Efficient Optimization of Loops and Limits with Randomized Telescoping Sums

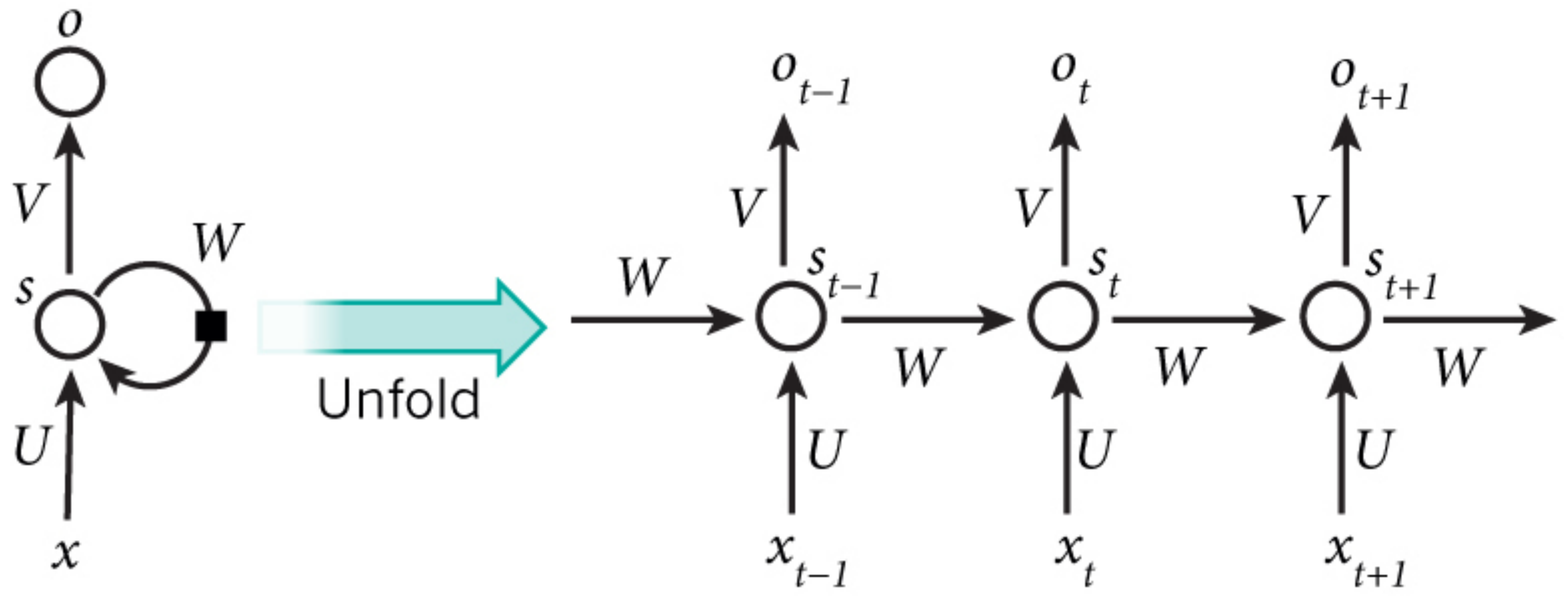
Alex Beatson and Ryan P. Adams

Princeton University
Department of Computer Science

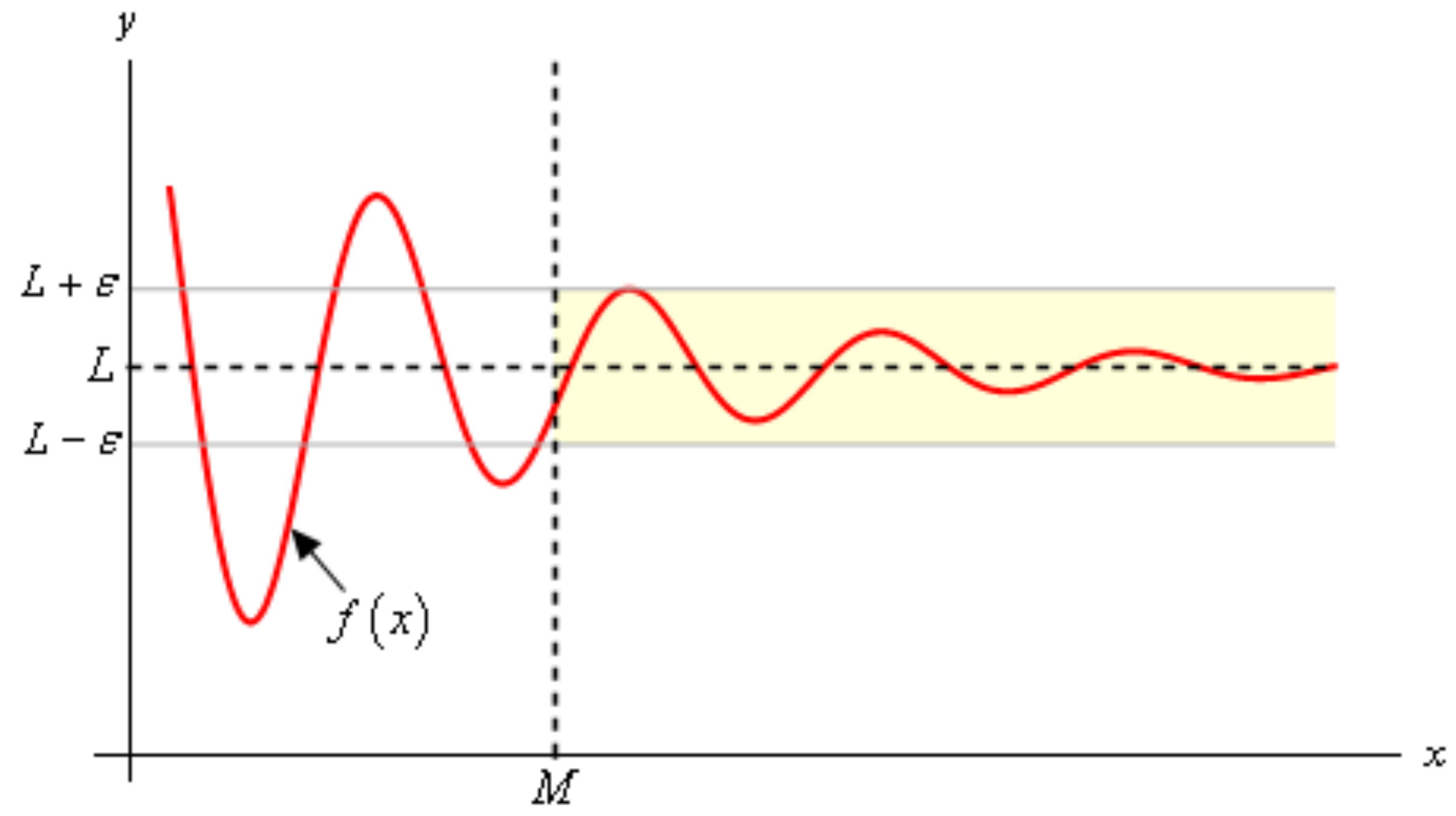
MOTIVATION

- Optimization with inner loops
  - Meta learning, hyperparameter optimization
  - Recurrent models
- Optimization with limits
  - Discretized numerical methods: PDEs, ODEs, …
  - Iterative methods: linear systems, inverses, eigenvalues,
  - Integration with Monte Carlo or quadrature
- In both cases..
  - cheap truncations/approximations cause **harmful bias**
  - accurate approximations are **computationally expensive**

$$\min_\theta \mathcal{L}(\theta) = \min_\theta \lim_{n\to H} \mathcal{L}_n(\theta)$$



http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/



http://tutorial.math.lamar.edu/Classes/CalcI/DefnOfLimit.aspx

RANDOMIZED TELESCOPES: UNBIASED ESTIMATION OF LIMITS

Consider: $Y_H := \lim_{n \to H} Y_n$

Then: $$Y_H = \sum_{n=1}^{H} \Delta_n \quad \text{where} \quad \Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Consider an estimator: $$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

This is unbiased iff: $$\sum_{N=n}^{H} W(n, N) q(N) = 1 \quad \forall n$$

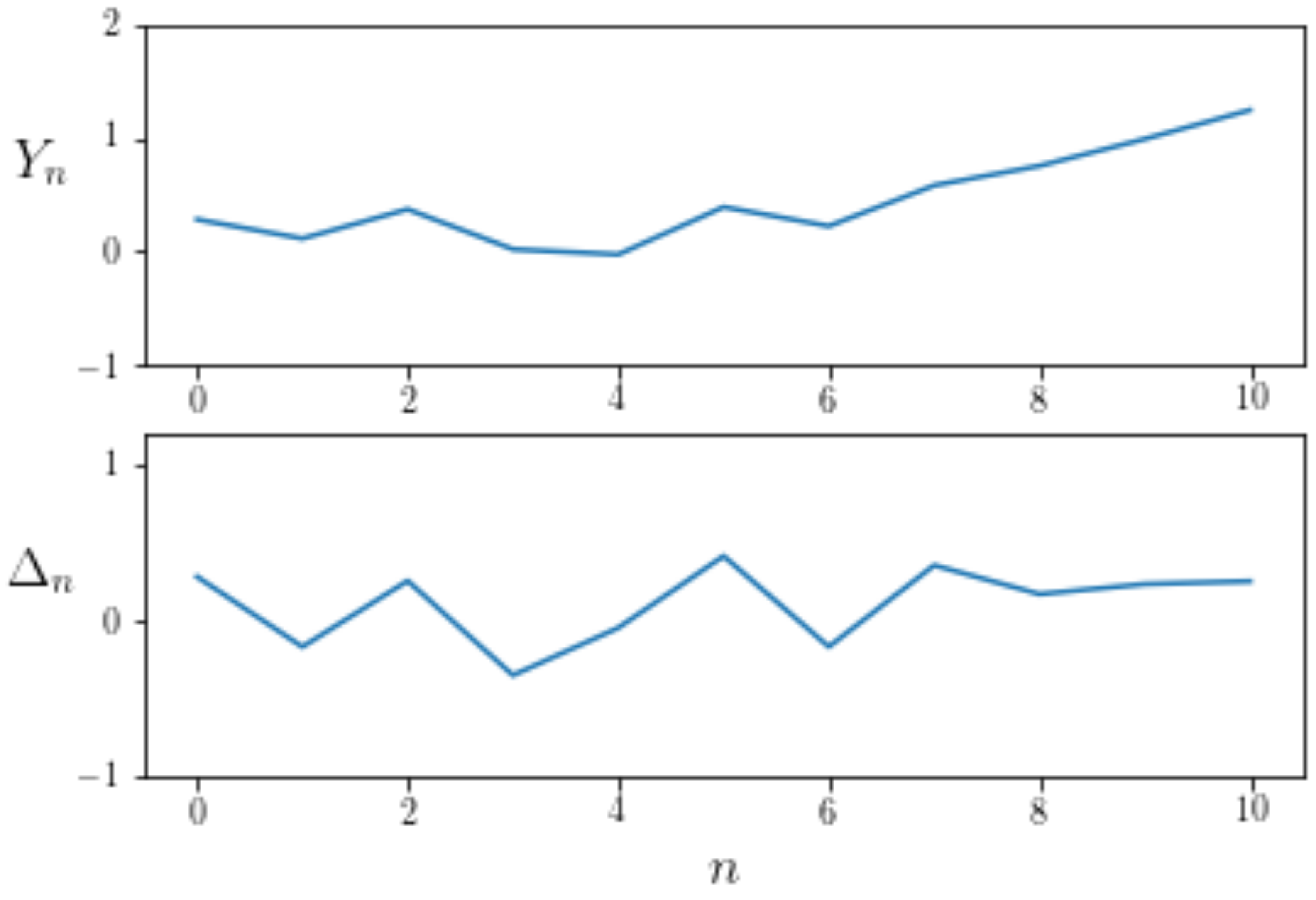
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



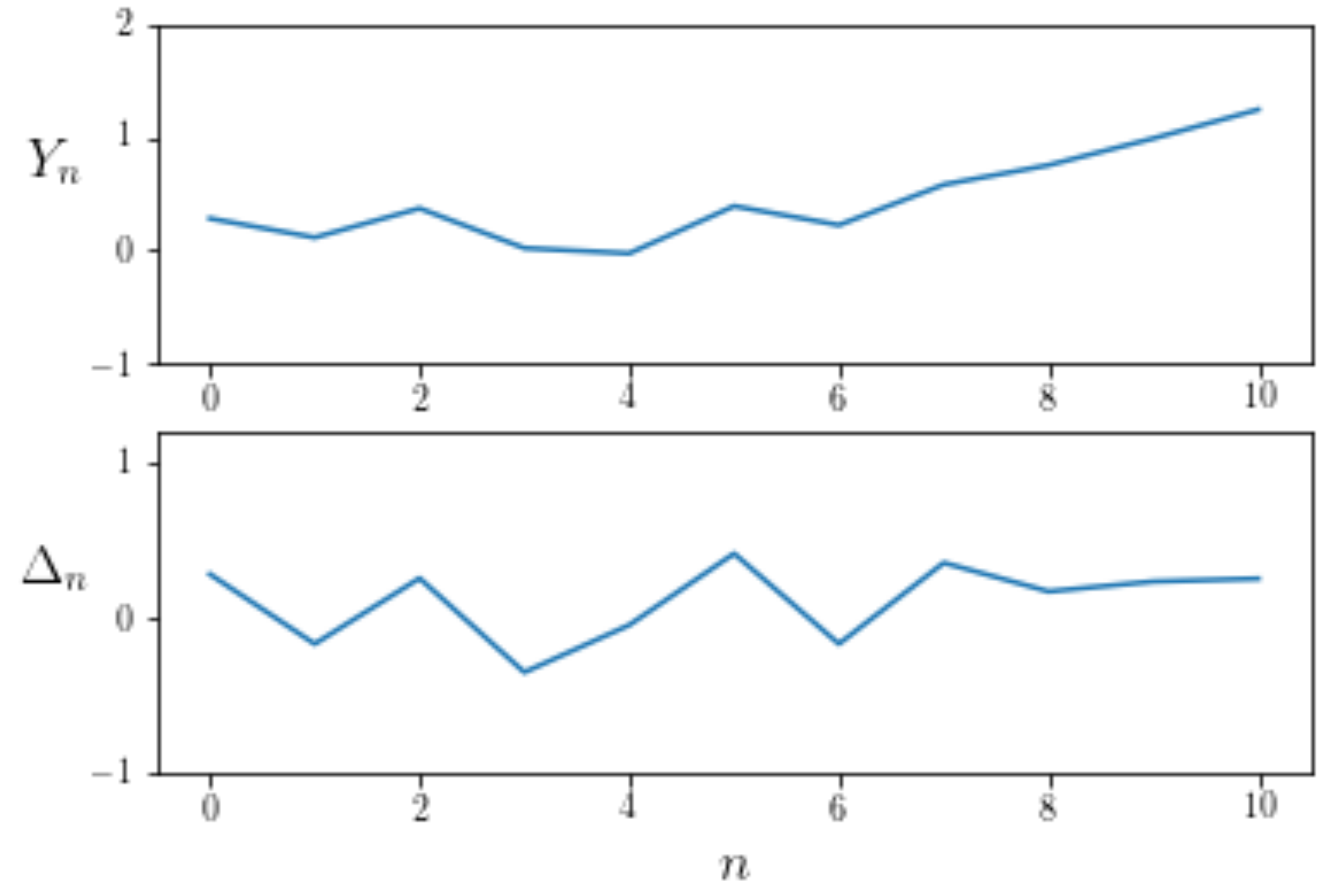
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$

