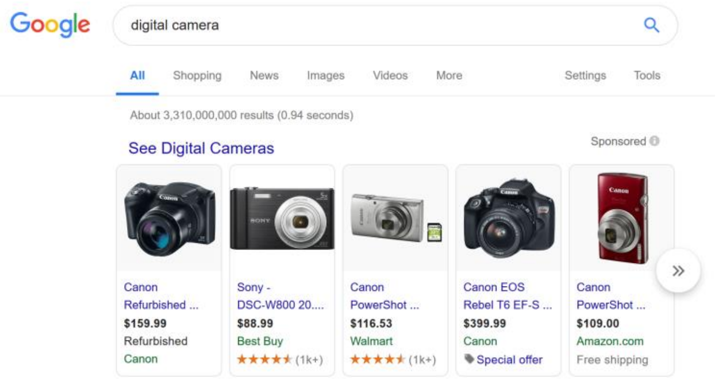
FUNDAMENTAL LIMITS OF TESTING THE INDEPENDENCE OF IRRELEVANT ALTERNATIVES IN DISCRETE CHOICE

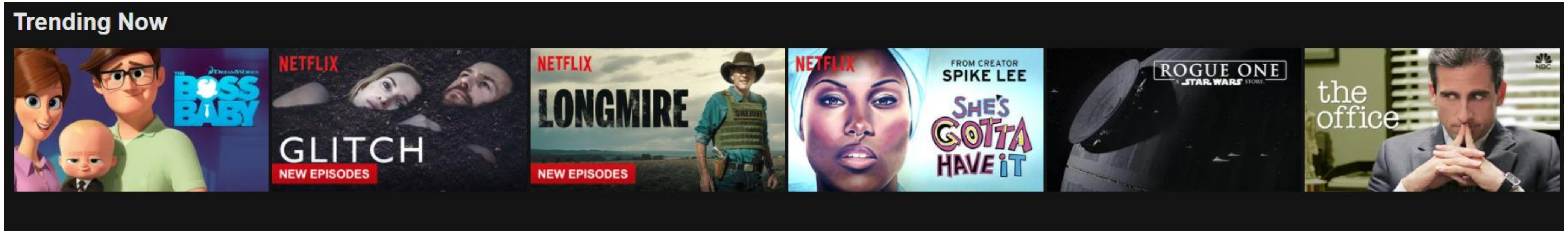
Arjun Seshadri, Stanford University
Johan Ugander, Stanford University

EC 2019

DISCRETE CHOICE

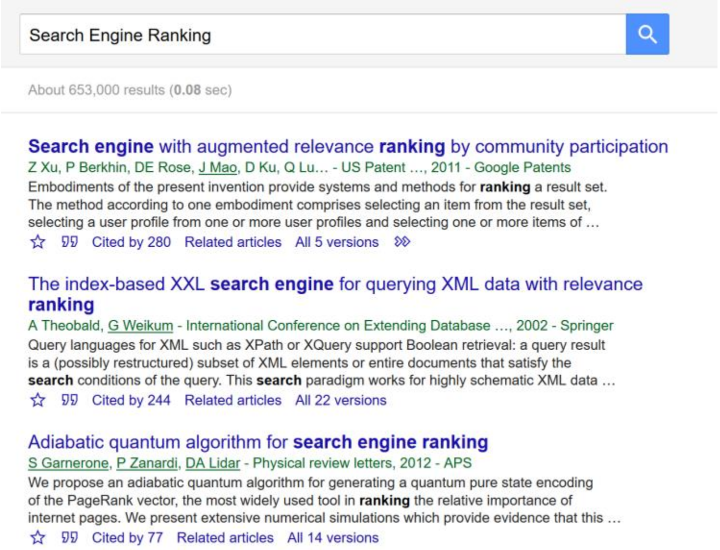
- Data of the form (x, C) where "alternative x is chosen from the set C" and C is a subset of $\mathcal{X}$, the universe of n alternatives
- Discrete choice settings are ubiquitous





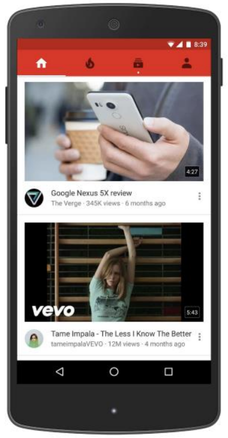
ESSENTIAL IN MACHINE LEARNING

Inverse reinforcement learning



Search Engine Ranking

Virtual Assistants



Recommender Systems

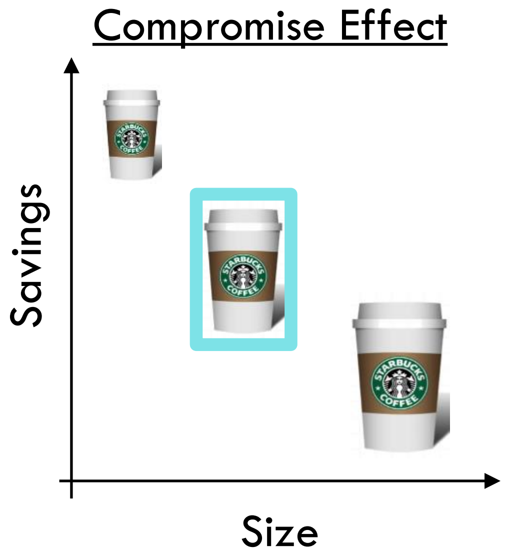
INDEPENDENCE OF IRRELEVANT ALTERNATIVES (IIA)

- Fully determines the workhorse Multinomial Logit (MNL) Model

$$\left.\begin{array}{l} x, y \in A \\ x, y \in B \end{array}\right\} \Rightarrow \frac{\Pr(x \text{ from } A)}{\Pr(y \text{ from } A)} = \frac{\Pr(x \text{ from } B)}{\Pr(y \text{ from } B)} \iff P_{x,C} = \frac{\gamma_x}{\sum_{z \in C} \gamma_z}, \gamma \in \Delta_n$$

<u>IIA</u> <u>MNL</u> or "Softmax"

- Cannot account for behavioral economics "anomalies" all over the place
  - Compromise Effect
  - Search Engine Ads (Ieong-Mishra-Sheffet '12, Yin et al. '14)
  - Google Web Browsing Choices (Benson-Kumar-Tomkins '16)
- Explosion of new online choice domains



Can we test IIA?

BUT FIRST, WHY HYPOTHESIS TESTING?

- Hypothesis tests provide an objective measure of inference (Johari et al, 2015)
  - Interpretable: rejecting at level α gives precise false positive control
  - Transparent: can apply personal tolerance of error to a p-value
- Starting point for understanding theoretical behavior of statistical problems
  - e.g. High dimensional models
  - A clean, precise framework
- Recent results in discrete distributions
  - Property tests require far fewer samples than estimates (Acharya et al, 2015; Valiant and Valiant, 2017)

TESTING IIA: AN AGE OLD PROBLEM

1959

"an idea of the sample sizes needed for the two alternative case" – Luce

1977

"a series of diagnostic tests for the property are developed" – McFadden et al.

