

# Supplementary Materials to Incorporating Group Correlations in Genome-Wide Association Studies Using Smoothed Group Lasso

JIN LIU\*

*School of Public Health, Yale University, New Haven, CT 06520, U.S.A.*

jin.liu.jl2329@yale.edu

JIAN HUANG

*Department of Statistics and Actuarial Science, and Department of Biostatistics, University of*

*Iowa, Iowa City, IA 52242, U.S.A.*

SHUANGGE MA

*School of Public Health, Yale University, New Haven, CT 06520, U.S.A.*

KAI WANG

*Department of Biostatistics, University of Iowa, Iowa City, IA 52242, U.S.A.*

## CANONICAL CORRELATION

With the SGL penalty, the canonical correlation is suitable to measure the associations between adjacent groups. Canonical correlation analyzes the correlation between a linear combination of variables in one set and a linear combination of variables in another set. It searches for coefficient

\*To whom correspondence should be addressed.

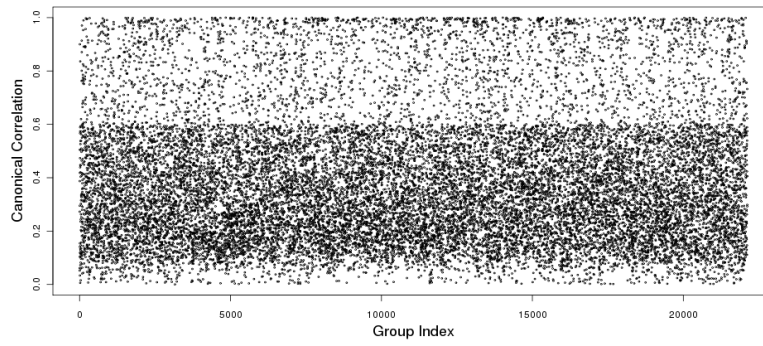
vectors  $\mathbf{a}$  and  $\mathbf{b}$  such that

$$\text{Corr}(U, V) = \frac{\mathbf{a}'\Sigma_{12}\mathbf{b}}{\sqrt{\mathbf{a}'\Sigma_{11}\mathbf{a}}\sqrt{\mathbf{b}'\Sigma_{22}\mathbf{b}}}$$

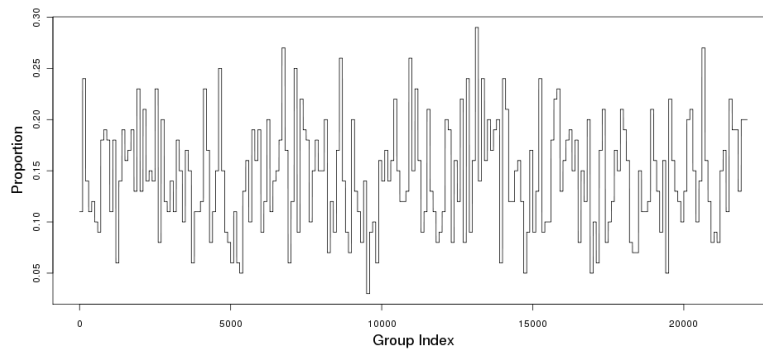
is maximized. Here  $U = \mathbf{a}'X^{(1)}$ ,  $V = \mathbf{b}'X^{(2)}$ ,  $\Sigma_{11} = \text{Cov}(X^{(1)}, X^{(1)})$ ,  $\Sigma_{22} = \text{Cov}(X^{(2)}, X^{(2)})$  and  $\Sigma_{12} = \text{Cov}(X^{(1)}, X^{(2)})$ . By the change of basis and Cauchy-Schwartz inequality, it can be shown that  $\max_{\mathbf{a}, \mathbf{b}} \text{Corr}(\mathbf{a}'X^{(1)}, \mathbf{b}'X^{(2)}) = \sqrt{\pi_1}$ , where  $\pi_1$  is the largest eigenvalue of  $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ .

Note that canonical correlation is always positive. This can be guaranteed by choosing an appropriate sign for  $\mathbf{b}$ . This property is desirable as SGL uses the canonical correlation as weight to smooth estimates.

COMPLEMENTARY FIGURE AND TABLE



(a) Correlation coefficient  $\zeta_j$



(b) Correlation coefficients larger than 0.6 averaged within non-overlapping 100-SNPs windows.

Fig. 3. Plots of absolute lag-1 autocorrelation  $\zeta_j$  on Chromosome 6 from Genetic Analysis Workshop 16 rheumatoid arthritis data.

