## Appendix A. Stability analysis for linear fractional viscoelastic materials

Lemmas 4 and 5 provide stability estimates for the linear viscoelastic mechanics problem introduced in Eq. (47). We note that the fractional derivative approximation follows the form discussed in Section 2.4. Here $N$ denotes the number of Prony terms, and $\beta_k, \tau_k \in \mathbb{R}^+$ denote the scaling and time scale of the $k^{th}$ Maxwell element in the approximation (with $\beta_0$ being the scaling of a pure dashpot element). This approximation relies on intermediate variables $\boldsymbol{Q}_k^n$ which obey the update formula in Eq. (A.1).

$$\hat{\mathrm{D}}_n^{\alpha}(\boldsymbol{D}\boldsymbol{u}_h) = \beta_0 \boldsymbol{D}\boldsymbol{v}_h^n + \sum_{k=1}^{N} \boldsymbol{Q}_k^n, \qquad \boldsymbol{Q}_k^n = e_k^2 \boldsymbol{Q}_k^{n-1} + \beta_k e_k \Delta_t \boldsymbol{D}\boldsymbol{v}_h^n \tag{A.1}$$

With this approximation in mind, we can derive the following stability estimate for the discrete solution showing that the discrete solution is unconditionally stable and bounded by given data.

**Lemma 4.** *Consider the linear incompressible fractional viscoelasticity problem shown in Eq. (47) and the update formula shown in Eq. (A.1). Assuming that* $\boldsymbol{v}_h^0, \boldsymbol{u}_h^0 \in \boldsymbol{\mathcal{V}}_0^h$, *that* $\boldsymbol{Q}_k^0 = \boldsymbol{0}$, *and*

$$\boldsymbol{b} \in \boldsymbol{L}^{\infty}([0,T]; \boldsymbol{L}^2(\Omega_0)), \qquad \sup_{t\in[0,T]} \|\boldsymbol{b}(t)\|_0 \leq K_{\boldsymbol{b}}$$

$$\varrho \|\boldsymbol{v}_h^n\|_0^2 + E\|\boldsymbol{D}\boldsymbol{u}_h^n\|_0^2 + \eta\beta_0 \sum_{m=1}^{n} \Delta_t \|\boldsymbol{D}\boldsymbol{v}_h^m\|_0^2 + \sum_{k=1}^{N} \frac{\eta}{\beta_k e_k} \|\boldsymbol{Q}_k^n\|_0^2 \leq \varrho \|\boldsymbol{v}_h^0\|_0^2 + E\|\boldsymbol{D}\boldsymbol{u}_h^0\|_0^2 + \frac{C_{\Omega} t_n K_{\boldsymbol{b}}}{\eta\beta_0}$$

*for each* $n = 1, \ldots, N_T$.

**Proof of Lemma 4.** To prove Lemma 4, we choose $\boldsymbol{w}_h = \boldsymbol{0}$, $\boldsymbol{y}_h = \boldsymbol{v}_h^n$ and $q_h = 0$ in Eq. (47) resulting in the equation

$$\left(\varrho \delta_t \boldsymbol{v}_h^n, \boldsymbol{v}_h^n\right) + \left(\boldsymbol{\Sigma}^n, \nabla_{\boldsymbol{X}} \boldsymbol{v}_h^n\right) - \left(\boldsymbol{b}^n, \boldsymbol{v}_h^n\right) = 0 \tag{A.2}$$

Rearranging the update formula for $\boldsymbol{Q}_k^n$ in Eq. (A.1), we note that

$$\boldsymbol{D}\boldsymbol{v}_h^n = \frac{1}{\beta_k e_k \Delta_t}\left(\boldsymbol{Q}_k^n - e_k^2 \boldsymbol{Q}_k^{n-1}\right). \tag{A.3}$$

