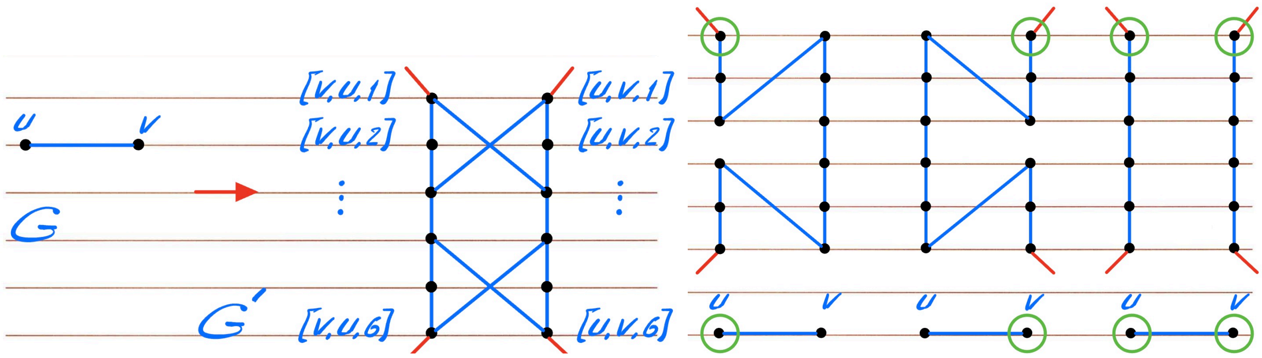
Traveling Salesman Problem & Hamiltonian Cycle

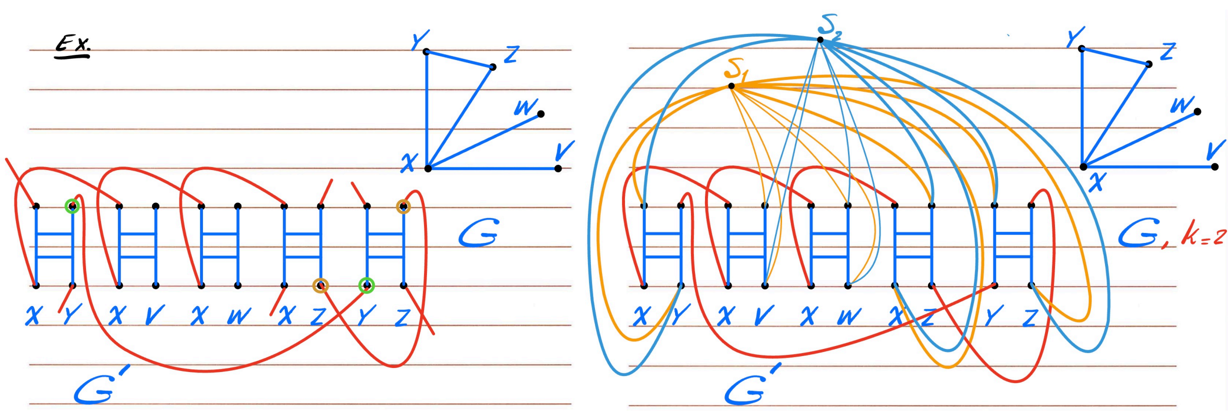
Julius April 10, 2024

1 Hamiltonian Cycle

Hamiltonian Cycle: A cycle C in G is a Hamiltonian Cycle if it visits each vertex exactly once.
Definition: Given an undirected graph G, is there a Hamiltonian Cycle in G?
Hamiltonian Cycle problem is NP-Complete.
Show that the problem is in NP.
Certificate: An ordered list of nodes on the Hamiltonian Cycle.
Certifier: Check all nodes appear on the list exactly once. Every pair of adjacent nodes in the given order must have an edge between them. The first and last nodes have an edge between them.
Vertex Cover $\leq_p$ Hamiltonian Cycle
1. Given any instance of Vertex Cover problem in an undirected graph G = (V, E) and an integer k, we construct G' = (V', E') that has a Hamiltonian Cycle iff G has a vertex cover of size at most k.
For each edge (u, v) in G, G' will have one gadget W_{uv} with following node labeling:



We can get the intuition behind the construction of this gadget: There are only 3 ways that a HC can go through all nodes of one gadget. These three ways correspond to the 3 ways that an edge can be covered in the vertex cover problem, i.e. only U or V or both UV. Note that the gadget is constructed such that if a HC enters the gadget on one side, it has to leave the gadget on the same side.
2. Add k Selector Vertices in G' $S_1, S_2, \ldots, S_k$.
3. For each vertex $X \in V$, add edges in G' to join all gadgets containing X, i.e. $[X, Y_1, 6] \to [X, Y_2, 1] \to [X, Y_2, 6] \to [X, Y_3, 1] \to \ldots [X, Y_{\deg(X)}, 1]$, to form a path going through all these gadgets.



