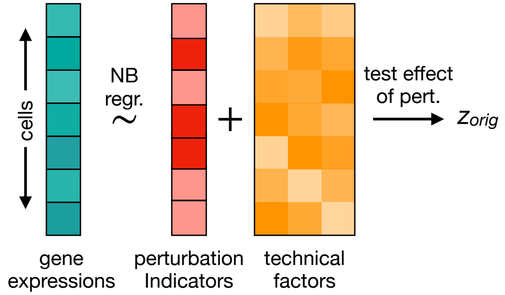
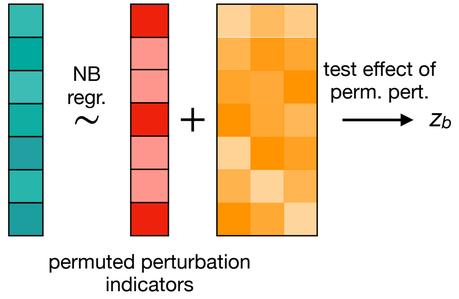
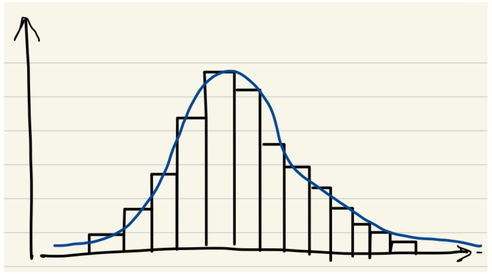
1 Obtain z-score from NB regression of gene expression on perturbation indicator and technical factors



2 `For b = 1,...,B:`
Permute perturbation indicator and rerun NB regression to obtain null z-score z_b

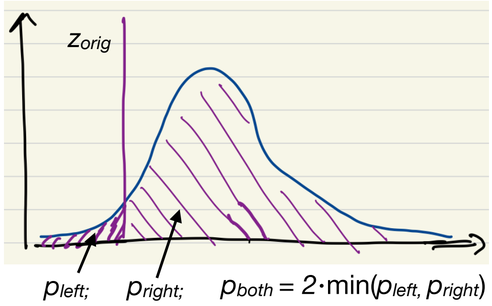


3 Fit smooth curve to null z-scores z_b

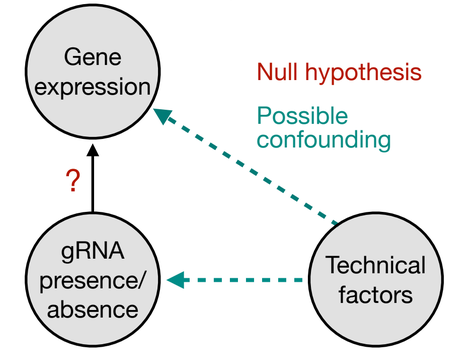


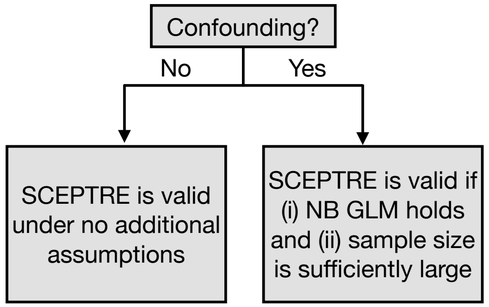
z_{orig}

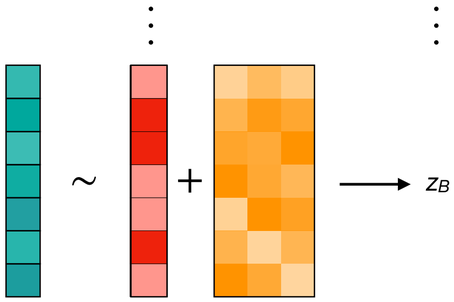
4 Obtain p-value by comparing original z-score z_{orig} to fitted null distribution

