Multichain Hidden Markov Models

Nathalie Peyrard
INRAE, Unité MIA, Toulouse

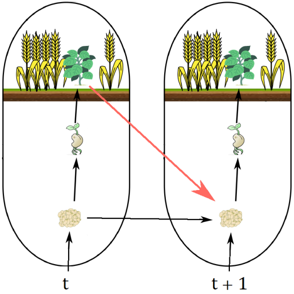
with the 'multichain' working group of ANR project HSMM-INCA
Hanna Bacave, JB Durand, Alain Franc, Sandra Plancade, Régis Sabbadin

MaSeMo workshop - July 2, 2025

Outline of the presentation

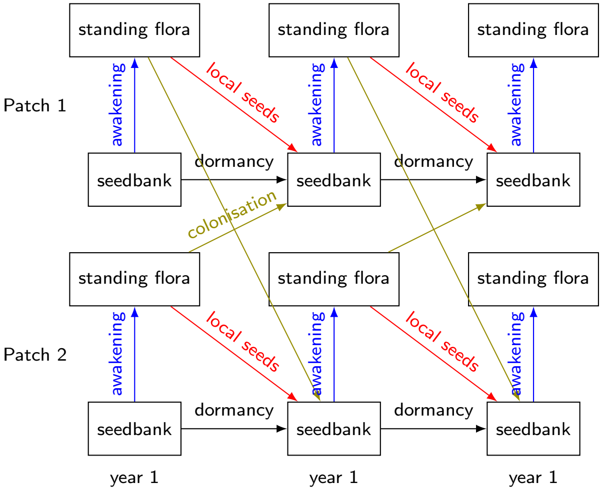
Annual plants dynamics: one patch

- How can we estimate the key parameters of annual plant dynamics based only on observations on standing flora?



- Dormancy: seed survival
- Seedbank: **not visible**
- Dynamics: germination, seed production, dormancy and colonisation

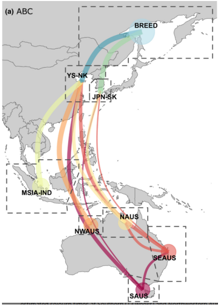
Annual plants dynamics: several patches



[with S. Le Coz at MIAT and P.-O. Cheptou at CEFE Montpellier]

Birds migration routes

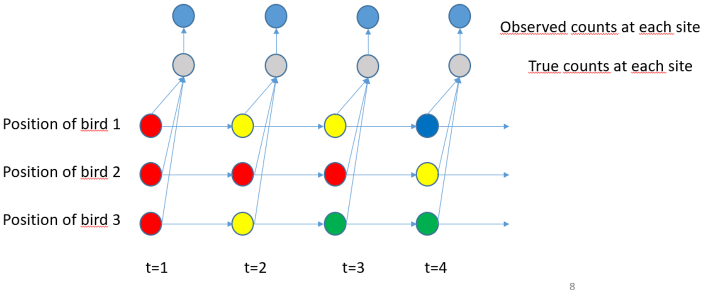
- Which are the main routes used by the shorebirds?



- several stop overs
- possible routes form a network
- bird trajectories **not available**
- data: imperfect counts (eBird data)

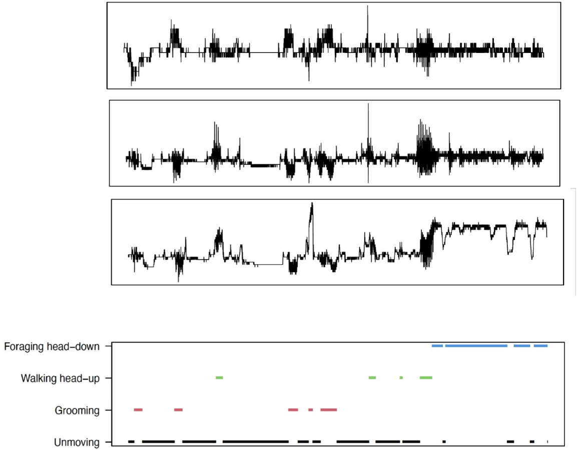
[with R. Sabbadin, R. Trépos and M.-J. Cros at MIAT, and S. Nicol at CSIRO Brisbane]

Birds migration routes



Deers behaviour

- How to reconstruct successive activity phases from accelerometry data?

