A conjecture (M. Steel, April 2001)

Suppose a $0, 1$ random variable evolves on a tree T with n labelled leaves (and the remaining vertices unlabelled and of degree 3) under the usual symmetric Markov model. Suppose that, for each edge of T, the probability of a net transition (from $0 \rightarrow 1$ or $1 \rightarrow 0$) between the endpoints of the edge lies between f and g, where $0 < f \leq g < 0.5$. Now, suppose we independently evolve k such variables under this model, and record just the values taken at the leaves of the tree (this gives us n binary sequences of length k). In (Erdös et al. 1999) it is shown that, without any knowledge about T, f or g, and for any $\epsilon > 0$, it is possible to correctly reconstruct T with probability at least $1 - \epsilon$ provided

$$k > \frac{c \log(n)}{f^2(1-2g)^{d(T)}},$$

where c depends only on ϵ and $d(T)$ is a quantity that typically is $O(\log(\log(n))$ and is always $O(\log(n))$. Furthermore there is a polynomial time algorithm for carrying out this reconstruction. Following a result of (Evans et al. 2000) we offer the following:

Conjecture Provided $g < \frac{1}{2}(1 - \frac{1}{\sqrt{2}})$ there is a method for correctly reconstructing T with probability at least $1 - \epsilon$ provided

$$k > \frac{c' \log(n)}{f^2},$$

where c' depends only on ϵ.