OAK RIDGE National Laboratory | COMPUTATIONAL & APPLIED MATHEMATICS

Foundations of Rigorous Mathematics for Uncertainty Quantification in Large Systems At The Extreme scale (FORMULATE)

Clayton G. Webster* & **Pablo Seleson**

W. Dahmen (Aachen), R. DeVore (TAMU), Q. Du (Columbia),
M. Gunzburger (FSU), N. Zabaras (Warwick), G. Zhang (ORNL)

*Department of Computational & Applied Mathematics (CAM)
Oak Ridge National Laboratory, **and**
Department of Mathematics, University of Tennessee

DARPA EQUiPS Kick-off Meeting

https://formulate.ornl.gov

Outline

1. The FORMULATE Team
2. Motivation, overview, and program goals
3. The necessary mathematical innovations
4. Applications to nonlocal models for materials fracture
5. Deliverables, risk, and mitigation plans

The FORMULATE Team

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

Mathematics Team:

Andrea Bornia, TAMU
Wolfgang Dahmen, RWTH (co-PI)
Ron DeVore, TAMU (co-PI)
Qiang Du, Columbia (co-PI)
Max Gunzburger, FSU (co-PI)
Kody Law, ORNL
Greta Petrova, TAMU
Clayton Webster, ORNL (PI)
Nicholas Zabaras, Warwick (co-PI)
Guannan Zhang, ORNL (co-PI)

Application Team:

Qiang Du, Columbia (co-PI)
Max Gunzburger, FSU (co-PI)
Pablo Seleson, ORNL
Miroslav Stoyanov, ORNL
Nicholas Zabaras, Warwick (co-PI)

Students & Postdocs:

Nick Dexter (UTK)
Peter Jantsch (UTK)
4 additional students (TBD)
Hoang Tran (ORNL)
3 additional PD's (TBD)

Motivation & goals

Uncertainty quantification for large-scale applications of materials fracture

The goals of FORMULATE

We will design a set of *fundamentally novel* mathematical approaches for UQ used to predict and control the probability of a deformation (i.e., fracture) of a material, leading to a catastrophic failure of a designed structure.

- Critical limitation in deciding the reliability and lifetime of a designed structure for application deployment to aerospace, marine and high-energy applications [Materials genome initiative, 2011].

"**The pentagon confronts new enemy: RUST**" [Wired, 2012].

"...fighting corrosion costs the DOD (Air Force and Navy) over 3% of the national GDP, or about $1,000 per person per year."

"**Recent blade failures underscore importance of risk mitigation**" [The Bakken, 2014]

Gas turbine blades: *"Failures in nickel-based superalloys can occur by various mechanisms that are operative at high temperatures...typically creep and stress rupture."*

Wind turbine blades: *"Rapid large-scale structural fracture is not a common failure mode and is often due to some sort of manufacturing defect."*

Motivation & goals

Uncertainty quantification for large-scale applications of materials fracture

The goals of FORMULATE

We will design a set of *fundamentally novel* mathematical approaches for UQ used to predict and control the probability of a deformation (i.e., fracture) of a material, leading to a catastrophic failure of a designed structure.

- Critical limitation in deciding the reliability and lifetime of a designed structure for application deployment to aerospace, marine and high-energy applications [Materials genome initiative, 2011].

"**The pentagon confronts new enemy: RUST**" [Wired, 2012].

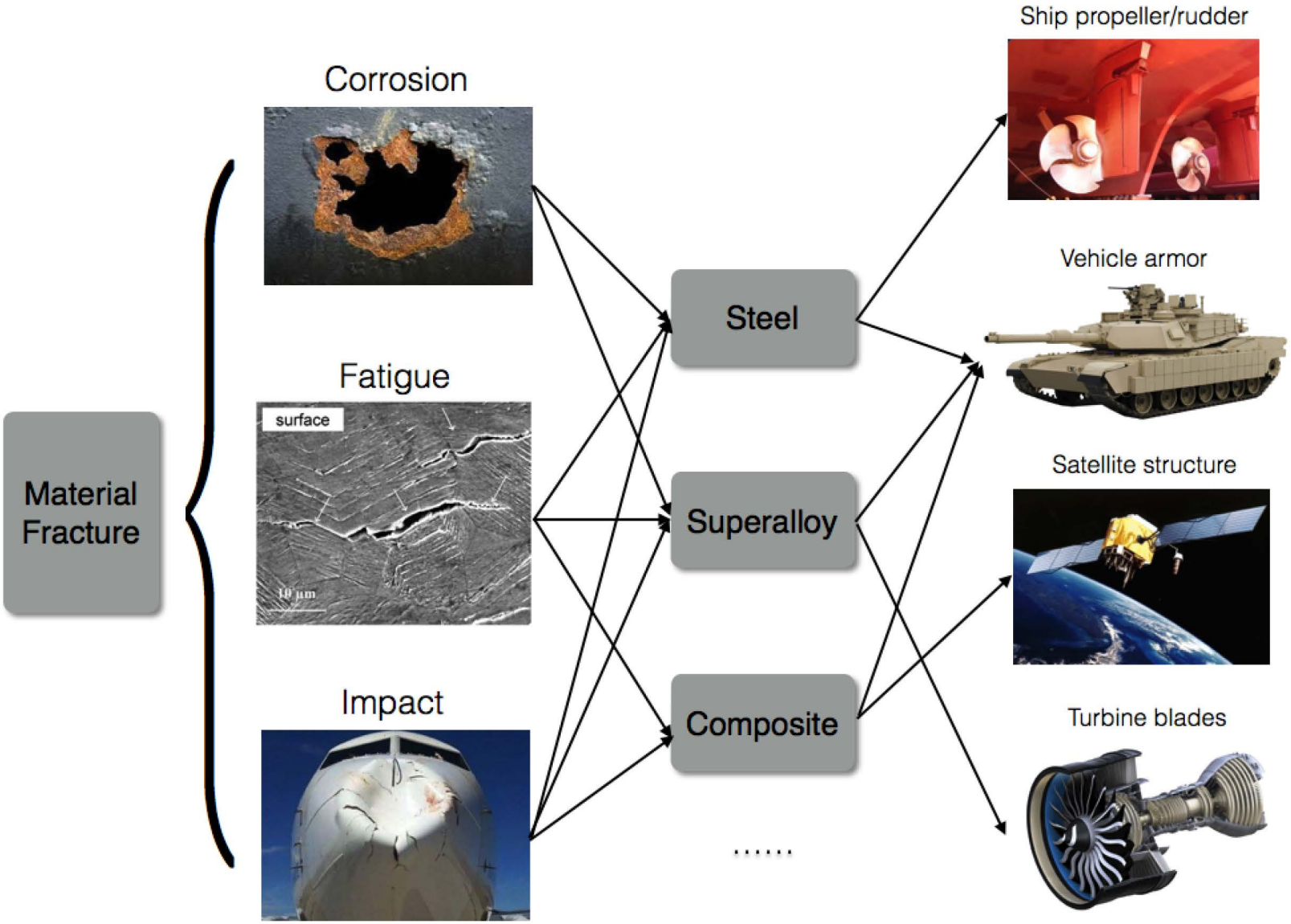
> *"...fighting corrosion costs the DOD (Air Force and Navy) over 3% of the national GDP, or about $1,000 per person per year."*

"**Recent blade failures underscore importance of risk mitigation**" [The Bakken, 2014]

> **Gas turbine blades**: *"Failures in nickel-based superalloys can occur by various mechanisms that are operative at high temperatures...typically creep and stress rupture."*
> **Wind turbine blades**: *"Rapid large-scale structural fracture is not a common failure mode and is often due to some sort of manufacturing defect."*

Why UQ for fracture mechanics?

