KU LEUVEN

Sensitivity calculations for Monte Carlo particle simulations of neutrals in the plasma edge

N. Horsten, S. Carli, W. Dekeyser

KU Leuven, Department of Mechanical Engineering, Leuven, Belgium

EUROfusion

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

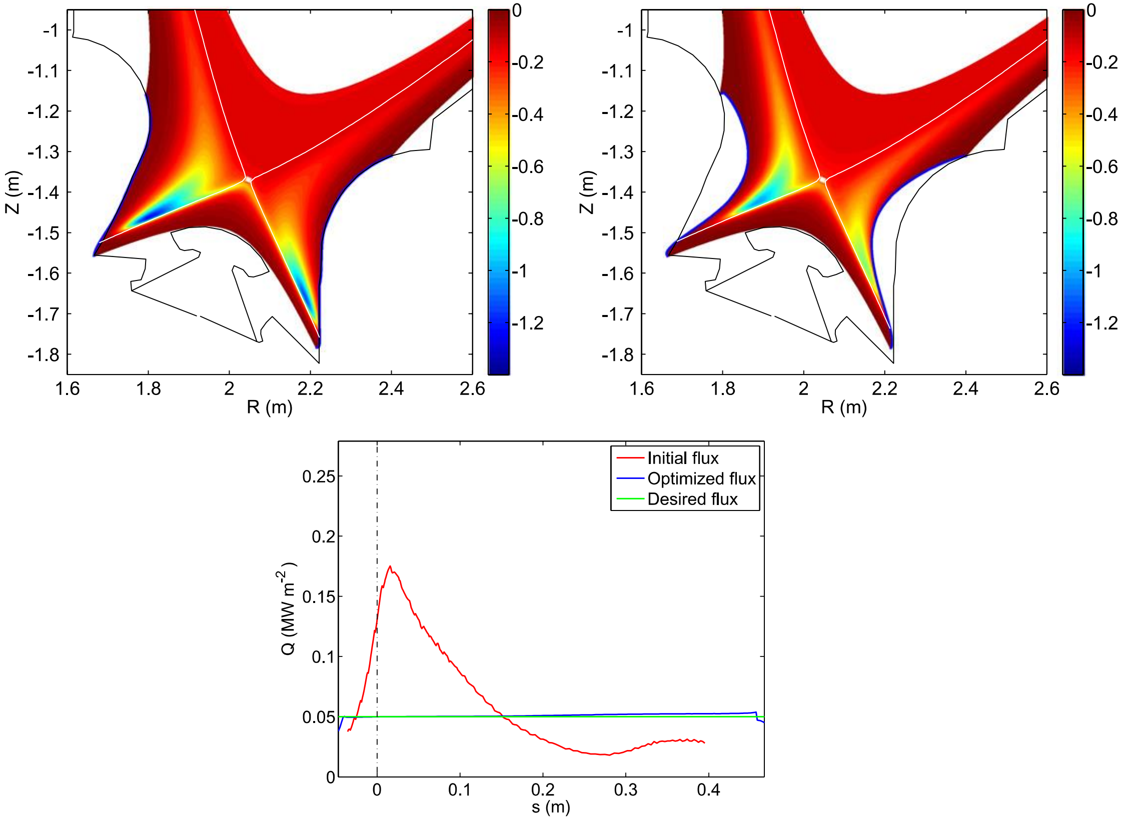
Gradients/sensitivities are extremely useful for efficient optimization calculations

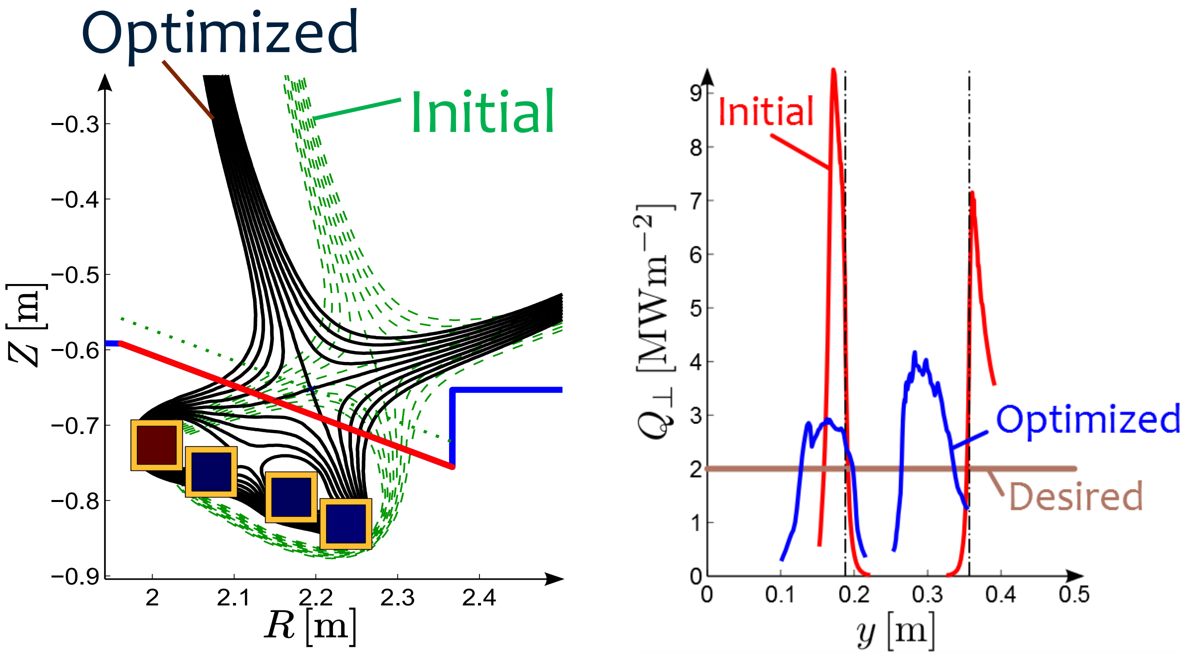
E.g. mitigation of heat load

Cost ≈ 10 x forward simulation

Shape optimization of the divertor

Magnetic field optimization



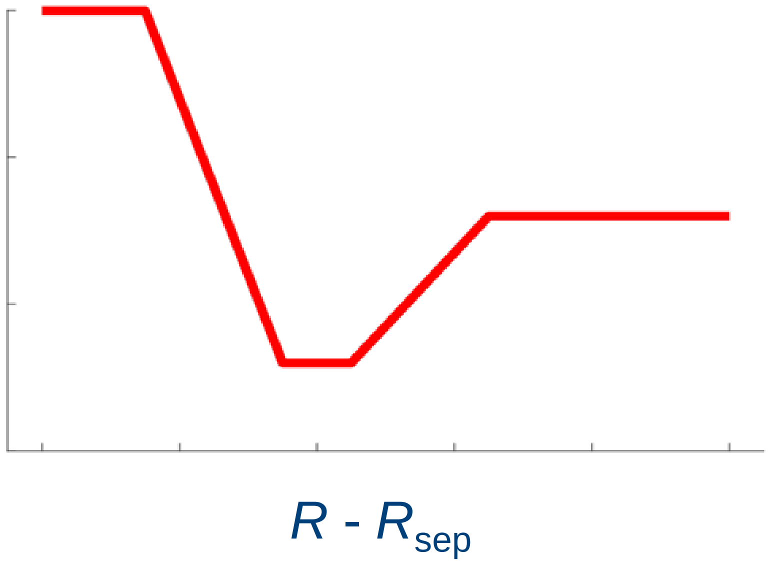


W. Dekeyser et al., NF **54** (2014)

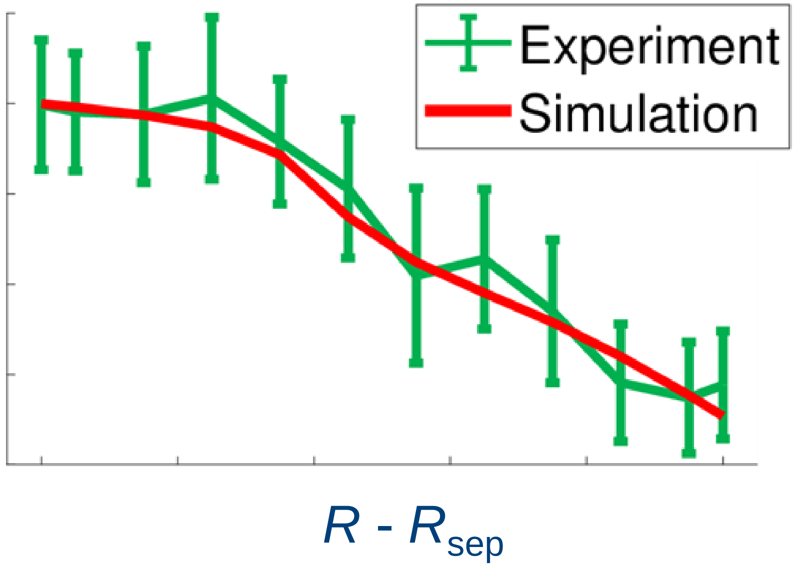
M. Blommaert et al., JNM **463** (2015)

Gradients for Uncertainty Quantification (UQ)

Transport coefficients



Outer midplane profiles



Sensitivities of output quantities of interest w.r.t. model parameters give useful insights in the physics and uncertainties of the model

S. Carli et al., CPP **62** (2022)

All previous work with fluid neutral model!

