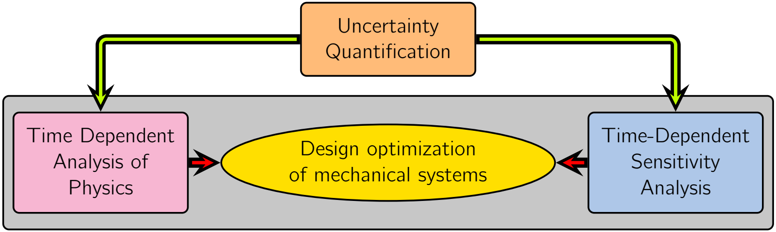
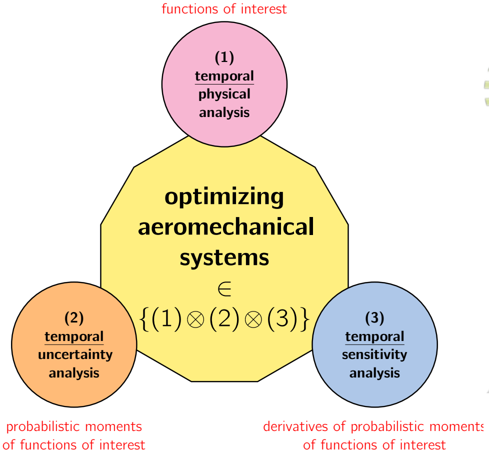
Semi-Intrusive Uncertainty Propagation and Adjoint Sensitivity Analysis Using the Stochastic Galerkin Method

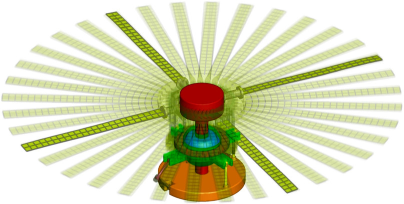
Reusing Deterministic Finite Element and Adjoint Implementations for Stochastic Galerkin Computations



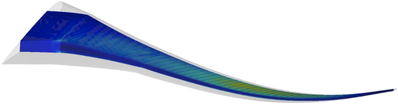
Komahan Boopathy & **Graeme Kennedy**
Georgia Institute of Technology

AIAA SciTech 2020 – Orlando, FL – January 8, 2020





flexible multibody dynamics

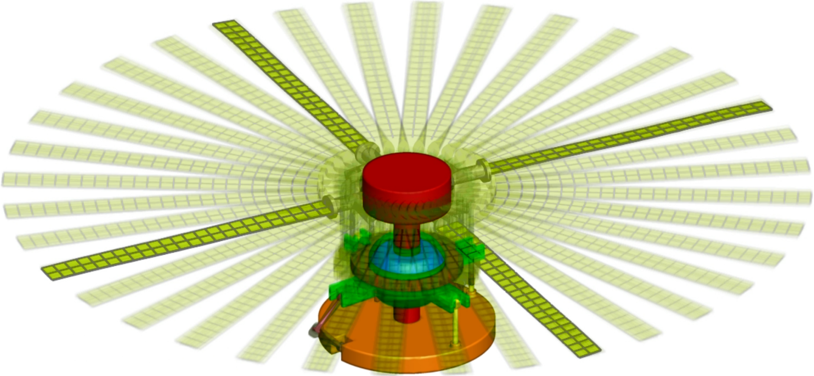


wing design

"Parallel Finite Element Framework for Rotorcraft Multibody Dynamics and Adjoint Sensitivities", AIAA Journal, Vol. 57, No. 8, pp. 3159–3172, 2019.

