Appendix

Remarks on the convergence property

The difficulty of showing the convergence property lies in the fact that the optimization problem expands over iterations. We provide *ad hoc* discussions here. The convergence property relies on both the implemented algorithm at each iteration (to converge to the true minimizer) and also the sign consistency of the penalized parameter estimates (so that the optimization problem can be fully and correctly expanded). Note that the argument may vary depending on the optimization method and the form of the penalty term. However, the strategy given in the following steps should be generically applicable.

(I) At any iteration of the proposed algorithm, the optimization in (8) is a convex problem defined on a convex domain. Hence the target minimizer of (8) in any iteration is unique.

(II) The optimization domain cannot remain unchanged for infinite many iterations unless all true main effects are activated.

(III) All true nonzero main effects must be contained in the active set within finite iterations. That is, $\{j : \beta_j \neq 0\} \subseteq \mathcal{A}_m = \{j : \beta_j^{(m)} \neq 0\}$ for some m large enough.

The convexity in Argument (I) is trivial, and the uniqueness property can be found in recent literature such as Tibshirani [2013], with the assumption that the predictor variables ($\mathbf{X}$ and $\mathbf{Z}$) are drawn from a continuous probability distribution. When the predictor variables are discrete, theoretical results are largely unknown. Our simulation suggests that when the predictor variables are continuous, the algorithm performs better.

Argument (II) is a crucial step, which relies both on the selection consistency of the penalization problem [Zhao and Yu, 2006; Zhang and Huang, 2008] and the finite-step convergence of the optimization algorithm [Saha and Tewari, 2010; Li and Osher, 2009]. Repeatedly applying argument

(II) allows new main effects to be activated, which correctly expands the optimization problem to the full domain. When the activated main effect set $\mathcal{A}_m$ contains all of the true nonzero parameters, the optimization problem becomes very similar to the original Lasso problem since all of the true nonzero main and interaction effects are being considered in the optimization domain. This leads to the overall consistency property of the proposed method.

Additional simulation results

Table 5: $G \times E$ Model with binary X

Scenario 1: 15 main effects, 5 interactions.

n	p	Method	pAUC (sd)	TP20 (sd)	TP40 (sd)	TP.best(sd)	FP.best(sd)	MSE.best(sd)
150	500	Proposed	0.66 (0.09)	9.94 (1.73)	12.90 (2.02)	15.07 (2.36)	65.91 (22.70)	2.74 (0.42)
		All-pairwise	0.57 (0.10)	8.89 (1.75)	11.36 (1.96)	13.15 (2.35)	65.49 (20.23)	2.90 (0.43)
		MT	0.31 (0.08)	4.68 (1.66)	6.29 (1.93)	-	-	-
		MT+	0.37 (0.08)	5.48 (1.89)	7.38 (1.84)	-	-	-
150	1000	Proposed	0.56 (0.10)	8.47 (1.99)	11.04 (2.09)	13.13 (2.61)	69.37 (26.16)	3.08 (0.47)
		All-pairwise	0.49 (0.09)	7.53 (1.98)	9.68 (2.15)	11.45 (2.55)	70.42 (28.93)	3.21 (0.47)
		MT	0.24 (0.07)	3.79 (1.68)	4.84 (1.72)	-	-	-
		MT+	0.30 (0.08)	4.55 (1.64)	5.98 (1.70)	-	-	-
250	500	Proposed	0.88 (0.06)	14.34 (1.65)	17.71 (1.73)	19.03 (1.00)	89.33 (27.25)	1.81 (0.22)
		All-pairwise	0.84 (0.07)	13.71 (1.64)	16.83 (1.54)	18.28 (1.43)	99.15 (32.76)	1.93 (0.24)
		MT	0.52 (0.09)	8.55 (1.88)	10.33 (1.83)	-	-	-
		MT+	0.54 (0.07)	8.10 (1.54)	10.85 (1.59)	-	-	-
250	1000	Proposed	0.83 (0.07)	13.06 (1.73)	16.39 (1.90)	18.60 (1.39)	97.79 (27.23)	2.06 (0.29)
		All-pairwise	0.79 (0.08)	12.53 (1.72)	15.50 (1.74)	17.87 (1.37)	106.29 (28.00)	2.17 (0.28)
		MT	0.45 (0.08)	7.36 (1.71)	8.92 (1.75)	-	-	-
		MT+	0.48 (0.07)	7.41 (1.51)	9.69 (1.47)	-	-	-

Scenario 2: 10 main effects, 10 interactions.

