Supplementary material for "Latent space models for network perception data"

Daniel K. Sewell

1 Gibbs sampler

1.1 Full conditional distributions

We first place the following priors on the unknowns.

$$\alpha|\sigma^2 \sim N(\mu_\alpha, \nu_\alpha \sigma^2), \qquad \beta_1 \sim \ell N(\boldsymbol{\mu}_\beta, \nu_\beta),$$
$$\sigma^2 \sim IG(\gamma_\sigma/2, \eta_\sigma/2), \qquad \omega \sim IG(\gamma_\omega/2, \eta_\omega/2),$$
$$\Sigma \sim IW(\gamma_\Sigma, \Gamma_\Sigma), \qquad \mathbf{Z}_i \stackrel{iid}{\sim} N(\mathbf{0}, \sigma_z^2 I_p),$$
$$\theta \sim \ell N(\mu_\theta, \nu_\theta), \qquad \sigma_z^2 \sim IG(\gamma_z/2, \eta_z/2),$$

where $IG(a, b)$ is the inverse gamma distribution with shape a and scale b, $IW(a, B)$ is the inverse Wishart distribution with degrees of freedom a and scale matrix B, and $\ell N(a, b)$ is the log normal distribution with log-mean a and log-variance b.

We have closed form solutions for most of the full conditional distributions. The full conditional distributions for the $A^*_{k,ij}$'s are trivially given by the truncated normal distribution with mean and variance as given in the main text. The full conditional for each α_k is given by

$$\begin{aligned} &\alpha_k|\cdot \sim N(\mu_{\alpha_k}, \sigma^2_{\alpha_k}), \\ &\frac{1}{\sigma^2_{\alpha_k}} = \frac{1}{\sigma^2} + \sum_{i \neq j} \tau_{k,ij}, \\ &\mu_{\alpha_k} = \frac{\alpha/\sigma^2 + \sum_{i\neq j} \tau_{k,ij}(A^*_{k,ij} - s_i - r_j + \|\mathbf{Z}_i - \mathbf{Z}_j\|)}{1/\sigma^2 + \sum_{i\neq j}\tau_{k,ij}}. \end{aligned} \tag{1}$$

For the sender and receiver effects we have

$$\begin{aligned} &\begin{pmatrix}\boldsymbol{s}' & \boldsymbol{r}'\end{pmatrix}'|\cdot \sim N(\boldsymbol{\mu}_{sr}, \Sigma_{sr}), \\ &\Sigma_{sr}^{-1} = \sum_{k\in S} \begin{pmatrix} \mathrm{Diag}(T_k\mathbb{1}) & T_k \\ T_k & \mathrm{Diag}(T_k\mathbb{1}) \end{pmatrix} + \Sigma^{-1} \otimes I_n, \\ &\boldsymbol{\mu}_{sr} = \Sigma_{sr} \sum_{k\in S} \begin{pmatrix} \mathbf{diag}(\widetilde{A}_k T_k) \\ \mathbf{diag}(\widetilde{A}_k' T_k) \end{pmatrix}, \\ &[\widetilde{A}_k]_{ij} = A^*_{k,ij} - \alpha_k + \|\mathbf{Z}_i - \mathbf{Z}_j\|, \\ &[T_k]_{ij} = \tau_{k,ij} 1_{\{j\neq i\}}, \end{aligned} \tag{2}$$

where $\mathrm{Diag}(\mathbf{a})$ is the square diagonal matrix with the vector $\mathbf{a}$ on the diagonal, and $\mathbf{diag}(A)$

is the vector equaling the diagonal entries of a square matrix A. For Σ we have

$$\begin{aligned}
&\Sigma|\cdot \sim IW(\widetilde{\gamma}_\Sigma, \widetilde{\Gamma}_\Sigma), \\
&\widetilde{\gamma}_\Sigma = \gamma_\Sigma + n, \\
&\widetilde{\Gamma}_\Sigma = \Gamma_\Sigma + \sum_{i=1}^n \begin{pmatrix} s_i \\ r_i \end{pmatrix} \begin{pmatrix} s_i & r_i \end{pmatrix}.
\end{aligned} \tag{3}$$

