An agent approach to finding local optima and other ideas about distributed resolution of optimization problems

D. Villanueva, *R. Le Riche*, *G. Picard*

University of Florida
Institut Henri Fayol, EMSE
CNRS LIMOS

12 April 2013

One presentation, three stories...

- Collaboration between an applied mathematician and a computer scientist
- Collaborative decision: an analytical model for a wide-ranging topic
- An agent-based algorithm for locating local optima

One presentation, three layers...

Context

ID4CS & MDO

Future works

Dialog

Incomprehension

Enrichment

Scientific work

Algorithm

Overall context: ID4CS project (1)

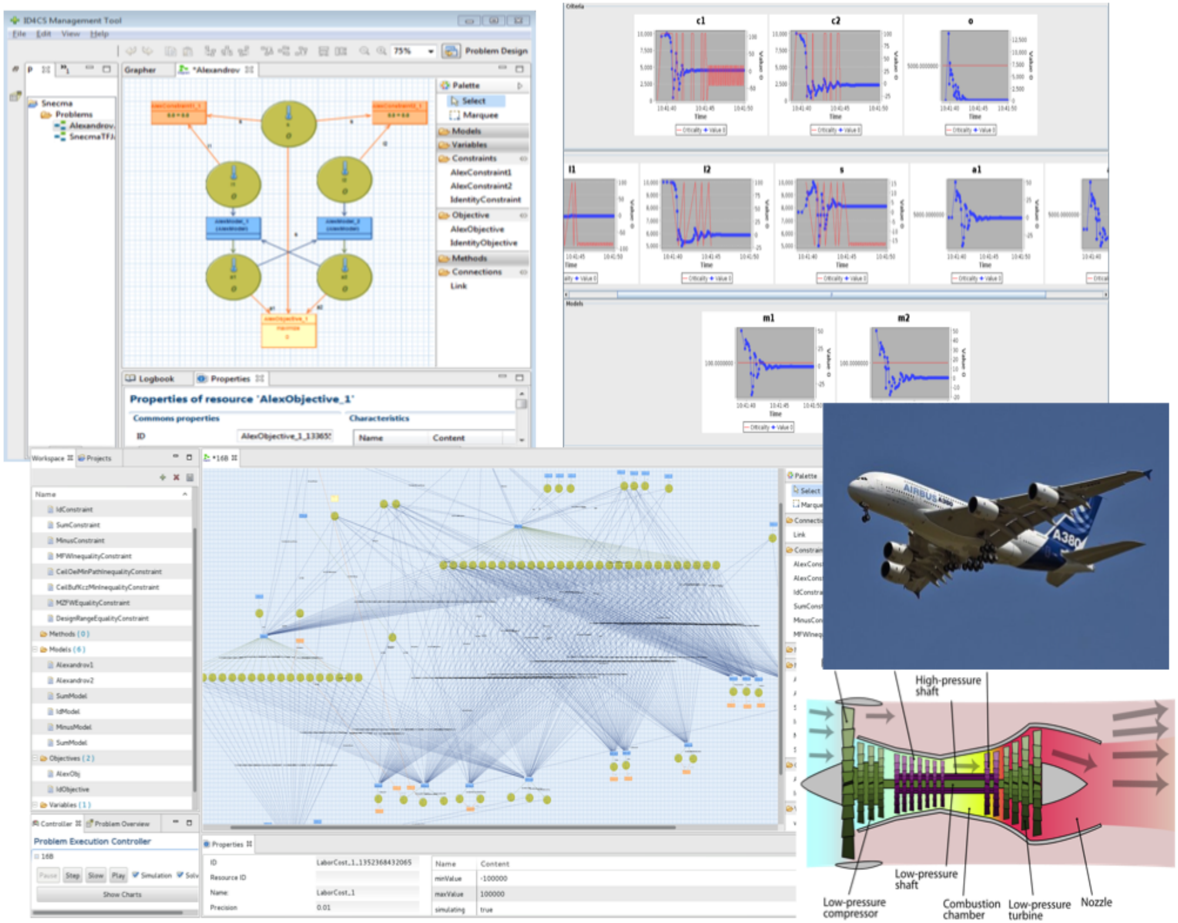
- ANR Project (2010-2013)
- Integrative Design for Complex Systems

  e.g. aircraft, motors
- Pluri-disciplinary consortium

http://www.irit.fr/id4cs

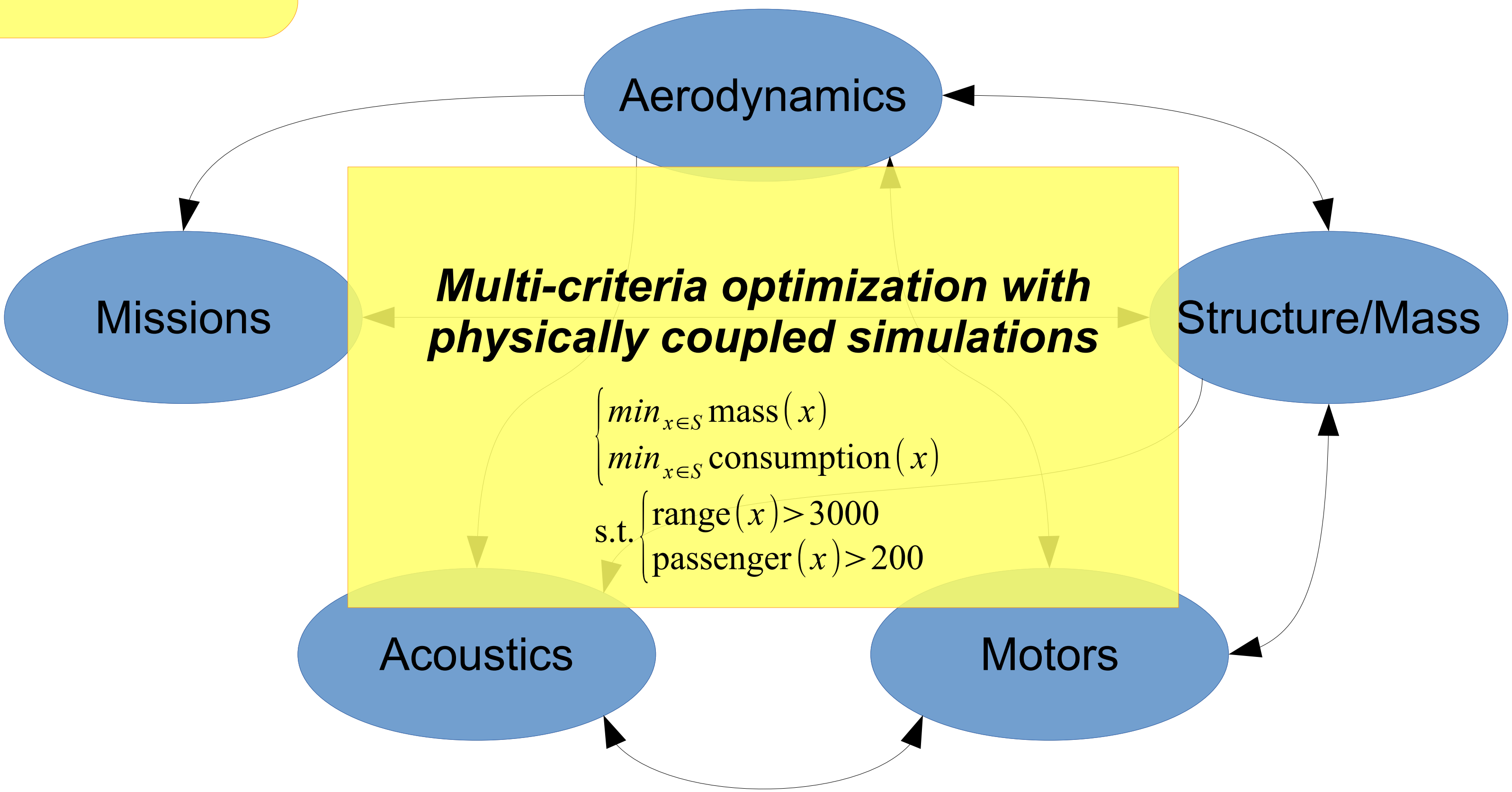
Overall context: ID4CS project (2)

- Several research directions:
  - Multi-disciplinary
  - Multi-fidelity
  - Multi-criteria
  - Multi-*
  - Uncertainties
- Integrative approach
  - multi-agent platform → Fine-grained *a priori*



Multi-disciplinary optimization (centralized – preliminary design)

