Paintboxes and probability functions for edge-exchangeable graphs

Diana Cai
Department of Statistics
University of Chicago

Joint work with:
Trevor Campbell and Tamara Broderick
(MIT CSAIL)
Network data (graphs): interactions between individuals
Network data (graphs): interactions between individuals

**social**: Facebook, Twitter, email
Network data (graphs): interactions between individuals

**social**: Facebook, Twitter, email

**biological**: ecological, protein, gene
Network data (graphs): interactions between individuals

social: Facebook, Twitter, email
biological: ecological, protein, gene
transportation: roads, railways
Network data (graphs): interactions between individuals

Probabilistic models for graphs

$\mathbb{P}(\text{social}: \text{Facebook, Twitter, email})$

$\mathbb{P}(\text{biological}: \text{ecological, protein, gene})$

$\mathbb{P}(\text{transportation}: \text{roads, railways})$
Network data (graphs): interactions between individuals

Probabilistic models for graphs

$p(\cdot)$

social: Facebook, Twitter, email

biological: ecological, protein, gene

transportation: roads, railways

Many probabilistic models assume vertex exchangeability: dense (too many edges).
Network data (graphs): interactions between individuals

Probabilistic models for graphs

$p(\cdot)$

- **social**: Facebook, Twitter, email
- **biological**: ecological, protein, gene
- **transportation**: roads, railways

Many probabilistic models assume *vertex exchangeability*: dense (too many edges). Real-world graphs are *sparse*. 
Network data (graphs): interactions between individuals

Probabilistic models for graphs

Many probabilistic models assume vertex exchangeability: dense (too many edges). Real-world graphs are sparse.

Under edge exchangeability, we can get sparse graphs.

[Broderick & Cai, 2015; Cai, Campbell, Broderick, 2016; Crane & Dempsey, 2015, 2016]
Network data (graphs): interactions between individuals

Probabilistic models for graphs

Many probabilistic models assume vertex exchangeability: dense (too many edges). Real-world graphs are sparse.

Under edge exchangeability, we can get sparse graphs. We want a representation theorem that characterizes all edge-exchangeable graphs.

[Broderick & Cai, 2015; Cai, Campbell, Broderick, 2016; Crane & Dempsey, 2015, 2016]
Network data (graphs): interactions between individuals

Probabilistic models for graphs

\[ p(\cdot) \]

social: Facebook, Twitter, email
biological: ecological, protein, gene
transportation: roads, railways

Many probabilistic models assume vertex exchangeability: dense (too many edges). Real-world graphs are sparse.

Under edge exchangeability, we can get sparse graphs. We want a representation theorem that characterizes all edge-exchangeable graphs. We also want to characterize models with easy inference.

[Broderick & Cai, 2015; Cai, Campbell, Broderick, 2016; Crane & Dempsey, 2015, 2016]
Network models should reflect **real-life scaling properties**
Network models should reflect **real-life scaling properties**

sequence of graphs
Network models should reflect real-life scaling properties

sequence of graphs

$G_1$
Network models should reflect **real-life scaling properties**

sequence of graphs

$G_1$ $G_2$
Network models should reflect **real-life scaling properties**
Network models should reflect real-life scaling properties

sequence of graphs

$G_1$  $G_2$  $G_3$  $G_4$
Network models should reflect **real-life scaling properties**

sequence of graphs

$G_1$ $G_2$ $G_3$ $G_4$ ...
Network models should reflect **real-life scaling properties**

sequence of graphs

\[ G_1 \quad G_2 \quad G_3 \quad G_4 \quad \ldots \]

**real-life scaling properties:**
Network models should reflect **real-life scaling properties**

sequence of graphs

\[ G_1 \quad G_2 \quad G_3 \quad G_4 \quad \ldots \]

**real-life scaling properties:**

Sparse:
Network models should reflect **real-life scaling properties**.

sequence of graphs

**real-life scaling properties:**
Sparse: \[\#\text{edges}(G_n) \in o(\#\text{vertices}(G_n)^2)\]
Network models should reflect **real-life scaling properties**

sequence of graphs

\[ G_1 \quad G_2 \quad G_3 \quad G_4 \quad \ldots \]