4. Join the first vertex $[X, Y_1, 1]$ and the last vertex $[X, Y_{\deg(X)}, 6]$ of above paths to all selector vertices $S_1, S_2, \ldots, S_k$.
5. Given the Hamiltonian Cycle in G', we can find the vertex cover in G by selecting the vertices connecting with the selector vertices in Hamiltonian Cycle. E.g. one hamiltonian cycle in G' could be from $S_1 \to [X, Y, 1] \to [X, Y, 6] \to [X, V, 1] \to [X, V, 3] \to [V, X, 1] \to [V, X, 6] \to [X, V, 4] \to [X, V, 6] \to [X, W, 1] \to [X, W, 3] \to [W, X, 1] \to [W, X, 6] \to [X, W, 4] \to [X, W, 6] \to [X, Z, 1] \to [X, Z, 3] \to [Z, X, 1] \to [Z, X, 6] \to [X, Z, 4] \to [X, Z, 6] \to S_2 \to [Y, Z, 1] \to [Y, Z, 3] \to [Z, Y, 1] \to [Z, Y, 6] \to [Y, Z, 4] \to [Y, Z, 6] \to [Y, X, 1] \to [Y, X, 6] \to S_1$.
We can get vertex cover set $\{X, Y\}$ because the edges $S_1 \to [X, Y, 1]$ and $S_2 \to [Y, Z, 1]$

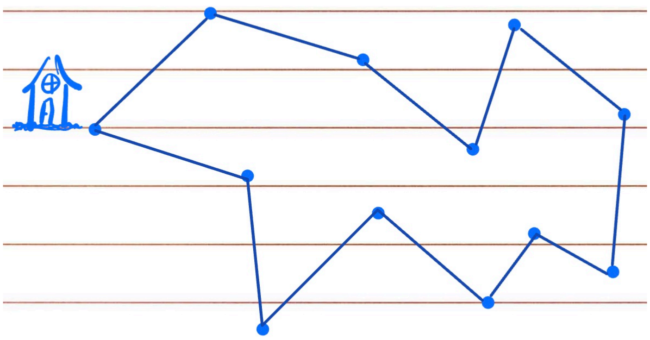
Proof. **Suppose that G = (V, E) has a vertex cover of size k.** Let the cover set be $S = \{U_1, U_2, \ldots, U_k\}$. Then we can identify the neighbors of U_i as $U_i^1, U_i^2, \ldots, U_i^{deg(U_i)}$. Then we can form a Hamiltonian Cycle in G' in this

order: If U_1^1 is not in S, $S_1 \to [U_1, U_1^1, 1] \to [U_1, U_1^1, 3] \to [U_1^1, U_1, 1] \to [U_1^1, U_1, 6] \to [U_1, U_1^1, 4] \to [U_1, U_1^1, 6] \to \dots$
If U_1^1 is in S, $S_1 \to [U_1, U_1^1, 1] \to [U_1, U_1^1, 6] \to \dots$ Go through all the edge gadgets of U_1 by one of the above two path choices. After reaching $[U_k, U_k^{deg(U_k)}, 6]$ return back to S_1 to complete the Hamiltonian Cycle.
Suppose that G' has a Hamiltonian Cycle. Then the set $S = \{U_j \in V : (S_j, [U_j, U_j^1, 1]) \in Cycle,\ 1 \le j \le k\}$ will be a vertex cover set in G, since the segments in Hamiltonian Cycle between $S_j \to S_{j+1}$ go through all gadgets corresponding to edges that are incident on U_j in G, indicating that node U_j covers all edges incident on it in G. And because the HC goes through all gadgets in G', then all corresponding edges will be covered by the nodes in the set S. □