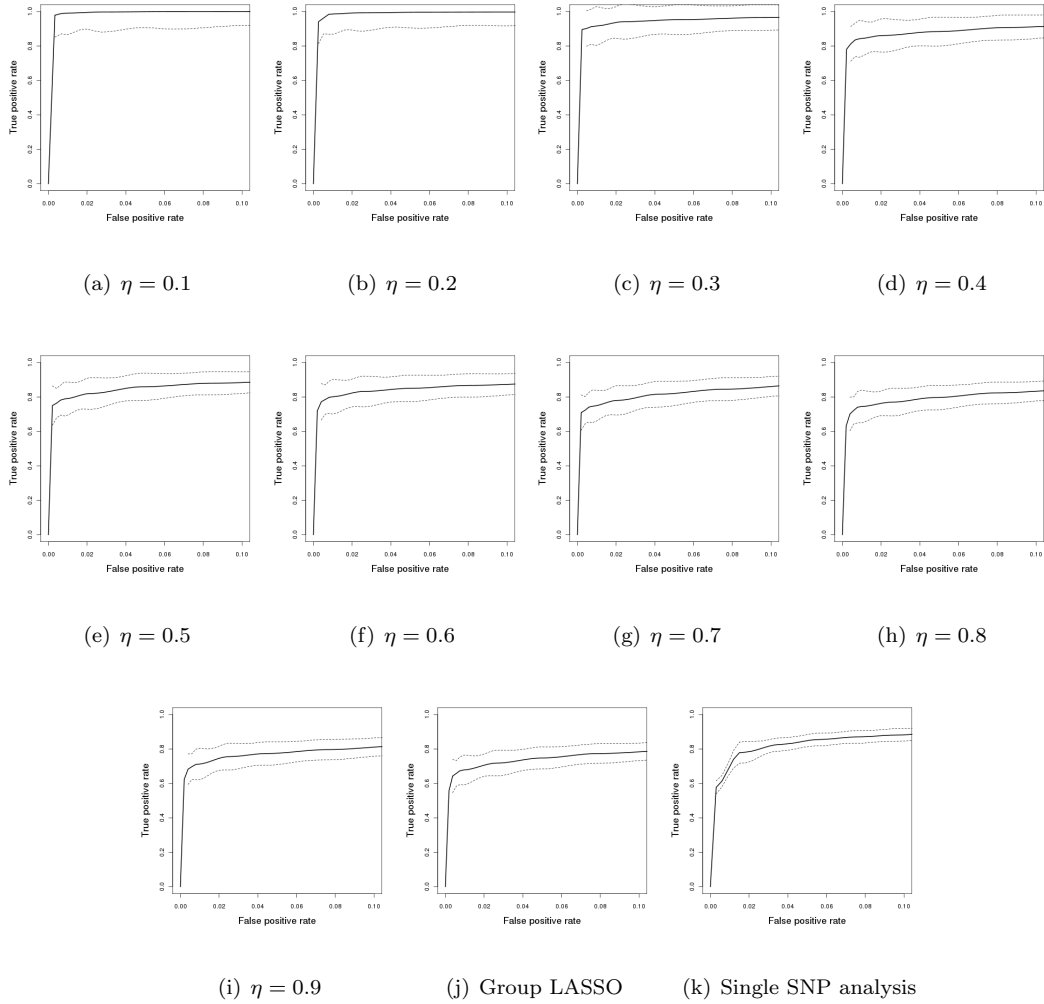


Fig. 4. ROC plots for example 1. Plots (a)-(i) are for SGL. Solid lines are the estimated ROC curve and dotted lines are 95% point-wise confidence intervals.

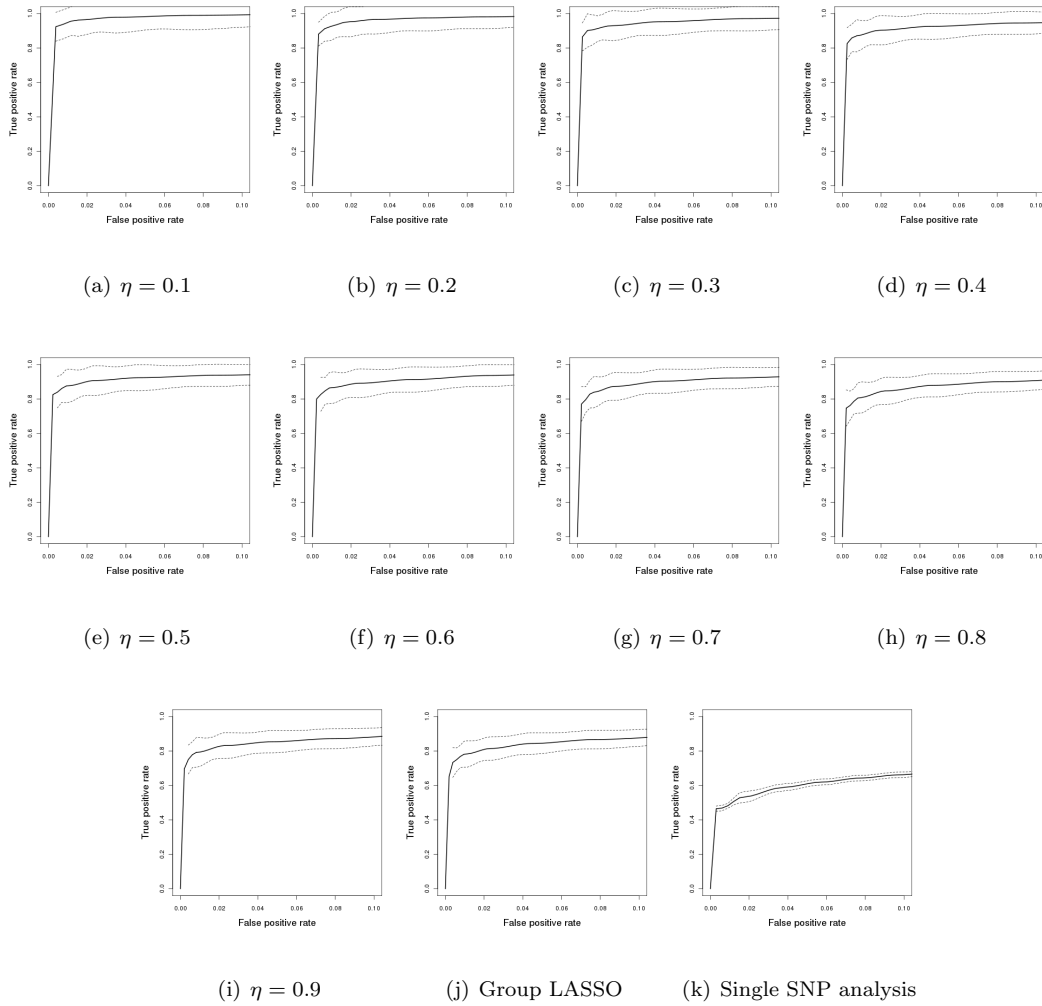


Fig. 5. ROC plots for example 2. Plots (a)-(i) are for SGL. Solid lines are the estimated ROC curve and dotted lines are 95% point-wise confidence intervals.

