This homework will step you through the proof of Lemma 1 in the Jerrum (1995) paper. I mostly use Jerrum’s notation, except that the number of colors is $q$, not $k$.

The setting
The set of available “colors” is $C = \{1, 2, \ldots, q\}$. We have a graph $G$ on a finite set of vertices $V = \{v_1, \ldots, v_n\}$ with edge set $E = \{e_1, \ldots, e_R\}$. The degree of a vertex $v$, denoted by $\deg(v)$, is the number of edges having $v$ as one of the endpoints. The neighborhood of vertex $v$ is the set of vertices connected to $v$ by an edge,

$$N(v) := \{w \in V : w \neq v \text{ and } \{v, w\} \in E\}.$$

A coloring of the graph is a map $\sigma : V \to C$. The coloring is proper if the two vertices that make up each edge are assigned different colors, that is, if

$$\sigma(v) \notin \{\sigma(w) : w \in N(v)\} \quad \text{for every } v \in V.$$

Denote the set of all proper colorings of the graph by $S$.

It is easy to see (by means of a greedy coloring method) that $S$ is nonempty if $q > \Delta$. For $q > 2\Delta$ Jerrum’s algorithm generates observations from the uniform distribution $\pi$ on $S$ by means of a Markov chain that converges very rapidly towards $\pi$.

Transition probabilities
The transition probabilities $P(\sigma, \tau)$ for the chain are defined implicitly by a random method for producing a new coloring $\tau$ from a coloring $\sigma$.

First define a function $\text{TRY}$ that generates a new coloring $\tau$ given a new color $c$, a vertex $v$, and a current coloring $\sigma$. Define $\tau = \text{TRY}(c, v, \sigma)$ by

(i) if $c \in \{\sigma(w) : w \in N(v)\}$ then $\tau = \sigma$

(ii) if $c \notin \{\sigma(w) : w \in N(v)\}$ then $\tau(v) = c$ and $\tau(w) = \sigma(w)$ for all $w$ not equal to $v$.

In other words, $\text{TRY}(c, v, \sigma)$ changes the color of $v$ to $c$, provided the resulting coloring is proper. If the proposed change would create an improper coloring, $\sigma$ is left unchanged.

---

You might need to access the site via http://sfx.library.yale.edu/sfx_local/azlist.
Here is the random procedure corresponding to $P(\sigma, \tau)$:

(a) Choose a vertex $V$ at random from (the uniform distribution on) $\mathcal{V}$.
(b) Choose a color $C$ at random from (the uniform distribution on) $\mathcal{C}$.
(c) Define $\tau = \text{TRY}(C, V, \sigma)$.

The coupling

Jerrum’s proof works by creating a coupling of a Markov chain $\{Y_t : t \geq 0\}$ with state space $\mathcal{S}$ and an arbitrary (but fixed) initial distribution $\mu$ and another Markov chain $\{X_t : t \geq 0\}$ with state space $\mathcal{S}$ and initial distribution $\pi$, using a method a little like the coupling used to prove the BLT.

Write $\text{info}_t$ for the information corresponding to everything that has happened up to the completion of step $t$. Initially $X_0 \sim \pi$ and $Y_0 \sim \mu$, independently. After the completion of step $t$ the $X_{t+1}$ and $Y_{t+1}$ values are coupled as follows

(a) Choose a new vertex $V_{t+1}$ at random from (the uniform distribution on) $\mathcal{V}$, independently of $\text{info}_t$.
(b) Independently of $V_{t+1}$ and of $\text{info}_t$, choose a color $C$ at random from (the uniform distribution on) $\mathcal{C}$.
(c) Define $X_{t+1} = \text{TRY}(C_{t+1}, V_{t+1}, X_t)$. Based on $\text{info}_t$ and $V_{t+1}$, construct a one-to-one function $g : \mathcal{C} \rightarrow \mathcal{C}$ then define $Y_{t+1} = \text{TRY}(g(C_{t+1}), V_{t+1}, Y_t)$.

Remark. The random color $g(C_{t+1})$ is also uniformly distributed on $\mathcal{C}$, independently of $\text{info}_t$ and $V_{t+1}$. The change from $Y_t$ to $Y_{t+1}$ still follows the $P$ transition probabilities; marginally, $Y$ is still just a Markov chain with initial distribution $\mu$ and transition matrix $P$.

