Applied Math, Statistics, and Data Science

GRAPH QUILTING

Giuseppe Vinci, Assistant Professor, University of Notre Dame

Genevera Allen, Rice University

Gautam Dasarathy, Arizona State University

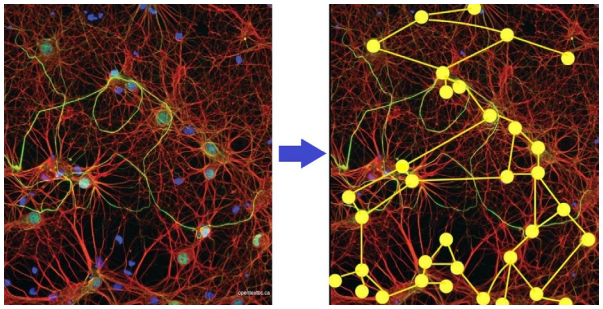
Estimating a conditional dependence graph is a seemingly impossible task when several pairs of variables are never jointly observed. Recovering the edges of the graph in such settings requires one to infer conditional dependencies between variables with no evidence of their marginal dependence. This unexplored statistical problem arises in several situations, such as in neuroimaging where, because of technology limitations, it is impossible to jointly record the activities of all neurons simultaneously. We call this statistical challenge the "Graph Quilting problem." In the Gaussian graphical model, the unavailability of parts of the covariance matrix translates into the nonidentifiability of the precision matrix, which specifies the graph. However, we demonstrate that, under mild conditions, it is possible to correctly identify not only the edges connecting the observed pairs of nodes, but also a minimal superset of those connecting the variables that are never observed jointly. To accomplish the latter task, we devise a novel technique that we call the "Recursive-Complement" algorithm. We propose a graph estimator based on partially observed sample covariances and l1-regularization, and establish its rates of convergence in high-dimensions. We illustrate the methodology using synthetic data, as well as data obtained from in vivo calcium imaging of ten thousand neurons in mouse visual cortex.

GRAPH QUILTING

Giuseppe Vinci[1], Gautam Dasarathy[2], Genevera I. Allen[3]
[1]Department of Applied and Computational Mathematics and Statistics, University of Notre Dame
[2]Department of Electrical and Computer Engineering, Arizona State University
[3]Department of Electrical and Computer Engineering, Rice University

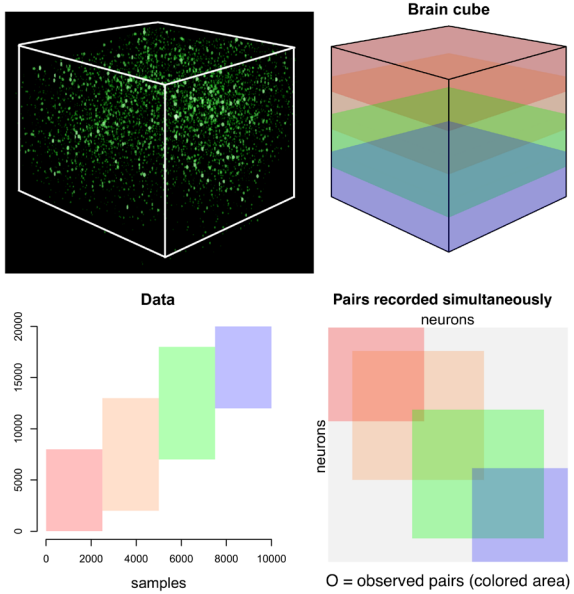
Contact Information:
Giuseppe Vinci
Department of Applied and Computational Mathematics and Statistics
Crowley Hall, University of Notre Dame
Notre Dame, IN 46556, USA
Email: gvinci@nd.edu

Functional Connectivity Graph



no edge ⇔ conditional independence

Large-Scale Nonsimultaneous Calcium Imaging Recordings

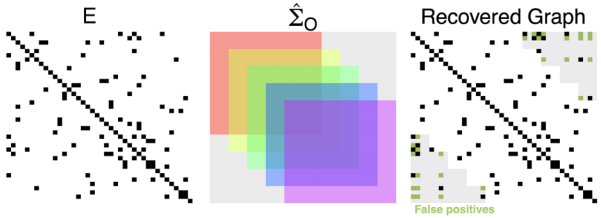


⇒ graph recovery impossible with existing methods

Graph Quilting

⇒ Recovering the full graph with a minimal number of false positives in O^c with high-probability.

