Evaluating Smoothing Methods for Eigenvalue-Weighted Word Similarity

Aritra Saha[1] and Rishika Goswami[1]

Heritage Institute of Technology, Kolkata, West Bengal 700107, India
aritra.s484@gmail.com, goswami.rishika67@gmail.com

1 Appendix

1.1 MEN Dataset

MEN Fig. 1 presents the most *volatile* behaviour among our five datasets, with the overall range straddling the zero line. All three smoothing techniques begin with *negative* correlations at $\alpha = 0$, indicating that eigenvalue weighting is crucial for extracting any useful signal from MEN. As α increases, **Bayesian smoothing** shows the steepest ascent: it overtakes the other methods by $\alpha = 0.6$, crosses into positive territory around $\alpha = 0.8$, and ultimately tops out at $\rho \approx 0.11$ when the principal components are fully emphasised ($\alpha = 1.0$). **Dirichlet smoothing** follows a gentler upward trajectory, climbing steadily yet remaining just below zero even at $\alpha = 1.0$. **Jelinek–Mercer** lags throughout; although it mirrors Dirichlet's shape, its gains are smaller and it never clears the negative band, ending around $\rho \approx -0.05$.

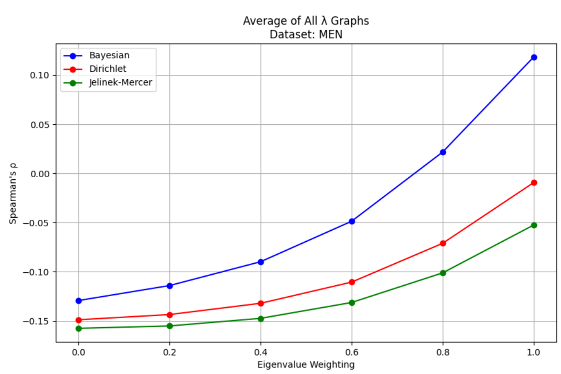

Fig. 1: MEN

The pronounced slope for Bayesian suggests that MEN benefits disproportionately from the heavier tail-probability correction that Bayesian smoothing provides once the leading eigenvectors are amplified. In contrast, Dirichlet's fixed prior and Jelinek–Mercer's linear interpolation appear less able to recover the fine-grained semantic judgements in MEN, even when combined with aggressive eigenvalue weighting. As with the other datasets, variation over λ (not shown) tweaks the exact rankings but does not alter the overall hierarchy: Bayesian > Dirichlet > Jelinek–Mercer for any $\alpha \geq 0.6$.

1.2 Graphs of each dataset for each tested λ value

Due to space constraints in the main text, we provide the detailed plots for each of the five datasets here. Each dataset has six graphs: the first five correspond to the Jelinek–Mercer parameter λ ranging from 0.1 to 0.9 (in increments of 0.2), and the sixth plot shows the average performance across these five λ values. The x-axis represents the eigenvalue weighting factor α, and the y-axis is the Spearman rank correlation. These plots allow a more granular look at how different smoothing methods (Bayesian, Dirichlet, and Jelinek–Mercer) interact with both α and λ.

Below, each figure is presented in landscape mode for enhanced legibility, with the dataset name explicitly mentioned in the caption. Refer to the main text for additional discussion and overall comparisons among the smoothing methods.

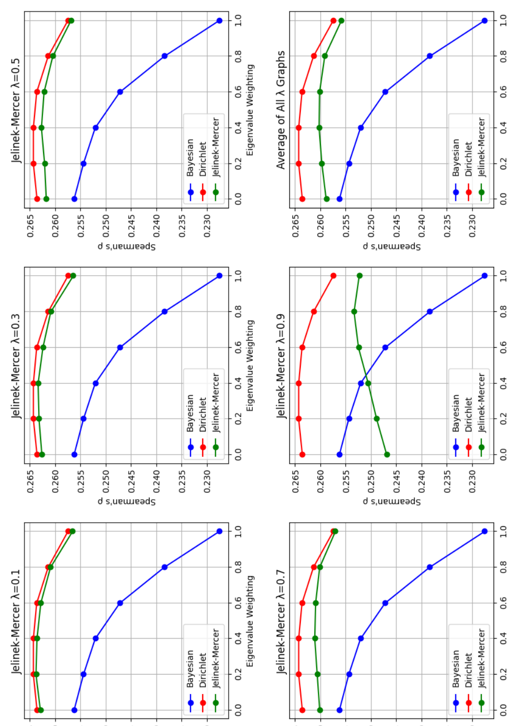

Fig. 2: SimLex-999: Six graphs for $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ plus the average.

