50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference
17th
4 - 7 May 2009, Palm Springs, California

AIAA 2009-2277

Risk Quantification in Unsteady Flow Simulations using Adjoint-based Approaches

Qiqi Wang*, Framk Ham, Gianluca Iaccarino and Parviz Moin

Center for Turbulence Research, Stanford University, Stanford, CA 94305, United States

In this paper, we study the unsteady flow past a circular cylinder subject to rotational oscillations. The uncertainties induced by the unknown frequency, magnitude and phase with respect to the vortex shedding are treated as sources of stochasticity, and we consider the problem of quantifying the tail probability distribution of the aerodynamic forces. Specifically we estimate the probability that the overall time-averaged drag exceeds a given critical value. First we compare adjoint-based and Monte Carlo based estimates. Then, we demonstrate that using an adjoint method, we can design a better statistical estimator and an importance sampling strategy and obtain accurate predictions with a computational cost orders of magnitude lower than brute force Monte-Carlo method.

I. Introduction

Discussed in this article is a method for failure and risk analysis. These types of analyses require calculating tail probabilities of an objective function. More specifically, we want to calculate the small probability that the objective function exceeds a certain critical value, usually representing the probability that a system fails or suffers catastrophic losses. Mathematically, let $\mathbf{J}(\xi)$ be the objective function, which depends on a vector of uncertain variables ξ. The probability distribution of ξ is known. Let $\mathbf{J_C}$ be a known constant. We want to calculate $P(\mathbf{J} > \mathbf{J_C})$. This probability can be as high as 10%, and as small as 10^{-5} or even smaller.

Calculating such tail probabilities is challenging. Polynomial chaos and collocation based methods do not represent tail probabilities accurately. Monte Carlo is the most commonly used method in this situation. However, the brute force Monte Carlo method can be very computationally expensive and inefficient; if the tail probability is small, only a small fraction of the samples would fall into the tail region, resulting in insufficient sampling. In fact, let $\xi_i, i = 1, \ldots, N$ be samples of ξ. The brute force Monte Carlo method approximates

$$P(\mathbf{J} > \mathbf{J_C}) \approx P_N = \frac{1}{N} \sum_{i=1}^{N} I(\mathbf{J}(\xi_i) > \mathbf{J_C}) \, ,$$

where $I(\mathbf{J}(\xi_i) > \mathbf{J_C})$ is the indicator function. Its value is 1 if $\mathbf{J}(\xi_i) > \mathbf{J_C}$ is true and zero otherwise. The variance of this estimator P_N is

$$\mathrm{Var}[P_N] = \frac{1}{N} \mathrm{Var}[I(\mathbf{J}(\xi_i) > \mathbf{J_C})] = \frac{P(\mathbf{J} > \mathbf{J_C}) - P(\mathbf{J} > \mathbf{J_C})^2}{N} \, . \tag{1}$$

The relative error of the Monte Carlo method can be characterized by the ratio of the standard deviation and the mean of the estimator P_N, which is

$$\frac{\sqrt{\mathrm{Var}[P_N]}}{P(\mathbf{J} > \mathbf{J_C})} = \sqrt{\frac{1 - P(\mathbf{J} > \mathbf{J_C})}{N \, P(\mathbf{J} > \mathbf{J_C})}}.$$

*Corresponding author.

This formula reveals the inefficiency of the brute force Monte Carlo method when the tail probability is small. For a fixed number of samples, the smaller the tail probability $P(\mathbf{J} > \mathbf{J_C})$, the larger the relative error; for a fixed target relative error, the smaller the tail probability, the more samples must be used to meet the target relative error. For example, if the tail probability $P(\mathbf{J} > \mathbf{J_C}) = 5\%$, and the target relative error is 10%, about 2,000 samples should be used. But when the tail probability to be calculated is 0.1%, 100,000 samples should be used. In computationally intensive engineering problems, calculating $\mathbf{J}(\xi_i)$ for each ξ_i is expensive. Therefore, it is impractical to use the brute force Monte Carlo method, and some method of accelerating its convergence rate must be used.

This article presents an adjoint based approach for accelerating the Monte Carlo method. Section II derives the adjoint equation for unsteady incompressible Navier-Stokes equations, and calculating sensitivity derivatives using the adjoint solution. This section also briefly discusses our approach of solving the adjoint equations, and presents an example of its numerical solution. Section III then discusses two methods of using the adjoint solution to reduce the variance of Monte Carlo methods, and accelerate their convergence. Section IV presents an numerical example of these accelerated Monte Carlo methods, and demonstrate their effectiveness. Section V then concludes this work.

II. Adjoint equations for unsteady incompressible Navier-Stokes Equations

II.A. Mathematical Formulation

In this section, we derive the adjoint sensitivity gradient for uncertainty quantification problems related to incompressible Navier-Stokes flow. The uncertainties affect both the boundaries and the flow in the interior. The former describe The former describes uncertain wall roughness, unmodeled object movement / oscillations etc. The latter describe numerical uncertainties and subgrid model uncertainties. The objective function is assumed to be a measurable aerodynamic quantity, such as lift, drag or moment of a solid object in the flowfield, which depends on the pressure and viscous stress on the boundary of the solid object. The adjoint equations for Navier-Stokes flows have been derived and solved by [1,] [2,] and [3,].

The incompressible unsteady Navier-Stokes and continuity equations with normalized density are

$$\begin{aligned} &\frac{\partial v}{\partial t} + v \cdot \nabla v - \nabla \cdot (\mu \nabla v) + \nabla p = f(\xi) \\ &\nabla \cdot v = 0 \,, \end{aligned}$$

in the spatial domain Ω and time period $[0, T]$. The source term $f(\xi)$ depends on ξ, the random variables describing the sources of uncertainties. The initial condition is

$$v = u_0, \quad t = 0$$

and the boundary condition are dependent on the sources of uncertainties

$$v = v_b(\xi), \quad x \in \partial\Omega.$$

The objective function is a time-integrated aerodynamic quantity, which depends on the pressure and viscous stresses at the boundary:

$$\mathbf{J}(v, \xi) = \int_0^T J(\mathbf{n} \cdot \tau \,|_{\partial\Omega},\, p \,|_{\partial\Omega};\, t)\, dt \,,$$

where the viscous stress tensor is $\tau = \mu \nabla v$, and $\mathbf{n}$ is the unit wall-normal.

Let ξ_0 be the point of linearization, v_0 and p_0 be the solution of the Navier-Stokes equations with boundary condition $v_b(\xi_0)$ and forcing term $f(\xi_0)$. We can linearize both the Navier-Stokes equations and the objective function. By defining $\delta\xi = \xi - \xi_0$, $\delta v = v - v_0$ and $\delta p = p - p_0$, the linearized incompressible LES equation is

$$\begin{aligned} &\frac{\partial\, \delta v}{\partial t} + \mathcal{L}_{v_0}\, \delta v + \nabla \delta p = \frac{\partial f}{\partial \xi}\, \delta\xi \\ &\nabla \cdot \delta v = 0 \end{aligned} \tag{2}$$

with boundary condition

$$\delta v = \frac{\partial v_b}{\partial \xi}\,\delta\xi, \quad x \in \partial\Omega,$$

and initial condition

$$\delta v = 0, \quad t = 0.$$

The linearized Navier-Stokes operator $\mathcal{L}_{v_0}$ is defined as

$$\mathcal{L}_{v_0}\delta v = \delta v \cdot \nabla v_0 + v_0 \cdot \nabla \delta v - \nabla \cdot (\mu \nabla \delta v)\ .$$

Note that the operator depends on the point of linearization v_0.

On the other hand, the linearized objective function is

$$\begin{aligned}\delta \mathbf{J}(v,\xi) &= \mathbf{J}(v,\xi) - \mathbf{J}(v_0,\xi_0) \\ &= \int_0^T \iint_{\partial\Omega} (a(x,t)\cdot(\mathbf{n}\cdot\delta\tau) + b(x,t)\ \delta p)\, ds\ dt \\ &+ \int_0^T \iiint_{\Omega} c(x,t)\cdot\delta v\ dx\ dt\ .\end{aligned} \tag{3}$$

where

$$\delta\tau = \tau - \tau_0 = \mu_0 \nabla \delta v + \delta\mu \nabla v_0$$

$a(x,t)$, $b(x,t)$ and $c(x,t)$ are the Frechet derivatives

$$a = \frac{\partial J}{\partial\,(\mathbf{n}\cdot\tau)}\ , \qquad b = \frac{\partial J}{\partial p}\ , \qquad c = \frac{\partial J}{\partial v}$$

a is a vector function at the boundaries, b is a scalar function at the boundaries, and c is a vector function in the interior.

Define the adjoint variables $\hat{v}$, $\hat{p}$ and $\hat{\mu}$ such that they satisfy the adjoint equation

$$\begin{aligned} -\frac{\partial \hat{v}}{\partial t} + \mathcal{L}^*_{v_0}\hat{v} + \nabla\hat{p} &= -c \\ \nabla\cdot\hat{v} &= 0 \end{aligned} \tag{4}$$

where $\mathcal{L}^*_{v_0}$ is the adjoint operator of $\mathcal{L}_{v_0}$:

$$\mathcal{L}^*_{v_0}\hat{v} = \nabla v_0\cdot\hat{v} - v_0\cdot\nabla\hat{v} - \nabla\cdot(\mu_0\nabla\hat{v})\ .$$

The adjoint variables also must satisfy the terminal condition

$$\hat{v} = 0, \quad t = T$$

and the adjoint boundary condition

$$\hat{v} = a - \mathbf{n}\,(b + a\cdot\mathbf{n})\,, \quad x \in \partial\Omega\ . \tag{5}$$

We now show that the sensitivity derivatives of the objective function $\mathbf{J}$ with respect to ξ can be calculated using ξ and adjoint variables.

