Supplementary Information
Appendix for
Analysis of the Wikipedia Network of Mathematicians

Bingsheng Chen, Zhengyu Lin, Tim S. Evans

Centre for Complexity Science, and Theoretical Physics Group,
Imperial College London, SW7 2AZ, U.K.
23rd December 2018

A Appendix

Additional information is provided in this appendix.

A.1 Variance in Degree in Noise Model

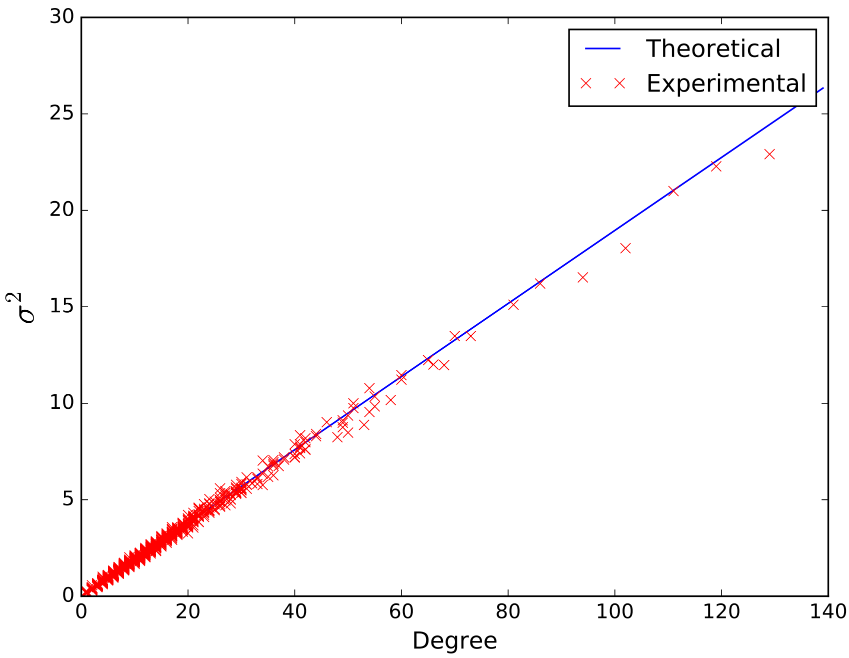

Figure A1: Each cross indicates the standard deviation in degree of one node after 1000 simulations. The theoretical result that $\sigma \approx 0.44\sqrt{k_{\text{orig}}}$ is compatible with this numerical result as the linear fit between variance and degree shows (an adjusted-r square value of 0.997).

A.2 Formal Definitions of Centrality Measures

The closeness c_v for a vertex v is defined to be (Bavelas, 1950; Hagberg et al., 2008; Newman, 2010)

$$c_v = \frac{(|\mathcal{C}_v| - 1)}{\sum_{u \in \mathcal{C}_v \setminus v} d(u,v)}\,, \tag{A1}$$

where $d(u, v)$ is the length of the shortest path between vertex u and some distinct vertex v which is in the same component, $\mathcal{C}_u$, as u. Note that we use a standard normalisation using N, the number of vertices, but this is irrelevant after our rescaling (see equation (4) in main text).

Our formal definition of betweenness b_v of a vertex v is (Freeman, 1977; Brandes, 2008; Hagberg et al., 2008; Newman, 2010)

$$b_v = \sum_{s,t \ \in \mathcal{C}_v} \frac{\sigma(s, t|v)}{\sigma(s, t)} \,. \tag{A2}$$

Here $\mathcal{C}_v$ is the set of vertices of the component containing vertex v, $\sigma(s, t)$ is the number of shortest paths available from vertex s to t, and $\sigma(s, t|v)$ is the number of shortest paths from s to t which pass through vertex v. This takes account of cases where there are two or more shortest paths between a pair of nodes s and t.

Eigenvalue centrality derived from the the adjacency matrix A, which we define such that A_{ij} is one (zero) if there is a link (no link) from vertex i to vertex j. The Eigenvector centrality for a vertex i is simply the i-th entry of the eigenvector of A associated with the largest eigenvalue (Newman, 2010; Hagberg et al., 2008). We perform our analysis on the largest component which then guarantees a unique value for each node.

PageRank is defined in terms of a transfer matrix, T where each entry, T_{ji} represents the probability of a random walker at vertex i moving to vertex j at the next time step. So we have that

$$T_{ji} = \frac{1}{s_i^{(\text{out})}} A_{ji} \,, \quad \text{where} \quad s_i^{(\text{out})} = \sum_j A_{ji} \,. \tag{A3}$$

An additional stochastic process also occurs. At each step, with probability α, the random walker follows a link chosen at random as given by the transfer matrix T but with probability $(1-\alpha)$ the current walk is deemed to end, or equivalently, we follow a new user or a new walk by starting at a randomly chosen vertex. The Markovian matrix G which describes this process is given by

$$G_{ij} = \alpha T_{ij} + (1-\alpha)\frac{1}{N} \tag{A4}$$

where N corresponds to total number of vertices and α is the damping factor, chosen to be $\alpha = 0.85$ in this work. The probability that a random walker is at vertex i in the long-time limit is proportional to the PageRank for that vertex and this is given by the i-th entry of the eigenvector associated with the largest Eigenvalue of the G. This makes PageRank similar to Eigenvector but different to the other centrality measures considered in that PageRank probes the whole network structure using walks of all types.