1984

"two sets of computationally convenient specification tests" – Hausman and McFadden

1985

"a modified [test]… to eliminate the asymptotic bias… yet avoids computational problems" – Small and Hsiao

1996

"…majority of tests based upon partitioning the choice set appear to have very poor size properties…" – Fry and Harris

2007

"tests of the IIA assumption… based on estimation of a choice set are unsatisfactory for applied work" – Cheng and Long

2016

"a battery of statistical tests that quantify IIA violations" – Benson et al.

FOLKLORE

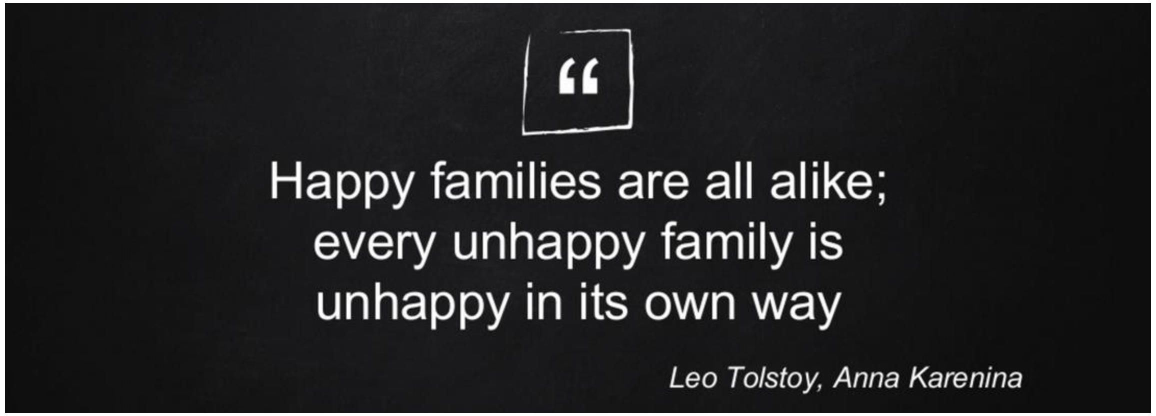
"to have anything like a sensitive test... it is clear that rather large sample sizes are required from each subset" — (Luce, 1959)

"We do not believe tests of IIA are useful...can almost always obtain some tests that...reject the null when using the same model with the same data" — (Long and Freese, 2014)

"It is likely that part of the problem arises through the poor size properties of the asymptotic procedures." — (Fry and Harris, 1998)

"I can't recommend these tests to anyone" — (Paul Allison, 2012)

ANNA KARENINA PRINCIPLE



A few ways to be “rational”, many ways to be “irrational”

The MNL model has a low dimensional representation

A model of arbitrary choice can behave arbitrarily on any single subset of alternatives

A combinatorial number of ways to deviate from IIA

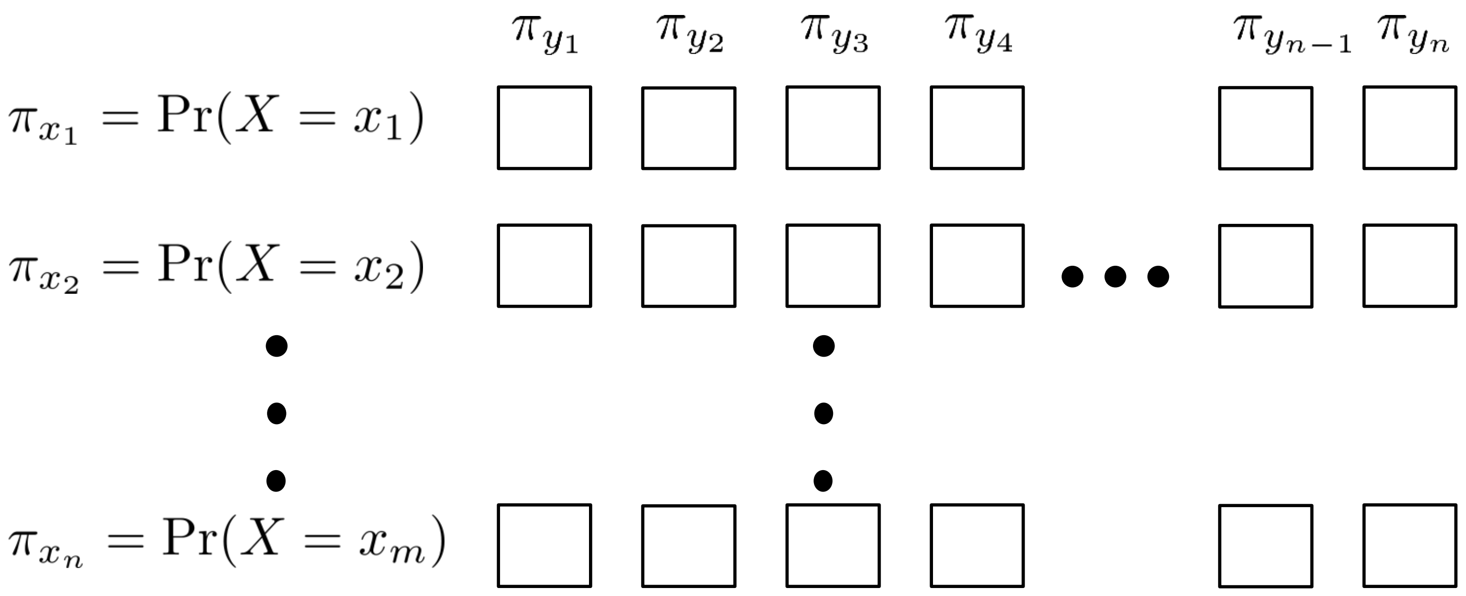
APPROACHES TO THE PROBLEM

- Classical Asymptotics **(Prior work)**
  - "Fixed cells" assumptions: $N \to \infty$ while d remains fixed
  - Likelihood Ratio Tests, χ^2 Tests are optimal in the minimax sense
  - **But in practice, N is small**
- High Dimensional Asymptotics
  - Both N and $d \to \infty$, use relative rate to get problem complexity
  - Unclear how to preserve comparison structure for this problem
- Finite Sample Analysis **(Our work)**
  - Many recent developments: (Acharya et al, 2015; Valiant and Valiant, 2017; Wei and Wainwright 2016)
  - The **good**: Comparison structure does not disappear + guidance on large N + 'special dimension dependence'
  - The **bad**: lower bound is hard, achievable upper bound is unknown

WHAT IS THE RELEVANT DIMENSION?

