a

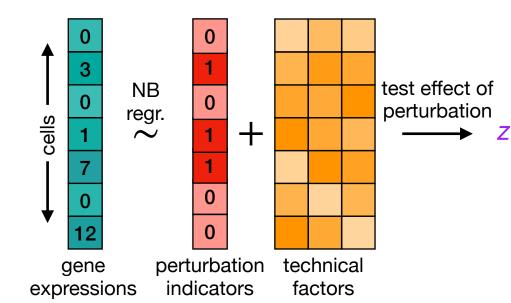**1** Obtain z-score from NB regression of gene expression on perturbation indicator and technical factors

cells

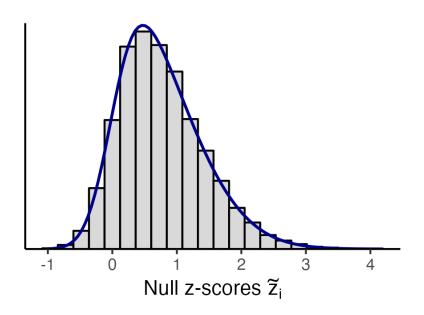gene expressions | perturbation indicators | technical factors

NB regr. $\sim$ + test effect of perturbation $\rightarrow$ z

2 For i = 1,...,M:
Permute perturbation indicator and rerun NB regression to obtain null z-score $\tilde{z}_i$

NB regr. $\sim$ + test effect of **permuted** perturbation $\rightarrow$ $\tilde{z}_i$

ith **permuted** perturbation indicator vector

4 Obtain p-value by comparing original z-score z to fitted null distribution

z

(right-tailed) SCEPTRE p-value

Null z-score distribution

3 Fit smooth curve to null z-scores $\tilde{z}_i$

Null z-scores $\tilde{z}_i$

b Relationships among variables in the analysis

Gene expression

Perturbation presence

Technical factors

Null hypothesis ?

Possible confounding

c

Confounding?

No / Yes

SCEPTRE is valid

SCEPTRE is valid if (i) NB GLM holds (up to dispersion parameter) and (ii) sample size is sufficiently large