priors on exchangeable directed graphs

Diana Cai
Department of Statistics
University of Chicago

Collaborators:
Nathanael Ackerman (Harvard)
Cameron Freer (Gamalon and MIT)

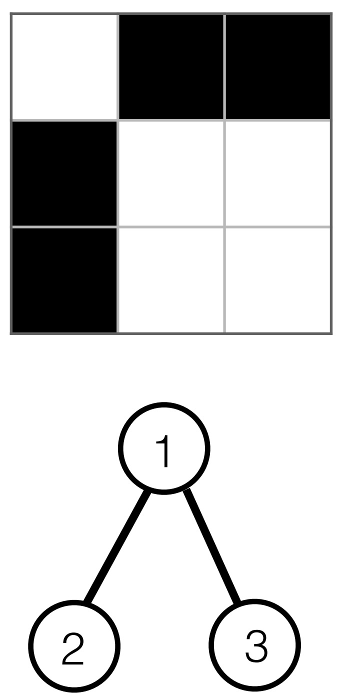
Overview

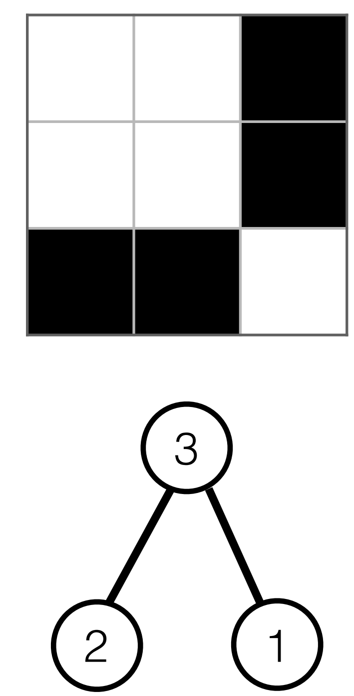
- We consider **exchangeable (dense) graphs**.
- Specifically, we are interested in **directed** graphs. However, most work has focused on undirected graphs.
- **Aldous–Hoover** has an analogous statement for directed graphs, which is more complicated than merely using an *asymmetric* function.
- Many natural nonparametric priors on exchangeable undirected graphs **extend to the directed case.**
- This perspective leads to natural priors on other exchangeable structures, such as **tournaments** and **directed acyclic graphs**.

Exchangeable graphs

exchangeability:

- order of vertices doesn’t affect distribution of graph

$\Pr($ 

$) = \Pr($ 

$)$

Exchangeable graphs

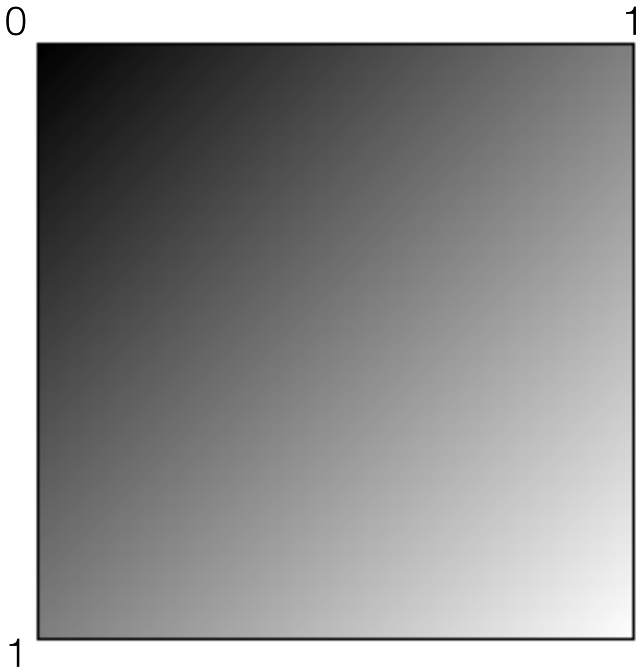
Undirected	Directed
Any exchangeable random infinite graph is obtained as a mixture of $\mathbb{G}(\infty, W)$. $\mathbb{G}(\infty, W)$ is a sampling procedure from the **graphon** W. $W : [0,1]^2 \to [0,1]$ $W(x,y) = W(y,x)$	

Aldous–Hoover

Exchangeable graphs

Graphon:

$$W : [0,1]^2 \to [0,1]$$
$$W(x,y) = W(y,x)$$

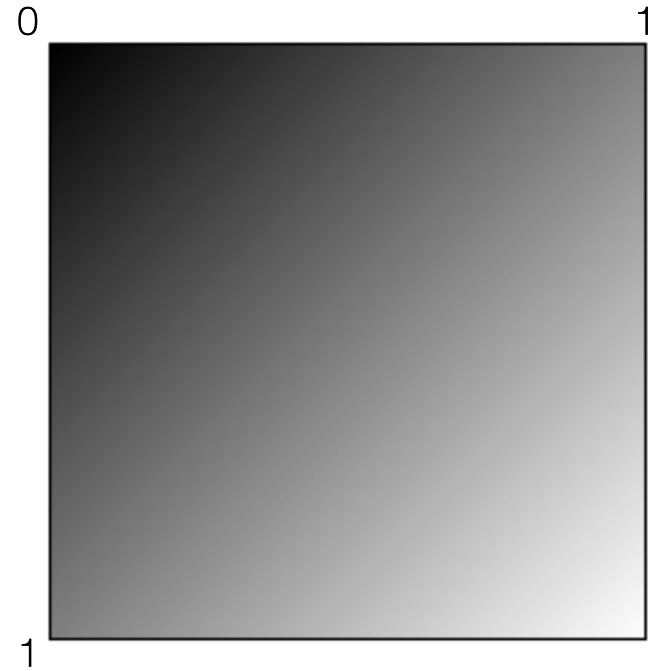


$$W(x,y) = \frac{(1-x)+(1-y)}{2}$$

Exchangeable graphs

Graphon:

$$W : [0,1]^2 \to [0,1]$$
$$W(x,y) = W(y,x)$$



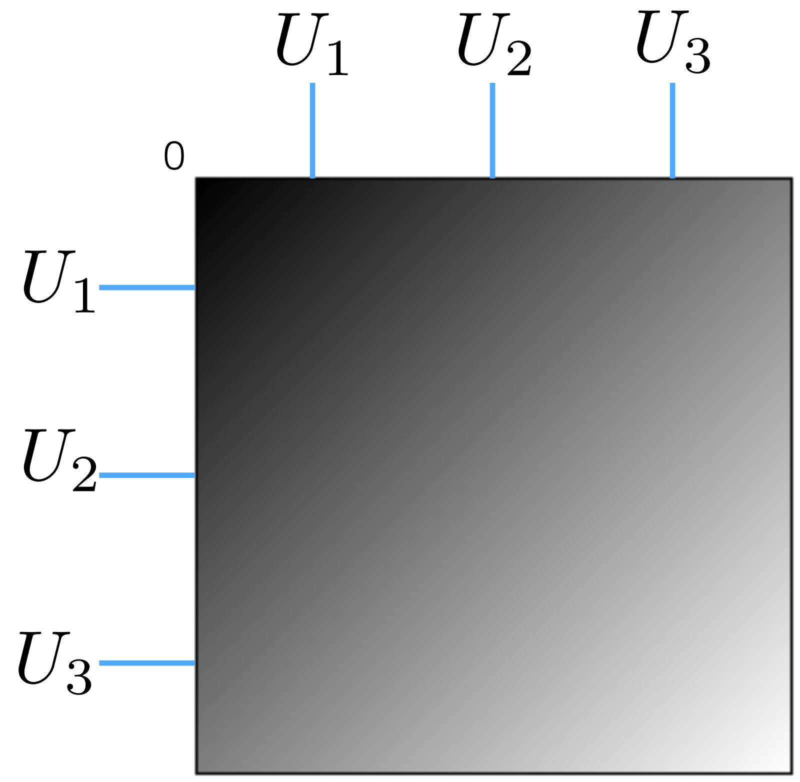
Sampling procedure:

$$U_i \sim \mathrm{Uniform}[0,1]$$
$$G_{ij} \sim \mathrm{Bern}(W(U_i, U_j)), i < j \quad (\text{Set } G_{ji} = G_{ij})$$

Exchangeable graphs

Graphon:

$$W : [0,1]^2 \to [0,1]$$
$$W(x,y) = W(y,x)$$



Sampling procedure:

$$U_i \sim \text{Uniform}[0,1]$$

$$G_{ij} \sim \text{Bern}(W(U_i, U_j)), i < j \quad (\text{Set } G_{ji} = G_{ij})$$

 [Cai–Ackerman–Freer, 2015]

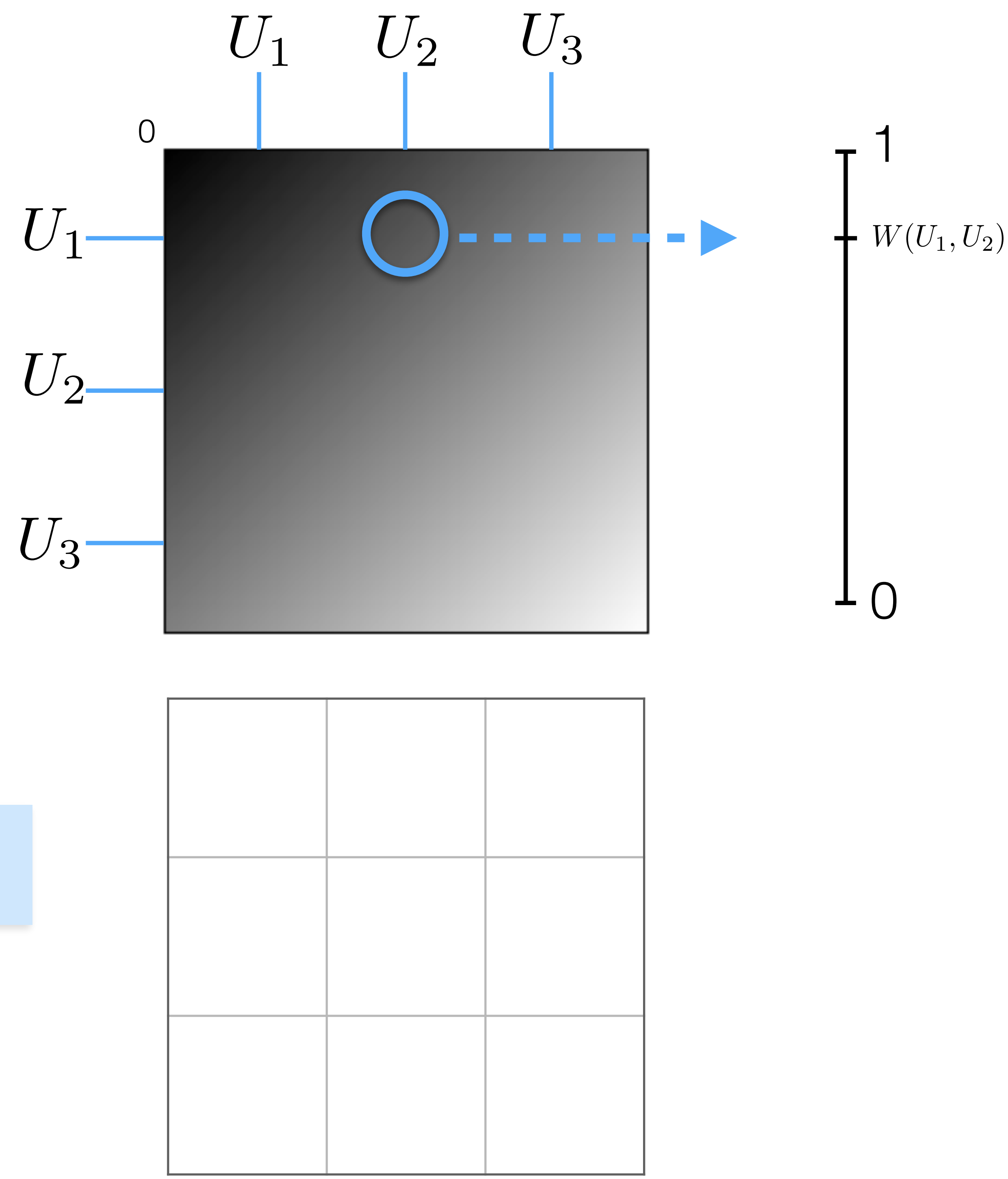
Exchangeable graphs

Graphon:

$$W : [0,1]^2 \to [0,1]$$
$$W(x,y) = W(y,x)$$

Sampling procedure:

$$U_i \sim \mathrm{Uniform}[0,1]$$
$$G_{ij} \sim \mathrm{Bern}(W(U_i, U_j)), i < j$$



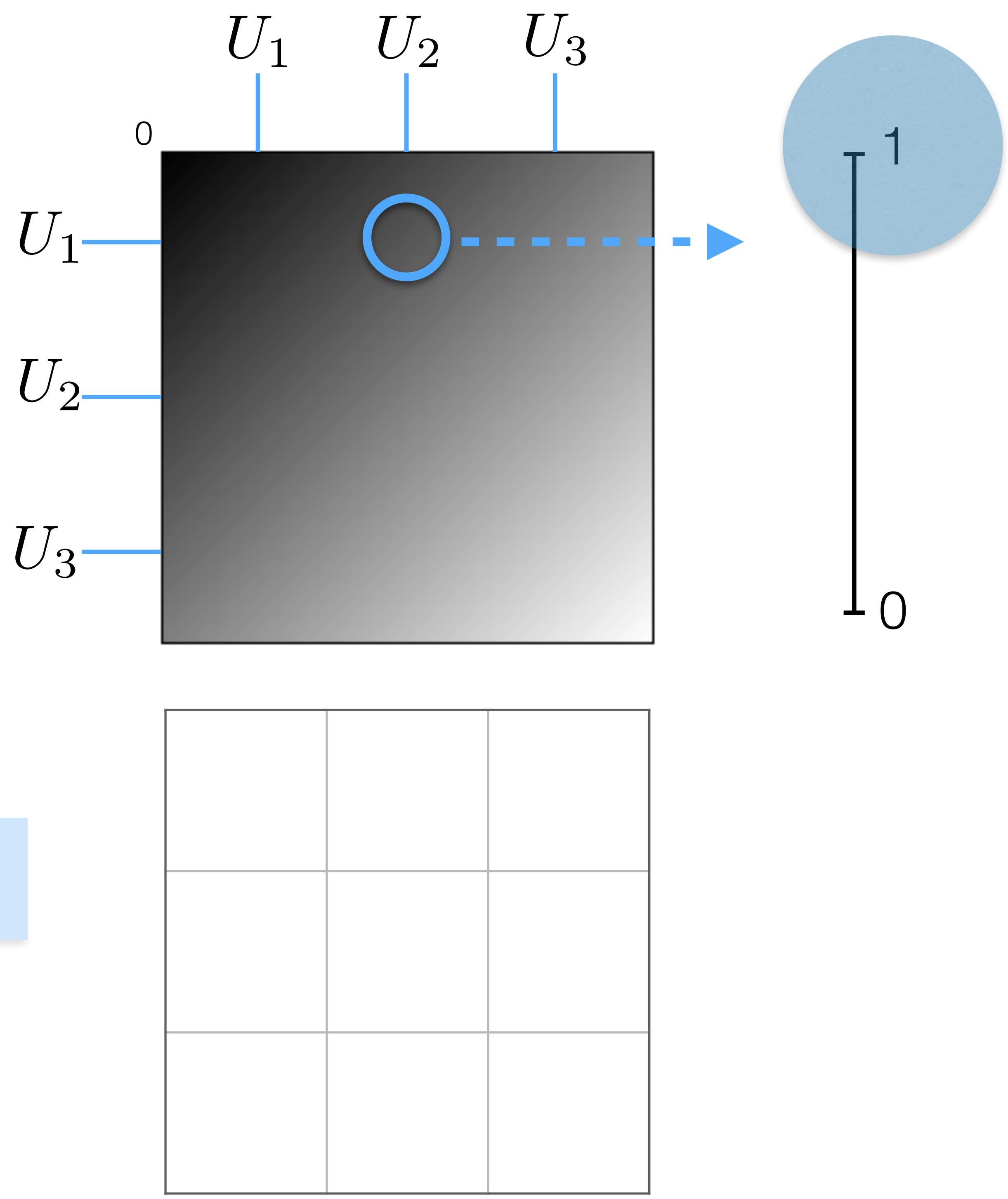
Exchangeable graphs

Graphon:

$$W : [0,1]^2 \to [0,1]$$
$$W(x,y) = W(y,x)$$

Sampling procedure:

$$U_i \sim \text{Uniform}[0,1]$$
$$G_{ij} \sim \text{Bern}(W(U_i, U_j)), i < j$$

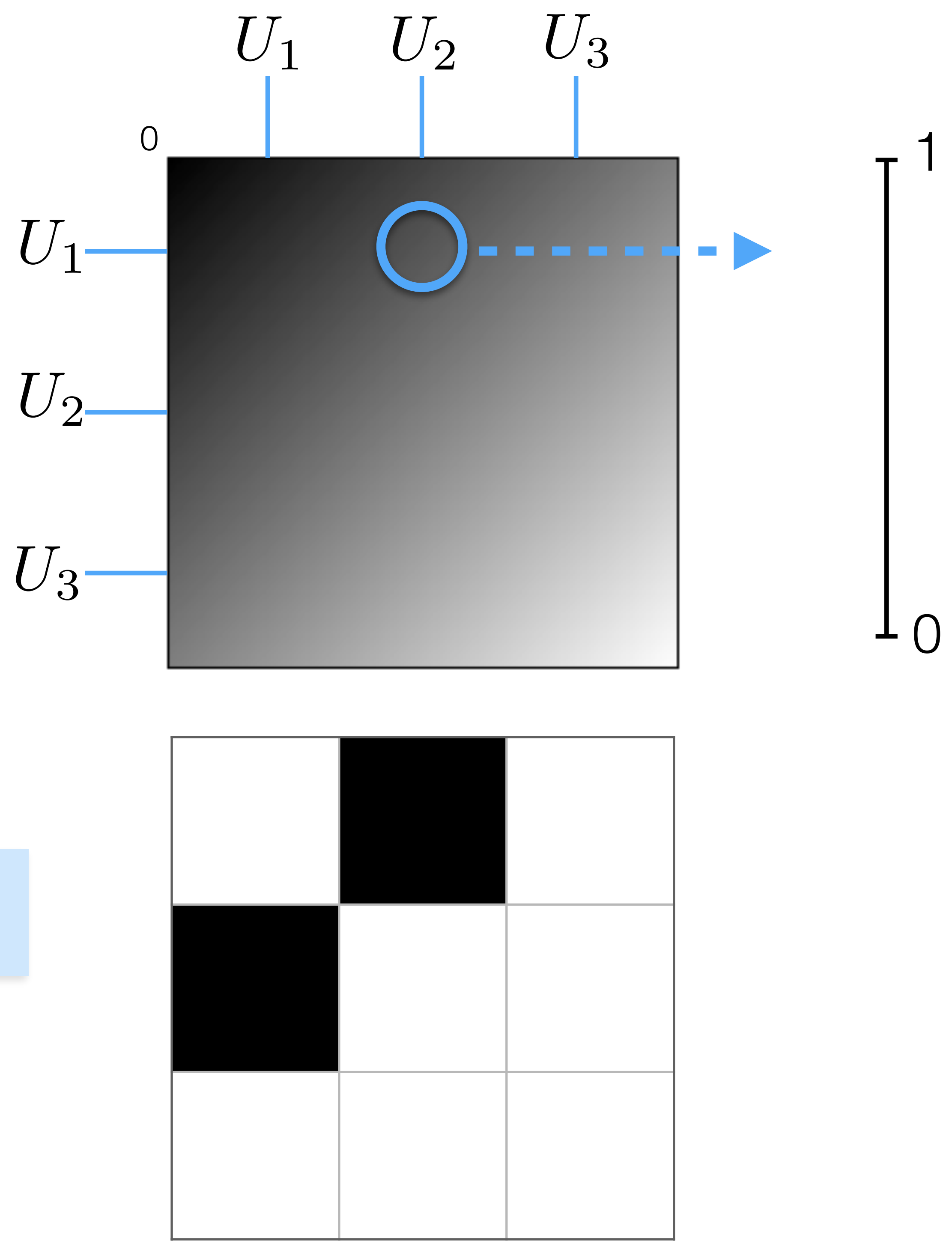


[Cai–Ackerman–Freer, 2015]

Exchangeable graphs

Graphon:

$$W : [0,1]^2 \to [0,1]$$
$$W(x,y) = W(y,x)$$



Sampling procedure:

$$U_i \sim \text{Uniform}[0,1]$$
$$G_{ij} \sim \text{Bern}(W(U_i, U_j)), i < j$$

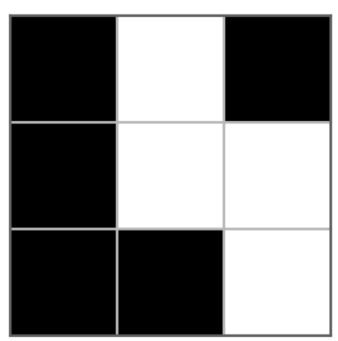
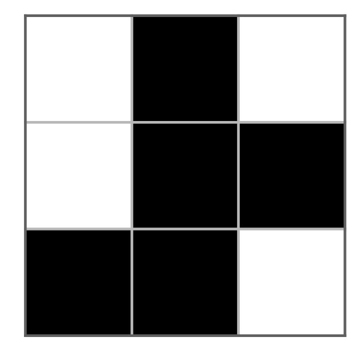
Exchangeable digraphs

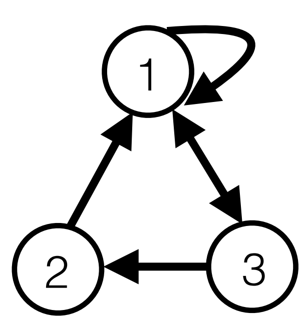
Many results for exchangeable **undirected** graphs extend to **directed** graphs.