The full conditionals for α and σ^2 are given as

$$\begin{aligned}
\alpha|\sigma^2, \cdot &\sim N(\widetilde{\mu}_\alpha, \widetilde{\nu}_\alpha \sigma^2), \\
\sigma^2|\cdot &\sim IG(\widetilde{\gamma}_\sigma/2, \widetilde{\eta}_\sigma/2), \\
\widetilde{\mu}_\alpha &= \frac{\mu_\alpha + \nu_\alpha \sum_{k\in S} \alpha_k}{1+\nu_\alpha n_1}, \\
\widetilde{\nu}_\alpha &= \frac{\nu_\alpha}{1+\nu_\alpha n_1}, \\
\widetilde{\gamma}_\sigma &= \gamma_\sigma + n_1, \\
\widetilde{\eta}_\sigma &= \eta_\sigma + \sum_{k\in S} \alpha_k^2 + \frac{\mu_\alpha^2}{\nu_\alpha} - \frac{\widetilde{\mu}_\alpha^2}{\widetilde{\nu}_\alpha}.
\end{aligned} \tag{4}$$

For β_{0k} the full conditional is given by

$$\begin{aligned}
&\beta_{0k} \sim Ga(\widetilde{\theta}_{1k}/2, \widetilde{\theta}_{2k}/2), \\
&\widetilde{\theta}_{1k} = \theta + n(n-1), \\
&\widetilde{\theta}_{2k} = \theta + \sum_{i\neq j} e^{-\beta_1 \left(\|\mathbf{Z}_k - \mathbf{Z}_i\| + \|\mathbf{Z}_k - \mathbf{Z}_j\|\right)} \cdot (A^*_{k,ij} - \alpha_k - s_i - r_j + \|\mathbf{Z}_i - \mathbf{Z}_j\|)^2.
\end{aligned} \tag{5}$$

For σ_z^2 we have

$$\begin{aligned}
&\sigma_z^2 \sim IG(\widetilde{\gamma}_z/2, \widetilde{\eta}_z/2), \\
&\widetilde{\gamma}_z = \gamma_z + np, \\
&\widetilde{\eta}_z = \eta_z + \sum_{i=1}^n \|\mathbf{Z}_i\|^2.
\end{aligned} \tag{6}$$

There is no closed form well known distribution for Z, β_1, or for θ. These require Metropolis-Hastings steps to obtain posterior draws. In our analyses we used a simple multivariate normal distribution with a spherical covariance matrix for Z, and a normal distribution for $\log \beta_1$ and $\log \theta$. Tuning was done in an automated manner up over the duration of a prespecified burn-in period to obtain an acceptance rate approximately 0.234, at which point the tuning parameters were fixed for the remainder of the MCMC sampling.

To perform estimation when covariates are incorporated into the model, we need to first assign priors to $\boldsymbol{B}_s$, $\boldsymbol{B}_r$, and B_z, derive their full conditional distributions, and alter the full conditional distributions above that are affected by the incorporation of covariates into the model. Let us assume that the prior distributions follow the form

$$\begin{pmatrix} \boldsymbol{B}_s \\ \boldsymbol{B}_r \end{pmatrix} \sim N(\mathbf{0}, \sigma^2_{B_{sr}} I_{2q}), \tag{7}$$

$$B_z \sim N_{q\times p}(0, I_p \otimes (\sigma^2_{B_z} I_q)). \tag{8}$$

Then the new full conditionals are given by

$$\begin{pmatrix} \boldsymbol{B}_s \\ \boldsymbol{B}_r \end{pmatrix} \sim N(\boldsymbol{\mu}_{B_{sr}}, \Sigma_{B_{sr}}) \tag{9}$$

