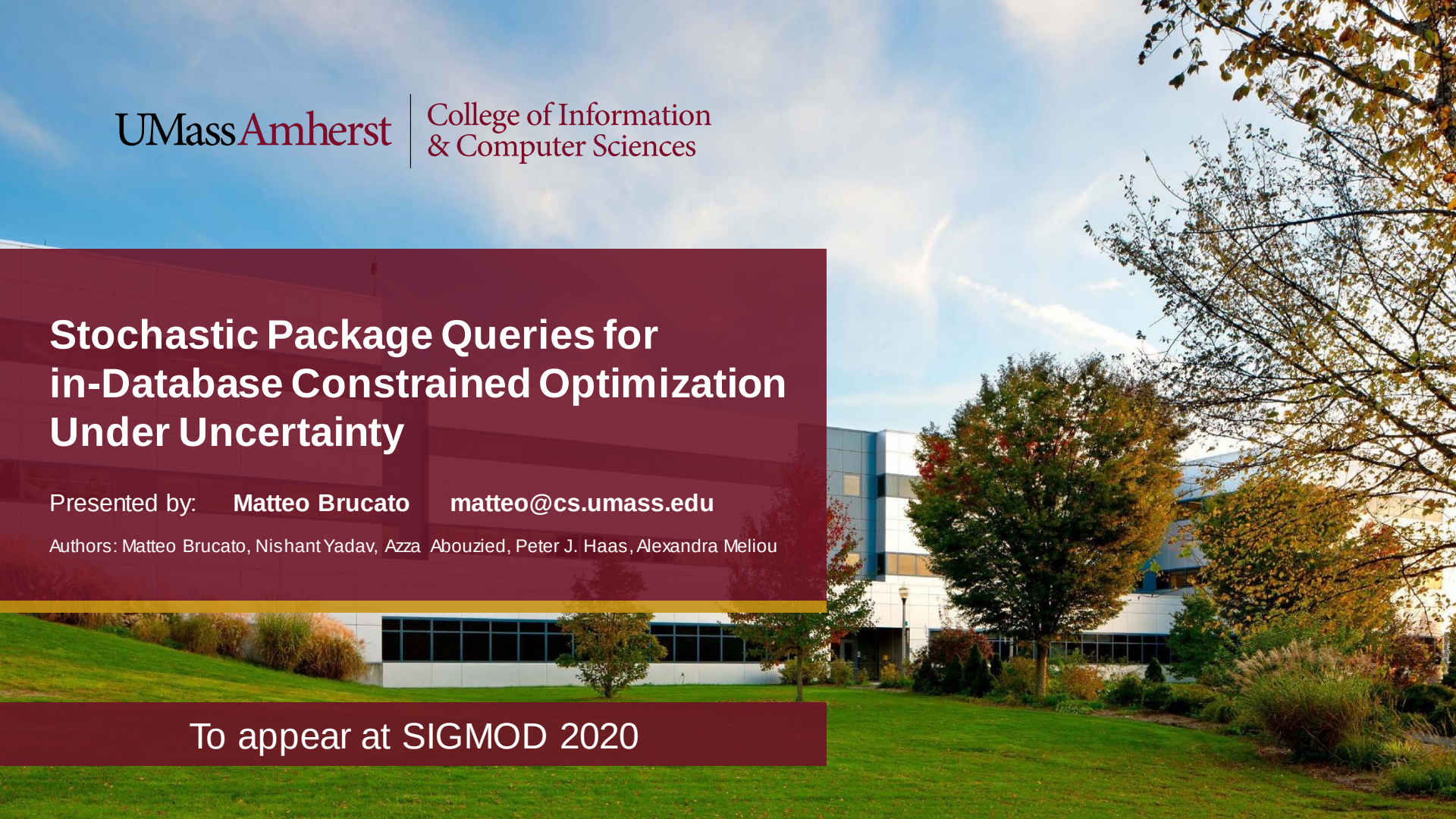

UMass Amherst | College of Information & Computer Sciences

Stochastic Package Queries for in-Database Constrained Optimization Under Uncertainty

Presented by: **Matteo Brucato** **matteo@cs.umass.edu**

Authors: Matteo Brucato, Nishant Yadav, Azza Abouzied, Peter J. Haas, Alexandra Meliou

To appear at SIGMOD 2020

Decision Making

- Example: **Portfolio of stocks (Finance)**
  - Which stocks to buy
  - When to sell them
  - Maximize revenue
  - Willing to risk 1% of investment, but with low probability (5%)
- Many domains:
  - Transportation
  - Travel industry
  - Robotics
  - Engineering
  - ...

Predictions

- Data used to make predictions is **unknown** today!
- Prediction models are widespread
  - Simulation models
  - Machine learning models
  - …
- Question:
  "How to **make use** of these predictions
  to generate **optimal decisions**?"

Stochastic Programming

- Generate many scenarios
- Put them together in one large problem
- Solve problem with existing solver (e.g., CPLEX, Gurobi, etc.)
- This apprach breaks down:
  - (1) Disconnect between data, software for predictions, and optimization software
  - (2) Solutions do not scale!
    "**Optimizer's curse**": *Random scenarios tend to make the solver "too optimistic"*
    The large the input dataset, the worst this is!

Our Solution: Summaries of Scenarios

- Don't rely on "luck of the draw"
- Craft "conservative" scenarios, called "**summaries**"
- Find the right balance between conservativeness and optimality

Results:

- Can generate feasible solutions!
- Orders of magnitude faster (when both methods can)
- Solutions are always of very good quality
- Provably $(1+\epsilon)$-approximate

UMass Amherst | College of Information & Computer Sciences

COMPUTING FOR THE COMMON GOOD

Matteo Brucato
matteo@cs.umass.edu