Trajectory Sensitivities and Parametric Uncertainty in Power System Dynamics

IEEE CDC 2021 Workshop: Uncertainty Management in Power System Dynamics

Daniel Adrian Maldonado

Mathematics and Computing Science
Argonne National Laboratory

December 11, 2021

Motivation

- Renewable generation will change how we manage the grid.
  - **Uncertainty**. Weather fluctuations. Higher complexity! (uDERs)
  - **Dynamics**. Inverter resources. Ancillary services.
- Uncertainty Quantification (UQ) in dynamics is challenging but exciting area of research.
- **However.** The problem needs better definition. What role does uncertainty play in power system dynamics? What are desirable properties of UQ techniques?
- I will present some two methods:
  1. Method of Moments.
  2. Worst-case trajectory via optimization.

Sensitivities are a great technique to have in your research tool belt.

About me

My bias:

- I was trained as an Electrical Engineer. Using high performance computing techniques to simulate large-scale power system dynamics.
- Conversations with industry partners. Parameters not known accurately. Especially load models and now, distributed energy resources.
- Focused on statistical techniques, inverse problems, etc.
- Joined Argonne's MCS. Work alongside mathematicians and statisticians. Optimization and Data Assimilation group. Adjoints, AD, etc. Focus on leveraging HPC.

Acknowledgements

Mihai Anitescu (ANL) , Emil Constantinescu (ANL), Vivak Patel (UWisconsin), Hong Zhang (ANL), Michel Schanen (ANL), Vishwas Rao (ANL), François Pacaud(ANL), Tyler Maltba (UCBerkeley), Junbo Zhao (UConn)

Power System Dynamics

- The electrical power system is a complex machine that transforms energy from diverse sources into electrical energy and then distributes it to end users.
- Transitory processes occur due to normal operations or perturbances. Power system dynamics studies are concerned with:
  1. The stability of the system after such process.
  2. The mitigation of the transient regime.
  3. The quality of the transition between steady-state regimes.
- A widely-used power system dynamics model uses a DAE:

$$\dot{x} = f(t, x, y, \theta)\,,$$
$$0 = g(t, x, y, \theta)\,.$$

- Owing to control mechanisms and physical constraints, functions f often present nonlinearities such as saturation, state limiters, switches.

Uncertainty in Dynamics

Let $\dot{x} = f(t, x, \theta)$ be the diff. equation that represents our dynamics.

- When parameters vary with time stochastically. We might use a forcing s. process

$$\dot{x} = f(t, x, \theta) + (x(t), t)dW(t).$$

- When parameters do not vary with time but they have a known p.d.f

$$\theta \sim \pi().$$

- When parameters do not vary with time and do not have known p.d.f

$$\theta \in [\theta_m, \theta_M].$$

Different methods depending on: nature of uncertainty, problem, requirements.

Sensitivities

- I will introduce MoM.
- Before that, we will survey the sensitivity equations.

Sensitivities

We consider an ODE for simplicity:

$$\frac{dx}{dt} = f(t, x, \theta), \quad t_0 \leq t \leq t_F, \quad x(t_0) = x_0(\theta)\,.$$

where $x \in \mathbf{R}^n, \theta \in \mathbf{R}^p$. Sensitivity with respect to θ_k, $u_k(t) = \frac{dx(t)}{d\theta_k}$

$$\begin{aligned} \frac{du_k}{dt} &= f_x(t, x, \theta)u_k + f_\theta(t, x, \theta)e_k\,, \\ u_k(t_0) &= \frac{dx(0)}{d\theta_k}\,. \end{aligned}$$

with $f_x \in \mathbf{R}^{n \times n}$ and $f_\theta \in \mathbf{R}^{n \times p}$.

Sensitivity equation, tangent linear model, variational equation.