- KU Leuven simplified in-house plasma edge code (divertor shape and magnetic field optimization)
- B2.5 standalone simulations with Advanced Fluid Neutral (AFN) model (parameter estimation)

Fully deterministic

Simulations with kinetic Monte Carlo neutrals (EIRENE)?

Outline

- Introduction
- **Gradient calculation with Algorithmic Differentiation (AD)**
- Verification with Finite Differences (FD)
- Resulting sensitivities & statistical errors
- Conclusions & outlook

How calculating gradients?

Finite Differences (FD)

- Cost ~ number of parameters
- Truncation + cancellation error

Adjoint equations [M. Baelmans et al., PPCF **56** (2014)]

- Cost independent of number of parameters
- Continuous developments → manual implementation not feasible

Algorithmic Differentiation (AD)

- Exact to machine precision
- *Tangent* AD (~finite differences), *adjoint* AD (~adjoint equations)
- **Correlation preserving for Monte Carlo simulations!**

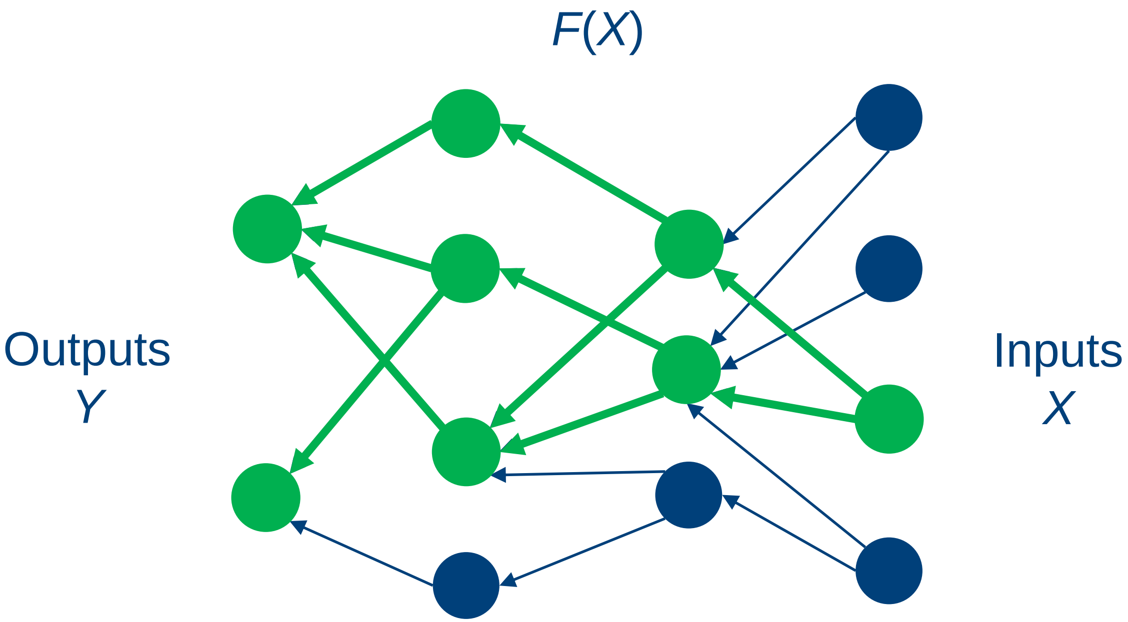
This presentation

TAPENADE generates differentiated code

TAPENADE tool detects all elementary operations, differentiates the source code line by line and creates a new source code with the gradient information

Gradient → $\dot{Y} = F'(X) \times \dot{X} = f_p'(X_{p-1}) \times f_{p-1}'(X_{p-2}) \times \cdots \times f_1'(X) \times \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$

Perturbed input vector



Repeat for each input

L. Hascoët et al., ACM Transaction of Mathematical Software **29** (2013)

Code changes to get AD working

- Started from EIRENE 3.0.8 *develop* branch on ITER repository merged in extended grid code (*feature/wg-release*)

- MsV version still future research ➔ problems due to object-oriented features not supported by TAPENADE?

- Some adaptions for correct interpretation by TAPENADE:
  - Entries removed in eirene.f
  - GOTO statements for throwing error messages in input.f, eirmod_locate.f, colatm.f, colmol.f, colion.f and escape.f replaced by separate subroutines
  - Some issues with pointers
  - Some additional small changes

Outline

- Introduction
- Gradient calculation with Algorithmic Differentiation (AD)
- **Verification with Finite Differences (FD)**
- Resulting sensitivities & statistical errors
- Conclusions & outlook

Sensitivities of total atom content w.r.t. scaling factors different reactions

Fixed JET L-mode background plasmas

Independent variables

Reaction scaling factors

$$X_i = F_i$$

Quantity of interest

Total atom content

$$Y = N_\mathrm{a} = \int\int f_\mathrm{a}(\mathbf{r},\mathbf{v})\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}$$

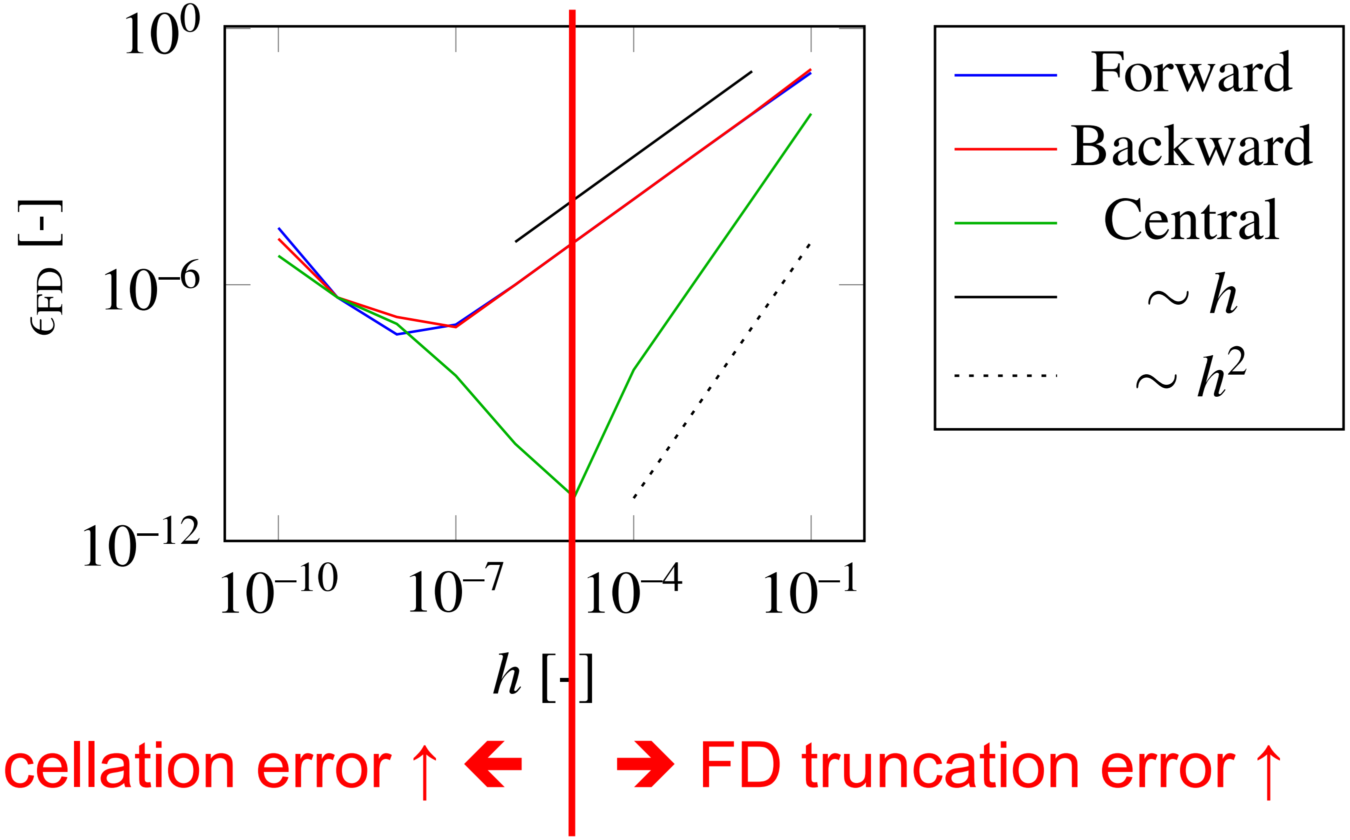
$$\sigma_i' = F_i \times \sigma_i(E)$$

Cross-section

	i	Reaction	Type
D	1	$D + e \rightarrow D^+ + 2e$	EI
	2	$D + D^+ \rightarrow D^+ + D$	CX
D_2	3	$D_2 + e \rightarrow D_2^+ + 2e$	EI
	4	$D_2 + e \rightarrow 2D + e$	DS
	5	$D_2 + e \rightarrow D + D^+ + 2e$	DS
	6	$D_2 + D^+ \rightarrow D_2 + D^+$	EL
	7	$D_2 + D^+ \rightarrow D_2^+ + D$	CX
D_2^+	8	$D_2^+ + e \rightarrow D + D^+ + e$	DS
	9	$D_2^+ + e \rightarrow 2D^+ + 2e$	EI
	10	$D_2^+ + e \rightarrow 2D$	DS

Verification for 1 particle (P = 1) ➔ correct AD gradient

Rel. difference AD & FD $\frac{\partial Y}{\partial F_1}$ (atom ionization)