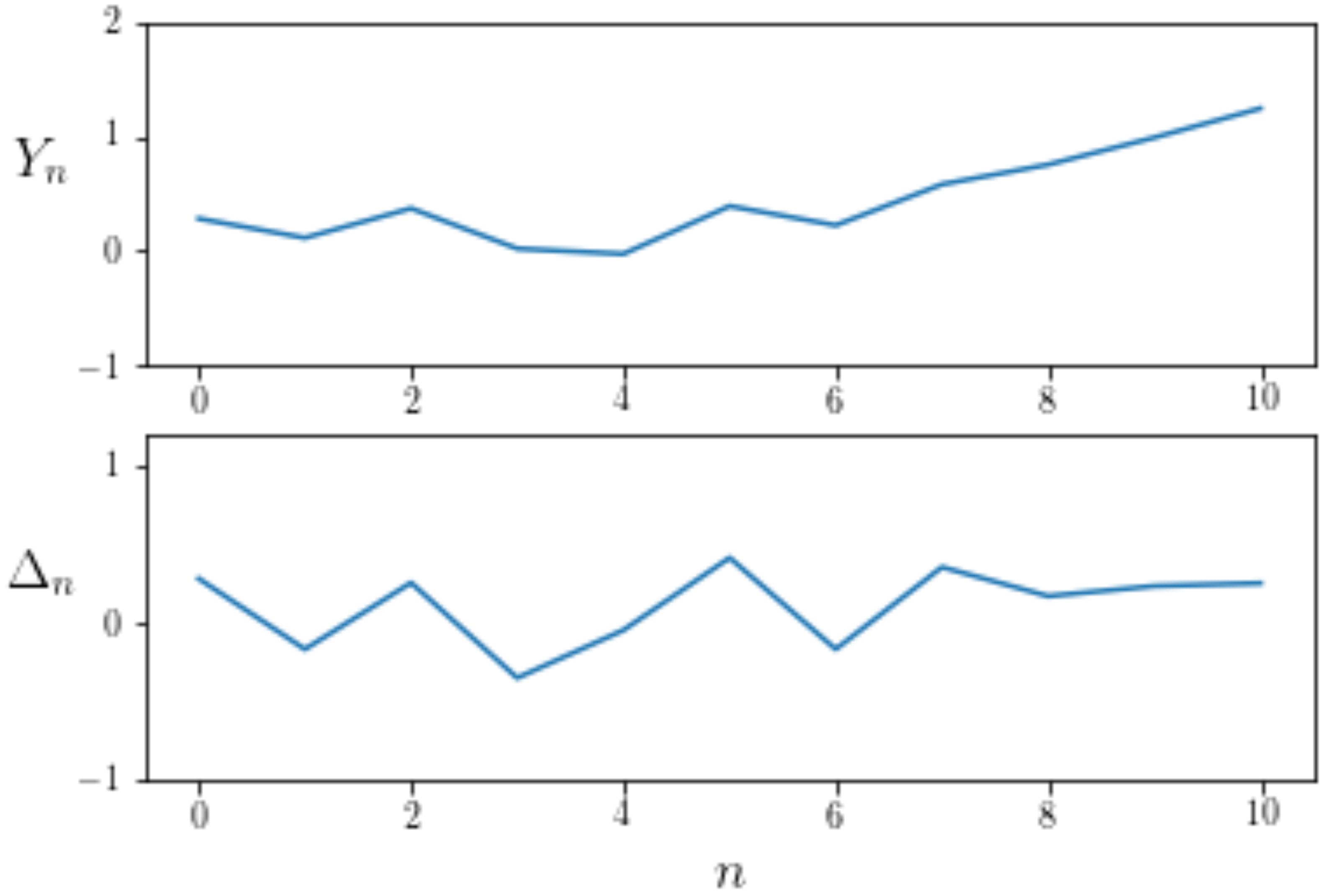
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

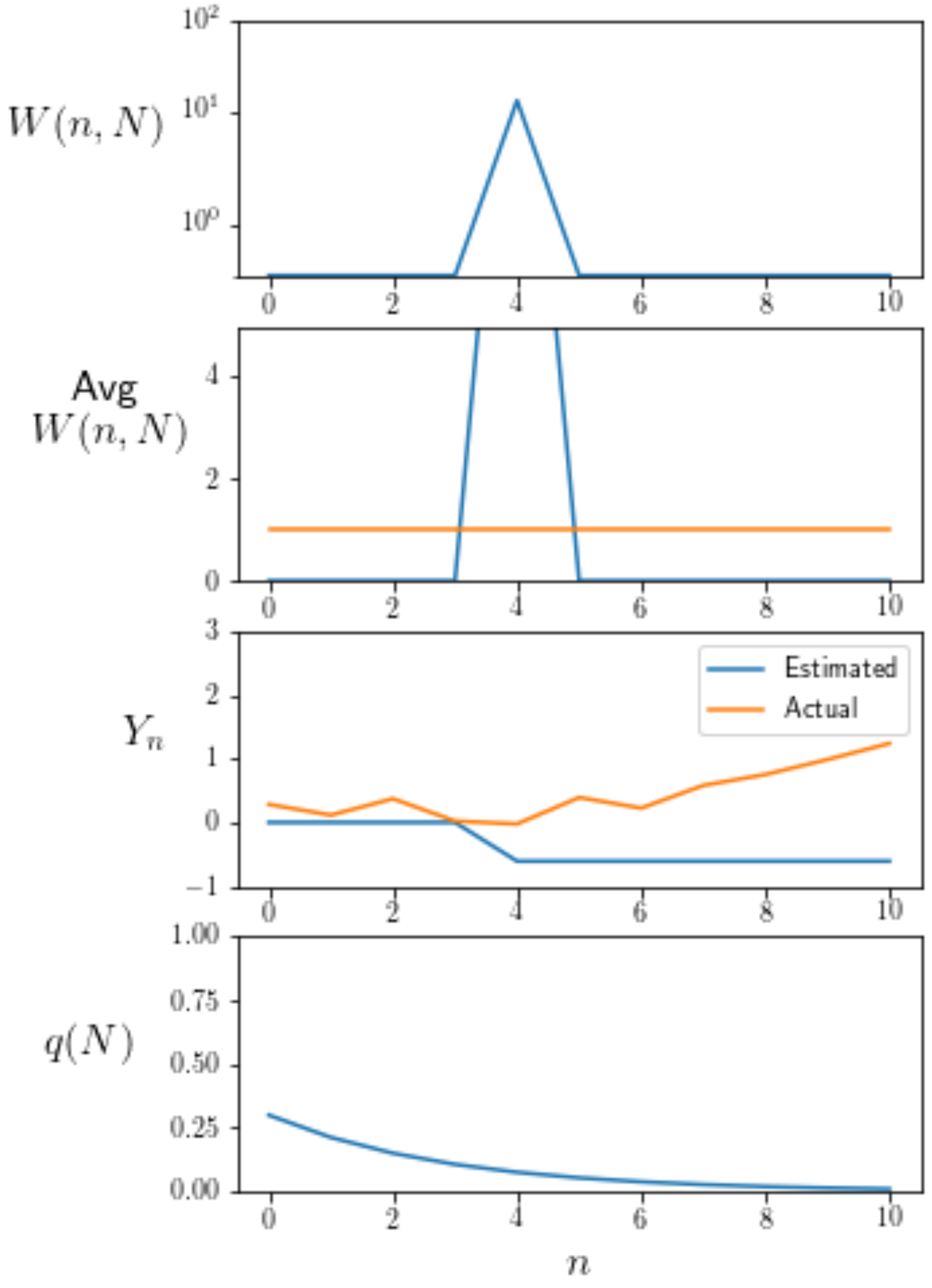


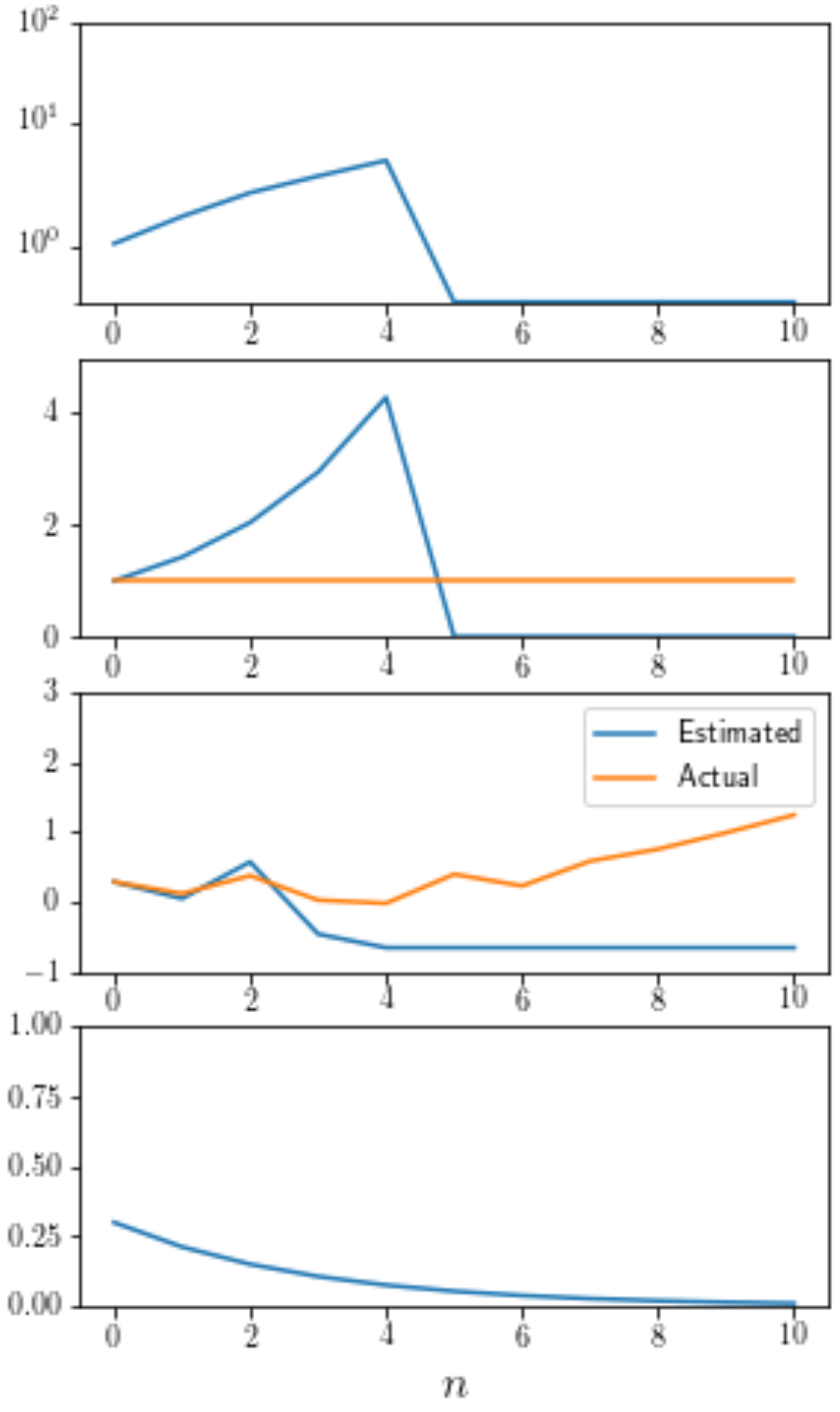
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$





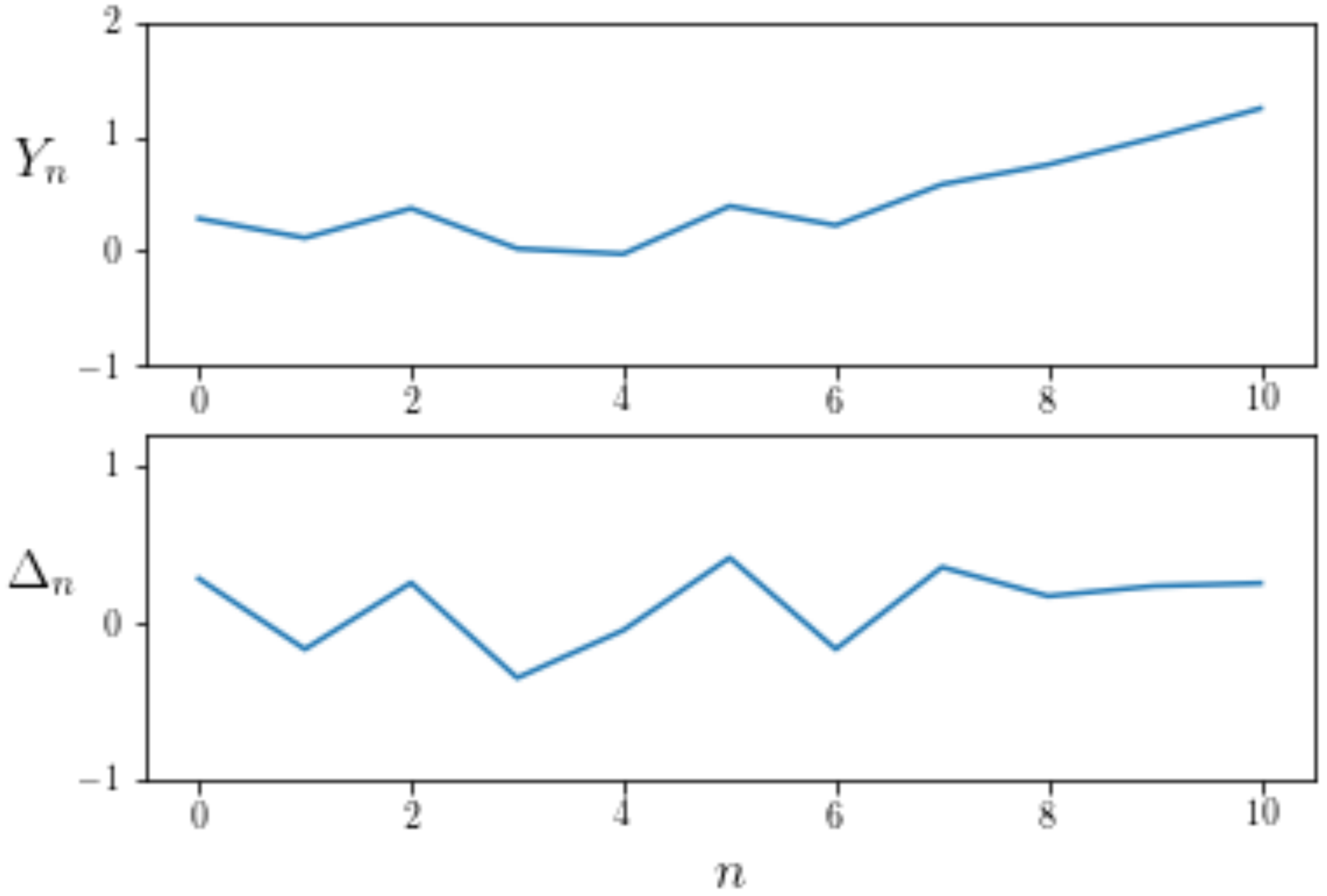
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

