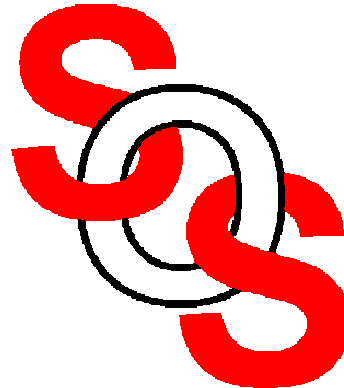


Trust Regions and Aggressive Space Mapping: New Illustrations and Insight

John W. Bandler and Sameh Dakroury

Simulation Optimization Systems Research Laboratory
McMaster University



Bandler Corporation, www.bandler.com
john@bandler.com



presented at

Workshop on "Optimization Engines for Wireless and Microwave Computer Aided Engineering"
Carleton University, Ottawa, ON, June 20, 2002



Outline

aggressive space mapping

trust region method

trust region and aggressive space mapping

TRASM algorithm

example: wedge cutting problem

conclusion



Aggressive Space Mapping

(Bandler et al., 1995)

find the optimal fine model solution by solving the system of nonlinear equations

$$\mathbf{f} = \mathbf{P}(\mathbf{x}) - \mathbf{z}^* = \mathbf{0}$$

using a quasi-Newton step to find the next iterate

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{h}^{(i)}$$

by solving a system of linear equations

$$\mathbf{B}^{(i)} \mathbf{h}^{(i)} = -\mathbf{f}^{(i)}$$

where $\mathbf{B}^{(i)}$ is an approximation to the Jacobian of $\mathbf{P}(\mathbf{x}^{(i)})$



Trust Region Method

(Levenberg 1944, Powell 1970)

at the current iteration, an approximate linear or quadratic model is used to approximate the objective function

this approximate model is trusted only within a region of a specific radius (trust region size)

if the model minimizer achieves sufficient actual reduction in the objective function, the trust region size is increased

if sufficient reduction is not achieved then the trust region is reduced or kept unchanged



Trust Region Aggressive Space Mapping

(Bakr et al., 2000)

find the next iterate

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{h}^{(i)}$$

by solving the trust region subproblem

$$\mathbf{h}^{(i)} = \arg \min_h \left\| \mathbf{B}^{(i)} \mathbf{h} + \mathbf{f}^{(i)} \right\|$$

subject to

$$\|\mathbf{h}\| \leq \delta^{(i)}$$



TRASM Algorithm (*Bakr et al., 2000*)

Step 1 initialize $\mathbf{x}^{(1)} = \mathbf{z}^*$, $\mathbf{B}^{(1)} = \mathbf{I}$, $i = 1$, $\delta^{(1)} = \delta_0$

Step 2 extract $\mathbf{z}^{(1)}$ such that $\mathbf{R}_c(\mathbf{z}^{(1)}) \approx \mathbf{R}_f(\mathbf{x}^{(1)})$

Step 3 evaluate $\mathbf{f}^{(1)} = \mathbf{z}^{(1)} - \mathbf{z}^*$, if $\|\mathbf{f}^{(1)}\| \leq \varepsilon$, stop

Step 4 find the minimizer $\mathbf{h}^{(i)} = \arg \min_h \|\mathbf{B}^{(i)} \mathbf{h} + \mathbf{f}^{(i)}\|$

subject to $\|\mathbf{h}\| \leq \delta^{(i)}$

Step 5 set $\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \mathbf{h}^{(i)}$



TRASM Algorithm (*Bakr et al., 2000*)

Step 6 evaluate $\mathbf{R}_f(\mathbf{x}^{(i+1)})$

Step 7 extract $\mathbf{z}^{(i+1)}$ such that $\mathbf{R}_c(\mathbf{z}^{(i+1)}) \approx \mathbf{R}_f(\mathbf{x}^{(i+1)})$

Step 8 evaluate $\mathbf{f}^{(i+1)} = \mathbf{z}^{(i+1)} - \mathbf{z}^*$, if $\|\mathbf{f}^{(i+1)}\| \leq \varepsilon$, stop

Step 9 find $\rho^{(i)} = \frac{\|\mathbf{f}^{(i)}\| - \|\mathbf{f}^{(i+1)}\|}{\|\mathbf{f}^{(i)}\| - \|\mathbf{f}^{(i)} + \mathbf{B}^{(i)}\mathbf{h}^{(i)}\|}$

Step 10 adjust the trust region size

if $\rho^{(i)} < \eta_1$ reject $\mathbf{x}^{(i+1)}$, select $\delta^{(i+1)} \in [\alpha_1\delta^{(i)}, \alpha_2\delta^{(i)}]$



TRASM Algorithm (*Bakr et al., 2000*)

else if $\eta_1 \leq \rho^{(i)} < \eta_2$, accept $\mathbf{x}^{(i+1)}$, $\delta^{(i+1)} \in [\alpha_2 \delta^{(i)}, \delta^{(i)}]$

else accept $\mathbf{x}^{(i+1)}$, take $\delta^{(i+1)} \geq \delta^{(i)}$

Comment $0 < \alpha_1 \leq \alpha_2 < 1$, $0 < \eta_1 \leq \eta_2 < 1$

(for example, $\eta_1 = 0.1$, $\eta_2 = 0.9$ and $\alpha_1 = \alpha_2 = 0.5$)

Step 11 update $\mathbf{B}^{(i+1)} = (\mathbf{J}_c^T \mathbf{J}_c)^{-1} \mathbf{J}_c^T \mathbf{J}_f$