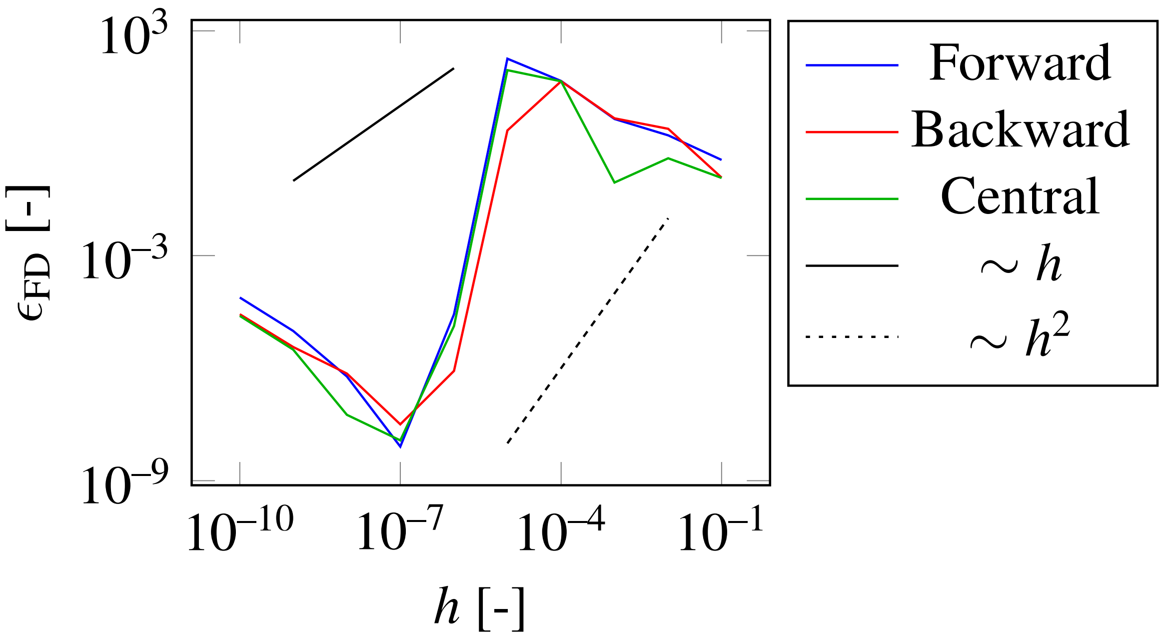


FD: same random seed for original and perturbed simulation

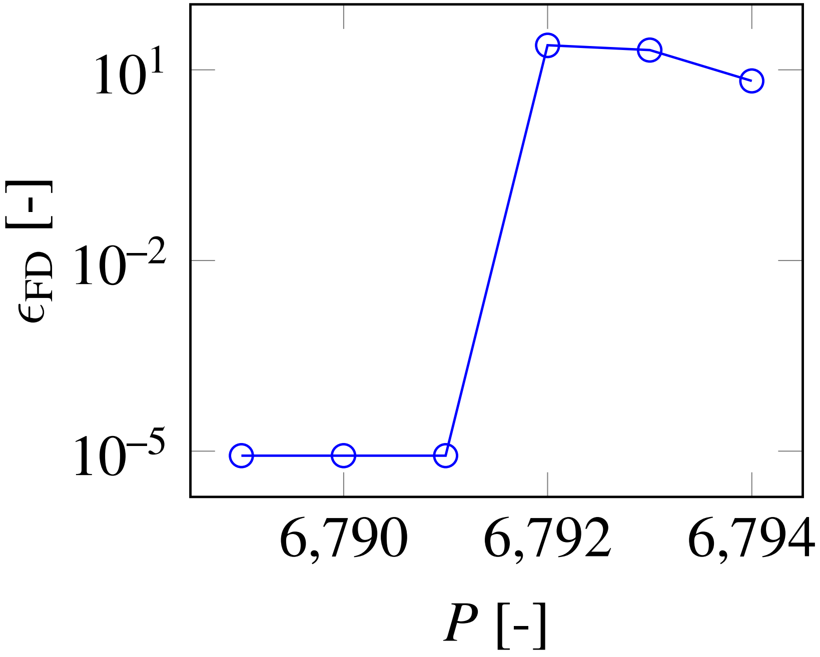
FD cancellation error ↑ 🡄 | ➔ FD truncation error ↑

Large differences between AD and FD when increasing *P*

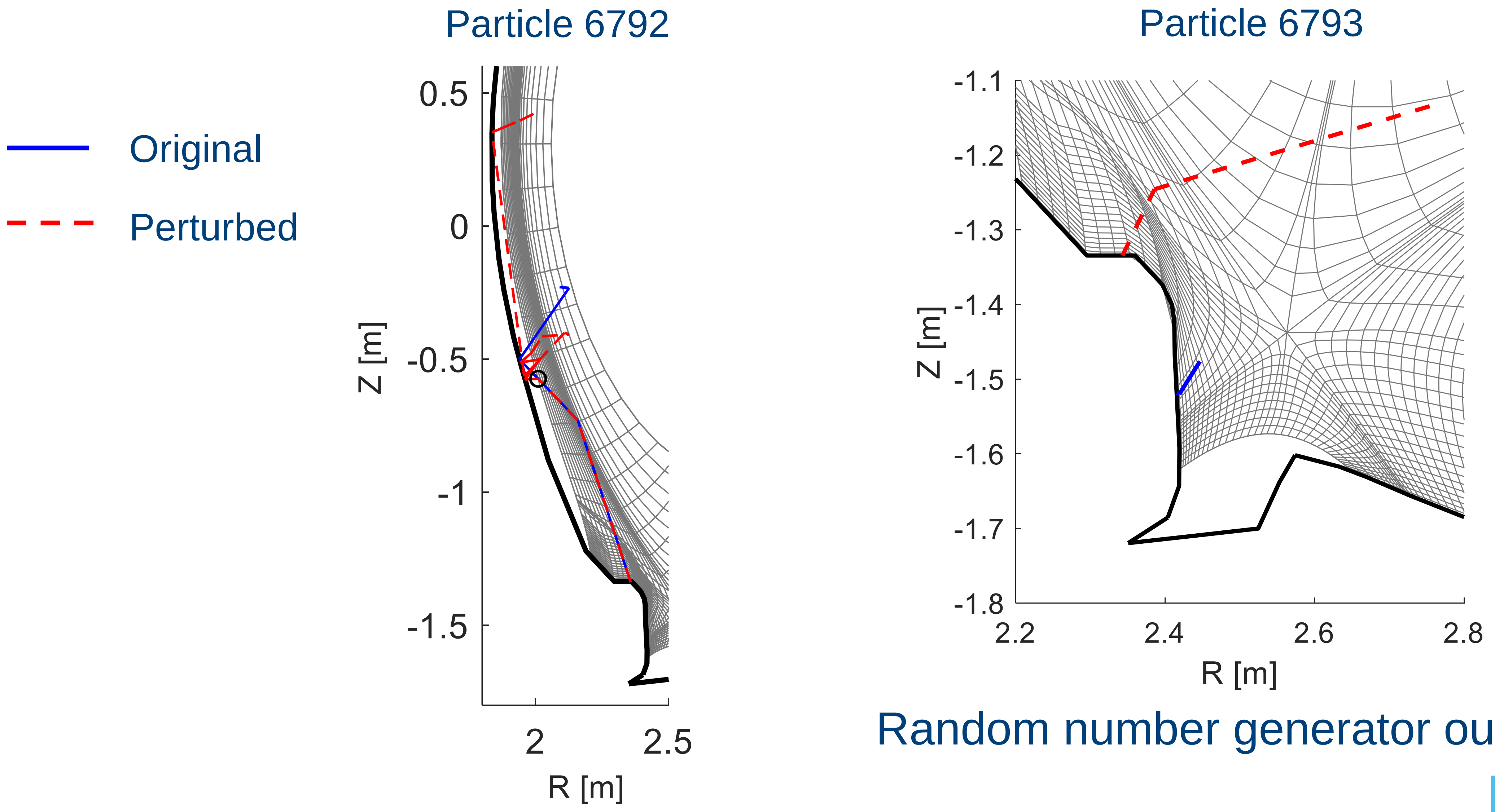
P = 10 000 particles



$h = 10^{-5}$



FD becomes inaccurate due to loss of correlation



Random number generator out of sync!