Gaussian Graphical Model: $X \sim N(\mu, \Sigma)$, where $\Theta = \Sigma^{-1}$ and $\Theta_{ij} = 0 \Leftrightarrow X_i, X_j$ conditional independent.



• $E = \{(i,j) : \Theta_{ij} \neq 0\}$ • $\hat{\Sigma}_O$ = available emp. covariances

Definition (MAD$_{GQlasso}$)

$$\widehat{\Theta} = \underset{\Theta \succ 0, \Theta_{O^c}=0}{\operatorname{argmax}} \log\det\Theta - \sum_{(i,j)\in O} \Theta_{ij}\widehat{\Sigma}_{ij} - \|\Lambda \odot \Theta\|_{1,\text{off}}$$

Theorem 1 If $\|\Theta_{O^c}\|_\infty$ is sufficiently small, then $\exists \tau > 0$ s.t. for sufficiently large sample size, with high-probability

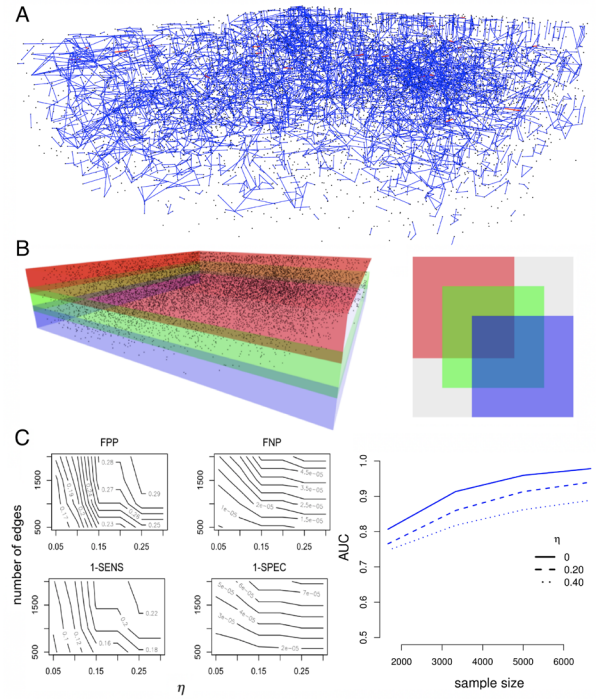
$$\widehat{E}^{(\tau)} := \{(i,j) : |\widehat{\Theta}_{ij}| > \tau\} = E_O$$

Recursive-Complement (RECO) Algorithm: exploits the pattern of the detected distortions between $\widehat{\Theta}$ and Θ (e.g. dropped edges are false positives) to identify all potential graphical structures in O^c automatically.

Theorem 2 Under appropriate conditions, with high probability, the RECO algorithm yields

$$\widehat{\mathscr{S}} = \mathscr{S} \supseteq E_{O^c}$$

where $\mathscr{S}$ is a minimal superset of E_{O^c}.



A: Graph of about 10,000 simultaneously recorded neurons in mouse visual cortex (data from [1]).

B: Nonsimultaneous recordings. **C**: Graph Quilting performance ($\eta = |O^c|/p^2$).

Acknowledgements

Funding: **NSF NeuroNex-1707400**, NSF DMS-1554821, and Rice Academy Postdoctoral Fellows. We thank Andreas Tolias, Xaq Pitkow, and Jacob Reimer at Baylor College of Medicine, and Krešimir Josić at the University of Houston, for useful discussions on calcium imaging recording procedures.

References

[1] Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. Science, 364(6437), 255-255.

[2] Vinci, G., Dasarathy G., Allen G. I. (2019). Graph Quilting: graphical model selection from partially observed covariances. Submit. to the Annals of Statistics.