A.3 Additional Results

A.3.1 MacTutor Results

Rank	Degree	Closeness	Betweenness	PageRank	O(2nd) Clustering	Word Count
1	Newton	Newton	Euclid	Euclid	Hilbert	Euler
2	Hilbert	Hilbert	Newton	Newton	Newton	Galileo
3	Euclid	Riemann	Euler	Laplace	Euclid	Leibniz
4	Riemann	Euler	Riemann	Hilbert	Riemann	Newton
5	Euler	Euclid	Van der Waerden	Lagrange	Klein	Laplace
6	Klein	Cauchy	Weierstrass	Euler	Euler	Nash
7	Weierstrass	Gauss	Hilbert	Riemann	Weierstrass	Ptolemy
8	Poincare	Klein	Dieudonne	Gauss	Descartes	Tait
9	Gauss	Dirichlet	Cartan Henri	Klein	Leibniz	Kepler
10	Einstein	Laplace	Cauchy	Aristotle	Gauss	Aristotle
11	Cauchy	Lagrange	Hardy	Cauchy	Einstein	Lax Anneli
12	Lagrange	Poincare	Leibniz	Leibniz	Huygens	Copernicus
13	Laplace	Fourier	Dirichlet	Einstein	Lagrange	Euclid
14	Leibniz	Weierstrass	Weil	Jacobi	Aristotle	Polya
15	Hardy	Legendre	Fermat	Weierstrass	Poincare	Escher

Table A1: Centrality results for the top fifteen mathematicians in the directed network based on the hyperlinks between biographies on the MacTutor (O'Connor and Robertson, 2017) database, data from 2011. Copy of Table 4 from the appendix of Clarke (2011).

A.3.2 Wikipedia 2013 Results

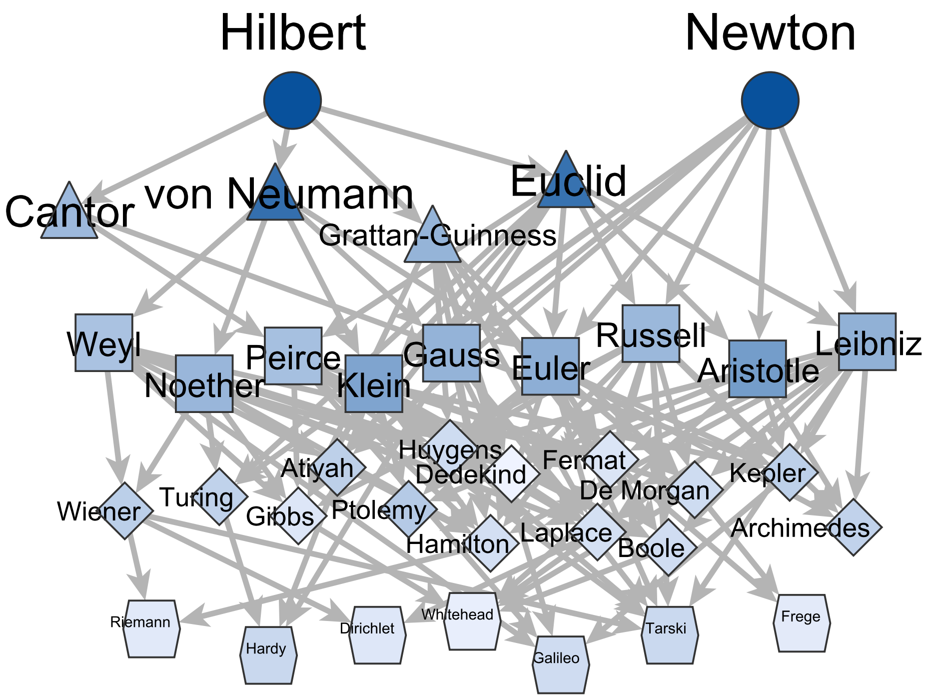

Figure A2: The top 36 mathematicians (2013 data) by a scaled average of the five ratings: Degree, Betweenness, Closeness, Eigenvector centrality and PageRank. Mathematician ‘A’ is connected placed higher than mathematician ‘B’ if each of their five ratings of ’A’ is higher than the corresponding rating for ‘B’. The arrow of the line points from the higher to the lower ranked mathematician but only those essential for the logical relationships are shown (a Hasse diagram of the corresponding poset, equivalently the transitively reduced form of the corresponding directed acyclic graph). The shape of a node, size of label and the vertical location reflects the ‘height’ of each node in the corresponding poset (see text for definition). The colour reflects the average scaled rating of each mathematician.

Name	Degree	Betweenness	Closeness	Eigenvector	PageRank	Average mark	Rank
David Hilbert	87.39	100	100	88.27	85.64	92.26	1
Isaac Newton	100	69.88	90.59	100	100	92.09	2
John von Neumann	74.79	92.67	97.07	58.73	81.55	80.96	3
Euclid	82.35	60.63	91.19	85.62	80.55	80.07	4
Aristotle	66.39	25.17	83.43	84.76	62.28	64.41	5
Felix Klein	67.23	33.82	89.5	51.9	65.78	61.65	6
Leonhard Euler	56.3	31.36	88.08	55.6	58.45	57.96	7
Gottfried Wilhelm Leibniz	51.26	24.55	88.27	70.74	50.29	57.02	8
Carl Friedrich Gauss	57.14	35.62	88.86	39.64	60.16	56.28	9
Ivor Grattan-Guinness	41.18	38.02	92.68	70.88	37.24	56	10
Emmy Noether	50.42	38.29	92.73	40.92	52.82	55.04	11
Bertrand Russell	46.22	25.77	90.22	66.09	44.69	54.6	12
Georg Cantor	42.86	30.49	92.51	65.06	39.24	54.03	13
Charles Sanders Peirce	40.34	26.26	90.63	61.68	39.08	51.6	14
Hermann Weyl	40.34	34.8	92.85	44.82	40.6	50.68	15
Ptolemy	53.78	9.82	75.1	50.47	49.56	47.75	16
Norbert Wiener	36.97	29	92.12	32.47	40.58	46.23	17
Michael Atiyah	43.7	36.68	85.45	10.21	52.71	45.75	18
Johannes Kepler	41.18	15.39	81.1	51.71	39.15	45.71	19
Alan Turing	36.13	29.35	89.64	27.83	41.93	44.98	20
Archimedes	44.54	9.22	77.76	53.15	39.76	44.88	21
G. H. Hardy	35.29	28.45	88.73	20.51	40.96	42.79	22
Alfred Tarski	36.13	21.1	84.42	32.26	38.13	42.41	23
Augustus De Morgan	31.09	12.96	85.58	45.63	32.1	41.47	24
Christiaan Huygens	35.29	10.63	82.11	43.32	35.23	41.32	25
Galileo Galilei	36.97	10.73	79.68	43.75	35	41.23	26
George Boole	31.09	11.72	84.77	47.2	29.34	40.82	27
William Rowan Hamilton	28.57	21.95	87.96	33.27	29.17	40.18	28
Pierre-Simon Laplace	33.61	13.39	84.3	34.12	34.67	40.02	29
Srinivasa Ramanujan	32.77	24.26	86.36	15.43	40.06	39.78	30
Nicolaus Copernicus	35.29	6.27	77.25	43.37	32.53	38.94	31
Pierre de Fermat	26.05	12.17	86.16	42.04	23.67	38.02	32
Josiah Willard Gibbs	26.05	13.41	89.22	33.77	25.85	37.66	33
Lejeune Dirichlet	31.93	8.37	83.54	32.29	30.11	37.25	34
Apollonius of Perga	31.93	6.81	78.37	40.02	28.78	37.18	35