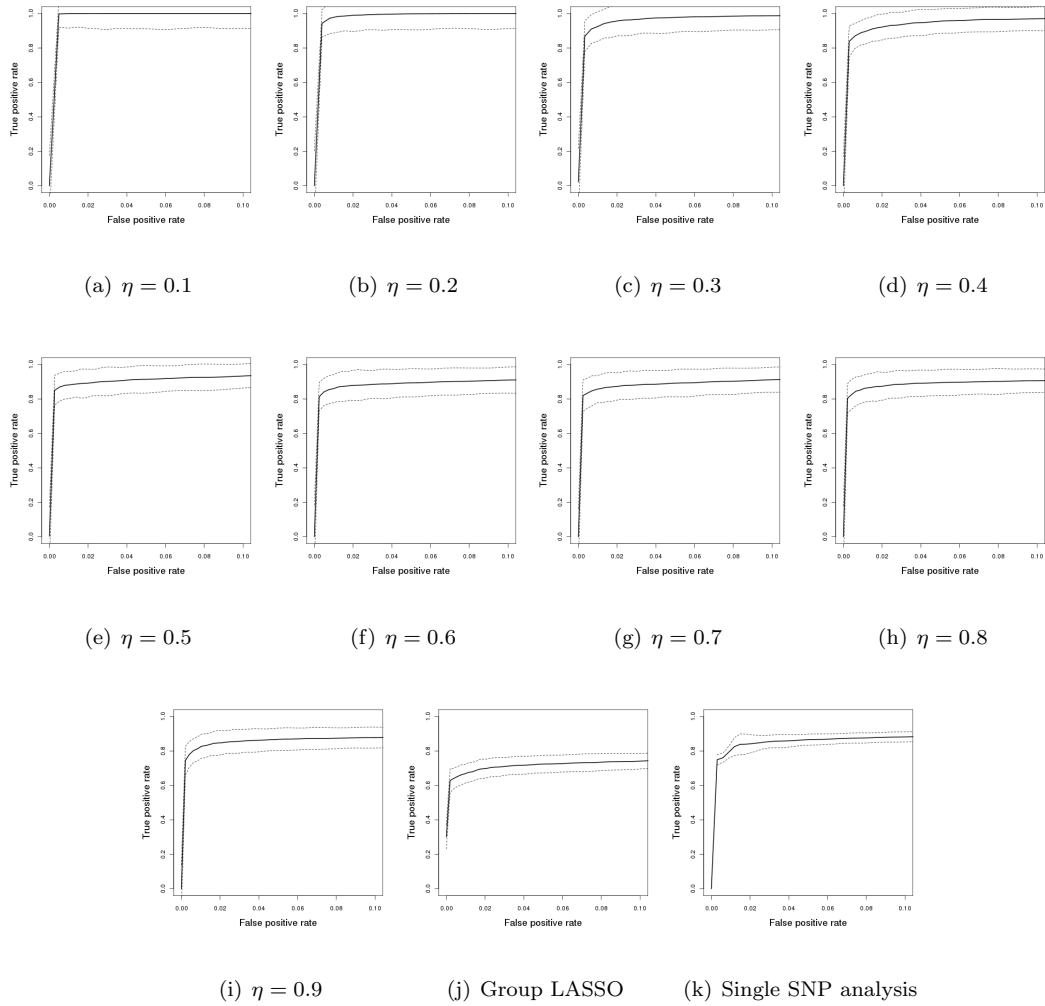


Fig. 6. ROC plots for example 3. Plots (a)-(i) are for SGL. Solid lines are the estimated ROC curve and dotted lines are 95% point-wise confidence intervals.

