

Supplementary materials for the paper: Context-Specific Nested Effects Models

Yuriy Sverchkov Yi-Hsuan Ho Audrey Gasch
Mark Craven

January 26, 2018

1 Detailed derivation of the likelihood of a k-CSNEM

In this section we derive Equation (??). Given that $\Gamma, \Gamma^1, \dots, \Gamma^k, \Theta, \zeta, \Theta^1, \dots, \Theta^k$, and Ψ are defined as in Section ??, we introduce two inverses to the index mapping ζ : $[\kappa(e) = i] \equiv [e \in \mathcal{E}_i]$ and $[\eta_i(e) = j] \equiv [\zeta(i, j) = e]$.

The log-likelihood of the CSNEM can be written as

$$\log \prod_{i=1}^k L(\Gamma^i, \Theta^i) \stackrel{\text{Using (??)}}{=} \sum_{i=1}^k \sum_{a \in \mathcal{A}} \sum_{j=1}^{|\mathcal{E}_i|} (\Gamma^i \Theta^i)_{aj} R_{ja}^i + \log L(N) \quad (1)$$

$$= \sum_{a \in \mathcal{A}} \sum_{e \in E} (\Gamma^{\kappa(e)} \Theta^{\kappa(e)})_{a\eta_{\kappa(e)}(e)} R_{\eta_{\kappa(e)}(e)a}^{\kappa(e)} + \log L(N) \quad (2)$$

$$= \sum_{a \in \mathcal{A}} \sum_{e \in E} (\Gamma^{\kappa(e)} \Theta^{\kappa(e)})_{a\eta_{\kappa(e)}(e)} R_{ea} + \log L(N) \quad (3)$$

$$= \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{A}} \sum_{e \in E} \Gamma_{ab}^{\kappa(e)} \Theta_{b\eta_{\kappa(e)}(e)}^{\kappa(e)} R_{ea} + \log L(N) \quad (4)$$

$$= \sum_{e \in E} \sum_{a \in \mathcal{A}} \sum_{b=|\mathcal{A}|(\kappa(e)-1)+1}^{|\mathcal{A}|\kappa(e)} \Gamma_{[(\kappa(e)-1)|\mathcal{A}|+a],b} \Theta_{be} R_{ea} + \log L(N) \quad (5)$$

$$= \sum_{e \in E} \sum_{a \in \mathcal{A}} \sum_{b=1}^{|\mathcal{A}|k} \Gamma_{[(\kappa(e)-1)|\mathcal{A}|+a],b} \Theta_{be} R_{ea} + \log L(N) \quad (6)$$

$$= \sum_{e \in E} \sum_{a \in \mathcal{A}} (\Gamma \Theta)_{[(\kappa(e)-1)|\mathcal{A}|+a],e} R_{ea} + \log L(N) \quad (7)$$

$$= \sum_{e \in E} \sum_{a \in \mathcal{A}} \sum_{i=1}^k (\Gamma \Theta)_{[(i-1)|\mathcal{A}|+a],e} R_{ea} + \log L(N) \quad (8)$$

$$= \sum_{e \in E} \sum_{a \in \mathcal{A}} (\Psi \Gamma \Theta)_{ae} R_{ea} + \log L(N) \quad (9)$$

$$= \sum_{a \in \mathcal{A}} (\Psi \Gamma \Theta R)_{aa} + \log L(N) \quad (10)$$

