Note that

$$\frac{\beta_j}{\sum_{\hat{\kappa}_{il}=1}\beta_l} = \frac{s_j}{\sum_{\hat{\kappa}_{il}=1}s_l} \text{ for all } (i,j) \in \{1,\ldots,m-1\} \times \{1,\ldots,n\}. \tag{17}$$

(16) implies

$$\hat{r}_i = \frac{\hat{\alpha}_i}{\sum_{\hat{\kappa}_{il}=1}s_l} \text{ for all } i = 1,\ldots,m-1. \tag{18}$$

Consequently,

$$\hat{\beta}_j = s_j \sum_{i=1}^{m-1} \frac{\hat{\kappa}_{ij}\hat{\alpha}_i}{\sum_{\kappa_{il}=1}s_l} \text{ for all } j = 1,\ldots,n. \tag{19}$$

(19) gives us the same solution as (9), taking into account (17). Therefore, Wagner conditioning and PME agree.

## 6. Conclusion

Wagner-type problems (but not obverse Majerník-type problems) can be solved using JUP and Wagner's ad hoc method. Obverse Majerník-type problems, and therefore all Wagner-type problems, can also be solved using PME and its established and integrated formal method. What at first blush looks like serendipitous coincidence, namely that the two approaches deliver the same result, reveals that JUP is safely incorporated in PME. Not to gain information where such information gain is unwarranted and to process all the available and relevant information is the intuition at the foundation of PME. My results show that this more fundamental intuition generalizes the more specific intuition that ratios of probabilities should remain constant unless they are affected by observation or evidence. Wagner's argument that PME conflicts with JUP is ineffective because it rests on assumptions that proponents of PME naturally reject.

## A. Appendix: PME generalizes Jeffrey Conditioning

A proof that PME generalizes standard conditioning is in [35]. A proof that PME generalizes Jeffrey conditioning is in [27]. I will give my own simple proofs here that are more in keeping with the notation in the paper. An interested reader can also apply these proofs to show that PME generalizes Wagner conditioning, but not without simplifications that compromise mathematical rigour. The more rigorous proof for the generalization of Wagner conditioning is in the body of the paper.

I assume finite (and therefore discrete) probability distributions. For countable and continuous probability distributions, the reasoning is largely analogous (for an introduction to continuous entropy see [12] (p.16ff); for an example of how to do a proof of this section for continuous probability densities see [27,34]; for a proof that the stationary points of the Lagrange function are indeed the desired extrema see [36] (p.55) and [3] (p.410); for the pioneer of the method applied in this section see [34] (p.241ff)).

### *A.1. Standard Conditioning*

Let $y_i$ (all $y_i \neq 0$) be a finite type II prior probability distribution summing to 1, $i \in I$. Let $\hat{y}_i$ be the posterior probability distribution derived from standard conditioning with $\hat{y}_i = 0$ for all $i \in I'$ and