Independence

- Let X and Y be two discrete random variables with m and n states respectively.
- Model a joint distribution: $\pi_{x_i,y_j} = \Pr(X = x_i \cap Y = y_i)$



$$H_0 : \pi_{x_i,y_j} = \pi_{x_i}\pi_{y_j}$$

$$\pi_x \in \Delta_m, \pi_y \in \Delta_n$$

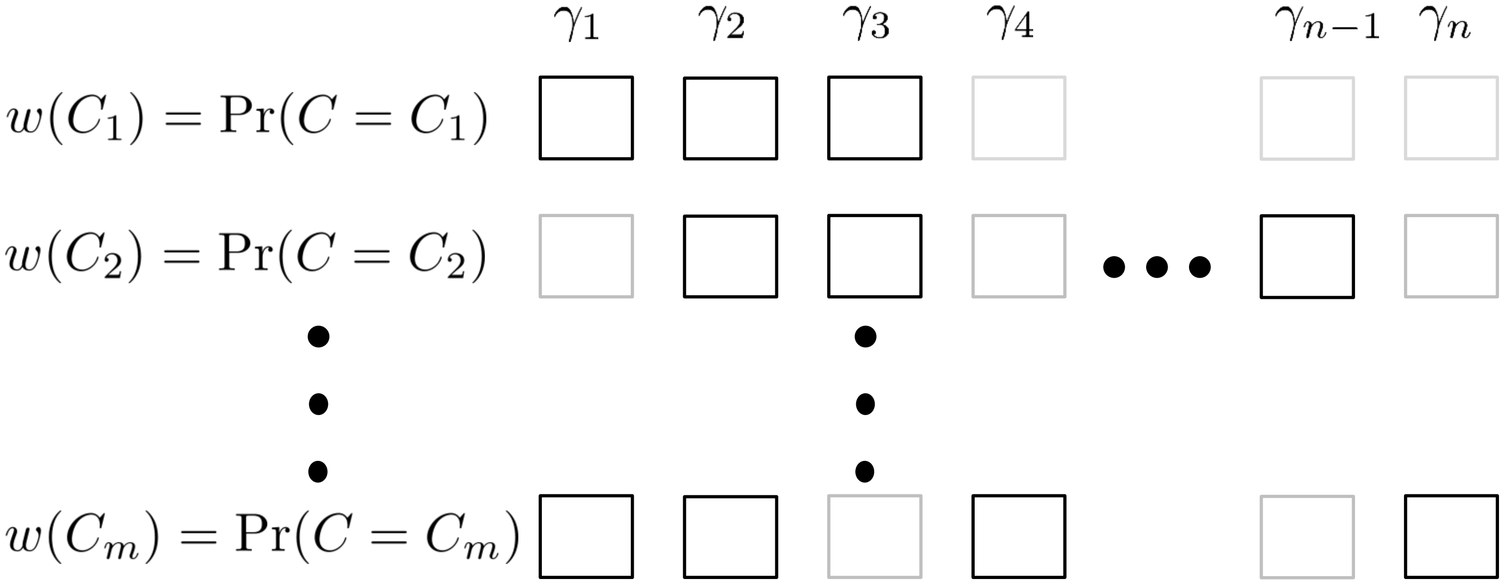
Although the null model has only $m + n - 2$ parameters, the full space of alternatives **has** $mn - 1$ **parameters**

WHAT IS THE RELEVANT DIMENSION?

Independence of Irrelevant Alternatives

- Model a “choice system”:

$$\pi_{i,C_j} = \Pr(C = C_i \cap x = i)$$



$$H_0 : \pi_{i,C_j} = w(C_j)\frac{\gamma_i}{\sum_{k \in C_j} \gamma_k}$$

$$w \in \Delta_m, \gamma \in \Delta_n$$

Although the null model has only $m + n - 2$ parameters, the full space of alternatives **has** $\mathrm{d} = \sum_C |C|$ **parameters**

PRESENT WORK: DEFINING THE PROBLEM

- *Choice System*: $q \in \Delta_d$, where $q_{(x,C)} = w(C)P_{x,C}$
  - defined differently from (Falmagne, 1978), who did not consider $w(C)$
- IIA constraints q only by restricting P
  - $P_{x,C} = \frac{\gamma_x}{\sum_{y \in C} \gamma_y}$
- $\mathcal{P}_{\mathcal{C}}$ - the space of all possible q
- $\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}} \subset \mathcal{P}_{\mathcal{C}}$ - the space of all q satisfying the IIA condition.
- Crucial assumptions:
  - C are of even size
  - every item appears an even number of times over $\mathcal{C}$

TESTING PROBLEM: SEPARATION

- Might be tempted to write:
$$\begin{cases} H_0 : (x, C) \sim p^N & \text{for some unknown } p \in \mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}} \\ H_1 : (x, C) \sim q^N & \text{for some unknown } q \in \mathcal{P}_{\mathcal{C}} \setminus \mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}} \end{cases}$$

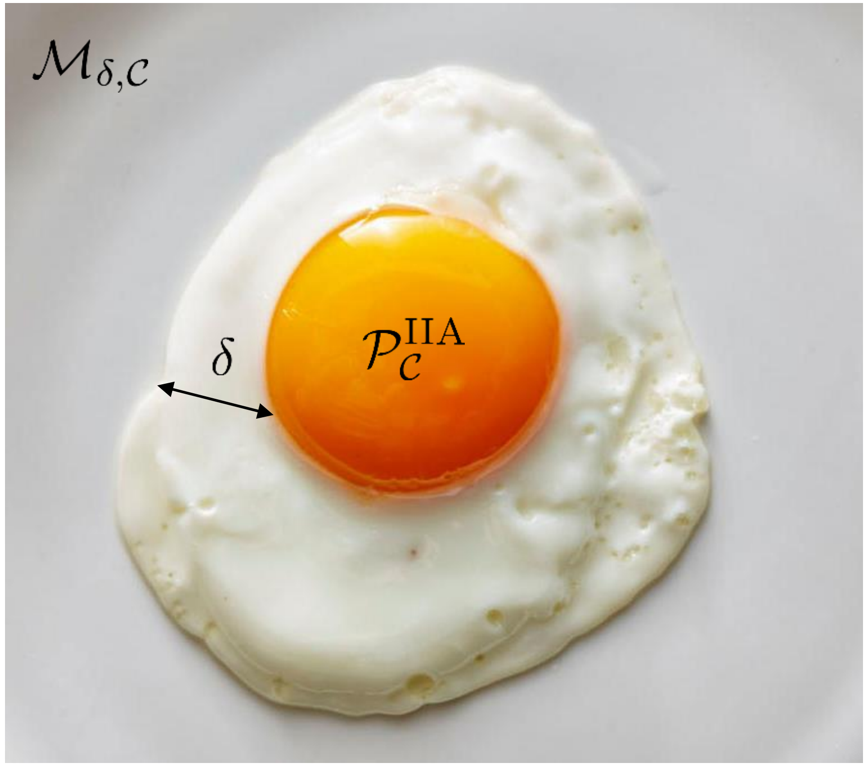
- But tests cannot be analyzed without a notion of **separation**
  - Indifference zone between null and alternative, beyond which false acceptance a serious error
  - Separation makes division of IIA vs non-IIA sharp

