# RecipEntropy Figures

November 22, 2013

## 1 Figures by Dataset

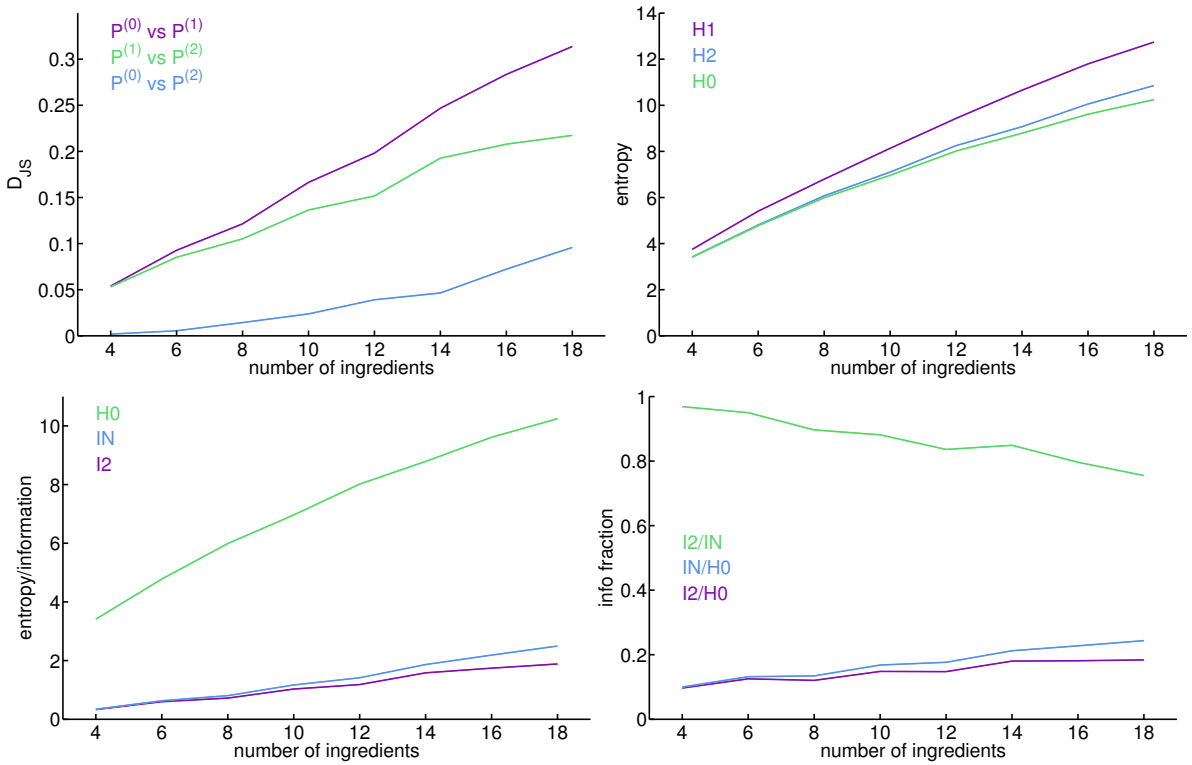


Figure 1: **Standard.** Includes all ingredients and recipes. These figures illustrate that, by the time we include just 18 ingredients, a pairwise maxent model already starts to fail, with $D_{\mathrm{JS}}\left[P^{(0)} \mid P^{(2)}\right] \to 0.1$ and the combination of $I_2/I_N \to 0.8$ and $I_N/H_0 \to .25$. All trends indicates that a pairwise model will only become worse as more ingredients are included. This suggests that higher order correlations play an important role in shaping recipes.

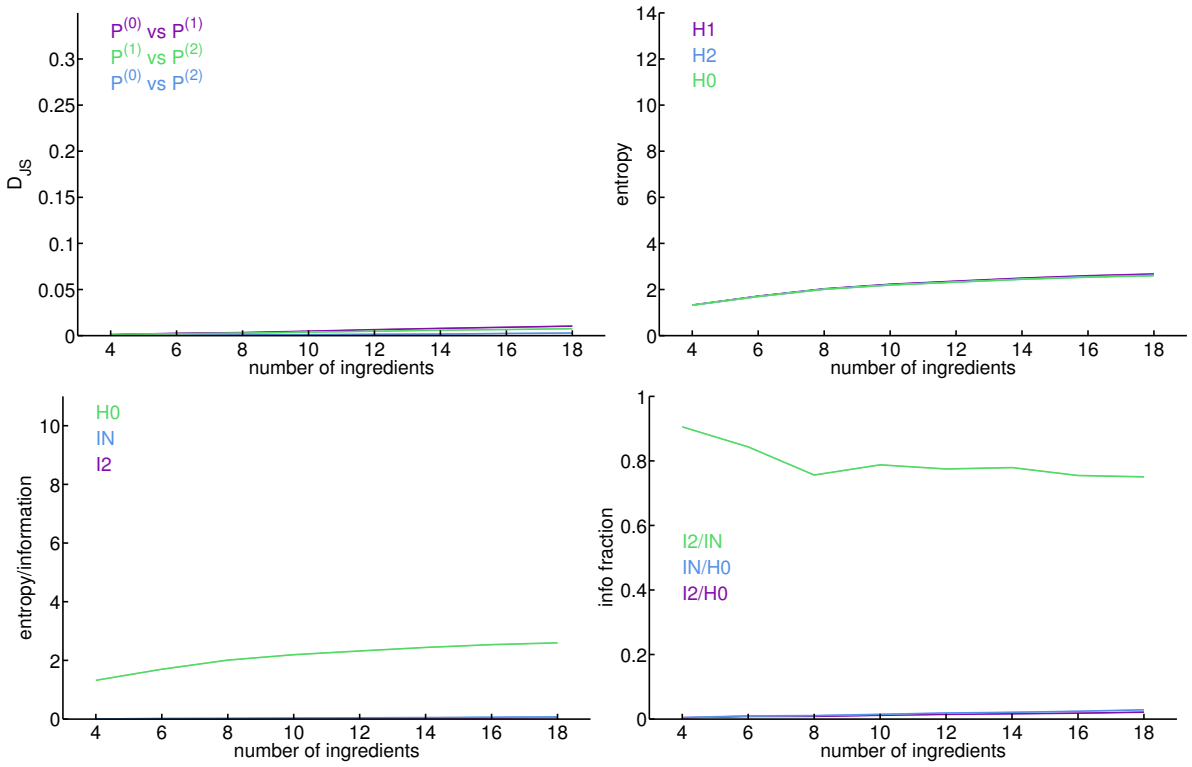


Figure 2: **Meats.** Includes all recipes but only meats for ingredients. In stark contrast to the full dataset (standard), the subset of meats exhibits minimal structure, with a first-order model performing essentially perfectly ($D_{\mathrm{JS}}\left[P^{(0)} \mid P^{(1)}\right] \approx 0$, $H_0 \approx H_1$, and $I_2 \approx I_N \approx 0$). Thus, not only are higher-order correlations unimportant, even *second-order* correlations appear to play a minimal role.