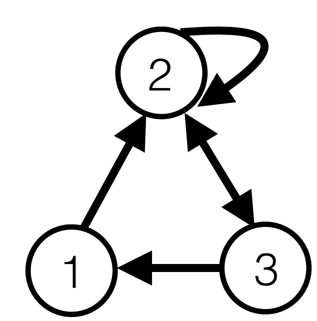
Exchangeable digraphs

Many results for exchangeable **undirected** graphs extend to **directed** graphs.

- order of vertices doesn't affect distribution of graph

$\Pr($  $) = \Pr($  $)$





Exchangeable graphs

Undirected	Directed
Any exchangeable random infinite graph is obtained as a mixture of $\mathbb{G}(\infty, W)$.	Any exchangeable random infinite **digraph** is obtained as a mixture of $\mathbb{G}(\infty, \boldsymbol{W})$.
$\mathbb{G}(\infty, W)$ is a sampling procedure from the **graphon** W.	$\mathbb{G}(\infty, \boldsymbol{W})$ is a sampling procedure from the **digraphon** $\boldsymbol{W}$.
$W : [0,1]^2 \to [0,1]$ $W(x,y) = W(y,x)$	$\boldsymbol{W} = ?$
Aldous–Hoover; cf. Lovász–Szegedy	implicit in A–H; cf. Diaconis–Janson

Exchangeable digraphs

Can we just use a single **asymmetric** measurable function?

Exchangeable digraphs

Can we just use a single **asymmetric** measurable function?

Yes, by independently choosing each edge direction.

Exchangeable digraphs

Can we just use a single **asymmetric** measurable function?

Yes, by independently choosing each edge direction.

But this does not cover *all* directed graphs:

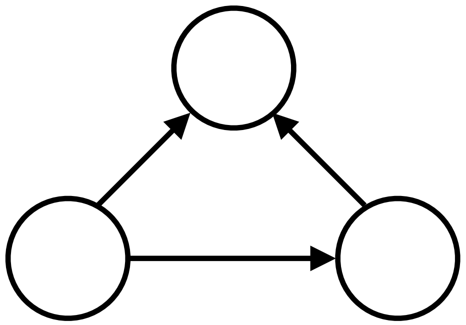
Exchangeable digraphs

Can we just use a single **asymmetric** measurable function?

Yes, by independently choosing each edge direction.

But this does not cover *all* directed graphs:

- **tournaments** (each pair of vertices has exactly one directed edge)



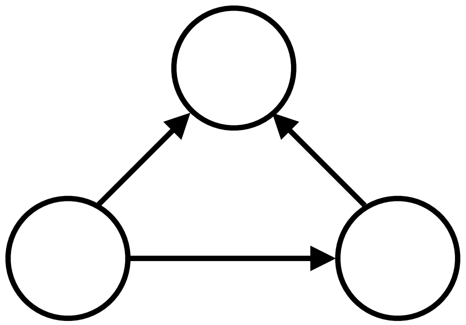
Exchangeable digraphs

Can we just use a single **asymmetric** measurable function?

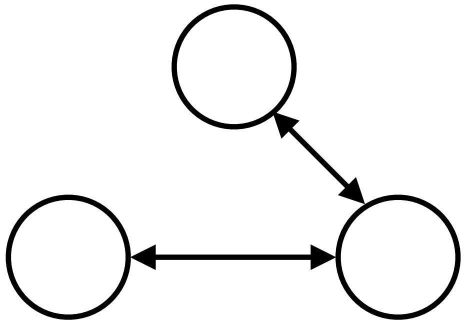
Yes, by independently choosing each edge direction.

But this does not cover *all* directed graphs:

- **tournaments** (each pair of vertices has exactly one directed edge)

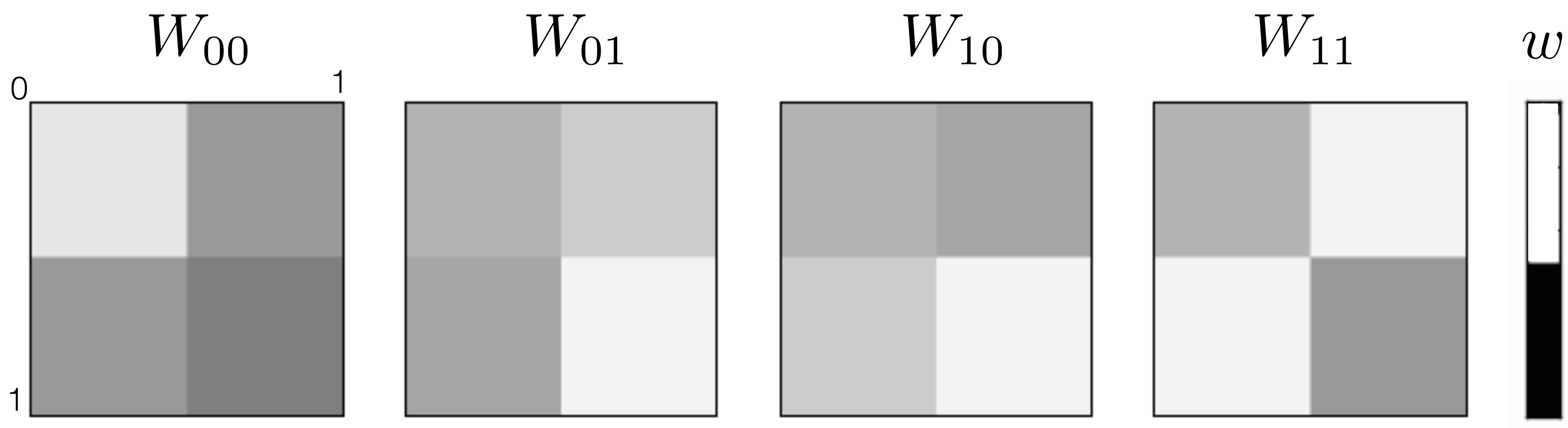


- **undirected** graphs (each pair has either both or no directed edges)



 [Cai–Ackerman–Freer, 2015]

Digraph representation



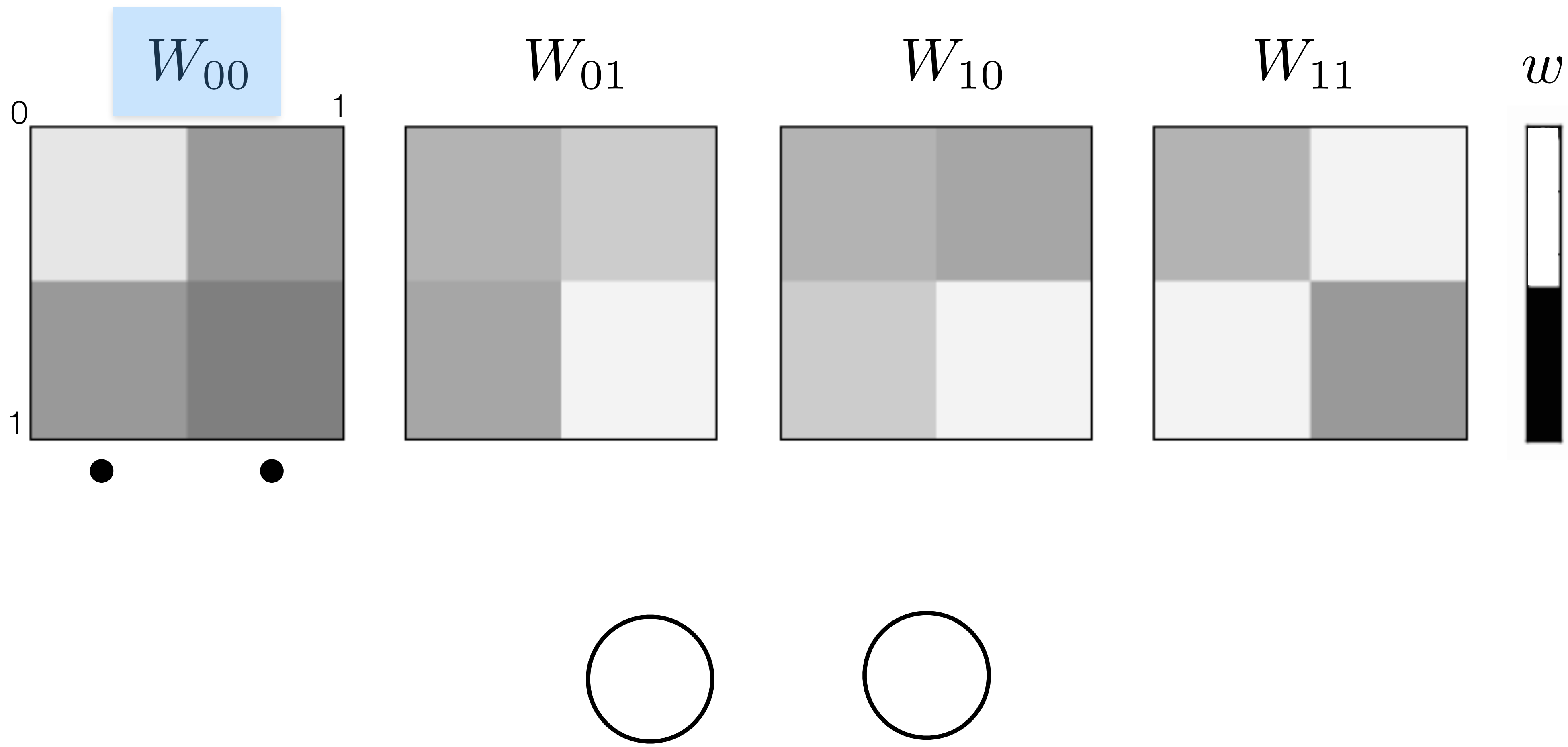
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

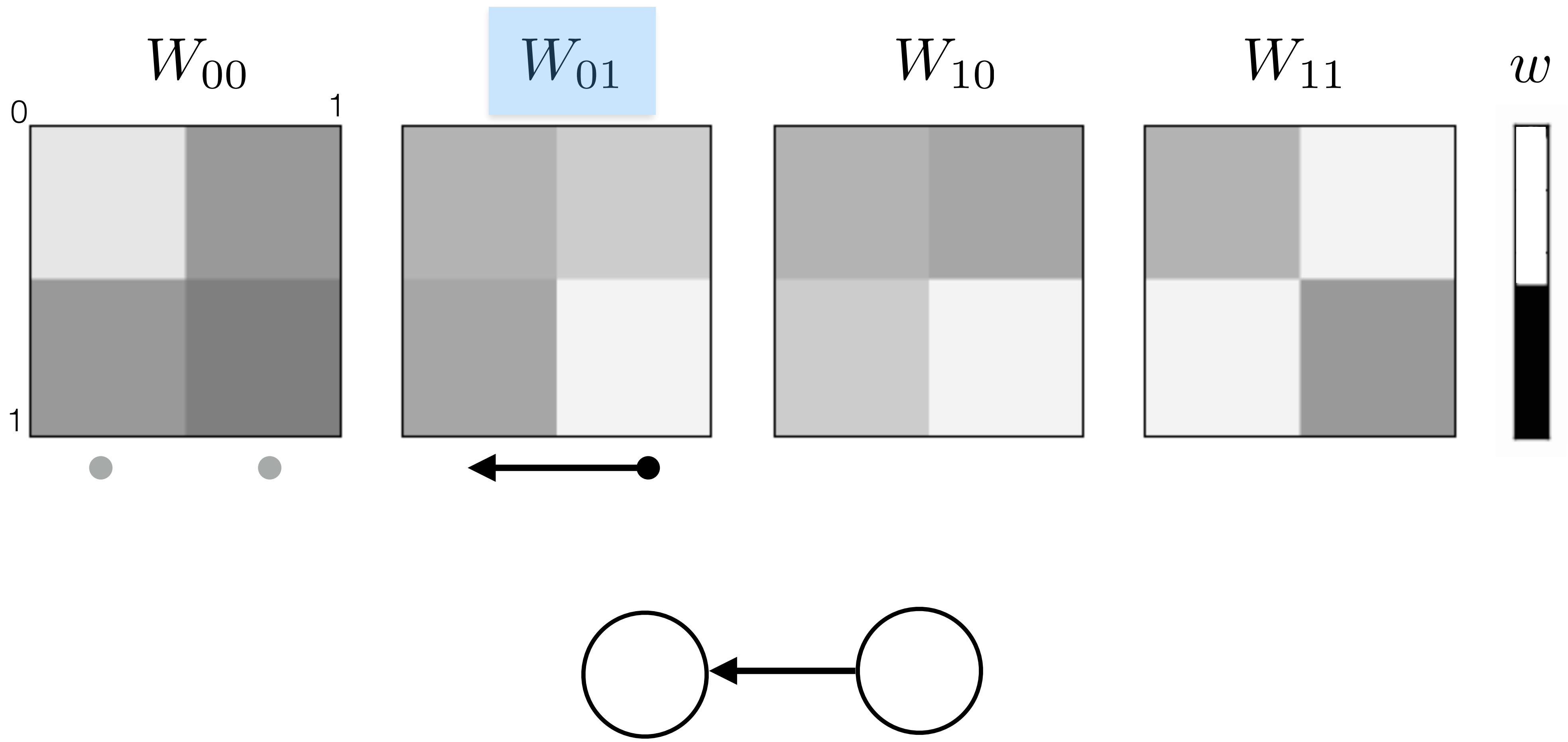
- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b$
  2. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