Table A2: Centrality scores for top 35 mathematicians from 2013 data (without noise) given on a common scale with 100 for the largest value (see Equation (4) in the main text). Ordered in terms of their average score rating.

A.3.3 Wikipedia 2017 Results

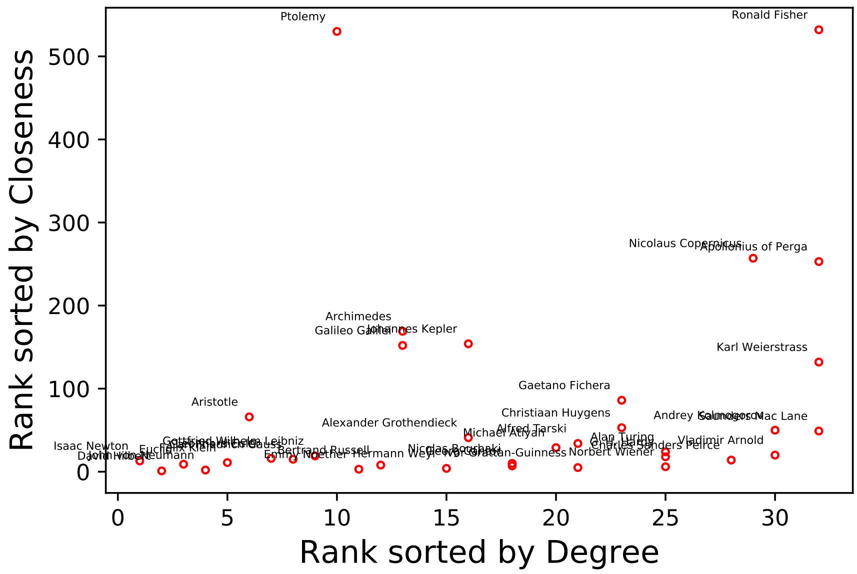

Figure A3: A comparison of the rank of mathematicians by degree and by closeness. the top 35 mathematicians by their average score in the 2017 Wikipedia data are shown under different centrality measures.

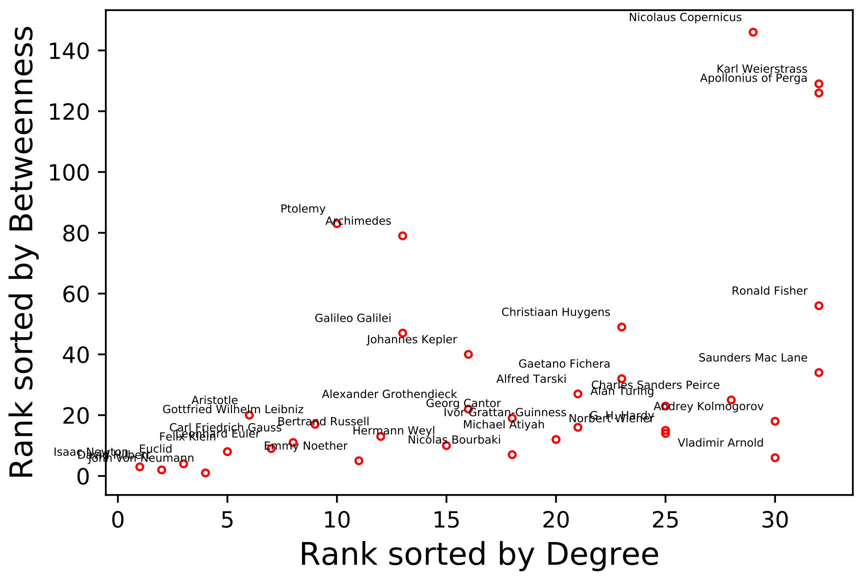

Figure A4: A comparison of the rank of mathematicians by degree and by betweenness. The top 35 mathematicians by their average score in the 2017 Wikipedia data are shown under different centrality measures.