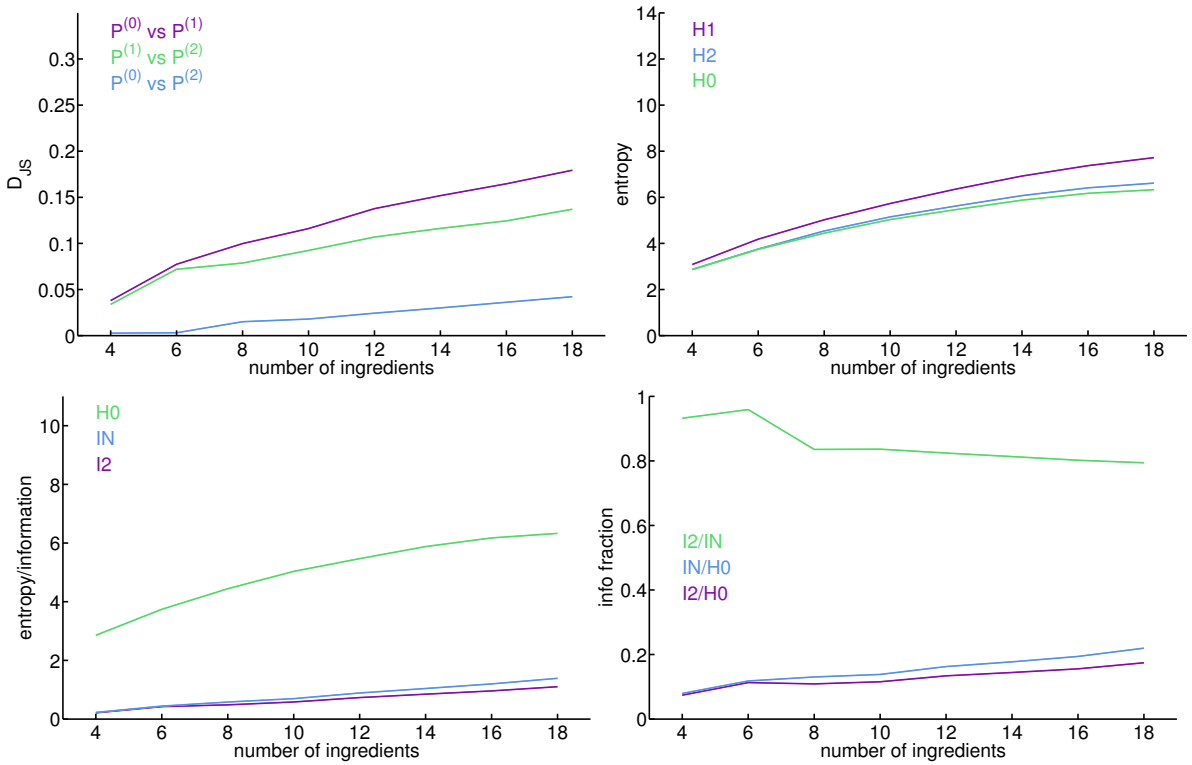


Figure 3: **Spices.** Includes all recipes but only spices for ingredients. These figures suggest higher-order structure, though perhaps not as much as the full dataset. While the information fractions $I_2/I_N$ and $I_N/H_0$ approach similar levels as when we used all the ingredients above, $D_{\mathrm{JS}}\left[P^{(0)} \mid P^{(2)}\right]$ is only half that of the full dataset. However, it still approaches a not small number ($\approx 0.05$) and looks as though it will continue to rise with more ingredients.

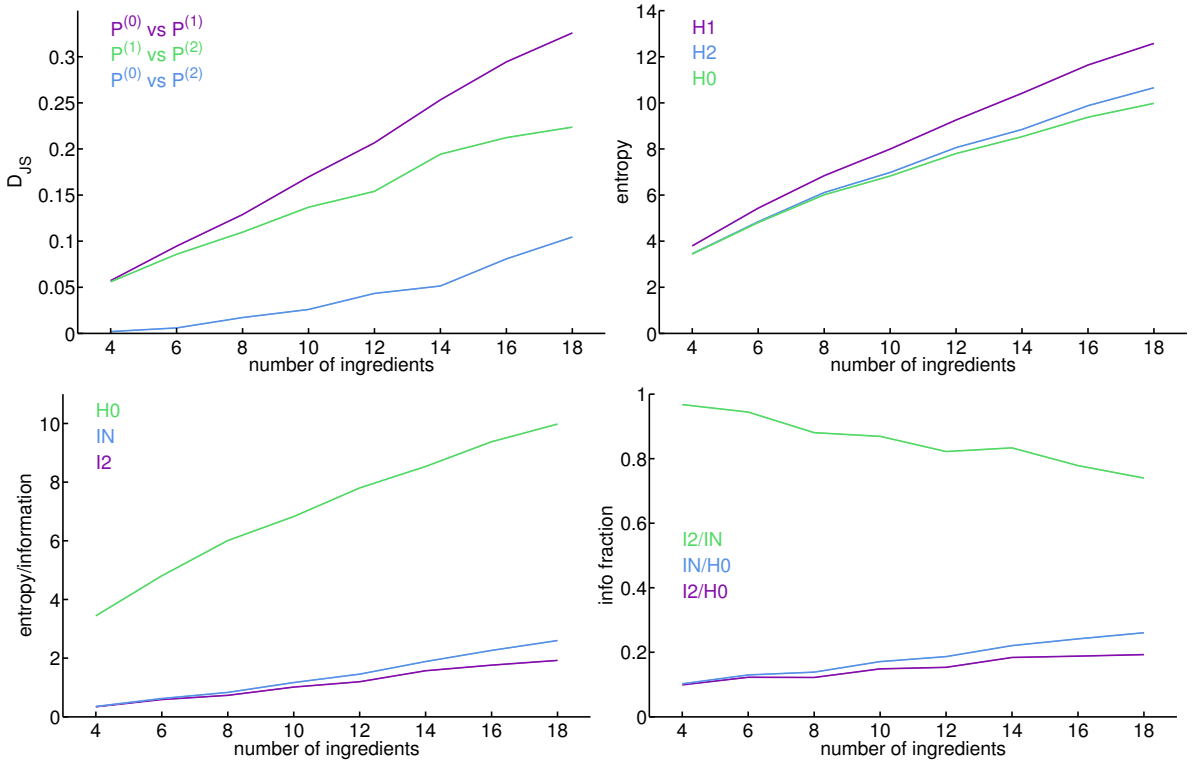


Figure 4: **North American.** Includes all ingredients but only North American recipes.