Uncertainty Propagation through Complex Time Dependent Models

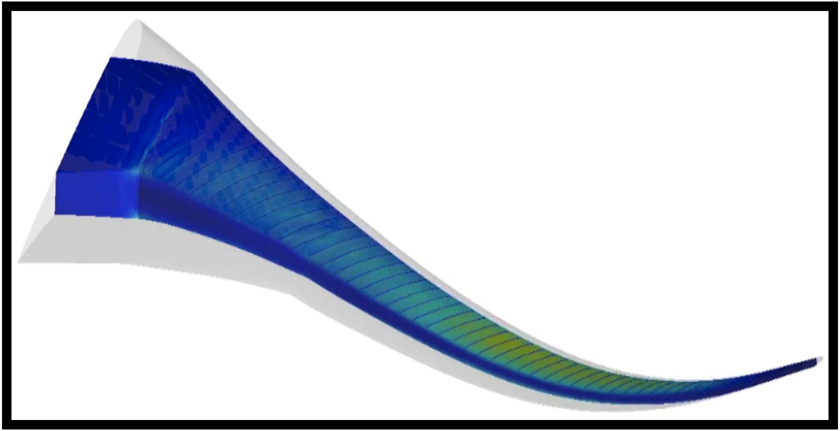
collective blade pitch



❶ deterministic physical states . . . . stochastic physical states
❷ deterministic adjoint states . . . . stochastic adjoint states

Uncertainty Propagation through Complex Time Dependent Models

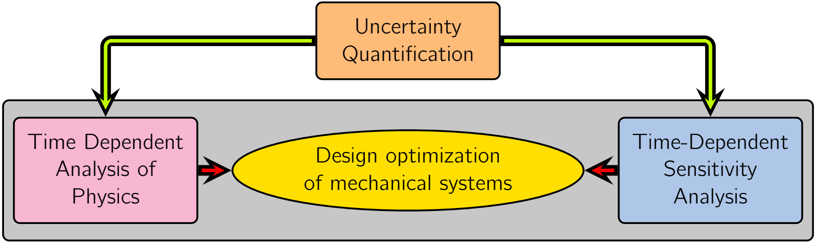
uCRM wing subject to gust loads



❶ deterministic physical states ....stochastic physical states

❷ deterministic adjoint states ....stochastic adjoint states

Scope of the work



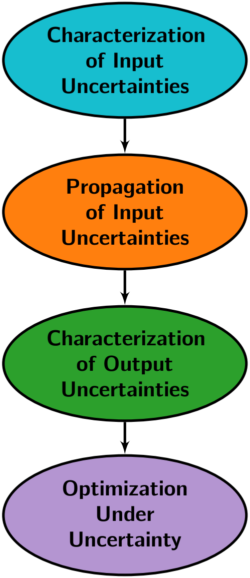
1. interested in design of time dependent mechanical systems using adjoint derivatives
2. we include **probabilistically modeled uncertainties**
3. propagate the uncertainties using **stochastic Galerkin projection method**
4. perform stochastic Galerkin computations by **reusing deterministic finite element and adjoint code implementation**
5. perform **optimization under uncertainty**

Systematic treatment of uncertainties for design optimization

> *A product should be designed in such a way that makes its* **performance insensitive to variation** *in variables beyond the control of the designer* — *Genichi Taguchi*

... we need a systematic process to achieve **robustness**, **reliability** and **optimality** of design

Systematic treatment of uncertainties for design optimization



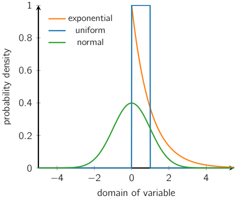
1. quantify input randomness
2. uncertainties through PDE models of physics
3. quantify output randomness

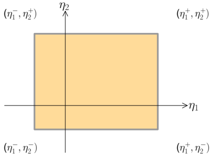
✈ formulate optimization problem statements

Stages I and III – Characterization of Input and Output Randomness

Characterization of Input Uncertainties

probabilistic ← → nonprobabilistic





↓

Propagation of Input Uncertainties

↓

Characterization of Output Uncertainties

probabilistic ← → nonprobabilistic

▲
▲
▲
current approach

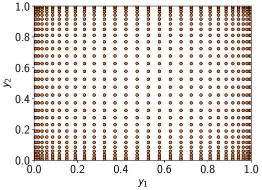
↓

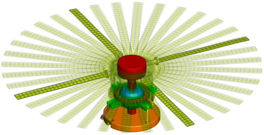
Optimization Under Uncertainty

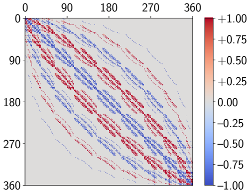
Stage II – Methods for Uncertainty Propagation

