Appendix S1- Showing equivalence between a test based on an animal centric model and a test based on a SNP effects fixed model

Consider the animal-centric model given by

$$\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{e} \tag{I}$$

where $\boldsymbol{a} \sim N(\mathbf{0}, \boldsymbol{ZZ'}\sigma_u^2)$ and $\boldsymbol{e} \sim N(\mathbf{0}, \boldsymbol{I}\sigma_e^2)$. In this case, solutions for animal effects are given by $\hat{\boldsymbol{a}} = \left(\boldsymbol{I}/\sigma_e^2 + (\boldsymbol{ZZ'})^{-1}/\sigma_u^2 \right)^{-1} \boldsymbol{y}/\sigma_e^2$. According to Strandén and Garrick (2009), Badke *et al*. (2014) and Gualdrón Duarte *et al*. (2014), model (I) is equivalent to the animal-centric model given by

$$\boldsymbol{y} = \boldsymbol{Zg} + \boldsymbol{e} \tag{II}$$

with $\boldsymbol{g} \sim N(\mathbf{0}, \boldsymbol{I}\sigma_u^2)$, $\boldsymbol{e} \sim N(0, \boldsymbol{I}\sigma_e^2)$, and where SNP effects can be estimated as $\hat{\boldsymbol{g}} = \left(\boldsymbol{ZZ'}/\sigma_e^2 + \boldsymbol{I}/\sigma_u^2 \right)^{-1} \boldsymbol{Z'y}/\sigma_e^2$. Taking into account model (I), SNP effects can be expressed as a linear transformation of breeding values, that is, $\hat{\boldsymbol{g}} = \boldsymbol{Z'}(\boldsymbol{ZZ'})^{-1}\hat{\boldsymbol{a}}$.

Also, prediction error variance of $\hat{\boldsymbol{g}}$ is defined as $\mathrm{Var}(\hat{\boldsymbol{g}} - \boldsymbol{g}) = \left(\boldsymbol{ZZ'}/\sigma_e^2 + \boldsymbol{I}/\sigma_u^2 \right)^{-1}$, and $\mathrm{Var}(\hat{\boldsymbol{g}}) = \boldsymbol{I}\sigma_u^2 - \left(\boldsymbol{ZZ'}/\sigma_e^2 + \boldsymbol{I}/\sigma_u^2 \right)^{-1}$. In this context, a *t*-statistic can be defined for *i*-th SNP effect in (II), that is:

$$tg_i = \frac{\hat{g}_i}{\sqrt{\mathrm{Var}(\hat{g}_i)}} \tag{III}$$

Now, consider the efficient model association eXpedited EMMAX (Kang *et al*. 2010; Zhang *et al*. 2010) given by:

$$\boldsymbol{y} = \boldsymbol{z}_i b_i + \boldsymbol{a} + \boldsymbol{e} \tag{IV}$$

with $\boldsymbol{a} \sim N\,(\mathbf{0}, \boldsymbol{ZZ'}\sigma_u^2)$ and $\boldsymbol{e} \sim N\,(\mathbf{0}, \boldsymbol{I}\sigma_e^2)$. Equivalently, model (IV) can be expressed as $\boldsymbol{y} = \boldsymbol{z}_i b_i + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim N\,(\mathbf{0},\ \boldsymbol{V})$ and $\boldsymbol{V} = \boldsymbol{ZZ'}\sigma_u^2 + \boldsymbol{I}\sigma_e^2$. For this model, solution of i-th SNP effect is given by

$$\hat{b}_i = \left(z_i'\boldsymbol{V}^{-1}z_i\right)^{-1} z_i'\boldsymbol{V}^{-1}y = \frac{z_i'\boldsymbol{V}^{-1}y}{z_i'\boldsymbol{V}^{-1}z_i} \qquad \text{(V)}$$

With $\mathrm{Var}\left(\hat{b}_i\right) = \left(z_i'\boldsymbol{V}^{-1}z_i\right)^{-1}$. In this case, the t-statistic derived from (V) is equal to

$tb_i = \dfrac{\hat{b}_i}{\sqrt{Var\left(\hat{b}_i\right)}}$. With which, in order to demonstrate that derived test tg_i is a computationally tractable solution to model in (IV), it has been showed that $tg_i = tb_i$. In this sense, two propositions will be verified:

1. $\mathrm{Var}\left(\hat{g}_i\right) = \dfrac{\left(\sigma_u^2\right)^2}{\mathrm{Var}\left(\hat{b}_i\right)}$

2. $\hat{g}_i = \dfrac{\sigma_u^2}{\mathrm{Var}\left(\hat{b}_i\right)}\hat{b}_i$

Proof of proposition 1.

