615). Subsequently, the study retrieved the metadata and citation network information of all relevant papers from the OpenAlex platform using the journal ISSNs. OpenAlex fully inherits and integrates the final version of the Microsoft Academic Graph (MAG) and supplements it with data from multiple authoritative databases, including Crossref, PubMed, and ORCID. Additionally, knowledge entities associated with each paper were extracted from OpenAlex. The platform uses an automated annotation system based

on Wikidata concepts to label the specific knowledge domains involved in the content of the papers.

The resulting dataset constructed in this study covers 547,890 papers published between 1980 and 2024, representing 274,728 unique scientists and encompassing a total of 4,798 knowledge concepts, which serves as the basis for constructing the scientific collaboration network. To ensure the continuity of research behavior in the empirical analysis and avoid potential distortions to the model's robustness from incidental academic activities, we adhered to established research practices and included only researchers who had published at least three papers and demonstrated sustained innovative behavior in the regression analysis. Following this filtering process, a total of 47,280 active researchers were included in the model.

Variables

Dependent variable

This study employs the CD-index, proposed by Funk and Owen-Smith (2016), to quantify the degree of disruptive innovation for each individual publication. The dependent variable is operationalized as the average CD-index value of all publications authored by scientist i in year t , denoted as $\text{Disruptiveness}_{it}$. The CD-index, leveraging the directed citation network of scientific papers, offers insights into how a scientific paper influences future scientific trajectories. The intuition is that if a paper is disruptive, the subsequent work that cites it is less likely to also cite its predecessors; for future researchers, the ideas that went into its production are less relevant. Conversely, if a paper is consolidating, subsequent works citing it are more likely to reference the viewpoints of its predecessors; for future researchers, the knowledge upon which that work is based remains (or perhaps becomes even more) relevant. The CD-index indicates the degree of consolidation or disruption, ranging from -1 (consolidating) to 1 (disruptive). Compared to traditional qualitative analysis methods, the CD-index serves as a post-hoc evaluation tool, objectively reflecting the academic community's recognition of disruptive innovation. The CD-index has undergone extensive validation and exploration through expert assessments and empirical analyses (Funk & Owen-Smith, 2017; L. Wu et al., 2019; Park et al., 2023; Wei et al., 2023). Figure 1 illustrates the overall calculation methodology and details.

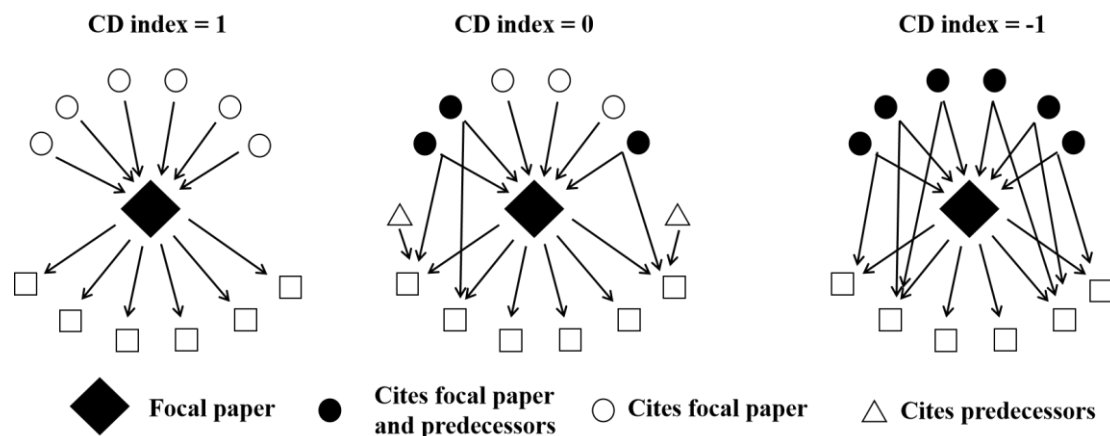


Figure 1. Overview of the CD index measurement approach

Funk and Owen-Smith (2017) defined the CD index as follows; Where $f_{it}=1$ if i cites the focal paper; 0 if not, $b_{it}=1$ if i cites the predecessors of the focal paper; 0 if not. n_t : number of forward cites to the focal work and/or its predecessors as of time t .

$$CD_t = \frac{1}{n_t} \sum_{i=1}^n -2f_{it}b_{it} + f_{it}$$

To precisely calculate the degree of disruptive innovation in scientific papers, we applied the following procedures: First, we excluded articles such as editorials, book reviews, and academic commentaries. These types of articles generally lack citation network structures, or their network structures are biased, making them ineffective in reflecting trends and directions in scientific innovation. Second, review articles, which typically focus on summarizing and revisiting previous work, are conservative rather than disruptive in nature (Wu et al., 2019). Therefore, these articles were also excluded from the analysis to avoid introducing systematic bias. Finally, to ensure the robustness of the indicator and accurately reflect the long-term impact of the papers, we calculated the CD-index based on the citation network five years after each paper's publication. To maintain a consistent time window for all papers, we excluded those published after 2019.

Independent variables

This study aims to examine the causal effect of the structural characteristics of scientific collaboration networks on subsequent innovative output. To address the simultaneity bias inherent in network causal analysis between the independent and dependent variables, this paper adopts a classic time-window lag design framework from network causal inference (Gallo & Plunket, 2020; Gonzalez-Brambila et al., 2013; Guan et al., 2015). Specifically, for a publication i published in period t , its explanatory variables are derived from the scientific collaboration network constructed in a preceding period (from $t-3$ to $t-1$). For instance, for a paper published in 2010 (period t), its independent variables are calculated based on collaboration network data from 2007-2009 (periods $t-3$ to $t-1$). The choice of a three-year time window enables us to effectively capture the dynamic characteristics of the network, while ensuring a sufficiently large sample size to support subsequent statistical analyses, thereby further strengthening the robustness of the research findings (Gonzalez-Brambila et al., 2013). In subsequent analyses, we also test different time window lengths for comparative validation.