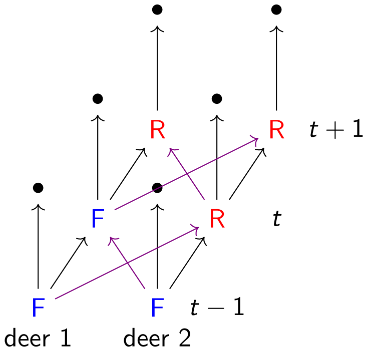


[S. Plancade at MIAT and N. Morellet, N. Ranc at CEFS]

Deers behaviour: several individuals

- Several deers in a same location: change in one animal's activity phase (**hidden**) may impact behaviours of other animals

$$\begin{cases} \mathbb{P}[Z^1_{t+1} = running | Z^1_t = foraging, Z^2_t = running] & = & 0.7 \\ \mathbb{P}[Z^1_{t+1} = running | Z^1_t = foraging, Z^2_t = foraging] & = & 0.1 \end{cases}$$



A common framework for dynamics of temporal processes in interaction with hidden states?

- Common features :
  - several individual dynamics
  - that are not independent
  - with two levels : hidden state (regime, class, ...) and observations (hidden+noise, proxy of hidden, ...)
- Differences : interaction does not always take place at the same level
- Other examples : epidemiology, earthquakes, signal processing...

Natural framework: HMM (already largely used for one individual dynamics)

$$\begin{array}{ccccccccccc}
Y_t & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet \\
 & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
Z_t & \circ & \longrightarrow & \circ & - - - - - & \circ & \longrightarrow & \circ & - - - - - & \circ & \longrightarrow & \circ \\
 & 0 & & 1 & & t & & t+1 & & T-1 & & T
\end{array}$$

This talk is about

- reviewing existing HMM based models for modeling interacting temporal processes with hidden states
- proposing a general definition of multichain HMM
- discussing estimation complexity

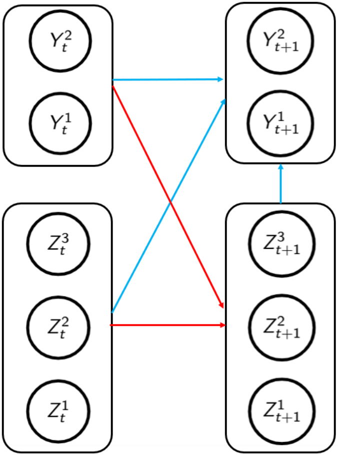
Starting point

Variables:

- Hidden: $\boldsymbol{Z}_t$ with C components (C chains): $\boldsymbol{Z}_t = (Z_t^1, \ldots, Z_t^C)$ with $Z_t^c \in \Omega_{Z^c}$.
- Observed: $\boldsymbol{Y}_t$ with O components (O observations): $\boldsymbol{Y}_t = (Y_t^1, \ldots, Y_t^O)$ with $Y_t^o \in \Omega_{Y^o}$.

Proposition of a hierarchy of definitions built by successively adding conditional independencies assumptions

General MHMM



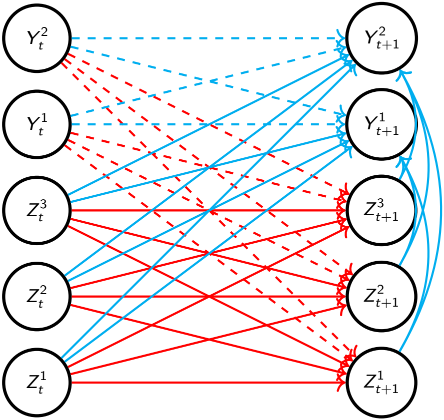
- $(\boldsymbol{Z}_t, \boldsymbol{Y}_t)_{t\in\mathbb{N}}$ is Markovian