A.3.4 Wikipedia 2018 Results

Quantity	2013	2018	% Increase	2018 After Rewiring
Mathematicians/Vertices	6050	8317	+37.4%	8317
Hyperlinks	15120	22669	+49.9%	22669
Undirected Edges	9701	14292	+47.3%	14291.42 ± 0.8
Average Degree	3.21	3.44	+7.2%	3.44 ± 0.0001
Vertices in largest component	4096	5829	+42.6%	5710.17 ± 17.7
Edges in largest component	9573	14115	+47.4%	14152.41 ± 10.0
Average Degree in largest component	4.71	4.84	+2.7%	4.96 ± 0.01
Network Diameter	13	15	+15.3%	14.12 ± 0.96
Average Path Length	5.07	5.14	+1.4%	4.90 ± 0.01
Clustering Coefficient	0.13	0.12	-7.7%	0.09 ± 0.002

Table A3: Network parameters for the 2013 and 2018 dataset, the percentage change between 2013 and 2017 data, and the mean values found for an ensemble of 1000 rewired 2018 data sets (with one standard deviation uncertainty quoted) as defined by our noise model of Section 2.3 with $p = 0.1$.

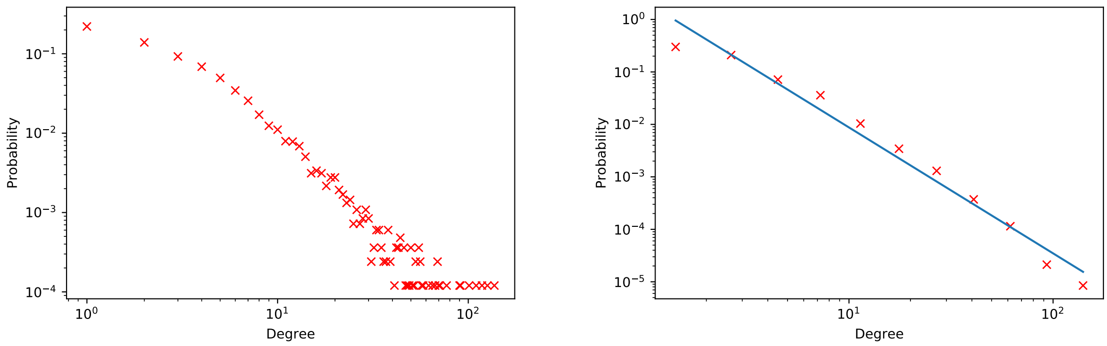

Figure A5: On the left, the degree distribution for the 2018 network of mathematicians. On the right the data is binned (using log binning with the ratio of consecutive bin edges set to be 1.5) and a best fit straight line to this data is shown added (slope is -2.70 ± 0.14).

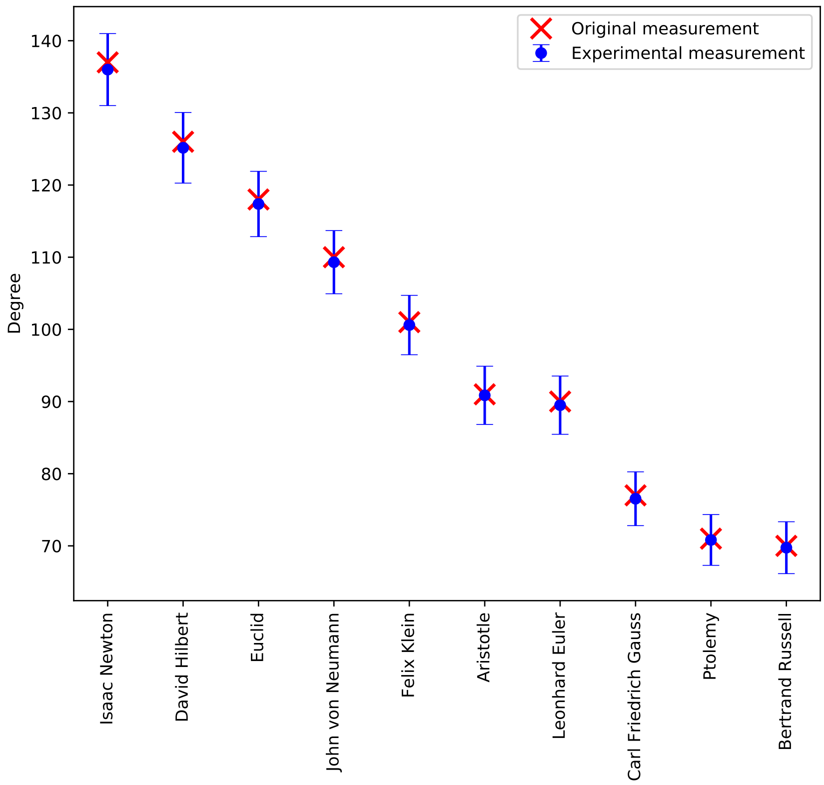

Figure A6: Degree distribution for the ten mathematicians whose Wikipedia biographies have the largest degree in the 2018 data (crosses). The circles give the mean degree for the same mathematicians as measured over 1000 simulations using our noise model of Section 2.3 where the error bars are specified by one standard deviation.

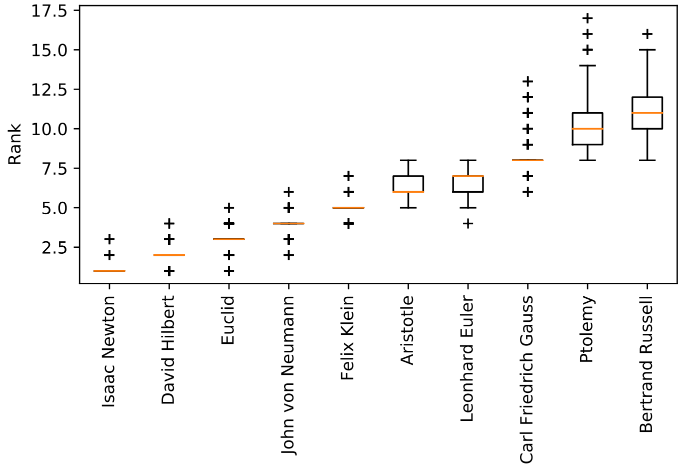

Figure A7: Whisper box plot for degree rank of mathematicians from 1000 simulations of our noise model from Section 2.3 applied to the 2017 data. The lower and upper edges of blue box show the 25 percentile (Q_1) and the 75 percentile (Q_3) of the rank of each mathematician, the red line in the middle of the box is the median. Given the small variation here, these lines often coincide. The black lines, at the end of the whiskers connected to the box, are defined to be at $Q_1 - 1.5(Q_3 - Q_1)$ and $Q_3 + 1.5(Q_3 - Q_1)$. The remaining black crosses beyond the whiskers indicate outliers beyond the whiskers.

