

Supplement S1 to “Simple penalties on maximum likelihood estimates of genetic parameters to reduce sampling variation”

1 Distribution of reduction in loss due to penalties on ordinary correlations

Figure S1 summarizes the distribution of PRIAL obtained by imposing penalties on ordinary genetic correlations. As for penalties on corresponding partial correlations, means and spread in PRIAL increased with assumed effective sample size. However, negative effects of penalization for estimates of the genetic covariance matrix are evident for the small sample even at low values of the ESS, in particular for penalty $\mathcal{P}_\rho(P)$ shrinking genetic correlations towards their phenotypic counterparts. These reflect a severe mismatch between assumed prior distribution and actual distribution. For $\mathcal{P}_\rho(P)$, all negative PRIAL in $\hat{\Sigma}_G$ occurred for correlation structure IV, which was deliberately chosen so that genetic and phenotypic correlations differed substantially. For $\mathcal{P}_\rho(0)$, the mismatch was most pronounced for correlation scenario VI, where genetic and phenotypic correlations ranged from 0.60 to 0.72 and 0.40 to 0.56, respectively, so that shrinkage towards zero can become ineffective (see also Figure S2 below) and lead to undesirable effects on estimates of the phenotypic covariance matrix, which indicates ill-formulated penalization. Differences between corresponding PAC were less pronounced, thus resulting in higher mean PRIAL and more robust penalties ($\mathcal{P}_\pi(P)$ and $\mathcal{P}_\pi(0)$).

2 Individual results for sets of population values

Reductions in loss for individual sets of population values showed considerable variation. This is illustrated in a heatmap type plot in Figure S2 for a medium sample size ($s = 400$ sire families) and an effective sample size of the prior of $\nu = 8$. The plot also clearly demonstrates the close relationship between penalties on the improvement in estimates of the genetic covariance matrix and the strong effect the penalty on canonical eigenvalues can have on estimates of the residual covariance matrix.

The pattern of PRIAL for $\hat{\Sigma}_G$ can be directly related to differences in the population values for canonical eigenvalues, summarized in Figure S3. Generally, PRIAL for $\hat{\Sigma}_G$ tended to be higher for lower levels of heritabilities and thus canonical eigenvalues. Low PRIAL for heritability sets A, B and F combined with correlation structures I, II or V are clearly a reflection of the limited spread in population eigenvalues, which for this sample size does not quite match the range of $[0, 1]$ assumed in deriving the penalty. As discussed in the manuscript, reducing this interval for penalty \mathcal{P}_λ increased PRIAL for these cases substantially, e.g. from 6 and 7% to 29 and 32% for $\hat{\Sigma}_G$ and $\hat{\Sigma}_E$, respectively, for case A-I and from 19 and 11% to 20 and 26% for case B-V. For the smaller sample ($s = 100$, not shown), however, sampling variation for these cases was substantial enough for such ‘range effects’ to be much less evident.

