d=dimension of the input and output data

| data1 |
| data2 |

↓

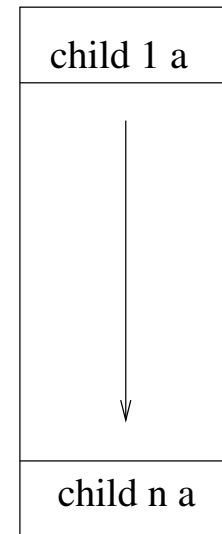| data 2n−1 |
| data 2n |
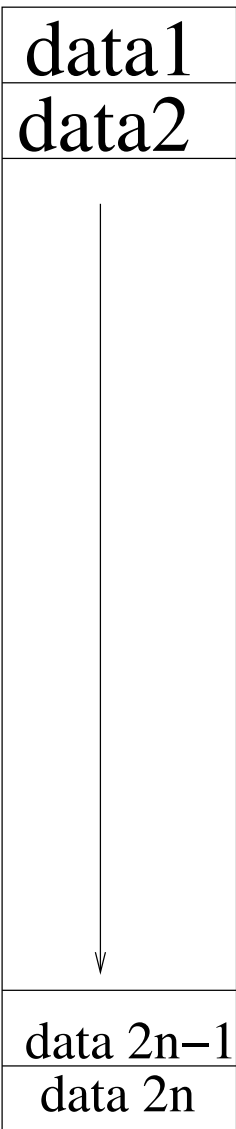
mean parent
variance parant

Sort the data
in direction 1<=dir<=d
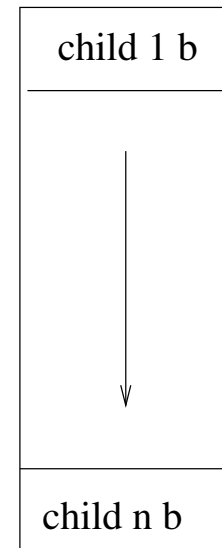and split according
to the median.

⟶
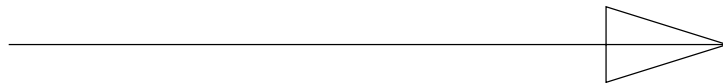
choose the critical direction

which maximizes the classification
variance reduction from parent to
children:

max (variance parent −
variance child a −
variance child b)

| child 1 a |

↓

| child n a |

mean child a
variance child a

| child 1 b |

↓

| child n b |

mean child b
variance child b