Supplementary Material for
Accelerating ADMM for Efficient Simulation and Optimization

1 BACKGROUND

In this supplementary material, we will verify the linear convergence theorems with a target function g that is locally Lipschitz differentiable (Theorems 3.3 and 3.4). We will use ADMM to solve the following optimization problem from [1] for physical simulation:

$$\min_{\mathbf{x},\mathbf{z}} \; f(\mathbf{x}) + g(\mathbf{z}) \quad \text{s.t. } \mathbf{W}(\mathbf{z} - \mathbf{D}\mathbf{x}) = 0, \tag{1}$$

Here $\mathbf{x}$ is the node positions of the discretized object. $f(\mathbf{x})$ is a momentum energy of the form

$$f(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \tilde{\mathbf{x}})^T \mathbf{G}(\mathbf{x} - \tilde{\mathbf{x}}),$$

with $\tilde{\mathbf{x}}$ being a constant vector, and $\mathbf{G}$ being a scaled mass matrix. $\mathbf{D}\mathbf{x}$ collects the deformation gradient of each element. $\mathbf{W}$ is a diagonal scaling matrix that improves conditioning. $g(\mathbf{z})$ is an elastic potential energy. Compared to the following form of optimization problems discussed in our paper:

$$\min_{\mathbf{x},\mathbf{z}} \; f(\mathbf{x}) + g(\mathbf{z}) \quad \text{s.t. } \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{z} = \mathbf{c}, \tag{2}$$

we can see that problems 1 and 2 are equivalent if $\mathbf{A} = \mathbf{W}\mathbf{D}, \mathbf{B} = \mathbf{W}, \mathbf{c} = \mathbf{0}$. In this report, we assume the simulation object to be a tetrahedral mesh and use the following potential energy:

$$g(\mathbf{z}) = \sum_{i=1}^{n_t} v_i \psi(\mathbf{F}^i),$$

where v_i is the volume of each tetrahedron, and $\mathbf{F}^i \in \mathbb{R}^{3\times3}$ is its deformation gradient with respect ot the rest shape, and ψ is the strain energy density function of StVK materials [2]:

$$\psi(\mathbf{F}) = \lambda_1 \mathbf{E} : \mathbf{E} + \frac{1}{2}\lambda_2 tr^2(\mathbf{E}), \tag{3}$$

where λ_1, λ_2 are given material parameters, and

$$\mathbf{E} = \frac{1}{2}(\mathbf{F}^T\mathbf{F} - \mathbf{I}) \in \mathbb{R}^{3\times3}$$

and $\mathbf{I}$ is the identity matrix.

The linear convergence theorems we will verify require a local Lipschitz constant for the gradient of g. Thus we need to analyze the Lipschitz differentiability of ψ. The gradient of ψ is the first Piola-Kirchhoff stress tensor:

$$\mathbf{P}(\mathbf{F}) = \mathbf{F}\mathbf{S}(\mathbf{F}). \tag{4}$$

where $\mathbf{S}$ is the second Piola-Kirchhoff stress tensor:

$$\mathbf{S}(\mathbf{F}) = 2\lambda_1\mathbf{E} + \lambda_2 tr(\mathbf{E})\mathbf{I},$$

Note that the value of $\mathbf{E}$ depends on $\mathbf{F}$. In the following, we will denote $\mathbf{E}_j = \mathbf{E}(\mathbf{F}_j)$ for a subscript j. Throughout this report, for a matrix $\mathbf{A}$, $\|\mathbf{A}\|$ denotes its Frobenius norm and $\|\mathbf{A}\|_2$ denotes its l_2 norm. Our task is to estimate the Lipschitz constant of $\mathbf{P}(\mathbf{F})$ with respect to $\mathbf{F}$.