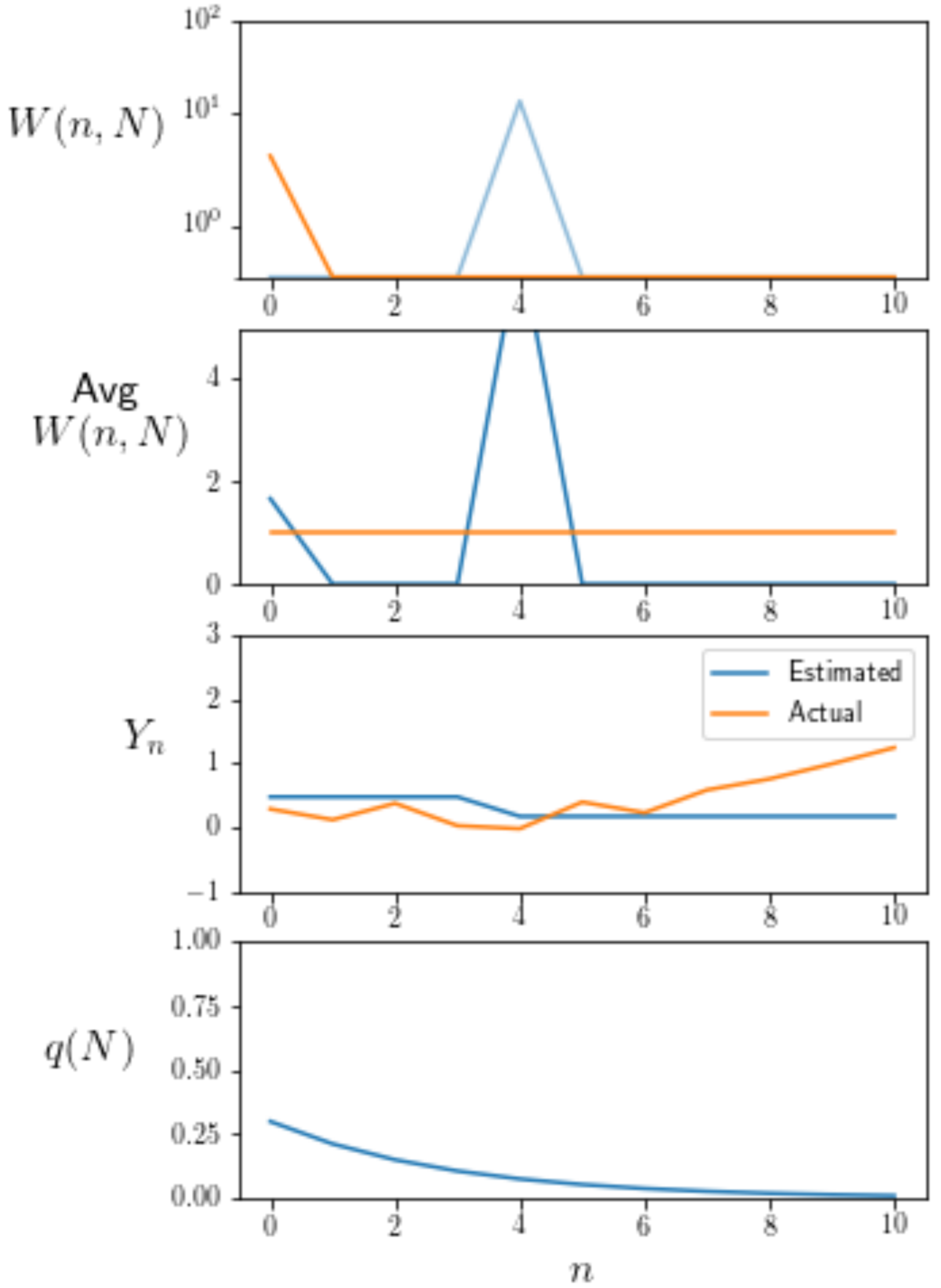
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



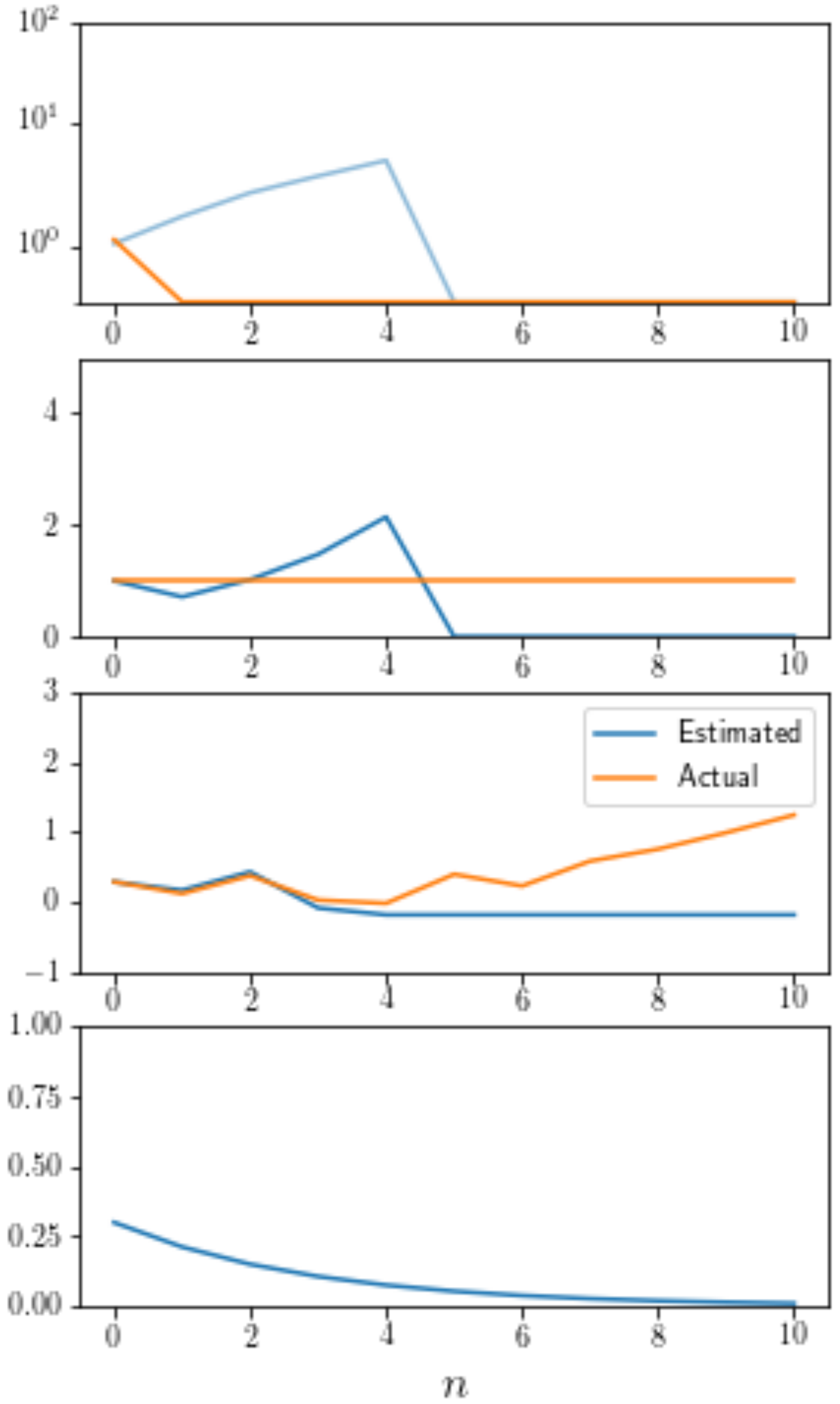
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



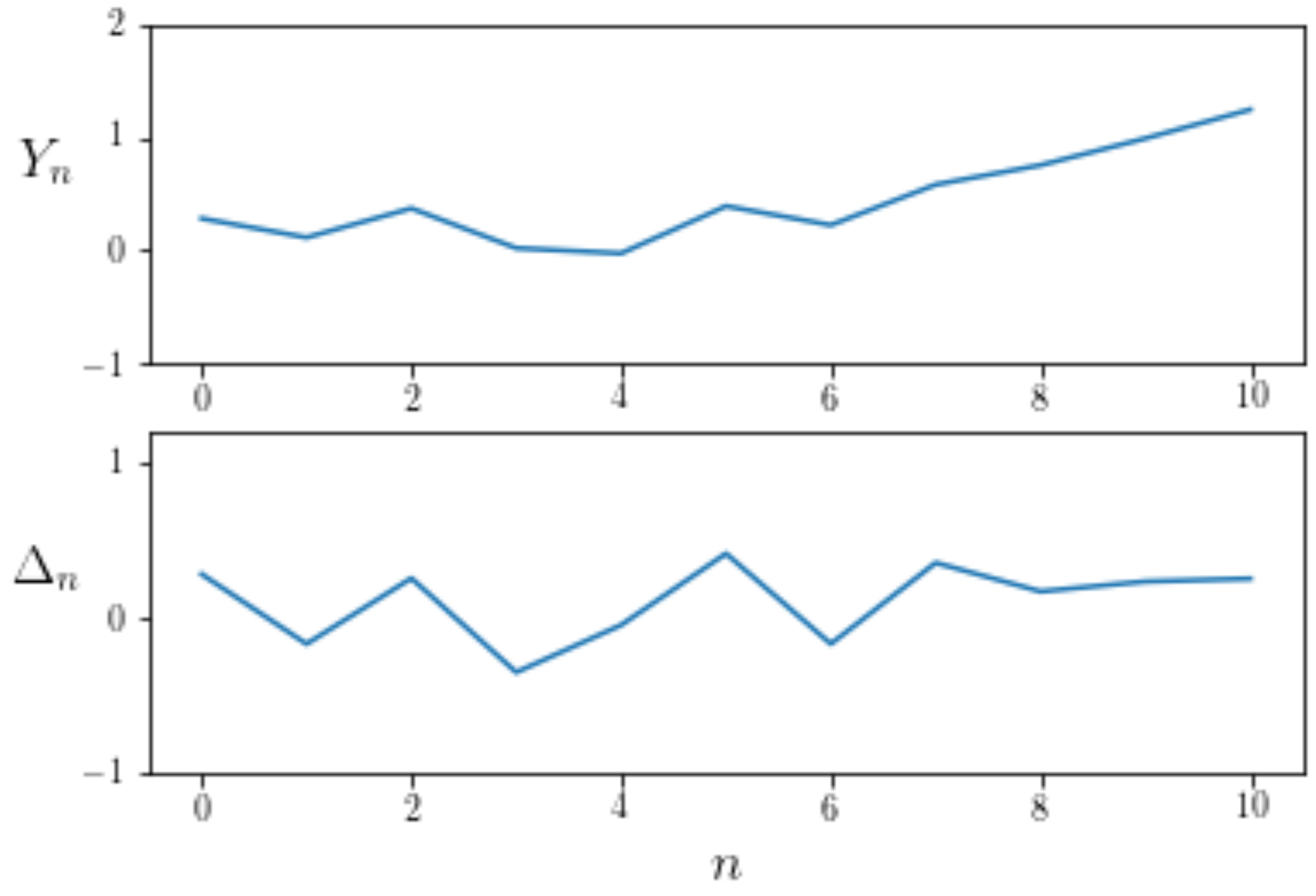
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