Business decision
e.g. 200 passengers, transatlantic

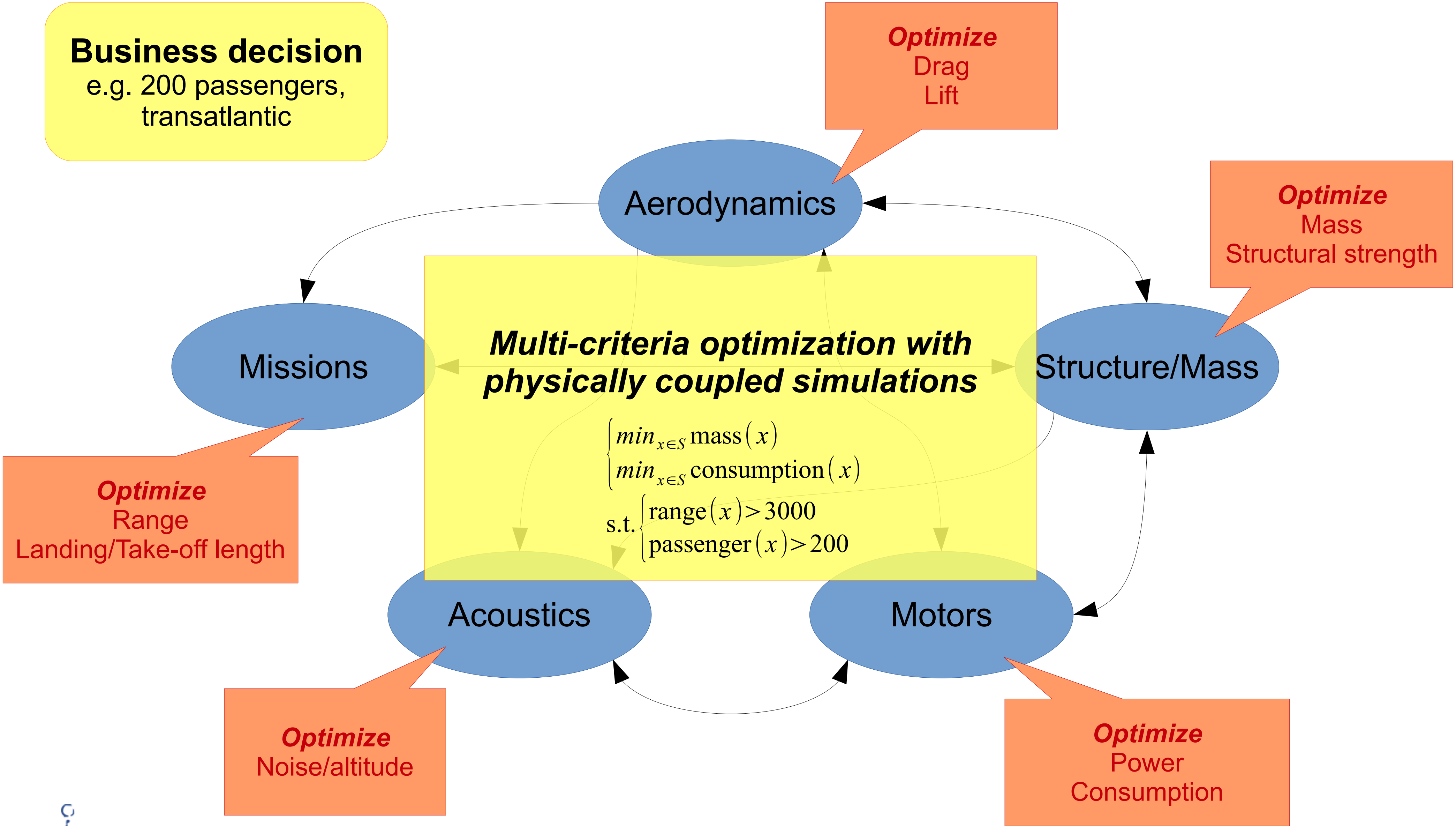


Multi-criteria optimization with physically coupled simulations

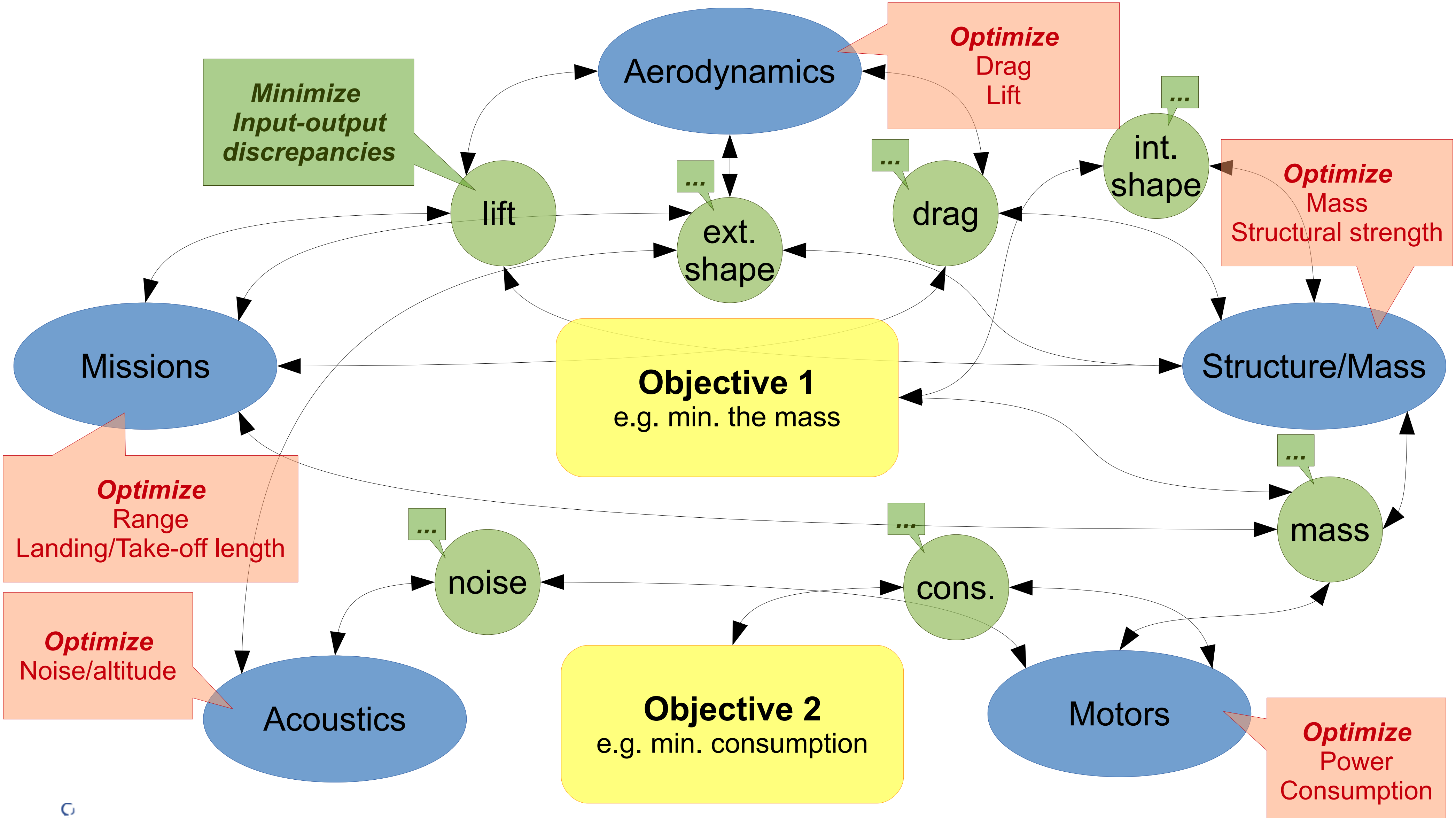
$$\begin{cases} min_{x \in S} \text{ mass}(x) \\ min_{x \in S} \text{ consumption}(x) \end{cases}$$

$$\text{s.t.} \begin{cases} \text{range}(x) > 3000 \\ \text{passenger}(x) > 200 \end{cases}$$

Multi-disciplinary optimization (distributed – consolidation)



Multi-disciplinary optimization (decentralized – finer agentification)



Dialog between the Computer Scientist (CS) and the Applied Mathematician (AM)

- AM (skeptical about the decomposition, particularly at low granularity): “What are agents ?”

Dialog between the Computer Scientist (CS) and the Applied Mathematician (AM)

- AM (skeptical about the decomposition, particularly at low granularity): "What are agents ?"
- CS : "They are a decomposition of a problem into autonomous tasks (agents) that collectively, through interaction mechanisms and protocols, solve the initial problem."
- AM (dubious, partial interest) : "hum ..."

Dialog between the Computer Scientist (CS) and the Applied Mathematician (AM)

- CS (somewhat skeptical about the application): "What is special about the optimization of such objects ?"

Dialog between the Computer Scientist (CS) and the Applied Mathematician (AM)

- CS : "What is special about the optimization of such objects ?
- AM : "An important issue is that realistic simulations are – and will always be – numerically costly. For the optimization, we use metamodels (statistical models of other numerical models)"

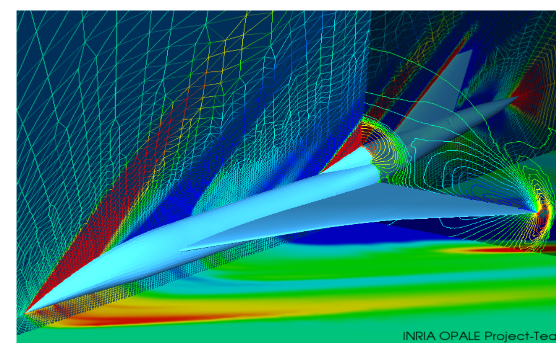
Goal : $min_{x \in S} f(x)$

OPTIMIZER ⟷ metamodel $\hat{f}(x)$ ⟷ database ⟷ physical simulation

database

variables	perfo
x^1	$f(x^1)$
...	...
x^m	$f(x^m)$

physical simulation



- CS (dubious about centralization, partial interest) : "hum ..."