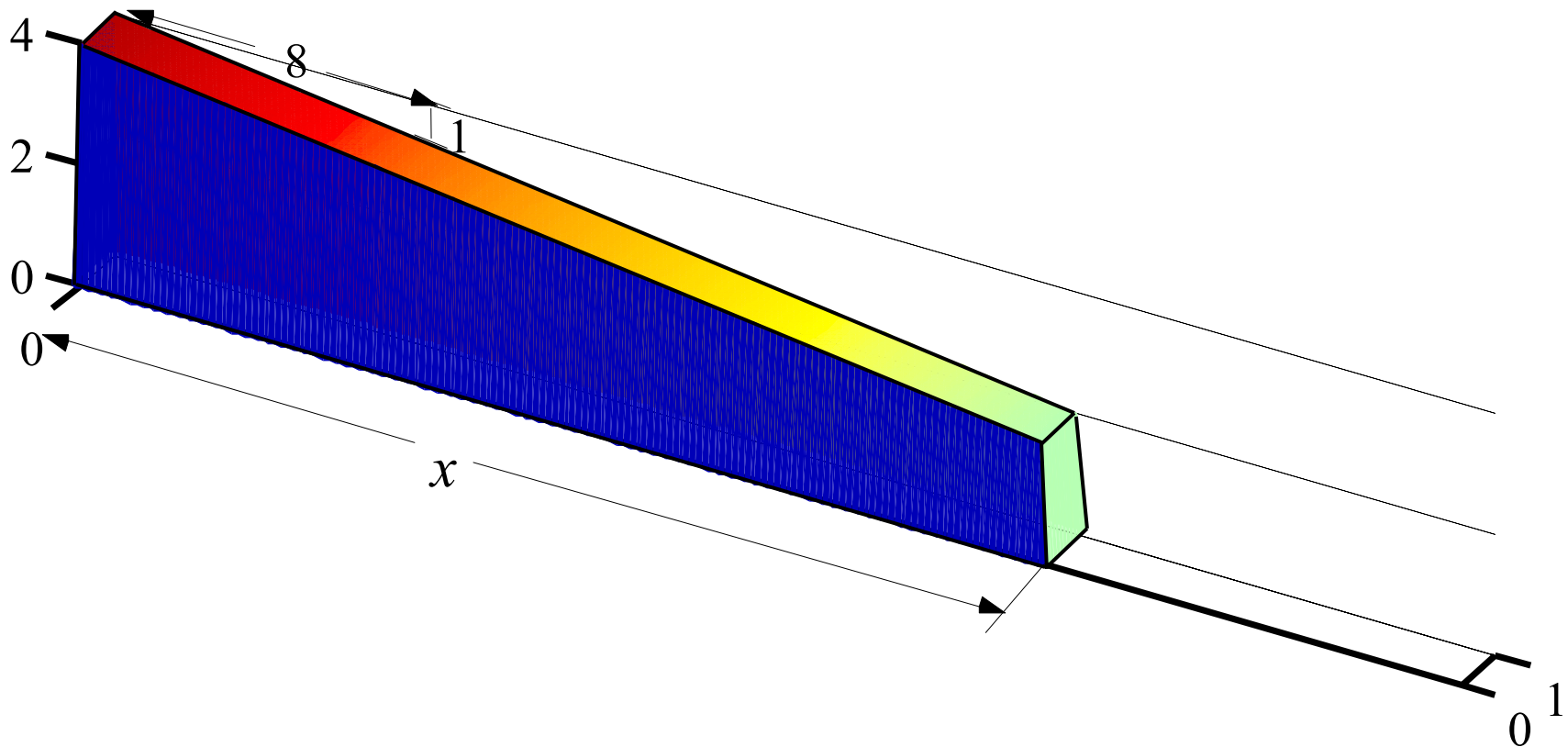
Comment \mathbf{J}_f , \mathbf{J}_c are evaluated at $\mathbf{x}^{(i+1)}$, $\mathbf{z}^{(i+1)}$

Step 12 set $i = i + 1$ and go to *Step 4*



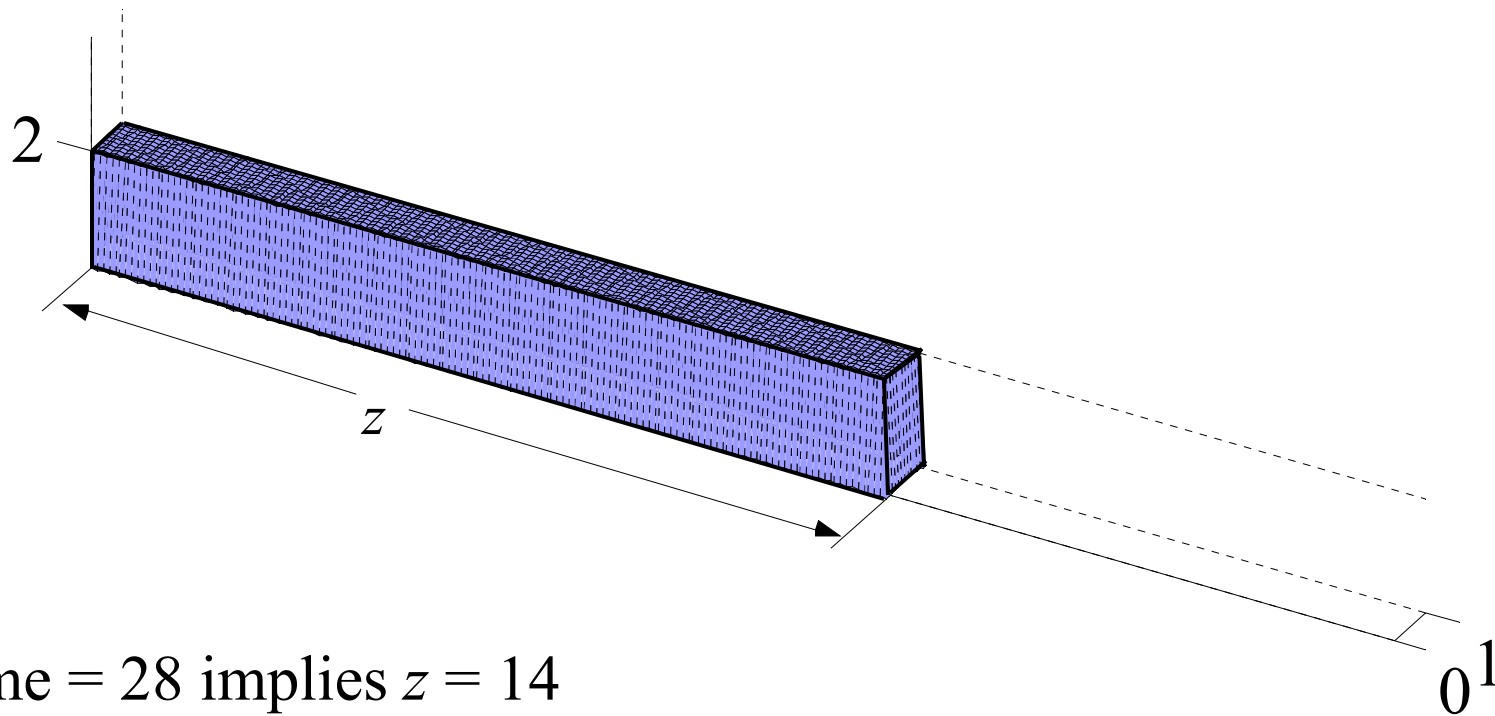
Wedge Cutting Problem (*Bandler, 2002*)

use space mapping to find the optimal position x of a cut such that the volume is equal to 28