$$\Sigma^{-1}_{B_{sr}} = \Sigma^{-1} \otimes (X'X) + \frac{1}{\sigma^2_{B_{sr}}} I_{2q}$$

$$\boldsymbol{\mu}_{B_{sr}} = \Sigma_{B_{sr}} (\Sigma^{-1} \otimes X)' \begin{pmatrix} \boldsymbol{s} \\ \boldsymbol{r} \end{pmatrix},$$

and

$$B_z \sim N_{q\times p}(M_z, I_p \otimes U_z) \tag{10}$$

$$U_z^{-1} = \frac{1}{\sigma_z^2} X'X + \frac{1}{\sigma^2_{B_z}} I_q$$

$$M_z = \frac{1}{\sigma_z^2} U_z X'Z.$$

In equation (2), we must change $\boldsymbol{\mu}_{sr}$ such that

$$\boldsymbol{\mu}_{sr} = \Sigma_{sr} \left[\sum_{k\in S} \begin{pmatrix} \mathbf{diag}(\widetilde{A}_k T_k) \\ \mathbf{diag}(\widetilde{A}'_k T_k) \end{pmatrix} + (\Sigma^{-1} \otimes X) \begin{pmatrix} \boldsymbol{B}_s \\ \boldsymbol{B}_r \end{pmatrix} \right], \tag{11}$$

and in equation (6) change $\widetilde{\eta}_z$ such that

$$\widetilde{\eta}_z = \eta_z + \sum_{i=1}^n \|\mathbf{Z}_i - \mathbf{X}_i B_z\|^2, \tag{12}$$

where $\mathbf{X}_i$ is the i^{th} row of X. In equation (3) we must replace $\boldsymbol{s}_i$ and $\boldsymbol{r}_i$ with $\tilde{\boldsymbol{s}}_i := \boldsymbol{s}_i - X_i \boldsymbol{B}_s$ and $\tilde{\boldsymbol{r}}_i := \boldsymbol{r}_i - X_i \boldsymbol{B}_r$

1.2 Initialization

To initialize the MCMC algorithm, I first ran a generalized linear model using all the data to estimate the α_k, $\boldsymbol{s}$, and $\boldsymbol{r}$; the average $\hat{\alpha}_k$ was used to initialize α. I then used multidimensional scaling on the matrix given by

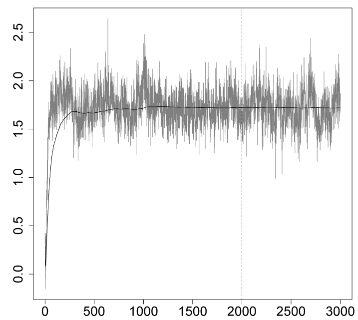
$$\frac{1}{|\mathcal{S}|} \sum_{k\in\mathcal{S}} (\alpha_k J_n + \boldsymbol{s}\mathbb{1}' + \mathbb{1}\boldsymbol{r}' - A_k)$$

where the entries of this matrix were shifted in order to ensure no off-diagonal element was less than 0.001. The β_{0k}'s were all initialized at 1, and β_1 was initialized at 0.1. $\boldsymbol{B}_s$, $\boldsymbol{B}_r$, and $\boldsymbol{B}_Z$ were initialized by finding the least squares estimate on the initialized $\boldsymbol{s}$, $\boldsymbol{r}$, and Z respectively. σ_Z^2 was then initialized by taking the sample variance of $\text{vec}(Z)$. Similarly Σ was initialized by taking the sample covariance of the initialized $\boldsymbol{s}$ and $\boldsymbol{r}$, and σ^2 was initialized by the sample variance of the initialized α_k's.

