Infinite Factorial Unbounded-Memory Finite State Machines

Francisco J. R. Ruiz, Isabel Valera, Fernando Perez-Cruz

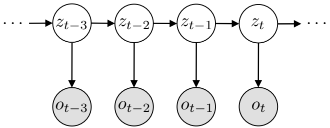
University Carlos III in Madrid

Outline

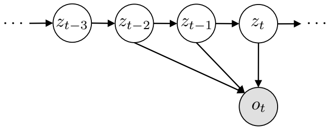
Outline

Introduction (1/3)

Flat fading channel (no ISI):

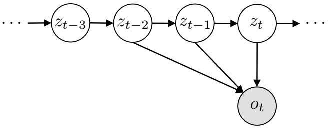


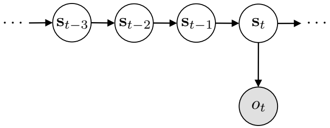
Frequency-selective fading channel (ISI):



Introduction (2/3)

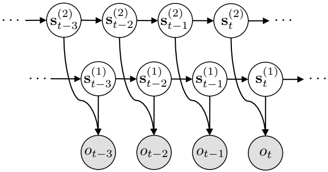
- The frequency-selective fading channel can be equivalently represented as

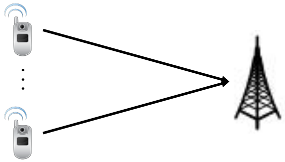




where $\mathbf{s}_t = [z_t,\ z_{t-1},\ \ldots,\ z_{t-r}]$.

- We focus on several transmitters (users) contributing to a single output sequence:





- The channel for each transmitter has its own memory r_m: $\mathbf{s}_t^{(m)} = [z_{tm},\ z_{(t-1)m},\ \ldots,\ z_{(t-r_m)m}]$.

Introduction (3/3)

Goal $\rightarrow$ Fully blind multi-user channel estimation:

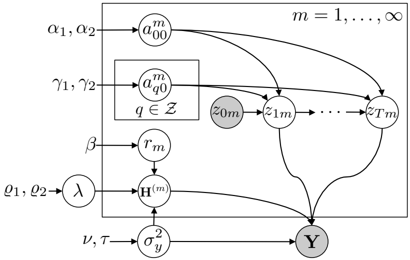
- # Transmitters.
- Channel between each transmitter-receiver pair (channel memory and coefficients).
- Transmitted symbols.

Our proposal:

- Each transmitter-receiver pair as an FSM.
- Bayesian nonparametric model with infinite number of parallel FSMs with unbounded memories.
- Similar to a factorial HMM but with forbidden state transitions.

Outline

Infinite Factorial Unbounded-Memory FSM (1/2)



- Bursts of symbols $\rightarrow$ First-order Markov behavior.
- Inputs (symbols) $z_{tm} \in \mathcal{Z} \bigcup \{0\}$ ($0 =$ "inactive input").
    - e.g., $\mathcal{Z} = \{\pm 1\}$ in BPSK.
- $M \rightarrow \infty$ (but only a finite subset of FSMs become active).
- Memories r_m follow infinite discrete prior distributions.
- σ_y^2 (noise variance) and λ (related to the SNR) are also random.

Infinite Factorial Unbounded-Memory FSM (2/2)

We can write some equations...

$$z_{tm}|z_{(t-1)m}, \mathbf{A}^m \sim \mathbf{a}^m_{z_{(t-1)m}}$$

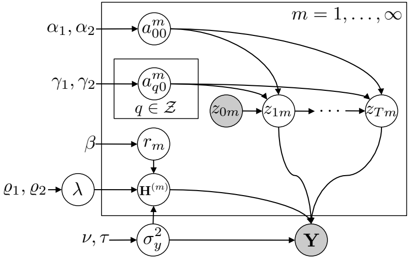
$$a^m_{00}|\alpha_1, \alpha_2 \sim \mathrm{Beta}\left(\alpha_1, \frac{\alpha_2\alpha_2}{M}\right)$$

$$a^m_{q0}|\gamma_1, \gamma_2 \sim \mathrm{Beta}(\gamma_1, \gamma_2)$$

$$r_m|\beta \sim \mathrm{Poisson}(\beta)$$

$$\lambda|\varrho_1, \varrho_2 \sim \mathrm{Gamma}(\varrho_1, \varrho_2)$$

$$\sigma^2_y|\nu, \tau \sim \text{Inverse-Gamma}(\nu, \tau)$$



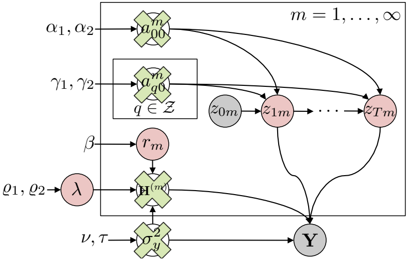
$$p(\mathbf{H}^{(m)}|\sigma^2_y, r_m, \lambda) = \frac{1}{(2\pi\sigma^2_y/\lambda)^{\frac{D(r_m+1)}{2}}} \exp\left\{-\frac{\lambda}{2\sigma^2_y}\mathrm{trace}\left[(\mathbf{H}^{(m)})^\top \mathbf{H}^{(m)}\right]\right\}.$$

$$\mathbf{y}_t = \sum_{m=1}^{M}\sum_{i=0}^{r_m} z_{(t-i)m}\mathbf{h}^{(m)}_i + \mathbf{n}_t = \sum_{m=1}^{M}\mathbf{s}^{(m)}_t\mathbf{H}^{(m)} + \mathbf{n}_t,$$

$$\mathbf{n}_t|\sigma^2_y \sim \mathcal{N}(\mathbf{0}, \sigma^2_y)$$

Outline

Inference (1/2)

