Math 673/4

# Multigrid Methods: A Mostly Matrix-Based Approach

## Chapter 11: The Axiomatic Method of Subspace Corrections

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu swise1@.utk.edu
University of Tennessee

F24/S25

# Chapter 11, Part 2 of 2

# The Axiomatic Method of Subspace Corrections

# The Fast Subspace Descent (FASD) Method

## SSO Corrections

Recall that, at every step of the SSO method, we need to solve, for $i = 1, \ldots, N$, the minimization problem

$$e_i^k = \operatorname*{argmin}_{w \in \mathcal{V}_i} J_i^k(w) \tag{1}$$

exactly. This requires, at least, the evaluation of the energy $E$ and its derivative $\mathrm{D}E$ in the space $\mathcal{V}_i$. Although the size of problem (1) is reduced to $\dim \mathcal{V}_i$, such evaluations are still in the original space of size $\dim \mathcal{V}$, which may be expensive.

## Injection and Restriction

To better explain the issue, we first need some notation. Let $I_i : \mathcal{V}_i \hookrightarrow \mathcal{V}$ be the natural inclusion and $R_i := I_i^\top : \mathcal{V}' \to \mathcal{V}_i'$ the natural restriction of functionals, where the superscript $\top$ indicates the adjoint in $\mathcal{V}$. Thus, for $k \geq 0$ and $i = 1, \ldots, N$, to solve (1), we must be able to evaluate, for $v, w \in \mathcal{V}_i$,

$$\begin{aligned} \langle \mathrm{D}E(I_i v + v_{i-1}^k), I_i w \rangle_{\mathcal{V}',\mathcal{V}} &= \langle I_i^\top \mathrm{D}E(I_i v + v_{i-1}^k), w \rangle_{\mathcal{V}_i',\mathcal{V}_i} \\ &= \langle R_i \mathrm{D}E(I_i v + v_{i-1}^k), w \rangle_{\mathcal{V}_i',\mathcal{V}_i}. \end{aligned}$$

So far, since this rarely leads to confusion, we have dropped the mappings $R_i$ and $I_i$, as their action can be assumed implicitly.

## Evaluating the Ambient Energy

In summary, to solve (1) we must be able to evaluate the derivative mapping

$$\mathcal{V}_i \ni v \mapsto R_i \mathrm{D}E(I_i v + v_{i-1}^k) \in \mathcal{V}_i',$$

and possibly second derivatives if a Newton-type method is used. As we see, even though $v \in \mathcal{V}_i$, a space of possibly small dimension, in order to evaluate we must inject (prolongate) $v$ into $\mathcal{V}$, which is assumed to be of large dimension.

## Subspace Energies

To circumvent the difficulty in evaluating derivatives, in this section we introduce the so-called Fast Subspace Descent method, originally proposed in (Chen, Hu, and Wise, 2000). The idea is that, instead of using the original energy $E$ in (1) we use, for $i = 1, \ldots, N$, a locally defined energy $E_i : \mathcal{V}_i \to \mathbb{R}$. In addition to prolongation and restriction operators, we shall also need, for $i = 1, \ldots, N$, projection operators $Q_i : \mathcal{V} \to \mathcal{V}_i$.

## Definition (FASD)

Let $E$ satisfy (E0) and $u \in \mathcal{V}$ be the unique solution to (**??**). The **Fast Subspace Descent (FASD)** method is an algorithm to approximate $u$ that computes a sequence $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ recursively,

$$u^{k+1} := \mathrm{FASD}\left(u^k\right),$$

and is defined as follows. Assume that, for $i = 1, \ldots, N$, we have Fréchet differentiable $E_i : \mathcal{V}_i \to \mathbb{R}$ and projections $Q_i : \mathcal{V} \to \mathcal{V}_i$. Let $u^0 \in \mathcal{V}$ be arbitrary.