From pluri- to inter-disciplinarity: will / time and pragmatism

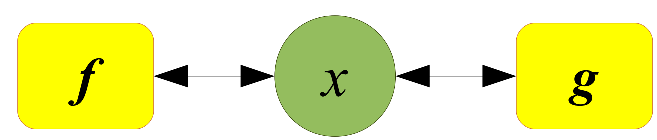
- At this point we have 1 multi-* problem and 2 points of view (agents vs. optimization)
- *Pragmatism*: A PhD is hired for the project (Diane Villanueva) → Need clear work directions
- *Enabler 1*: will / time. One hour meeting per week for a year
- *Enabler 2*: a joined PhD with the US and a student not trapped in formal disciplines (French CNU sections)

Research directions: how to agentify an optimization problem ?

$$\begin{cases} min_{x \in S} f(x) \\ g(x) \leq 0 \end{cases}$$

search space partition : synchronize *n* optimizers dividing work in S

variables and criteria decomposition



Research directions: how to agentify an optimization problem ?

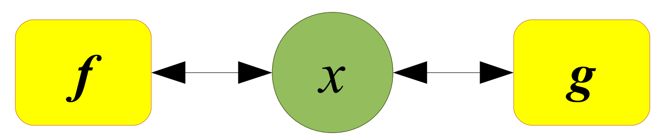
$$\begin{cases} min_{x \in S} f(x) \\ g(x) \leq 0 \end{cases}$$

main direction for us

search space partition : synchronize *n* optimizers dividing work in S

secondary direction

variables and criteria decomposition



Agent-based dynamic partitioning algorithm

1 subregion
+ 1 surrogate
+ 1 local constrained optimizer
+ 1 simulator

=

1 agent

search space *S*

Agents work in parallel to collectively solve the optimization problem :

$$\min_{x \in S \subset \mathbb{R}^n} f(x)$$
$$g(x) \leq 0$$

Agent coordination through :

- update of the partition
- agent creation
- agent deletion

(let's say 1 agent is affected to a set of computing nodes)

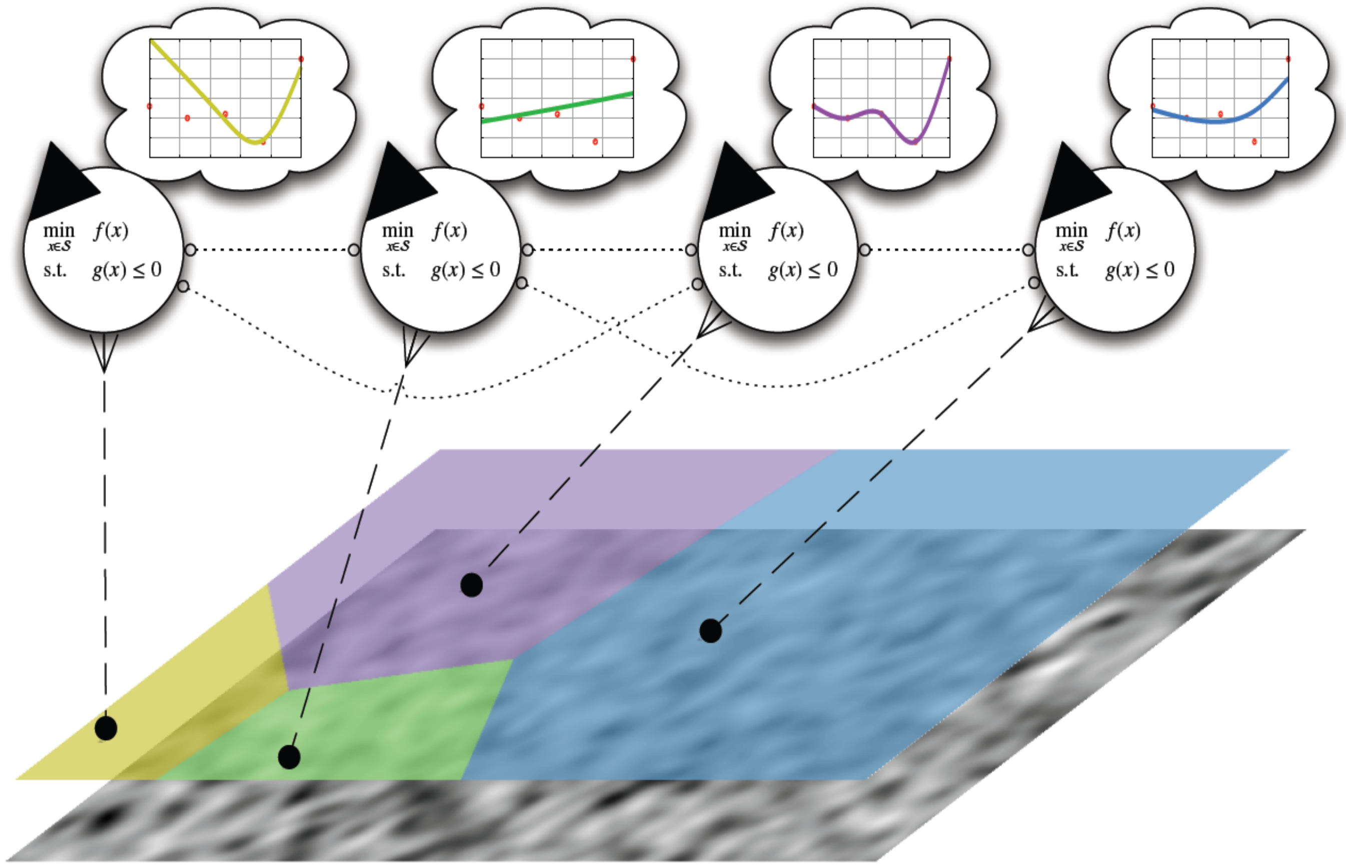
Agent-based dynamic partitioning algorithm: Goals

Solve a global optimization problem AND locate local optima
A method that can be used for expensive problems (thanks to the surrogates)

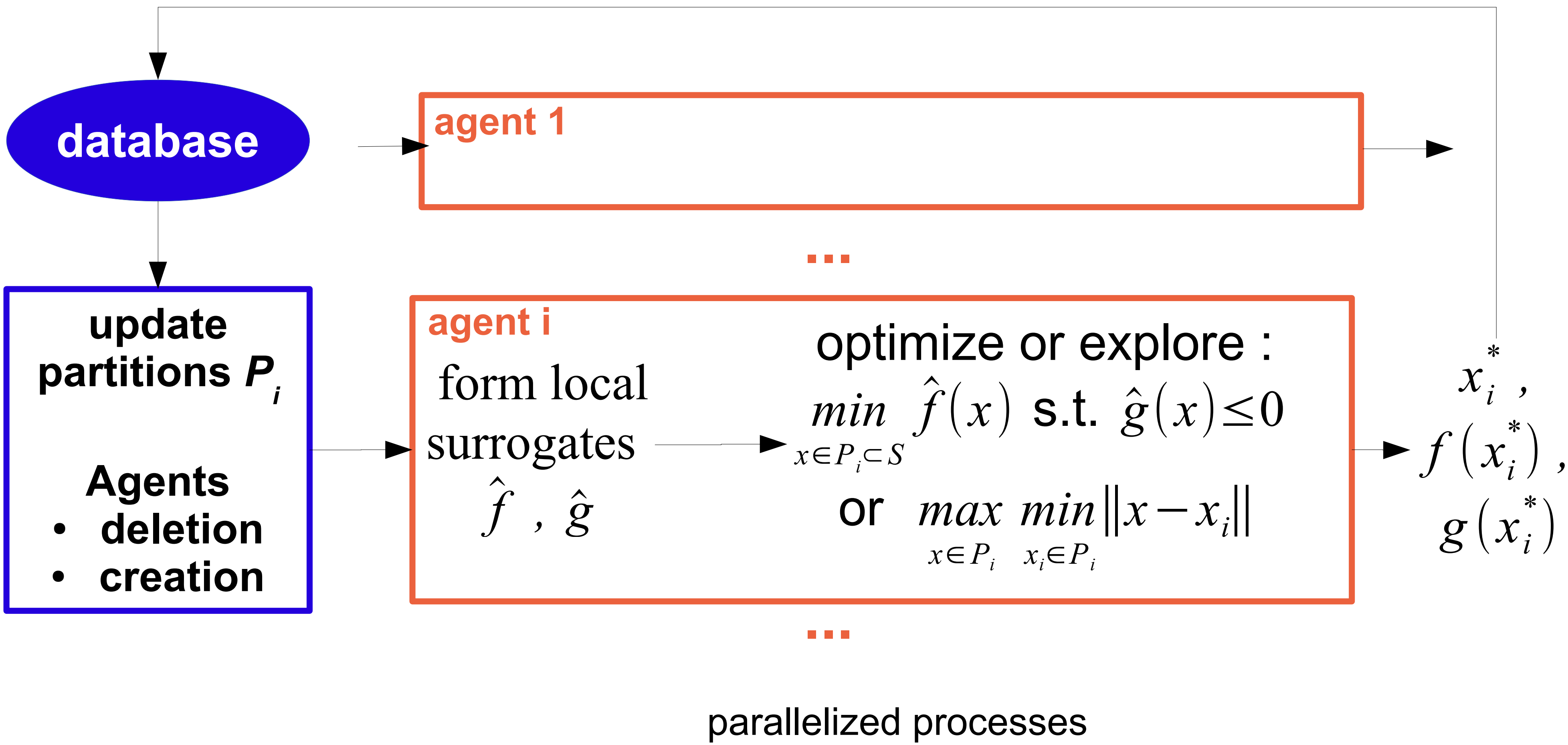
The search space partitioning allows:

1) to share the effort of finding local optima

2) to have surrogates defined locally (better for non stationary problems)



Agent-based dynamic partitioning algorithm: Global flow chart

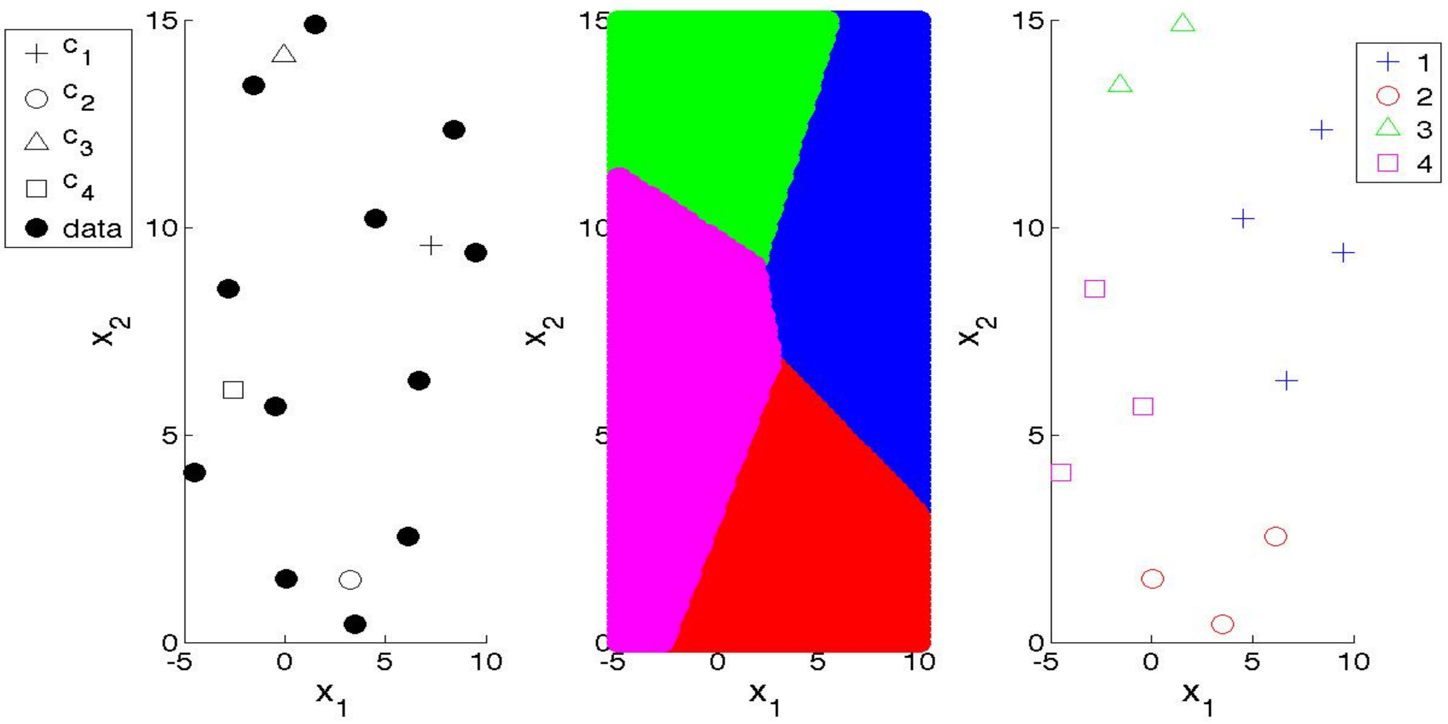


optimize : SQP.

surrogates : polynomial response surface (orders 1, 2 and 3), kriging (linear or quad. trend), chosen based on cross-validation error

Subregion definition

Subregions P_i are essentially defined by the centers c_i of the subregions: P_i is the set of points closer to c_i than to other centers. P_i are Voronoi cells



Dynamic partitioning

The partitioning is updated by moving the centers to the best point in their subregion:

current = current center
new = point added to P_i at the last iteration and not on boundary of P_i

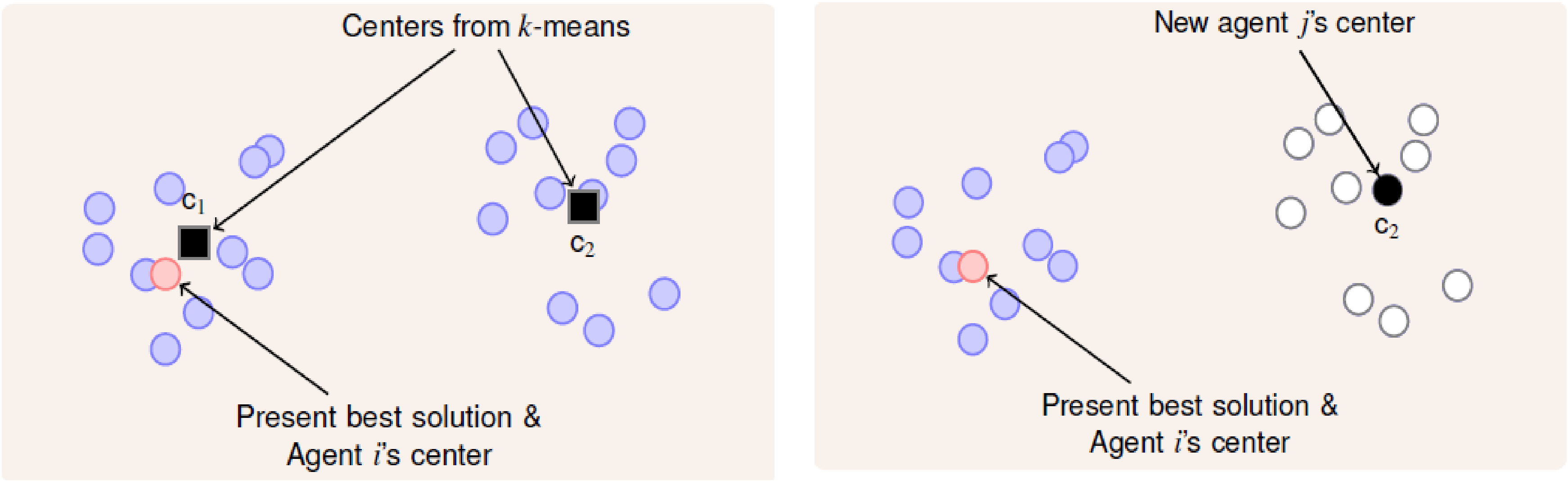
```
if current is infeasible then
   if new is less infeasible then move to new
elseif current is feasible then
   if new is feasible & has better f then move to new
end
```