Sensitivities II

For a DAE:

$$\begin{aligned} \dot{x} &= f(x, y, \theta)\,, \\ 0 &= g(x, y, \theta)\,. \end{aligned}$$

Define $\alpha = \theta_k$, $f_\alpha = f_\theta e_k$, $u_x^\alpha = \begin{bmatrix} \frac{\partial x_1}{\partial \alpha} & \dots \end{bmatrix}$. Then, we have:

$$\begin{bmatrix} \dot{u}_x^\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} f_x & f_y \\ g_x & g_y \end{bmatrix} \begin{bmatrix} u_x^\alpha \\ u_y^\alpha \end{bmatrix} + \begin{bmatrix} f_\alpha \\ g_\alpha \end{bmatrix} .$$

Integrating this with backward Euler:

$$\begin{bmatrix} \Delta t f_x - I & \Delta t f_y \\ g_x & g_y \end{bmatrix} \begin{bmatrix} \{u_x^\alpha\}^t \\ \{u_y^\alpha\}^t \end{bmatrix} = \begin{bmatrix} -\Delta t(f_\alpha) - \{u^\alpha\}^{t-1}x \\ -g_\alpha \end{bmatrix} .$$

Hybrid systems, discrete events [Hiskens and Pai, 2000].

Sensitivities III

2nd-order ([Choi et al., 2017], [Geng and Hiskens, 2019]). Define $\alpha = \theta_k$, $\beta = \theta_i$, $v^{\alpha\beta} = \begin{bmatrix} \frac{\partial x_1}{\partial \alpha \beta} & \dots \end{bmatrix}$.

$$\begin{bmatrix} \dot{v}_x^{\alpha\beta} \\ 0 \end{bmatrix} = \begin{bmatrix} f_x & f_y \\ g_x & g_y \end{bmatrix} \begin{bmatrix} v_x^{\alpha\beta} \\ v_y^{\alpha\beta} \end{bmatrix} + \xi(x, y, u)\,,$$

where the forcing term, ξ, is:

$$\begin{bmatrix} f_{\alpha\beta} \\ g_{\alpha\beta} \end{bmatrix} + \begin{bmatrix} f_{\alpha x} & f_{\alpha y} \\ g_{\alpha x} & g_{\alpha y} \end{bmatrix} \begin{bmatrix} u_x^{\alpha} \\ u_y^{\alpha} \end{bmatrix} \\ + \begin{bmatrix} f_{\beta x} & f_{\beta y} \\ g_{\beta x} & g_{\beta y} \end{bmatrix} \begin{bmatrix} u_x^{\beta} \\ u_y^{\beta} \end{bmatrix} + (I_m \otimes (u^{\alpha})^T)\mathcal{H} u^{\beta}\,.$$

For $\theta \in \mathbf{R}^p$, we will have p $n-$dimensional vectors of first-order sensitivities, p $n-$dimensional vectors of second-order self-sensitivities, and $\frac{p^2-p}{2}$ $n-$dimensional vectors of second-order mixed sensitivities.
Notice, however, that we only need to factorize one matrix.

Sensitivities

Canonical form. Some times useful to see parameters as initial conditions:

$$\frac{d}{dt}\begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} f(t, x, u) \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x(t_0) \\ u(t_0) \end{bmatrix} = \begin{bmatrix} x_0 \\ \theta \end{bmatrix}$$

Method of Moments

Method of Moments

Motivation

- We would like to obtain $p(x(t))$ given $x(0) \sim \pi_0(x_0)$.
- Obtaining the whole density can be challenging and costly.
- What if we look at first two moments of $p(x(t))$.

The evolution of such density is determined by the *map* of the diff. eq:

$$x_{k+1} = \phi(x_k)$$

Hence, we seek the moments of

$$\mathbb{E}\left[x_{k+1}\right] = \mathbb{E}\left[\phi(x_k)\right]$$
$$\mathbb{V}\left[x_{k+1}\right] = \mathbb{V}\left[(\phi(x_k) - \mu_k)^2\right]$$

Old idea ([Pugachev, 1965]): $\mathbb{E}\left[g(x)\right] \approx g(\mathbb{E}\left[x\right])$.

Approximation of Moments via Taylor series

Given x a r.v. with known moments we want to obtain $\mathbb{E}[g(x)]$. We Taylor expand $g(x)$ around μ and take expectations:

$$\mathbb{E}[g(x)] = g(\mu) + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 g}{\partial x_i \partial x_j}\mathbb{E}\left[(x_i-\mu_{x_i})(x_j-\mu_{x_j})\right]$$
$$+\frac{1}{3!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^3 g}{\partial x_i \partial x_j \partial x_k}\mathbb{E}\left[(x_i-\mu_{x_i})(x_j-\mu_{x_j})(x_k-\mu_{x_k})\right]$$
$$+\frac{1}{4!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\frac{\partial^4 g}{\partial x_i \partial x_j \partial x_k \partial x_l}\mathbb{E}\left[(x_i-\mu_{x_i})(x_j-\mu_{x_j})(x_k-\mu_{x_k})(x_l-\mu_{x_l})\right],$$

- In the case of the map $\phi()$, the derivatives are the sensitivities (also, [Hiskens and Alseddiqui, 2006])!
- If we start with known pdf for x_0, we can do this for $x_1 = \phi(x_0)$, but what about $x_2 = \phi(\phi(x_0))$?
- Closure problem.