Focusing on the stress term in Eq. (A.2), noting the symmetry of $\boldsymbol{\Sigma}^n$, that $\boldsymbol{v}_h^n$ is weakly divergence free, the modified update formula of Eq. (A.3), and the identity

$$\begin{aligned}(\boldsymbol{a}-\boldsymbol{b})\cdot\boldsymbol{a} &= \frac{1}{2}\left(|\boldsymbol{a}|^2 - |\boldsymbol{b}|^2 + |\boldsymbol{a}-\boldsymbol{b}|^2\right), \quad \text{for any } \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d \\ (\boldsymbol{A}-\boldsymbol{B}):\boldsymbol{A} &= \frac{1}{2}\left(|\boldsymbol{A}|^2 - |\boldsymbol{B}|^2 + |\boldsymbol{A}-\boldsymbol{B}|^2\right), \quad \text{for any } \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{d\times d}\end{aligned} \tag{A.4}$$

we can observe that

$$\begin{aligned}\left(\boldsymbol{\Sigma}^n, \nabla_{\boldsymbol{X}} \boldsymbol{v}_h^n\right) &= \left(\boldsymbol{\Sigma}^n, \boldsymbol{D}\boldsymbol{v}_h^n\right) \\ &= \left(E\,\boldsymbol{D}\boldsymbol{u}_h^n + \eta \hat{\mathrm{D}}_n^{\alpha}(\boldsymbol{D}\boldsymbol{u}_h) + p_h^n \boldsymbol{I}, \boldsymbol{D}\boldsymbol{v}_h^n\right) \\ &= \frac{E}{\Delta_t}\left(\boldsymbol{D}\boldsymbol{u}_h^n, \boldsymbol{D}[\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1}]\right) + \eta\left(\beta_0 \boldsymbol{D}\boldsymbol{v}_h^n + \sum_{k=1}^{N} \boldsymbol{Q}_k^n, \boldsymbol{D}\boldsymbol{v}_h^n\right) \\ &= \frac{E}{\Delta_t}\left(\boldsymbol{D}\boldsymbol{u}_h^n, \boldsymbol{D}[\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1}]\right) + \eta\beta_0 \|\boldsymbol{D}\boldsymbol{v}_h^n\|_0^2 + \sum_{k=1}^{N} \frac{\eta}{\beta_k e_k \Delta_t}\left(\boldsymbol{Q}_k^n, \boldsymbol{Q}_k^n - e_k^2 \boldsymbol{Q}_k^{n-1}\right) \\ &= \frac{E}{2\Delta_t}\left(\|\boldsymbol{D}\boldsymbol{u}_h^n\|_0^2 - \|\boldsymbol{D}\boldsymbol{u}_h^{n-1}\|_0^2 + \|\boldsymbol{D}(\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1})\|_0^2\right) + \eta\beta_0 \|\boldsymbol{D}\boldsymbol{v}_h^n\|_0^2 \\ &\qquad + \sum_{k=1}^{N} \frac{\eta}{2\beta_k e_k \Delta_t}\left(\|\boldsymbol{Q}_k^n\|_0^2 - \|e_k^2 \boldsymbol{Q}_k^{n-1}\|_0^2 + \|\boldsymbol{Q}_k^n - e_k^2 \boldsymbol{Q}_k^{n-1}\|_0^2\right)\end{aligned} \tag{A.5}$$