Propagation of Input Uncertainties

nonintrusive | intrusive







quadrature samples | projected Jacobian

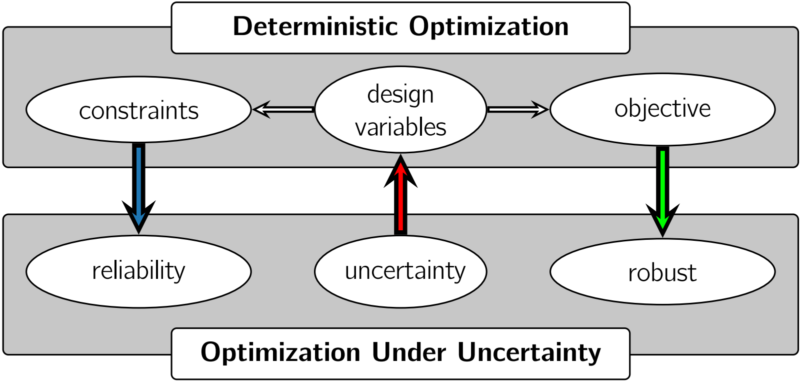
repeated solutions smaller system	principle	one solution of bigger nonlinear system
no modifications (black box)	code	requires modifications

how to reuse deterministic FEA and adjoint code for projection?

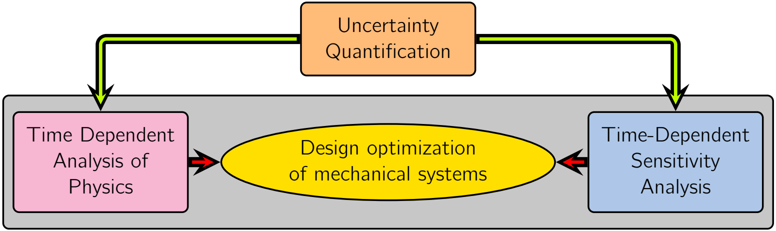
Stage IV – Optimization Under Uncertainty

Deterministic Optimization Problem

$$\begin{aligned} \underset{\xi}{\text{minimize}} \quad & F(\xi) \\ \text{subject to} \quad & G(\xi) \leq 0 \end{aligned}$$



Stage IV – Optimization Under Uncertainty



Optimization Under Uncertainty Problem

$$\underset{\xi}{\text{minimize}} \quad (1-\alpha)\cdot \mathbb{E}\left[F(\xi)\right] + \alpha \cdot \mathbb{S}\left[F(\xi)\right]$$

$$\text{subject to} \quad \mathbb{E}\left[G(\xi)\right] + \beta \cdot \mathbb{S}\left[G(\xi)\right] \leq 0$$

- α – objective robustness, β – constraint reliability
- need derivatives to solve gradient based optimization problem

Section 2

Semi-Intrusive Stochastic Galerkin Projection

Essentials for SGM from a Computational Standpoint

① **Forward Physical Analysis** (moments of functions)

- ✔ form stochastic states from deterministic states
- ✔ form stochastic residuals from deterministic residuals
- ✔ form stochastic Jacobians from deterministic Jacobians
- form stochastic initial & boundary conditions from deterministic initial (boundary) conditions

② **Adjoint Sensitivity Analysis** (deriv. of moments of functions)

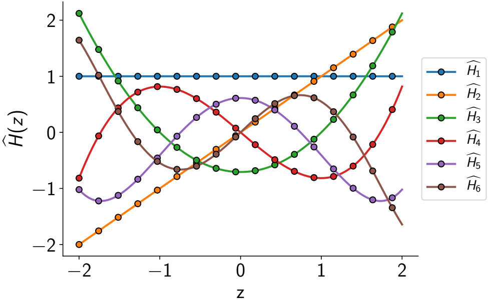
- ✔ form stochastic adjoint states from deterministic adjoint states
- form stochastic adjoint right hand terms from deterministic adjoint right hand terms
- form stochastic (transposed) Jacobians from deterministic (transposed) Jacobians?