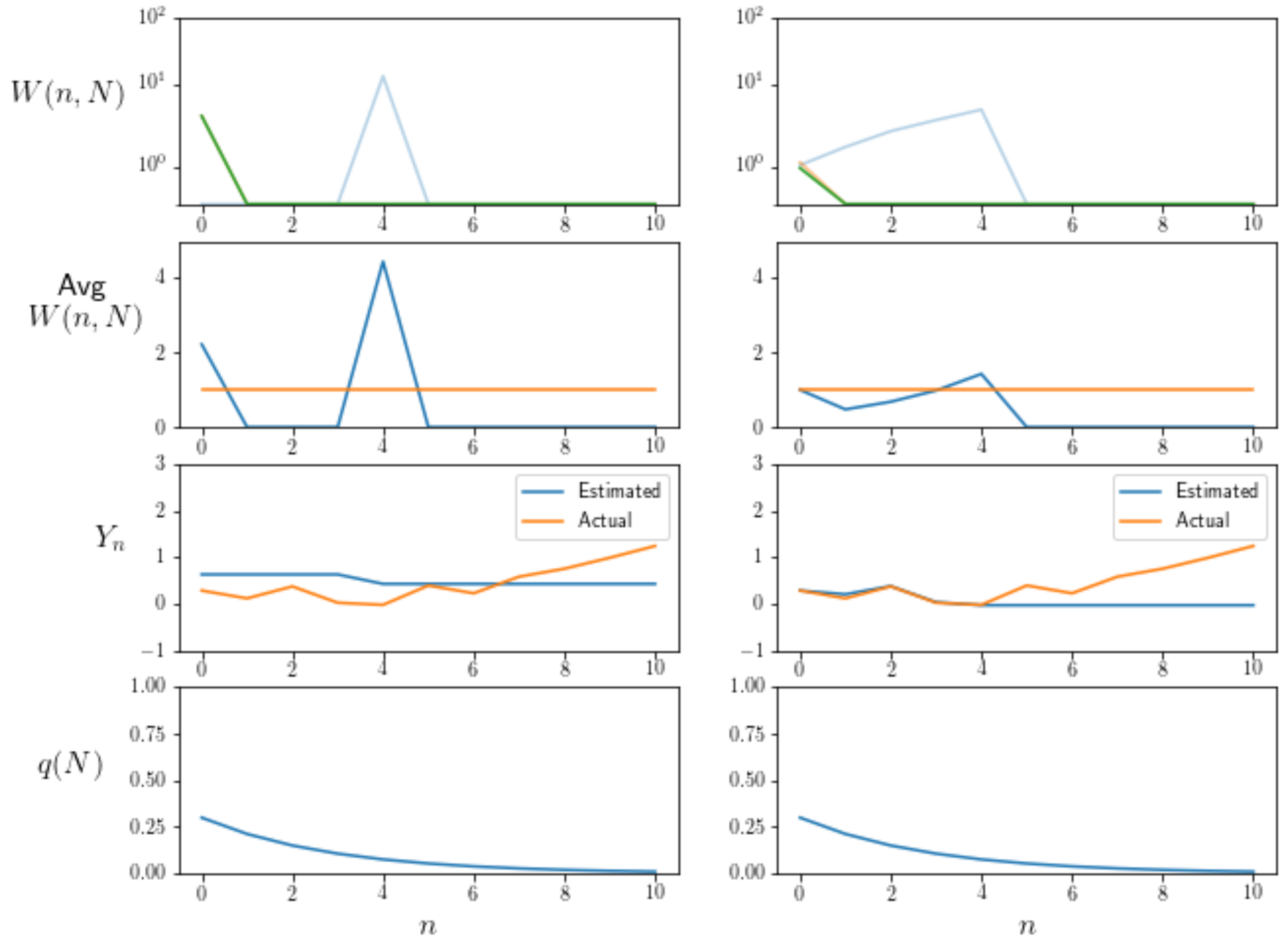


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



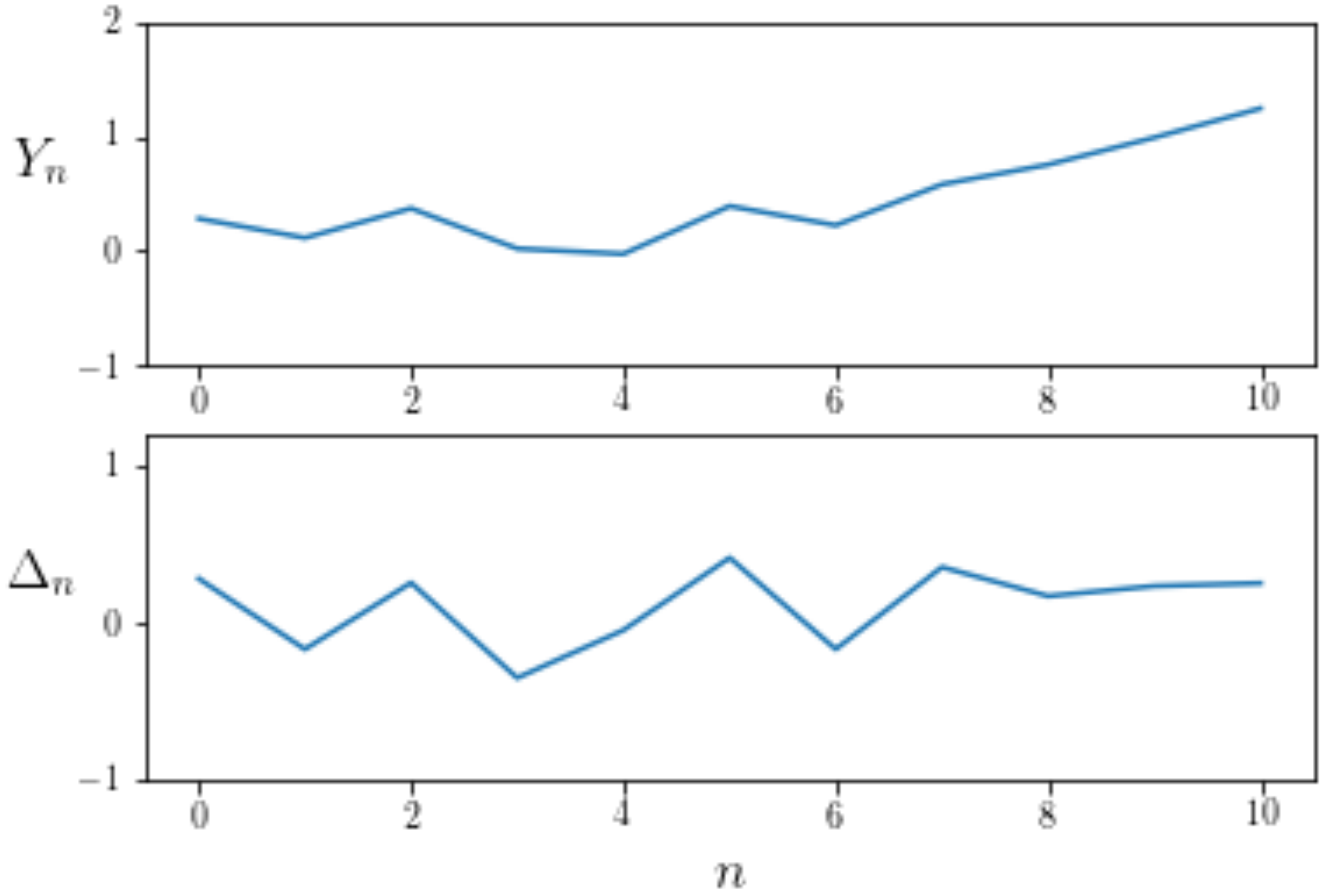
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

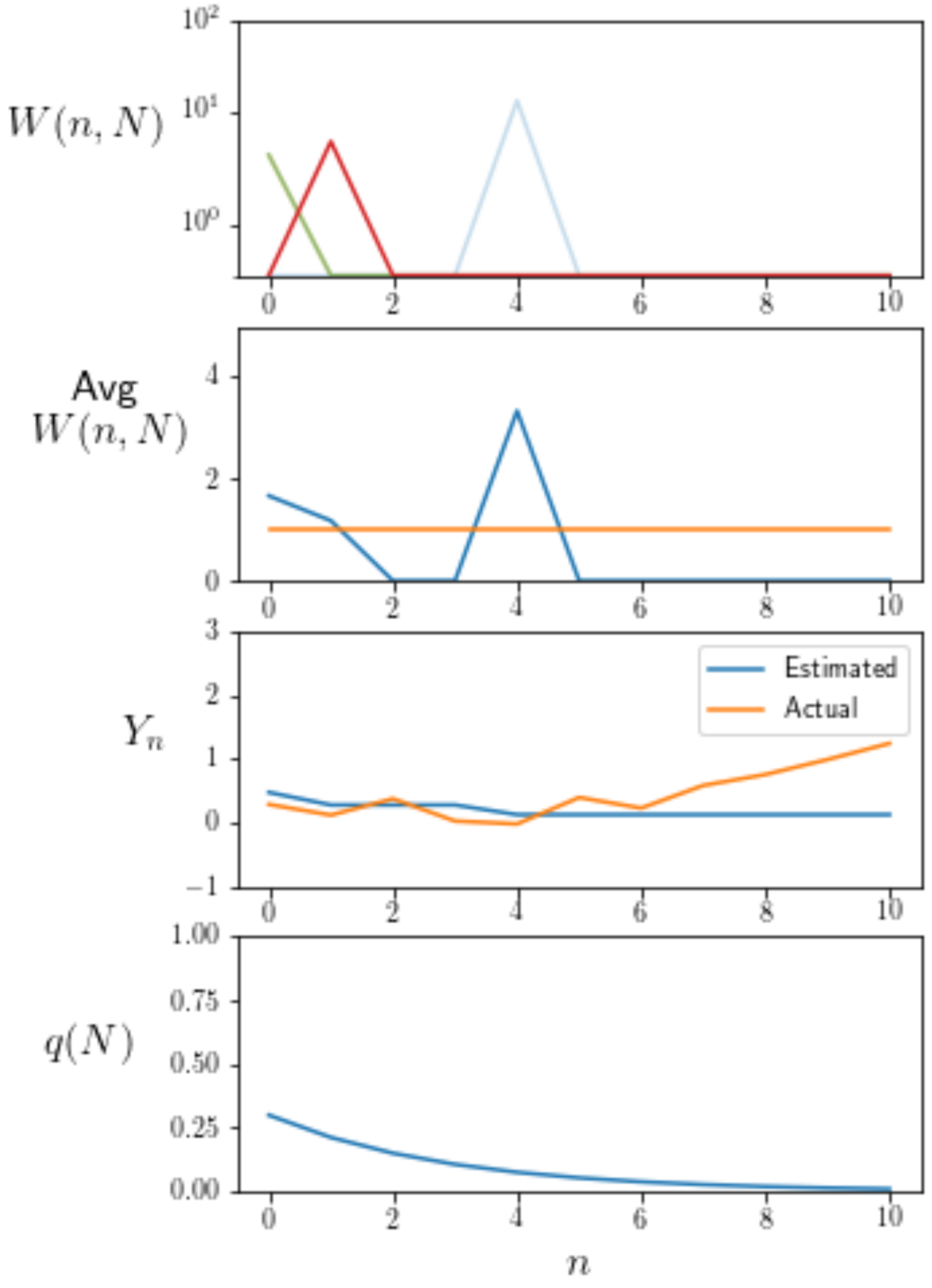
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



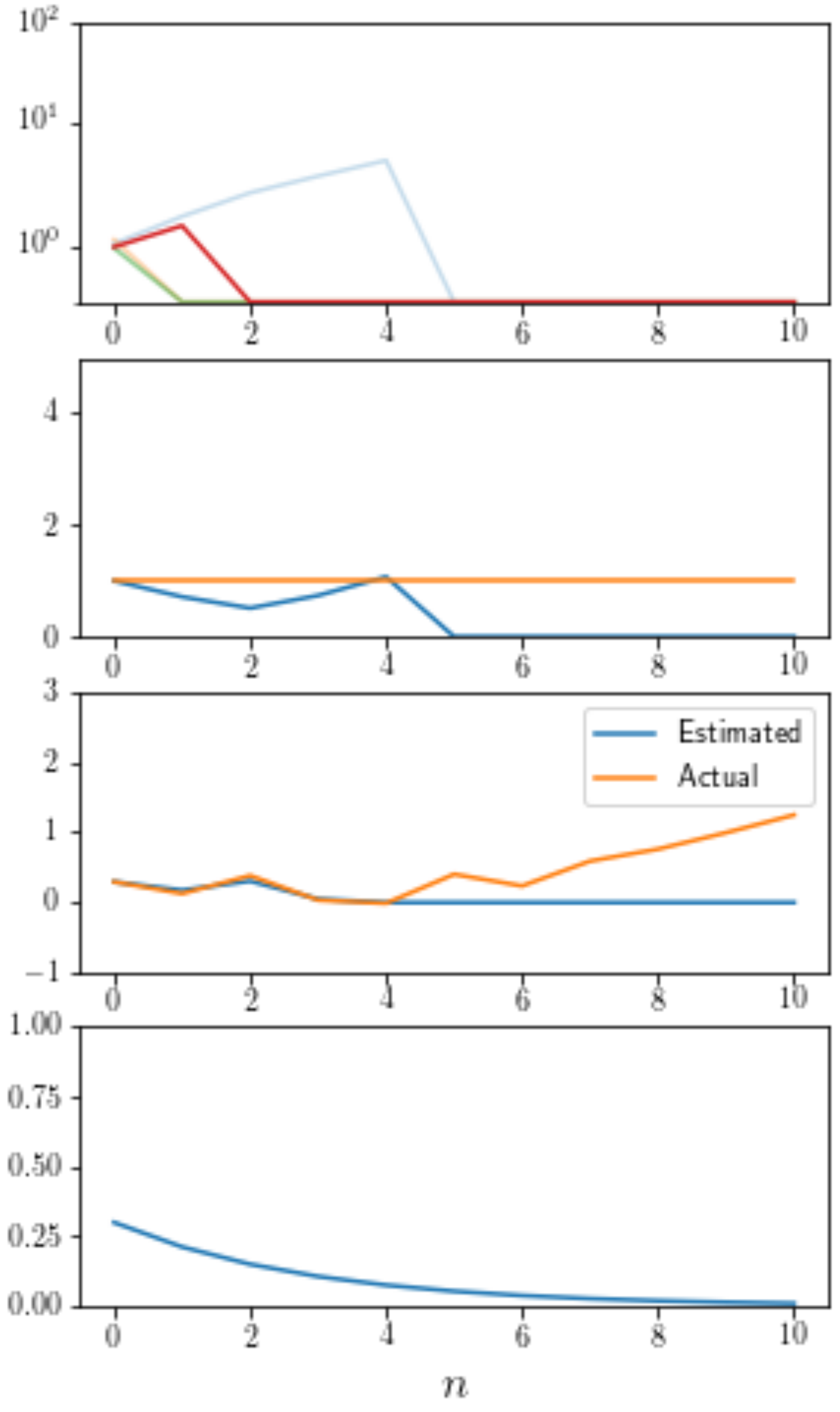
“Single sample”

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



“Russian roulette”