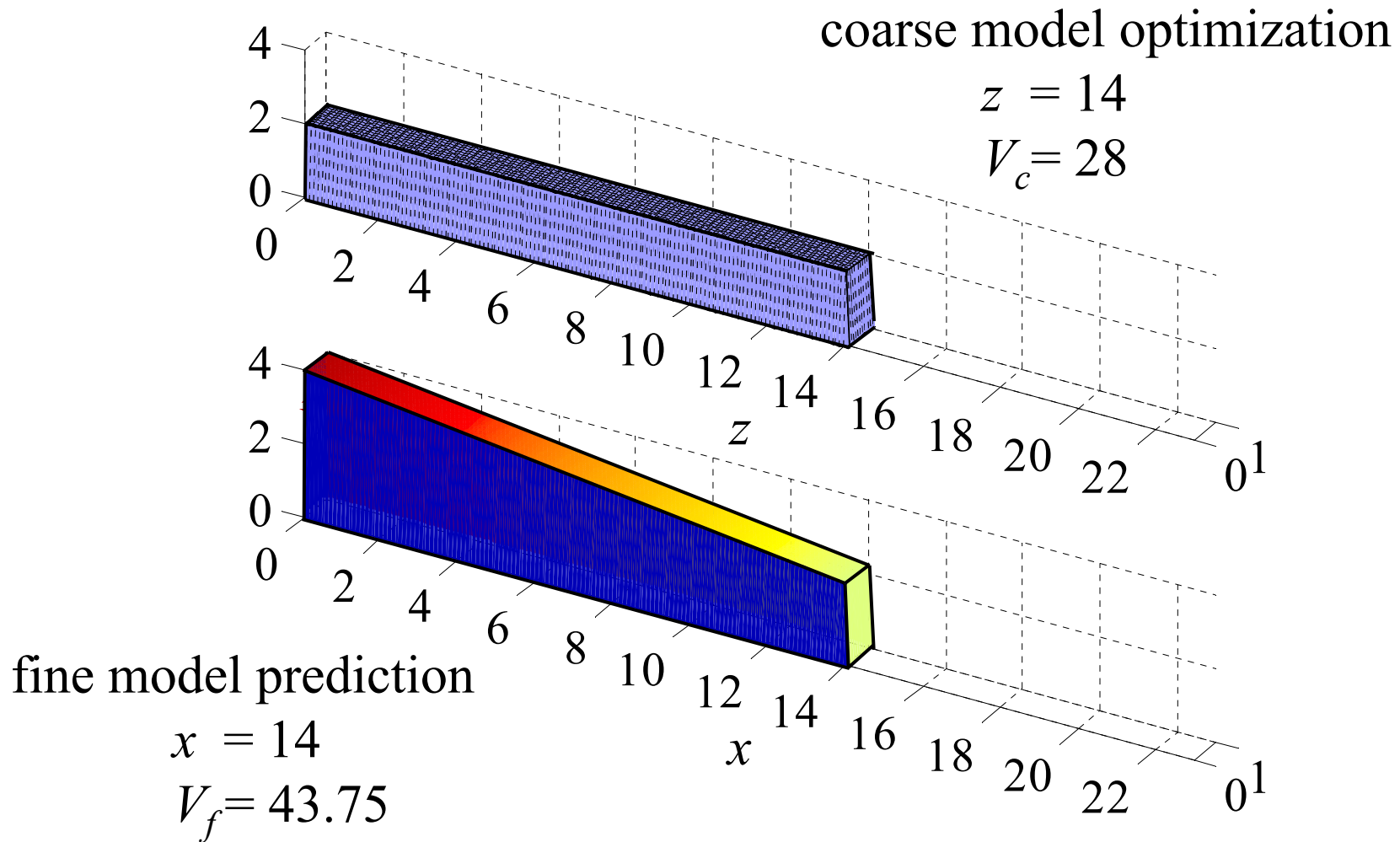
Proposed Coarse Model



volume = 28 implies $z = 14$

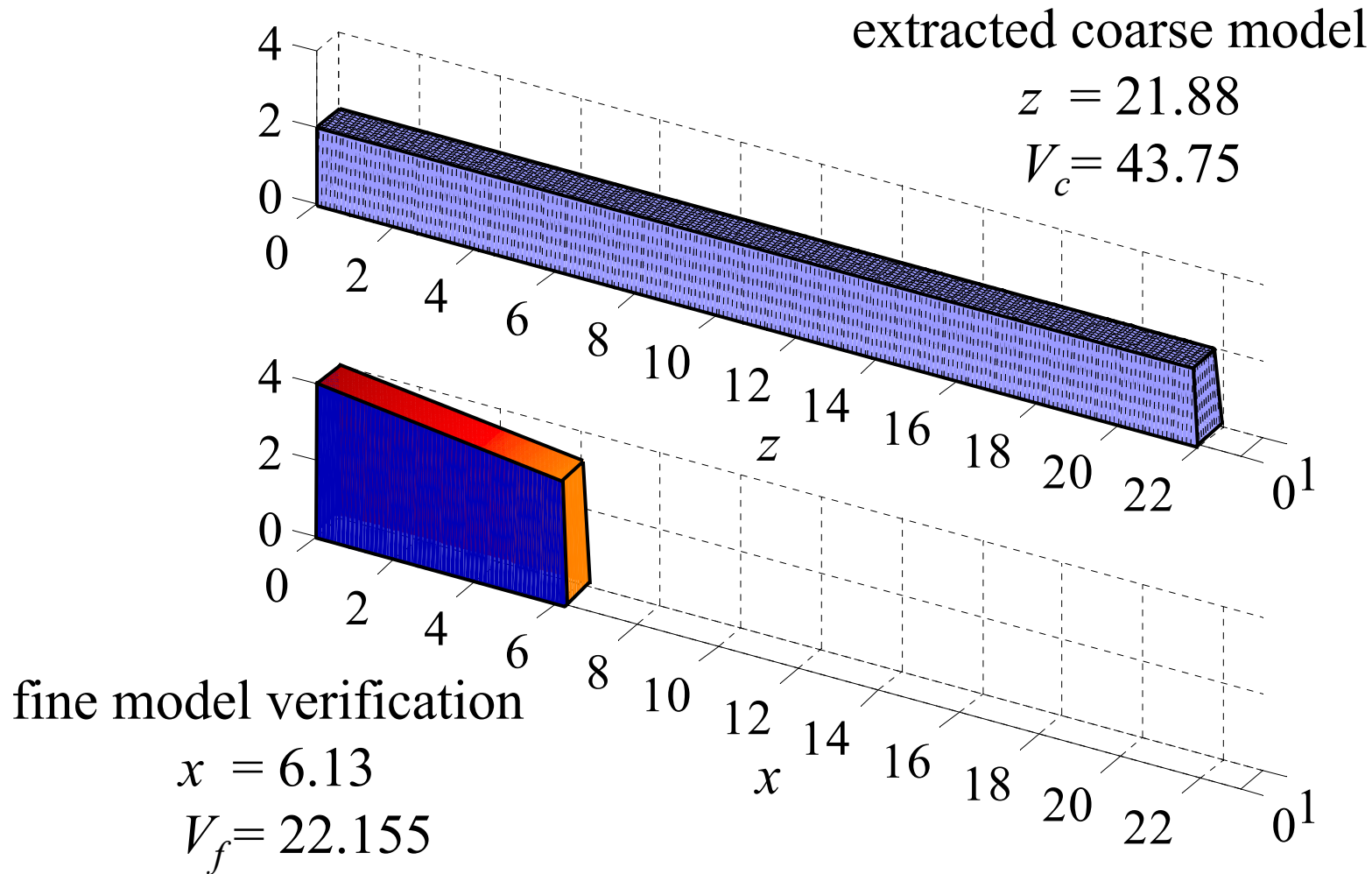


Initialization: Conventional ASM