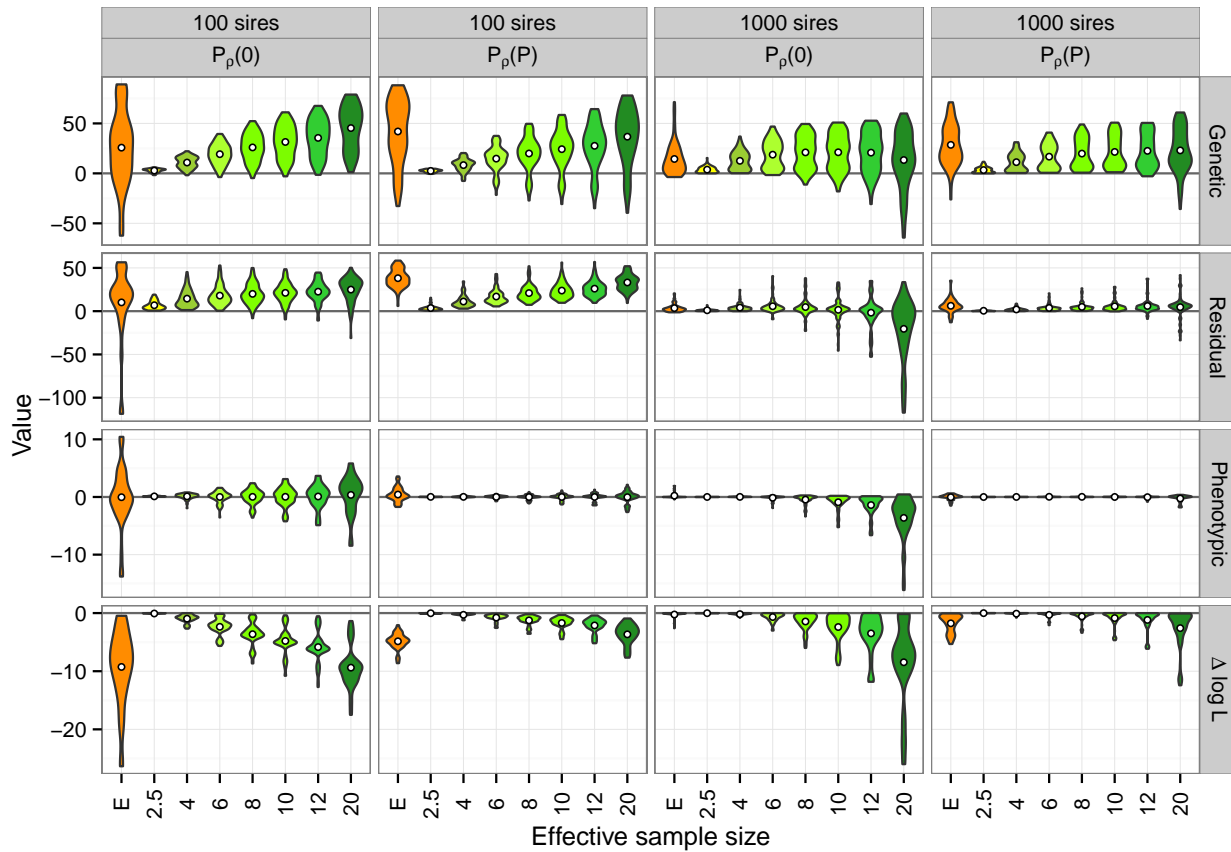


Figure S1 Distribution of percentage reduction in average loss for estimates of genetic, residual and phenotypic covariance matrices, together with corresponding change in log likelihood ($\Delta \mathcal{L}$) for penalties on genetic correlations, shrinking towards zero ($\mathcal{P}_\rho(0)$) or phenotypic values ($\mathcal{P}_\rho(P)$). Central circles give mean values. Numeric values on the x-axis are fixed, effective sample sizes while “E” denotes the use of a value estimated from the data for each replicate.

For the remaining cases the spread of eigenvalues appeared sufficient not to clash with the assumption of a distribution over the range $[0, 1]$. While differences shown in Figure S3 do not seem to be very pronounced, there are two other outliers, namely cases D-VI and E-VI. Both of these have a wide spread in population eigenvalues, with mean 0.33, the highest value approximately 0.8 and the three smallest values less than 0.1. This constellation appeared to lead to overshrinkage. For more stringent penalties, i.e. $\nu \geq 12$ these were the only two cases for which PRIAL for $\hat{\Sigma}_G$ due to penalty \mathcal{P}_λ was negative. It is not clear what particular feature of the population values was responsible for this behaviour – other cases had larger ranges or variances among canonical eigenvalues without being afflicted.

Similar patterns of PRIAL for all penalties (and high correlations shown in the manuscript) suggest that problematic constellations of canonical eigenvalues translate into corresponding difficulties for penalties on correlations. Corresponding modifications of the interval on with PAC are assumed to be distributed may be determined, but have not been investigated so far.

		Genetic												Residual															
VI	I	47	46	31	4	12	40	30	23	40	29	36	39	26	8	12	59	36	34	9	9	11	14	43	9	9	12	P_{λ}	
		6	10	33	57	58	14	49	56	54	56	41	36	52	7	11	63	34	37	7	9	12	15	41	9	9	13		
		50	55	55	33	43	52	43	38	46	38	49	52	42	39	49	40	61	60	12	12	22	26	51	12	13	35		