Once the time window is determined, it is crucial to accurately identify the specific collaboration network in which the authors are embedded. We construct the global scientific collaboration network from periods $t-3$ to $t-1$ and employ the Louvain algorithm for community detection. By maximizing modularity, the global network is partitioned into sub-communities with distinct structural characteristics. In a scientific context, these subgroups correspond to stable collaborative ecosystems based on shared research paradigms, institutional affiliations, or geographic proximity, representing the local knowledge network and collaborative constraints embedded by scientists before producing new work (Lambiotte & Panzarasa, 2009; Sciabolazza et al., 2017; Yang & Wang, 2024). To ensure the reproducibility of our results, we fixed the random seed to guarantee consistency in community partitioning. Furthermore, we extracted the giant connected component from each scientist's subgroup to eliminate interference from isolated nodes on the network connectivity measure.

Algebraic connectivity was selected to quantify the global connectivity of the scientific collaboration network due to its spectral properties, which are directly linked to information diffusion efficiency, its sensitivity to collaboration frequency weights, and its ability to detect critical structural changes. Specifically, for a network comprising n scientists, its weighted adjacency matrix $W = [w_{ij}]$ has elements w_{ij} representing the collaboration frequency between scientists i and j , where $w_{ij} = w_{ji} \geq 0$, and $w_{ii} = 0$. The degree matrix D is a diagonal matrix where $d_{ii} = \sum_{j=1}^n w_{ij}$. The weighted Laplacian matrix is defined as $L = D - W$. Let $\lambda_n(L)$ be the largest eigenvalue, and $\lambda_2(L)$ the second smallest eigenvalue of the weighted Laplacian matrix L is the algebraic connectivity. To mitigate the influence of network size n on the results, we used relative algebraic connectivity as the core metric, calculated as follows:

$$\text{Algebraic_connectivity (AC)} = \frac{\lambda_2(L)}{\lambda_n(L)}$$

The theoretical range of this coefficient is [0,1]. A higher coefficient indicates stronger algebraic connectivity, which suggests that the scientific collaboration network is more efficient and stable in terms of knowledge diffusion.

Control variables

Innovative inertia of scientists: To control the impact of scientists' short-term innovation patterns on subsequent disruptive innovation, this study analyzes the average disruptive innovation index of the papers published during the previous time window, which serves as a measure of their innovation habits in the short term.

Reputation indicators of scientists: To control for the potential confounding effect of individual scientist reputation on innovative output, two proxy variables were introduced: 1) the total number of publications by the scientist in the preceding time window, and 2) the scientist's H-index at time t . These variables collectively reflect their academic reputation and productivity.

Network position indicators of scientists: To eliminate potential confounding effects arising from the scientist's position in the collaboration network, and to ensure that the observed effects originate from the broader network environment rather than from individual structural characteristics, this study utilizes two network centrality measures: the normalized eigenvector centrality and the clustering coefficient of the node. Specifically, eigenvector centrality reflects the global influence of a node within its network neighborhood, while the clustering coefficient measures the proportion of closed triangle relationships formed between the node and its neighboring nodes, thereby influencing the node's exposure to locally concentrated and mutually reinforced information. The specific calculation formulas are as follows:

$$C_i = \frac{2T(i)}{k_i(k_i - 1)}$$

Where $T(i)$ represents the number of actual edges between scientist i and its neighboring nodes, i.e., the number of triangles formed by node i . k_i denotes the degree of node i , $k_i(k_i - 1)$ representing the maximum number of edges that could potentially be formed between node i and all of its neighbors.

Empirical model

This study employs a two-way fixed effects model for empirical analysis. This model simultaneously controls for time-invariant individual heterogeneity and common temporal trends, thereby enhancing the reliability of parameter estimates. The specific formulation is as follows:

$$Disruptiveness_{it} = \alpha + \beta_1 AC_{i,(t-1,t-3)} + \beta_2 AC_{i,(t-1,t-3)}^2 + \chi Control_{i,(t-1,t-3)} + \lambda_t + \mu_i + \varepsilon_{it}$$

Where i and t denote the individual scientist and year, respectively. The dependent variable, $Disruptiveness_{it}$ represents the average disruptive innovation index (CD-index) of publications produced by scientist i in year t . $AC_{i,(t-1,t-3)}$ is the core independent variable, representing the algebraic connectivity of the scientific collaboration network in which scientist i was embedded during the prior time window (periods $t-3$ to $t-1$). $AC_{i,(t-1,t-3)}^2$ denotes its quadratic term, introduced to examine the potential nonlinear relationship between network connectivity and disruptive innovative output. $Control_{i,(t-1,t-3)}$ is a vector of control variables measured in the prior time window. μ_i and λ_t represent individual and year fixed effects, respectively, and ε_{it} stands for the residual term. To mitigate the impact of individual-level heteroscedasticity and serial correlation on standard error estimation, clustered standard errors at the individual scientist level are applied in this study.

Results

Baseline model

Table 1 reports the baseline results. Column (1) introduces both the core independent variable—the algebraic connectivity of the scientific collaboration network—and its quadratic term. The empirical results indicate that the coefficient for the linear term of algebraic connectivity is significantly positive at the 1% level ($\beta_1 = 0.210, p < 0.01$), while the coefficient for the quadratic term is significantly negative at the 1% level ($\beta_2 = -0.602, p < 0.01$). This preliminarily supports the existence of an inverted U-shaped relationship between algebraic connectivity and disruptive innovation. Columns (2) and (3) incrementally incorporate control variables at the scientist individual level and network level, respectively. The results show that after controlling for these variables, the coefficients of both the linear and quadratic terms of algebraic connectivity remain statistically significant, with their signs consistent with those in Column (1). The inverted U-shaped relationship continues to hold robustly. This further reinforces the reliability and validity of the baseline regression conclusions.