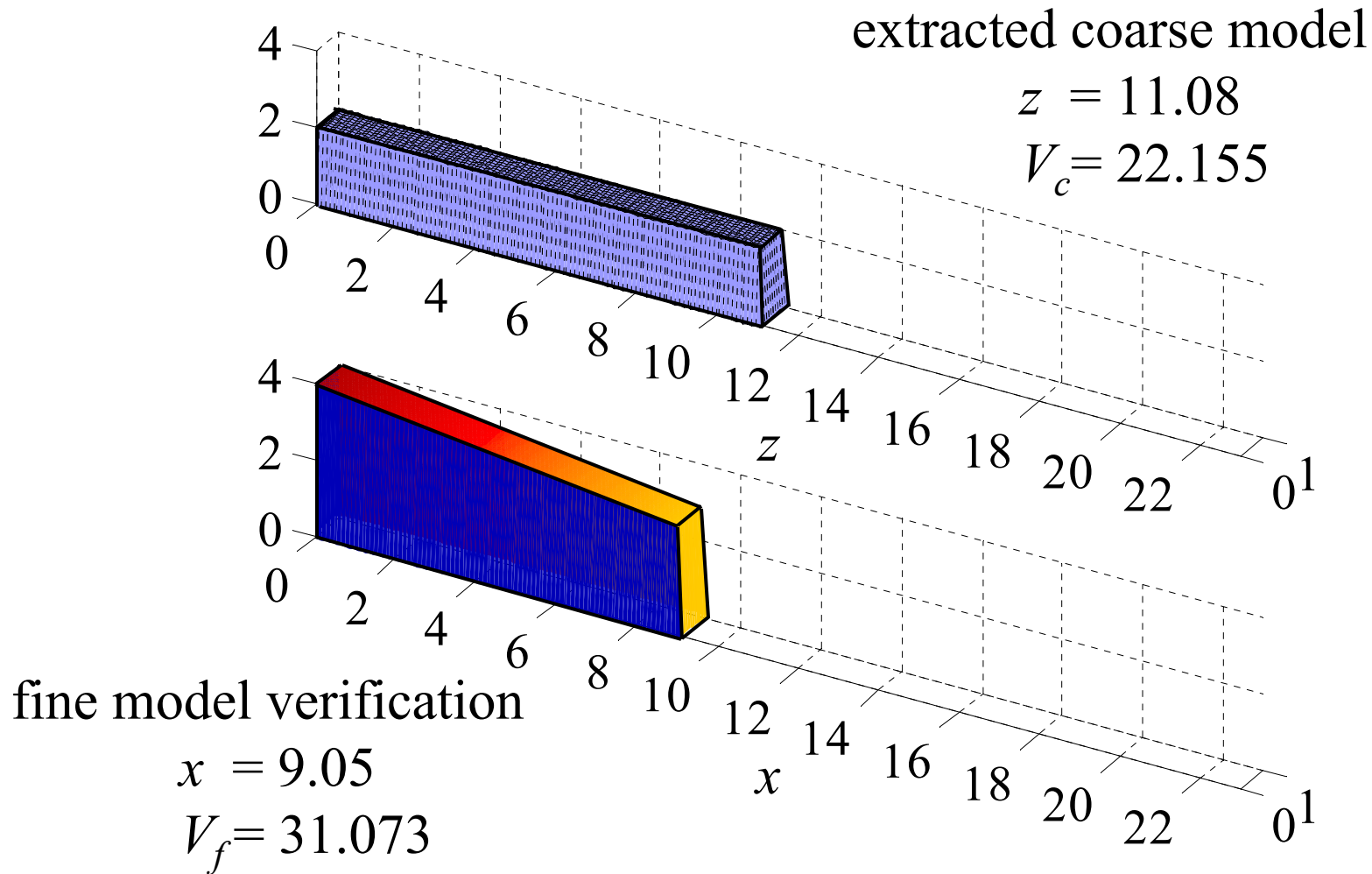


Iteration 1: Conventional ASM



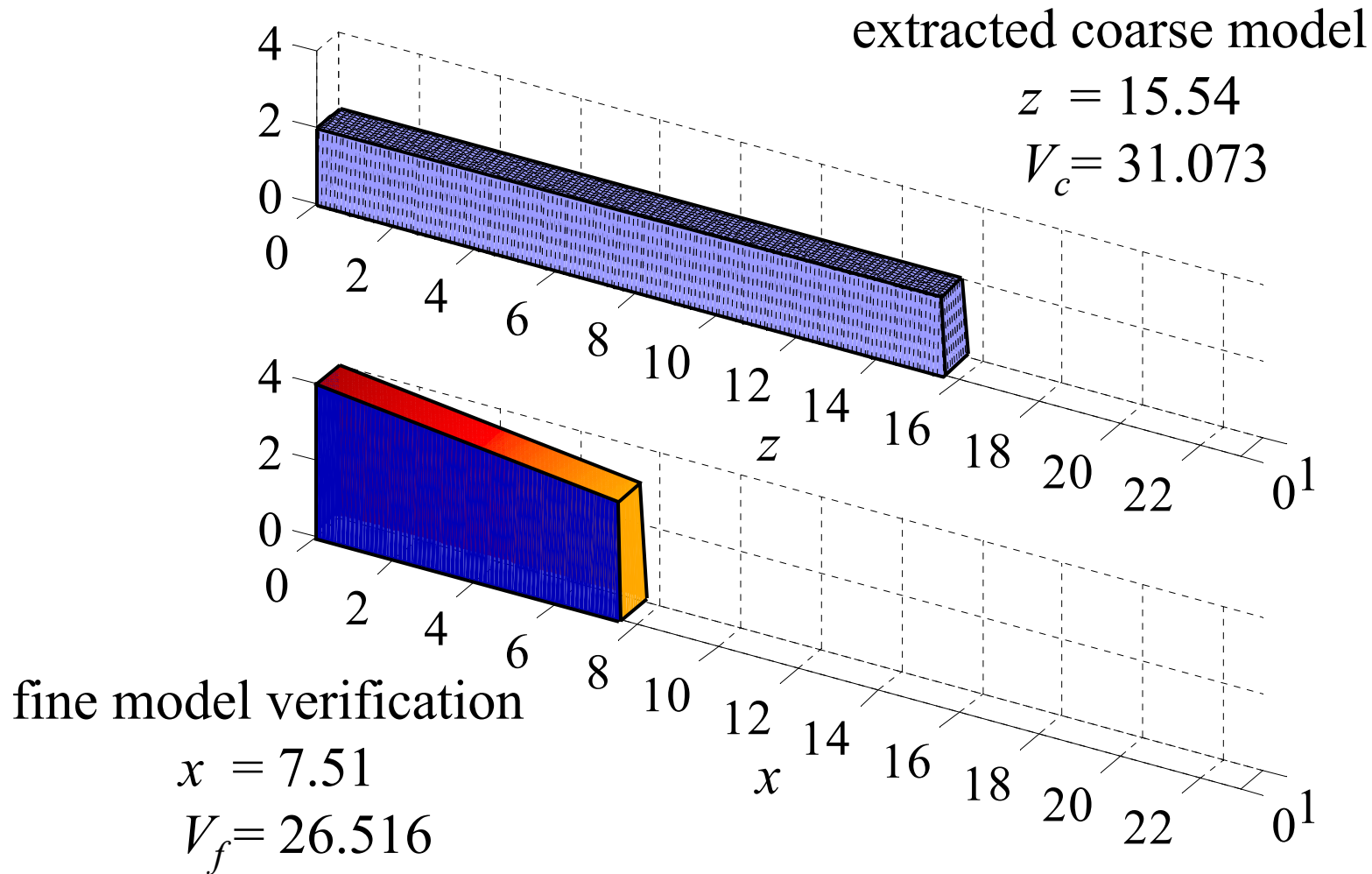


Iteration 2: Conventional ASM



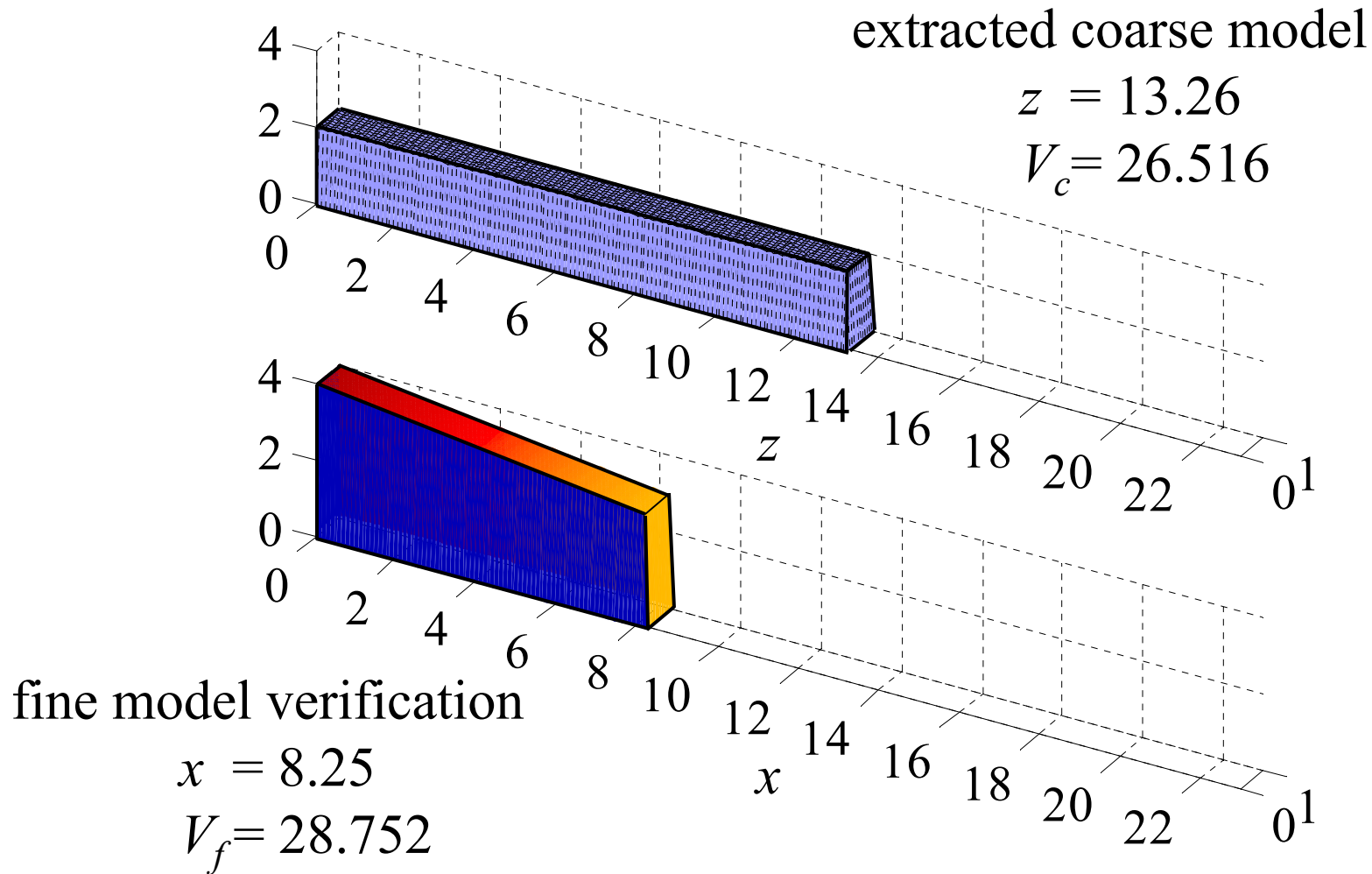