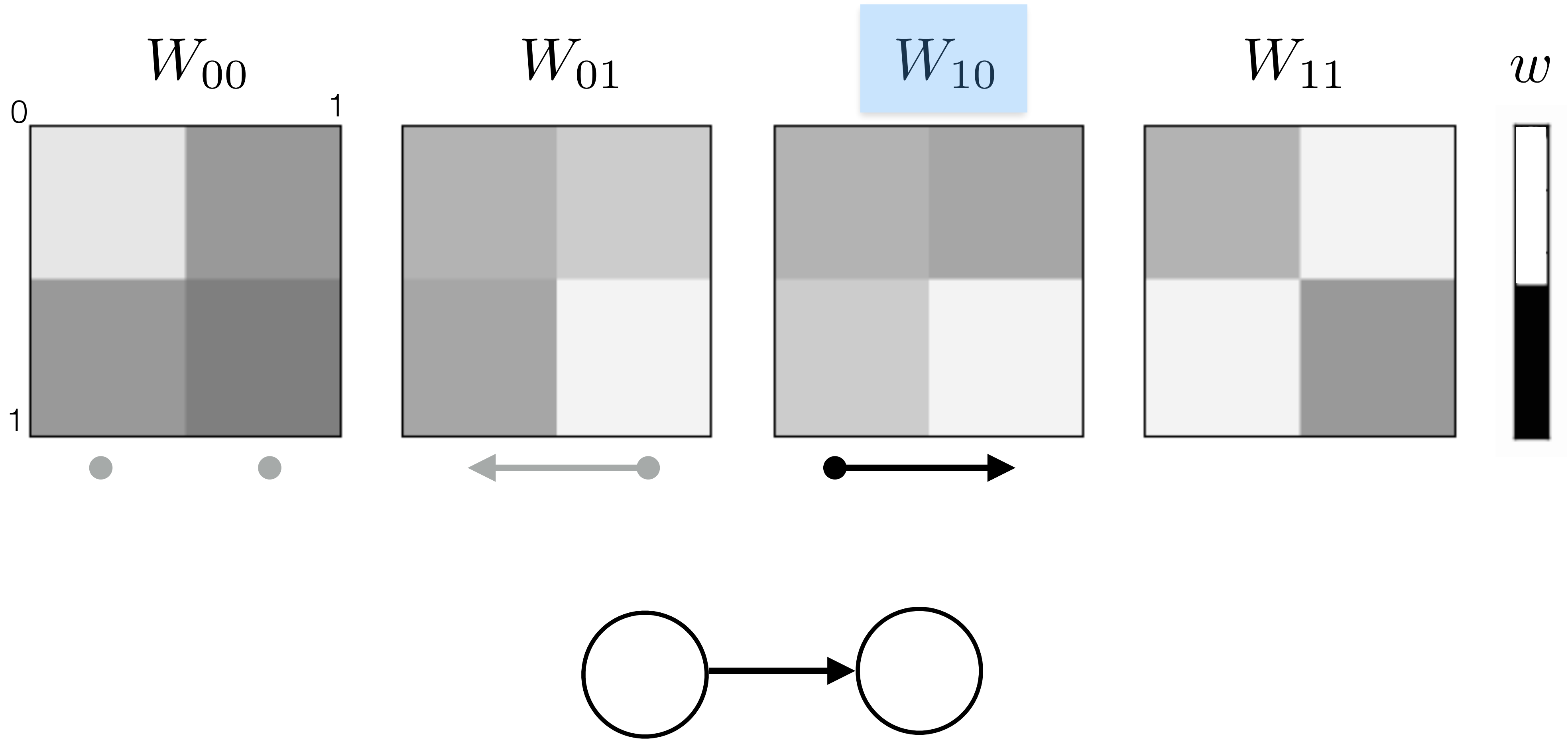
Digraph representation



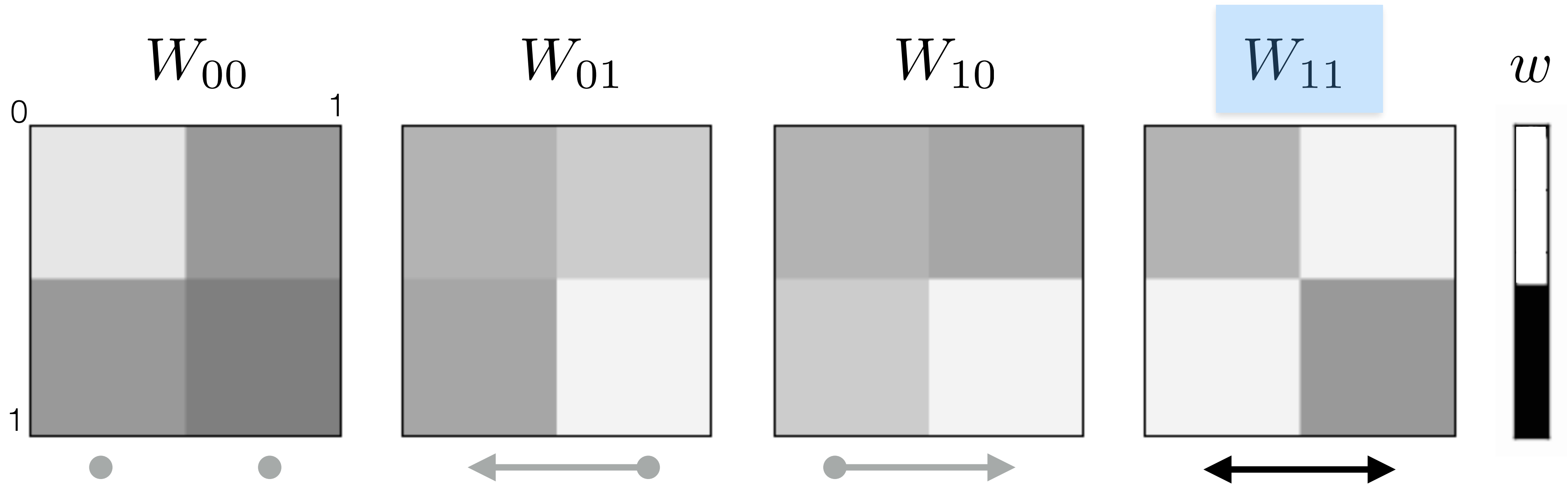
Digraph representation

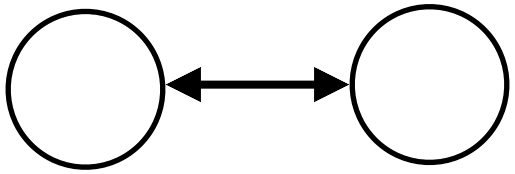


Digraph representation



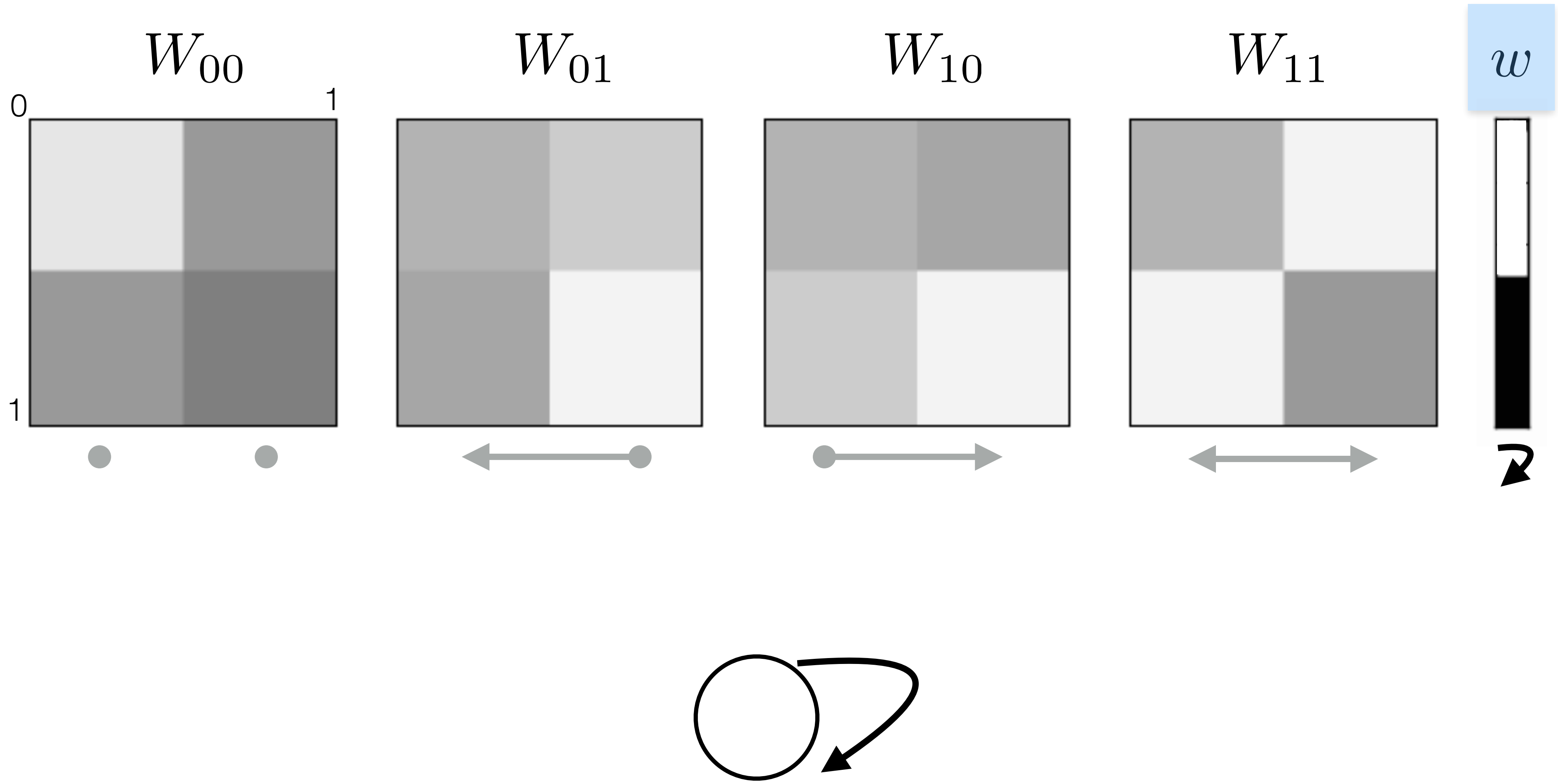
Digraph representation



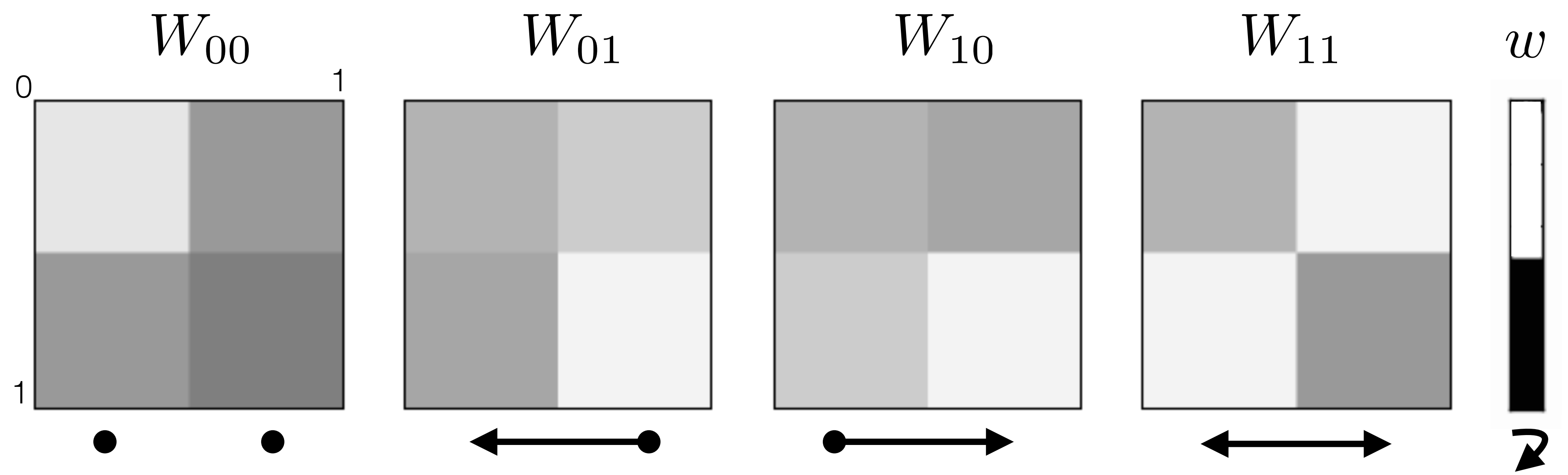


[Cai–Ackerman–Freer, 2015]

Digraph representation



Digraph representation



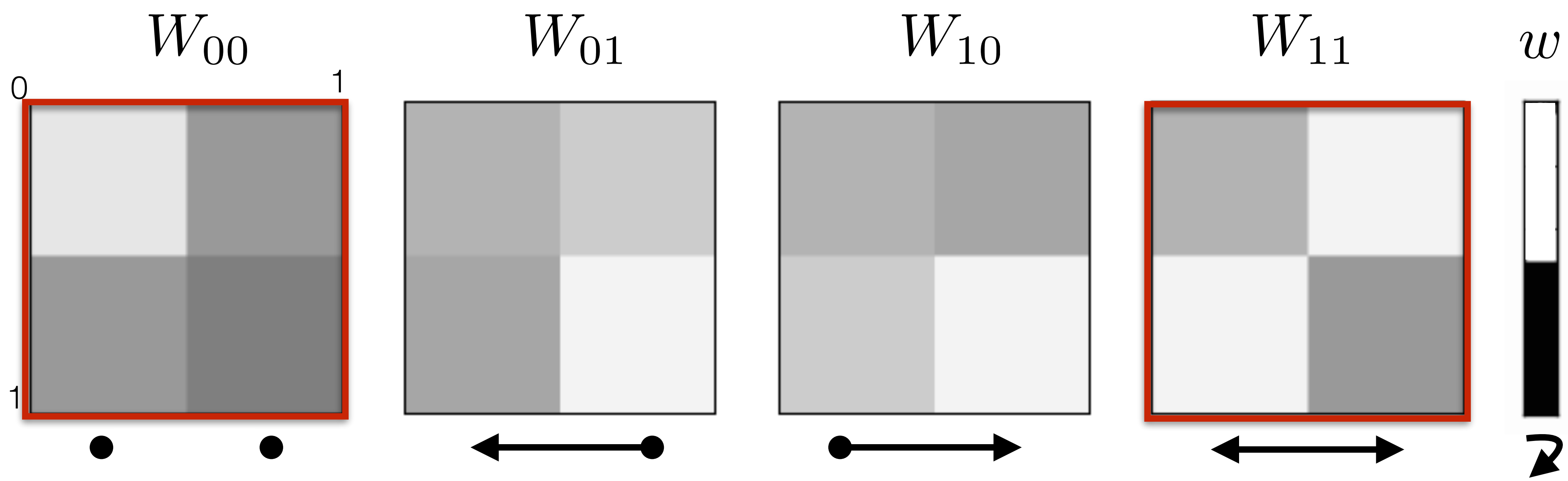
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b$
  2. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

Digraph representation



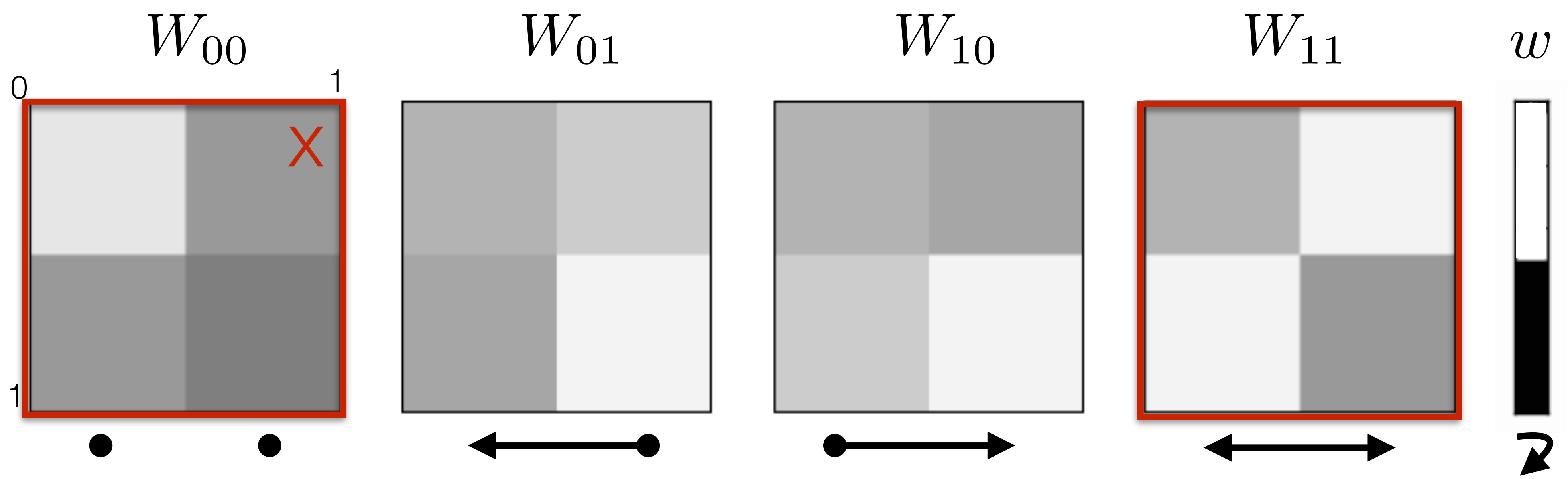
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $\boxed{W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b}$
  2. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

Digraph representation



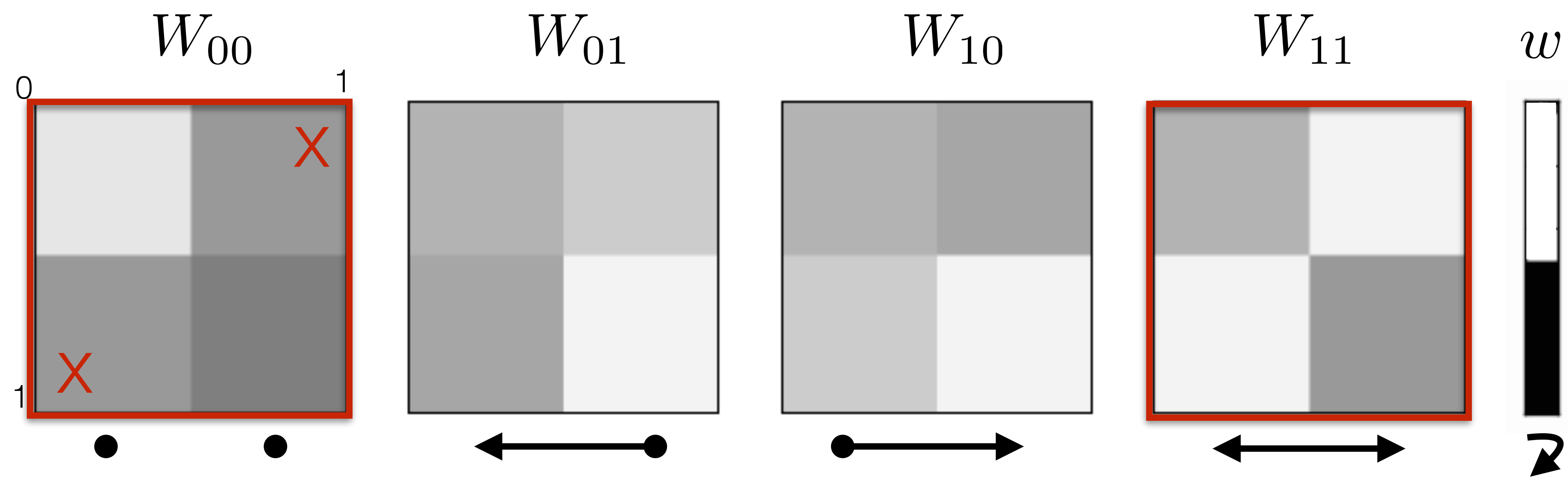
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $\boxed{W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b}$
  2. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

Digraph representation



- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

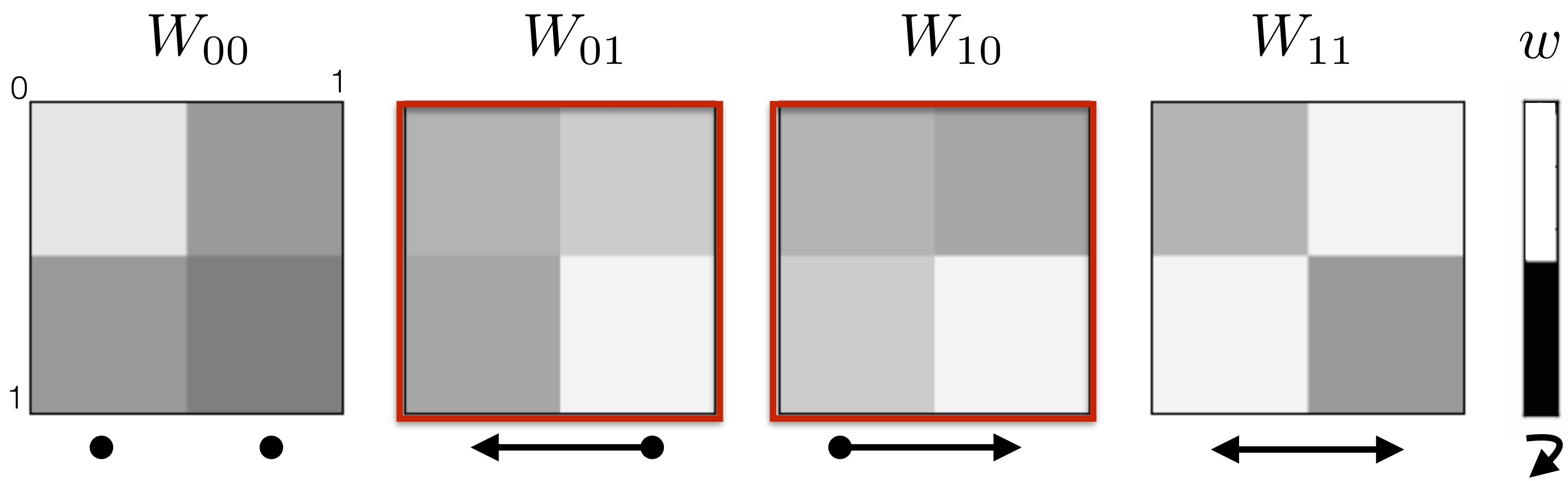
  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $\boxed{W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b}$
  2. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

[Cai–Ackerman–Freer, 2015]

Digraph representation



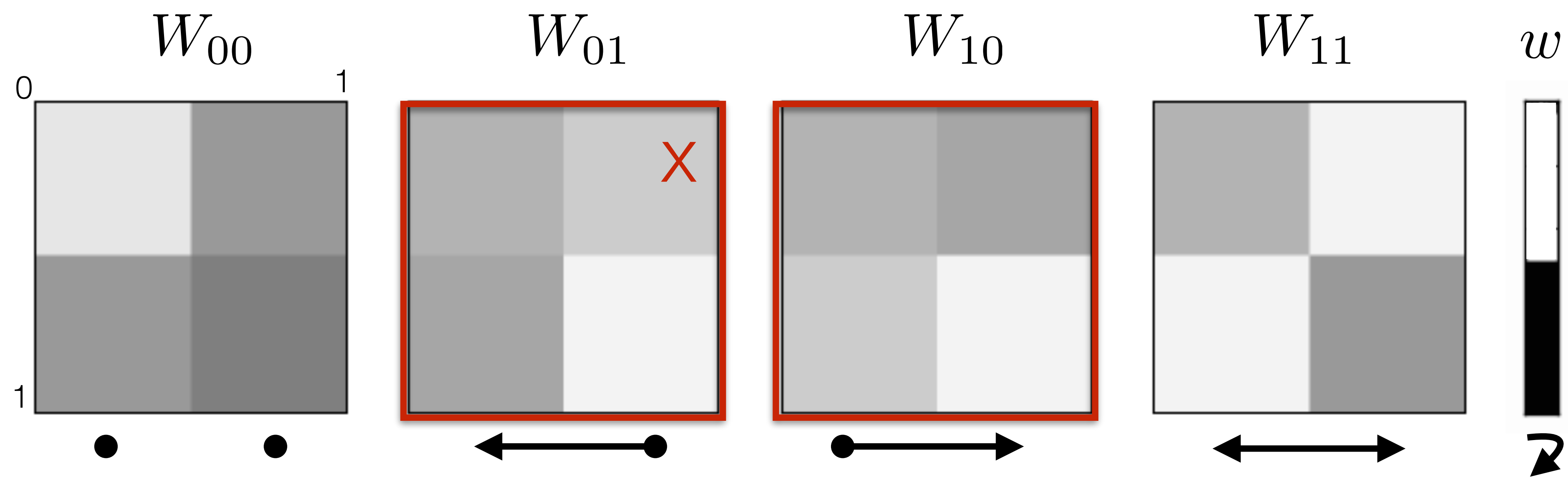
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b$
  2. $\boxed{W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b}$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

Digraph representation



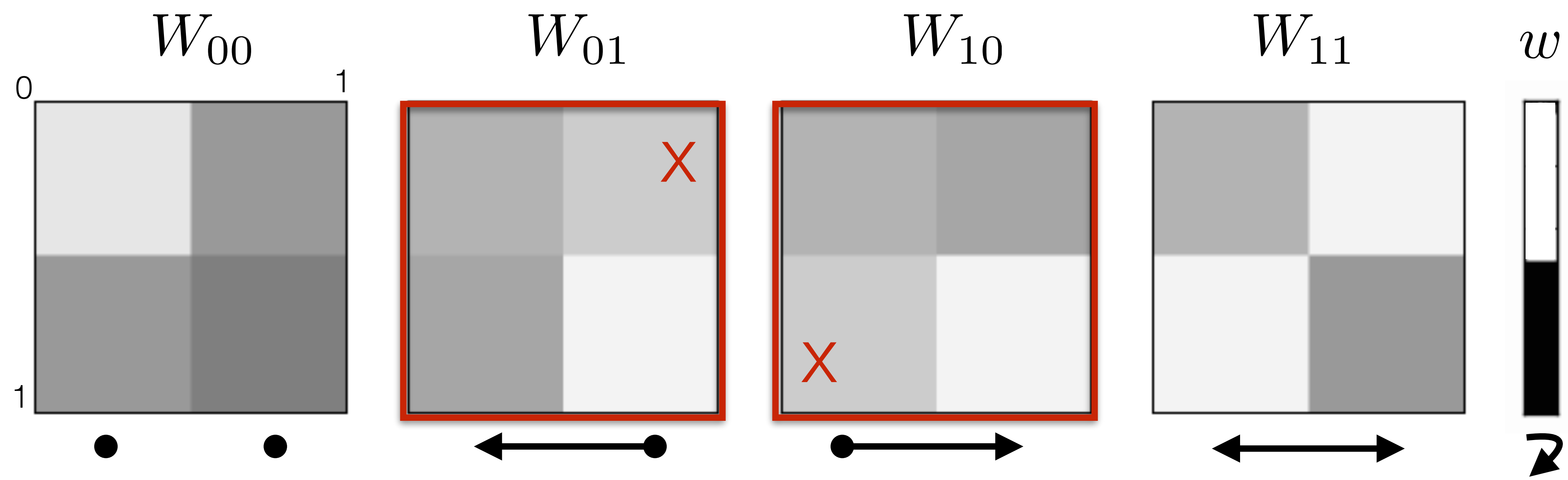
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b$
  2. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