## Definition (Cont.)

Then, for $k \geq 0$,

- $v_0^k = u^k$.
- For $i = 1$ to $N$:
  - Compute the so-called subspace $\tau$-**perturbation**: let $\xi_i^k := Q_i v_{i-1}^k$ and
  $$\tau_i^k := \mathrm{D}E_i(\xi_i^k) - R_i \mathrm{D}E(v_{i-1}^k) \in \mathcal{V}_i'.$$
  - Solve the subspace residual problem: Find $\eta_i^k \in \mathcal{V}_i$, such that
  $$\langle \mathrm{D}E_i(\eta_i^k), w \rangle_{\mathcal{V}_i', \mathcal{V}} = \langle \tau_i^k, w \rangle_{\mathcal{V}_i', \mathcal{V}_i}, \quad \forall\, w \in \mathcal{V}_i. \tag{2}$$
  - Compute the search direction:
  $$s_i^k := \eta_i^k - \xi_i^k \in \mathcal{V}_i.$$

## Definition (Cont.)

- Orthogonalize the subspace correction via line search: Find

$$\alpha_i^k := \operatorname*{argmin}_{\alpha\in\mathbb{R}} E(v_{i-1}^k + \alpha s_i^k)$$
$$= \operatorname*{argzero}_{\alpha\in\mathbb{R}} \langle \mathrm{D}E(v_{i-1}^k + \alpha s_i^k), s_i^k \rangle;$$

  and define

$$\varepsilon_i^k := \alpha_i^k s_i^k. \tag{3}$$

- Apply the subspace correction:

$$v_i^k := v_{i-1}^k + \varepsilon_i^k.$$

- $u^{k+1} := v_N^k$.

**Remark**

*The reader should notice the similarities with the FAS method of the last section and also some differences. The biggest difference is that, in the FAS method, once the coarse-grid correction is found, it is applied. In FASD, there is an additional orthogonalization step.*

## The Cost of Ortogonalization

Notice that, in the orthogonalization step of the FASD method, we perform a line search to find the optimal step size $\alpha_i^k$. This still requires the evaluation of the original energy $E$, which is defined in the ambient space, i.e., $E(v_{i-1}^k + \alpha s_i^k)$. The cost, however, is reduced when compared to the evaluation of $E(v_{i-1}^k + w)$ for arbitrary $w \in \mathcal{V}_i$. In the next section we shall analyze a simplified version of this scheme, one where the step size is found approximately. This is closer to the FAS scheme.

## Remark (Reduction to SSO)

*We note that FASD reduces to SSO when the local energies are defined as*

$$E_i(\eta) := E(v_{i-1}^k - Q_i v_{i-1}^k + \eta), \quad \forall\, \eta \in \mathcal{V}_i,$$

*provided $Q_i$ is a projection. As a consequence of this choice, $\tau_i^k \equiv 0$ and, for all $w \in \mathcal{V}_i$,*

$$\begin{aligned}\langle \mathrm{D}E(v_{i-1}^k + s_i^k), w\rangle &= \langle \mathrm{D}E(v_{i-1}^k - Q_i v_{i-1}^k + \eta_i^k), w\rangle \\ &= \langle \mathrm{D}E_i(\eta_i^k), w\rangle = 0.\end{aligned}$$

*With this choice in FASD, the orthogonalization step is redundant because, upon taking $w = s_i^k$.*

$$\langle \mathrm{D}E(v_{i-1}^k + s_i^k), s_i^k\rangle = 0.$$

*In other words, the orthogonality is valid with $\alpha_i^k = 1$.*

FASD generalizes SSO!

## Convergence Strategy

Let us now analyze the FASD scheme. We shall follow the same path as we did for the SSO method. In other words, all that is needed is to verify the lower and upper bounds on corrections, respectively,

$$E(u^k) - E(u^{k+1}) \geq C_L \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2, \tag{4}$$

and

$$E(u^{k+1}) - E(u) \leq C_U \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2. \tag{5}$$

Verifying the lower bound is relatively easy due to the line search and the convexity of $E$.

