Statistical inference of eQTL sharing among a high number of tissues

Timothée Flutre, Sarah Urbut, Xiaoquan Wen, Matthew Stephens

September 10, 2014

This document describes the "type" model, an extension of the "config" model in details in the document "config_model.tex/pdf" available in the private repository paper-eQtlBma on GitHub.

1 Motivation

- *Most importantly, we assume that, given a type, the activity of an eQTL in tissue r is independent from its activity in other tissues*
- However, does this still allow us to exploit the shared effect among tissues? For example, in the **Likeilhood of the Whole Data Set pdf**, equation (5) and (6) show that the prior covariance matrix is indexed by configurations.

$$\boldsymbol{b}_{gp}|U_0 \sim \mathcal{N}_R(\mathbf{0}, U_0) \tag{1}$$

where, following Wen (2014), U_0 is parametrized as $(\Gamma_{gp}, \Delta_{gp})$:

$$p(U_0) = p(\Delta_{gp}|\Gamma_{gp})\mathsf{P}(\Gamma_{gp}) \tag{2}$$

so that Γ_{gp} is a binary matrix consisting of entry-wise non-zero indicators and is identical in size and layout to U_0, and Δ_{gp} is an indexed set of numerical values quantifying each non-zero entry in Γ_{gp}. The skeleton Γ_{gp} has $\boldsymbol{\gamma}_{gp}$ on the diagonal. Each off-diagonal entry $\Gamma_{gp,ij}$ is equal to 1 has long as diagonal elements $\Gamma_{gp,ii}$ and $\Gamma_{gp,jj}$ are both equal to 1.

- Doesn't saying that given a type, the activity of an eQTL in tissue is independent from its activity in other tissues directly negate the covariance in effects?
- In the config model, given a configuration $\boldsymbol{\gamma}$,

$$b_{gpr}|\gamma_{gpr}, \bar{b}_{gp}, \phi \sim \gamma_{gpr}\mathcal{N}(\bar{b}_{gp}, \phi^2) + (1 - \gamma_{gpr})\delta_0 \tag{3}$$

But here, according to Solution (1) of the document,

The target distribution $\boldsymbol{b}|\boldsymbol{q}_k$ can thus be approximated by the $\mathcal{N}_R(0, U_k)$ where:

$$U_k = \begin{pmatrix} q_{k1}(\phi_l^2 + q_{k1}\omega_l^2) & \cdots & q_{k1}q_{kR}\omega_l^2 \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & q_{kR}(\phi_l^2 + q_{kR}\omega_l^2) \end{pmatrix}$$

Since the fact that there are non-zero entries on the diagonal means that the convariance is non-zero and beta is a multivariate normal, doesn't b—q mean that the effects are *not independent conditional on type*? So are we always left with the BF conditional on q, or do we ever integrate out q?

I see in solution 2, we are effectively doing as in the supplement of the AOAS paper (i.e., equation A.6)

2 Question

In the type-based model, we use a latent indicator K-dimensional vector $\boldsymbol{t}_{gp}$ to denote the actual type. In case the SNP is not an eQTL,

$$\mathsf{P}(\boldsymbol{t}_{gp} = \mathbf{0}|v_{gp} = 0) = 1. \tag{4}$$

Otherwise, we assume the gene-SNP pair belongs to the k-th type with prior probability

$$\mathsf{P}(t_{gpk} = 1|v_{gp} = 1) = \pi_k \tag{5}$$

with the constraints $\forall k\ \pi_k \geq 0$ and $\sum_k \pi_k = 1$. All column vectors $\boldsymbol{t}_{gp}$ for all m_g SNPs are gathered into a $K \times m_g$ matrix T_g.

In the type-based model, we also index all tissues in which the eQTL is active via a latent indicator R-dimensional vector $\boldsymbol{\gamma}_{gp}$, which hence corresponds to its configuration. However, compare to the "config" model where $\boldsymbol{\gamma}_{gp}$ simply corresponds to the latent variable $\boldsymbol{c}_{gp}$ indexing the configuration, we now put a prior on $\boldsymbol{\gamma}_{gp}$. In case the SNP is not an eQTL,

$$\mathsf{P}(\boldsymbol{\gamma}_{gp} = \mathbf{0}|\boldsymbol{t}_{gp} = \mathbf{0}) = 1. \tag{6}$$

Otherwise, we assume the eQTL is active in the r-th tissue with prior probability depending on the type

$$\mathsf{P}(\gamma_{gpr} = 1|t_{gpk} = 1) = q_{kr}, \tag{7}$$

for which we could also parametrize the q_{kr} in terms of tissue-specific annotations, e.g. DNase peaks, and therefore have q_{pkr}. Joining the column vectors $\boldsymbol{\gamma}_{gp}$ for all K types, we obtain a latent $R \times K$ matrix Γ_{gp}. All these matrices are gathered into $\boldsymbol{\Gamma}_g = (\Gamma_{g1}, \ldots, \Gamma_{gm_g})$. As a result, this bypasses the need for the J-dimensional vectors $\boldsymbol{c}_{gp}$ where $J = 2^R - 1$ and the corresponding prior probabilities (the η_j's).

- I understand that the $K \times m_g$ matrix T_g represents the identity of each SNP in the columns, so that the column sum is equal to 1 and the row sum is equal to the number of SNPs in the class.
- Furthermore, the $R \times K$ matrix Γ_{gp} represents the tissue-specific patterns of expression for each class across tissues, such that the column sums will represent the number of tissues a SNP of type 'k' might be active and the rows reprsent the total number of types from which a tissue derives activity.
- However, since the matrix Γ_{gp} is the same for all SNPs of a particular type, why is it necessary to stack the Γ_{gp} into $\boldsymbol{\Gamma}_g$?

3 Focus on a single gene-SNP pair

See the corresponding section in the document "config_model.tex/pdf".

3.1 Introducing types in the Bayes factor

From the hierarchical model above, we can write the Bayes factor measuring the support in the data for the pair made of gene g and SNP p to be an eQTL in at least one tissue, BF_{gp}.

