Supplementary Notes

In the main text of the manuscript, we have provided a proof of the equivalence between "GWAS by GBLUP" and the "P3D" approximated Q+K model in the case that each individual has only one phenotypic record. Here in this note, we generalized the proof to the case that some individuals may have multiple records (e.g. data from a multi-environment trial in which some individuals were tested in more than one environment). We also generalize the proof from the single SNP-based test to the window-based test in which the additive effects for a group of SNPs are tested together. In fact, the generalization is straightforward and hence it largely parallels the proof in the main text. In the proof, we will use the two lemmas in the Appendix of the manuscript, which we do not repeat here.

Throughout the note, we assume that n is the number of individuals, m is the number of phenotypic records ($m \geq n$), k is the number of covariates, p is the total number of markers and s is the number of SNPs in the window-based test. When $s = 1$, it simplifies to the single SNP-based test.

1 GWAS by the Q+K model with the P3D approaximation

The model has the following form:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}(\boldsymbol{W}\boldsymbol{a}_w + \boldsymbol{g}) + \boldsymbol{e}. \tag{1}$$

The notations are the following: $\boldsymbol{y}$ is the m-dimensional vector of phenotypic records. $\boldsymbol{\beta}$ is the k-dimensional vector of covariate effects. $\boldsymbol{X}$ is the corresponding $m \times k$ design matrix. $\boldsymbol{Z}$ is the $m \times n$ dimensional design matrix allocating the phenotypic records to each individuals. $\boldsymbol{a}_w$ is the vector of additive effect of s markers in the window being tested and $\boldsymbol{W}$ is the corresponding $n \times s$ dimensional matrix of marker profiles. $\boldsymbol{g}$ denotes the n-dimensional vector of polygenic background effects. $\boldsymbol{e}$ is the residual term. In the model, $\boldsymbol{\beta}$ and $\boldsymbol{a}_w$ are assumed to be fixed parameters and $\boldsymbol{g} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{G}\sigma_g^2)$, where $\boldsymbol{G} = \boldsymbol{M}\boldsymbol{M}'/c$ is a genomic relationship matrix, $\boldsymbol{M}$ is the $n \times p$ dimensional matrix of all marker profiles and c is a scaling factor (e.g. the VanRaden G-matrix [2]). The remaining assumptions are $\boldsymbol{e} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}\sigma_e^2)$ and $\mathrm{Cov}(\boldsymbol{e}, \boldsymbol{g}) = 0$.

The model can be rewritten as follows:

$$\boldsymbol{y} = \tilde{\boldsymbol{X}}\tilde{\boldsymbol{\beta}} + \boldsymbol{Z}\boldsymbol{g} + \boldsymbol{e}, \tag{2}$$

where $\tilde{X} = (X|ZW)$, $\tilde{\beta} = (\beta', a_w')'$.

From Henderson's mixed model equations [1], we know that the best linear unbiased estimation of the fixed effects are the following:

$$\hat{\tilde{\beta}} = (\tilde{X}'V^{-1}\tilde{X})^{-1}\tilde{X}'V^{-1}y, \quad \text{Var}(\hat{\tilde{\beta}}) = \tilde{C}_{11}\sigma_e^2, \tag{3}$$

where $V = I + \lambda ZGZ'$, $\lambda = \sigma_g^2/\sigma_e^2$ and $\tilde{C}_{11}$ is defined via the following:

$$\begin{pmatrix} \tilde{X}'\tilde{X} & \tilde{X}'Z \\ Z'\tilde{X} & Z'Z + \lambda^{-1}G^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} \tilde{C}_{11} & \tilde{C}_{12} \\ \tilde{C}_{12}' & \tilde{C}_{22} \end{pmatrix} \tag{4}$$

Using Lemma A.1, we can calculate that

$$\begin{aligned} \tilde{C}_{11} &= (\tilde{X}'\tilde{X} - \tilde{X}'Z(Z'Z + \lambda^{-1}G^{-1})^{-1}Z'\tilde{X})^{-1} \\ &= (\tilde{X}'(I - Z(Z'Z + \lambda^{-1}G^{-1})^{-1}Z')\tilde{X})^{-1} \\ &= (\tilde{X}'V^{-1}\tilde{X})^{-1}. \end{aligned} \tag{5}$$

