Accuracy of genomic prediction by

Singular Value Decomposition of the Genotype Matrix

L. Ayres* M.P.L. Calus* J. Ødegård[‡] T. Meuwissen[‡]

Aim

Evaluate the accuracy of prediction when varying the number of components in the truncated-SVD approximation of the genotype matrix.

Results 0.8 0.4 0.2 0.2 0.0 Components

Simulation

- 1,000 QTL along 16,454 SNPs
 (1,927 Atlantic Salmon individuals)
- QTL effects sampled from normal distribution $b_1, ..., b_{1000}$ i.i.d. $b_i \sim N(0, V_m)$

$$V_m = \frac{h^2}{2\sum_{i=1}^{1000} p_i (1 - p_i)}$$

- QTLs masked from genotype matrix X
- True breeding values g calculated

$$g = Xb$$

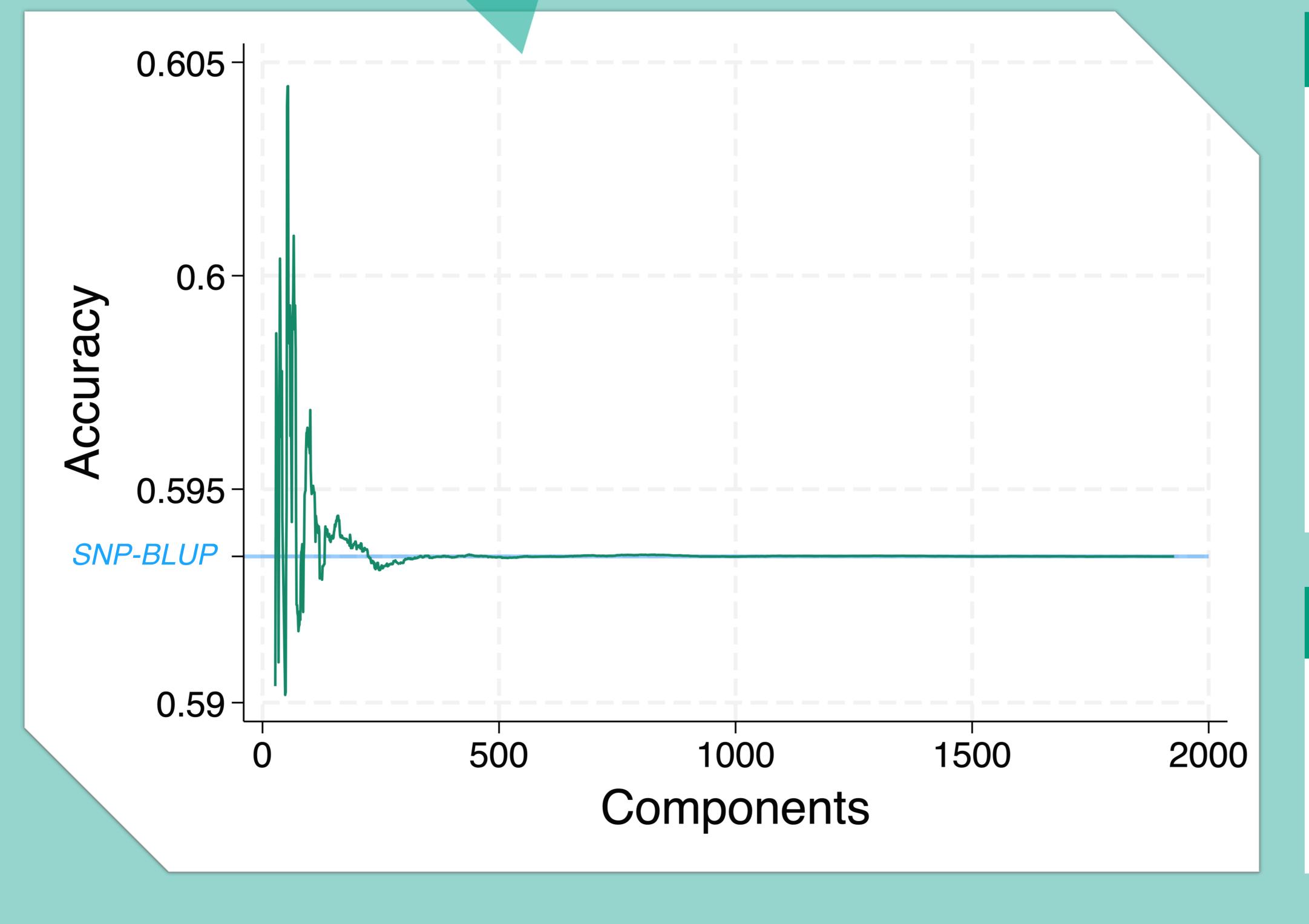
Breeding values predicted by PCRR

$$\hat{\mathbf{g}} = \mathbf{X}\hat{\mathbf{b}} = \mathbf{X}\mathbf{V}_k\hat{\mathbf{s}}_k$$

(10-fold cross-validation)

Correlation coefficients estimated

$$\hat{r} = \frac{Cov(g, \hat{g})}{\hat{\sigma}_a \hat{\sigma}_{\hat{g}}}$$



Conclusions

- SVD is useful for data reduction of genotype matrices
- Good accuracies with few components
- SVD can yield higher accuracies than SNP-BLUP

QR code for supplement













^{*} Wageningen University & Research, Animal Breeding and Genomics, Droevendaalsesteeg 1, 6700 AH Wageningen, the Netherlands

[‡] Norwegian University of Life Sciences, Department of Animal and Aquacultural Sciences, Oluf Thesens vei 6, 1433 Ås, Norway

[♦] AquaGen AS, Postboks 1240, Torgard, 7462 Trondheim, Norway