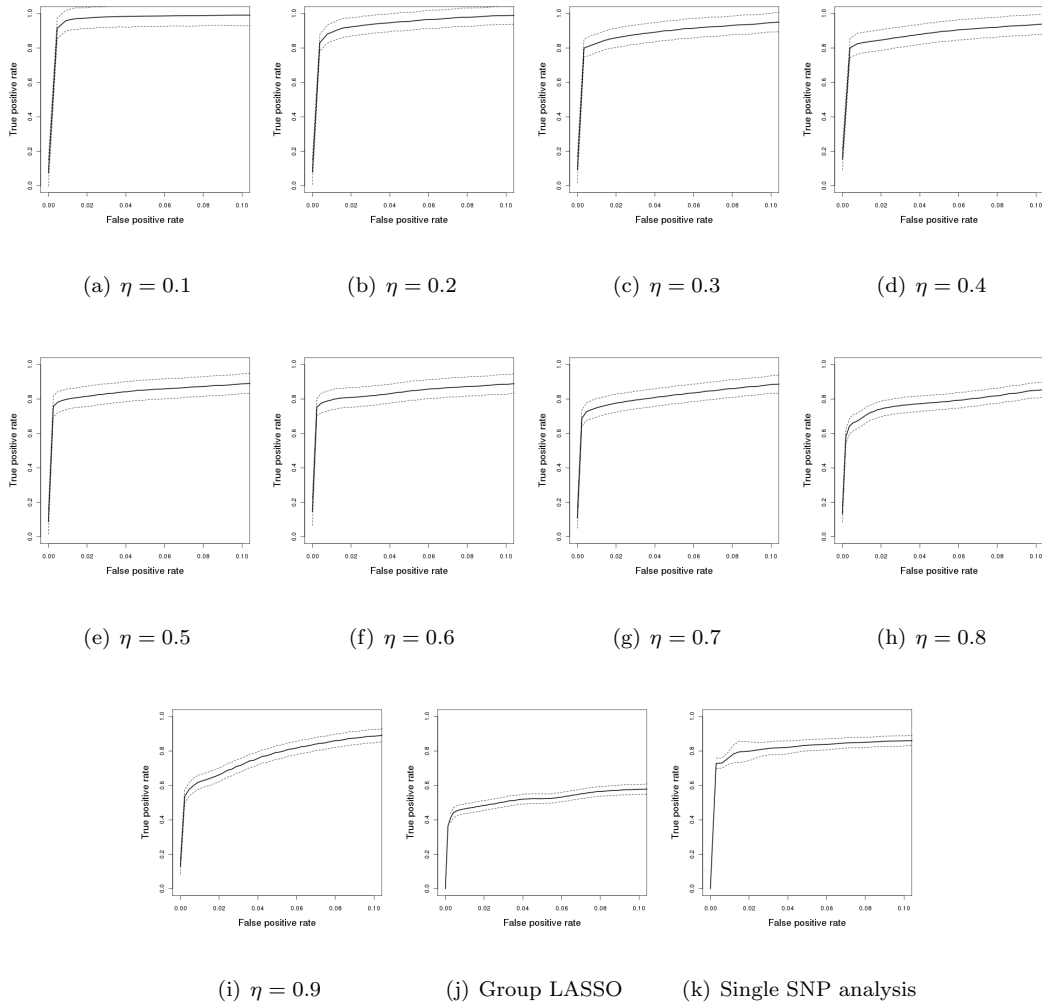
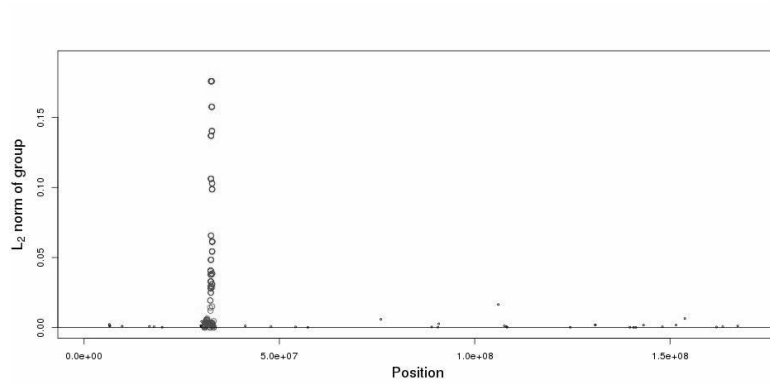
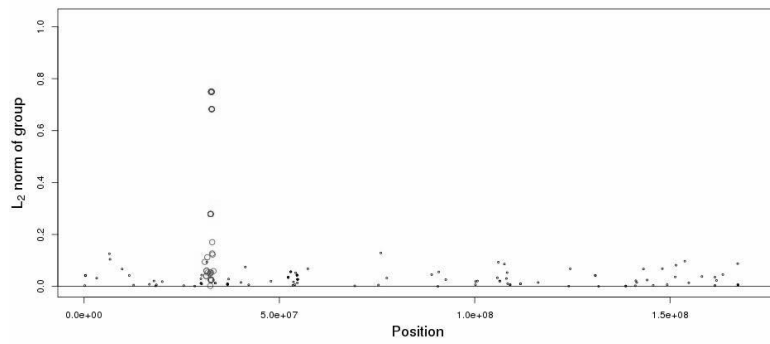


Fig. 7. ROC plots for example 4. Plots (a)-(i) are for SGL. Solid lines are the estimated ROC curve and dotted lines are 95% point-wise confidence intervals.

(a) SGL  $\eta = 0.1$ 

(b) group Lasso

Fig. 8. Plots of  $\|\beta\|$  for SGL and group Lasso with threshold = 0.2 to form groups for rheumatoid arthritis data. The smaller dots are estimates with insignificant  $p$ -values and the larger dots stand for estimates with significant  $p$ -values.



Table 3. Mean (standard deviation) of true positive (TP) number of groups, false discovery rate (FDR) and false negative rate (FNR) for simulated data with incorrect block sizes,  $s = 15$ .

$\eta$	Example 1			Example 2		
	TP	FDR	FNR	TP	FDR	FNR
0.1	8.90(0.30)	0.41(0.02)	0.01(0.03)	7.24(0.69)	0.52(0.05)	0.20(0.08)
0.2	8.66(0.48)	0.42(0.03)	0.04(0.05)	7.36(0.69)	0.51(0.05)	0.18(0.08)
0.3	8.38(0.64)	0.44(0.04)	0.07(0.07)	7.02(0.65)	0.53(0.04)	0.22(0.07)
0.4	8.02(0.89)	0.47(0.06)	0.11(0.10)	6.62(0.92)	0.56(0.06)	0.26(0.10)
0.5	7.44(0.84)	0.50(0.06)	0.17(0.09)	6.36(0.88)	0.58(0.06)	0.29(0.10)
0.6	6.74(0.75)	0.55(0.05)	0.25(0.08)	5.94(0.93)	0.60(0.06)	0.34(0.10)
0.7	6.76(0.66)	0.55(0.04)	0.25(0.07)	5.80(0.83)	0.61(0.06)	0.36(0.09)
0.8	6.34(0.59)	0.58(0.04)	0.30(0.07)	5.98(0.82)	0.60(0.05)	0.34(0.09)
0.9	6.02(0.71)	0.60(0.05)	0.33(0.08)	5.78(0.86)	0.61(0.06)	0.36(0.10)
1	5.78(0.65)	0.61(0.04)	0.36(0.07)	5.52(0.84)	0.63(0.06)	0.39(0.09)

$\eta$	Example 3			Example 4		
	TP	FDR	FNR	TP	FDR	FNR
0.1	7.58(1.11)	0.49(0.07)	0.16(0.12)	6.02(1.25)	0.60(0.08)	0.33(0.14)
0.2	6.72(0.70)	0.55(0.05)	0.25(0.08)	5.56(1.18)	0.63(0.08)	0.38(0.13)
0.3	6.52(0.68)	0.57(0.05)	0.28(0.08)	5.34(0.89)	0.64(0.06)	0.41(0.10)
0.4	6.56(0.67)	0.56(0.05)	0.27(0.08)	5.10(0.97)	0.66(0.06)	0.43(0.11)
0.5	6.46(0.61)	0.57(0.04)	0.28(0.07)	4.82(1.10)	0.68(0.07)	0.46(0.12)
0.6	6.54(0.73)	0.56(0.05)	0.27(0.08)	4.22(1.09)	0.72(0.07)	0.53(0.12)
0.7	6.28(0.57)	0.58(0.04)	0.30(0.06)	4.14(1.13)	0.72(0.08)	0.54(0.13)
0.8	6.12(0.63)	0.59(0.04)	0.32(0.07)	3.50(1.07)	0.77(0.07)	0.61(0.12)
0.9	5.70(0.79)	0.62(0.05)	0.37(0.09)	2.58(0.99)	0.83(0.07)	0.71(0.11)
1	4.34(0.85)	0.71(0.06)	0.52(0.09)	1.46(0.84)	0.90(0.06)	0.84(0.09)

<sup>a</sup> The optimal  $\eta$  for linear models (Example 1—3) is 0.3.