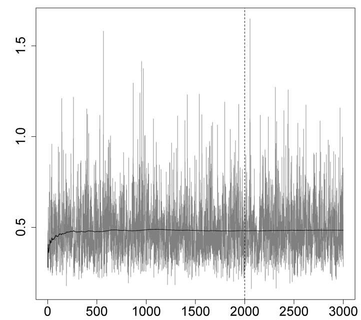
2 Advice-seeking network

2.1 MCMC plots

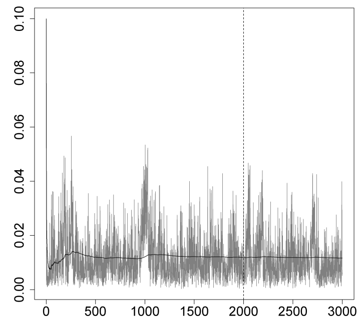
2.1.1 Trace plots



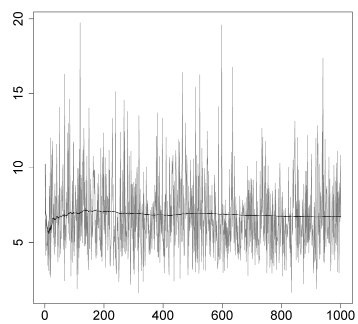
(a) α



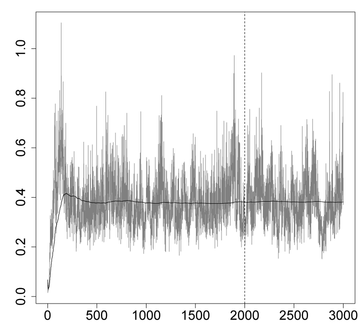
(b) σ^2



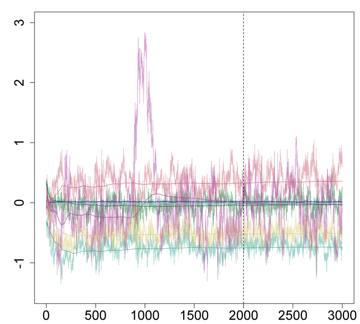
(c) β_1



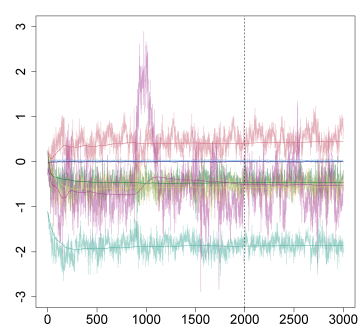
(d) θ



(e) σ_Z^2

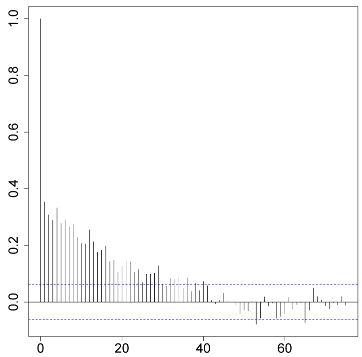