Digraph representation



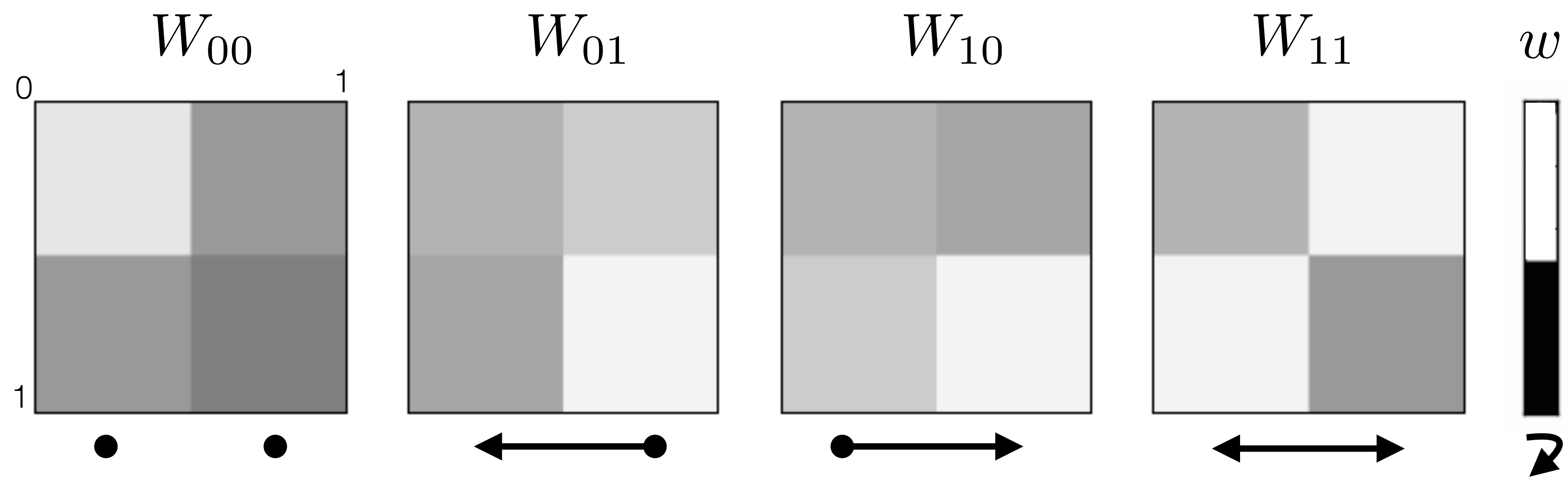
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a = b$
  2. $\boxed{W_{ab}(x,y) = W_{ba}(y,x), \text{ when } a \neq b}$
  3. $\sum_{a,b} W_{a,b}(x,y) = 1$

Digraph representation



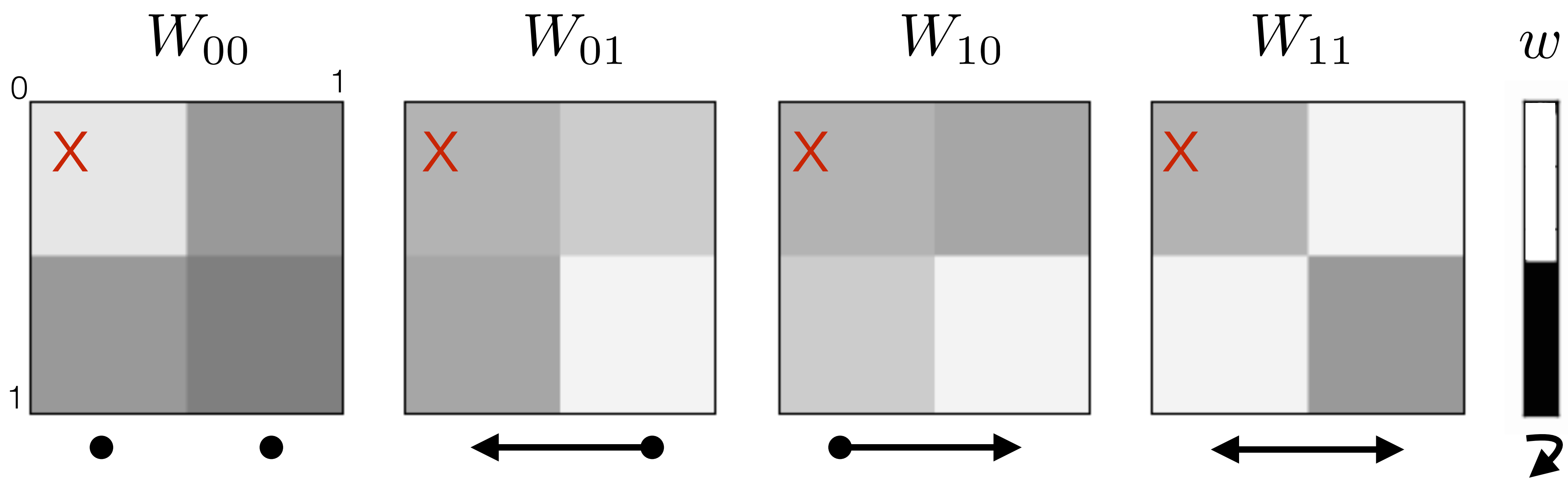
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

1. $W_{ab}(x,y) = W_{ba}(y,x), \ \text{when } a = b$
2. $W_{ab}(x,y) = W_{ba}(y,x), \ \text{when } a \neq b$
3. $\boxed{\sum_{a,b} W_{a,b}(x,y) = 1}$

Digraph representation



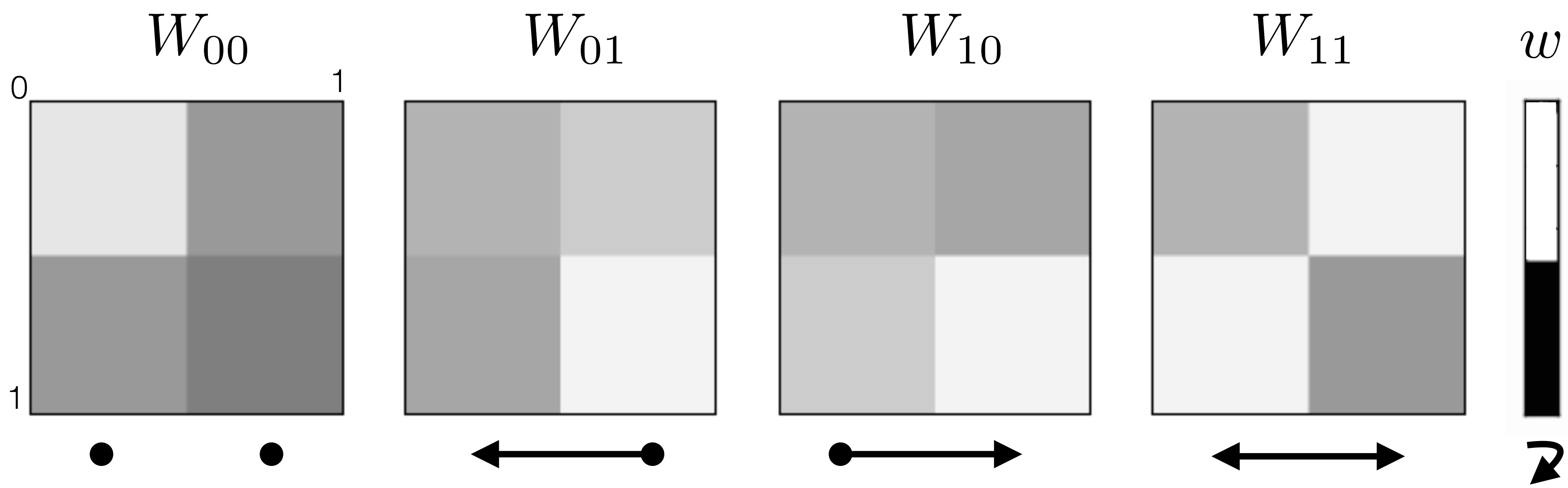
- Digraphon: $\boldsymbol{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$

  $W_{ab} : [0,1]^2 \to [0,1]$ and $w : [0,1] \to \{0,1\}$ measurable functions

- For $a, b \in \{0,1\}$ and $x, y \in [0,1]$,

  1. $W_{ab}(x,y) = W_{ba}(y,x), \;$ when $a = b$
  2. $W_{ab}(x,y) = W_{ba}(y,x), \;$ when $a \neq b$
  3. $\boxed{\sum_{a,b} W_{a,b}(x,y) = 1}$

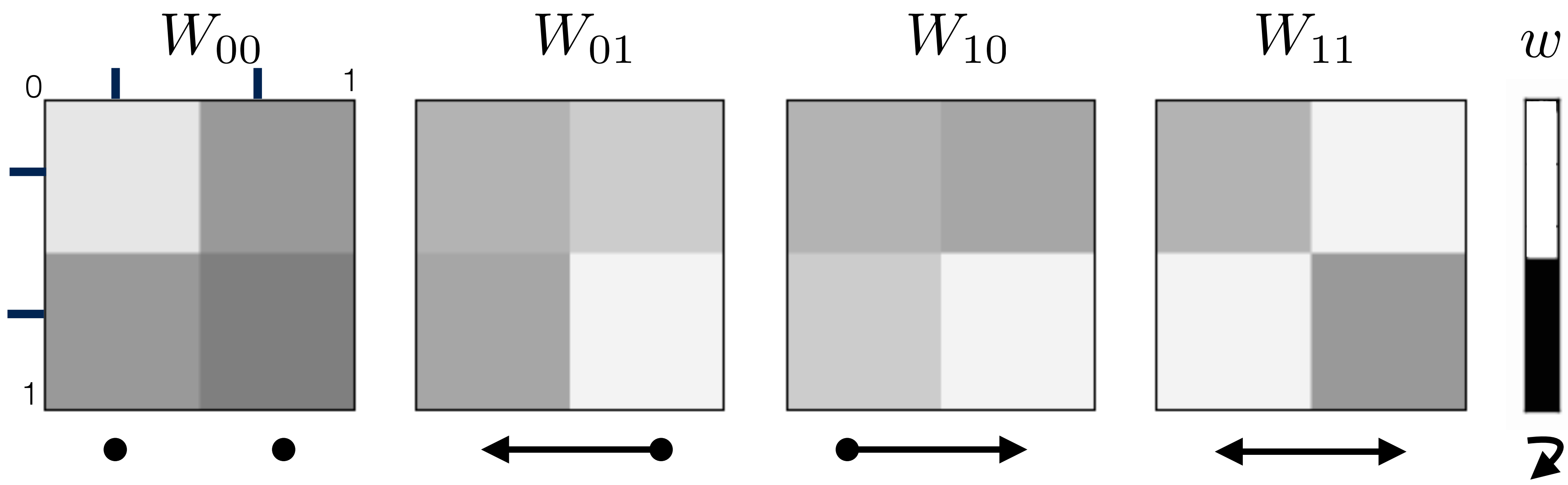
Sampling procedure



- Define $\boldsymbol{W}_4(x,y) := (W_{00}(x,y), W_{01}(x,y), W_{10}(x,y), W_{11}(x,y))$

$$U_i \sim \mathrm{Uniform}[0,1]$$
$$E_{ij} \sim \mathrm{Categorical}(\boldsymbol{W}_4(U_i,U_j)),\ \text{for } i<j$$
$$G_{ij}, G_{ji} = \text{directed edge(s) corresponding to } E_{ij}, \text{ for } i<j$$
$$G_{ii} = w(U_i)$$

Sampling procedure



- Define $\boldsymbol{W}_4(x,y) := (W_{00}(x,y), W_{01}(x,y), W_{10}(x,y), W_{11}(x,y))$

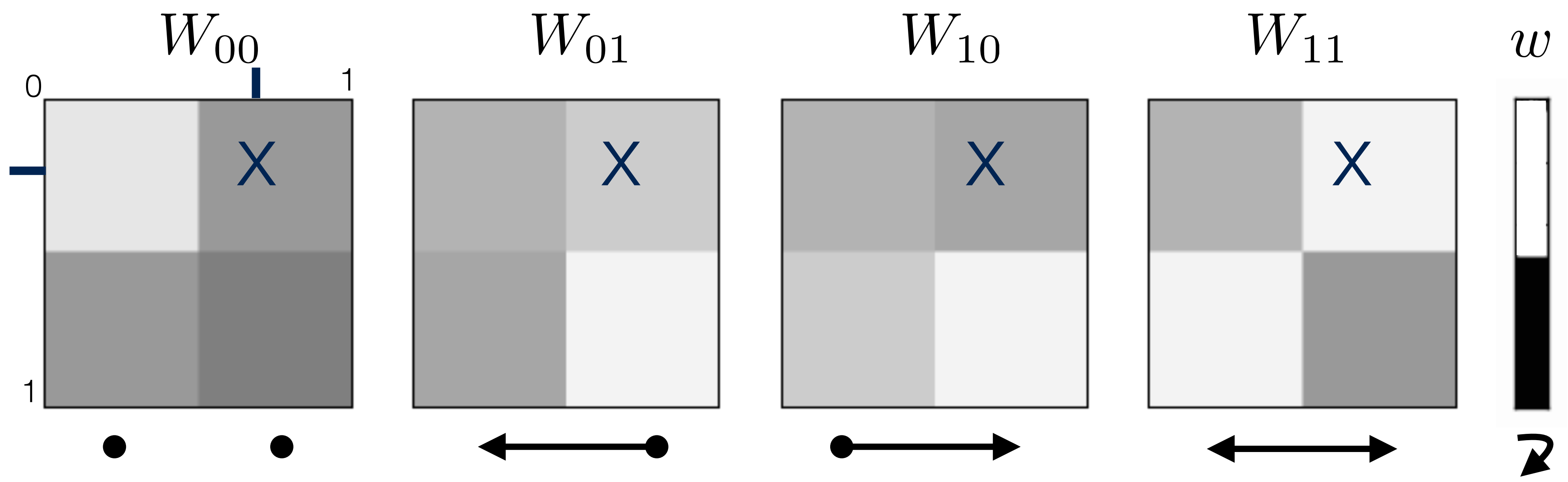
$$U_i \sim \text{Uniform}[0,1]$$

$$E_{ij} \sim \text{Categorical}(\boldsymbol{W}_4(U_i,U_j)), \text{ for } i<j$$

$$G_{ij}, G_{ji} = \text{directed edge(s) corresponding to } E_{ij}, \text{ for } i<j$$

$$G_{ii} = w(U_i)$$

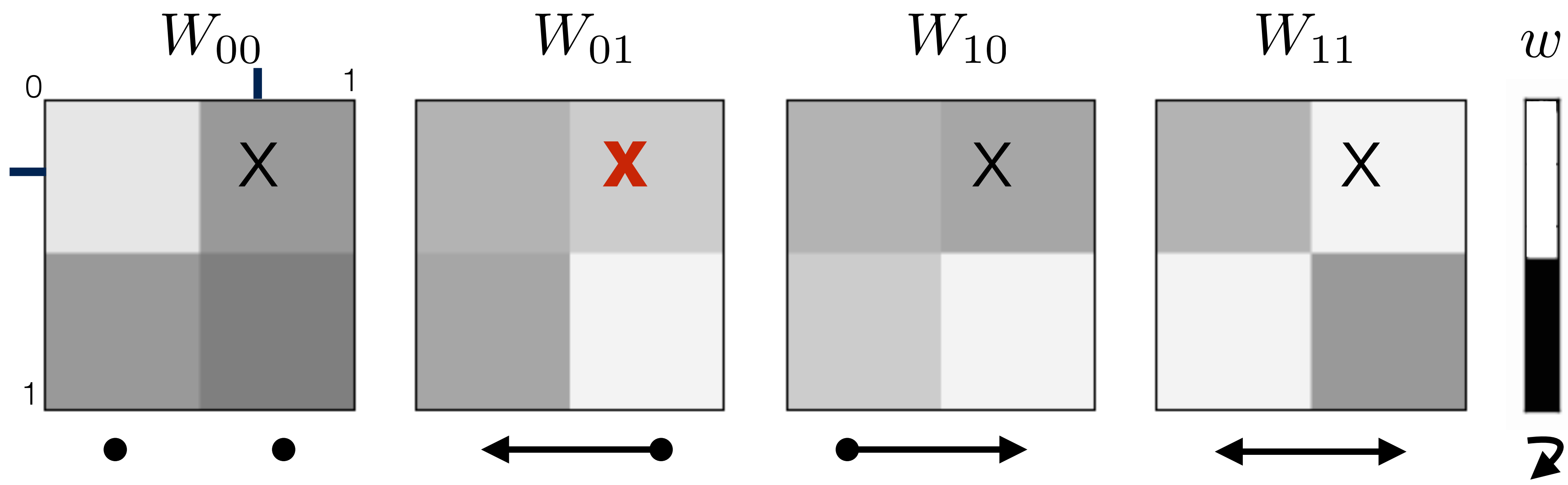
Sampling procedure



- Define $\boldsymbol{W}_4(x,y) := (W_{00}(x,y), W_{01}(x,y), W_{10}(x,y), W_{11}(x,y))$

$$
\begin{aligned}
U_i &\sim \mathrm{Uniform}[0,1] \\
E_{ij} &\sim \mathrm{Categorical}(\boldsymbol{W}_4(U_i,U_j)), \text{ for } i<j \\
G_{ij}, G_{ji} &= \text{directed edge(s) corresponding to } E_{ij}, \text{ for } i<j \\
G_{ii} &= w(U_i)
\end{aligned}
$$

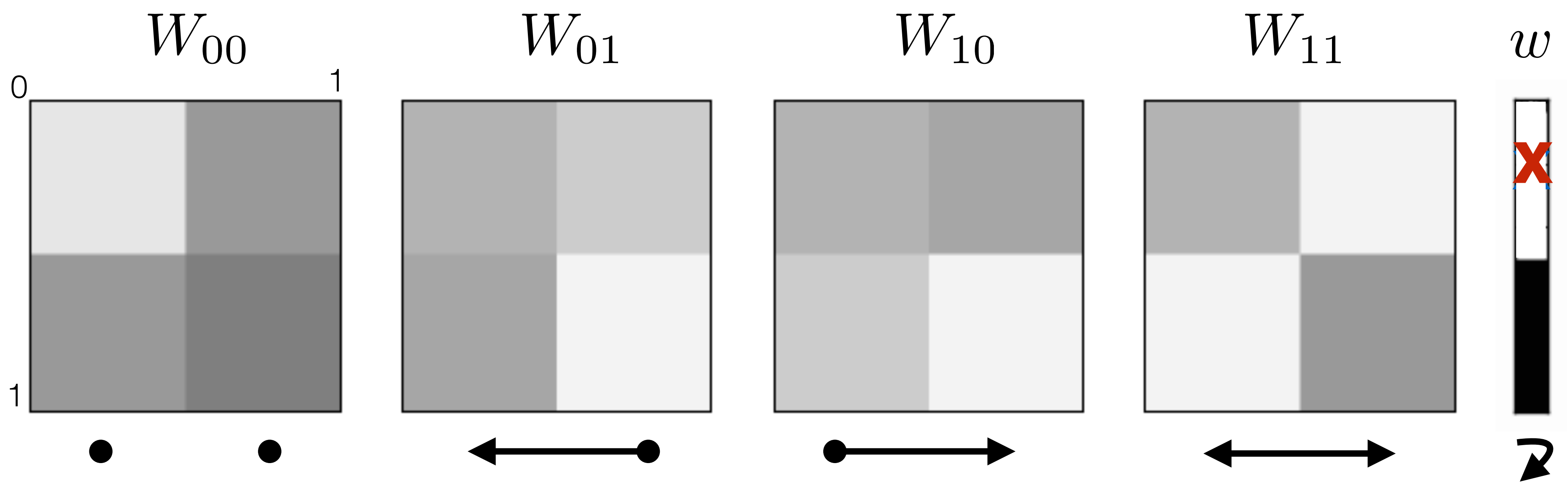
Sampling procedure



- Define $\boldsymbol{W}_4(x,y) := (W_{00}(x,y), W_{01}(x,y), W_{10}(x,y), W_{11}(x,y))$

$$U_i \sim \text{Uniform}[0,1]$$
$$E_{ij} \sim \text{Categorical}(\boldsymbol{W}_4(U_i, U_j)), \text{ for } i<j$$
$$G_{ij}, G_{ji} = \text{directed edge(s) corresponding to } E_{ij}, \text{ for } i<j$$
$$G_{ii} = w(U_i)$$

Sampling procedure

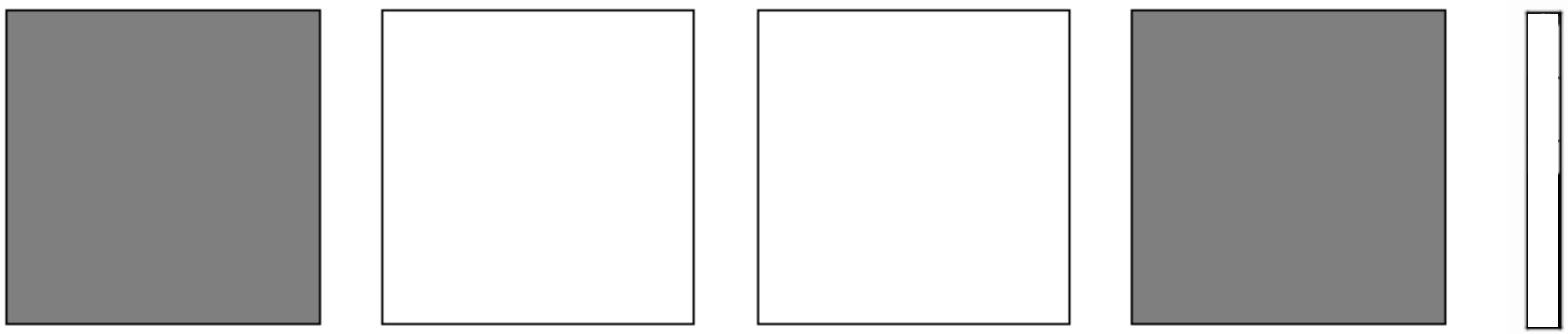


- Define $\boldsymbol{W}_4(x,y) := (W_{00}(x,y), W_{01}(x,y), W_{10}(x,y), W_{11}(x,y))$