<sup>b</sup> The optimal  $\eta$  for logistic regression model (Example 4) is 0.1.

<sup>c</sup> When  $\eta = 1$ , SGL becomes group Lasso.

Table 4. Mean (standard deviation) of true positive (number of groups), false discovery rate (FDR) and false negative rate (FNR) for simulated data,  $s = 10$ .

$\eta$	Example 1			Example 2		
	TP	FDR	FNR	TP	FDR	FNR
0.1	8.88(0.39)	0.11(0.04)	0.01(0.04)	8.00(0.57)	0.20(0.06)	0.11(0.06)
0.2	8.80(0.45)	0.12(0.05)	0.02(0.05)	8.08(0.44)	0.19(0.04)	0.10(0.05)
0.3	8.38(0.73)	0.16(0.07)	0.07(0.08)	7.98(0.43)	0.20(0.04)	0.11(0.05)
0.4	7.68(0.79)	0.23(0.08)	0.15(0.09)	7.64(0.63)	0.24(0.06)	0.15(0.07)
0.5	7.42(0.88)	0.26(0.09)	0.18(0.10)	7.46(0.65)	0.25(0.06)	0.17(0.07)
0.6	7.20(0.57)	0.28(0.06)	0.20(0.06)	7.38(0.64)	0.26(0.06)	0.18(0.07)
0.7	6.84(0.82)	0.32(0.08)	0.24(0.09)	6.96(0.67)	0.30(0.07)	0.23(0.07)
0.8	6.62(0.73)	0.34(0.07)	0.26(0.08)	7.00(0.70)	0.30(0.07)	0.22(0.08)
0.9	6.52(0.65)	0.35(0.06)	0.28(0.07)	6.76(0.74)	0.32(0.07)	0.25(0.08)
1	6.20(0.57)	0.38(0.06)	0.31(0.06)	6.74(0.92)	0.33(0.09)	0.25(0.10)

$\eta$	Example 3			Example 4		
	TP	FDR	FNR	TP	FDR	FNR
0.1	7.06(0.42)	0.29(0.04)	0.22(0.05)	5.80(1.26)	0.42(0.13)	0.36(0.14)
0.2	7.10(0.58)	0.29(0.06)	0.21(0.06)	5.44(1.30)	0.46(0.13)	0.40(0.14)
0.3	7.16(0.82)	0.28(0.08)	0.20(0.09)	5.04(1.26)	0.50(0.13)	0.44(0.14)
0.4	6.90(0.61)	0.31(0.06)	0.23(0.07)	5.12(1.42)	0.49(0.14)	0.43(0.16)
0.5	6.86(0.64)	0.31(0.06)	0.24(0.07)	4.74(1.16)	0.53(0.12)	0.47(0.13)
0.6	6.82(0.66)	0.32(0.07)	0.24(0.07)	4.70(0.91)	0.53(0.09)	0.48(0.10)
0.7	6.60(0.64)	0.34(0.06)	0.27(0.07)	3.92(1.29)	0.61(0.13)	0.56(0.14)
0.8	6.12(0.80)	0.39(0.08)	0.32(0.09)	3.42(1.16)	0.66(0.12)	0.62(0.13)
0.9	5.50(0.65)	0.45(0.06)	0.39(0.07)	3.02(1.13)	0.70(0.11)	0.66(0.13)
1	4.60(0.86)	0.54(0.09)	0.49(0.10)	2.28(0.73)	0.77(0.07)	0.75(0.08)

<sup>a</sup> The optimal  $\eta$  for linear models (Example 1—3) is 0.3.

<sup>b</sup> The optimal  $\eta$  for logistic regression model (Example 4) is 0.1.

<sup>c</sup> When  $\eta = 1$ , SGL becomes group Lasso.

Table 5. Mean (standard deviation) of true positive (TP) number of groups, false discovery rate (FDR) and false negative rate (FNR) for simulated data with incorrect block sizes,  $s = 10$ .

$\eta$	Example 1			Example 2		
	TP	FDR	FNR	TP	FDR	FNR
0.1	8.88(0.39)	0.11(0.04)	0.01(0.04)	7.04(0.73)	0.30(0.07)	0.22(0.08)
0.2	8.58(0.54)	0.14(0.05)	0.05(0.06)	6.84(0.62)	0.32(0.06)	0.24(0.07)
0.3	8.32(0.77)	0.17(0.08)	0.08(0.09)	6.72(0.83)	0.33(0.08)	0.25(0.09)
0.4	7.68(0.87)	0.23(0.09)	0.15(0.10)	6.22(0.71)	0.38(0.07)	0.31(0.08)
0.5	7.04(0.70)	0.30(0.07)	0.22(0.08)	6.02(0.87)	0.40(0.09)	0.33(0.10)
0.6	6.60(0.64)	0.34(0.06)	0.27(0.07)	5.80(0.78)	0.42(0.08)	0.36(0.09)
0.7	6.42(0.61)	0.36(0.06)	0.29(0.07)	5.64(0.92)	0.44(0.09)	0.37(0.10)
0.8	6.18(0.63)	0.38(0.06)	0.31(0.07)	5.54(0.79)	0.45(0.08)	0.38(0.09)
0.9	6.02(0.55)	0.40(0.06)	0.33(0.06)	5.52(0.91)	0.45(0.09)	0.39(0.10)
1	5.78(0.68)	0.42(0.07)	0.36(0.08)	5.38(0.78)	0.46(0.08)	0.40(0.09)