(f) $\boldsymbol{B}_s$

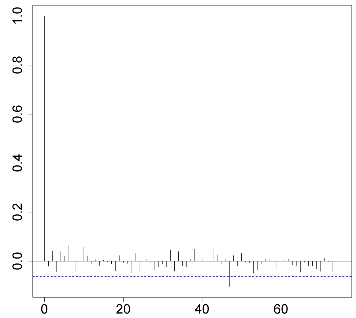


(g) $\boldsymbol{B}_r$

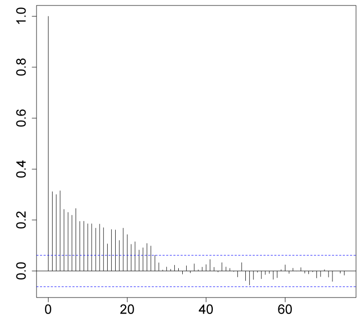
2.1.2 ACF plots



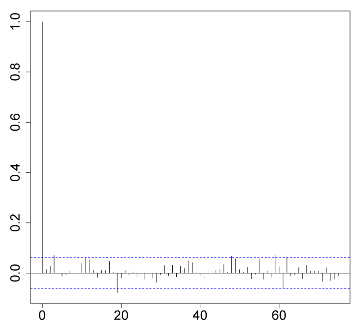
(a) α



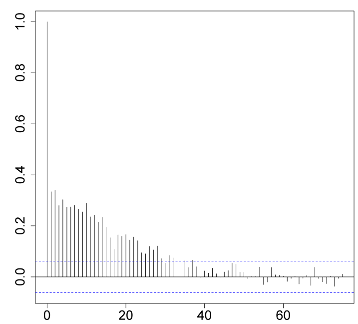
(b) σ^2



(c) β_1



(d) θ



(e) σ_Z^2

2.1.3 CCF Plots

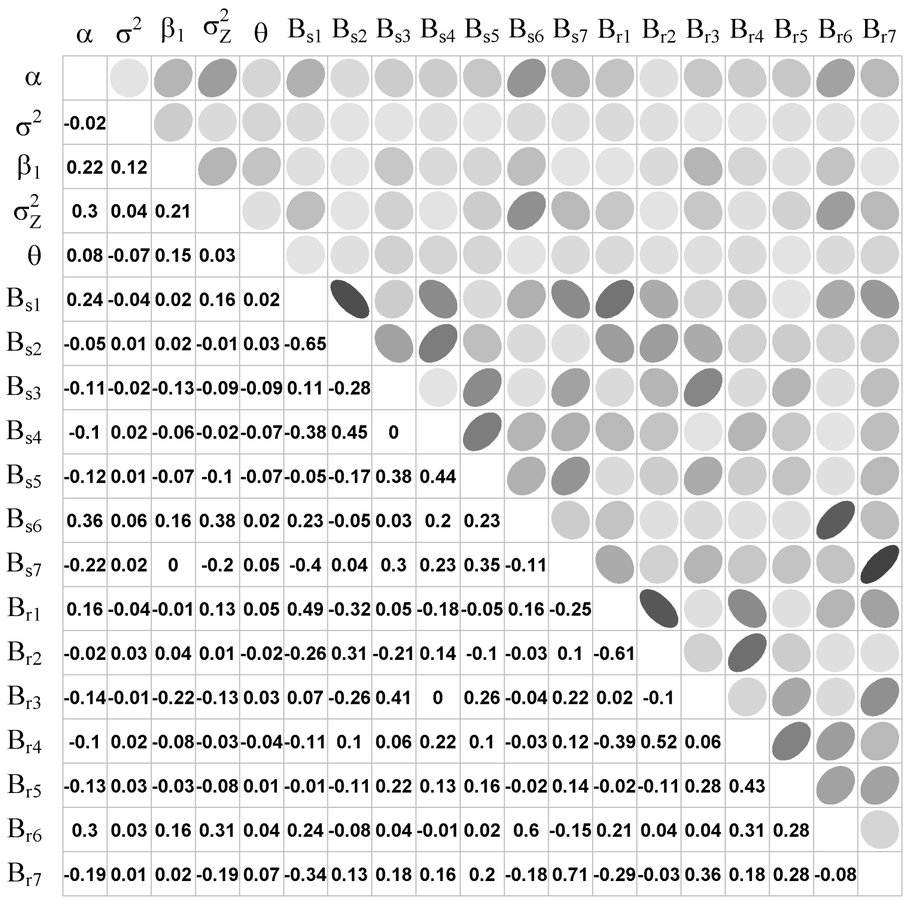
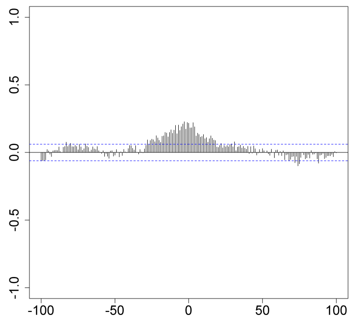
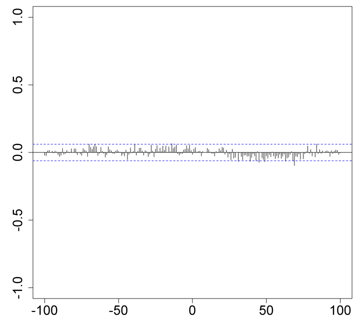


