

# Supplementary Material:

## Interaction Screening for Ultra-High Dimensional Data

### 1 Additional Simulation Examples

We present additional examples to illustrate the performance of the iFOR methods for interaction selection. We consider six procedures: forward-based joint analysis (FS2), the proposed forward-based methods iFORT, iFORM, iFORT-w, iFORM-w, and two interaction selection methods iMART1 and iMART2 which do not use forward selection. Various correlation structures are considered, including the compound symmetry (CS) and the complex correlation case from Fan and Song (2010), to be referred to as FS2010.

*Example 7:* (High dimensional: CS) All predictors are equally correlated with each other as  $\text{Corr}(X_j, X_k) = 0.5$  for any  $j \neq k$ . The rest model setup is the same as Example 3.

*Example 8:* (High dimensional: FS2010) We consider the covariance structure setup in Fan and Song (2010). First, we generate  $X_j, j = 1, \dots, 50$  independently from the standard normal distribution. Then, we define

$$X_k = \sum_{j=1}^s X_j (-1)^{j+1} / 5 + \sqrt{25 - s} / 5 \epsilon_k, \quad k = p - 50, \dots, p,$$

with  $s = 10$  and  $\{\epsilon_k\}_{k=p-49}^{50}$  follow  $N(0, 1)$ . The rest setup is the same as Example 3.

*Example 9:* (Weak Heredity: CS) All the predictors are equally correlated with  $\text{Corr}(X_j, X_k) = 0.5$  for any  $j \neq k$ . The rest setup is the same as Example 6.

For each example, we run  $M = 100$  Monte Carlo simulations and report the average performance of each procedure in terms of main effect selection (Cov, Cor0, Inc0, Ext), interaction selection (iCov, iCor0, iInc0, iExt), model size, mean squared error (MSE), and the out-of-sample  $R^2$ ; see their definitions in the main paper. Recall that  $n$  is the sample size,  $p$  is the number of covariates,  $p_0$  is the number of important main effects, and  $q_0$  is the number of important interaction effects. The results are summarized in Tables S1-S3.

Table S1: Results of Example 7,  $(n, p, p_0, q_0) = (400, 5000, 10, 10)$ , CS correlation.

	Linear Term Selection				Interaction Selection				Size and Prediction			
	Cov	Cor0	Inc0	Ext	iCov	iCor0	iInc0	Iext	size	MSE	Rsq	sdR
$\sigma = 2$												
iMART1	0.00	1.00	0.74	0.00	0.00	1.00	0.94	0.00	23.51	10.41	81.56	0.53
iMART2	0.00	1.00	0.73	0.00	0.00	1.00	0.97	0.00	14.75	10.70	78.43	0.55
iFORT-w	0.00	1.00	0.76	0.00	0.00	1.00	0.86	0.00	9.64	11.84	81.00	0.69
iFORT	0.00	1.00	0.76	0.00	0.00	1.00	0.91	0.00	9.85	11.57	77.09	0.83
iFORM-w	1.00	1.00	0.00	0.59	0.56	1.00	0.07	0.02	24.33	1.36	98.76	0.04
iFORM	1.00	1.00	0.00	0.30	0.93	1.00	0.01	0.05	25.04	1.00	98.85	0.04
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	0.84	98.87	0.04
$\sigma = 3$												
iMART1	0.00	1.00	0.75	0.00	0.00	1.00	0.94	0.00	22.71	10.60	80.20	0.52
iMART2	0.00	1.00	0.74	0.00	0.00	1.00	0.97	0.00	14.77	10.81	77.41	0.53
iFORT-w	0.00	1.00	0.77	0.00	0.00	1.00	0.88	0.00	9.27	12.15	79.34	0.68
iFORT	0.00	1.00	0.77	0.00	0.00	1.00	0.92	0.00	9.63	11.79	75.46	0.84
iFORM-w	0.91	1.00	0.01	0.91	0.00	1.00	0.35	0.00	22.61	3.18	96.91	0.12
iFORM	0.96	1.00	0.03	0.95	0.02	1.00	0.23	0.00	23.18	2.59	96.68	0.46
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	1.08	97.83	0.05
$\sigma = 4$												
iMART1	0.00	1.00	0.75	0.00	0.00	1.00	0.95	0.00	22.73	10.70	78.78	0.53
iMART2	0.00	1.00	0.75	0.00	0.00	1.00	0.97	0.00	14.67	10.88	76.08	0.54
iFORT-w	0.00	1.00	0.78	0.00	0.00	1.00	0.90	0.00	8.87	12.41	77.32	0.69
iFORT	0.00	1.00	0.78	0.00	0.00	1.00	0.92	0.00	9.21	11.95	73.85	0.80
iFORM-w	0.19	1.00	0.18	0.19	0.00	1.00	0.68	0.00	17.48	6.19	92.51	0.30
iFORM	0.80	1.00	0.08	0.80	0.00	1.00	0.40	0.00	21.43	4.03	93.83	0.63
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	1.35	96.41	0.06

