CAS 741 (Development of Scientific Computing Software)

Winter 2025

Mixed-Precision Iterative Solver

Xunzhou (Joe) Ye

Faculty of Engineering, McMaster University

January 27, 2025

Linear Solver: the Good Old $\mathbf{Ax} = \mathbf{b}$ Problem

- Direct method: Gaussian elimination

$$\mathbf{Ax} = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 11 \\ 3 \\ 3 \end{bmatrix} = \mathbf{b}$$

$$\mathbf{A}|\mathbf{b} = \left[\begin{array}{ccc|c} 1 & -1 & 3 & 11 \\ 1 & 1 & 0 & 3 \\ 3 & -2 & 1 & 3 \end{array}\right] \begin{array}{cc} \times 1 & \times 3 \\ \downarrow & \\ & \downarrow \end{array}$$

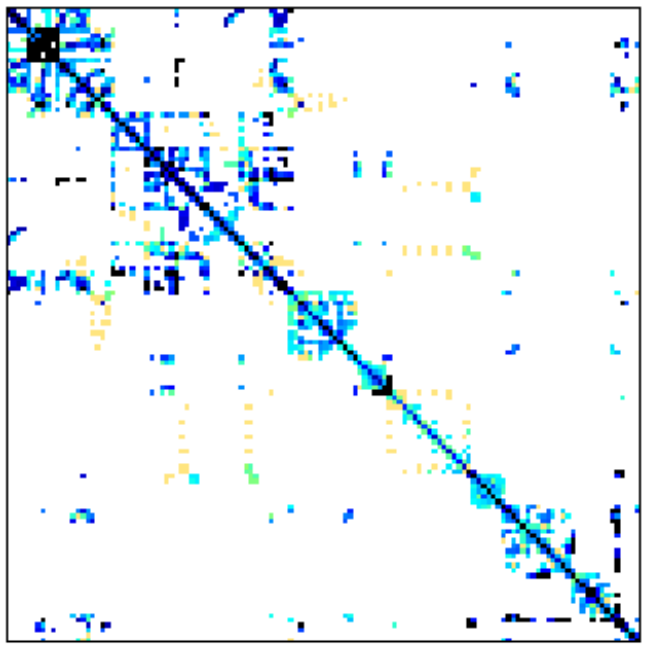
$$\mathbf{A}|\mathbf{b} \leftarrow \left[\begin{array}{ccc|c} 1 & -1 & 3 & 11 \\ 0 & 2 & -3 & -8 \\ 0 & 1 & -8 & -30 \end{array}\right]$$

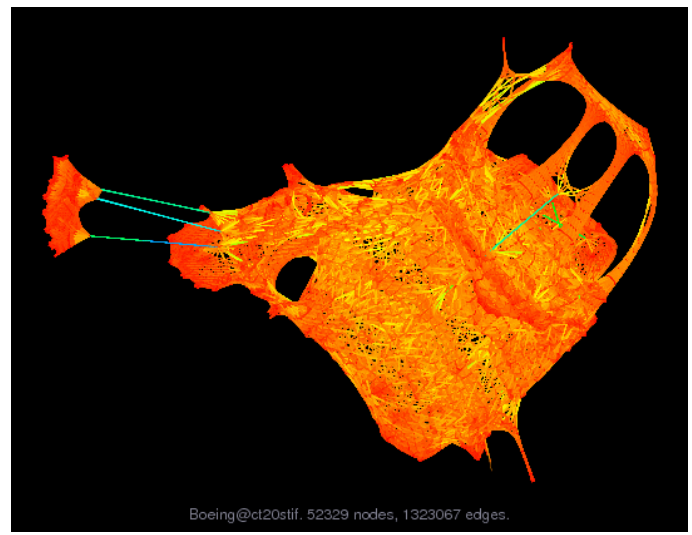
Think Big, Think Sparse

- In structural simulations (e.g., finite element analysis), structures are modeled based on real-world physics, where forces and constraints are often localized.
- Each element or node in a structure is typically connected only to a few nearby elements or nodes.
- System of equations in structural simulations is assembled from local contributions.
- In contrast, in neural networks, dense weight matrices are used in fully connected layers, particularly in large networks like transformers or dense layers after convolutional layers.