Probabilistic Space and Inner Product

- probabilistic function space is approximated with N basis entries

$$\mathcal{Y} \approx \text{span}\{\widehat{\psi}_1(y), \widehat{\psi}_2(y), \ldots, \widehat{\psi}_N(y)\}$$

- polynomial type based on the probability distribution type
  - Hermite, Legendre, Laguerre
  - Normal, Uniform, Exponential



- orthogonality + normality
- tensor product for multivariate basis

Formation of Stochastic Physical States

The stochastic state vector is

$$u(t, y) \approx \sum_{i=1}^{N} U_i(t)\widehat{\psi}_i(y)$$

Core principles at play:

1. principle of variable separation – time and probabilistic domains
2. principle of superposition .... summation

- the state vector coefficients: $U(t) = [U_1(t), U_2(t), \ldots, U_N(t)]$ are available as guessed values from iterative solution
- the length of stochastic state vector is N times the length of deterministic state vector
- time derivatives $\dot{u}(t, y)$ and $\ddot{u}(t, y)$ are appoximated similarly

Formation of Stochastic Residual

quadrature loop

$$\mathcal{Y} \approx \text{span}\{\widehat{\psi}_1(y), \widehat{\psi}_2(y), \ldots, \widehat{\psi}_N(y)\}$$

$$\mathcal{R} = \begin{bmatrix} \mathcal{R}_1 \\ \mathcal{R}_2 \\ \vdots \\ \mathcal{R}_N \end{bmatrix} \qquad \mathcal{R}_i \approx \sum_{q=1}^{Q} \underbrace{\alpha_q \widehat{\psi}_i(y_q)}_{\text{scalar}} \times \underbrace{R(t, y_q, u(t, y_q), \dot{u}(t, y_q), \ddot{u}(t, y_q))}_{\text{deterministic residual for } y_q}$$

- quadrature over deterministic residual implementations
- the length of stochastic residual vector is N times the length of deterministic residual vector
- need ability to update elements with new parameter values
- residuals can be fully assembled or elementwise ones

Formation of Stochastic Jacobian

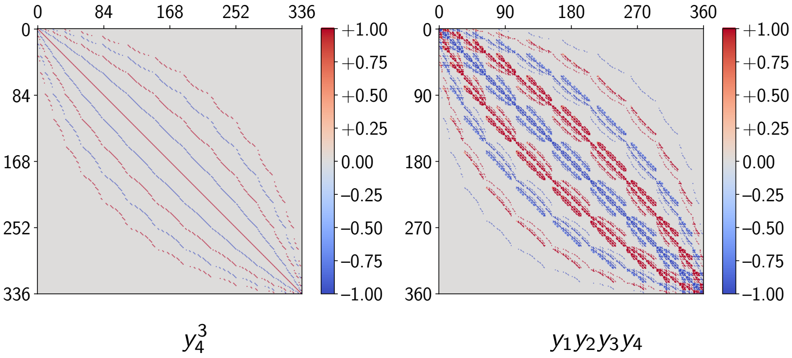
The stochastic Jacobian matrix is

$$\mathcal{J} = \begin{bmatrix} \mathcal{J}_{1,1} & \mathcal{J}_{1,2} & \cdots & \mathcal{J}_{1,N} \\ \mathcal{J}_{2\,1} & \mathcal{J}_{2,2} & \cdots & \mathcal{J}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{J}_{N,1} & \mathcal{J}_{N,2} & \cdots & \mathcal{J}_{N,N} \end{bmatrix}.$$

$$\mathcal{J}_{i,j} \approx \sum_{q=1}^{Q} \underbrace{\alpha_q \widehat{\psi}_i(y_q)\widehat{\psi}_j(y_q)}_{\text{scalar}} \times \underbrace{J(t, y_q, u(t, y_q), \dot{u}(t, y_q), \ddot{u}(t, y_q))}_{\text{deterministic Jacobian for } y_q}$$

- quadrature over deterministic jacobian implementations
- need ability to update element with new parameter values
- the size of stochastic Jacobian is N times the size of deterministic Jacobian
- applies to assembled and elementwise Jacobians