		31	52	58	46	55	59	52	47	55	47	56	58	58	36	43	64	60	60	13	13	14	15	40	13	13	36		
		7	13	45	58	61	31	50	55	57	57	48	46	56	7	11	60	35	33	9	9	12	14	39	9	9	12		
		6	9	26	54	45	13	47	55	54	57	36	29	41	7	10	63	26	24	7	8	11	13	27	8	8	11		
VI	II	48	46	31	9	16	47	30	22	44	30	38	42	23	6	5	9	8	10	12	9	7	8	12	9	9	6	$P_{\lambda}(\rho)$	
		4	5	29	53	54	22	53	62	61	64	46	41	52	3	2	1	2	2	9	6	5	6	4	7	6	3		
		54	59	57	29	39	58	39	31	42	32	51	55	34	18	22	8	15	16	11	10	10	12	21	11	11	17		
		33	52	59	46	56	68	56	47	57	49	61	64	60	9	11	3	9	8	9	8	7	6	7	9	8	9		
		6	10	42	57	59	44	58	63	68	67	58	55	57	5	4	5	6	7	13	9	7	8	10	10	10	5		
		4	4	21	52	41	23	54	63	64	67	45	37	42	4	2	1	2	2	10	7	6	7	5	8	7	3		
VI	III	48	47	29	3	8	48	24	15	39	24	35	39	20	13	13	26	29	26	17	11	7	7	24	13	12	9	$P_{\lambda}(0)$	
		5	5	28	54	56	23	55	62	61	64	48	41	54	4	4	6	9	16	9	8	7	8	10	9	8	7		
		53	57	57	33	43	60	44	37	47	38	53	57	42	24	32	17	30	29	11	10	11	13	21	10	11	20		
		34	52	60	46	57	67	56	49	59	51	61	64	60	15	21	11	22	19	9	9	8	8	12	9	9	9		
		11	13	44	57	61	50	60	64	70	67	61	58	60	11	9	20	21	22	18	11	9	10	15	13	12	9		
		5	5	19	51	42	23	53	63	64	66	45	37	44	4	2	2	2	2	10	7	6	7	5	8	7	3		
VI	IV	7	7	10	3	6	14	10	5	13	8	13	13	11	8	10	26	25	23	11	7	6	6	28	9	8	9	$P_{\lambda}(\rho)$	
		3	3	19	37	35	13	36	33	30	34	28	24	32	3	2	3	4	6	6	4	4	4	5	5	4	3		
		45	34	31	6	16	24	8	-2	2	-7	20	23	2	25	36	20	34	36	7	7	10	13	30	7	9	23		
		15	35	42	31	39	31	29	23	30	24	31	35	32	14	15	5	15	14	7	8	7	6	8	8	8	9		
		6	9	35	42	44	28	37	37	38	36	37	36	39	8	7	17	15	17	12	9	7	8	18	10	9	8		
		4	4	18	43	35	17	42	46	46	48	34	28	36	4	2	1	2	2	8	6	5	6	4	7	6	3		
VI	V	19	7	1	-13	-9	23	2	-3	11	2	6	7	-2	11	-4	31	18	11	5	-17	-27	-25	21	-15	-16	-17	$P_{\lambda}(0)$	
		3	1	13	38	42	8	28	29	20	28	19	13	31	2	4	13	12	21	5	2	0	1	6	2	1	4		
		43	40	34	15	22	30	18	10	15	8	27	30	20	30	39	26	41	51	3	7	7	11	30	3	5	23		
		18	37	40	30	41	32	26	24	30	25	31	32	38	18	28	15	31	38	4	7	5	3	15	4	7	4		
		15	9	31	44	47	41	37	39	42	40	39	38	42	16	5	31	19	22	15	-1	-4	-2	20	1	-1	0		
		4	4	18	46	36	17	42	48	47	49	35	30	34	4	2	1	2	2	8	6	5	6	5	6	6	3		
		A	B	C	D	E	F	G	I	H	J	K	L	M	A	B	C	D	E	F	G	I	H	J	K	L	M		