Gaussian Closure

We assume that x_0 has a normal distribution and that x_i remains Gaussian. This has two consequences:

- Odd moments are zero.
- Isserlis theorem. Higher-order even moments determined with first two moments.

For instance, taking $c_{ij} = \mathbb{E}\left[(x_i - \mu_i)(x_j - \mu_j)\right]$, we have:

$$\mathbb{E}\left[(x_i - \mu_i)(x_j - \mu_j)(x_k - \mu_k)(x_l - \mu_l)\right] = c_{ij}c_{kl} + c_{il}c_{jk} + c_{ik}c_{lj}$$

Higher-order moment terms in the expansions can be computed with first and second-order moments.

Gaussian Closure Expressions

Propagation of the mean and covariance:

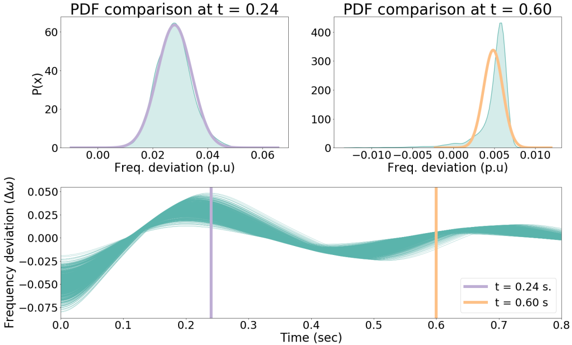
$$\mathbb{E}\left[g(x)\right] = g(\mu) + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 g}{\partial x_i \partial x_j} c_{ij} + \frac{1}{4!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\frac{\partial^4 g}{\partial x_i \partial x_j \partial x_k \partial x_l}\left(c_{ij} c kl + c_{il} c_{jk} + c_{ik} c_{lj}\right),$$

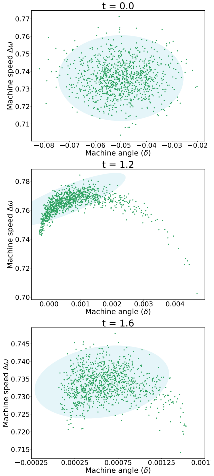
$$\begin{aligned}
c_{pq}^{g} = {} & \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial g_p}{\partial x_j}\frac{\partial g_q}{\partial x_i} + \frac{\partial g_q}{\partial x_j}\frac{\partial g_p}{\partial x_i}\right) c_{ij} + \frac{1}{3!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{\partial g_p}{\partial x_k}\frac{\partial^2 g_q}{\partial x_i x_j} + \frac{\partial g_p}{\partial x_k}\frac{\partial^2 g_q}{\partial x_i x_j} + \frac{\partial g_p}{\partial x_i}\frac{\partial^2 g_q}{\partial x_j x_k} + \frac{\partial g_q}{\partial x_k}\frac{\partial^2 g_p}{\partial x_i x_j} + \frac{\partial g_q}{\partial x_j}\frac{\partial^2 g_p}{\partial x_i x_k} + \frac{\partial g_q}{\partial x_k}\frac{\partial^2 g_p}{\partial x_i x_j}\right) c_{ijk} \\
& + \frac{1}{4!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\left(\frac{\partial g_p}{\partial x_i}\frac{\partial^3 g_q}{\partial x_j x_k x_l} + \frac{\partial g_p}{\partial x_j}\frac{\partial^3 g_q}{\partial x_i x_k x_l} + \frac{\partial g_p}{\partial x_k}\frac{\partial^3 g_q}{\partial x_i x_j x_l} + \frac{\partial g_p}{\partial x_l}\frac{\partial^3 g_q}{\partial x_i x_j x_k} + \frac{\partial^2 g_p}{\partial x_i x_j}\frac{\partial^2 g_q}{\partial x_k x_l} + \frac{\partial^2 g_p}{\partial x_i x_k}\frac{\partial^2 g_q}{\partial x_j x_l}\right. \\
& \left. + \frac{\partial^2 g_p}{\partial x_i x_l}\frac{\partial^2 g_q}{\partial x_j x_k} + \frac{\partial^2 g_p}{\partial x_j x_k}\frac{\partial^2 g_q}{\partial x_i x_l} + \frac{\partial^2 g_p}{\partial x_j x_l}\frac{\partial^2 g_q}{\partial x_i x_k} + \frac{\partial^2 g_p}{\partial x_k x_l}\frac{\partial^2 g_q}{\partial x_i x_j} + \frac{\partial^3 g_p}{\partial x_j x_k x_l}\frac{\partial g_q}{\partial x_i} + \frac{\partial^3 g_p}{\partial x_i x_k x_l}\frac{\partial g_q}{\partial x_j} + \frac{\partial^3 g_p}{\partial x_i x_j x_l}\frac{\partial g_q}{\partial x_k} + \frac{\partial^3 g_p}{\partial x_i x_j x_k}\frac{\partial g_q}{\partial x_l}\right) c_{ijkl} \\
& - \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\left(\frac{\partial^2 g_p}{\partial x_i x_j}\frac{\partial^2 g_q}{\partial x_k x_l}\right) c_{ij} c_{kl}\,.
\end{aligned}$$