In the configuration model, we average over the configurations and the grid:

$$\mathrm{BF}_{gp}(\boldsymbol{\eta}, \boldsymbol{\lambda}) = \sum_{j,l} \eta_j \ \lambda_l \ \mathrm{BF}_{gpjl}(\omega_l, \phi_l) \tag{8}$$

where

$$\begin{aligned} \mathrm{BF}_{gpjl}(\omega_l, \phi_l) &= \frac{p(Y_g | \boldsymbol{X}_{gp}, c_{gpj} = 1, d_{gpl} = 1)}{p(Y_g | \boldsymbol{X}_{gp}, z_g = 0)} \\ &= \frac{\int p(Y_g | \boldsymbol{X}_{gp}, \boldsymbol{\mu}, \boldsymbol{b}, \Sigma) \ p(\boldsymbol{\mu}, \Sigma) \ p(\boldsymbol{b} | \boldsymbol{\gamma}, \omega_l, \phi_l) \ \mathrm{d}\boldsymbol{b} \ \mathrm{d}\boldsymbol{\mu} \ \mathrm{d}\Sigma}{\int p(Y_g | \boldsymbol{X}_{gp}, \boldsymbol{\mu}, \boldsymbol{b} = \boldsymbol{0}, \Sigma) \ p(\boldsymbol{\mu}, \Sigma) \ \mathrm{d}\boldsymbol{\mu} \ \mathrm{d}\Sigma} \end{aligned} \tag{9}$$

This Bayes factor can be calculated analytically using the Laplace approximation from Wen & Stephens (AoAS 2014, if individuals are different between tissues) or from Wen (Biometrics 2014, if individuals are shared between tissues).

In the type model, we would first average over the types and the grid:

$$\mathrm{BF}_{gp}(\boldsymbol{\pi}, \boldsymbol{\lambda}, Q) = \sum_{k,l} \pi_k \ \lambda_l \ \mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \omega_l, \phi_l) \tag{10}$$

where

$$\begin{aligned} \mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \omega_l, \phi_l) &= \frac{p(Y_g | \boldsymbol{X}_{gp}, t_{gpk} = 1, d_{gpl} = 1)}{p(Y_g | \boldsymbol{X}_{gp}, z_g = 0)} \\ &= \frac{\int p(Y_g | \boldsymbol{X}_{gp}, \boldsymbol{\mu}, \boldsymbol{b}, \Sigma) \ p(\boldsymbol{\mu}, \Sigma) \ p(\boldsymbol{b} | \boldsymbol{q}_k, \omega_l, \phi_l) \ \mathrm{d}\boldsymbol{b} \ \mathrm{d}\boldsymbol{\mu} \ \mathrm{d}\Sigma}{\int p(Y_g | \boldsymbol{X}_{gp}, \boldsymbol{\mu}, \boldsymbol{b} = \boldsymbol{0}, \Sigma) \ p(\boldsymbol{\mu}, \Sigma) \ \mathrm{d}\boldsymbol{\mu} \ \mathrm{d}\Sigma} \end{aligned} \tag{11}$$

But then, in the "naive" version of the type model, we also have to average over all configurations possibly generated by type k (via its $\boldsymbol{q}_k$) in the numerator of the Bayes factor:

$$p(\boldsymbol{b} | \boldsymbol{q}_k, \omega_l, \phi_l) = \sum_{j=1}^{J} \mathsf{P}(\boldsymbol{\gamma} | \boldsymbol{q}_k) \ p(\boldsymbol{b} | \boldsymbol{\gamma}, \omega_l, \phi_l) \tag{12}$$

which is clearly intractable when $J = 2^R - 1$ is large.

Note that the first version of the "type" model (currently implemented in the eQtlBma package) drastically simplifies this sum by only considering the consistent configuration.

3.2 Solution 1: average over all configurations per type

William proposed to approximate $\boldsymbol{b}|\boldsymbol{q}_k$ (12) with a multivariate Normal by matching the first two moments (as we can calculate them analytically, see below). We already know that the priors are $\forall r\ b_r|\bar{b},\gamma_r \sim \gamma_r\mathcal{N}(\bar{b},\phi_l^2) + (1-\gamma_r)\delta_0$, $\bar{b} \sim \mathcal{N}(0,\omega_l^2)$ and $\gamma_r \sim \mathcal{B}(q_{kr})$. Using the laws of total expectations, variances and covariances, we first average over the possible values of $\boldsymbol{\gamma}$ as parametrized by $\boldsymbol{q}_k$, conditional on $\bar{b}$:

$$\mathbb{E}[b_r \mid q_{kr},\bar{b}] = \mathbb{E}[\mathbb{E}[b_r|q_{kr},\bar{b},\gamma_r]] = \mathbb{E}[\gamma_r\bar{b}|q_{kr}] = \bar{b}q_{kr}$$
$$\mathbb{V}[b_r \mid q_{kr},\bar{b}] = \mathbb{E}[\mathbb{V}[b_r|q_{kr},\bar{b},\gamma_r]] + \mathbb{V}[\mathbb{E}[b_r|q_{kr},\bar{b},\gamma_r]] = \phi_l^2 q_{kr} + \bar{b}^2 q_{kr}(1-q_{kr})$$
$$\mathrm{Cov}[b_r, b_{r'} \mid q_{kr}, q_{kr'}, \bar{b}] = 0$$

And then we integrate out $\bar{b}$:

$$\mathbb{E}[b_r \mid q_{kr}] = 0$$
$$\mathbb{V}[b_r \mid q_{kr}] = \phi_l^2 q_{kr} + q_{kr}^2\omega_l^2$$
$$\mathbb{C}[b_r, b_{r'} \mid q_{kr}, q_{kr'}] = q_{kr}q_{kr'}\omega_l^2$$

The target distribution $\boldsymbol{b}|\boldsymbol{q}_k$ can thus be approximated by the $\mathcal{N}_R(0,U_k)$ where:

$$U_k = \begin{pmatrix} q_{k1}(\phi_l^2 + q_{k1}\omega_l^2) & \cdots & q_{k1}q_{kR}\omega_l^2 \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & q_{kR}(\phi_l^2 + q_{kR}\omega_l^2) \end{pmatrix}$$

This approximation should be more and more accurate as R increases.