Outline

- Introduction
- Gradient calculation with Algorithmic Differentiation (AD)
- Verification with Finite Differences (FD)
- **Resulting sensitivities & statistical errors**
- Conclusions & outlook

Definition of sensitivities

Sensitivity resulting from single simulation with
P = 50 000:

$$S_i = \frac{1}{<Y>} \frac{\partial Y}{\partial F_i}$$

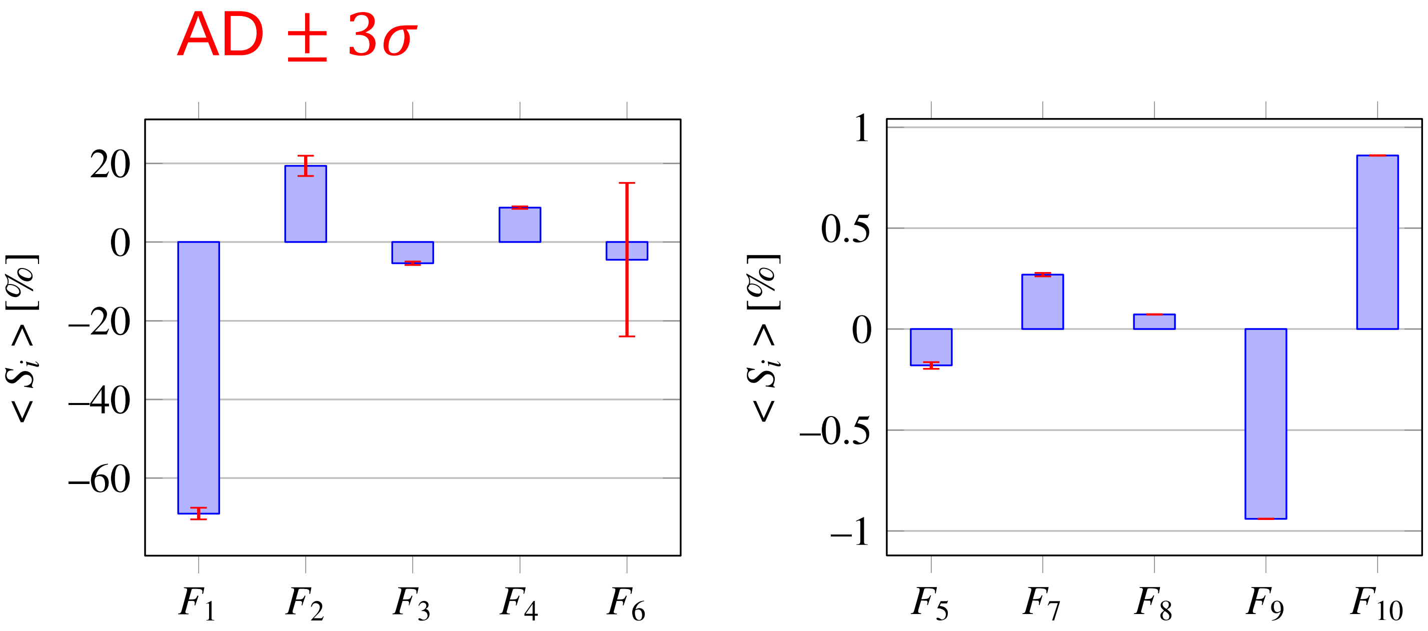
Average sensitivity of 1 000 simulations:

$$< S_i >= \frac{1}{<Y>} \left\langle \frac{\partial Y}{\partial F_i} \right\rangle$$

	i	**Reaction**	**Type**
D	1	D + e → D^+ + 2e	EI
	2	D + D^+ → D^+ + D	CX
D_2	3	D_2 + e → D_2^+ + 2e	EI
	4	D_2 + e → 2D + e	DS
	5	D_2 + e → D + D^+ + 2e	DS
	6	D_2 + D^+ → D_2 + D^+	EL
	7	D_2 + D^+ → D_2^+ + D	CX
D_2^+	8	D_2^+ + e → D + D^+ + e	DS
	9	D_2^+ + e → 2D^+ + 2e	EI
	10	D_2^+ + e → 2D	DS

-69% and +19% sensitivity w.r.t. atom ionization and charge exchange, respectively

Low recycling: $T_{e,ot} \approx 60$ eV



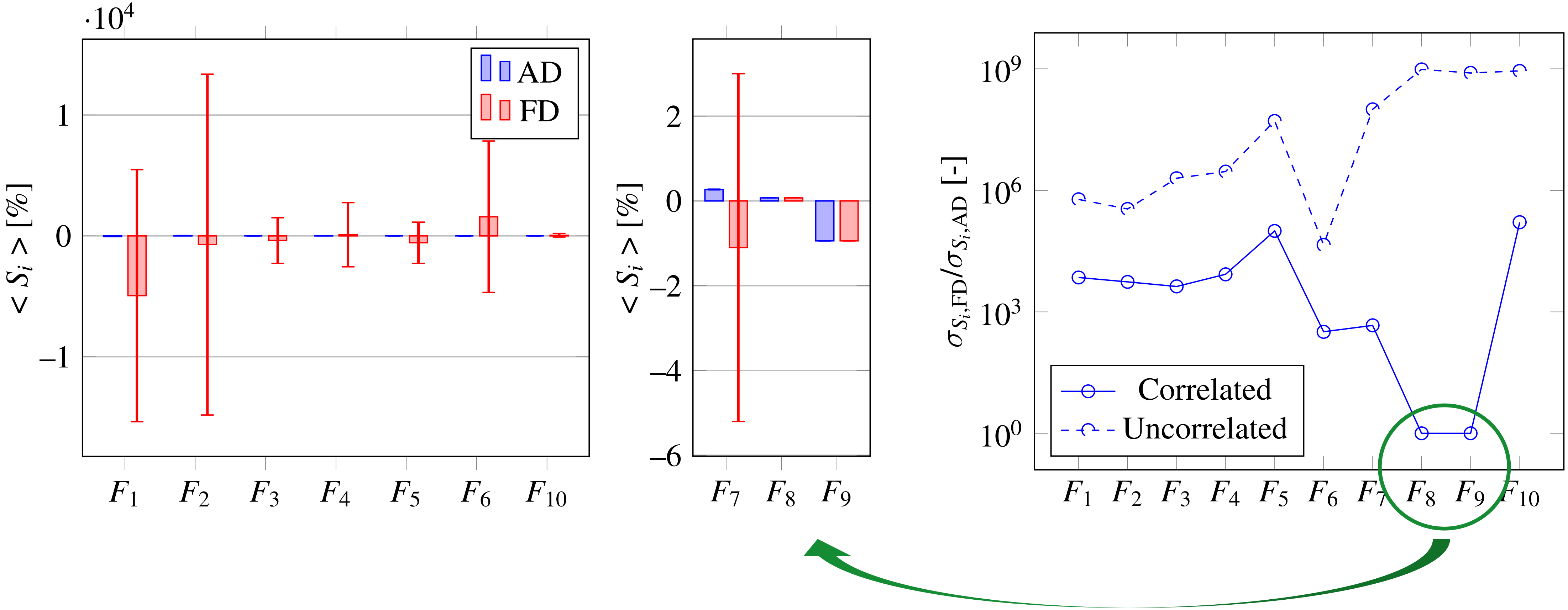
$D + e \rightarrow D^+ + 2e$

$D + D^+ \rightarrow D^+ + D$

$D_2 + D^+ \rightarrow D_2 + D^+$

	i	Reaction	Type
	1	$D + e \rightarrow D^+ + 2e$	EI
	2	$D + D^+ \rightarrow D^+ + D$	CX
D_2	3	$D_2 + e \rightarrow D_2^+ + 2e$	EI
	4	$D_2 + e \rightarrow 2D + e$	DS
	5	$D_2 + e \rightarrow D + D^+ + 2e$	DS
	6	$D_2 + D^+ \rightarrow D_2 + D^+$	EL
	7	$D_2 + D^+ \rightarrow D_2^+ + D$	CX
D_2^+	8	$D_2^+ + e \rightarrow D + D^+ + e$	DS
	9	$D_2^+ + e \rightarrow 2D^+ + 2e$	EI
	10	$D_2^+ + e \rightarrow 2D$	DS

Up to factor 10^5 statistical error reduction AD compared to FD

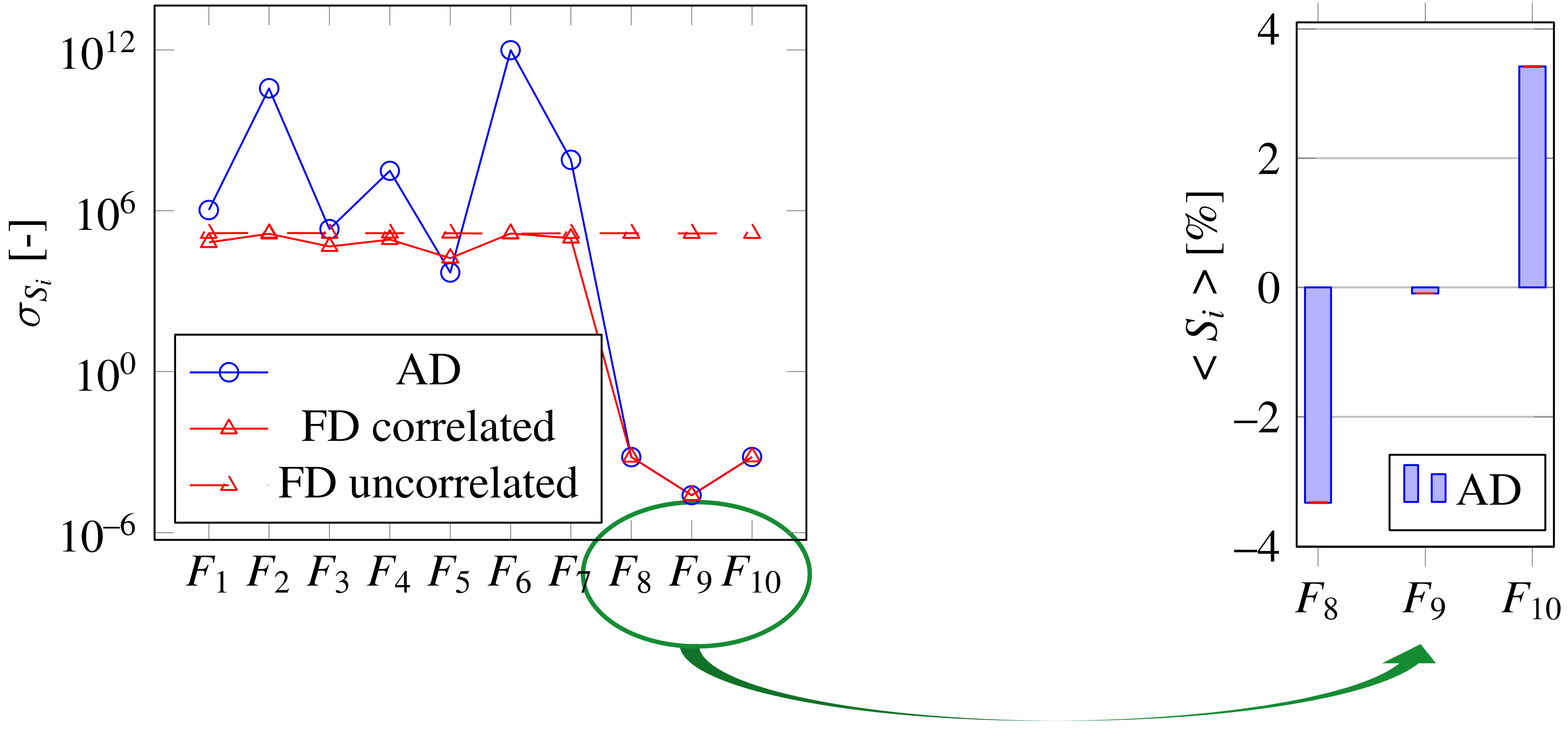


F_8: $D_2^+ + e \rightarrow D + D^+ + e$

F_9: $D_2^+ + e \rightarrow 2D^+ + 2e$

Massive increase in AD statistical error for several sensitivities when moving to detachment

