THE GHOST MCMC ALGORITHM: rare event sampling and applications to power systems reliability

Alessandro Zocca
California Institute of Technology
Increasing penetration of renewables

Progressive transport electrification

Can extreme fluctuations in load and generation cause failures?
German transmission network model with $n = 585$ nodes and $m = 852$ edges

renewables: blue = offshore wind, light blue = onshore wind, yellow = solar
Weighted graph with sources and sinks $\rightarrow$ power injections $p$
power grids as stochastic network

power injections are modeled as random variables

\[ \mathbb{E} \mathbf{p} = \mu \quad \text{and} \quad \text{Cov}(\mathbf{p}) = \Sigma_p \]

heterogeneous variances for the power injections fluctuations as well as correlations between them

→ non-trivial covariance matrix \( \Sigma_p \)

quantities of interest

• line power flows \( \mathbf{f} = V \mathbf{p} \)

• nodal frequencies

\[
M_j \dot{\omega}_j = -D_j \omega_j + (p_j - \mu_j) - \sum_{i: (i,j) \in E} f_{i,j} \\
\dot{f}_{i,j} = B_{i,j}(\omega_i - \omega_j)
\]
investigating power grids reliability

**Failure event** = line power flow overloads

\[ C^{\text{line}} = \bigcap_{\ell \in E} \{ |f_\ell| \geq 1 \} \]

**Failure event** = nodal frequencies or rocof violations

\[ C^{\text{freq}} = \bigcap_{j \in G} \left\{ \max_{t \in [0, T]} |\omega_j(t)| \geq \theta_{\text{max}} \right\} \]

\[ C^{\text{rocof}} = \bigcap_{j \in G} \left\{ \max_{t \in [0, T]} |\dot{\omega}_j(t)| \geq r_{\text{max}} \right\} \]

how does the system looks like upon failure?
how do failures most likely happen?
introduce the parameter $\varepsilon > 0$ to describe the magnitude of the system’s noise

$$\text{Cov}(p) = \varepsilon \Sigma_p$$

small-noise limit $\varepsilon \to 0^+$

$$\lim_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{P}_\mu(|f_\ell| \geq 1) = -I_\ell(\mu)$$

$$\lim_{\varepsilon \downarrow 0} \varepsilon \mathbb{P}_\mu(C^{\text{line}}) = \lim_{\varepsilon \downarrow 0} \varepsilon \mathbb{P}_\mu(\max_\ell |f_\ell| \geq 1) = -\min_\ell I_\ell(\mu)$$

in the Gaussian case

$$I_\ell(\mu) = \inf_{p \in \mathbb{R}^n : |e_\ell^T V p| \geq 1} \frac{1}{2} (p - \mu)^T \Sigma_p^{-1} (p - \mu) = \frac{(1 - |\nu_\ell|)^2}{2\sigma_\ell^2}$$
conditioning on failure event

\[ p^{(\ell)} = \arg \inf_{p \in \mathbb{R}^n : |e^T_{\ell} V p| \geq 1} \frac{1}{2} (p - \mu)^T \Sigma_p^{-1} (p - \mu) = \mathbb{E} [p | f_\ell = 1] = \mu + \frac{(1 - \nu_\ell)}{\sigma^2_\ell} \Sigma_p V^T e_\ell \]

however, this close-form expression for power injections realization upon failure strongly depends on linearized dynamics and Gaussian assumption
how to deal with...?

• non-Gaussian & mixed distribution?
• non-linear dynamics?

\[
M_j \dot{\omega}_j = -D_j \omega_j + (p_j - \mu_j) - \sum_{i: (i,j) \in E} f_{i,j}
\]
\[
\dot{f}_{i,j} = B_{i,j} (\omega_i - \omega_j)
\]

→ MCMC method to sample conditionally on failure event
devise a method to efficiently sample from the conditional distribution $\pi = \frac{\rho 1_C}{\rho(C)}$ where $C$ is a rare event, i.e., $\rho(C) \ll 1$
how to sample?

Some naïve ideas...

- Sampling from $\rho$ and rejecting when in $C^c$
- Sampling from $\pi$ using MCMC
**Algorithm 1:** Ghost Random Walk Metropolis algorithm (n-th step)

1. Generate a SRWM proposal $Y_{n+1}$ distributed according to the density $q(y - X_n)dy$;
2. Calculate direction $\varphi_n = Y_{n+1} - X_n$;
3. Calculate all the intersection points (which are at most two) $T := \{ t > 0 : X_n + t\varphi \in \delta C\}$;
4. **if** $T = \{t_1, t_2\}$ and $\min\{t_1, t_2\} < 1$ **then**
   5. $Z_{n+1} = Y_{n+1} + (t_2 - t_1)\varphi$;
   6. **else**
   7. $Z_{n+1} = Y_{n+1}$;
8. **end**
9. Evaluate the acceptance probability:

$$\alpha(X_n, Z_{n+1}) = \min\left(1, \frac{\pi(Z_{n+1})1_C(Z_{n+1})}{\pi(X_n)1_C(X_n)}\right)$$

interpreted as one if $\pi(X_n)1_C(X_n) = 0$;
10. Generate a uniform random variable $U$ on $[0, 1]$;
11. **if** $U \leq \alpha(X_n, Z_{n+1})$ **then**
12. $X_{n+1} = Z_{n+1}$;
13. **else**
14. $X_{n+1} = X_n$;
15. **end**
16. **return** $X_{n+1}$
ghost algorithm in action

classical MCMC algorithm

ghost MCMC algorithm
**a technical condition (ray-boundedness)**

Define \( l_x^\varphi(r) := \int_0^r 1_C(x + t\varphi)\,dt \) and \( G_x^C : C \to \mathbb{R}^d \) as

\[
G_x^C(x + r\varphi) := x + l_x^\varphi(r)\varphi
\]