Figure S2 Percentage reduction in average loss for estimates of genetic and residual covariance matrices for 78 sets of population values for $s = 400$ sires, applying penalties with an effective prior sample size of $\nu = 8$

3 Detailed results for selected sets of population values

The effect of penalization of estimates of canonical eigenvalues and estimates of genetic correlations is illustrated for two sets of population values, obtained combining heritability set L with the compound symmetric correlation structure, II, (L-II) and combining heritability set H with correlation scenario III (H-III).

3.1 Case L-II

Heritabilities in set L are low to moderate, yielding population canonical eigenvalues ranging from 0.38 to 0.03 with a mean of 0.17. With all genetic correlations in II set to

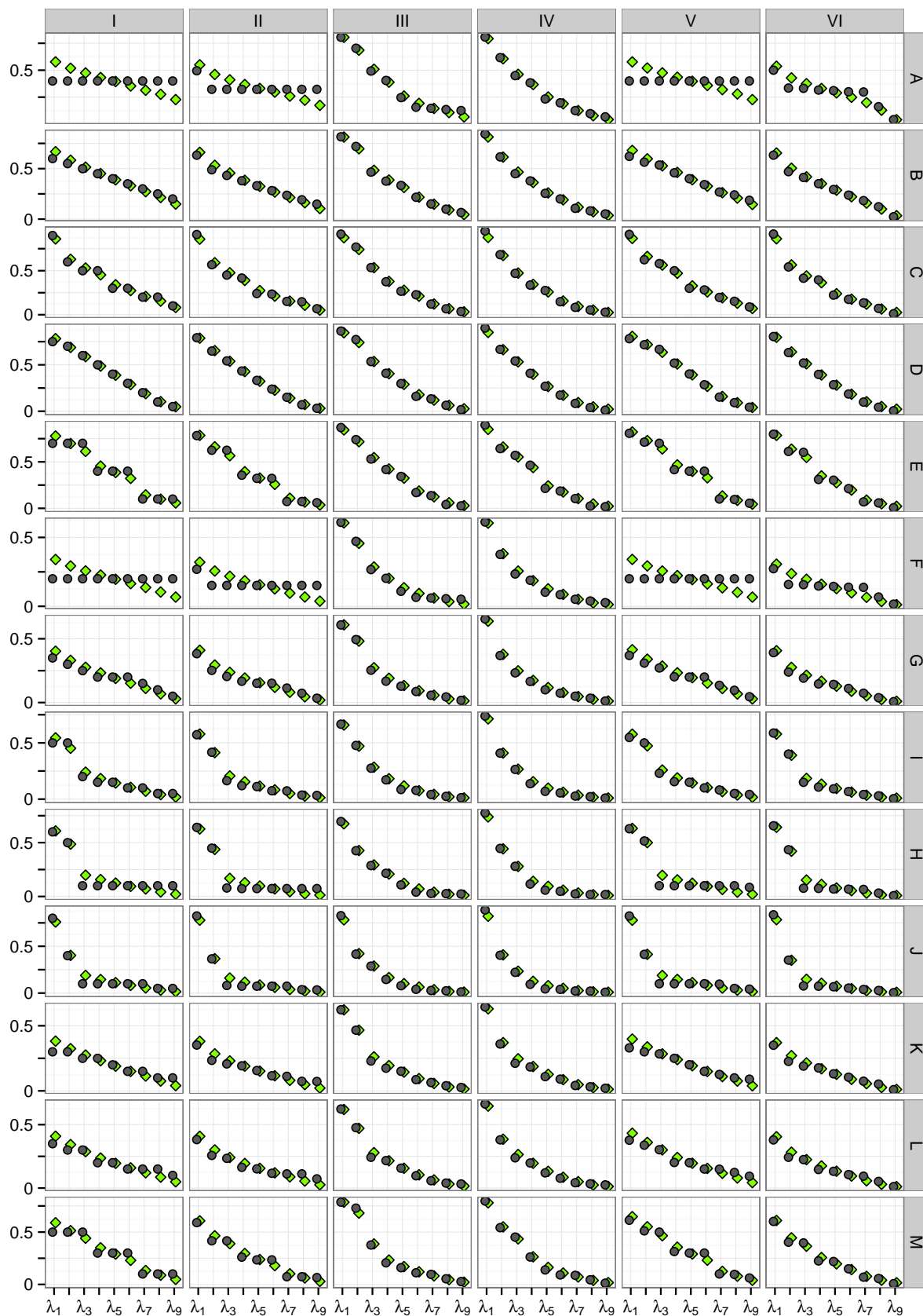


Figure S3 Population values for canonical eigenvalues (●) for 78 sets together with corresponding mean estimates across replicates (◆) for $s = 400$ applying a penalty on the canonical eigenvalues with an effective prior sample size of $\nu = 8$

0.50, PACs at the genetic level depend on the difference between trait numbers only, with values equal to 0.50, 0.33, 0.25, 0.20, 0.17, 0.14, 0.13 and 0.11 for 0 to 8 intervening traits, respectively. Phenotypic correlations for II range from 0.30 to 0.36 with corresponding PAC from 0.36 to 0.07.

Mean estimates of canonical eigenvalues across replicates from unpenalized and penalized analyses with increasingly stringent penalties are summarized in Table S1, together with the respective PRIAL. As expected from sampling theory, unpenalized estimates of eigenvalues are overdispersed with the largest values systematically biased upwards and the smallest values biased downwards. All penalties reduce the overdispersion but in a slightly different fashion. The most notable difference is that the penalty on partial auto-correlations increases the mean eigenvalue slightly, i.e. increases the lowest values more than it reduces the highest ones. For this case, penalties \mathcal{P}_λ and $\mathcal{P}_\pi(P)$ up $\nu = 20$ are advantageous. As phenotypic PAC are greater than zero, penalty $\mathcal{P}_\pi(0)$ is more stringent than $\mathcal{P}_\pi(P)$. This is evidenced by larger PRIAL and changes in likelihood at the same ESS. However, a decreasing PRIAL for $\hat{\Sigma}_E$ for $\nu > 12$ provides some indication that such values are too stringent – this is supported by the mean estimates of the highest and lowest eigenvalues being below and above, respectively, the corresponding population values.

Figure S4 gives the distribution of estimates of genetic correlations, contrasting penalized and unpenalized estimates. Distributions of unpenalized estimates in the three panels are not exactly the same as they represent different simulation runs. Correlations are given in order of the lower triangle of the correlation matrix row-wise. As population heritabilities decrease with trait number, this means that correlations are also shown in increasing ordering of sampling variances. Again, the overall pattern is very similar for all three penalties considered. The effects of penalization are most apparent for correlations for the higher trait numbers. For these, unpenalized estimates are most afflicted by constraints on the parameter space and thus have mean below the population value of 0.5.

3.2 Case H-III

Set H is comprised of two high heritabilities, 0.5 and 0.6, with the remainder equal to 0.1. Genetic correlations for III follow an auto-regressive pattern, from 0.7 to $0.7^8 = 0.06$. This gives phenotypic correlations for H-III ranging from -0.11 to 0.50, and population canonical eigenvalues from 0.70 to 0.02. The auto-regressive genetic correlation structure gives genetic PAC for H-III that are high for ‘adjacent’ traits (0.7) and zero otherwise. Corresponding phenotypic values range from 0.50 (traits 1 and 2) to 0.11, creating differences between genetic and phenotypic PAC up to 0.81.