First, the adjoint boundary condition (5) implies that

$$\delta p\ \mathbf{n}\cdot\hat{v} = b\,\delta p \tag{6}$$

and because $\mathbf{n}\cdot\tau\cdot\mathbf{n} \equiv 0$ on the boundaries, (5) further implies

$$\mathbf{n}\cdot(\mu\nabla\delta v)\cdot\hat{v} = \mathbf{n}\cdot\delta\tau\cdot\hat{v} = -\mathbf{n}\cdot\delta\tau\cdot a\ . \tag{7}$$

In addition, from the adjoint equation,

$$c \cdot \delta v = -\delta v \cdot \left(-\frac{\partial \hat{v}}{\partial t} + \mathcal{L}^*_{v_0} \hat{v} + \nabla \hat{p} \right) \tag{8}$$

and

$$0 = \delta p \, \nabla \cdot \hat{v} \tag{9}$$

Incorporating (6), (7), (8) and (9) into the linearized objective function (3), we get

$$\begin{aligned} \delta \mathbf{J}(v, \xi) = & \int_0^T \iint_{\partial\Omega} (\mathbf{n} \cdot (\mu \nabla \delta v) \cdot \hat{v} - \delta p \, \mathbf{n} \cdot \hat{v}) \, ds \, dt \\ & - \int_0^T \iiint_{\Omega} \delta v \cdot \left(-\frac{\partial \hat{v}}{\partial t} + \mathcal{L}^*_{v_0} \hat{v} + \nabla \hat{p} \right) dx \, dt \\ & + \int_0^T \iiint_{\Omega} \delta p \, \nabla \cdot \hat{v} \, dx \, dt \end{aligned} \tag{10}$$

Furthermore, using integration by parts, and applying the initial and boundary conditions for both the linearized Navier-Stokes equations and the adjoint equation, we get the following four equalities:

$$\int_0^T \hat{v} \cdot \frac{\partial \, \delta v}{\partial t} dt = - \int_0^T \delta v \cdot \frac{\partial \hat{v}}{\partial t} dt \tag{11}$$

$$\begin{aligned} \iiint_{\Omega} \hat{v} \cdot \mathcal{L}_{v_0} \delta v \, dx = & \iiint_{\Omega} \delta v \cdot \mathcal{L}^*_{v_0} \hat{v} \, dx \\ & - \iint_{\partial\Omega} \mathbf{n} \cdot (\mu \nabla \delta v) \cdot \hat{v} \, ds \\ & + \iint_{\partial\Omega} \mathbf{n} \cdot (\mu \nabla \hat{v} - v_w \hat{v}) \cdot \delta v \, ds \end{aligned} \tag{12}$$

$$\iiint_{\Omega} \hat{v} \cdot \nabla \delta p \, dx = - \iiint_{\Omega} \delta p \, \nabla \cdot \hat{v} \, dx + \iint_{\partial\Omega} \delta p \, \mathbf{n} \cdot \hat{v} \, ds \tag{13}$$

$$\iiint_{\Omega} \hat{p} \, \nabla \cdot \delta v \, dx = - \iiint_{\Omega} \delta v \cdot \nabla \hat{p} \, dx + \iint_{\partial\Omega} \hat{p} \, \mathbf{n} \cdot \delta v \, ds \, . \tag{14}$$

Incorporating these four equations (6), (7), (8) and (9) into (10), we get

$$\begin{aligned} \delta \mathbf{J}(v, \xi) = & \int_0^T \iint_{\partial\Omega} (\mathbf{n} \cdot (\mu \nabla \hat{v}) - \mathbf{n} \cdot v_w \, \hat{v} - \hat{p} \, \mathbf{n}) \cdot \delta v \, ds \\ & - \int_0^T \iiint_{\Omega} \hat{v} \cdot \left(\frac{\partial \, \delta v}{\partial t} + \mathcal{L}_{v_0} \delta v + \nabla \delta p \right) dx \, dt \\ & + \int_0^T \iiint_{\Omega} \hat{p} \, \nabla \cdot \delta v \, dx \, dt \\ = & \int_0^T \iint_{\partial\Omega} (\mathbf{n} \cdot (\mu \nabla \hat{v}) - \mathbf{n} \cdot v_w \, \hat{v} - \hat{p} \, \mathbf{n}) \cdot \frac{\partial v_b}{\partial \xi} \, \delta \xi \, ds \\ & - \int_0^T \iiint_{\Omega} \hat{v} \cdot \frac{\partial f}{\partial \xi} \, \delta \xi \, dx \, dt \, . \end{aligned}$$

Therefore, the sensitivity gradient is

$$\begin{aligned} \frac{D\mathbf{J}}{D\xi} = & \int_0^T \iint_{\partial\Omega} (\mathbf{n} \cdot (\mu \nabla \hat{v}) - \mathbf{n} \cdot v_w \, \hat{v} - \hat{p} \, \mathbf{n}) \cdot \frac{\partial v_b}{\partial \xi} \, ds \\ & - \int_0^T \iiint_{\Omega} \hat{v} \cdot \frac{\partial f}{\partial \xi} \, dx \, dt \, . \end{aligned} \tag{15}$$

It can be calculated by using the adjoint solution at each time step. This is a key motivation for obtaining the adjoint solution in the first place.

II.B. The adjoint Navier-Stokes solver

This section describes the numerical scheme used to solve the adjoint Navier-Stokes equations. In designing the numerical scheme, we follow the "discrete adjoint" approach. The numerical scheme we obtain by this approach is not only a consistent discretization of the continuous adjoint Navier-Stokes equations, it is also the direct discrete adjoint of the numerical scheme for solving the unsteady Navier-Stokes equations. This approach has an important advantage[a]: the sensitivity gradient we obtained is not affected by discretization error of either the Navier-Stokes equation or the adjoint Navier-Stokes equations. Since the discrete adjoint scheme is dependent on the scheme for solving the Navier-Stokes equations, we start by describing CDP, our unsteady Navier-Stokes solver on an unstructured hybrid mesh.

The 3-D unsteady Navier-Stokes solver is called CDP [4,]. The solver uses a fractional-step method to enforce the divergence-free condition. An implicit BDF-2 time integration scheme [5,] is used for time discretization. The spatial discretization is a node based finite volume method. Velocity and pressure are stored in the dual-volume corresponding to each node. They are denoted as u without subscript, and p, respectively. A separate velocity field is stored on the surfaces of the dual-volumes, precisely corresponding to the face edges, to enforce the divergence-free condition. This auxiliary velocity field is denoted as u_m. The simplex superposition scheme [6,] is used for spatial discretization of the gradient, divergence, Laplacian and convection operators. The scheme is second-order accurate both in space and time for arbitrarily shaped elements.

A detailed description of the numerical scheme for solving the discrete adjoint equations can be found in [7,]. The adjoint equation evolves backward in time; its terminal condition is set at the last time step of the Navier-Stokes time integration, and each evaluation of the adjoint scheme brings the adjoint solution one time step backward. The adjoint scheme resembles a reversed prediction-correction scheme. Each time step of the adjoint solver involves the following operations:

1. Obtain the Navier-Stokes state variables u^{k-1}, u^{k-2}, u_m^{k-1}, u_m^{k-2} and p^{k-1} from the checkpointing scheme [8,].

2. Solve the prediction part (first three lines) of the Navier-Stokes step.

3. Retrograde the correction part of the adjoint step.

4. Retrograde the prediction part of the adjoint step.

Since the adjoint equation is calculated retrograde in time, the Navier-Stokes state variables are needed in a reverse-time order. The dynamic checkpointing scheme developed by [8,] is used to fulfill this requirement. Using the dynamic checkpointing scheme, the adjoint equation can be solved at a computational cost of around 4 to 8 times that of the Navier-Stokes equations, with moderate increase in memory usage.

II.C. Numerical example of an adjoint solution

The last section discussed how to solve the adjoint Navier-Stokes equations for general objective functions. This section gives an example of the process. We consider a Newtonian incompressible fluid flowing past an infinitely long circular cylinder. The Reynolds number with respect to the far-field flow velocity and the cylinder diameter is 100. At this Reynolds number, the flowfield is 2-D, unsteady and periodic in time [9,].

The incompressible Navier-Stokes equations were solved in a 2-D domain of size 60 cylinder diameters (flow direction) by 80 cylinder diameters (crossflow direction). The domain was discretized with approximately 10,000 unstructured quadrangle mesh elements. By taking advantage of the unstructured grid, most of the mesh points were concentrated around the cylinder and in a band of length 10 cylinder diameters

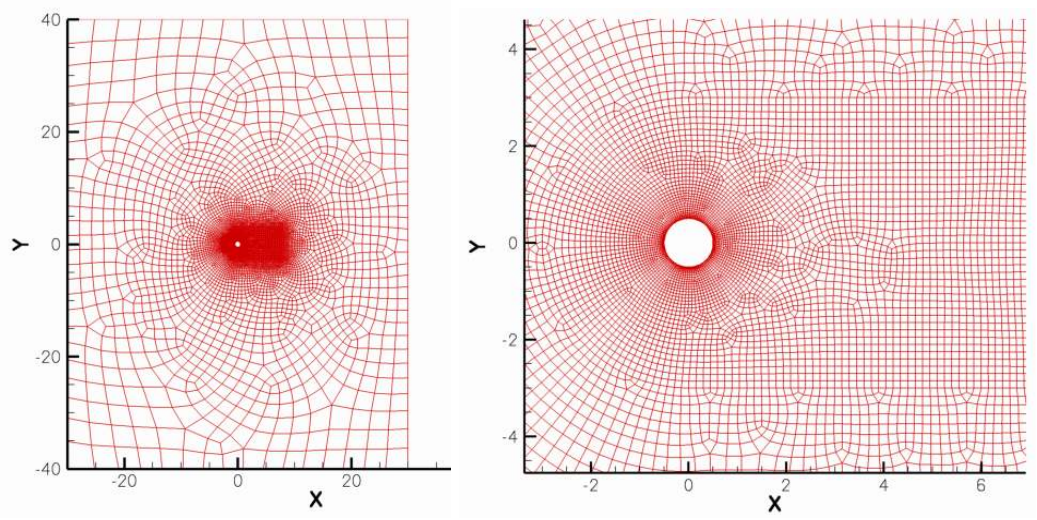

Figure 1. The mesh used for calculating the flowfield at Reynolds number 100. The left picture shows the entire computational domain of 60×80; the right picture zooms in to a small region around the cylinder.

downstream of the cylinder. Figure 1 shows the unstructured mesh used. The flow velocity on the left boundary of the domain was set to the freestream velocity, the upper and lower boundaries of the domain were set to periodic, and an convective outlet boundary condition was used at the downstream boundary. An algebraic multigrid solver was used to solve the Poisson equation in the corrector steps. We used a fixed time step size, $\Delta t = 0.1$, in this calculation. The corresponding CFL number was approximately 3 in the region near the cylinder. All physical quantities were normalized with respect to the fluid density, the freestream velocity and the cylinder diameter.