Formation of Stochastic Jacobian



- sparsity patterns depend on the nonlinearity parameter y
- can optimize the number of quadrature evaluations
- can determine the sparsity apriori

Formation of Stochastic Adjoint States

recall... stochastic physical state vector

$$U(t) = [U_1(t), U_2(t), \dots, U_N(t)]$$

$$u(t, y) \approx \sum_{i=1}^{N} U_i(t)\widehat{\psi}_i(y)$$

Stochastic Adjoint state vector is formed in a similar manner

$$\Lambda(t) = [\Lambda_1(t), \Lambda_2(t), \dots, \Lambda_N(t)]$$

$$\lambda(t, y) \approx \sum_{i=1}^{N} \Lambda_i(t)\widehat{\psi}_i(y)$$

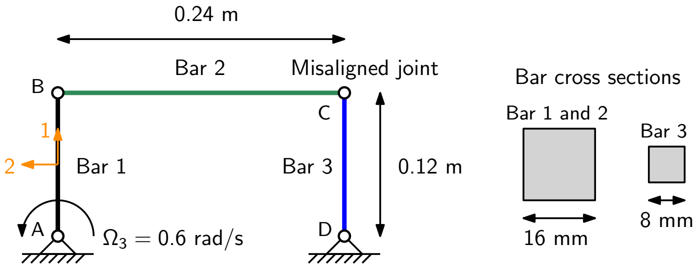
- ✔ transposed Jacobian matrix is formed similar to forward solve
- ✔ the right hand sides of the adjoint linear system are formed in a manner similar to residuals

Section 3

Four Bar Mechanism Benchmark

Four Bar Mechanism – Problem Setup

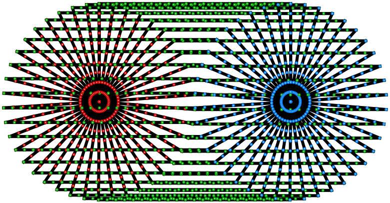
standard finite-element benchmark case

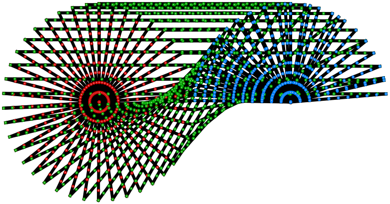


- a time-dependent benchmark problem
- ① three Timoshenko bars, ② an actuator driving the mechanism, ③ three revolute constraints
- the revolute joint at C is misaligned by an angle of 5°

Four Bar Mechanism – Problem Physics

The motion of mechanism differs considerably with the alignment angle of Joint C

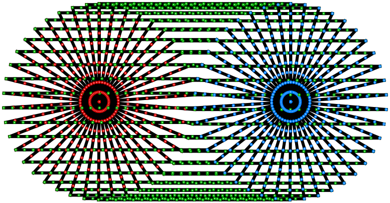

no misalignment


5° misalignment of joint

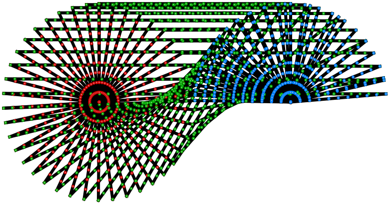
✈ we treat the misalignment angle as a source of uncertainty and model it as a random variable $\theta(y) \sim \mathcal{N}(\mu = 5^\circ, \sigma = 2.5^\circ)$

Four Bar Mechanism – Problem Physics

The motion of mechanism differs considerably with the alignment angle of Joint C



no misalignment

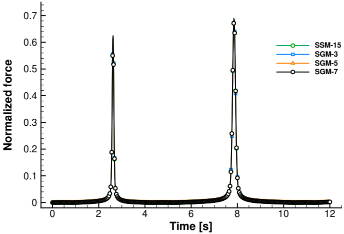


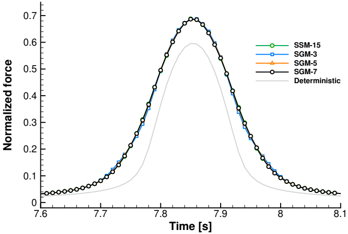
5° misalignment of joint

- we treat the misalignment angle as a source of uncertainty and model it as a random variable $\theta(y) \sim \mathcal{N}(\mu = 5^\circ, \sigma = 2.5^\circ)$