Mean estimates of canonical eigenvalues for different penalties for this set of population values are given in Table S2 and the distribution of estimates of genetic correlations is shown in Figure S5. In this case, correlations are arranged according to their population values, with groups equal to the elements of successive subdiagonal of the lower triangle of the correlation matrix. While the general pattern of results is similar to that observed for population values L-II, the spread of canonical eigenvalues is larger with more eigenvalues close to zero and the effects of sampling variation are thus more pronounced. In particular, bias in unpenalized estimates of genetic correlations between traits i and $i + 1$ due to constraints on the parameter space is notable for the traits with low heri-

Table S1 Mean estimates of canonical eigenvalues ($\times 1000$) together with corresponding mean change in unpenalized likelihood for population values obtained combining heritability set L with correlation scenario II for $s = 400$.

ESS ^a	Canonical eigenvalues ^b										$\Delta\mathcal{L}^c$	PRIAL ^d	
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	$\bar{\lambda}$		G	E
	<i>Population values</i>												
	382	257	234	163	152	117	112	112	73	178			
	<i>Unpenalized estimates</i>												
	422	310	250	198	157	120	85	49	13	178	0.000		
	<i>Penalty on canonical eigenvalues</i>												
4	417	308	247	197	156	120	85	51	20	178	-0.070	32	4
6	411	303	245	195	156	120	86	54	23	177	-0.163	40	7
8	409	302	244	195	156	119	87	55	26	177	-0.275	45	10
12	402	298	241	193	155	120	89	59	33	177	-0.478	52	13
16	399	296	240	193	155	121	91	63	37	177	-0.721	57	16
20	393	293	237	192	155	122	93	66	40	177	-0.998	61	19
	<i>Penalty on partial correlations: shrink towards phenotypic</i>												
4	420	308	248	197	157	121	88	55	25	180	-0.175	42	4
6	417	309	248	197	157	122	89	59	32	181	-0.369	51	7
8	412	307	247	197	157	122	90	62	36	181	-0.567	55	10
12	409	307	247	196	158	123	94	67	43	183	-0.952	61	13
16	404	304	244	195	158	126	97	72	49	183	-1.320	65	17
20	400	300	244	196	157	126	99	75	53	183	-1.752	68	19
	<i>Penalty on partial correlations: shrink towards zero</i>												
4	412	309	251	199	157	122	88	55	26	180	-0.238	45	6
6	407	310	250	199	160	124	92	62	33	182	-0.546	53	10
8	399	309	251	200	159	125	94	65	38	182	-0.898	58	13
12	388	306	249	200	163	128	100	73	47	184	-1.863	64	15
16	379	304	250	200	162	129	101	75	52	184	-3.318	66	11
20	370	303	248	199	161	130	103	78	55	183	-4.910	66	7
	<i>Penalty on correlations: shrink towards phenotypic</i>												
4	407	301	242	196	155	118	82	45	8	172	-0.076	15	4
6	406	302	242	196	157	121	83	46	11	174	-0.280	27	7
8	401	301	243	197	159	122	84	48	14	174	-0.553	36	9
12	392	301	244	198	161	126	90	53	19	176	-1.131	48	12
16	384	299	244	201	162	127	91	56	22	176	-1.669	53	12
20	381	298	241	200	163	129	94	58	25	177	-2.177	55	14

^aEffective sample size of prior

^bIn descending order with $\bar{\lambda}$ denoting their mean

^cChange in log likelihood compared to unpenalized analysis

^dPercentage reduction in average loss for genetic (G) and residual (E) covariance matrix