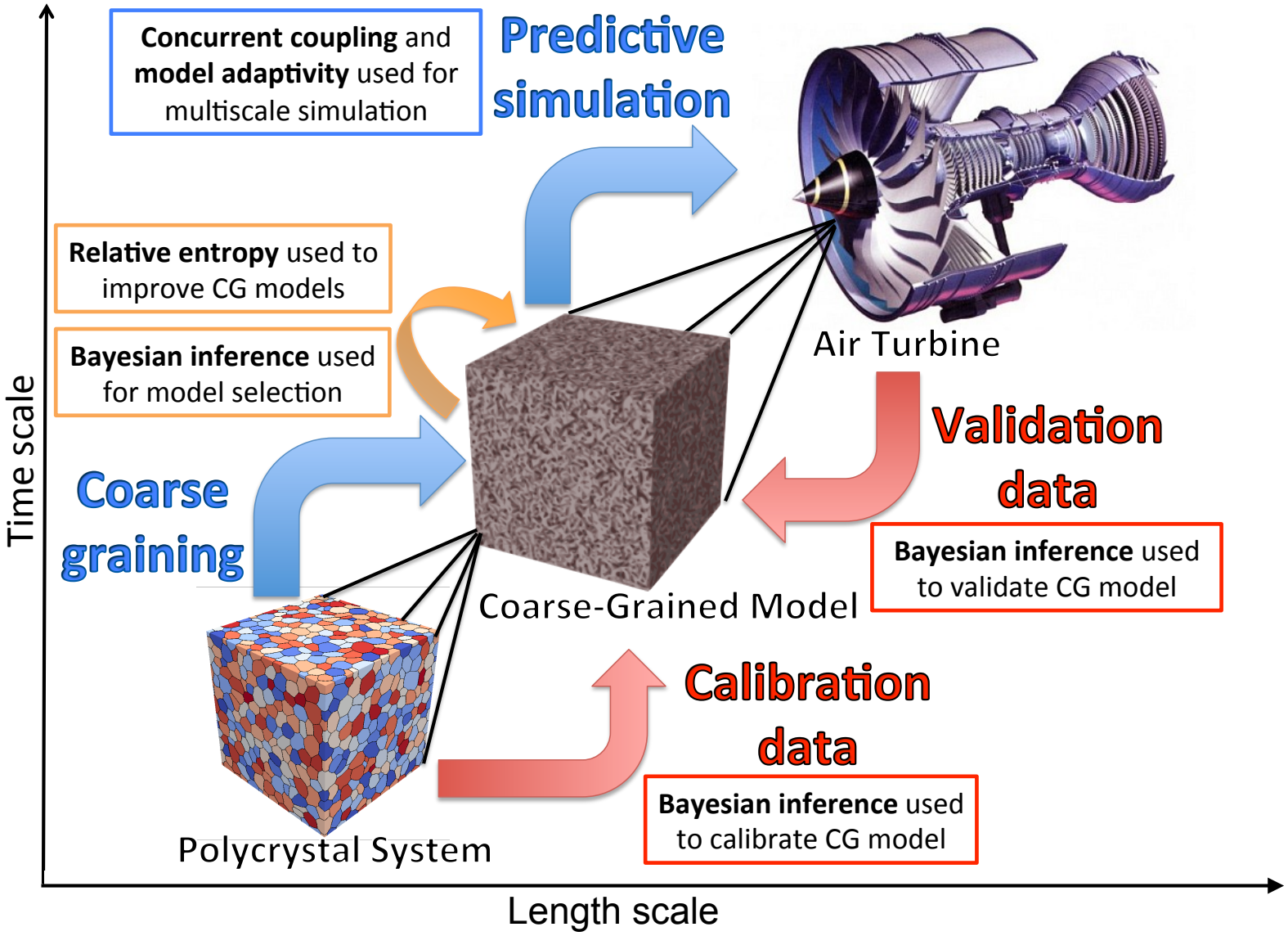
Broad impact on DOD applications



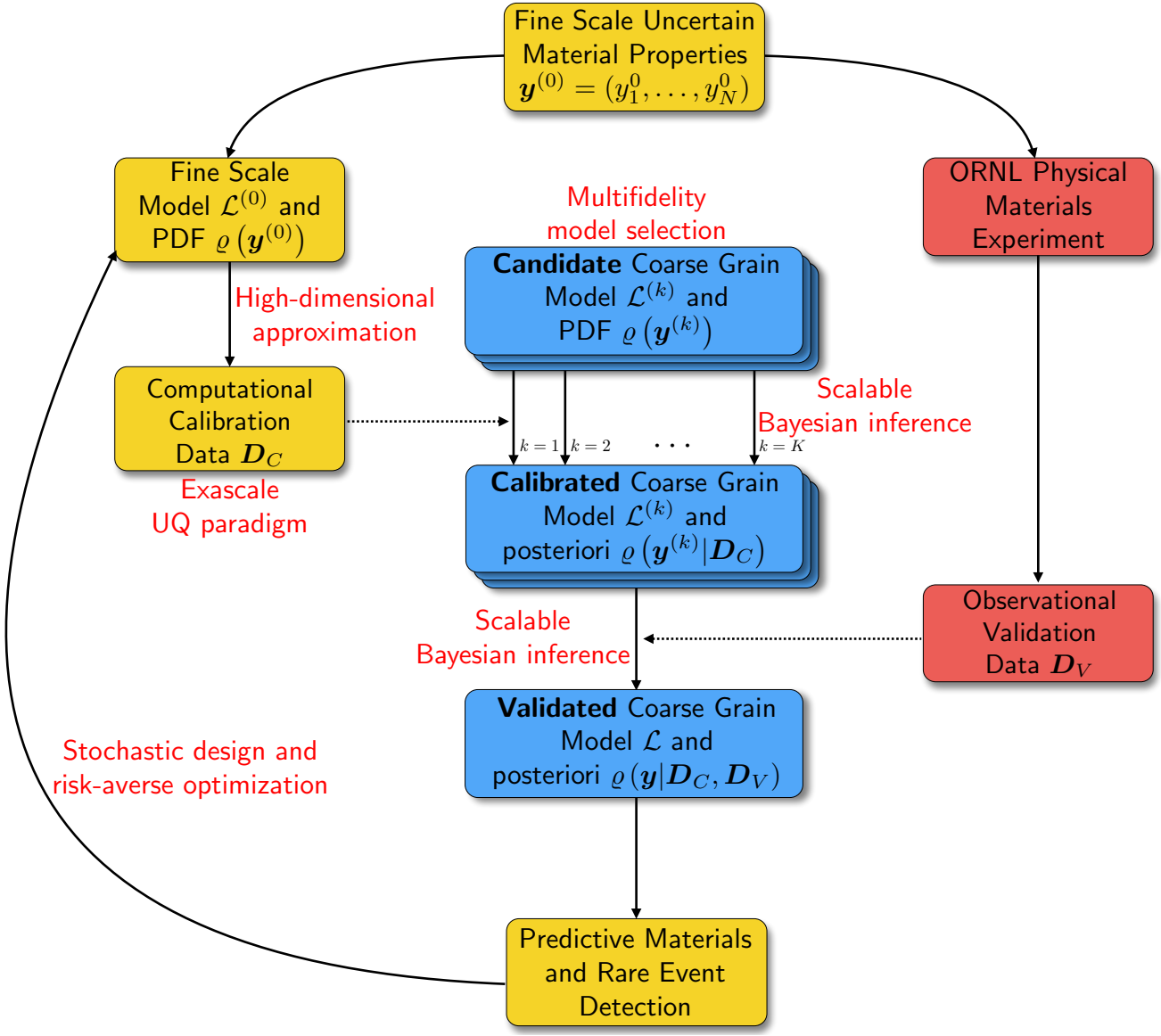
The basic idea...

To improve the predictability of *fracture* models we will extract knowledge from both fine-scale simulations and experimental data, which when combined with extreme-scale computing, will be used to conduct simulation code verification, model calibration, validation, bias correction, and even control of high-fidelity solutions.

The basic idea...



An overview of the tasks within FORMULATE



A nonlocal model for fracture prediction

We propose a generalized **continuum** theory based on **spatial integration**, that employs a **nonlocal model** of force interaction. Key components: 1 can recover classical fracture mechanics or peridynamics through the definition of the constitutive relationship; 2 no differentiability assumption of displacement fields.

Nonlocal equation of motion

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x,t) = \int_D \left\{\mathcal{T}\left[x,t\right]\left\langle x'-x\right\rangle - \mathcal{T}\left[x',t\right]\left\langle x-x'\right\rangle\right\} dV_{x'} + b(x,t)$$

ρ material density, u displacement field, b body force density, $D \subset \mathbb{R}^{1,2,3}$.

Force vector state

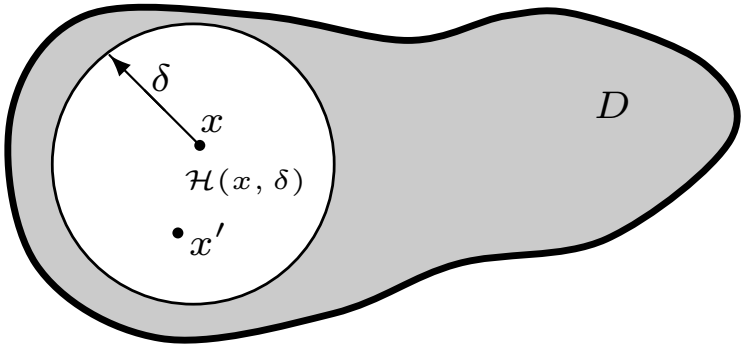
$\mathcal{T}\left[x,t\right]\left\langle\cdot\right\rangle$: "bond" $\rightarrow$ force per volume squared

Neighborhood

$\mathcal{H}(x,\delta) := \left\{x' \in \mathbb{R}^d : \|x'-x\| \leq \delta\right\}$

$\Rightarrow \mathcal{T}\left[x,t\right]\left\langle x'-x\right\rangle = \mathbf{0}$, for $\|x'-x\| > \delta$

δ is the horizon (length scale)



A nonlocal model for fracture prediction with uncertainty

Nonlocal equation of motion

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x,t) = \int_D \left\{\mathcal{T}\left[x,t\right]\langle\xi\rangle - \mathcal{T}\left[x',t\right]\langle-\xi\rangle\right\} dV_{x'} + b(x,t) \tag{1}$$

- Random vector $\boldsymbol{y}_1 \in \mathcal{U}_{d_1} \subset \mathbb{R}^{d_1}$ coming from the external loadings
- Random vector $\boldsymbol{y}_2 \in \mathcal{U}_{d_2} \subset \mathbb{R}^{d_2}$ coming from the material distribution
- Random vector $\boldsymbol{y}_3 \in \mathcal{U}_{d_3} \subset \mathbb{R}^{d_3}$ coming from the constitutive relation
- In addition, we may have uncertainty in the initial and boundary conditions, e.g., displacement, traction.

Let $\mathcal{U} = \mathcal{U}_{d_1} \times \mathcal{U}_{d_2} \times \cdots \times \mathcal{U}_{d_n} \subset \mathbb{R}^d$, $(\Omega, \mathcal{F}, P)$ a probability space, and $\boldsymbol{y} : \Omega \to \mathcal{U}$ a random vector with joint probability density function $\varrho : \mathcal{U} \to \mathbb{R}_+$

- In this setting $(\Omega, \mathcal{F}, P)$ is equivalent to $(\Gamma, \mathcal{B}(\Gamma), \varrho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\Gamma)$ denotes the Borel σ-algebra on Γ and $\varrho(\boldsymbol{y})d\boldsymbol{y}$ is the probability measure of $\boldsymbol{y}$.

Goal: given a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximate the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,, \boldsymbol{y})$.

A nonlocal model for fracture prediction with uncertainty

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

Nonlocal equation of motion with uncertainty

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x,t,\boldsymbol{y}) = \int_D \left\{\mathcal{T}\left[x,t\right]\langle\xi\rangle - \mathcal{T}\left[x',t\right]\langle-\xi\rangle\right\} dV_{x'} + b(x,t,\boldsymbol{y}_1) \qquad (1)$$

- Random vector $\boldsymbol{y}_1 \in \mathcal{U}_{d_1} \subset \mathbb{R}^{d_1}$ coming from the external loadings
- Random vector $\boldsymbol{y}_2 \in \mathcal{U}_{d_2} \subset \mathbb{R}^{d_2}$ coming from the material distribution
- Random vector $\boldsymbol{y}_3 \in \mathcal{U}_{d_3} \subset \mathbb{R}^{d_3}$ coming from the constitutive relation
- In addition, we may have uncertainty in the initial and boundary conditions, e.g., displacement, traction.

