Imagine that we have N people, and for each of them we have their answers to p yes/no questions. E.g. with $p = 5$, one person's answers could be $\boldsymbol{x} = (1, 0, 1, 1, 1)$. There are then 2^p possible sets of answers (i.e. there are 2^p 'cells'). If, e.g. $p = 100$ and $N = 10^6$, then $N \ll 2^p$, and so the data points will be sprinkled very sparsely among the cells.

We could consider the 'histogram' with a separate bin for each cell, with the distribution being assumed multinomial, with likelihood:

$$P(\boldsymbol{m}|\boldsymbol{\gamma}) = \frac{N!}{\prod_{j=1}^{2^p} m_j!} \prod_{j=1}^{2^p} \gamma_j^{m_j}$$

where m_j is the number of points in cell j, (with $\sum m_j = N$), and γ_j is the probability of a point falling in cell j. But if we look at the ML estimate for $\boldsymbol{\gamma}$ it just puts all the probability in the few cells which have data points, and doesn't tell us about the probability of a data point falling into other 'nearby' cells. So instead define a histogram with much coarser binning, by chopping the 'box' containing all 2^p cells into 2 boxes, one with all the $x_1 = 0$ cells, and the other all the $x_1 = 1$ cells, and we can further chop these boxes along other dimensions, possibly chopping many times, but stopping with the number of boxes, K, being $\ll 2^p$. Let λ_i be the number of cells in the i^{th} box, (with $\sum_{i=1}^{K} \lambda_i = 2^p$). If we let the probability of the i^{th} box be β_i, and all of the cells within a box have equal probability, then $\gamma_j = \beta_i/\lambda_i$, when the j^{th} cell is in the i^{th} box. Then the probability of the data $\boldsymbol{m}$ given partition $\mathfrak{P}$ and box-probabilities $\boldsymbol{\beta}$ becomes:

$$P(\boldsymbol{m}|\mathfrak{P}, \boldsymbol{\beta}) = \frac{N!}{\prod_{j=1}^{2^p} m_j!} \prod_{i=1}^{K} (\beta_i/\lambda_i)^{n_i}$$

The posterior density is then obtained by multiplying by a prior density for the β_i, which let us take to be $Dir(1, 1, 1, ...1)$, i.e. a uniform density of $\Gamma(K)$ over it's support (the $(K-1)$ simplex).

$$P(\mathfrak{P}, \boldsymbol{\beta}|\boldsymbol{m}) = P(\boldsymbol{m}|P, \boldsymbol{\beta})P(\mathfrak{P}, \boldsymbol{\beta})/P(\boldsymbol{m}) = [\frac{1}{P(\boldsymbol{m})} \frac{N!}{\prod_{j=1}^{2^p} m_j!}] P(\mathfrak{P})\Gamma(K) \prod_{i=1}^{K} (\beta_i/\lambda_i)^{n_i}$$

The factor in square brackets doesn't depend on $\mathfrak{P}, \boldsymbol{\beta}$, so we can ignore it when running an M-H Markov chain; for now just call it $C(\boldsymbol{m})$. This leaves:

$$P(\mathfrak{P}, \boldsymbol{\beta}|\boldsymbol{m}) = C(\boldsymbol{m})P(\mathfrak{P}) \prod_{i=1}^{K} \lambda_i^{-n_i} \Gamma(K) \prod_{i=1}^{K} \beta_i^{n_i}$$

Next integrate out the β_i (integrating over the $(K-1)$ simplex) to get:

$$P(\mathfrak{P}|\boldsymbol{m}) = C(\boldsymbol{m})P(\mathfrak{P})\prod_{i=1}^{K}\lambda_i^{-n_i}\prod_{i=1}^{K}n_i!\frac{\Gamma(K)}{\Gamma(N+K)}$$

Now consider starting with a partition of size K, and splitting one of its boxes which contains λ cells, and n data points (dropping the subscripts for the moment), and let the numbers of data points in the two new boxes be k and $n-k$. We need the ratio of split to unsplit posterior probabilities. The product of factorials leads to a factor of $k!(n-k)!/n!$, in the ratio. The product involving λ's leads to a factor of 2^n in the ratio, because each of the new boxes formed in the split has $\lambda/2$ cells, so the factor of λ^{-n} gets replaced by $(\lambda/2)^{-k}(\lambda/2)^{-(n-k)} = 2^n\lambda^{-n}$. And $\Gamma(K)/\Gamma(N+K)$ leads to a factor of $K/(N+K)$ in the ratio. The ratio is then:

$$\frac{P(\mathfrak{P}_s|\boldsymbol{m})}{P(\mathfrak{P}_u|\boldsymbol{m})} = \frac{P(\mathfrak{P}_\mathfrak{s})}{P(\mathfrak{P}_\mathfrak{u})}\frac{2^n k!(n-k)!}{n!}\frac{K}{(N+K)}$$