n	p	Method	pAUC (sd)	TP20 (sd)	TP40 (sd)	TP.best(sd)	FP.best(sd)	MSE.best(sd)
150	500	Proposed	0.59 (0.11)	8.90 (2.31)	11.64 (2.42)	13.69 (2.81)	63.01 (25.56)	2.84 (0.49)
		All-pairwise	0.52 (0.11)	8.15 (2.20)	10.31 (2.47)	12.03 (2.62)	65.60 (24.80)	3.01 (0.49)
		MT	0.30 (0.08)	4.83 (1.64)	5.93 (1.72)	-	-	-
		MT+	0.41 (0.10)	6.42 (1.93)	8.11 (2.17)	-	-	-
150	1000	Proposed	0.51 (0.13)	8.10 (2.26)	10.25 (2.64)	11.64 (3.25)	61.06 (28.70)	3.05 (0.54)
		All-pairwise	0.46 (0.11)	7.28 (1.85)	9.22 (2.18)	10.27 (2.66)	58.69 (25.72)	3.21 (0.53)
		MT	0.27 (0.07)	4.18 (1.59)	5.42 (1.66)	-	-	-
		MT+	0.38 (0.10)	5.98 (1.86)	7.63 (2.18)	-	-	-
250	500	Proposed	0.83 (0.08)	13.39 (1.93)	16.39 (1.93)	18.38 (1.36)	87.16 (30.83)	1.90 (0.29)
		All-pairwise	0.79 (0.08)	12.92 (1.82)	15.71 (1.79)	17.57 (1.46)	93.21 (23.90)	2.04 (0.30)
		MT	0.50 (0.08)	8.46 (1.67)	9.75 (1.85)	-	-	-
		MT+	0.61 (0.08)	9.45 (1.60)	12.03 (1.95)	-	-	-
250	1000	Proposed	0.78 (0.09)	12.37 (2.03)	15.44 (1.98)	17.66 (1.58)	88.65 (28.75)	2.05 (0.27)
		All-pairwise	0.75 (0.08)	11.89 (1.92)	14.63 (1.99)	16.86 (1.71)	94.87 (27.79)	2.17 (0.29)
		MT	0.44 (0.08)	7.55 (1.74)	8.88 (1.82)	-	-	-
		MT+	0.58 (0.09)	9.12 (1.83)	11.65 (1.90)	-	-	-

The pAUC is calculated up to 80 false positive nonzero parameters. TP20 and TP40 denote the true positives under model size 20 and 40, respectively. TP.best, FP.best, and MSE.best correspond to the model that yields the best prediction MSE.

Table 6: $G \times G$ Model with continuous X

Scenario 1: 15 main effects, 5 interactions.

n	p	Method	pAUC (sd)	TP20 (sd)	TP40 (sd)	TP.best(sd)	FP.best(sd)	MSE.best(sd)
150	500	Proposed	0.77 (0.13)	11.65 (2.48)	15.31 (3.04)	17.38 (2.94)	83.81 (32.31)	5.80 (2.12)
		All-pairwise	0.20 (0.08)	3.10 (1.44)	4.10 (1.81)	4.12 (2.62)	44.64 (34.32)	12.07 (1.52)
		hierNet	0.66 (0.10)	9.20 (1.80)	12.54 (2.40)	18.31 (1.57)	188.94 (26.68)	5.75 (1.39)
150	1000	Proposed	0.68 (0.14)	10.19 (2.33)	13.94 (3.11)	16.09 (3.12)	88.32 (31.84)	6.95 (2.22)
		All-pairwise	0.11 (0.05)	1.59 (1.01)	2.40 (1.22)	2.02 (1.89)	29.53 (29.44)	12.57 (1.79)
		hierNet	0.56 (0.09)	7.87 (1.69)	10.77 (1.77)	16.49 (2.14)	196.77 (40.82)	7.16 (1.54)
250	500	Proposed	0.97 (0.04)	17.48 (1.72)	19.82 (0.78)	19.97 (0.30)	87.53 (41.87)	1.93 (0.44)
		All-pairwise	0.56 (0.11)	8.60 (1.97)	11.10 (2.28)	14.75 (2.35)	169.16 (51.40)	8.15 (1.56)
		hierNet	0.94 (0.03)	14.99 (1.50)	19.23 (1.01)	19.99 (0.10)	237.68 (56.47)	2.63 (0.43)
250	1000	Proposed	0.94 (0.07)	16.18 (1.98)	19.36 (1.63)	19.68 (1.25)	99.52 (41.05)	2.37 (0.98)
		All-pairwise	0.39 (0.11)	5.79 (2.24)	7.80 (2.32)	10.28 (3.01)	119.52 (57.00)	9.86 (1.16)
		hierNet	0.87 (0.06)	13.51 (1.75)	17.30 (1.71)	19.92 (0.28)	273.60 (52.43)	3.25 (0.55)