Let $\mathcal{U} = \mathcal{U}_{d_1} \times \mathcal{U}_{d_2} \times \cdots \times \mathcal{U}_{d_n} \subset \mathbb{R}^d$, $(\Omega, \mathcal{F}, P)$ a probability space, and $\boldsymbol{y} : \Omega \to \mathcal{U}$ a random vector with joint probability density function $\varrho : \mathcal{U} \to \mathbb{R}_+$

- In this setting $(\Omega, \mathcal{F}, P)$ is equivalent to $(\Gamma, \mathcal{B}(\Gamma), \varrho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\Gamma)$ denotes the Borel σ-algebra on Γ and $\varrho(\boldsymbol{y})d\boldsymbol{y}$ is the probability measure of $\boldsymbol{y}$.

Goal: given a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximate the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,,\boldsymbol{y})$.

A nonlocal model for fracture prediction with uncertainty

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

Nonlocal equation of motion with uncertainty

$$\rho(x, \boldsymbol{y}_2)\frac{\partial^2 u}{\partial t^2}(x,t,\boldsymbol{y}) = \int_D \left\{\mathcal{T}\left[x,t\right]\langle\xi\rangle - \mathcal{T}\left[x',t\right]\langle-\xi\rangle\right\} dV_{x'} + b(x,t,\boldsymbol{y}_1) \qquad (1)$$

- Random vector $\boldsymbol{y}_1 \in \mathcal{U}_{d_1} \subset \mathbb{R}^{d_1}$ coming from the external loadings
- Random vector $\boldsymbol{y}_2 \in \mathcal{U}_{d_2} \subset \mathbb{R}^{d_2}$ coming from the material distribution
- Random vector $\boldsymbol{y}_3 \in \mathcal{U}_{d_3} \subset \mathbb{R}^{d_3}$ coming from the constitutive relation
- In addition, we may have uncertainty in the initial and boundary conditions, e.g., displacement, traction.

Let $\mathcal{U} = \mathcal{U}_{d_1} \times \mathcal{U}_{d_2} \times \cdots \times \mathcal{U}_{d_n} \subset \mathbb{R}^d$, $(\Omega, \mathcal{F}, P)$ a probability space, and $\boldsymbol{y} : \Omega \to \mathcal{U}$ a random vector with joint probability density function $\varrho : \mathcal{U} \to \mathbb{R}_+$

- In this setting $(\Omega, \mathcal{F}, P)$ is equivalent to $(\Gamma, \mathcal{B}(\Gamma), \varrho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\Gamma)$ denotes the Borel σ-algebra on Γ and $\varrho(\boldsymbol{y})d\boldsymbol{y}$ is the probability measure of $\boldsymbol{y}$.

Goal: given a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximate the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,, \boldsymbol{y})$.

A nonlocal model for fracture prediction with uncertainty

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

Nonlocal equation of motion with uncertainty

$$\rho(x, \boldsymbol{y}_2)\frac{\partial^2 u}{\partial t^2}(x,t,\boldsymbol{y}) = \int_D \left\{\mathcal{T}\left[x,t,\boldsymbol{y}_3\right]\langle\xi\rangle - \mathcal{T}\left[x',t,\boldsymbol{y}_3\right]\langle-\xi\rangle\right\} dV_{x'} + b(x,t,\boldsymbol{y}_1) \tag{1}$$

- Random vector $\boldsymbol{y}_1 \in \mathcal{U}_{d_1} \subset \mathbb{R}^{d_1}$ coming from the external loadings
- Random vector $\boldsymbol{y}_2 \in \mathcal{U}_{d_2} \subset \mathbb{R}^{d_2}$ coming from the material distribution
- Random vector $\boldsymbol{y}_3 \in \mathcal{U}_{d_3} \subset \mathbb{R}^{d_3}$ coming from the constitutive relation
- In addition, we may have uncertainty in the initial and boundary conditions, e.g., displacement, traction.

Let $\mathcal{U} = \mathcal{U}_{d_1} \times \mathcal{U}_{d_2} \times \cdots \times \mathcal{U}_{d_n} \subset \mathbb{R}^d$, $(\Omega, \mathcal{F}, P)$ a probability space, and $\boldsymbol{y} : \Omega \to \mathcal{U}$ a random vector with joint probability density function $\varrho : \mathcal{U} \to \mathbb{R}_+$

- In this setting $(\Omega, \mathcal{F}, P)$ is equivalent to $(\Gamma, \mathcal{B}(\Gamma), \varrho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\Gamma)$ denotes the Borel σ-algebra on Γ and $\varrho(\boldsymbol{y})d\boldsymbol{y}$ is the probability measure of $\boldsymbol{y}$.

Goal: given a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximate the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,, \boldsymbol{y})$.

A nonlocal model for fracture prediction with uncertainty

Nonlocal equation of motion with uncertainty

$$\rho(x, \boldsymbol{y}_2)\frac{\partial^2 u}{\partial t^2}(x,t,\boldsymbol{y}) = \int_D \left\{\mathcal{T}\left[x,t,\boldsymbol{y}_3\right]\langle\xi\rangle - \mathcal{T}\left[x',t,\boldsymbol{y}_3\right]\langle-\xi\rangle\right\} dV_{x'} + b(x,t,\boldsymbol{y}_1) \tag{1}$$

- Random vector $\boldsymbol{y}_1 \in \mathcal{U}_{d_1} \subset \mathbb{R}^{d_1}$ coming from the external loadings
- Random vector $\boldsymbol{y}_2 \in \mathcal{U}_{d_2} \subset \mathbb{R}^{d_2}$ coming from the material distribution
- Random vector $\boldsymbol{y}_3 \in \mathcal{U}_{d_3} \subset \mathbb{R}^{d_3}$ coming from the constitutive relation
- In addition, we may have uncertainty in the initial and boundary conditions, e.g., displacement, traction.

Let $\mathcal{U} = \mathcal{U}_{d_1} \times \mathcal{U}_{d_2} \times \cdots \times \mathcal{U}_{d_n} \subset \mathbb{R}^d$, $(\Omega, \mathcal{F}, P)$ a probability space, and $\boldsymbol{y} : \Omega \to \mathcal{U}$ a random vector with joint probability density function $\varrho : \mathcal{U} \to \mathbb{R}_+$

- In this setting $(\Omega, \mathcal{F}, P)$ is equivalent to $(\Gamma, \mathcal{B}(\Gamma), \varrho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\Gamma)$ denotes the Borel σ-algebra on Γ and $\varrho(\boldsymbol{y})d\boldsymbol{y}$ is the probability measure of $\boldsymbol{y}$.

Goal: given a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximate the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,, \boldsymbol{y})$.

A nonlocal model for fracture prediction with uncertainty

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

Nonlocal equation of motion with uncertainty

$$\rho(x, \boldsymbol{y}_2)\frac{\partial^2 u}{\partial t^2}(x,t,\boldsymbol{y}) = \int_D \left\{\mathcal{T}\left[x,t,\boldsymbol{y}_3\right]\langle\xi\rangle - \mathcal{T}\left[x',t,\boldsymbol{y}_3\right]\langle-\xi\rangle\right\} dV_{x'} + b(x,t,\boldsymbol{y}_1) \tag{1}$$

- Random vector $\boldsymbol{y}_1 \in \mathcal{U}_{d_1} \subset \mathbb{R}^{d_1}$ coming from the external loadings
- Random vector $\boldsymbol{y}_2 \in \mathcal{U}_{d_2} \subset \mathbb{R}^{d_2}$ coming from the material distribution
- Random vector $\boldsymbol{y}_3 \in \mathcal{U}_{d_3} \subset \mathbb{R}^{d_3}$ coming from the constitutive relation
- In addition, we may have uncertainty in the initial and boundary conditions, e.g., displacement, traction.

Let $\mathcal{U} = \mathcal{U}_{d_1} \times \mathcal{U}_{d_2} \times \cdots \times \mathcal{U}_{d_n} \subset \mathbb{R}^d$, $(\Omega, \mathcal{F}, P)$ a probability space, and $\boldsymbol{y} : \Omega \to \mathcal{U}$ a random vector with joint probability density function $\varrho : \mathcal{U} \to \mathbb{R}_+$

- In this setting $(\Omega, \mathcal{F}, P)$ is equivalent to $(\Gamma, \mathcal{B}(\Gamma), \varrho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\Gamma)$ denotes the Borel σ-algebra on Γ and $\varrho(\boldsymbol{y})d\boldsymbol{y}$ is the probability measure of $\boldsymbol{y}$.

Goal: given a query $\boldsymbol{y} \in \mathcal{U}$, quickly approximate the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,, \boldsymbol{y})$.

Grand challenges

We must overcome several challenges which arise when applying UQ methodologies to the high-dimensional stochastic model for fracture prediction:

1. **Thrust area 1**: Rigorous mathematical analysis and massively scalable algorithms which alleviate the curse of dimensionality (CoD) and reduce the computational burden of accurate stochastic solutions;

2. **Thrust area 2**: Robust approaches that are capable of reducing model uncertainties and improving model predictability, by assimilating experimental data and exploiting multifidelity models;

3. **Thrust area 3**: A theoretical framework for stochastic design and decision-making under uncertainty, that uses multiple, high volume data sources and multi-index hierarchical modeling capabilities, and;

4. **Thrust area 4**: Scalable algorithms for UQ that exploit greater levels of parallelism provided by emerging many-core architectures.