## Theorem (Lower Bound on Corrections)

*Assume that the energy $E$ satisfies* (E1), *and that the sequence* $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ *is obtained via the FASD scheme. Then, for* $k \geq 0$,

$$E(u^k) - E(u^{k+1}) \geq \frac{\mu}{2} \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2.$$

## Proof.

We apply a similar technique as in the proof for the SSO case. Due to the line search, we still have a, rather useful, orthogonality property. Namely,

$$\langle \mathrm{D}E(v_i^k), w \rangle = 0, \quad \forall\, w \in \mathrm{span}\{s_i^k\}.$$

## Proof (Cont.)

We may then apply a previous lemma on energy sections, with the subspace $\mathcal{W} := \operatorname{span}\{s_i^k\}$, and use that

$$v_i^k - v_{i-1}^k = \varepsilon_i^k = \alpha_i^k s_i^k \in \operatorname{span}\{s_i^k\},$$

to conclude that

$$\begin{aligned} E(v_{i-1}^k) - E(v_i^k) &\geq \frac{\mu}{2}\|v_{i-1}^k - v_i^k\|_{\mathcal{V}}^2 \\ &= \frac{\mu}{2}\|\varepsilon_i^k\|_{\mathcal{V}}^2. \end{aligned}$$

Adding this inequality for $i = 1, \ldots, N$ we then obtain

$$\begin{aligned} E(u^k) - E(u^{k+1}) &= \sum_{i=1}^{N} \Big( E(v_{i-1}^k) - E(v_i^k) \Big) \\ &\geq \frac{\mu}{2} \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2. \end{aligned} \qquad \square$$

Here is the result from the last chapter of which we made careful use.

## Lemma (Convexity of Sections)

*Let $\mathcal{V}$ be a Hilbert space and $E : \mathcal{V} \to \mathbb{R}$ be strongly convex and locally Lipschitz smooth. Fix $\alpha \in \mathbb{R}$ and set $\mathcal{B} = \mathrm{level}_\alpha(E)$. Let $\xi \in \mathcal{B}$ be arbitrary, and $\mathcal{W} \subseteq \mathcal{V}$ be a subspace. Define the section*

$$J_\xi(w) := E(\xi + w), \quad \forall\, w \in \mathcal{W}.$$

*Then, $J_\xi : \mathcal{W} \to \mathbb{R}$ is differentiable, strongly convex, and locally Lipschitz smoother. Furthermore, there exists a unique element $\eta \in \mathcal{W}$ such that $\xi + \eta \in \mathcal{B}$, $\eta$ is the unique global minimizer of $J_\xi$, and*

$$\langle \mathrm{D}E(\xi + \eta), w \rangle = \langle \mathrm{D}J_\xi(\eta), w \rangle = 0, \quad \forall\, w \in \mathcal{W}.$$

*As a consequence, for all $w \in \mathcal{W}$ with $w + \xi \in \mathcal{B}$,*

$$\frac{\mu}{2} \|w - \eta\|_{\mathcal{V}}^2 \le J(w) - J(\eta) = E(\xi + w) - E(\xi + \eta) \le \frac{L}{2} \|w - \eta\|_{\mathcal{V}}^2 .$$

*The lower bound holds for any $w \in \mathcal{W}$, without restriction.*

## Inflating the Energy Set

The derivation of the upper bound on corrections is more complicated, and it requires some assumptions on the local energies $E_i : \mathcal{V}_i \to \mathbb{R}$. To properly state them, from now on, we set

$$\mathcal{B} := \mathrm{level}_{E(u^0)}(E). \tag{6}$$

The last theorem implies that the energy is always decreasing between iterates. Thus $\{u^k\}_{k=0}^{\infty} \subset \mathcal{B}$. However, the search region and, e.g., the point $\xi_i^k + s_i^k$, may not be contained in $\mathcal{B}$. To be able to use Lipschitz smoothness, we introduce a larger domain

$$\mathcal{B}^+ := \{v \in \mathcal{V} \mid \mathrm{dist}(v, \mathcal{B}) \leq \sqrt{\chi}\}, \tag{7}$$

where

$$\chi := \frac{2L^2}{\mu \min_{i=1}^N \mu_i^2}(E(u^0) - E(u)).$$

With this at hand we assume that, for $i = 1, \ldots, N$, the local energies $E_i$ satisfy the following assumptions.

### Definition (Assumption (E2))

Let $\mathcal{V}$ be a Hilbert space, $\{\mathcal{V}_i\}_{i=1}^N$ a subspace decomposition, and, for $i = 1, \ldots, N$, $E_i : \mathcal{V}_i \to \mathbb{R}$. We say that **Assumption (E2) holds** iff the local energies $\{E_i\}_{i=1}^N$ satisfy the following:

(E2a) The local energy $E_i$ is Fréchet differentiable on $\mathcal{V}_i$.