Applying Eq. (A.4) and the equality in Eqs. (A.5) to (A.2),

$$\begin{aligned}&\frac{\varrho}{2}\left(\|\boldsymbol{v}_h^n\|_0^2 - \|\boldsymbol{v}_h^{n-1}\|_0^2 + \|\boldsymbol{v}_h^n - \boldsymbol{v}_h^{n-1}\|_0^2\right) + \frac{E}{2}\left(\|\boldsymbol{D}\boldsymbol{u}_h^n\|_0^2 - \|\boldsymbol{D}\boldsymbol{u}_h^{n-1}\|_0^2 + \|\boldsymbol{D}(\boldsymbol{u}_h^n - \boldsymbol{u}_h^{n-1})\|_0^2\right) \\ &\quad + \eta\beta_0 \Delta_t \|\boldsymbol{D}\boldsymbol{v}_h^n\|_0^2 + \sum_{k=1}^{N} \frac{\eta}{2\beta_k e_k}\left(\|\boldsymbol{Q}_k^n\|_0^2 - \|e_k^2 \boldsymbol{Q}_k^{n-1}\|_0^2 + \|\boldsymbol{Q}_k^n - e_k^2 \boldsymbol{Q}_k^{n-1}\|_0^2\right) - \Delta_t\left(\boldsymbol{b}^n, \boldsymbol{v}_h^n\right) = 0.\end{aligned} \tag{A.6}$$

Examining the final term in Eq. (A.6), noting through both Korn and Poincaré inequalities there exists a $C_\Omega > 0$ such that,

$$\|\boldsymbol{v}_h^n\|_0 \le C_\Omega \|\boldsymbol{D}\boldsymbol{v}_h^n\|_0$$

and applying Young's inequality (with $\epsilon = \eta\beta_0 / C_\Omega$), then

$$\begin{aligned}\left(\boldsymbol{b}^n, \boldsymbol{v}_h^n\right) &\le \frac{1}{2\epsilon}\|\boldsymbol{b}^n\|_0^2 + \frac{\epsilon}{2}\|\boldsymbol{v}_h^n\|_0^2 \\ &\le \frac{C_\Omega}{2\eta\beta_0}\|\boldsymbol{b}^n\|_0^2 + \frac{\eta\beta_0}{2}\|\boldsymbol{D}\boldsymbol{v}_h^n\|_0^2\end{aligned} \tag{A.7}$$

Combining Eqs. (A.6) and (A.7) and re-arranging terms, we observe that

$$\begin{aligned}&\varrho\|\boldsymbol{v}_h^n\|_0^2 + E\|\boldsymbol{D}\boldsymbol{u}_h^n\|_0^2 + \eta\beta_0\Delta_t\|\boldsymbol{D}\boldsymbol{v}_h^n\|_0^2 + \sum_{k=1}^{N}\frac{\eta}{\beta_k e_k}\|\boldsymbol{Q}_k^n\|_0^2 \\ &\qquad \le \varrho\|\boldsymbol{v}_h^{n-1}\|_0^2 + E\|\boldsymbol{D}\boldsymbol{u}_h^{n-1}\|_0^2 + \sum_{k=1}^{N}\frac{\eta}{\beta_k e_k}\|e_k^2 \boldsymbol{Q}_k^{n-1}\|_0^2 + \Delta_t \frac{C_\Omega}{\eta\beta_0}\|\boldsymbol{b}^n\|_0^2\end{aligned}$$

Noting that $e_k \leq 1$, applying induction, noting $\boldsymbol{Q}_k^0 = \boldsymbol{0}$ and utilizing the boundedness of $\boldsymbol{b}$, we arrive at the stability estimate.

To show stability for the pressure, $p_h^n$, we assume that the spaces $\boldsymbol{\mathcal{V}}_0^h$ and $\mathcal{P}^h$ are inf–sup stable, satisfying the condition ($\beta > 0$) [96],

$$\forall q_h \in \mathcal{P}^h \;\; \exists\, \boldsymbol{w}_h \in \boldsymbol{\mathcal{V}}_0^h, \;\; \boldsymbol{w}_h \neq \boldsymbol{0} : \quad (q_h, \nabla_{\boldsymbol{X}} \cdot \boldsymbol{w}_h) \geq \beta \|q_h\|_0 \|\boldsymbol{w}_h\|_1. \tag{A.8}$$

