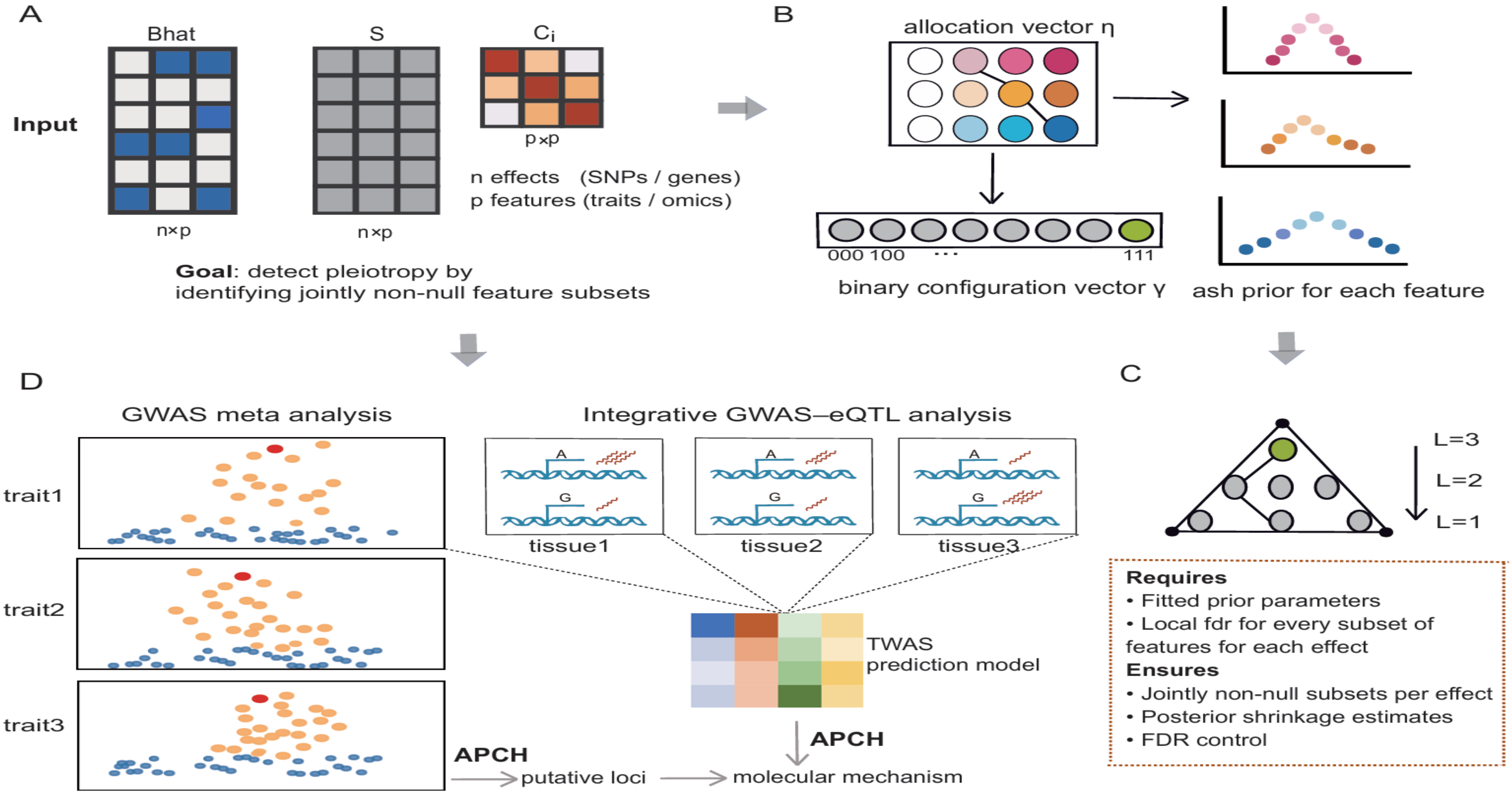
Adaptive Partial Conjunction Hypothesis for Identifying Pleiotropy Across Heterogeneous Effect Units

Yuxin Li[1], Zicheng Lu[1], and Xiaolei Lin*[1]

[1]School of Data Science, Fudan University, Shanghai, China


A
Input
Bhat
n×p
S
n×p
Ci
p×p
n effects (SNPs / genes)
p features (traits / omics)
Goal: detect pleiotropy by identifying jointly non-null feature subsets
B
allocation vector η
000 100 ··· 111
binary configuration vector γ
ash prior for each feature
C
L=3
L=2
L=1
Requires
• Fitted prior parameters
• Local fdr for every subset of features for each effect
Ensures
• Jointly non-null subsets per effect
• Posterior shrinkage estimates
• FDR control
D
GWAS meta analysis
trait1
trait2
trait3
Integrative GWAS–eQTL analysis
A
G
tissue1
tissue2
tissue3
TWAS prediction model
APCH
APCH
putative loci
molecular mechanism

Two kinds of correlation in joint-effect modeling
Biology/ true effect vs Error / residual