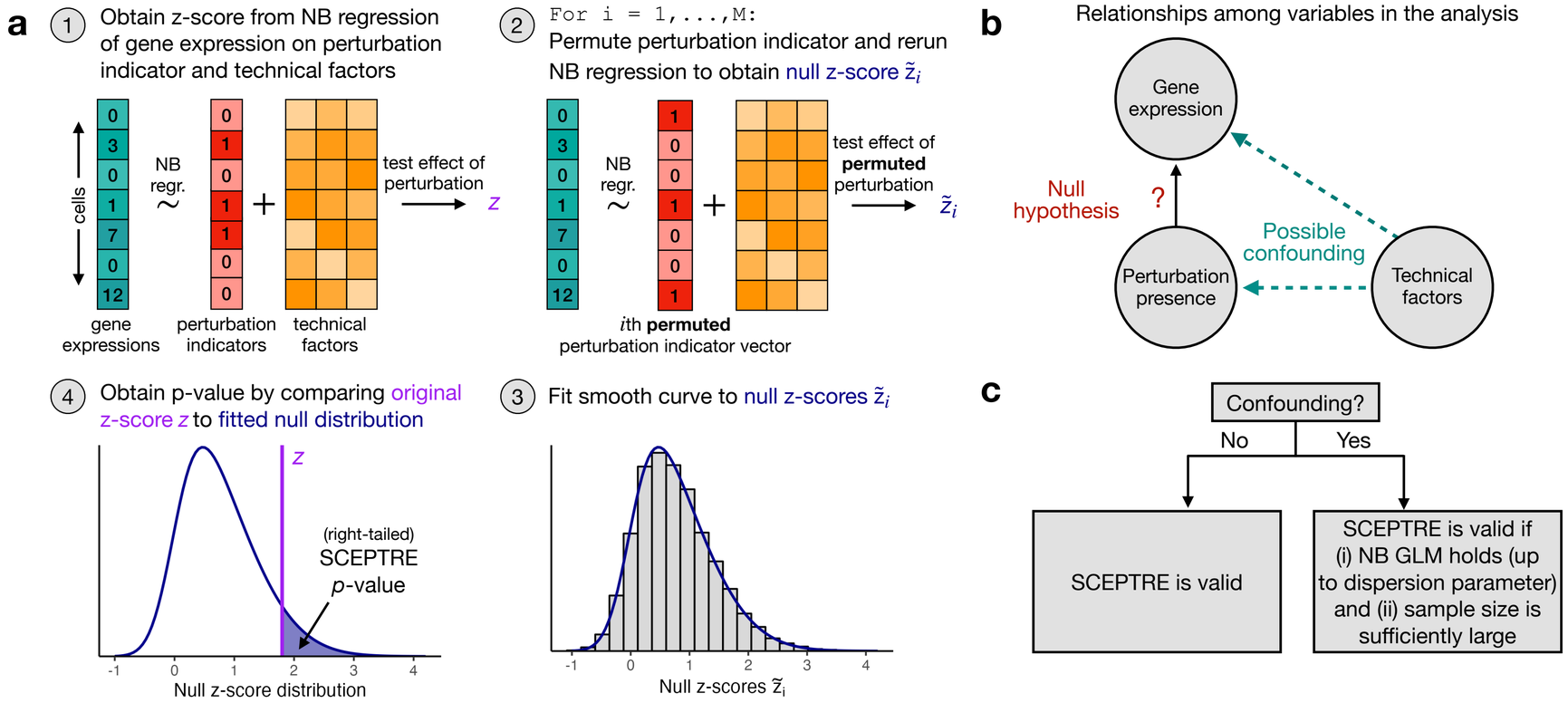


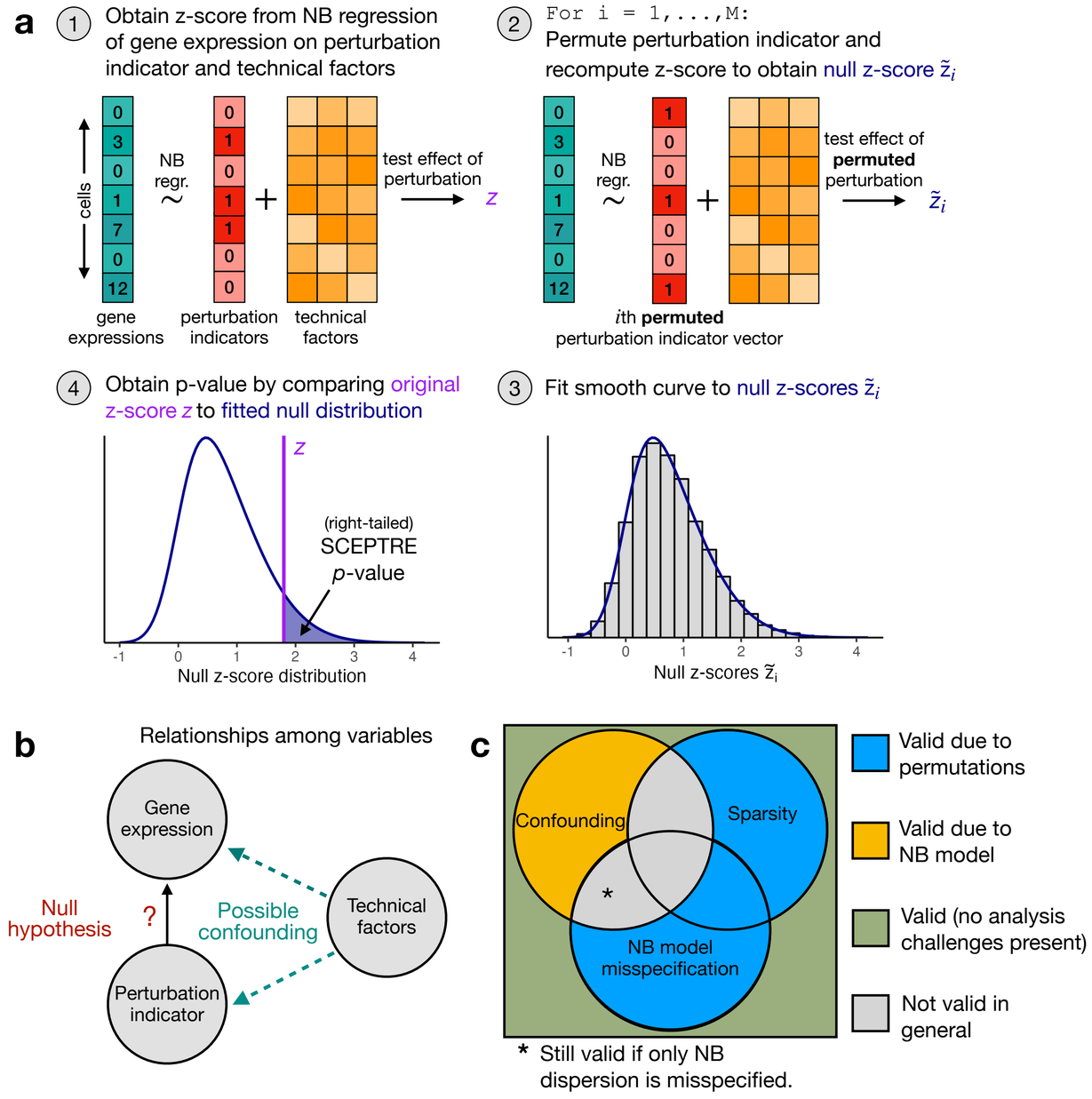
Relationship between variables in analysis