With this prior for $\boldsymbol{b}|\boldsymbol{q}_k$, we can directly use the approximated Bayes factor from Wen (Biometrics 2014):

$$\mathrm{BF}_{gpkl}(\boldsymbol{q}_k,\omega_l,\phi_l) = |I + \hat{V}^{-1}W_k|^{-\frac{1}{2}} \exp\left(\frac{1}{2}\hat{\boldsymbol{\beta}}^T\hat{V}^{-1}\left[W_k(I+\hat{V}^{-1}W_k)^{-1}\right]\hat{V}^{-1}\hat{\boldsymbol{\beta}}\right) \tag{13}$$

where $\hat{\boldsymbol{\beta}}$ is the MLE for $\boldsymbol{\beta}$ and $\hat{V}$ its variance. In the scale-invariance formulation, we have $\boldsymbol{b} = diag(\Sigma)^{-\frac{1}{2}}\boldsymbol{\beta}$, so that $\boldsymbol{b}|\boldsymbol{q}_k \sim \mathcal{N}_R(\boldsymbol{0},U_k)$ induces the prior $\boldsymbol{\beta}|\boldsymbol{q}_k \sim \mathcal{N}_R(\boldsymbol{0},W_k)$ where $W_k = diag(\Sigma)^{-\frac{1}{2}}U_k diag(\Sigma)^{-\frac{1}{2}}$.

However, the remaining issue is to calculate the derivative of the log of this Bayes factor w.r.t. $\boldsymbol{q}_k$ in the M step of the EM algorithm (see below).

3.3 Solution 2: condition on the average effect instead of its prior variance

Matthew proposed to calculate a new Bayes factor conditional on $(\bar{b}_{gpl},\phi_l)$ instead of (ω_l,ϕ_l):

$$\mathrm{BF}_{gpkl}(\boldsymbol{q}_k,\omega_l,\phi_l) = \frac{\int p(\bar{b}_{gpl})\ p(Y_g|\boldsymbol{X}_{gp},t_{gpk}=1,d_{gpl}=1,\bar{b}_{gpl})\ \mathrm{d}\bar{b}_{gpl}}{p(Y_g|\boldsymbol{X}_{gp},z_g=0)} \tag{14}$$

The idea comes from the assumptions that, given the type, the γ_r's are independent and, given $\bar{b}$, the b_r are also independent (note that we also have to assume that the errors are independent between tissues):

$$\begin{aligned}
\mathrm{BF}_{gpkl}(\boldsymbol{q}_k,\omega_l,\phi_l) &= \frac{\int p(\bar{b}_{gpl})\ \prod_r p(\boldsymbol{y}_{gr}|\boldsymbol{X}_{gp},t_{gpk}=1,d_{gpl}=1,\bar{b}_{gpl})\ \mathrm{d}\bar{b}_{gpl}}{\prod_r p(\boldsymbol{y}_{gr}|\boldsymbol{X}_{gp},z_g=0)} \\
&= \frac{\int p(\bar{b}_{gpl})\ \prod_r [q_{kr}p(\boldsymbol{y}_{gr}|\boldsymbol{X}_{gp},d_{gpl}=1,\bar{b}_{gpl},\gamma_{gpr}=1) + (1-q_{kr})p(\boldsymbol{y}_{gr}|\boldsymbol{X}_{gp},\gamma_{gpr}=0)]\ \mathrm{d}\bar{b}_{gpl}}{\prod_r p(\boldsymbol{y}_{gr}|\boldsymbol{X}_{gp},z_g=0)} \\
&= \int_{-\infty}^{+\infty} f(\bar{b}_{gpl},\boldsymbol{q}_k,\omega_l,\phi_l)\ \mathrm{d}\bar{b}_{gpl}
\end{aligned} \tag{15}$$

where the function f is defined as:

$$f(\bar{b}_{gpl}, \boldsymbol{q}_k, \omega_l, \phi_l) = \frac{1}{\omega_l\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\bar{b}_{gpl}}{\omega_l}\right)^2\right) \prod_r \left[q_{kr}\mathrm{BF}_{gplr}(\bar{b}_{gpl}, \phi_l) + (1 - q_{kr})\right] \tag{16}$$

This integral can't be analytically calculated but it is unidimensional and should thus be efficiently approximated numerically, e.g. with QAGI as implemented in the GSL.

The goal now is to calculate BF_{gplr} which effectively test $b_{gplr} \sim \mathcal{N}(\bar{b}_{gpl}, \phi_l^2)$ versus $b_{gplr} \sim \mathcal{N}(0, 0)$. As we need to focus only on a single gene-snp pair in a single tissue, we will ignore indices g, p, l and r in the remaining of this section. Moreover, we will heavily borrow from the appendices of two papers, Wen & Stephens (2014) and Wen (2014).

For one individual, we model its expression level y_i at the gene as

$$y_i = \mu + \beta g_i + \epsilon_i \text{ with } \epsilon_i \sim \mathcal{N}(0, \tau^{-1}) \tag{17}$$

where g_i is the genotype at the SNP.

To have a notation similar to the articles mentioned above, we will replace the scalar μ by β_c, and similarly β by β_g. As there can be other covariates (e.g. sex), we will rather use the vector notation $\boldsymbol{\beta}_c := \beta_c$ and $\boldsymbol{\beta}_g := \beta_g$. These vectors are gathered into $\boldsymbol{\beta} := (\boldsymbol{\beta}_c\ \boldsymbol{\beta}_g)^T$. Concerning the design matrix, we will use the $N \times 2$ matrix $X := (X_c\ X_g)$ where $X_c := \mathbf{1}$ and $X_g := \boldsymbol{g}$.