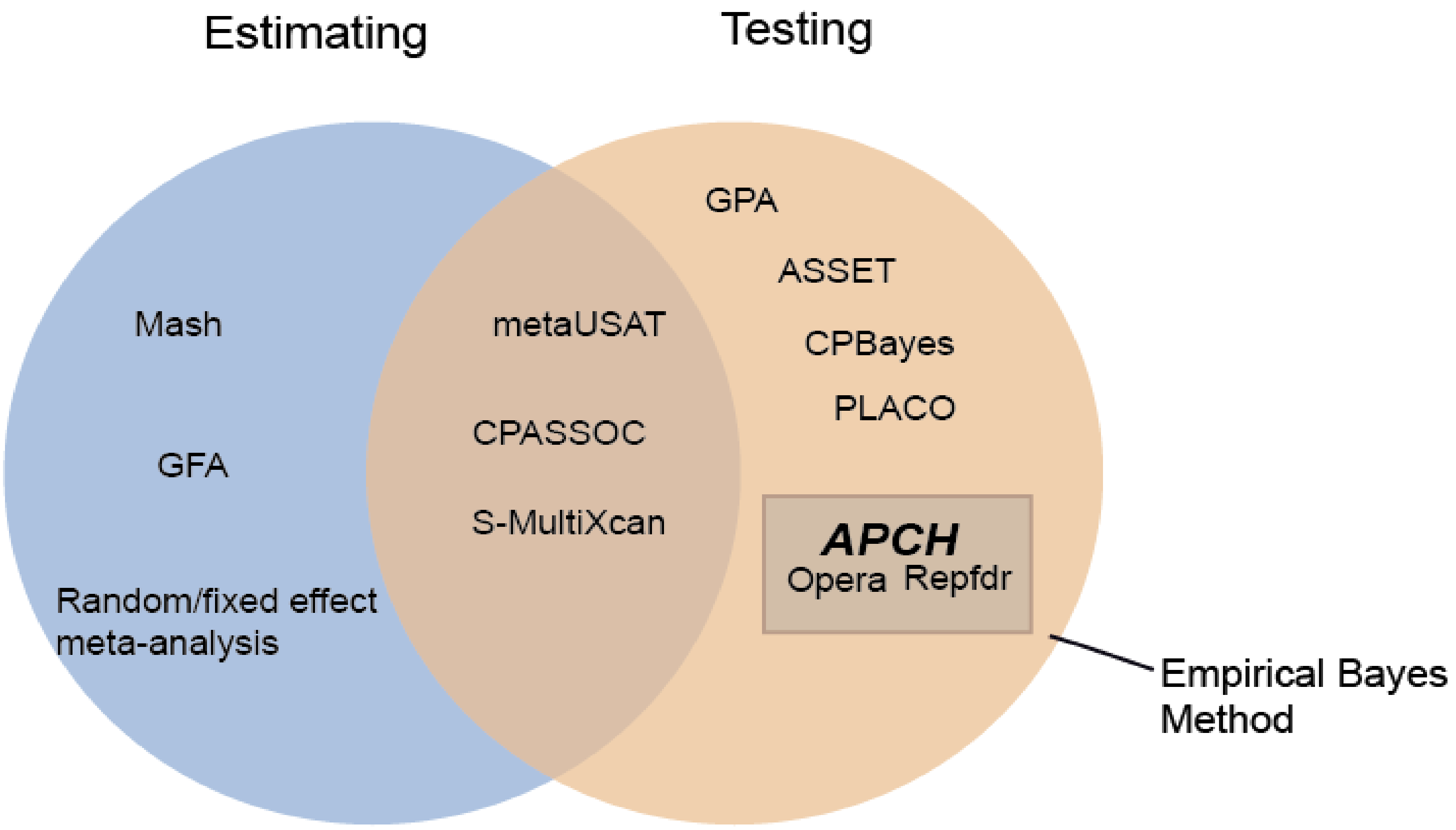
Biology Correlation:
1.Co-expression across tissues because of tissue similarity
2. the same SNP/gene perturbs a shared pathway, producing effects on multiple related traits

Biological correlation → two viewpoints: estimating shared effect patterns vs testing which subsets of features are jointly non-null for each effect unit.

Estimation Error Correlation:
1. When GWAS for Trait A and Trait B share individuals, their z-scores are correlated even if the true cross-trait effect is zero.
2. Tissue-specific TWAS stats are different linear combinations of the same cis *Z* under shared LD

Estimation error correlation must be modeled carefully in both estimation and testing of joint effects.

Related method



1. Estimating shared structure

2. Using structure to build powerful global tests

3. Primary focus on subset-level testing and error control

Empirical Bayes approach for multiple testing

Representative: Bradley Efron; Matthew Stephens (adaptive shrinkage, **ash**)

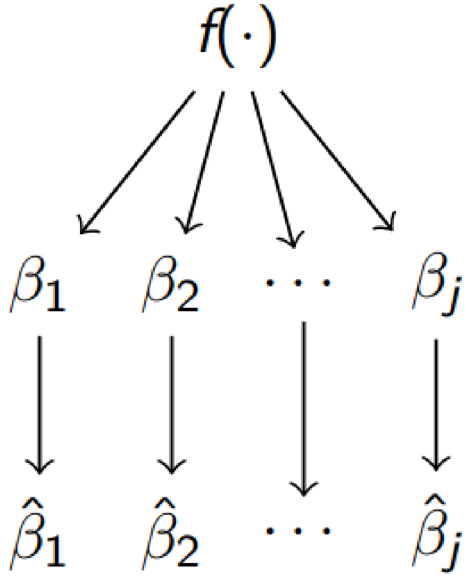
Two-group model

$$f(z) = \pi_0 f_0(z) + \pi_1 f_1(z)$$

- $f_0(z)$: density under the null
- $f_1(z)$: density under the alternative
- π_0: prior probability an effect is null
- $\pi_1 = 1 - \pi_0$: prior probability an effect is non-null