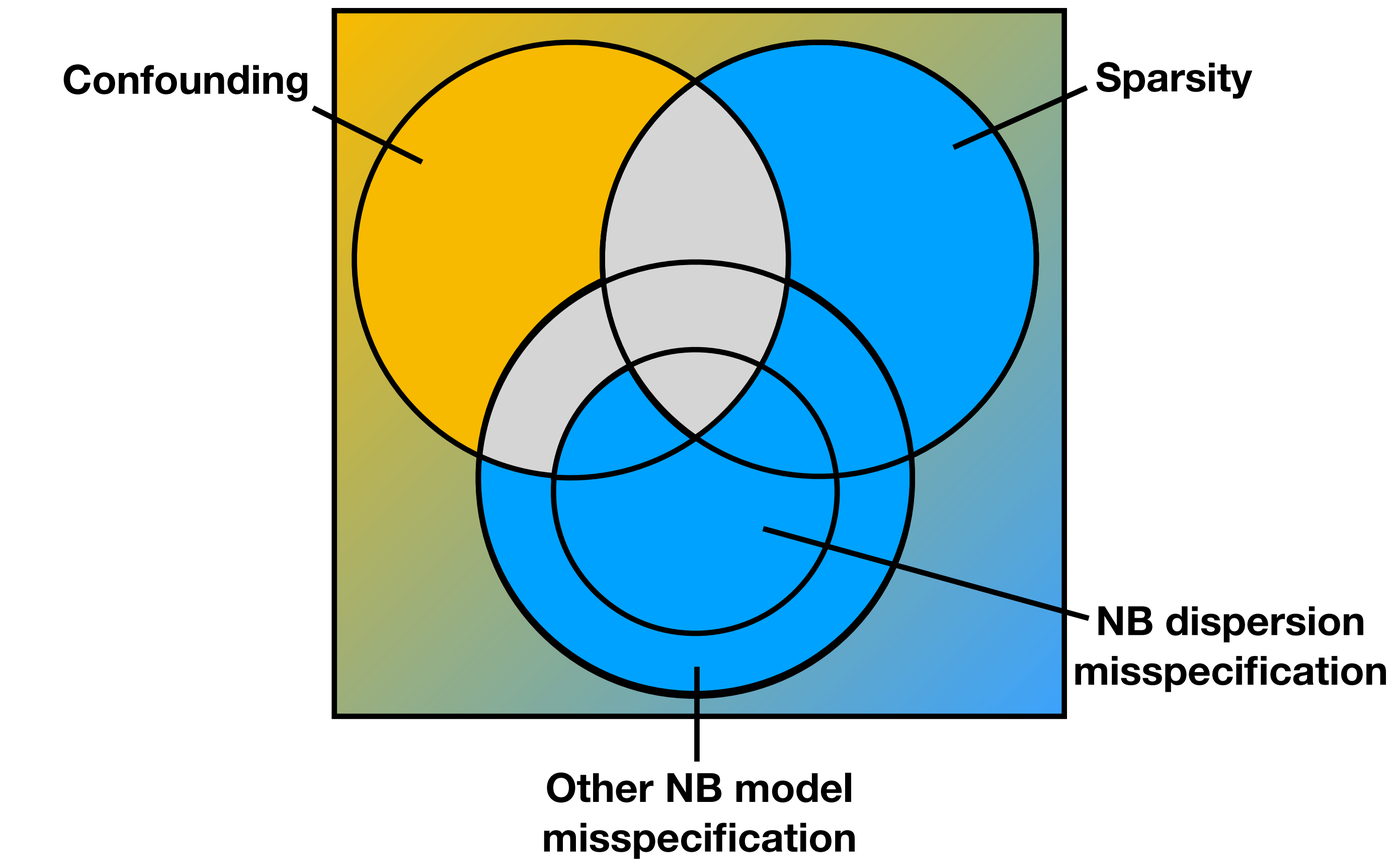


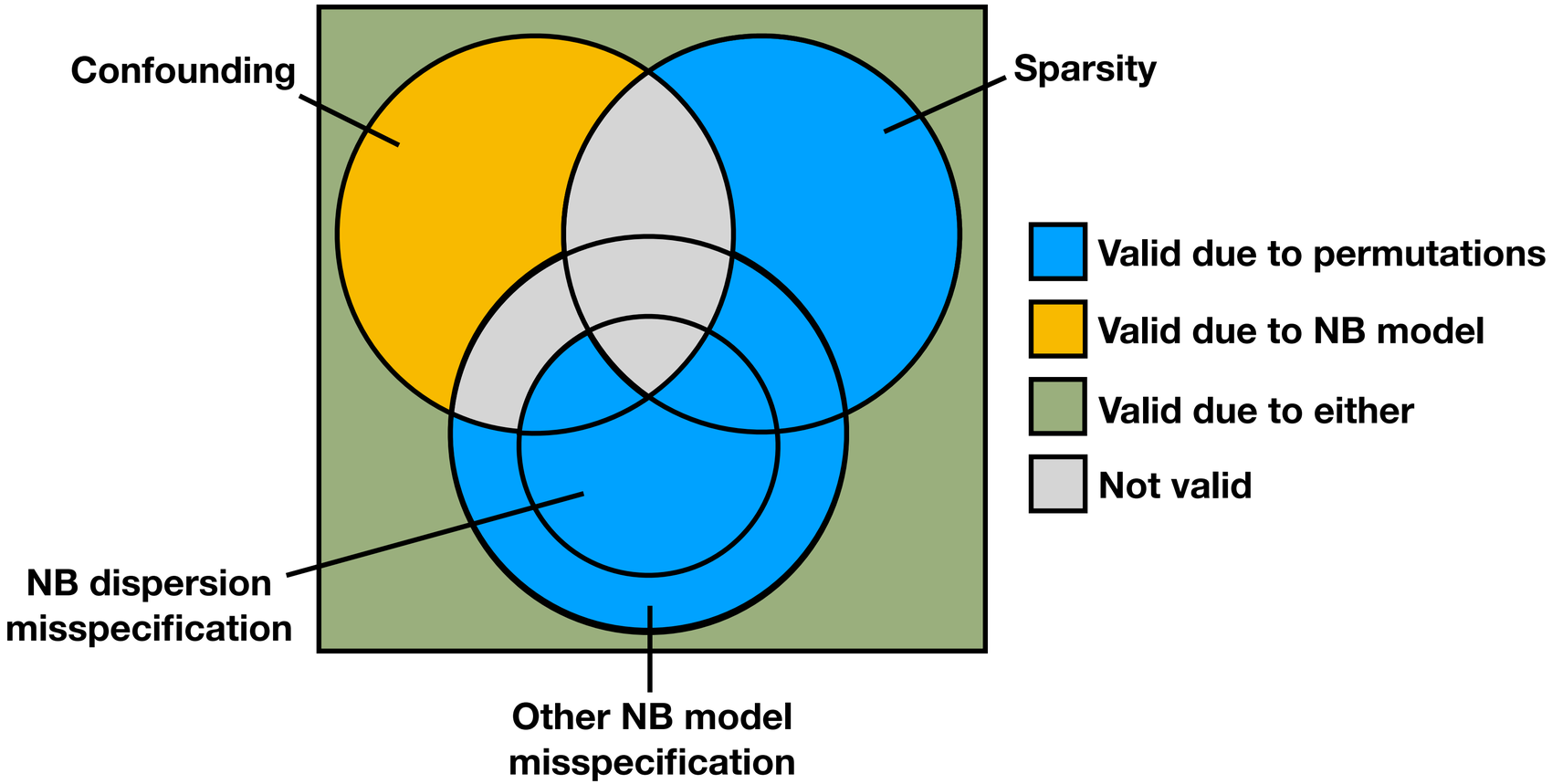


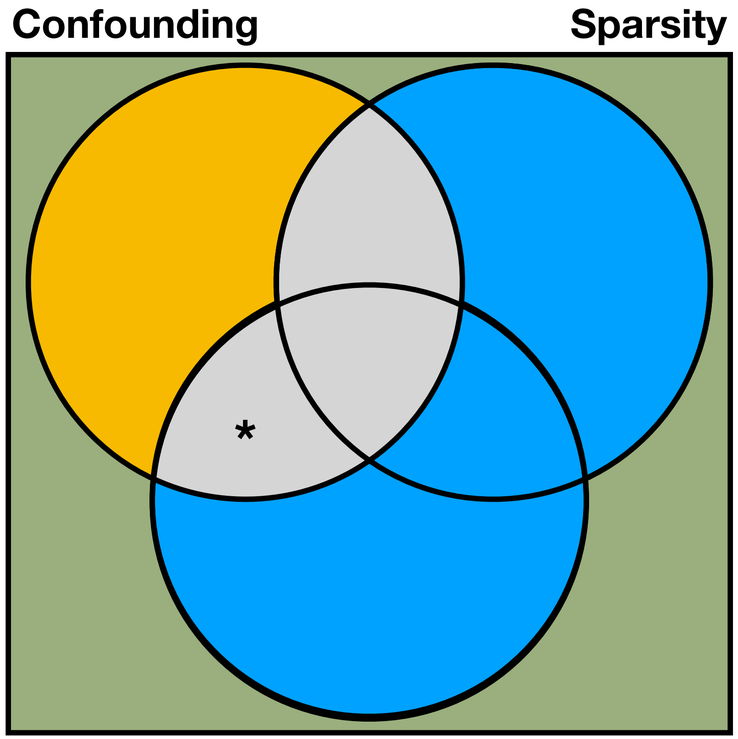



a
1
Obtain z-score from NB regression of gene expression on perturbation indicator and technical factors
cells
NB regr.
test effect of perturbation
z
gene expressions
perturbation indicators
technical factors
2
For i = 1,...,M:
Permute perturbation indicator and rerun NB regression to obtain null z-score $\tilde{z}_i$
test effect of permuted perturbation
$\tilde{z}_i$
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
Null hypothesis
?
Possible confounding
Perturbation presence
Technical factors
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


a
1
Obtain z-score from NB regression of gene expression on perturbation indicator and technical factors