Dependent variable: Disruptiveness			
VARIABLES	FE	FE	FE
	(1)	(2)	(3)
Algebraic_connectivity	0.210*** (0.00747)	0.201*** (0.00729)	0.194*** (0.00713)
Algebraic_connectivity ²	-0.602*** (0.00166)	-0.569*** (0.00170)	-0.558*** (0.00175)
Number of papers		-8.49e-05 (5.27e-05)	-7.85e-05 (5.15e-05)
H_index		-7.29e-06 (1.37e-05)	-5.12e-06 (1.34e-05)
Innovation_inertia		0.189*** (0.00217)	0.184*** (0.00212)
Eigenvector_centrality			0.233*** (0.00338)
Clustering_coefficient			-0.000348 (0.00104)
Constant	0.0919*** (0.000951)	0.0845*** (0.00102)	0.0755*** (0.00118)
Year fixed effects	√	√	√
Individual fixed effects	√	√	√
Observations	228,179	228,179	228,179
Individual	47280	47280	47280
Adj R-squared	0.423	0.450	0.473

NOTE: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 1. Baseline results

To further validate the postulated inverted U-shaped relationship between scientific collaboration network connectivity and disruptive innovation, this study adopts the rigorous U-test procedure proposed by Lind and Mehlum (2009). This test consists of three primary steps: First, confirming that the coefficient of the quadratic term is significantly negative. The regression results in columns (3) show that the quadratic term (Algebraic_connectivity²) coefficient is -0.558, significant at the 1% level. Second, verifying that the inflection point lies within the data range of the independent variable. Based on the regression results from Column (3), the calculated inflection point is 0.174, which falls within the 95% Fieller confidence interval [0.161, 0.186] (*p-value* < 0.01), providing further support for the existence of the inverted U-shaped relationship. Finally, testing that the slopes on both sides of the inflection point are significant and of opposite signs. The results indicate that, on the left side of the inflection point (lower range), the slope is positive

($\beta=0.160$, $p\text{-value} < 0.001$), whereas on the right side of the inflection point (upper range), the slope is negative ($\beta=-0.096$, $p\text{-value} < 0.001$). The slopes on both sides are of opposite signs and are highly significant at the 1% level, providing strong evidence for the existence of an inverted U-shaped relationship.

For intuitive visualization of this nonlinear relationship, we plotted the marginal effect curve of algebraic connectivity on disruptive innovation based on the regression results from Column (3). As shown in Figure 2, on the left side of the inflection point, an increase in the connectivity of the scientific collaboration network promotes disruptive innovation; the effect reaches its maximum at the inflection point of 0.174; beyond this point, further increases in connectivity exert a suppressive effect on disruptive innovation. Mechanistically, a moderate level of network connectivity helps mitigate the risks associated with collaborative fragmentation, broadening and deepening knowledge flows, thereby creating favorable conditions for disruptive innovation. However, excessively high connectivity may lead to group polarization and path dependence, stifling the cross-fertilization and creative collision of diverse knowledge domains, consequently hindering the emergence of disruptive innovation.

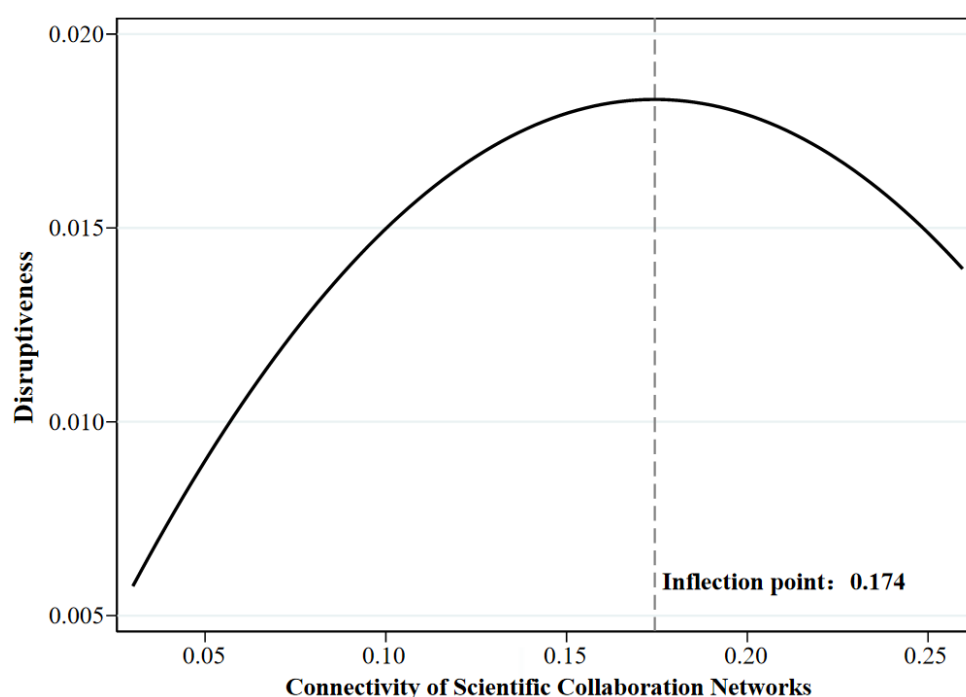


Figure 2. Overview of the CD index measurement approach

The moderating effects of knowledge density

The efficiency and stability of knowledge dissemination serve as the core mechanism by which connectivity fosters disruptive innovation. To empirically test the validity of this mechanism, this paper introduces the concept of 'knowledge density' as a moderating variable, in order to analyze the differentiated impact of connectivity on disruptive innovation at varying levels of knowledge aggregation. Knowledge density is defined as the ratio of the total number of knowledge entities to the number of scientists within a scientific collaboration network. Compared to directly using the total amount of knowledge, knowledge density more accurately reflects the abundance of knowledge resources that can be accessed and integrated by individual scientists. It also aligns more closely with the inherent logic of knowledge flow and spill-over in real-world contexts, making it more suitable for capturing the moderating role of connectivity in the knowledge dissemination process.