Empirical Bayes idea

- Estimate (π_0, f_0, f_1) directly from the observed z
- Plug in these estimates to compute posterior quantities



Local false discovery rate (lfdr)

$$\mathrm{lfdr}(z) = P(\text{null} \mid Z = z) = \frac{\pi_0 f_0(z)}{f(z)}$$

From univariate tests to partial conjunction hypotheses

Subset-level inference ⇄ Partial conjunction hypothesis

PCH: $H_{0,i}^{U}: \exists\, j \in U: \ b_{ij} = 0, \qquad H_{A,i}^{U}: \forall\, j \in U: \ b_{ij} \neq 0.$

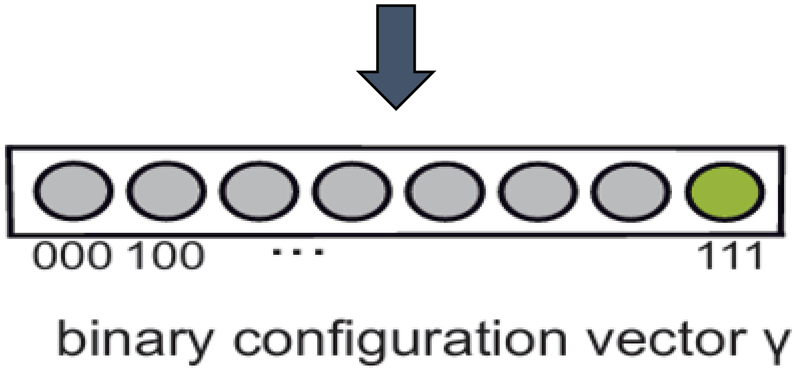
U: the subset of p features(traits/omics)
b_{ij}： the true effect of the ith effect of jth feature

➡ Jointly non-null subset per effect with FDR control

Technic goal： derive the multivariate lfdr for each PCH per effect

From univariate tests to partial conjunction hypotheses

$$f(z) = \pi_0 f_0(z) + \pi_1 f_1(z)$$



$$\mathrm{BF}_i(\boldsymbol{r}) = \frac{p(\hat{\boldsymbol{b}}_i \mid \boldsymbol{\gamma}_i = \boldsymbol{r})}{p(\hat{\boldsymbol{b}}_i \mid \boldsymbol{\gamma}_i = \boldsymbol{0})}$$

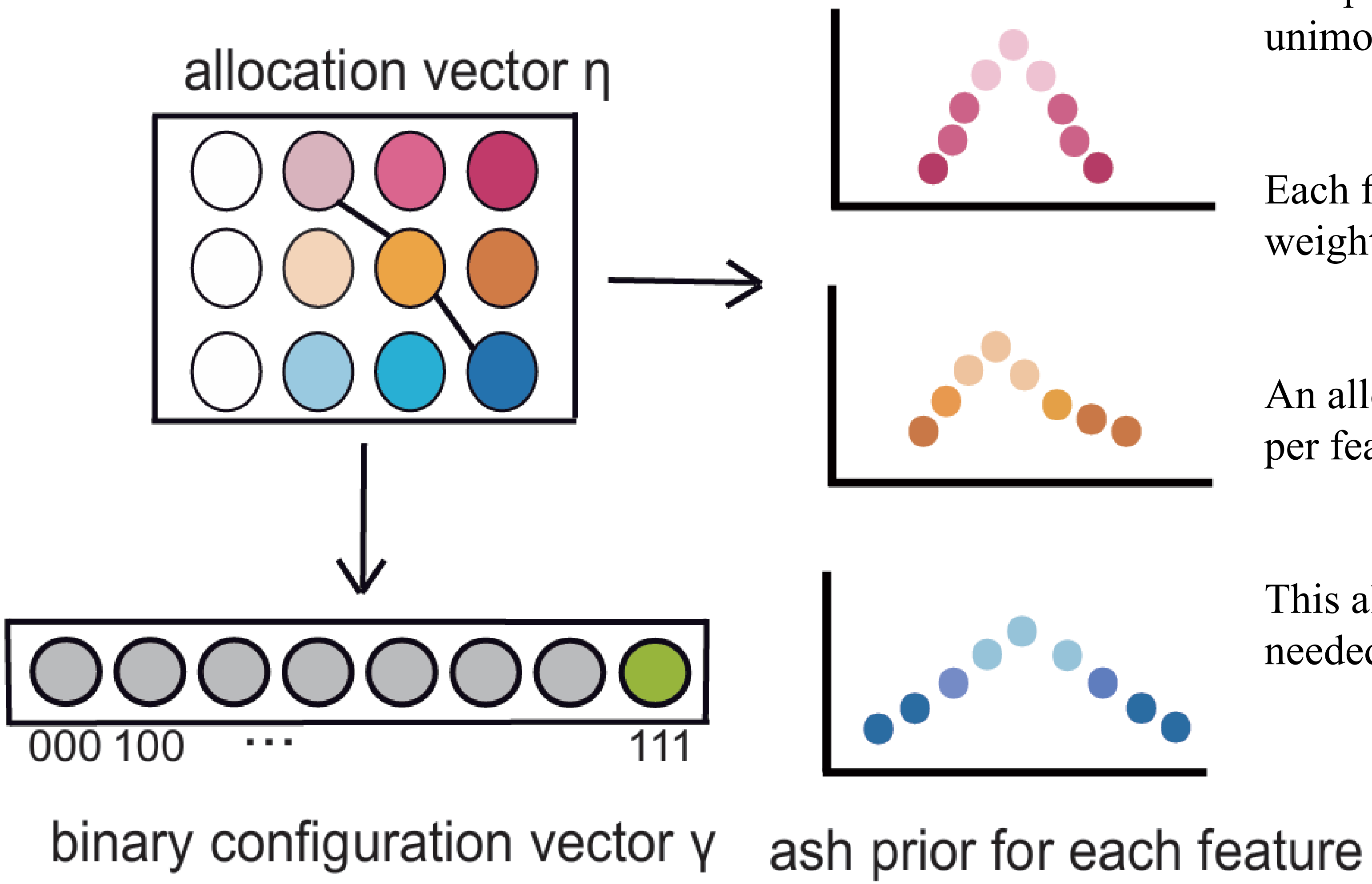
$$\mathrm{PPC}_i(\boldsymbol{r}) = \Pr(\boldsymbol{\gamma}_i = \boldsymbol{r} \mid \hat{\boldsymbol{b}}_i) = \frac{\varphi_{\boldsymbol{r}}\,\mathrm{BF}_i(\boldsymbol{r})}{\sum_{\boldsymbol{r}' \in \mathcal{R}} \varphi_{\boldsymbol{r}'}\,\mathrm{BF}_i(\boldsymbol{r}')}$$