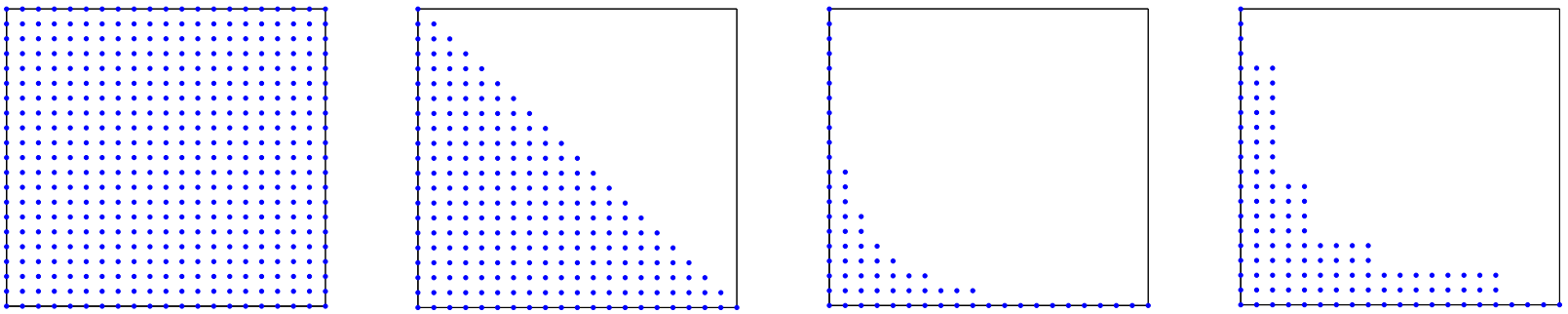
Task 1

Best s-term linear and nonlinear methods

Let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \in \Lambda \subset \mathbb{N}^d$ a multi-index set, and $\boldsymbol{\Psi}_{\boldsymbol{\nu}}$ be multivariate polynomials. Approximate the solution u by:

$$u_\Lambda(x, t, \boldsymbol{y}) = \sum_{\boldsymbol{\nu} \in \Lambda} c_{\boldsymbol{\nu}}(x, t) \boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}) \quad \in \mathbb{P}_\Lambda(\mathcal{U}) = span\{\boldsymbol{\Psi}_{\boldsymbol{\nu}}(\boldsymbol{y}), \boldsymbol{\nu} \in \Lambda\} \tag{2}$$

- The efficiency of such methods heavily depends on the selection of Λ.
- Standard approaches: impose index sets Λ *a priori*. The cardinality of the polynomial space $\mathbb{P}_\Lambda(\mathcal{U})$ can grow exponentially with respect to the dimension d.



Tensor Product
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \max_{1 \le i \le N} \nu_i \le w\}$

Total Degree
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum \nu_i \le w\}$

Hyperbolic Cross
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \prod (\nu_i + 1) \le w + 1\}$

Smolyak
$\Lambda(w) = \{\nu \in \mathbf{N}^N : \sum f(\nu_i) \le f(w)\}$,
with $f(\nu) = \lceil \log_2(\nu) \rceil$, $\nu \ge 2$.

Goal: construct the "optimal" $\Lambda \subset \mathbb{N}^d$ has minimal cardinality and enables the approximation of $\boldsymbol{y} \mapsto u(\boldsymbol{y})$ with maximum accuracy for a given computational cost.

Best s-term approximations using $L^2(\mathcal{U}, \varrho)$ projections

Goal: Construct the **optimal** set Λ_s^{best} of s most effective multi-indices with respect to sharp estimates $B(\boldsymbol{\nu})$ of $\|c_{\boldsymbol{\nu}}\|_{\mathcal{V}}$ [Cohen, DeVore, Schwab '12, '13], [Tran, Webster, Zhang '15]:

$$\sup_{\boldsymbol{y}\in\Gamma}\left\|u - \sum_{\boldsymbol{\nu}\in\Lambda_s^{\text{best}}} c_{\boldsymbol{\nu}}\boldsymbol{\Psi}_{\boldsymbol{\nu}}\right\|_{\mathcal{V}} \leq \sum_{\boldsymbol{\nu}\notin\Lambda_s^{\text{best}}} \|c_{\boldsymbol{\nu}}\|_{\mathcal{V}} \leq C_\epsilon s\, e^{-\kappa(\boldsymbol{r},d)s} \leq \|B(\boldsymbol{\nu})\|_{\ell^p(\mathbb{N}^d)} s^{1-\frac{1}{p}}$$

- Provides the **sharpest** realizable best s-term (sub-exponential) convergence rate with optimal κ in very high-dimensions.
- **Challenge:** Estimates of $B(\boldsymbol{\nu})$ are difficult and somewhat problem dependent.

Task 1.1: To develop *a priori* and *a posteriori* procedures for building Λ_s^{best} in the context of the nonlocal model (1).

- *a priori:* are derived using knowledge from the uncertainty in the input data
- *a posteriori:* adaptively build nested sequences $(\Lambda_s)_{s\geq 0}$ of Λ_s^{best} at a cost that scales linearly in $\#(\Lambda_s)$: given Λ_s, construct Λ_{s+1} by enriching Λ_s with the most effective indices $\boldsymbol{\nu}$ in its neighborhood, which results in the best residual reduction.

Non-intrusive recovery of the best s-term approximation

1. Sparse interpolation in Λ_s^{best}

Goal: Construct a *hierarchical interpolant* in $\mathbb{P}_{\Lambda_s^{\text{best}}}(\mathcal{U})$, on a set of *distinct collocation points*:

$$\mathcal{I}_{\Lambda_s^{\text{best}}}[u] = \sum_{\boldsymbol{\nu}\in\Lambda_s^{\text{best}}} \bigotimes_{n=1}^{d} \left(\mathcal{J}^{m(\nu_n)} - \mathcal{J}^{m(\nu_n-1)}\right)[u] \quad \text{and} \quad \mathcal{G}_{\Lambda_s^{\text{best}}} = \bigcup_{\boldsymbol{\nu}\in\Lambda_s^{\text{best}}} \bigotimes_{n=1}^{d} \{y_{n,k}\}_{k=1}^{m(\nu_n)}$$

- $\mathcal{J}^{m(\nu_n)}$ is a sequence of one-dimensional Lagrange interpolation operators using abscissas $\{y_{n,k}\}_{k=1}^{m(\nu_n)} \subset \mathcal{U}_n$, with $m(\nu) : \mathbb{N}_+ \to \mathbb{N}_+$ a strictly increasing function.
- **Challenge:** To recover the best approximation: ❶ the Lebesgue constant $\mathbb{L}_{\Lambda_s^{\text{best}}}$ of the underlying abscissas must grow slowly, and; ❷ $\#(\mathcal{G}_{\Lambda_s^{\text{best}}}) = \dim(\mathbb{P}_{\Lambda_s^{\text{best}}}(\mathcal{U})) = s$ $\Rightarrow$ abscissas satisfy $\{y_{n,k}\}_{k=1}^{m(\nu_n-1)} \subset \{y_{n,k}\}_{k=1}^{m(\nu_n)}$ and $m(\nu) = \nu + 1$.

Task 1.2: To explore greedy approaches, wherein, given the sequence $\{y_k\}_{k=1}^{K-1}$, with $K = m(\nu_n)$, the Kth abscissa is chosen as the extrema of some functional, e.g., a residual (Leja points), Lebesgue function, etc.

- Computationally construct the multi-dimensional interpolant, and theoretically prove a sharp bound on the corresponding multi-dimensional Lebesgue constant.

Non-intrusive recovery of the best s-term approximation

2. Compressed sensing reconstruction using weighted ℓ_1 and hard thresholding

Goal: Given a set Λ_0 of N multi-indices (could be far from optimal), find an approximation of (2) comparable to the best s-term ($s \ll N$).

- $c_\nu = c_\nu(x,t)$ is a function that belongs to a Banach space $\mathcal{V} \Rightarrow \mathbf{c} \in \mathcal{V}^N$, equipped with the norm: $\|\boldsymbol{c}\|_{\mathcal{V},p} = (\sum_{i=1}^N \|c_i\|_{\mathcal{V}}^p)^{1/p}$.
- **Challenge:** Here compressed sensing reconstruction involves randomly sampling $\mathbf{u} = (u(\boldsymbol{y}_k))_{k=1,\dots,m} \in \mathcal{V}^m$, and solving the **underdetermined** linear system

$$\mathbf{u} = \boldsymbol{\Psi}\mathbf{c}, \qquad \text{where } \boldsymbol{\Psi} = (\Psi_\nu(\boldsymbol{y}_i)) \text{ is an } m \times N \text{ sampling matrix.}$$

$$\Rightarrow \mathbf{c} = \text{argmin}\|\hat{\mathbf{c}}\|_{\mathcal{V},1} \quad \text{subject to } \mathbf{u} = \boldsymbol{\Psi}\hat{\mathbf{c}}.$$

- Best available estimate to guarantee uniform recovery (with high probability) is $m \geq CK^2\delta^{-2}s\log^2(s)\log(N)$ [Chkifa, Tran, Webster '15]

Task 1.3: To guarantee the algorithms overcome the *curse of dimensionality* we will develop innovative:

- weighted ℓ_1 (e.g., $\omega_\nu = \|\Psi_\nu\|_\infty$) techniques for optimal recovery that exploit the *a priori* and *a posteriori* decay of $(\|c_\nu\|_{\mathcal{V}})_{\nu\in\Lambda_0}$.
- tractable "hard thresholding" operators that iteratively exploit the structure (e.g., lower or downward closed sets) of the best s-term index set.

Task 2

A hierarchical Bayesian framework for inverse problems

Goal: Construct a hierarchical Bayesian framework to simultaneously quantify both parameter and model-form uncertainties.

- In a hierarchical Bayesian framework, we have

$$\text{Hyper-prior: } \boldsymbol{\theta} \sim \mu_{\boldsymbol{\theta}}, \quad \text{Prior: } \boldsymbol{y}|\boldsymbol{\theta} \sim \mu_{\boldsymbol{y}|\boldsymbol{\theta}}, \quad \text{Data density: } \boldsymbol{d}|(\boldsymbol{y},\boldsymbol{\theta}) \sim \varrho_{\boldsymbol{d}|(\boldsymbol{y},\boldsymbol{\theta})}.$$

- The QoI $Q : \mathcal{U} \times \Theta \to \mathbb{R}$ can be estimated by its posterior expectation

$$\mathbb{E}(Q|\boldsymbol{d}) = \int_{\mathcal{U}\times\Theta} Q(\boldsymbol{y},\boldsymbol{\theta})\, \mu_{(\boldsymbol{y},\boldsymbol{\theta})|\boldsymbol{d}}(d\boldsymbol{y}, d\boldsymbol{\theta}) \quad \text{with} \quad \frac{d\mu_{(\boldsymbol{y},\boldsymbol{\theta})|\boldsymbol{d}}}{d\mu_0}(\boldsymbol{y},\boldsymbol{\theta}) = \frac{1}{Z(\boldsymbol{d})}\varrho_{\boldsymbol{d}|(\boldsymbol{y},\boldsymbol{\theta})}(\boldsymbol{d}),$$

  where $Z(\boldsymbol{d}) = \int \varrho_{\boldsymbol{d}|(\boldsymbol{y},\boldsymbol{\theta})}(\boldsymbol{d}) d\mu_0$ and $\mu_0(d\boldsymbol{y}, d\boldsymbol{\theta}) = \mu_{\boldsymbol{y}|\boldsymbol{\theta}}(d\boldsymbol{y})\mu_\theta(d\boldsymbol{\theta})$.