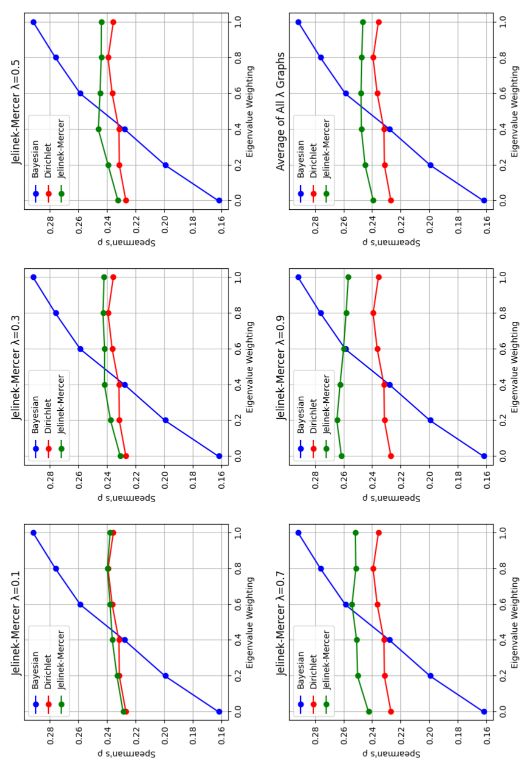

Fig. 3: WordSim-353: Six graphs for $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ plus the average.

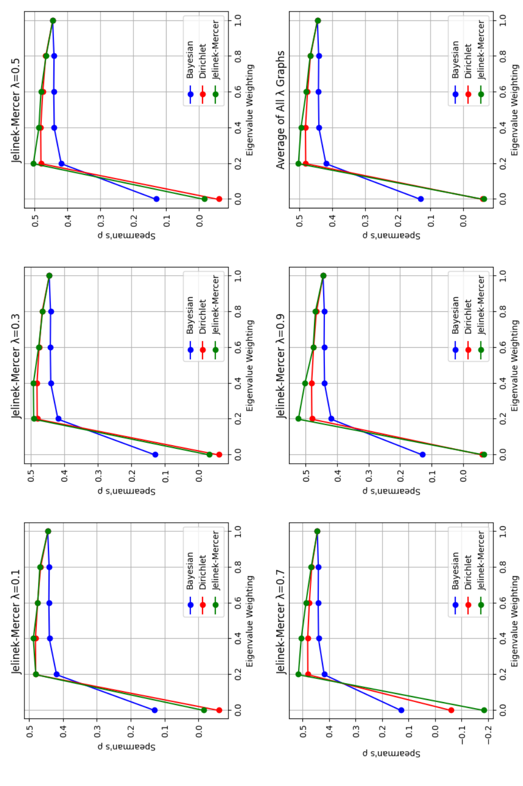

Fig. 4: RG-65: Six graphs for $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ plus the average.

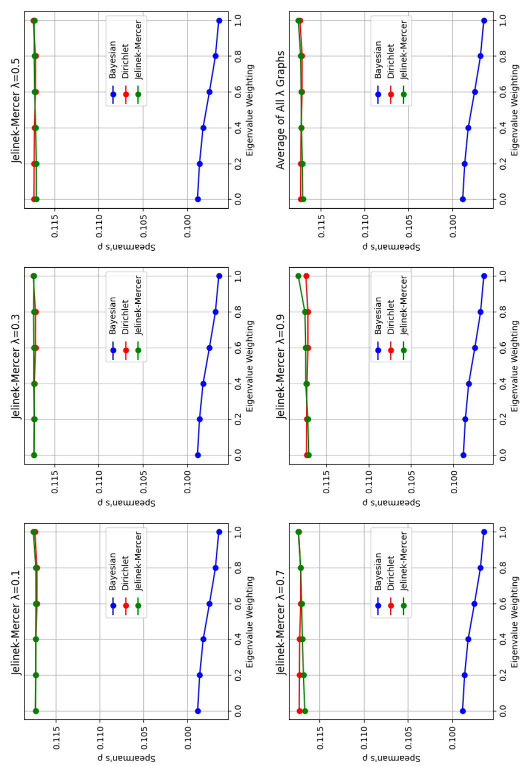

Fig. 5: RW (Rare Words): Six graphs for $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ plus the average.

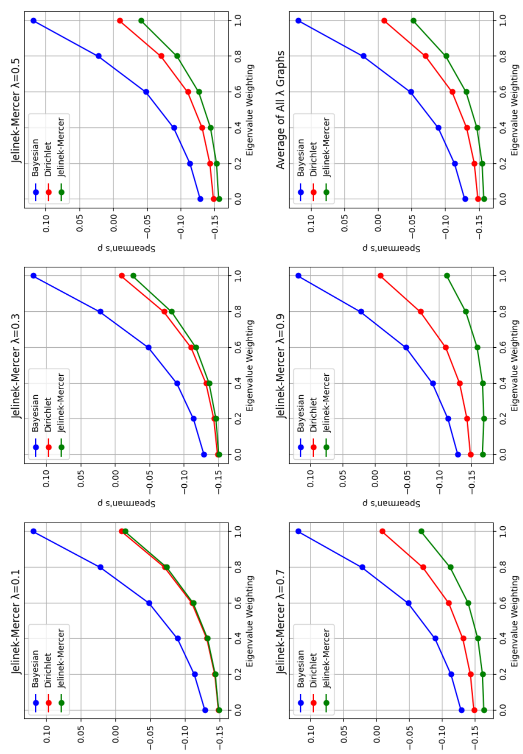

Fig. 6: MEN: Six graphs for $\lambda \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ plus the average.