As a result, equation (17) can be written as:

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon} \text{ with } \boldsymbol{\epsilon} \sim \mathcal{N}_N(0, \tau^{-1}I) \tag{18}$$

For reasons detailed in the protocol S1 of Servin & Stephens (2007), we now shift to the "exchangeable-standardized"model (ES model in Wen & Stephens, 2014). Let's define $\tilde{\boldsymbol{y}} := \sqrt{\tau}\boldsymbol{y}$ and $\boldsymbol{b} = \sqrt{\tau}\boldsymbol{\beta}$. Under the full model, the likelihood is

$$p(\boldsymbol{y}|X, \tau, \boldsymbol{\beta}) = p(\boldsymbol{y}|X, \tau, \boldsymbol{b}) \propto \exp\left(-\frac{1}{2}(\tilde{\boldsymbol{y}} - X\boldsymbol{b})^T(\tilde{\boldsymbol{y}} - X\boldsymbol{b})\right) \tag{19}$$

Using the maximum likelihood estimate $\hat{\boldsymbol{b}} := (\hat{b}_c\ \hat{b}_g)^T$ allows to reformulate (19) using

$$(\tilde{\boldsymbol{y}} - X\boldsymbol{b})^T(\tilde{\boldsymbol{y}} - X\boldsymbol{b}) = (\tilde{\boldsymbol{y}} - X\hat{\boldsymbol{b}})^T(\tilde{\boldsymbol{y}} - X\hat{\boldsymbol{b}}) + (\boldsymbol{b} - \hat{\boldsymbol{b}})^T X^T X(\boldsymbol{b} - \hat{\boldsymbol{b}}) \tag{20}$$

Similarly, for the null model (i.e. full model without the genotype), the likelihood is

$$p(\boldsymbol{y}|X_c, \tau, \boldsymbol{\beta}_c) = p(\boldsymbol{y}|X_c, \tau, \boldsymbol{b}_c) \propto \exp\left(-\frac{1}{2}(\tilde{\boldsymbol{y}} - X_c\boldsymbol{b}_c)^T(\tilde{\boldsymbol{y}} - X_c\boldsymbol{b}_c)\right) \tag{21}$$

and defining $\hat{\boldsymbol{b}}_0 := (\hat{b}_0)^T$ gives

$$(\tilde{\boldsymbol{y}} - X_c\boldsymbol{b}_c)^T(\tilde{\boldsymbol{y}} - X_c\boldsymbol{b}_c) = (\tilde{\boldsymbol{y}} - X_c\hat{\boldsymbol{b}}_0)^T(\tilde{\boldsymbol{y}} - X_c\hat{\boldsymbol{b}}_0) + (\boldsymbol{b}_c - \hat{\boldsymbol{b}}_0)^T X_c^T X_c(\boldsymbol{b}_c - \hat{\boldsymbol{b}}_0) \tag{22}$$

However, note that $\hat{b}_0 \neq \hat{b}_c$. Indeed, for the full model, the Normal equation is $X^T X\hat{\boldsymbol{b}} = X^T\tilde{\boldsymbol{y}}$ whereas for the null model the Normal equation is $X_c^T X_c\hat{\boldsymbol{b}}_0 = X_c^T\tilde{\boldsymbol{y}}$, yielding

$$\hat{b}_0 = \hat{b}_c + (X_c^T X_c)^{-1} X_c^T X_g \hat{b}_g \tag{23}$$

To test the effect of the SNP, we use the Bayes factor

$$\mathrm{BF}(\bar{b},\phi) = \frac{\int p(\tau)\left(\int p(\boldsymbol{y}|X,\tau,\boldsymbol{b})\; p(\boldsymbol{b}|\tau,\bar{\boldsymbol{b}},\phi)\;\mathrm{d}\boldsymbol{b}\right)\mathrm{d}\tau}{\int p(\tau)\left(\int p(\boldsymbol{y}|X_c,\tau,\boldsymbol{b}_c)\; p(\boldsymbol{b}_c|\tau)\;\mathrm{d}\boldsymbol{b}_c\right)\mathrm{d}\tau} \tag{24}$$

for which the prior on $\boldsymbol{b}$ under the alternative is

$$p(\boldsymbol{b}|\tau,\bar{\boldsymbol{b}},\phi) = (2\pi)^{-1}|\Phi|^{-1/2}\exp\left(-\frac{1}{2}(\boldsymbol{b}-\bar{\boldsymbol{b}})^T\Phi^{-1}(\boldsymbol{b}-\bar{\boldsymbol{b}})\right) \tag{25}$$

where $\Phi := diag(v^2,\phi^2)$. In terms of notation, we will also use $v^2 := \Phi_c$ and $\phi^2 := W_g$. The prior on τ is conjuguate

$$p(\tau) = \frac{(\frac{l}{2})^{\frac{m}{2}}}{\Gamma(\frac{m}{2})}\tau^{\frac{m}{2}-1}\exp\left(-\frac{l}{2}\tau\right) \tag{26}$$

See the articles cited above for the other, less important priors.

The first step to compute the Bayes factor (24) is to integrate $\boldsymbol{b}$ using its prior (25):

$$F_a = \int M_a\;\mathrm{d}\boldsymbol{b} \text{ where } M_a = p(\boldsymbol{b}|\tau,\bar{\boldsymbol{b}},\phi)\; p(\boldsymbol{y}|X,\tau,\boldsymbol{b}) \tag{27}$$

Keeping only the terms in $\boldsymbol{b}$

$$M_a \propto \exp(\boldsymbol{b}^T\Phi^{-1}\boldsymbol{b} - \boldsymbol{b}^T\Phi^{-1}\bar{\boldsymbol{b}} - \bar{\boldsymbol{b}}^T\Phi^{-1}\boldsymbol{b}) \times \exp(\boldsymbol{b}^TX^TX\boldsymbol{b} - \boldsymbol{b}^TX^TX\hat{\boldsymbol{b}} - \hat{\boldsymbol{b}}^TX^TX\boldsymbol{b}) \tag{28}$$

Factorizing:

$$M_a \propto \exp(\boldsymbol{b}^T(\Phi^{-1}+X^TX)\boldsymbol{b} - (\bar{\boldsymbol{b}}^T\Phi^{-1} + \hat{\boldsymbol{b}}^TX^TX)\boldsymbol{b} - \boldsymbol{b}^T(\Phi^{-1}\bar{\boldsymbol{b}} + X^TX\hat{\boldsymbol{b}})) \tag{29}$$

Introducing the matrix $\Omega = (\Phi^{-1}+X^TX)^{-1}$ as in Servin & Stephens (2007), which is symmetric (i.e. $\Omega^T = \Omega \Rightarrow \Omega^{-1}\Omega^T = I$), and "completing the square" gives:

$$\begin{aligned} M_a \propto \exp((\boldsymbol{b} - \Omega(\Phi^{-1}\bar{\boldsymbol{b}} + X^TX\hat{\boldsymbol{b}}))^T\Omega^{-1}(\boldsymbol{b} - \Omega(\Phi^{-1}\bar{\boldsymbol{b}} + X^TX\hat{\boldsymbol{b}}))) \\ \times \exp(-(\Phi^{-1}\bar{\boldsymbol{b}} + X^TX\hat{\boldsymbol{b}})^T\Omega(\Phi^{-1}\bar{\boldsymbol{b}} + X^TX\hat{\boldsymbol{b}})) \end{aligned} \tag{30}$$