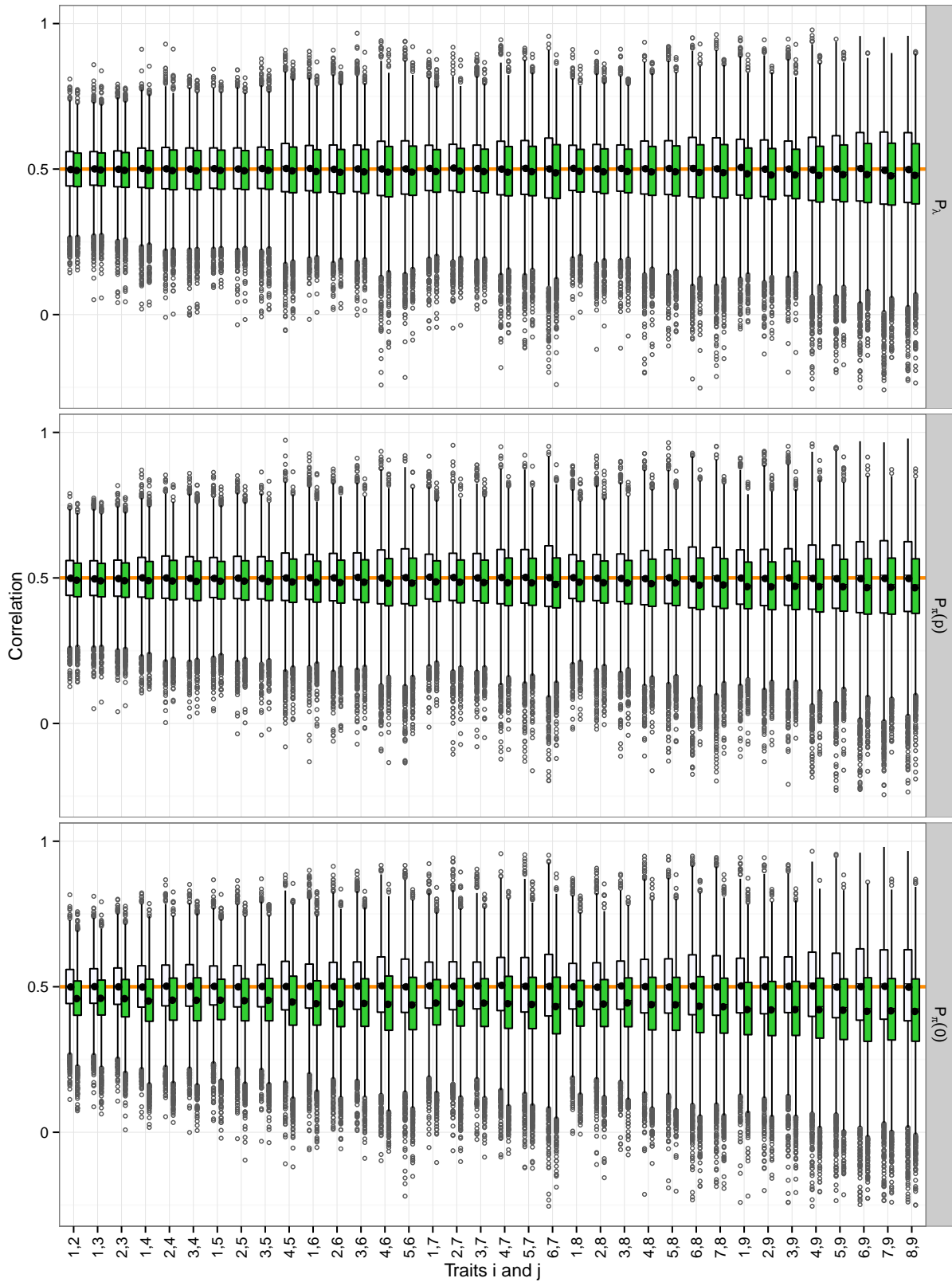


Figure S4 Distribution of estimates of genetic correlations between trait i and j across replicates for population values obtained by combining heritability set L with correlation scenario II for $s = 400$, applying penalties with an effective prior sample size of $\nu = 8$ (□ unpenalized, ■ penalized estimates; ● mean values)

Table S2 Mean estimates of canonical eigenvalues ($\times 1000$) together with corresponding mean change in unpenalized likelihood for population values obtained combining heritability set H with correlation scenario III for $s = 400$.

ESS ^a	Canonical eigenvalues ^b										$\Delta\mathcal{L}^c$	PRIAL ^d	
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	$\bar{\lambda}$		G	E
	<i>Population values</i>												
	696	425	287	215	108	41	29	26	23	206			
	<i>Unpenalized estimates</i>												
	715	445	305	214	121	67	29	5	1	211			
	<i>Penalty on canonical eigenvalues</i>												
4	699	436	302	209	119	69	34	12	5	209	-0.326	40	10
6	688	433	296	211	120	70	37	17	8	209	-0.712	49	13
8	678	431	296	209	118	71	41	21	12	208	-1.115	54	15
12	664	422	290	207	124	75	46	26	17	208	-1.974	58	17
16	648	415	289	207	123	76	49	30	21	206	-2.888	59	15
20	632	412	289	206	125	81	54	36	25	206	-3.698	57	12
	<i>Penalty on partial correlations: shrink towards phenotypic</i>												
4	719	445	302	211	122	69	38	18	10	215	-0.739	52	3
6	706	438	300	209	123	74	44	25	16	215	-1.447	57	4
8	712	436	302	208	123	76	50	31	21	218	-2.026	57	6
12	706	429	301	203	123	81	56	38	28	218	-3.260	57	8
16	702	427	297	205	123	84	61	44	34	220	-4.361	54	11
20	697	425	294	202	126	88	66	50	40	221	-5.343	51	11
	<i>Penalty on partial correlations: shrink towards zero</i>												
4	710	438	303	213	120	69	38	18	10	213	-0.683	51	3
6	709	444	304	207	121	73	44	25	15	216	-1.340	57	5
8	701	440	306	205	124	75	47	29	20	216	-1.997	59	7
12	697	438	303	199	124	79	54	37	27	218	-3.331	58	10
16	689	440	299	194	125	81	58	41	31	218	-4.863	57	10
20	679	441	303	195	125	84	60	45	35	219	-6.267	55	10
	<i>Penalty on correlations: shrink towards phenotypic</i>												
4	712	444	305	214	126	74	39	11	1	214	-0.178	12	2
6	705	443	306	209	125	79	46	17	3	215	-0.587	23	3
8	711	438	301	204	128	84	52	24	6	216	-1.076	30	6
12	700	438	300	202	130	88	59	31	11	218	-2.216	38	8
16	700	436	296	197	129	91	63	37	17	218	-3.422	42	10
20	698	428	297	195	130	92	66	42	21	219	-4.528	43	11