Scenario 2: 10 main effects, 10 interactions.

n	p	Method	pAUC (sd)	TP20 (sd)	TP40 (sd)	TP.best(sd)	FP.best(sd)	MSE.best(sd)
150	500	Proposed	0.67 (0.20)	10.05 (3.28)	13.24 (4.19)	15.66 (4.51)	76.58 (36.40)	6.55 (2.57)
		All-pairwise	0.20 (0.09)	3.14 (1.59)	4.07 (1.90)	4.37 (2.48)	44.49 (30.80)	11.57 (1.69)
		hierNet	0.59 (0.11)	9.06 (2.15)	11.71 (2.44)	14.47 (2.33)	128.09 (40.49)	7.59 (1.75)
150	1000	Proposed	0.60 (0.17)	8.36 (2.73)	11.91 (3.58)	14.57 (4.48)	78.63 (39.02)	7.96 (2.22)
		All-pairwise	0.12 (0.07)	1.73 (1.22)	2.44 (1.44)	2.20 (2.07)	30.27 (37.91)	12.37 (1.54)
		hierNet	0.50 (0.11)	7.48 (2.26)	9.92 (2.36)	12.70 (2.20)	134.73 (46.86)	9.10 (1.55)
250	500	Proposed	0.92 (0.11)	15.64 (3.12)	18.84 (2.63)	19.57 (1.74)	88.00 (45.77)	2.41 (1.26)
		All-pairwise	0.54 (0.12)	8.28 (2.19)	10.48 (2.29)	14.3 (2.68)	158.10 (44.59)	8.24 (1.67)
		hierNet	0.88 (0.06)	14.56 (1.88)	17.72 (1.60)	19.18 (0.81)	163.06 (45.39)	4.75 (1.06)
250	1000	Proposed	0.88 (0.14)	13.99 (3.76)	18.20 (3.14)	19.20 (2.73)	112.05 (54.74)	3.06 (1.93)
		All-pairwise	0.36 (0.12)	5.61 (1.97)	7.16 (2.31)	9.35 (3.73)	117.30 (57.15)	10.45 (1.29)
		hierNet	0.82 (0.09)	13.00 (2.32)	16.24 (1.87)	18.65 (1.04)	192.50 (39.67)	5.26 (1.15)

The pAUC is calculated up to 80 false positive nonzero parameters. TP20 and TP40 denote the true positives under model size 20 and 40, respectively. TP.best, FP.best, and MSE.best correspond to the model that yields the best prediction MSE.

Table 7: Analysis of lung cancer data: identified main and interaction gene effects (p-values) using MT.

		Interactions					
Gene	Main Effects	Age	Gender	Smoke	Chemo	Stage I	Stage II
RPS4Y1	3.91E-07	1.96E-07	1.34E-07				
CPS1	9.15E-06						
CPS1	2.26E-06	5.88E-06					
DUSP4	7.78E-06	7.61E-06					
PSPH		8.30E-06					
DDX3Y		9.12E-06	5.41E-08				
TESC		1.48E-05					
UBD			3.04E-06				
ID1	5.90E-06	6.02E-06					
PBK	1.15E-07	1.22E-07					
SLC1A1	9.92E-06	9.16E-06	6.20E-06				
JARID1D	9.69E-07	6.22E-07	2.59E-08				
GCNT3			9.16E-06				
PFN2	1.40E-05	9.17E-06					
KIAA0101	5.57E-10	8.51E-10	7.09E-08				
GINS1	3.29E-06	3.57E-06					
FZD10		1.39E-05					
HIST1H2BC	9.81E-06	3.33E-06					
IRS2	6.83E-06	2.25E-06	6.75E-06				
EIF1AY	4.66E-07	2.67E-07	9.04E-09				
IRS2	4.50E-06	2.37E-06	5.27E-06				
STS	1.95E-06	9.53E-07	1.45E-05				
PRSS21	2.00E-07	1.37E-07	3.68E-07				
SAA1			8.43E-06				
HEATR1	1.10E-05						
AIDA	6.12E-07	3.91E-07					
SOAT1	1.94E-05	6.82E-06					
PLEKHB1			9.38E-06				
NPTX2	2.90E-06	1.71E-06	4.37E-06				
TMEM106B		1.59E-05	9.76E-06				
KDR	1.11E-05	4.17E-06					
RHOBTB3			5.97E-07				
WARS			1.28E-05				
EFHD1	8.98E-06	2.03E-06	2.61E-06				
C1orf103	3.56E-06	3.40E-06	1.62E-06				
CXADR			3.41E-07				
ENDOD1	5.35E-09	4.84E-09	1.37E-07				
IFITM1			4.56E-06				
PCNA	6.60E-08	3.78E-08	3.00E-06				
CLEC5A	1.47E-05	1.78E-05	9.48E-06				
RABGAP1L		5.18E-06	1.99E-06				
NANOG			1.16E-05				
WARS			1.05E-05				
ATP5I			1.00E-05				
NBN	4.76E-06						