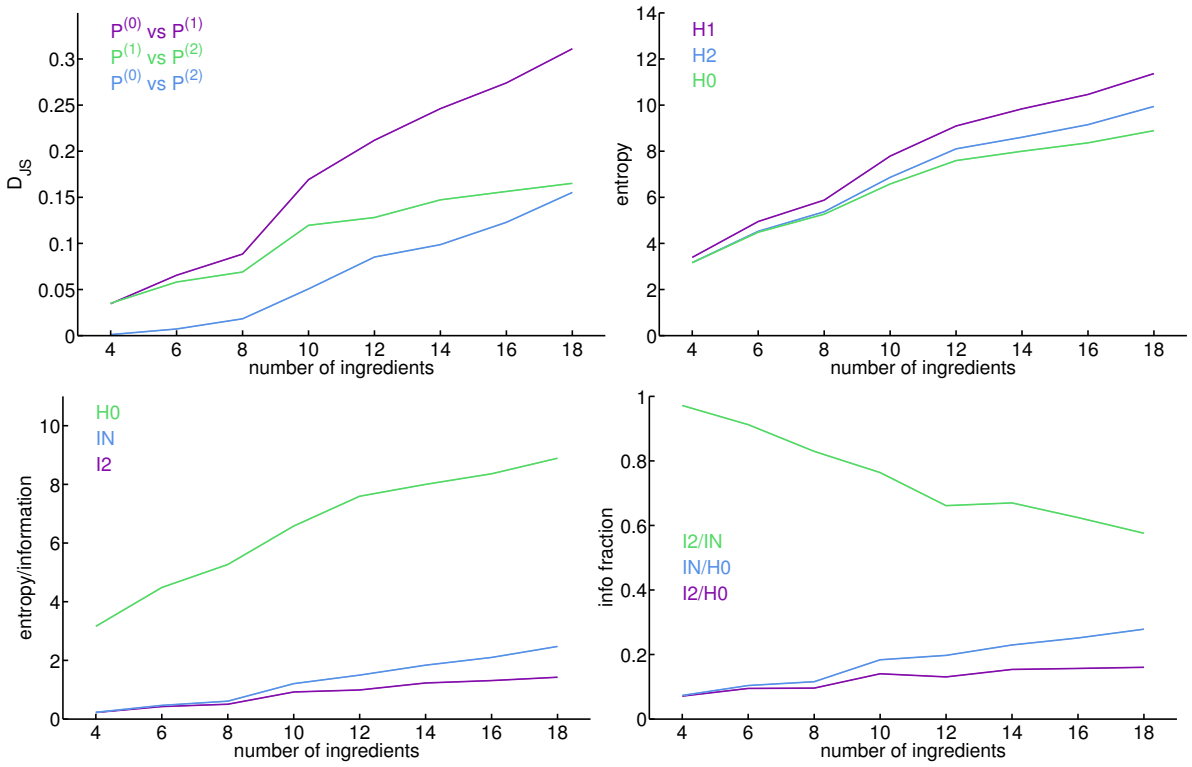


Figure 5: **Southern European.** Includes all ingredients but only Southern European recipes.

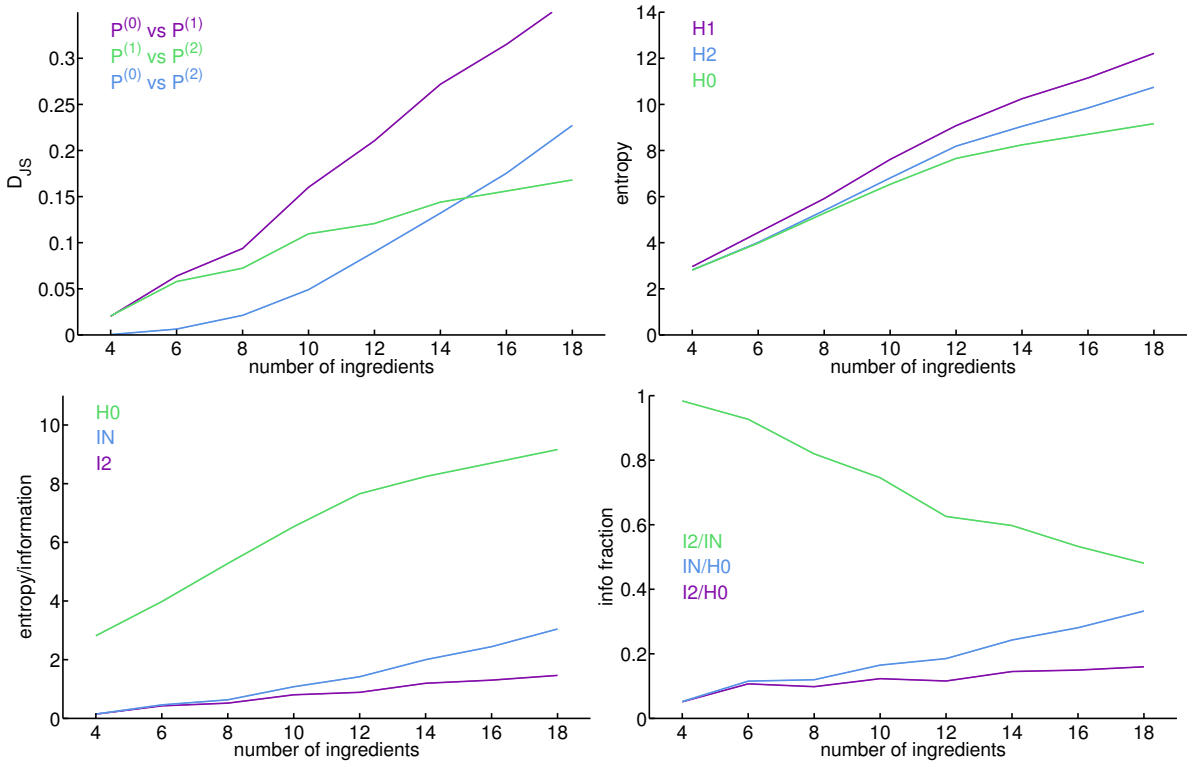


Figure 6: **Latin American.** Includes all ingredients but only Latin American recipes.

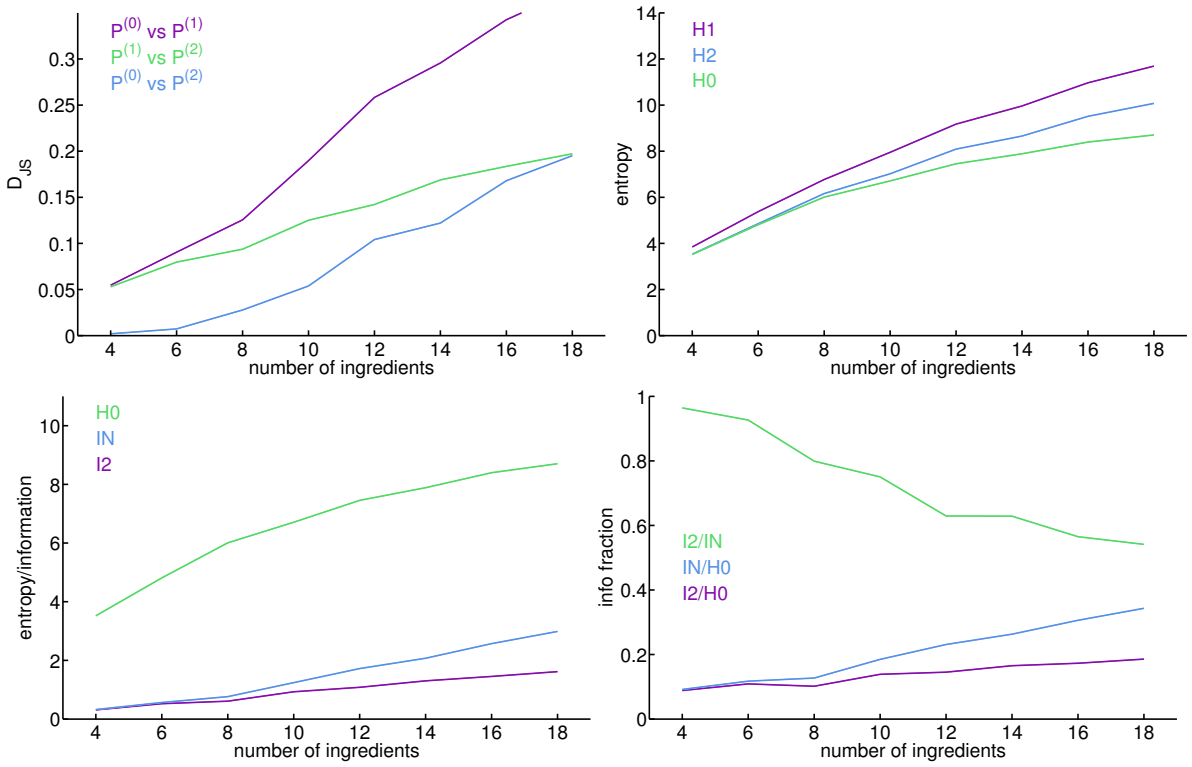


Figure 7: **Western European.** Includes all ingredients but only Western European recipes.