Proposition 1. $\|\mathbf{P}(\mathbf{F}_1) - \mathbf{P}(\mathbf{F}_2)\| \leq ((\lambda_1 + \frac{\sqrt{3}}{2}\lambda_2)\|\mathbf{F}_1\|(\|\mathbf{F}_1\| + \|\mathbf{F}_2\|) + (2\lambda_1 + 3\lambda_2)\|\mathbf{E}_2\|)\|\mathbf{F}_1 - \mathbf{F}_2\|$.

Proof. We have:

$$
\begin{aligned}
\mathbf{E}_1 - \mathbf{E}_2 &= \frac{\mathbf{F}_1^T\mathbf{F}_1 - \mathbf{F}_2^T\mathbf{F}_2}{2} \\
&= \frac{1}{2}(\mathbf{F}_1^T\mathbf{F}_1 - \mathbf{F}_1^T\mathbf{F}_2 + \mathbf{F}_1^T\mathbf{F}_2 - \mathbf{F}_2^T\mathbf{F}_2) \\
&= \frac{1}{2}(\mathbf{F}_1^T(\mathbf{F}_1 - \mathbf{F}_2) + (\mathbf{F}_1^T - \mathbf{F}_2^T)\mathbf{F}_2) \qquad (5)
\end{aligned}
$$

Hence:

$$
\|\mathbf{E}_1 - \mathbf{E}_2\| \le \frac{1}{2}(\|\mathbf{F}_1\| + \|\mathbf{F}_2\|)\|\mathbf{F}_1 - \mathbf{F}_2\| \qquad (6)
$$

And:

$$
\begin{aligned}
\|tr(\mathbf{E}_1)\mathbf{I} - tr(\mathbf{E}_2)\mathbf{I}\| &= \sqrt{3}|tr(\mathbf{E}_1) - tr(\mathbf{E}_2)| \\
&= \sqrt{3}|tr(\mathbf{E}_1 - \mathbf{E}_2)| \\
&= \frac{\sqrt{3}}{2}|tr(\mathbf{F}_1^T(\mathbf{F}_1 - \mathbf{F}_2) + (\mathbf{F}_1^T - \mathbf{F}_2^T)\mathbf{F}_2)| \\
&\le \frac{\sqrt{3}}{2}(\|\mathbf{F}_1\| + \|\mathbf{F}_2\|)\|\mathbf{F}_1 - \mathbf{F}_2\| \qquad (7)
\end{aligned}
$$

For $\mathbf{P}(\mathbf{F})$ we have:

$$
\begin{aligned}
\mathbf{P}(\mathbf{F}_1) - \mathbf{P}(\mathbf{F}_2) &= \mathbf{F}_1\mathbf{S}(\mathbf{F}_1) - \mathbf{F}_2\mathbf{S}(\mathbf{F}_2) \\
&= \mathbf{F}_1\mathbf{S}(\mathbf{F}_1) - \mathbf{F}_1\mathbf{S}(\mathbf{F}_2) + \mathbf{F}_1\mathbf{S}(\mathbf{F}_2) - \mathbf{F}_2\mathbf{S}(\mathbf{F}_2) \\
&= \mathbf{F}_1(\mathbf{S}(\mathbf{F}_1) - \mathbf{S}(\mathbf{F}_2)) + (\mathbf{F}_1 - \mathbf{F}_2)\mathbf{S}(\mathbf{F}_2) \qquad (8)
\end{aligned}
$$

Therefore:

$$
\|\mathbf{P}(\mathbf{F}_1) - \mathbf{P}(\mathbf{F}_2)\| \le \|\mathbf{F}_1\|\|\mathbf{S}(\mathbf{F}_1) - \mathbf{S}(\mathbf{F}_2)\| + \|\mathbf{F}_1 - \mathbf{F}_2\|\|\mathbf{S}(\mathbf{F}_2)\| \qquad (9)
$$