Iteration 3: Conventional ASM



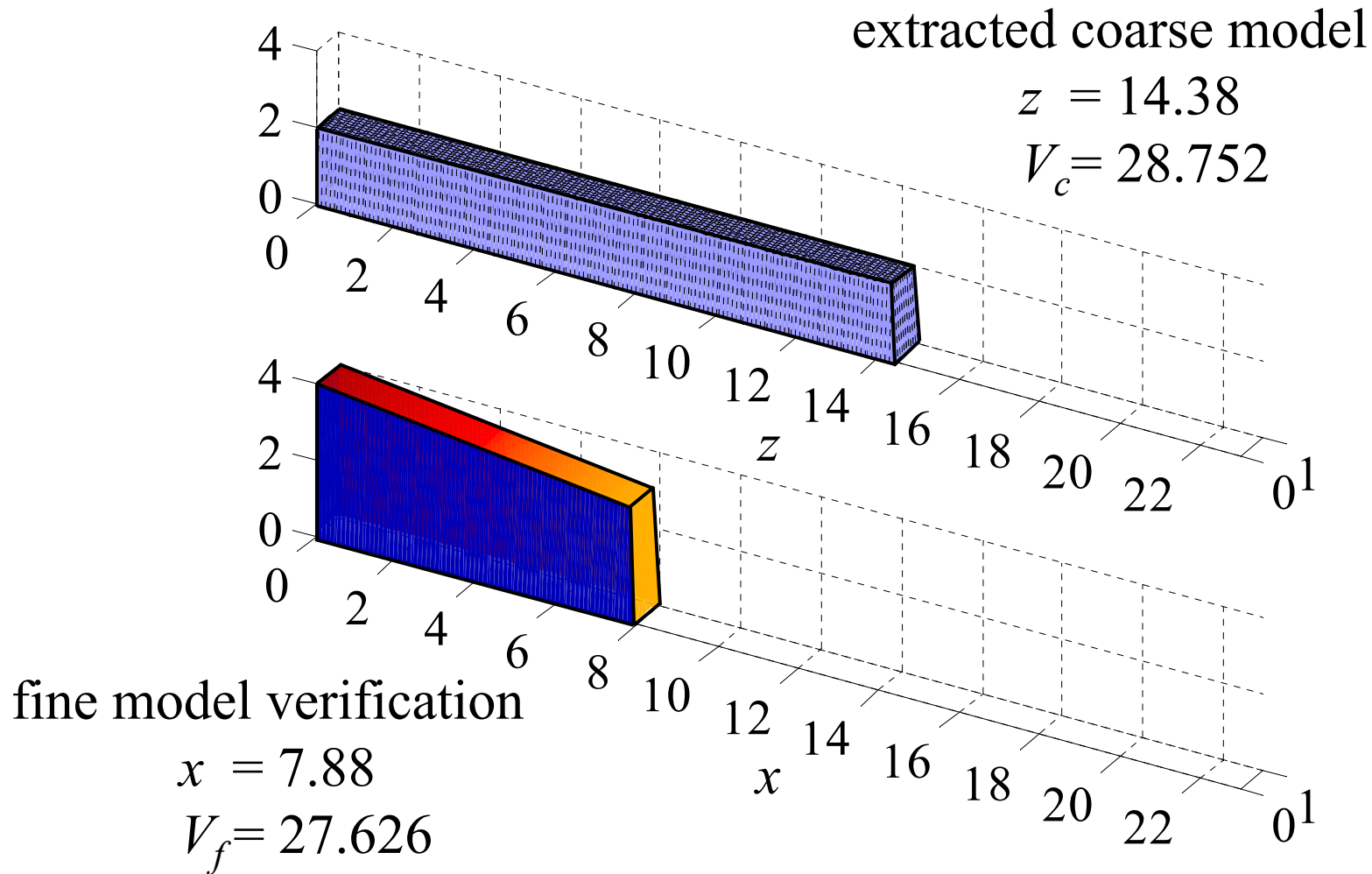


Iteration 4: Conventional ASM



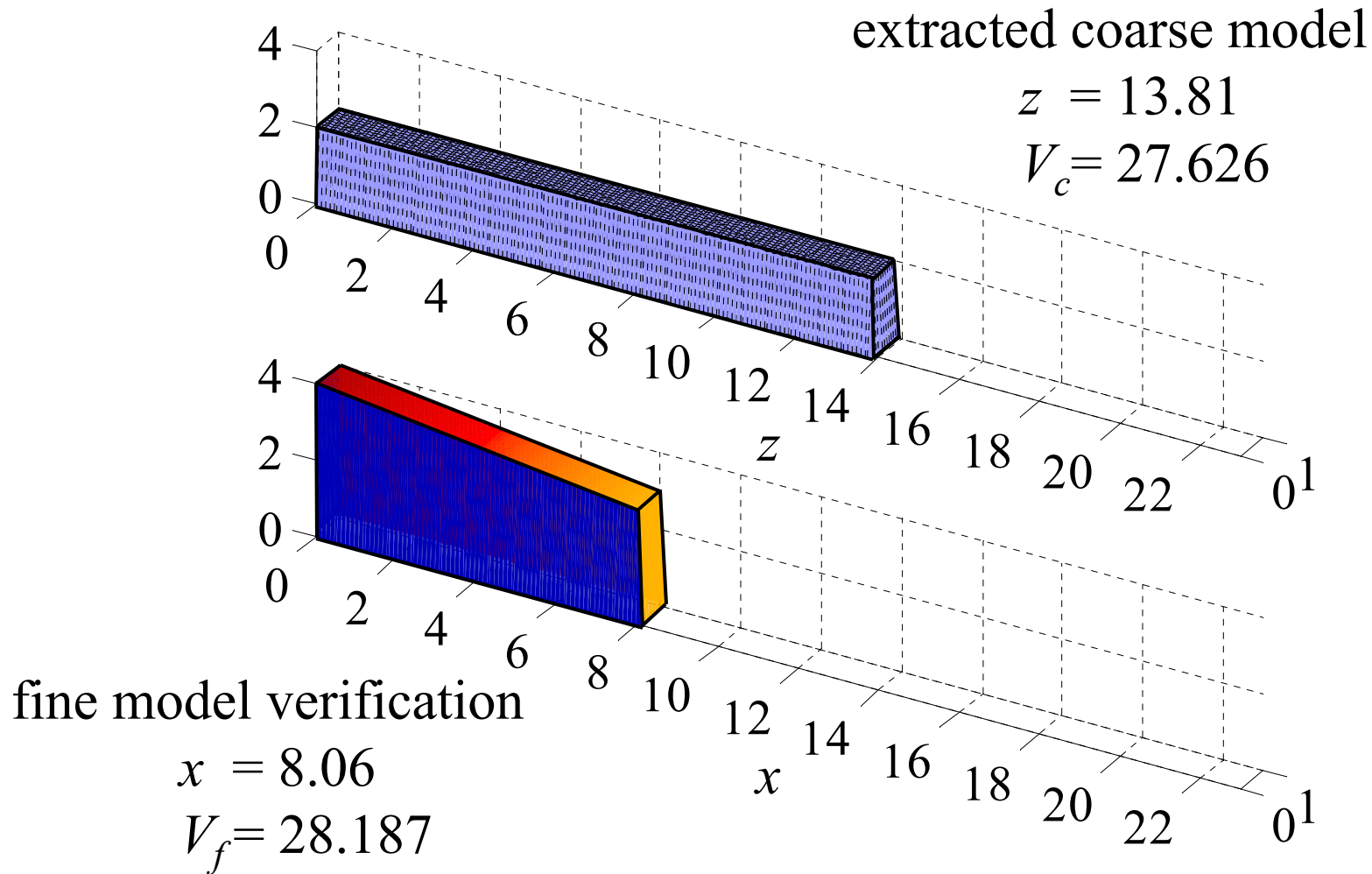


Iteration 5: Conventional ASM



