Bayesian Nonparametrics For Time Series Modeling

DOCTORAL THESIS

Francisco Jesús Rodríguez Ruiz

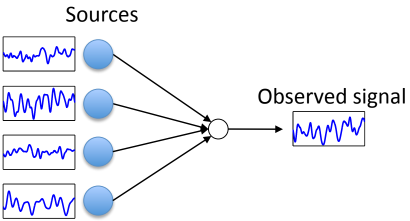
June 30th, 2015

Outline

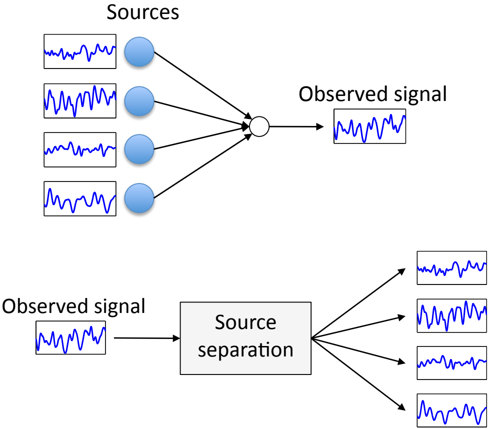
Outline

1. Introduction
2. Bayesian Nonparametrics
3. Contributions
   - Infinite Factorial Unbounded-State HMM
   - Infinite Factorial Finite State Machine
4. Conclusions

Motivation



Motivation


Sources
Observed signal
Observed signal
Source separation

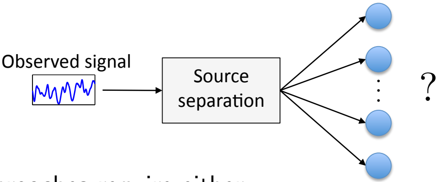
Motivation

Applications:

- Power disaggregation.
- Multiuser communication systems.
- Speech separation.
- Multi-target tracking.
- Electroencephalography (EEG).
- ...

Motivation

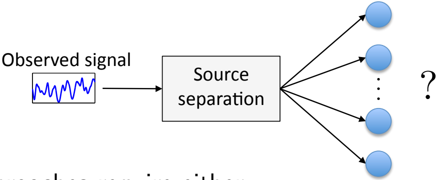
How many hidden sources?



- Classical approaches require either
  - known number of sources.
  - upper bound.
  - model selection.

Motivation

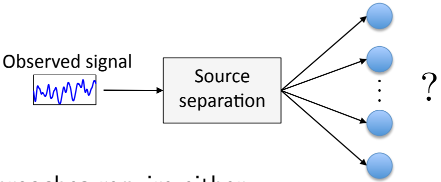
How many hidden sources?



- Classical approaches require either
  - known number of sources.
  - upper bound.
  - model selection.
- Bayesian nonparametrics can
  - infer the number of latent sources from the data.
  - avoid model selection.

MOTIVATION

How many hidden sources?

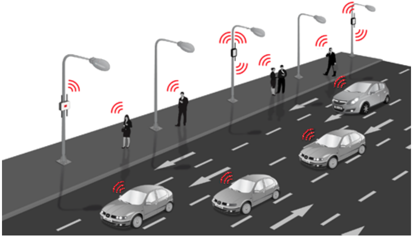


- Classical approaches require either
  - known number of sources.
  - upper bound.
  - model selection.
- Bayesian nonparametrics can
  - infer the number of latent sources from the data.
  - avoid model selection.

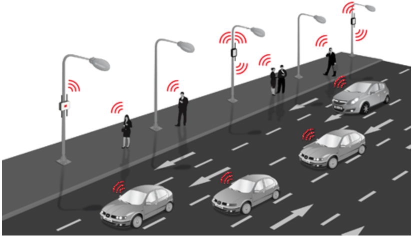
OUR APPROACH

Bayesian nonparametric modeling of
discrete-time series for source separation problems

Why BNP?

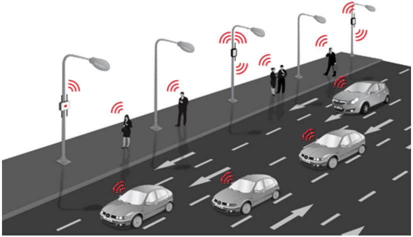


Why BNP?



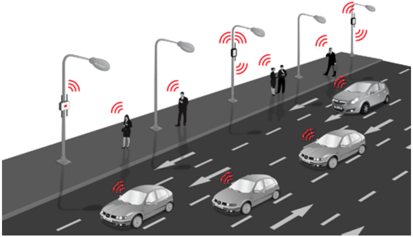
- Pick a large enough #sources.

Why BNP?



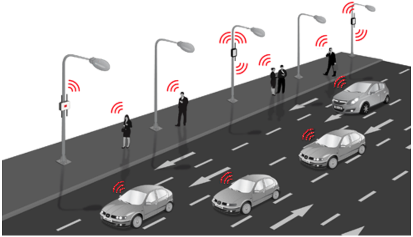
- Pick a large enough #sources.
- Model selection (AIC, BIC).

Why BNP?



- Pick a large enough #sources.
- Model selection (AIC, BIC).
- Bayesian model selection.

Why BNP?



- Pick a large enough #sources.
- Model selection (AIC, BIC).
- Bayesian model selection.
- BNP:
  - Model complexity grows with data size.
  - Unbounded #sources.

State Of The Art

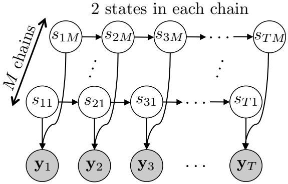
- Many BNP models for discrete-time series.
  - e.g., infinite HMM.
- Not many BNP models for source separation.