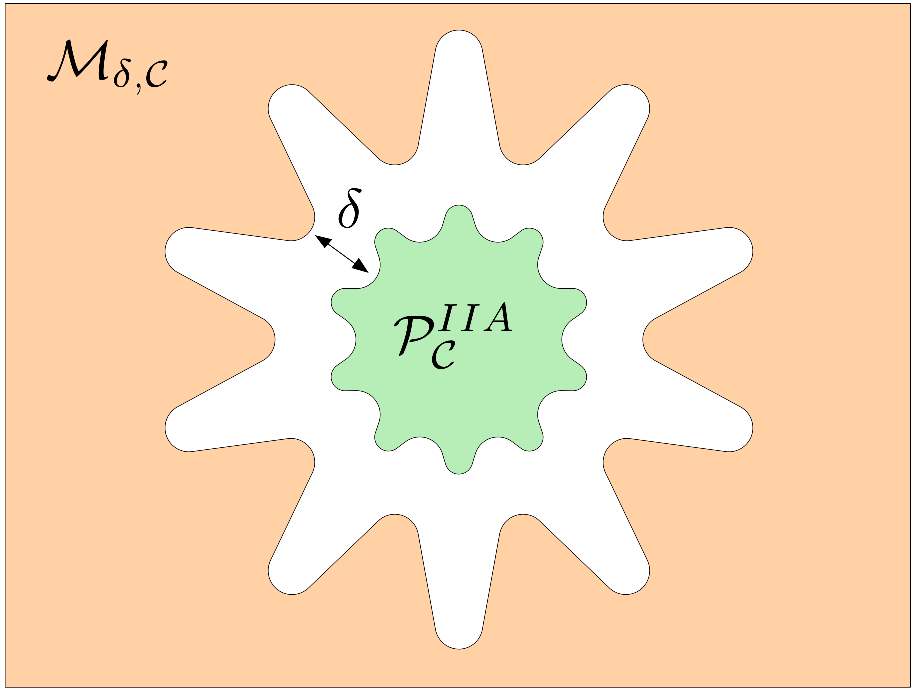
- Why TV?
  - Interpretable: all events are δ apart
  - Many other measure of separation

$$\mathcal{M}_{\delta,\mathcal{C}} = \{q : q \in \mathcal{P}_{\mathcal{C}},\ \inf_{p \in \mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}} ||q - p||_{\mathrm{TV}} \geq \delta\}.$$

- Actual testing problem:
$$\begin{cases} H_0 : (x, C) \sim p^N & \text{for some unknown } p \in \mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}} \\ H_1 : (x, C) \sim q^N & \text{for some unknown } q \in \mathcal{M}_{\delta,\mathcal{C}}. \end{cases}$$

SEPARATION





"Inside the yolk, or outside the egg"

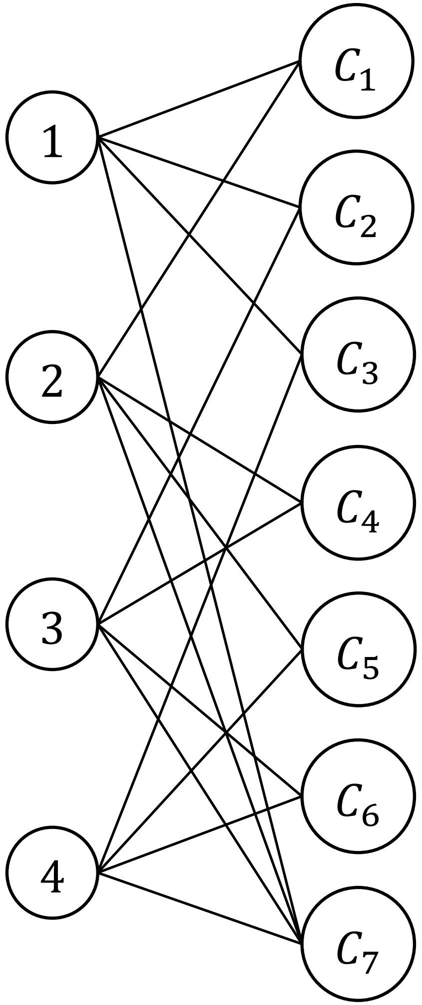
COMPARISON INCIDENCE GRAPH

- $G_{\mathcal{C}} = (\mathcal{X}, \mathcal{C}, E)$ - bipartite undirected graph with n nodes for each item, and with m nodes for each set, and d edges denoting membership of item $x \in \mathcal{X}$ in a set $C \in \mathcal{C}$.
- $G_{\mathcal{C}}$ is Eulerian if all $C \in \mathcal{C}$ are of even size, and all $x \in \mathcal{X}$ appears in an even number of sets.

A reoccurring example:

$$\mathcal{C} = \{\underbrace{\{1,2\}}_{C_1}, \underbrace{\{1,3\}}_{C_2}, \underbrace{\{1,4\}}_{C_3}, \underbrace{\{2,3\}}_{C_4}, \underbrace{\{2,4\}}_{C_5}, \underbrace{\{3,4\}}_{C_6}, \underbrace{\{1,2,3,4\}}_{C_7}\}$$

$$n = 4; m = 7; d = 16$$



The resulting $G_{\mathcal{C}}$

TESTING PROBLEM: RISK

- *Hypothesis test of interest*: Given N samples from a q for a known $\mathcal{C}$, determine if $q \in \mathcal{P}_{\mathrm{IIA}}$ or $q \in \mathcal{M}_\delta$
- $\phi : \underbrace{(x_1, C_1), ..., (x_N, C_N)}_{\mathcal{D}_N} \mapsto \{0, 1\}$ - a tester, i.e. ϕ defined as a map from data to decision: IIA or not
- $R_{N,\delta}(\mathcal{P}_0, \phi) = \sup_{p \in \mathcal{P}_0, q \in \mathcal{M}_\delta} \frac{1}{2} p^N(\phi(\mathcal{D}_N) = 1) + \frac{1}{2} q^N(\phi(\mathcal{D}_N) = 0)$ - the "worst case error" of a ϕ
- **Objective**: Lower bound $R_{N,\delta}(\mathcal{P}_0) = \inf_\phi R_{N,\delta}(\mathcal{P}_0, \phi)$, the error of the *best possible test.*

MAIN RESULT

Theorem. *Up to some constant* c_1 *and problem parameters* $\mu(\sigma)$ *and* $\alpha(\sigma)$ *that depend on the Eulerian comparison incidence graph* $G_{\mathcal{C}}$*, the minimax risk* $R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{IIA})$ *is lower bounded as*

$$\frac{1}{2} - \frac{1}{4}\Big(\exp\Big(\frac{c_1\mu(\sigma)^4\alpha(\sigma)N^2\delta^4}{d} - 1\Big)^{\frac{1}{2}} \leq R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{IIA}).$$

The testing radius then scales as $\delta_N(\mathcal{P}_{\mathcal{C}}^{IIA}) \asymp \frac{d^{\frac{1}{4}}}{\mu(\sigma)\alpha(\sigma)^{\frac{1}{4}}\sqrt{N}}$*, and the sample complexity as* $N_\delta(\mathcal{P}_{\mathcal{C}}^{IIA}) \asymp \frac{\sqrt{d}}{\sqrt{\mu(\sigma)^4\alpha(\sigma)\delta^2}}$.