The uncertainty in the problem we consider comes from small rotational oscillations of the cylinder in the flow. The specific form of the oscillation is unknown, and the rate of rotation ω is modeled as a random time process. A no-slip boundary condition is used on the wall, which is

$$v_x = \omega y, \quad v_y = -\omega x.$$

Oscillatory rotation of the cylinder has been shown to have significant effect on the drag coefficient [10,]. The objective of this example is to analyze the effect of the small random rotational oscillations on the drag coefficient. Specifically, we want to obtain the probability distribution of the time-averaged drag coefficient of the cylinder from the random rotation. We describe this process as ''propagating" the uncertainties from the sources, the random rotation in this example, to the objective quantity, the time-averaged drag coefficient $\bar{c_d}$.

The objective function is

$$\mathbf{J} = \overline{c_d} = \frac{1}{T}\int_0^T c_d \, dt \, ,$$

where c_d is the instantaneous drag coefficient,

$$c_d = \frac{D(t)}{\frac{1}{2}\rho v_\infty^2 d} \, .$$

$D(t)$ is the instantaneous drag on the cylinder per spanwise length, ρ is fluid density, v_∞ is freestream velocity and d is the cylinder diameter. Since all quantities are normalized with ρ, v_∞ and d, $\rho v_\infty^2 d = 1$, and

$$c_d = 2D(t) \, .$$

[a]The discrete adjoint also have disadvantages compared to the continuous adjoint, such as its dependency on the primal scheme, and generally greater complexity.

The instantaneous unit-span drag $D(t)$ consists of pressure drag D_p and viscous drag D_v, which are

$$D_p = -\oint p\mathbf{n}\cdot\mathbf{e_x}ds$$

$$D_v = \oint \mathbf{n}\cdot\tau\cdot\mathbf{e_x}ds$$

and

$$c_d = 2\oint(\mathbf{n}\cdot\tau - p\mathbf{n})\cdot\mathbf{e_x}\ ds$$

where $\mathbf{n}$ is wall-normal unit vector, μ is viscosity, $\mathbf{e_x} = (1,0,0)$ is the unit vector in the x-direction, and the integration is on the circumference of the cylinder cross-section.

A numerical approximation of the time-averaged drag coefficient is

$$\mathbf{J} = \frac{2}{T}\sum_{k=1}^{T}\Delta t\sum_{i\in sfc}(\mathbf{n}_i\cdot\tau_i - p_i\mathbf{n}_i)\cdot\mathbf{e_x}ds_i\ . \tag{16}$$

In the formula, sfc represents mesh nodes on the wall boundary of the cylinder, p_i is the pressure, and τ_i is the viscous stress at node i, which can be directly calculated from the velocity of its neighboring nodes:

$$\tau_i = \frac{1}{Re}\nabla v|_i = \frac{1}{Re}\sum_{j\in nbr(i)}\kappa_{i\,j}v_j\ ,$$

where $\kappa_{i\,j}$ is the Simplex-Superposition discretization of the gradient operator, its specific value depends on the cell types and geometry. We note that this objective function is linear with respect to the velocity and pressure. The source term of the adjoint equation corresponding to this objective function is

$$\begin{aligned} J_{v\,i} &= \sum_{j\in sfc\cup nbr(i)}\mathbf{n}_j\cdot\kappa_{j\,i}\mathbf{e_x} \\ J_{p\,i} &= \begin{cases}\mathbf{n}_i\cdot\mathbf{e_x} & i\in sfc \\ 0 & i\notin sfc\end{cases}. \end{aligned} \tag{17}$$

We solve the discrete adjoint equations with objective function being the time-averaged drag coefficient (16). Several snapshots of the adjoint solution are depicted in Figure 2. The left side of the figure shows the streamwise velocity u_x at four time instances; the right side shows the corresponding adjoint streamwise velocity $\hat{u}_x$. These adjoint fields reveals the sensitivity derivative of the objective function with respect to changes in the flowfield. In these plots, red is positive values of $\hat{u}_x$, indicating that a positive change in the streamwise velocity would increase the drag of the cylinder. Such regions include the upstream of the cylinder, where an increase in u_x would generate more skin friction drag, and unsteady bands in the shear layers on the sides of the cylinder, where an increase in u_x would temporarily widen the wake, creating more pressure drag on the cylinder. In contrast, blue is negative values of $\hat{u}_x$, indicating that a positive change in the streamwise velocity would reduce the drag of the cylinder. These regions are unsteady bands in the shear layers where an increase in u_x would temporarily narrow the wake, decreasing the pressure drag.

By integrating the adjoint solution using (15), we obtain the sensitivity gradient of the objective function with respect to the random variables describing the sources of uncertainty. In this case, it is the sensitivity gradient of the time-averaged drag coefficient with respect to the rate of rotation ω. Since ω is a random time process, the sensitivity gradient obtained is a function of time. This function $\hat{\omega} = \left.\dfrac{\partial\mathbf{J}}{\partial\omega}\right|_t$ is plotted in Figure 3, along with the time history of the instantaneous lift and drag coefficients on the cylinder.

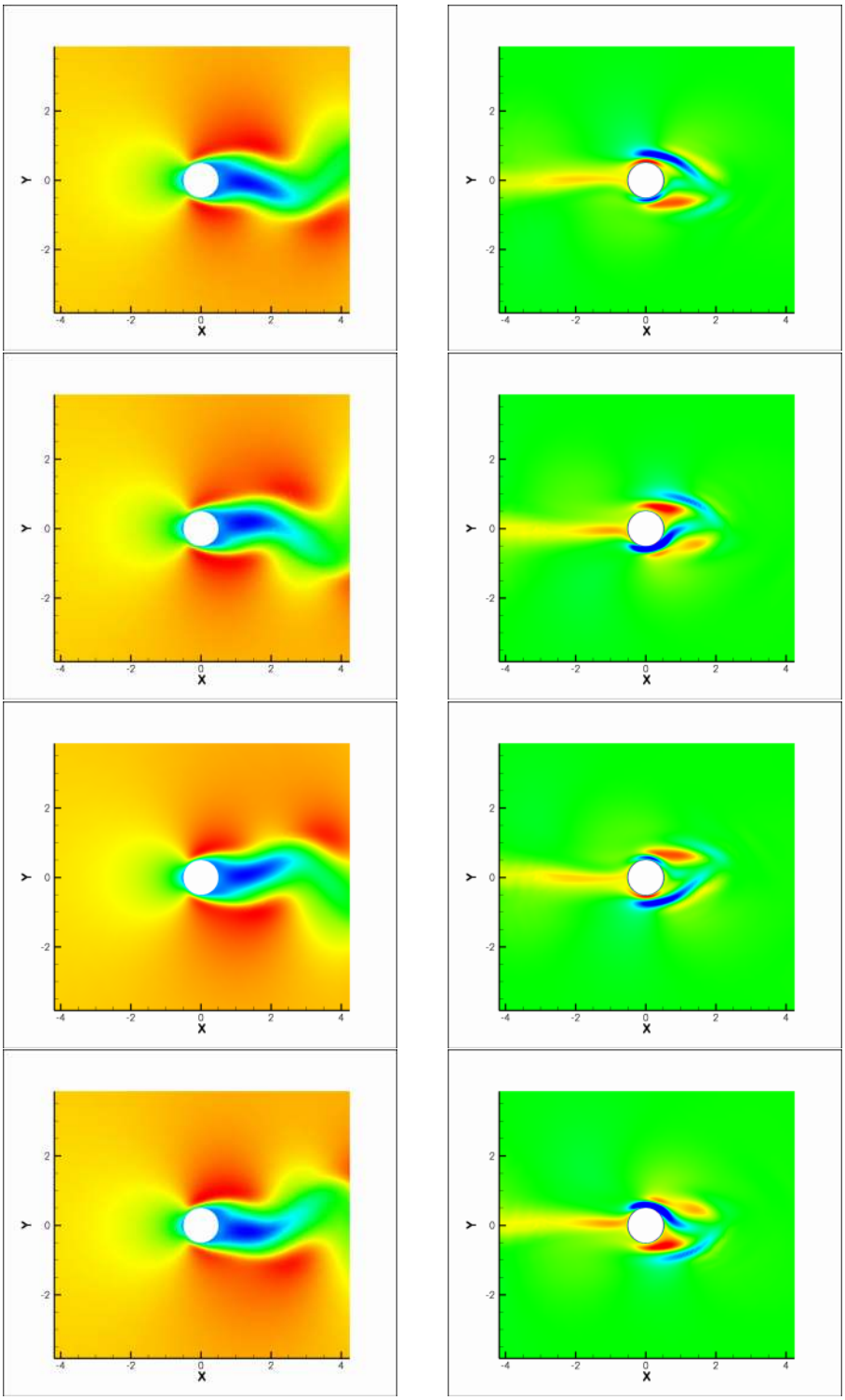

Figure 2. Flow and adjoint solutions at $t = 125.0, 126.5, 128.0, 129.5$ (upper-left, upper-right, lower-left, lower-right)