State Of The Art

- Many BNP models for discrete-time series.
  - e.g., infinite HMM.
- Not many BNP models for source separation.
  - Infinite ICA.

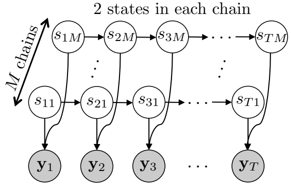
State Of The Art

- Many BNP models for discrete-time series.
  - e.g., infinite HMM.
- Not many BNP models for source separation.
  - Infinite ICA.
  - Infinite Factorial HMM (IFHMM).



State Of The Art

- Many BNP models for discrete-time series.
  - e.g., infinite HMM.
- Not many BNP models for source separation.
  - Infinite ICA.
  - Infinite Factorial HMM (IFHMM).



  - ICA-IFHMM.

State Of The Art

Lack of BNP models for source separation:

- Infinite factorial HMM with **non-binary** hidden states.
  - e.g., power disaggregation.

State Of The Art

Lack of BNP models for source separation:

- Infinite factorial HMM with **non-binary** hidden states.
  - e.g., power disaggregation.
- Model that accounts for **multipath propagation**.
  - e.g., multiuser communications systems.

State Of The Art

Lack of BNP models for source separation:

- Infinite factorial HMM with **non-binary** hidden states.
  - e.g., power disaggregation.
- Model that accounts for **multipath propagation**.
  - e.g., multiuser communications systems.
- Model with continuous-valued states that captures **temporal dependencies**.
  - e.g., speech separation.

Contributions

Infinite Factorial Unbounded-State HMM

❶ Non-binary IFHMM.

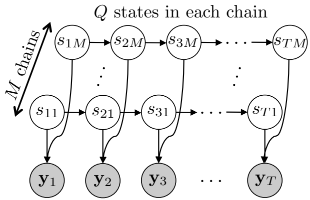
- Can infer the number of HMMs in a factorial model.

❷ IFUHMM.

- Can additionally infer the cardinality of the state space.

Applications:

- Power disaggregation.
- Multiuser communication systems.



Contributions

Infinite Factorial Unbounded-State HMM

1. Non-binary IFHMM.
   - Can infer the number of HMMs in a factorial model.
2. IFUHMM.
   - Can additionally infer the cardinality of the state space.

Applications:

- Power disaggregation.
- Multiuser communication systems.

Infinite Factorial Finite State Machine

- Can infer the number of FSMs in a factorial model.
- Naturally account for multipath, echo, . . .

Applications:

- Multiuser communication systems.

Outline

Bayesian Nonparametrics

- Bayesian framework for **model selection**.
- Prior over **infinite-dimensional** parameter space.
- Only a **finite subset** of the parameters is used for any finite dataset.
- The model complexity is allowed to grow with data size.
- Rely on **stochastic processes**:
  - Gaussian process.
  - Dirichlet process.
  - Beta process.
  - ...

Bayesian Nonparametrics

- Bayesian framework for **model selection**.
- Prior over **infinite-dimensional** parameter space.
- Only a **finite subset** of the parameters is used for any finite dataset.
- The model complexity is allowed to grow with data size.
- Rely on **stochastic processes**:
  - Gaussian process.
  - Dirichlet process.
  - Beta process.
  - ...

Indian Buffet Process

- Prior over binary matrices with infinite columns.
- Rows $\equiv$ Data points. Columns $\equiv$ Features.
- $\mathbf{S} \sim \mathrm{IBP}(\alpha)$.
- α: Concentration parameter.
- Each element $s_{tm} \in \{0, 1\}$ indicates whether the m-th feature contributes to the t-th data point.
- Only a finite number of columns M_+ active for any finite number of rows.

$$\mathbf{S} = \underbrace{\left[\begin{array}{ccccccc} s_{11} & s_{12} & \cdots & s_{1M_+} & 0 & 0 & \cdots \\ s_{21} & s_{22} & \cdots & s_{2M_+} & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ s_{T1} & s_{T2} & \cdots & s_{TM_+} & 0 & 0 & \cdots \end{array}\right]}_{M \text{ columns (features)}} T \text{ observations}$$

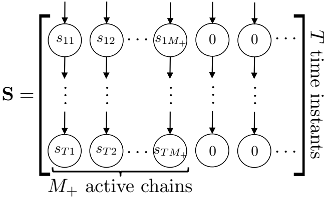
(M_+ non-zero columns)

Markov Indian Buffet Process

- Prior over binary matrices with infinite columns.
- Each column follows a Markov process.
- For any T, only M_+ chains become active.
- The probability $p(\mathbf{S})$ vanishes, but $p([\mathbf{S}]) > 0$.
  - $[\mathbf{S}]$: set of matrices equivalent to $\mathbf{S}$.
- Useful to build a (binary) infinite factorial HMM.

$$\mathbf{S} = \underbrace{\left[\begin{array}{cccccccc} s_{11} & s_{12} & \cdots & s_{1M_+} & 0 & 0 & \cdots \\ s_{21} & s_{22} & \cdots & s_{2M_+} & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ s_{T1} & s_{T2} & \cdots & s_{TM_+} & 0 & 0 & \cdots \end{array}\right]}_{M \text{ columns (chains)}} \; T \text{ time instants}$$

(M_+ non-zero columns)



Markov Indian Buffet Process

$$\mathbf{S} \sim \mathrm{MIBP}(\alpha, \beta_0, \beta_1)$$

- Can be obtained by defining the transition probabilities

$$\mathbf{A}^m = \begin{bmatrix} a^m & 1-a^m \\ b^m & 1-b^m \end{bmatrix} \qquad \begin{aligned} a^m &= p(s_{tm}=0|s_{(t-1)m}=0) \\ b^m &= p(s_{tm}=0|s_{(t-1)m}=1) \end{aligned}$$