[3] Vinci, G., Ventura, V., Smith, M. A., & Kass, R. E. (2018). Adjusted regularization of cortical covariance. Journal of computational neuroscience, 45(2), 83-101.

[4] Vinci, G., Ventura, V., Smith, M. A., & Kass, R. E. (2018). Adjusted regularization in latent graphical models: Application to multiple-neuron spike count data. The Annals of Applied Statistics, 12(2), 1068–1095.

[5] Vinci, G., Ventura, V., Smith, M. A., & Kass, R. E. (2016). Separating spike count correlation from firing rate correlation. Neural computation, 28(5), 849–881.

PANDEMIC PULSE

Jennifer Schnur, Graduate Student, University of Notre Dame

Steven Krieg, University of Notre Dame

Nitesh Chawla, University of Notre Dame

The COVID-19 outbreak has altered the routine of U.S. society at an unparalleled scale. The Lucy Family Institute for Data and Society at Notre Dame has been collecting a great variety of data, including raw case numbers, population mobility trends, government interventions, online news, Google searches, mask usage surveys, and more, to understand the broad effects of both the pandemic on society, and social behaviors on the trajectory of the pandemic. Collectively, this data represents the pulse and vital statistics of our society. Analyzing it collectively rather than separately has enabled us to glean new insights, such as the fact that right-leaning sources are overrepresented in the corpus of online COVID-19 news, and that scientific sources are underrepresented. Other key findings include that mask usage and mobility statistics across states are strongly correlated with partisan voting in the 2020 presidential election. Our work suggests that studying these potent data sets together can produce new insights that would have remained hidden if the data remained isolated. To enable further exploration and discovery, we have published an interactive dashboard at http://cnds.nd.edu/pandemicpulse.

Pandemic Pulse

Modeling Society's Symptoms and Interventions during the COVID-19 Pandemic

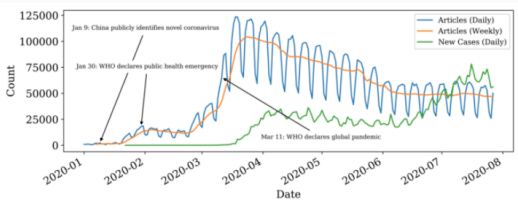
Lucy Family Institute for Data & Society

University of Notre Dame | Research

Interactive dashboard available at https://cnds.nd.edu/pandemicpulse/

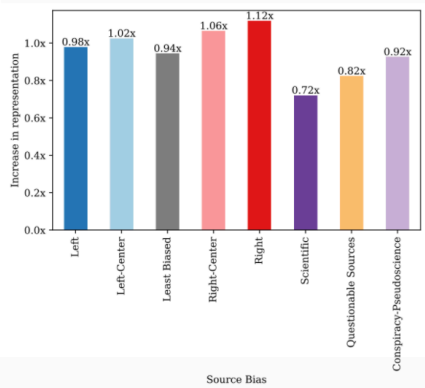
Pandemic Pulse represents an effort to understand the broad effects of COVID-19 on U.S. society by aggregating and analyzing data from many sources, including raw case numbers, population mobility trends, government interventions, online news, Google searches, mask usage surveys, and more.

COVID-19 in the News



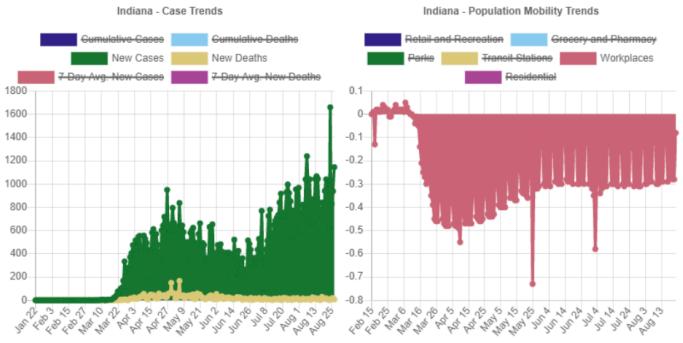
The most COVID-19 news was published at the beginning of the pandemic, while the quantity of news does not mirror case trends.

