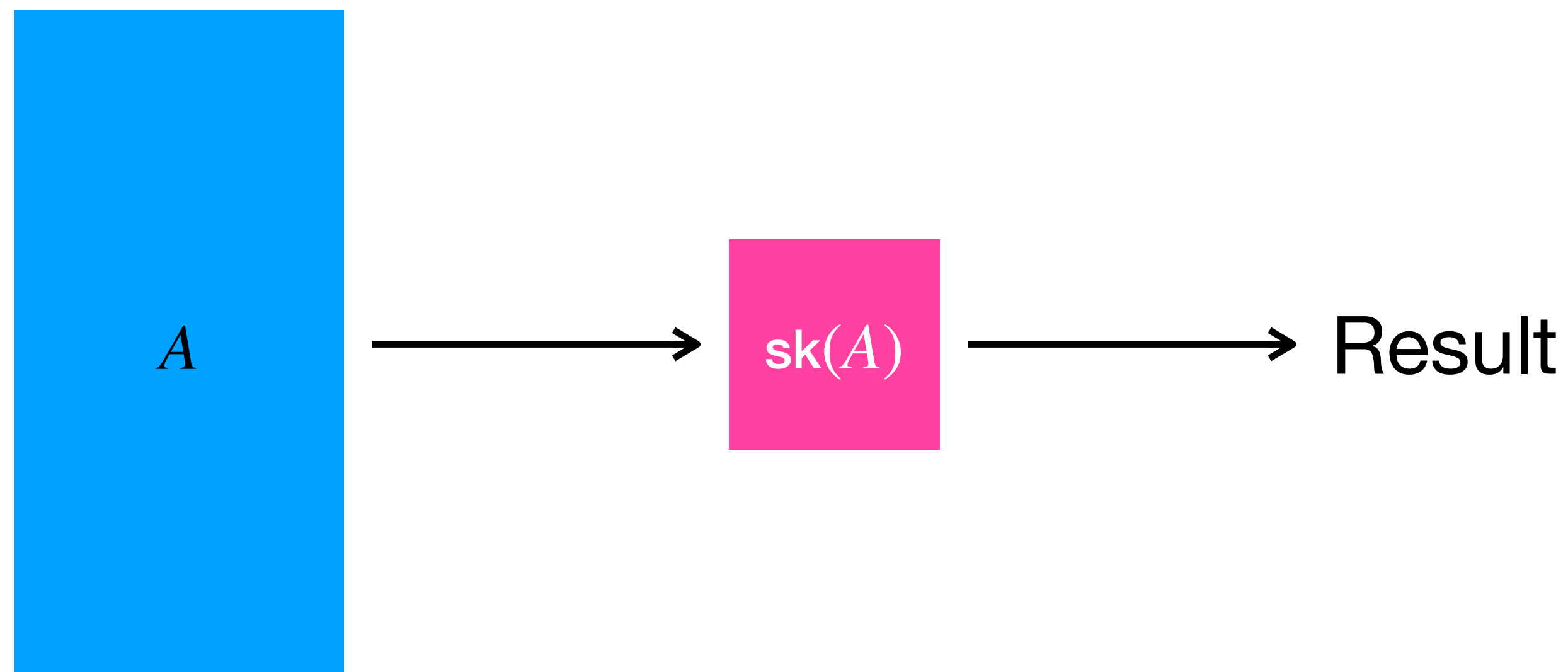


# Sketching

- A broad set of techniques developed over last 20 years to handle large datasets
- Shrink the large dataset into a small dataset
- Efficiently perform computation on the small dataset



# Linear Sketching

- Apply a randomized linear transform
  - CountSketch [Clarkson, Woodruff '13], OSNAP [Nelson, Nguyen '13], SRHT [Tropp '10], ...

$$\begin{array}{ccccc} \text{Yellow Box } S & \times & \text{Blue Box } A & = & \text{Pink Box } \text{sk}(A) \end{array}$$

**Sketching Matrix**

- A good sketching matrix  $S$  consolidates information into the “sketch”