We show empirically that the benefits of the latter method are more significant as the number of variables increases. We derive a related variant of belief propagation which is better suited for information propagation in social networks.

We consider influential subset selection in social networks. They show that entropy and mutual information (under mild conditions) are monotone functions.

The key results of this paper are:

- Complexity Reduction due to Sparsity
- Incremental Belief Propagation (IncrBP)

**Motivation**

- Information planning considers the problem of finding the most informative subset of measurements for some inference problem.
- Optimal selection is generally intractable as the number of latent and measurement nodes increases.
- For submodular problems, approximate techniques provide nearly optimal solutions.
- These techniques are usually one-step greedy approaches that select the best measurement at each step given past selections.
- Prior analyses often assume an oracle-value model; i.e., the reward of any set is computable in O(1) time.
- In many settings, the oracle-value model does not hold.
- Additionally, the reward varies depending on the order in which subsets are considered.
- Consequently, we consider efficient methods for computing information rewards over multiple orderings.

**Contributions**

- We derive an approach that leads to substantial complexity reduction for information planning in Gaussian models under the assumption that a measurement depends only on a few latent variables.
- We demonstrate experimentally the advantages of this approach as the number of measurements and sparsity increases.
- We derive a related variant of belief propagation which is better suited for information propagation in social networks.
- We show empirically that the benefits of the latter method are more significant as the distance between consecutive elements of the visitation order decreases.

**Related work**

- [7] consider the problem of selecting an optimal k-element subset of measurements from a ground set V of size W that maximizes some reward f.
  - They provide a (1 − 1/e)-worst-case bound.
- [7] provides a 50% lower bound to matroidal structures.
- [7] considers information planning in the batch setting, where all measurements are available.
  - They show that entropy and mutual information (under mild conditions) are submodular.
- All previous works assume the oracle value model; a universal "oracle" provides the function value for any input set in constant time.
- Evaluation of rewards usually presents a significant computational challenge.
- [7] note that the evaluation of entropies can be prohibitive.
- [3] acknowledge the complexity of evaluating the influence function.

**Gaussian HMMs**

- We restrict ourselves to HMMs (results easily extended to trees and polytrees).
- \( x_t = A_t x_{t-1} + v_{t-1} \)
- \( y_t = C_t x_t + w_t \)
- where \( V \sim N(0, Q) \), \( W \sim N(0, R) \), and \( A_t, \ C_t \) are the dynamics and measurement matrices, respectively.

**Notation**

- \( X = \{X_1, \ldots, X_T\} \): T latent variables that are the focus of inference (\( X_t \in \mathbb{R}^D \)).
- \( \{Y_1, \ldots, Y_T\} \): T observation sets (\( Y_t \in \mathbb{R}^d \)).
- \( Y = \{Y_1, \ldots, Y_T\} \): Each \( Y_t \) is an \( N \times m \times 1 \) vector holding the m measurements of set \( Y_t \) (\( Y_{t} \in \mathbb{R}^{m \times 1} \)).

**Problem statement**

- Select up to \( k_t \) measurements from each set \( Y_t \) that maximize some reward of interest.
- In the Gaussian case, the reward function can be expressed as:
  \[
  R(Y_{t}, A_t) = \frac{1}{2} A_t Y_t^\top Y_t A_t^\top + \sum_{j=1}^{m} d_j \log |d_j| - \log |\det A_t|
  \]

**Greedy selection and theoretical guarantees**

For a specified visitation order \( \{w_1, w_2, \ldots, w_t\} \) such that \( w_t^{j} > w_t^{t-j} \), select \( g_t \) that maximizes
\[
  g_t = \arg \max_{g_t \in \mathbb{R}^d} f_{\{w_t, \ldots, w_t^{j}\}}
\]

**Update**

Update the covariance at the current point after the incorporation of measurement \( g_t \).
\[
  \Sigma_{t+1} = \Sigma_{t} + C_{t} E_{\{w_t, \ldots, w_t^{j}\}} C_{t}^\top
\]

**Complexity Reduction due to Sparsity**

For a chosen measurement \( u \), update the covariance as
\[
  \Sigma = \Sigma - E_{\{w_t, \ldots, w_t^{j}\}} \left( C_{t} Y_t^\top C_{t}^\top + R_{w-t} \right) E_{\{w_t, \ldots, w_t^{j}\}}
\]

**References**