Recognizing the first exponential as the kernel of a Normal distribution which must integrate to 1, re-adding the terms not depending on $\boldsymbol{b}$ and factorizing, we get:

$$\begin{aligned} F_a = \left(\frac{2\pi}{\tau}\right)^{-N/2}|\Phi|^{-1/2}|\Omega|^{1/2}\exp\left[-\frac{1}{2}(\tilde{\boldsymbol{y}} - X\hat{\boldsymbol{b}})^T(\tilde{\boldsymbol{y}} - X\hat{\boldsymbol{b}})\right] \\ \times \exp\left[-\frac{1}{2}(\hat{\boldsymbol{b}}^TX^TX((X^TX)^{-1} - \Omega)X^TX\hat{\boldsymbol{b}} + \bar{\boldsymbol{b}}^T\Phi^{-1}(\Phi-\Omega)\Phi^{-1}\bar{\boldsymbol{b}} - 2\bar{\boldsymbol{b}}^T\Phi^{-1}\Omega X^TX\hat{\boldsymbol{b}})\right] \end{aligned} \tag{31}$$

Similarly for the null

$$\begin{aligned} F_0 = \left(\frac{2\pi}{\tau}\right)^{-N/2}|\Phi_0|^{-1/2}|\Omega_0|^{1/2}\exp\left[-\frac{1}{2}(\tilde{\boldsymbol{y}} - X_c\hat{\boldsymbol{b}}_0)^T(\tilde{\boldsymbol{y}} - X_c\hat{\boldsymbol{b}}_0)\right] \\ \times \exp\left[-\frac{1}{2}(\hat{\boldsymbol{b}}_0^TX_c^TX_c((X_c^TX_c)^{-1} - \Omega_0)X_c^TX_c\hat{\boldsymbol{b}}_0)\right] \end{aligned} \tag{32}$$

The second step to compute the Bayes factor (24) is to integrate τ using its prior (26) while taking the limits $v^2 \to +\infty$ and $l, m \to 0$, as in Servin & Stephens (2007):

$$\mathrm{BF}(\bar{b},\phi) = \lim\frac{\int p(\tau)\; F_a\;\mathrm{d}\tau}{\int p(\tau)\; F_0\;\mathrm{d}\tau} = \frac{\int K_a\;\mathrm{d}\tau}{\int K_0\;\mathrm{d}\tau} \;\times\; \lim\frac{|\Phi|^{-1/2}}{|\Phi_0|^{-1/2}}\;\frac{|\Omega|^{1/2}}{|\Omega_0|^{1/2}} \tag{33}$$

Let us first take care of the second term. Given the definition of Φ, we have

$$|\Phi|^{-1/2} = \frac{1}{v}\,\frac{1}{\phi}$$

and

$$|\Phi_0|^{-1/2} = \frac{1}{v}$$

Given the definition of Ω, we have

$$\begin{aligned}\Omega &= \left(\begin{pmatrix}\frac{1}{v^2} & 0\\ 0 & \frac{1}{\phi^2}\end{pmatrix} + \begin{pmatrix}N & N\bar{g}\\ N\bar{g} & \boldsymbol{g}^T\boldsymbol{g}\end{pmatrix}\right)^{-1}\\ &= \frac{1}{(N+\frac{1}{v^2})(\boldsymbol{g}^T\boldsymbol{g}+\frac{1}{\phi^2}) - N^2\bar{g}^2}\begin{pmatrix}\boldsymbol{g}^T\boldsymbol{g}+\frac{1}{\phi^2} & -N\bar{g}\\ -N\bar{g} & N+\frac{1}{v^2}\end{pmatrix}\end{aligned} \tag{34}$$

so that

$$|\Omega|^{1/2} = \sqrt{\frac{1}{(N+\frac{1}{v^2})(\boldsymbol{g}^T\boldsymbol{g}+\frac{1}{\phi^2}) - N^2\bar{g}^2}}$$

Similarly

$$\Omega_0 = \left(\frac{1}{v^2} + N\right)^{-1}$$

so that

$$|\Omega_0|^{1/2} = \sqrt{\frac{1}{\frac{1}{v^2}+N}}$$

Recognizing the standard error of $\hat{b}_g$

$$\delta^2 = \frac{1}{\boldsymbol{g}^T\boldsymbol{g} - N\bar{g}^2}$$

we can obtain

$$\lim_{v\to+\infty}\frac{|\Phi|^{-1/2}}{|\Phi_0|^{-1/2}}\,\frac{|\Omega|^{1/2}}{|\Omega_0|^{1/2}} = \frac{1}{\phi}\sqrt{\frac{1}{\frac{1}{\delta^2}+\frac{1}{\phi^2}}} = \sqrt{\frac{\delta^2}{\delta^2+\phi^2}} \tag{35}$$

Going back to the Bayes factor as in (33), let's first assume that τ is known. We then need a relationship

between (20) and (22) (see equation 16 in the supplement of Wen, 2014):