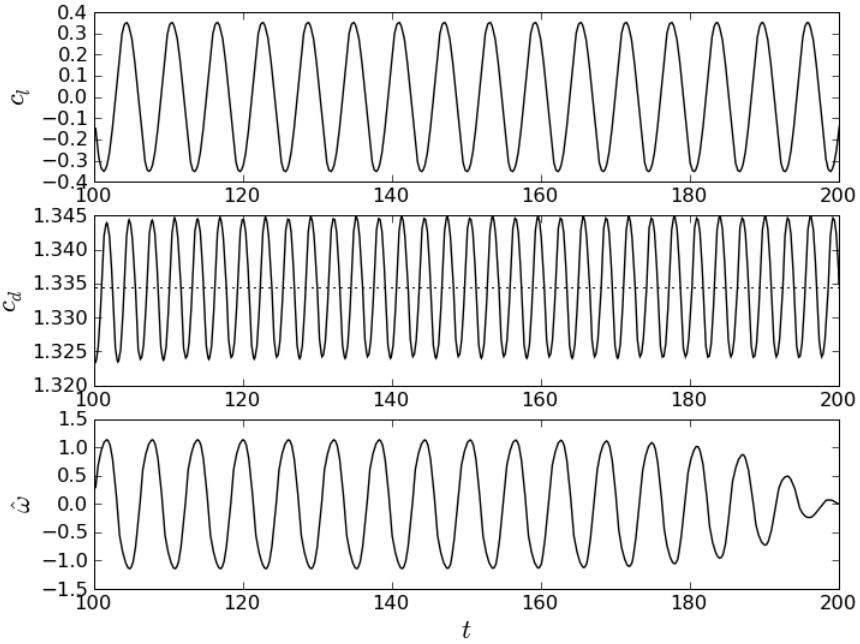

Figure 3. The time history of c_l, c_d and $\hat{\omega} = \frac{\partial \mathbf{J}}{\partial \omega}$ of flow past a non-rotating circular cylinder at $Re = 100$. The objective function for the adjoint equation is the time-integrated drag on the cylinder. The time averaged drag coefficient 1.3345 is plotted as a dotted line in the c_d plot.

III. Accelerating Monte Carlo using adjoint sensitivity gradient

Acceleration of convergence for Monte Carlo methods can be achieved by reducing the variance of its estimator. For a fixed target approximation error, the number of samples required is proportional to the variance of an unbiased estimator. This section discusses variance-reduction techniques for calculating the tail probability $P(\mathbf{J} > \mathbf{J_C})$ based on a single adjoint calculation. The adjoint calculation can be used to calculate the sensitivity gradient, which can be used to approximate the objective function as a linear function of the random variables describing the sources of uncertainty. Suppose the adjoint is evaluated at ξ_0, and the sensitivity gradient $\mathbf{J}'(\xi_0)$ is calculated from the adjoint solution. The objective function can be approximated in the vicinity of ξ_0 as

$$\mathbf{J}(\xi) \approx \mathbf{J}_L(\xi) = \mathbf{J}(\xi_0) + \nabla \mathbf{J}(\xi_0) \cdot (\xi - \xi_0) \ .$$

The information provided by this linear approximation enables us to reduce the variance of the Monte Carlo method using control variate and importance sampling. Section III.A discusses the control variate technique, and Section III.B combines control variate with importance sampling to further reduce the variance.

III.A. Control variates

The idea of control variate [11,] explores the similarity between the real objective function $\mathbf{J}$ and its linear approximation $\mathbf{J}_L$. In this method, we approximate the target tail probability $P(\mathbf{J} > \mathbf{J_C})$ with the estimator

$$P(\mathbf{J} > \mathbf{J_C}) \approx P_N^{CV} = P(\mathbf{J}_L > \mathbf{J_C}) + \frac{1}{N} \sum_{i=1}^{N} \left(I(\mathbf{J}(\xi_i) > \mathbf{J_C}) - I(\mathbf{J}_L(\xi_i) > \mathbf{J_C}) \right) \ .$$

This estimator P_N^{CV} is unbiased for $P(\mathbf{J} > \mathbf{J_C})$ because

$$\begin{aligned}\mathrm{E}\left[P_N^{CV}\right] &= P(\mathbf{J}_L > \mathbf{J_C}) + \mathrm{E}[I(\mathbf{J}(\xi_i) > \mathbf{J_C})] - \mathrm{E}[I(\mathbf{J}_L(\xi_i) > \mathbf{J_C})] \\ &= P(\mathbf{J}_L > \mathbf{J_C}) + P(\mathbf{J} > \mathbf{J_C}) - P(\mathbf{J}_L > \mathbf{J_C}) \\ &= P(\mathbf{J} > \mathbf{J_C})\,.\end{aligned}$$

The variance of this new estimator is

$$\begin{aligned}\mathrm{Var}\left[P_N^{CV}\right] =& \frac{1}{N}\,\mathrm{Var}\left[I(\mathbf{J} > \mathbf{J_C}) - I(\mathbf{J}_L > \mathbf{J_C})\right] \\ =& \frac{1}{N}\left(P(\mathbf{J} > \mathbf{J_C} > \mathbf{J}_L \text{ or } \mathbf{J} < \mathbf{J_C} < \mathbf{J}_L) - (P(\mathbf{J}_L > \mathbf{J_C}) - P(\mathbf{J} > \mathbf{J_C}))^2\right).\end{aligned} \tag{18}$$

From this formula, we can see that the variance of the control variate estimator depends on the accuracy the linear approximation $\mathbf{J}_L$. If the approximation is good, then $\mathbf{J} \approx \mathbf{J}_L$. As a result, $P(\mathbf{J} > \mathbf{J_C} > \mathbf{J}_L \text{ or } \mathbf{J} < \mathbf{J_C} < \mathbf{J}_L)$, which is the probability that $\mathbf{J_C}$ lies within the small interval $[\mathbf{J}, \mathbf{J}_L]$, is small. Since $\mathrm{Var}\left[P_N^{CV}\right] \leq \frac{1}{N}\,P(\mathbf{J} > \mathbf{J_C} > \mathbf{J}_L \text{ or } \mathbf{J} < \mathbf{J_C} < \mathbf{J}_L)$, the variance is small. In fact, if the approximation is exact, i.e., $\mathbf{J}_L = \mathbf{J}$, then the variance is zero. However, if the approximation is not accurate at all, this estimator may not reduce the variance. For example, if the $\mathbf{J}_L < \mathbf{J_C}$ is a constant, then $P(\mathbf{J} > \mathbf{J_C} > \mathbf{J}_L \text{ or } \mathbf{J} < \mathbf{J_C} < \mathbf{J}_L) = P(\mathbf{J} > \mathbf{J_C})$, and $\mathrm{Var}\left[P_N^{CV}\right] = \frac{1}{N}\,\left(P(\mathbf{J} > \mathbf{J_C}) - P(\mathbf{J}_L > \mathbf{J_C})^2\right)$ is the same as the variance of the Naive Monte Carlo method. For this reason, this method offers the most improvement in convergence when the linear approximation $\mathbf{J}_L$ is at least a fairly accurate approximation.

In Section IV, we will apply the control variate technique in an unsteady fluid flow problem, and demonstrate its variance-reduction effectiveness.

III.B. Importance sampling

The brute force Monte Carlo method is ineffective when the target failure probability is small because only a small fraction of samples lay in the region where failure might occur. The importance sampling technique [12,] addresses this ineffectiveness by concentrating the samples in regions where the failure is more likely to occur.

In our adjoint-based importance sampling, we select the concentration of sampling based on how likely the objective function $\mathbf{J}$ and its linear approximation $\mathbf{J}_L$ are on different sides of the critical value $\mathbf{J_C}$. Therefore, the first step of our method is to estimate the error of this linear approximation; since it is likely to grow as $O(|\xi - \xi_0|^2)$, we model the normalized approximation error $\dfrac{\mathbf{J} - \mathbf{J}_L}{||\xi - \xi_0||^2}$ as a random variable, whose probability distribution is estimated by plotting the histogram of $\dfrac{\mathbf{J}(\xi_i) - \mathbf{J}_L(\xi_i)}{||\xi_i - \xi_0||^2}$ for a small number of samples ξ_i. This step quantifies how much the objective function $\mathbf{J}$ may deviate from its linear approximation $\mathbf{J}_L$. For each ξ, without going through the expensive process of calculating the objective function $\mathbf{J}(\xi)$, we can estimate its prior probability distribution based only on its linear approximation $\mathbf{J}_L(\xi)$ and $|\xi - \xi_0|^2$. Note that this probability distribution is defined in the Bayesian sense, in contrast to the tail probability we intend to calculate, which is defined as a frequency probability. For the sake of clarity, we denote the Bayesian probability as p and the frequency probability as P.

Now for each ξ, we can calculate the prior probability

$$p(\mathbf{J}(\xi) > \mathbf{J_C} \mid \mathbf{J}_L(\xi)) = p\left(\frac{\mathbf{J}(\xi) - \mathbf{J}_L(\xi)}{||\xi - \xi_0||^2} > \frac{\mathbf{J_C} - \mathbf{J}_L(\xi)}{||\xi - \xi_0||^2} \;\middle|\; \mathbf{J}_L(\xi)\right)$$

based on the value of $\mathbf{J}_L(\xi)$ and the estimated probability distribution of $\dfrac{\mathbf{J}(\xi) - \mathbf{J}_L(\xi)}{||\xi - \xi_0||^2}$. We call it *a prior* probability because this probability is estimated prior to calculating the real value of the objective function $\mathbf{J}(\xi)$. Furthermore, we can calculate the prior probability that the objective function and its linear

approximation are on different sides of $\mathbf{J_C}$. We denote this probability as p_D,

$$\begin{aligned} p_D(\xi) &= p(\mathbf{J}(\xi) > \mathbf{J_C} > \mathbf{J}_L(\xi) \text{ or } \mathbf{J}(\xi) < \mathbf{J_C} < \mathbf{J}_L(\xi) \mid \mathbf{J}_L(\xi)) \\ &= \begin{cases} p(\mathbf{J}(\xi) > \mathbf{J_C} \mid \mathbf{J}_L(\xi)), & \mathbf{J_C} > \mathbf{J}_L(\xi) \\ 1 - p(\mathbf{J}(\xi) > \mathbf{J_C} \mid \mathbf{J}_L(\xi)), & \mathbf{J_C} < \mathbf{J}_L(\xi) \, . \end{cases} \end{aligned}$$

Our importance sampling technique concentrates the samples based on this prior probability. For each proposed sample ξ_i obtained according to the distribution of ξ, we calculate $\mathbf{J}_L(\xi_i)$ and $p_D(\xi)$. This probability determines the rate at which the proposed sample is approved or rejected. Specifically, an independent uniform pseudo-random number is generated and compared with this prior probability, and the sample is rejected if the pseudo-random number is greater of the two. Calculations for $\mathbf{J}(\xi_i)$ are done for approved ξ_is only. This mechanism concentrates the approved samples in regions where the linear approximation $\mathbf{J}_L$ is likely to give a false indication of whether the objective function exceeds the critical value $\mathbf{J_C}$.