estimate the proportion of null and non-null

$$\mathrm{PPA}_i(U) = \Pr(H^U_{A,i} \mid \hat{\boldsymbol{b}}_i) = \Pr\left(\forall j \in U : \gamma_{ij} = 1 \mid \hat{\boldsymbol{b}}_i\right) = \sum_{\substack{\boldsymbol{r} \in \mathcal{R} \\ r_j = 1\, \forall j \in U}} \mathrm{PPC}_i(\boldsymbol{r})$$

estimate the proportion of PCHs assigned to each configuration

Model intuition/DGP

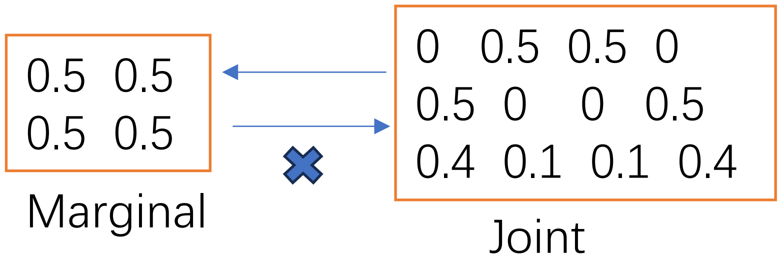


Ash prior: a normal mixture that can approximate any unimodal distribution

Each feature has a marginal ash prior whose grid weights act as hyper-parameters

An allocation vector chooses one Gaussian component per feature, spanning all combinations

This allocation induces the binary configuration vector needed for PPA / lfdr



Marginal

0.5	0.5
0.5	0.5

Joint

0	0.5	0.5	0
0.5	0	0	0.5
0.4	0.1	0.1	0.4

Parameters estimation and computation strategy

The likelihood is conceptually simple: a finite mixture of multivariate Gaussians,
so in principle we can fit it with an EM algorithm.

The real difficulty is the huge number of mixture components / parameters.

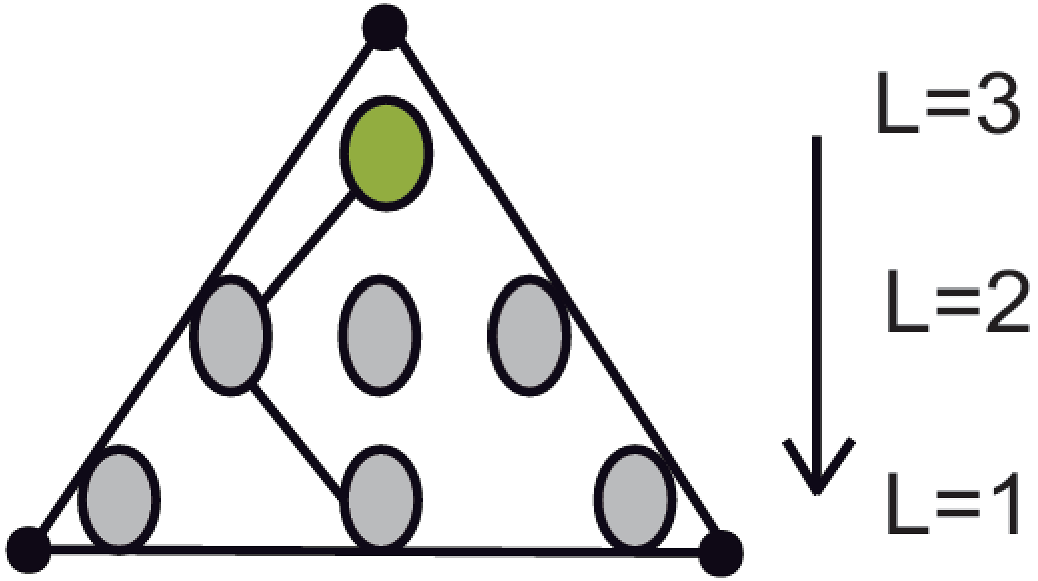
- Focus estimation on 2^p configuration probabilities instead of all fine-grained weights.