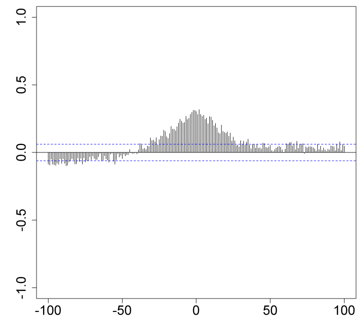
Figure 3: Posterior correlation



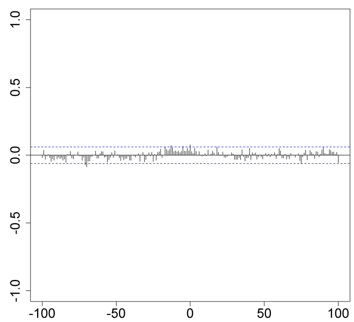
(a) α and β_1



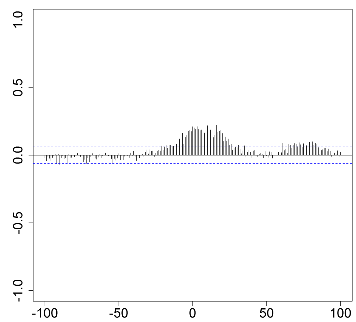
(b) α and σ^2



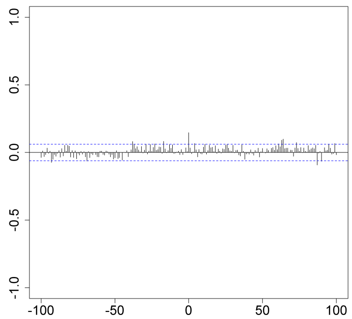
(c) α and σ_Z^2



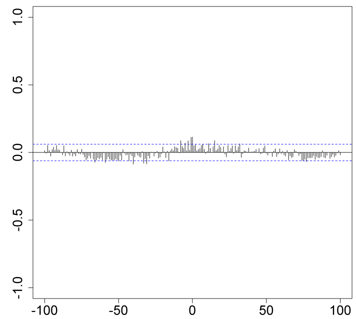
(d) α and θ



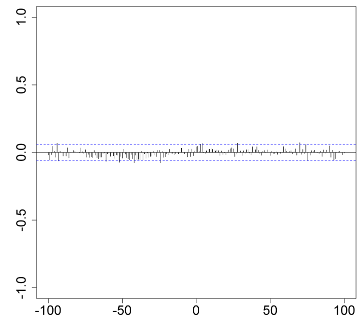
(e) $beta_1$ and σ_Z^2



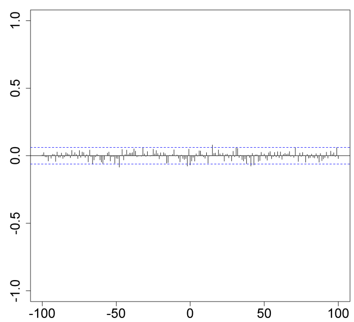
(a) β_1 and θ



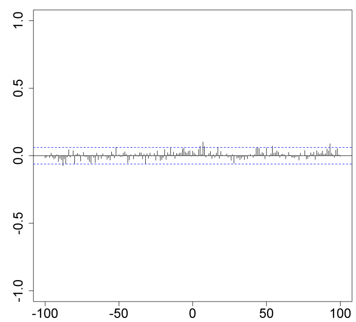
(b) β_1 and σ^2



(c) σ^2 and σ_Z^2



(d) σ^2 and θ



(e) σ_Z^2 and θ

2.2 Better starting positions

We analyzed the data of Section 5 of the main text using the initialization described above in Section 1.2. We also analyzed the same data after initializing the MCMC algorithm with the MAP estimators of the parameters and latent positions. This was done via an Iterated Conditional Modes algorithm, using Lagrange multipliers to constrain the means of s and r to be zero and of the β_{0k}'s to be one. Figure 6 shows the boxplots for the posterior draws of the model parameters comparing the two initialization schemes. From this figure it is apparent that the posterior distributions are negligibly different.