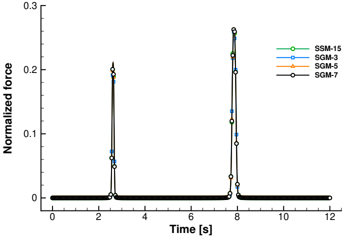
Four Bar Mechanism – Probabilistic Moments Verification

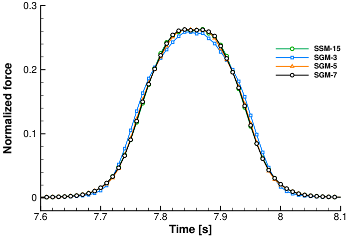
mean of normalized axial force in bar *AB* as a function of time



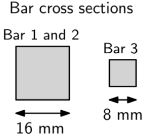


variance of normalized axial force in bar *AB* as a function of time





Optimization Problem Formulation



OUU with Four Bar Mechanism

minimize	$\mathbb{E}$[mass]
design variable	width of bars
random parameter	revolute axis $\theta \sim \mathcal{N}(\mu = 5^\circ, \sigma = 2.5^\circ)$
subject to	$\mathbb{E}$[failure index] $+ \beta \cdot \mathbb{S}$[failure index] ≤ 1
	$\mathbb{E}$[displacement] $+ \beta \cdot \mathbb{S}$[displacement] $\leq 5mm$
bounds	$5mm \leq$ width $\leq 25mm$

- Deterministic optimization with $\theta = 5^\circ$
- OUU with $\beta = 0, 1, 2, 3$

Gradient Verification

derivatives of probabilistic moments of functions of interest

quantity	mass	failure	displacement
adjoint $d\mathbb{E}[F]/d\xi$	1078.2720000000...	22.5748...	-0.1067834...
complex $d\mathbb{E}[F]/d\xi$	1078.2720000000...	22.5748...	-0.1067834...
error	4.5×10^{-11}	8.7×10^{-6}	5.5×10^{-8}
adjoint $d\mathbb{V}[F]/d\xi$	–	22.45792...	$-1.57732\ldots \times 10^{-4}$
complex $d\mathbb{V}[F]/d\xi$	–	22.45792...	$-1.57732\ldots \times 10^{-4}$
error	–	6.7×10^{-6}	1.1×10^{-10}

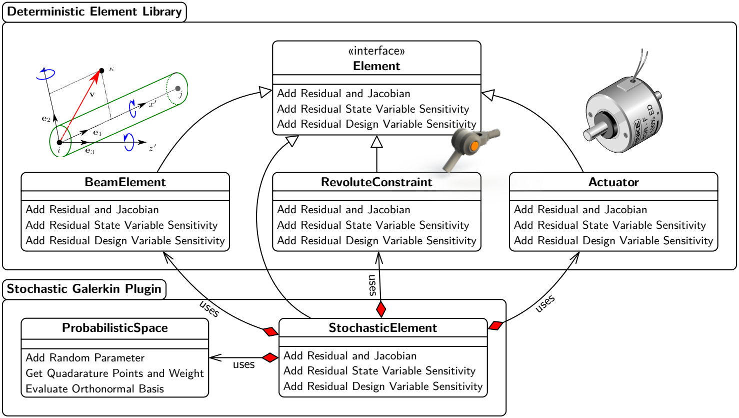
- ✔ used complex-step method to verify the consistency of adjoint derivatives
- ✔ no variance derivative for mass due to the choice of random parameter

Optimization Results

Quantity	deterministic	$\beta = 0$	$\beta = 1$	$\beta = 2$	$\beta = 3$
width 1	5.0	5.0	13.0	18.3	23.5
width 2	5.0	5.0	5.0	6.0	7.0
mass	1.1	1.1	6.02	11.7	19.3
failure constraint	47%	55%	90%	100%	100%
displacement constraint	78.22%	78.72%	100%	100%	100%