$\eta$	Example 3			Example 4		
	TP	FDR	FNR	TP	FDR	FNR
0.1	6.72(0.67)	0.33(0.07)	0.25(0.07)	5.30(1.45)	0.47(0.14)	0.41(0.16)
0.2	6.34(0.56)	0.37(0.06)	0.30(0.06)	4.92(1.19)	0.51(0.12)	0.45(0.13)
0.3	6.48(0.58)	0.35(0.06)	0.28(0.06)	4.44(1.42)	0.56(0.14)	0.51(0.16)
0.4	6.26(0.53)	0.37(0.05)	0.30(0.06)	4.46(1.15)	0.55(0.11)	0.50(0.13)
0.5	6.26(0.44)	0.37(0.04)	0.30(0.05)	3.80(1.28)	0.62(0.13)	0.58(0.14)
0.6	6.28(0.54)	0.37(0.05)	0.30(0.06)	4.02(0.91)	0.60(0.09)	0.55(0.10)
0.7	6.16(0.58)	0.38(0.06)	0.32(0.06)	3.64(1.29)	0.64(0.13)	0.60(0.14)
0.8	6.12(0.72)	0.39(0.07)	0.32(0.08)	3.08(1.10)	0.69(0.11)	0.66(0.12)
0.9	5.42(0.76)	0.46(0.08)	0.40(0.08)	2.28(1.03)	0.77(0.10)	0.75(0.11)
1	3.96(0.81)	0.60(0.08)	0.56(0.09)	1.46(0.89)	0.85(0.09)	0.84(0.10)

<sup>a</sup> The optimal  $\eta$  for linear models (Example 1—3) is 0.3.

<sup>b</sup> The optimal  $\eta$  for logistic regression model (Example 4) is 0.1.

<sup>c</sup> When  $\eta = 1$ , SGL becomes group Lasso.

Table 6. Multi-split  $p$ -values for a simulated dataset in example 3. The true positives are from 705 to 708 and from 710 to 714.

Group index	Group Info.		Group LASSO		SGL	
	Start index	End index	$\ \hat{\beta}\ $	$p$ -value	$\ \hat{\beta}\ $	$p$ -value
150	462	463	0.008	1		
179	551	551	0.043	1		
634	2025	2025	0.016	1		
654	2120	2123	0.011	1		
664	2152	2155	0.002	1		
695	2269	2270			0.002	1
703	2283	2285	0.009	1	0.0005	1
704	2286	2286	0.072	1	0.018	1
705	2287	2290	0.291	8.2e-08	0.046	5.3e-08
706	2291	2296			0.026	9.4e-05
708	2299	2299			0.058	1
709	2300	2303	0.006	2.9e-06	0.196	9.3e-09
710	2304	2306	0.578	2.5e-09	0.176	8.9e-10
711	2307	2307	0.078	0.060	0.139	0.003
712	2308	2310			0.254	6.9e-10
713	2311	2312	0.544	3.1e-07	0.191	5.9e-08
714	2313	2318			0.222	5.5e-08
715	2319	2319			0.058	1
716	2320	2321			0.014	4.9e-04
773	2528	2531	0.028	1		
782	2558	2559	0.005	1		
1038	3462	3462	0.043	1	0.001	1

Table 7: SNPs selected on chromosome 6 by SGL with  $\eta = 0.1$  for rheumatoid arthritis data. The threshold to form groups is 0.6.

SNP name	Position	Band	Gene <sup>a</sup>	$L_2$ norm	$p$ -value
rs12203592	341321	6p25.3	IRF4	0.014	0.170
rs3804488	6587594	6p25.1	LY86	4.7E-04	1
rs3807031	30141863	6p21.33	PPP1R11	0.001	2.5E-03
rs9261535	30235302	6p21.33	TRIM10	0.000	2.6E-04
rs261946	30379313	6p21.33	LOC100133303	0.001	3.9E-05
rs1076832	30527810	6p21.33	MICC	1.8E-04	0.011
rs12665339	30709211	6p21.33	C6orf134	0.011	1.2E-04
rs1140809	30719655	6p21.33	C6orf134	0.013	2.0E-06
rs12190167	30874173	6p21.33	IER3	0.002	4.0E-05
rs12197154	30874223	6p21.33	IER3	0.002	4.0E-05
rs4713391	30900214	6p21.33	IER3	0.001	2.4E-05
rs12192704	30900249	6p21.33	IER3	0.001	2.4E-05
rs12198723	30904713	6p21.33	IER3	0.004	9.2E-05
rs9262632	31132787	6p21.33	LOC729792	0.009	0.773
rs2523849	31133030	6p21.33	LOC729792	0.009	0.001
rs2535311	31160663	6p21.33	HCG22	0.002	1
rs3778638	31200103	6p21.33	PSORS1C1	5.9E-05	1
rs1265098	31214156	6p21.33	PSORS1C2	0.001	5.7E-04
rs9295957	31265572	6p21.33	PSORS1C3	0.008	1.8E-07
rs2074488	31348410	6p21.33	HLA-C	0.011	3.4E-11
rs9461680	31351326	6p21.33	HLA-C	0.001	2.9E-06
rs2524123	31373293	6p21.33	HLA-C	0.008	1
rs2596501	31429190	6p21.33	HLA-B	0.002	1
rs2596542	31474574	6p21.33	DHFRP2	0.001	1
rs2256175	31488428	6p21.33	XXbac-BPG181B23.1	0.001	4.0E-05
rs9266825	31490861	6p21.33	XXbac-BPG181B23.1	0.001	4.0E-05
rs2524279	31500885	6p21.33	MICA	0.003	4.2E-08
rs2244839	31546347	6p21.33	HCP5	0.015	0.003
rs12660382	31551302	6p21.33	3.8-1	0.002	1
rs2844509	31618903	6p21.33	ATP6V1G2	0.018	8.7E-08
rs3093662	31652168	6p21.33	TNF	0.002	0.076
rs2844463	31723146	6p21.33	BAT3	0.001	0.001
rs3830076	32204222	6p21.32	FKBPL	0.003	1.0E-05
rs204999	32217957	6p21.32	FKBPL	0.003	6.2E-10
rs2070600	32259421	6p21.32	PBX2	0.004	1.3E-12
rs3830041	32299317	6p21.32	NOTCH4	0.002	0.007
rs375244	32299435	6p21.32	NOTCH4	0.009	1.1E-07
rs9267873	32307330	6p21.32	NOTCH4	0.007	<1.0E-18
rs3130299	32311515	6p21.32	NOTCH4	0.003	<1.0E-18
rs9267911	32313088	6p21.32	NOTCH4	0.003	<1.0E-18
rs382259	32317005	6p21.32	NOTCH4	0.002	<1.0E-18
rs507778	32317839	6p21.32	NOTCH4	0.002	<1.0E-18
rs412657	32319063	6p21.32	NOTCH4	0.002	<1.0E-18