Using (5), the mixed model solution (3) can be written as follows:

$$\begin{aligned} \begin{pmatrix} \hat{\beta} \\ \hat{a}_w \end{pmatrix} &= \begin{pmatrix} X'V^{-1}X & X'V^{-1}ZW \\ W'Z'V^{-1}X & W'Z'V^{-1}ZW \end{pmatrix}^{-1} \begin{pmatrix} X'V^{-1}y \\ W'Z'V^{-1}y \end{pmatrix} \\ \text{Var}\begin{pmatrix} \hat{\beta} \\ \hat{a}_w \end{pmatrix} &= \begin{pmatrix} X'V^{-1}X & X'V^{-1}ZW \\ W'Z'V^{-1}X & W'Z'V^{-1}ZW \end{pmatrix}^{-1} \sigma_e^2. \end{aligned} \tag{6}$$

Using (6) and Lemma A.1, we can resolve $\hat{a}_w$ and $\text{Var}(\hat{a}_w)$:

$$\begin{aligned} \hat{a}_w &= (W'Z'V^{-1}ZW - W'Z'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}ZW)^{-1}W'Z'V^{-1}y \\ &\quad - (W'Z'V^{-1}ZW - W'Z'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}ZW)^{-1}\cdot \\ &\quad\quad W'Z'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}y \\ &= (W'Z'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})ZW)^{-1}\cdot \\ &\quad\quad W'Z'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})y \\ &= (W'Z'TZW)^{-1}W'Z'Ty, \\ \text{Var}(\hat{a}_w) &= (W'Z'V^{-1}ZW - W'Z'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}ZW)^{-1}\sigma_e^2 \\ &= (W'Z'(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})ZW)^{-1}\sigma_e^2 \\ &= (W'Z'TZW)^{-1}\sigma_e^2, \end{aligned}$$

where $T = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}$.

2

Thus, to test the hypothesis $H_0 : \boldsymbol{a}_w = 0$, we can use the following Wald statistic:

$$\begin{aligned}
\mathcal{W}_{Q+K} &= \hat{\boldsymbol{a}}'_w \operatorname{Var}(\hat{\boldsymbol{a}}_w)^{-1} \hat{\boldsymbol{a}}_w \\
&= (\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{y})'(\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{Z}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{y}\sigma_e^{-2} \\
&= \boldsymbol{y}'\boldsymbol{T}\boldsymbol{Z}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{Z}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{y}\sigma_e^{-2}
\end{aligned} \tag{7}$$

2 GWAS by GBLUP

The GBLUP model has the following form:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{g} + \boldsymbol{e}, \tag{8}$$

where all notations are the same as in (1). It is equivalent to the following RR-BLUP model:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{M}\boldsymbol{a} + \boldsymbol{e}, \tag{9}$$

where $\boldsymbol{a}$ is the p-dimensional vector of additive effects of all markers, $\boldsymbol{a} \sim \mathcal{N}(0, \boldsymbol{I}\sigma_a^2)$ and $\boldsymbol{M}$ is an $n \times p$ matrix of marker profiles. For the equivalence between (9) and (8), it is required that $\sigma_a^2 = \sigma_g^2/c$.

According to Henderson [1], the best linear unbiased prediction of random effects $\boldsymbol{a}$ and its variance for the model (9) is the following:

$$\begin{aligned}
\hat{\boldsymbol{a}} &= \rho \boldsymbol{M}'\boldsymbol{Z}'\tilde{\boldsymbol{V}}^{-1}(\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}'\tilde{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\tilde{\boldsymbol{V}}^{-1}\boldsymbol{y}) \\
\operatorname{Var}(\hat{\boldsymbol{a}}) &= (\rho\boldsymbol{I} - \boldsymbol{C}_{22})\sigma_e^2
\end{aligned} \tag{10}$$

where $\rho = \sigma_a^2/\sigma_e^2$, $\boldsymbol{C}_{22}$ is the defined as follows:

$$\begin{pmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{Z}\boldsymbol{M} \\ \boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{X} & \boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{M} + \rho^{-1}\boldsymbol{I} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}'_{12} & \boldsymbol{C}_{22} \end{pmatrix} \tag{11}$$

and $\tilde{\boldsymbol{V}} = \boldsymbol{I} + \rho\boldsymbol{Z}\boldsymbol{M}\boldsymbol{M}'\boldsymbol{Z}'$.