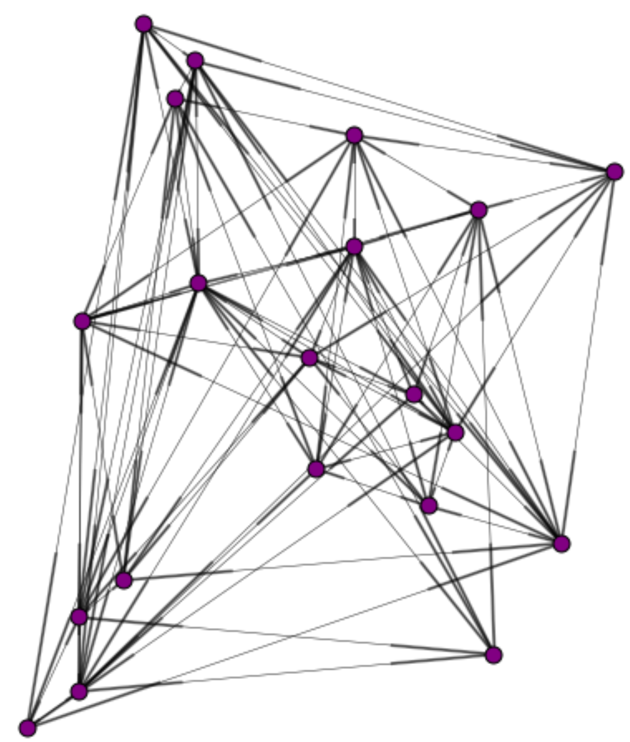
$$U_i \sim \mathrm{Uniform}[0,1]$$
$$E_{ij} \sim \mathrm{Categorical}(\boldsymbol{W}_4(U_i,U_j)),\ \text{for } i<j$$
$$G_{ij}, G_{ji} = \text{directed edge(s) corresponding to } E_{ij},\ \text{for } i<j$$
$$G_{ii} = w(U_i)$$

Special cases

- Undirected graphs (graphon) (e.g., ER(½)) $W : [0,1]^2 \to [0,1]$ $W(x,y) = W(y,x)$

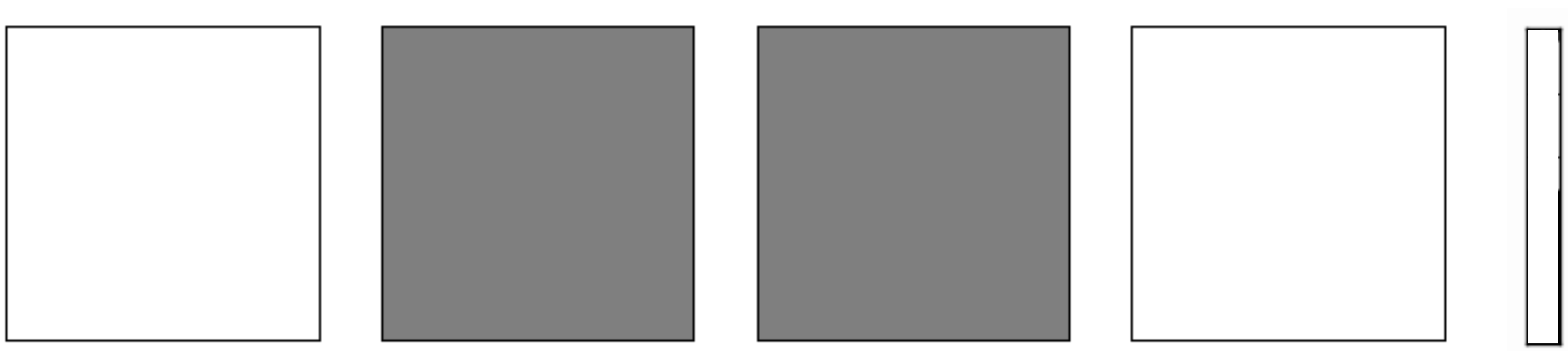


W_{00} W_{01} W_{10} W_{11} w

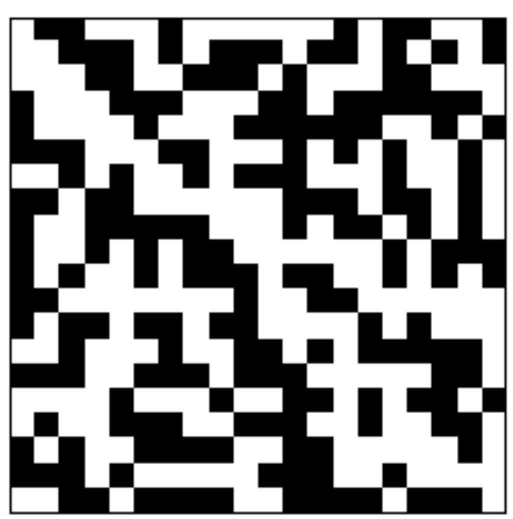


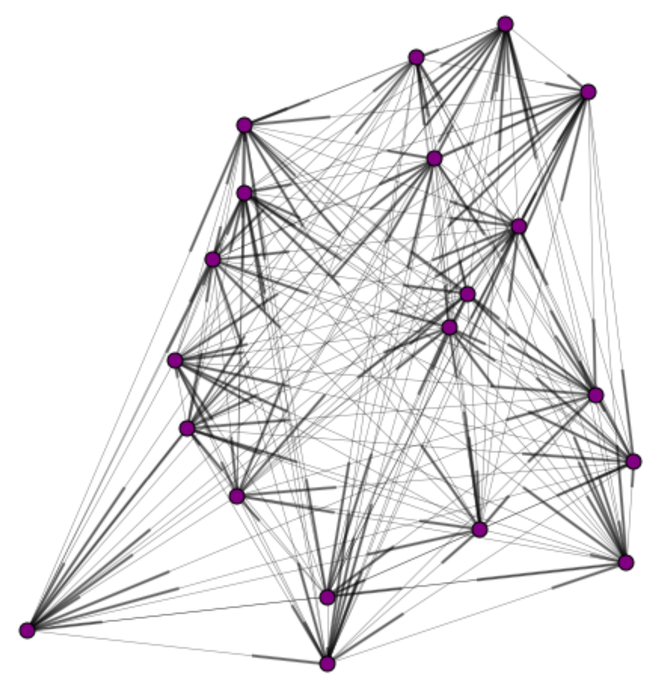
Special cases

- Tournaments (e.g., generic tournament)



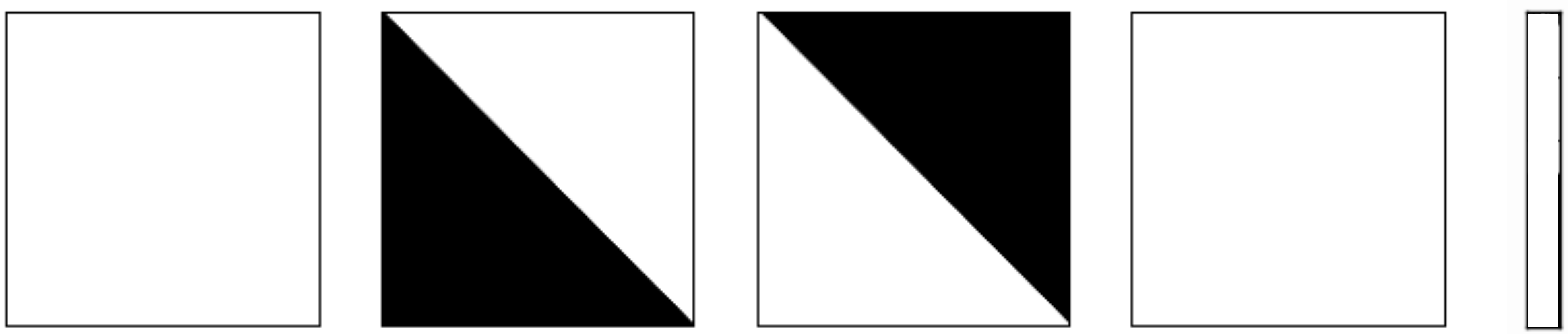
W_{00} W_{01} W_{10} W_{11} w



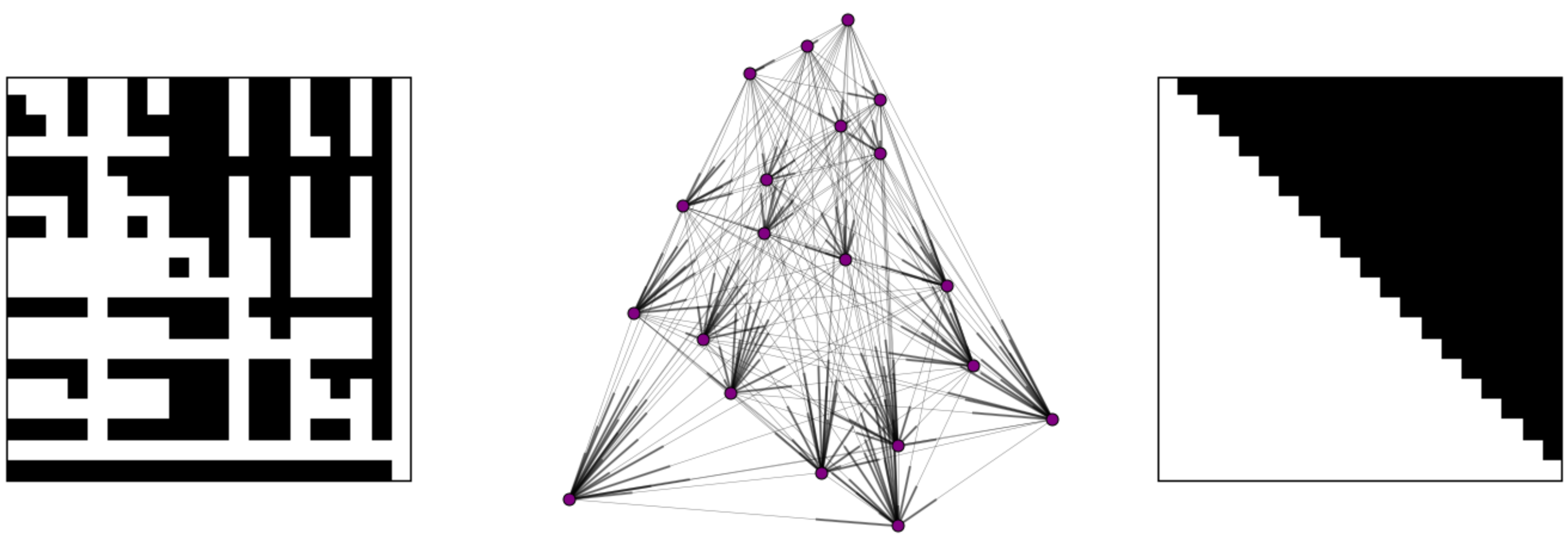


Special cases

- Linear ordering (the only one, by Glasner–Weiss)

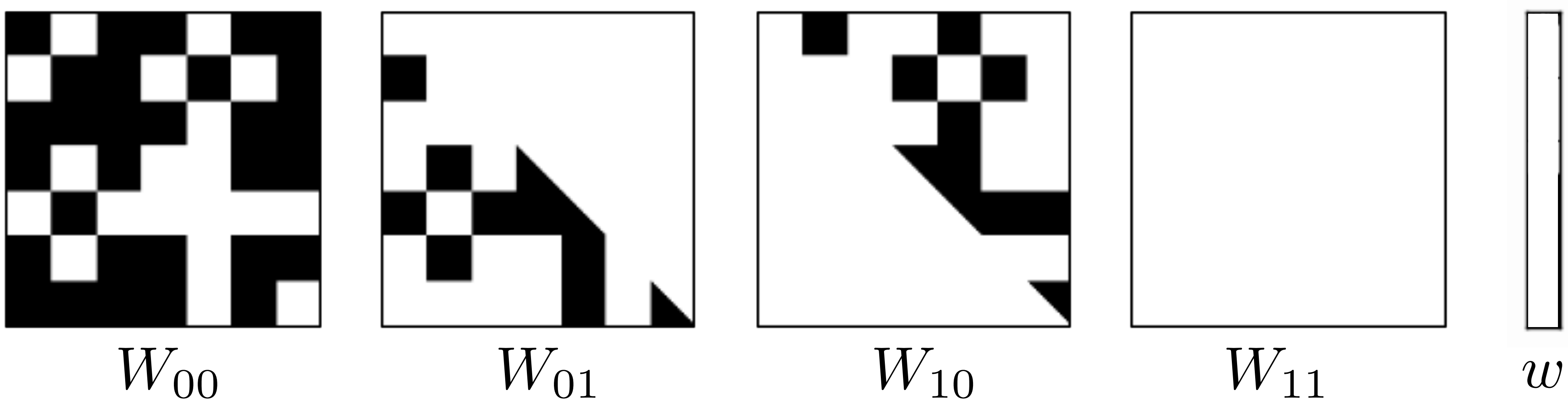


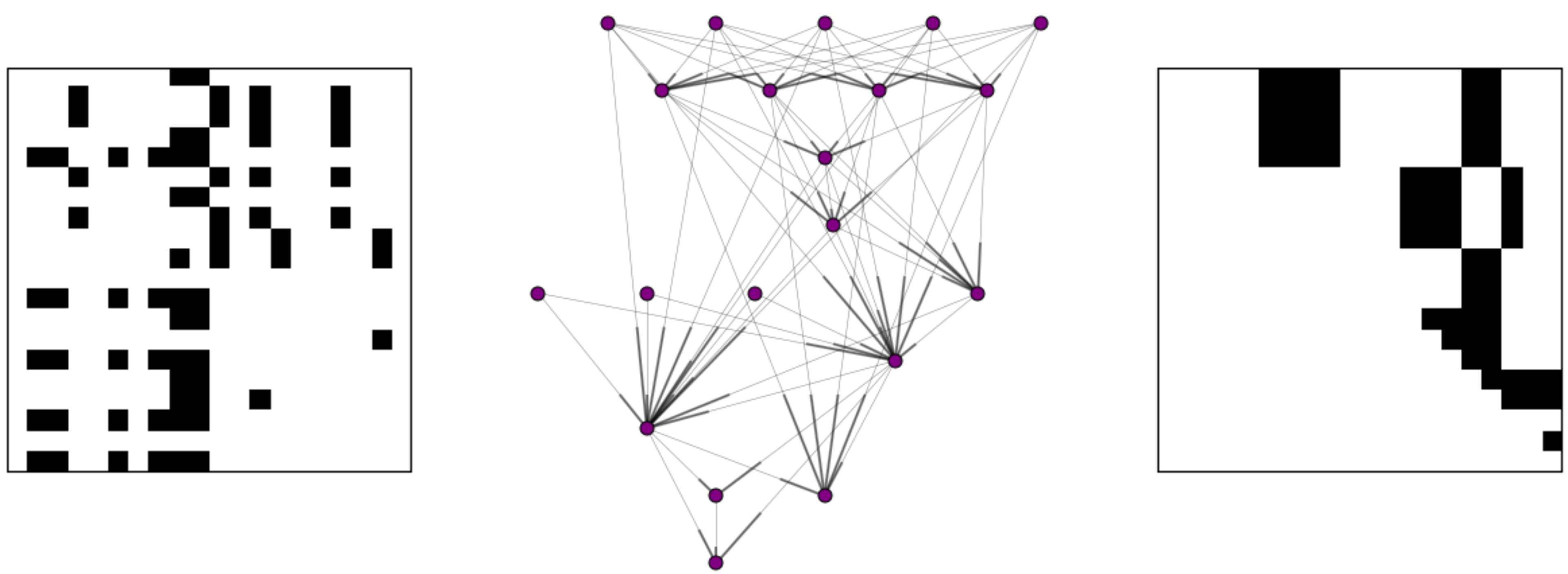
W_{00} W_{01} W_{10} W_{11} w



Special cases

- Directed acyclic graphs (DAGs)

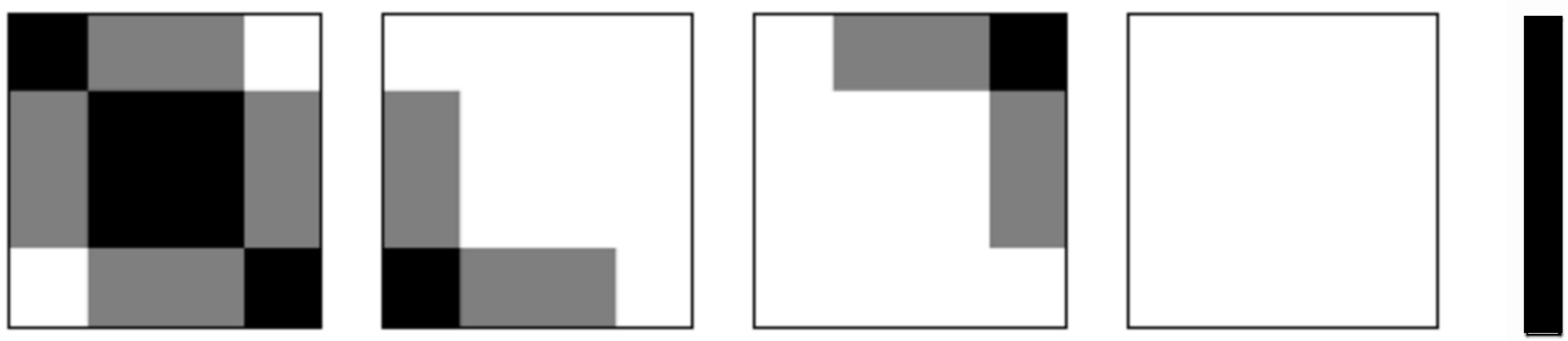




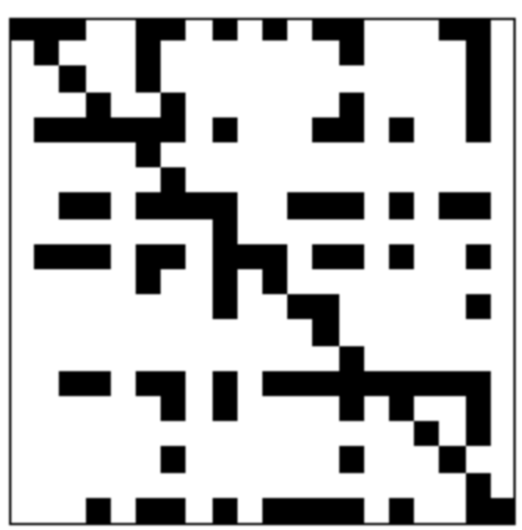
[Cai–Ackerman–Freer, 2015]

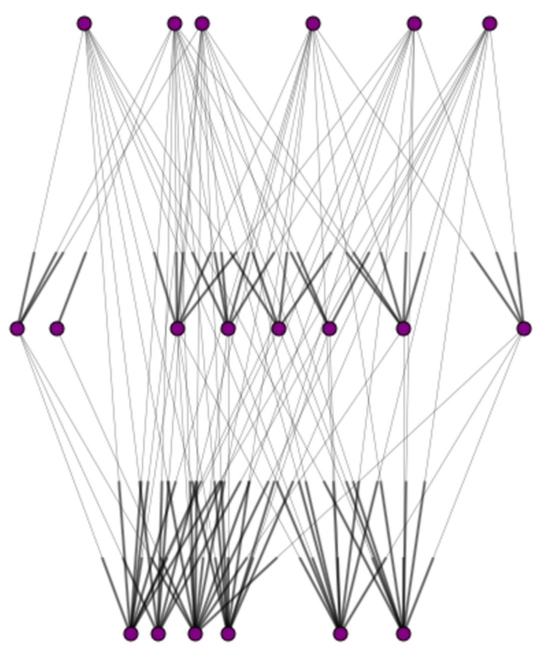
Special cases

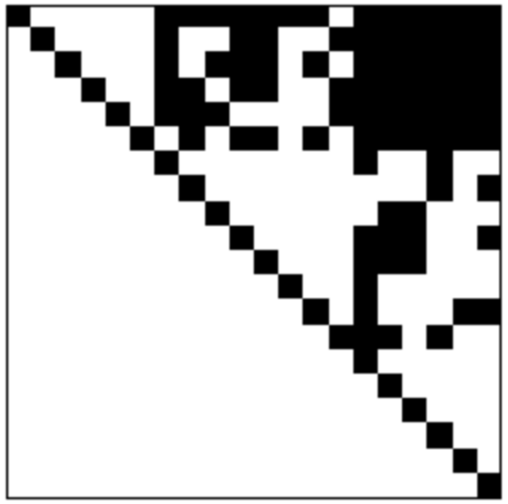
- Partial ordering (poset)