(E2b) (Strong convexity) There exists a constant $\mu_i$ such that, for all $v, w \in \mathcal{V}_i$,

$$\langle \mathrm{D}E_i(w) - \mathrm{D}E_i(v), w - v \rangle \geq \mu_i \|w - v\|_{\mathcal{V}}^2.$$

(E2c) (Lipschitz continuity of the derivative) The derivatives of the energies $E_i$ are locally Lipschitz continuous. Let $\mathcal{B}$ be defined in (6), and $\mathcal{B}^+$ in (7). Set $\mathcal{B}_i = Q_i \mathcal{B}^+$. We denote the local constant of $E_i$ over $\mathcal{B}_i$ as $L_i$.

The following result will be useful in the analysis that follows.

## Proposition (Convexity)

*The sets $\mathcal{B}^+$ and $\mathcal{B}_i$ are convex.*

## Proof.

Observe that

$$\mathcal{B}^+ = \mathcal{B} + \left\{ v \in \mathcal{V} \,\middle|\, \|v\|_{\mathcal{V}} \leq \sqrt{\chi} \right\}.$$

The second set is, of course, the ball of radius $\sqrt{\chi}$ centered at $0 \in \mathcal{V}$. Since both sets are convex, so is their sum.

Next, recall that $\mathcal{B}_i$ is the image of a convex set, namely $\mathcal{B}^+$, through a linear map $Q_i$. □

Later, during the course of the proof of convergence, we will show that, for any $u^0 \in \mathcal{V}$ and all $k \geq 0$, $\xi_i^k + s_i^k \in \mathcal{B}_i$. Thus, (E2c) shall become relevant. It is interesting to note that, at this stage, there is no relation between the energy $E$ and the local energies $E_i$ besides satisfying similar properties.

## Lemma (Local Quadratic Trap)

*Assume that, for $i = 1, \dots, N$, the local energy $E_i$ satisfies (E2b) and (E2c). Then, for any $v, w \in \mathcal{B}_i$,*

$$\mu_i \|w - v\|_{\mathcal{V}}^2 \leq \langle \mathrm{D}E_i(w) - \mathrm{D}E_i(v), w - v \rangle \leq L_i \|w - v\|_{\mathcal{V}}^2 ,$$

*and*

$$\frac{\mu_i}{2} \|w - v\|_{\mathcal{V}}^2 + \langle \mathrm{D}E_i(v), w - v \rangle \leq E_i(w) - E_i(v) \leq \langle \mathrm{D}E_i(v), w - v \rangle + \frac{L_i}{2} \|w - v\|_{\mathcal{V}}^2 .$$

*Moreover, if we denote by $u_i \in \mathcal{V}_i$ the minimizer of $E_i$, and we have that $u_i \in \mathcal{B}_i$, then, for all $w \in \mathcal{B}_i$,*

$$\frac{\mu_i}{2} \|w - u_i\|_{\mathcal{V}}^2 \leq E_i(w) - E_i(u_i) \leq \frac{L_i}{2} \|w - u_i\|_{\mathcal{V}}^2 .$$

*The lower bounds above hold for all $w \in \mathcal{V}_i$.*

## Why the $\tau$-Perturbation?

The next point to be noticed is that, in (2), we are not minimizing $E_i$, which would entail solving $\mathrm{D}E_i(s_i^k + Q_i v_{i-1}^k) = 0$. Instead, a so-called $\tau$-perturbation is added to the right hand side. The following considerations build up to an explanation of why this is needed.