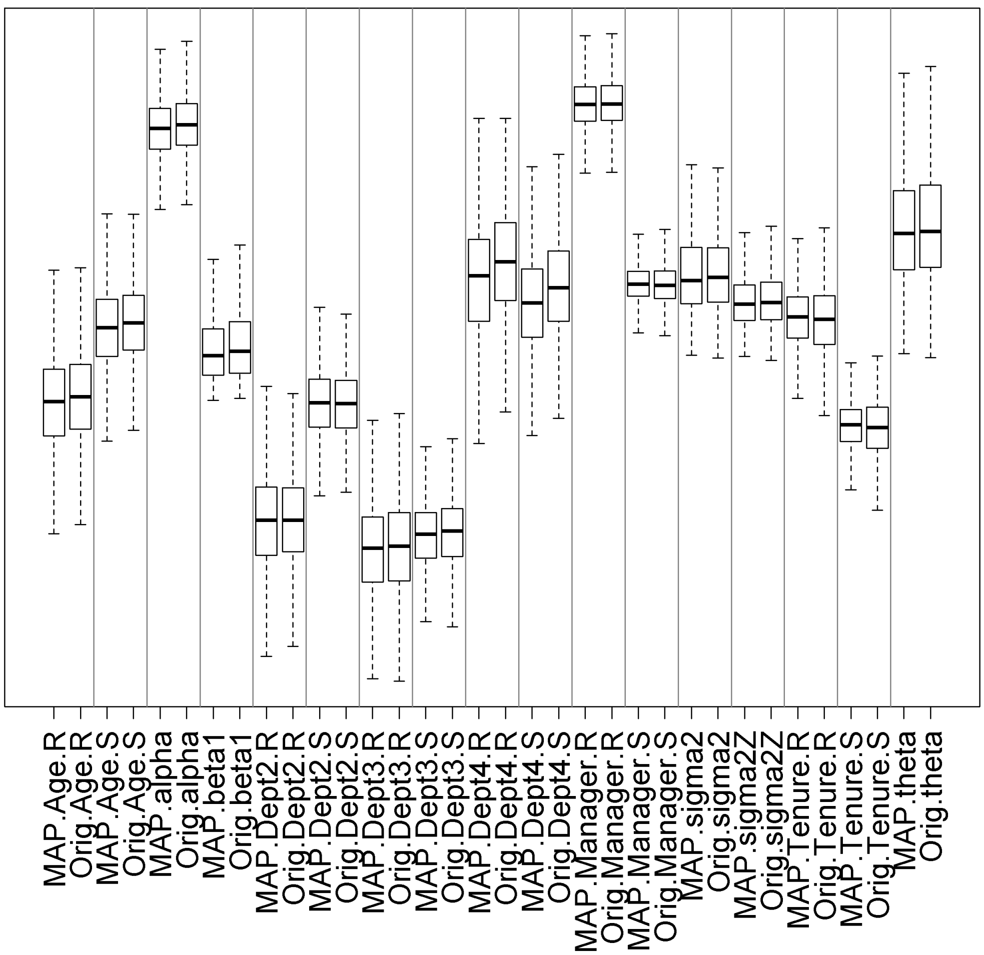
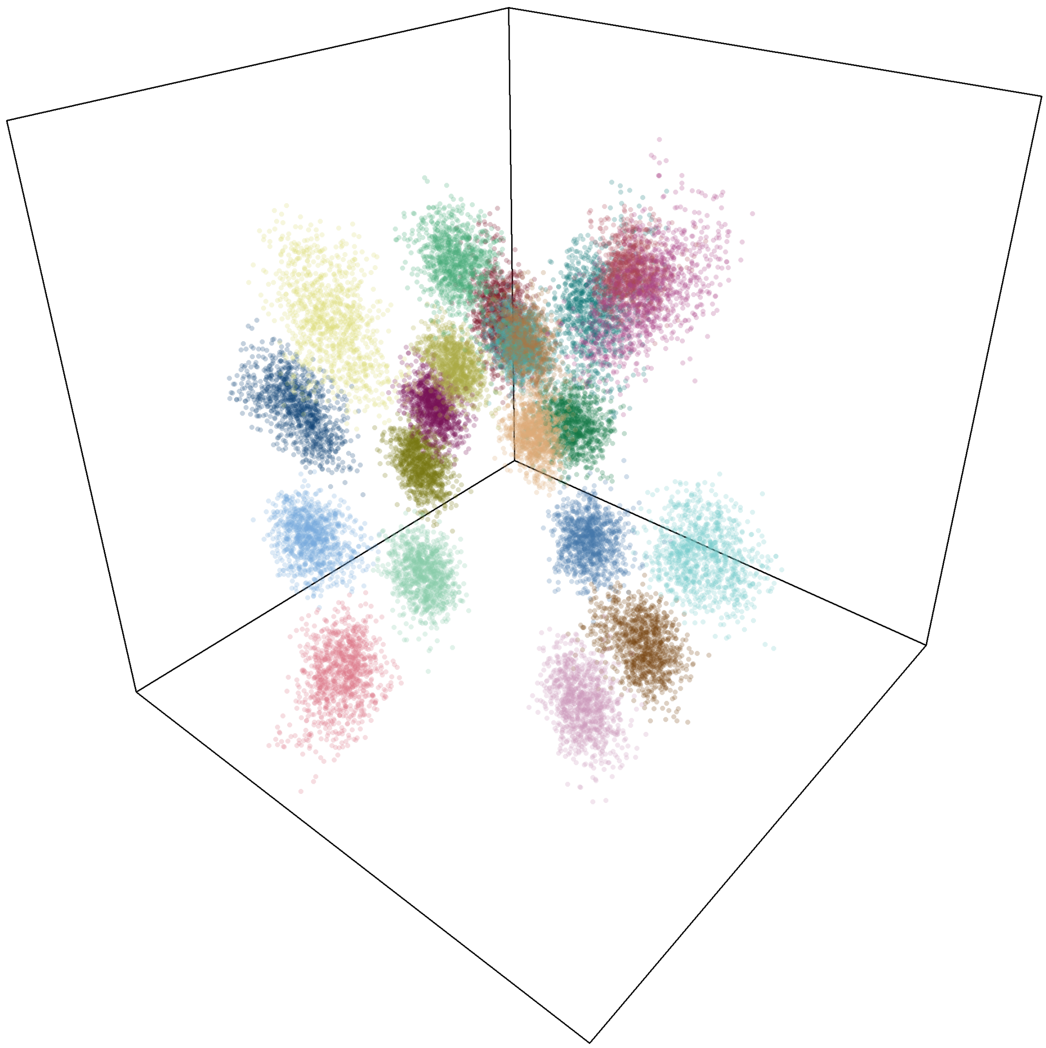


Figure 6: Comparing the posterior distributions of the model parameters under two different initialization schemes. The parameter draws have been rescaled by the same scalar for both initialization schemes but different scalars for each parameter in order to increase the clarity of the visualization.

2.3 Color figure of latent position uncertainties



2.4 Color figure of sender/receiver effect uncertainties