In addition, we assume there is a $\boldsymbol{v}_h^{-1} \in \boldsymbol{\mathcal{V}}_0^h$ satisfying

$$\|\Delta_t^{-1}(\boldsymbol{v}_h^0 - \boldsymbol{v}_h^{-1})\| \leq K \tag{A.9}$$

for some $K \geq 0$ (independent of $\Delta_t$) and that the initial conditions satisfy the discrete weakform (with $\boldsymbol{Q}_k^{-1} = \boldsymbol{0}$ for $k = 1, \ldots, N$),

$$\left(\varrho \delta_t \boldsymbol{v}_h^0, \boldsymbol{y}_h\right) + (\boldsymbol{\Sigma}^0, \nabla_{\boldsymbol{X}} \boldsymbol{y}_h) - (\boldsymbol{b}^0, \boldsymbol{y}_h) = 0 \tag{A.10}$$

for all weakly divergence free $\boldsymbol{w}_h \in \boldsymbol{\mathcal{V}}_{0,Div}^h$, with the space

$$\boldsymbol{\mathcal{V}}_{0,Div}^h = \left\{ \boldsymbol{w}_h \in \boldsymbol{\mathcal{V}}_0^h \;\middle|\; (q_h, \nabla_{\boldsymbol{X}} \cdot \boldsymbol{v}_h^n) = 0, \quad \forall q_h \in \mathcal{P}^h \right\}.$$

With these assumptions, we can prove unconditional stability for the discrete model pressure.

**Lemma 5.** *Suppose the assumptions of Lemma 4 hold. Assuming there is a $\boldsymbol{v}_h^{-1} \in \boldsymbol{\mathcal{V}}_0^h$ satisfying Eq. (A.9), that $\boldsymbol{Q}_k^{-1} = \boldsymbol{0}$, that*

$$\boldsymbol{b} \in \boldsymbol{W}^{1,\infty}([0,T]; \boldsymbol{L}^2(\Omega_0)), \qquad \sup_{t\in[0,T]} \|\boldsymbol{b}(t)\|_0 + \sup_{t\in[0,T]} \|\partial_t \boldsymbol{b}\|_0 \leq K_{\boldsymbol{b}}',$$

*and that the spaces $\boldsymbol{\mathcal{V}}_0^h$ and $\mathcal{P}^h$ are inf–sup stable, then $p_h^n$ satisfies the stability estimate,*

$$\|p_h^n\|_0 \leq \frac{C}{\beta}\left(K + K_{\boldsymbol{b}} + t_n^{1/2} K_{\boldsymbol{b}}' + \|\boldsymbol{v}_h^0\|_0 + \|\boldsymbol{D}\boldsymbol{u}_h^0\|_0 + \|\boldsymbol{D}\boldsymbol{v}_h^0\|_0\right)$$

*for a constant $C > 0$ (independent of $h$ and $\delta_t$) and for each $n = 1, \ldots, N_T$.*

**Proof of Lemma 5.** Looking at Eq. (47), we first separate components of $\boldsymbol{\Sigma}^n$, choose $\boldsymbol{w}_h = \boldsymbol{0}$ and $q_h = 0$ and rearrange the equation. Applying Cauchy–Schwarz inequality, we can arrive at the following inequality.

$$\begin{aligned}
(p_h^n, \nabla_{\boldsymbol{X}} \cdot \boldsymbol{y}_h) &= (\boldsymbol{b}^n, \boldsymbol{w}_h) - \varrho\left(\delta_t \boldsymbol{v}_h^n, \boldsymbol{y}_h\right) - (E\boldsymbol{D}\boldsymbol{u}_h^n + \eta \hat{\mathrm{D}}_n^{\alpha}(\boldsymbol{D}\boldsymbol{u}_h), \nabla_{\boldsymbol{X}} \boldsymbol{y}_h) \\
&\leq \|\boldsymbol{b}^n\|_0 \|\boldsymbol{y}_h\|_0 + \varrho \|\delta_t \boldsymbol{v}_h^n\|_0 \|\boldsymbol{y}_h\|_0 + \|E\boldsymbol{D}\boldsymbol{u}_h^n + \eta \hat{\mathrm{D}}_n^{\alpha}(\boldsymbol{D}\boldsymbol{u}_h)\|_0 \|\nabla_{\boldsymbol{X}} \boldsymbol{y}_h\|_0 \\
&\leq \left(\|\boldsymbol{b}^n\|_0 + \varrho\|\delta_t \boldsymbol{v}_h^n\|_0 + E\|\boldsymbol{D}\boldsymbol{u}_h^n\|_0 + \eta\beta_0 \|\boldsymbol{D}\boldsymbol{v}_h^n\|_0 + \sum_{k=1}^{N} \eta \|\boldsymbol{Q}_k^n\|_0\right) \|\boldsymbol{y}_h\|_1
\end{aligned} \tag{A.11}$$