Table 8: Analysis of lung cancer data: identified main and interaction gene effects (p-values) using MT+.

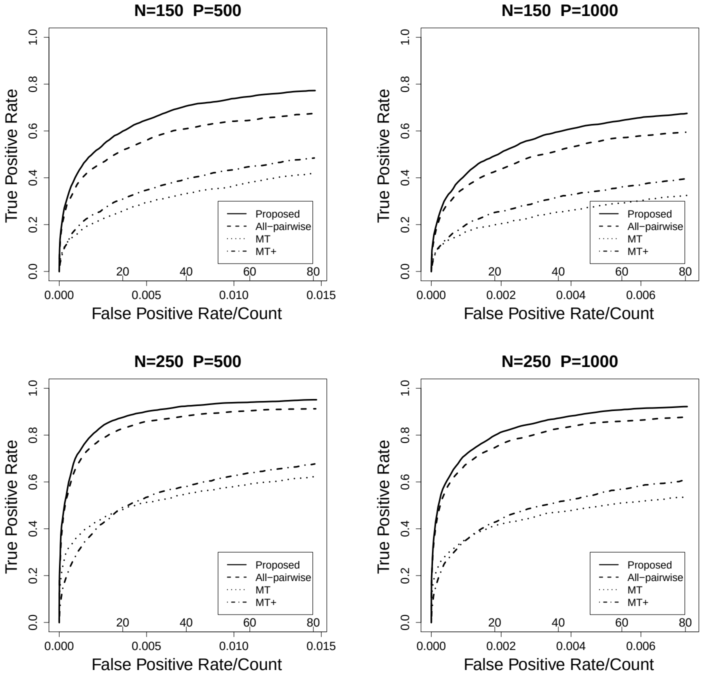
		Interactions					
Gene	Main Effects	Age	Gender	Smoke	Chemo	Stage I	Stage II
M10098_3_at				5.40E-05			
M10098_M_at				1.25E-04			
M10098				5.99E-05			
TFF1	8.02E-05						
RPS4Y1	3.91E-07						
CPS1	9.15E-06						
CPS1	2.26E-06						
FGA		4.76E-06					
DUSP4	7.78E-06						
GPR116	2.60E-05						
PSPH	6.50E-05	9.34E-05					
DDX3Y	2.14E-05						
GPX2	2.78E-05						
TESC	2.00E-05						
UBD	1.13E-04						
DKK1	1.21E-04						
APOD	1.02E-04						
QPCT					4.14E-05		
KRT17	7.65E-05						
HBA2		3.43E-05					
S100A7A					4.89E-05		
ID1	5.90E-06						
GPR116	4.88E-05						
AUTS2		6.65E-05					
HBA2		1.11E-04					
PBK	1.15E-07						
IFT57	4.21E-05						
SLC1A1	9.92E-06						
JARID1D	9.69E-07						
TM4SF1					1.18E-04		
PFN2	1.40E-05						
TUBB2A	1.17E-04						
KIAA0101	5.57E-10						
GINS1	3.29E-06						
GALNT6		7.76E-05					
FZD10	1.95E-05						
HIST1H2BC	9.81E-06						
IRS2	6.83E-06						
EIF1AY	4.66E-07						
SERPINE1		1.14E-04					
TM4SF1					9.11E-05		
IRS2	4.50E-06						
PAH							6.93E-05
CPD	6.04E-05						
STS	1.95E-06						
PRSS21	2.00E-07						
HEATR1	1.10E-05						

Table 9: Analysis of lung cancer data: identified main and interaction gene effects (p-values) using MT+ (continued).