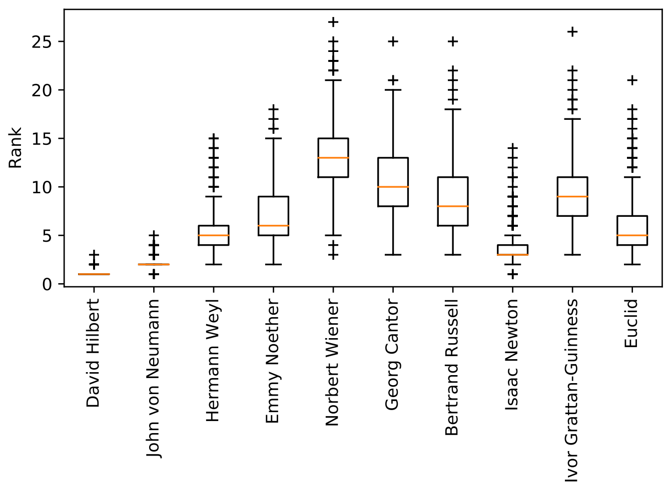

Figure A8: Whisper box plot for the rank of mathematicians by closeness, for the ten mathematicians with largest closeness. The closeness centrality is calculated for the largest component of the 2018 data and the uncertainties are estimated using 1000 simulations using the noise model of Section 2.3 with $p = 0.1$. The criteria used to place the boxes and other features of the plot are as in Fig. 3.

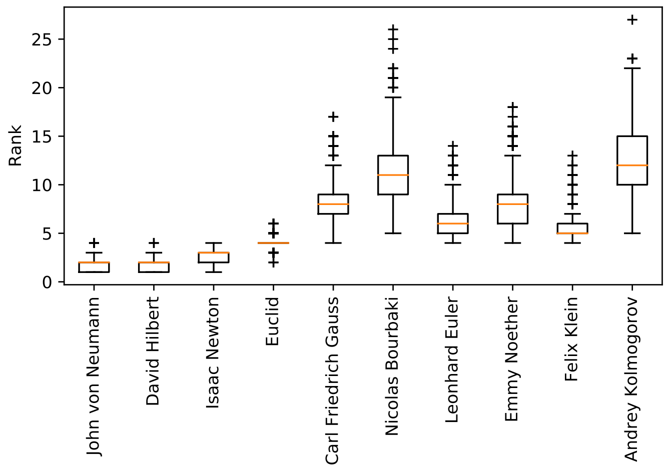

Figure A9: Whisper box plot for rank by betweenness of the ten mathematicians with highest betweenness. This is for the largest component of the 2018 data based on 1000 simulations using the noise model of Section 2.3. The criteria used to place the boxes and other features of the plot are as in Fig. 3.

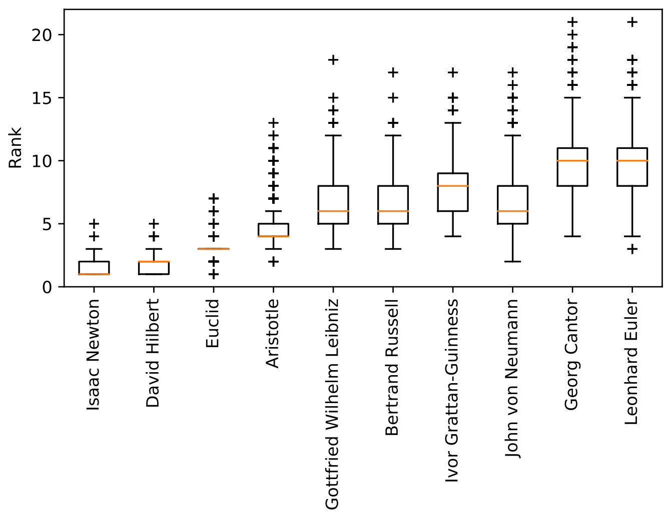

Figure A10: Whisper box plot for rank of mathematicians derived from their Eigenvalue centrality. This is for the largest component of the 2017 data based on 1000 simulations using the noise model of Section 2.3. The criteria used to place the boxes and other features of the plot are as in Fig. 3.

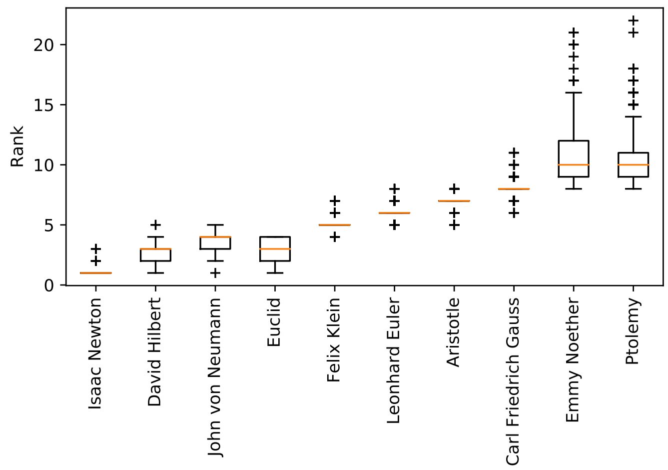

Figure A11: Whisper box plot for rank of mathematicians derived from their PageRank ratings. This is for the largest component of the 2018 data based on 1000 simulations using the noise model of Section 2.3. The criteria used to place the boxes and other features of the plot are as in Fig. 3 .