Noting that the spaces of discrete solutions satisfy the inf–sup condition, the inequality in Eq. (A.11) can be simplified to provide an upper bound on the pressure, $p_h^n$.

$$\|p_h^n\|_0 \leq \frac{1}{\beta}\left(\|\boldsymbol{b}^n\|_0 + \varrho\|\delta_t \boldsymbol{v}_h^n\|_0 + E\|\boldsymbol{D}\boldsymbol{u}_h^n\|_0 + \eta\beta_0 \|\boldsymbol{D}\boldsymbol{v}_h^n\|_0 + \sum_{k=1}^{N} \eta \|\boldsymbol{Q}_k^n\|_0\right) \tag{A.12}$$

On the RHS of Eq. (A.12), the terms involving $\boldsymbol{b}$ remain bounded due to the implicit assumption that $\boldsymbol{b} \in \boldsymbol{W}^{1,\infty}([0,T]; \boldsymbol{L}^2(\Omega_0))$. Moreover, from Lemma 4, we know there exists a $C_0 > 0$ (independent of $h$ and $\Delta_t$) such that,

$$E\|\boldsymbol{D}\boldsymbol{u}_h^n\|_0 + \sum_{k=1}^{N} \eta \|\boldsymbol{Q}_k^n\|_0 \leq C_0\left(\|\boldsymbol{v}_h^0\|_0 + \|\boldsymbol{D}\boldsymbol{u}_h^0\|_0 + t_n^{1/2} K_{\boldsymbol{b}}\right). \tag{A.13}$$

The stability estimate in Eq. (A.12) relies on the boundedness of the discrete time derivative of $\boldsymbol{v}_h^n$ as well as the symmetric derivative of $\boldsymbol{v}_h^n$. To ensure these quantities remain bounded independent of $\Delta_t$, we return to Eq. (47). Subtracting the equation for $t_{n-1}$ from that for $t_n$, letting $\delta(\cdot)^n = (\cdot)^n - (\cdot)^{n-1}$, and selecting the test functions $\boldsymbol{w}_h = \boldsymbol{0}$, $\boldsymbol{y}_h = \delta \boldsymbol{v}_h^n$ and $q_h = 0$ yields

$$\left(\varrho\delta(\delta_t \boldsymbol{v}_h^n), \delta \boldsymbol{v}_h^n\right) + (\delta \boldsymbol{\Sigma}^n, \nabla_{\boldsymbol{X}} \delta \boldsymbol{v}_h^n) - (\delta \boldsymbol{b}^n, \delta \boldsymbol{v}_h^n) = 0 \tag{A.14}$$

Applying the same approach as in Lemma 4 we can arrive at the following bound.