^aEffective sample size of prior

^bIn descending order with $\bar{\lambda}$ denoting their mean

^cChange in log likelihood compared to unpenalized analysis

^dPercentage reduction in average loss for genetic (G) and residual (E) covariance matrix

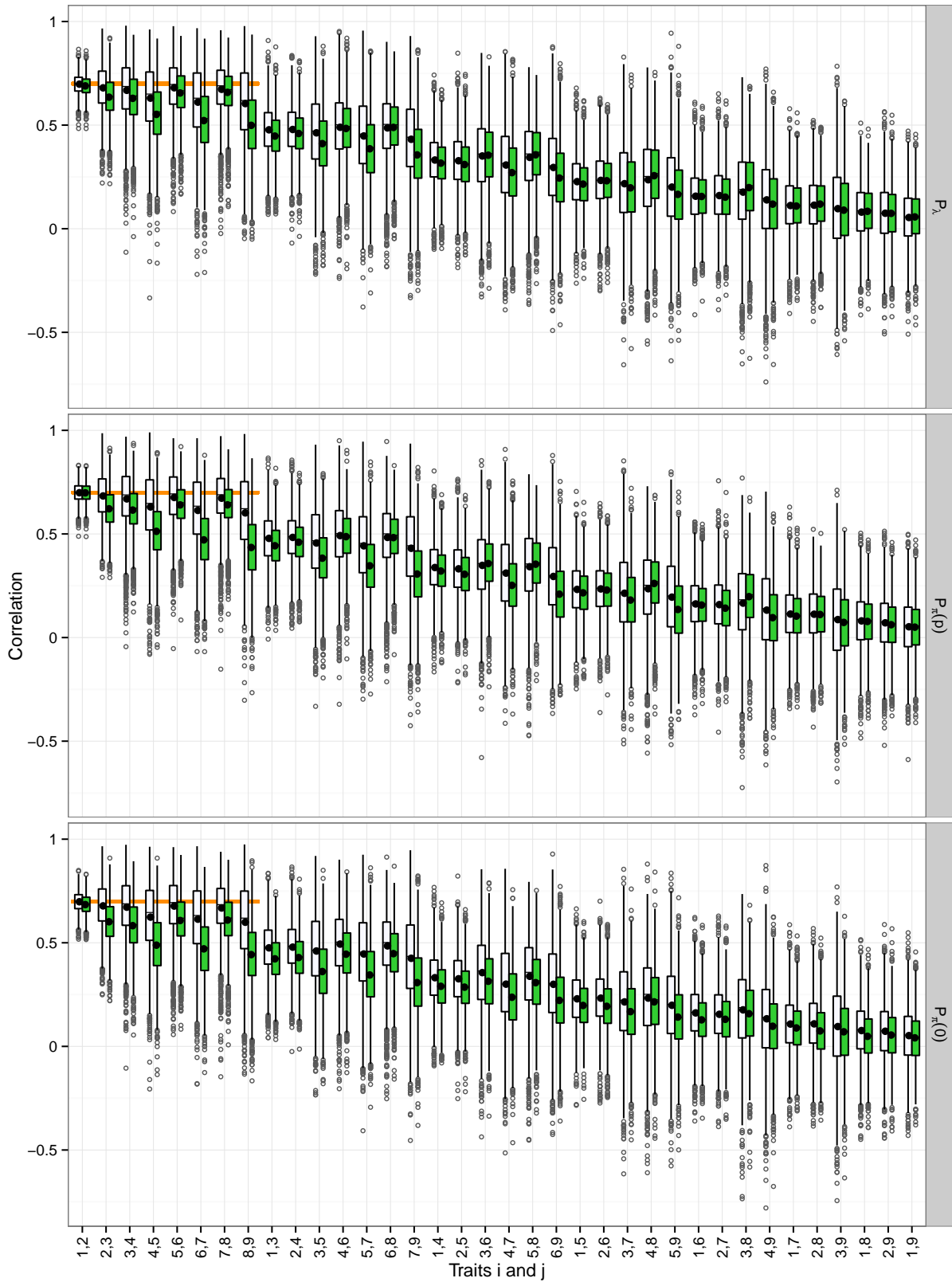


Figure S5 Distribution of estimates of genetic correlations between trait i and j across replicates for population values obtained by combining heritability set H with correlation scenario III for $s = 400$, applying penalties with an effective prior sample size of $\nu = 8$ (□ unpenalized, ■ penalized estimates; ● mean values). Horizontal line shows population value for traits i and $i + 1$)

