**[Probabilistic Knowledge Graph Embeddings](https://openreview.net/forum?id=rJ4qXnCqFX)**

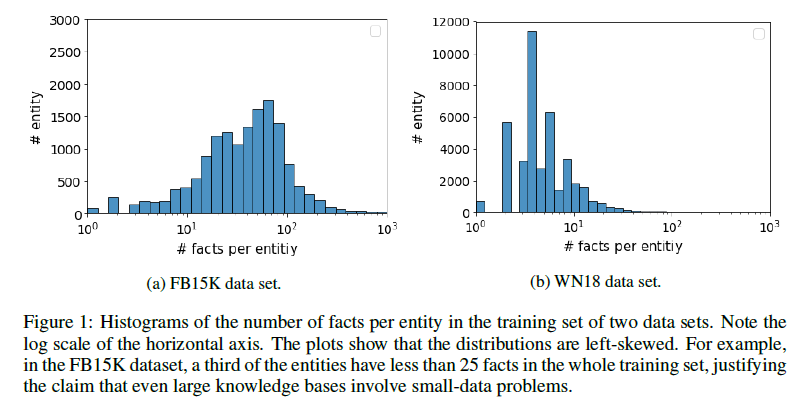
[Farnood Salehi](https://openreview.net/profile?email=farnood.salehi%40epfl.ch), [Robert Bamler](https://openreview.net/profile?email=robert.bamler%40gmail.com), [Stephan Mandt](https://openreview.net/profile?email=stephan.mandt%40gmail.com)

‎28‎ ‎Sep‎ ‎2018 (modified: ‎21‎ ‎Dec‎ ‎2018) ICLR 2019 Conference Blind Submission Readers: Everyone 7 Replies

[Hide details](https://openreview.net/group?id=ICLR.cc/2019/Conference#rJ4qXnCqFX-details-871)

* **Abstract:** We develop a probabilistic extension of state-of-the-art embedding models for link prediction in relational knowledge graphs. Knowledge graphs are collections of relational facts, where each fact states that a certain relation holds between two entities, such as people, places, or objects. We argue that knowledge graphs should be treated within a Bayesian framework because even large knowledge graphs typically contain only few facts per entity, leading effectively to a small data problem where parameter uncertainty matters. We introduce a probabilistic reinterpretation of the DistMult (Yang et al., 2015) and ComplEx (Trouillon et al., 2016) models and employ variational inference to estimate a lower bound on the marginal likelihood of the data. We find that the main benefit of the Bayesian approach is that it allows for efficient, gradient based optimization over hyperparameters, which would lead to divergences in a non-Bayesian treatment. Models with such learned hyperparameters improve over the state-of-the-art by a significant margin, as we demonstrate on several benchmarks.
* **Keywords:** knowledge graph, variational inference, probabilistic models, representation learning
* **TL;DR:** Scalable hyperparameter learning for knowledge graph embedding models using variational EM

We develop a probabilistic extension of state-of-the-art embedding models for link prediction in relational knowledge graphs. Knowledge graphs are collections of relational facts, where each fact states that a certain relation holds between two entities, such as people, places, or objects. We argue that knowledge graphs should be treated within a Bayesian framework because even large knowledge graphs typically contain only few facts per entity, leading effectively to a small data problem where parameter uncertainty matters.



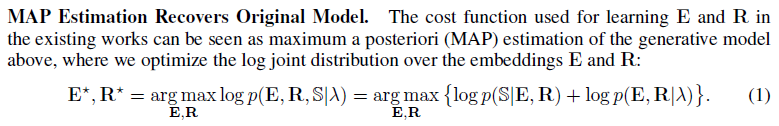
We introduce

1）a probabilistic reinterpretation of the DistMult (Yang et al., 2015) and

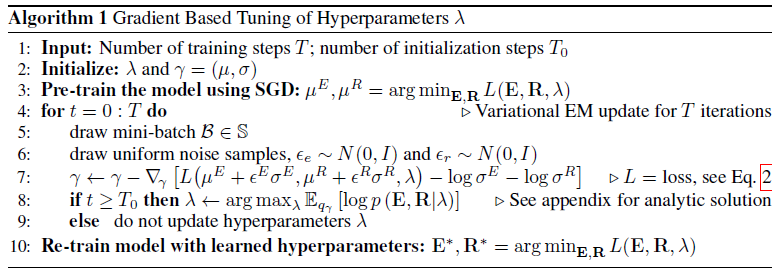
ComplEx (Trouillon et al., 2016) models and

2）We apply stochastic variational inference to scalably estimate an approximate posterior for each entity and relation embedding in the knowledge graph.

This approach not only allows us to estimate uncertainty, but more importantly allows for gradient-based hyperparameter optimization by stochastic gradient descent on the optimized variational bound (variational EM).



GRADIENT-BASED HYPERPARAMETER TUNING VIA VARIATIONAL EM



Implementing our methodology in existing models requires only a few lines of additional code, and it does not increase the computational complexity compared to a non-Bayesian training approach.

We find that the main benefit of the Bayesian approach is that it allows for

efficient, gradient based optimization over hyperparameters, which would lead to divergences in a non-Bayesian treatment.

Models with such learned hyperparameters improve over the state-of-the-art by a significant margin, as we demonstrate on several benchmarks.