The measurement of knowledge entities is based on the distinct knowledge concepts associated with the publications produced by scientists during the collaboration period. Data on the knowledge concepts corresponding to each publication are sourced from the OpenAlex database. Empirically, we test this moderating effect by constructing interaction terms between knowledge density and both the linear and quadratic terms of algebraic connectivity. These interaction terms are incorporated into the baseline two-way fixed effects model. The regression results are presented in Table 2. In Column (1) of Table 2, the coefficients for the interaction term between algebraic connectivity and knowledge density ($AC \times KD$) and its quadratic term ($AC^2 \times KD$) are both negative and statistically significant at the 1% level. After further including a set of control variables in Column (2), these two interaction terms remain significant. The results indicate that knowledge density exerts a negative moderating effect: as knowledge density increases, the inverted U-shaped relationship between algebraic connectivity and disruptive innovation becomes more pronounced, characterized by a steeper curve and greater curvature.

Dependent variable: Disruptiveness		
VARIABLES	FE	FE
	(1)	(2)
Algebraic_connectivity (AC)	0.170*** (0.00808)	0.158*** (0.00770)
Algebraic_connectivity ² (AC ²)	-0.308*** (0.00227)	-0.285*** (0.00228)
Knowledge_density (KD)	-0.0148*** (0.000438)	-0.0139*** (0.000419)
AC×KD	-0.0106*** (0.00352)	-0.00934*** (0.00337)
AC ² ×KD	-0.282*** (0.00214)	-0.271*** (0.00211)
Number of papers		-5.03e-05 (4.41e-05)
H_index		-7.03e-06 (1.14e-05)
Innovation_inertia		0.131*** (0.00176)
Eigenvector_centrality		0.212*** (0.00301)
Clustering_coefficient		-0.00112 (0.000885)
Constant	0.131*** (0.00103)	0.116*** (0.00118)
Year fixed effects	√	√
Individual fixed effects	√	√
Observations	228,179	228,179
Individual	47280	47280
Adj R-squared	0.481	0.516

NOTE: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2: Results of Moderating Effects

Figure 3 illustrates the marginal effect of scientific collaboration network connectivity on disruptive innovation at different levels of knowledge density. The high knowledge density group

is defined as ‘mean + one standard deviation,’ and the low knowledge density group as ‘mean - one standard deviation’. The results show that in scientific collaboration networks with higher knowledge density, the inflection point of connectivity's impact on disruptive innovation occurs earlier, shifting forward from 0.105 under mean conditions to approximately 0.079. This suggests that higher knowledge density accelerates the rate of change in the marginal effect of connectivity, thereby lowering the ‘optimal connectivity threshold’ required for maximizing disruptive innovation. Consequently, collaboration networks can rapidly establish an innovation advantage at relatively lower levels of connectivity. Conversely, in networks with lower knowledge density, the inflection point shifts to the right, implying that a higher level of connectivity is necessary to achieve optimal innovative outcomes.