**real-life scaling properties:**
Sparse: \[ \#\text{edges}(G_n) \in o(\#\text{vertices}(G_n)^2) \] sub-quadratic
Network models should reflect **real-life scaling properties**

sequence of graphs

real-life scaling properties:
Sparse: \[
\#\text{edges}(G_n) \in o((\#\text{vertices}(G_n))^2)
\] sub-quadratic

popular models:
Network models should reflect **real-life scaling properties**

sequence of graphs

![Graph sequence](image)

real-life scaling properties:
- Sparse: \[ \#\text{edges}(G_n) \in o(\#\text{vertices}(G_n)^2) \] sub-quadratic

popular models:
- dense:
Network models should reflect real-life scaling properties:

sequence of graphs:

\[ G_1, G_2, G_3, G_4, \ldots \]

real-life scaling properties:
- Sparse: \[ \#\text{edges}(G_n) \in o((\#\text{vertices}(G_n))^2) \] sub-quadratic
- Dense: \[ \#\text{edges}(G_n) \in \Omega((\#\text{vertices}(G_n))^2) \]

popular models:
Network models should reflect **real-life scaling properties**

sequence of graphs

real-life scaling properties:
- **Sparse:** \[ \#\text{edges}(G_n) \in o(\#\text{vertices}(G_n)^2) \] \textit{sub-quadratic}

popular models:
- **Dense:** \[ \#\text{edges}(G_n) \in \Omega(\#\text{vertices}(G_n)^2) \] \textit{quadratic}
Network models should reflect **real-life scaling properties**

sequence of graphs

real-life scaling properties:

- Sparse: \[\#\text{edges}(G_n) \in o\left(\left[\#\text{vertices}(G_n)\right]^2\right)\] *sub-quadratic*

- Dense: \[\#\text{edges}(G_n) \in \Omega\left(\left[\#\text{vertices}(G_n)\right]^2\right)\] *quadratic*

**probabilistic models dense w.p. 1:**
- stochastic block model,
- mixed membership stochastic block model,
- infinite relational model, and many more
Network models should reflect **real-life scaling properties**

sequence of graphs

real-life scaling properties:
- Sparse: \( \#\text{edges}(G_n) \in o(\#\text{vertices}(G_n)^2) \) sub-quadratic
- Dense: \( \#\text{edges}(G_n) \in \Omega(\#\text{vertices}(G_n)^2) \) quadratic

popular models: **vertex exchangeability**

probabilistic models dense w.p. 1:
- stochastic block model,
- mixed membership stochastic block model,
- infinite relational model, and many more
Network models should reflect **real-life scaling properties**

sequence of graphs

real-life scaling properties: **edge exchangeability**

Sparse: \[
\#\text{edges}(G_n) \in o(\#\text{vertices}(G_n)^2)
\] sub-quadratic

popular models: **vertex exchangeability**

dense: \[
\#\text{edges}(G_n) \in \Omega(\#\text{vertices}(G_n)^2)
\] quadratic

**probabilistic models dense w.p. 1:**

stochastic block model,
mixed membership stochastic block model,
infinite relational model, and many more
Network models should reflect **real-life scaling properties**

sequence of graphs

real-life scaling properties: **edge exchangeability**
Sparse: \[
\#\text{edges}(G_n) \in o\left(\left[\#\text{vertices}(G_n)\right]^2\right)
\]

popular models: **vertex exchangeability**
dense: \[
\#\text{edges}(G_n) \in \Omega\left(\left[\#\text{vertices}(G_n)\right]^2\right)
\]

**probabilistic models dense w.p. 1:**
stochastic block model, mixed membership stochastic block model, infinite relational model, and many more
Vertex-exchangeable graph sequences (are always dense)
$G_1$
$G_1$
$G_1$

$G_2$

$G_3$

$G_4$

and so on …
\[ p(G_1) = p(G_2) = p(G_3) = p(G_4) \]
\[ p(G_1) = p(G_2) = p(G_3) = p(G_4) \]
$p(\begin{array}{cc} 1 & 2 \\ 2 & 3 \end{array}) = p(\begin{array}{cc} 1 & 4 \\ 2 & 3 \end{array})$
The Aldous-Hoover theorem implies that vertex-exchangeable graphs are dense or empty with probability 1.
Edge-exchangeable graph sequences
$G_1$
$G_1$