$$\begin{aligned}
&\varrho\|\delta \boldsymbol{v}_h^n\|_0^2 + E\|\boldsymbol{D}(\delta \boldsymbol{u}_h^n)\|_0^2 + \eta\beta_0\Delta_t\|\boldsymbol{D}(\delta \boldsymbol{v}_h^n)\|_0^2 + \sum_{k=1}^{N}\frac{\eta}{\beta_k e_k}\|\delta \boldsymbol{Q}_k^n\|_0^2 \\
&\quad\le \varrho\|\delta \boldsymbol{v}_h^{n-1}\|_0^2 + E\|\boldsymbol{D}(\delta \boldsymbol{u}_h^{n-1})\|_0^2 + \sum_{k=1}^{N}\frac{\eta}{\beta_k e_k}\|e_k^2\delta \boldsymbol{Q}_k^{n-1}\|_0^2 + \Delta_t\frac{C_\Omega}{\eta\beta_0}\|\delta \boldsymbol{b}^n\|_0^2 \\
&\quad\le \varrho\|\delta \boldsymbol{v}_h^0\|_0^2 + E\|\boldsymbol{D}(\delta \boldsymbol{u}_h^0)\|_0^2 + \sum_{m=1}^{n}\Delta_t\frac{C_\Omega}{\eta\beta_0}\|\delta \boldsymbol{b}^m\|_0^2
\end{aligned} \tag{A.15}$$

Dividing by $\Delta_t^2$ and noting $\delta_t \boldsymbol{u}_h^n = \boldsymbol{v}_h^n$ for any $n$, we observe that the remaining terms in Eq. (A.12) are bounded by given data, *e.g.*

$$\varrho\|\delta_t \boldsymbol{v}_h^n\|_0^2 + E\|\boldsymbol{D}\boldsymbol{v}_h^n\|_0^2 \le \varrho K^2 + E\|\boldsymbol{D}\boldsymbol{v}_h^0\|_0^2 + \sum_{m=1}^{n}\Delta_t\frac{C_\Omega}{\eta\beta_0}\|\delta_t \boldsymbol{b}^m\|_0^2. \tag{A.16}$$

Combining these results, along with bounds on $\boldsymbol{b}$ and its derivative, we arrive at the stability estimate.

## Appendix B. Optimal prony series parameters for approximating the Caputo fractional derivative

In Section 2.5, we presented the method for optimizing the parameters of the Prony series for a defined time scale proportional to the duration of the simulation. In practice, we observed in our examples (Sections 5 and 6) that better convergence rates can be achieved by scaling the time interval for which the Prony series parameters were mapped to. In every case, the convergence rates were significantly improved by lengthening the time scale by