- deterministic optimum is at the lower bounds of variables
- OUU solutions are in the interior of design space and near the constraint manifolds
- mass reduction by thinning of bar 3

Software Architecture for Stochastic Galerkin Method



- ✔ `StochasticElement` –
  - **is an** `Element` by inheritance and also
  - **has an** *element* (deterministic) by composition
- ✔ `ProbabilisticSpace` – prob. quadrature and basis evaluations

Section 4

Conclusions

Summary

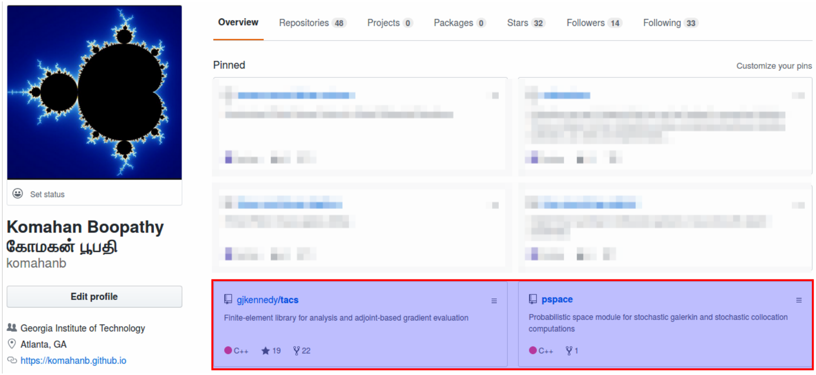
Semi-Intrusive Galerkin Projection Technique

- ✔ reuse determistic element residuals, Jacobians like black-box
- ✔ time dependent discrete adjoint sensitivities for gradient-based optimization
- ✔ implicit formation of stochastic governing equations from deterministic governing equations

Applications

- ✔ extended finite element library TACS to perform stochastic Galerkin computations
- ✔ verified the moments and derivatives using stochastic sampling and complex-step techniques
- ✔ demonstrated on four bar mechanism problem and simple dynamical systems

Questions?



1. **TACS** ..... finite element library, flexible multibody dynamics, temporal adjoint (handles spatial, temporal and design parameter domains)
2. **PSPACE** ...... library for probabilistic quadrature and basis evaluation (handles probabilistic domain)

Section 5

Backup Slides

Derivatives of Expectation of Functions of Interest

$$\frac{d\mathbb{E}[F(y,u(y),\lambda(y))]}{d\xi} = \left\langle \widehat{\psi}_1(y) \;\middle|\; \frac{dF(y,u(y),\lambda(y))}{d\xi} \right\rangle_{\rho(y)}^{\mathcal{Y}}$$

$$\approx \sum_{q=1}^{Q} \alpha_q \times \widehat{\psi}_1(y_q) \times \underbrace{\frac{dF(y,u(y_q),\lambda(y_q))}{d\xi}}_{\text{deterministic adjoint for } y_q}$$

- ✔ quadrature over the deterministic adjoint code to compute mean derivative
- ✔ variance derivative can be computed similarly using

$$\frac{d\mathbb{E}[F(y,u(y),\lambda_F(y))]}{d\xi} \text{ and } \frac{d\mathbb{E}[F^2(y,u(y),\lambda_{F^2}(y))]}{d\xi}$$

The Big Picture of Uncertainty Propagation through Physical Models

Nonlinear Algebraic System

$\underset{u(\xi)}{\text{solve}}$ $R(\xi, u(\xi)) = 0$

evaluate $F(\xi, u(\xi))$

↓

Nonlinear ODE System

$\underset{u(t,\xi),\dot{u}(t,\xi),\ddot{u}(t,\xi)}{\text{solve}}$ $R(t, \xi, u(t, \xi), \dot{u}(t, \xi), \ddot{u}(t, \xi)) = 0$

