This summarizes work done in February 2014. Sudelbuch around page 900. The problem is that we have an $m \times n$ joint prior probability matrix p_{ij} with some zeroes in it (Wagner's constraints) that we want to transform into a joint posterior probability matrix $(q_{ij}$.

(A) For p_{ij}, we use the constraints and MAXENT. (B) For q_{ij}, we use the observation and *Infomin*.

(A) Let $K = K_\omega \times K_\theta \subset \{1, \ldots, m\} \times \{1, \ldots, n\}$ be the set for which $p_{ij} = 0$ (iff $(i, j) \in K$). First let $K = \emptyset$.

Then leonbloy on page 914 in Sudelbuch, using Cover and Thomas theorems 2.2.1. and 2.6.5., shows that $p_{ij} = P(\omega_i)P(\theta_i)$.

To show this manually let $\alpha_i = P(\omega_i)$ and $\beta_j = P(\theta_j)$. Then the maximum entropy for (p_{ij}) is

$$\begin{aligned} -H = \quad & \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} x_{ij}\log x_{ij} + \sum_{i=1}^{m-1}\left(\alpha_i - \sum_{j=1}^{n-1} x_{ij}\right) log \left(\alpha_i - \sum_{j=1}^{n-1} x_{ij}\right) + \\ & \sum_{j=1}^{n-1}\left(\beta_j - \sum_{i=1}^{m-1} x_{ij}\right) \log \sum_{j=1}^{n-1}\left(\beta_j - \sum_{i=1}^{m-1} x_{ij}\right) + X\log X \end{aligned} \tag{1}$$

with

$$X = \left(1 - \sum_{i=1}^{m-1}\alpha_i - \sum_{j=1}^{n-1}\beta_j + \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} x_{ij}\right) \tag{2}$$

To find the extremum we differentiate with respect to x_{ij} (i and j now fixed) and set to 0 (no Lagrange multipliers necessary, page 919).

$$\begin{aligned} 0 = \quad & \log x_{ij} - \log\left(\alpha_i - \sum_{j=1}^{n-1} x_{ij}\right) - \log\left(\beta_j - \sum_{i=1}^{m-1} x_{ij}\right) + \\ & \log\left(1 - \sum_{i=1}^{m-1}\alpha_i - \sum_{j=1}^{n-1}\beta_j + \sum_{i=1}^{m-1}\sum_{j=1}^{n-1} x_{ij}\right) \end{aligned} \tag{3}$$

It is easy to see that $x_{ij} = \alpha_i\beta_j$ solves this. Therefore, $p_{ij} = P(\omega_i)P(\theta_i)$ is the maximum entropy solution for $K = \emptyset$.