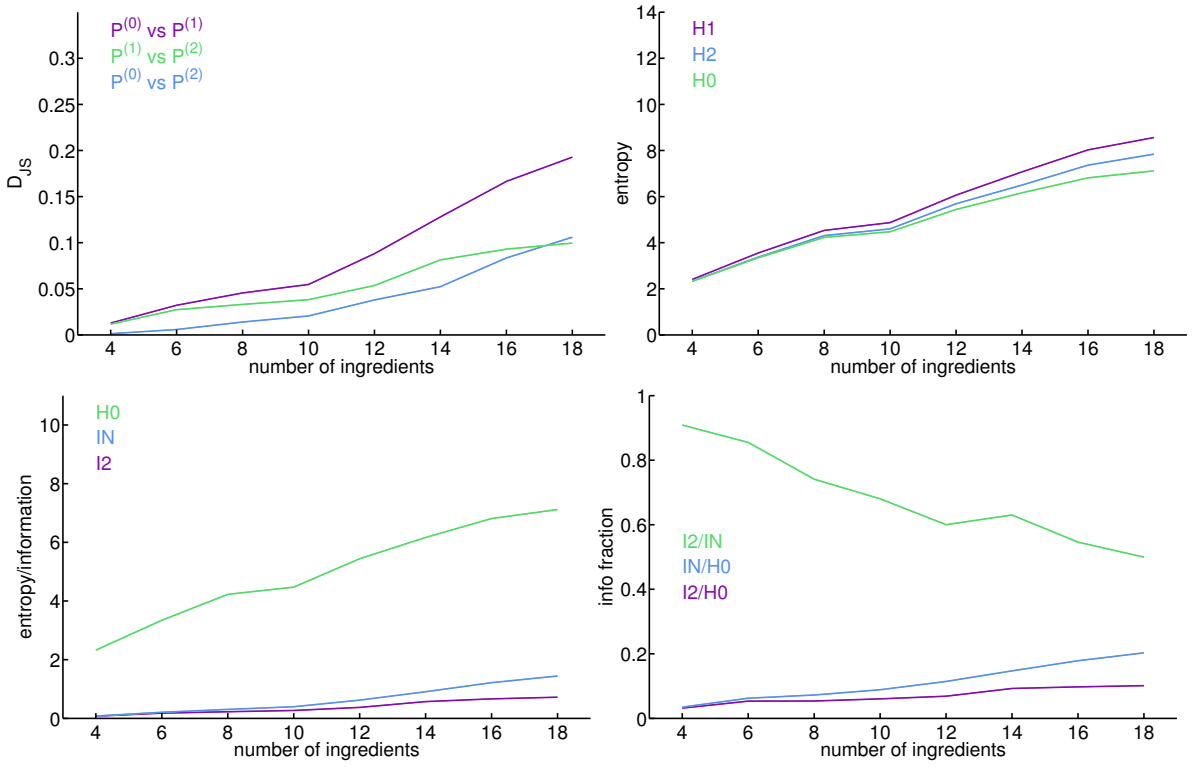


Figure 8: **East Asian.** Includes all ingredients but only East Asian recipes.

## 2 Figures by Type

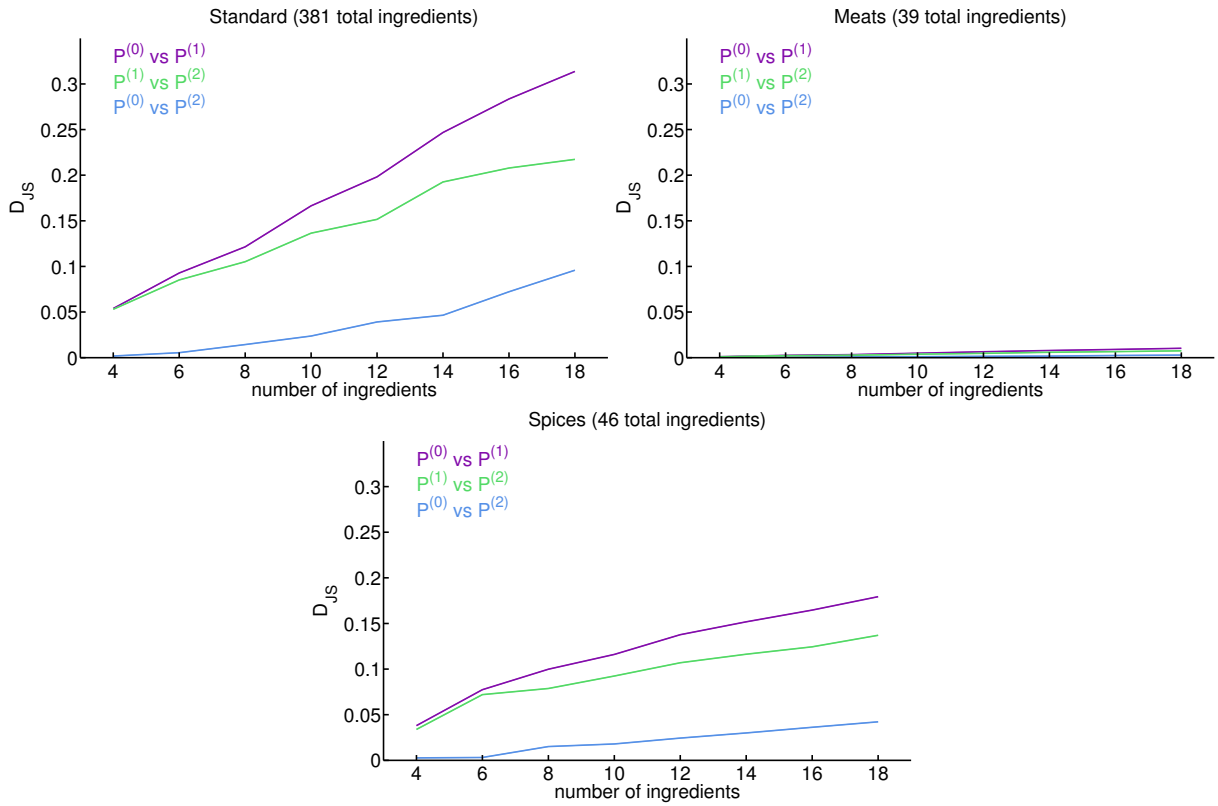


Figure 9: **Jensen-Shannon divergences for different ingredient subsets.** In general, the maxent models perform better on subsets of the ingredients. For the meats, a first-order model performs essentially perfectly. For the spices, first- and second-order models perform roughly twice as well as they do on the full ingredient set. However, their declining performance with the number of ingredients included suggests the importance of higher-order correlations there as well.