$$
\begin{aligned}
&(\tilde{\boldsymbol{y}} - X_c\hat{\boldsymbol{b}}_0)^T(\tilde{\boldsymbol{y}} - X_c\hat{\boldsymbol{b}}_0) - (\tilde{\boldsymbol{y}} - X\hat{\boldsymbol{b}})^T(\tilde{\boldsymbol{y}} - X\hat{\boldsymbol{b}}) \\
&= -\tilde{\boldsymbol{y}}^T X_c\hat{\boldsymbol{b}}_0 - \hat{\boldsymbol{b}}_0^T X_c^T\tilde{\boldsymbol{y}} + \hat{\boldsymbol{b}}_0^T X_c^T X_c\hat{\boldsymbol{b}}_0 + \tilde{\boldsymbol{y}}^T X\hat{\boldsymbol{b}} + \hat{\boldsymbol{b}}^T X^T\tilde{\boldsymbol{y}} - \hat{\boldsymbol{b}}^T X^T X\hat{\boldsymbol{b}} \\
&\text{use formula (23) to replace } \hat{\boldsymbol{b}}_0 \\
&= -\tilde{\boldsymbol{y}}^T X_c(\hat{\boldsymbol{b}}_c + (X_c^T X_c)^{-1}X_c^T X_g\hat{\boldsymbol{b}}_g) - (\hat{\boldsymbol{b}}_c^T + \hat{\boldsymbol{b}}_g^T X_g^T X_c(X_c^T X_c)^{-1})X_c^T\tilde{\boldsymbol{y}} \\
&+ (\hat{\boldsymbol{b}}_c^T + \hat{\boldsymbol{b}}_g^T X_g^T X_c(X_c^T X_c)^{-1})X_c^T X_c(\hat{\boldsymbol{b}}_c + (X_c^T X_c)^{-1}X_c^T X_g\hat{\boldsymbol{b}}_g) \\
&+ \tilde{\boldsymbol{y}}^T(X_c\hat{\boldsymbol{b}}_c + X_g\hat{\boldsymbol{b}}_g) + (\hat{\boldsymbol{b}}_c^T X_c^T + \hat{\boldsymbol{b}}_g^T X_g^T)\tilde{\boldsymbol{y}} \\
&- (\hat{\boldsymbol{b}}_c^T X_c^T + \hat{\boldsymbol{b}}_g^T X_g^T)(X_c\hat{\boldsymbol{b}}_c + X_g\hat{\boldsymbol{b}}_g) \\
&\text{develop and simplify} \\
&= -\tilde{\boldsymbol{y}}^T X_c(X_c^T X_c)^{-1}X_c^T X_g\hat{\boldsymbol{b}}_g - \hat{\boldsymbol{b}}_g^T X_g^T X_c(X_c^T X_c)^{-1}X_c^T\tilde{\boldsymbol{y}} + \hat{\boldsymbol{b}}_g^T X_g^T X_c(X_c^T X_c)^{-1}X_c^T X_g\hat{\boldsymbol{b}}_g \\
&+ \tilde{\boldsymbol{y}}^T X_g\hat{\boldsymbol{b}}_g + \hat{\boldsymbol{b}}_g^T X_g^T\tilde{\boldsymbol{y}} - \hat{\boldsymbol{b}}_g^T X_g^T X_g\hat{\boldsymbol{b}}_g \\
&\text{use the Normal equation to replace } \tilde{\boldsymbol{y}} \\
&= \hat{\boldsymbol{b}}_g^T\left(X_g^T X_g - X_g^T X_c(X_c^T X_c)^{-1}X_c^T X_g\right)\hat{\boldsymbol{b}}_g
\end{aligned}
\tag{36}
$$

Here again we can recognize the standard error of $\hat{b}_g$, this time with the vector notation

$$V_g^{-1} = X_g^T X_g - X_g^T X_c(X_c^T X_c)^{-1}X_c^T X_g \tag{37}$$

Using these relationships in (33) (when assuming τ known)

$$
\begin{aligned}
\frac{\int K_a \,\mathrm{d}\tau}{\int K_0 \,\mathrm{d}\tau} &= \exp\left(\frac{1}{2}\left(\hat{\boldsymbol{b}}_g^T V_g^{-1}\hat{\boldsymbol{b}}_g\right)\right) \\
&\times \exp\left(\frac{1}{2}\left(\hat{\boldsymbol{b}}_0^T X_c^T X_c((X_c^T X_c)^{-1} - \Omega_0)X_c^T X_c\hat{\boldsymbol{b}}_0\right)\right) \\
&\times \exp\left(-\frac{1}{2}\left(\hat{\boldsymbol{b}}^T X^T X((X^T X)^{-1} - \Omega)X^T X\hat{\boldsymbol{b}} - \bar{\boldsymbol{b}}^T\Phi^{-1}(\Phi - \Omega)\Phi^{-1}\bar{\boldsymbol{b}} + 2\bar{\boldsymbol{b}}^T\Phi^{-1}\Omega X^T X\hat{\boldsymbol{b}}\right)\right)
\end{aligned}
\tag{38}
$$

When taking the limit on v, $\Omega_0 = (X_c^T X_c)^{-1}$ and the second exponential disappears. For the third exponential, we can use results in the appendix of Wen (2014) which, after taking the limit, give:

$$\hat{\boldsymbol{b}}^T X^T X((X^T X)^{-1} - \Omega)X^T X\hat{\boldsymbol{b}} = \hat{\boldsymbol{b}}_g^T D\hat{\boldsymbol{b}}_g \tag{39}$$

where $D = V_g^{-1} - V_g^{-1}(V_g^{-1} + W_g^{-1})^{-1}V_g^{-1}$. This allows the simplification

$$
\begin{aligned}
\hat{\boldsymbol{b}}_g^T V_g^{-1}\hat{\boldsymbol{b}}_g - \hat{\boldsymbol{b}}^T X^T X((X^T X)^{-1} - \Omega)X^T X\hat{\boldsymbol{b}} &= \hat{\boldsymbol{b}}_g^T V_g^{-1}(V_g^{-1} + W_g^{-1})^{-1}V_g^{-1}\hat{\boldsymbol{b}}_g \\
&= \frac{\hat{b}_g^2}{\delta^2}\frac{\phi^2}{\delta^2 + \phi^2}
\end{aligned}
\tag{40}
$$

When taking the limit, we also have

$$\bar{\boldsymbol{b}}^T\Phi^{-1}\bar{\boldsymbol{b}} = \frac{\bar{b}_g^2}{\phi^2} \tag{41}$$

Using the following block matrix inversion formula

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix} \tag{42}$$

for Ω, and taking the limit, we get

$$\begin{aligned}\bar{\boldsymbol{b}}^T\Phi^{-1}\Omega\Phi^{-1}\bar{\boldsymbol{b}} &= \frac{\bar{b}_g^2}{\phi^4}((X_g^TX_g + W_g^{-1}) - X_g^TX_c(X_c^TX_c)^{-1}X_c^TX_g)^{-1} \\ &= \frac{\bar{b}_g^2}{\phi^2}\frac{\delta^2}{\delta^2+\phi^2}\end{aligned} \tag{43}$$

Similarly

$$\bar{\boldsymbol{b}}^T\Phi^{-1}\Omega X^TX\hat{\boldsymbol{b}} = \frac{\bar{b}_g\hat{b}_g}{\delta^2+\phi^2} \tag{44}$$

In the end, the Bayes factor is

$$\mathrm{BF}(\bar{b},\phi) = \sqrt{\frac{\delta^2}{\delta^2+\phi^2}}\exp\left(\frac{1}{2}\left(\frac{\hat{b}^2}{\delta^2}\frac{\phi^2}{\delta^2+\phi^2} + \frac{\bar{b}^2}{\phi^2}\frac{\phi^2}{\delta^2+\phi^2} - 2\frac{\bar{b}\,\hat{b}}{\delta^2+\phi^2}\right)\right) \tag{45}$$

When τ is unknown, integrating it can't be achieved analytically. We hence use the Laplace's method to approximate the Bayes factor (33). Following Wen (2014), we introduce α as mixture proportion of the weighted sum of ...