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



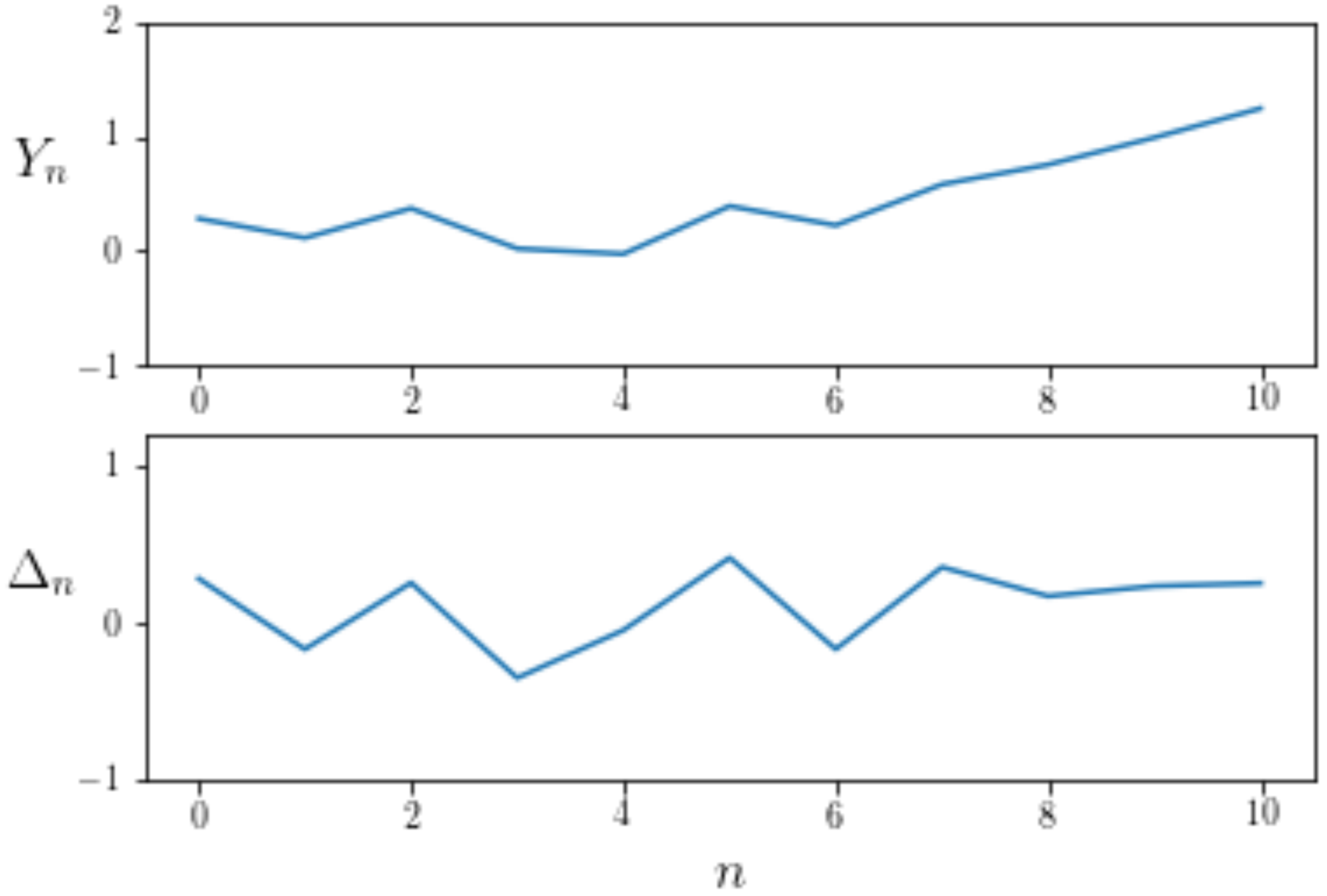
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

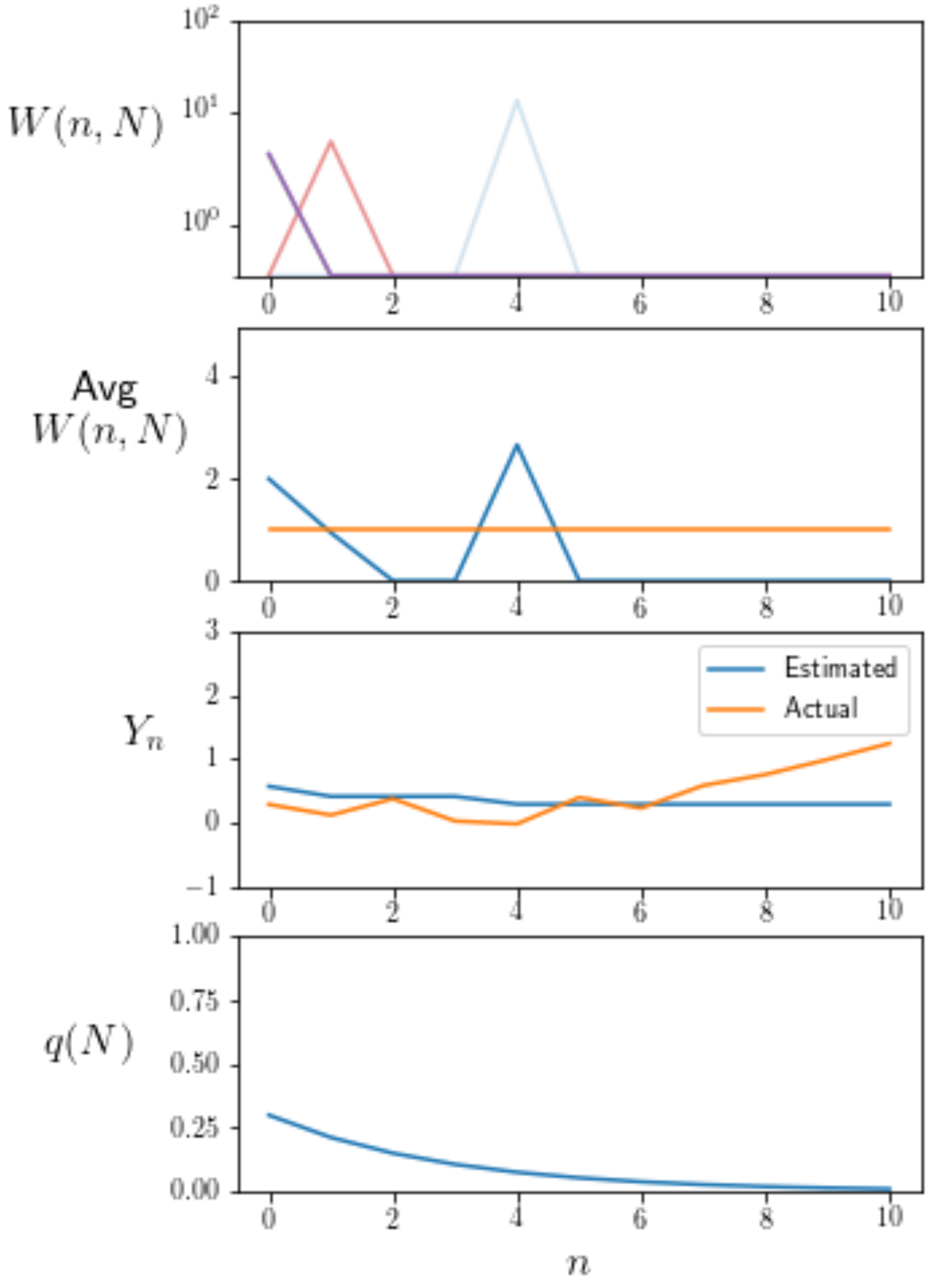
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



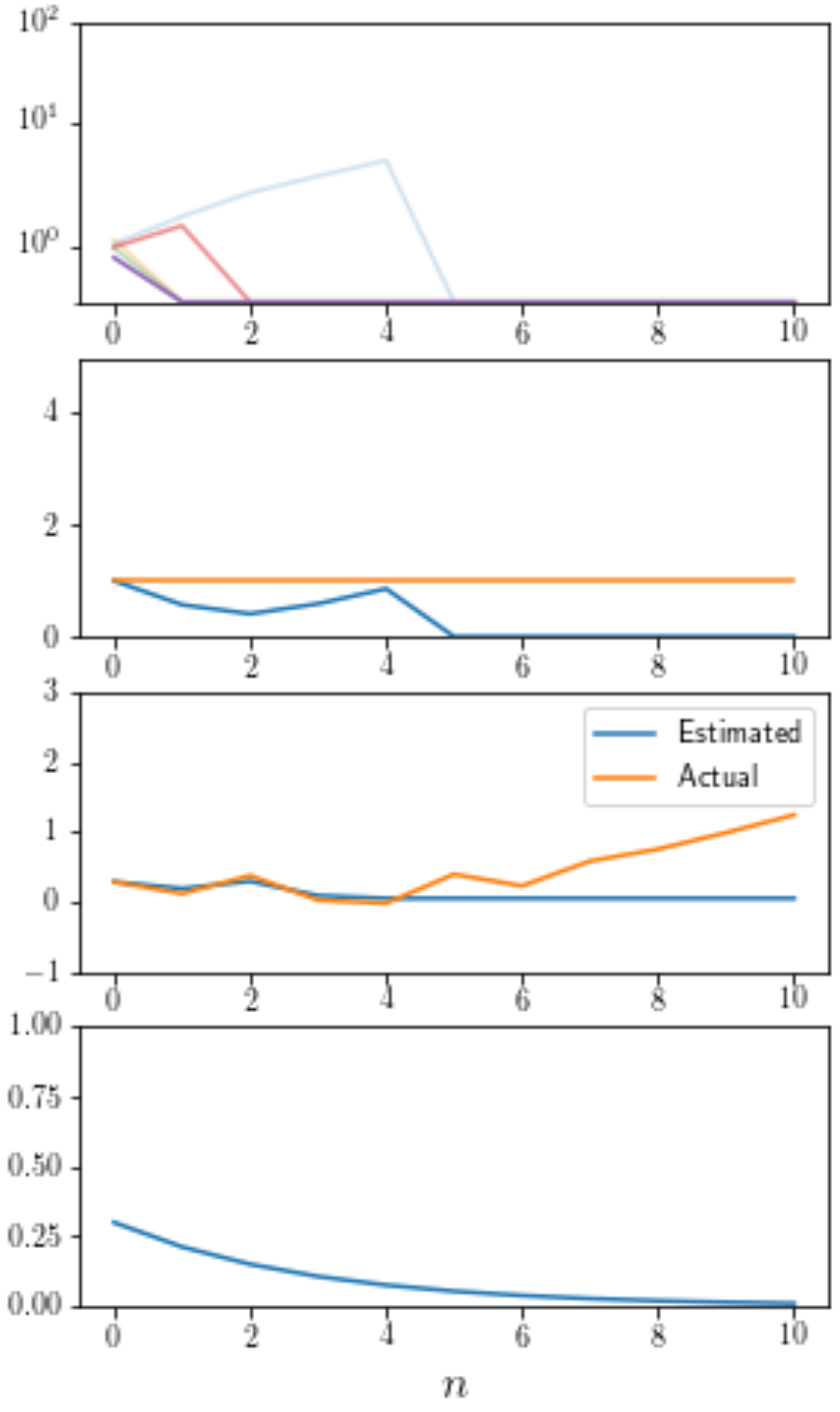
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



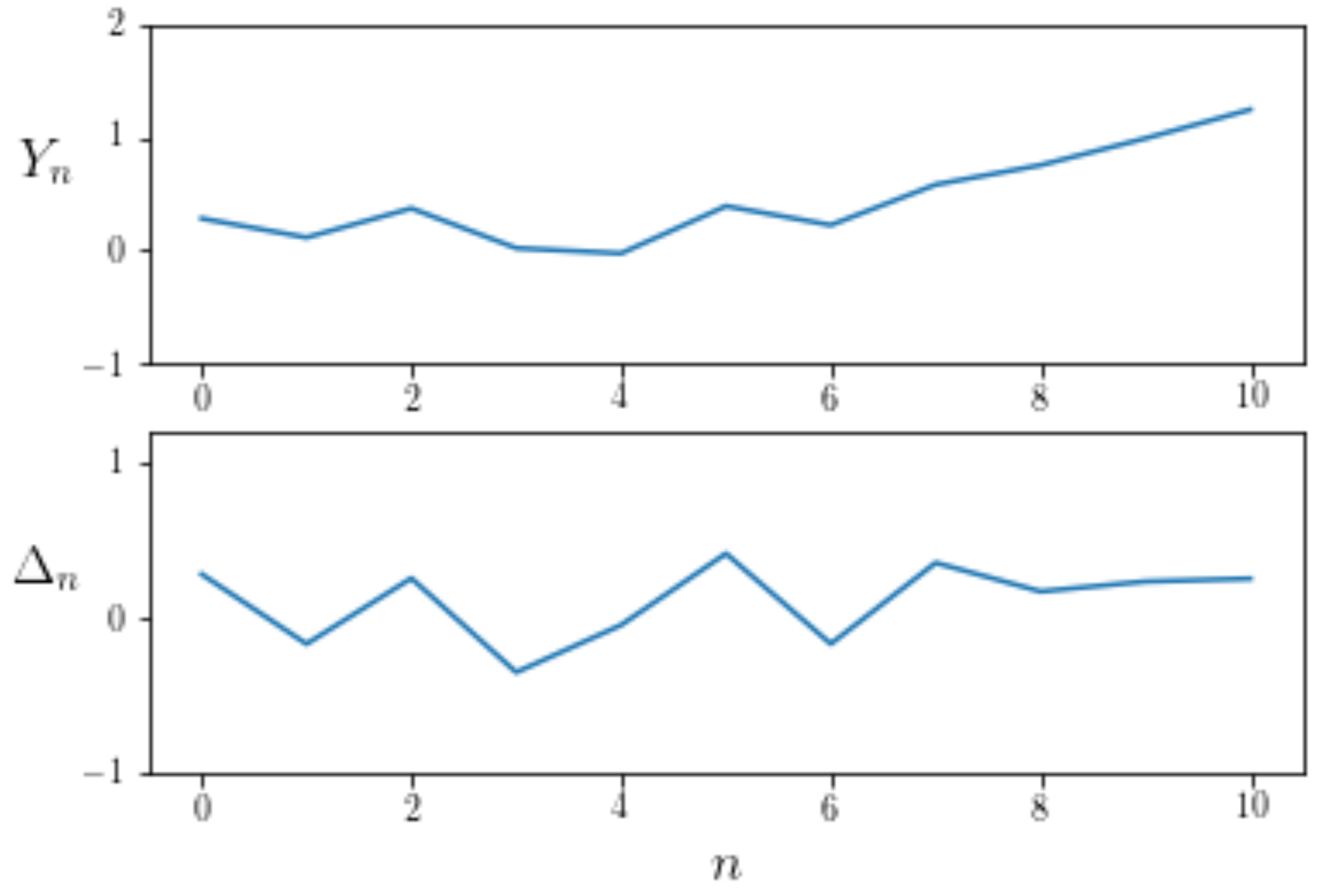
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