- $\mu(\sigma)$ and $\alpha(\sigma)$ are structural problem parameters dependent on $G_{\mathcal{C}}$
  - universal constants for dense graphs, at most $2n$ for any graph
  - leads to weakened universal bound:

$$\frac{1}{2} - \frac{1}{4}\Big(\exp\Big(\frac{c_1 n^5 N^2\delta^4}{d} - 1\Big)^{\frac{1}{2}} \leq R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}).$$

CONSEQUENCES (HIGH LEVEL)

- Pessimism: a lower bound for the setting of all subsets:

$$R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\text{IIA}}) \geq \frac{1}{2} - \frac{1}{4}\Big(\exp\Big(\frac{c_1 n^4 N^2 \delta^4}{2^{n-2}}\Big) - 1\Big)^{\frac{1}{2}}.$$

  - best possible test for IIA has worst case error of $\frac{1}{2}$ until the samples are exponentially large in n
  - Anna Karenina principle at work

- Optimism: a lower bound for the setting of all pairs:

$$R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\text{IIA}}) \geq \frac{1}{2} - \frac{1}{4}\Big(\exp\Big(\frac{c_1 N^2 \delta^4}{n(n-1)}\Big) - 1\Big)^{\frac{1}{2}}.$$

  - Packing argument reduces $\alpha(\sigma)$ and $\mu(\sigma)$ to constants
  - Rationality is much easier to test if you restrict the number of irrationalities it is tested against

CONSEQUENCES (HIGH LEVEL)

- A simple "cycle" among n items: e.g. $\{i,j\},\{j,k\},\{k,l\},...\{z,i\}$

$$R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \geq \frac{1}{2} - \frac{1}{4}\Big(\exp\Big(c_1 n^4 N^2 \delta^4\Big) - 1\Big)^{\frac{1}{2}}.$$

  – Lower bound falls away fast, regardless of n - "dimension free"

  – Researcher priors are valuable in very low data settings

THE DETAILS

REDUCING THE PROBLEM

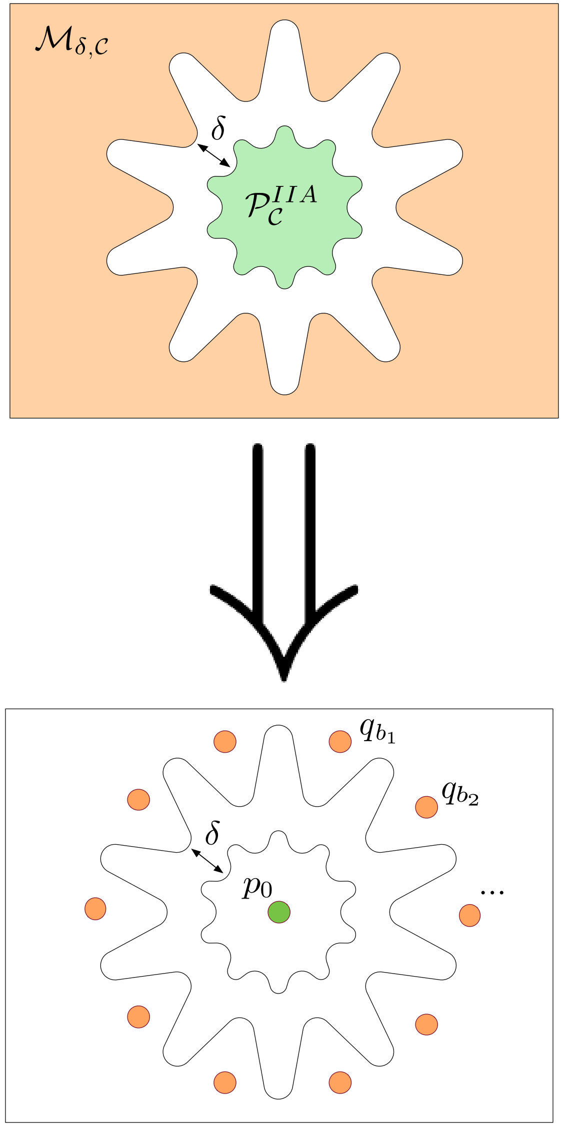
- Two stages of reductions
  - Reduction 1: $\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}$ vs $\mathcal{M}_{\delta,\mathcal{C}} \rightarrow$ uniform (p_0) vs $\mathcal{M}_{\delta,\mathcal{C}}$
  - Reduction 2: uniform (p_0) vs $\mathcal{M}_{\delta,\mathcal{C}} \rightarrow p_0$ vs M-mixture of perturbations $\bar{q}_\epsilon = \frac{1}{M}\sum_b q_{b,\epsilon}$
    - where:

$$q_{b,\epsilon} = p_0 + \frac{\epsilon b}{d}$$

    - and $b \in \{-1, 1\}^d$ satisfies

$$\sum_{\substack{C \in \mathcal{C} \\ C \ni x}} b_{x,C} = 0 \ \ \forall x \in \mathcal{X}, \qquad \sum_{y \in C} b_{y,C} = 0 \ \ \forall C \in \mathcal{C}.$$

    - Reduction 2 gives a *binary* hypothesis test
- Perturbations exit $\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}$ for any ϵ
  - IIA allowed to vary uniformly over items and sets $\rightarrow b$ travel "orthogonally" to that
  - Another way to motivate is that the MLE is always uniform to this point
- We use a mixture beacuse it models, in a single distribution, the hardness of both resolving the non-IIA perturbation and distinguishing it from an IIA point



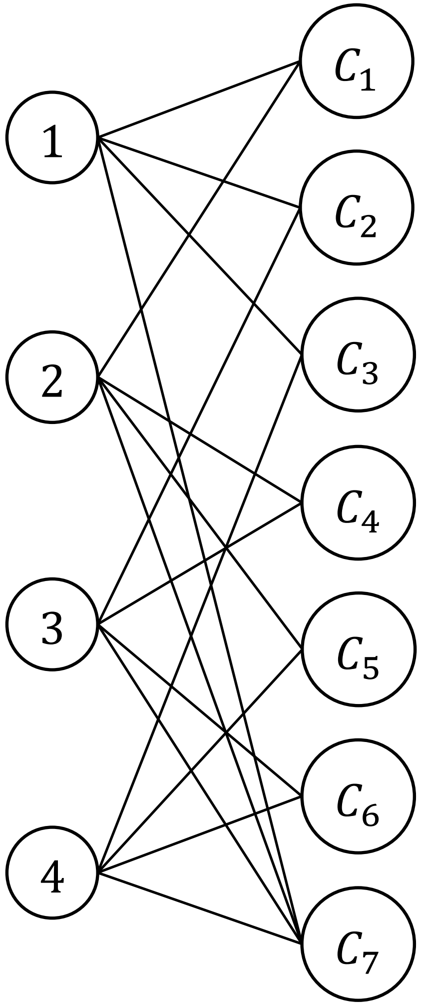
CONSTRUCTING PERTURBATIONS

❖ Are there any such b? How do we construct them?