$G_2$
$G_1$  \hspace{1cm} $G_2$  \hspace{1cm} $G_3$  \hspace{1cm} $G_4$  \\

\[ p(1 \rightarrow 2 \rightarrow 4) \]
$p(1, 2, 4) = p(2, 4, 1)$

[Broderick & Cai, 2015; Cai, Campbell, Broderick, 2016; Crane & Dempsey, 2015, 2016]
\[ p(G_2) = p(G_3) \]

[Broderick & Cai, 2015; Cai, Campbell, Broderick, 2016; Crane & Dempsey, 2015, 2016]

Want to characterize *all* sparse, edge-exchangeable graphs.
Characterizing edge-exchangeable graphs: the graph paintbox
Clustering (a.k.a. partitions)
Clustering (a.k.a. partitions)

“clusters”
Clustering (a.k.a. partitions)

<table>
<thead>
<tr>
<th></th>
<th>Cat</th>
<th>Dog</th>
<th>Mouse</th>
<th>Lizard</th>
<th>Sheep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**exchangeable**: permuting data doesn’t change distribution of the random partition
**Vertex allocations**

- **Cat**
- **Dog**
- **Mouse**
- **Lizard**
- **Sheep**

<table>
<thead>
<tr>
<th>Edge 1</th>
<th>Cat</th>
<th>Dog</th>
<th>Mouse</th>
<th>Lizard</th>
<th>Sheep</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Edge 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Edge 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Edge 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Edge 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Edge 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Edge 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

**exchangeable**

Permuting the edges doesn’t change the distribution of the random vertex allocation (graph).
dust
these clusters only appear in a single data point
This clustering is *exchangeable.*
Theorem (Kingman, 1978). A random clustering is exchangeable iff it has a Kingman paintbox representation.

This clustering is exchangeable.
any point intersects at most 2 subsets of (0,1)

“cat and dog just interacted”
[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
“regular” (colored) vs “dust” (gray) vertices

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
"regular" (colored) vs "dust" (gray) vertices
“regular” (colored) vs “dust” (gray) vertices

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
at most two overlapping subsets
at most two overlapping subsets

one dust vertex, one regular vertex

dust vertex
cat vertex
dog vertex
mouse vertex
sheep vertex

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
no other overlapping subsets creates 2 dust vertices

dust vertex

cat vertex
dog vertex
mouse vertex
sheep vertex

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
Step 1
Step 2
Step 3
Step 4
Step 5
Step 6
Step 7
Step 8
Step 9

Dust vertex
cat vertex
dog vertex
mouse vertex
sheep vertex
step 1
step 2
step 3
step 4
step 5
step 6
step 7
step 8
step 9

1
dust vertex
cat vertex
dog vertex
mouse vertex
sheep vertex

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
step 1
step 2
step 3
step 4
step 5
step 6
step 7
step 8
step 9

C. [Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
step 1
step 2
step 3
step 4
step 5
step 6
step 7
step 8
step 9

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
step 1
step 2
step 3
step 4
step 5
step 6
step 7
step 8
step 9

[dust vertex, cat vertex, dog vertex, mouse vertex, sheep vertex]

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
step 1
step 2
step 3
step 4
step 5
step 6
step 7
step 8
step 9

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
This random graph is edge-exchangeable.

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
The graph paintbox relates edges that connect to the same vertex, and allows us to control the topology of the graph.

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
The graph paintbox relates edges that connect to the same vertex, and allows us to control the topology of the graph.
The graph paintbox relates edges that connect to the same vertex, and allows us to control the topology of the graph.

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
The graph paintbox relates edges that connect to the same vertex, and allows us to control the topology of the graph.

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
Theorem.
A random graph is edge-exchangeable iff it has a graph paintbox representation.

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
Graph frequency models

= Exchangeable vertex probability function
The graph paintbox is expressive but complex.
The graph paintbox is expressive but complex.

graph frequency model

[Cai, Campbell, Broderick, 2016a]
The graph paintbox is expressive but complex.

graph frequency model

1. Draw rates ($w_i$) from some distribution

[Cai, Campbell, Broderick, 2016a]
The graph paintbox is expressive but **complex**.