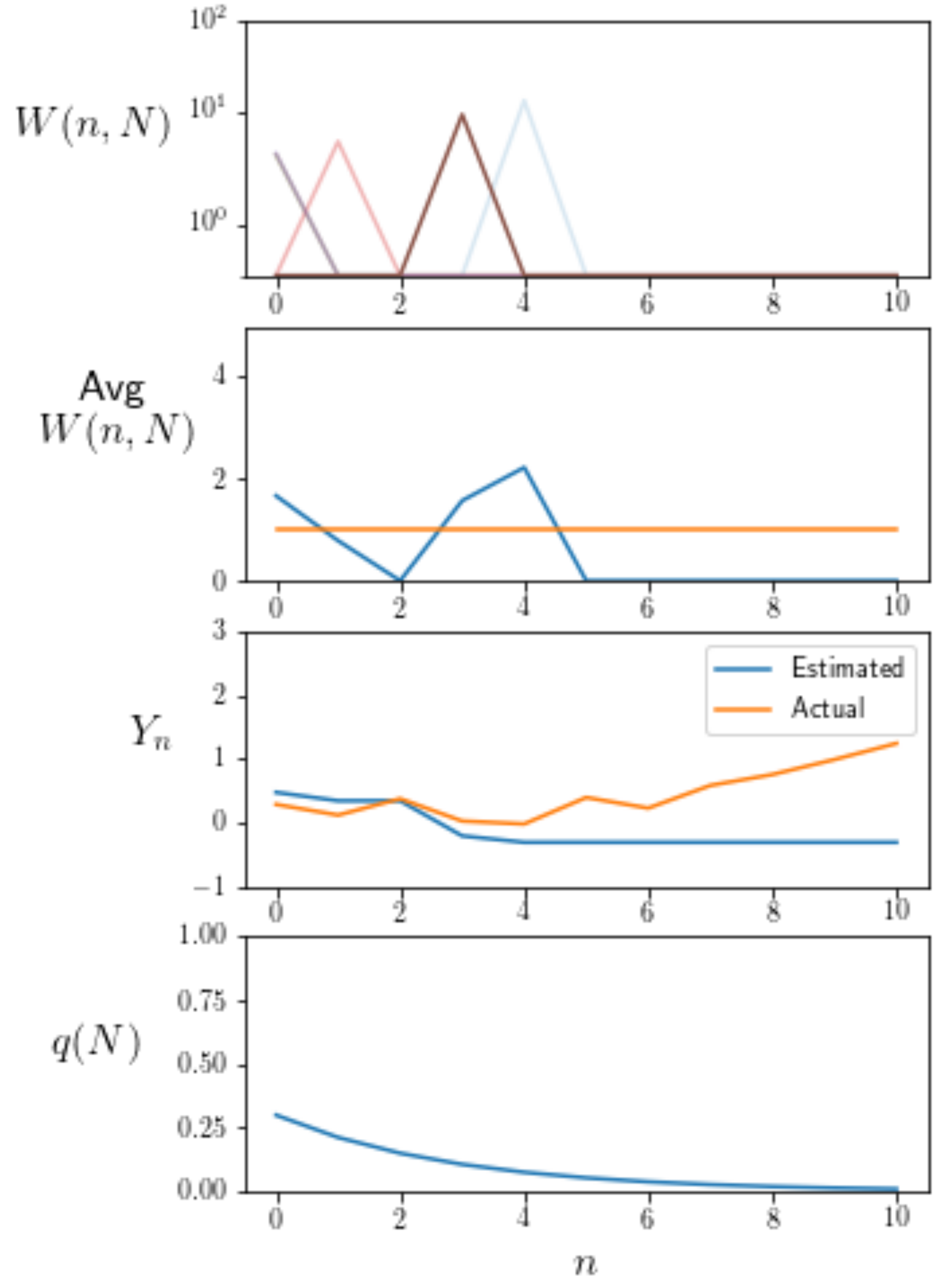


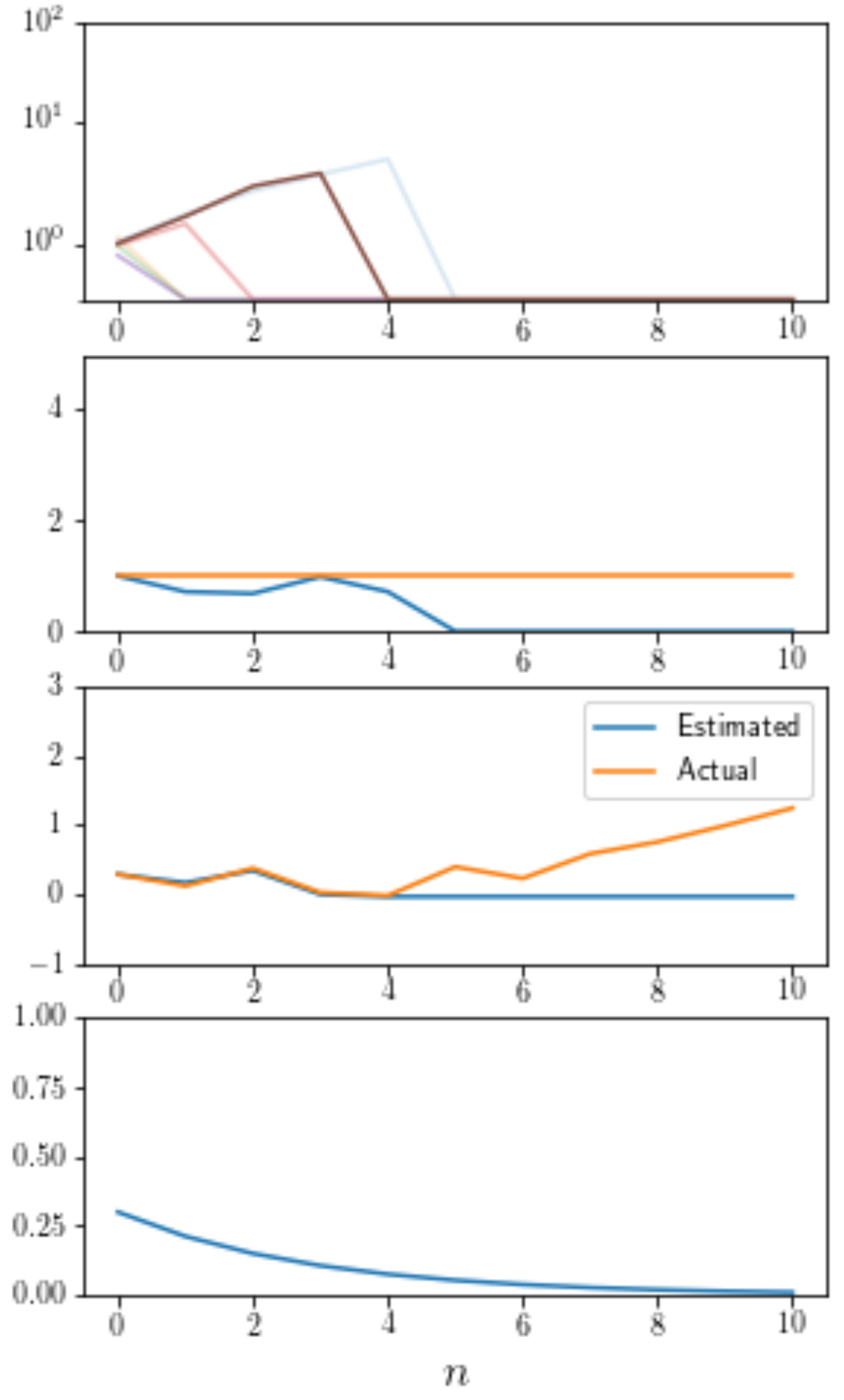
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$





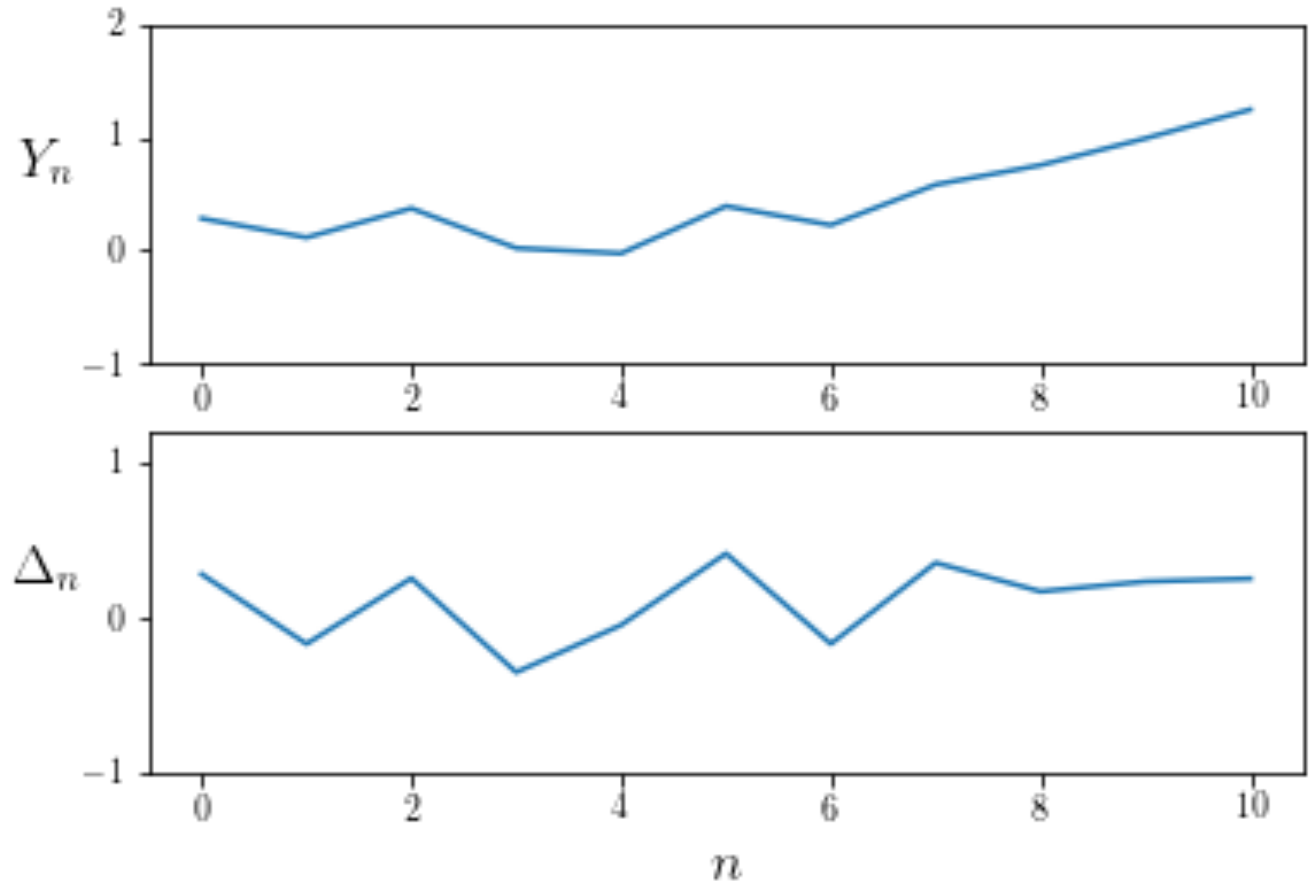
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

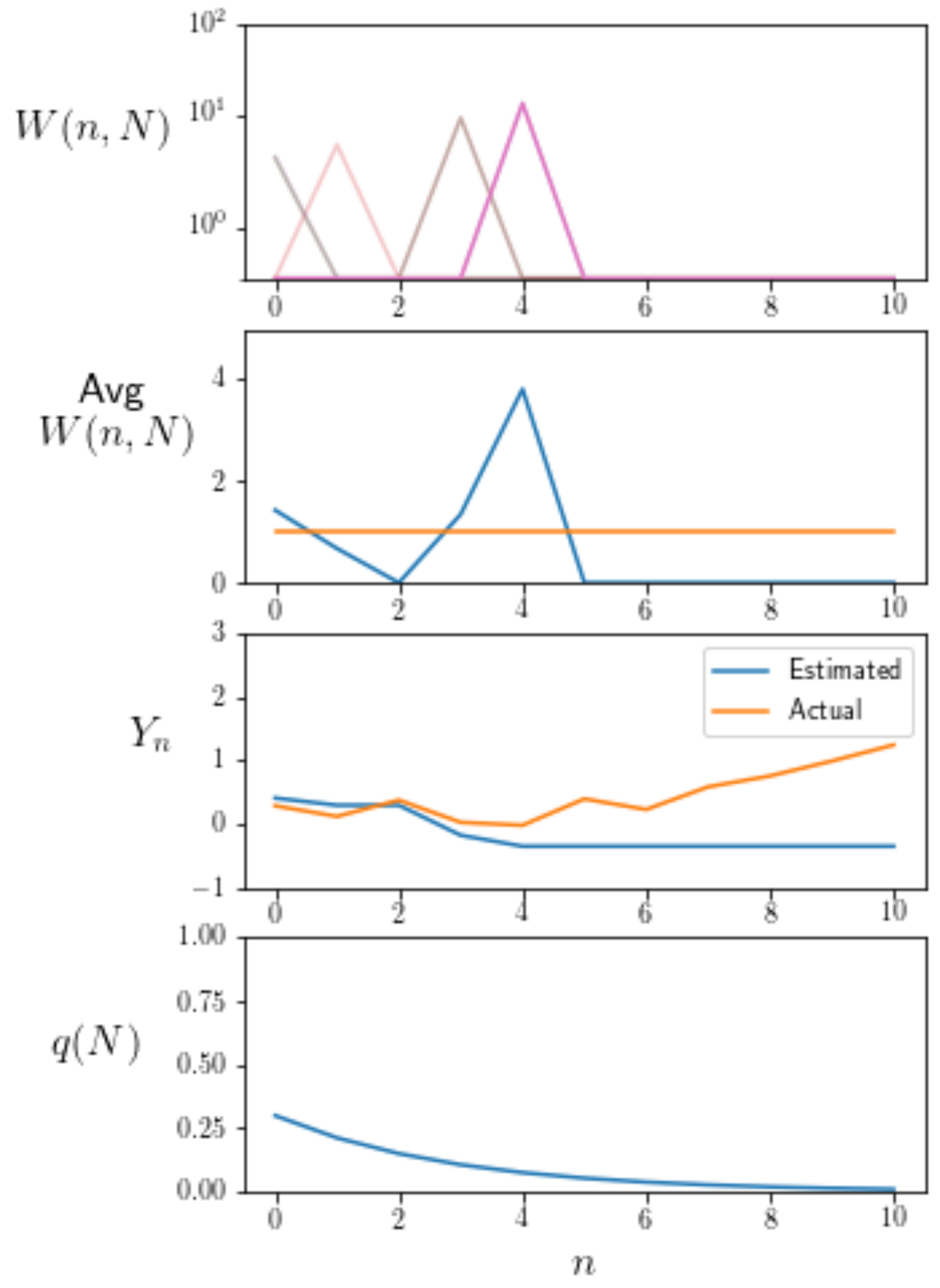
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