Table S2: Results of Example 8,  $(n, p, p_0, q_0) = (400, 5000, 10, 10)$ , FS2010 correlation

	Linear Term Selection				Interaction Selection				Size and Prediction			
	Cov	Cor0	Inc0	Ext	iCov	iCor0	iInc0	Iext	size	MSE	Rsqr	sdR
$\sigma = 2$												
iMART1	0.40	1.00	0.08	0.39	0.04	1.00	0.15	0.31	19.78	2.09	89.57	0.67
iMART2	0.40	1.00	0.08	0.40	0.00	1.00	0.38	0.00	16.36	2.82	86.49	0.58
iFORT-w	0.78	1.00	0.05	0.70	0.81	1.00	0.06	0.79	19.32	1.47	92.55	0.57
iFORT	0.78	1.00	0.05	0.70	0.78	1.00	0.08	0.33	20.00	1.55	92.10	0.68
iFORM-w	1.00	1.00	0.00	0.82	1.00	1.00	0.00	0.93	20.63	0.71	95.19	0.06
iFORM	0.99	1.00	0.00	0.82	0.99	1.00	0.01	0.99	20.37	0.78	94.86	0.35
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	0.68	95.21	0.06
$\sigma = 3$												
iMART1	0.35	1.00	0.09	0.25	0.03	1.00	0.18	0.13	20.44	2.47	83.77	0.66
iMART2	0.35	1.00	0.09	0.35	0.00	1.00	0.39	0.00	17.22	3.06	80.62	0.61
iFORT-w	0.64	1.00	0.08	0.64	0.24	1.00	0.23	0.22	17.45	2.45	84.78	0.66
iFORT	0.64	1.00	0.08	0.62	0.64	1.00	0.13	0.26	19.28	2.22	84.98	0.86
iFORM-w	0.94	1.00	0.02	0.84	0.32	1.00	0.16	0.29	19.00	1.71	87.84	0.41
iFORM	0.94	1.00	0.02	0.77	0.65	1.00	0.09	0.64	19.73	1.48	88.33	0.59
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	0.84	90.12	0.12
$\sigma = 4$												
iMART1	0.25	1.00	0.11	0.09	0.02	1.00	0.25	0.02	20.95	3.06	75.80	0.66
iMART2	0.25	1.00	0.11	0.25	0.00	1.00	0.42	0.00	17.69	3.45	72.74	0.70
iFORT-w	0.37	1.00	0.13	0.37	0.00	1.00	0.46	0.00	14.60	3.55	74.43	0.73
iFORT	0.37	1.00	0.13	0.34	0.37	1.00	0.23	0.15	17.67	3.01	76.11	0.94
iFORM-w	0.82	1.00	0.06	0.82	0.02	1.00	0.44	0.02	15.68	3.03	77.36	0.66
iFORM	0.83	1.00	0.06	0.83	0.03	1.00	0.38	0.03	16.31	2.84	77.91	0.84
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	1.03	83.81	0.18

Table S3: Results of Example 9 for the weak heredity,  $(n, p, p_0, q_0) = (400, 5000, 10, 10)$ , CS.