In order to achieve an unbiased estimator, this concentration of samples must be compensated for by weighing each sample differently. Let $f(\xi)$ be the probability density function of ξ, then the probability density function of the approved samples is $\dfrac{f(\xi)p_D(\xi)}{\int p_D(\xi)dP(\xi)}$; if we denote this probability distribution as $Q(\xi)$, then the Radon-Nikodym derivative is given by

$$\frac{dQ}{dP} = \frac{p_D(\xi)}{\int p_D(\xi)dP(\xi)} \, . \tag{19}$$

With the approved samples $\xi_1, \xi_2, \ldots, \xi_N$ distributed according to the new probability distribution Q, a different estimator P_N^{IS} is used.

$$P(\mathbf{J}(\xi) > \mathbf{J_C}) \approx P_N^{IS} = P(\mathbf{J}_L > \mathbf{J_C}) + \frac{1}{N}\sum_{i=1}^{N} \frac{dP}{dQ} \left(I(\mathbf{J}(\xi_i) > \mathbf{J_C}) - I(\mathbf{J}_L(\xi_i) > \mathbf{J_C})\right).$$

This estimator P_N^{IS} is unbiased for $P(\mathbf{J} > \mathbf{J_C})$ because

$$\begin{aligned} \mathrm{E}_Q\left[P_N^{IS}\right] &= P(\mathbf{J}_L > \mathbf{J_C}) + \mathrm{E}_Q\left[\frac{dP}{dQ}\left(I(\mathbf{J}(\xi_i) > \mathbf{J_C}) - I(\mathbf{J}_L(\xi_i) > \mathbf{J_C})\right)\right] \\ &= P(\mathbf{J}_L > \mathbf{J_C}) + \mathrm{E}_P\left[I(\mathbf{J}(\xi_i) > \mathbf{J_C}) - I(\mathbf{J}_L(\xi_i) > \mathbf{J_C})\right] \\ &= P(\mathbf{J} > \mathbf{J_C}) \, . \end{aligned}$$

The new estimator P_N^{IS} with importance sampling can further reduce the variance of the Monte Carlo method and accelerate its convergence. The variance of this importance sampling estimator is

$$\begin{aligned} \mathrm{Var}_Q\left[P_N^{IS}\right] =& \frac{1}{N}\,\mathrm{Var}_Q\left[\frac{dP}{dQ}\left(I(\mathbf{J}(\xi_i) > \mathbf{J_C}) - I(\mathbf{J}_L(\xi_i) > \mathbf{J_C})\right)\right] \\ =& \frac{1}{N}\,\mathrm{E}_Q\left[\left(\frac{dP}{dQ}\right)^2\left(I(\mathbf{J}(\xi_i) > \mathbf{J_C}) - I(\mathbf{J}_L(\xi_i) > \mathbf{J_C})\right)^2\right] \\ & - \frac{1}{N}\left(P(\mathbf{J} > \mathbf{J_C}) - P(\mathbf{J}_L > \mathbf{J_C})\right)^2 \\ =& \frac{1}{N}\,\mathrm{E}_P\left[\frac{dP}{dQ} I(\mathbf{J} > \mathbf{J_C} > \mathbf{J}_L \text{ or } \mathbf{J} < \mathbf{J_C} < \mathbf{J}_L)\right] \\ & - \frac{1}{N}\left(P(\mathbf{J} > \mathbf{J_C}) - P(\mathbf{J}_L > \mathbf{J_C})\right)^2 \\ =& \frac{1}{N}\,\mathrm{E}_P\left[\frac{I(\mathbf{J} > \mathbf{J_C} > \mathbf{J}_L \text{ or } \mathbf{J} < \mathbf{J_C} < \mathbf{J}_L)}{p(\mathbf{J} > \mathbf{J_C} > \mathbf{J}_L \text{ or } \mathbf{J} < \mathbf{J_C} < \mathbf{J}_L)}\right] \int p_D(\xi)dP(\xi) \\ & - \frac{1}{N}\left(P(\mathbf{J} > \mathbf{J_C}) - P(\mathbf{J}_L > \mathbf{J_C})\right)^2 . \end{aligned} \tag{20}$$

In the next section, we will demonstrate that the importance sampling technique indeed reduces the variance.

IV. Application to an unsteady fluid flow problem

We consider the laminar flow around a circular cylinder at a Reynolds number of 100. We use the same mesh, problem setup and numerical methods as described in Section II.C. The objective function is the time-averaged drag coefficient of the cylinder, and we want to calculate the probability that it is greater than a critical value $\mathbf{J_C} = 1.345$.

The uncertainties in the problem come from small, unknown rotational oscillations of the cylinder in the flow. We assume that this random rotation consists of 10 different frequencies, including the vortex-shedding frequency and its subharmonics. The rotation at each frequency is described by two random numbers, making each frequency component of the oscillation random both in magnitude and random in phase. Specifically, let f_V be the frequency of the vortex-shedding, the angular speed of the cylinder rotating about its symmetric axis is

$$w(t) = \sum_{i=1}^{10} \xi_{2i-1} \cos(2\pi i f_V t) + \xi_{2i} \sin(2\pi i f_V t) \ , \tag{21}$$

where $\xi_1, \ldots, \xi_{20}$ are Gaussian random variables with mean 0 and standard deviation 0.01.

Although the magnitude of the random rotational oscillations are small, they have significant impact on the flowfield. As a result, the forces exerted on the cylinder by the fluid can be significantly affected. Figure 4 plots the time history of the drag coefficient c_d and lift coefficient c_l of the cylinder for 1000 samples of the random oscillations. These non-dimensionalized quantities c_d and c_l are defined as

$$c_d = \frac{D}{\frac{1}{2}\rho v_\infty^2 D\, b} \qquad c_l = \frac{L}{\frac{1}{2}\rho v_\infty^2 D\, b}$$

where D is the drag of the cylinder, i.e., the component of the aerodynamic force that is parallel to the freestream; L is the lift of the cylinder, i.e the component of the aerodynamic force that is perpendicular to the freestream; ρ is the density of the fluid; v_∞ is the speed of the freestream flow; D is the diameter of the cylinder, and b is the spanwise length of the cylinder.

In Figure 4, the white lines represent the lift and drag coefficients with no random oscillations. The black dotted lines show the lift and drag coefficients of cylinders undergoing rotational oscillations described by equation (21) with 1000 samples of the random vector $(\xi_1, \ldots, \xi_{20})$. As can be seen, the drag coefficient can be significantly changed by the random rotational oscillations. Therefore, our objective function, the time-averaged drag coefficient $\mathbf{J}$, should also depend on the specific form of the rotational oscillation, which is specified by the random vector $(\xi_1, \ldots, \xi_{20})$.

Figure 5 shows that the objective function is indeed modified by the random rotational oscillations. This histogram shows the empirical density function for the time-averaged drag coefficient, our objective function $\mathbf{J}$. The dotted vertical line represents the critical value $\mathbf{J_C} = 1.345$. As can be seen, there is a small but non-trivial probability that our objective function is higher than this critical value. We can also roughly estimate this tail probability from the histogram. Since the objective function exceeds the critical value in 48 of the 1000 samples we calculated, an unbiased estimate of the tail probability is $P(\mathbf{J} > \mathbf{J_C}) \approx 4.8\%$. However, the standard deviation of this estimate is 0.007, making the 3σ confidence interval of the tail probability

$$P(\mathbf{J} > \mathbf{J_C}) = 4.8 \pm 2.0\%. \tag{22}$$

These numbers reveal the inefficiency of the brute force Monte Carlo method. Despite of calculating the objective function for 1000 samples, the fraction of the samples whose objective function is higher than the critical value is small. As a result, the probability calculated by this method has a large variance and therefore is not accurate.

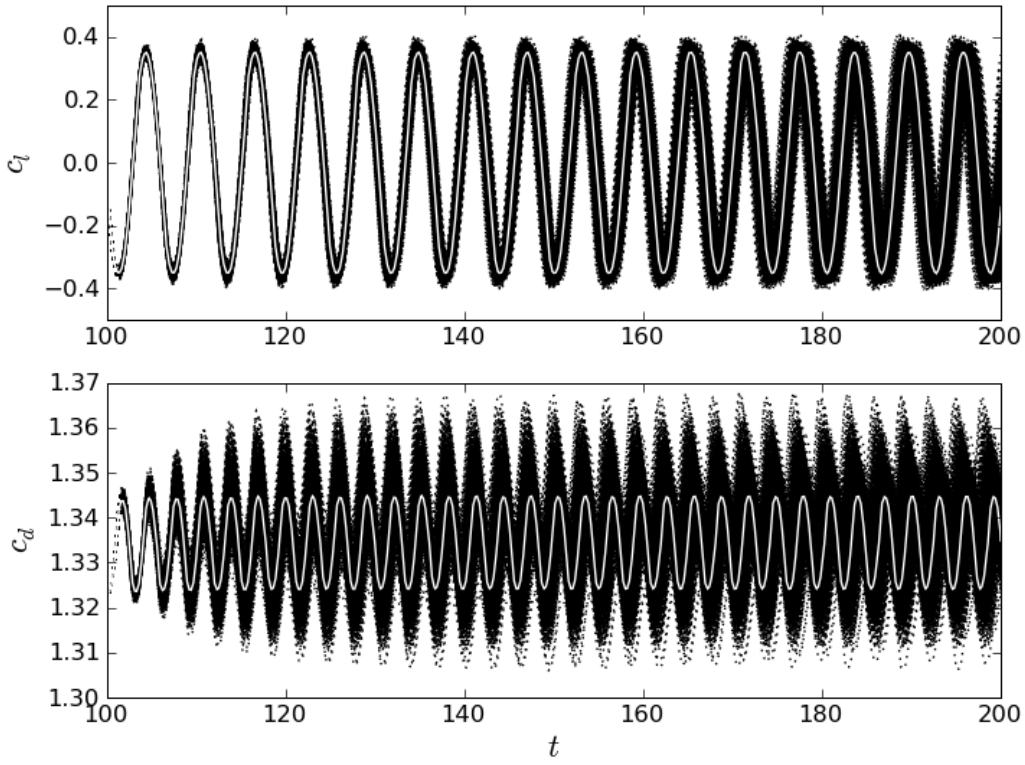

Figure 4. Realizations of the c_l (the top figure) and c_d (the bottom figure). The white lines indicate c_l and c_d with no rotations.