If $X_t(V_{t+1}) \neq Y_t(V_{t+1})$ then $g$ is taken to be the identity map (that is, $g(c) = c$ for all $c$ in $\mathcal{C}$) and $Y_{t+1} = \text{TRY}(C_{t+1}, V_{t+1}, Y_t)$. The cleverness in Jerrum’s algorithm comes from the choice of $g$ for the cases where $X_t(V_{t+1}) = Y_t(V_{t+1})$.

We need some notation for the case where $X_t$ and $Y_t$ agree on the color given to vertex $V_{t+1}$. After step $t$ of the algorithm the vertex set $\mathcal{V}$ finds itself partitioned into two subsets: $A_t = \{v \in \mathcal{V} : X_t(v) = Y_t(v)\}$, the set of vertices where the two colorings agree; and $D_t = \{v \in \mathcal{V} : X_t(v) \neq Y_t(v)\}$, the set where they disagree.

For the sake of notational clarity let me omit some subscripts $t$ and $t+1$, writing $V$ for $V_{t+1}$ and so on. Write $\mathcal{K}_X(V)$ for $\{X_t(w) : w \in \mathcal{N}(V)\}$, the set of colors used by $X_t$ for the neighbors of $V$. Define $\mathcal{K}_Y(V)$ analogously. Then $\mathcal{K}_{X\cap Y}(V) := \mathcal{K}_X(V) \cap \mathcal{K}_Y(V)$ is the set of colors used only by $X_t$ for the
neighbors of $V$ and $\mathcal{C}_Y(V) := K_Y(V) \cap K^c_X(V)$ is the set of colors used only by $Y_t$ for the neighbors of $V$.

The $g$ is chosen in a way that discourages changes that would turn an agreement at vertex $V$ into a disagreement.

If either $\mathcal{C}_X(V)$ or $\mathcal{C}_Y(V)$ is empty let $g$ be the identity map on $\mathcal{C}$.

If both $\mathcal{C}_X(V)$ and $\mathcal{C}_Y(V)$ are nonempty the smaller of the two sets of colors drives the construction. For concreteness suppose $0 < \#\mathcal{C}_X(V) \leq \#\mathcal{C}_Y(V)$, so that the distinct colors in $\mathcal{C}_X(V)$ can be enumerated as $c_1, c_2, \ldots, c_r$ and the distinct colors in $\mathcal{C}_Y(V)$ can be enumerated as $c'_1, c'_2, \ldots, c'_{r'}$, with $r \leq r'$. Define $g$ by
\[ g(c_i) = c'_i \quad \text{and} \quad g(c'_i) = c_i \quad \text{for } i = 1, 2, \ldots, r \]
\[ g(c) = c \quad \text{otherwise} \]

The picture on the left represents the partition of $\mathcal{C}$, with the dependence on $v = V_{t+1}$ omitted, for the case $0 < \#\mathcal{C}_X(v) \leq \#\mathcal{C}_Y(v)$. The picture on the right represents $V$, showing only the the neighbors of $v$ when $X_t(v) = Y_t(v)$, with $t$ subscripts omitted.

If $\#\mathcal{C}_X(V) > \#\mathcal{C}_Y(V) > 0$, reverse the roles of $X$ and $Y$ in the construction, so that $g$ maps $\mathcal{C}_Y(V)$ onto a subset of $\mathcal{C}_X(V)$.

The meeting of the chains
At the random time $T := \min\{t : D_t = \emptyset\}$ the colorings agree for all vertices. At that time both $\mathcal{C}_X(v)$ and $\mathcal{C}_Y(v)$ are empty for every $v$ in $V$. Subsequently, the algorithm merely modifies the color scheme without creating any new disagreements.

The details of Jerrum’s argument appear in the following Problems.

References

Homework problems

Write \( f_t \) for \( \#D_t \), the number of vertices where the colorings disagree after completion of step \( t \). The main idea is to show that \( \mathbb{E} f_{t+1} \leq (1 - \alpha) \mathbb{E} f_t \) where

\[
\alpha := \frac{q - 2\Delta}{nq} > 0,
\]

which will produce a rapidly decreasing bound on the total variation distance between \( \pi \) and the distribution of \( Y_t \).

\[1\] (5 points) For each \( \sigma \) and \( \tau \) in \( S \) show that \( P(\sigma, \tau) = P(\tau, \sigma) \). Deduce that the uniform distribution \( \pi \) on \( S \) is the stationary distribution.