Additionally, right-leaning sources are overrepresented in the corpus of COVID-19 news, while scientific sources are underrepresented.



Public Dashboard

To enable further exploration, we developed an interactive dashboard that visualizes many of these trends, and allows for simple analysis. For example, the figures below show the negative correlation between daily COVID-19 cases and changes in travel to workplaces from Jan 22 through Aug 13, 2020, in the state of Indiana.



Indiana - Case Trend - New Cases
vs
Indiana - Mobility Trend - Workplaces

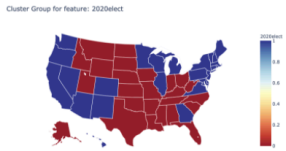
Correlation Coefficient: -0.3622
P-Value: 2.82e-7

Krieg, S. J., Schnur, J. J., Marshall, J. D., Schoenbauer, M. M., & Chawla, N. V. (2020). Pandemic Pulse: Unraveling and Modeling Social Signals during the COVID-19 Pandemic. Digital Government: Research and Practice, 2(2), 1-9.

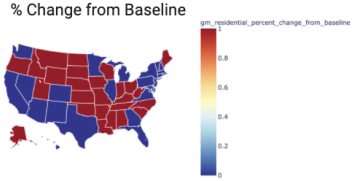
Clustering Preventative Behaviors

Social distancing and mask usage trends align well with the 2020 electoral map, implying that pandemic behavior is a partisan issue.

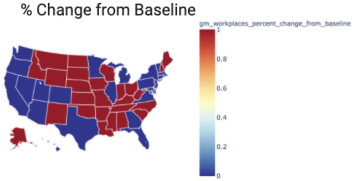
2020 Election Results



Google Maps Residential Mobility % Change from Baseline



Google Maps Workplace Mobility % Change from Baseline



Mask Usage Frequency = Always Proportion of Population

KIDNEY TRANSPLANT IN AFRICAN AMERICANS AND ASSOCIATED RISK OF INFERIOR ALLOGRAFT SURVIVAL

Peter Umukoro, Renal Fellow,
Indiana University School of Medicine

Oluwafisayo Adebiyi,
Indiana University School of Medicine

Asif Sharfuddin,
Indiana University School of Medicine

Tim Taber,
Indiana University School of Medicine

Jeanne Chen,
Indiana University School of Medicine

Kathleen Lane,
Indiana University School of Medicine

Xiaochun Li,
Indiana University School of Medicine

Williams Goggins,
Indiana University School of Medicine

Muhammad Yaqub,
Indiana University School of Medicine

Objective: To compare whether African Americans (AA) have worse graft/survival outcomes compared to other races.

Hypothesis: No difference between AA and Non-AA in graft survival and function, acute rejection rates, and patient survival one year after a kidney transplant.

Methods: Retrospective cohort, single center. 2097 study participants, mean (SD) age 51.0 (14.4), 456 (21.8%) AA who received kidney transplant between 2005 -2016 were followed up for one year; comparing patient/graft survival and function at 1 year; 1-year rates of biopsy-proven acute rejection in AA versus non-AA. Used Cox proportional hazards models to estimate hazard ratios.

Results: No differences in patient survival and surviving graft function (estimated by GFR); death censored graft survival (HR of 2.56, $p=0.04$) and one-year acute rejection (HR of 1.73, $p<0.01$) were worse in AA compared to Non-AA after adjusting for age, sex, and donor type.

Limitation: Potential for residual confounding by socioeconomic status.

Conclusion: In this sizeable cohort with ethnic diversity, African Americans are at higher risk for worse graft outcomes even in the early post-transplant period compared to other races. Further studies evaluating factors including socio-economic determinants of health are needed to mitigate against poorer outcomes.