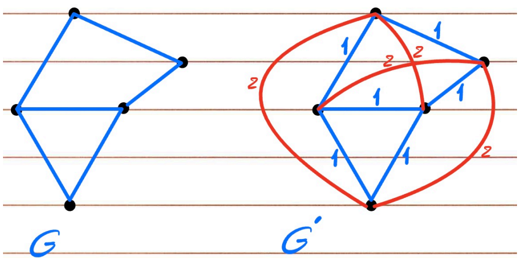
2 Traveling Salesman Problem (TSP)

2.1 Traveling Salesman Problem is NP-Complete.

Definition: Given a Complete Graph with cities and the distances between them, the traveling salesman problem is to find the shortest possible route that visits each city exactly once and returns to the origin city. i.e. order n cities in a tour $V_{i_1}, V_{i_2}, \dots, V_{i_n}$ with $i_1 = 1$ so it minimizes $\sum_{j=1}^{n-1} d(V_{i_j}, V_{i_{j+1}}) + d(V_{i_n}, V_{i_1})$.



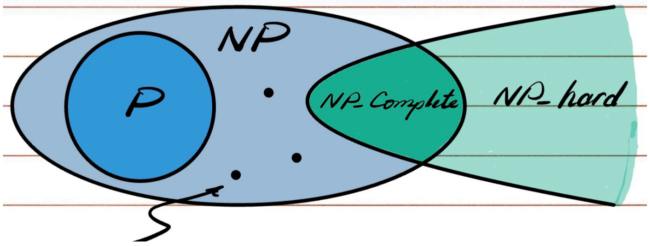
Decision Version: Given a set of distances on n cities and a bound D, is there a tour of length/cost at most D?
Show that the problem is in NP.
Certificate: An ordered list of cities on the tour of cost at most D.
Certifier: Check all we did for Hamiltonian Cycle + The sum of distances between adjacent cities in the given order is at most D.
Hamiltonian Cycle $\le_p$ Traveling Salesman Problem
Given an instance of Hamiltonian Cycle problem on graph G = (V, E), we will construct G' such that G has a Hamiltonian Cycle iff G' has a tour of cost at most $|V|$.
G' should have same set of nodes as G, and G' is fully connected (Complete) with edges already in G cost of 1 and all other new added edges cost of 2.



Proof. **Suppose that G has a Hamiltonian Cycle.** Then the tour in G' will be the same as the Hamiltonian Cycle in G, with the cost of the tour being $|V|$.
Suppose that G' has a tour of cost at most $|V|$. Then the tour in G' will be the Hamiltonian Cycle in G, because all edges in the tour are of cost 1, so these edges must be in G. □

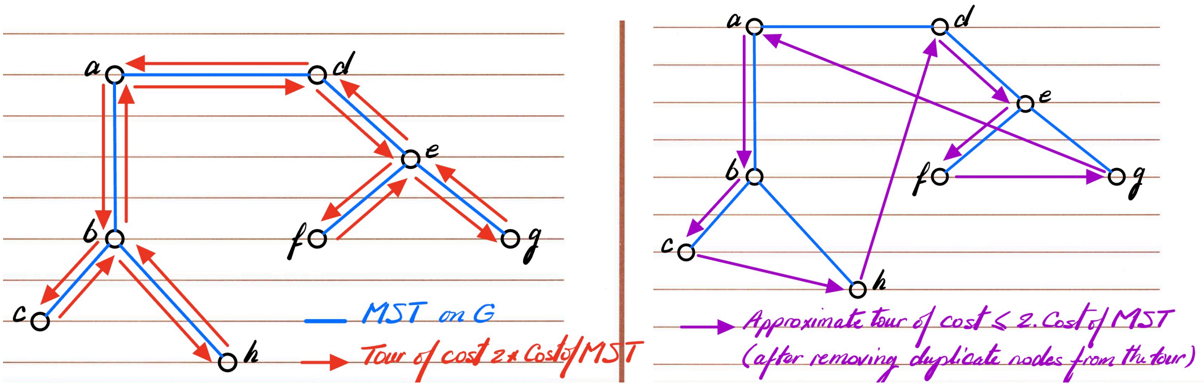
2.2 Known NP-Complete and NP-Intermediate Problems

3-SAT $\le_p$ Independent Set $\le_p$ Vertex Cover $\le_p$ Hamiltonian Cycle (Hamiltonian Path) $\le_p$ TSP
Vertex Cover $\le_p$ Set Cover Independent Set $\le_p$ Set Packing Decidison Version of 0-1 Knapsack $\le_p$ Subset Sum
NP-Intermediate: Problems in NP (Decision Problem) that are neither proven to be NP-Complete nor do we have a polynomial time solution for, e.g. Graph Isomorphism, Integer Factoring, Discrete Logarithm, etc.