Taking into account that $\mathrm{Var}\left(\hat{g}_i\right)$ are the diagonal elements of

$\boldsymbol{I}\sigma_u^2 - \left(\boldsymbol{ZZ'}/\sigma_e^2 + \boldsymbol{I}/\sigma_u^2\right)^{-1}$, the Woodbury-Morrison formula can be applied to the second term to simplify this expression. In general, the Woodbury-Morrison formula establishes that

$$\left(\boldsymbol{A} + \boldsymbol{UCU'}\right)^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{U}\left(\boldsymbol{C}^{-1} + \boldsymbol{U'A}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{U'A}^{-1} \qquad \text{(VI)}$$

In order to apply it to the previous identity, let $\boldsymbol{A} = \boldsymbol{I}/\sigma_u^2$, $\boldsymbol{U} = \boldsymbol{Z}'$, $\boldsymbol{U}' = \boldsymbol{Z}$ and $\boldsymbol{C} = \boldsymbol{I}\sigma_e^2$. Thus,

$$\left(\boldsymbol{ZZ}'/\sigma_e^2 + \boldsymbol{I}/\sigma_u^2 \right)^{-1} = \boldsymbol{I}\sigma_u^2 - \sigma_u^2\boldsymbol{Z}'\left(\boldsymbol{I}\sigma_e^2 + \boldsymbol{ZZ}'\sigma_u^2\right)^{-1}\boldsymbol{Z}\sigma_u^2 \qquad \text{(VII)}$$

Notice that $\boldsymbol{V} = \boldsymbol{ZZ}'\sigma_u^2 + \boldsymbol{I}\sigma_e^2$ is the central term on the right of expression (VII), so that

$$\left(\boldsymbol{ZZ}'/\sigma_e^2 + \boldsymbol{I}/\sigma_u^2 \right)^{-1} = \boldsymbol{I}\sigma_u^2 - \left(\sigma_u^2\right)^2 \boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z} \qquad \text{(VIII)}$$

Thus, the variance-covariance matrix of the estimated SNP effects under the SNP centric-model is equal to

$$\begin{aligned} \boldsymbol{I}\sigma_u^2 - \left(\boldsymbol{ZZ}'/\sigma_e^2 + \boldsymbol{I}/\sigma_u^2 \right)^{-1} &= \boldsymbol{I}\sigma_u^2 - \boldsymbol{I}\sigma_u^2 + \left(\sigma_u^2\right)^2 \boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z} \\ &= \left(\sigma_u^2\right)^2 \boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z} \end{aligned} \qquad \text{(IX)}$$

On noting that the diagonal element of (IX) is $\left(\sigma_u^2\right)^2 z_i'\boldsymbol{V}^{-1}z_i$, which is equal to $\dfrac{\left(\sigma_u^2\right)^2}{\mathrm{Var}\left(\hat{b}_i\right)}$, proposition 1 has been verified.

Proof of proposition 2.

Now we have that

$$
\begin{aligned}
\frac{\sigma_u^2}{\mathrm{Var}\left(\hat{b}_i\right)}\hat{b}_i &= \sigma_u^2 \frac{\dfrac{z_i'\boldsymbol{V}^{-1}y}{z_i'\boldsymbol{V}^{-1}z_i}}{\dfrac{1}{z_i'\boldsymbol{V}^{-1}z_i}} \\
&= \sigma_u^2 z_i'\boldsymbol{V}^{-1}\boldsymbol{y} \\
&= \sigma_u^2 z_i'\left(\boldsymbol{ZZ}'\sigma_u^2 + \boldsymbol{I}\sigma_e^2\right)\boldsymbol{y} \\
&= z_i'\left(\boldsymbol{ZZ}' + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_u^2}\right)^{-1}\boldsymbol{y}
\end{aligned}
\qquad \text{(X)}
$$