- Let $\boldsymbol{\theta}$ to be a discrete random variable representing a set of alternative models:
  - $\mu_{\boldsymbol{\theta}|\boldsymbol{d}}(d\boldsymbol{\theta}) \propto \left(\int_{\mathcal{U}} \varrho_{\boldsymbol{d}|(\boldsymbol{y},\boldsymbol{\theta})}\mu_{\boldsymbol{y}|\boldsymbol{\theta}}(d\boldsymbol{y})\right) \mu_{\boldsymbol{\theta}}(d\boldsymbol{\theta})$ allows model selection.
  - $\mu_{\boldsymbol{y}|\boldsymbol{d}}$ is the model-averaged posterior.

- The overall task is to draw samples from the posterior $\mu_{(\boldsymbol{y},\boldsymbol{\theta})|\boldsymbol{d}}$ or the marginals $\mu_{\boldsymbol{y}|\boldsymbol{d}} := \int_\Theta \mu_{(\boldsymbol{y},\boldsymbol{\theta})|\boldsymbol{d}}(\,\cdot\,, d\boldsymbol{\theta})$ or $\mu_{\boldsymbol{\theta}|\boldsymbol{d}} := \int_{\mathcal{U}} \mu_{(\boldsymbol{y},\boldsymbol{\theta})|\boldsymbol{d}}(d\boldsymbol{y}, \,\cdot\,)$.

Derivative-free dimension-independent sampling algorithms

Goal: Develop black-box Markov chain Monte Carlo (MCMC) algorithms to sample complex nonlinear and non-Gaussian posteriors over high-dimensional parameter spaces.

- **Challenge:** Most MCMC methods suffer in certain limits, as the parameter dimension $\dim\mathcal{U}+\dim\Theta =: N \to \infty$ or the likelihood variance $\sigma \to 0$.
- Dimension-independent (DI), likelihood-informed (LI) sampling algorithms:
  - DI. By deriving the algorithm in the limiting function space, stability is ensured in the limit $N \to \infty$;
  - LI. Use of posterior concentration information from the likelihood *Hessian* ensures stability in the limit $\sigma \to 0$.

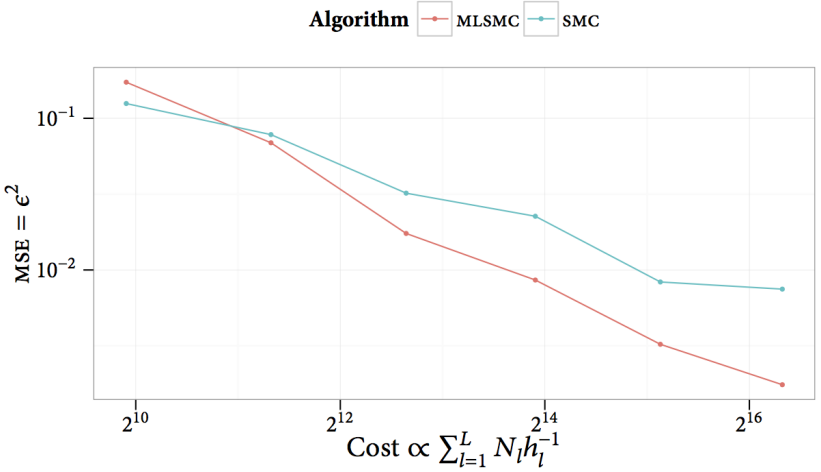
Task 2.1: Develop and analyze **Hessian-free** (black-box) DILI algorithms .

- Likelihood-informed space will be identified from $m-$sample covariance with $m = \mathcal{O}(N)$ (e.g. using robust regularized m-estimators);
- Will be used within DILI-type, adaptive Metropolis, or map-based proposals;
- Application within hierarchical models, for model selection and averaging, etc.

Multilevel sequential Monte Carlo (MLSMC) samplers

Goal: Develop algorithms to sample complex nonlinear and non-Gaussian posteriors with expensive forward model connecting parameter to data.

- **Challenge**: the cost of sampling is the product of (a) the cost to sample a scalar random variable and (b) the cost of a single deterministic solve at given accuracy.
- Recently developed MLSMC samplers reduce cost for posterior sampling to the minimum of (a) and (b).



Task 2.2: Develop, analyze, and use ML and multi-index (MI) SMC for hierarchical Bayesian computation.

- Current version assumes finite parameter. Extension will approximate *continuous parameter* – multilevel in parameter also.
- MI is a natural generalization of ML which will make the latter obsolete.

Assimilating data for optimal recovery of the inverse map

Goal: Recover an element u of a Hilbert space $\mathcal{H}$ from the measurement data of the form $\ell_j(u), j = 1, \ldots, m$, where ℓ_j are known linear functionals on $\mathcal{H}$.

- We assume that all possible states of u are described by a solution manifold

$$\mathcal{M} = \{u(\boldsymbol{y}) \,:\, \boldsymbol{y} \in \mathcal{U}\}.$$

- The data $\ell_j(u)$ for $j = 1, \ldots, m$ constitute a subspace $\mathcal{W} \subset \mathcal{H}$, and the measurements determines the projection $P_{\mathcal{W}}[u] \in \mathcal{W}$.

- The task is to construct a mapping $A : \mathcal{W} \to \mathcal{M}$, such that an approximation $\hat{u}$ of u can be defined by

$$\hat{u} := A(P_{\mathcal{W}}[u]).$$

- **Challenge:** it is usually expensive to find the optimal recovery of u in $\mathcal{M}$, due to the computational complexity of the underlying integro-PDEs.

Reduced-basis methods for optimal recovery

Task 2.3: To construct a *near optimal recovery* of u by exploiting a finite sequence of reduced-basis subspaces $V_0 \subset V_1 \subset \cdots \subset V_n \subset \mathcal{M}$, where V_k approximates $\mathcal{M}$ with accuracy ε_k for $k = 0, \ldots, n$.

- Develop novel iterative algorithms based on *alternating projection techniques* to construct the optimal map $A : \mathcal{W} \to \mathcal{K}$, where $\mathcal{K}$ is defined by

$$\mathcal{K} := \bigcap_{j=0}^{n} \mathcal{K}_j = \bigcap_{j=0}^{n} \{u \in \mathcal{M} : \text{dist}(u, V_j) \leq \varepsilon_j\}.$$

- Develop an a posterior estimate for measuring the performance of the new algorithms, which can serve both as a stopping criteria in the algorithms and a method to derive convergence rates.

- Theoretically analyze the optimality of the proposed method by proving that the map $A : \mathcal{W} \to \mathcal{K}$ constructed using our algorithms has the following error bound

$$\|u - A(P_{\mathcal{W}}[u])\| \leq C_A \, \text{dist}(u, \mathcal{K}),$$

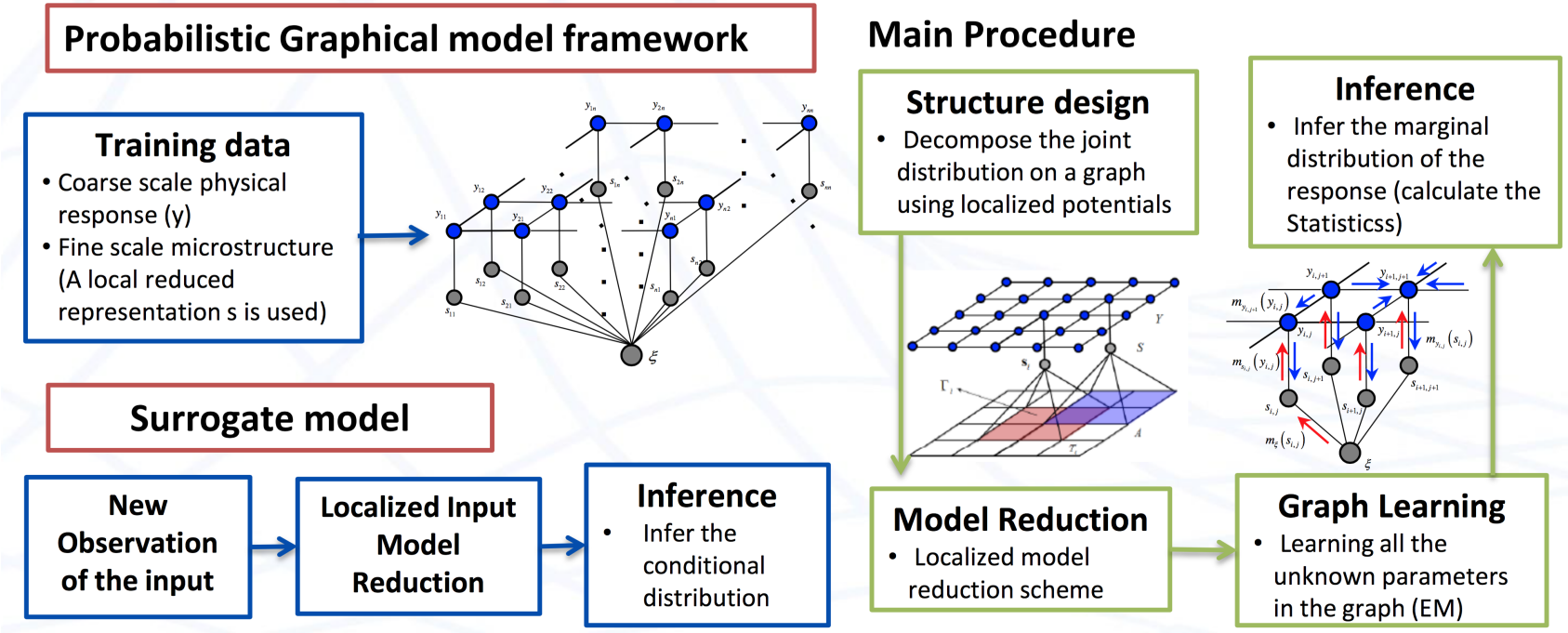
with the *smallest* constant C_A.

