# Multiobjective Genetic Algorithms for Multiscaling Excited-State Direct Dynamics in Photochemistry

Kumara Sastry[1], D.D. Johnson[2], A. L. Thompson[3],
D. E. Goldberg[1], T. J. Martinez[3], J. Leiding[3], J. Owens[3]

[1]Illinois Genetic Algorithms Laboratory, Industrial and Enterprise Systems Engineering
[2]Materials Science and Engineering
[3]Chemistry and Beckman Institute
University of Illinois at Urbana-Champaign


Supported by AFOSR F49620-03-1-0129, NSF/DMR at MCC DMR-03-76550

# Chemical Reaction Dynamics Over Multiple Timescales


Solve Schrödinger's equations
Accurate but slow (hours-days)

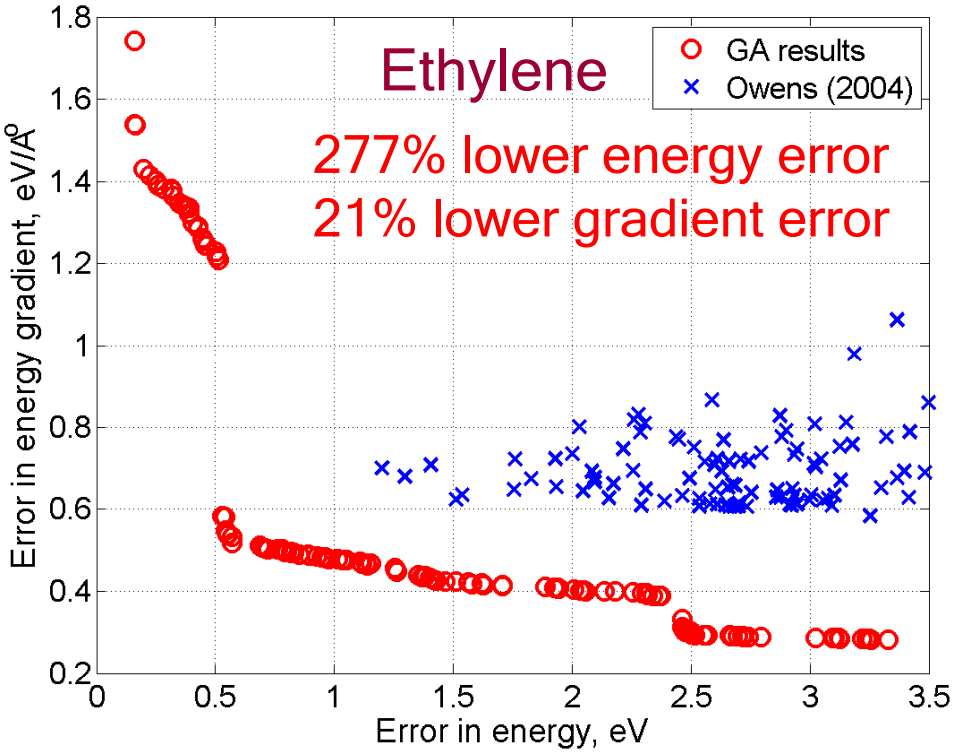
Solve approximate Schrödinger's equations.
Fast (secs-mins). Accuracy depends on semiempirical potentials

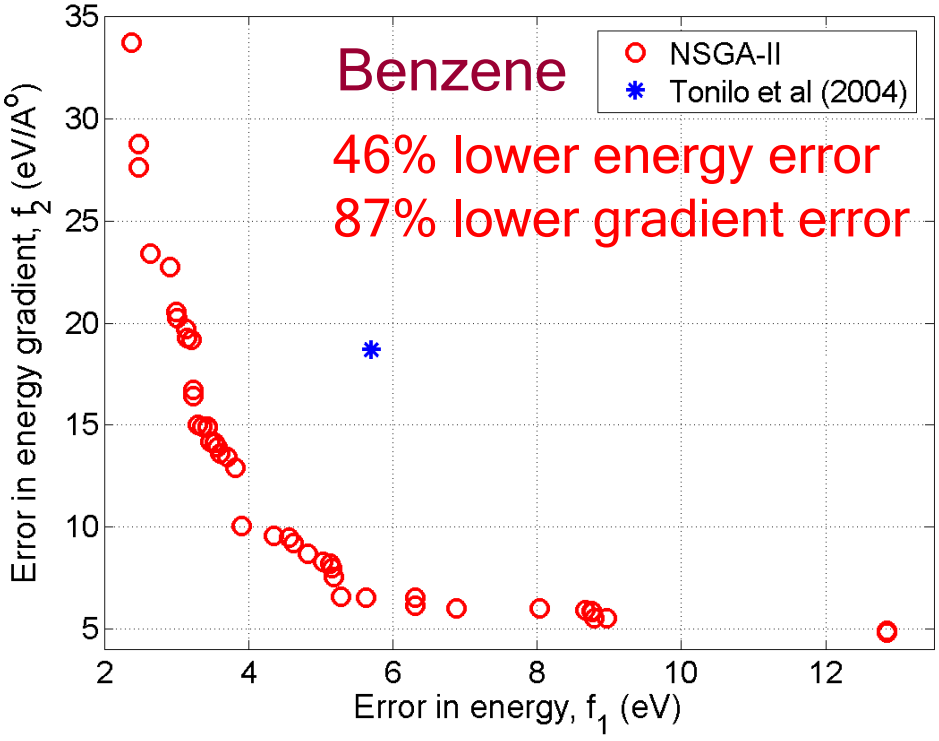
- Fitting/Tuning semiempirical potentials is non-trivial
- Energy & shape of energy landscape matter
  - Both around ground states and excited states
- Two objectives at the bare minimum
  - Minimizing errors in energy and energy gradient