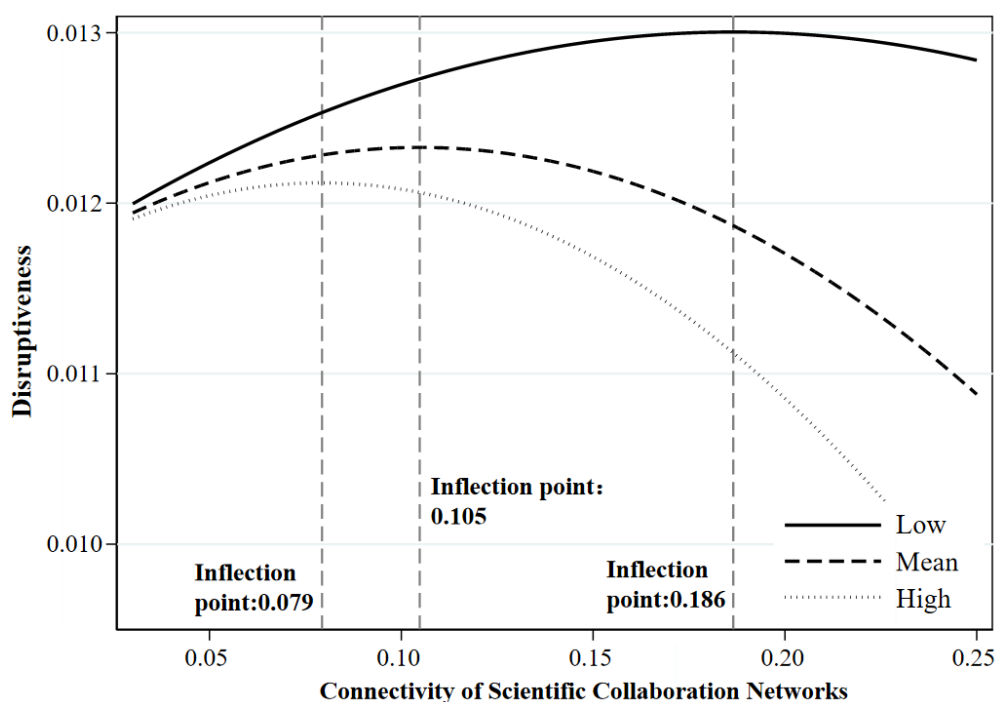


Figure 2. Moderating Effects of Knowledge Density in Collaborative Networks

Robustness check

Measurement based on CI^{nok}

Current criticism of CD-index focuses on the use of predecessors of the focal paper's cited references. The impact of the focal paper's cited references may disproportionately affect the final score when it is used since the core issue is the degree to which new papers cite the focus paper, but not the papers that the focus paper is based on. Another problem with the term predecessor is that it is often very large compared to the other terms in the formula, which leads to small absolute differences between the final scores and hence is not conducive to further observations. To avoid such problems, Bornmann et al. (2020) simply ignores predecessors and proposes a new form CI^{nok} ($I=5$) to solve the problems. Therefore, we used this improved metric and generated new results according to the baseline model. The results are depicted in Table 3. The directions of all coefficients are consistent with those in the baseline model, and all are statistically significant at the 0.01 level. This result confirms the robustness of our findings.

Dependent variable: Value of C^{mok} ($J=5$)			
VARIABLES	FE	FE	FE
	(1)	(2)	(3)
Algebraic_connectivity	1.982*** (0.0439)	1.694*** (0.0406)	1.577*** (0.0390)
Algebraic_connectivity ²	-5.991*** (0.00994)	-5.058*** (0.0128)	-4.967*** (0.0123)
Number of papers		0.000289 (0.000291)	0.000324 (0.000280)
H_index		7.80e-05 (7.58e-05)	5.31e-05 (7.28e-05)
Innovation_inertia		0.258*** (0.00186)	0.238*** (0.00180)
Eigenvector_centrality			1.790*** (0.0187)
Clustering_coefficient			-0.00525 (0.00611)
Constant	0.927*** (0.00561)	0.766*** (0.00582)	0.708*** (0.00689)
Year fixed effects	✓	✓	✓
Individual fixed effects	✓	✓	✓
Observations	228,179	228,179	228,179
Individual	47280	47280	47280
Adj R-squared	0.523	0.581	0.624

NOTE: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3: Results based on the C^{mok}

Extending the lagged time window

In network causal analysis, constructing network variables based on a preceding time window is a standard approach to mitigate potential simultaneity bias between the dependent and independent variables (Gallo & Plunket, 2020; Gonzalez-Brambila et al., 2013; Guan et al., 2015). Accordingly, this study initially constructed the scientific collaboration network using a three-year preceding window. To further enhance the robustness of our findings, we extended the observation window to four and five years, respectively, and re-ran the regression analyses. The results, presented in Table 4, demonstrate that the direction and statistical significance of the coefficients for the core explanatory variables remain consistent with the baseline results across different window lengths. This indicates that the main findings of this study are insensitive to the choice of time window, thereby underscoring their robustness.

Dependent variable: Disruptiveness		
VARIABLES	4-year window	5-year window
	(1)	(2)
Algebraic_connectivity	0.290*** (0.00283)	0.296*** (0.00266)
Algebraic_connectivity ²	-0.676*** (0.00137)	-0.693*** (0.00140)
Number of papers	3.65e-06 (4.29e-05)	-5.39e-05 (4.17e-05)
H_index	1.02e-05 (1.12e-05)	-1.00e-06 (1.07e-05)
Innovation_inertia	0.0907*** (0.00163)	0.119*** (0.00169)
Eigenvector_centrality	0.182*** (0.00274)	0.277*** (0.00274)
Clustering_coefficient	0.000640 (0.000922)	-0.000345 (0.000895)
Constant	0.0816*** (0.000845)	0.0792*** (0.000819)
Year fixed effects	✓	✓
Individual fixed effects	✓	✓
Observations	228,179	228,179
Individual	47280	47280
Adj R-squared	0.450	0.473

NOTE: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 4: Results based on the lagged time window

Controlling for scientist-specific temporal trends

Although the baseline model incorporates two-way fixed effects to account for time-invariant individual characteristics and common temporal trends, individual scientists may still exhibit heterogeneous behavioral patterns that evolve over time (Foster et al., 2015). For instance, some scientists may continuously adjust their research directions, adapt to emerging fields, or seek breakthroughs, while others may increasingly reinforce risk-averse tendencies, entrenching themselves in established research trajectories. To mitigate potential confounding effects arising from such scientist-specific temporal trends on the core causal estimates, we further augment the model by including interactions between individual scientists and a linear time trend. The extended model is specified as follows:

$$Disruptiveness_{it} = \alpha + \beta_1 AC_{i,(t-1,t-3)} + \beta_2 AC_{i,(t-1,t-3)}^2 + \chi Control_{i,(t-1,t-3)} + \lambda_t \times \mu_i + \varepsilon_{it}$$

Here, the term $\lambda_t \times \mu_i$ controls for a linear time trend specific to scientist i . The results, presented in Table 5. After incorporating this specification, the direction, statistical significance, and economic magnitude of the coefficients for the core explanatory variables remain highly consistent with the baseline results. This provides further evidence supporting the robustness of our conclusions.

Dependent variable: <i>Disruptiveness</i>			
VARIABLES	FE-IX	FE-IX	FE-IX
	(1)	(2)	(3)
Algebraic_connectivity	0.210*** (0.00839)	0.201*** (0.00819)	0.194*** (0.00801)
Algebraic_connectivity ²	-0.602*** (0.00187)	-0.569*** (0.00191)	-0.558*** (0.00197)
Number of papers		-8.50e-05 (5.92e-05)	-7.87e-05 (5.79e-05)
H_index		-7.04e-06 (1.54e-05)	-4.95e-06 (1.51e-05)
Innovation_inertia		0.189*** (0.00244)	0.184*** (0.00238)
Eigenvector_centrality			0.233*** (0.00380)
Clustering_coefficient			-0.000364 (0.00116)
Year fixed effects	✓	✓	✓
Individual fixed effects	✓	✓	✓
Individual × Year	✓	✓	✓
Observations	228,179	228,179	228,179
Individual	47280	47280	47280
Adj R-squared	0.424	0.452	0.472

NOTE: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 5: Results Controlling for Scientist-Specific Linear Time Trends

Discussion

This study provides the first empirical evidence, drawn from computer science collaboration networks, of a significant inverted U-shaped relationship between algebraic connectivity and disruptive innovation. This finding extends current theoretical frameworks on scientific

collaboration and innovation and offers practical insights for research organization and policy design.