Contribution of second-order sensitivities can be substantial.

Case Study I

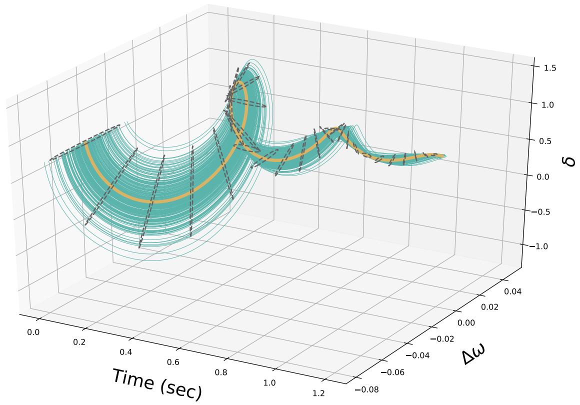
OMIB, 14 states. Gaussian approximation and normal deviation. Tails.





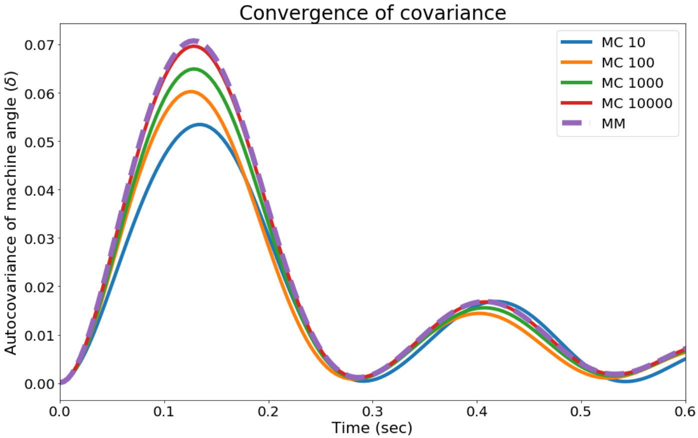
Case Study II

Plotting the uncertainty ellipsoid for $\Delta\omega$, δ.



Case Study III

Monte Carlo convergence.



Case Study IV

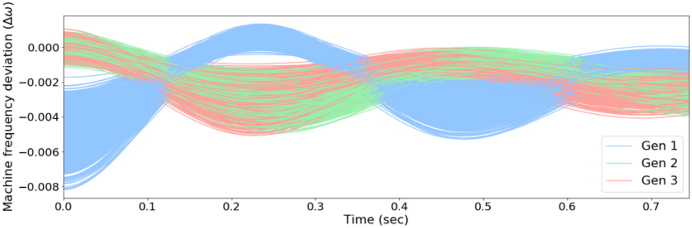
Table: Approximation errors using higher order degree derivatives in the Method of Moments (MM). Comparison with Monte Carlo experiment using $10,000$ samples. We show the approximated mean at $t = 0.6$ sec.