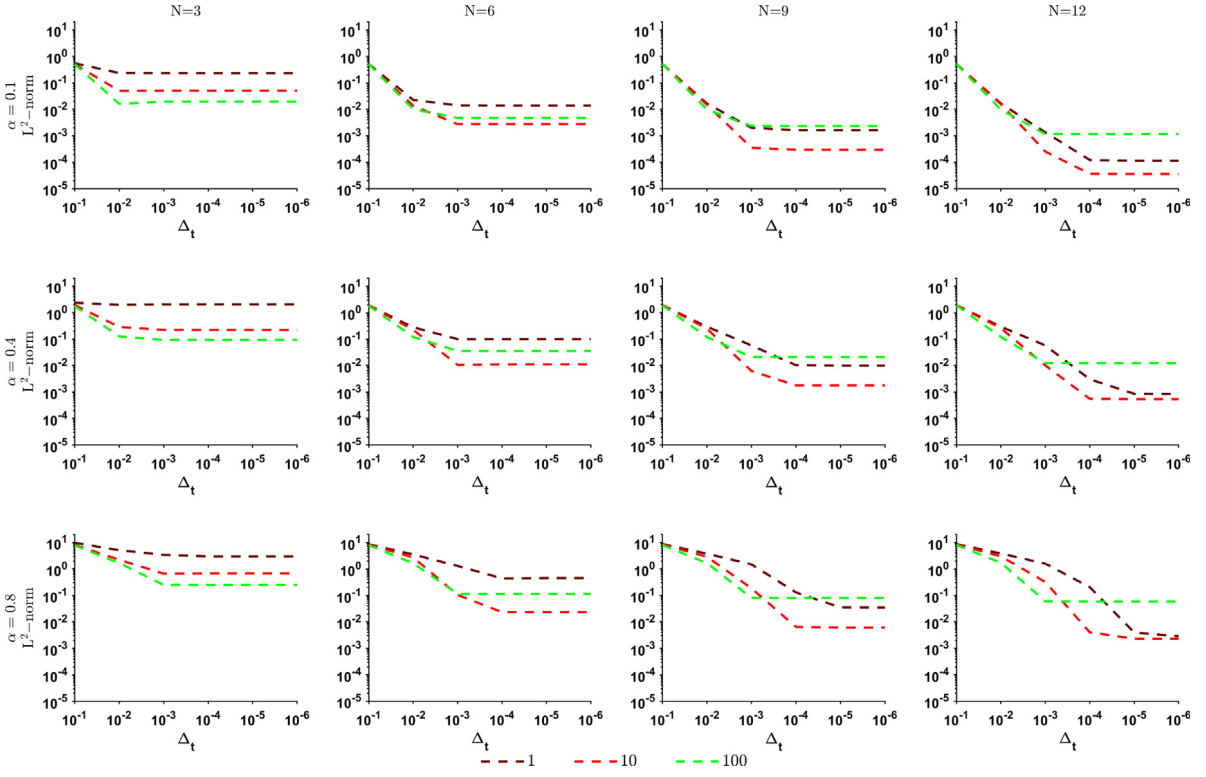


**Fig. B.11.** Comparing convergence of the Prony-based approximation Eq. (12) in the polynomial example with refinement in time, $\alpha = 0.1$, 0.4, 0.8, and 3, 6, 9, and 12 Prony terms when scaling the Prony series parameters to time intervals that are 1 times, 10 times and 100 times the size of the actual time interval.