In fact, $\tilde{\boldsymbol{V}}$ is the same as $\boldsymbol{V}$ defined in section 1. Due to the equivalence between (8) and (9), we have $\sigma_a^2 = \sigma_g^2/c$. Hence, we know that $\rho = \sigma_g^2/c\sigma_e^2 = \lambda/c$. Recalling that $\boldsymbol{G} = \boldsymbol{M}\boldsymbol{M}'/c$, we have

$$\tilde{\boldsymbol{V}} = \boldsymbol{I} + \rho\boldsymbol{Z}\boldsymbol{M}\boldsymbol{M}'\boldsymbol{Z}' = \boldsymbol{I} + \lambda\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}' = \boldsymbol{V}.$$

Replacing $\tilde{\boldsymbol{V}}$ by $\boldsymbol{V}$ in (10), we have

$$\begin{aligned}
\hat{\boldsymbol{a}} &= \rho\boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}) \\
&= \rho\boldsymbol{M}'\boldsymbol{Z}'(\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1})\boldsymbol{y} \\
&= \rho\boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{y}.
\end{aligned} \tag{12}$$

where $\boldsymbol{T} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}$ as in the last subsection.

Using Lemma A.1, we can derive $\boldsymbol{C}_{22}$ as follows:

$$
\begin{aligned}
\boldsymbol{C}_{22} &= (\boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{Z}\boldsymbol{M} + \rho^{-1}\boldsymbol{I} - \boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}\boldsymbol{M})^{-1} \\
&= (\rho^{-1}\boldsymbol{I} + \boldsymbol{M}'\boldsymbol{Z}'(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{Z}\boldsymbol{M})^{-1} \\
&= (\frac{\sigma_e^2}{\sigma_a^2}\boldsymbol{I} + \boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{S}\boldsymbol{Z}\boldsymbol{M})^{-1}
\end{aligned}
\tag{13}
$$

where $\boldsymbol{S} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$.

Thus, replacing $\boldsymbol{C}_{22}$ by (13) in (10), we have

$$
\mathrm{Var}(\hat{\boldsymbol{a}}) = \sigma_a^2\boldsymbol{I} - (\frac{\sigma_e^2}{\sigma_a^2}\boldsymbol{I} + \boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{S}\boldsymbol{Z}\boldsymbol{M})^{-1}\sigma_e^2
\tag{14}
$$

Using (12), we can easily extract the estimation of marker effects within the window being tested:

$$
\hat{\boldsymbol{a}}^w = \rho\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{y},
\tag{15}
$$

where W is the submatrix of M consisting of the columns corresponding to the markers within the testing window. Note that in order to distinguish from the marker effects estimated in the Q+K model, which has been denoted by $\hat{\boldsymbol{a}}_w$, we used $\hat{\boldsymbol{a}}^w$ to denote the marker effects estimated by the GBLUP model.

On the other hand, it is not straightforward to derive an explicit form of $\mathrm{Var}(\hat{\boldsymbol{a}}^w)$ from (14). But we know that $\mathrm{Var}(\hat{\boldsymbol{a}})$ is an $p \times p$ matrix and $\mathrm{Var}(\hat{\boldsymbol{a}}^w)$ is an $s \times s$ submatrix of $\mathrm{Var}(\hat{\boldsymbol{a}})$ corresponding to the markers within the testing window. So we may denote it by $\mathrm{Var}(\hat{\boldsymbol{a}})_{w,w}$.