## Lemma (Descent Direction)

*Assume that the local energies satisfy* (E2). *Assume that, for* $k \geq 0$, $s_i^k$ *is computed with the FASD method. Then* $s_i^k$ *is a descent direction in the sense that*

$$\langle -R_i \mathrm{D}E(v_{i-1}^k), s_i^k \rangle_{\mathcal{V}_i', \mathcal{V}_i} \geq \mu_i \|s_i^k\|_{\mathcal{V}}^2,$$

*or, equivalently,*

$$\langle \mathrm{D}E(v_{i-1}^k), I_i s_i^k \rangle \leq -\mu_i \|s_i^k\|_{\mathcal{V}}^2.$$

## Proof.

By the definition of the $\tau$-perturbation, the local problem (2) can be rewritten as follows. Find $\eta_i^k \in \mathcal{V}_i$ s.t.

$$\langle \mathrm{D}E_i(\eta_i^k) - \mathrm{D}E_i(\xi_i^k), w \rangle_{\mathcal{V}_i', \mathcal{V}_i} = -\langle R_i \mathrm{D}E(v_{i-1}^k), w \rangle_{\mathcal{V}_i', \mathcal{V}_i}, \quad \forall\, w \in \mathcal{V}_i, \tag{8}$$

here recall that $\xi_i^k = Q_i v_{i-1}^k \in \mathcal{V}_i$ and $\eta_i^k = \xi_i^k + s_i^k$.

## Proof (Cont.)

Choose $w = s_i^k$ and use the strong convexity of $E_i$ to obtain

$$\begin{aligned}\langle -R_i \mathrm{D}E(v_{i-1}^k), s_i^k \rangle &= \langle \mathrm{D}E_i(\eta_i^k) - \mathrm{D}E_i(\xi_i^k), s_i^k \rangle \\ &\geq \mu_i \|s_i^k\|_{\mathcal{V}}^2.\end{aligned}$$

$\square$

In order to better understand the choice of the step size, we introduce the scalar function $f_{i,k}$, which is a one dimensional section of the energy $E$.

$$f_{i,k}(\alpha) := E(v_{i-1}^k + \alpha s_i^k). \tag{9}$$

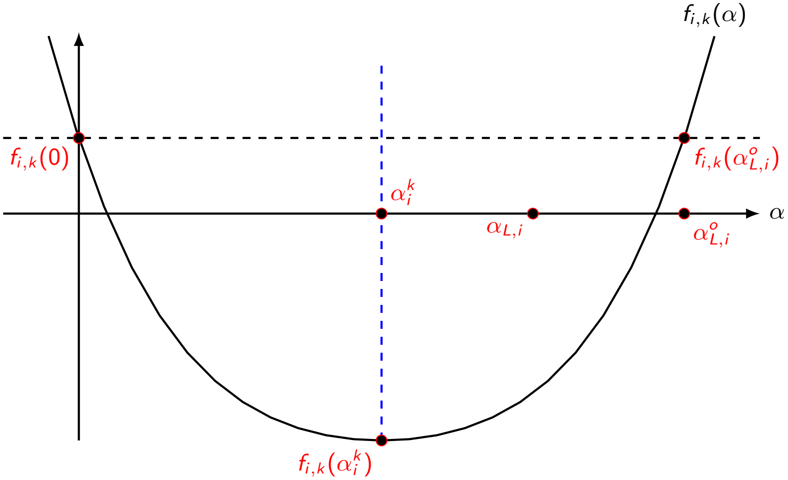


Figure: The function $f_{i,k}$ defined in (9). $f_{i,k}$ is a one-dimensional energy section. It is straightforward to prove that its minimizer, $\alpha^k_i$ is positive.

## Proposition (Energy Section I)

*Assume that $E$ satisfies* (E1) *and the local energies satisfy* (E2)*. The one-dimensional energy section $f_{i,k}$, defined in* (9)*, satisfies*

$$f'_{i,k}(0) = \langle \mathrm{D}E(v^k_{i-1}), s^k_i \rangle \le -\mu_i \left\| s^k_i \right\|^2_{\mathcal{V}} .$$

*Furthermore, $\alpha^k_i > 0$ and, for all $\alpha \in (0, \alpha^k_i]$, $f_{i,k}(\alpha) < f_{i,k}(0)$.*

## Lemma (Energy Section II)

*Assume that $E$ satisfies* (E1) *and the local energies satisfy* (E2)*. Then $f_{i,k}$, defined in* (9)*, is differentiable and strongly convex in the following sense: for all $\alpha, \beta \in \mathbb{R}$,*

$$(f'_{i,k}(\alpha) - f'_{i,k}(\beta))(\alpha - \beta) \ge (\alpha - \beta)^2 \mu \| s^k_i \|^2_{\mathcal{V}}.$$

