**A useful fact.** The solution to (b) involves a sum of terms (the sum of node degrees) that we want to show is asymptotically sub-quadratic. Here's a fact that's useful in this type of situation.

> *Lemma: Let $a_1, a_2, \ldots, a_n$ be integers, each between 0 and $n$, such that $\sum_i a_i \geq \varepsilon n^2$. Then at least $\frac{1}{2}\varepsilon n$ of the $a_i$ have value at least $\frac{1}{2}\varepsilon n$.*

To prove this lemma, let $k$ denote the number of $a_i$ whose value is at least $\frac{1}{2}\varepsilon n$. Then we have $\varepsilon n^2 \leq \sum_i a_i \leq kn + \frac{1}{2}(n-k)\varepsilon n \leq kn + \frac{1}{2}\varepsilon n^2$, from which we get $k \geq \frac{1}{2}\varepsilon n$.

**(a)** For each edge $e = (u, v)$, there is a path $P_{uv}$ in $H$ of length at most $3\ell_e$ — indeed, either $e \in F$, or there was such a path at the moment $e$ was rejected. Now, given an pair of nodes $s, t \in V$, let $Q$ denote the shortest $s$-$t$ path in $G$. For each edge $(u, v)$ on $Q$, we replace it with the path $P_{uv}$, and then short-cut any loops that arise. Summing the length edge-by-edge, the resulting path has length at most 3 times that of $Q$.

**(b)** We first observe that $H$ can have no cycle of length $\leq 4$. For suppose there were such a cycle $C$, and let $e = (u, v)$ be the last edge added to it. Then at the moment $e$ was considered, there was a $u$-$v$ path $Q_{uv}$ in $H$ of at most three edges, on which each edge had length at most $\ell_e$. Thus $\ell_e$ is not less than a third the length of $Q_{uv}$, and so it should not have been added.

This constraint implies that $H$ cannot have $\Omega(n^2)$ edges, and there are several different ways to prove this. One proof goes as follows. If $H$ has at least $\varepsilon n^2$ edges, then the sum of all degrees is $2\varepsilon n^2$, and so by our lemma above, there is a set $S$ of at least $\varepsilon n$ nodes each of whose degrees is at least $\varepsilon n$. Now, consider the set $Q$ of all pairs of edges $(e, e')$ such $e$ and $e'$ each have an end equal to the same node in $S$. We have $|Q| \geq \varepsilon n\binom{\varepsilon n}{2}$, since there are at least $\varepsilon n$ nodes in $S$, and each contributes at least $\binom{\varepsilon n}{2}$ such pairs. For each edge pair $(e, e') \in Q$, they have one end in common; we *label* $(e, e')$ with the pair of nodes at their other ends. Since $|Q| > \binom{n}{2}$ for sufficiently large $n$, the pigeonhole principle implies that some two pairs of edges $(e, e'), (f, f') \in Q$ receive the same label. But then $\{e, e', f, f'\}$ constitutes a four-node cycle.

For a second proof, we observe that an $n$-node graph $H$ with no cycle of length $\leq 4$ must contain a node of degree at most $\sqrt{n}$. For suppose not, and consider any node $v$ of $H$. Let $S$ denote the set of neighbors of $v$. Notice that there is no edge joining two nodes of $S$, or we would have a cycle of length 3. Now let $N(S)$ denote the set of all nodes with a neighbor in $S$. Since $H$ has no cycle of length 4, each node in $N(S)$ has exactly one neighbor in $S$. But $|S| > \sqrt{n}$, and each node in $S$ has $\geq \sqrt{n}$ neighbors other than $v$, so we would have $|N(S)| > n$, a contradiction. Now, if we let $g(n)$ denote the maximum number of edges in an $n$-node graph with no cycle of length 4, then $g(n)$ satisfies the recurrence $g(n) \leq g(n-1) + \sqrt{n}$ (by deleting the lowest-degree node), and so we have $g(n) \leq n^{3/2} = o(n^2)$.

[1] ex616.972.80