APPENDIX A
Proof of Theorem 1
First, we introduce a new way to derive the expected value of mutual information in case of random partitions and under hyper-geometric distribution assumption and then we use the expected value to prove (13). Consider a pair of clusters P_i and G_j. The probability that an object in P_i exists in G_j is m_i / N. Accordingly, the number of objects in both P_i and G_j is simplified as: n_j = n_i × (m_j / N). Then, the expected value can be calculated according to (7) as:

\[ E(\text{MI}) = E \left( \sum_i \sum_j n_i \log \left( \frac{N \times (n_i \times m_j / N)}{n_i \times n_j} \right) \right) \]

(24)

According to (2), AMI=NM1 which confirms the result from [9]. Applying max(MI)=H(P)+H(G))/2 as an option for normalization [22], [17], we can write:

\[ \text{AMI} = \frac{2 \times \text{MI}(P,G)}{H(P)+H(G)} \]  

(25)

Since E(H(P)) = H(P) and E(H(G)) = H(G) under hyper-geometric distribution assumption, the expected value of VI (8) is derived as:

\[ E(\text{VI}) = H(P) + H(G) \]  

(26)

VI is a dissimilarity measure and min(VI) = 0 when the two partitions are equal. Therefore, the adjusted variation of information according to (2) is:

\[ \text{AVI} = \frac{VI}{H(P)+H(G)} \]  

(27)

An upper bound for VI is H(P) + H(G) and therefore (27) also represents the normalized variation of information. We simplify AVI_s and NVI_s using (8) as follows:

\[ \text{AVI}_s = \text{NVI}_s = \frac{2 \times \text{MI}(P,G)}{H(P)+H(G)} \]  

(28)

From (25) and (28), we see that the adjusted mutual information and adjusted variation of information are equal to their normalized forms, and thus, theorem 1 is proven.

APPENDIX B
Proof of Theorem 2
Suppose that in a matching, m_i is paired to n_i < n_j and n_j is paired to m_j (case a). We show that if we change the matching so that m_i is paired to n_1 and m_j is paired to n_2 (case b), higher similarity is achieved. The total similarities for these two cases (a and b) are:

\[ S_a = \frac{m_i \times (n_i / N)}{\max(m_i,n_i)} + \frac{m_j \times (n_j / N)}{\max(m_j,n_j)} \]

\[ S_b = \frac{m_i \times (n_j / N)}{\max(m_i,n_i)} + \frac{m_j \times (n_i / N)}{\max(m_j,n_j)} \]  

(29)

where S_a is the original pairing and S_b is the new pairing after changing the pairs for m_i and m_j. Six different situations may happen:

1. \( m_1 > m_j > m_i > n_i \)

\[ S_a = \frac{1}{N} (n_i + n_j) \]

\[ S_b = \frac{1}{N} (n_j + n_i) \]

2. \( m_1 > n_i > m_j > n_i \)

\[ S_a = \frac{1}{N} (n_i + m_j) \]

\[ S_b = \frac{1}{N} (n_j + n_i) \]

3. \( m_1 > n_i > n_j > m_j \)

\[ S_a = \frac{1}{N} (n_i + m_j) \]

\[ S_b = \frac{1}{N} (n_j + m_j) \]

4. \( n_1 > m_i > n_j > m_j \)

\[ S_a = \frac{1}{N} (m_i + m_j) \]

\[ S_b = \frac{1}{N} (n_j + n_i) \]

5. \( n_1 > n_i > m_j > m_j \)

\[ S_a = \frac{1}{N} (m_i + m_j) \]

\[ S_b = \frac{1}{N} (n_j + n_i) \]

6. \( n_1 > n_i > n_j > m_j \)

\[ S_a = \frac{1}{N} (m_i + n_j) \]

\[ S_b = \frac{1}{N} (n_i + n_j) \]  

(30)

Considering all the above situations, pairings (m_i, n_i) and (n_i, m_j) must be changed to (n_1, m_1) and (m_j, n_j) to achieve higher similarity. We can apply this proof recursively to all the smaller clusters as well. Hence, the two largest clusters must be always paired and then the next two largest and so on to achieve maximum total similarity with a random partition. This proves the theorem 2.