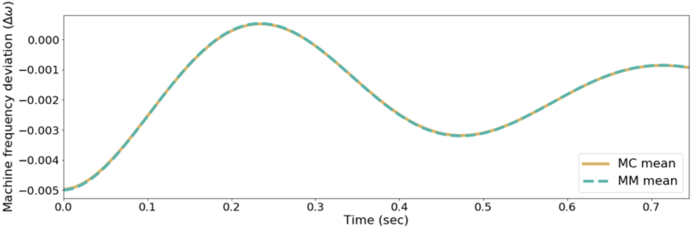
Method	Mean	Error
Monte Carlo	0.004860	-
MM, 1st order derivatives	0.005822	19,796%
MM, 2nd order derivatives	0.004918	1,211%
MM, 3rd order derivatives	0.004864	0,098%

Approximation is quite good. Second order sensitivities have a significant impact on accuracy.

Case Study VI

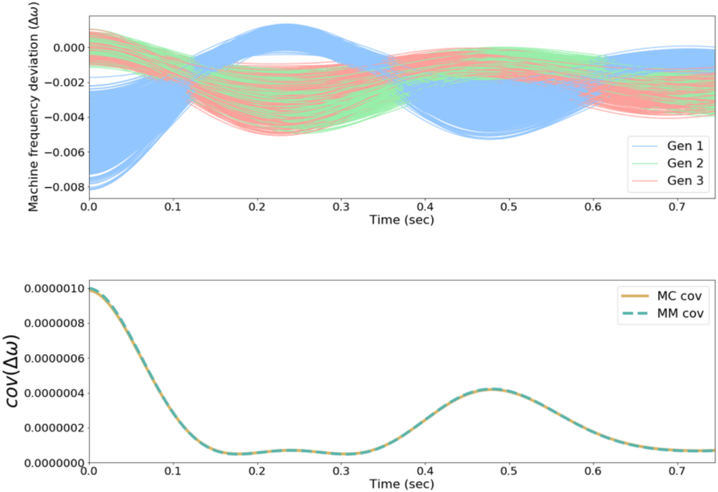
Same set-up with a 9 bus system. Mean of $\Delta\omega_1$.





Case Study VII

Same set-up with a 9 bus system.



Thoughts

Challenges

- Gaussian hypothesis. What about other distributions?
- Does not capture tails.
- Misleading if distribution is bimodal (e.g. bifurcation)
- Closure. Common approach but no guarantees. Approximation may degrade over time (good performance might depend on system being dissipative).

Thoughts

Advantages and extension

- Second-order give us increased accuracy.
- No sampling involved. Potential to scale to large systems.
- Moments can be used with Chebyshev's inequality to determine probability of x_k exceeding value.
- Complex distributions can be approximated by Gaussian mixtures: $p(x) = \sum_{i=0}^{n} w_i N(x; \sigma_i, P_i)$ with $\sum w_i = 0$.
- It can be combined with sampling techniques (think control variates in MC).

Depending on application, this could be a good tool. It can also serve as an ingredient to more general approaches. We leverage sensitivities.

Trust Region Optimization

Approximation of Trajectory Extremes with Trust Region Optimization

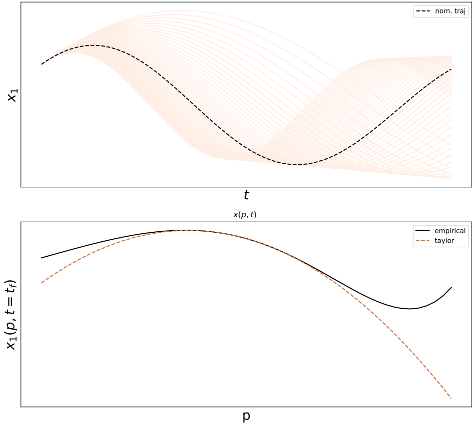
Motivation

- Requirements for system performance (e.g. voltage and frequency within bounds).
- Parameters not known but we have ranges (unknown-but-bounded [Schweppe, 1973])
- Consider the DAE $M\dot{z} = h(z, \theta, t)$. Find the extremes given a bounded set for θ.
- Reachability, interval analysis, etc.
- One idea:

$$z_i(p_m + s, t) \approx z_i(p_m, t) + s^T \left.\frac{\partial z_i}{\partial p}\right|_{(p_m,t)} + \frac{1}{2} s^T \left.\frac{\partial^2 z_i}{\partial p^2}\right|_{(p_m,t)} s\,.$$

[Hiskens and Alseddiqui, 2006] (1st order), and [Choi et al., 2017] (2nd).