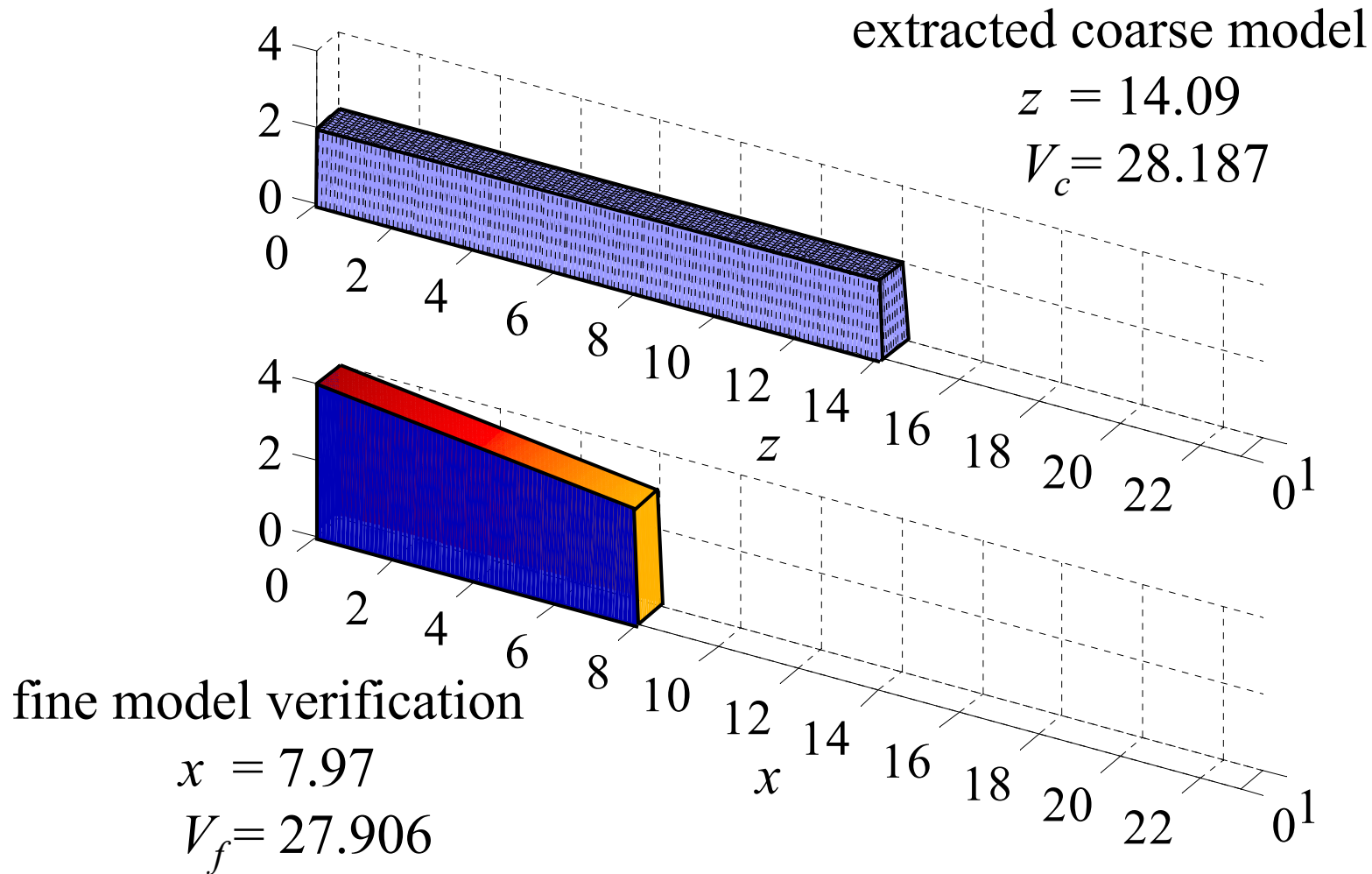


Iteration 6: Conventional ASM





Iteration 7: Conventional ASM





Wedge Cutting Problem with TRASM (*Bandler et al., 2002*)