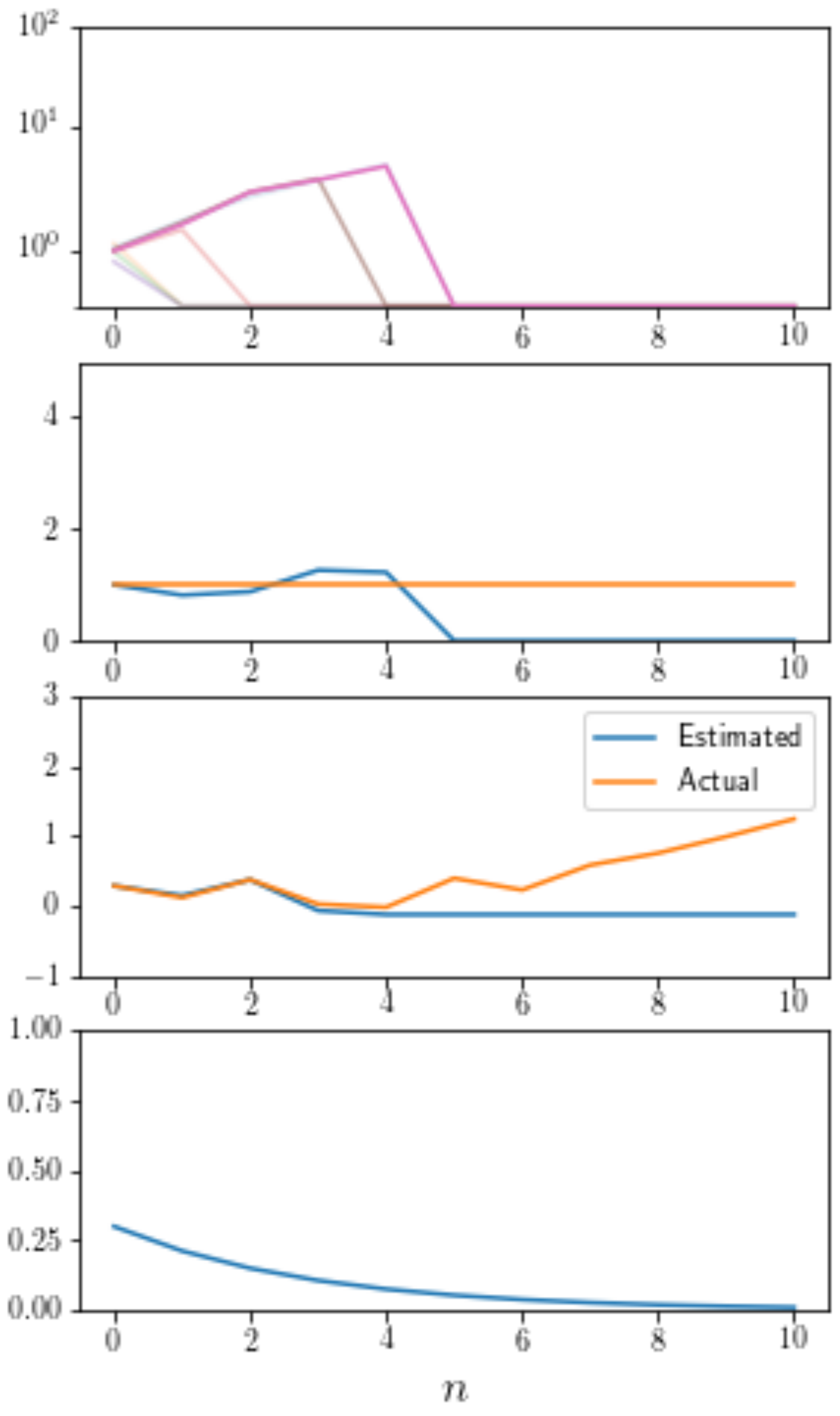
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



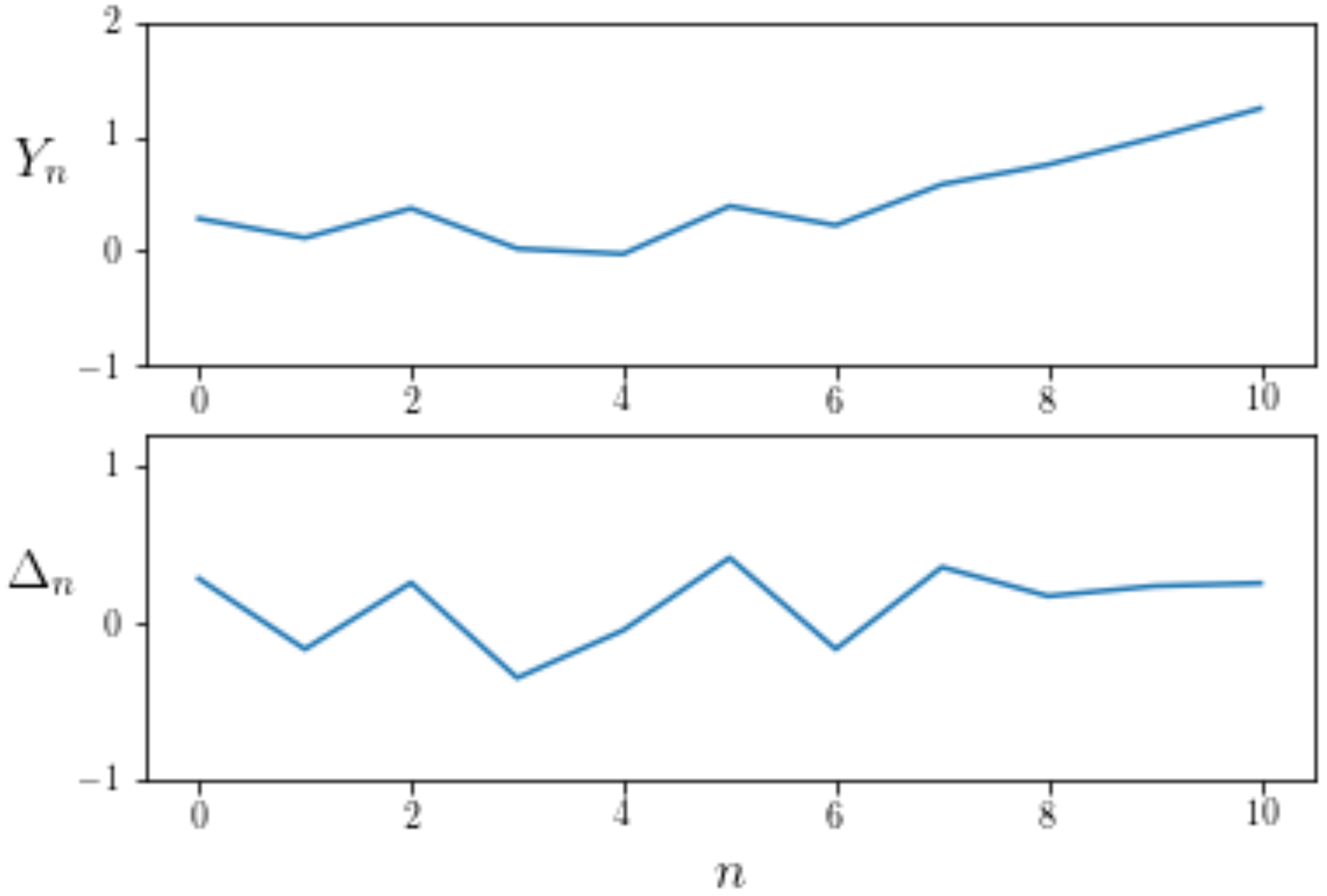
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

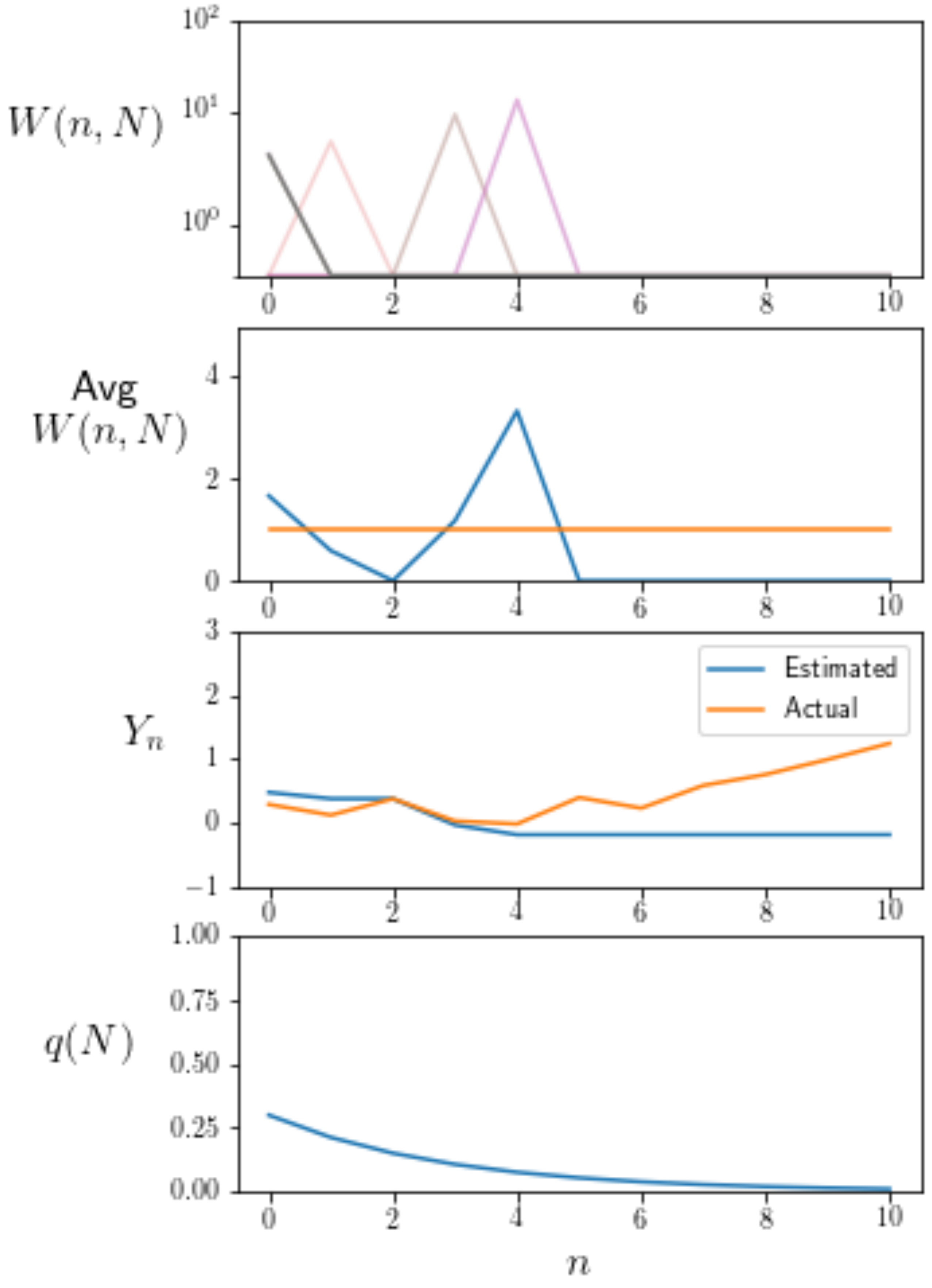
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



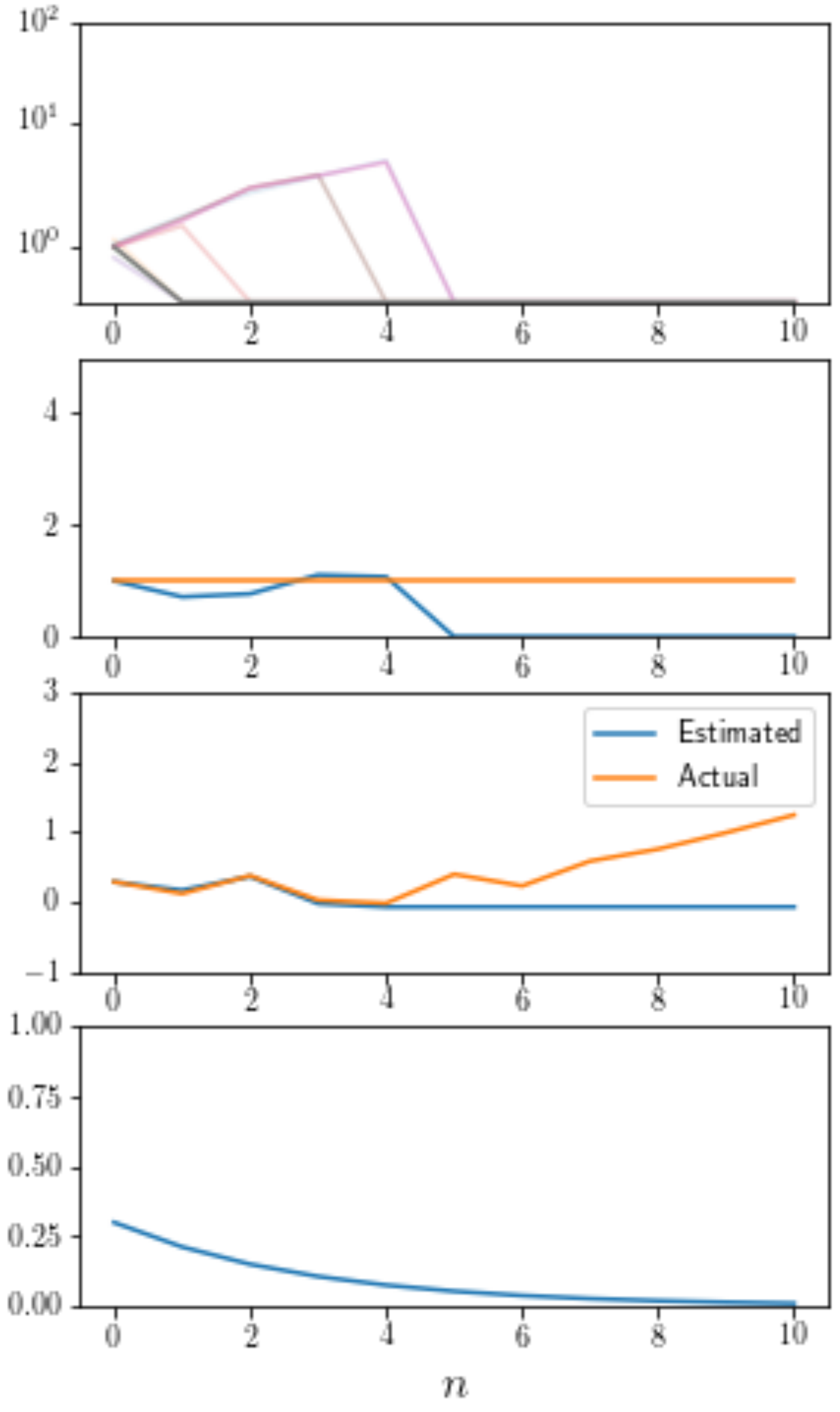
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