\( C \) closed set and \( x \in C \) then \( G_x^C : C \to \mathbb{R}^d \) is injective

\( C^c \) **ray-bounded** if \( G_x^C : C \to \mathbb{R}^d \) is surjective for all \( x \in C \)

\( C^c \) **convex polytope**: ray boundedness simply means that any ray starting in \( C \) and intersecting \( C^c \) will also exit from it
Theorem (ghost sampling) (MVZ18)

Assume $C$ is closed and $C^c$ ray-bounded subset.

The ghost sampling RW is a MH algorithm with proposal density symmetric on $C$ given by

$$q_{GS}(x, x + r\varphi) := q(l_x^\varphi(r)\varphi) \left(\frac{l_x^\varphi(r)}{r}\right)^{d-1} 1_C(x + r\varphi).$$

Moreover, the GS algorithm is $\pi$-irreducible and the SLLN holds, that is, for every $\pi$-integrable function $f$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n} f(X_i) = \pi(f)$$
back to frequency violations

assume the system is affected by a step disturbance

$$u(t) = u 1_{\{t \geq 0\}}$$

where $u = p - \mu$ is a $d$-dimensional r.v. with distribution $\rho$

the system evolution is described by

$$\dot{x} = Ax \quad x = \begin{bmatrix} \dot{\omega} \\ \omega \end{bmatrix}, \quad A = \begin{bmatrix} -M^{-1}D & -M^{-1}L \\ I & 0 \end{bmatrix}, \quad x(0) = \begin{bmatrix} M^{-1}u \\ 0 \end{bmatrix}$$

we approximate the event «rocof violation» in the interval $[0, \varepsilon]$ with $\varepsilon = 0.5\text{sec}$ as

$$C^{\text{rocof}}(N) = \bigcap_{j \in G} \bigcap_{n=0}^{N} \left\{ u \in \mathbb{R}^d : \left| \dot{\omega}_j \left( \frac{n}{N} \varepsilon \right) \right| \leq r_{\text{max}} \right\}$$

$$= \bigcap_{j \in G} \bigcap_{n=0}^{N} \left\{ u \in \mathbb{R}^d : \left| \exp \left( \frac{n}{N} \varepsilon \right) j \begin{bmatrix} M^{-1}u \\ 0 \end{bmatrix} \right| \leq r_{\text{max}} \right\}$$
a case study

IEEE 39-bus test network

\[ \rho(u_1, u_2) \sim \frac{1}{1 + (30(u_1 - u_2/2))^4} \cdot \frac{1}{1 + (30(u_2 - u_1/2))^4} \]

+ the other components i.i.d. zero-mean Gaussian

<table>
<thead>
<tr>
<th>N</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>G4</th>
<th>G5</th>
<th>G6</th>
<th>G7</th>
<th>G8</th>
<th>G9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28.9</td>
<td>80.3</td>
<td>0.5</td>
<td>0.6</td>
<td>0.9</td>
<td>0.4</td>
<td>1.6</td>
<td>6.5</td>
<td>1.6</td>
</tr>
<tr>
<td>5</td>
<td>27.6</td>
<td>81.5</td>
<td>12.4</td>
<td>1.1</td>
<td>2.1</td>
<td>0</td>
<td>1.2</td>
<td>9.8</td>
<td>1.7</td>
</tr>
<tr>
<td>20</td>
<td>27.5</td>
<td>79.5</td>
<td>11.5</td>
<td>1.9</td>
<td>3.0</td>
<td>0.1</td>
<td>2.4</td>
<td>15.5</td>
<td>2.0</td>
</tr>
<tr>
<td>50</td>
<td>28.5</td>
<td>78.8</td>
<td>12.2</td>
<td>1.1</td>
<td>2.7</td>
<td>0.1</td>
<td>2.4</td>
<td>17.1</td>
<td>2.6</td>
</tr>
<tr>
<td>100</td>
<td>28.6</td>
<td>79.8</td>
<td>12.2</td>
<td>1.7</td>
<td>2.4</td>
<td>0.1</td>
<td>1.9</td>
<td>15.6</td>
<td>2.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N = 1</th>
<th>N = 5</th>
<th>N = 20</th>
<th>N = 50</th>
<th>N = 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p_d)</td>
<td>15.2%</td>
<td>22.4%</td>
<td>24.0%</td>
<td>25.0%</td>
<td>25.0%</td>
</tr>
<tr>
<td>(\bar{d})</td>
<td>1.21</td>
<td>1.37</td>
<td>1.44</td>
<td>1.46</td>
<td>1.44</td>
</tr>
<tr>
<td>(\bar{L})</td>
<td>596</td>
<td>701</td>
<td>735</td>
<td>744</td>
<td>736</td>
</tr>
</tbody>
</table>
extension to more general events
some preliminaries

\[ \mathcal{K} = \{ K_\varphi \mid \varphi \in S^{d-1} \} \text{ is a halting regime if} \]