$$\text{Step 1} \quad x^{(1)} = z^* = 14, B^{(1)} = 1, \delta^{(1)} = 2$$

$$\text{Step 2} \quad V_f(x^{(1)}) = 4x^{(1)} - \frac{(x^{(1)})^2}{16} = 43.75 \xrightarrow{\text{PE}} z^{(1)} = 21.875$$

$$\text{Step 3} \quad f^{(1)} = 21.875 - 14 = 7.875$$

$$\text{Step 4} \quad h^{(1)} = \arg \min_h \|7.875 + 1 \cdot h\| \text{ subject to } \|h\| < 2$$

$$h^{(1)} = -2$$

$$\text{Step 5} \quad x^{(2)} = 14 + (-2) = 12$$



Wedge Cutting Problem with TRASM (*Bandler et al., 2002*)

Step 6 $V_f(x^{(2)}) = 39$

Step 7 PE $z^{(2)} = 19.5$

Step 8 $f^{(2)} = 19.5 - 14 = 5.5$

Step 9 $\rho^{(1)} = \frac{7.875 - 5.5}{7.875 - (7.875 + 1(-2))} = 1.18$

Step 10 adjust trust region size $\delta^{(2)} = 2\delta^{(1)} = 4$

Step 11 $J_f = 4 - x^{(2)} / 8, J_c = 2 \rightarrow B^{(2)} = \frac{J_f}{J_c} = 1.25$



Wedge Cutting Problem with TRASM (*Bandler et al., 2002*)

