

---

# Heterogeneous Continuous Dynamic Bayesian Networks with Flexible Structure and Inter-Time Segment Information Sharing

## Supplementary Material

---

**Frank Dondelinger**

FRANKD@BIOSS.AC.UK

Biomathematics and Statistics Scotland, JCMB, EH9 3JZ, Edinburgh, UK  
Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh

**Sophie Lèbre**

SOPHIE.LEBRE@LSIIT-CNRS.UNISTRA.FR

Université de Strasbourg, LSIIT - UMR 7005, 67412 Illkirch, France

**Dirk Husmeier**

DIRK@BIOSS.AC.UK

Biomathematics and Statistics Scotland, JCMB, EH9 3JZ, Edinburgh, UK

## 1. Introduction

This document contains extra information for the ICML 2010 paper “Heterogeneous Continuous Dynamic Bayesian Networks with Flexible Structure and Inter-Time Segment Information Sharing”. The figures and tables contained herein had to be omitted from the main paper due to space constraints. In the following sections you will find:

- A graphical representation of the sequential information-sharing model (HetDBN-SI).
- A table with all p-values from the paired t-tests between different information sharing methods.
- Diagrams of the reconstructed graphs from the *Drosophila* muscle development dataset.

Please refer to the main paper for any questions about the methods or datasets.

## 2. Graphical representation of the sequential information sharing model

### 3. Paired t-tests for method comparisons

Table 1 shows paired t-tests between the AUROC scores of the different information sharing methods. We see that when changing the number of changes between segments, all differences in AUROC scores are significant except for the difference between HetDBN-SI and HetDBN-GI when the number of changes is 1.

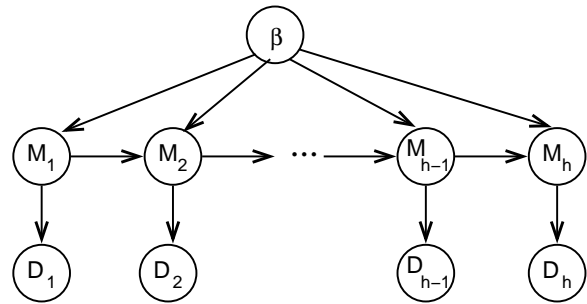


Figure 1. Sequential information sharing scheme, whereby each network  $M_h$  depends on the previous network  $M_{h-1}$ , through an exponential prior on the number of differences between the two networks, regularized by a hyperparameter  $\beta$ .

For the AUPRC score, the results are similar, except that when the number of changes is 2, the difference between HetDBN-0 and HetDBN-GI is no longer significant. For the TFP5 score, the difference between HetDBN-0 and HetDBN-GI is never significant when the number of changes is  $> 0$ .

### 4. Reconstruction of *Drosophila* Muscle Development Gene Network

In this section, we present a graphical representation of the networks recovered using the heterogeneous DBN mode with sequential information sharing (HetDBN-SI). These networks have been constructed by discarding all edges with marginal posterior probability  $< 0.25$ . Due to the enforced sparsity introduced by

Table 1. P-values from paired t-tests for AUROC scores of HetDBN-0, HetDBN-SI and HetDBN-GI for full DBN model when varying the mean number of changes between segments (TOP) AUROC Scores (MIDDLE) AUPRC Scores (BOTTOM) TFP5 Scores.

(a)

CHANGE NUM	0	1	2
HETDBN-0 vs -SI	$< 1e-5$	$< 1e-5$	$< 1e-5$
HETDBN-SI vs -GI	$< 1e-5$	0.92	$< 1e-2$
HETDBN-GI vs -0	$< 1e-5$	$< 1e-5$	$< 1e-4$

(b)

CHANGE NUM	0	1	2
HETDBN-0 vs -SI	$< 1e-5$	$< 1e-4$	$< 1e-3$
HETDBN-SI vs -GI	$< 1e-5$	0.04	$< 1e-3$
HETDBN-GI vs -0	$< 1e-5$	$< 1e-3$	0.26

(c)

CHANGE NUM	0	1	2
HETDBN-0 vs -SI	$< 1e-5$	$< 1e-3$	0.04
HETDBN-SI vs -GI	$< 1e-5$	$< 1e-3$	$< 1e-3$
HETDBN-GI vs -0	$< 1e-5$	0.19	0.26

our prior on the number of parents, a higher threshold would have led to overly sparse networks, which might miss the subtle distinctions between the phases.

The recovered networks are presented in Figure 2. An interpretation of these findings, as well as a comparison to networks recovered by previous methods, can be found in the main paper.

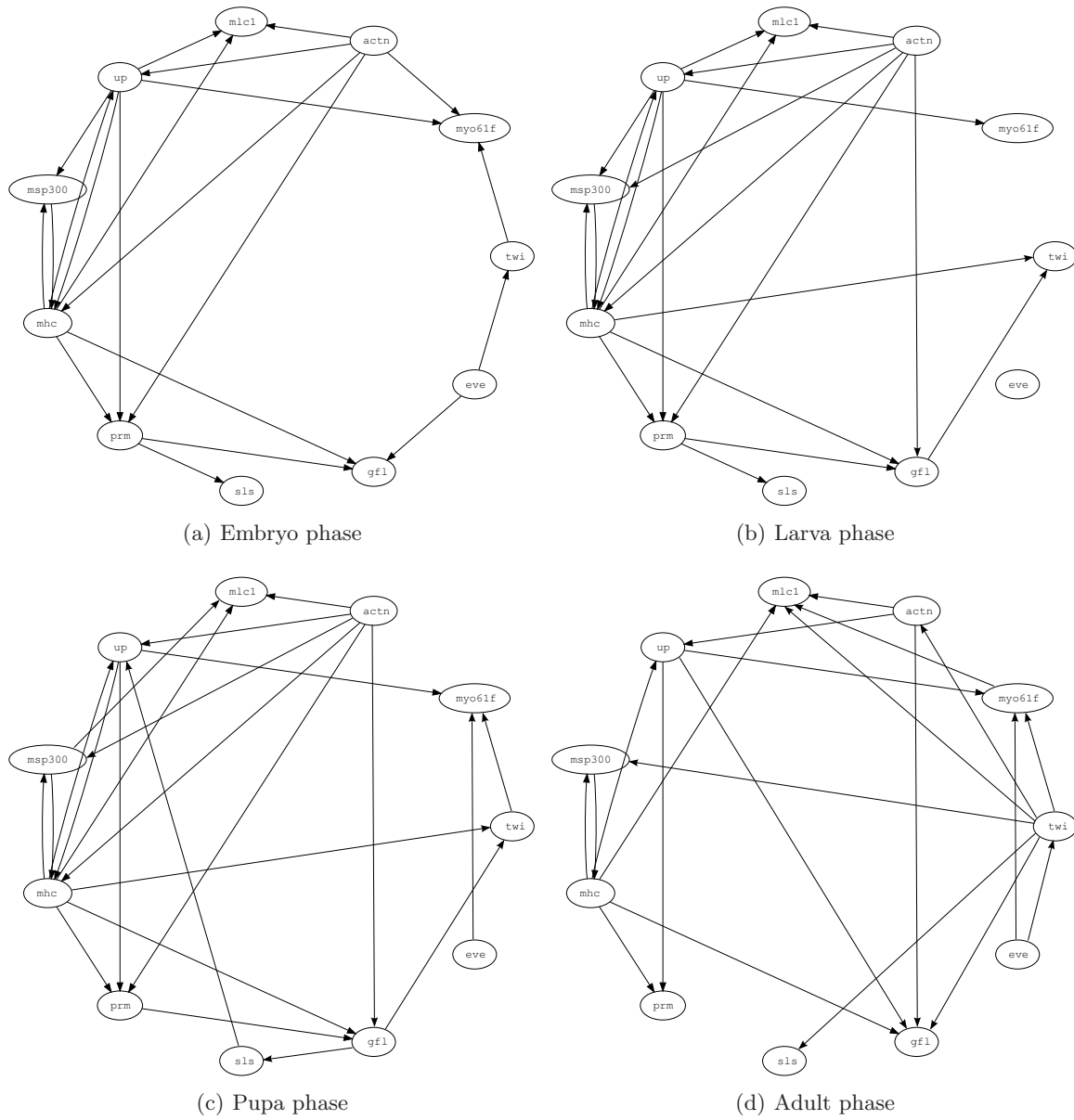


Figure 2. Recovered networks for each of the morphological phases in the development of *Drosophila melanogaster*, using the HetDBN-SI method.