Probabilistic graphical model approach

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

Task 2.4:

- Develop probabilistic graphical model framework for recovering the desired posterior distribution or its marginals.
- Develop novel nonparametric belief propagation algorithms by employing loopy belief propagation algorithms.



Task 3

Nonlocal constrained stochastic optimization under uncertainty

Goal: Find the material properties that match the desired material behavior
Find state $u^*(\boldsymbol{y})$ and constitutive relationship $\mathcal{T}^*(\boldsymbol{y})$ that solve

$$\min_{\mathcal{T},u} \mathcal{K}(\mathcal{T},u;g) \qquad \text{subject to (1)}$$

- $\mathcal{K}(\mathcal{T},u;g) = \mathbb{E}[\|u-g\|] + \delta(\mathcal{T}) = \int_{\mathcal{U}} \|u(\boldsymbol{y}) - g(\boldsymbol{y})\|\rho(\boldsymbol{y})d\boldsymbol{y} + \delta(\mathcal{T})$.
- $\delta(\mathcal{T})$ is the "cost" of $\mathcal{T}$, e.g., the cost of manufacturing material with constitutive relationship $\mathcal{T}$.
- $g(\boldsymbol{y})$ is the desired state of the system (target).

Challenge: minimizing the statistics of the difference $\|u-g\|$ does not result in best match between the statistics of $u(\boldsymbol{y})$ and $g(\boldsymbol{y})$.

New optimization paradigm:

$$\min_{\mathcal{T},u} \mathcal{M}(\mathcal{T},u;g) \qquad \text{subject to (1)}$$

where $\mathcal{M}(\mathcal{T},u;g) = \sum_{q=1}^{Q} \|\mathbb{E}[u^q] - \mathbb{E}[g^q]\| + \delta(\mathcal{T})$, and $Q \in \mathbb{N}_+$

Nonlocal constrained stochastic optimization

Stochastic optimization under equality and inequality state constraint

Task 3.1: To address the following fundamental problems

- Analysis of the existence and uniqueness of the optimal solution
- Derive optimality system for $\mathcal{T}^*$ and $u^*(\boldsymbol{y})$
- Develop convergent finite element discretization schemes
- Implementation of the resulting methods for select model problems
- Assess the robustness of the solution, i.e., the sensitivity of $u^*(\boldsymbol{y})$ to perturbations of the constitutive relationship
- Incorporate previously developed methods, for the underlying forward approximation, to develop efficient approaches for stochastic optimization

Risk-averse optimization

Optimization under inequality state constraint

Goal: devise optimization procedure that avoids catastrophic scenarios

Let $F(u)$ be a functional and define undesirable situation by $F(u) \geq \alpha$.

Challenge: directly imposing $-F(u) + \alpha > 0$ is not always feasible

Example: Suppose $F(u)$ indicates damage to the material from impact and $F(u) \geq \alpha$ indicates critical damage that leads to system failure. We want to minimize the expected damage under events appearing with probability $1 - \beta$, i.e., we accept risk β.

Task 3.2: To incorporate risk-averse approach to PDE context

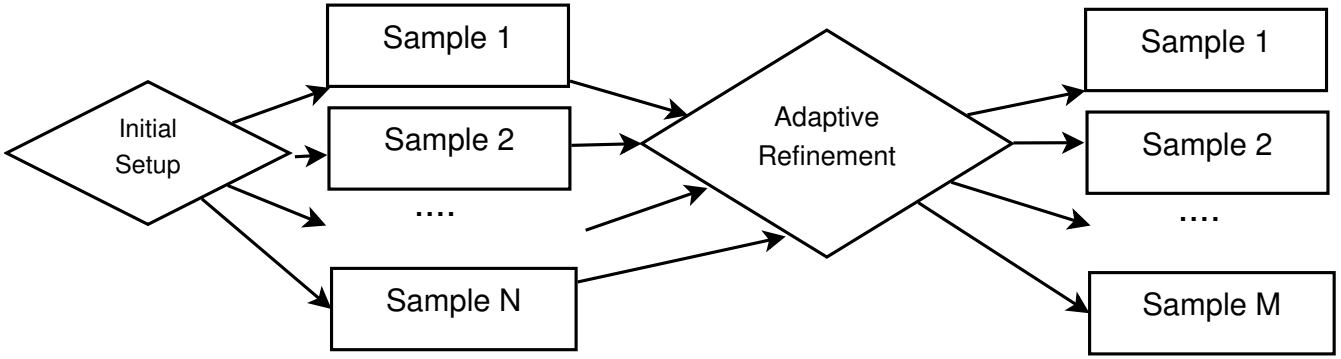
$$\min \mathbb{E}[F(u)] = \min(1-\beta)^{-1} \int_{F(u) \geq \hat{\alpha}_\beta(u)} F(u) \rho d\boldsymbol{y}$$

where $\hat{\alpha}_\beta(u) = \min\{\hat{\alpha} \in \mathbb{R} : \Psi(u, \hat{\alpha}) \geq \beta\}$ and $\Psi(u, \hat{\alpha}) = \int_{F(u) \leq \hat{\alpha}} \rho(\boldsymbol{y}) d\boldsymbol{y}$

Task 4

Scalable sampling paradigm

Goal: develop sampling strategy that exploits extreme parallel architectures



- Recovering the best s-term approximation requires adaptive sampling
- The current paradigm of adaptive sampling employs iterative approach
- The next iteration cannot begin unless the entire current set has been computed
- Multilevel methods, multifidelity models, and random hardware faults result in large variability of runtime associated with each sample
- The current sampling strategies are very sensitive to runtime fluctuations

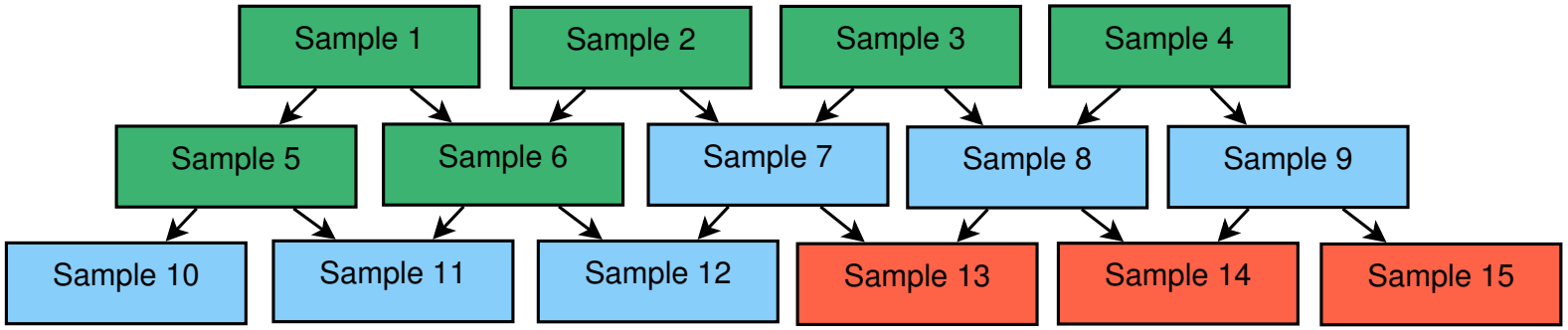
Challenge: Develop a sampling strategy that avoids global iterations and exploits massive parallelism

Scalable sampling paradigm

Utilizing latest advanced in work-flow based on directed acyclic graphs

Task 4.1: Asynchronous sampling procedure

- The directed acyclic graph (DAG) approach does not require a global decision
- DAG-based job scheduling is the current practice for massive parallelism
- Nodes of the graph correspond to samples
- Edges correspond to dependence, i.e., information required to decide whether this sample is needed to construct the best s-term approximation
- The nodes and edges have to be constructed "on-the-fly", unlike standard applications of the DAG approach



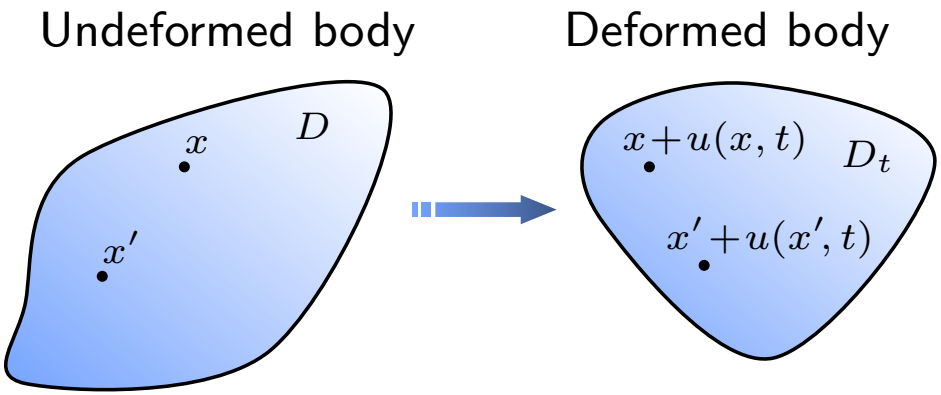
Example: Samples 5 and 6 take less time to compute and samples 8 and 9 take more time, then samples 10 and 11 can start without waiting for the entire previous level.