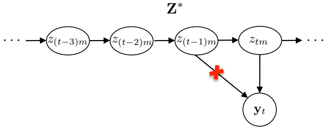


- We analytically integrate out $\mathbf{H}^{(m)}$, σ_y^2, a_{00}^m and a_{q0}^m.
- Algorithm:
  - **Step 1:** Sample all the inputs z_{tm}.
  - **Step 2:** Sample the memories r_m of the active FSMs.
  - **Step 3:** Consider splitting a FSM into two or merging two into one.
  - **Step 4:** Sample the hyper-parameter λ.
- Gibbs sampling (Step 1) and MH (Steps 2-4).

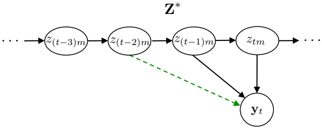
Inference (2/2)

In Step 2 (sample r_m), we can perform a time-shift of the input sequence.

- Examples of memory sampling without time-shift:

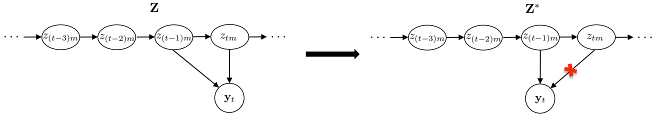


$r^*_m = 0 < r_m = 1$

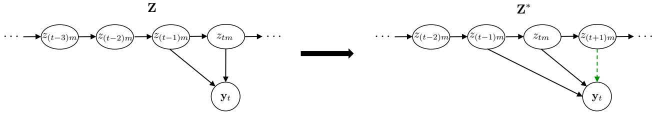


$r^*_m = 2 > r_m = 1$

- Examples of memory sampling with time-shift:



$r^*_m = 0 < r_m = 1$



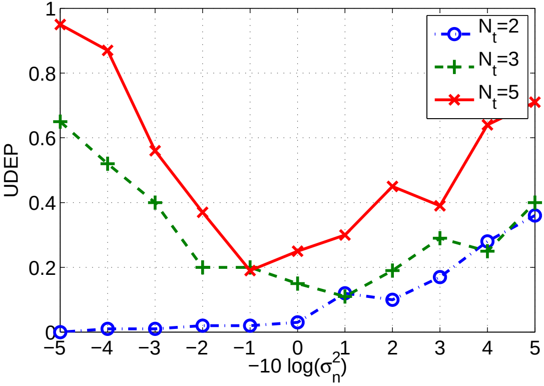
$r^*_m = 2 > r_m = 1$

Outline

Experiments (1/4)

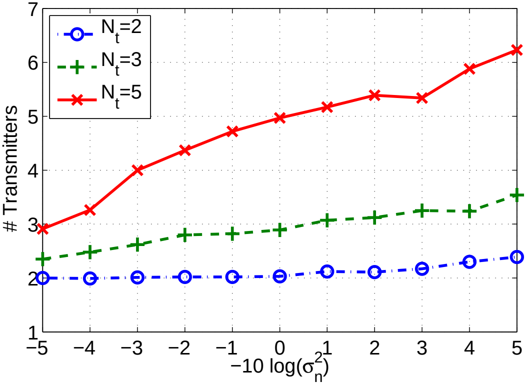
- BPSK system.
- Observation period of $T = 400$.
- Each transmitter sends a burst of symbols (random initial instant and random duration $\in [T/5, T/2]$).
- $N_r = 10$ receiving antennas.
- Three scenarios:
  - *Scenario A*: $N_t = 2$ transmitters (memories $r_1 = r_2 = 1$).
  - *Scenario B*: $N_t = 3$ transmitters (memories $r_1 = 2$, $r_2 = 1$, $r_3 = 0$).
  - *Scenario C*: $N_t = 5$ transmitters (memories $r_1 = r_2 = 3$, $r_3 = 2$, $r_4 = r_5 = 1$).
- Variation of the SNR.
  - 100 independent simulations for each value.

Experiments (2/4)

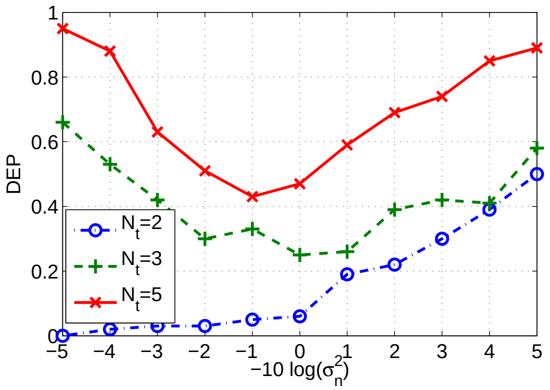


- User detection error probability (UDEP): Error probability of detecting the number of transmitters.

Experiments (3/4)



Experiments (4/4)



- Detection error probability (DEP): Error probability of detecting both the number of transmitters and the channel memories.

Outline

Conclusions

- First attempt to learn both the number of transmitters and the channel memory.
- BNP model that blindly (without pilot symbols) learns the network structure.
- Negligible error when the network topology has been correctly identified.
- Open issues:
  - Considering overhead of each transmitter.
  - The IBP prior and MCMC inference algorithm underperform when the noise level drops.
- Applicable to other problems dealing with time series.