$T_{\mathrm{e,ot}} < 1$ eV

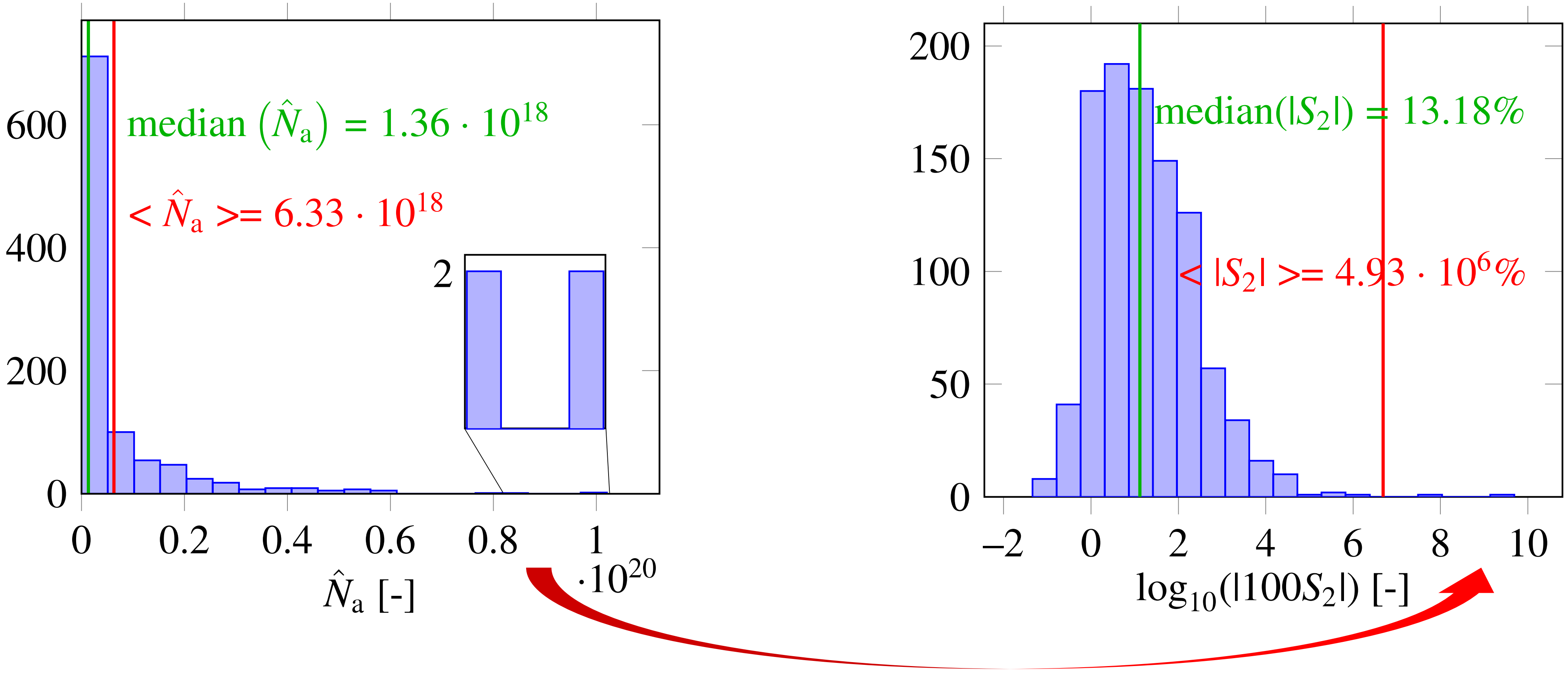


D_2^+ reactions

Statistical error increase caused by a few long-lived particles

Individual contributions
to $\widehat{N}_\text{a}$ (P = 1 000)

Individual contributions
to sensitivity (S_2) w.r.t. CX



Well-known issues with AD for integrators
[J. Hückelheim et al., Understanding AD Pitfalls]

Study of the issue for a uniform infinite plasma

- D atoms only
- Only 1 absorption and 1 scattering rate, R_a and R_s, respectively
- Spatially uniform reaction rates
- No interactions with boundaries

➡ **Feasible to make analytic derivations**

Different combinations of simulation and estimator types

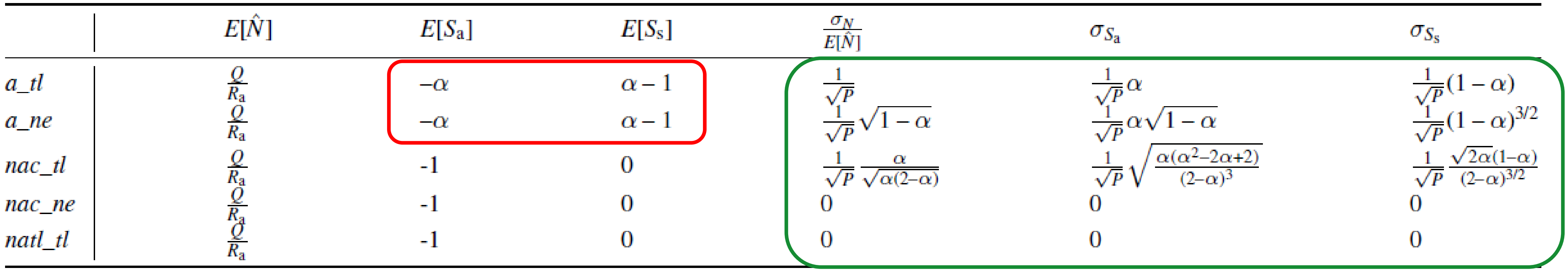
Simulation type	Estimator type
Analog (*a*) ➔ particle weight is kept constant between collisions, particle disappears at absorption events	**Track-length (*tl*)** ➔ Estimation during trajectory
Non-analog collision type (*nac*) ➔ Every collision is a scattering event and particle weight rescaled with $R_s/(R_a+R_s)$	**Next-event (*ne*)** ➔ Estimation at collision + integration until boundary is reached
Non-analog track-length type (*natl*) ➔ Every collision is a scattering event and weight rescaled during trajectory with $\exp(-R_a L/v)$	

6 combinations: *a_tl*, *a_ne*, *nac_tl*, *nac_ne*, *natl_tl* and *natl_ne*

Diverging sensitivities in EIRENE for all tested combinations: *a_tl*, *a_ne* and *nac_tl*

B. Mortier et al., PhD thesis (2020)

Expected values and statistical errors



	$E[\hat{N}]$	$E[S_a]$	$E[S_s]$	$\frac{\sigma_N}{E[\hat{N}]}$	σ_{S_a}	σ_{S_s}
a_tl	$\frac{Q}{R_a}$	$-\alpha$	$\alpha - 1$	$\frac{1}{\sqrt{P}}$	$\frac{1}{\sqrt{P}}\alpha$	$\frac{1}{\sqrt{P}}(1-\alpha)$
a_ne	$\frac{Q}{R_a}$	$-\alpha$	$\alpha - 1$	$\frac{1}{\sqrt{P}}\sqrt{1-\alpha}$	$\frac{1}{\sqrt{P}}\alpha\sqrt{1-\alpha}$	$\frac{1}{\sqrt{P}}(1-\alpha)^{3/2}$
nac_tl	$\frac{Q}{R_a}$	-1	0	$\frac{1}{\sqrt{P}}\frac{\alpha}{\sqrt{\alpha(2-\alpha)}}$	$\frac{1}{\sqrt{P}}\sqrt{\frac{\alpha(\alpha^2-2\alpha+2)}{(2-\alpha)^3}}$	$\frac{1}{\sqrt{P}}\frac{\sqrt{2\alpha}(1-\alpha)}{(2-\alpha)^{3/2}}$
nac_ne	$\frac{Q}{R_a}$	-1	0	0	0	0
natl_tl	$\frac{Q}{R_a}$	-1	0	0	0	0