If \( P(\sigma, \tau) > 0 \) with \( \sigma \neq \tau \) then there must exist exactly one vertex \( v \) for which \( \sigma(v) \neq \tau(v) \). Of course neither \( \sigma(v) \) nor \( \tau(v) \) can be included in the set of colors used by both \( \sigma \) and \( \tau \) for \( N(v) \). To get from \( \sigma \) to \( \tau \) the algorithm must select vertex \( v \) (probability \( n^{-1} \)) and color \( \tau(v) \) (probability \( q^{-1} \)); to get from \( \tau \) to \( \sigma \) the algorithm must select vertex \( v \) and color \( \sigma(v) \). Both pairs of choices have the same probability, \( (nq)^{-1} \).

Time reversibility forces the uniform distribution to be the stationary distribution, as in HW3.2(d).

\[2\] (5 points) Explain why

\[
TV_t := \max_{B \subseteq S} |P\{Y_t \in B\} - \pi(B)| \leq P\{T > t\}
\]

Hint: The argument is similar to the one used for the proof of the BLT.

For each \( B \subseteq S \),

\[
|P\{Y_t \in B\} - \pi(B)| = |P\{Y_t \in B\} - P\{X_t \in B\}|
\leq |P\{Y_t \in B, T \leq t\} - P\{X_t \in B, T \leq t\}|
+ |P\{Y_t \in B, T > t\} - P\{X_t \in B, T > t\}|
\]

From time \( T \) onwards the two chains stay the same; they give the same color to each vertex. Thus the first term after the inequality equals zero. For the other term note that

\[
-P\{T > t\} \leq P\{Y_t \in B, T > t\} - P\{X_t \in B, T > t\} \leq P\{T > t\},
\]

which gives \( |P\{Y_t \in B, T > t\} - P\{X_t \in B, T > t\}| \leq P\{T > t\} \), as in the proof of the BLT. Some of you derived only the weaker bounder \( 2P\{T > t\} \).

Some of you only derived an upper bound for \( P\{Y_t \in B\} - \pi(B) \). (Actually, when paired with the analogous bound for \( B \) replaced by \( B^c \), the one-sided bound also leads to the desired two-sided bound.)
[3] (10 points) Explain why \( T = \min\{t : f_t = 0\} \)

\[
\mathbb{P}\{T > t\} \leq \mathbb{P}\{f_t \neq 0\} \leq \mathbb{E}f_t.
\]

**Hint:** The random variable \( f_t \) takes only nonnegative integer values.

By definition, \( D_t \neq \emptyset \) (implying \( f_t > 0 \)) for \( t < T \). And, because \( X_t = Y_t \) for all \( t \geq T \), the set \( D_t \) stays empty from time \( T \) onwards. Consequently, \( f_t \neq 0 \) if and only if \( t < T \). The first inequality is actually an equality: \( \mathbb{P}\{T > t\} \leq \mathbb{P}\{f_t \neq 0\} \). The second inequality follows by taking expectations on both sides of \( \mathbb{I}\{f_t \geq 1\} \leq f_t \). Some of you failed to notice that \( f_t = 0 \) for \( t \geq T \).

[4] (5 points) Explain why \( |f_{t+1} - f_t| \leq 1 \) always, with \( f_{t+1} \leq f_t \) if \( V_{t+1} \in D_t \) and \( f_{t+1} \geq f_t \) if \( V_{t+1} \in A_t \).

Once again abbreviate \( V_{t+1} \) to \( V \). At each step the \( X \)- and \( Y \)-colors can changed only at vertex \( V \). For an increase in the number of disagreements, it must be that \( X_t(V) = Y_t(V) \) but \( X_{t+1}(V) \neq Y_{t+1}(V) \). For a decrease in the number of disagreements, it must be that \( X_t(V) \neq Y_t(V) \) but \( X_{t+1}(V) = Y_{t+1}(V) \). Problems [6] and [7] provide more details about when \( f_{t+1} = f_t \pm 1 \).