cells
NB regr.
test effect of perturbation
z
gene expressions
perturbation indicators
technical factors
2
For i = 1,...,M:
Permute perturbation indicator and recompute z-score to obtain null z-score $\tilde{z}_i$
test effect of permuted perturbation
$\tilde{z}_i$
ith permuted perturbation indicator vector
4
Obtain p-value by comparing original z-score z to fitted null distribution
(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
Gene expression
Technical factors
Perturbation indicator
Null hypothesis
?
Possible confounding
c
Confounding
Sparsity
NB model misspecification
Valid due to permutations
Valid due to NB model
Valid (no analysis challenges present)
Not valid in general
* Still valid if only NB dispersion is misspecified.


Confounding
Sparsity
NB dispersion misspecification
Other NB model misspecification
Valid due to permutations
Valid due to NB model
Not valid


Confounding
Sparsity
NB dispersion misspecification
Other NB model misspecification
Valid due to permutations
Valid due to NB model
Valid due to either
Not valid



NB model misspecification

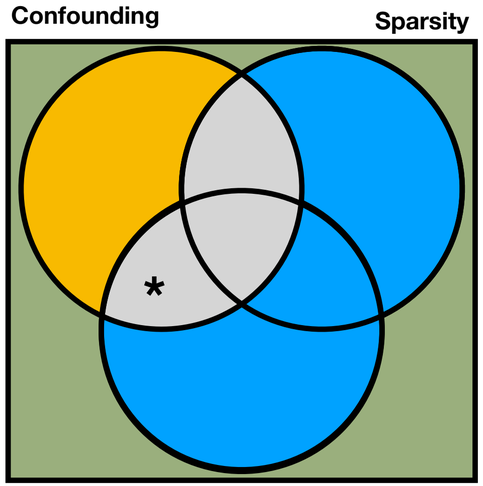
Valid due to permutations

Valid due to NB model

Valid due to either

Not valid in general

*Still valid if only NB dispersion is misspecified.



NB model misspecification

Valid due to permutations

Valid due to NB model

Valid due to either

Not valid in general

*Still valid if only NB dispersion is misspecified.