evaluate $F(t, \xi, u(t, \xi), \dot{u}(t, \xi), \ddot{u}(t, \xi))$

↓

Nonlinear Stochastic ODE System

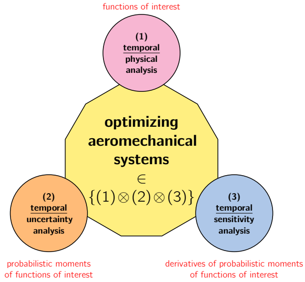
$\underset{u(t,y(\xi)),\dot{u}(t,y(\xi)),\ddot{u}(t,y(\xi))}{\text{solve}}$ $R(t, y(\xi), u(t, y(\xi)), \dot{u}(t, y(\xi)), \ddot{u}(t, y(\xi))) = 0$

evaluate

$\mathbb{E}\left[F(t, y(\xi), u(t, y(\xi)), \dot{u}(t, y(\xi)), \ddot{u}(t, y(\xi)))\right]$

$\mathbb{V}\left[F(t, y(\xi), u(t, y(\xi)), \dot{u}(t, y(\xi)), \ddot{u}(t, y(\xi)))\right]$

$\mathbb{S}\left[F(t, y(\xi), u(t, y(\xi)), \dot{u}(t, y(\xi)), \ddot{u}(t, y(\xi)))\right]$



The Big Picture of Sensitivity Analysis Problem

Linear Algebraic System

define $\mathcal{L}(\circ, \lambda(\xi)) = R(\circ) + \lambda(\xi) F(\circ)$

where $\circ = \xi, u(\xi)$

$\underset{\lambda(\xi)}{\text{solve}}$ $\dfrac{\partial \mathcal{L}(\circ, \lambda(\xi))}{\partial u} = 0$

evaluate $\dfrac{dF(\circ, \lambda(\xi))}{d\xi} = \dfrac{\partial F(\circ)}{\partial \xi} + \lambda(\xi) \dfrac{\partial R(\circ)}{\partial \xi}$

Linear ODE System

define $\mathcal{L}(\circ, \lambda(t,\xi)) = R(\circ) + \lambda(t,\xi) F(\circ)$

where $\circ = t, \xi, u(t,\xi), \dot{u}(t,\xi), \ddot{u}(t,\xi)$

$\underset{\lambda(t,\xi)}{\text{solve}}$ $\dfrac{\partial \mathcal{L}(\circ, \lambda(t,\xi))}{\partial u} - \dfrac{d}{dt}\left(\dfrac{\partial \mathcal{L}(\circ, \lambda(t,\xi))}{\partial \dot{u}}\right) + \dfrac{d^2}{dt^2}\left(\dfrac{\partial \mathcal{L}(\circ, \lambda(t,\xi))}{\partial \ddot{u}}\right) = 0$

evaluate $\dfrac{dF(\circ, \lambda(t,\xi))}{d\xi} = \dfrac{\partial F(\circ)}{\partial \xi} + \lambda(t,\xi) \dfrac{\partial R(\circ)}{\partial \xi}$

Linear Stochastic ODE System

define $\mathcal{L}(\circ, \lambda(t,y(\xi))) = R(\circ) + \lambda(t,y(\xi)) F(\circ)$

where $\circ = t, y(\xi), u(t,y(\xi)), \dot{u}(t,y(\xi)), \ddot{u}(t,y(\xi))$

$\underset{\lambda(t,y(\xi))}{\text{solve}}$ $\dfrac{\partial \mathcal{L}(\circ, \lambda(t,y(\xi)))}{\partial u} - \dfrac{d}{dt}\left(\dfrac{\partial \mathcal{L}(\circ, \lambda(t,y(\xi)))}{\partial \dot{u}}\right) + \dfrac{d^2}{dt^2}\left(\dfrac{\partial \mathcal{L}(\circ, \lambda(t,y(\xi)))}{\partial \ddot{u}}\right) = 0$

evaluate $\dfrac{d\mathbb{E}\left[F(\circ, \lambda(t,y(\xi)))\right]}{d\xi}, \dfrac{d\mathbb{V}\left[F(\circ, \lambda(t,y(\xi)))\right]}{d\xi}, \dfrac{d\mathbb{S}\left[F(\circ, \lambda(t,y(\xi)))\right]}{d\xi}$