

Application of Buneman Graphs: We now show that r, p can be found efficiently. To prove this we need some tools from the theory of Buneman graphs [24]. Let M be a set of taxa defined by character set C of size m . A Buneman graph F for M is a vertex induced subgraph of the m -cube. Graph F contains vertices v if and only if for every pair of characters $i, j \in C$, $(v[i], v[j]) \in G_{i,j}$. Recall that $G_{i,j}$ is the set of gametes (or projection of M on dimensions i, j). Each edge of the Buneman graph is labeled with the character at which the adjacent vertices differ.

We will use the Buneman graph to show how to incrementally extend a set of taxa M by adding characters that share exactly two gametes with some existing character. As before, we can assume without loss of generality that the all zeros taxon is present in M . Therefore if a pair of characters share exactly two gametes then they are identical. Assume that we want to add character i to M and $i' \in M$ is identical to i . We extend M to M' by first adding the states on character i' for all taxa. For the rest of the discussion let $G_{i,j}$ be the set of gametes shared between characters i, j in matrix M' . We extend M' to M'' by adding a taxon t s.t. $t[i] = 0, t[i'] = 1$ and for all other characters j , if $(0, 1) \notin G_{j,i}$ then $t[j] = 1$ else $t[j] = 0$. Since we introduced a new gamete on i, i' , no pair of characters share exactly two gametes in M'' . Therefore a Buneman graph G'' for M'' can be constructed as before. A Buneman graph is a median graph [24] and clearly a subgraph of the $m + 1$ -cube, where $m + 1$ is the number of characters in M'' . Every taxon in M' is present in G'' by construction. Using the two properties, we have the following lemma.

Lemma 4.3: Every optimum phylogeny for the taxa in M' defined over the $m + 1$ characters is contained in G'' (See Section 5.5, [24] for more details).

We now show the following important property on the extended matrix M'' .

Lemma 4.4: If a pair of characters c, c' conflict in M'' then they conflict in M' .

Proof: For the sake of contradiction, assume not. Clearly i, i' share exactly three gametes in M'' . Now consider any character j and assume that j, i shared exactly three gametes in M' . For the newly introduced taxon t , $t[i] = 0$. If $t[j] = 1$, then j, i cannot share $(0, 1)$ gamete in M'' and therefore they do not conflict. If $t[j] = 0$, then the newly introduced taxon creates the $(0, 0)$ gamete which should be present in all pairs of characters. Now consider the pair of characters (j, i') . If $t[j] = 1$, then in any taxon t' of M' , if $t'[j] = 1$ then $t'[i] = 1$ and therefore $t[i] = 1$ (since i, i' are identical on all taxa except t) and therefore $(1, 1)$ cannot be a newly introduced