[5] (5 points) For each vertex \( v \), define

\[
d_t(v) := \begin{cases} 
\#(N(v) \cap A_t) & \text{if } v \in D_t \\
\#(N(v) \cap D_t) & \text{if } v \in A_t.
\end{cases}
\]

Equivalently, \( d_t(v) \) is the number of edges (with \( v \) as one of the endpoints) that join a point in \( A_t \) to a point in \( D_t \). Explain why the total number of edges that join a point in \( A_t \) to a point in \( D_t \) equals

\[
m_t := \sum_{v \in A_t} d_t(v) = \sum_{v \in D_t} d_t(v).
\]

The first expression for \( m_t \) comes from working through the list of all the vertices \( v \) in \( A_t \) while counting how many vertices in \( D_t \) are connected to each \( v \) on the list. The second expression comes from working through the list of all the vertices \( v \) in \( D_t \) while counting how many vertices in \( A_t \) are connected to each \( v \) on the list.

[6] Consider the case where \( V_{t+1} \in A_t \). Abbreviate \( V_{t+1} \) to \( V \) and \( C_{t+1} \) to \( C \).

(i) (5 points) Explain why \( f_{t+1} = f_t \) if \( \mathcal{C}_X(V) = \mathcal{C}_Y(V) = \emptyset \).

If both \( \mathcal{C}_X(V) \) and \( \mathcal{C}_Y(V) \) are empty then \( K_X(V) = K_Y(V) \), that is, the \( X \)- and \( Y \)-colorings agree for all the neighbors of \( V \). The color \( C_{t+1} \) is either accepted or rejected by both chains. Both \( X \) and \( Y \) still give the same color to vertex \( V \).
Without loss of generality suppose \( \#C_X(V) \leq \#C_Y(V) \) and \( \mathcal{C}_Y(V) \neq \emptyset \) for the rest of this problem. It is possible that \( \mathcal{C}_X(V) \) is empty, which would make (ii) a bit easier.

As I explained in class, the “Without loss of generality” is harmless if we already know that \( V_{t+1} = v \), for a specific \( v \) in \( A_t \). The analysis for the case where \( \#C_X(v) > \#C_Y(v) \) is essentially the same as for the case \( \#C_Y(v) \geq \#C_X(v) \): one merely has to interchage the roles of \( X \) and \( Y \). However, the way the question was written did suggest that \( \#C_Y(v) \geq \#C_X(v) \) for every \( v \) in \( A_t \), which need not be true.

The following three parts should refer to \( V_{t+1} \) equal to a (temporarily) fixed \( v \) for which \( 0 < \#C_X(v) \leq \#C_Y(v) \).

(ii) (15 points) If \( C \notin \mathcal{C}_Y(V) \), show that \( f_{t+1} = f_t \). Hint: Consider separately the cases where \( C \in (K_X(V) \cup K_Y(V))^c \) or \( K_X(V) \cap K_Y(V) \) or \( \mathcal{C}_X(V) \).

If \( C \in (K_X(V) \cup K_Y(V))^c \) then neither chain has color \( C \) used for any of the neighbors of \( v \) and \( g(C) = C \), resulting in \( X_{t+1}(v) = C = Y_{t+1}(v) \).

If \( C \in K_X(V) \cap K_Y(V) \) then \( g(C) = C \) and neither chain accepts color \( C \), leaving \( X_{t+1}(v) = X_t(v) = Y_t(v) = Y_{t+1}(v) \).

If \( C \in \mathcal{C}_X(v) \) then chain \( X \) rejects \( C \) as one of the colors it already has for a neighbor of \( v \). Chain \( Y \) rejects the color \( g(C) \), which belongs to \( \mathcal{C}_Y(v) \) because of the way the \( g \) was constructed.

(iii) (5 points) Explain why \( f_{t+1} \) might equal \( 1 + f_t \) if \( C \in \mathcal{C}_Y(V) \).

If \( C = c_i \) for some \( i \leq r \) then \( g(C) = c_i \in \mathcal{C}_X(v) \). Both chains accept the new color for vertex \( v \) and \( f_{t+1} = f_t \). However if \( C = c'_i \) for some \( r < i \leq r' \) then the \( Y \)-chain rejects the color \( C \) but the \( X \)-chain accepts \( g(C) = c'_i \) because none of the neighbors of \( v \) have \( X \)-color \( c'_i \). The result is a different coloring of vertex \( v \) by the two chains.
(iv) (5 points) Explain why $\#C_Y(V) \leq d_t(V)$ when $V \in A_t$.