Markov Indian Buffet Process

$$\mathbf{S} \sim \mathrm{MIBP}(\alpha, \beta_0, \beta_1)$$

- Can be obtained by defining the transition probabilities

$$\mathbf{A}^m = \begin{bmatrix} a^m & 1-a^m \\ b^m & 1-b^m \end{bmatrix} \qquad \begin{aligned} a^m &= p(s_{tm}=0 | s_{(t-1)m}=0) \\ b^m &= p(s_{tm}=0 | s_{(t-1)m}=1) \end{aligned}$$

- ... with priors

$$a^m \sim \mathrm{Beta}\left(1, \frac{\alpha}{M}\right) \qquad b^m \sim \mathrm{Beta}(\beta_0, \beta_1)$$

- ... and let $M \to \infty$

Markov Indian Buffet Process

$$\mathbf{S} \sim \mathrm{MIBP}(\alpha, \beta_0, \beta_1)$$

- Can be obtained by defining the transition probabilities

$$\mathbf{A}^m = \begin{bmatrix} a^m & 1-a^m \\ b^m & 1-b^m \end{bmatrix} \qquad \begin{aligned} a^m &= p(s_{tm} = 0 | s_{(t-1)m} = 0) \\ b^m &= p(s_{tm} = 0 | s_{(t-1)m} = 1) \end{aligned}$$

- ... with priors

$$a^m \sim \mathrm{Beta}\left(1, \frac{\alpha}{M}\right) \qquad b^m \sim \mathrm{Beta}(\beta_0, \beta_1)$$

- ... and let $M \to \infty$
- After integrating out a^m and b^m:

$$\lim_{M\to\infty} p([\mathbf{S}]) = \frac{\alpha^{M_+}}{\prod_{h=1}^{2^T} M_h!} e^{-\alpha H_T} \prod_{m=1}^{M_+} \frac{(n_{01}^m - 1)!(n_{00}^m)!\Gamma(\beta_0+\beta_1)\Gamma(\beta_0+n_{10}^m)\Gamma(\beta_1+n_{11}^m)}{(n_{00}^m + n_{01})!\Gamma(\beta_0)\Gamma(\beta_1)\Gamma(\beta_0+\beta_1+n_{10}^m+n_{11}^m)}$$

- Markov exchangeable in the rows.

Outline

Outline

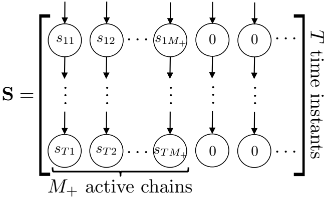
1. Introduction
2. Bayesian Nonparametrics
3. Contributions
   - Infinite Factorial Unbounded-State HMM
   - Infinite Factorial Finite State Machine
4. Conclusions

Non-Binary Infinite Factorial HMM

- Generalization of the MIBP for non-binary matrices.
- Each state $s_{tm} \in \{0, 1, \ldots, Q-1\}$.
- Inactive state ($s_{tm} = 0$).

$$\mathbf{S} = \underbrace{\left[\begin{array}{cccccccc} s_{11} & s_{12} & \cdots & s_{1M_+} & 0 & 0 & \cdots \\ s_{21} & s_{22} & \cdots & s_{2M_+} & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ s_{T1} & s_{T2} & \cdots & s_{TM_+} & 0 & 0 & \cdots \end{array} \right]}_{M \text{ columns (chains)}} \; T \text{ time instants}$$

(M_+ non-zero columns)



Non-Binary Infinite Factorial HMM

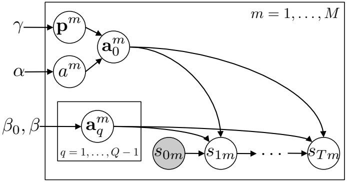
- Can be obtained by defining the transition probabilities

$$\mathbf{A}^m = \begin{bmatrix} a_{00}^m & a_{01}^m & \cdots & a_{0(Q-1)}^m \\ a_{10}^m & a_{11}^m & \cdots & a_{1(Q-1)}^m \\ \vdots & \vdots & \ddots & \vdots \\ a_{(Q-1)0}^m & a_{(Q-1)1}^m & \cdots & a_{(Q-1)(Q-1)}^m \end{bmatrix}$$

Non-Binary Infinite Factorial HMM

- Can be obtained by defining the transition probabilities

$$\mathbf{A}^m = \begin{bmatrix} a^m_{00} & a^m_{01} & \cdots & a^m_{0(Q-1)} \\ a^m_{10} & a^m_{11} & \cdots & a^m_{1(Q-1)} \\ \vdots & \vdots & \ddots & \vdots \\ a^m_{(Q-1)0} & a^m_{(Q-1)1} & \cdots & a^m_{(Q-1)(Q-1)} \end{bmatrix}$$



- Prior distribution:

$$a^m \sim \text{Beta}\left(1, \frac{\alpha}{M}\right) \qquad \mathbf{p}^m \sim \text{Dirichlet}(\gamma)$$

$$\mathbf{a}^m_0 = \begin{bmatrix} a^m & (1-a^m)\mathbf{p}^m \end{bmatrix}$$

$$\mathbf{a}^m_q \sim \text{Dirichlet}(\beta_0, \beta, \ldots, \beta), \quad q = 1, \ldots, Q-1$$