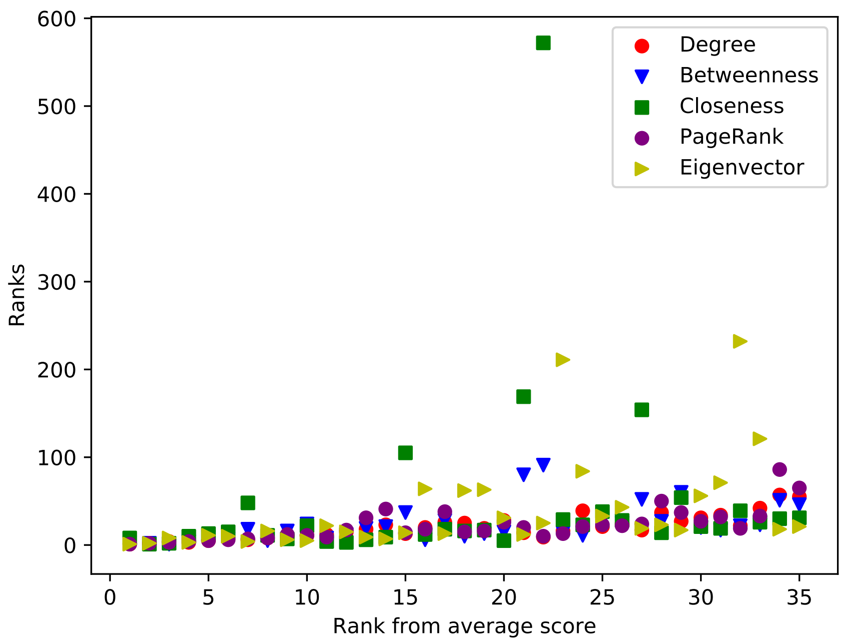

Figure A12: A comparison of the rank of mathematicians under different centrality measures. The horizontal axis is the rank of each mathematician by their average score; the top 35 are shown. Note that as the rank gets higher, there is a small but increasing variation in the ranks by different centrality measures for each mathematician.

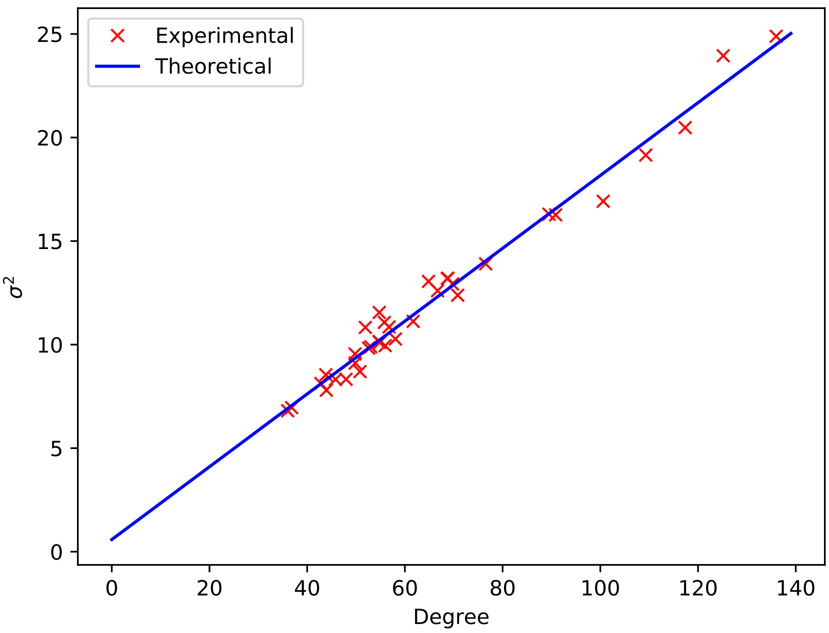

Figure A13: Each cross indicates the standard deviation in degree of one node after 1000 simulations for top 35 mathematician . The theoretical result that $\sigma \approx 0.42\sqrt{k_{\rm orig}}$ is compatible with this numerical result as the linear fit between variance and degree shows (an adjusted-r square value of 0.981).

Name	Degree	Betweenness	Closeness	Eigenvector	PageRank	Average mark	Rank
Isaac Newton	100	77.81	92.46	100	100	99.84	1
David Hilbert	92	91.76	100	87.29	94.15	91.99	2
Euclid	86.1	65.24	92.32	84.88	83.27	86.28	3
John von Neumann	80.3	100	97.19	87.28	62.25	80.35	4
Felix Klein	73.7	42	91.65	71.48	55.62	73.95	5
Aristotle	66.4	26.93	85.24	62.14	76.55	66.78	6
Leonhard Euler	65.7	43.89	91.35	68.52	60.06	65.78	7
Carl Friedrich Gauss	56.2	44.41	92.02	58.59	49.02	56.26	8
Ptolemy	51.82	8.82	75.29	48.67	39.15	52.05	9
Bertrand Russell	51.09	29.41	92.47	47.46	69.87	51.27	10
Emmy Noether	50.36	43.82	93.62	50.3	42.41	50.58	11
Gottfried Wilhelm Leibniz	50.36	24.9	89.43	48.36	70.02	50.49	12
Galileo Galilei	48.91	16.45	82.78	46.02	53.27	49.02	13
Archimedes	47.45	9.45	80.31	41.72	55.51	47.66	14
Hermann Weyl	45.26	38.98	94.93	44.7	49.74	45.35	15
Michael Atiyah	42.34	33.76	87.89	46.76	11.71	42.68	16
Johannes Kepler	41.61	13.35	80.79	40.03	43.29	41.73	17
G. H. Hardy	40.88	35.26	90.28	45.61	22.92	41.14	18
Georg Cantor	40.88	26.24	92.48	37.35	61.61	41.01	19
Alfred Tarski	40.15	21.44	86.52	40.63	32.19	40.3	20
Nicolas Bourbaki	40.15	43.93	91.82	42.87	22.58	40.25	21
Alexander Grothendieck	40.15	25.73	86.29	42.31	10.82	40.25	22
Alan Turing	38.69	23.23	87.94	40.73	29.29	38.98	23
Ivor Grattan-Guinness	38.69	25.75	92.44	34.33	66.9	38.63	24
Andrey Kolmogorov	37.96	40.01	90.5	45.82	23.5	38.15	25
Charles Sanders Peirce	37.23	21.76	89.98	34.98	54.02	37.35	26
Christiaan Huygens	36.5	11.74	85.01	35.37	47.6	36.63	27
Norbert Wiener	36.5	32.58	92.69	39.1	34.02	36.6	28
Richard Courant	35.04	25.76	89.47	38.19	23.77	35.25	29
Emil Artin	33.58	28.58	89.88	37.22	21.58	33.67	30
Vladimir Arnold	32.12	39.89	89.31	41.16	19.61	32.29	31
Bernhard Riemann	32.12	21.79	91.61	29.56	41.88	32.21	32
Srinivasa Ramanujan	31.39	25.22	88.38	36.97	16.17	31.46	33
Alfred North Whitehead	27.01	14.23	87.37	25.73	42.95	27.07	34
Pierre de Fermat	26.28	13.53	87.89	23.31	46.77	26.47	35