Then, to test the hypothesis $H_0 : \boldsymbol{a}^w = 0$, we can use the following Wald statistic:

$$
\begin{aligned}
\mathcal{W}_{GBLUP} &= (\hat{\boldsymbol{a}}^w)' \mathrm{Var}(\hat{\boldsymbol{a}})_{w,w}^{-1}\hat{\boldsymbol{a}}^w \\
&= \rho^2\boldsymbol{y}'\boldsymbol{T}\boldsymbol{Z}\boldsymbol{W} \mathrm{Var}(\hat{\boldsymbol{a}})_{w,w}^{-1}\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{y}
\end{aligned}
\tag{16}
$$

3 The equivalence between the two GWAS approaches

We need to prove $\mathcal{W}_{Q+K} = \mathcal{W}_{GBLUP}$. Comparing (7) and (16), it is necessary to prove the following:

$$
\mathrm{Var}(\hat{\boldsymbol{a}})_{w,w} = \rho^2\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{Z}\boldsymbol{W}\sigma_e^2.
\tag{17}
$$

As $\rho = \sigma_a^2/\sigma_e^2$, it is sufficient to prove

$$
\mathrm{Var}(\hat{\boldsymbol{a}})_{w,w} = \frac{\sigma_a^4}{\sigma_e^2}\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{Z}\boldsymbol{W}.
\tag{18}
$$

To achieve our goal, we need the singular value decomposition (SVD) of the matrix $\boldsymbol{X}$. Assume that the SVD of $\boldsymbol{X}$ is $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{W}$. In the decomposition, $\boldsymbol{U}$ is an $m \times m$ orthogonal

matrix, $\mathbf{\Sigma} = (\boldsymbol{D}\ \boldsymbol{0}_{k\times(m-k)})'$, where $\boldsymbol{D}$ is an $k \times k$ diagonal matrix whose diagonal entries are the singular values of $\boldsymbol{X}$ and $\boldsymbol{0}_{k\times(m-k)}$ is a $k \times (m-k)$ matrix of zeros, $\boldsymbol{W}$ is a $k \times k$ orthogonal matrix. We can write $\boldsymbol{U} = (\boldsymbol{U}_1\ \boldsymbol{U}_2)$, where $\boldsymbol{U}_1$ is the left $m \times k$ block and $\boldsymbol{U}_2$ is the right $m \times (m-k)$ block of $\boldsymbol{U}$. Then we have:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{W} = (\boldsymbol{U}_1\ \boldsymbol{U}_2)\begin{pmatrix} \boldsymbol{D} \\ \boldsymbol{0}_{k\times(m-k)} \end{pmatrix}\boldsymbol{W} = \boldsymbol{U}_1\boldsymbol{D}\boldsymbol{W}. \tag{19}$$

The orthogonality of $\boldsymbol{U}$ gives the following:

$$\begin{aligned} \boldsymbol{U}_2'\boldsymbol{U}_1 &= \boldsymbol{0}_{(m-k)\times k}, \quad \boldsymbol{U}_1'\boldsymbol{U}_2 = \boldsymbol{0}_{k\times(m-k)}, \\ \boldsymbol{U}_1'\boldsymbol{U}_1 &= \boldsymbol{I}_k, \quad \boldsymbol{U}_2'\boldsymbol{U}_2 = \boldsymbol{I}_{m-k}, \quad \boldsymbol{U}_1\boldsymbol{U}_1' + \boldsymbol{U}_2\boldsymbol{U}_2' = \boldsymbol{I}_m. \end{aligned} \tag{20}$$

Using (19) and (20), we have

$$\begin{aligned} \boldsymbol{S} &= \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \boldsymbol{I} - \boldsymbol{U}_1\boldsymbol{D}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{D}\boldsymbol{U}_1'\boldsymbol{U}_1\boldsymbol{D}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{D}\boldsymbol{U}_1' \\ &= \boldsymbol{I} - \boldsymbol{U}_1\boldsymbol{D}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{D}^{-2}\boldsymbol{W})\boldsymbol{W}'\boldsymbol{D}\boldsymbol{U}_1' \\ &= \boldsymbol{I} - \boldsymbol{U}_1\boldsymbol{U}_1' = \boldsymbol{U}_2\boldsymbol{U}_2' \end{aligned} \tag{21}$$