*Furthermore, $f'_{i,k}$ is locally Lipschitz in the following sense: for all $0 \le \alpha, \beta \le \alpha_{L,i}$,*

$$|f'_{i,k}(\alpha) - f'_{i,k}(\beta)| \le L \| s^k_i \|^2_{\mathcal{V}} |\alpha - \beta|,$$

*where $\alpha_{L,i} := (1 + \sqrt{\mu/L})\alpha^k_i$.*

## Lemma (Lower Bound on $\alpha_i^k$)

*Assume that $E$ satisfies* (E1) *and the local energies satisfy* (E2). *In this setting, we have the lower bound*

$$\frac{\mu_i}{L} \le \alpha_i^k. \tag{10}$$

*Consequently,*

$$\alpha_{L,i}^o > \frac{\mu_i}{L}.$$

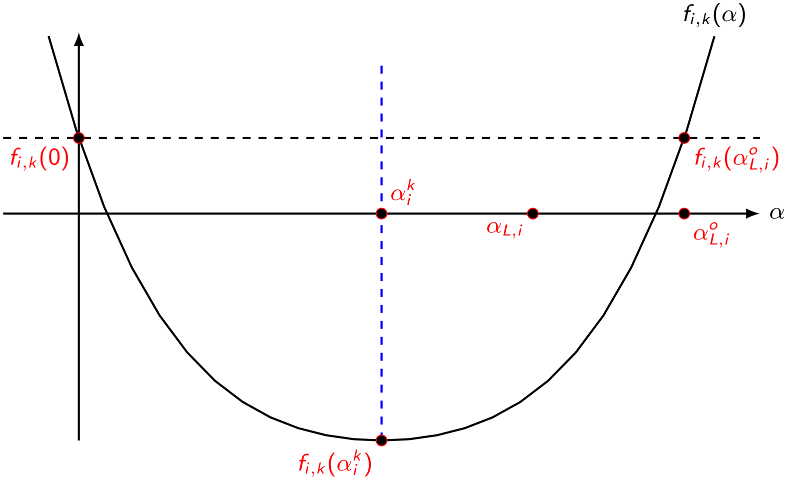


Figure: The function $f_{i,k}$ defined in (9). $f_{i,k}$ is a one-dimensional energy section. It is straightforward to prove that its minimizer, $\alpha^k_i$ is positive.

## Lemma (Improved Upper Bound)

*Assume that $E$ satisfies* (E1) *and the local energies satisfy* (E2). *Then, we have the upper bound*

$$\alpha_i^k \leq \frac{L_i}{\mu}.$$

## Theorem (Upper Bound on Corrections)

*Assume that the space decomposition satisfies (NS1) and (NS2), that the energy $E$ satisfies* (E1) *and the local energies, $E_i$, satisfy* (E2). *Then, the sequence $\{u^k\}_{k=0}^{\infty}$, produced by the FASD method of Definition 1, satisfies the upper bound*

$$E(u^{k+1}) - E(u) \leq C_U \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2,$$

*where*

$$C_U := \frac{C_A^2}{2\mu}\left[C_S + L\left(1 + \max_{i=1}^{N}\left\{\frac{L_i}{\mu_i}\right\}\right)\right]^2.$$

**Proof.**

Let $w \in \mathcal{V}$ be arbitrary and we choose a stable decomposition $w = \sum_{i=1}^N w_i$. Then

$$\begin{aligned}\langle \mathrm{D}E(u^{k+1}), w\rangle &= \sum_{i=1}^N \langle \mathrm{D}E(u^{k+1}), w_i\rangle \\ &= \sum_{i=1}^N \Big(\langle \mathrm{D}E(u^{k+1}) - \mathrm{D}E(v_i^k), w_i\rangle + \langle \mathrm{D}E(v_i^k), w_i\rangle\Big) \\ &= \mathrm{I}_1 + \mathrm{I}_2,\end{aligned}$$

where

$$\mathrm{I}_1 := \sum_{i=1}^N \langle \mathrm{D}E(u^{k+1}) - \mathrm{D}E(v_i^k), w_i\rangle,$$

$$\mathrm{I}_2 := \sum_{i=1}^N \langle \mathrm{D}E(v_i^k), w_i\rangle.$$

## Proof (Cont.)

Using the stability of the decomposition (NS1) and the strengthened Cauchy-Schwartz inequality (NS2), the term $\mathrm{I}_1$ can be estimated in exactly the same way as for the SSO method. Therefore,

$$\mathrm{I}_1 \leq C_S C_A \left( \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2 \right)^{1/2} \|w\|_{\mathcal{V}}.$$