- Distill per-feature grids to a few adaptive components to shrink the state space.

- Exploit factorized likelihood (independent / block-diagonal noise) for two-stage fitting.

- Accelerate all EM updates with SQUAREM.

Inference



Now we have an lfdr for every subset of features for each effect unit.
Our goal is to report one jointly significant subset for each effect.

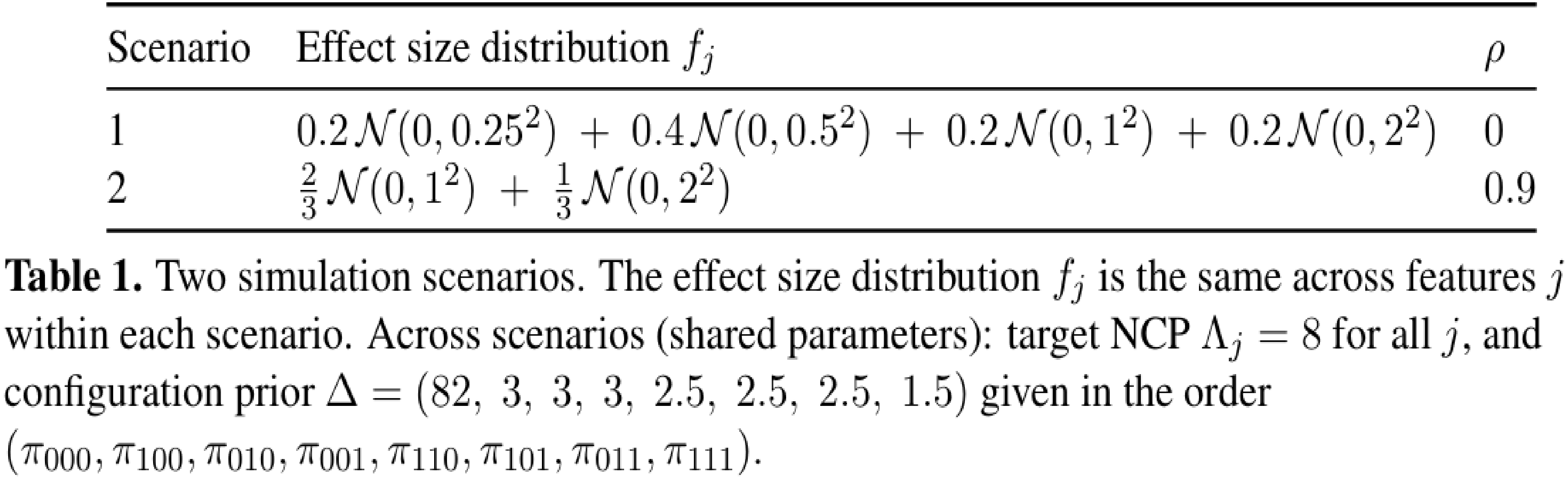
Key observation: lfdr is monotone in subset size ($U1 \subset U2 \;\Rightarrow\; \text{lfdr}(U1) \le \text{lfdr}(U2)$).

⇒ Use a level-by-level search over L (from larger subsets down to singletons).

Take home:
Jointly non-null subsets per effect
Posterior shrinkage estimates
FDR control