Table A4: Centrality scores for top 35 mathematicians from 2018 data (without noise) given on a common scale with 100 for the largest value according to equation (3). Ordered in terms of their average score rating.

Name	Degree	Betweenness	Closeness	Eigenvector	PageRank	Average	Rank
Isaac Newton	99.84 ± 0.87	88.5 ± 9.12	95.95 ± 1.37	99.97 ± 0.39	95.99 ± 6.39	96.05 ± 2.96	1
David Hilbert	91.99 ± 4.68	93.96 ± 7.34	100 ± 0.04	87.38 ± 4.95	94.81 ± 7.11	93.63 ± 3.93	2
John von Neumann	80.35 ± 4.33	94.28 ± 7.09	97.37 ± 1.12	84.1 ± 4.8	64.15 ± 6.57	84.05 ± 3.8	3
Euclid	86.28 ± 4.51	65.56 ± 7.88	94.41 ± 1.23	86.45 ± 4.84	79.48 ± 7.73	82.44 ± 4.22	4
Felix Klein	73.95 ± 4.04	51.32 ± 7.14	93.36 ± 1.25	72.17 ± 4.18	57.14 ± 6.16	69.59 ± 3.7	5
Leonhard Euler	65.78 ± 3.77	48 ± 6.62	92.91 ± 1.33	68.1 ± 4.07	57.94 ± 5.81	66.55 ± 3.53	6
Aristotle	66.78 ± 3.68	37.76 ± 5.86	90.44 ± 1.47	64.2 ± 3.86	70.39 ± 7.24	65.91 ± 3.53	7
Carl Friedrich Gauss	56.26 ± 3.44	43.92 ± 5.98	92.76 ± 1.27	57.83 ± 3.68	47.83 ± 5.11	59.72 ± 3.19	8
Bertrand Russell	51.27 ± 3.19	33.47 ± 5.39	93.49 ± 1.24	48.17 ± 3.17	64.87 ± 6.17	58.25 ± 3.11	9
Gottfried Wilhelm Leibniz	50.49 ± 3.17	30.17 ± 4.76	91.74 ± 1.21	48.91 ± 3.27	64.08 ± 5.87	57.08 ± 2.98	10
Emmy Noether	50.58 ± 3.2	44.23 ± 6.51	94.07 ± 1.22	49.8 ± 3.25	44.08 ± 5.22	56.55 ± 3.26	11
Hermann Weyl	45.35 ± 2.88	37.65 ± 5.68	94.65 ± 1.13	44.28 ± 2.96	49.35 ± 5.2	54.26 ± 2.93	12
Georg Cantor	41.01 ± 2.85	26.2 ± 4.37	92.98 ± 1.13	38.01 ± 2.81	57.65 ± 5.61	51.17 ± 2.76	13
Galileo Galilei	49.02 ± 3.09	23.16 ± 4.21	87.52 ± 1.48	47.44 ± 3.19	48.54 ± 5.44	51.14 ± 2.78	14
Ivor Grattan-Guinness	38.63 ± 2.63	25.96 ± 4.61	93.38 ± 1.13	34.85 ± 2.52	61.48 ± 5.47	50.86 ± 2.71	15
Archimedes	47.66 ± 3.13	17.55 ± 3.89	86.16 ± 1.61	43.65 ± 3.06	50.47 ± 5.92	49.1 ± 2.87	16
Ptolemy	52.05 ± 3.15	18.21 ± 3.93	83.47 ± 1.91	50.37 ± 3.22	36.73 ± 5.01	48.17 ± 2.77	17
Charles Sanders Peirce	37.35 ± 2.53	23.5 ± 4.12	91.54 ± 1.24	35.5 ± 2.52	49.84 ± 5.34	47.55 ± 2.57	18
Nicolas Bourbaki	40.25 ± 2.9	38.06 ± 5.35	91.63 ± 1.24	41.52 ± 3.05	25.35 ± 3.9	47.36 ± 2.74	19
G. H. Hardy	41.14 ± 2.71	34.34 ± 4.84	90.34 ± 1.35	44.09 ± 2.98	24.5 ± 3.64	46.88 ± 2.51	20
Andrey Kolmogorov	38.15 ± 2.75	35.8 ± 4.91	90.4 ± 1.34	43.34 ± 3.15	24.71 ± 3.44	46.48 ± 2.52	21
Norbert Wiener	36.6 ± 2.55	30.45 ± 4.5	92.28 ± 1.2	38.06 ± 2.72	34.31 ± 4.1	46.34 ± 2.49	22
Johannes Kepler	41.73 ± 2.84	18.74 ± 3.81	86.14 ± 1.66	40.92 ± 2.92	39.88 ± 4.88	45.48 ± 2.64	23
Michael Atiyah	42.68 ± 2.78	32.69 ± 4.96	88.73 ± 1.46	45.29 ± 3.01	15.73 ± 3.14	45.02 ± 2.53	24
Alfred Tarski	40.3 ± 2.69	23.26 ± 3.78	87.75 ± 1.35	40.38 ± 2.82	31.56 ± 4.28	44.65 ± 2.46	25
Alan Turing	38.98 ± 2.73	23.44 ± 4.09	88.53 ± 1.35	40.01 ± 2.88	29.47 ± 4	44.09 ± 2.44	26
Christiaan Huygens	36.63 ± 2.55	15.06 ± 3.18	87.39 ± 1.39	35.95 ± 2.59	43.16 ± 4.59	43.64 ± 2.29	27
Bernhard Riemann	32.21 ± 2.38	20.69 ± 3.64	91.88 ± 1.13	29.94 ± 2.32	40.48 ± 4.34	43.04 ± 2.27	28
Alexander Grothendieck	40.25 ± 2.78	29.47 ± 4.98	88.17 ± 1.57	41.1 ± 2.92	15.16 ± 3.07	42.83 ± 2.53	29
Richard Courant	35.25 ± 2.47	25.87 ± 4.06	89.44 ± 1.25	37.06 ± 2.67	25.38 ± 3.48	42.6 ± 2.27	30
Vladimir Arnold	32.29 ± 2.32	31.19 ± 4.48	89.4 ± 1.49	38.7 ± 2.77	20.74 ± 3.33	42.46 ± 2.31	31
Emil Artin	33.67 ± 2.47	26.6 ± 4.12	89.74 ± 1.31	35.9 ± 2.67	23.28 ± 3.55	41.84 ± 2.35	32
Srinivasa Ramanujan	31.46 ± 2.39	23.45 ± 3.43	88.12 ± 1.34	35.41 ± 2.69	17.56 ± 3.09	39.2 ± 2.14	33
Pierre de Fermat	26.47 ± 2.09	13.18 ± 2.89	89.29 ± 1.3	24.16 ± 2.02	42.52 ± 4.28	39.12 ± 2.07	34
Alfred North Whitehead	27.07 ± 2.18	14.93 ± 2.95	88.61 ± 1.22	25.92 ± 2.19	38.91 ± 4.29	39.09 ± 2.07	35