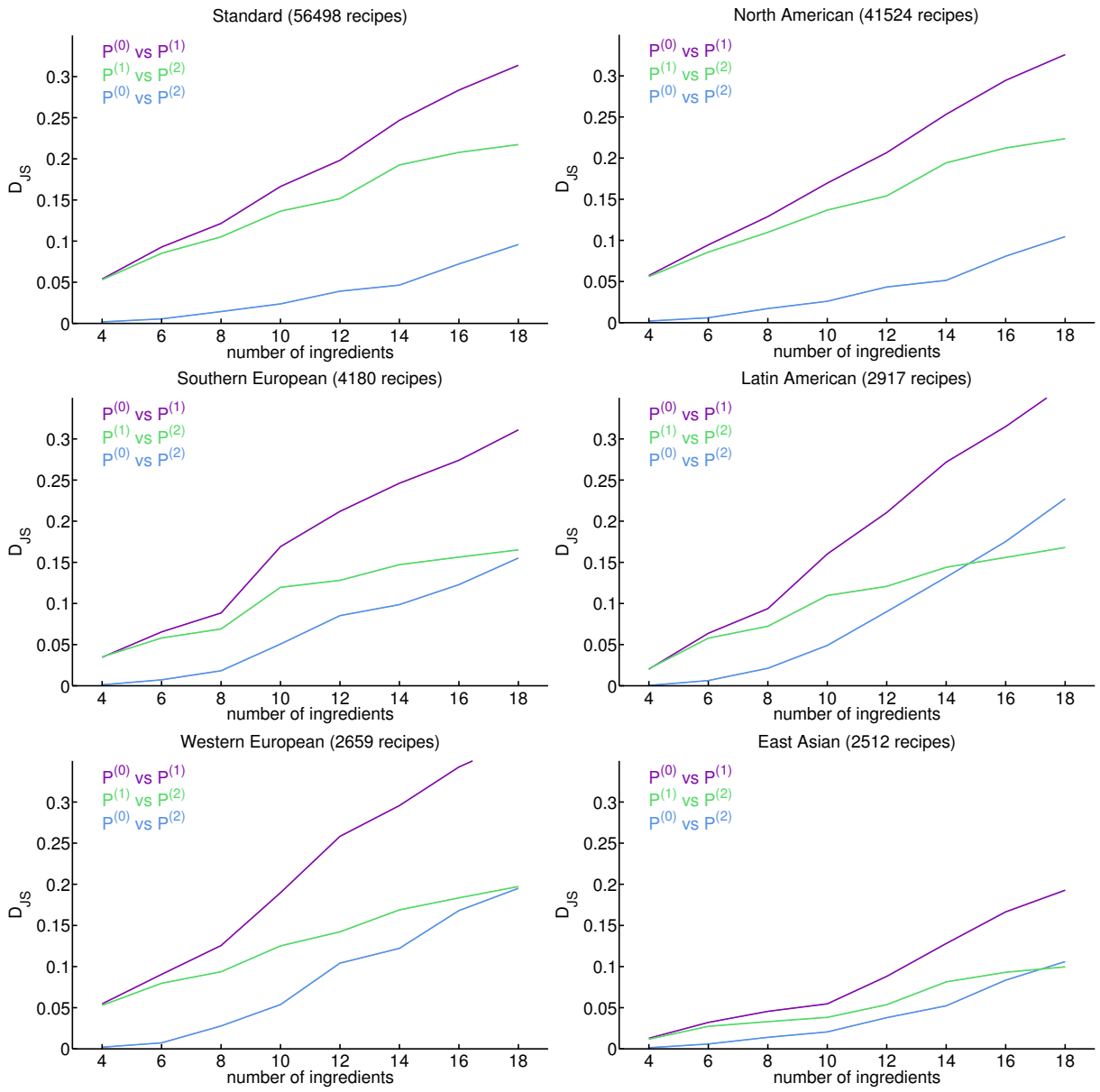


Figure 10: **Jensen-Shannon divergences for different recipe subsets.** As North American recipes dominate our dataset, we see that the plots for North American are essentially the same as when one includes all of the recipes. In both cases, higher-order correlations play an important role with $D_{\rm JS}\left[P^{(0)} \mid P^{(2)}\right] \to 0.1$ when 18 ingredients are included. For Southern European recipes, $D_{\rm JS}\left[P^{(0)} \mid P^{(2)}\right] \to 0.15$, for Latin America, $D_{\rm JS}\left[P^{(0)} \mid P^{(2)}\right] \to 0.225$, for Western Europe, $D_{\rm JS}\left[P^{(0)} \mid P^{(2)}\right] \to 0.18$, and for East Asia, $D_{\rm JS}\left[P^{(0)} \mid P^{(2)}\right] \to 0.1$. This suggests a "complexity ordering" of cuisines (from most to least) with Latin America at the top, followed by Western Europe, then Southern Europe, and with North America and East Asia rounding out the bottom. However, this only holds for the top 18 ingredients of each cuisine, and the results could change as more ingredients are included.