Limit of $p([\mathbf{S}])$

$$\lim_{M\to\infty} p([\mathbf{S}]) = \frac{(Q-1)!}{(Q-N_Q)!N_f} \frac{\alpha^{M_+}}{\prod_{h=1}^{Q^T-1} M_h!} e^{-\alpha H_T}$$
$$\times \prod_{m=1}^{M_+} \left[\frac{\Gamma(n_{00}^m+1)\Gamma\left(\sum_{i=1}^{Q-1} n_{0i}^m\right)}{\Gamma(n_{0\bullet}^m+1)} \frac{\Gamma\left((Q-1)\gamma\right)\prod_{i=1}^{Q-1}\Gamma(n_{0i}^m+\gamma)}{\Gamma\left(\sum_{i=1}^{Q-1}(n_{0i}^m+\gamma)\right)(\Gamma(\gamma))^{Q-1}} \right.$$
$$\left. \times \prod_{q=1}^{Q-1}\left(\frac{\Gamma\left(\beta_0+(Q-1)\beta\right)}{\Gamma(\beta_0)\left(\Gamma(\beta)\right)^{Q-1}} \frac{\Gamma(n_{q0}^m+\beta_0)\prod_{i=1}^{Q-1}\Gamma(n_{qi}^m+\beta)}{\Gamma\left(n_{q\bullet}^m+\beta_0+(Q-1)\beta\right)} \right)\right].$$

Culinary Metaphor

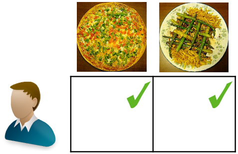
$Q = 3$ states (1 inactive + 2 active)

$t = 1$

Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)

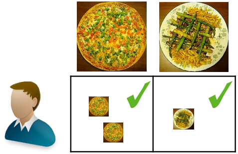
$t = 1$



Culinary Metaphor

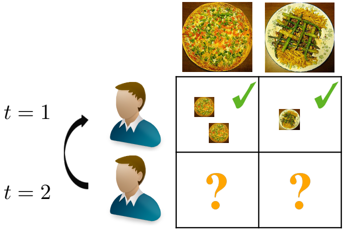
$Q = 3$ states (1 inactive + 2 active)

$t = 1$



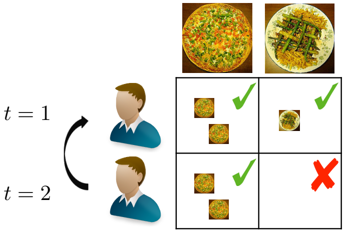
Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)



Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)

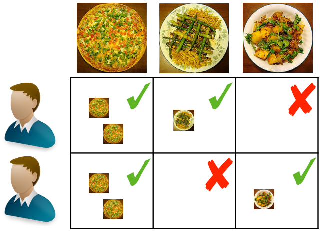


Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)

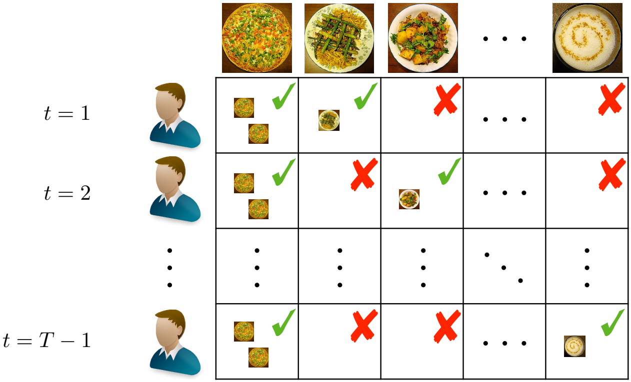
$t = 1$

$t = 2$



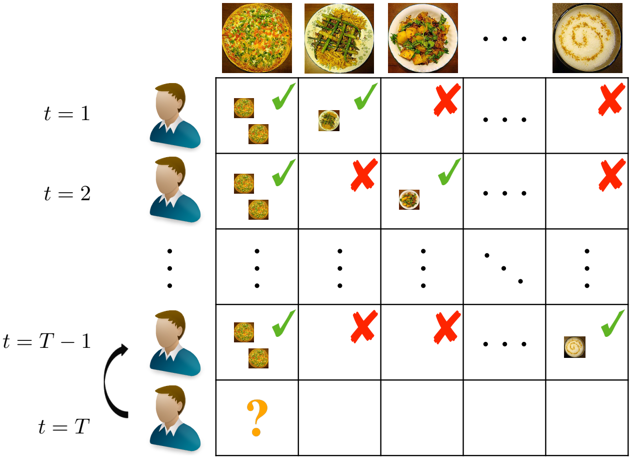
Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)



Culinary Metaphor

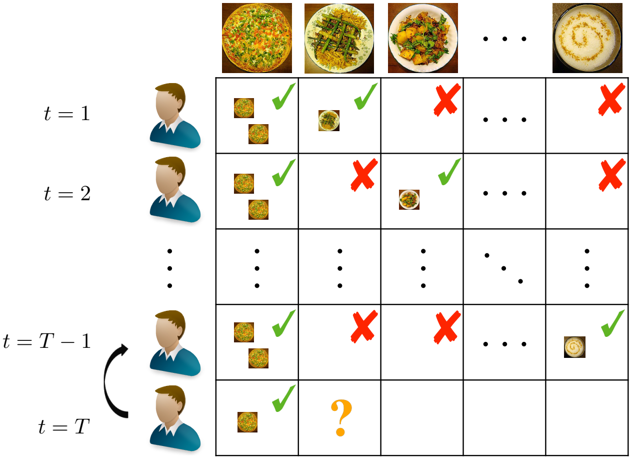
$Q = 3$ states (1 inactive + 2 active)



$$p(s_{T1} = 0) \propto \beta_0 + n_{20}^{1}$$
$$p(s_{T1} = 1) \propto \beta + n_{21}^{1}$$
$$p(s_{T1} = 2) \propto \beta + n_{22}^{1}$$

Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)

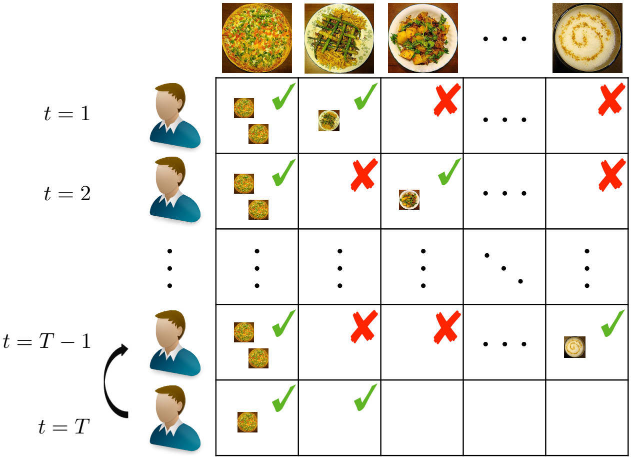


$p(s_{T2} = 0) \propto 1 + n_{00}^{2}$

$p(s_{T2} \neq 0) \propto n_{01}^{2} + n_{02}^{2}$

Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)

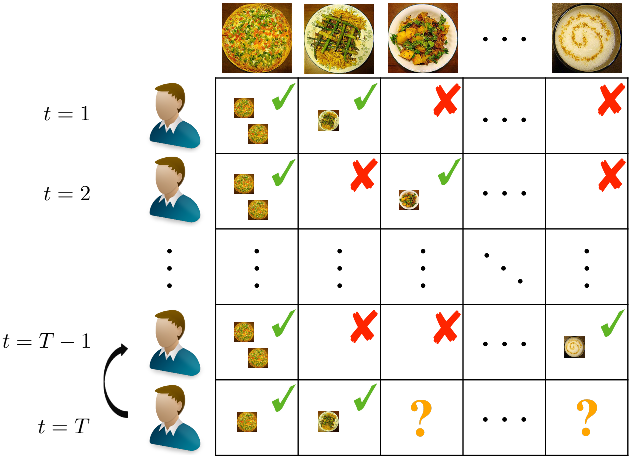


$p(s_{T2} = 1) \propto \gamma + n_{01}^{2}$

$p(s_{T2} = 2) \propto \gamma + n_{02}^{2}$

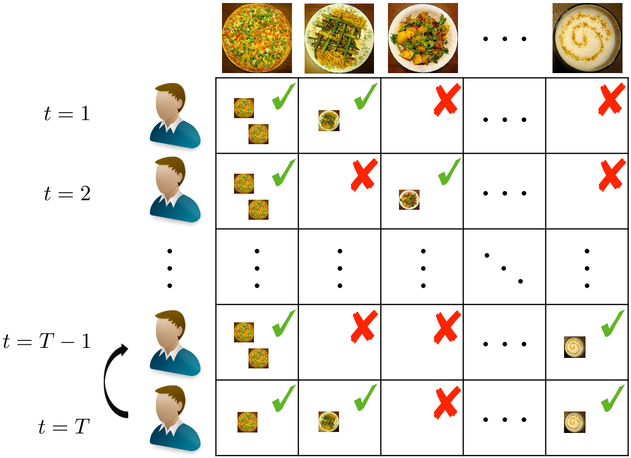
Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)



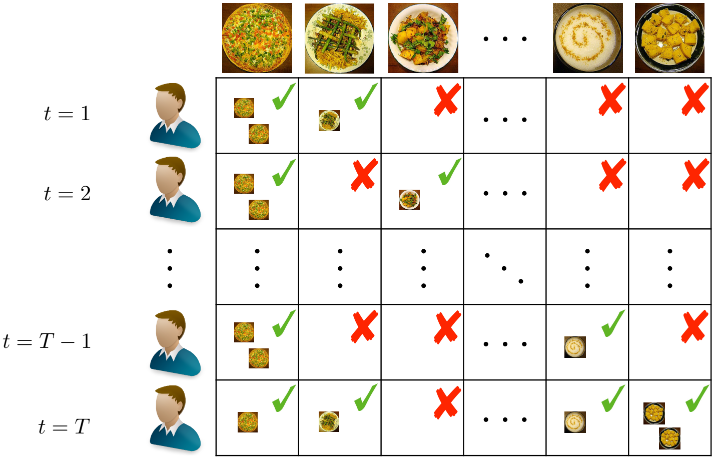
Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)



Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)



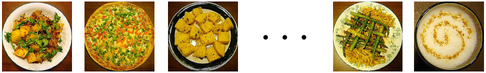
Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)

$t = 1$	2	1	0	$\cdots$	0	0
$t = 2$	2	0	1	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$t = T - 1$	2	0	0	$\cdots$	1	0
$t = T$	1	1	0	$\cdots$	1	2

Culinary Metaphor

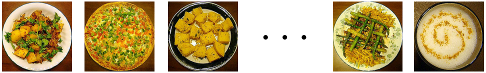
$Q = 3$ states (1 inactive + 2 active)



$t = 1$	0	2	0	$\cdots$	1	0
$t = 2$	1	2	0	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$t = T - 1$	0	2	0	$\cdots$	0	1
$t = T$	0	1	2	$\cdots$	1	1

Culinary Metaphor

$Q = 3$ states (1 inactive + 2 active)



$t = 1$	0	1	0	$\cdots$	2	0
$t = 2$	2	1	0	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$t = T-1$	0	1	0	$\cdots$	0	2
$t = T$	0	2	1	$\cdots$	2	2

Inference

Fixed Q

1. MCMC:
   - Sample from the posterior.
   - Blocked sampling approach.
   - Slice sampling $\rightarrow$ Stick-breaking construction.
   - FFBS for each Markov chain.
2. Variational:
   - Approximate the posterior.
   - Structured approach.
   - Involves a forward-backward algorithm.