TODO

4 Augmented likelihood

See the corresponding section in the document "config_model.tex/pdf".

Based on the hierarchical model described in the previous section, we can now write the augmented log-likelihood as follows,

$$l_a(\Theta;\boldsymbol{Y},\boldsymbol{z},\boldsymbol{V},\boldsymbol{T},\boldsymbol{D}|\boldsymbol{X}) = \sum_g \log p(\boldsymbol{Y}_g, z_g, \boldsymbol{v}_g, T_g, D_g|\boldsymbol{X}_g,\Theta) \tag{46}$$

Expanding the term inside the sum as done in the document "config_model.tex/pdf", we obtain:

$$\begin{aligned}l_a(\Theta;\boldsymbol{Y},\boldsymbol{z},\boldsymbol{V},\boldsymbol{T},\boldsymbol{D}|\boldsymbol{X}) = &\sum_g (1-z_g)\log\pi_0 + \sum_g z_g\log(1-\pi_0) + \sum_g \log p(Y_g|\boldsymbol{X}_g, z_g=0) \\ &+ \sum_{g,p} z_g v_{gp}\log\nu_p + \sum_{g,p,k} z_g v_{gp}t_{gpk}\log\pi_k + \sum_{g,p,l} z_g v_{gp}d_{gpl}\log\lambda_l \\ &+ \sum_{g,p,k,l} z_g v_{gp}t_{gpk}d_{gpl}\log\mathrm{BF}_{gpkl}\end{aligned} \tag{47}$$

where BF_{gpkl} is defined above (11).

5 EM algorithm

See the corresponding section in the document "config_model.tex/pdf".

The objective function, noted Q, for the EM algorithm, is:

$$Q(\Theta|\boldsymbol{Y},\boldsymbol{X},\Theta^{(i)}) = \mathbb{E}_{\boldsymbol{z},\boldsymbol{V},\boldsymbol{T},\boldsymbol{D}|\boldsymbol{Y},\boldsymbol{X},\Theta}[l_a(\Theta)|\boldsymbol{Y},\boldsymbol{X},\Theta^{(i)}] \tag{48}$$

Starting from randomly-initialized parameters $\Theta^{(0)}$, in the E-step for the $(i+1)^{\text{th}}$ iteration, we evaluate the objective function (48).

$$\mathbb{E}[z_g|\boldsymbol{Y},\boldsymbol{X},\Theta^{(i)}] = \frac{(1-\pi_0^{(i)})\text{BF}_g^{(i)}}{\pi_0^{(i)} + (1-\pi_0^{(i)})\text{BF}_g^{(i)}}. \tag{49}$$

where

$$\begin{aligned}\text{BF}_g^{(i)} &= \frac{p(Y_g|\boldsymbol{X}_g,\Theta^{(i)},z_g=1)}{p(Y_g|\boldsymbol{X}_g,z_g=0)} \\ &= \sum_{p,k,l}\nu_p^{(i)}\pi_k^{(i)}\lambda_l^{(i)}\text{BF}_{gpkl}(\boldsymbol{q}_k^{(i)},\phi_l,\omega_l)\end{aligned} \tag{50}$$

Similarly,

$$\mathbb{E}[z_g v_{gp}|\boldsymbol{Y},\boldsymbol{X},\Theta^{(i)}] = \frac{(1-\pi_0^{(i)})\nu_p^{(i)}\text{BF}_{gp}^{(i)}}{\pi_0^{(i)} + (1-\pi_0^{(i)})\text{BF}_g^{(i)}}, \tag{51}$$

where

$$\text{BF}_{gp}^{(i)} = \sum_{k,l}\pi_k^{(i)}\lambda_l^{(i)}\text{BF}_{gpkl}(\boldsymbol{q}_k^{(i)},\phi_l,\omega_l) \tag{52}$$

And

$$\mathbb{E}[z_g v_{gp} t_{gpk}|\boldsymbol{Y},\boldsymbol{X},\Theta^{(i)}] = \frac{(1-\pi_0^{(i)})\nu_p^{(i)}\pi_k^{(i)}\sum_l\lambda_l^{(i)}\text{BF}_{gpkl}(\boldsymbol{q}_k^{(i)},\phi_l,\omega_l)}{\pi_0^{(i)} + (1-\pi_0^{(i)})\text{BF}_g^{(i)}}, \tag{53}$$

$$\mathbb{E}[z_g v_{gp} d_{gpl}|\boldsymbol{Y},\boldsymbol{X},\Theta^{(i)}] = \frac{(1-\pi_0^{(i)})\nu_p^{(i)}\lambda_l^{(i)}\sum_k\pi_k^{(i)}\text{BF}_{gpkl}(\boldsymbol{q}_k^{(i)},\phi_l,\omega_l)}{\pi_0^{(i)} + (1-\pi_0^{(i)})\text{BF}_g^{(i)}}, \tag{54}$$

$$\mathbb{E}[z_g v_{gp} t_{gpk} d_{gpl}|\boldsymbol{Y},\boldsymbol{X},\Theta^{(i)}] = \frac{(1-\pi_0^{(i)})\nu_p^{(i)}\pi_k^{(i)}\lambda_l^{(i)}\text{BF}_{gpkl}(\boldsymbol{q}_k^{(i)},\phi_l,\omega_l)}{\pi_0^{(i)} + (1-\pi_0^{(i)})\text{BF}_g^{(i)}}. \tag{55}$$

In the M-step for the $(i+1)^{\text{th}}$ iteration, we estimate a new set of parameters, $\Theta^{(i+1)}$, by maximizing the objective function (48).