- Recall: C are of even size, and every item appears an even number of times over $\mathcal{C}$, and so $G_{\mathcal{C}}$ is Eulerian
- Values in vector b can be thought as directing the edges of $G_{\mathcal{C}}$
- The constraints:

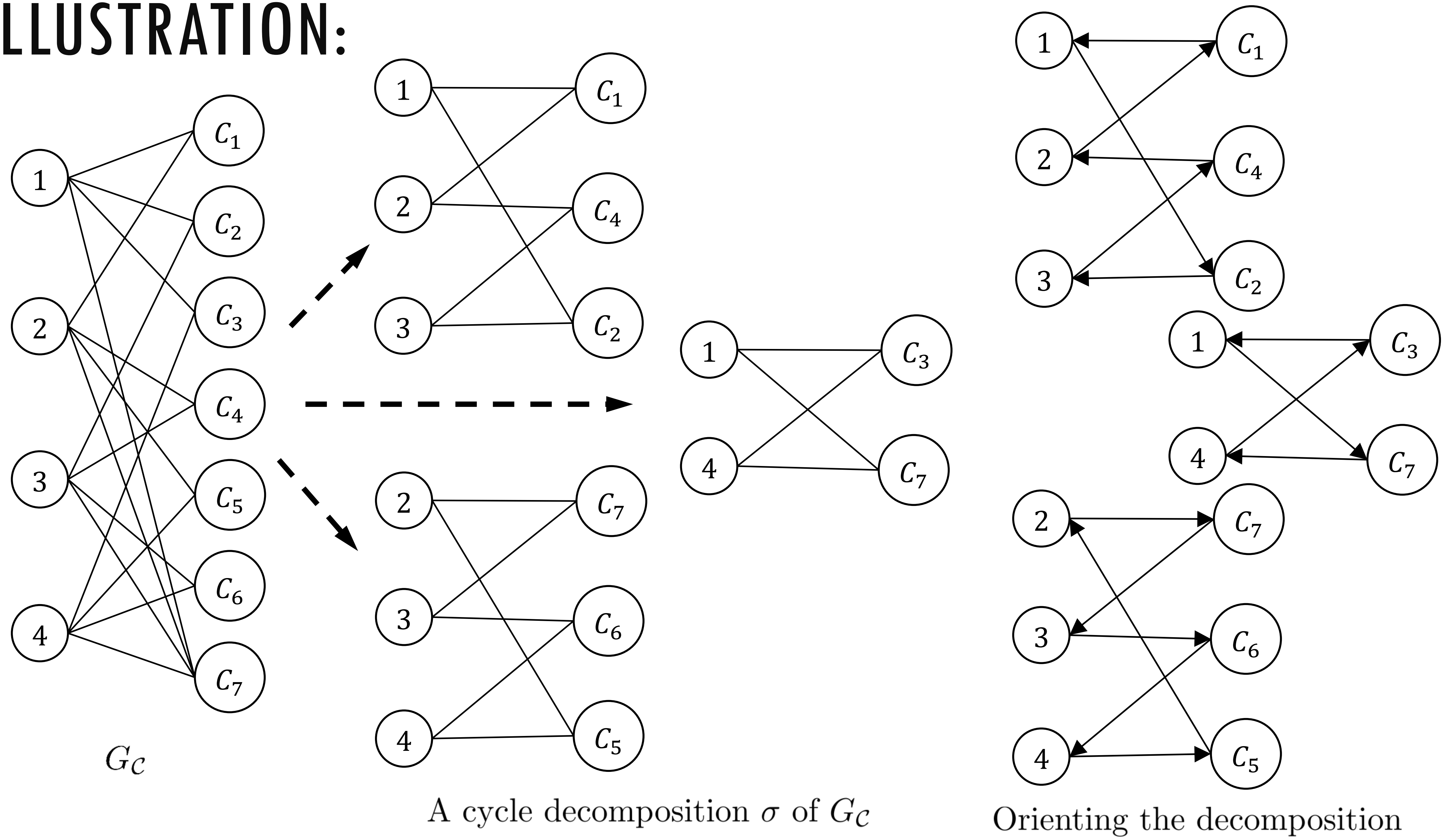
$$\sum_{\substack{C \in \mathcal{C} \\ C \ni x}} b_{x,C} = 0 \quad \forall x \in \mathcal{X}, \qquad\qquad \sum_{y \in C} b_{y,C} = 0 \quad \forall C \in \mathcal{C}.$$

  - Can be thought as ensuring *indegree* of every node equals *outdegree*
  - That is, b are ***Eulerian Orientations*** of $G_{\mathcal{C}}$!
- The process:
  - Find some simple cycle decomposition $\sigma \in \Sigma$, where Σ is the collection of all decompositions of $G_{\mathcal{C}}$ and $|\sigma|$ is the number of cycles in decomposition
    - ∗ Eulerian graphs can always be decomposed into simple cycles
      - · Since $G_{\mathcal{C}}$ bipartite, also bounded size cycles
  - Construct $2^{|\sigma|}$ b's by orienting each cycle clockwise and counter clockwise and toggling