2.3 Approximation Algorithm for Metric TSP

Approximation Algorithm: An algorithm that finds a solution to an optimization problem that is guaranteed to be within a factor of the optimal solution.
Metric Traveling Salesman Problem: A Fully Connected Graph with distances that holds triangle inequality, i.e. $d(u,v) \leq d(u,w) + d(w,v)$ for all u, v, w.



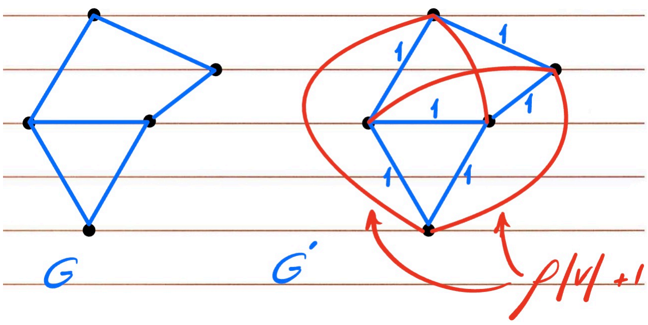
The Cost of Approximation Solution to TSP is within a factor of 2 of the cost of the optimal tour.

Proof. Cost of left initial tour = 2 * cost of MST, since triangle inequalities hold in Metric TSP, after removing duplicate nodes from left initial tour, we have: cost of right approximate tour $\leq$ cost of left initial tour $\leq$ 2 * cost of MST $\leq$ 2 * cost of Optimal tour. So our solution is called a 2-approximation since it guarantees to come to within a factor of 2 of the optimal solution. □

2.4 General TSP

Theorem: If P $\neq$ NP, then there is no Polynomial time Approximation algorithm with approximation ratio $\rho \geq 1$ for General TSP. Intuitively, in Non-Euclidean (Metric) graphs, there there could be some alternative tour that's much shorter than the one we get from MST, so there's no (known) bound on how bad the MST Approximation algorithm can be.

Proof. We will assume that such an approximation algorithm exists then we can use it to solve the Hamiltonian Cycle problem in polynomial time, i.e. P = NP.
Given an instance of Hamiltonian Cycle problem on graph G = (V, E), we will construct G' as follows: G' has the same set of nodes as G, and G' is fully connected with edges already in G in cost of 1 and all other new added edges cost of $\rho|V| + 1$.

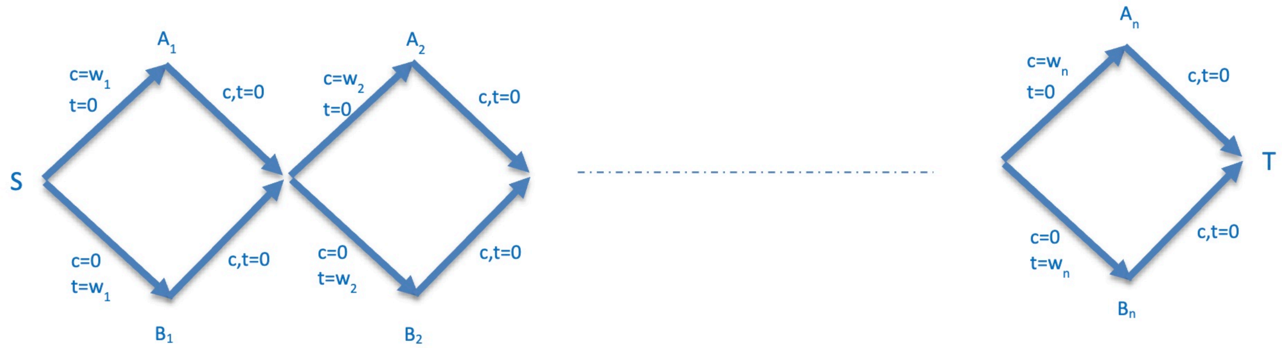


If we have a Hamiltonian Cycle in G, there will be a tour of cost |V| in G'; If we have a tour of cost at most $\rho|V|$ in G', there will be a Hamiltonian Cycle in G.
Now we can run the approximation algorithm on G'. If G has a HC, the algorithm must return a tour of cost no

more than $\rho|V|$, because it guarantees to find a tour of cost no more than a facotr of ρ from the optimal solution |V|. This tour in G' is just a Hamiltonian Cycle in G, i.e. we can solve HC problem in polynomial time. □