# Why Does This Matter?

- Multiscaling speeds all modeling of physical problems:
  - Solids, fluids, thermodynamics, kinetics, etc.,
  - Example: GP used for multi-timescaling Cu-Co alloy kinetics [Sastry, et al (2006), *Physical Review B*]
- Here we use MOGA to enable fast and accurate modeling
  - Retain *ab initio* accuracy, but exponentially faster
- Enabling technology: Science and Synthesis
  - Fast, accurate models permit larger quantity of scientific studies
  - Fast, accurate models permit synthesis via repeated analysis
- This study potentially enables:
  - Biophysical basis of vision
  - Biophysical basis of photosynthesis
  - Protein folding and drug design
  - Rapid design of functional materials (zeolites, LCDs, etc.,)

# GA Produces Physical and Accurate Potentials (PES)

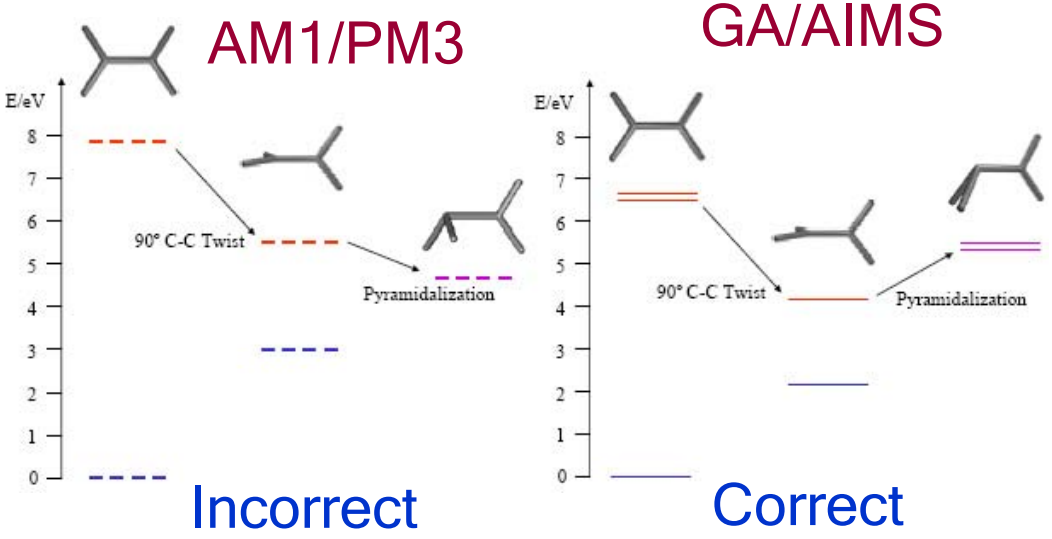


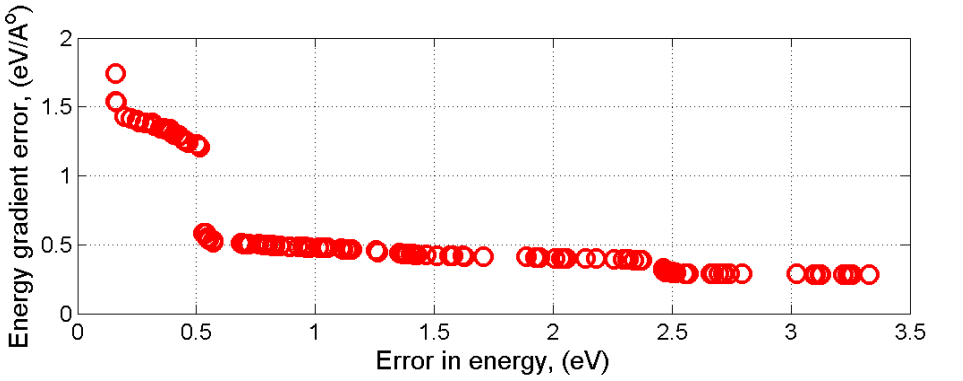



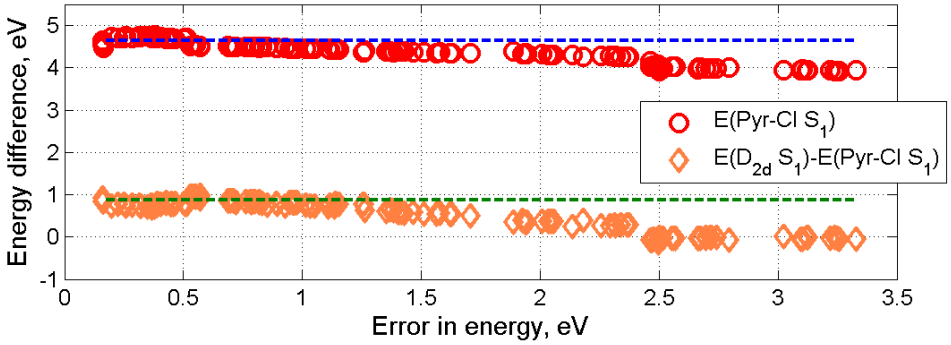

- Significant reduction in errors
- Globally accurate potential energy surfaces
  - Resulting in physical reaction dynamics
- Evidence of transferability: "Holy Grail" in molecular dynamics

# GA Optimized SE Potentials are Physical

- Dynamics agree with *ab initio* results
- Validates expermental results for both benzene & ethylene
- Example: *cis-trans* isomerization in ethylene
  - AM1, PM3, and other parameter sets yield wrong energetics
  - GA yields results consistent with AIMS and experiments



# Human Competitive Claims: Criteria B, C, D, E

- Criterion B: The result is equal to or better than a result that was accepted as a new scientific result at the time when it was published in a peer-reviewed scientific journal.
- Criterion C: The result is equal to or better than a result that was placed into a database or archive of results maintained by an internationally recognized panel of scientific experts.
- Criterion D: The result is publishable in its own right as a new scientific result 3/4 independent of the fact that the result was mechanically created.
- Criterion E: The result is equal to or better than the most recent human-created solution to a long-standing problem for which there has been a succession of increasingly better human-created solutions.

# Criterion B: Better Than Result Accepted As A New Scientific Result

- Current best published results
  - *Journal of American Chemical Society* (2nd), *Journal of Chemical Physics* (3rd), *Journal of Physical Chemistry* (4th), and *Chemical Physics Letters* (8th)
  - 13,417+ citations of top 10 papers
- Multiobjective GA results
  - Parameter sets with up to 277% lower energy error and 87% lower gradient error
    - Semiempirical potentials with results well beyond previous attempts, or expectation of human experts
  - Efficient and yields multiple potentials with accurate PES
    - Up to 1000 times faster than current methods
  - Evidence of transferability