Kidney Transplant in African Americans and Associated Risk of Inferior Allograft Survival

Peter E. Umukoro[1], Oluwafisayo Adebiyi[1], Asif Sharfuddin[1], Tim Taber[1], Jeanne Chen[1], Kathleen A. Lane[2], Xiaochun Li[2], Williams Goggins[1]. Muhammad S Yaqub[1]

1. Indiana University Health Transplant. Indianapolis, Indiana 2. Indiana University School of Medicine, Department of Biostatistics and Health Data, Indianapolis Indiana

BACKGROUND

- Disparities exist within the health system and structure that account for a significant proportion of differences in transplant outcomes among different races.
- African Americans have been shown to have worse graft and survival outcomes.
- There are increasing number of studies evaluating social and structural factors that persist in African Americans who complete the rigorous process of qualifying for a kidney transplant having met the clinical, social, and psychological requirements.
- It is unclear how soon after a successful kidney transplant, disparities in health outcomes (graft and patient survival) occur in African Americans compared to other races.

HYPOTHESIS

No difference in graft survival and function, acute rejection rates, and patient survival one year after a kidney transplant compared to other races.

SETTINGS AND PARTICIPANTS

Conducted in Indiana University, 2096 study participants, mean (SD) age 51.0 (14.4), 456 (21.8%) African Americans who received a kidney transplant between between 2005 -2016 and were followed up for one year after their kidney transplant.

METHODS

Study Design: Retrospective cohort, single-center

Exposure: African American versus other races

Outcome(s): Our primary outcome was patient survival at 1 year. Our secondary outcomes were graft survival, death censored graft survival,1-year rates of biopsy-proven acute rejection, graft function at 3 months, 6 months and 1 year.

Analytical Approach: We used a cox proportional hazards model to estimate hazard ratios.

RESULTS

Patient and Graft Survival Outcomes in AA versus Non-AA

*Outcome/Model	HR	95% CI	p-value
Patient Survival	1.17	0.54 – 2.53	0.69
Graft Failure or Death	1.62	0.91 – 2.88	0.10
Death Censored Graft Failure	2.56	1.03 – 6.35	0.04
Acute rejection	1.73	1.35 – 2.21	<0.01

*All models were controlled for age at transplant, sex, and donor type (living donor versus deceased donor)

Graft Function at 3, 6 and 12 months in AA and Non-AA

	AA (n=456)		Non-AA (n=1641)		
eGFR	N	Mean±SD/ Count (%)	N	Mean±SD/ Count (%)	*P-value
3 months	416	60.2 ± 18.8	1484	55.9 ± 17.1	<0.01
6 months	394	59.7 ± 19.7	1370	55.6 ± 17.0	<0.01
1 year	356	60.4 ± 21.8	1263	56.3 ± 18.1	<0.01

*Statistically significant but not clinically significant.

RESULTS

- There were no difference in patient survival at 1 year with Hazard Ratio (HR) of 1.17 (95% CI: 0.54 -2.53, p=0.69)
- Death censored graft survival with HR of 2.56 (95% CI: 1.03 – 6.35, p=0.04) between AA and other races after adjusting for age, sex, and donor type.
- One-year acute rejection (AR) occurred in 22.1% in African Americans and was worse compared to other races (13.3%) with HR of 1.73 (95% CI: 1.35 – 2.21, p<0.01) after adjusting for age, sex and donor type.
- Surviving raft function at 3 months, 6 months and 1 year were statistically significant but not clinically different between the races (all p<0.01).

CONCLUSIONS

- In this sizeable cohort with ethnic diversity, African Americans are at higher risk for worse graft outcomes even in the early post-transplant period compared to other races.
- Further studies evaluating factors including socio-economic determinants of health are needed to mitigate against poorer outcomes.

REFERENCES

- Naylor, K.L., et al., Socioeconomic Status and Kidney Transplant Outcomes in a Universal Healthcare System: A Population-based Cohort Study. Transplantation, 2019. 103(5): p. 1024- 1035.
- Taber, D.J., et al., Twenty years of evolving trends in racial disparities for adult kidney transplant recipients. Kidney Int, 2016. 90(4): p. 878-87

CONTACT

peter_umukoro@mail.harvard.edu or pumukoro@iu.edu