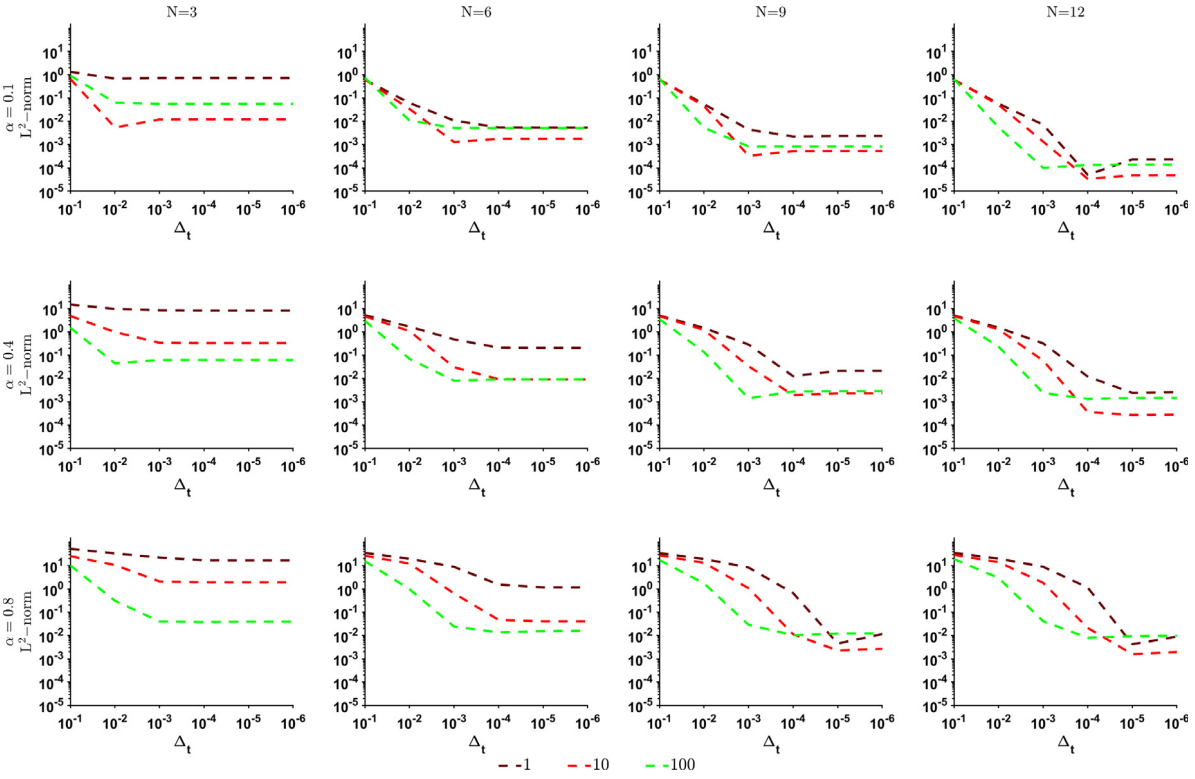


**Fig. B.12.** Comparing convergence of the Prony-based approximation Eq. (12) in Appendix B exponential function example (Example 2) with refinement in time, $\alpha = 0.1$, 0.4, 0.8, and 3, 6, 9, and 12 Prony terms when scaling the Prony series parameters to time intervals that are 1 times, 10 times and 100 times the size of the actual time interval.

a factor of 10 or even 100. This was first observed as a consequence of the Geo et al. example (Section 6), where the convergence response can be worse initially as the number of Prony terms increases. This is a result of the increase in value of $C(\boldsymbol{\beta}, \boldsymbol{\tau})$, which results in slow convergence until the approximation error hits the bounds of the error estimate in Theorem 3. Numerically, lengthening the time interval for mapping the Prony series parameters increases the time constants $\tau_k$ and decreases the weights $\beta_k$. We can observe from Theorem 3 that the error bound, specifically $\beta_0$ and $C(\boldsymbol{\beta}, \boldsymbol{\tau})$, reduces as a result, decreasing the overall error of the approximation.

As this effect appears to be problem dependent (more significant in Gao et al. Example 1 [67]), we tested this in more detail using two polynomial examples: (1) the decaying oscillating function like polynomial presented in Fig. 4, which behaves with short variable base frequencies and (2) a polynomial that is monotonically increasing with highly exponential behavior, which behaves with single long base frequency (i.e. fitted to $e^{6*t} - 1$). The convergence response was tested for $\alpha \in \{0.1, 0.4, 0.8\}$ and $N \in \{3, 6, 9, 12\}$ with time scaling of 1, 10, and 100.

The overall response is quite similar. Most important is that lengthening the time scale by a factor of 10 always results in better convergence and overall error (Fig. B.11 and Fig. B.12), sometimes by two order of magnitude. Results from scaling by a factor of 100 is more variable. Most noticeably, the lower-bound with a scaling of 100 can be much worse at larger $N$. However, the convergence at larger $\Delta_t$ can be significantly better. Most noticeably, the $L^2$-norm in the example 2 (Fig. B.12) can be 3 magnitude better at $\Delta_t = 10^{-3}$ for $\alpha = 0.8$ and N = 12 with a scaling of 100. However, also note that with a scaling of 10, the $L^2$-norm eventually surpasses the scaling of 100 after it has plateaued. Here, the gain in convergence at large $\Delta_t$ surpasses the 1 magnitude gain in the lowerbound as a result of the truncation error (Lemma 1), which requires 1.5 more magnitude in $\Delta_t$ refinement. In contrast, for example 1 (Fig. B.11), we observe that the gain in convergence initially is much less significant, and the truncation error is around 2 magnitudes larger.

Clearly, these results are also problem dependent. Part of this dependency can be observed from Theorem 3, by the weights of $\left[|f'(0)| + \|f''\|_{0,1}\right]$ and $\|f\|_{W^{3,\infty}(0,T)}$ on $\varepsilon$ and $C(\boldsymbol{\beta}, \boldsymbol{\tau})$ respectively. Example 1 (Fig. B.11) starts at