Inference

Fixed Q

1. MCMC:
   - Sample from the posterior.
   - Blocked sampling approach.
   - Slice sampling $\rightarrow$ Stick-breaking construction.
   - FFBS for each Markov chain.
2. Variational:
   - Approximate the posterior.
   - Structured approach.
   - Involves a forward-backward algorithm.

Infinite Factorial Unbounded-State HMM

- Prior over the number of states:

$$Q = 2 + Q', \qquad Q' \sim \text{Poisson}(\lambda)$$

Inference

Fixed Q

1. MCMC:
   - Sample from the posterior.
   - Blocked sampling approach.
   - Slice sampling $\rightarrow$ Stick-breaking construction.
   - FFBS for each Markov chain.
2. Variational:
   - Approximate the posterior.
   - Structured approach.
   - Involves a forward-backward algorithm.

Infinite Factorial Unbounded-State HMM

- Prior over the number of states:

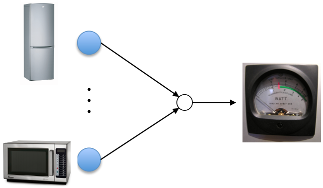
$$Q = 2 + Q', \qquad Q' \sim \text{Poisson}(\lambda)$$

Unknown Q

1. MCMC:
   - Based on reversible jump MCMC.
   - Integrate out dimension-changing variables.
   - Updating variables:
     - Q: Split/merge, birth/death.
     - M_+: Slice sampling.

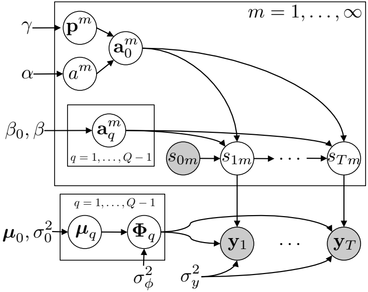
Power Disaggregation

- Estimate the power consumption of each device.
- Non-invasive measurements.
  - Improve efficiency of consumers.
  - Detect faulty equipment.
- Two datasets.
  - REDD (1 day, 5 houses, 6 devices).
  - AMP (2 days, 1 house, 8 devices).



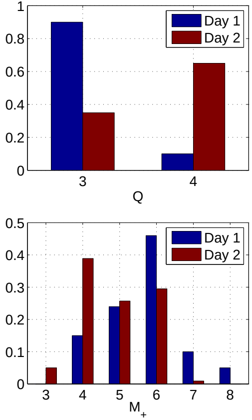
Power Disaggregation

Gaussian observation model



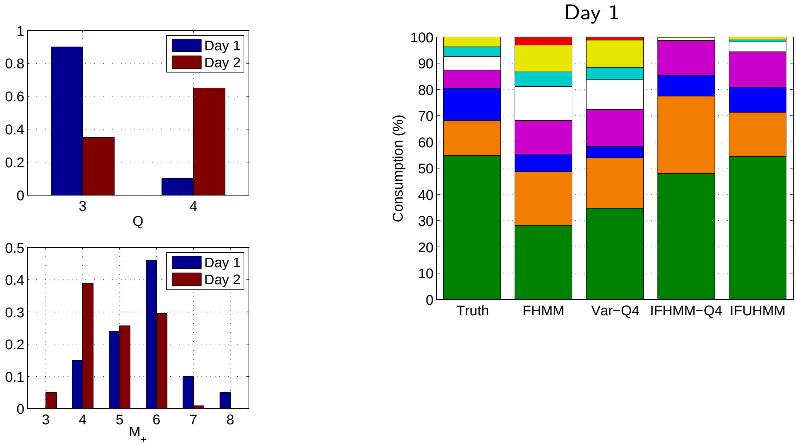
Power Disaggregation

Results for the AMP database (2 days, 8 devices):



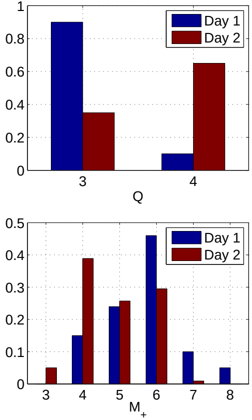
Power Disaggregation

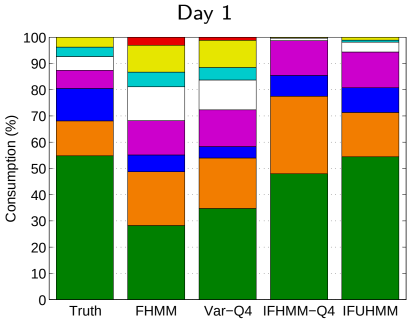
Results for the AMP database (2 days, 8 devices):



Power Disaggregation

Results for the AMP database (2 days, 8 devices):

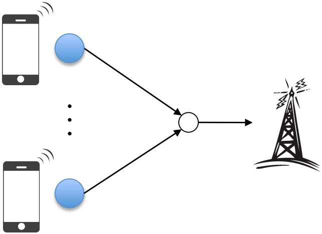




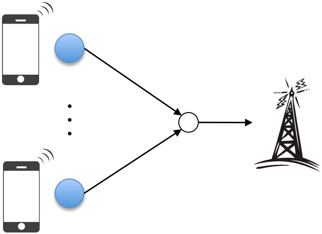
$$\text{accuracy} = 1 - \frac{\sum_{t=1}^{T}\sum_{m=1}^{M}|y_t^{(m)} - \hat{y}_t^{(m)}|}{2\sum_{t=1}^{T}\sum_{m=1}^{M} y_t^{(m)}}$$