- for every \( \varphi \in S^{d-1} \) \( K_\varphi \) is a r.v. on \( \mathbb{N} \cup \{\infty\} \)
- \( K_\varphi \sim K_{-\varphi} \) for every direction \( \varphi \in S^{d-1} \)

assume the proposal density can be written as

\[ q(y - x) = q_{r|\varphi}(r|\varphi)q_\varphi(\varphi) \]

where \( r = |y - x| \) and \( \varphi = (y - x)/|y - x| \)
if the proposal density is $q(x - y) \sim \mathcal{N}(0, \Sigma)$, then

$$q_\varphi(\varphi) = \frac{\Gamma\left(\frac{d}{2}\right)}{2\pi^{d/2} \cdot \sqrt{\det \Sigma} \cdot (\varphi^T \Sigma^{-1} \varphi)^{d/2}}$$

$$q_{r|\varphi}(r) = \frac{\left(\varphi^T \Sigma^{-1} \varphi\right)^{d/2}}{2^{d/2-1} \Gamma\left(\frac{d}{2}\right)} e^{-\left(\varphi^T \Sigma^{-1} \varphi\right) \frac{r^2}{2} r^{d-1}}$$

$q_{r|\varphi}(r)$ is **generalized gamma distribution** with zero location parameter, scale parameter

$$\beta = \left(\frac{2}{\varphi^T \Sigma^{-1} \varphi}\right)^{1/2}$$

and shape parameters

$$\alpha = d/2 \quad \gamma = 2$$
**Algorithm 2**: Skipping Random Walk Metropolis algorithm ($n$-th step)

**Input**: The $n$-th sample $X_n \in \mathbb{R}^d$

1. Set $X := X_n$;
2. Generate the initial SRWM proposal $Y$ distributed according to the density $q(u - X)du$;
3. Calculate the direction $\phi = (Y - X)/|Y - X|$;
4. Generate a halting index $K \sim K_\phi$;
5. Set $k = 1$ and $Z_1 := Y$;
6. **while** $Z_k \in C^c$ **and** $k < K$ **do**
   7. Generate a distance increment $R$ distributed according to $q_{r|\phi}(\cdot | \phi)$;
   8. Set $Z_{k+1} = Z_k + \phi R$;
   9. Increase $k$ by one;
10. **end**
11. Set $Z := Z_k$;
12. Evaluate the acceptance probability:

\[
\alpha(X, Z) = \begin{cases} 
\min \left(1, \frac{\pi(Z)}{\pi(X)} \right) & \text{if } \pi(X) \neq 0, \\
1, & \text{otherwise},
\end{cases}
\]

13. **if** $U \leq \alpha(X, Z)$ **then**
14. \hspace{1em} $X_{n+1} = Z$;
15. **else**
16. \hspace{1em} $X_{n+1} = X$;
17. **end**
18. return $X_{n+1}$
another technical condition

let $T_C$ be the first entry time of the skipping chain into $C$, that is $T_C := \min\{k \geq 1 : Z_k \in C\}$.

assume the halting regime $\mathcal{K} = \{K_\varphi \mid \varphi \in S^{d-1}\}$ and the subset $C$ are such that

$$\forall x \in \mathbb{R}^d \quad \sup_{\varphi \in S^{d-1}} \mathbb{E}_x [T_C \wedge K_\varphi] < \infty$$
Theorem (skipping sampling) (MVZ19?)

Assume the halting regime $\mathcal{K} = \{K_\varphi \mid \varphi \in \mathbb{S}^{d-1}\}$, the subset $\mathcal{C}$ and the proposal density are as before.

Then the skipping RWM algorithm:

- is a MH algorithm for a proposal density $q_{SK}$ which is symmetric on $\mathcal{C}$;
- is $\pi$-irreducible and has unique stationary probability measure $\pi$;
- the SLLN holds;
- is provably faster than the classical MH algorithm.
future work

• extensive numerics for power grids
  - large-scale transmission network
  - realistic distributions for disturbances

• finish writing the paper about the skipping sampler…

• find new applications for skipping sampling
  - queueing systems?
  - random matrices?
  - random graphs?
Alessandro Zocca
azocca@caltech.edu
https://sites.google.com/site/zoccaale/


**NZZ18** T. Nesti, A. Zocca, B. Zwart «Emergent failures and cascades in power grids: a statistical physics perspective», In Physical Review Letters 120, 258301
Alessandro Zocca
azocca@caltech.edu
https://sites.google.com/site/zoccaale/


absolute values power flows

line failure decay rates

our large deviations framework enables the correct identification of network vulnerabilities!