## Proof (Cont.)

For $\mathrm{I}_2$ we insert $\tau_i^k - \mathrm{D}E_i(\xi_i^k + s_i^k)$, which is zero in $\mathcal{V}_i'$, and estimate as follows,

$$
\begin{aligned}
\mathrm{I}_2 &= \sum_{i=1}^{N} \langle \mathrm{D}E(v_i^k) - \mathrm{D}E(v_{i-1}^k) - \mathrm{D}E_i(\xi_i^k + s_i^k) + \mathrm{D}E_i(\xi_i^k), w_i \rangle \\
&\overset{(E1)+(E2c)}{\leq} \sum_{i=1}^{N} \left( L\|\varepsilon_i^k\|_{\mathcal{V}} + L_i\|s_i^k\|_{\mathcal{V}} \right) \|w_i\|_{\mathcal{V}} \\
&\overset{(3)}{=} \sum_{i=1}^{N} \left( L\|\varepsilon_i^k\|_{\mathcal{V}} + \frac{L_i}{\alpha_i^k}\|\varepsilon_i^k\|_{\mathcal{V}} \right) \|w_i\|_{\mathcal{V}} \\
&\overset{(10)}{\leq} L \sum_{i=1}^{N} \left( 1 + \frac{L_i}{\mu_i} \right) \|\varepsilon_i^k\|_{\mathcal{V}} \|w_i\|_{\mathcal{V}} \\
&\leq L \left( 1 + \max_{i=1}^{N} \frac{L_i}{\mu_i} \right) \left( \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2 \right)^{1/2} \left( \sum_{i=1}^{N} \|w_i\|_{\mathcal{V}}^2 \right)^{1/2} \\
&\overset{(NS1)}{\leq} L C_A \left( 1 + \max_{i=1}^{N} \frac{L_i}{\mu_i} \right) \left( \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2 \right)^{1/2} \|w\|_{\mathcal{V}}.
\end{aligned}
$$

## Proof (Cont.)

Putting the estimates for $\mathrm{I}_1$ and $\mathrm{I}_2$ together we have, for any $w \in \mathcal{V}$,

$$\langle \mathrm{D}E(u^{k+1}), w \rangle \leq C_A \left[ C_S + L \left( 1 + \max_{i=1}^{N} \frac{L_i}{\mu_i} \right) \right] \left( \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2 \right)^{1/2} \|w\|_{\mathcal{V}},$$

which implies that

$$\left\| \mathrm{D}E(u^{k+1}) \right\|_{\mathcal{V}'}^2 \leq C_A^2 \left[ C_S + L \left( 1 + \max_{i=1}^{N} \frac{L_i}{\mu_i} \right) \right]^2 \sum_{i=1}^{N} \|\varepsilon_i^k\|_{\mathcal{V}}^2.$$

The result follows, as in the proof for SSO convergence. $\square$

## Corollary (Convergence)

*Assume that the space decomposition satisfies (NS1) and (NS2), that the energy $E$ satisfies* (E1) *and the local energies, $E_i$, satisfy* (E2)*. Let $\{u^k\}_{k=0}^{\infty} \subset \mathcal{V}$ be the sequence generated by the FASD method of Definition 1. Then, we have*

$$E(u^{k+1}) - E(u) \leq \rho(E(u^k) - E(u)),$$

*with*

$$\rho := \frac{C_A^2 \left[C_S + L\left(1 + \max_i\{L_i/\mu_i\}\right)\right]^2}{C_A^2 \left[C_S + L\left(1 + \max_i\{L_i/\mu_i\}\right)\right]^2 + \mu^2}.$$