Simulation1

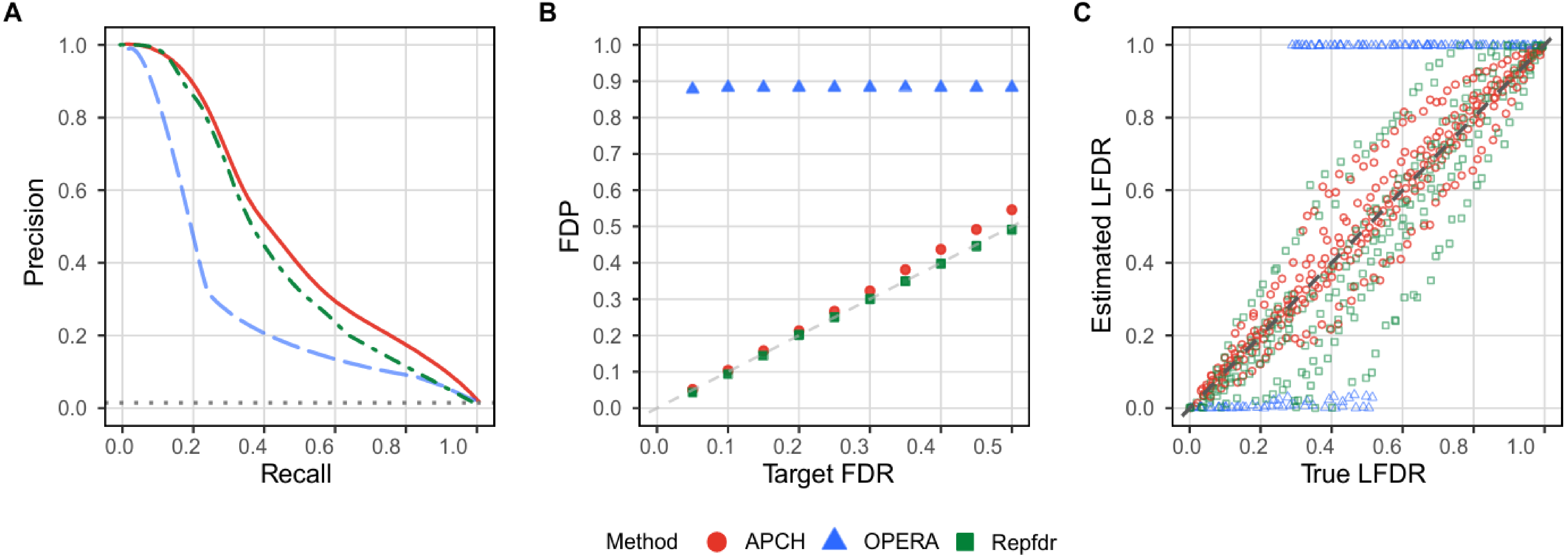
3 features 15000 effects



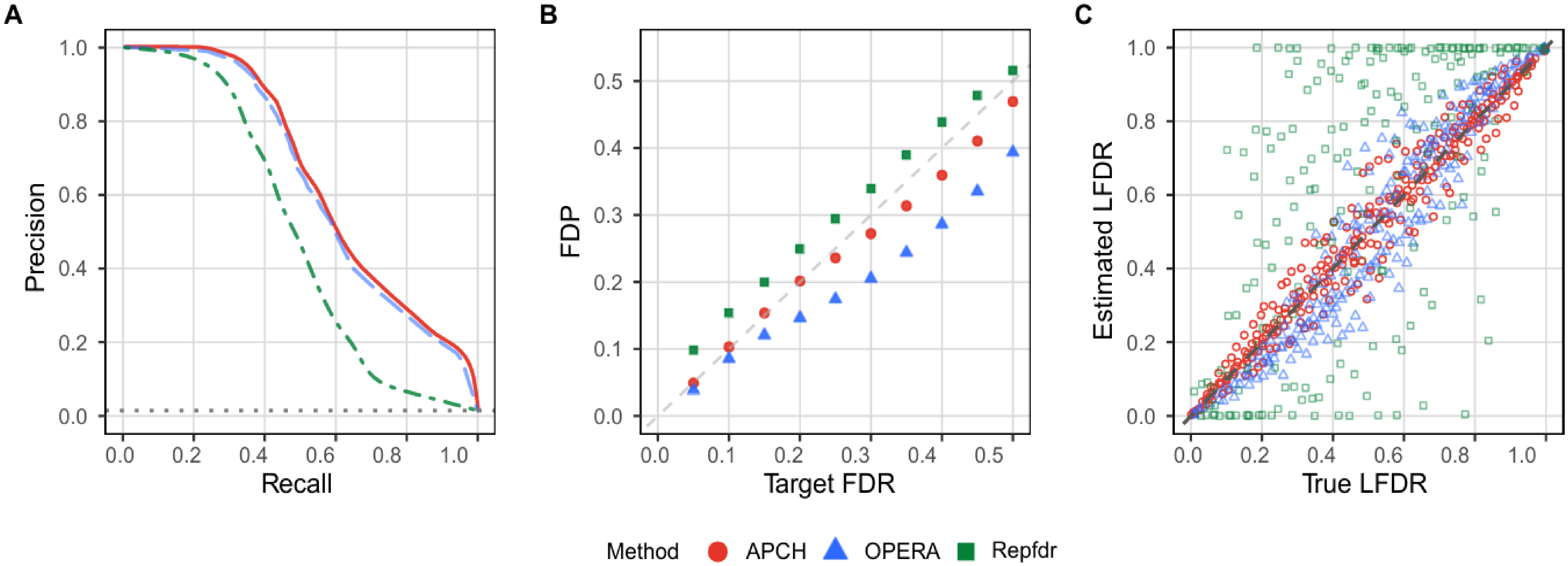
Scenario	Effect size distribution f_j	ρ
1	$0.2\,\mathcal{N}(0, 0.25^2)\ +\ 0.4\,\mathcal{N}(0, 0.5^2)\ +\ 0.2\,\mathcal{N}(0, 1^2)\ +\ 0.2\,\mathcal{N}(0, 2^2)$	0
2	$\frac{2}{3}\,\mathcal{N}(0, 1^2)\ +\ \frac{1}{3}\,\mathcal{N}(0, 2^2)$	0.9

Table 1. Two simulation scenarios. The effect size distribution f_j is the same across features j within each scenario. Across scenarios (shared parameters): target NCP $\Lambda_j = 8$ for all j, and configuration prior $\Delta = (82,\ 3,\ 3,\ 3,\ 2.5,\ 2.5,\ 2.5,\ 1.5)$ given in the order $(\pi_{000}, \pi_{100}, \pi_{010}, \pi_{001}, \pi_{110}, \pi_{101}, \pi_{011}, \pi_{111})$.

