You can check others' work more quickly than doing it yourself

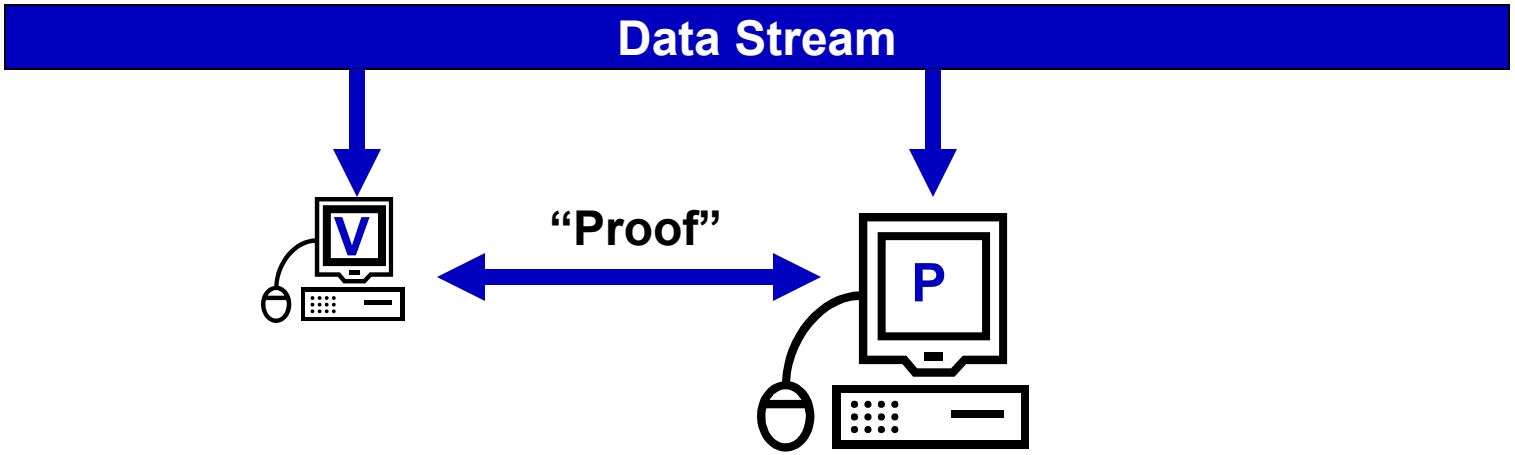
Graham Cormode

G.Cormode@warwick.ac.uk

Chris Hickey (Warwick)

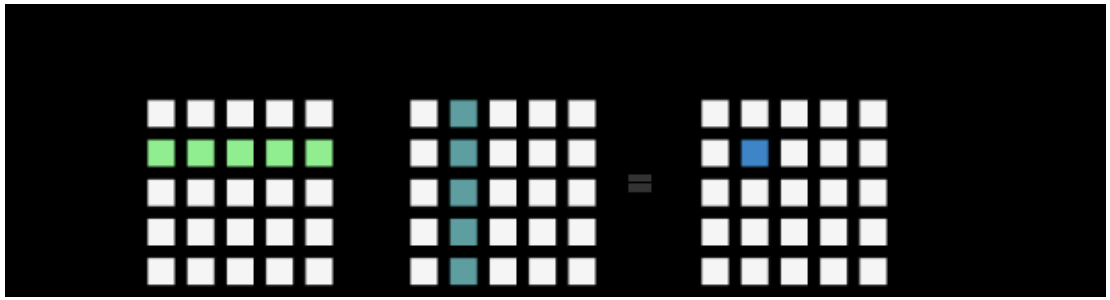
Checksums for Computation

- Checksums on data are used to ensure correct transmission
  - If the checksums agree then (almost certainly) the data matches
- What could we do if we had checksums for computation?
  - Check that an algorithm has provided the expected answer
  - Check that a hardware accelerator has not made a mistake
  - Check that the cloud has not tried to cheat us!



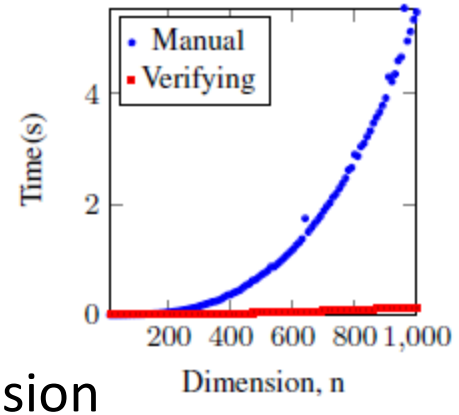
Checksums for Computation Do Exist!

- There are techniques to quickly check arbitrary computation, but:
  - They need the computation to be written as an arithmetic circuit
  - They can be quite slow and require a lot of rounds of interaction
- There are faster techniques to check specific computations
- Example: Matrix multiplication
  - Given n x n matrices A and B, compute checksums for A, B and AB
  - Computing h(A), h(B) takes time linear in number of nonzero entries
  - Computing h(AB) from h(A), h(B) takes time O(n)
  - Compared to computing AB, takes time ~$O(n^{2.8})$



Verifying Data Analysis

- Recent work [C, Hickey 18] shows how to apply this model to:
  - (Least Squares) Regression
  - Principal Component Analysis
  - Linear Discriminant Analysis Classifier



- Technical challenges:
  - Have to tolerate rounding errors to finite precision
  - Need to verify that vectors are approximate eigenvectors
  - Build primitives to check matrix inversion, matrix decomposition

Challenges to Data Engineering

- Incorporate checksums for computation into real systems
  - Outsourced computations return mathematical proof of correctness?
  - Internal checks within systems?
- Generalize these techniques for a wider range of problems
  - Check Machine learning models are (approximately) optimal
  - Verify result of database queries (see [vSQL 2017])
- Optimize, extend and simplify
  - When can proof be provided as a byproduct of computation?
  - Allow efficient composition of computations?
  - Other models: interactive proofs, multiple provers?

THE UNIVERSITY OF WARWICK

THE ROYAL SOCIETY

erc
European Research Council
Established by the European Commission