By (6) and (7) we have:

$$
\|\mathbf{S}(\mathbf{F}_1) - \mathbf{S}(\mathbf{F}_2)\| \le (\lambda_1 + \frac{\sqrt{3}}{2}\lambda_2)(\|\mathbf{F}_1\| + \|\mathbf{F}_2\|)\|\mathbf{F}_1 - \mathbf{F}_2\| \qquad (10)
$$

We next estimate $\|\mathbf{S}(\mathbf{F}_2)\|$:

$$
\begin{aligned}
\|\mathbf{S}(\mathbf{F}_2)\| &\le 2\lambda_1\|\mathbf{E}_2\| + \lambda_2\sqrt{3}|tr(\mathbf{E}_2)| \\
&\le (2\lambda_1 + 3\lambda_2)\|\mathbf{E}_2\| \qquad (11)
\end{aligned}
$$

The result comes from (6), (7), (9) and (11). □

Proposition 2. *Assume* $\|\mathbf{F}\|^2 \ge 27$*, then we have:*

$$
\psi(\mathbf{F}) \ge (\frac{16}{729}\lambda_1 + \frac{72}{729}\lambda_2)\|\mathbf{F}\|^4 \qquad (12)
$$

Proof. Assume the singular values of $\mathbf{F}$ are $\sigma_1 \geq \sigma_2 \geq \sigma_3$. Then we have:

$$\|\mathbf{F}\|^2 = \sum_{i=1}^{3} \sigma_i^2 \tag{13}$$

$$\|\mathbf{E}\|^2 = \frac{1}{4}\sum_{i=1}^{3}(\sigma_i^2 - 1)^2 \tag{14}$$

$$tr^2(\mathbf{E}) = \frac{1}{4}(\sum_{i=1}^{3} \sigma_i^2 - 3)^2 \tag{15}$$

Since $\|\mathbf{F}\|^2 \geq 27$ we have $\sigma_1 \geq 3$, and we have:

$$(\sigma_1^2 - 1)^2 = \max_{i=1,2,3}(\sigma_i^2 - 1)^2 \tag{16}$$

Hence:

$$\|\mathbf{E}\|^2 \geq \frac{1}{4}(\sigma_1^2 - 1)^2 \geq \frac{1}{4}(\frac{8}{9}\sigma_1^2)^2 \geq \frac{1}{4}(\frac{8}{27}\|\mathbf{F}\|^2)^2 \tag{17}$$

$$tr^2(\mathbf{E}) \geq \frac{1}{4}(\frac{24}{27}\|\mathbf{F}\|^2)^2 \tag{18}$$

The result comes from (16) and (17). □

Proposition 3. $\text{conv}(\mathscr{L}_\psi^a) \subset \{\mathbf{F} \mid \|\mathbf{F}\| \leq b\}$, *where* $b = \max\{3\sqrt{3}, \sqrt[4]{\frac{a}{\frac{16}{729}\lambda_1 + \frac{72}{729}\lambda_2}}\}$.

Proof. Since $\{\mathbf{F} \mid \|\mathbf{F}\| \leq b\}$ is convex, it suffice to show $\mathscr{L}_\psi^a \subset \{\mathbf{F} \mid \|\mathbf{F}\| \leq b\}$. Now assume $\phi(\mathbf{F}) \leq a$. If $\|\mathbf{F}\| \leq 3\sqrt{3}$ this is trivial, otherwise by Proposition 2 we have:

$$(\frac{16}{729}\lambda_1 + \frac{72}{729}\lambda_2)\|\mathbf{F}\|^4 \leq a \tag{19}$$

which completes the proof. □

Proposition 4. *(1): The value* $((2\lambda_1 + \sqrt{3}\lambda_2)b^2 + (\lambda_1 + \frac{3}{2}\lambda_2)\sqrt{b^4 + 3})$ *is a Lipschitz constant of* $\mathbf{P}(\mathbf{F})$ *over the set* $\text{conv}(\mathscr{L}_\psi^a)$, *where b is defined in Proposition 3.*