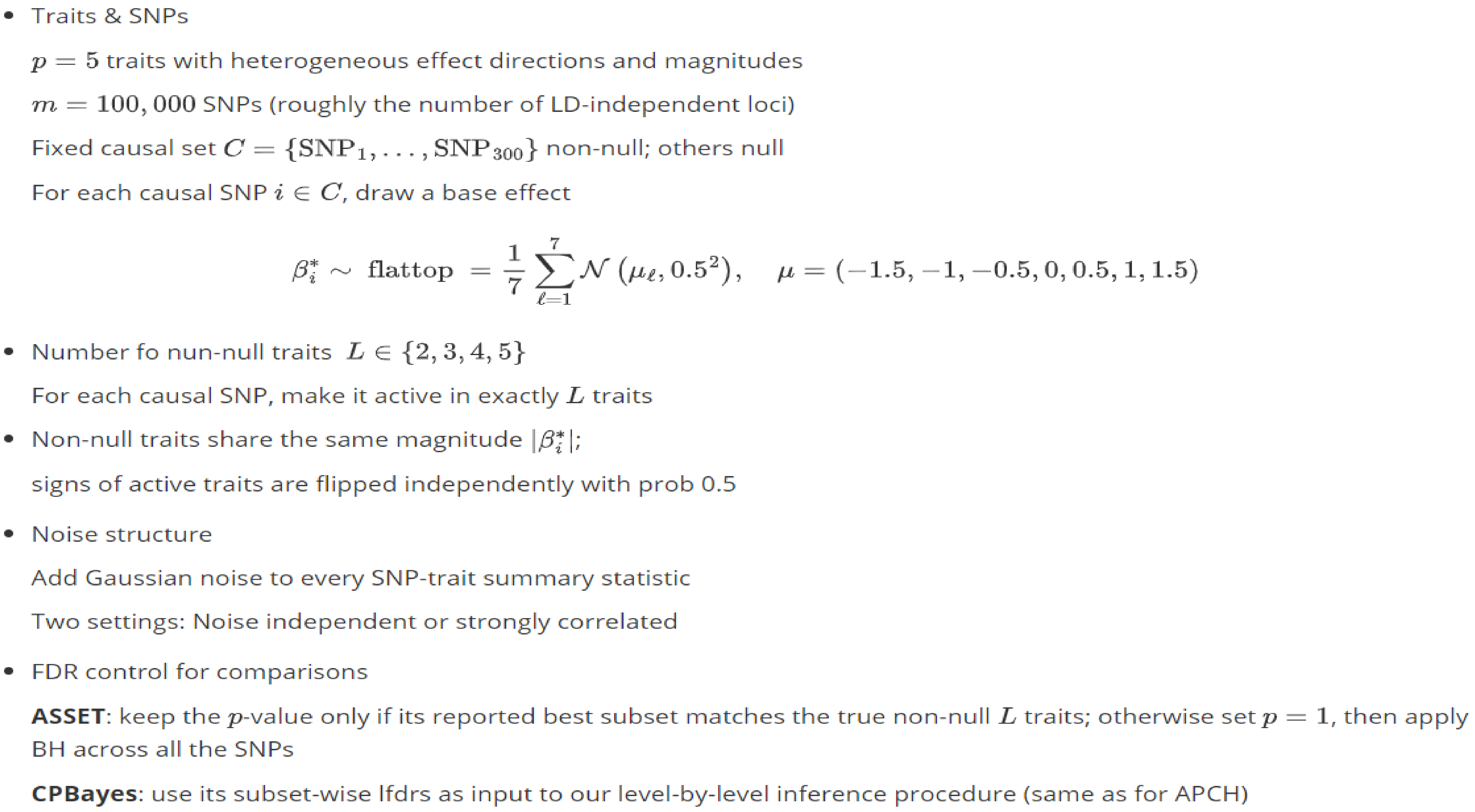
Scenario1


A
Precision
Recall
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
B
FDP
Target FDR
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.1
0.2
0.3
0.4
0.5
C
Estimated LFDR
True LFDR
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Method
● APCH
▲ OPERA
■ Repfdr

Scenario2


A
Precision
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Recall
B
FDP
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.1
0.2
0.3
0.4
0.5
Target FDR
C
Estimated LFDR
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
True LFDR
Method
● APCH
▲ OPERA
■ Repfdr

Simulation2



- Traits & SNPs

  $p = 5$ traits with heterogeneous effect directions and magnitudes

  $m = 100,000$ SNPs (roughly the number of LD-independent loci)

  Fixed causal set $C = \{\mathrm{SNP}_1, \ldots, \mathrm{SNP}_{300}\}$ non-null; others null

  For each causal SNP $i \in C$, draw a base effect

$$\beta_i^* \sim \text{ flattop } = \frac{1}{7}\sum_{\ell=1}^{7} \mathcal{N}\left(\mu_\ell, 0.5^2\right), \quad \mu = (-1.5, -1, -0.5, 0, 0.5, 1, 1.5)$$