Q: source strength
P: number of particles
$\alpha = R_a/(R_a+R_s)$

S_a: sensitivity w.r.t. R_a
S_s: sensitivity w.r.t. R_s

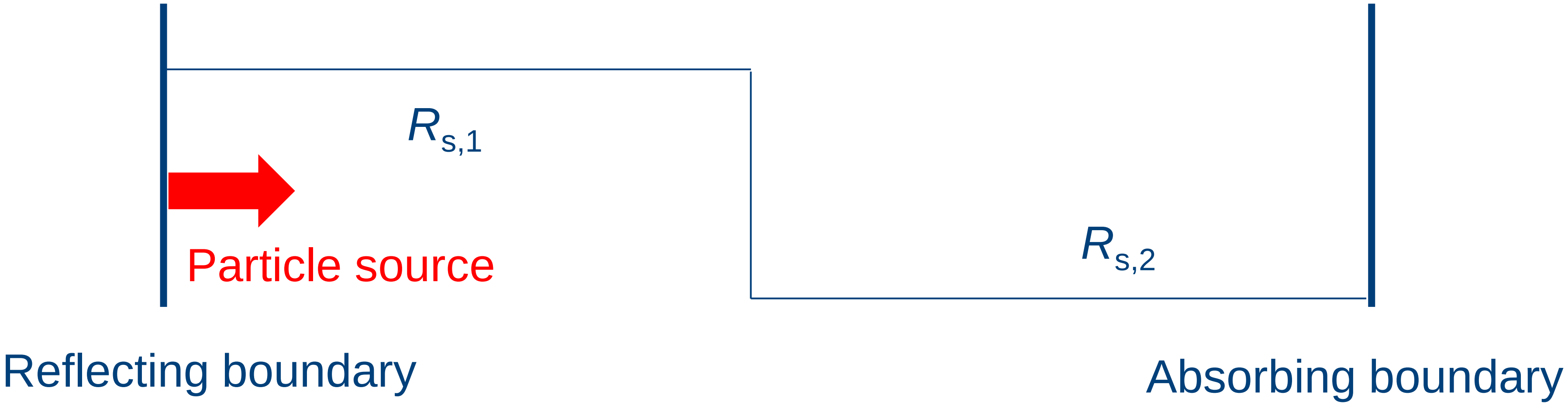
Expected value of gradient ≠ gradient of expected value

Reason: expected value of gradient on number of collisions is zero, whereas the gradient of the expected value is nonzero

Analog simulation type cannot be used!

No declaration for diverging sensitivities

Most simplified case for which I see diverging sensitivities



- 1D model
- Mono-energetic particles
- $R_a = 0$
- Large difference between $R_{s,1}$ and $R_{s,2}$
- Long particle trajectories

Outline

- Introduction
- Gradient calculation with Algorithmic Differentiation (AD)
- Verification with Finite Differences (FD)
- Resulting sensitivities & statistical errors
- **Conclusions & outlook**

Conclusions & outlook

- AD is a promising method for MC particle simulations because it guarantees correlation between the primal and perturbed trajectories

- Statistical error reduction of up to a factor 10^5 compared to FD for several sensitivities for low-recycling conditions

- Problems with high-collisional conditions
  1. Still an issue for coupled fluid plasma – kinetic neutral simulations? [W. Dekeyser et al., CPP **58** (2018); E. Løvbak et al. (2023)]
  2. Try to understand the origin of seemingly diverging sensitivities for simplified settings

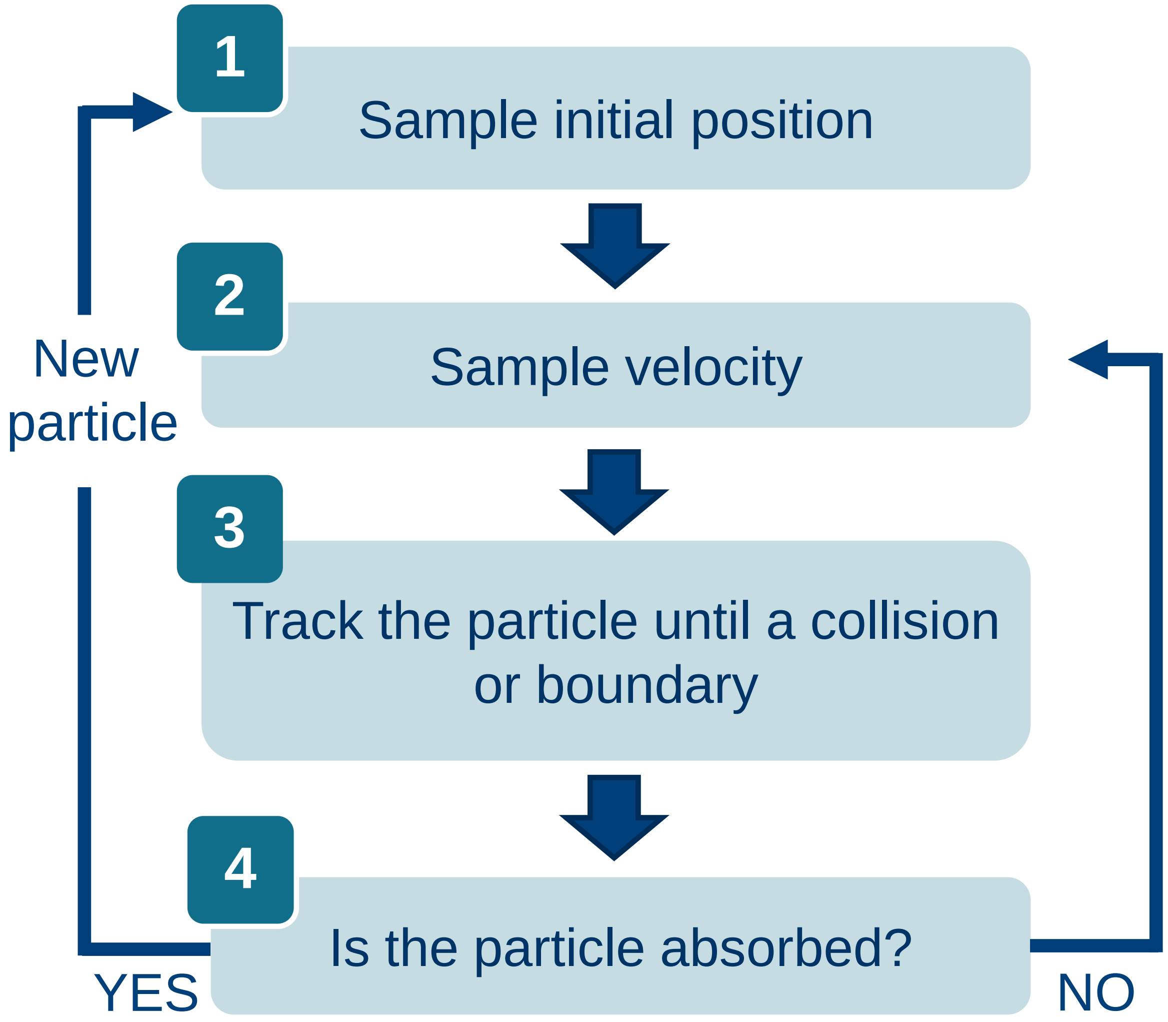
KU LEUVEN

Thank you! Questions?

KU LEUVEN

Back-up slides

Particle tracing MC procedure for neutrals

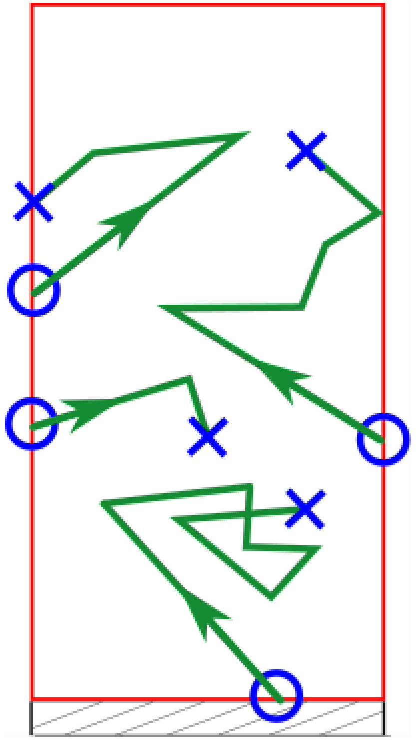


1 Sample initial position

2 Sample velocity

3 Track the particle until a collision or boundary

4 Is the particle absorbed?



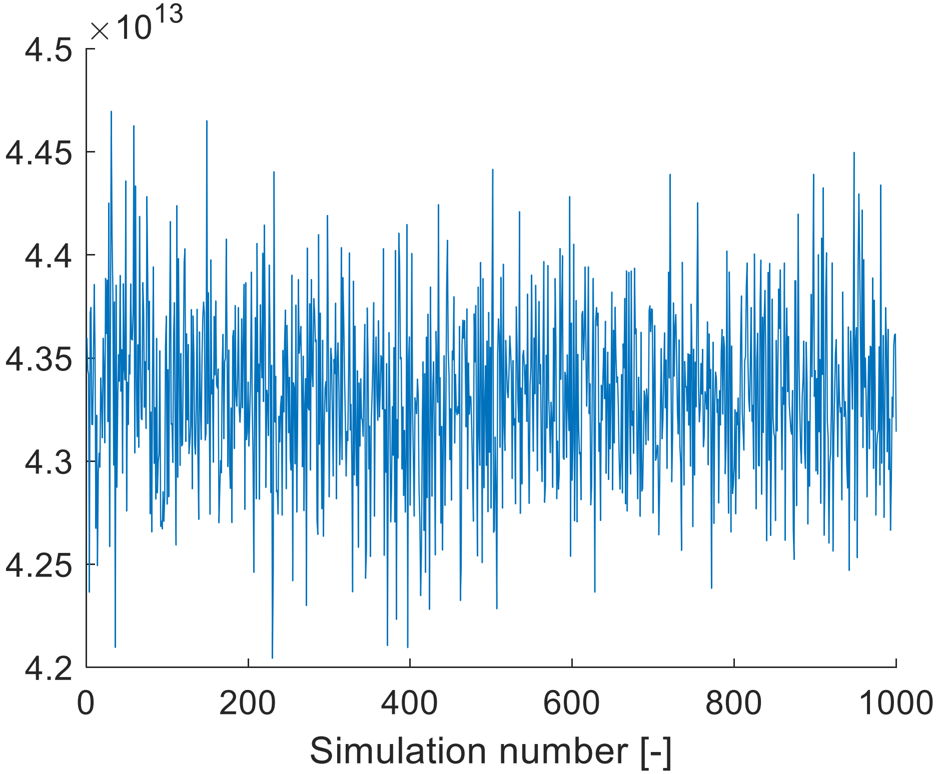
Track-length estimator:

$$\hat{N}_{\mathrm{a}} = Q \sum_{i=1}^{N} w_i \frac{l_i}{v_i}$$

w_i	Particle weight
l_i	Traveled distance between 2 collisions
v_i	Particle speed

