$O(\sqrt{T})$ upperbound on T_3

March 19, 2023

Lemma 0.1 ($O(\sqrt{T})$ upperbound on T_3)**.**

$$T_3 = \sum_{t=1}^{T} \frac{\eta}{2} \cdot g_t^T X_t g_t \leq \frac{4nG_\infty \eta}{\hat{\epsilon}^3} \sqrt{T}$$

where, $\|g_t\|_\infty \leq G_\infty$ *and parameters* $\epsilon = \hat{\epsilon} G_\infty \sqrt{T}$, $\lambda_t = G_\infty \sqrt{t}$ *in Algorithm 1.*

Proof. Using Theorem 3.1, nonzero entries of X_t can be written as follows:

$$(X_t)_{ii} = \frac{1}{H_{ii}} \left(1 + \sum_{(i,j) \in E_{\mathcal{G}}} \frac{H_{ij}^2}{H_{ii} H_{jj} - H_{ij}^2} \right)$$
$$(X_t)_{ii+1} = -\frac{H_{ii+1}}{H_{ii} H_{i+1i+1} - H_{ii+1}^2}$$

where, $E_{\mathcal{G}}$ denote the set of edges of the chain graph $\mathcal{G}$ in Theorem 3.1. Also, for brevity, the subscript is dropped for H_t. Let $\hat{X}_t = \sqrt{\mathrm{diag}(H)} X_t \sqrt{\mathrm{diag}(H)}$, then $\hat{X}_t$ can be written as

$$(\hat{X}_t)_{ii} = \left(1 + \sum_{(i,j) \in E_{\mathcal{G}}} \frac{\hat{H}_{ij}^2}{1 - \hat{H}_{ij}^2} \right),$$
$$(\hat{X}_t)_{ii+1} = -\frac{\hat{H}_{ii+1}}{1 - \hat{H}_{ii+1}^2},$$

where, $\hat{H}_{ij} = H_{ij}/\sqrt{H_{ii} H_{jj}}$. Note that $\hat{X}_t \preceq \|\hat{X}_t\|_2 I_n \preceq \|\hat{X}_t\|_\infty I_n$, using $\max\{|\lambda_1(\hat{X}_t)|, \ldots, |\lambda_n(\hat{X}_t)|\} \leq \|\hat{X}_t\|_\infty$. So we upperbound $\|\hat{X}_t\|_\infty = \max_{i \in [n]}\{|(\hat{X}_t)_{11}| + |(\hat{X}_t)_{12}|, \ldots, |(\hat{X}_t)_{ii-1}| + |(\hat{X}_t)_{ii}| + |(\hat{X}_t)_{ii+1}|, \ldots, |(\hat{X}_t)_{nn}| + |(\hat{X}_t)_{nn-1}|\}$. Individual terms $|(\hat{X}_t)_{ii-1}| + |(\hat{X}_t)_{ii}| + |(\hat{X}_t)_{ii+1}|$ can be written as follows:

$$
\begin{aligned}
\sum_{(i,j)\in E_{\mathcal{G}}} |(\hat{X}_t)_{ij}| &= 1 + \sum_{(i,j)\in E_{\mathcal{G}}} \frac{\hat{H}_{ij}^2}{1-\hat{H}_{ij}^2} + \frac{|\hat{H}_{ij}|}{1-\hat{H}_{ij}^2} \\
&= 1 + \sum_{(i,j)\in E_{\mathcal{G}}} \frac{|\hat{H}_{ij}|}{1-|\hat{H}_{ij}|} \\
&\leq 2 \max_{i\in[n-1]} \frac{1}{1-|\hat{H}_{ii+1}|}
\end{aligned}
$$