Replacing $\boldsymbol{S}$ in (14) and using Lemma A.1, we can simplify the formula for $\mathrm{Var}(\hat{\boldsymbol{a}})$ as follows:

$$\begin{aligned} \mathrm{Var}(\hat{\boldsymbol{a}}) &= \sigma_a^2\boldsymbol{I} - (\frac{\sigma_e^2}{\sigma_a^2}\boldsymbol{I} + \boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{U}_2\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{M})^{-1}\sigma_e^2 \\ &= \sigma_a^2\boldsymbol{I} - (\rho\boldsymbol{I} - \rho^2\boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{U}_2(\boldsymbol{I} + \rho\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{M}\boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{M})\sigma_e^2 \\ &= \frac{\sigma_a^4}{\sigma_e^2}\boldsymbol{M}'\boldsymbol{Z}'\boldsymbol{U}_2(\boldsymbol{I} + \lambda\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}'\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{M}, \end{aligned} \tag{22}$$

where in the last equality we used $\rho = \lambda/c$ and $\boldsymbol{G} = \boldsymbol{M}\boldsymbol{M}'/c$. With the above formula, we now have a direct formula for $\mathrm{Var}(\hat{\boldsymbol{a}})_{w,w}$:

$$\mathrm{Var}(\hat{\boldsymbol{a}})_{w,w} = \frac{\sigma_a^4}{\sigma_e^2}\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{U}_2(\boldsymbol{I} + \lambda\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}'\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{W} \tag{23}$$

Now, replacing $\mathrm{Var}(\hat{\boldsymbol{a}})_{w,w}$ in the left hand side of (18) by (22), we only need to prove

$$\boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{T}\boldsymbol{Z}\boldsymbol{W} = \boldsymbol{W}'\boldsymbol{Z}'\boldsymbol{U}_2(\boldsymbol{I} + \lambda\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}'\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{W} \tag{24}$$

Using (19) and (20), we can calculate that

$$\begin{aligned} \boldsymbol{T} &= \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1} \\ &= \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{U}_1\boldsymbol{D}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{D}\boldsymbol{U}_1'\boldsymbol{V}^{-1}\boldsymbol{U}_1\boldsymbol{D}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{D}\boldsymbol{U}_1'\boldsymbol{V}^{-1} \\ &= \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{U}_1\boldsymbol{D}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{D}^{-1}(\boldsymbol{U}_1'\boldsymbol{V}^{-1}\boldsymbol{U}_1)^{-1}\boldsymbol{D}^{-1}\boldsymbol{W})\boldsymbol{W}'\boldsymbol{D}\boldsymbol{U}_1'\boldsymbol{V}^{-1} \\ &= \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{U}_1(\boldsymbol{U}_1'\boldsymbol{V}^{-1}\boldsymbol{U}_1)^{-1}\boldsymbol{U}_1'\boldsymbol{V}^{-1} \end{aligned} \tag{25}$$

5

Note that $\boldsymbol{V}^{-1}$ is positive-definite matrix, $\boldsymbol{U}_1'\boldsymbol{U}_2 = 0$. Thus, we can apply Lemma A.2 to the above formula, yielding

$$\begin{aligned}
\boldsymbol{T} &= \boldsymbol{U}_2(\boldsymbol{U}_2'\boldsymbol{V}\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2' \\
&= \boldsymbol{U}_2(\boldsymbol{U}_2'(\boldsymbol{I} + \lambda\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}')\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2' \\
&= \boldsymbol{U}_2(\boldsymbol{I} + \lambda\boldsymbol{U}_2'\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}'\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2',
\end{aligned} \tag{26}$$

which completes the proof.

References

[1] Charles R Henderson. Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31(2):423–447, 1975.

[2] Paul M VanRaden. Efficient methods to compute genomic predictions. *Journal of dairy science*, 91(11):4414–4423, 2008.