Noisy results complicate sensitivity calculation

Total atom content [-] ➡ Obtained by integrating particle trajectories



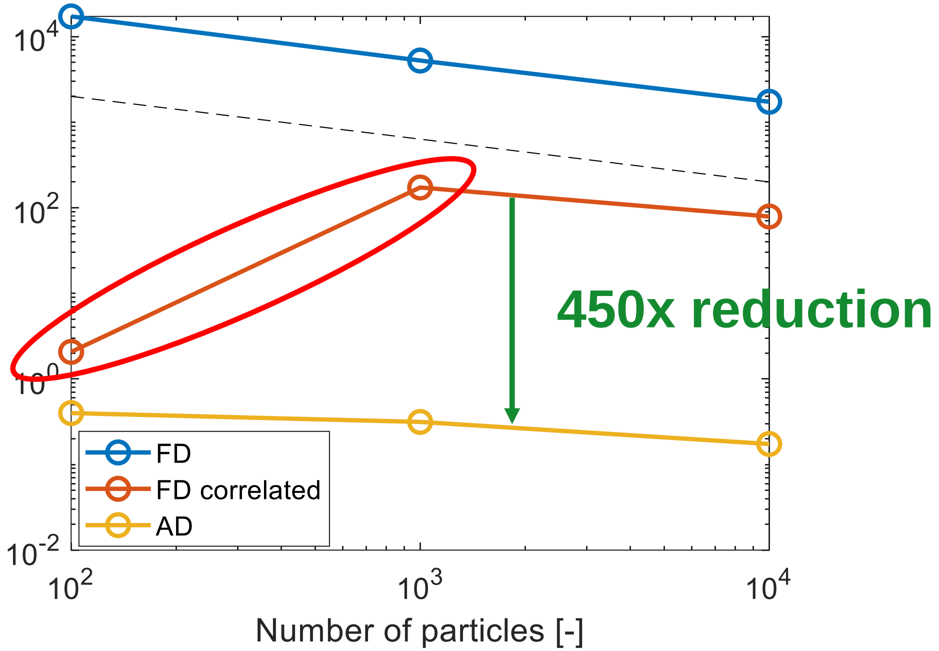
Critical to have **correlated particle trajectories** in the primal and perturbed simulation!

$$\frac{\mathrm{d}Y}{\mathrm{d}X} = \frac{Y(X+\Delta X) - Y(X-\Delta X)}{2\Delta X}$$

$$\begin{aligned}\mathrm{Var}\left[\frac{\mathrm{d}Y}{\mathrm{d}X}\right] &= \frac{1}{4(\Delta X)^2}(\mathrm{Var}[Y(X+\Delta X)]\\ &+\mathrm{Var}[Y(X-\Delta X)]\\ &-2\mathrm{Cov}[Y(X+\Delta X), Y(X-\Delta X)])\end{aligned}$$

Standard deviation AD factor 450 lower than FD

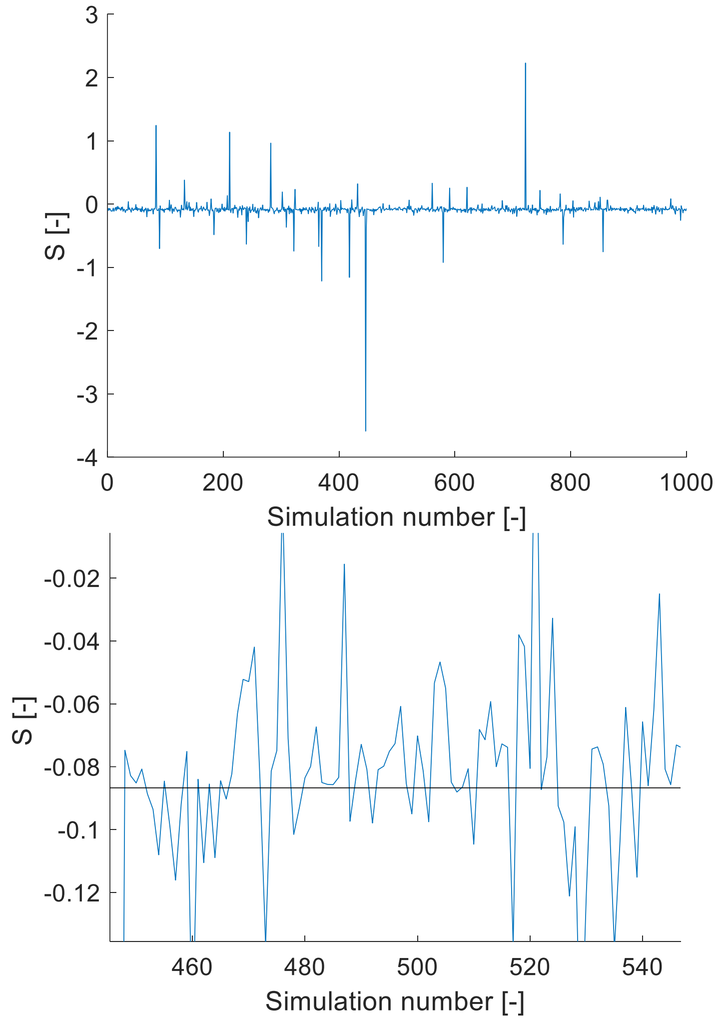
Standard deviation sensitivity [-]



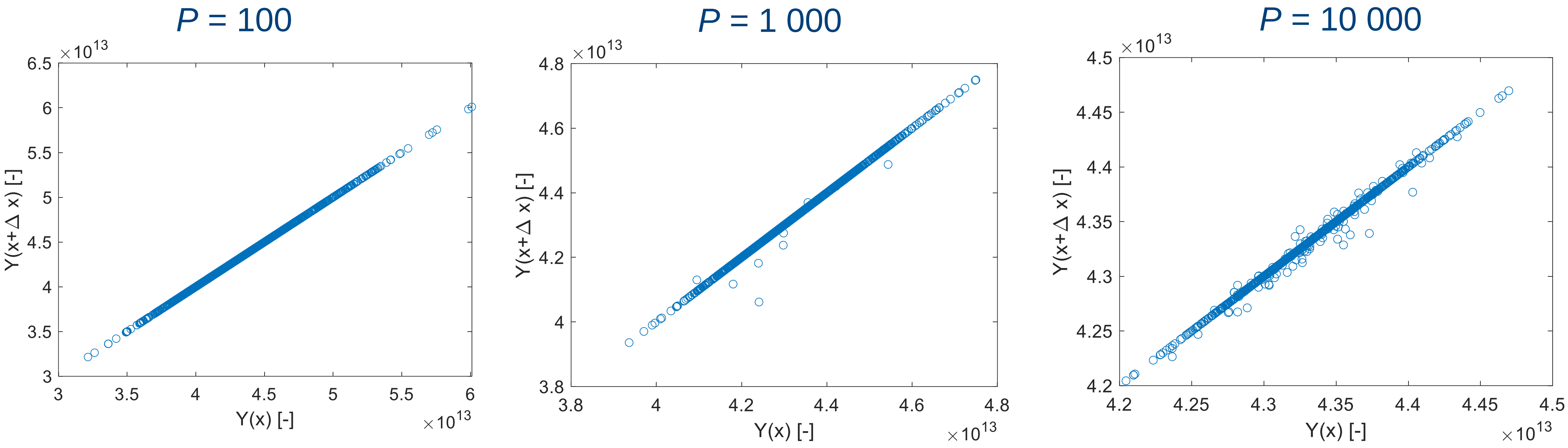
Increased probability for loss of correlation

Strong spikes in sensitivity

AD with P = 10 000

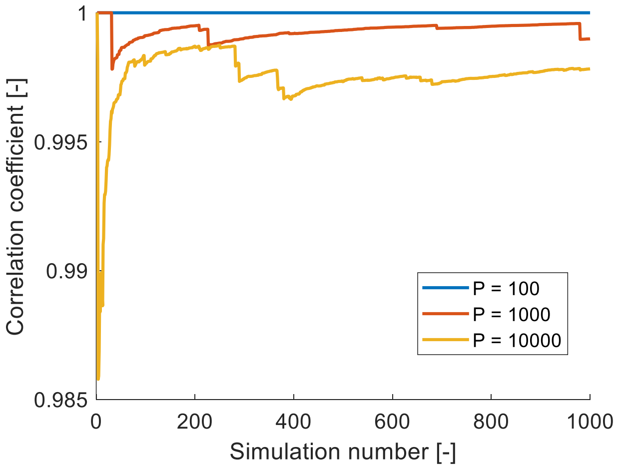


Larger probability for loss of correlation in FD for higher *P*


P = 100

P = 1 000

P = 10 000



Correlation coefficient could be improved by reinitializing the random seed for each particle