The last inequality is because $|\hat{H}_{ij}| \leq 1$. Thus, $\|\hat{X}_t\|_\infty \leq 2\max_{i\in[n-1]} \frac{1}{1-|\hat{H}_{ii+1}|}$. Now

$$
\begin{aligned}
g_t^T X_t g_t &\leq g_t^T \mathrm{diag}(H_t)^{-1} \hat{X}_t \mathrm{diag}(H_t)^{-1} g_t \\
&\leq \|\hat{X}_t\|_\infty \|\mathrm{diag}(H_t)^{-1/2} g_t\|_2^2 \\
&\leq 2 \max_{i\in[n-1]} \frac{1}{1-|\hat{H}_{ii+1}|} g_t^T \mathrm{diag}(H_t)^{-1} g_t.
\end{aligned}
$$

Using $\mathrm{diag}(H_t) \succeq \epsilon I_n$ (step 1 in Algorithm 1), where $\epsilon = \hat{\epsilon} G_\infty \sqrt{T}$ as set in Lemma A.8, gives

$$
\begin{aligned}
g_t^T X_t g_t &\leq 2 \max_{i\in[n-1]} \frac{1}{1-|\hat{H}_{ii+1}|} \frac{\|g_t\|_2^2}{\hat{\epsilon} G_\infty \sqrt{T}} \\
&\leq 2 \max_{i\in[n-1]} \frac{n G_\infty}{\hat{\epsilon}(1-|\hat{H}_{ii+1}|)\sqrt{T}} \\
&\leq \frac{2nG_\infty}{\hat{\epsilon}(1-\beta)\sqrt{T}} \qquad (1)
\end{aligned}
$$

Now we show that $1/(1-\beta)$ is $O(1)$.

$$
\begin{aligned}
1/(1-\beta) &= \frac{1+\beta}{1-\beta^2} \\
&\leq \max_t \max_{i\in[n]} \frac{(H_t)_{ii}(H_t)_{i+1i+1}}{(H_t)_{ii}(H_t)_{i+1i+1} - (H_t)_{ii+1}^2} \\
&\leq \max_t \max_{i\in[n]} \frac{(H_t)_{ii}(H_t)_{i+1i+1}}{\det\left(\begin{bmatrix} (H_t)_{ii} & (H_t)_{ii+1} \\ (H_t)_{i+1i} & (H_t)_{i+1i+1} \end{bmatrix}\right)} \qquad (2)
\end{aligned}
$$

Note that $\begin{bmatrix} (H_t)_{ii} & (H_t)_{ii+1} \\ (H_t)_{i+1i} & (H_t)_{i+1i+1} \end{bmatrix} \succeq \epsilon \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (using step 1 in Algorithm 1), thus $\det\left(\begin{bmatrix} (H_t)_{ii} & (H_t)_{ii+1} \\ (H_t)_{i+1i} & (H_t)_{i+1i+1} \end{bmatrix}\right) \geq \det\left(\epsilon \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \epsilon^2$. summing up over t gives the result in the Lemma. The numerator last inequality can be upperbounded by bounding $(H_t)_{ii}$ individually as follows:

$$\begin{aligned}
(H_t)_{ii} &= \sum_{s=1}^{t} (g_s)_i^2/\lambda_s \\
&= \sum_{s=1}^{t} (g_s)_i^2/\lambda_s \\
&= \sum_{s=1}^{t} (g_s)_i^2/(G_\infty\sqrt{s}) \\
&\le \sum_{s=1}^{t} G_\infty^2/(G_\infty\sqrt{s}) \\
&\le \sum_{s=1}^{t} \frac{G_\infty}{\sqrt{s}} \\
&\le 2G_\infty\sqrt{t}
\end{aligned}$$

Substituting the above in (2) gives

$$\begin{aligned}
1/(1-\beta) &\le \frac{4G_\infty^2 t}{\hat{\epsilon}^2 G_\infty^2 T} \\
&\le \frac{4}{\hat{\epsilon}^2}
\end{aligned}$$

Substituting this in (1) and summing up over t gives

$$\begin{aligned}
\sum_t \frac{\eta}{2} g_t^T X_t g_t &\le \sum_t \frac{4nG_\infty\eta}{\hat{\epsilon}^3\sqrt{T}} \\
&\le \frac{4nG_\infty\eta}{\hat{\epsilon}^3}\sqrt{T}
\end{aligned}$$

□