Table A5: Centrality scores for top 35 mathematicians derived from the the noise model described of Section 2.3 applied to the 2018 data with $p = 0.1$ for 1000 simulations. The mean value and one standard deviation is quoted for each centrality measure for each mathematician. As the scores for each run are always rescaled so that the largest value is 100, explaining why the value quoted for any one centrality measure is always less than 100. The column marked average gives the average over the five named centrality measures with associated standard deviation. Mathematicians are ordered in terms of this average and the ranks given are in terms of this average over centrality values.

Top 35	Degree	PageRank	Eigenvector	Betweenness	Closeness	Average
Degree	1.00	0.98	0.75	0.78	0.36	0.96
PageRank	*0.92*	1.00	0.64	0.67	0.32	0.95
Eigenvector	*0.66*	*0.44*	1.00	0.20	0.35	0.79
Betweenness	*0.50*	*0.85*	*0.44*	1.00	0.74	0.87
Closeness	*0.31*	*0.41*	*0.30*	*0.78*	1.00	0.56
Average	*0.87*	*0.81*	*0.75*	*0.69*	*0.62*	1.00

Table A6: The correlation values for the 35 top mathematicians as defined by the average of centrality scores in the 2018 data. The upper right triangle contains the Pearson correlation values (in blue) while the lower left triangle contains the Spearman correlation values (in red italics). Note that for both cases the degree and PageRank are particularly well correlated as are Betweenness and Closeness.

largest component	Degree	PageRank	Eigenvector	Betweenness	Closeness	Average
Degree	1.00	0.98	0.82	0.86	0.57	0.95
PageRank	*0.95*	1.00	0.74	0.92	0.51	0.91
Eigenvector	*0.63*	*0.42*	1.00	0.49	0.94	0.86
Betweenness	*0.88*	*0.87*	*0.70*	1.00	0.39	0.82
Closeness	*0.70*	*0.52*	*0.56*	*0.59*	1.00	0.78
Average	*0.85*	*0.69*	*0.90*	*0.69*	*0.96*	1.00

Table A7: The correlation values for mathematicians in largest component as defined by the average of centrality scores in the 2018 data. The upper right triangle contains the Pearson correlation values (in blue) while the lower left triangle contains the Spearman correlation values (in red italics). Note that for both cases the degree and PageRank are particularly well correlated as are Betweenness and Closeness.

References

Bavelas, A., 1950. Communication patterns in task-oriented groups. The Journal of the Acoustical Society of America 22, 725–730.

Brandes, U., 2008. On variants of shortest-path betweenness centrality and their generic computation. Social Networks 30, 136–145. doi:10.1016/j.socnet.2007.11.001.

Clarke, C., 2011. The Network of Mathematical Innovation. Master's thesis. Imperial College London.

Freeman, L.C., 1977. A set of measures of centrality based on betweenness. Sociometry , 35–41.

Hagberg, A., Swart, P., S Chult, D., 2008. Exploring network structure, dynamics, and function using NetworkX. Technical Report. Los Alamos National Laboratory (LANL).

Newman, M., 2010. Networks: an introduction. Oxford University Press.

O'Connor, J.J., Robertson, E.F., 2017. Mactutor history of mathematics archive.