The primary theoretical contribution of this work is its shift from micro-level individual features to macro-level network properties in understanding innovation capacity. Algebraic connectivity, as a spectral measure of global network cohesion, effectively captures how collaborative structures either constrain or facilitate knowledge integration. Our results show that moderate connectivity offers an optimal balance: it enables global knowledge recombination without causing cognitive homogenization and redundancy characteristic of overly dense networks. The inverted U-shaped pattern reflects fundamental mechanisms of knowledge integration and innovation. Low connectivity isolates groups, limiting cross-domain synthesis despite preserving local diversity. Increased connectivity improves access to non-redundant knowledge, supporting disruptive idea generation. Beyond a certain point, however, high connectivity induces information redundancy, group consensus, and path dependence, which suppress radical exploration. In the context of digital scholarship, excessive connectivity may also amplify echo chambers and reduce critical engagement with divergent ideas, undermining the inclusivity and enlightenment goals of digital knowledge societies.

These conclusions carry important implications for research management. Rather than indiscriminately promoting large-scale collaborations, research agencies and managers should aim for a balance between modularity and connectivity. Medium- to long-term collaboration plans and interdisciplinary platforms can be structured to avoid excessive inter-team coupling, thereby preserving structural flexibility for exploration and recombination. Such an approach not only enhances innovation potential but also fosters a more inclusive research ecology where diverse voices and knowledge traditions can interact authentically. In project review and team assembly, evaluations should include overall network topology to avoid high homogenization and foster environments conducive to disruptive innovation. From a digital society perspective, supporting structurally balanced collaboration networks can help cultivate information-literate communities capable of discerning, integrating, and advancing knowledge in a credible and socially meaningful manner.

While this study reveals a nonlinear relationship between the algebraic connectivity of scientific collaboration networks and disruptive innovation, several limitations remain, pointing to directions for future research. First, our empirical sample focuses on the field of computer science. Although this field serves as an ideal research context due to its characteristics of knowledge recombination and trend toward network fragmentation, systematic differences may exist across disciplines in terms of knowledge structures, collaboration norms, and innovation mechanisms. Second, this study employs static collaboration networks based on co-authorship data. While algebraic connectivity effectively captures the overall connection strength of a network, scientific collaboration is inherently a dynamic, evolving process. Future research could incorporate temporal network analysis to explore how dynamic changes in algebraic connectivity influence the evolution of innovative outputs and to identify critical network restructuring phases preceding breakthrough innovations. Third, the measurement of disruptive innovation relies on the CD index, which is based on subsequent citation patterns. Although the CD index is widely used, it is essentially an ex-post measure and may be influenced by disciplinary citation practices and the choice of the time window. Future work could integrate peer assessment, topic modeling analysis, or patent conversion data to develop multidimensional, forward-looking measures of disruptive innovation, thereby enhancing the robustness of the findings.

Conclusion

This study systematically examines the nonlinear relationship between the algebraic connectivity of scientific collaboration networks and disruptive innovation, using the computer science domain as an empirical setting. It confirms an inverted U-shaped relationship, indicating that moderate

connectivity promotes cross-group knowledge integration and recombination, thereby fostering disruptive innovation. Either too little or too much connectivity inhibits innovation: low connectivity restricts knowledge flow, while high connectivity leads to cognitive homogenization and path dependence. Moreover, the moderating effect of network knowledge density shows that high knowledge density reduces the connectivity level required for optimal innovation, as it accelerates the marginal effect change of connectivity and helps establish an innovation advantage more rapidly.

Theoretically, this work shifts the focus from micro-level individual features to macro-level network properties in studying collaboration and innovation. Algebraic connectivity offers a holistic metric for analysing how collaborative structures shape knowledge flow and innovation emergence. This research elucidates the non-monotonic impact of connectivity and reveals the dynamic interplay between knowledge integration and cognitive diversity within the innovation mechanism. Importantly, it aligns with broader discussions on how digital technologies and network structures can be harnessed to support information literacy, authentic knowledge exchange, and inclusive participation in science-core aspirations of an enlightened digital society.

Practically, it provides insights for research management and policy. Funding agencies and innovation managers should avoid overemphasizing connectivity or scale, and instead strive for structural balance between modularity and connectivity. Introducing evaluations of network topology in team building and project assessment could encourage collaborative structures that balance local exploration and global integration. Particularly in high knowledge density fields, controlling collaboration intensity helps preserve structural diversity and mitigate innovation risks. These strategies not only optimize innovation outcomes but also contribute to building a more equitable, critically engaged, and socially responsive digital research ecosystem.

About the authors

Yingqun Li is a Doctoral candidate at the School of Information Management, Nanjing University, specializing in scientific collaboration, disruptive innovation, and complex network analysis. Contact: yingqunli@smail.nju.edu.cn

Chuhui Zhang is a PhD candidate at the School of Information Management, Nanjing University, Nanjing, China. His research interests include data ethics and data governance. He can be contacted at chh_zhang@smail.nju.edu.cn.

Zhiwei Hu is a PhD student at the School of Information Management, Nanjing University, China. His research interests include the science of team science and knowledge management. He can be contacted at huzhiwei@smail.nju.edu.cn.