FHMM	Var-Q4	IFHMM-Q4	**IFUHMM**
0.36 ± 0.05	0.48 ± 0.06	0.58 ± 0.11	$\mathbf{0.69 \pm 0.10}$

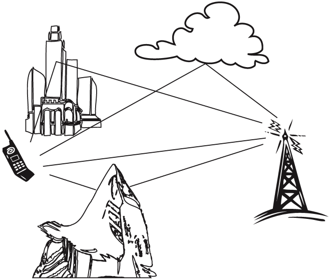
Multiuser Communication System



Multiuser Communication System



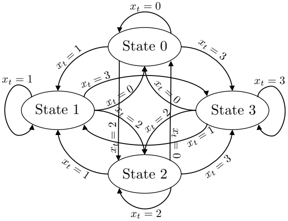
Multipath propagation



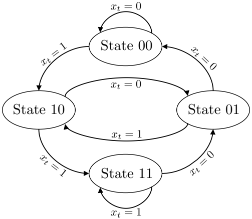
Outline

Finite-Memory Finite State Machine

Finite-Memory FSM: The state depends on the last L inputs x_t.



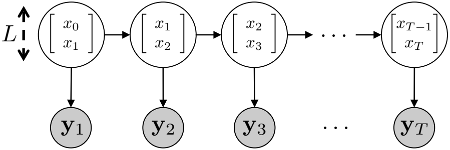
HMM with $Q = 4$ states.
Dense transition probability matrix.



FSM with memory length $L = 2$.
Sparse transition probability matrix.

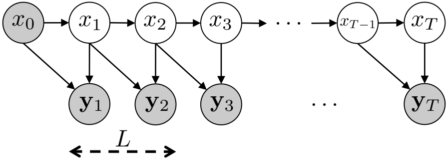
INFINITE FACTORIAL FINITE STATE MACHINE

- HMM representation of an FSM:



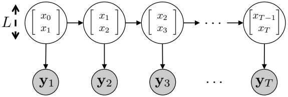
State space cardinality: $|\mathcal{X}|^L$.

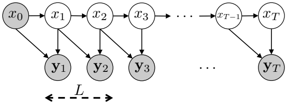
- Alternative representation (likelihood accounts for the memory):



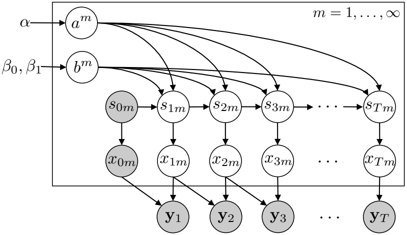
Infinite Factorial Finite State Machine

- The likelihood accounts for the memory.





- Infinite Factorial FSM:
  - $M \to \infty$ parallel FSMs.
  - $\mathbf{S} \sim \mathrm{MIBP}(\alpha, \beta_0, \beta_1)$.
  - Auxiliary variables s_{tm} indicate activity/inactivity.
  - $x_{tm} = 0$ if $s_{tm} = 0$ and $x_{tm} \in \mathcal{A}$ otherwise.



Inference

Inference

MCMC inference algorithm:

1. Propose new parallel FSMs.
   - Slice sampling.
   - Stick-breaking construction.
2. Update hidden states x_{tm}, s_{tm}.
   - Particle Gibbs with ancestor sampling.
3. Remove inactive FSMs.
4. Sample global variables.

Inference

Inference

MCMC inference algorithm:

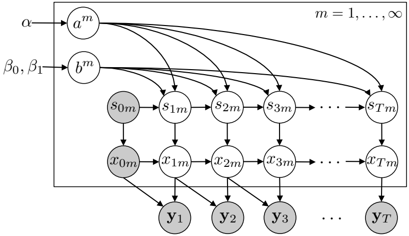
1. Propose new parallel FSMs.
   - Slice sampling.
   - Stick-breaking construction.
2. Update hidden states x_{tm}, s_{tm}.
   - Particle Gibbs with ancestor sampling.
3. Remove inactive FSMs.
4. Sample global variables.

Particle Gibbs with ancestor sampling

- Combines MCMC and SMC.
- Better mixing properties than FFBS.
- Outperforms FFBS:
  - Quadratic complexity with memory L.
  - Can handle more general models.

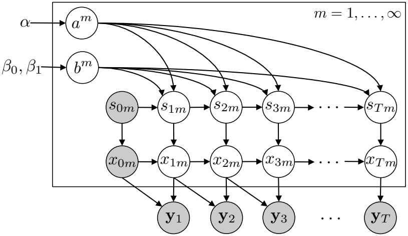
Generalization Of The Model

- Extensions that we can easily handle:
  - States x_{tm} do not necessarily belong to finite set.
  - The state x_{tm} depends on $x_{(t-1)m}$.



Generalization Of The Model

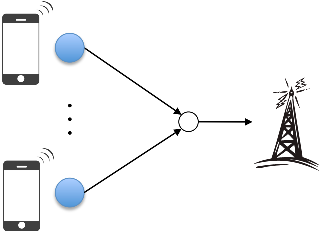
- Extensions that we can easily handle:
  - States x_{tm} do not necessarily belong to finite set.
  - The state x_{tm} depends on $x_{(t-1)m}$.



- Applications:
  - Multi-target tracking.
  - Speech separation.
  - ...