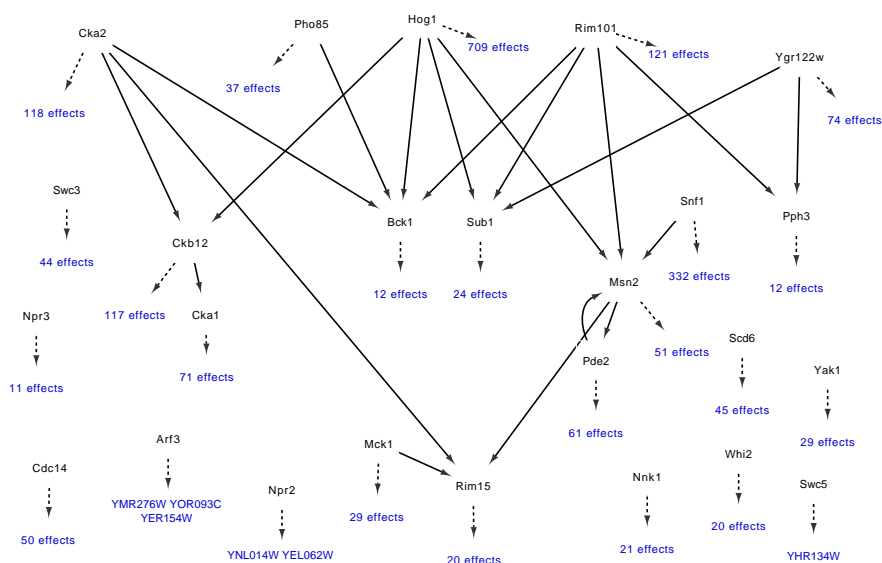
$$= \text{tr}(\Psi \Gamma \Theta R) + \log L(N) \quad (11)$$

$$= \text{tr}(\Gamma \Theta (R \Psi)) + \log L(N) \quad (12)$$

Where we get (6) because Θ_{be} is necessarily 0 for $b < (\kappa(e) - 1)|\mathcal{A}| + 1$ or $b > \kappa(e)|\mathcal{A}|$. Similarly, we get (8) because $(\Gamma \Theta)_{xe}$ is necessarily 0 for $x < (\kappa(e) - 1)|\mathcal{A}| + 1$ or $x > \kappa(e)|\mathcal{A}|$ because of the block diagonal nature of Γ and because Θ can have at most one nonzero in each column.

2 NEM Lerner from the *S. cerevisiae* NaCl stress knockout microarray data

The NEM network learned from *S. cerevisiae* NaCl stress knockout microarray data. Action nodes are in black and effect nodes are in blue. Effects are grouped according to the actions to which they are attached. Where the number of effects in a group is less than 10, the effects are listed. Where it is 10 or more, the number of effects in the group is shown.



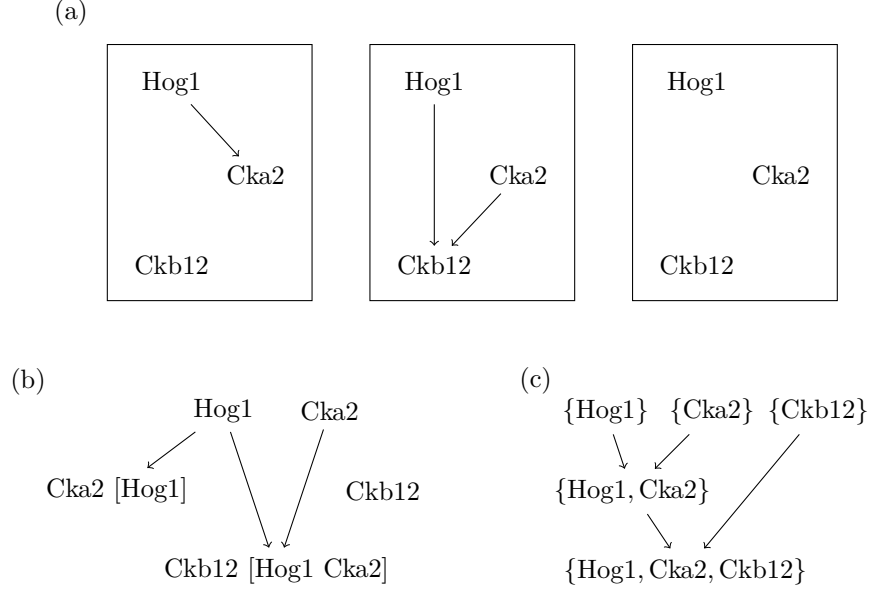


Figure 1: Building a CSNEM from a mixture of NEMs. (a) Three NEMs that compose a mixture. (b) A single graph obtained by an edge-preserving merge of the three NEMs. (c) The corresponding action-set graph.

3 Compact Visualization and Identifiability of a k -CSNEM

Having obtained k NEMs and the corresponding partitioning of the effect set, a single graph can be composed by merging all action nodes across the graphs that have the same ancestors (are reachable from the same set of actions). Figure 1 is an example: Fig. 1 (a) shows three graphs that describe the structures of three NEMs that compose a mixture, and Fig. 1 (b) is the result of merging them. Note that Hog1 is reachable from no nodes but itself in all three NEMs. Consequently, in the compact CSNEM, there is only one version of Hog1. In contrast, Cka2 is reachable from Hog1 in one of the NEMs, and is only reachable from itself in the others, which is why it has two contexts in the CSNEM. Similarly, Ckb14 is reachable from both Hog1 and Cka2 in one of the three NEMs, but not the others, and has two contexts as well. To keep track of the various contexts we append the list of genes from which a context is reachable when displaying the graph, e.g. the context of Cka2 that is reachable from Hog1 is labeled ‘Cka2 [Hog1]’, while the context that is not reachable from other nodes is labeled simply ‘Cka2.’ This is particularly helpful when viewing graphs with many nodes and many contexts.

The merged graph in Fig. 1 (b) preserves the edges that were present in

the mixture of NEMs, but it is not necessarily a unique maximizer of the likelihood, but a member of an equivalence class of equally likely CSNEMs. What characterizes the equivalence class is the set of *inclusive ancestries* of the nodes in the CSNEM. The inclusive ancestry of a node is a set of actions containing the action at the node and all actions from which it is reachable: e.g. the inclusive ancestry of the Cka2 node in the leftmost NEM in Fig. 1 (a) is {Hog1, Cka2}, while the inclusive ancestry of the Cka2 node in the middle NEM is simply {Cka2}. The set inclusive ancestries for the example of Fig. 1 is therefore $\{\{Hog1\}, \{Cka2\}, \{Ckb12\}, \{Hog1, Cka2\}, \{Hog1, Cka2, Ckb12\}\}$. This set of sets can be uniquely graphically represented in an *action-set graph*, an inverted Hasse Diagram of the sets ordered by inclusion (Hasse Diagrams are arranged with the smallest elements at the bottom), an example of which is in Fig. 1 (c). Any two CSNEMs with identical sets of inclusive ancestries necessarily have the same set of unique accessibility matrix columns $\Gamma_{\cdot a}^i : i \in \{1, \dots, k\}, a \in \mathcal{A}$ columns, and consequently, have the same likelihood for likelihood-maximizing attachments Θ . The characterization of equivalence classes in terms of inclusive ancestry sets relates to previous results about NEM identifiability: for transitively closed Γ , cycles form fully connected components that can be merged into single nodes, [markowetz2007nested] yielding a graph that is isomorphic to the action-set graph, with the (often singleton) set of actions S at a node in the NEM being the set difference of the set of actions A in the corresponding node in the action-set graph and the union of the sets of actions at its parent nodes $S = A \setminus \cup \text{pa}(A)$. This can also be extended to the case of non-transitive Γ and the result on identifiability of non-transitive NEMs up to cycle reversals, [tresch2008structure] the full discussion of which we omit here for brevity.

Note that while set of ancestries characterizes the likelihood equivalence class, the posterior maximized by MC-EMiNEM would be, for example, higher for CSNEMs with less edges in Γ under a sparsifying edge prior. There are advantages and disadvantages to each of the possible representations of a CSNEM in terms of the information preserved about the solution reached by the inference algorithm and the interpretability of the graph. In our presentation of the network learned from real data below we use the edge-preserving merged CSNEM graph.