Spatial resolution of sensitivities w.r.t. reaction rates

Different independent variable:

Independent variable

Ionization rate scaling factor

$$x = F_{\text{ion}}$$

→

Objective function

Atom density

$$J = n_{\text{a}}(\mathbf{r}) = \int f_{\text{a}}(\mathbf{r}, \mathbf{v}) \mathrm{d}\mathbf{v}$$

Next slide: P = 50 000

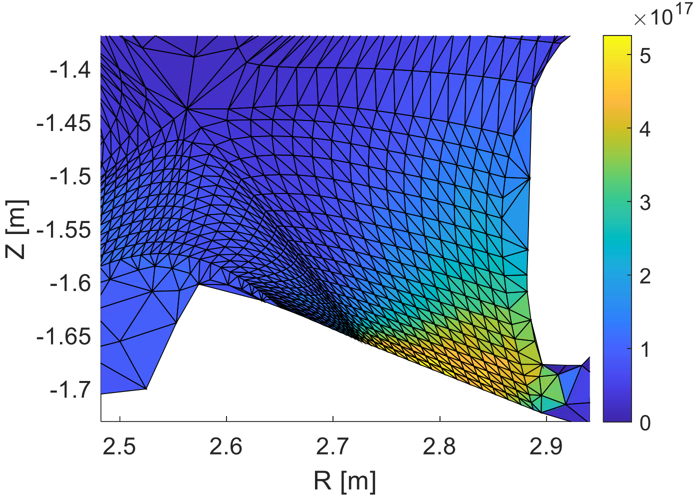
$\mathrm{d}n_{\text{a}}/\mathrm{d}F_{\text{ion}}$ capped between -10^{18} and 10^{18}

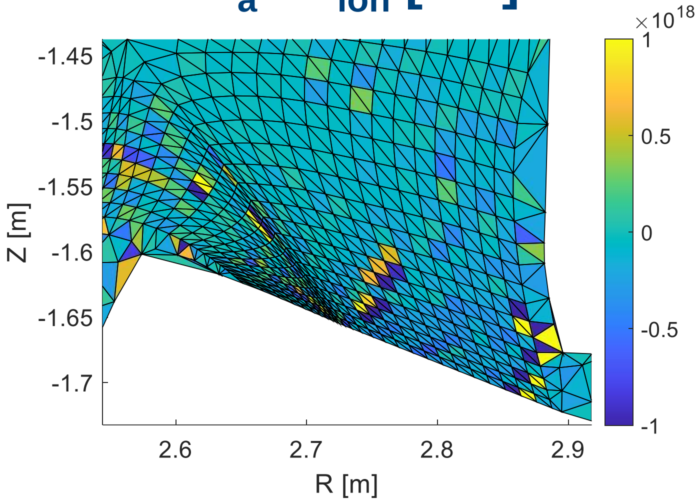
Very noisy spatially resolved AD sensitivities

n_a [m^{-3}]

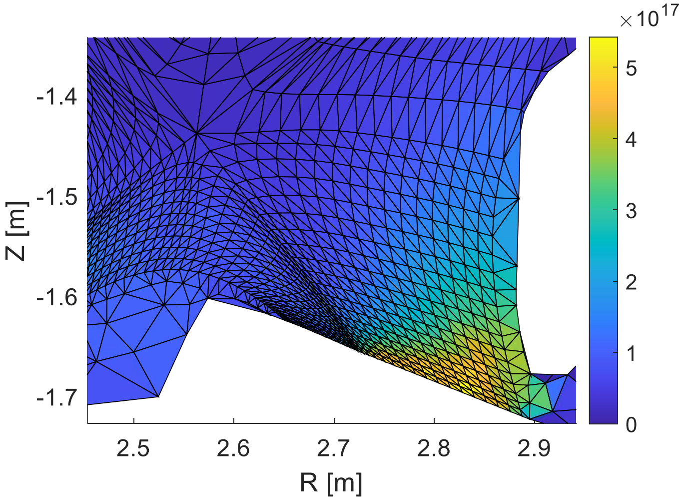
dn_a/dF_{ion} [m^{-3}]

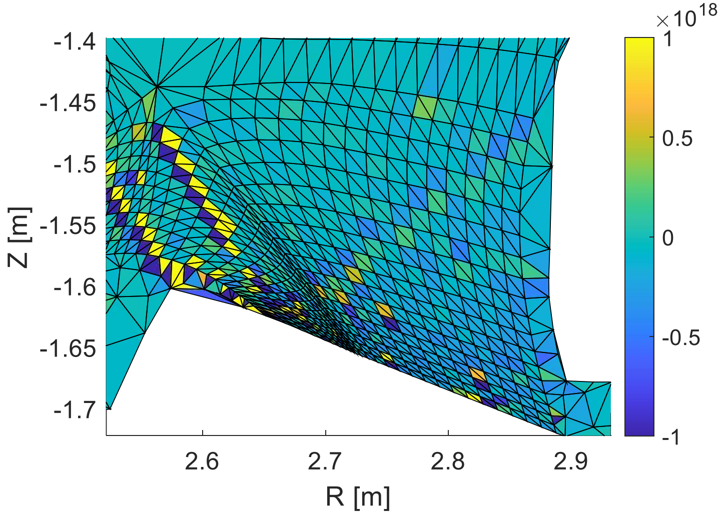
Random seed 1



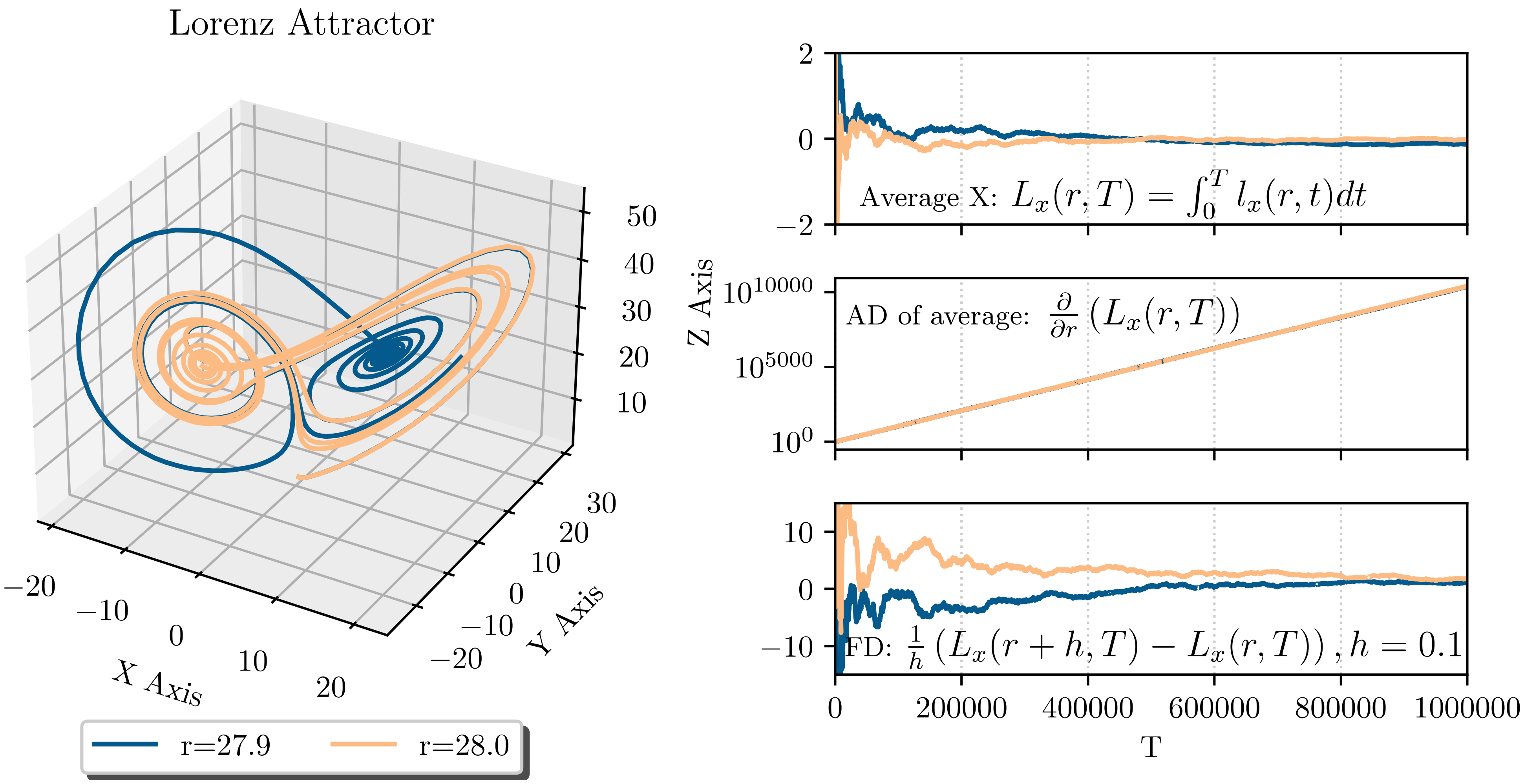


Random seed 2





Similar issue as for sensitivities of chaotic systems?



Problems

P	***S***
50	-5.56e3 %
100	-3.45e5 %