	Linear Term Selection				Interaction Selection				Size and Prediction			
	Cov	Cor0	Inc0	Ext	iCov	iCor0	iInc0	Iext	size	MSE	Rsqr	sdR
$\sigma = 2$												
iMART1	0.00	1.00	0.72	0.00	0.00	1.00	0.95	0.00	23.05	10.41	81.84	0.91
iMART2	0.00	1.00	0.72	0.00	0.00	1.00	0.98	0.00	14.52	10.71	78.89	0.48
iFORT-w	0.00	1.00	0.76	0.00	0.00	1.00	0.89	0.00	9.65	11.91	80.27	0.65
iFORT	0.00	1.00	0.76	0.00	0.00	1.00	0.95	0.00	9.90	11.67	76.29	0.73
iFORM-w	1.00	1.00	0.00	0.51	0.64	1.00	0.05	0.01	24.63	1.23	98.78	0.04
iFORM	0.98	1.00	0.01	0.98	0.00	1.00	0.53	0.00	21.30	4.51	96.50	0.25
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	0.86	98.86	0.04
$\sigma = 3$												
iMART1	0.00	1.00	0.73	0.00	0.00	1.00	0.96	0.00	23.28	10.47	80.99	0.46
iMART2	0.00	1.00	0.73	0.00	0.00	1.00	0.98	0.00	14.15	10.78	77.63	0.51
iFORT-w	0.00	1.00	0.77	0.00	0.00	1.00	0.91	0.00	9.26	12.22	78.30	0.67
iFORT	0.00	1.00	0.77	0.00	0.00	1.00	0.95	0.00	9.55	11.84	74.63	0.77
iFORM-w	0.94	1.00	0.01	0.94	0.00	1.00	0.35	0.00	22.95	3.08	96.94	0.12
iFORM	0.76	1.00	0.07	0.76	0.00	1.00	0.61	0.00	20.14	5.27	94.10	0.55
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	1.10	97.80	0.05
$\sigma = 4$												
iMART1	0.00	1.00	0.74	0.00	0.00	1.00	0.96	0.00	22.57	10.60	79.52	0.48
iMART2	0.00	1.00	0.73	0.00	0.00	1.00	0.98	0.00	14.59	10.80	76.51	0.50
iFORT-w	0.00	1.00	0.78	0.00	0.00	1.00	0.92	0.00	9.00	12.43	76.45	0.72
iFORT	0.00	1.00	0.78	0.00	0.00	1.00	0.95	0.00	9.30	12.00	73.09	0.75
iFORM-w	0.22	1.00	0.19	0.22	0.00	1.00	0.68	0.00	18.00	6.20	92.03	0.35
iFORM	0.37	1.00	0.17	0.37	0.00	1.00	0.69	0.00	17.96	6.31	90.93	0.63
ORACL	1.00	1.00	0.00	1.00	1.00	1.00	0.00	1.00	20.00	1.36	96.37	0.06

The results in the above tables are quite consistent to those in the main paper. In Examples 7 and 8, where the strong heredity condition holds, the iFORM performs the best in terms of both model selection and prediction performance. The iFORM method has the smallest MSE, the highest out-of-the sample  $R^2$ , and the largest exact coverage probability for main effects and interactions. iMART1 and iMART2 do not perform very well due to complex covariance structure among covariates.

In Example 9, where the weak heredity condition holds, we expect that the iFOR methods under the weak heredity constraint perform better than those under the strong heredity. This is confirmed by the results in Table S3. It is observed that the iFORM-w (or iFORT-w) gives better variable selection and model estimation results than the corresponding method under the strong heredity such as iFORM (or iFORT).