APPENDIX C
Triangular Inequality Proof for the Simplified form of PSI
Let P_1, P_2 and P_3 be three partitions with K_1, K_2 and K_3 clusters, and K_12=max(K_1,K_2), K_23=max(K_2,K_3), K_13=max(K_1,K_3). Let n_i, n_j and n_k be the number of objects in clusters i, j and k in P_1, P_2 and P_3 respectively. We denote the number of shared objects between clusters by n_ij, n_ik and n_k. The simplified distance form of PSI, for P_1 and P_2, according to (20) is:

\[ D_{12} = \frac{K_{12} - S_{12}}{K_{12} - 1} \]  

(31)

Lemma. \( D_{12} + D_{23} \geq D_{13} \)

Proof. We define D'_1 = K_1 - S_1, D'_2 = K_2 - S_2 and D'_3 = K_3 - S_3 and prove first that: D'_1 + D'_2 \geq D'_3 which is equivalent to

\[ K_1 - S_1 + K_2 - S_2 \geq K_3 - S_3 \]  

(32)

We consider three possible situations and simplify (32):

1. \( K_1 \geq K_2 \geq S_2 \geq K_3 \geq S_3 \)

2. \( K_2 \geq K_1 \geq S_1 \geq K_3 \geq S_3 \)

3. \( K_3 \geq K_2 \geq S_1 \geq K_1 \geq S_2 \)

In the case (3), since \( K_2 \geq K_3 \), it is sufficient to prove \( S_1 + S_2 \leq K_3 + S_3 \). Since \( K_3 \geq K_2 \) and \( K_2 \geq K_3 \), for all cases it is
sufficient to prove:
\[ S_{12} + S_{23} \leq K_2 + S_{13} \]  
(33)

According to the definitions (14) and (15), we divide the inequality (33) into sub-inequalities by considering each cluster \( j \) in \( P_2 \) on the left. Each sub-inequality is of the form:
\[
\frac{n_{ij}}{\max(n_{ij}, n_{ik})} + \frac{n_{jk}}{\max(n_{jk}, n_{ki})} \leq 1 + \frac{n_{ij} + n_{jk} - n_i}{\max(n_i, n_k)}
\]
(34)

Clusters \( i \) and \( k \) from \( P_1 \) and \( P_3 \) which are the pairs for cluster \( j \) are not necessarily a pair in comparing \( P_1 \) and \( P_3 \). Since \( S_{13} \) is derived according to perfect matching, we can consider another matching of \( P_1 \) and \( P_3 \) in which \( i \) and \( k \) are paired. If (33) holds in this case, it will also be true for \( S_{13} \) which is the maximum possible similarity.

If the cluster \( j \) has a pair cluster only in \( P_1 \) or \( P_3 \), it is trivial to prove (34). If it has pair clusters both in \( P_1 \) and \( P_3 \), and \( n_{ij} + n_{ik} \leq n_i \), proving (34) is trivial as well since the left side of the inequality is smaller than one. Note that if the clusters \( i \) and \( k \) do not have any shared objects, \( n_{ij} + n_{jk} \leq n_i \). So we prove (34) when \( n_{ij} + n_{jk} > n_i \). Considering a minimum value for \( n_{ik} \) as \( n_{ij} + n_{jk} - n_i \), we rewrite (34) as follows:
\[
\frac{n_{ij}}{\max(n_{ij}, n_{ik})} + \frac{n_{jk}}{\max(n_{jk}, n_{ki})} \leq 1 + \frac{n_{ij} + n_{jk} - n_i}{\max(n_i, n_k)}
\]
(35)

Three possible cases are:
(1) \( n_i \geq \max(n_i, n_k) \): By replacing \( \max(n_i, n_k) \) and \( \max(n_i, n_k) \) by \( n_i \) and after simplifications, we have:
\[
(n_{ij} + n_{jk} - n_i)(n_i - \max(n_i, n_k)) \geq 0
\]
which is always true in this case.
(2) \( n_i \geq \max(n_i, n_k) \): We replace \( \max(n_i, n_k) \) and \( \max(n_i, n_k) \) by \( n_i \). Since \( \max(n_i, n_k) \geq n_i \), it is sufficient to prove (35) by replacing \( \max(n_i, n_k) \) by \( n_i \). The equivalent inequality derived after simplification:
\[
(n_i - n_i)(n_i - n_i) \geq 0
\]
is always true.
(3) \( n_k \geq \max(n_i, n_k) \): The same proof in the case (2) can be applied.

The lemma (31) can now be represented as:
\[
\frac{K_{12} - S_{12}}{K_{12} - 1} + \frac{K_{23} - S_{23}}{K_{23} - 1} \geq \frac{K_{13} - S_{13}}{K_{13} - 1}
\]
(36)

We consider three possible cases:
(1) \( K_2 \geq K_{23} \): It is sufficient to prove (36) if \( K_{23} \) in denominator is replaced by \( K_1 \). So we simplify (36) as follows:
\[
\frac{K_{12} - S_{12}}{K_1 - 1} + \frac{K_{23} - S_{23}}{K_1 - 1} \geq \frac{K_{13} - S_{13}}{K_1 - 1}
\]

Since \( K_2 \geq 2 \), The denominators can be canceled and the inequality is true according to (32).
(2) \( K_3 \geq K_{12} \): The same inference as the case (1) can be performed by replacing \( K_{12} \) with \( K_3 \).