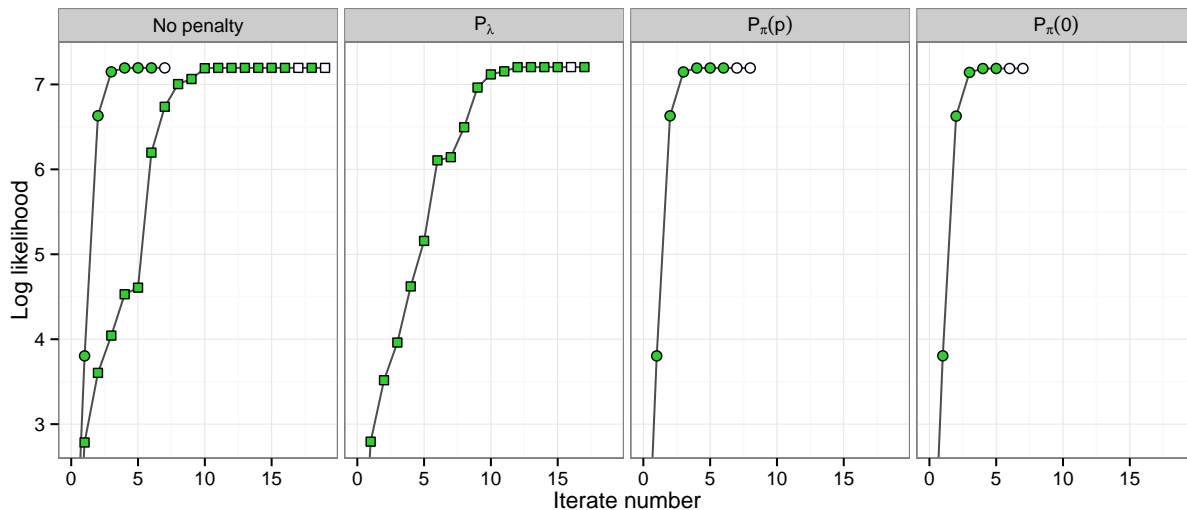


Figure S6 Effect of penalization and parameterisation on convergence of REML analyses for a single replicate of simulated data obtained by combining heritability set H with correlation scenario III for $s = 400$, applying penalties with an effective prior sample size of $\nu = 8$ (Parameterisation: ● Elements of Cholesky factors of covariance matrices, ■ Elements of canonical decomposition; filled symbols: average information, open symbols: derivative-free maximization steps; Likelihood values are given as deviation from value starting values and scaled by dividing by 1000)

tability ($i \geq 3$). This bias is larger if the corresponding residual correlation is negative. This results in a substantial difference between genetic and phenotypic values so that penalization increases bias in estimates of those genetic correlations quite markedly.

4 Effect of penalization on convergence of REML analyses

We illustrate the effects of penalization on the convergence behaviour of REML analyses for a data set simulated for population values H-III and $s = 400$ sires and effective sample sizes of $\nu = 8$ and $\nu = 20$. Analyses were carried out using our mixed model package WOMBAT, using an average information algorithm to locate the maximum of the likelihood, followed by derivative-free search steps (and possibly additional average information and derivative-free steps) to ensure convergence. Data and parameter files are available as worked Example 19 for the package.

Figure S6 shows the convergence behaviour for mild penalization ($\nu = 8$) and additional characteristics are summarized in Table S3. For a simple model with means as the only fixed effects and all records recorded on all animals, convergence of the standard, unpenalized analysis is rapid in spite of the substantial number of parameters to be estimated (45 genetic covariances and 45 residual covariances). This uses a parameterisation to the elements of the Cholesky factors of the two covariance matrices to be estimated. Performing the same analysis parameterising to estimate the canonical eigenvalues and elements of the canonical transformation instead increases the number of iterates required. The effect of parameterisation carries through to penalized analyses. While estimation on the canonical simplifies implementation, it is relatively slow to converge. Convergence appeared to be to the maximum of the penalized likelihood, with derivative-free search

Table S3 Characteristics of convergence behaviour and changes in likelihood for a single, simulated data set for $s = 400$ sires with population parameters obtained combining heritability set H with correlation scenario III.

ESS ^a		Unpenalized		\mathcal{P}_λ	$\mathcal{P}_\pi(0)$	$\mathcal{P}_\pi(P)$	$\mathcal{P}_\rho(P)$
		Cholesky ^b	Canonical ^c	Canonical	Cholesky	Cholesky	Cholesky
8	Likelihood ^d	0	-0.002	-1.397	-1.742	-1.578	-1.319
	AI iterates ^e	10	25	30	10	8	8
	DF iterates ^f	0.000	0.000	0.001	0.058	0.060	0.012
20	Likelihood			-8.252	-5.979	-4.294	-4.523
	AI iterates			25	11	9	9
	DF iterates			0.000	0.079	0.109	0.068

^aEffective sample size ν

^bStandard parameterisation to elements of the Cholesky factor of covariance matrices to be estimated

^cParameterisation to elements of the canonical transformation

^dUnpenalized log likelihood, expressed as deviation from standard, unpenalized analysis

^eNumber of average information algorithm REML iterates carried out

^fChange in penalized log likelihood in subsequent derivative-free maximization steps, expressed as deviation from value for last AI step

steps not increasing it further, even for more stringent penalization for $\nu = 20$. In contrast, penalties on correlations are implemented using the standard parameterisation to elements of the Cholesky factor. Thus penalized estimation has little impact on the number of average information algorithm iterates required. However, for these cases subsequent derivative-free search steps yield additional, small increases in the likelihood, especially for the more stringent penalisation for $\nu = 20$.