

ATAGAGGTACG  
GACCGCATGAT  
CGGACATCAGA  
TCAGACTAAT

A G A  
T G A  
A C A  
A A T

A	0.75	0.25	0.75
C	0	0.25	0
G	0	0.50	0
T	0.25	0	0.25

# Remove 3rd string and recalculate matrix

ATAGAGGTACG  
GACCGCATGAT  
CGGACATCAGA  
TCAGACTAAT

A	G	A
T	G	A
A	A	T

A	0.66	0.33	0.66
C	0	0	0
G	0	0.66	0
T	0.33	0	0.33

# Find best match to profile within 3rd string

ATAGAGGTACG  
GACCGCATGAT  
CGGACATCAGA  
TCAGACTAAT

A	G	A
T	G	A
A	G	A
A	A	T

A	0.75	0.25	0.75
C	0	0	0
G	0	0.75	0
T	0.25	0	0.25

# Remove 4th string

ATAGAGGTACG  
GACCGCATGAT  
CGGACATCAGA  
TCAGACTAAAT

A	G	A
T	G	A
A	G	A

A	0.66	0	1
C	0	0	0
G	0	1	0
T	0.33	0	0

# Replace with best match from 4th string

ATAGAGGTACG  
GACCGCATGAT  
CGGACATCAGA  
TCAGACTAAAT

A	G	A
T	G	A
A	G	A
A	G	A

A	0.75	0	1
C	0	0	0
G	0	1	0
T	0.25	0	0

# Remove 2nd string

ATAGAGGTACG  
GACCGCATGAT  
CGGACATCAGA  
TCAGACTAAAT

A G A

A G A

A G A

A	1	0	1
C	0	0	0
G	0	1	0
T	0	0	0

# Replace with best match from 2nd string

ATAGAGGTACG	A	G	A
GACCGCATGAT	T	G	A
CGGACATCAGA	A	G	A
TCAGACTAAAT	A	G	A

A	0.75	0	1
C	0	0	0
G	0	1	0
T	0.25	0	0

Note: without pseudo-counts all k-mers on 2nd string are equally likely (probability 0)

Entropy?

$2 - \sum(-p \log p)$