(2): $\sup_{\mathbf{F}\in\mathscr{L}_\psi^a} \|\mathbf{P}(\mathbf{F})\|^2 \leq (2\lambda_1 + 3\lambda_2)^2 b^2(\frac{1}{4}b^4 + \frac{3}{4})$

Proof. Assume the singular values of F are $\sigma_1 \geq \sigma_2 \geq \sigma_3$, by (14) we have:

$$\|\mathbf{E}\|^2 \leq \frac{1}{4}\sum_{i=1}^{3}\sigma_i^4 + \frac{3}{4} \leq \frac{1}{4}(\sum_{i=1}^{3}\sigma_i^2)^2 + \frac{3}{4} = \frac{1}{4}\|\mathbf{F}\|^4 + \frac{3}{4} \tag{20}$$

Now assume $\mathbf{F}_1, \mathbf{F}_2 \in \text{conv}(\mathscr{L}_\psi^a)$, by Proposition 3 we have:

$$\|\mathbf{F}_1\| \leq b, \|\mathbf{F}_2\| \leq b \tag{21}$$

By Proposition 1 and (21) we have:

$$\|\mathbf{P}(\mathbf{F}_1) - \mathbf{P}(\mathbf{F}_2)\| \leq ((2\lambda_1 + \sqrt{3}\lambda_2)b^2 + (\lambda_1 + \frac{3}{2}\lambda_2)\sqrt{b^4 + 3})\|\mathbf{F}_1 - \mathbf{F}_2\| \tag{22}$$

which proves (1).

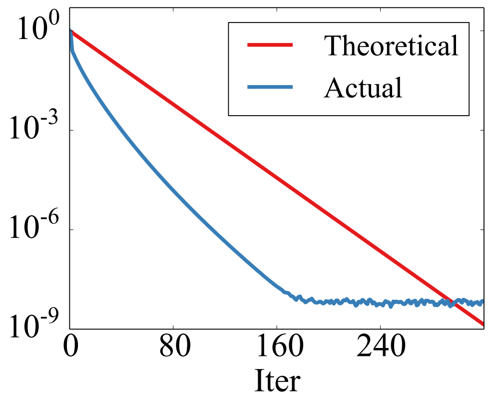

Fig. 1. Comparison between theoretical and actual shrinkage of $\|\mathbf{B}\mathbf{z}^{k+1} - \mathbf{B}\mathbf{z}^k\|$ for verification of Theorem 3.3. The values are normalized by $\|\mathbf{B}\mathbf{z}^1 - \mathbf{B}\mathbf{z}^0\|$ and plotted in logarithmic scale. The actual shrinkage ratio stays below the theoretical upper bound, before it oscillates due to numerical error when close enough to the solution.

For (2), by (11), Proposition 3 and (20):

$$\|\mathbf{P}(\mathbf{F})\|^2 \leq b^2\|\mathbf{S}(\mathbf{F})\|^2 \leq b^2(2\lambda_1 + 3\lambda_3)^2\|\mathbf{E}\|^2 \leq (2\lambda_1 + 3\lambda_2)^2 b^2(\frac{1}{4}b^4 + \frac{3}{4}) \tag{23}$$

□

2 VERIFICATION FOR THE X-Z-U ITERATION

In this subsection, we construct an example to verify the linear convergence theorem for the **x**-**z**-**u** iteration (Theorem 3.3). Assume the function $J : \mathbb{R}^9 \to \mathbb{R}^{3\times3}$ assembles a vector into its matrix form. Assume $\mathbf{z} \in \mathbb{R}^{9n_t}$, $\mathbf{z}_i \in \mathbb{R}^9$ is its component, where $i \in [1, n_t]$. Then $g(\mathbf{z})$ becomes:

$$g(\mathbf{z}) = \sum_{i=1}^{n_t} v_i \psi(J(\mathbf{z}_i)). \tag{24}$$

The next proposition is just a corollary from the proofs of previous propositions so we omit its proof.