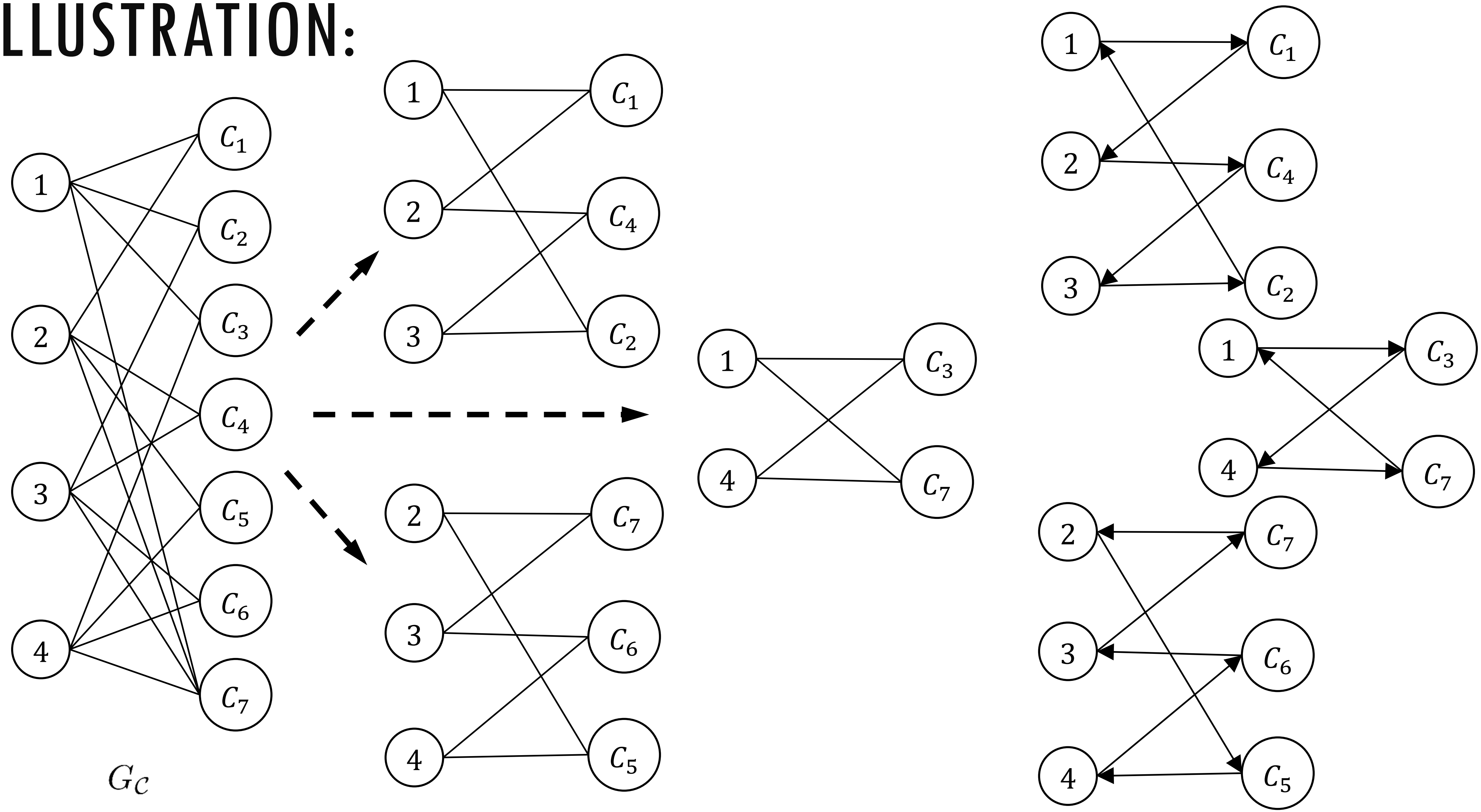


Our example from before

AN ILLUSTRATION:



AN ILLUSTRATION:



$G_{\mathcal{C}}$

A cycle decomposition σ of $G_{\mathcal{C}}$

Orienting the decomposition

PUTTING IT ALL TOGETHER

- A variation on LeCam's Method for binary hypothesis tests
  - Mild variations on a traditional chain of inequalities
- Two main quantities: $\mu(\sigma)$ (average cycle length) and $\alpha(\sigma)$ (average squared cycle length
- Main intuition: smaller the cycle decomposition, the better the result
  - Can always use worst case n result if all else fails

EXPERIMENTAL INTUITIONS

- The bound, as a function of the cycle decomposition, is revealing for experimental design
- Tradeoff: want to broaden scope of test, but more sets means the minimum error of any tester goes up; either due to bigger d or due to changes in $\alpha(\sigma)$and $\mu(\sigma)$ above a threshold
- **Goal:** maximize coverage of choice sets given a fixed number of samples and a risk threshold
- With this in mind,

REVISITING THE CONSEQUENCES

- All subsets:

$$R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \geq \frac{1}{2} - \frac{1}{4}\Big(\exp\Big(\frac{c_1 n^4 N^2 \delta^4}{2^{n-2}}\Big) - 1\Big)^{\frac{1}{2}}.$$

  – Any true test of IIA must encompass all subsets, as it is a property defined of the complete choice system

  – Lower bound is constructed by considering all even sized subsets

    ∗ a simple calculation reveals every item appears an even number of times

  – Global "n" lower bound suffices since d exponential in n

  – Sample complexity: $N_\delta(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \asymp \frac{\sqrt{2^{n-2}}}{n^2\delta^2}$.

  – Testing Radius: $\delta_N(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \asymp \frac{2^{\frac{n-2}{4}}}{n\sqrt{N}}$

REVISITING THE CONSEQUENCES

- All pairs:

$$R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \geq \frac{1}{2} - \frac{1}{4}\left(\exp\left(\frac{c_1 N^2 \delta^4}{n(n-1)}\right) - 1\right)^{\frac{1}{2}}.$$

  – $G_{\mathcal{C}}$ is Eulerian only when n is odd
  – An old result by Kirkman: for $n = 6x+1$ and $n = 6x+3$, any complete graph can be decomposed into triangles
    * Corollary is that $G_{\mathcal{C}}$ can be decomposed into 6-cycles
    * $\alpha_(\sigma)$ and $\mu(\sigma)$ are both constants
  – Feder et al show a similar result for $n = 6x+5$: triangles + one 4-cycle
  – Sample complexity: $N_\delta(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \asymp \frac{n}{\delta^2}$; Testing Radius: $\delta_N(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \asymp \frac{\sqrt{n}}{\sqrt{N}}$
  – In many matchups, pairs are the only relevant sets - tests scale linearly with items!

A SPECIAL CASE

- A simple "cycle" among n items: e.g. $\{i, j\}, \{j, k\}, \{k, l\}, ...\{z, i\}$

$$R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \geq \frac{1}{2} - \frac{1}{4}\Big(\exp\Big(c_1 n^4 N^2 \delta^4\Big) - 1\Big)^{\frac{1}{2}}.$$

  - Upper bound on $\alpha(\gamma)$ is sharp, since only cycle decomposition of $G_{\mathcal{C}}$ is $G_{\mathcal{C}}$ itself
  - Lower bound falls away fast, regardless of n - "dimension free"
    - feature reminiscent of property testing for cyclicality

- In settings of highly limited samples, bound suggests choosing a simple cycle C
  - In choice systems rife with IIA violations, most cycles would contain some
  - The benefit of picking a cycle, over, say, all pairs is a low rate of test error, since pair tests will necessarily have high test error
    - Trades off a result guarenteed to be errant for a more conservative result likely to be veritable

MODEL BASED TESTS

- Another notion of a prior stems from a valid model for departures from IIA.
- These “model based” tests are scant
  - the models themselves are often uninterpreble, inferentially intractable, or both
- Recent work proposes the CDM (Seshadri et al.)
  - Capable of modelling departures from IIA
  - Exhibiting ease of optimization (convexity), tractable finite sample uniform convergence guarantees, and parametric efficiency
- Natural limitation: that departures in the blindspots of the model will remain untested.