On multiplying and dividing by a matrix (***G*** = ***ZZ***’) and its inverse

$$
\begin{aligned}
\frac{\sigma_u^2}{\mathrm{Var}\left(\hat{b}_i\right)}\hat{b}_i &= z_i'\boldsymbol{G}^{-1}\boldsymbol{G}\left(\boldsymbol{ZZ}' + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_u^2}\right)^{-1}\boldsymbol{y} \\
&= z_i'\boldsymbol{G}^{-1}\left(\boldsymbol{G}^{-1}\right)^{-1}\left(\boldsymbol{ZZ}' + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_u^2}\right)^{-1}\boldsymbol{y}
\end{aligned}
\qquad \text{(XI)}
$$

And after applying the rule of the inverse of a product, we have:

$$
\frac{\sigma_u^2}{\mathrm{Var}\left(\hat{b}_i\right)}\hat{b}_i = z_i'\boldsymbol{G}^{-1}\left(\boldsymbol{G}^{-1}\right)^{-1}\left(\left(\boldsymbol{ZZ}' + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_u^2}\right)\boldsymbol{G}^{-1}\right)^{-1}\boldsymbol{y} \qquad \text{(XII)}
$$

Distribute now the product as follows

$$
\begin{aligned}
\frac{\sigma_u^2}{\mathrm{Var}\left(\hat{b}_i\right)}\hat{b}_i &= z_i'\boldsymbol{G}^{-1}\left(\left(\boldsymbol{ZZ}'\boldsymbol{G}^{-1} + \boldsymbol{G}^{-1}\frac{\sigma_e^2}{\sigma_u^2}\right)\boldsymbol{G}^{-1}\right)^{-1}\boldsymbol{y} \\
&= z_i'\boldsymbol{G}^{-1}\left(\boldsymbol{I} + \boldsymbol{G}^{-1}\frac{\sigma_e^2}{\sigma_u^2}\right)^{-1}\boldsymbol{y}
\end{aligned}
\qquad \text{(XIII)}
$$

However, notice that

$$
\begin{aligned}
\left(\boldsymbol{I} + \boldsymbol{G}^{-1}\frac{\sigma_e^2}{\sigma_u^2}\right)^{-1}\boldsymbol{y} &= \left(\boldsymbol{I}\big/\sigma_e^2 + \left(\boldsymbol{ZZ}'\right)^{-1}\big/\sigma_u^2\right)^{-1}\boldsymbol{y}\big/\sigma_e^2 \\
&= \hat{\boldsymbol{a}}
\end{aligned}
\qquad \text{(XIV)}
$$

And thus:

$$\frac{\sigma_u^2}{\mathrm{Var}\left(\hat{b}_i\right)}\hat{b}_i = z_i'\boldsymbol{G}^{-1}\hat{\boldsymbol{a}} \tag{XV}$$

Which is an indirect way of deriving $\hat{g}_i$ (Strandén & Garrick 2009; Hayes & Goddard 2010; Wang *et al*. 2012; Wang *et al*. 2014; Gualdrón Duarte *et al*. 2014) as we wanted to prove. Finally, equivalence between tg_i and tb_i follows from propositions 1) and 2):

$$tg_i = \frac{\hat{g}_i}{\sqrt{Var\left(\hat{g}_i\right)}} = \frac{\frac{\sigma_u^2}{\mathrm{Var}\left(\hat{b}_i\right)}\hat{b}_i}{\sqrt{\frac{\left(\sigma_u^2\right)^2}{\mathrm{Var}\left(\hat{b}_i\right)}}} = \frac{\hat{b}_i}{\sqrt{\mathrm{Var}\left(\hat{b}_i\right)}} = tb_i \tag{XVI}$$

This demonstrates that tg_i is actually a computationally tractable solution to SNP effects fixed model as EMMAX. The result presented in (XVI) still holds true even other fixed effects are included. Thus, the analytical solution developed dissipates the possibility of “agreement by chance” between the results of both tests.