Again argue for $V$ equal to a fixed $v$ in $A_t$. The colors in $C_Y(v)$ can appear only on (a subset of the) neighbors of $v$ that lie in $D_t$. That is, for each color in $C_Y(v)$ there is at least one edge from $v$ to a neighbor in $D_t$. The number of colors in $C_Y(v)$ cannot exceed $d_t(V)$, the total number of edges from $v$ to neighbors in $D_t$. (A similar argument gives $\#C_X(v) \leq d_t(v)$.)

(v) (10 points) Deduce that

$$\mathbb{P}\{f_{t+1} = 1 + f_t \mid \text{info}_t\} \leq \mathbb{P}\{V \in A_t, C \in C_Y(V) \mid \text{info}_t\} \leq \frac{m_t}{nq}.$$  

Explain your reasoning in detail.

It is better to argue in the following way. First condition on $V_{t+1}$. By Problem [4], only if $v \in A_t$ can we have $f_{t+1} = f_t + 1$.

$$\mathbb{P}\{f_{t+1} = 1 + f_t \mid \text{info}_t\}$$

$$= \mathbb{P}\{f_{t+1} = 1 + f_t, V_{t+1} \in A_t \mid \text{info}_t\}$$

$$= \sum_{v \in A_t} \mathbb{P}\{V_{t+1} = v \mid \text{info}_t\} \mathbb{P}\{f_{t+1} = 1 + f_t \mid \text{info}_t, V_{t+1} = v\}$$

For every $v$ in $V$ the probability $\mathbb{P}\{V_{t+1} = v \mid \text{info}_t\}$ is equal to $n^{-1}$. For the analysis of $\mathbb{P}\{f_{t+1} = 1 + f_t \mid \text{info}_t, V_{t+1} = v\}$ suppose, “without loss of generality”, that $0 < \#C_X(v) \leq \#C_Y(v)$. Then

$$\mathbb{P}\{f_{t+1} = 1 + f_t \mid \text{info}_t, V_{t+1} = v\}$$

$$\leq \mathbb{P}\{C_{t+1} \in C_Y(v) \mid \text{info}_t, V_{t+1} = v\}$$

using without loss of generality . . .

$$\leq d_t(v)/q$$  by (iv).

Substitute into <1> to deduce that

$$\mathbb{P}\{f_{t+1} = 1 + f_t \mid \text{info}_t\} \leq \sum_{v \in A_t} n^{-1}d_t(v)/q = (nq)^{-1}m_t.$$  

Some of you claimed that the events $\{V_{t+1} \in A_t\}$ and $\{C_{t+1} \in C(V_{t+1})\}$ are independent, which is false. Conditioning on $V_{t+1} = v \in A_t$ does let you focus only on the variability in $C_{t+1}$. Some of you also claimed that the
event \( \{ V_{t+1} \in A_t, C_{t+1} \in C(V_{t+1}) \} \) doesn’t depend on info\(_t\), which is also false because the sets \( A_t \) and \( D_t \) are defined using info\(_t\).

[7] Consider the case where \( V_{t+1} \in D_t \). Abbreviate \( V_{t+1} \) to \( V \) and \( C_{t+1} \) to \( C \).

(i) (5 points) For each \( v \) in \( D_t \) explain why

\[
P(f_{t+1} = f_t - 1 \mid V = v, \text{info}_t) \geq P(C \notin K_X(v) \cup K_Y(v) \mid V = v, \text{info}_t)
\]

If \( C_{t+1} \) is in neither the \( K_X(v) \) nor the \( K_Y(v) \) list then \( g(C_{t+1}) = C_{t+1} \) and both chains accept the new color for vertex \( v \). If \( v \in D_t \) the number of disagreements is decreased by 1.

(ii) (5 points) Explain why

\[
\#(K_X(v) \cup K_Y(v)) \leq 2\Delta - d_t(v) \quad \text{for all } v \in D_t.
\]

Hint: What is the largest number of distinct colors that can be contributed by \( N(v) \cap D_t \)?