CONCLUSIONS

- First formal results on the complexity of testing IIA, resolving an decades-old question
- We are left to wonder: exactly when has IIA been rejected with veracity?
  - Lays the groundwork for a rigorous rethinking of the IIA testing problem
- Relationships between the comparison structure and testing complexity open several new directions for experimental design
  - **Exciting Future Direction:** An optimal procedure
- Simultaneously brings both:
  - Machine Learning rigor to Econometrics models (finite sample minimax rates)
  - Econometrics models (IIA composite nulls) into Machine Learning research

EXTRA SLIDES

PRESENT WORK: DEFINING THE PROBLEM

- $\mathcal{X}$ - universe set of n items
- $\mathcal{C}$ - collection of comparison sets, i.e. unique subsets of $\mathcal{X}$; $m = |\mathcal{C}|$; $d = \sum_{C \in \mathcal{C}} |C|$.
- $P_{\cdot,C} \in \Delta_{|C|}$, where $P_{x,C}$ is probability that item x is chosen from C
- $w \in \Delta_m$, where $w(C) > 0$ is the probability of seeing choice set $C \in \mathcal{C}$
- $q \in \Delta_d$, where $q_{(x,C)} = w(C)P_{x,C}$, is the choice system (defined differently from Falmagne)
- IIA constraints q only by restricting P so that $P_{x,C} = \frac{\gamma_x}{\sum_{y \in C} \gamma_y}$ for some $\gamma \in \Delta_n$; w remains arbitrary.
- $\mathcal{P}_{\mathcal{C}}$ - the space of all possible q; $\mathcal{P}_{\mathcal{C}}^{\text{IIA}} \subset \mathcal{P}_{\mathcal{C}}$ - the space of all q satisfying the IIA condition.
- Crucial assumption: C are of even size, and every item appears an even number of times over $\mathcal{C}$

MAIN RESULT: REFRAMING FOR LEVEL-α TESTS

- Define $\Phi_{N,\alpha} = \{\phi : \sup_{p \in \mathcal{P}_0} p^N(\phi(\mathcal{D}_N) = 1) \leq \alpha\}$ as the set of all level α tests.
- Since the type I error is always controlled, the risk of interest is now only the type II error. Thus, the minimax risk is:

$$R_{N,\delta,\alpha}(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) = \inf_{\phi \in \Phi_{N,\alpha}} \sup_{q \in \mathcal{M}_{\delta,\mathcal{C}}} q^N(\phi(\mathcal{D}_N) = 0).$$

- The minimax risk $R_{N,\delta,\alpha}(\mathcal{P}_0)$ is lower bounded as

$$1 - \alpha - \frac{1}{2}\Big(\exp\Big(\frac{c_1 \mu(\sigma)^4 \alpha(\sigma) N^2 \delta^4}{d} - 1\Big)^{\frac{1}{2}} \leq R_{N,\delta,\alpha}(\mathcal{P}_0),$$

- When $\frac{c_1 \mu(\sigma)^4 \alpha(\sigma) N^2 \delta^4}{d}$ is small, the power is just α
  - test is simply a coin with probability α in the worst case

BOUNDING SEPARATION

- We have perturbations in terms of ϵ, but for what ϵ does $q_{b,\epsilon} \in \mathcal{M}_{\delta,\mathcal{C}}$?

- Easy to show $||q_{b,\epsilon} - p_0||_{\mathrm{TV}} = \frac{\epsilon}{2}$
  - *But*, this is not the closest point in $\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}$
- Can show that $\inf_{p \in \mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}} ||q_{b,\epsilon} - p||_{\mathrm{TV}} \geq \frac{\epsilon|\sigma|}{2d} = \frac{\epsilon}{2\mu(\sigma)}$
  - where $\mu(\sigma)$ is the average cycle length, which is at most $2n$
- Thus, setting $\epsilon = 2\mu(\sigma)$ guarentees membership in $\mathcal{M}_{\delta,\mathcal{C}}$ for a σ, and $\epsilon = 4n\mu(\sigma)$ guarentees membership for any σ.
- Proof follows partitioning the indices into cycles, and showing that each cycle has a lower bound on its contribution to the TV distance

LECAM'S METHOD: A LOWER BOUND

- We have reduced the problem into a <u>binary hypothesis test</u>: $\begin{cases} H_0 : (x, C) \sim \mathbb{P}_0 = {p_0}^N \\ H_1 : (x, C) \sim \mathbb{P}_1 = \bar{q}_\epsilon^N. \end{cases}$

- Using $\gamma(\mathbb{P}_0, \mathbb{P}_1)$ to denote the average of type I and type II errors of the best possible test, we have,

$$\gamma(\mathbb{P}_0, \mathbb{P}_1) \geq \frac{1}{2} - \frac{1}{2}||\mathbb{P}_0 - \mathbb{P}_1||_{\mathrm{TV}}$$

- Consequently,

$$R_{N,\delta}(\mathcal{P}_{\mathcal{C}}^{\mathrm{IIA}}) \geq \gamma(\mathbb{P}_0, \mathbb{P}_1) \geq \frac{1}{2} - \frac{1}{2}||\mathbb{P}_0 - \mathbb{P}_1||_{\mathrm{TV}}.$$

- We have a lower bound!

BOUNDING TV

- All that remains is bounding the TV in a useful manner
  - Consider that

$$||p-q||_{TV}^2 \leq \frac{1}{4}\chi^2(p,q)$$

  - Then, we have,

Lemma. *3,4*

$$\chi^2(\mathbb{P}_1,\mathbb{P}_0)+1 \leq \frac{1}{M^2}\sum_{b,b'\in\mathcal{B}_C}\exp\left(\frac{N\epsilon^2}{d}b^Tb'\right) \leq \exp\left(\frac{N^2\epsilon^4}{2d}\alpha(\sigma)\right),$$

where $\mathcal{B}_C$ is a set of arbitrary perturbations b of size $|\mathcal{B}_C| = M$ satisfying $b \in \{-1,1\}^d$, $\sum_{\substack{C\in\mathcal{C}\\ C\ni x}} b_{x,C} = 0, \forall x,$ and $\sum_{y\in C} b_{y,C} = 0, \forall C$.

  - Important is $\alpha(\sigma) = \frac{1}{d}\sum_{\sigma_i\in\sigma}|\sigma_i|^2$
    - serves as a normalized measure of the "energy" of the cycle decomposition σ