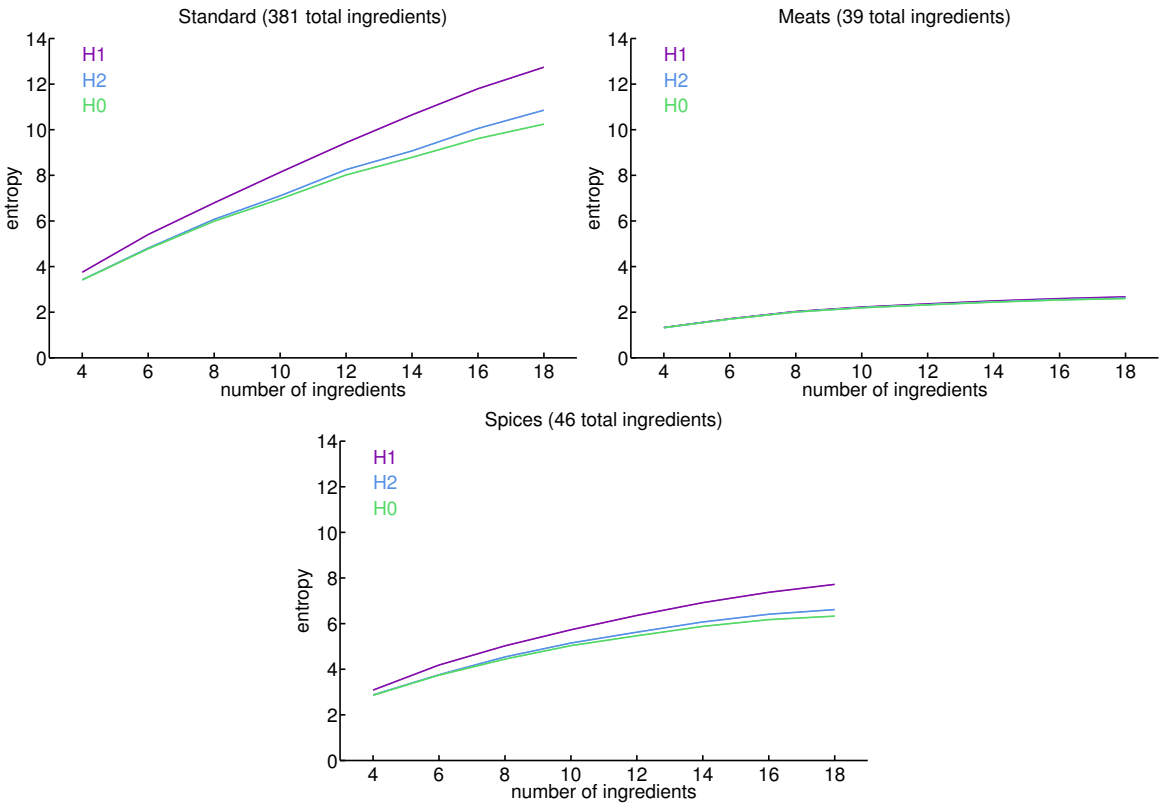


Figure 11: **Entropies for different ingredient subsets.**

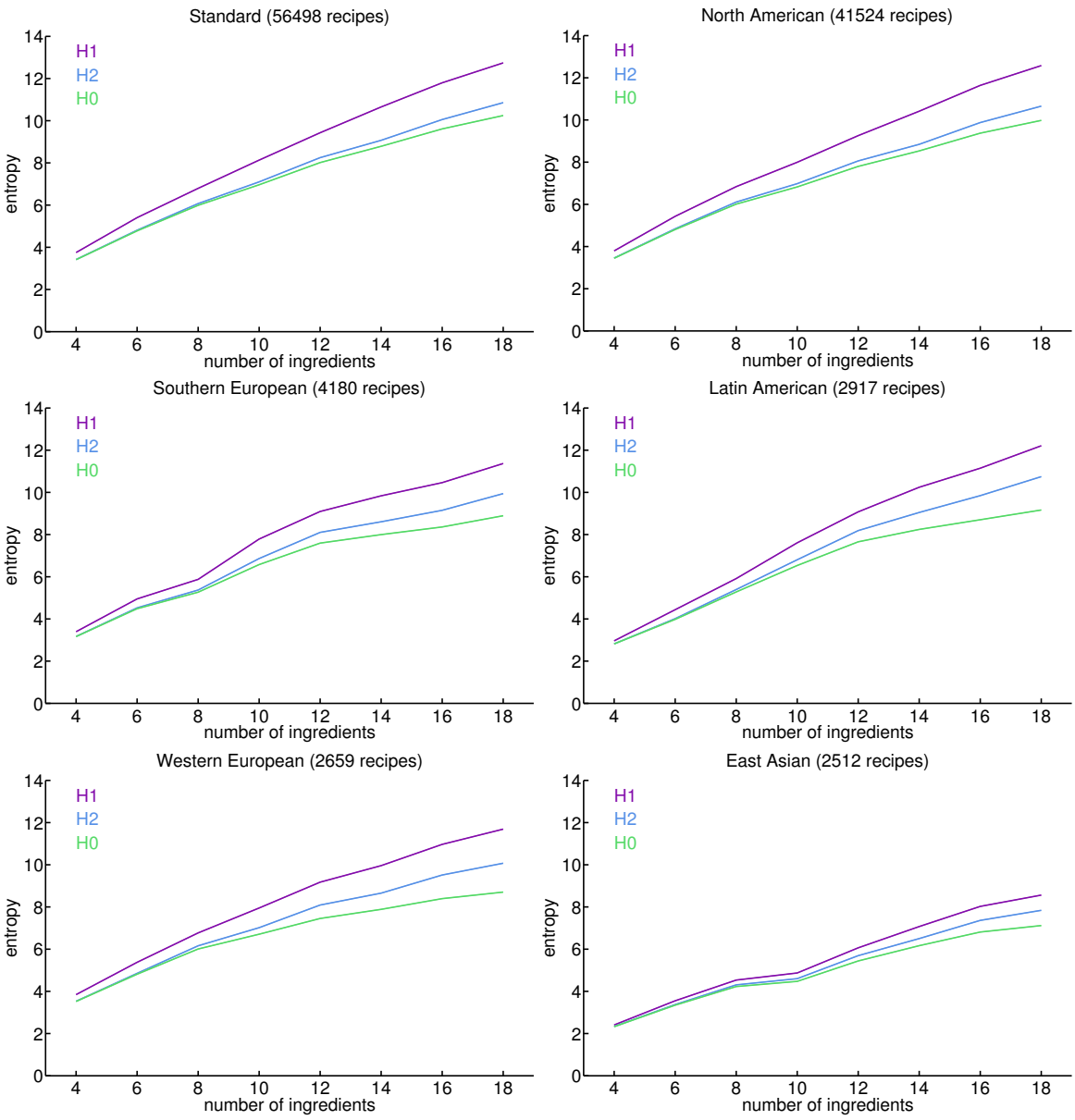


Figure 12: **Entropies for different recipe subsets.**

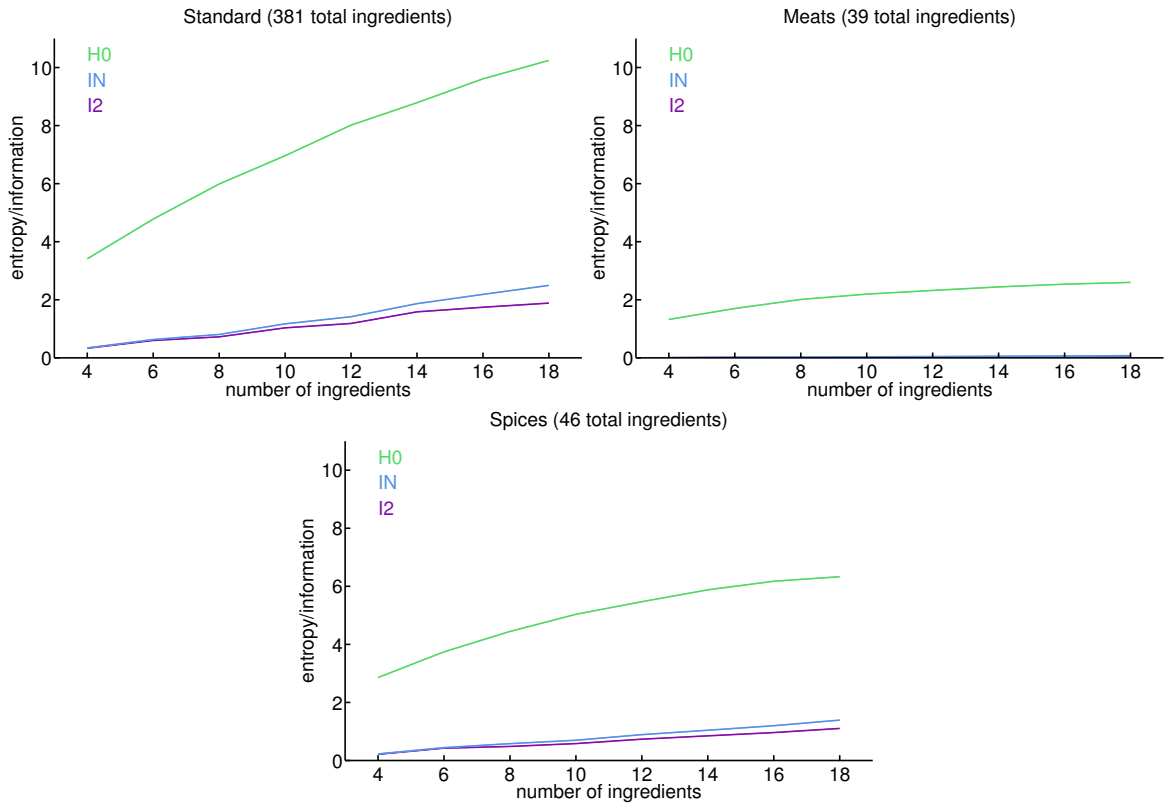


Figure 13: **Informations for different ingredient subsets.**

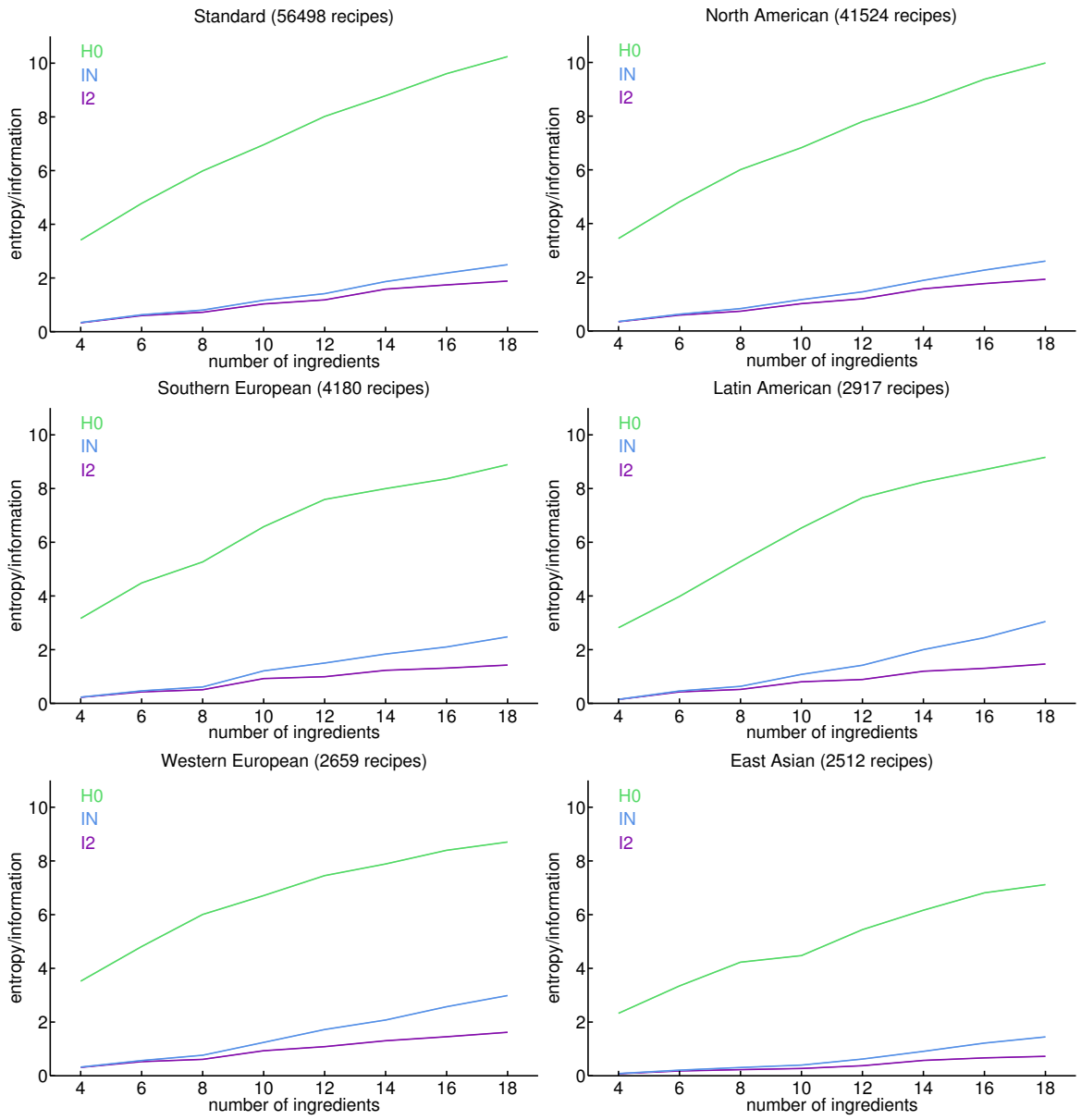