Nonlocal models for fracture prediction

1. Material deformation

Definitions

- x: reference position
- u: displacement
- $x + u$: current position



$$[x' + u(x',t)] - [x + u(x,t)] = \underbrace{(I + \nabla u(x,t))}_{\text{deformation gradient}} (x' - x) + O(\|x' - x\|^2)$$

Deformation gradient is a local linear approximation of the true deformation

$$\boxed{\text{Nonlocal constitutive models depend on the actual deformation}}$$

Nonlocal models for fracture prediction

2. Connections to classical (local) continuum mechanics

Nonlocal equation of motion

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x,t) = \int_D \left\{\mathcal{T}\left[x,t\right]\left\langle x'-x\right\rangle - \mathcal{T}\left[x',t\right]\left\langle x-x'\right\rangle\right\} dV_{x'} + b(x,t)$$

If: (a) deformation is twice continuously differentiable in space and time
(b) $\mathcal{T}$ is a continuously differentiable function of the deformation and x

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x,t) = \nabla\cdot\nu(x,t) + b(x,t)$$

with the nonlocal stress tensor [Silling, Lehoucq '08]

$$\nu(x,t) = \int_{\mathcal{S}}\int_0^\delta\int_0^\delta (w+z)^2\mathcal{T}\left[x-zm,t\right]\left\langle (w+z)m\right\rangle \otimes m\, dz dw d\Omega_m$$

$$\Big\Downarrow \quad \delta\to 0$$

$$\boxed{\text{Piola-Kirchhoff stress tensor}}$$

Nonlocal models for fracture prediction

3. Connections to higher-order gradient theories

One-dimensional example [Seleson, Parks, Gunzburger, Lehoucq '09]

$$\rho \frac{\partial^2 u}{\partial t^2}(x,t) = \int_{-\delta}^{\delta} \frac{c}{|\xi|}(u(x+\xi,t) - u(x,t))d\xi + b(x,t)$$

$$c = \frac{2K}{\delta^2}$$

If u smooth enough

Many length scales

$$\rho \frac{\partial^2 u}{\partial t^2} = K\left[\frac{\partial^2 u}{\partial x^2} + \frac{\delta^2}{24}\frac{\partial^4 u}{\partial x^4} + \frac{\delta^4}{1080}\frac{\partial^6 u}{\partial x^6} + \dots\right] + b(x,t)$$

$\delta \to 0$

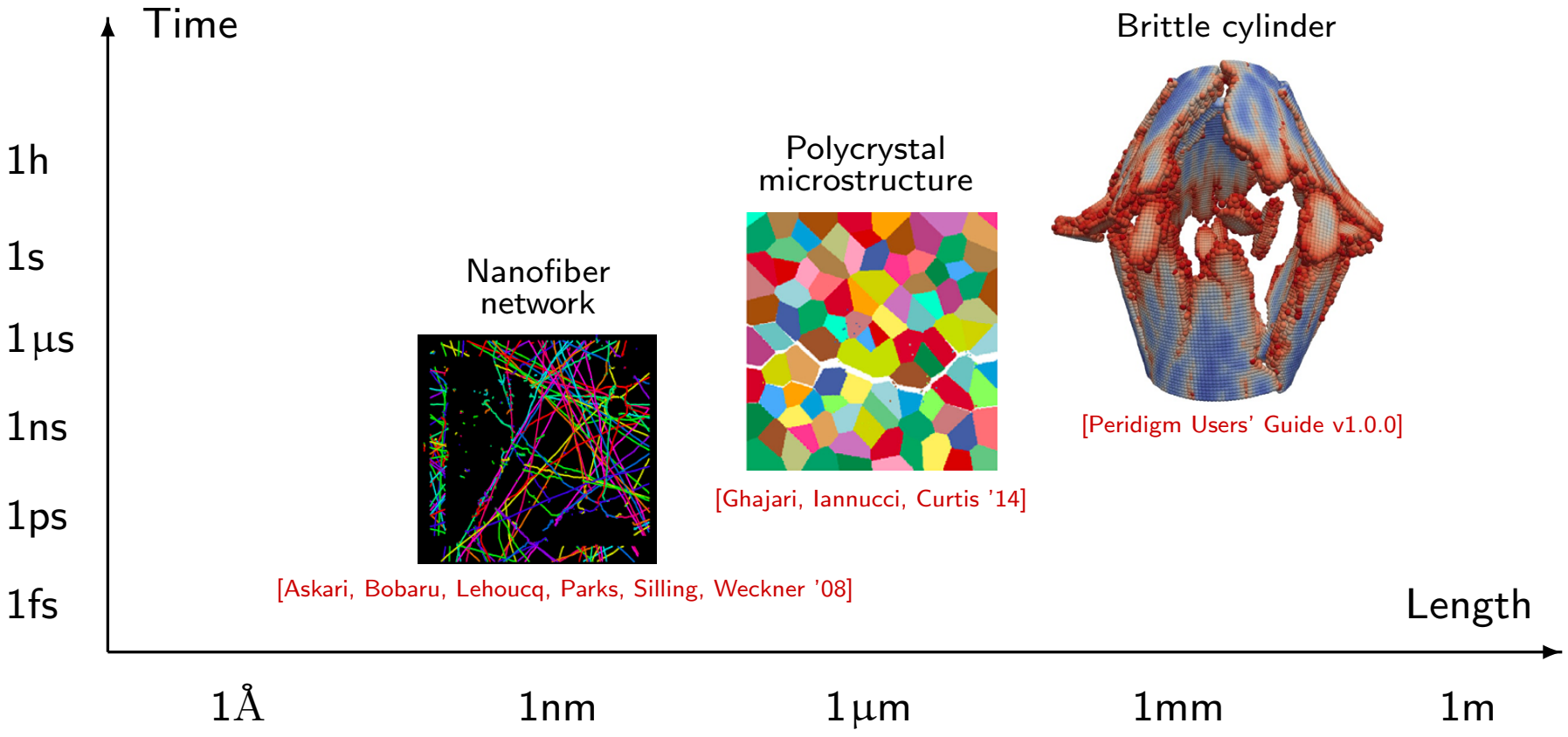
No length scale

$$\rho \frac{\partial^2 u}{\partial t^2}(x,t) = K\frac{\partial^2 u}{\partial x^2}(x,t) + b(x,t)$$

Nonlocal models for fracture prediction

4. Applications

Applications over wide range of length/time scales

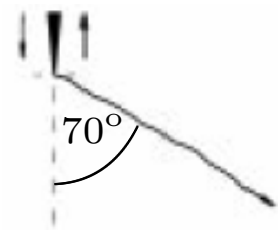
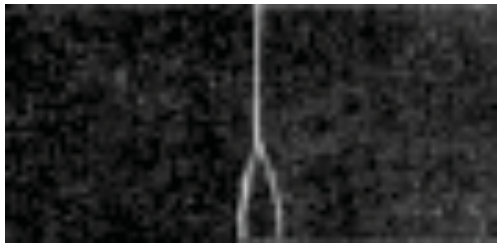
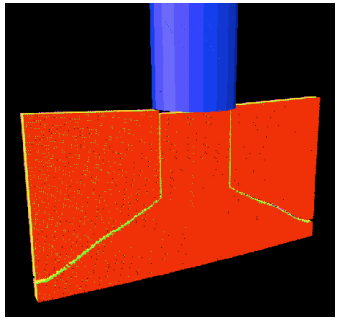
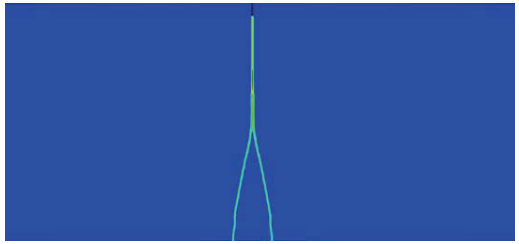


Nonlocal models for fracture prediction

4. Applications

OAK RIDGE National Laboratory
COMPUTATIONAL & APPLIED MATHEMATICS

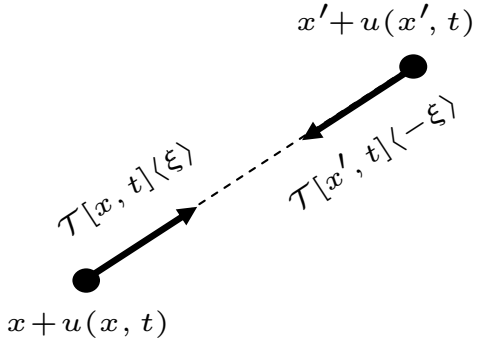
Comparison with experiments

	Fracture in steel (Kalthoff-Winkler)	Crack branching in soda-lime glass	Taylor impact test with 6061-T6 aluminium
Experiment	 [Kalthoff '00]	 [Bowden, Brunton, Field, Heyes '67]	[Anderson, Nicholls, Chocron, Ryckman, '06]
Simulation	 [Silling '03]	 [Simulation by Parks in PDLAMMPS]	[Foster, Silling, Chen, '10]

Nonlocal models for fracture prediction

5. Constitutive models

Ordinary materials [Silling, Epton, Weckner, Xu, Askari '07]



$$\mathcal{T} = \tau M,$$

$$\mathcal{T}[x,t]\langle\xi\rangle = \underbrace{\tau[x,t]\langle\xi\rangle}_{\text{scalar-valued}} \underbrace{\frac{\xi+\eta}{|\xi+\eta|}}_{\text{bond direction}}$$

$\xi = x' - x$: relative position ; $\eta = u(x',t) - u(x,t)$: relative displacement

Example: Elastic Material - Linear Peridynamic Solid

$$\tau[x,t]\langle\xi\rangle = \left[(3K - 5G)\theta[x,t] + 15G\frac{\|\xi+\eta\| - \|\xi\|}{\|\xi\|}\right]\frac{\omega\langle\xi\rangle\,\|\xi\|}{m}$$

K: bulk modulus, G: shear modulus, ω: influence function

Weighted volume

$$m = \int_{\mathcal{H}(0,\delta)} \omega\langle\xi\rangle\,\|\xi\|^2\,dV_\xi$$

Dilatation

$$\theta[x,t] = \frac{3}{m}\int_{\mathcal{H}(x,\delta)} \omega\langle\xi\rangle\,\|\xi\|\,\|\xi+\eta\|\,dV_{x'}$$

Nonlocal models for fracture prediction

5. Constitutive models

Non-ordinary materials [Silling, Epton, Weckner, Xu, Askari '07]

Example: Classical correspondence models

$$\mathcal{T}[x,t]\langle\xi\rangle = \omega\langle\xi\rangle\sigma(\bar{F})(x,t)K^{-1}\xi$$

ω: influence function, σ: classical stress tensor

$\xi = x' - x$: relative position ; $\eta = u(x',t) - u(x,t)$: relative displacement

