Gaussian Processes are a Bayesian method for regression
- Prior on function values: joint Gaussian distribution
- Specified by mean and covariance
- For Gaussian noise & observations, predictions are straightforward (conditional Gaussian distributions)

Study learning curves: average (over data set and teacher) mean square error as function of number of examples $N$
- Relatively well understood in continuous spaces but what about discrete spaces?
- Use GPs to learn a function defined on the vertices of a graph
In previous work (NIPS 2009) we studied how learning curve approximations from continuous spaces performed in the discrete graph case, for:

- Random regular graphs (all vertices have same degree)
- Random walk kernel

While approximation worked o.k. for continuous spaces, in our discrete space:

- Learning curve approximation performs badly in the midsection of the learning curves
- Becomes worse as noise level decreases

This approximation does not account for the additional structure in random graphs: locally tree-like
In this paper we
- exploit graph structure to produce far more accurate results
- derive approximations for a wide range of graph ensembles
- obtain predictions that become exact in large graph limit

Write mean square error as
\[ \epsilon_g = -\frac{2}{V} \lim_{\lambda \to 0} \frac{\partial}{\partial \lambda} \langle \ln Z \rangle \{n_i\}, G \]

\[ Z = \int df \exp \left( -\frac{1}{2} f^T C^{-1} f - \frac{1}{2\sigma^2} \sum_{i=1}^{V} n_i f_i^2 - \frac{\lambda}{2} \sum_{i=1}^{V} f_i^2 \right) \]

Want to apply the cavity method but first need to:
- Eliminate the inverse of the covariance matrix \( C \)
- \( C \propto ((1 - a^{-1}) I + a^{-1}D^{-1/2}AD^{-1/2})^p \) correlates vertices \( p \) steps apart – reduce to nearest neighbour correlations by introducing \( 2p \) additional variables at each vertex

End up with a complex valued Gaussian graphical model
Even for graphs with $V = 500$ vertices graphs our new approximation (triangles) appears exact.

More on the poster: other types of random graph, comparisons with previous approximation, ...