Property : agents will stabilize at local optima

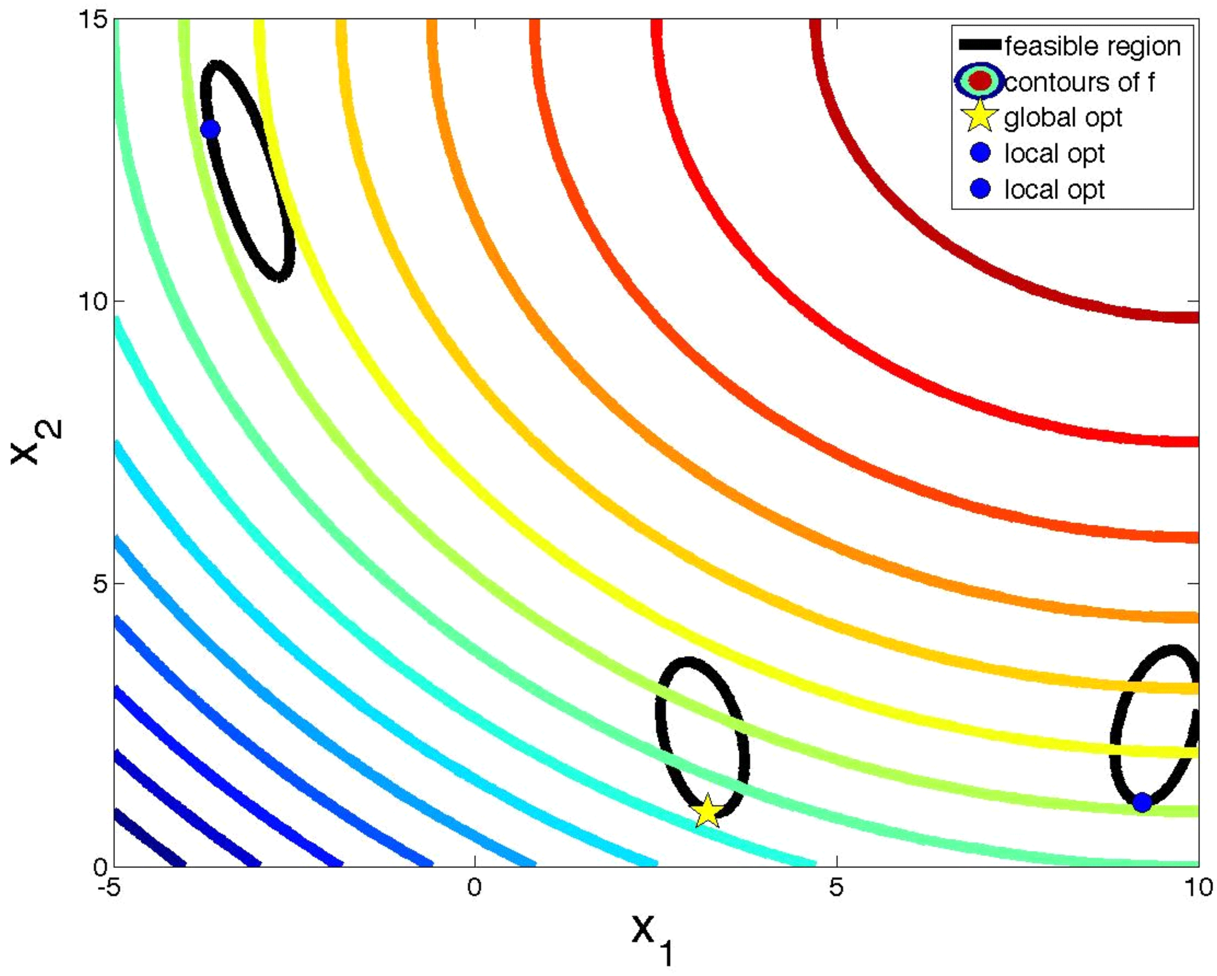
Agent deletion and creation

Deletion
If two agent centers are getting too close to each other, delete the worst

Creation
Principle 1: the existence of 2 clusters in a subregion is a sign of at least 2 basins of attraction → split the subregion by creating a new agent
Principle 2: when an agent is stagnant for 3 iterations → split the subregion by creating a new agent
Implementation : K-means + check on inter vs. intra class inertia + move centers at data points (farthest from existing centers)



Let's look at the behavior in 2D...



Let's look at the behavior in 2D...

Two Examples

- Examined two problems to study the success of this method
- Compared **multiple agents with partitioning** to a **single global agent** for an **equal number of expensive function evaluations**
  - Single Global Agent: Single surrogate acting over the entire design space
    - Exploration due to points being too near to each other
- Dynamics
  - Minimum of 1 region
  - Initially 1 region

Modified Hartman 6: Problem Description

- Hartman 6 is a popular benchmark test problem for surrogate-based global optimization algorithms
  - 6 dimensional multi-modal problem

$$\underset{x}{\text{minimize}} \quad f_{hart}(x) = -\sum_{i=1}^{q} a_i exp\left(-\sum_{j=1}^{m} b_{ij}(x_j - d_{ij})^2\right)$$

$$\text{subject to} \quad 0 \le x_j \le 1, j = 1, 2, \ldots, m = 6$$

- Modified Hartman 6 includes two Gaussian holes “drilled” into the design space to create 4 clear optima

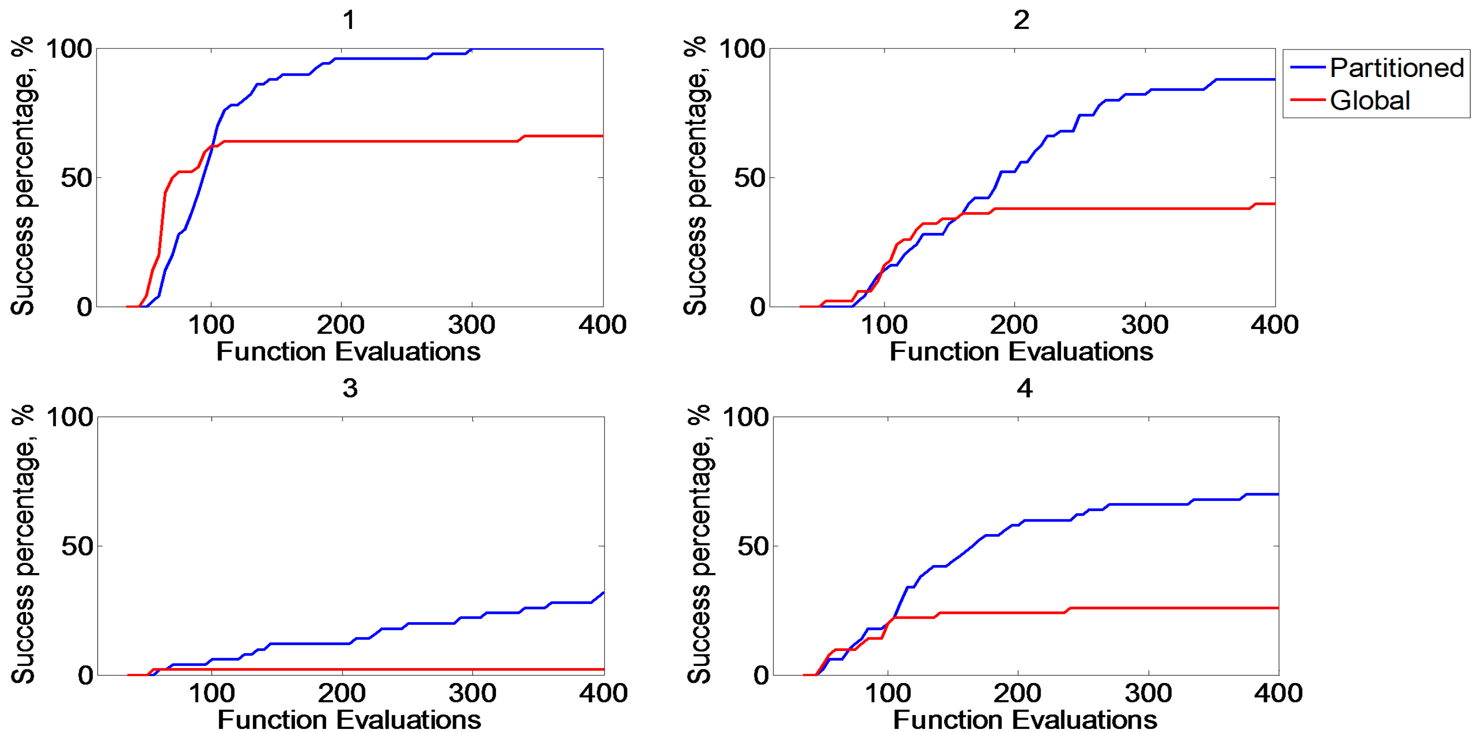
- Measured volume of basins of attraction by percentage of starts with gradient based optimizer at random locations in design space that found each optimum

Optimum	f	% of starts
1	-3.33	50.4
2	-3.21	21.1
3	-3.00	8.7
4	-2.90	19.8

Should be the hardest to located

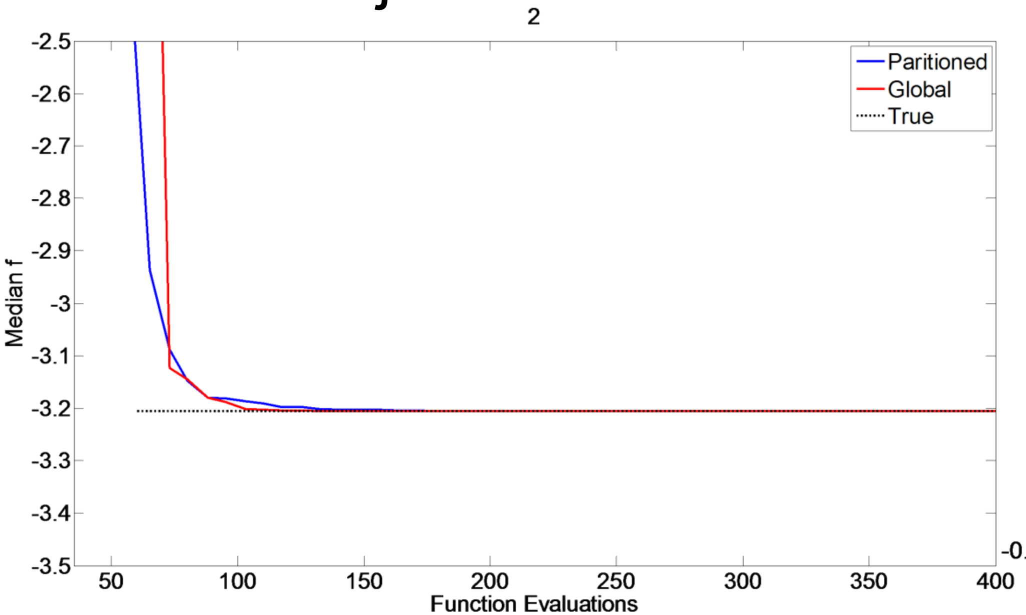
Modified Hartman 6: Success in Locating Optima