IV.A. Adjoint linear approximation

The adjoint equation for this problem is solved with the method described in Chapters 3 and 4. From the adjoint solution, we can calculate the gradient of the objective function with respect to the rate of rotation of the cylinder.

$$\hat{\omega}(t) = \frac{\partial \mathbf{J}}{\partial \omega(t)}$$

Figure 3 shows the time history of $\hat{\omega}$. With $\hat{\omega}$ calculated, we can obtain the sensitivity gradient of the objective function with respect to $\xi_1, \ldots, \xi_{20}$, the random variables describing the rotational oscillation.

$$\begin{aligned}\frac{\partial \mathbf{J}}{\partial \xi_i} &= \int_{T_0}^{T_1} \frac{\partial \mathbf{J}}{\partial \omega(t)} \frac{\partial \omega(t)}{\partial \xi_i}\, dt \\ &= \begin{cases} \displaystyle\int_{T_0}^{T_1} \hat{\omega} \cos\left(2\pi \frac{i+1}{2} f_V t\right) dt, & i \text{ is odd} \\ \displaystyle\int_{T_0}^{T_1} \hat{\omega} \sin\left(2\pi \frac{i}{2} f_V t\right) dt, & i \text{ is even} \end{cases}\end{aligned}$$

Calculating this sensitivity derivatives for each $i = 1, \ldots, 20$ generates the sensitivity gradient vector of $\mathbf{J}$ as a function of the random variables $\xi = (\xi_1, \ldots, \xi_{20})$.

$$\nabla \mathbf{J} = \left(\frac{\partial \mathbf{J}}{\partial \xi_1}, \ldots, \frac{\partial \mathbf{J}}{\partial \xi_{20}}\right)$$

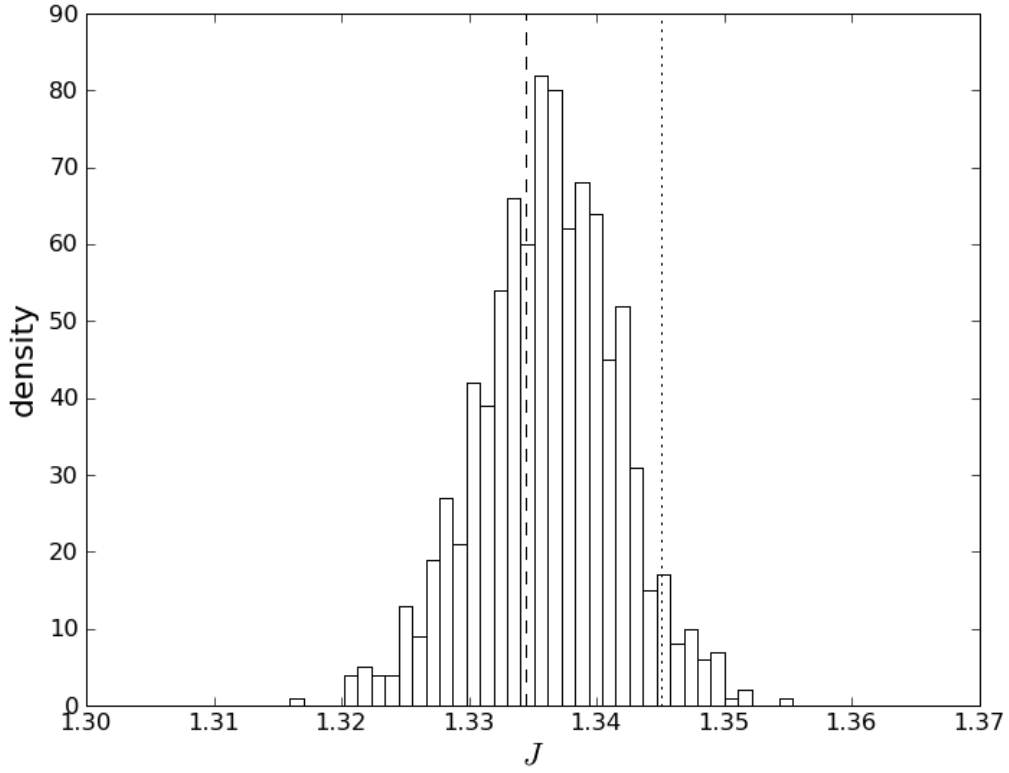

Figure 5. The histogram of the objective function J = $\overline{c_d}$. The dashed vertical line indicates J with no rotation. The dotted vertical line indicates $\mathbf{J_C} = 1.345$.

With this sensitivity gradient, we can construct a linear approximation of the objective function

$$\mathbf{J}(\xi) \approx \mathbf{J}_L(\xi) = \mathbf{J}_0 + \nabla\mathbf{J} \cdot \xi \tag{23}$$

where $\mathbf{J}_0 = \mathbf{J}(0)$ is the value of the objective function when $\xi = 0$, i.e., when the cylinder is not rotating. This adjoint linear approximation can be obtained with one Navier-Stokes solution, from which $\mathbf{J}_0$ is calculated, and one adjoint solution, from which the sensitivity gradient $\nabla\mathbf{J}$ is calculated. Once we have $\mathbf{J}_0$ and $\nabla\mathbf{J}$, the adjoint linear approximation $\mathbf{J}_L(\xi)$ can be calculated using (23) at essentially no additional computational cost.

Figure 6 shows the true value of the objective function $\mathbf{J}(\xi_i)$ against its adjoint linear approximation $\mathbf{J}_L(\xi_i)$ for 1000 randomly sampled $\xi_1, \ldots, \xi_{1000}$. Each cross on the plot represents one sample. As can be seen, although some samples deviate from the diagonal line, indicating large approximation errors, the adjoint linear approximation is a sufficiently accurate approximation to the objective function.

The adjoint linear approximation can be directly used to obtain a very efficient first-order estimate of $P(\mathbf{J} > \mathbf{J_C})$, the probability we want to calculate. Because $\mathbf{J}_L(\xi)$ can be evaluated with essential no computational cost, a very large number of Monte Carlo samples can be used to accurately calculate $P(\mathbf{J}_L > \mathbf{J_C})$. This probability that the linear approximation exceeds the critical value can be used to approximate the probability that the objective function exceeds the critical value. $P(\mathbf{J} > \mathbf{J_C}) \approx P(\mathbf{J}_L > \mathbf{J_C})$.

Figure 7 shows the application of this approach to our cylinder problem. The solid line is the empirical distribution function obtained from 10,000,000 samples of the adjoint linear approximation $\mathbf{J}_L$. For comparison, the bars are the same empirical distribution function obtained from the 1000 samples of $\mathbf{J}$. The

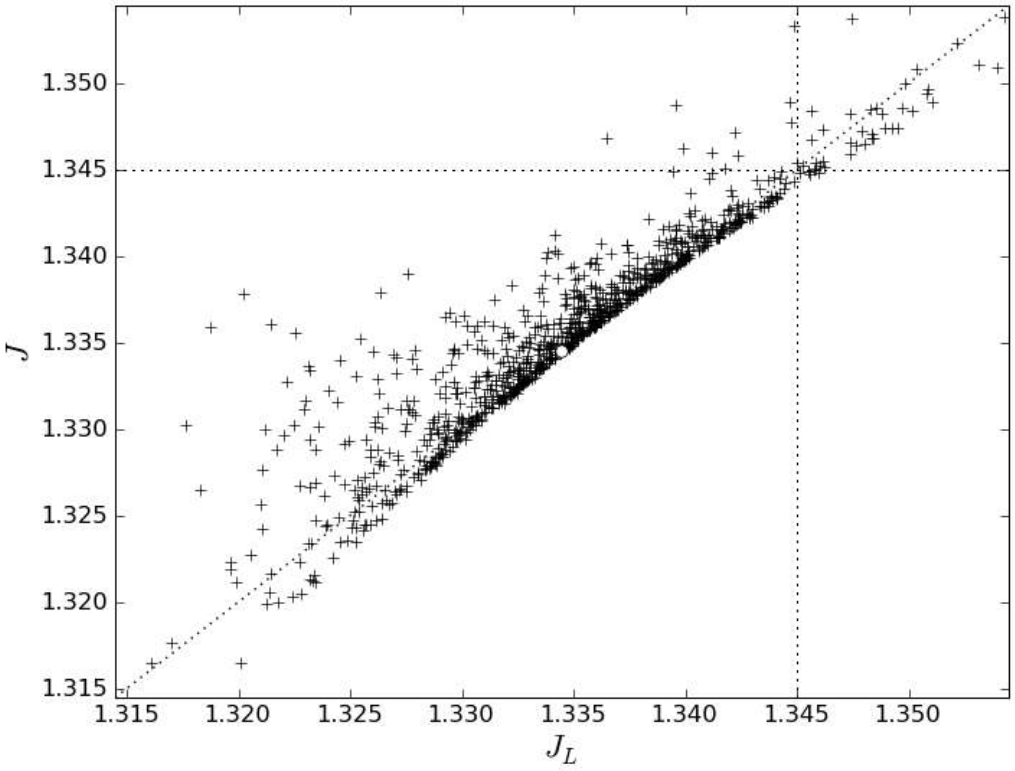

Figure 6. The correlation between adjoint approximation (horizontal axis) and true objective function (vertical axis). The horizontal and vertical dotted lines indicates the critical value $\mathbf{J_C}$. The circular symbol at the center indicates the objective function without rotation.

