S2 Appendix

Lower bound to the joint likelihood and the VB distributions.

If we assume that the prior distribution for \( \alpha \) and \( \beta \) are multivariate Gaussian distributions (denoted as \( \pi(\alpha, \beta) \)) with parameters \( \mu^0_\alpha, \Sigma^0_\alpha \) and \( \mu^0_\beta, \Sigma^0_\beta \) respectively then;

\[
p(y, z, \alpha, \beta) = \pi(\alpha, \beta) \prod_{i=1}^{n} o_i^{z_i} (1 - o_i)^{1 - z_i} \prod_{j=1}^{K_j} (z_i d_{i,j})^{y_{i,j}} (1 - z_i d_{i,j})^{1 - y_{i,j}}
\]

where we assume that we visit \( n \) sites \( K = (K_1, K_2, \ldots, K_n)^T \) times respectively. Let \( b(x) = \ln(1 + \exp(x)) \). \( \ln p = \ln p(y, z, \alpha, \beta) \) is equal to

\[
\ln p = \sum_{i:z_i=1} \sum_{j=1}^{K_j} (y_{i,j} w_{i,j} \alpha - b(w_{i,j} \alpha)) + z^T X \beta - 1_n^T b(X \beta) + \ln \pi(\alpha, \beta)
\]

\[
= y^T \text{diag}(\tilde{z}) W \alpha - \tilde{z}^T b(W \alpha) + z^T X \beta - 1_n^T b(X \beta) + \ln \pi(\alpha, \beta).
\] (1)

Equation (1) follows after some simplification and since when \( z_i = 0 \) we have \( y_i = 0 \). Here \( \tilde{z} = (z_1 1_{K_1}^T, \ldots, z_n 1_{K_n}^T)^T \).

A lower bound on the marginal log-likelihood can be obtained using

\[
\ln p(y) \geq \sum_{z} \int q(z, \alpha, \beta) \ln \left( \frac{p(y, z, \alpha, \beta)}{q(z, \alpha, \beta)} \right) d\alpha d\beta,
\] (2)

where \( q(z, \alpha, \beta) = q(z)q(\alpha)q(\beta) \). VB methods approximate the joint distribution of \( \pi(z, \alpha, \beta|y) \) by maximising the lower bound of the marginal log-likelihood and providing update equations for \( q(z) \), \( q(\alpha) \) and \( q(\beta) \) in turn. Note that the dependence on \( y \) has been excluded from the notation for the variational distributions in order to save space.
Note that the variational posterior distribution of $\alpha$ follows since

$$\ln q(\alpha) \propto \mathbb{E}_\beta (\ln p)$$

$$\propto \mathbb{E}_\beta (y^T \text{diag}(\hat{z})W\alpha - \hat{z}^T b(W\alpha) + z^T X\beta - 1_n^T b(X\beta) + \ln \pi(\alpha, \beta))$$

$$\propto y^T \text{diag}(\hat{z})W\alpha - \hat{z}^T b(W\alpha) + \ln \pi(\alpha).$$

Similarly

$$\ln q(\beta) \propto \mathbb{E}_\alpha (\ln p)$$

$$\propto \mathbb{E}_\alpha (y^T \text{diag}(\hat{z})W\alpha - \hat{z}^T b(W\alpha) + z^T X\beta - 1_n^T b(X\beta) + \ln \pi(\alpha, \beta))$$

$$\propto z^T X\beta - 1_n^T b(X\beta) + \ln \pi(\beta).$$