- Measured success in locating solution 1% distance away from optimum for 50 repetitions (50 different initial DOEs)
  - Distance is Euclidean distance normalized by largest possible distance in space



Modified Hartman 6: Convergence to Each Optimum

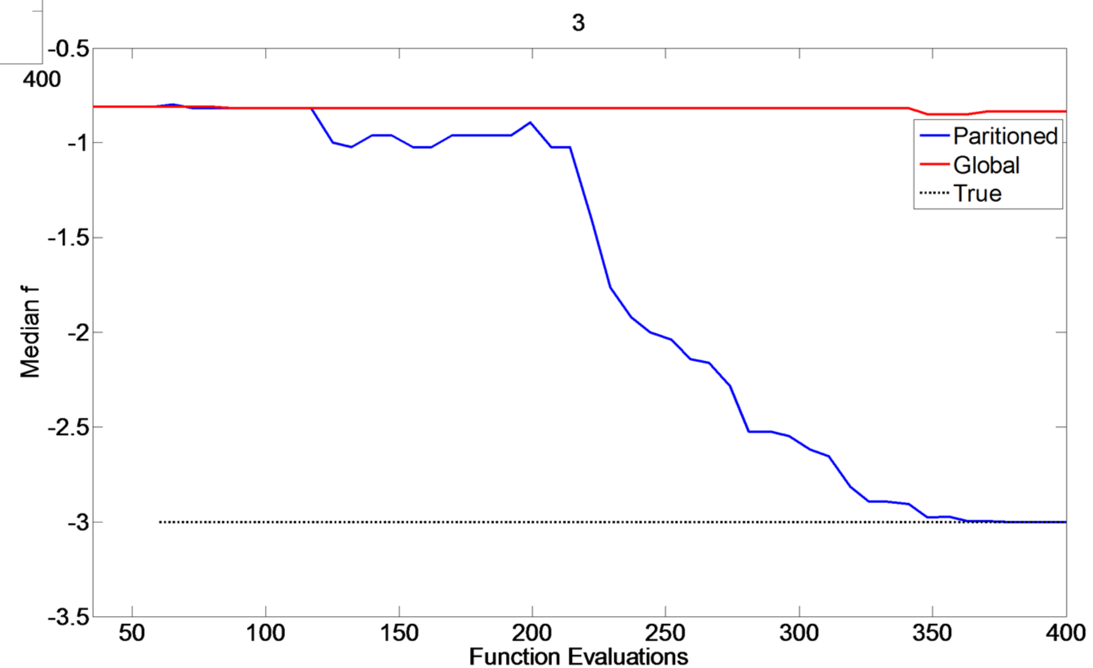
- Median objective function with increasing function evaluations



- For most optima, not a significant difference in convergence rates

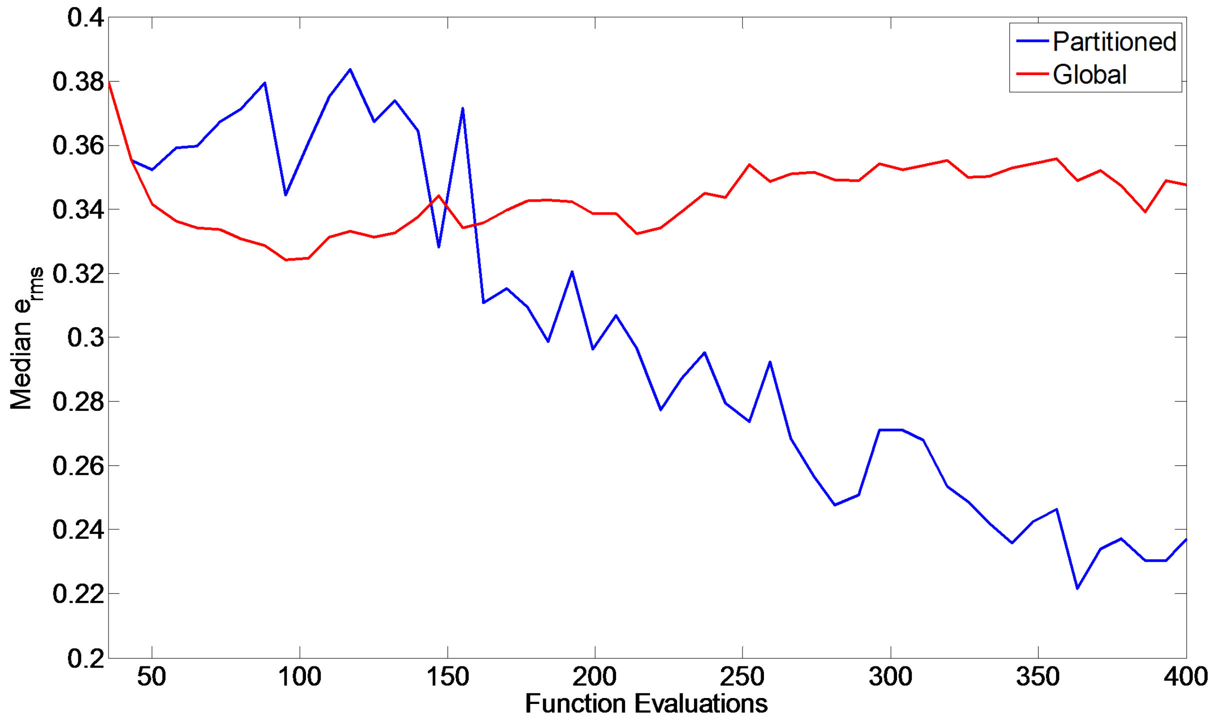
… but this was not the case for optima in smaller basins

- Slow convergence to optimum 3
- Multiple agents with partitioning were able to find these optima



Modified Hartman 6: Surrogate Error at Test Points

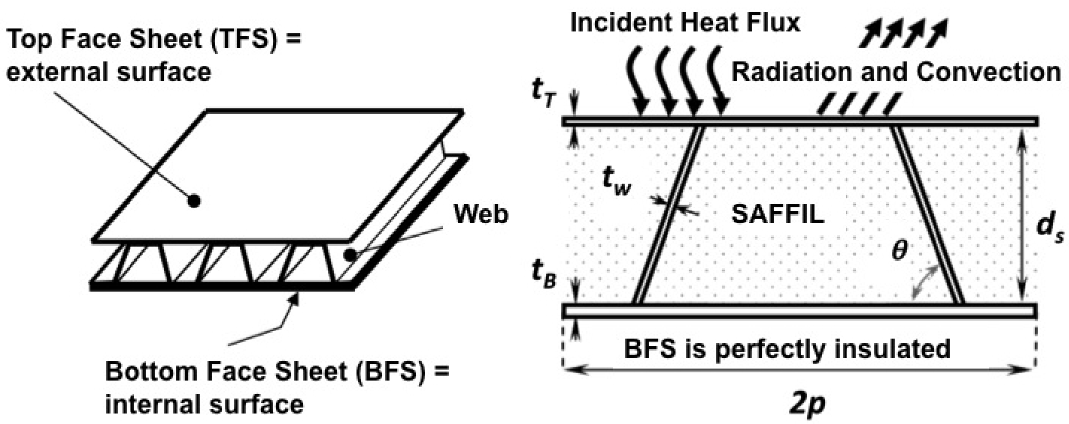
- Measured the error of the surrogate approximations of *f* at 1000 test points (LHS sampling) by e_{rms}



- Error is reduced in the case with partitioning
- Error for single global agent stays nearly constant