		Interactions					
Gene	Main Effects	Age	Gender	Smoke	Chemo	Stage I	Stage II
LEF1		8.42E-05					
PLA2G16	6.97E-05						
AIDA	6.12E-07						
PTX3		1.45E-05					
CDC2	1.13E-04						
RBPMS	4.74E-05						
SOAT1	1.94E-05						
PLEKHB1	5.10E-05						
GADD45A				1.07E-04			
NPTX2	2.90E-06						
NR2F1		6.89E-06					
KLF5	9.79E-05						
TMEM106B	4.23E-05						
KDR	1.11E-05						
HSD17B2					6.16E-05		
RHOBTB3	3.09E-05						
TACSTD1	1.20E-04						
WARS	3.07E-05						
EFHD1	8.98E-06						
VGLL1	1.05E-04						
C1orf103	3.56E-06						
CXADR	2.20E-05						
CDH1	9.25E-05						
LBH		8.72E-05					
CRY1				2.14E-05			
ADCY9	1.02E-04						
ENDOD1	5.35E-09						
IFITM1	9.35E-05						
PCNA	6.60E-08						
CLEC5A	1.47E-05						
MCM4	3.56E-05						
ZNHIT6	3.20E-05						
MET		2.69E-06					
RABGAP1L	2.05E-05						
MTMR11							1.35E-05
WARS	4.00E-05						
FOSL1	7.48E-05						
ATP5I			1.11E-04				
NBN	6.79E-05						
NBN	4.76E-06						

Figure 3: $G \times E$ Model with binary X

Scenario 1



Scenario 2

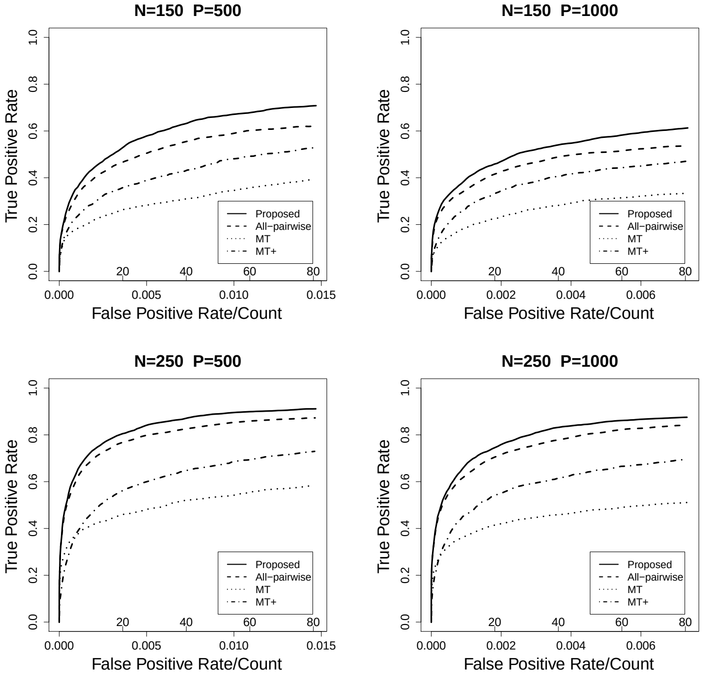
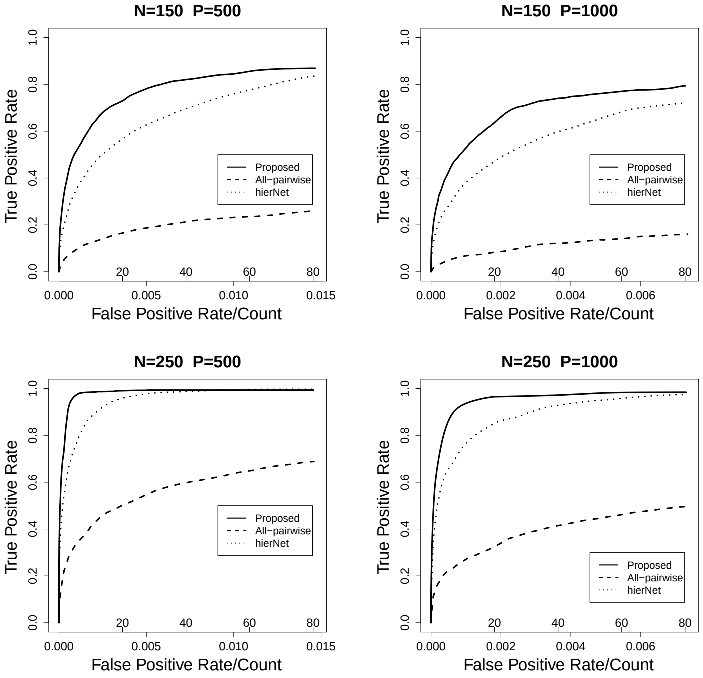


Figure 4: $G \times G$ Model with continuous X

Scenario 1



Scenario 2