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{Z}_{t+1}, \boldsymbol{Y}_{t+1} \mid \boldsymbol{Z}_t, \boldsymbol{Y}_t) \\
&= \mathbb{P}(\boldsymbol{Y}_{t+1} \mid \boldsymbol{Z}_{t+1}, \boldsymbol{Z}_t, \boldsymbol{Y}_t) \times \mathbb{P}(\boldsymbol{Z}_{t+1} \mid \boldsymbol{Z}_t, \boldsymbol{Y}_t) \\
&= \textit{emission} \times \textit{transition}
\end{aligned}$$

- limit: space needed for representation $|\Omega_Z|^{2C} \times |\Omega_Y|^{2O}$

MHMM with conditional independencies in emission and transition

- in emission (resp. transition) terms, additional assumption of conditional independencies of the Y^o_{t+1} (the Z^c_{t+1})
- factorisation
  - $\mathbb{P}(\boldsymbol{Z}_{t+1} \mid \boldsymbol{Z}_t, \boldsymbol{Y}_t) = \prod_{c=1}^{C} \mathbb{P}(Z^c_{t+1} \mid \boldsymbol{Z}_t, \boldsymbol{Y}_t)$
  - $\mathbb{P}(\boldsymbol{Y}_{t+1} \mid \boldsymbol{Z}_{t+1}, \boldsymbol{Z}_t, \boldsymbol{Y}_t) = \prod_{o=1}^{O} \mathbb{P}(Y^o_{t+1} \mid \boldsymbol{Z}_{t+1}, \boldsymbol{Z}_t, \boldsymbol{Y}_t)$



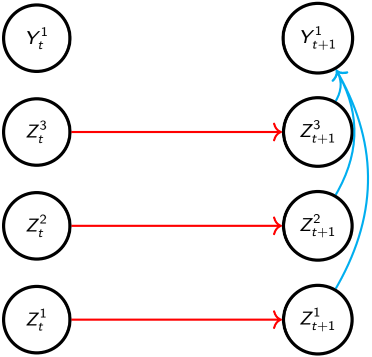
MHMM with conditional independencies in emission and transition

- gain in representation space
  - $C|\Omega_Z|^{C+1} \times |\Omega_Z|^{O}$ terms for transition
  - $O|\Omega_Z|^{2C} \times |\Omega_Z|^{O+1}$ terms for emission
- factorisation meaningful in many applications (weeds, birds, deers)
- some dynamical graphical models from literature are examples MHMM-CI

Subcase 1 of MHMM-CI

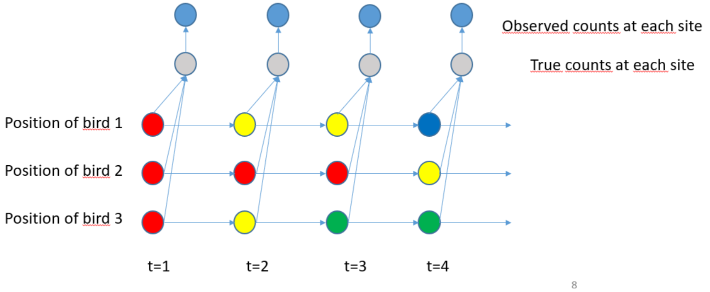
Factorial HMM (FHMM, [Ghahramani and Jordan, 1997])