Integrated Thermal Protection System: Problem Description

- Design of an integrated thermal protection system
  - Structure on launch vehicle that provides structural support and heating protection
    - Two failure modes: thermal and stress
    - 5 design variables: x = *tw*, *tB*, *dS*, *tT*, θ



$$\min_x m(x)$$

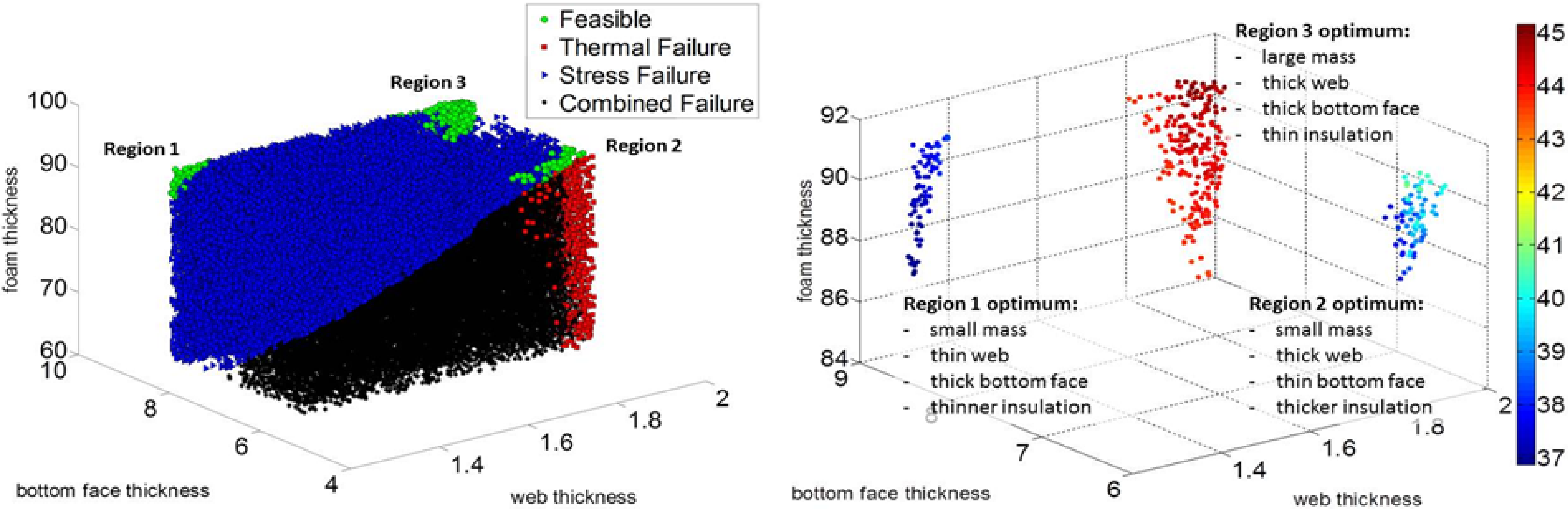
$$\text{s.t. } g_1(x) = \frac{T_{BFS}(x) + S_T}{T^{allow}} - 1 \le 0$$

$$g_2(x) = \frac{\sigma_{web}(x) S_S}{\sigma^{allow}} - 1 \le 0$$

Approximate both constraints with surrogates

Errors at test points for both surrogates were small over the iterations (~10^{-10})

Integrated Thermal Protection System: design trade-offs



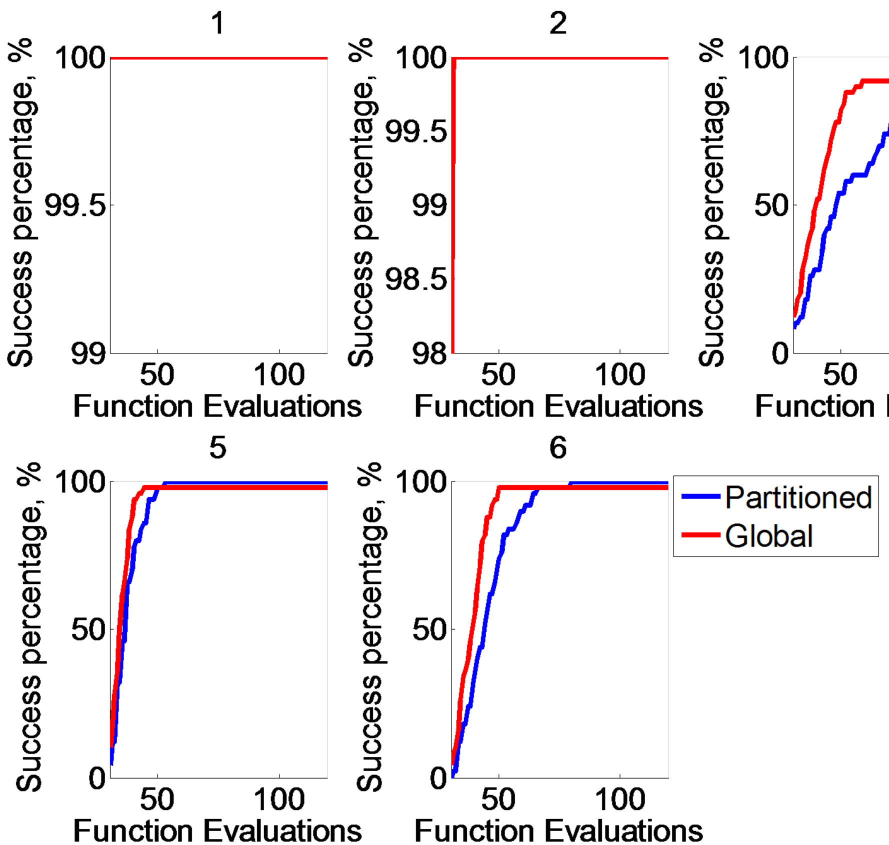
$$\min_x m(x)$$

$$\text{s.t. } g_1(x) = \frac{T_{BFS}(x) + S_T}{T^{allow}} - 1 \le 0$$

$$g_2(x) = \frac{\sigma_{web}(x) S_S}{\sigma^{allow}} - 1 \le 0$$

ITPS Example: Success in Locating Optima

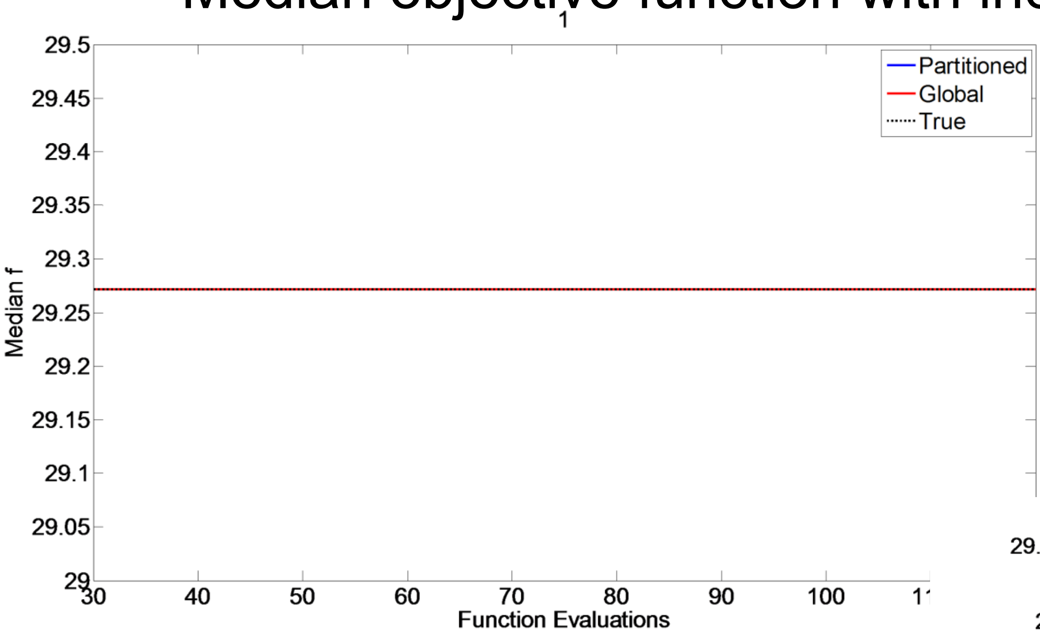
- Measured success in locating a **feasible** solution 0.01 distance from optimum for 50 repetitions (50 different initial DOEs)



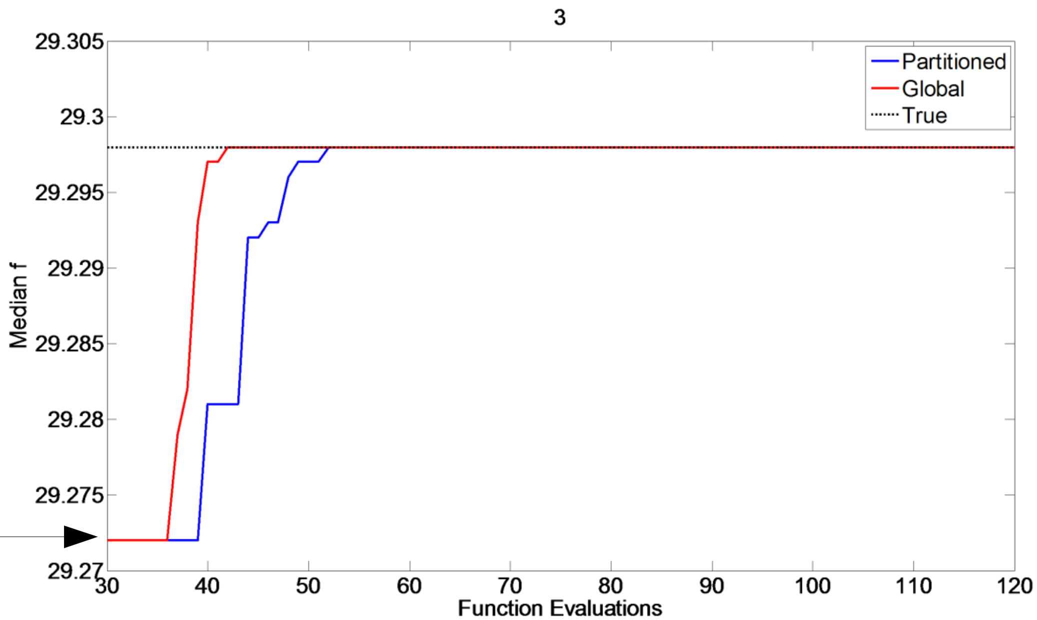
All cases show similar success percentages