An optimization problem



Formulation:

$$\begin{aligned} \underset{\theta}{\text{maximize}} \quad & z_i(\theta, t) \\ \text{subject to} \quad & \theta \in \Theta \end{aligned}$$

- We are using a **surrogate** of z_i
- Nonlinear optimization problem

Trust Region

Find $\{x_k\}$ that converge to a point x_*, where $\nabla f(x_*) = 0$.
We use a quadratic surrogate model - Taylor expansion at x_k:

$$m_k(x_k + s) = f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s$$

We need to ensure surrogate is *suitable* approximation in a given region $\mathcal{B}_k$, where

$$\mathcal{B}_k = \{x \in \mathbf{R}^n \mid \|x - x_k\| \leq \Delta_k\}.$$

We seek s that minimizes $m_k()$ and $\|s_k\|_2 \leq \Delta_k$ (the subproblem). Then, we check agreement:

$$\rho_k = \frac{\textit{actual reduction}}{\textit{predicted reduction}} = \frac{f_i(x_k) - f_i(x_k + s_k)}{m_k(x_k) - m_k(x_k + s_k)}$$

If ρ_k is small: reject and reduce trust region radius. If accepted, set x_{k+1} and create new surrogate.

Algorithm 1 Trust Region

```
procedure TRUST REGION
    Given Δ̂ > 0, initialize Δ_0 ∈ (0, Δ̂), and η ∈ [0, η_1)
    for k = 0, 1, 2, ··· do
        obtain s_k by reducing m_k(x)
        evaluate ρ_k from (8)
        if ρ_k < η_1 then
            Δ_{k+1} = 1/4 Δ_k
        else
            if ρ_k > η_2 and ||s_k|| = Δ_k then
                Δ_{k+1} = min(2Δ_k, Δ̂)
            else
                Δ_{k+1} = 1/4 Δ_k
        if ρ_k > η then
            x_{k+1} = x_k + s_k
        else
            x_{k+1} = x_k
```

Constants η_1 and η_2 typically set to $\frac{1}{4}$ and $\frac{3}{4}$.

Trust Region II. Extreme trajectories

Again, consider $M\dot{z} = h(z, \theta, t)$ and $\theta \in \Theta$. To approximate the extreme trajectories we:

- Choose a point on the parameter interval. Usually the center.
- Integrate the DAE and its sensitivity equations from t_0 to t_j.
- Construct surrogate and solve Trust Region:

$$\min_{d} \quad m_k(p_k + s) = z_i + u_i^T s + \frac{1}{2} s^T V_i s \,, \quad \text{s.t } \|s\| \leq \Delta_k \,,$$

- Optimization problem is solved for each time step.
- We can "hot start" but no guarantee optimal p is close.

Trust Region III. Uncertainty Model

We set up a system with a generator, governor, exciter with saturation. The uncertainty comes from the load composition, α.

$$P_{inj}(V_0, t_0) = \alpha P_z + (1-\alpha) P_{mot}\,,$$
$$Q_{inj}(V_0, t_0) = \alpha Q_z + (1-\alpha) Q_{mot}\,.$$

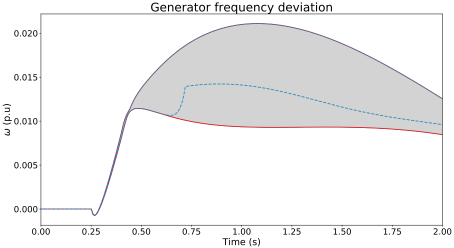
where $P_z = \left(\frac{V_i}{V_0}\right)^2 P_0$ and the motor:

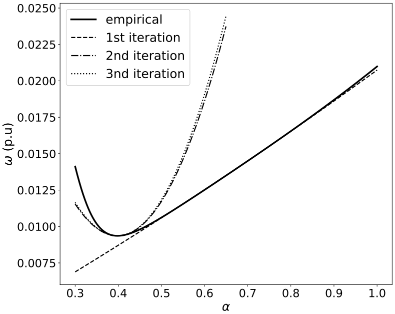
$$\dot{e}'_d = -\frac{1}{T_p}(e'_d + (x_0 - x')i_q) + s\omega_s e'_q\,,$$
$$\dot{e}'_q = -\frac{1}{T_p}(e'_q - (x_0 - x')i_d) - s\omega_s e'_d\,,$$
$$\dot{s} = \frac{1}{2H}(\tau_m - e'_d i_d - e'_q i_q)$$

$$0 = r_a i_d - x' i_q + e'_d + V sin(\theta)\,,$$
$$0 = r_a i_q - x' i_d + e'_q - V cos(\theta)\,,$$
$$P_{mot} = -V sin(\theta) i_d + V cos(\theta) i_q\,,$$
$$Q_{mot} = V cos(\theta) i_d + V sin(\theta) i_q\,.$$