Step 4b $h^{(2)} = \arg \min_h \|5.5 + 1.25h\|$ subject to $\|h\| < 4$

$$h^{(2)} = -4$$

Step 5b set $x^{(3)} = 12 + (-4) = 8$

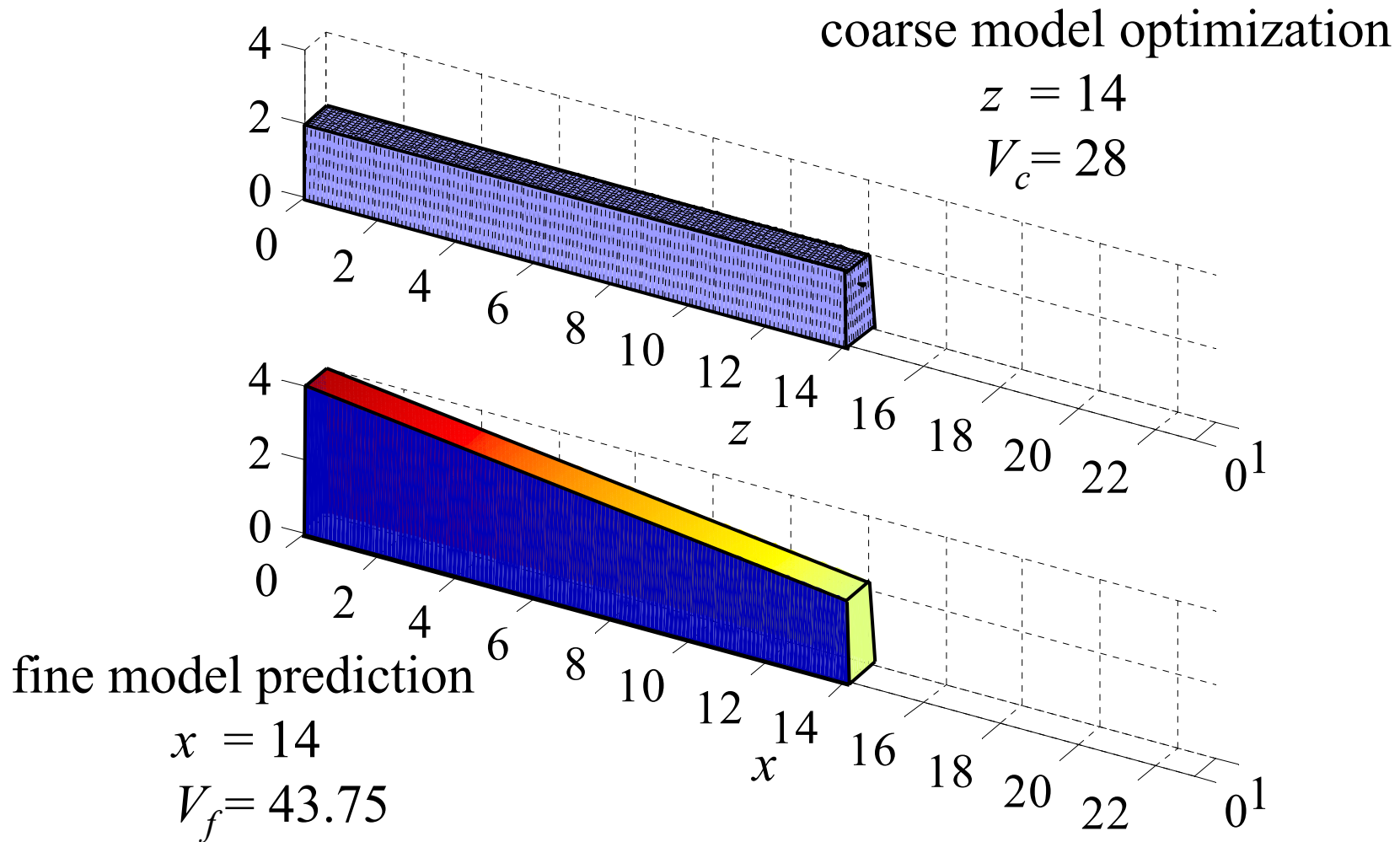
Step 6b $V_f(x^{(3)}) = 28$

Step 7b PE $z^{(3)} = 14$

Step 8b $f^{(3)} = 14 - 14 = 0$ stop the algorithm

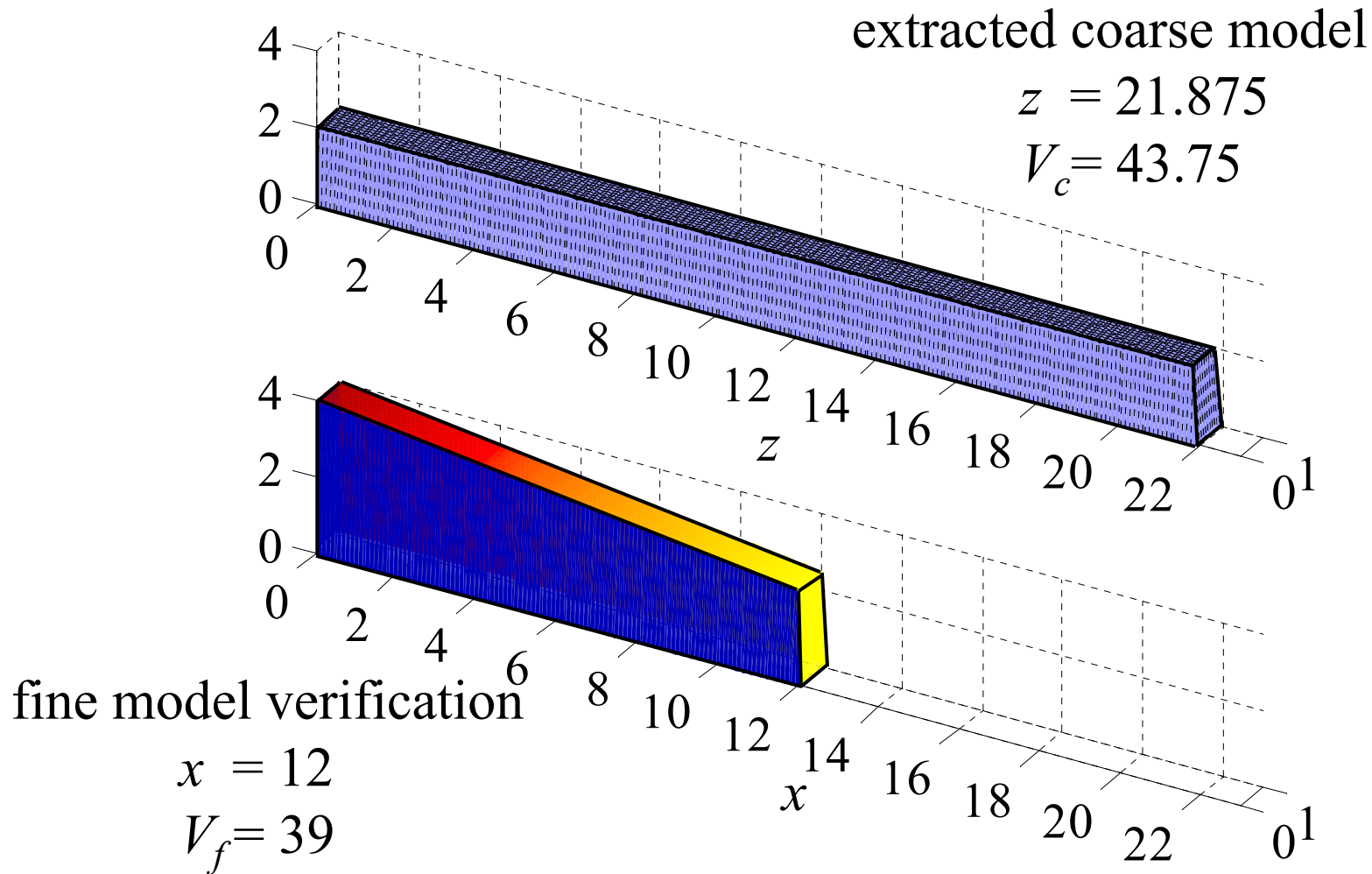


Initial Step: TRASM



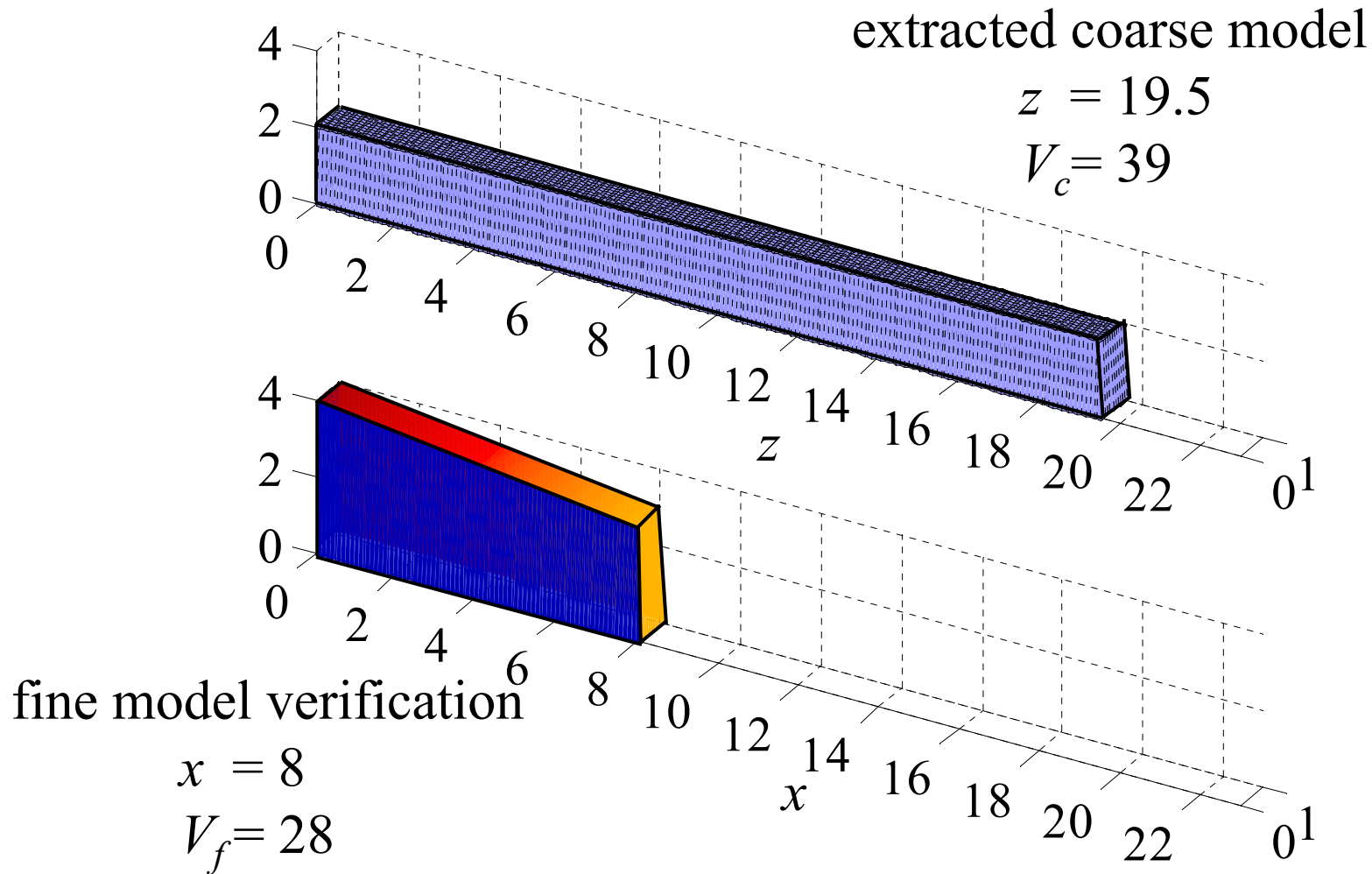


Iteration 1: TRASM





Iteration 2: TRASM





Change of Initial Trust Region Size

$\delta^{(1)}$	x^*	V_f	number of iterations
1	7.99905	27.99715	4
2	8	28	2
3	7.99905	27.99715	3
4	7.99983	27.99948	3



Conclusions

when the coarse model is severely different from the fine model
ASM may diverge or oscillate around the optimal solution

TRASM allows us to control step size and hence add damping to the
system to enhance stability and reduce the oscillations around the
solution