ITPS Example: Convergence to Each Optimum

- Median objective function with increasing function evaluations



For most optima, incredibly quick convergence (within 5 function evaluations, not including the initial DOE)



For some optima, we observe that the nearest best points are nearby optima until locating the other basin

Optimum 1
(nearest optimum)

Problem Dependent Success

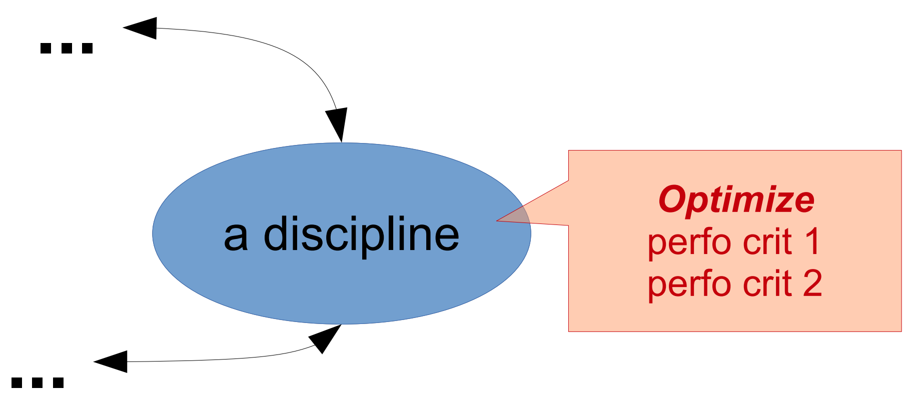
- Why is there a difference in the success and efficiency of partitioning between both problems?
  - Behavior in the ITPS problem is easy to approximate globally
    - Observed smaller error at test points with single surrogate
  - Hartman 6 is more complex, requiring more accurate surrogates to approximate the behavior
- Partitioning may be **dependent on the need for higher accuracy surrogates**
- Otherwise, simpler methods are sufficient

To sum up

- Limited expensive function calls (thanks to metamodels)
- Local optima are found
- Partitioning may be more efficient than random exploration
- Potential for distribution (thanks to agents)

Back to ID4CS

- This optimization algorithm will be used in the ID4CS platform to solve local optimization problems



- Asset: find local optima, which might become global as the overall problem formulation changes (new constraints)

Back to the interdisciplinary dialog

- CS plus : new knowledge useful for the future. Surrogate-based reasoning should be useful in other multi-agent applications

- CS minus : contribution somewhat unbalanced towards the applied math / mechanical engineering side (due to Diane's background)

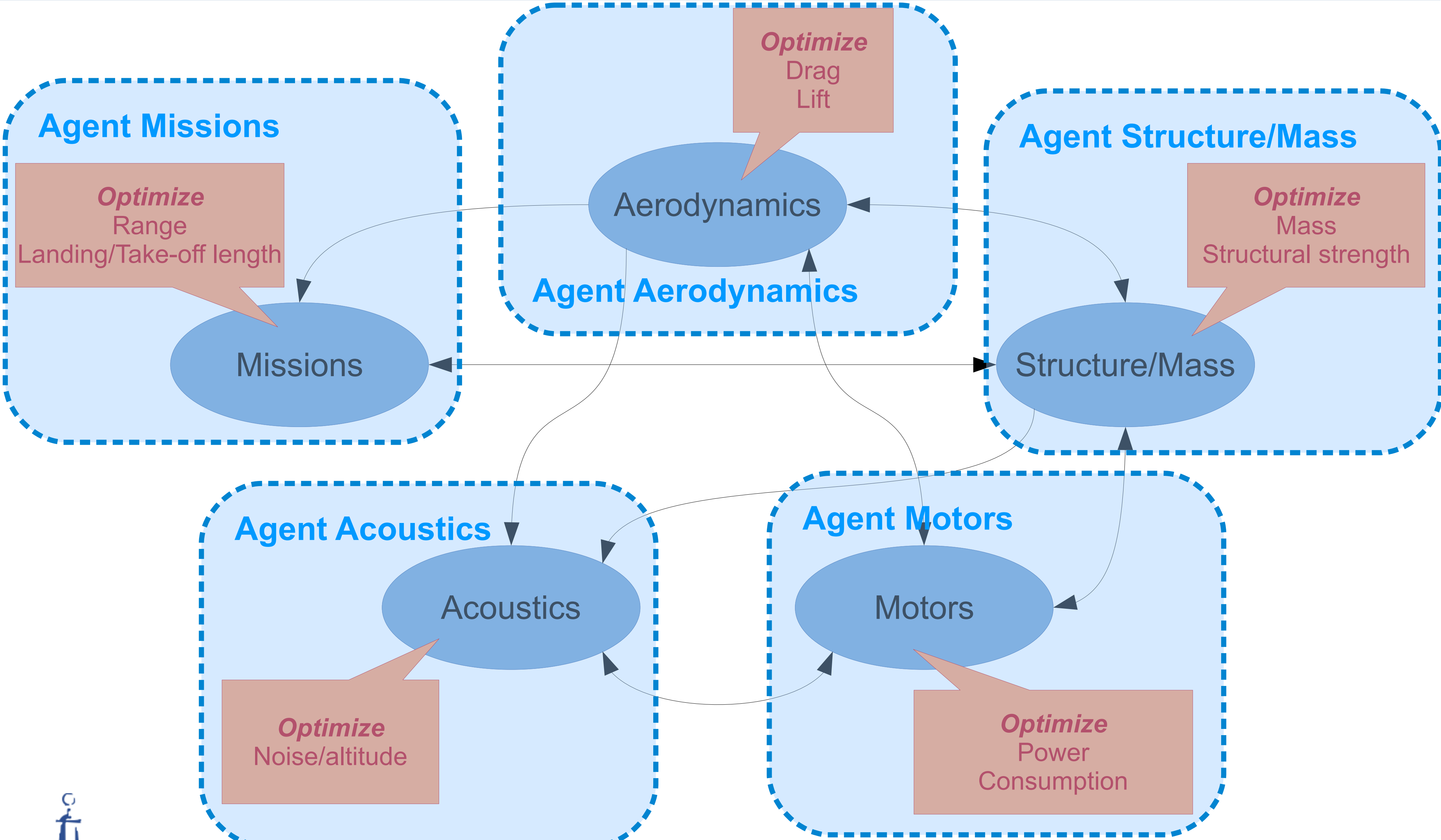
Back to the interdisciplinary dialog

- AM plus : towards multi-optimizers for distributed computing and/or collaborative decision. Would not have done it otherwise since autonomy is suboptimal in terms of centralized information
- AM minus : would like to see middle grain agents, either emerging from low grain or from a priori decomposition (according to the organization structure). Would like convergence analysis

Back to decision, agents and optimization

- Formalized decision model based on multi-agent and optimization
- There still exist solutions to explore, between fully centralized MDO and fully agentified MDO

Multi-disciplinary optimization (discipline-to-agent mapping)



Decision, agents and optimization

- Some reflexions to integrate PLM in ID4CS
  - To exploit the integrative properties of such platforms
  - But additionally requires to handle multi-fidelity and to integrate more models (at least)
- Is such an approach applicable to human organizations (*à la* Airbus)?

Bibliography

[1] Villanueva, D., Le Riche, R., Picard, G., and Haftka, R. Dynamic design space partitioning for optimization of an integrated thermal protection system. In 9th AIAA Multidisciplinary Design Optimization Specialist Conference co-located with the 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference (SDM'13) (2013), AIAA.

[2] Villanueva, D., Le Riche, R., Picard, G., and Haftka, R. Self-organized space partitioning for multi-agent optimization. In *6th International Workshop on Optimisation in Multi-Agent Systems (OPTMAS 2013, in conjunction with AAMAS 2013 6th-7th May 2013)* (2013).

[3] Villanueva, D., Le Riche, R., Picard, G., and Haftka, R. Surrogate-based agents for constrained optimization. In *14th AIAA Non-Deterministic Approaches Conference, Honolulu, HI* (2012), AIAA.

Questions ?

This work has benefited from support from Agence Nationale de la Recherche (French National Research Agency) with ANR-09-COSI-005 reference.