3 Two Upper Bounds Path Problem

Given a directed graph G = (V, E) along with positive integer times t_e and positive costs c_e on each edge. The goal is to determine if there is a path P from s to t such that the total time on the path is at most T and the total cost is at most C (both T and C are parameters to the problem). Prove that this problem is NP-complete.
Show that the problem is in NP.
Certificate: an s-t path with total cost at most C and total time at most T.
Certifier: Check that the path is from s to t, the total cost is at most C, and the total time is at most T. These can be easily done in polynomial time.
Subset Sum $\leq_p$ Two Upper Bounds Path Problem
Decision Version of Subset Sum Problem: Given n items with weight w_i, if there is a subset of items whose total weight is greater than M and less than W. (Lower and Upper Bound)
Given an instance of Subset Sum problem, we need to build a graph G such that there is a subset of items whose total weight is between M and W iff there is a s-t path in G with $\sum cost \leq W$ and $\sum time \leq \sum w_i - M$.
Use gadgets to represent each item with two paths through the gadget, with $cost = w_i$ $time = 0$ for the upper path and $cost = 0$ $time = w_i$ for the lower path. If the s-t path in G goes through the upper path of the gadget, we interpret it as selecting the item. If the s-t path goes through the lower path of the gadget, we interpret it as NOT selecting the item. String up the gadgets and set cost and time to edges as described above.



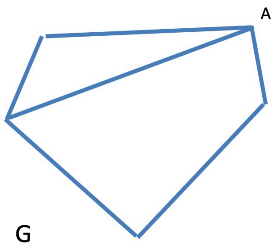
Note that $\sum cost + \sum time = \sum w_i$ for any s-t path in G, so if there is a s-t path in G with $\sum cost \leq W$ and $\sum time \leq \sum w_i - M \rightarrow \sum w_i - \sum cost \leq \sum w_i - M$, we can get $M \leq \sum cost \leq W$. So we consider the upper path of the gadget as selecting the item.

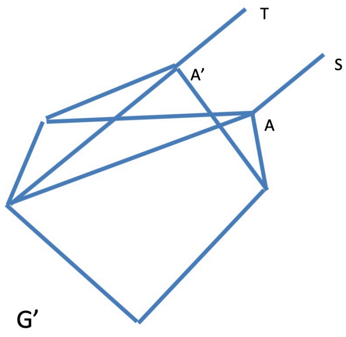
Proof. **Suppose that there is a subset of items whose total weight is between M and W.** Then we can choose the path through each gadget based on whether the item is part of the subset or not. If we go through upper path for object i, this object contributes w_i to the total cost of the path; If we go through lower path for object i, this object contributes w_i to the total time of the path. So the $\sum cost$ for the path will be the total weight of selected objects which we know $\leq W$ and the $\sum time$ of the path is actually the total weight of unselected objects that is $\leq \sum w_i - M$ because the total weight of selected objects is $\geq M$.
Suppose that there is a s-t path in G with $\sum cost \leq W$ and $\sum time \leq \sum w_i - M$. Then we can select the objects based on the path through the gadgets. If we go through the upper path for object i, we select object i; If we go through the lower path for object i, we do not select object i. Since the total cost of path is $\leq W$, then the total weight of selected objects is $\leq W$. And since the total time of the path is $\leq \sum w_i - M$, then the total weight of unselected objects is $\leq \sum w_i - M$, meaning the total weight of selected objects is $\geq M$. □

4 Hamiltonian Path

Hamiltonian Path: A path that visits each vertex exactly once, and isn't required to return to its starting point.
Prove that Hamiltonian Path is in NP.
Certificate: an ordering of nodes in G that forms a Hamiltonian Path.
Certifier: there is an edge between each pair of adjacent vertices in the given order and all nodes in G are visited by the path.
Hamiltonian Cycle $\leq_p$ Hamiltonian Path
Given any instance of Hamiltonian Cycle problem in an undirected graph G = (V, E), we will construct G' = (V', E') that has a Hamiltonian Path iff G has a Hamiltonian Cycle. The tricky part is how to identify the Cycle in G if

we only know there is a Path in G' because Hamiltonian Path does not necessarily begin and end with the adjacent node in G. To do it, we can split any one node, say node A. A and A' will have the same connecting edges as the original node A in G. Besides, we need to add two more nodes S and T connecting with A and A' respectively.