W_{00} W_{01} W_{10} W_{11} w

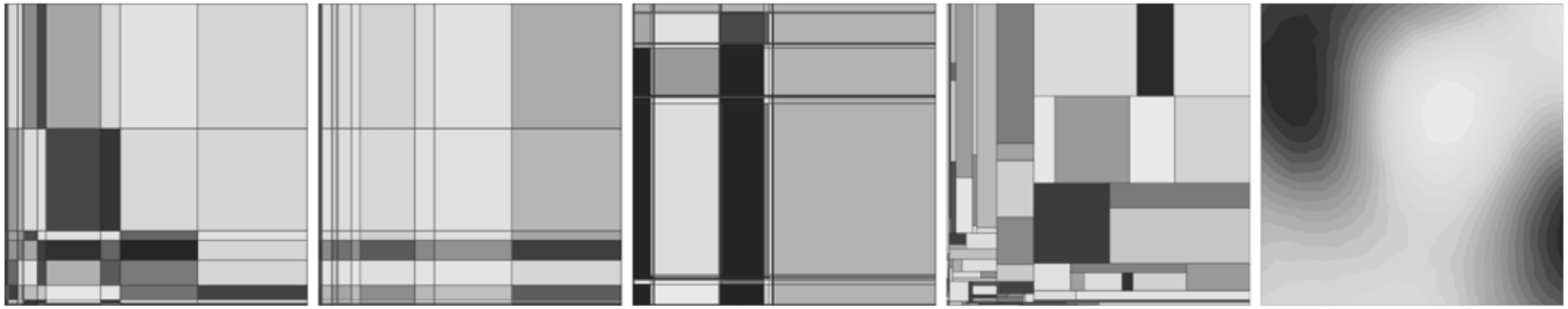






Priors on digraphons

- Can extend literature for graphon priors (cf. Orbanz–Roy) to directed graphs, e.g., infinite relational model



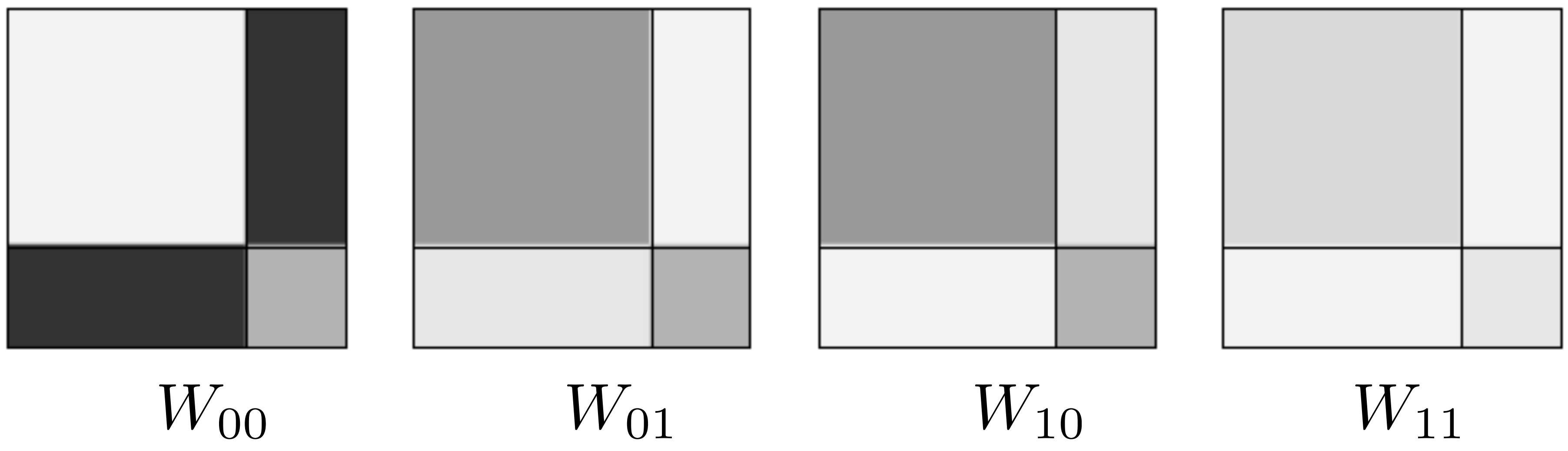
[Orbanz–Roy, 2015]

- Some of these models are already intended for directed graphs, via an **asymmetric** measurable function (to describe independent edge directions).

Directed block models

- In a block model, pairs of regions can vary in how tournament-like, as well as how dense they are:
  - e.g., Directed stochastic block model [Wang–Wong, 1987]

Example of SBM digraphon, 0.7 division

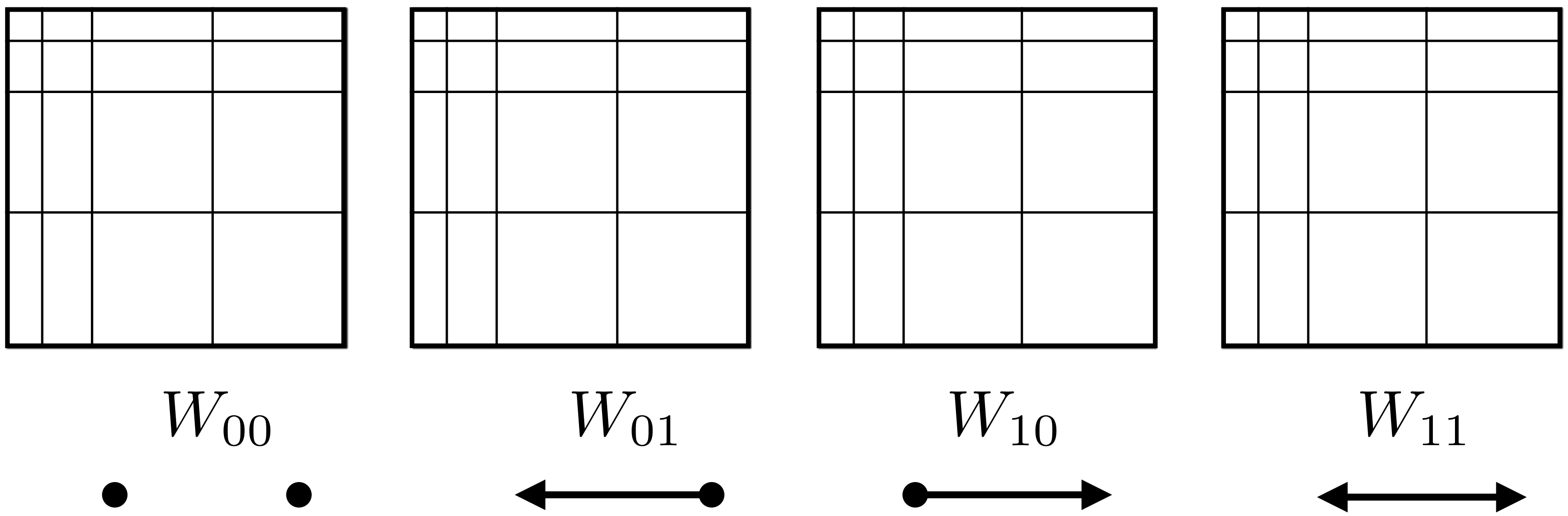


[Cai–Ackerman–Freer, 2015]

Infinite digraph block model

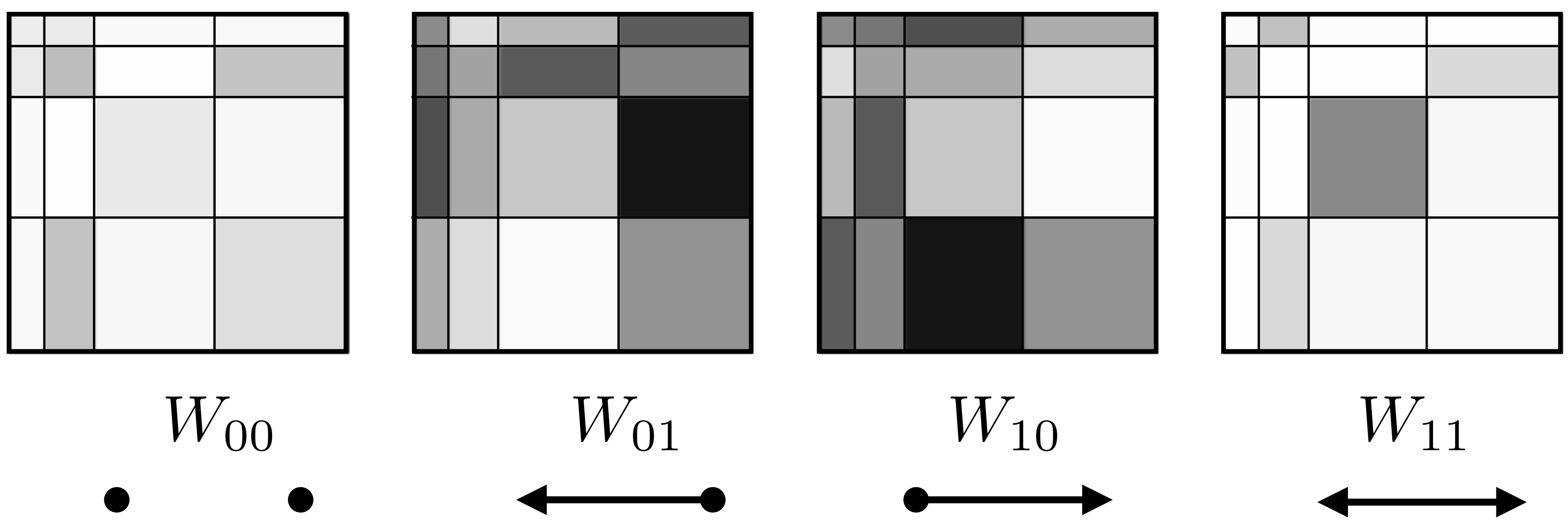
W_{00} $\quad$ W_{01} $\quad$ W_{10} $\quad$ W_{11}

Infinite digraph block model



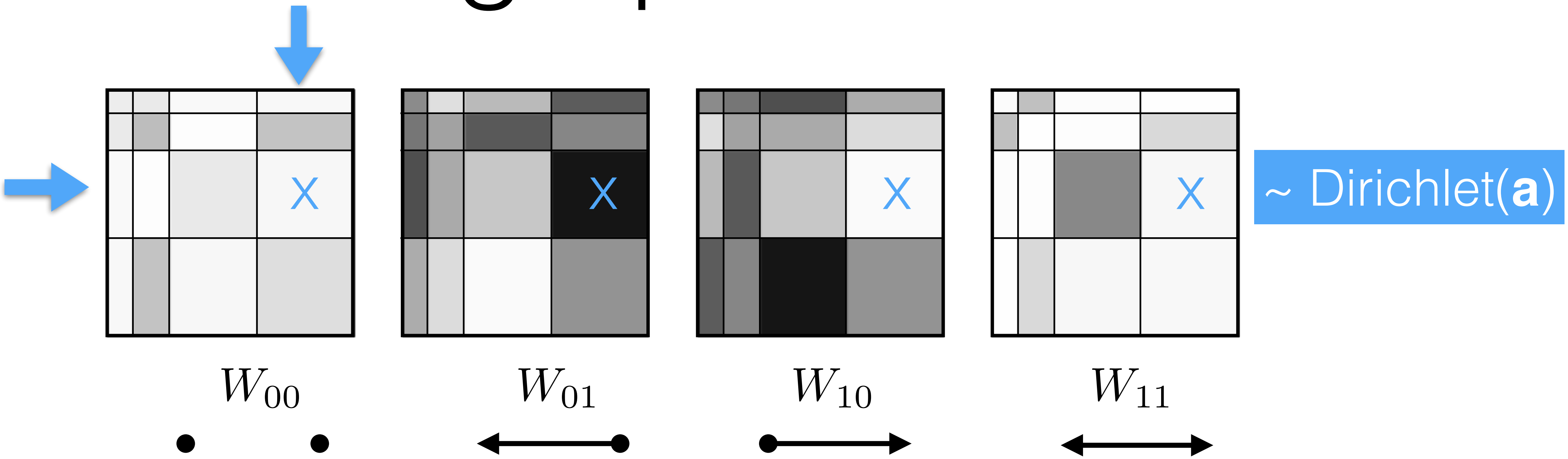
1. Draw partition ~ DP-Stick(alpha)

Infinite digraph block model



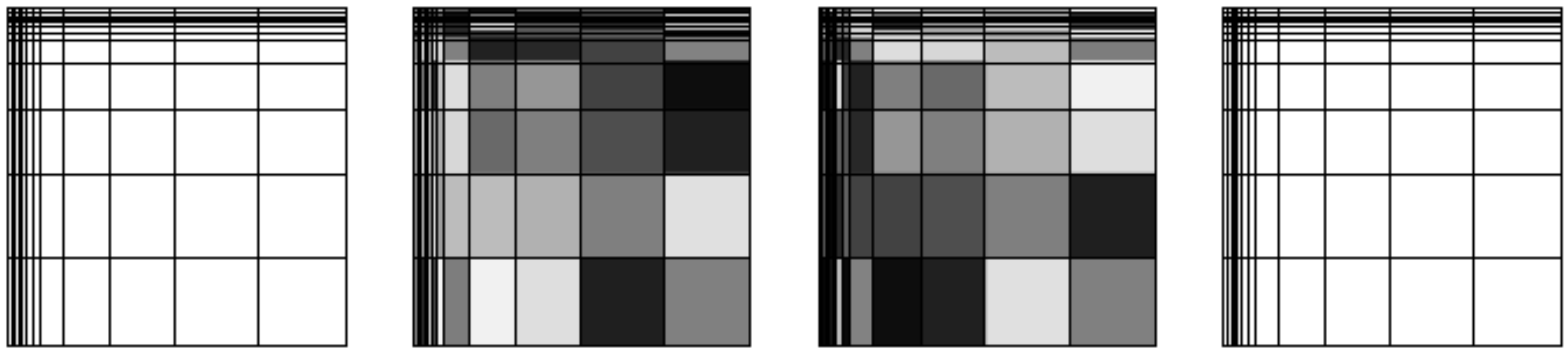
1. Draw partition ~ DP-Stick(alpha)
2. Draw weights ~ Dirichlet(**a**)

Infinite digraph block model



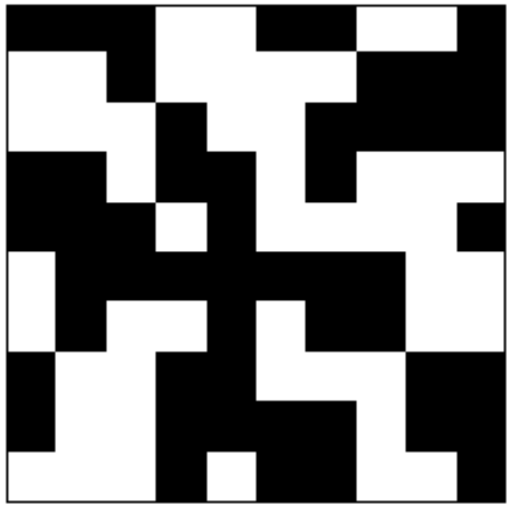
1. Draw partition ~ DP-Stick(alpha)
2. Draw weights ~ Dirichlet(**a**)

Infinite digraph block model

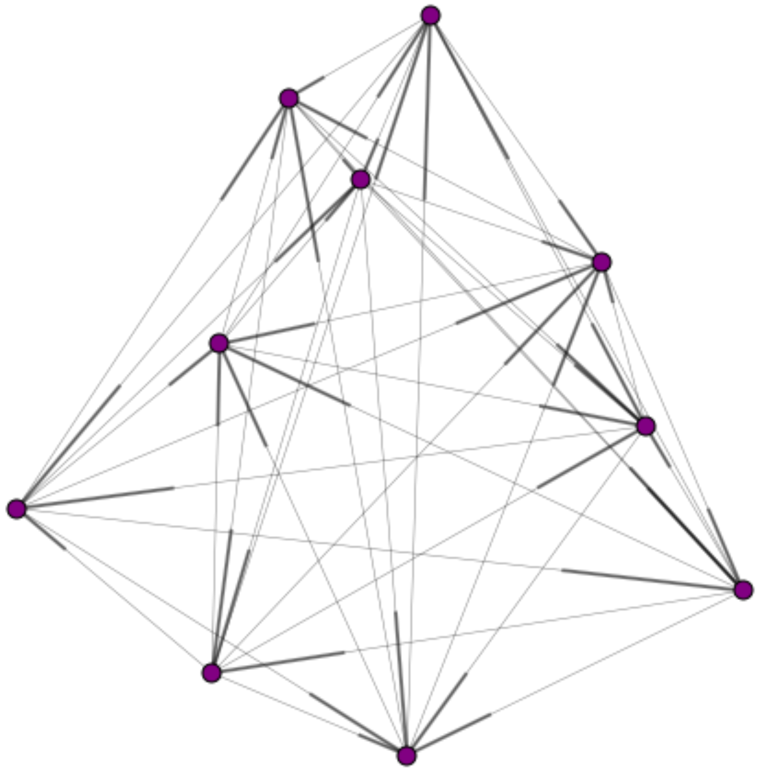


Tournament

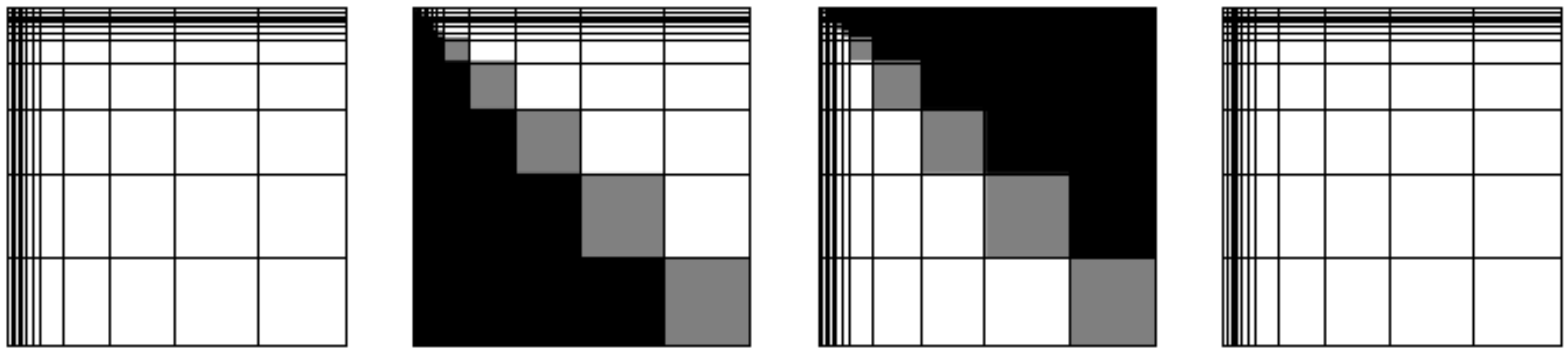
Dirichlet parameter a = [0, 2, 1, 0]



sample

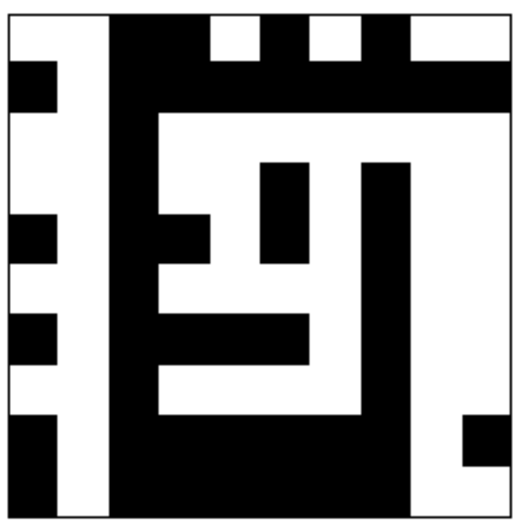


Infinite digraph block model

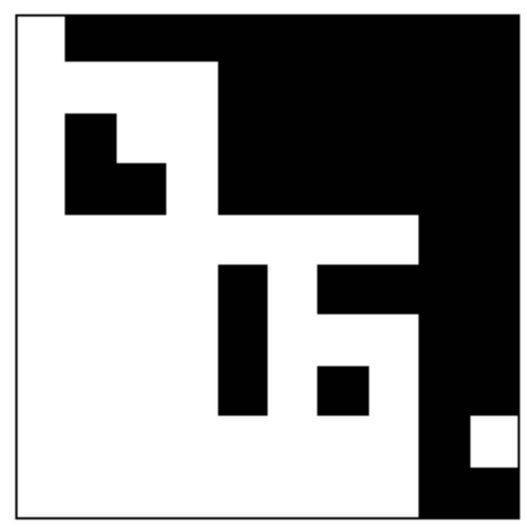


Almost totally ordered

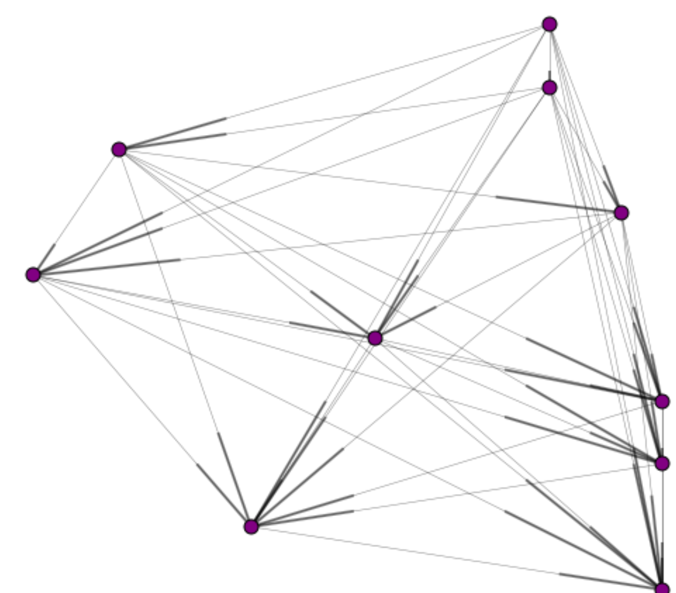
Dirichlet parameter a = [0, 0, 1, 0]