![Diagram](image.png)

By definition, \( d_t(v) = \#(N(v) \cap A_t) \) and \( \#N(v) \leq \Delta \). Thus

\[
\#(N(v) \cap D_t) = \#N(v) - \#(N(v) \cap A_t) \leq \Delta - d_t(v).
\]

The vertices in \( N(v) \cap D_t \) can contribute at most \( 2 \times \#(N(v) \cap D_t) \) distinct colors to \( K_X(v) \cup K_Y(v) \). The vertices in \( N(v) \cap A_t \) can contribute at most \( d_t(v) \) distinct colors. Together they contribute at most \( 2 \times (\Delta - d_t(v)) + d_t(v) \leq 2\Delta - d_t(v) \) colors.

Many of you argued that the vertices in \( N(v) \cap A_t \) all carry colors that appear in \( K_X(v) \cap K_Y(v) \). However, there is no guarantee that each vertex in \( N(v) \cap A_t \) has a distinct color. In the extreme case, both \( X \) and \( Y \) could use a single color for every vertex in \( N(v) \cap A_t \). The number of colors used for vertices in \( N(v) \cap A_t \) need not equal \( d_t(v) = \#(N(v) \cap A_t) \).

(iii) (5 points) Deduce that

\[
P(f_{t+1} = f_t - 1 \mid \text{info}_t) \geq \alpha f_t + \frac{m_t}{nq}
\]

Again condition on \( V_{t+1} \). If \( V_{t+1} = v \in D_t \) and \( C_{t+1} \notin K_X(v) \cup K_Y(v) \) then both \( X \) and \( Y \) accept the new color, creating one fewer disagreement.
Thus

\[ P(f_{t+1} = f_t - 1 | \text{info}_t) \]
\[ \geq P\{V \in D_t, C \notin K_X(V) \cup K_Y(V) | \text{info}_t\} \]
\[ = \sum_{v \in D_t} P\{V = v | \text{info}_t\} P\{C \notin K_X(v) \cup K_Y(v) | \text{info}_t, V = v\} \]
\[ \geq n^{-1} \sum_{v \in D_t} \left(1 - \frac{\#(K_X(v) \cup K_Y(v))}{q}\right) \]
\[ \geq \frac{1}{n} \sum_{v \in D_t} \left(1 - \frac{2\Delta - d_t(v)}{q}\right) \]
\[ \geq \frac{1}{n} \left(1 - \frac{2\Delta}{q}\right) f_t + \frac{m_t}{nq} \quad \text{because } f_t = \#D_t \]
\[ = \alpha f_t + \frac{m_t}{nq} \]

Some of you failed to condition properly, treating \( V \) (or \( v \)) as random in one line and fixed in the next, with summations over \( v \in D_t \) appearing out of nowhere. Some of you also confused \( I\{C \notin K_X(v) \cup K_Y(v)\} \) with \( \#(K_X(v) \cup K_Y(v)) \).

[8] Combine the last two results.

(i) (5 points) Show that

\[ E(f_{t+1} - f_t | \text{info}_t) = P\{f_{t+1} = 1 + f_t | \text{info}_t\} - P(f_{t+1} = f_t - 1 | \text{info}_t) \]
\[ \leq -\alpha f_t \]

From [6](v) and [7](iii),

\[ P\{f_{t+1} = 1 + f_t | \text{info}_t\} - P(f_{t+1} = f_t - 1 | \text{info}_t) \leq \frac{m_t}{nq} - \left(\alpha f_t + \frac{m_t}{nq}\right) \]

(ii) (5 points) Deduce that \( E f_{t+1} \leq (1 - \alpha) E f_t \) for all \( t \).

Multiply both sides of \( E(f_{t+1} - f_t | \text{info}_t) \leq -\alpha f_t \) by \( P\{\text{info}_t\} \) then sum over all possible values for \( \text{info}_t \) to deduce that \( E(f_{t+1} - f_t) \leq -\alpha E f_t \). Some of you failed to distinguish between \( E(f_{t+1} - f_t | \text{info}_t) \) and \( E(f_{t+1} - f_t) \).

(iii) (5 points) Deduce that

\[ E f_t \leq (1 - \alpha)^t E f_0 \leq n(1 - \alpha)^t. \]

Notice that the upper bound does not depend on \( \mu \).

Use \( f_0 \leq \#V = n \) together with recursive appeals to the bound from (ii).


\[ TV_t \leq n(1 - \alpha)^t \]

Remark. You could solve to find how large \( t \) must be in order to make
\[ TV_t \leq \epsilon, \] for any given \( \epsilon > 0 \).