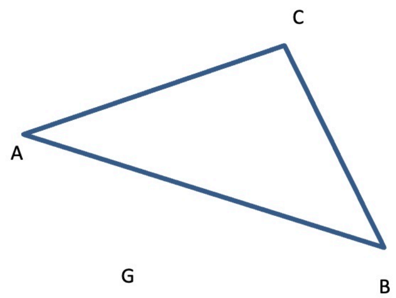


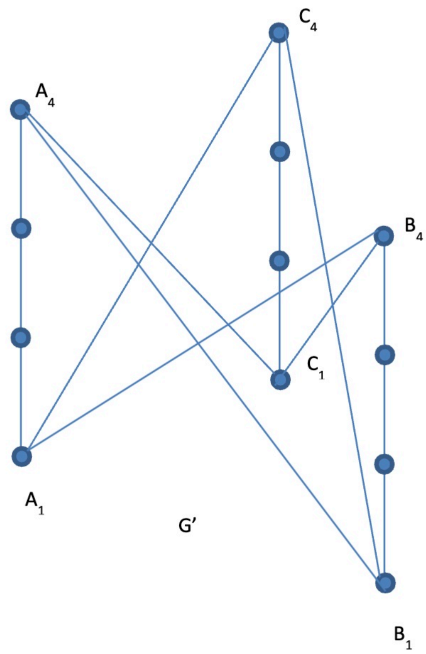
Proof. **Suppose that G has a Hamiltonian Cycle.** We can create a Hamiltonian Path in G' by splitting the cycle at node A and get a path from A to A', then a Hamiltonian Path from S to A, following the hamiltonian cycle to A' and end at T.
Suppose that G' has a Hamiltonian Path. Since S and T has a degree of 1, the path must go from S to T. Ignoring the two new edges SA and TA', this path must give us a hamiltonian cycle in G since A and A' are the same node in G. □

5 Hamiltonian Cycle in Bipartite Graph

Some NP-Complete problems are polynomial time solvable on special types of graphs, such as bipartite graphs. Others are still NP-Complete. Prove Hamiltonian Cycle in a Bipartite graph is still NP-complete.
Prove that Hamiltonian Cycle in Bipartite Graph is in NP.
Certificate: an ordering of nodes in G that forms a Hamiltonian Cycle.
Certifier: there is an edge between each pair of adjacent vertices in the given order and all nodes in G are visited by the path. There is an edge between the first and last nodes in the given order.
Hamiltonian Cycle $\leq_p$ Hamiltonian Cycle in Bipartite Graph
Given any instance of Hamiltonian Cycle problem in an undirected graph G = (V, E), we will construct Bipartite Graph G' that has a Hamiltonian Cycle iff G has a Hamiltonian Cycle.
For each node in G, we use a gadget with four nodes in G' to represent it. If two nodes A and B are connected in G, we connect nodes (A_1, B_4) (A_4, B_1) in G'.

Gadget in G' representing the node A in G





Then G' is bipartite since we place all nodes with odd index V_1, V_3 into set X and all nodes with even index

V_2, V_4 into set Y. Then all edges in G' are between nodes in X and Y. And G' has a Hamiltonian Cycle iff G has a Hamiltonian Cycle.

Proof. **Suppose that G has a Hamiltonian Cycle that goes through nodes** $V, U, \ldots, V$**.** We can create a Hamiltonian Cycle in G' by going through the gadgets corresponding to these nodes, i.e. $V_1V_2V_3V_4U_1U_2U_3U_4\ldots V_1$ since if there is a connection between node V and U, there must be a connection between nodes V_4 and U_1 in G'. **Suppose that G' has a Hamiltonian Cycle.** Then it must be of the form $V_1V_2V_3V_4U_1U_2U_3U_4\ldots V_1$ since there is no other way for the Hamiltonian Cycle to go through the nodes of each gadget. We can use the same sequence of nodes $V, U, \ldots, V$ to form a Hamiltonian Cycle in G since if there is a connection between V_4 and U_1 in G', there must be a connection between node V and U in G. □