Trust Region III

Fault applied from $t = 0.25$ to $t = 0.45$. In red, Trust Region. In blue, optimize for the Taylor expansion around nominal point. Compared with grid-sampling, we observe very small errors (e.g. for minimum trajectory, relative error is $1.406e^{-6}$ over whole trajectory)





Note corners not necessarily correspond to extremes.

Thoughts

- We tried this on New England case (19 parameters). Good performance.
- Small number of TR iterations (average 5-10).
- This is still an approximation. Problems might arise with local minima.
- But the field of NLOpt is very mature. Lots of room for better performance.
- We might turn to global optimization techniques (sampling might be involved)
- Trust-region can be used to accelerate global optimization techniques

UQGrid

```python
# load data
psys = load_psse(raw_filename="IEEE39_v33.raw")
add_dyr(psys, "IEEE39.dyr")

# add fault and create initial data structures
v, Sinj = runpf(psys, verbose=True)
psys.add_busfault(1, zfault, 0.01)

# set up parameters
pmax = np.ones(psys.nloads)
pmin = np.zeros(psys.nloads)

pnom = pmin + 0.5*(pmax - pmin)
psys.set_load_parameters(pnom)

# integrate
results = integrate_system(psys, verbose=True,
    comp_sens=True, dt=dt, tend=10.0)
```

- Trajectory sensitivities useful tool in UQ.
- Implementation barrier.
- Achieving performance might be challenging.
- We are developing a new library, `UQGrid`.
- PETSc. Discrete sensitivities and adjoints.
- AD for f_θ.
- Parallelization (GPU).
- High performance and flexibility.

We look for users, collaborators.

Thank You

References I

[Choi et al., 2017] Choi, H., Seiler, P. J., and Dhople, S. V. (2017).
Propagating Uncertainty in Power-System DAE Models with Semidefinite Programming.
IEEE Transactions on Power Systems, 32(4):3146–3156.

[Geng and Hiskens, 2019] Geng, S. and Hiskens, I. A. (2019).
Second-order trajectory sensitivity analysis of hybrid systems.
IEEE Transactions on Circuits and Systems I: Regular Papers, 66(5):1922–1934.

[Hiskens and Pai, 2000] Hiskens, I. and Pai, M. (2000).
Trajectory sensitivity analysis of hybrid systems.
IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 47(2):204–220.

References II

[Hiskens and Alseddiqui, 2006] Hiskens, I. A. and Alseddiqui, J. (2006).
Sensitivity, approximation, and uncertainty in power system dynamic simulation.
IEEE Transactions on Power Systems, 21(4):1808–1820.

[Pugachev, 1965] Pugachev, V. (1965).
Theory of random functions and its applications to control problems.
International series of monographs on automation and automatic control. Addison Wesley.

[Schweppe, 1973] Schweppe, F. C. (1973).
Uncertain dynamic systems.
Prentice-Hall, hardcover edition.

Sensitivities

Let $f : S \to \mathbf{R}^m, S \subseteq \mathbf{R}^n$ be a vector function, and let z be an interior point of S the second-degree Taylor expansion:

$$f(z+d) = f(z) + \mathcal{J}d + \frac{1}{2}(I_m \otimes d^T)\mathcal{H}d + \mathbf{O}(\|d\|^3)$$

Here $\mathcal{J}$ is the Jacobian matrix, $\mathcal{H}$ is the Hessian matrix. The quadratic term:

$$(I_m \otimes d^T)\mathcal{H}d = \begin{bmatrix} d^T\mathcal{H}_1 d \\ d^T\mathcal{H}_2 d \\ \vdots \\ d^T\mathcal{H}_m d \end{bmatrix}^T$$

Often, we are interested in sensitivities with respect to a vector parameter. For instance, given $f(t,x,\theta)$, then $f_\theta(t,x,\theta) = \frac{\partial f(t,x,\theta)}{\partial \theta}$.