- $O = 1$
- independent hidden Markov chains
- joint emission of the observation



Subcase 1 of MHMM-CI

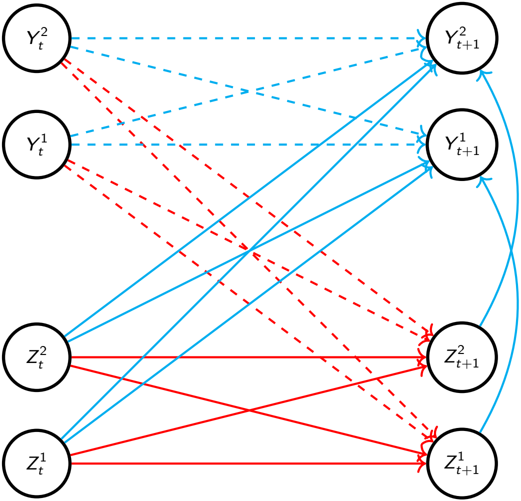
Application: birds migration routes



Subcase 2 of MHMM-CI

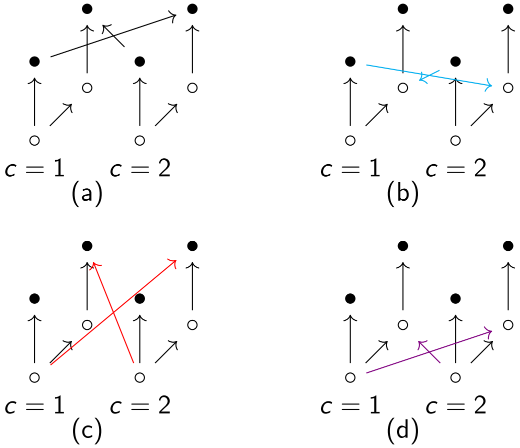
1to1-MHMM-CI: one observation per hidden variable

- $O = C$
- $\mathbb{P}(Y^c_{t+1} \mid \boldsymbol{Z}_{t+1}, \boldsymbol{Z}_t, \boldsymbol{Y}_t) = \mathbb{P}(Y^c_{t+1} \mid Z^c_{t+1}, \boldsymbol{Z}_t, \boldsymbol{Y}_t)$



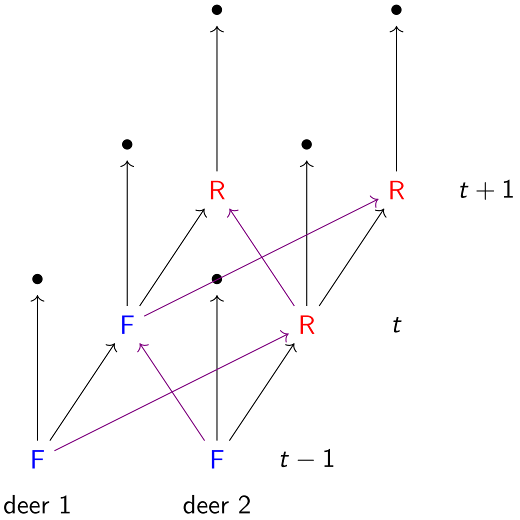
Subcase 2 of MHMM-CI

Building **1to1-MHMM-CI** by coupling HMM structures



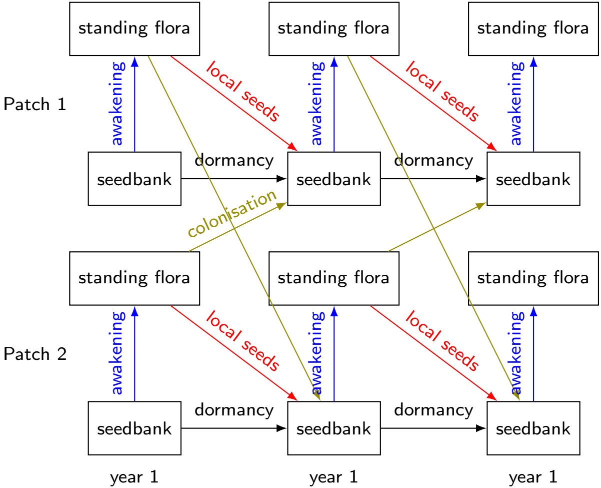
Subcase 2 of MHMM-CI

Application: deers behaviours as a CHMM [Brand, 1997]

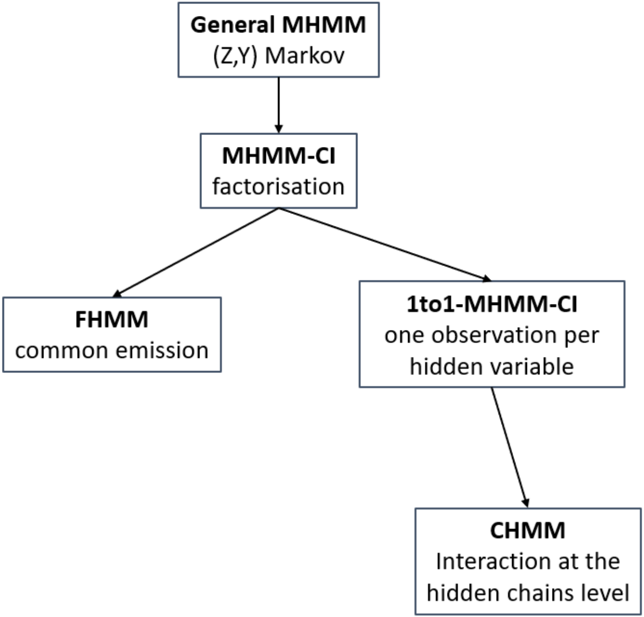


Subcase 2 of MHMM-CI

Application: weeds dynamics [Le Coz et al., 2019]