**graph frequency model**

1. Draw rates \((w_i)\) from some distribution
2. Draw edge \({i,j}\) with probability proportional to \(w_iw_j\)

[Cai, Campbell, Broderick, 2016a]
The graph paintbox is expressive but **complex**.

**graph frequency model**

1. Draw rates ($w_i$) from some distribution
2. Draw edge $\{i,j\}$ with probability proportional to $w_i w_j$

[Cai, Campbell, Broderick, 2016a]
The graph paintbox is expressive but complex.

### Graph Frequency Model

1. Draw rates ($w_i$) from some distribution
2. Draw edge $\{i,j\}$ with probability proportional to $w_i w_j$

---

[Cai, Campbell, Broderick, 2016a]
The graph paintbox is expressive but **complex**.

### Graph Frequency Model

1. Draw rates ($w_i$) from some distribution
2. Draw edge $\{i,j\}$ with probability proportional to $w_i w_j$

---

![Graph Frequency Model Diagram](image)

- **Degrees**: $\{1, 3, 2\}$, $3$
- **# of edges**: $3$

---

[Cai, Campbell, Broderick, 2016a]
The graph paintbox is expressive but complex.

**graph frequency model**

1. Draw rates \((w_i)\) from some distribution
2. Draw edge \(\{i,j\}\) with probability proportional to \(w_i w_j\)

\[
f(\{1,3,2\}, 3)
\]

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
The graph paintbox is expressive but complex.

Graph frequency model:
1. Draw rates ($w_i$) from some distribution
2. Draw edge $\{i,j\}$ with probability proportional to $w_i w_j$

Exchangeable vertex probability function (EVPF)

Graph frequency model:
$f(\{1,3,2\}, 3)$

# of edges
The graph paintbox is expressive but complex.

**graph frequency model**

1. Draw rates \((w_i)\) from some distribution
2. Draw edge \(\{i,j\}\) with probability proportional to \(w_i w_j\)

---

**Exchangeable vertex probability function (EVPF)**

Reminiscent of exchangeable partition probability functions for clustering for efficient Gibbs sampling and variational inference algorithms.

---

\[ f(\{1,3,2\}, 3) \]

\# of edges

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
The graph paintbox is expressive but \textit{complex}.

**Theorem.**
An edge-exchangeable graph has a graph frequency model iff it has an EVPF.

Reminiscent of exchangeable partition probability functions for clustering for efficient Gibbs sampling and variational inference algorithms.

[Campbell, Cai, Broderick, 2016; Cai, Campbell, Broderick, 2016b]
conclusions
✓ characterized the class of edge-exchangeable graphs

edge-exchangeable
  = graph paintbox

graph frequency models
  = EVPF
conclusions
✓ characterized the class of edge-exchangeable graphs
✓ characterized the class of graph frequency models
conclusions

✓ characterized the class of edge-exchangeable graphs
✓ characterized the class of graph frequency models

future work:

edge-exchangeable = graph paintbox

graph frequency models = EVPF
conclusions

✓ characterized the class of edge-exchangeable graphs
✓ characterized the class of graph frequency models

future work:

▷ characterize sparse, edge-exchangeable graph models
conclusions

✓ characterized the class of edge-exchangeable graphs
✓ characterized the class of graph frequency models

future work:

⦁ characterize sparse, edge-exchangeable graph models
⦁ characterize various types of sparse power laws (e.g., degrees, triangles)
conclusions

✓ characterized the class of edge-exchangeable graphs
✓ characterized the class of graph frequency models

future work:

- characterize sparse, edge-exchangeable graph models
- characterize various types of sparse power laws (e.g., degrees, triangles)
- truncation and practical posterior inference algorithms (frequency models and EVPF)
references


Alternate versions in:

  • NIPS 2016 Workshop on Practical Bayesian Nonparametrics.


Preliminary versions in:

  • NIPS 2015 Workshop on Networks in the Social and Information Sciences.
  • NIPS 2015 Workshop, Bayesian Nonparametrics: The Next Generation.