In particular, for π_0,

$$\frac{\partial Q}{\partial \pi_0}(\pi_0^{(i+1)}) = 0 \Leftrightarrow \pi_0^{(i+1)} = \frac{1}{G}\sum_g \frac{\pi_0^{(i)}}{\pi_0^{(i)} + (1-\pi_0^{(i)})\text{BF}_g^{(i)}}. \tag{56}$$

For the grid points, using a Lagrange multiplier, L_a, to enforce the constraint,

$$\frac{\partial Q}{\partial \lambda_l}(\lambda_l^{(i+1)}) = 0 \Leftrightarrow \lambda_l^{(i+1)} = \frac{\sum_{g,p,k}\frac{\nu_p^{(i)}\pi_k^{(i)}\text{BF}_{gpkl}(\boldsymbol{q}_k^{(i)},\phi_l,\omega_l)}{\pi_0^{(i)}+(1-\pi_0^{(i)})\text{BF}_g^{(i)}}\cdot\lambda_l^{(i)}}{\sum_{l'}\left(\sum_{g,p,k}\frac{\nu_p^{(i)}\pi_k^{(i)}\text{BF}_{gpkl'}(\boldsymbol{q}_k^{(i)},\phi_{l'},\omega_{l'})}{\pi_0^{(i)}+(1-\pi_0^{(i)})\text{BF}_g^{(i)}}\cdot\lambda_{l'}^{(i)}\right)} \tag{57}$$

Similarly, for the type proportions,

$$\pi_k^{(i+1)} = \frac{\sum_{g,p,l} \frac{\nu_p^{(i)} \lambda_l^{(i)} \mathrm{BF}_{gpkl}(\boldsymbol{q}_k^{(i)}, \phi_l, \omega_l)}{\pi_0^{(i)} + (1-\pi_0^{(i)}) \mathrm{BF}_g^{(i)}} \cdot \pi_k^{(i)}}{\sum_{k'} \left(\sum_{g,p,l} \frac{\nu_p^{(i)} \lambda_l^{(i)} \mathrm{BF}_{gpk'l}(\boldsymbol{q}_{k'}^{(i)}, \phi_l, \omega_l)}{\pi_0^{(i)} + (1-\pi_0^{(i)}) \mathrm{BF}_g^{(i)}} \cdot \pi_{k'}^{(i)} \right)}. \tag{58}$$

Finally, for the "tissues per type",

$$\frac{\partial Q}{\partial \boldsymbol{q}_k}(\boldsymbol{q}_k) = \sum_{g,p,l} \mathbb{E}[z_g v_{gp} t_{gpk} d_{gpl} | \boldsymbol{Y}, \boldsymbol{X}, \Theta^{(i)}] \frac{\partial \log \mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \phi_l, \omega_l)}{\partial \boldsymbol{q}_k} \tag{59}$$

5.1 Calculate the derivative of BF_{gpkl} w.r.t. q_k

If we follow William's idea, we need to derive (13) w.r.t. $\boldsymbol{q}_k$ using rules from matrix calculus.

$$\begin{aligned} \frac{\partial \log \mathrm{BF}_{gpkl}}{\partial \boldsymbol{q}_k} &= -\frac{1}{2} \mathrm{tr} \left[(I + V^{-1} W)^{-1} \frac{\partial V^{-1} W}{\partial \boldsymbol{q}_k} \right] \\ &\quad + \frac{1}{2} \dots \end{aligned} \tag{60}$$

TODO

5.2 Calculate the derivative of BF_{gpkl} w.r.t. q_{kr}

If we follow Matthew's idea, we can derive Q w.r.t q_{kr} instead of $\boldsymbol{q}_k$:

$$\frac{\partial Q}{\partial q_{kr}}(q_{kr}) = \sum_{g,p,l} \mathbb{E}[z_g v_{gp} t_{gpk} d_{gpl} | \boldsymbol{Y}, \boldsymbol{X}, \Theta^{(i)}] \frac{\partial \log \mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \phi_l, \omega_l)}{\partial q_{kr}} \tag{61}$$

where

$$\frac{\partial \log \mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \phi_l, \omega_l)}{\partial q_{kr}} = \frac{1}{\mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \phi_l, \omega_l)} \frac{\partial \mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \phi_l, \omega_l)}{\partial q_{kr}} \tag{62}$$

Using Leibniz's rule:

$$\frac{\partial \log \mathrm{BF}_{gpkl}(\boldsymbol{q}_k, \phi_l, \omega_l)}{\partial q_{kr}} = \frac{1}{\int f(\bar{b}_{gpl}, \boldsymbol{q}_k, \phi_l, \omega_l) \, \mathrm{d}\bar{b}_{gpl}} \int \frac{\partial f(\bar{b}_{gpl}, \boldsymbol{q}_k, \phi_l, \omega_l)}{\partial q_{kr}} \, \mathrm{d}\bar{b}_{gpl} \tag{63}$$

However, what should we do with the $q_{kr'}$'s where $r' \neq r$?

TODO

6 Posteriors on latent variables

See the corresponding section in the document "config_model.tex/pdf".

Moreover, we may still want to assess the magnitude of eQTL sharing between tissues in terms of configurations. We can do that as in the config model by obtaining point estimates of each configuration probability:

$$\hat{\eta}_j = p(\boldsymbol{\gamma}_j|\hat{\boldsymbol{\pi}}, \hat{\boldsymbol{q}}) = \sum_{k=1}^{K} \hat{\pi}_k \left[\prod_{r=1}^{R} \hat{q}_{kr}^{\gamma_r} (1 - \hat{q}_{kr})^{1-\gamma_r} \right] \tag{64}$$

7 Posteriors on genotype effect sizes

See the corresponding section in the document "config_model.tex/pdf".

8 Prior on types

In order to simulate data, we may find it useful to have a prior on types. Similarly to the F model in population genetics (Balding and others), we can assume that all types are related to an "ancestral" type via a star tree which edges differ. By changing the length of the edges, we can make the types more or less different from each other. Tim has some R code to compare "config" and "type" models on simulated data.