Figure 14: **Informations for different recipe subsets.**

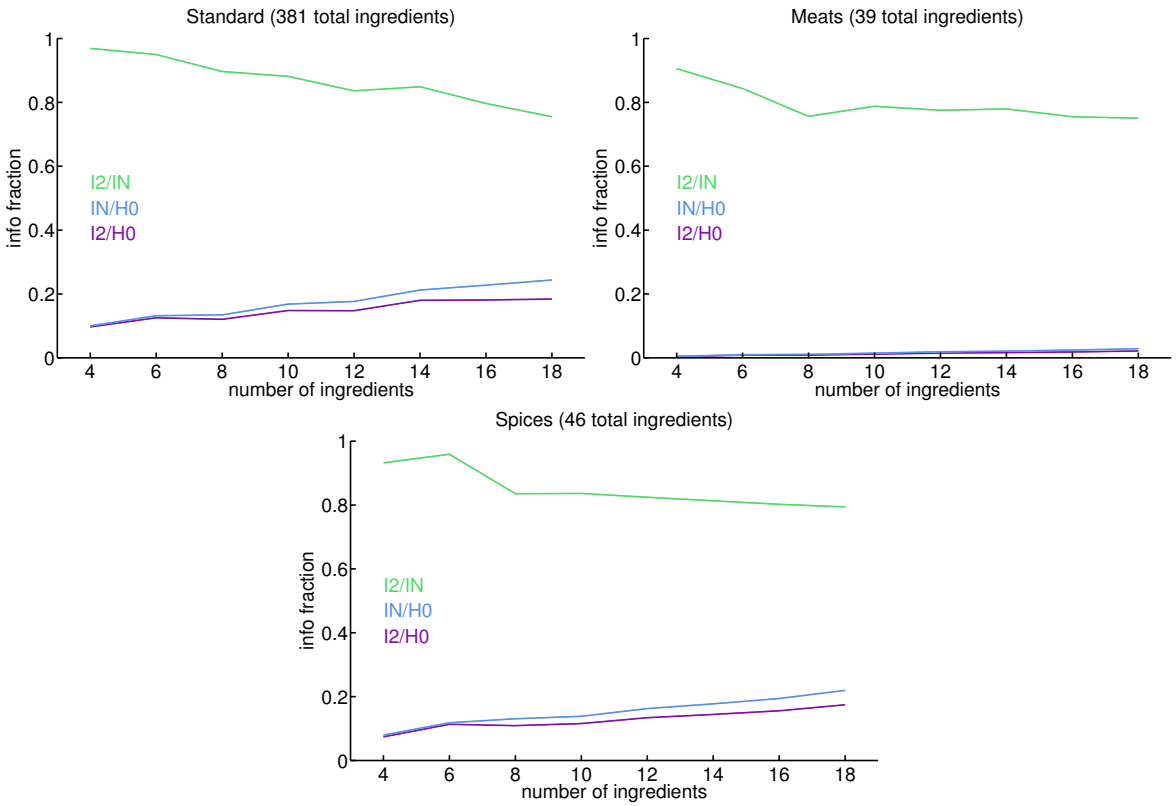


Figure 15: **Information fractions for different ingredient subsets.**

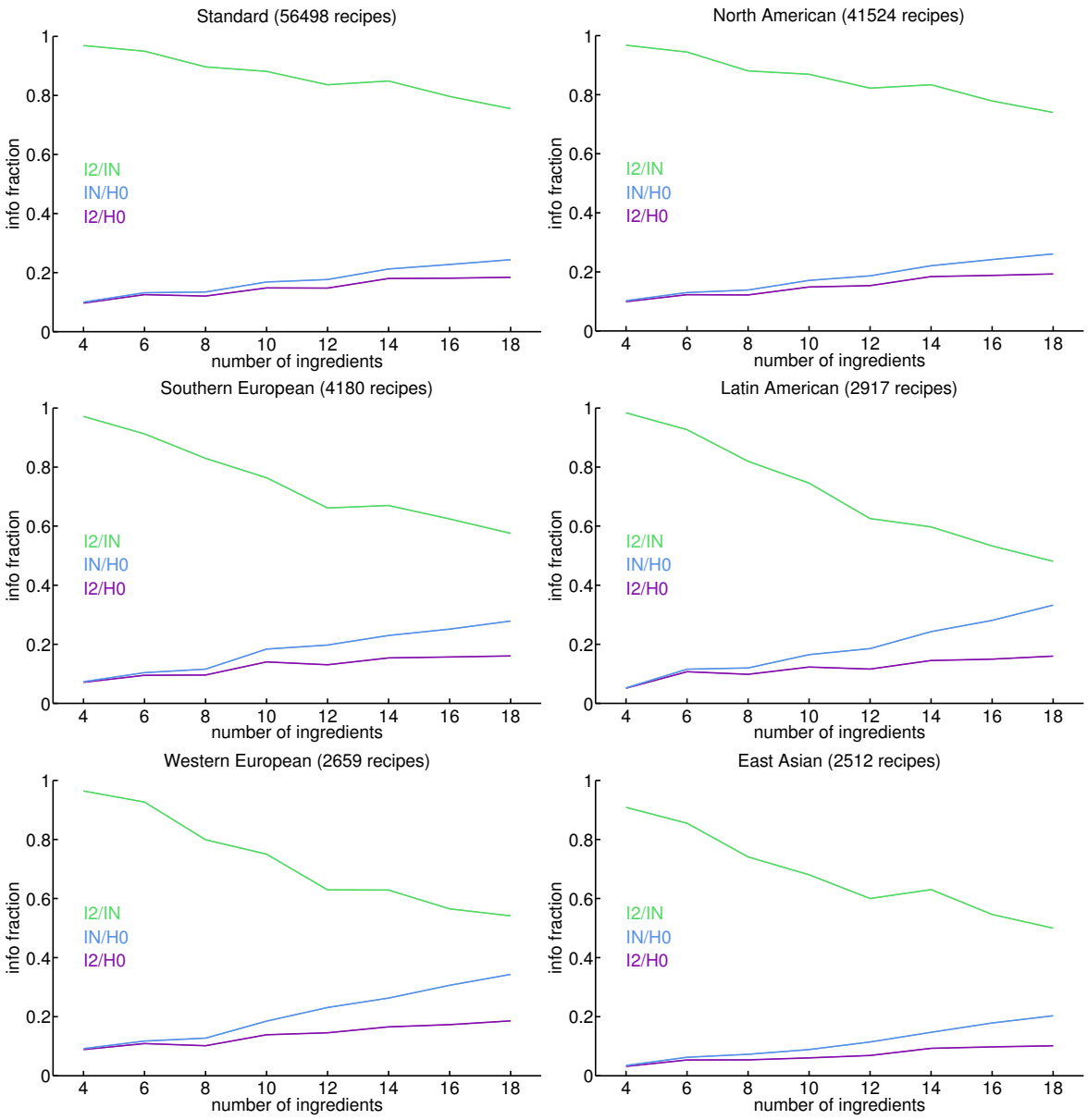


Figure 16: **Information fractions for different recipe subsets.**

# 3 Figures by Statistic

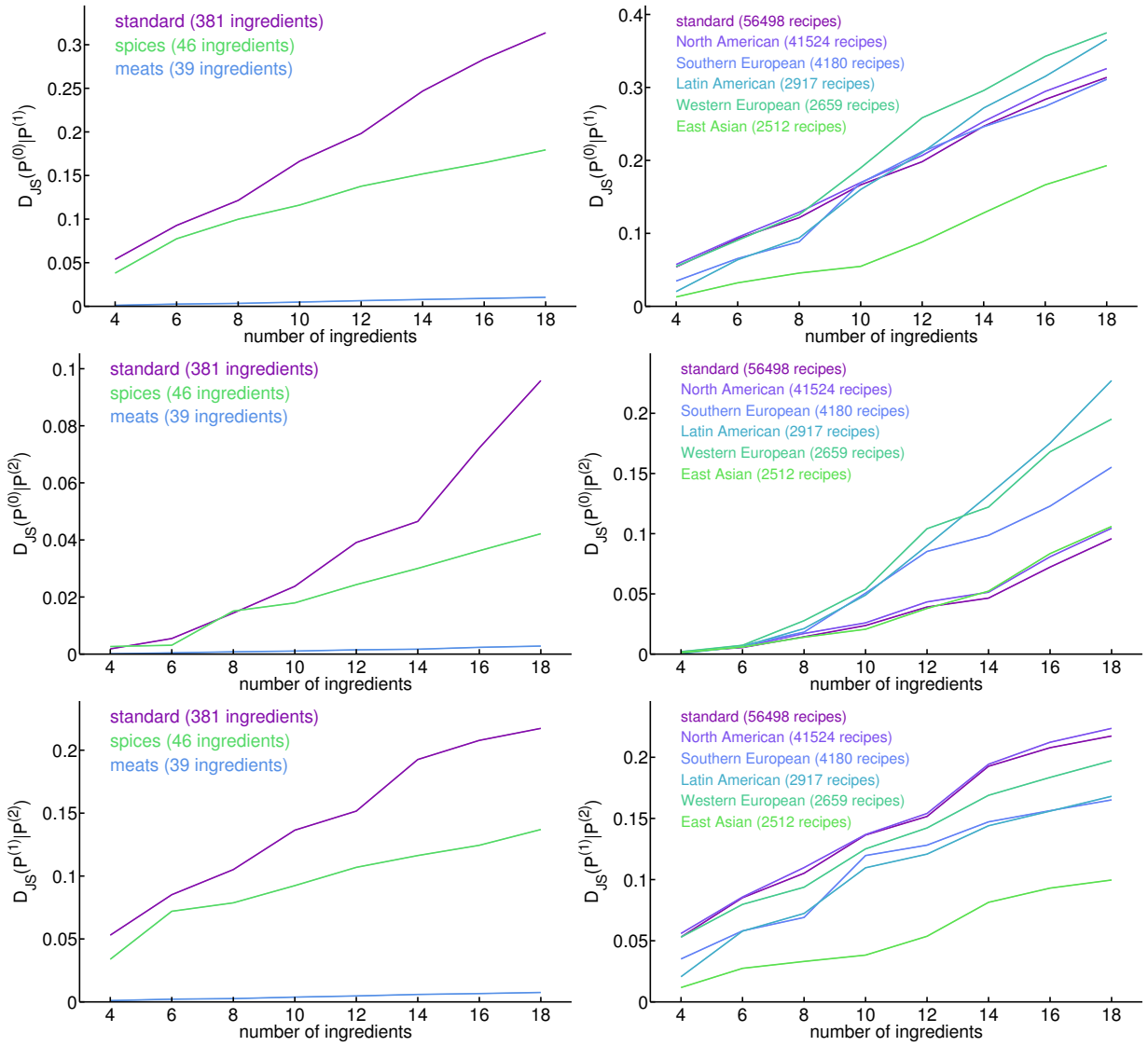


Figure 17: **JS divergences.** Left column compares divergences across ingredient subsets; right column compares divergences across recipe subsets (cultures). The right column shows us how much diversity there is across cultures. Some (e.g. Latin American) display strong higher-order structure, whereas others (e.g. East Asian) display very little. A "complexity ordering" across cultures is easy to read off from these plots.