vertical dotted line indicates the critical value $\mathbf{J_C}$. As can be seen, the distribution function of the adjoint linear approximation is close to the distribution of the true objective function. In addition, because the distribution of the adjoint approximation is obtained with many more samples, it is much smoother than the empirical distribution of the true objective function. At the same time, the tail probability for the adjoint approximation $P(\mathbf{J}_L > \mathbf{J_C})$ can be obtained with little variance. In this case, we calculated

$$P(\mathbf{J} > \mathbf{J_C}) \approx P(\mathbf{J}_L > \mathbf{J_C}) = 3.6\%. \tag{24}$$

We compare this result with the value calculated using the brute force Monte Carlo method (22). The accuracy of their results are similar, because the true value of $P(\mathbf{J} > \mathbf{J_C})$ is $4.1 \pm 0.5\%$ (as calculated in Equation (26) later in this article). Nevertheless, the error of this adjoint approximation method is different in nature from the error of the brute force Monte Carlo method. The error of the brute force Monte Carlo method comes from its large variance of the estimator, while the result of the adjoint approximation method has very little variance due to the large number of samples used. Instead, the error of the adjoint approximation method is a consequence of approximating $P(\mathbf{J} > \mathbf{J_C})$ with $P(\mathbf{J}_L > \mathbf{J_C})$. The difference between these two probabilities can be illustrated using Figure 6. In the plot, the vertical axis is $\mathbf{J}$, and the horizontal axis is $\mathbf{J}_L$. The probability $P(\mathbf{J} > \mathbf{J_C})$ we want to calculate is the proportion of samples above the horizontal dotted line. In contrast, $P(\mathbf{J}_L > \mathbf{J_C})$, the probability we use to approximate $P(\mathbf{J} > \mathbf{J_C})$, is the proportion of samples to the right of the vertical line. By using this approximation, we underestimate the samples to the upper-left of the intersection of the two dotted lines, and overestimate the samples to the lower-right of the intersection.

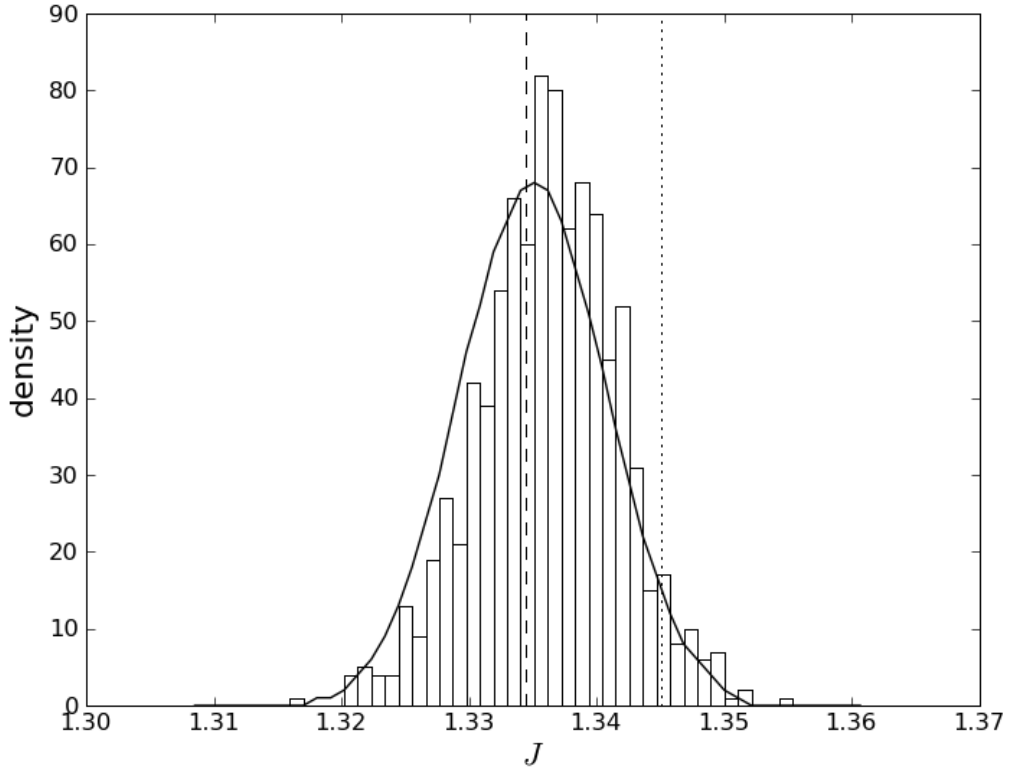

Figure 7. The empirical density function of the objective function obtained using the brute-force Monte Carlo (vertical bars), and the empirical density function of the adjoint approximation obtained using adjoint method (solid line). The dashed vertical line indicates J with no rotation. The dotted vertical line indicates $\mathbf{J_C} = 1.345$.

Although the adjoint approximation method has a similar large error to the brute force Monte Carlo method in our example, it is by far less expensive. The brute force Monte Carlo method requires 1000 Navier-Stokes calculations, while the adjoint approximation method involves only one Navier-Stokes calculation and one adjoint calculation. Since the adjoint calculation requires about 4 times the computational resources of a Navier-Stokes calculation, the brute force Monte Carlo method is 200 times more expensive than the adjoint approximation method. This extreme efficiency makes the adjoint method the first choice for a rough estimate of the tail probability.

IV.B. Accelerated Monte Carlo

Although the adjoint approximation method of estimating $P(\mathbf{J} > \mathbf{J_C})$ is computationally economical, it does not produce very accurate results. In this section, we use the methods described in the previous section to reduce the variance of the Monte Carlo method. With these variance-reduction techniques, we produce more accurate results with the same cost as the brute force Monte Carlo method.

We first apply the control variates technique to this problem, as described in Section III.A. With this technique, we use the same samples as in the brute force Monte Carlo method, but change the estimator to

$$P_N^{CV} = P(\mathbf{J}_L > \mathbf{J_C}) + \frac{1}{N}\sum_{i=1}^{N} \left(I(\mathbf{J}(\xi_i) > \mathbf{J_C}) - I(\mathbf{J}_L(\xi_i) > \mathbf{J_C}) \right).$$

We note that the first term in the estimator is simply the result we obtain from the adjoint approximation

method, and the second term is simply an unbiased estimator of the difference between the true tail probability $P(\mathbf{J} > \mathbf{J_C})$ and the estimated tail probability by adjoint approximation $P(\mathbf{J}_L > \mathbf{J_C})$. In Figure 6, this estimator is counting the proportion of samples to the upper-left of the intersection, minus the proportion of samples to the lower-right of the intersection. With this estimator and the same 1000 samples as the brute force Monte Carlo method, we calculate

$$P(\mathbf{J} > \mathbf{J_C}) = 4.1 \pm 1.2\%, \tag{25}$$

where the 3σ confidence interval is obtained from the variance formula (18). Comparing to (22), the variance of the Monte Carlo method is significantly reduced. To achieve the same variance-reduction using the brute force Monte Carlo method, 1778 more Navier-Stokes equations must be solved. In contrast, we achieved the same reduction in variance with merely one additional Navier-Stokes solution and one adjoint solution. This implies that one single adjoint solution, whose cost is 4 times that of a Navier-Stokes solution, achieves the same effect as 1,777 Navier-Stokes equations in this example.

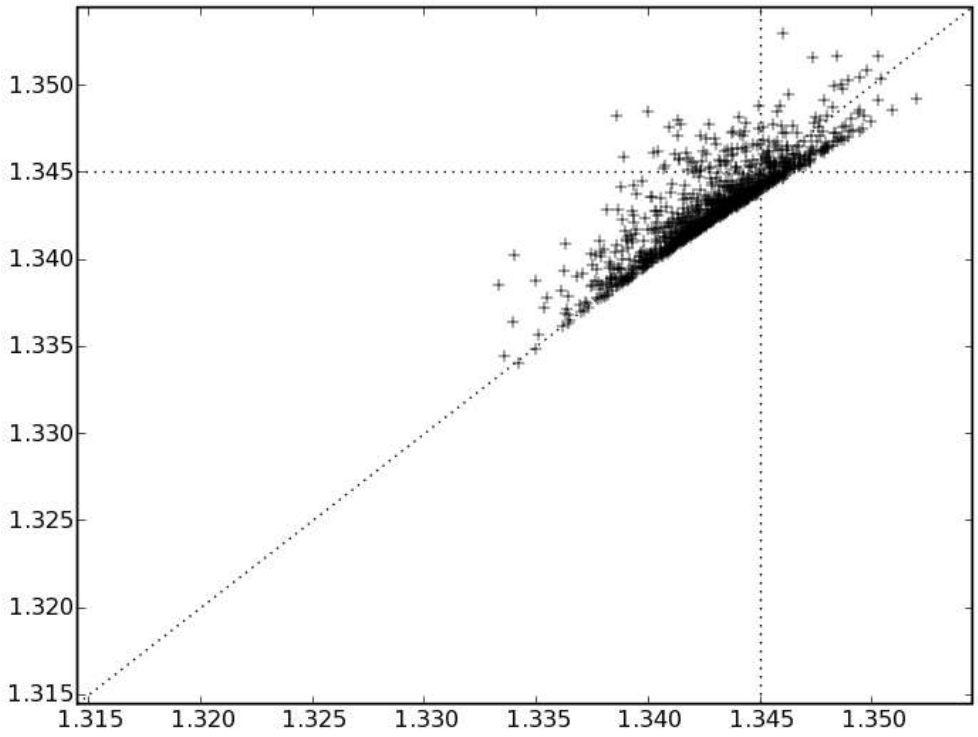

Figure 8. Monte Carlo samples with importance sampling. The horizontal axis is the adjoint approximation; the vertical axis is true objective function. The circular symbol at the center indicates the objective function without rotation.