Proposition 5. *(1): The value* $\max_i v_i((2\lambda_1 + \sqrt{3}\lambda_2)b_i^2 + (\lambda_1 + \frac{3}{2}\lambda_2)\sqrt{b_i^4 + 3})$ *is a Lipschitz constant of* $\nabla g(z)$ *over the set* $\mathrm{conv}(\mathscr{L}_g^a)$*, where* $b_i = \max\{3\sqrt{3}, \sqrt[4]{\frac{a/v_i}{\frac{16}{729}\lambda_1 + \frac{72}{729}\lambda_2}}\}$.

(2): $\sup_{\mathbf{z}\in\mathscr{L}_g^a} \|\nabla g(\mathbf{z})\|^2 \leq \sum_{i=1}^m v_i^2(2\lambda_1 + 3\lambda_2)^2 b_i^2(\frac{1}{4}b_i^4 + \frac{3}{4})$.

For the sake of simplicity, we suppose $\mathbf{B} = \mathbf{I}$ in Eq. 2. The optimization problem for the verification example is constructed via the following procedure:

1. Choose the initial value as suggested by Assumption 3.5.
2. Let $a = T^0 + 1$. Compute L_c and $\sup_{\mathbf{z}\in\mathscr{L}_g^a} \|\nabla g(\mathbf{z})\|^2$ using Proposition 5.
3. Choose a matrix $\mathbf{G}$ that is large enough such that $\rho(\mathbf{K}) \leq \frac{1}{2L_c}$.

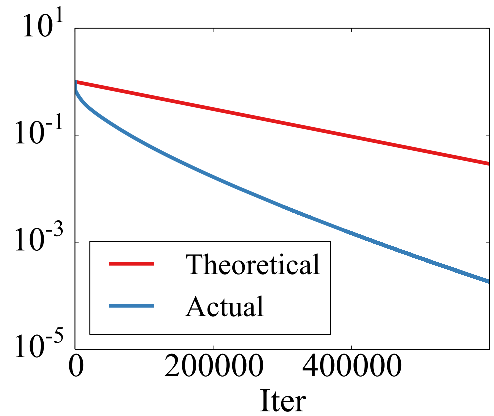

Fig. 2. Comparison between theoretical and actual shrinkage of $\|\mathbf{v}^{k+1} - \mathbf{v}^k\|$ for verification of Theorem 3.4. The values are normalized by $\|\mathbf{v}^1 - \mathbf{v}^0\|$ and plotted in logarithmic scale. The actual shrinkage ratio stays below the theoretical upper bound.

4. Choose a μ that is large enough such that $c_1 \leq 1$ as defined in Assumption 3.5, $\frac{\mu}{2} - \frac{L_c^2}{\mu} \geq \frac{L_c}{2}$ and $\mu > \max\{\frac{1}{\frac{1}{2L_c} - \rho(\mathbf{K})}, \frac{1}{L_c}\}$.

Fig. 1 plots in logarithmic scale the value of $\|\mathbf{Bz}^{k+1} - \mathbf{Bz}^k\|$ throughout the iterations, as well as a straight line where the value changes according to the constant upper bound of shrinkage ratio given in Theorem 3.3. We can see that the actual shrinkage ratio stays below the theoretical upper bound before convergence.

3 VERIFICATION FOR THE Z-X-U ITERATION

We now construct an example to verify the linear convergence theorem for the **z**-**x**-**u** iteration (Theorem 3.4). The first problem is to compute η. We first compute the SVD for $\mathbf{A}$. Suppose $\mathbf{A} = \mathbf{UDV}$ and let r be the rank of $\mathbf{A}$. Let $\mathbf{U}_r$ be a sub-matrix of $\mathbf{U}$ consisting of its first r columns. Then we know $R(\mathbf{A}) = R(\mathbf{U}_r)$. Moreover, $\forall \mathbf{y} \in R(\mathbf{A})$, suppose $\mathbf{y} = \mathbf{U}_r\mathbf{z}$, then we have: $\|\mathbf{y}\| = \|\mathbf{z}\|$. Thus we choose η to be the minimal eigenvalue of $\mathbf{U}_r^T\mathbf{K}\mathbf{U}_r$.