Sparse Matrix Example: Boeing/ct20stif

- $52\,329 \times 52\,329$
- $2\,600\,295$ non-zeros
- ≈ 33 MB in file size

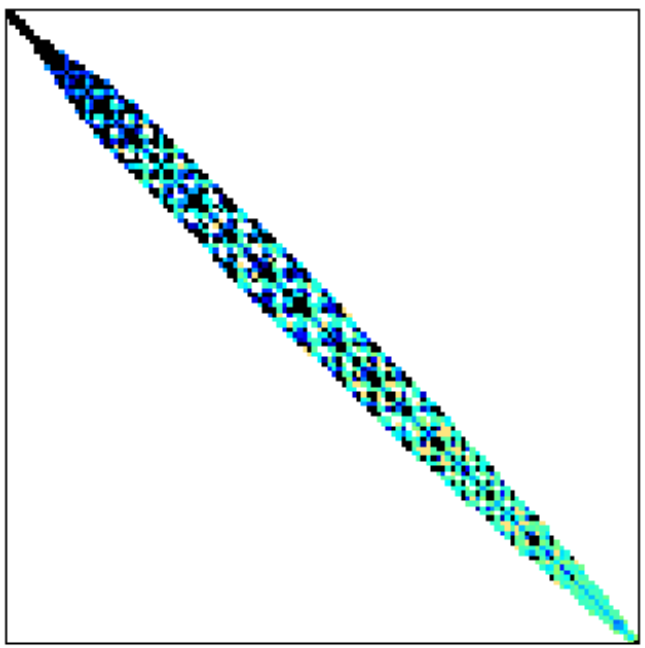




Boeing/ct20stif: CT20 Engine Block – Stiffness matrix

Sparse Matrix Example: Janna/Serena

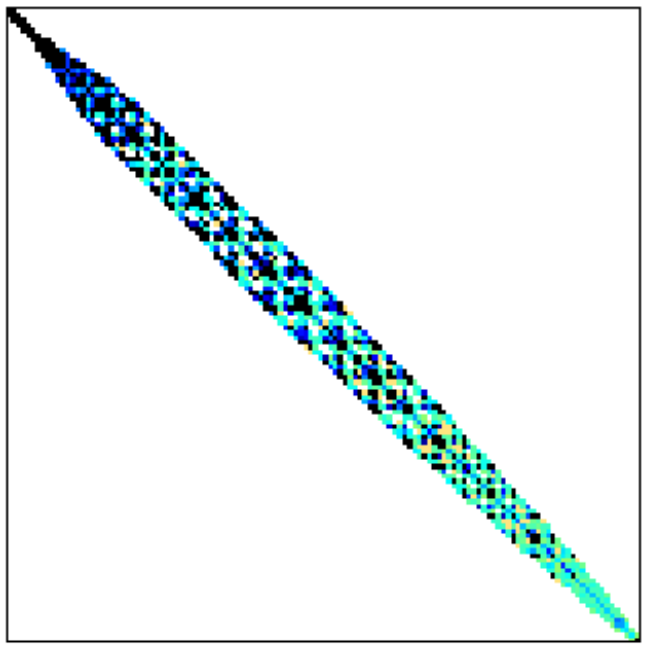
- $1\,391\,349 \times 1\,391\,349$
- $64\,131\,971$ non-zeros
- ≈ 847 MB in file size

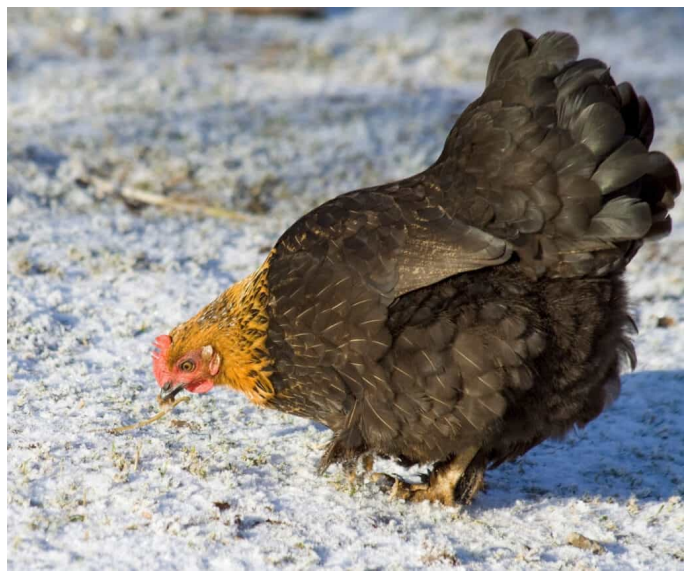


Janna/Serena: gas resevoir simulation for CO_2 sequestration

The Problem with Gaussian Elimination

- Store the whole matrix
- Memory bounded
- Sparsity preservation





Janna/Serena: gas resevoir simulation for CO_2 sequestration