Shape tensor

$$K = \int_{\mathcal{H}(0,\delta)} \omega\langle\xi\rangle\xi\otimes\xi dV_\xi$$

Approximate deformation gradient

$$\bar{F}(x,t) = \left(\int_{\mathcal{H}(x,\delta)} \omega\langle\xi\rangle(\xi+\eta)\otimes\xi dV_{x'}\right)K^{-1}$$

We can incorporate existing classical complex material models

Nonlocal models for fracture prediction

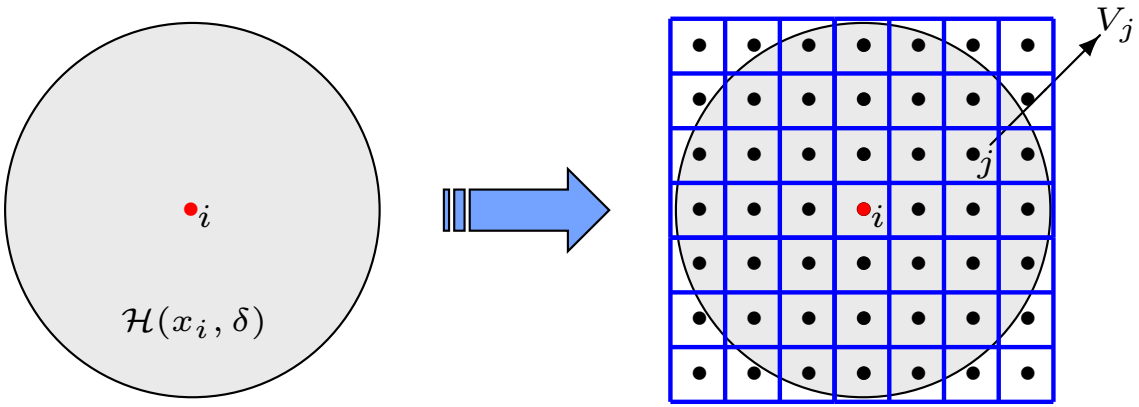
6. Simulation example

Meshfree discretization [Silling, Askari '05]

Given the nonlocal equation of motion

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x,t) = \int_D \left\{\mathcal{T}\left[x,t\right]\left\langle x'-x\right\rangle - \mathcal{T}\left[x',t\right]\left\langle x-x'\right\rangle\right\} dV_{x'} + b(x,t)$$

we discretize the body D with a set of nodes forming a cubic lattice



to get

$$\rho_i \frac{d^2 u_i}{dt^2} = \sum_{j\in\mathcal{F}_i} \left\{\mathcal{T}\left[x_i,t\right]\left\langle x_j - x_i\right\rangle - \mathcal{T}\left[x_j,t\right]\left\langle x_i - x_j\right\rangle\right\} V_j + b_i$$

$$\mathcal{F}_i = \{j \ : \ \|x_j - x_i\| \le \delta,\ j \ne i\}$$

Nonlocal models for fracture prediction

6. Simulation example

Criterion for Material Failure [Silling, Askari '05]

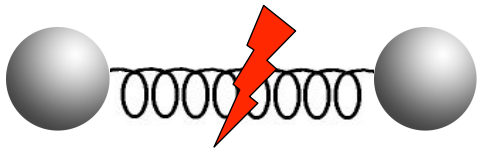
Let the stretch be

$$s(x, x', t) = (\|\xi + \eta\| - \|\xi\|) / \|\xi\|$$

$\xi = x' - x \quad ; \quad \eta = u(x', t) - u(x, t)$

Breaking bonds

$$\mu(x, x', t) = \begin{cases} 1 & s(x, x', \tilde{t}) < s_0 \text{ for all } 0 \le \tilde{t} \le t \\ 0 & \text{otherwise} \end{cases}$$



Example: Prototype Microelastic Brittle (PMB) Model

$$\mathcal{T}[x, t]\langle\xi\rangle = \frac{1}{2} \underbrace{\mu(x, x', t)}_{\text{Boolean variable}} \underbrace{s(x, x', t)}_{\text{Force magnitude}} \underbrace{\frac{\xi + \eta}{\|\xi + \eta\|}}_{\text{Force direction}}$$

Damage

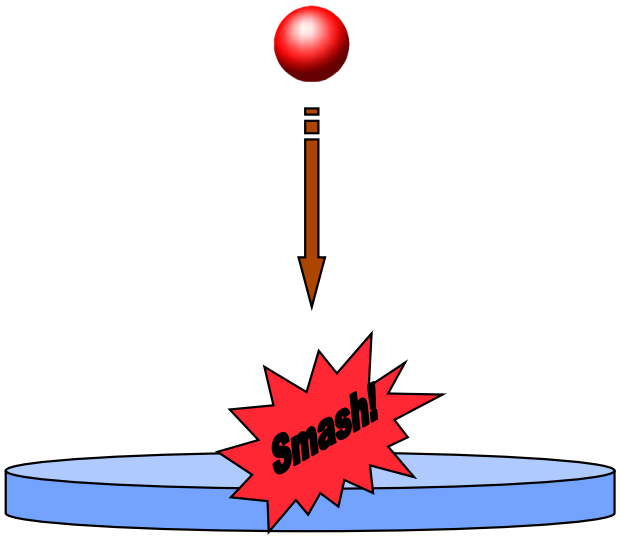
$$\varphi(x, t) := 1 - \frac{\int_{\mathcal{H}(x,\delta)} \mu(x, x', t) dV_{x'}}{\int_{\mathcal{H}(x,\delta)} dV_{x'}} \quad ; \quad 0 \le \varphi(x, t) \le 1$$

$\varphi(x, t) = 0 \rightarrow$ pristine material ; $\boxed{\text{maximum damage is one}}$

Nonlocal models for fracture prediction

6. Simulation example

Example: 3D fracture dynamics



• Projectile	• Target
- Sphere: diameter 14.80 mm	- diameter 74 mm - thickness 1.25 mm
- Mesh spacing ∼0.5 mm - Horizon 3.1 mm - 8,497 particles	- Mesh spacing ∼1 mm - Horizon 3.1 mm - 22,953 particles
- Bulk modulus 160.90 GPa - Shear modulus 78.30 GPa - Density 7700 kg/m^3	- Bulk modulus 14.90 GPa - Shear modulus 8.94 GPa - Density 2200 kg/m^3

6 month deliverables

For $0 < \alpha < d_x + 2$, we intend to start with a random field $c(x, x', t, \boldsymbol{y})$:

$$\mathcal{T}\left[x, t, \boldsymbol{y}\right]\langle\xi\rangle = \frac{c(x, x', t, \boldsymbol{y})}{\|\xi\|^{\alpha}}(\xi \otimes \xi)\,\eta, \quad \text{with } \xi = x' - x, \quad \eta = u(x', t, \boldsymbol{y}) - u(x, t, \boldsymbol{y})$$

1. Develop the theory for the regularity of the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,, \boldsymbol{y})$ of (1) given the constitutive relationship above. Develop *a priori* and *a posteriori* estimates for bound on the polynomial coefficients $c_{\boldsymbol{\nu}}$.
2. Computationally recover the multi-dimensional solution $u(\,\cdot\,, \boldsymbol{y})$, using best s-term interpolation and weighted ℓ_1 optimization for problems with $d \leq 200$.
   - Initial DAG implementation on TITAN at ORNL.
3. As the number of uncertainty variables increases, we will theoretical and numerically study the influence of the stochasticity in the material properties on material response.
4. Develop derivative-free approximations of likelihood-informed subspaces, and use this to develop efficient MCMC sampling algorithms.

Risks and mitigation plans

Risk: We expect that the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,,\boldsymbol{y})$ will maintain some regularity but it is conceivable that the strong nonlinear phenomena will induce irregularities.

Mitigation: To develop nonlinear reconstruction methods when $\boldsymbol{y} \mapsto u(\,\cdot\,,\boldsymbol{y})$ becomes irregular, which will exploit local variations in smoothness of u.

- Extend the best s-term approximation, from a sparse linear space (projection) to include a dictionary [DeVore '98].
- Employ suitable versions of *nonlinear N-widths* to construct adaptive partitioning of the parameter domain as to balance out the variations of local smoothness.

Risk: The quality of the experimental data is not sufficiently accurate or noisy, so that the developed coarse-grained stochastic nonlocal models cannot be properly validated.

Mitigation:

- De-noise the data with filtering and compressed sensing techniques, and exploit advanced design of experiments techniques to use both legacy data as well as conduct a new set of improved experiments to generate satisfactory data.
- Possibly generate data from expensive molecular dynamics simulations.

Risks and mitigation plans

Risk: We expect that the solution map $\boldsymbol{y} \mapsto u(\,\cdot\,,\boldsymbol{y})$ will maintain some regularity but it is conceivable that the strong nonlinear phenomena will induce irregularities.

Mitigation: To develop nonlinear reconstruction methods when $\boldsymbol{y} \mapsto u(\,\cdot\,,\boldsymbol{y})$ becomes irregular, which will exploit local variations in smoothness of u.

- Extend the best s-term approximation, from a sparse linear space (projection) to include a dictionary [DeVore '98].
- Employ suitable versions of *nonlinear N-widths* to construct adaptive partitioning of the parameter domain as to balance out the variations of local smoothness.

Risk: The quality of the experimental data is not sufficiently accurate or noisy, so that the developed coarse-grained stochastic nonlocal models cannot be properly validated.

Mitigation:

- De-noise the data with filtering and compressed sensing techniques, and exploit advanced design of experiments techniques to use both legacy data as well as conduct a new set of improved experiments to generate satisfactory data.
- Possibly generate data from expensive molecular dynamics simulations.

Risks and mitigation plans

Risk: Maintaining focus can be challenging due to the diversity of expertise (mathematics, computational and material science, engineering, and extreme-scale computing), as well as a diversity of institutions and locations (universities and labs in both US and Europe).

Mitigation:

- Centrally locating the administrative tasks at ORNL will leverage the lab's considerable resources for, and experience with, managing large-scale projects, allowing the scientists to spend the majority of their time conducting research.
- Currently scheduling monthly teleconferences with the entire team (of PIs) and bi-weekly meetings with the ORNL mathematicians and material scientists.