Proposition 6. $\sup_{x \in \mathscr{L}_f^a} \|\mathbf{Ax} - \mathbf{A\tilde{x}}\|^2 \leq 2\|\mathbf{A}\|_2^2 a/q$, *where q is the minimal eigenvalue of* $\mathbf{G}$.

Proof. By the definition of $f(\mathbf{x})$ we have:

$$\mathscr{L}_f^a = \{\mathbf{x} : \frac{1}{2}\|\mathbf{x} - \tilde{\mathbf{x}}\|_{\mathbf{G}}^2 \leq a\} \subset \{\|\mathbf{x} - \tilde{\mathbf{x}}\|^2 \leq 2a/q\} \tag{25}$$

Moreover, we have:

$$\|\mathbf{Ax} - \mathbf{A\tilde{x}}\|^2 \leq \|\mathbf{A}\|_2^2\|\mathbf{x} - \tilde{\mathbf{x}}\|^2 \tag{26}$$

which completes the proof. □

To determine c_2 defined in Assumption 3.6 we run one dimensional line-search for t and compute the upper bound for $\sup_{\mathbf{x} \in \mathscr{L}_f^t} \frac{2}{\eta^2\mu}\|\mathbf{Ax} - \mathbf{A\tilde{x}}\|^2 + \sup_{\mathbf{z} \in \mathscr{L}_g^{a-t}} (\frac{2\rho(\mathbf{K})^2}{\mu\eta^2} + \frac{1}{\mu})\|\mathbf{B}^{-T}\nabla g(\mathbf{z})\|^2$ by using Proposition 6 and Proposition 5.

The optimization problem used of verification is constructed via the following procedure:

1. Choose initial value as suggested by Assumption 3.6. Then compute η.

2. Let $a = T^0 + 1$. Compute L_d and $\sup_{\mathbf{z} \in \mathcal{L}_g^a} \|\nabla g(\mathbf{z})\|^2$ using Proposition 5.
3. Choose a matrix $\mathbf{G}$ that is large enough such that $\rho(\mathbf{K}) \leq \frac{1}{L_d}$.
4. Choose a μ that is large enough such that $c_2 + c_3 \leq 1$ defined in Assumption 3.6, $\frac{\mu}{2} \geq \frac{4}{\eta^2 \mu}$, $\frac{\mu - L_d}{2} \geq (\frac{4\rho(\mathbf{K})^2 L_d^2}{\mu \eta^2} + \frac{2L_d^2}{\mu})$ and $\mu > \max\{\frac{1}{\frac{2}{L_d} - \rho(\mathbf{K})}, \frac{1}{L_d}\}$.

Fig. 2 plots in logarithmic scale the value $\|\mathbf{v}^{k+1} - \mathbf{v}^k\|$ throughout the iterations, as well as a straight line where the value shrinks according to the constant theoretical upper bound given in Theorem 3.4. We can see that the actual shrinkage ratio stays below the upper bound.

REFERENCES

[1] M. Overby, G. E. Brown, J. Li, and R. Narain. 2017. ADMM ⊇ projective dynamics: Fast simulation of hyperelastic models with dynamic constraints. *IEEE Transactions on Visualization and Computer Graphics* 23, 10 (2017), 2222–2234.

[2] Eftychios Sifakis and Jernej Barbič. 2012. FEM simulation of 3D deformable solids: A practitioner's guide to theory, discretization and model reduction. In *ACM SIGGRAPH 2012 Courses*. 20:1–20:50.