Continued on next page

Table 7 – continued from previous page

SNP name	Position	Band	Gene <sup>a</sup>	$L_2$ norm	$p$ -value
rs17422797	32373506	6p21.32	C6orf10	0.012	0.005
rs6457536	32381743	6p21.32	C6orf10	0.017	2.2E-05
rs547261	32390011	6p21.32	C6orf10	0.059	<1.0E-18
rs6910071	32390832	6p21.32	C6orf10	0.059	<1.0E-18
rs498422	32394739	6p21.32	C6orf10	0.032	7.7E-06
rs547077	32397296	6p21.32	C6orf10	0.024	<1.0E-18
rs570963	32397572	6p21.32	C6orf10	0.022	2.4E-06
rs485774	32398932	6p21.32	C6orf10	0.016	<1.0E-18
rs9368713	32405315	6p21.32	C6orf10	0.016	<1.0E-18
rs10484560	32406115	6p21.32	C6orf10	0.017	0.015
rs3117099	32466248	6p21.32	C6orf10	0.001	9.7E-10
rs2076530	32471794	6p21.32	BTNL2	0.005	<1.0E-18
rs3793126	32479597	6p21.32	BTNL2	0.005	<1.0E-18
rs10947262	32481290	6p21.32	BTNL2	0.015	1.1E-06
rs3806156	32481676	6p21.32	BTNL2	0.028	<1.0E-18
rs3763309	32483951	6p21.32	HLA-DRA	0.028	<1.0E-18
rs3763312	32484326	6p21.32	HLA-DRA	0.028	<1.0E-18
rs6932542	32488240	6p21.32	HLA-DRA	0.003	<1.0E-18
rs2395163	32495787	6p21.32	HLA-DRA	0.020	<1.0E-18
rs3135363	32497626	6p21.32	HLA-DRA	0.004	<1.0E-18
rs3135353	32500855	6p21.32	HLA-DRA	0.004	<1.0E-18
rs3135338	32509195	6p21.32	HLA-DRA	0.005	<1.0E-18
rs2027856	32510683	6p21.32	HLA-DRA	0.008	2.6E-07
rs9268615	32510867	6p21.32	HLA-DRA	0.006	<1.0E-18
rs2395173	32512837	6p21.32	HLA-DRA	0.013	<1.0E-18
rs2395174	32512856	6p21.32	HLA-DRA	0.019	1
rs2395175	32513004	6p21.32	HLA-DRA	0.072	<1.0E-18
rs3129871	32514320	6p21.32	HLA-DRA	0.066	<1.0E-18
rs3129882	32517508	6p21.32	HLA-DRA	0.055	<1.0E-18
rs2239804	32519501	6p21.32	HLA-DRA	0.079	<1.0E-18
rs7192	32519624	6p21.32	HLA-DRA	0.079	<1.0E-18
rs2395182	32521295	6p21.32	HLA-DRA	0.079	<1.0E-18
rs3129890	32522251	6p21.32	HLA-DRA	0.082	<1.0E-18
rs9268832	32535767	6p21.32	HLA-DRA	0.082	<1.0E-18
rs6903608	32536263	6p21.32	HLA-DRB9	0.093	<1.0E-18
rs2395185	32541145	6p21.32	HLA-DRB9	0.114	<1.0E-18
rs477515	32677669	6p21.32	HLA-DRB1	0.114	<1.0E-18
rs2516049	32678378	6p21.32	HLA-DRB1	0.114	<1.0E-18
rs2858870	32680229	6p21.32	HLA-DRB1	0.120	1.4E-07
rs602875	32681607	6p21.32	HLA-DRB1	0.132	<1.0E-18
rs660895	32685358	6p21.32	HLA-DRB1	0.190	<1.0E-18
rs532098	32686030	6p21.32	HLA-DRB1	0.190	<1.0E-18
rs9271366	32694832	6p21.32	HLA-DRB1	0.081	2.5E-09
rs9271568	32698441	6p21.32	HLA-DRB1	0.060	<1.0E-18
rs17533090	32698700	6p21.32	HLA-DRB1	0.019	<1.0E-18