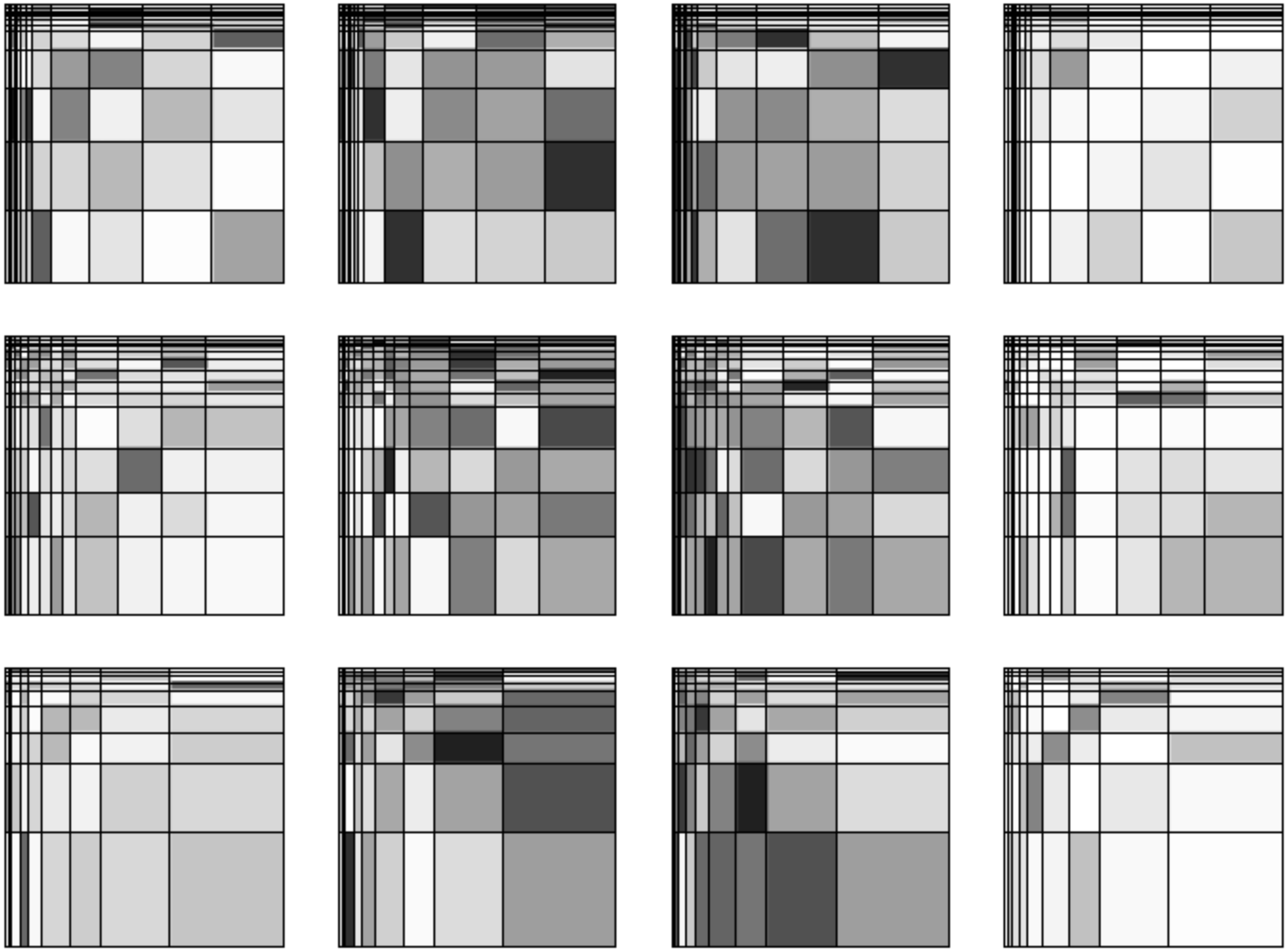
sample



reordered



Infinite digraph block model



Dirichlet parameter a = [0.9, 2.0, 1.0, 0.5]

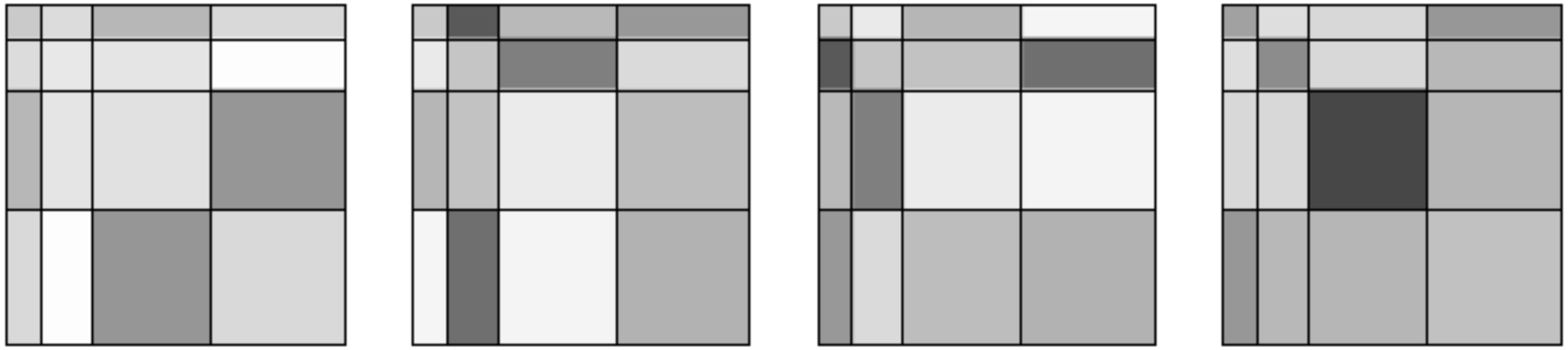
Infinite digraph block model

- Inference via collapsed Gibbs sampling of cluster assignments

$$p(z_i|z_{-i}, G) \propto p(z_i|z_{-i})p(G|z_i, z_{-i})$$

cluster assignment of vertex i   cluster assignments of all vertices except i

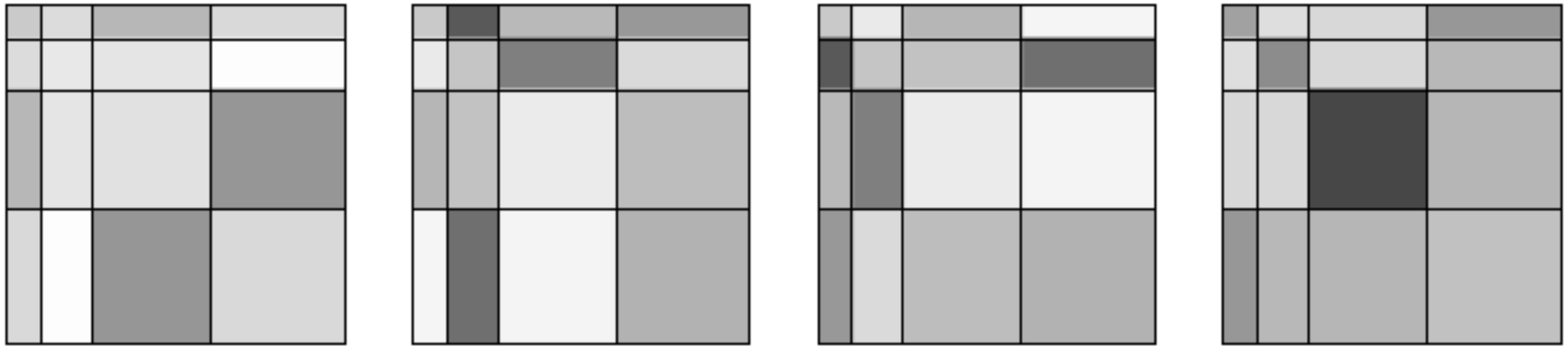
- Experiments using synthetic data



Dirichlet parameter a = [1, 1, 1, 1]

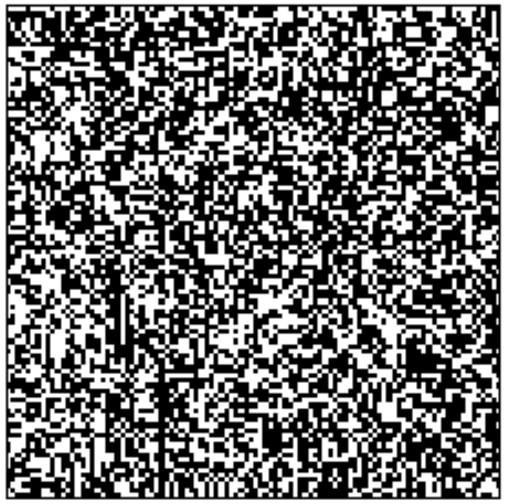
Infinite digraph block model

- Random digraphon

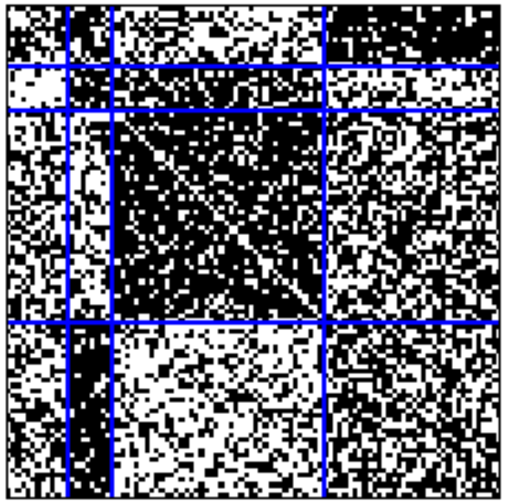


Dirichlet parameter a = [1, 1, 1, 1]

- Random sample, 100 vertices



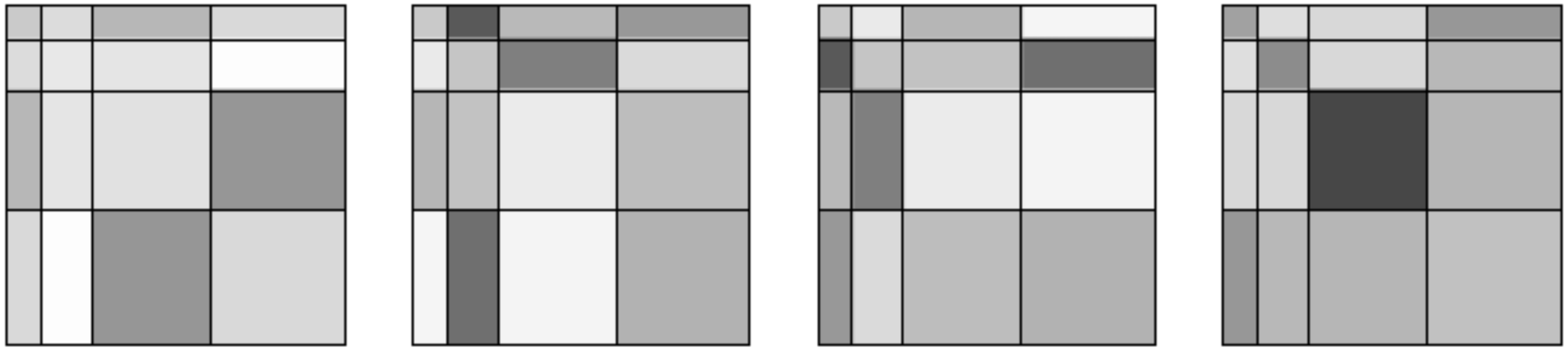
original ordering



resorted by increasing
uniform random variables

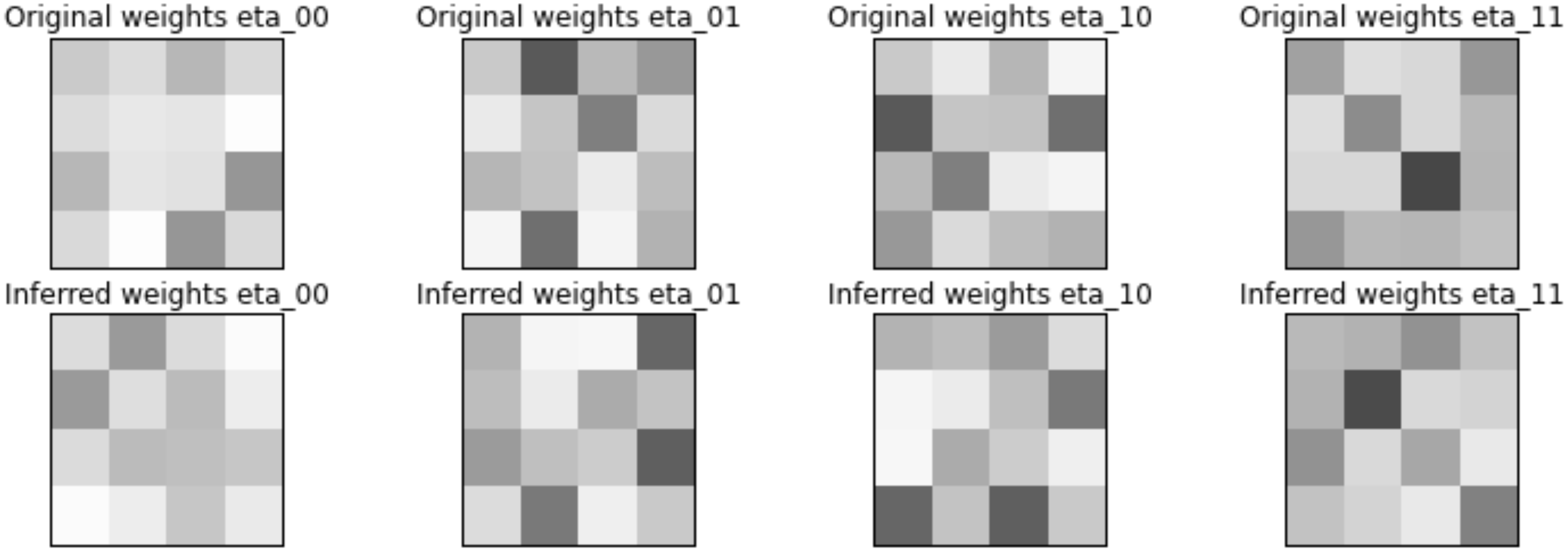
Infinite digraph block model

- Random digraphon



Dirichlet parameter a = [1, 1, 1, 1]

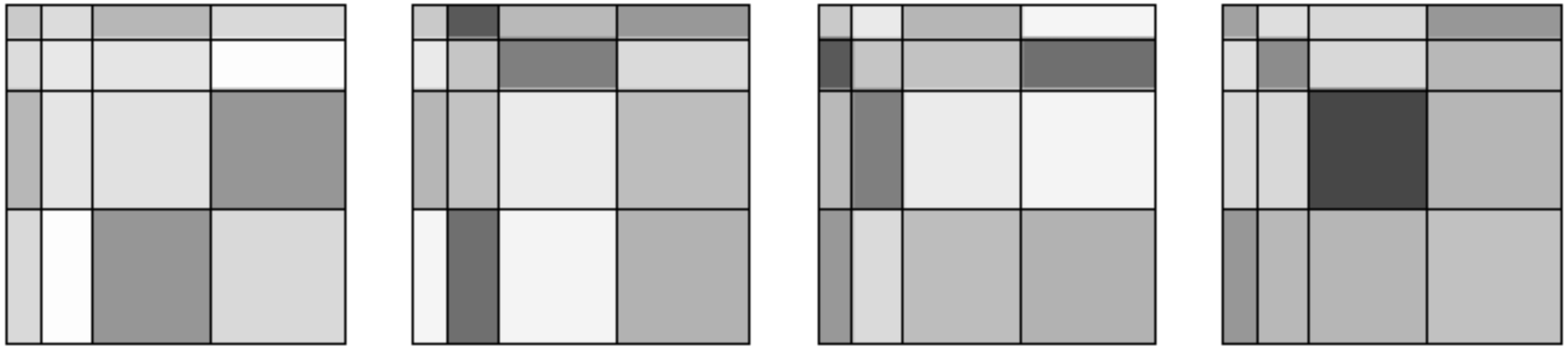
- Collapsed Gibbs sampling



[Cai–Ackerman–Freer, 2015]

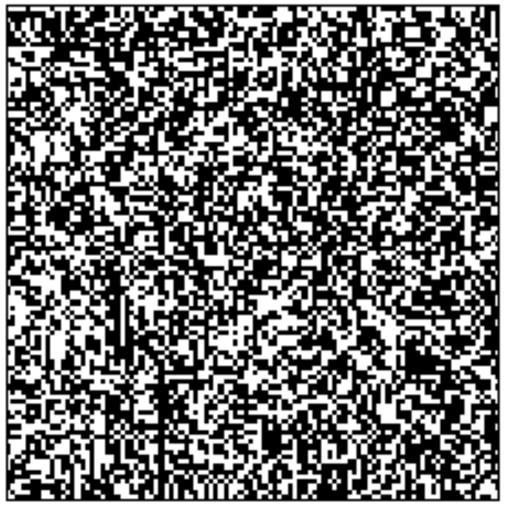
Infinite digraph block model

- Random digraphon

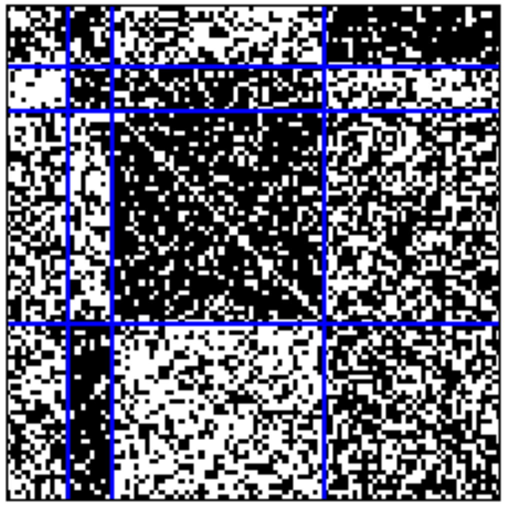


Dirichlet parameter a = [1, 1, 1, 1]