- Number fo nun-null traits $L \in \{2, 3, 4, 5\}$

  For each causal SNP, make it active in exactly L traits

- Non-null traits share the same magnitude $|\beta_i^*|$;

  signs of active traits are flipped independently with prob 0.5

- Noise structure

  Add Gaussian noise to every SNP-trait summary statistic

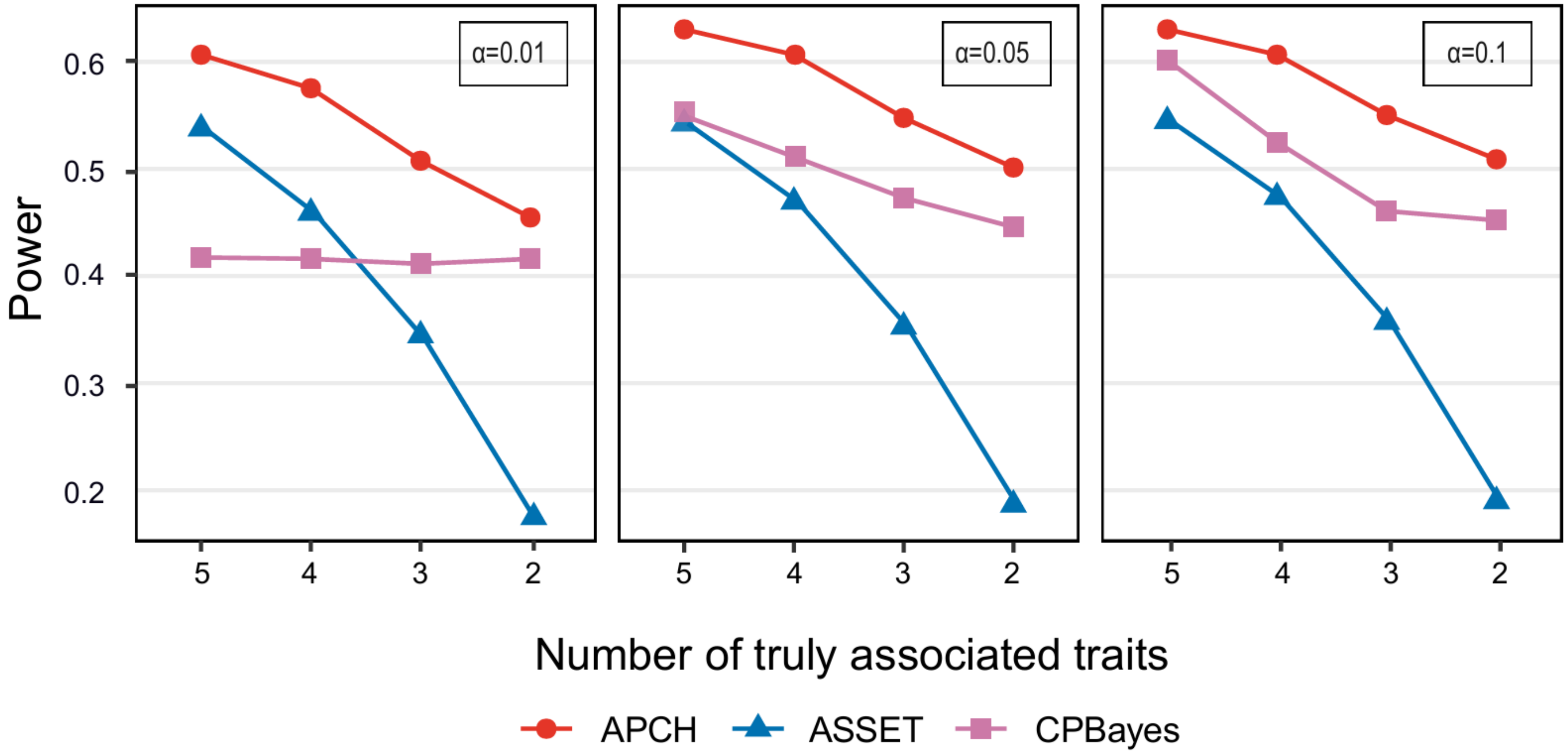
  Two settings: Noise independent or strongly correlated

- FDR control for comparisons

  ASSET: keep the p-value only if its reported best subset matches the true non-null L traits; otherwise set $p = 1$, then apply BH across all the SNPs

  CPBayes: use its subset-wise lfdrs as input to our level-by-level inference procedure (same as for APCH)

NO estimation error correlation



“Noise traits” make the space of jointly significant patterns sparser, which leads to power loss for all methods.

APCH attains the highest power and the mildest loss as more noise traits are added.

Strong estimation error correlation

L	Method	0.01			0.05			0.10		
		nTP	nFP	FDP	nTP	nFP	FDP	nTP	nFP	FDP
5	APCH	221.8	0.4	0.00	226.8	1.4	0.01	226.8	1.4	0.01
	CPBayes	85.5	0.8	0.01	96.4	4.7	0.04	104.3	9.4	0.08
	ASSET	12.4	0.0	0.00	12.7	0.0	0.00	12.9	0.0	0.00
4	APCH	211.9	0.3	0.00	220.7	2.0	0.01	220.7	2.0	0.01
	CPBayes	90.8	1.9	0.02	103.4	5.4	0.05	102.3	11.9	0.10
	ASSET	27.5	0.0	0.00	28.4	0.0	0.00	28.7	0.0	0.00
3	APCH	197.4	1.9	0.01	200.3	4.7	0.02	200.3	4.7	0.02
	CPBayes	103.0	5.4	0.05	112.7	16.3	0.13	111.5	27.3	0.20
	ASSET	44.5	0.0	0.00	46.7	0.0	0.00	47.7	0.0	0.00
2	APCH	140.4	1.3	0.01	148.8	5.2	0.03	148.8	5.2	0.03
	CPBayes	110.7	14.1	0.11	117.1	32.1	0.21	117.1	44.2	0.27
	ASSET	47.3	0.0	0.00	51.8	0.1	0.00	53.1	0.1	0.00

The other two methods fail under strong error correlation, whereas APCH benefits from it.