Continued on next page

Table 7 – continued from previous page

SNP name	Position	Band	Gene <sup>a</sup>	$L_2$ norm	$p$ -value
rs7775228	32766057	6p21.32	HLA-DQB1	0.004	4.5E-09
rs9275224	32767856	6p21.32	HLA-DQB1	0.005	<1.0E-18
rs5000634	32771542	6p21.32	HLA-DQB1	0.005	<1.0E-18
rs6457617	32771829	6p21.32	HLA-DQB1	0.005	<1.0E-18
rs2647012	32772436	6p21.32	HLA-DQB1	0.005	<1.0E-18
rs10484561	32773398	6p21.32	HLA-DQB1	0.009	0.002
rs2856725	32774716	6p21.32	HLA-DQB1	0.006	<1.0E-18
rs2858305	32778442	6p21.32	HLA-DQB1	0.023	<1.0E-18
rs9275425	32778852	6p21.32	HLA-DQB1	0.037	<1.0E-18
rs9275427	32778893	6p21.32	HLA-DQB1	0.037	<1.0E-18
rs2856705	32778934	6p21.32	HLA-DQB1	0.039	1.4E-09
rs9275428	32778956	6p21.32	HLA-DQB1	0.041	<1.0E-18
rs13192471	32779081	6p21.32	HLA-DQB1	0.041	<1.0E-18
rs1794275	32779226	6p21.32	HLA-DQB1	0.041	<1.0E-18
rs9275439	32779499	6p21.32	HLA-DQB1	0.029	<1.0E-18
rs9275555	32785066	6p21.32	HLA-DQB1	0.029	<1.0E-18
rs16898264	32785130	6p21.32	HLA-DQB1	0.012	1
rs9275572	32786977	6p21.32	HLA-DQB1	0.032	<1.0E-18
rs7765379	32788906	6p21.32	HLA-DQB1	0.050	<1.0E-18
rs7745656	32788948	6p21.32	HLA-DQB1	0.046	<1.0E-18
rs2858332	32789139	6p21.32	HLA-DQB1	0.046	<1.0E-18
rs2858331	32789255	6p21.32	HLA-DQB1	0.046	<1.0E-18
rs9275595	32789333	6p21.32	HLA-DQB1	0.036	<1.0E-18
rs9275596	32789609	6p21.32	HLA-DQB1	0.019	<1.0E-18
rs9275601	32790642	6p21.32	HLA-DQB1	0.039	1.3E-07
rs3873444	32790702	6p21.32	HLA-DQB1	0.044	8.3E-12
rs9275602	32790790	6p21.32	HLA-DQB1	0.008	0.123
rs10484565	32903010	6p21.32	TAP2	0.008	2.4E-07
rs3819721	32912776	6p21.32	TAP2	0.009	<1.0E-18
rs241425	32912887	6p21.32	TAP2	0.003	9.7E-10
rs241424	32912912	6p21.32	TAP2	0.003	9.7E-10
rs9276831	32940011	6p21.32	PPP1R2P1	0.006	4.1E-05
rs9500927	33069339	6p21.32	HLA-DOA	0.001	3.3E-04
rs3130215	33182941	6p21.32	HLA-DPB2	0.004	0.021
rs986521	33244123	6p21.32	COL11A2	0.008	0.105
rs439205	33281820	6p21.32	HSD17B8	0.003	1
rs421446	33282761	6p21.32	HSD17B8	0.003	1
rs10948693	52261958	6p12.2	LOC647163	0.014	0.892
rs11964270	107589649	6q21	PDSS2	0.000	1

<sup>a</sup> Gene names that SNPs belong to or are closest to.

Table 8. SNPs selected by SGL with  $\eta = 0.1$  but not by single-SNP analysis for rheumatoid arthritis data.

SNP name	Position	Band	Gene <sup>a</sup>	$L_2$ norm	$p$ -value
rs12203592	341321	6p25.3	IRF4	0.014	0.170
rs3804488	6587594	6p25.1	LY86	4.7E-04	1
rs9262632	31132787	6p21.33	LOC729792	0.009	0.773
rs2535311	31160663	6p21.33	HCG22	0.002	1
rs3778638	31200103	6p21.33	PSORS1C1	5.9E-05	1
rs2524123	31373293	6p21.33	HLA-C	0.008	1
rs2596501	31429190	6p21.33	HLA-B	0.002	1
rs2596542	31474574	6p21.33	DHFRP2	0.001	1
rs9266825	31490861	6p21.33	XXbac-BPG181B23.1	0.001	4.0E-05
rs12660382	31551302	6p21.33	3.8-1	0.002	1
rs2395174	32512856	6p21.32	HLA-DRA	0.019	1
rs17533090	32698700	6p21.32	HLA-DRB1	0.019	<1.0E-18
rs13192471	32779081	6p21.32	HLA-DQB1	0.041	<1.0E-18
rs1794275	32779226	6p21.32	HLA-DQB1	0.041	<1.0E-18
rs16898264	32785130	6p21.32	HLA-DQB1	0.012	1
rs2858332	32789139	6p21.32	HLA-DQB1	0.046	<1.0E-18
rs241425	32912887	6p21.32	TAP2	0.003	9.7E-10
rs241424	32912912	6p21.32	TAP2	0.003	9.7E-10
rs439205	33281820	6p21.32	HSD17B8	0.003	1
rs10948693	52261958	6p12.2	LOC647163	0.014	0.892
rs11964270	107589649	6q21	PDSS2	3.5E-05	1

<sup>a</sup> Gene names that SNPs belong to or are closest to.

Table 9. SNPs selected by SGL with  $\eta = 0.3$  under different thresholds for GAW 17 data.

SNP name	Chromosome	Position	Gene	Estimate	Estimate
				Threshold = 0.01	Threshold = 0.05
C1S8606	1	177344038	ABL2	1.5E-03	
C2S2355	2	112864155	RGPD8		0.002
C2S2527	2	120155094	TMEM177	8.7E-12	3.6E-14
C2S5989	2	219310862	TTLL4	8.4E-13	2.3E-15
C4S207	4	3107824	HTT	4.0E-13	7.0E-16
C5S4668	5	175944518	PCLKC	4.9E-14	7.8E-17
C6S1136	6	28377928	PGBD1	8.2E-15	
C11S6685	11	124128967	ESAM	4.0E-14	
C13S348	13	27789648	FLT1		9.9E-12
C13S522	13	27899910	FLT1	0.088	0.108
C13S523	13	27899912	FLT1	0.088	0.108
C13S524	13	27899915	FLT1	0.088	0.108
C14S2070	14	72707354	PSEN1	9.5E-17	2.0E-11
C16S1046	16	16152087	ABCC6		3.0E-12
C17S3993	17	54047655	TEX14	4.0E-12	8.3E-13