The variance can be further reduced using the importance sampling technique as discussed in Section III.B. We use 25 samples to calculate the variance of the normalized approximation error $\dfrac{\mathbf{J} - \mathbf{J}_L}{||\xi - \xi_0||^2}$, and approximate it with a Gaussian distribution with zero mean and same variance. This Gaussian distribution is used to calculate $p(\mathbf{J}(\xi) > \mathbf{J_C})$ for each ξ, the probability that a proposed sample is approved. The resulting samples distribute according to probability Q (19), and concentrate in areas where the adjoint approximation $P(\mathbf{J}_L > \mathbf{J_C}) \approx P(\mathbf{J} > \mathbf{J_C})$ is likely to be erroneous. This includes the regions where $\mathbf{J}$ is close to $\mathbf{J_C}$, and areas where $\mathbf{J}$ may be significantly different from its adjoint approximation $\mathbf{J}_L$.

Figure 8 shows the concentration of the approved samples used in importance sampling. We plot the true value of the objective function $\mathbf{J}(\xi_i)$ against its adjoint linear approximation $\mathbf{J}_L(\xi_i)$ for $\xi_1, \ldots, \xi_{1000}$ sampled from probability distribution Q. Each cross on the plot represents one sample. The diagonal dotted line indicates where the horizontal axis is equal to the vertical axis; the horizontal and vertical dotted lines indicate the critical value $\mathbf{J_C}$. As can be seen, most of the samples concentrate near this critical value. Compared to Figure 6, there are many more samples to the upper-left and to the lower-right of the intersection. These additional samples allows a much-reduced variance in estimating the errors made by the adjoint approximation $P(\mathbf{J}_L > \mathbf{J_C})$, and returns a significantly more accurate results. With this method, we calculated

$$P(\mathbf{J} > \mathbf{J_C}) = 4.1 \pm 0.5\%, \tag{26}$$

where the 3σ confidence interval is based on the variance formula (20). Comparing to the result obtained by the brute force Monte Carlo method (22), the variance of the Monte Carlo method is significantly reduced by 75%. To achieve the same variance-reduction using the brute force Monte Carlo method, 15,000 more Navier-Stokes equations would have to be solved. In contrast, we achieved the same reduction in variance with 26 additional Navier-Stokes solutions and one adjoint solution. This implies that one single adjoint solution, whose cost is 4 times that of a Navier-Stokes solution, achieves the same effect as 14974 Navier-Stokes equations in this example.

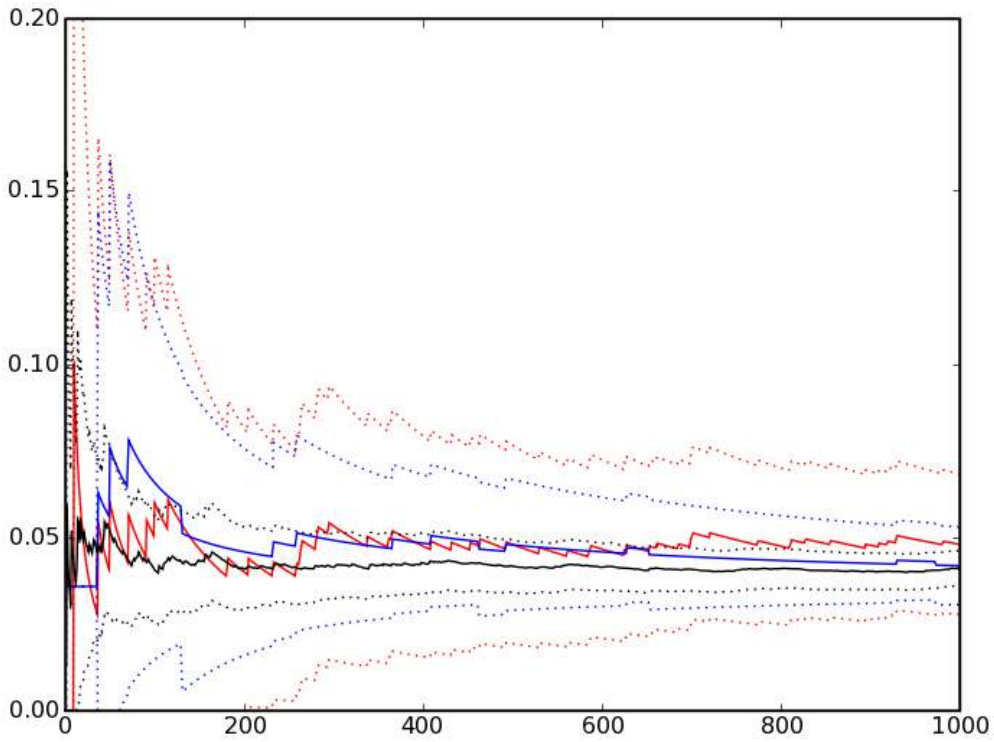

Figure 9. Convergence history of three different Monte Carlo methods: Red is brute-force Monte Carlo method; blue is Monte Carlo with control variate; black is Monte Carlo method with control variates and importance sampling. The horizontal axis indicates the number of samples; the solid lines are the $P(\mathbf{J} > \mathbf{J_C})$ calculated by the estimators of each method; the dotted lines are the 3σ confidence interval bounds of the estimators.

Figure 9 compares the convergence of three different Monte Carlo methods. Different colors represent different methods: red is brute-force Monte Carlo, blue is with redesigned estimator but without importance

sampling, and black is with importance sampling. It can be seen that using the better estimator with importance sampling reduces the standard deviation by a factor of 4, which implies that the number of samples required is reduced by a factor of 16.

V. Conclusion

Table 1 compares the computational cost, accuracy and efficiency of the four methods discussed in this article. As can be seen, using the adjoint equation can significantly increase the accuracy and reduce the computational time in quantification of margins and risk. The adjoint approximation method is by far less expensive than any Monte Carlo method, and is often sufficiently accurate for a first estimate. However, it is difficult to evaluate the accuracy of its result. Both the control variate method and the importance sampling method are accelerated Monte Carlo with adjoint solution. They are marginally more expensive than the brute force Monte Carlo method, but the accuracy is significantly improved.

Table 1. Comparison of the methods for estimating $P(\mathbf{J} > \mathbf{J_C})$.

	Computational time	3σ	Equivalent samples
Brute force Monte Carlo	240 hours	2.0%	1,000
Adjoint approximation	1 hour	—	—
Control variate	241 hours	1.2%	2,778
Importance sampling	247 hours	0.5%	16,000

Acknowledgments

This work was funded by the United States Department of Energy's ASC and PSAAP Programs at Stanford University.

References

[1]Jameson, A., "Aerodynamic Design via Control Theory," *J. of Scientific Computing*, Vol. 3, 1988, pp. 233–260.

[2]Bewley, T. R., Moin, P., and Temam, R., "DNS-based predictive control of turbulence: an optimal target for feedback algorithms," *J. Fluid Mech.*, Vol. 447, 2001, pp. 179–225.

[3]Abergel, F. and Temam, R., "On some control problems in fluid mechanics," *Theoretical and Computational Fluid Dynamics*, Vol. 1, No. 6, 1990, pp. 303–325.

[4]Ham, F., Mattsson, K., Iaccarino, G., and Moin, P., *Towards Time-Stable and Accurate LES on Unstructured Grids. Complex Effects in Large Eddy Simulation*, Vol. 56, Springer, 2007.

[5]Hosea, M. E. and Shampine, L. F., "Analysis and implementation of TR-BDF2," *Appl. Numer. Math.*, Vol. 20, No. 1-2, 1996, pp. 21–37.

[6]Ham, F., "Improved scalar transport for unstructured finite volume methods using grids based on simplex superposition," *Annual Research Briefs*, Center for Turbulence Research, Stanford, CA, 2008.

[7]Wang, Q., *Uncertainty Quantification for Unsteady Fluid Flow using Adjoint-based Approaches*, Ph.D. thesis, Stanford University, Stanford, CA, 2009.

[8]Wang, Q. and Moin, P., "Minimal Repetition Dynamic Checkpointing Algorithm for Unsteady Adjoint Calculation," to appear in SIAM Journal on Scientific Computing.

[9]Fey, U., Konig, M., and Eckelmann, H., "A new Strouhal–Reynolds-number relationship for the circular cylinder in the range $47 \leq Re \leq 2 \times 10^5$," *Physics of Fluids*, Vol. 10, No. 7, 1998, pp. 1547–1549.

[10]Tokumaru, P. T. and Dimotakis, P. E., "Rotary oscillation control of a cylinder wake," *Journal of Fluid Mechanics*, Vol. 224, April 2006, pp. 77–90.

[11]Glynn, P. W. and Szechtman, R., "Some New Perspectives on the Method of Control Variates," *Monte Carlo and Quasi-Monte Carlo Methods 2000*, Springer-Verlag, 2002, pp. 27–49.

[12]Glynn, P. W. and Iglehart, D. L., "Importance sampling for stochastic simulations," *Manage. Sci.*, Vol. 35, No. 11, 1989, pp. 1367–1392.

[13]Wang, Q., Gleich, D., Saberi, A., Etemadi, N., and Moin, P., "A Monte Carlo method for solving unsteady adjoint equations," *Journal of Computational Physics*, Vol. 227, No. 12, June 2008, pp. 6184–6205.

[14]Wang, Q., Moin, P., and Iaccarino, G., "A Rational Interpolation Scheme with Super-Polynomial Rate of Convergence," *Annual Research Briefs*, Center for Turbulence Research, Stanford, CA, 2008, pp. 31–54.

[15]Wang, Q., Moin, P., and Iaccarino, G., "A High-Order Multi-Variate Approximation Scheme for Arbitrary Data Sets," submitted.