Typology


General MHMM
(Z,Y) Markov
MHMM-CI
factorisation
FHMM
common emission
1to1-MHMM-CI
one observation per
hidden variable
CHMM
Interaction at the
hidden chains level

Tractability of MHMM estimation

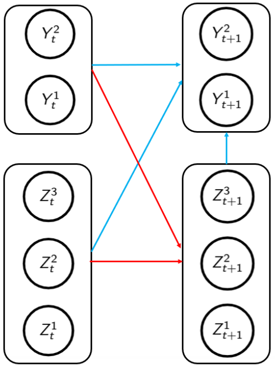
Estimation of the emission and the transition

Context

- EM: classical algorithm for computing the Maximum Likelihood Estimator in models with hidden (latent) variables
- EM for HMM
  - standard tool (many packages)
  - relies on the forward-backward algorithm
  - time complexity $\mathcal{O}(TK^2)$ (T, length of observation sequence, K, number of hidden states)

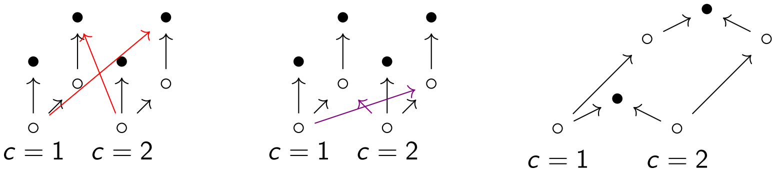
Question: is it still stractable for MHMM? Does it depends on the dependency structure?

Complexity of EM for a General MHMM



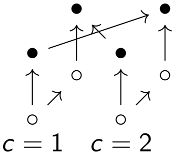
- a MHMM is an HMM with large state space (up to some adaptation)
- state space of $\boldsymbol{Z}_t$ is of dimension K^C
- EM complexity is $\mathcal{O}(TK^{2C})$: **untractable**

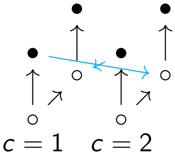
Complexity of EM for some MHMM-CI



- Hidden chains are not independent conditionally to the observations
- The forward-backward recursions must be run on the multidimentional hidden state
- EM complexity is $\mathcal{O}(TK^{2C})$: **untractable**

Complexity of EM for some MHMM-CI





- Hidden chains are independent conditionally to the observations
- A forward-backward algorithm is run independently for each chain
- EM complexity is $\mathcal{O}(CTK^2)$: **tractable**

Motivation

Multichain HMM: a typology

Multichain HMM: tractability of estimation

Conclusion

Conclusion

- A general framework that includes and generalizes well known structures of multichain HMMs (FHMM, CHMM)
- Various dependencies structures = can cover many different applications
- Parameter estimation
  - exact EM in general exponential in number of chains but can be linear for some structures
  - if exponential complexity, classical options for approximate inference are Monte Carlo, variational approximation or simplifying assumptions
- Beyond: multichain HSMM
  - the classical representation by state and duration not adapted anymore
  - we propose a rigorous definition based on hazard rate ... but that would be another talk

References

Brand, M. (1997).

Coupled hidden Markov models for modeling interacting processes.

Technical Report 405, MIT Media Laboratory, Cambridge.

Ghahramani, Z. and Jordan, M. (1997).

Factorial hidden Markov models.

Machine Learning, 29(2-3):245–273.

Le Coz, S., Cheptou, P.-O., and Peyrard, N. (2019).

A spatial Markovian framework for estimating regional and local dynamics of annual plants with dormancy.

Theoretical Population Biology.