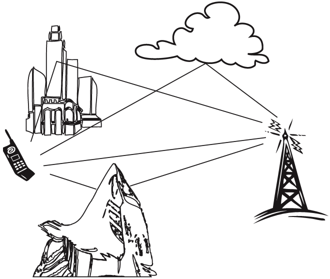
Multiuser Communication System

- Estimate the number of users and the transmitted symbols.
- Machine-to-machine communications:
  - Transmitters switching on and off asynchronously.
  - Short bursts of symbols.
  - Reduce message overhead.
  - 5G systems.



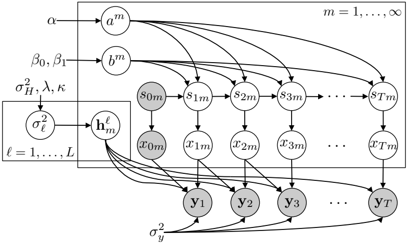
Multiuser Communication System

Multipath propagation



Gaussian observation model

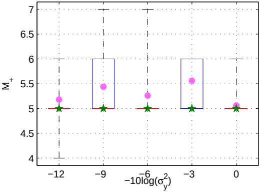
$$\mathbf{y}_t = \sum_{m=1}^{M_+} \sum_{\ell=1}^{L} \mathbf{h}_m^{\ell} x_{(t-\ell+1)m} + \mathbf{n}_t$$



Multiuser Communication System

Synthetic experiment with 5 transmitters and 20 receivers.

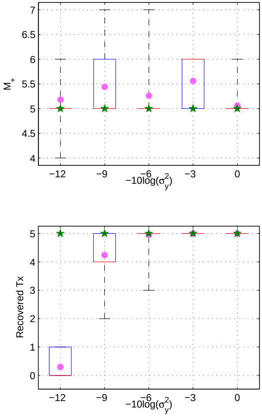
$L = 1$



Multiuser Communication System

Synthetic experiment with 5 transmitters and 20 receivers.

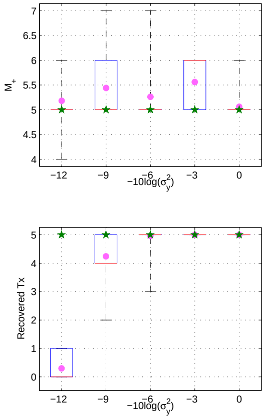
$L = 1$



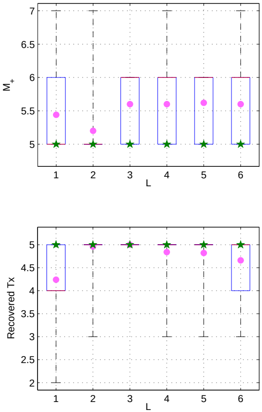
Multiuser Communication System

Synthetic experiment with 5 transmitters and 20 receivers.

$L = 1$



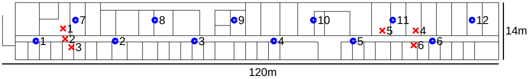
Varying L (−9 dB)



Multiuser Communication System

Wi-Fi experiment:

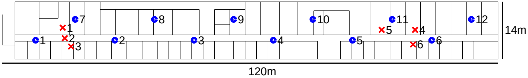
- Ray-tracing software (WISE).
- 6 transmitters, 12 receivers.
- Office at Bell Labs Crawford Hill.



Multiuser Communication System

Wi-Fi experiment:

- Ray-tracing software (WISE).
- 6 transmitters, 12 receivers.
- Office at Bell Labs Crawford Hill.



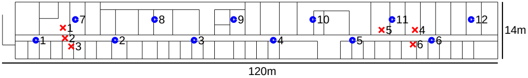
- Recovered transmitters / Inferred M_+:

Algorithm	L = 1	2	3	4	5
PGAS	6/6	6/6	6/6	6/6	6/6
FFBS	3/11	3/11	3/8	1/10	—

Multiuser Communication System

Wi-Fi experiment:

- Ray-tracing software (WISE).
- 6 transmitters, 12 receivers.
- Office at Bell Labs Crawford Hill.



- Recovered transmitters / Inferred M_+:

Algorithm	L				
	1	2	3	4	5
PGAS	6/6	6/6	6/6	6/6	6/6
FFBS	3/11	3/11	3/8	1/10	—

- MSE ($\times 10^{-6}$) of the first channel tap ($\ell = 1$):

Algorithm	L				
	1	2	3	4	5
PGAS	2.58	2.51	0.80	0.30	0.16
FFBS	2.79	1.38	5.53	1.90	—

(noise variance is $\sim 10^{-8}$)

Outline

Conclusions

Contributions

1. Non-Binary Infinite Factorial HMM.
    - MCMC inference.
    - Variational inference.
2. Infinite Factorial Unbounded-State HMM.
    - MCMC inference.
3. Infinite Factorial Finite State Machine.
    - Particle MCMC inference.

Conclusions

Contributions

1. Non-Binary Infinite Factorial HMM.
   - MCMC inference.
   - Variational inference.
2. Infinite Factorial Unbounded-State HMM.
   - MCMC inference.
3. Infinite Factorial Finite State Machine.
   - Particle MCMC inference.

Future Work

- Doubly nonparametric IFHMM.
- Semi-Markov approaches.
- Inference:
  - Scalability.
  - Mixing of MCMC.
  - Online.
- Other applications.
- Time-varying channels.

Thanks for your attention!

Binary IFHMM For Power Disaggregation

- REDD dataset (5 houses, 1 day, 6 devices).
- Binary IFHMM ($Q = 2$).
- Histogram of inferred M_+: