Constructing a selection function first, rank the latent variables according to the affinity computed in Eq. (4) to output the H* most relevant latent states \( \mathbf{K}_0 \) for the data point \( \mathbf{y}^{(n)} \).

**GP-select:** Learn affinity with GP regression

**Algorithm**

For EM iterations \( t = 1, \ldots, T \) do

1. Compute affinity of all latent variables \( \mathbf{p}^{(t)} \): (5)
2. Compute truncated posterior \( \mathbf{q}^{(t)}(\mathbf{s}) \), E-step: (2)
3. Update model parameters in M-step
4. Store \( \mathbf{p}^{(t)} \) for \( \mathbf{p}^{(t)} \) in EM iteration \( t + 1 \)
5. End for

**Experiments**

**Sparse coding models**

<table>
<thead>
<tr>
<th>Binary SC</th>
<th>Spike &amp; Slab SC</th>
<th>Nonlinear Spike &amp; Slab SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s \sim \mathcal{N}(0, \delta^2) )</td>
<td>( s \sim \mathcal{N}(0, \mu, \Sigma) )</td>
<td>( s \sim \mathcal{N}(0, \delta^2) )</td>
</tr>
<tr>
<td>( y \sim \mathcal{N}(\mathbf{y}</td>
<td>\mathbf{h}, \mathbf{r}^2) )</td>
<td>observations</td>
</tr>
</tbody>
</table>

**Data:** \( N = 2, 000 \) with \( D = 5 \) x 5 obs dims & \( H = 10 \) latent dims/bars gen. by each model, with GP-select to preselect \( H^* = 5 \) dims

**Show:** final EM fit; GP-select converges to GT params, \( W_{GP-select} \)

**Gaussian mixture model**

**Data:** \( C = 3 \) clusters, GP-select to preselect \( C^* = 2 \) clusters

**Show:** using the wrong selection function can do harm (i.e. miss patterns); sel. funcs need to be flexible and possibly nonlinear

**Translation invariant occlusive models [1]**

**Problem:** locate objects in scene (A), with massive latent space complexity \# of obj. locations exponentiated by \# of objects.

**Speed:** partial incomplete Cholesky approx to for faster GP regression computation, update GP hyperparams every 5 EM its

**Show:** all 3 variants of GP-selection learn all objects (B) with accuracy equivalent to hand-crafted selection (C & D)