    - Enables accurate simulations of photochemistry in complex environments without the need for complete reoptimization.

Sources: Most frequently referenced in Chemical Abstracts. Web of Science

# Criterion C: Better Than Result Placed Into a Database/Archive of Results.

- Standard semiempirical potentials:
  - AM1 (16,031+ cit.), INDO(4,583+ cit.), PM3 (4,416+ cit.), MNDO (1,919+ cit.), CNDO (1,120+ cit.)
  - Used in commercial software (MOLCAS, MOPAC, MOLPRO)
  - Globally inaccurate PES yields wrong chemistry
  - No evidence transferability, nor any physical insight
- Multiobjective GA results:
  - Globally accurate PES yields accurate chemistry
    - Never been obtained by any previous attempt at optimizing the semiempirical forms of MNDO, AM1, and PM3.
  - Evidence of transferability
    - "Holy Grail" for two decades in chemistry & materials science.
  - Physical insight from Pareto analysis using rBOA and symbolic regression via GP.

# Criterion D: GA Results are Publishable

- Paper at GECCO in Real World Applications track
  - Nominated for best paper award
- Preparing journal version highlighting new chemistry results the methodology revealed.
  - Target Journal: *Journal of Chemical Physics*
- Observed transferability is a very important to chemists
  - Enables accurate simulations without the need for complete reoptimization
- Pareto analysis reveals interactions between parameters
  - Semiempirical potentials have physical interpretability
  - Gave new insight into multiplicity of models and why they should exist.

# Criterion E: GA Wins MacArthur "Genius" Award

- Human created solutions:
  - Todd Martinez is the recipient of the MacArthur "Genius" award for his work on "combining effective strategies for computing the quantum mechanical properties of complex molecules with a deep intuition for their underlying chemical behavior"
- Multiobjective GA results:
  - Parameters sets that are up to 277% lower energy error and 87% lower gradient error
  - Interpretable semiempirical potentials
  - Enables orders of magnitude ($10^2$-$10^5$) increase in simulation time even for simple molecules
  - Orders of magnitude (10-$10^3$) faster than the current methodology for developing semiempirical potentials

# Why This is the "Best" Among Other Humies Submissions?

- Broadly applicable in chemistry and materials science
  - Analogous applicability when multiscaling phenomena is involved: Solids, fluids, thermodynamics, kinetics, etc.
- Facilitates fast and accurate materials modeling
  - Alloys: Kinetics simulations with *ab initio* accuracy. $10^4$-$10^7$ times faster than current methods.
  - Chemistry: Reaction-dynamics simulations with *ab initio* accuracy.$10^2$-$10^5$ times faster than current methods.
- Lead potentially to new drugs, new materials, fundamental understanding of complex chemical phenomena
  - Science: Biophysical basis of vision, and photosynthesis
  - Synthesis: Pharmaceuticals, functional materials