Lei Pei is a Professor at the School of Information Management, Nanjing University. He received their PhD from Wuhan University, focusing on information management and policy analysis. Contact: plei@nju.edu.cn

Reference

- Ahuja, G. (2000). Collaboration Networks, Structural Holes, and Innovation: a longitudinal study. *Administrative Science Quarterly*, 45(3), 425-455. <https://doi.org/10.2307/2667105>
- Arthur, W. B. (2007). The structure of invention. *Research Policy*, 36(2), 274-287. <https://doi.org/10.1016/j.respol.2006.11.005>

- Asplund, F., Björk, J., Magnusson, M., & Patrick, A. J. (2020). The genesis of public-private innovation ecosystems: Bias and challenges. *Technological Forecasting and Social Change*, 162, 120378. <https://doi.org/10.1016/j.techfore.2020.120378>
- Bornmann, L., Devarakonda, S., Tekles, A., & Chacko, G. (2020). Are disruption index indicators convergently valid? The comparison of several indicator variants with assessments by peers. *Quantitative Science Studies*, 1(3), 1242-1259. https://doi.org/10.1162/qss_a_00068
- Buldú, J. M., Busquets, J., Echegoyen, I., & Seirullo, F. (2019). Defining a historic football team: Using Network Science to analyze Guardiola's F.C. Barcelona. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-49969-2>
- Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, 110(2), 349-399. <https://doi.org/10.1086/421787>
- Cao, L., Chen, Z., & Evans, J. (2022). Destructive creation, creative destruction, and the paradox of innovation science. *Sociology Compass*, 16(11). <https://doi.org/10.1111/soc4.13043>
- Chan, H., & Akoglu, L. (2016). Optimizing network robustness by edge rewiring: a general framework. *Data Mining and Knowledge Discovery*, 30(5), 1395-1425. <https://doi.org/10.1007/s10618-015-0447-5>
- Cole, S. (1970). Professional standing and the reception of scientific discoveries. *American Journal of Sociology*, 76(2), 286-306. <https://doi.org/10.1086/224934>
- Crespo, J., Suire, R., & Vicente, J. (2013). Lock-in or lock-out? How structural properties of knowledge networks affect regional resilience. *Journal of Economic Geography*, 14(1), 199-219. <https://doi.org/10.1093/jeg/lbt006>
- Dallmann, A., Teutonico, D., Schaller, S., Burghaus, R., & Frechen, S. (2024). In - Depth analysis of the selection of PBPK modeling tools: Bibliometric and social network analysis of the Open Systems Pharmacology community. *The Journal of Clinical Pharmacology*, 64(9), 1055-1067. <https://doi.org/10.1002/jcph.2453>
- Di Lorenzo, P., & Barbarossa, S. (2014). Distributed estimation and control of algebraic connectivity over random graphs. *IEEE Transactions on Signal Processing*, 62(21), 5615-5628. <https://doi.org/10.1109/tsp.2014.2355778>
- Driskell, T., Funke, G., Tolston, M. T., Capiola, A., & Driskell, J. (2024). Supporting fluid teams: a research agenda. *Frontiers in Psychology*, 15, 1327885. <https://doi.org/10.3389/fpsyg.2024.1327885>
- Fares, J., Chung, K. S. K., & Abbasi, A. (2021). Stakeholder theory and management: Understanding longitudinal collaboration networks. *PLoS ONE*, 16(10), e0255658. <https://doi.org/10.1371/journal.pone.0255658>
- Foster, J. G., Rzhetsky, A., & Evans, J. A. (2015). Tradition and innovation in scientists' research strategies. *American Sociological Review*, 80(5), 875-908. <https://doi.org/10.1177/0003122415601618>
- Freitas, S., Yang, D., Kumar, S., Tong, H., & Chau, D. H. (2022). Graph Vulnerability and Robustness: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 1. <https://doi.org/10.1109/tkde.2022.3163672>
- Funk, R. J., & Owen-Smith, J. (2016). A dynamic network measure of technological change. *Management Science*, 63(3), 791-817. <https://doi.org/10.1287/mnsc.2015.2366>

- Gallo, J. L., & Plunket, A. (2020). Regional gatekeepers, inventor networks and inventive performance: Spatial and organizational channels. *Research Policy*, 49(5), 103981. <https://doi.org/10.1016/j.respol.2020.103981>
- Gamidullaeva, L., Tolstykh, T., Bystrov, A., Radaykin, A., & Shmeleva, N. (2021). Cross-Sectoral digital platform as a tool for innovation ecosystem development. *Sustainability*, 13(21), 11686. <https://doi.org/10.3390/su132111686>
- Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G., & Van Den Oord, A. (2008). Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy*, 37(10), 1717-1731. <https://doi.org/10.1016/j.respol.2008.08.010>
- Gonzalez-Brambila, C. N., Veloso, F. M., & Krackhardt, D. (2013). The impact of network embeddedness on research output. *Research Policy*, 42(9), 1555-1567. <https://doi.org/10.1016/j.respol.2013.07.008>
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360-1380. <https://doi.org/10.1086/225469>
- Guan, J., Zhang, J., & Yan, Y. (2015). The impact of multilevel networks on innovation. *Research Policy*, 44(3), 545-559. <https://doi.org/10.1016/j.respol.2014.12.007>
- Ibrahimi, G., Merioumi, W., & Bencheikroun, B. (2023). Fostering innovation through collective intelligence: a literature review. *Data & Metadata*, 2, 149. <https://doi.org/10.56294/dm2023149>
- Jacob, M., & Hellström, T. (2023). Affording excellence: What does excellence funding do for researchers? *Policy Studies*, 1-19. <https://doi.org/10.1080/01442872.2023.2267458>
- Janssen, M. J., & Frenken, K. (2019). Cross-specialisation policy: rationales and options for linking unrelated industries. *Cambridge Journal of Regions Economy and Society*, 12(2), 195-212. <https://doi.org/10.1093/cjres/rsz001>
- Kuhn, T. S. (1996). The structure of scientific revolutions. <https://doi.org/10.7208/chicago/9780226458106.001.0001>
- Lambiotte, R., & Panzarasa, P. (2009). Communities, knowledge creation, and information diffusion. *Journal of Informetrics*, 3(3), 180-190. <https://doi.org/10.1016/j.joi.2009.03.007>
- Lazega, E., & Burt, R. S. (1995). Structural holes: the social structure of competition. *Revue Française De Sociologie*, 36(4), 779. <https://doi.org/10.2307/3322456>
- Lind, J. T., & Mehlum, H. (2009). With or without U? the appropriate test for a U-Shaped relationship*. *Oxford Bulletin of Economics and Statistics*, 72(1), 109-118. <https://doi.org/10.1111/j.1468-0084.2009.00569.x>
- Lu, W., Yu, X., Li, Y., Cao, Y., Chen, Y., & Hua, F. (2024). Artificial Intelligence-Related Dental Research: Bibliometric and altmetric analysis. *International Dental Journal*. <https://doi.org/10.1016/j.identj.2024.08.004>
- Manso, G., & Pourbabaee, F. (2022). The impact of connectivity on the production and diffusion of knowledge. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2202.00729>
- Mayo, A. T. (2022). Syncing Up: A Process Model of Emergent Interdependence in Dynamic Teams. *Administrative Science Quarterly*, 67(3), 821-864. <https://doi.org/10.1177/00018392221096451>

- Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature*, 613(7942), 138-144. <https://doi.org/10.1038/s41586-022-05543-x>
- Perc, M. (2010). Growth and structure of Slovenia's scientific collaboration network. *Journal of Informetrics*, 4(4), 475-482. <https://doi.org/10.1016/j.joi.2010.04.003>
- Ren, R., & He, J. (2023). Network traits driving knowledge evolution in open collaboration systems. *PLoS ONE*, 18(11), e0291097. <https://doi.org/10.1371/journal.pone.0291097>
- Savona, R., Modena, A., Alessi, L., Alberini, C. M., Baussano, I., Guerra, R., Pecorelli, S., Rasi, G., Siviero, P. D., Dellaportas, P., Khozin, S., & Stein, R. M. (2025). Towards a framework for a new research ecosystem. *Humanities and Social Sciences Communications*, 12(1). <https://doi.org/10.1057/s41599-025-05281-1>
- Sciabolazza, V. L., Vacca, R., Okraku, T. K., & McCarty, C. (2017). Detecting and analyzing research communities in longitudinal scientific networks. *PLoS ONE*, 12(8), e0182516. <https://doi.org/10.1371/journal.pone.0182516>
- Staudt, J., Yu, H., Light, R. P., Marschke, G., Börner, K., & Weinberg, B. A. (2018). High-impact and transformative science (HITS) metrics: Definition, exemplification, and comparison. *PLoS ONE*, 13(7), e0200597. <https://doi.org/10.1371/journal.pone.0200597>
- Stephen, A. T., Zubcsek, P. P., & Goldenberg, J. (2015). Lower Connectivity is Better: The Effects of Network Structure on Redundancy of Ideas and Customer Innovativeness in Interdependent Ideation Tasks. *Journal of Marketing Research*, 53(2), 263-279. <https://doi.org/10.1509/jmr.13.0127>
- Stolarczyk, S., Bhardwaj, M., Bassler, K. E., Ji, W., MA, & Josić, K. (2017). Loss of information in feedforward social networks. *Journal of Complex Networks*, 6(3), 448-469. <https://doi.org/10.1093/comnet/cnx032>
- Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly*, 31(3), 439. <https://doi.org/10.2307/2392832>
- Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157), 468-472. <https://doi.org/10.1126/science.1240474>
- Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8), 1416-1436. <https://doi.org/10.1016/j.respol.2017.06.006>
- Wei, C., Li, J., & Shi, D. (2023). Quantifying revolutionary discoveries: Evidence from Nobel prize-winning papers. *Information Processing & Management*, 60(3), 103252. <https://doi.org/10.1016/j.ipm.2022.103252>
- Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744), 378-382. <https://doi.org/10.1038/s41586-019-0941-9>
- Xu, L., Zhou, Y., & Chen, L. (2024). Digital transformation and breakthrough innovation in Chinese manufacturing firms: based on Ability-Motivation-Opportunity (AMO) framework of human capital. *SAGE Open*, 14(3). <https://doi.org/10.1177/21582440241264348>
- Yang, W., & Wang, Y. (2024). Higher-order structures of local collaboration networks are associated with individual scientific productivity. *EPJ Data Science*, 13(1). <https://doi.org/10.1140/epjds/s13688-024-00453-6>

- Zanetti, M. S., Scholtes, I., Tessone, C. J., & Schweitzer, F. (2013). The rise and fall of a central contributor: Dynamics of social organization and performance in the GENTOO community. 2013 6th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE), 49-56., 49-56. <https://doi.org/10.1109/chase.2013.6614731>
- Zhang, X., Lei, S., Sun, J., & Kou, W. (2023). Robustness of Multi-Project knowledge collaboration network in open-source community. *Entropy*, 25(1), 108. <https://doi.org/10.3390/e25010108>
- Zheng, J., Gao, M., Lim, E., Lo, D., Jin, C., & Zhou, A. (2022). On measuring network robustness for weighted networks. *Knowledge and Information Systems*, 64(7), 1967-1996. <https://doi.org/10.1007/s10115-022-01670-z>

© [CC-BY-NC 4.0](#) The Author(s). For more information, see our [Open Access Policy](#).