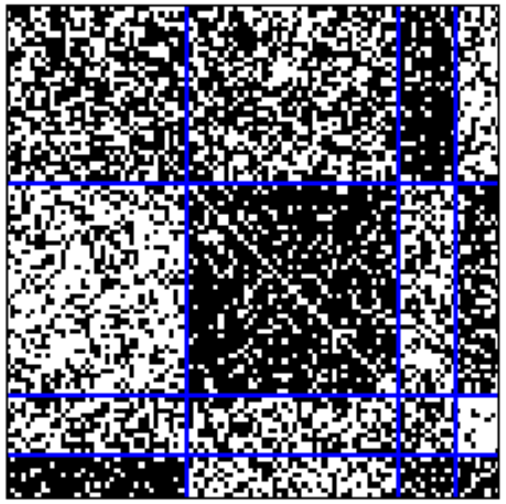
- Collapsed Gibbs sampling



original ordering



resorted by increasing uniform random variables

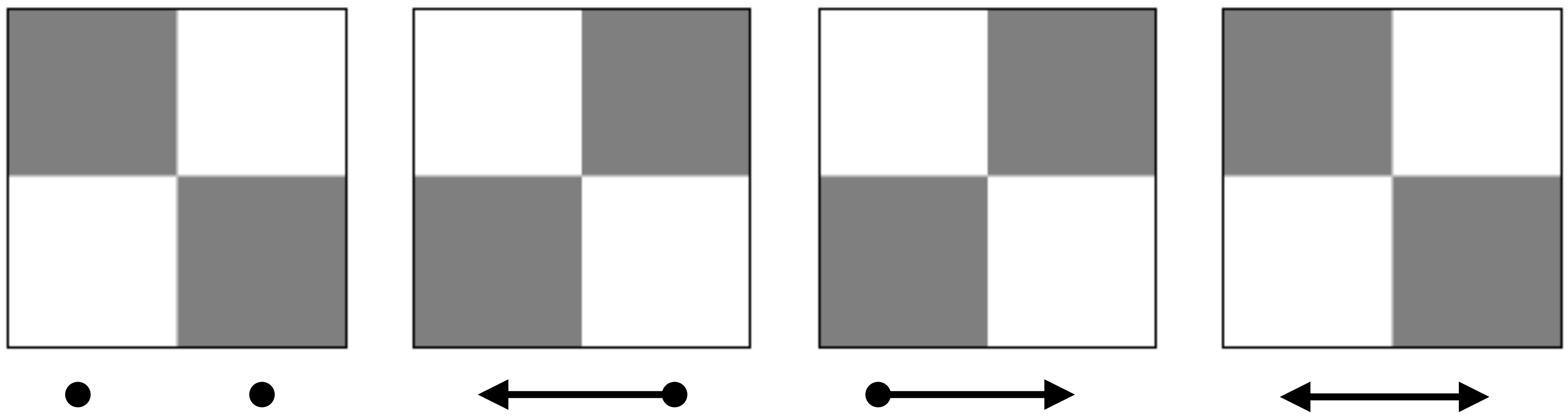


resorted by inferred cluster

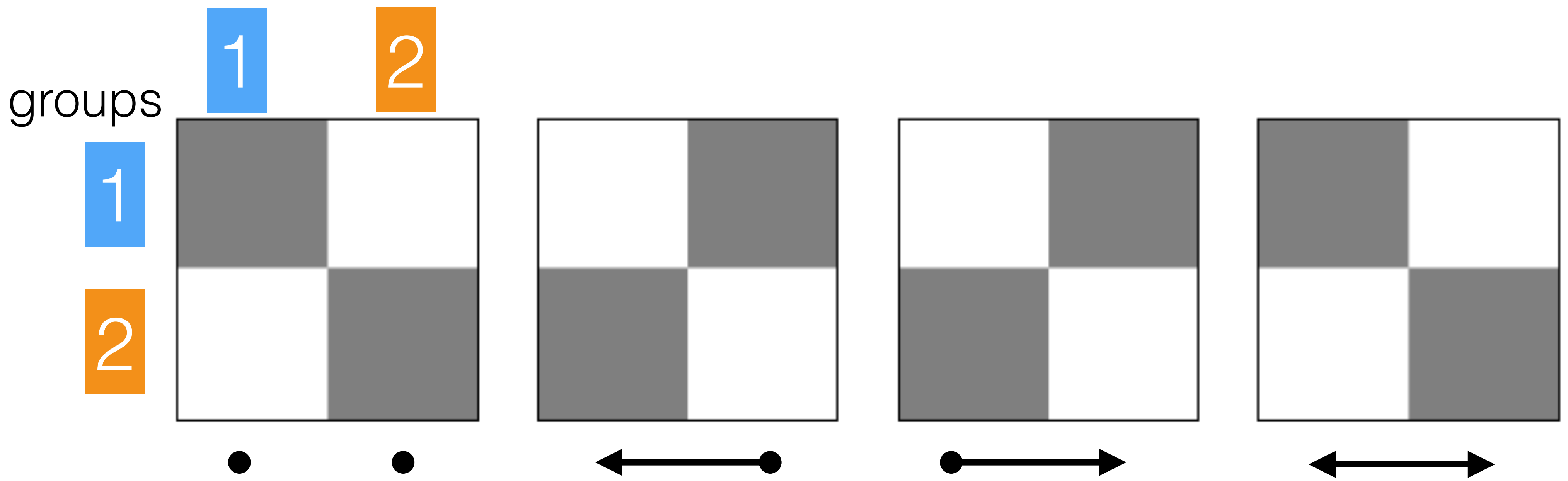
 [Cai–Ackerman–Freer, 2015]

Infinite digraph block model

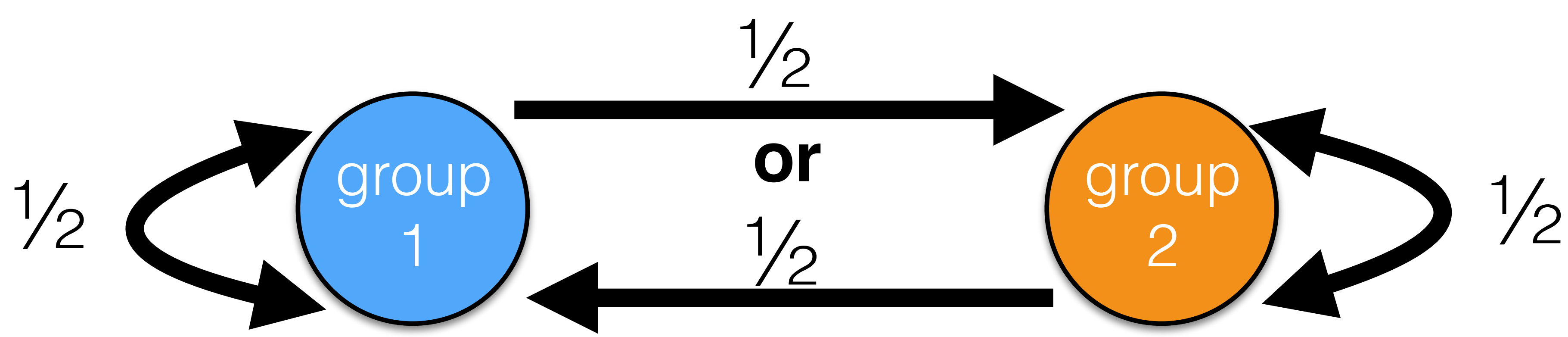
- Random digraphon: ER + tournament



Infinite digraph block model

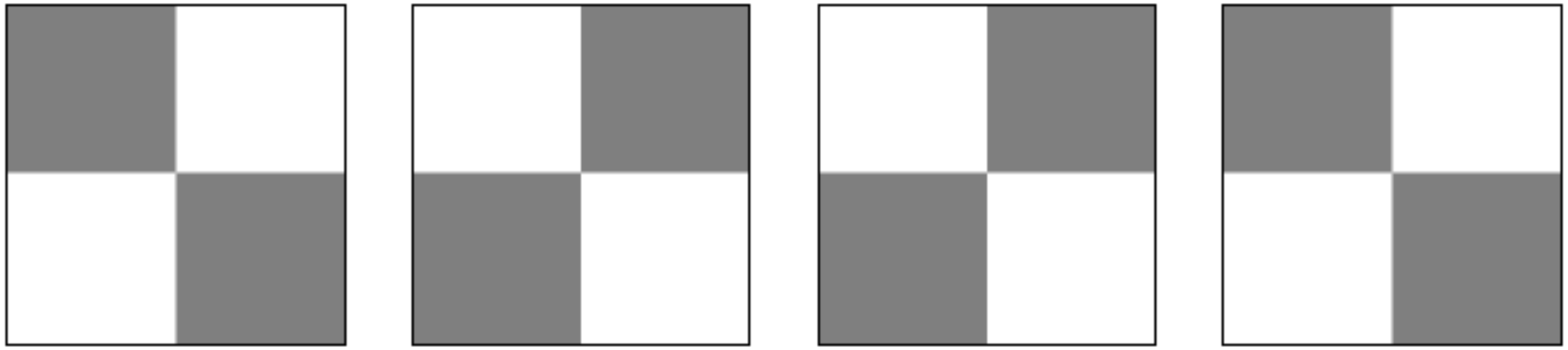


- Schematic of the sampled digraph:

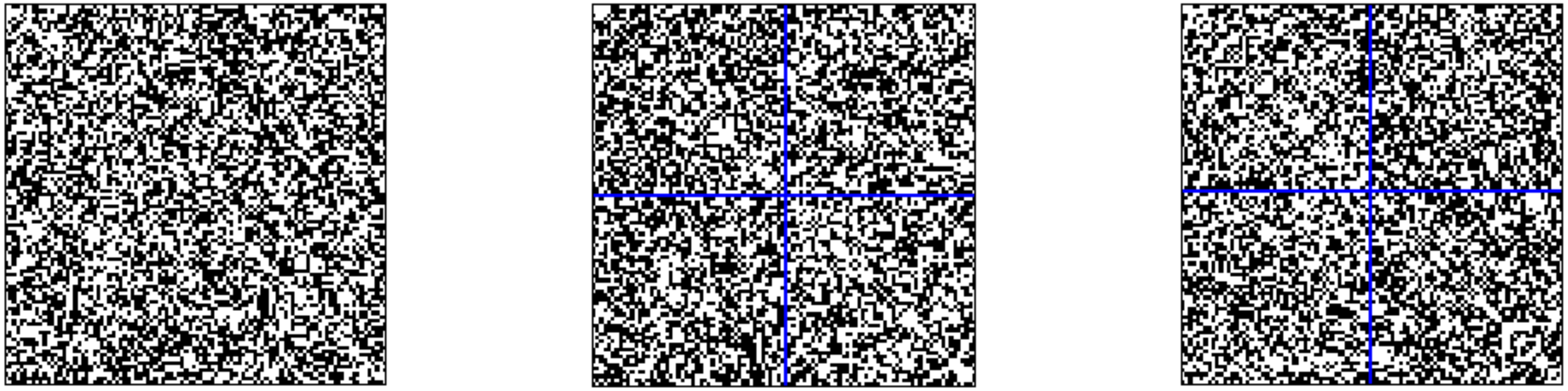


Infinite digraph block model

- Random digraphon: ER + tournament



- Collapsed Gibbs sampling for **this model**



original ordering

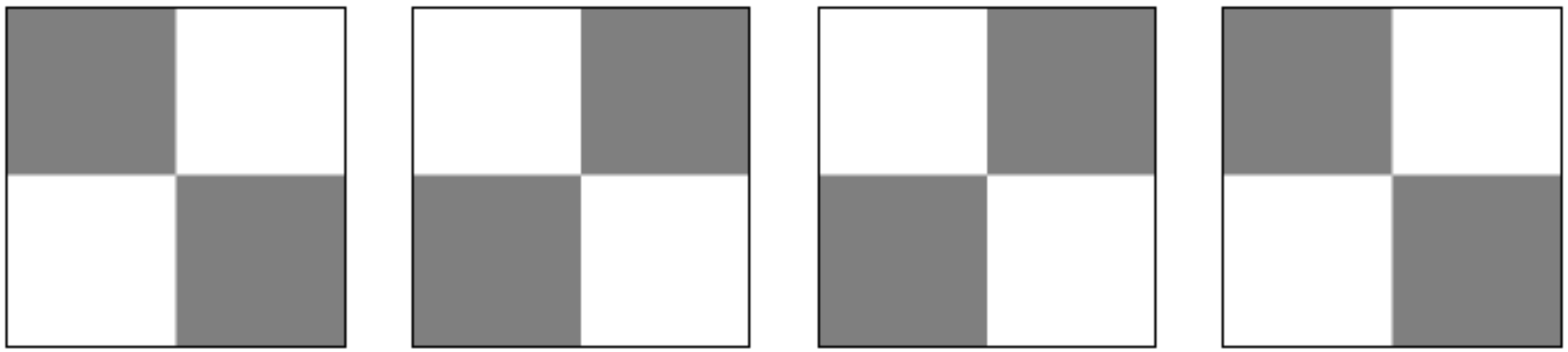
resorted by increasing uniform random variables

resorted by inferred cluster

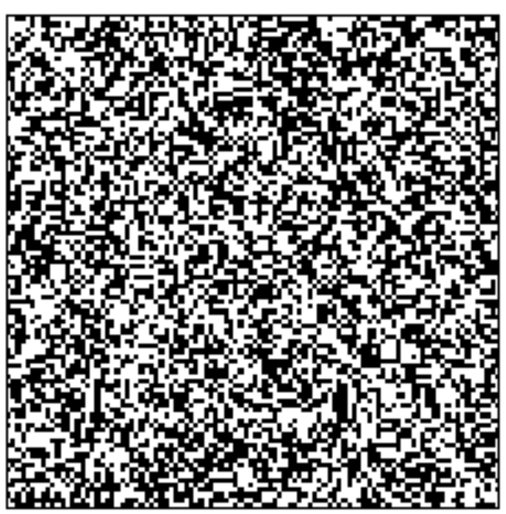
[Cai–Ackerman–Freer, 2015]

Infinite digraph block model

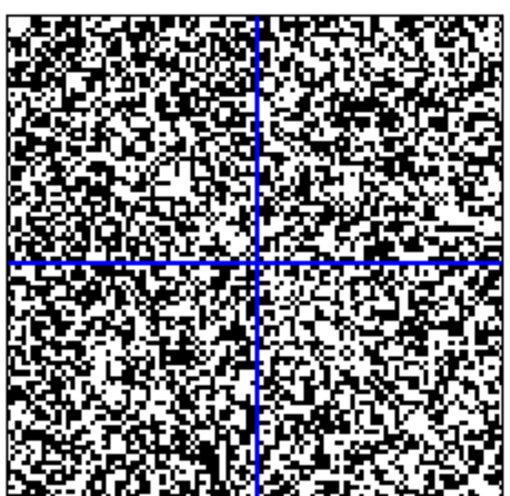
- Random digraphon: ER + tournament



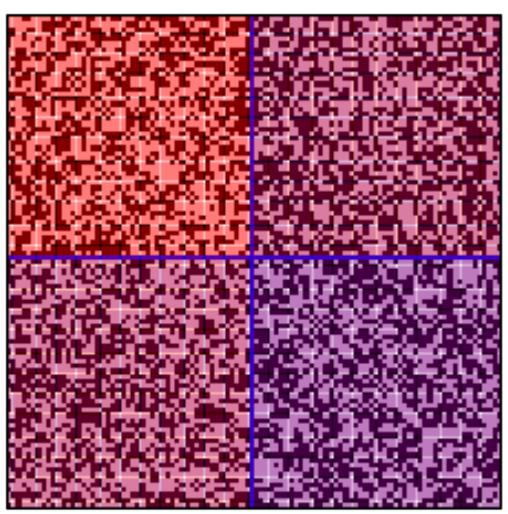
- Collapsed Gibbs sampling for **this model**



original ordering



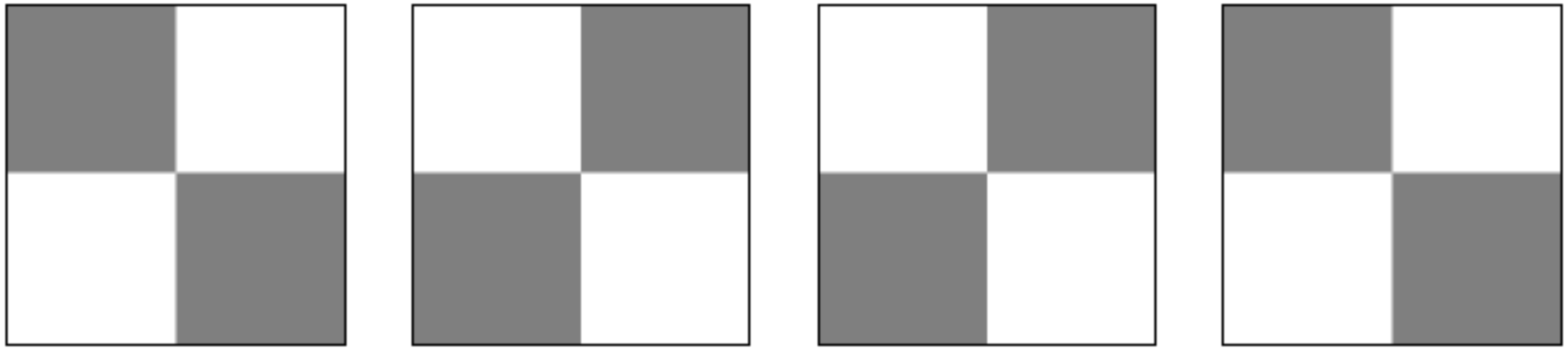
resorted by increasing uniform random variables



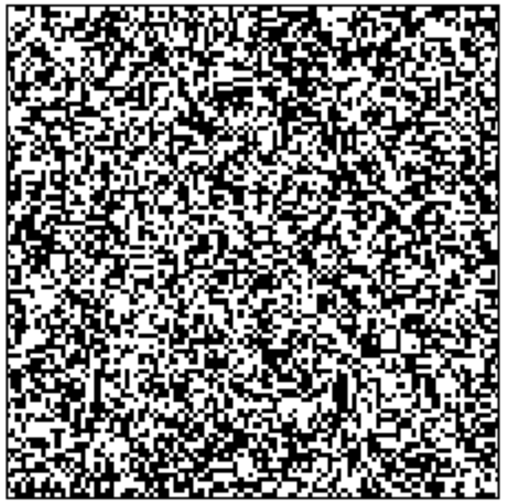
resorted by inferred cluster

Infinite digraph block model

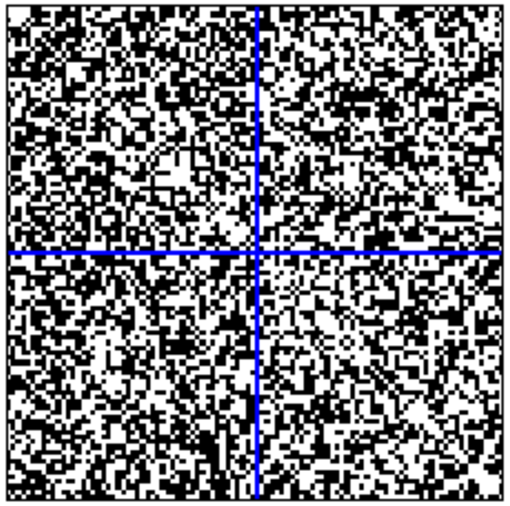
- Random digraphon: ER + tournament



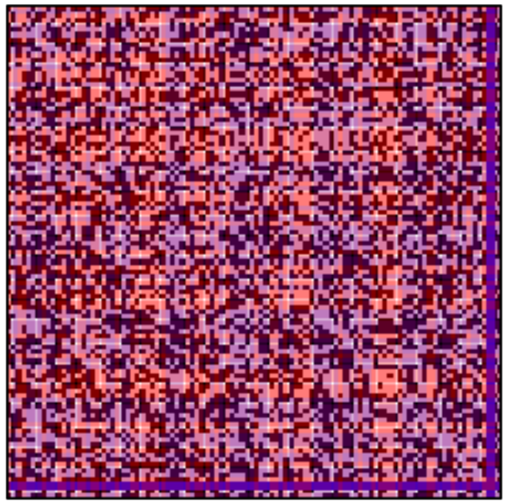
- Collapsed Gibbs sampling for the **infinite relational model**



original ordering



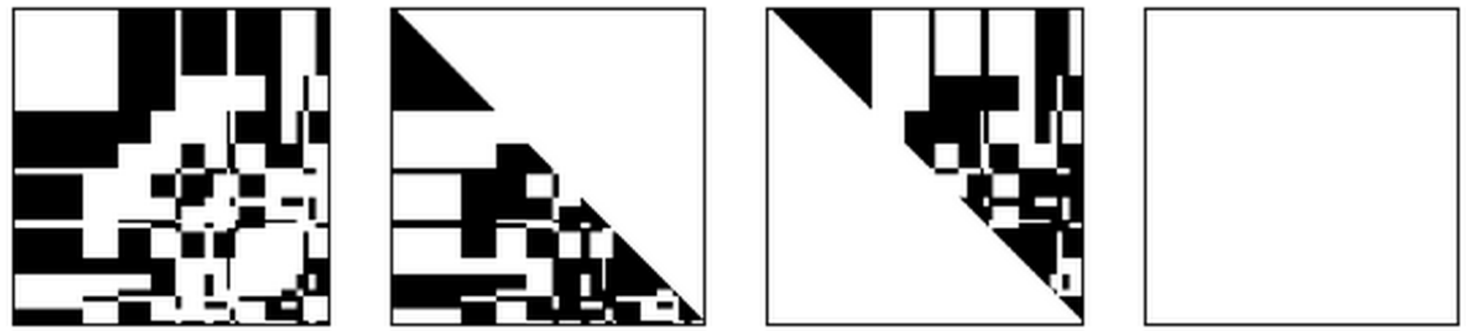
resorted by increasing uniform random variables



resorted by inferred cluster

Conclusions

- Summary
- Discussion:



  - other types of block models:
    - could consider other partitions, e.g., **Pitman–Yor**
  - other priors, e.g., **Gaussian process** (as in Lloyd et al.)
  - Aldous–Hoover already describes exchangeable **hypergraphs**
  - **sparsity**

References

- D. J. Aldous. Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.,* 11(4): 581–598, 1981.
- P. Diaconis and S. Janson. Graph limits and exchangeable random graphs. *Rend. Mat. Appl.* (7), 28(1):33–61, 2008.
- E. Glasner and B. Weiss, Minimal actions of the group S(Z) of permutations of the integers, *Geom. Funct. Anal.* 12(5):964–988, 2002.
- D. N. Hoover. Relations on probability spaces and arrays of random variables. Preprint, Institute for Advanced Study, Princeton, NJ, 1979.
- C. Kemp, J. B. Tenenbaum, T. L. Griffiths, T. Yamada, and N. Ueda. Learning systems of concepts with an infinite relational model. In *Proc. 21st Nat. Conf. Artificial Intelligence (AAAI–06),* 2006.
- J. R. Lloyd, P. Orbanz, Z. Ghahramani, and D. M. Roy. Random function priors for exchangeable arrays with applications to graphs and relational data. In *Adv. Neural Inform. Process. Syst. (NIPS)* 25:1007–1015, 2012.
- L. Lovász. *Large networks and graph limits,* vol. 60 of American Math. Soc. Colloq. Publ. American Math. Soc., Providence, RI, 2012.
- P. Orbanz and D. M. Roy. Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Trans. Pattern Anal. Mach. Intell.,* 37(2):437–461, 2015.
- Y. J. Wang, and G. Y. Wong. Stochastic blockmodels for directed graphs. *J. American Statist. Assoc.,* 82(397), 8–19, 1987.