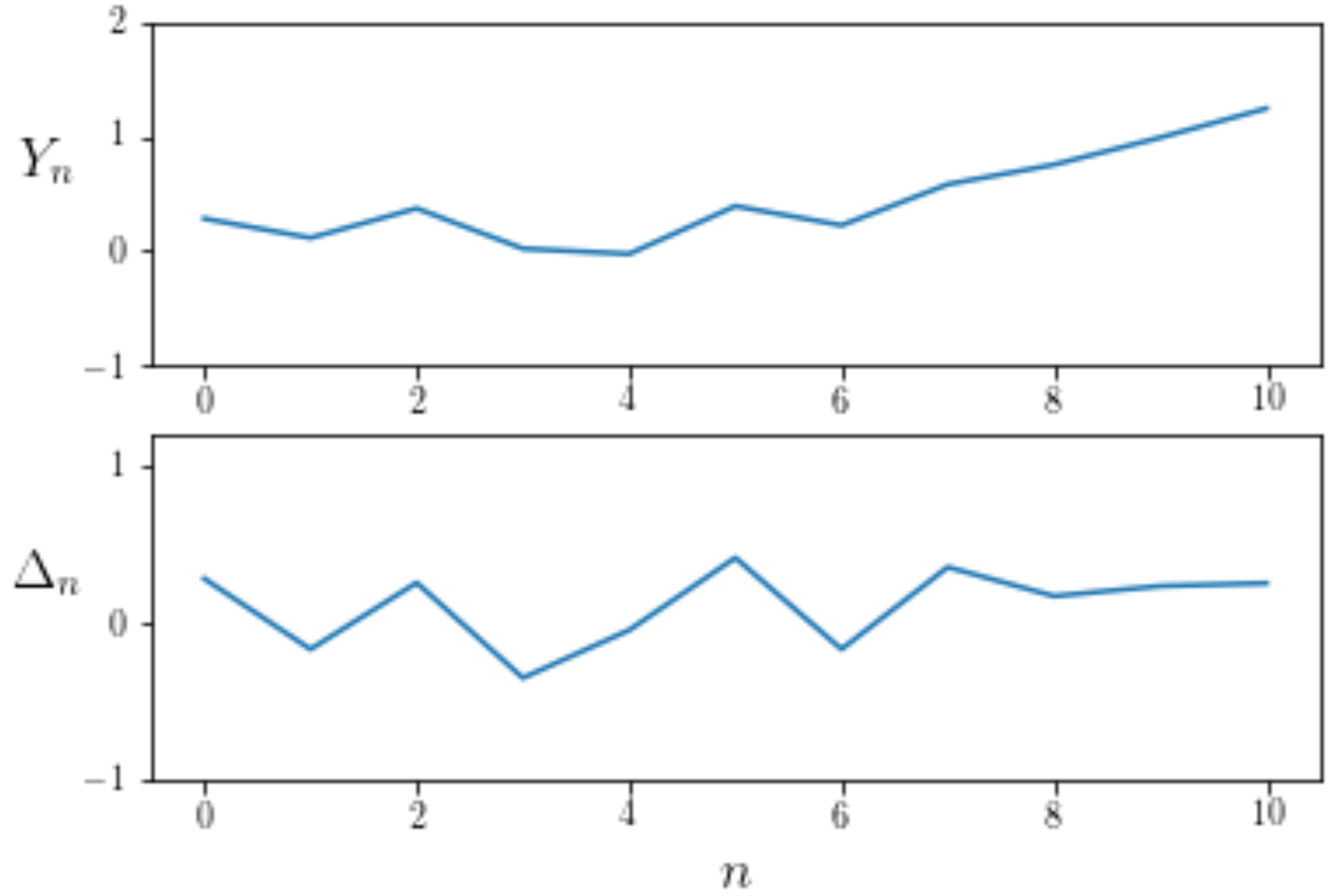
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

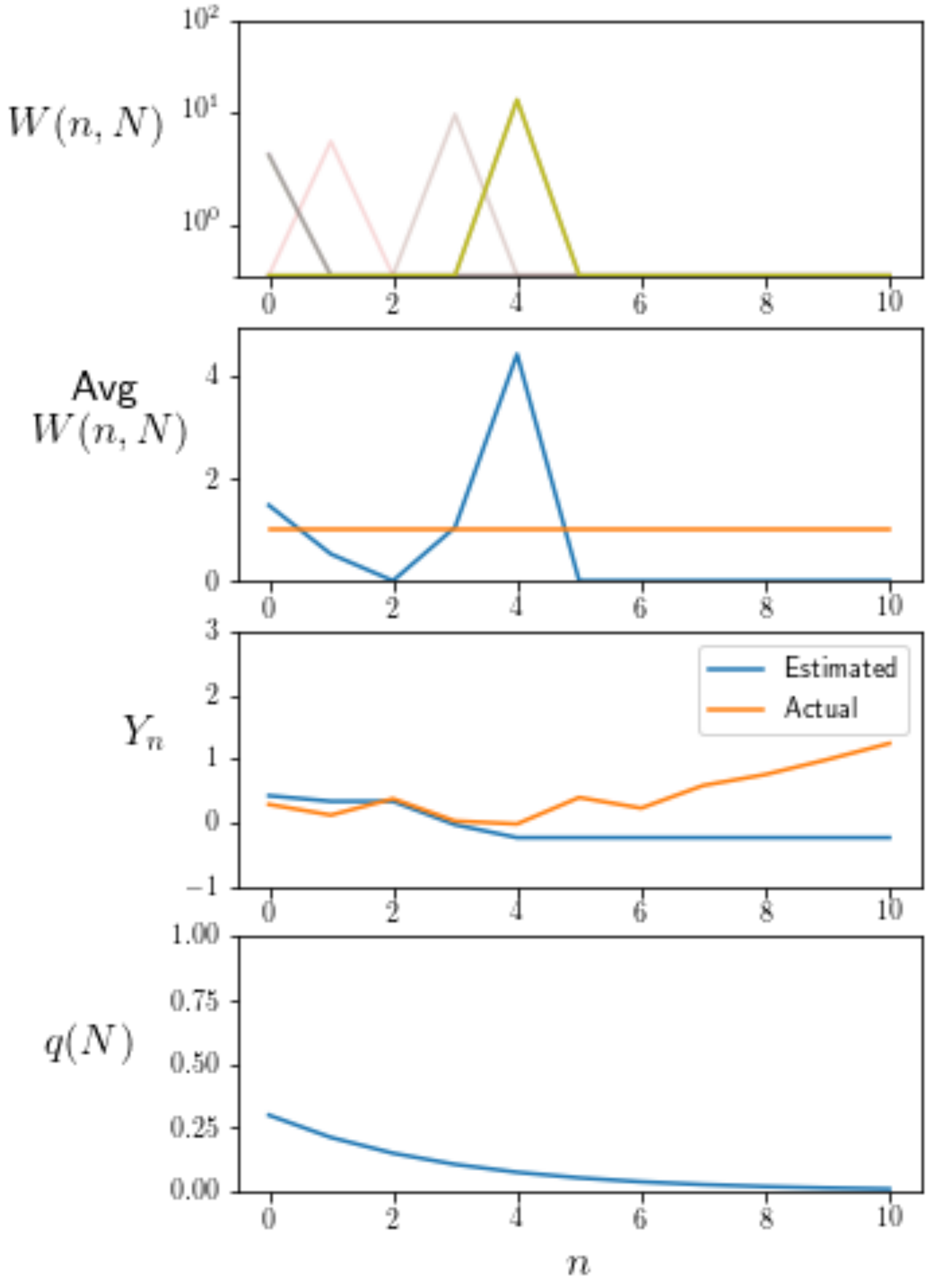
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



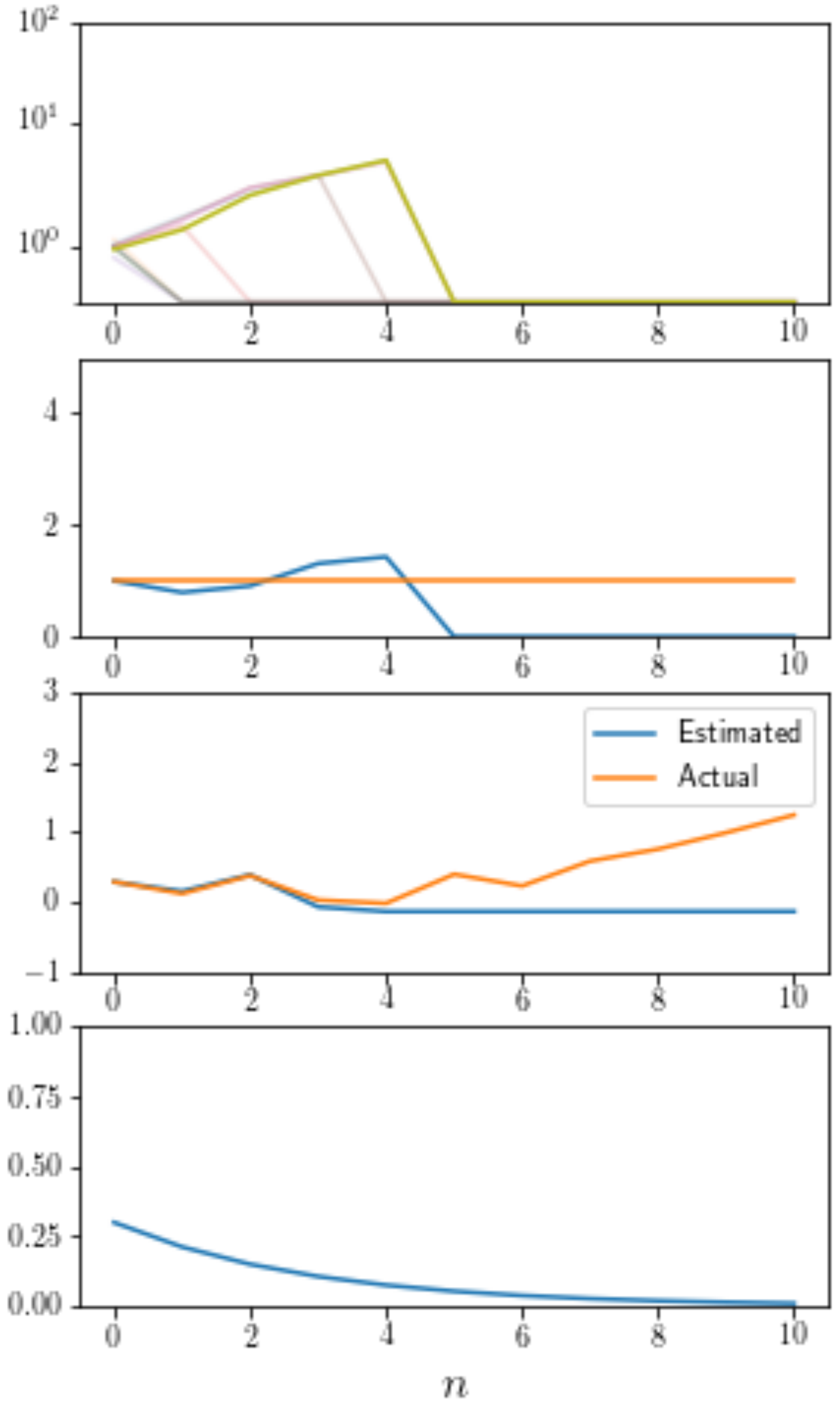
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



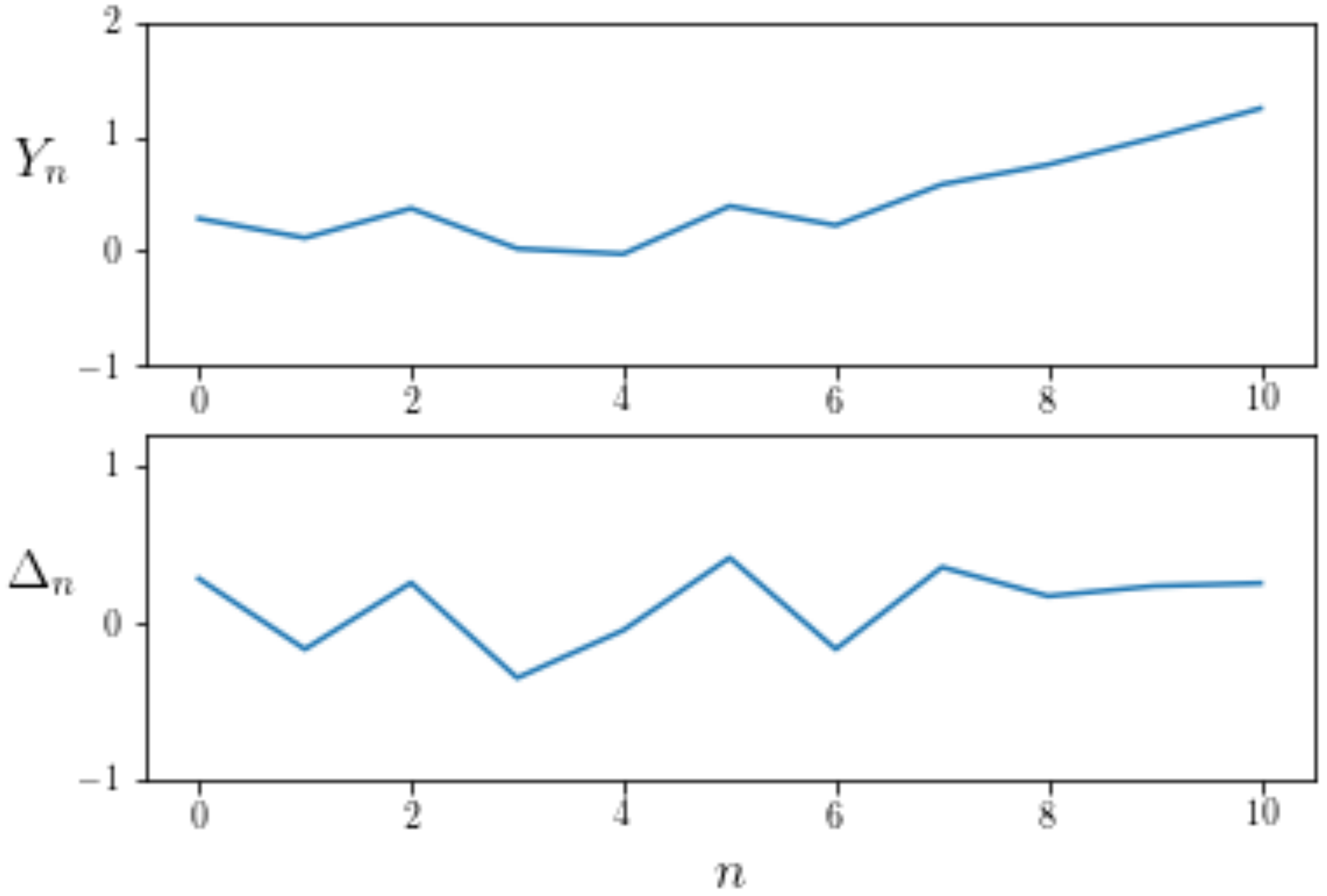
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

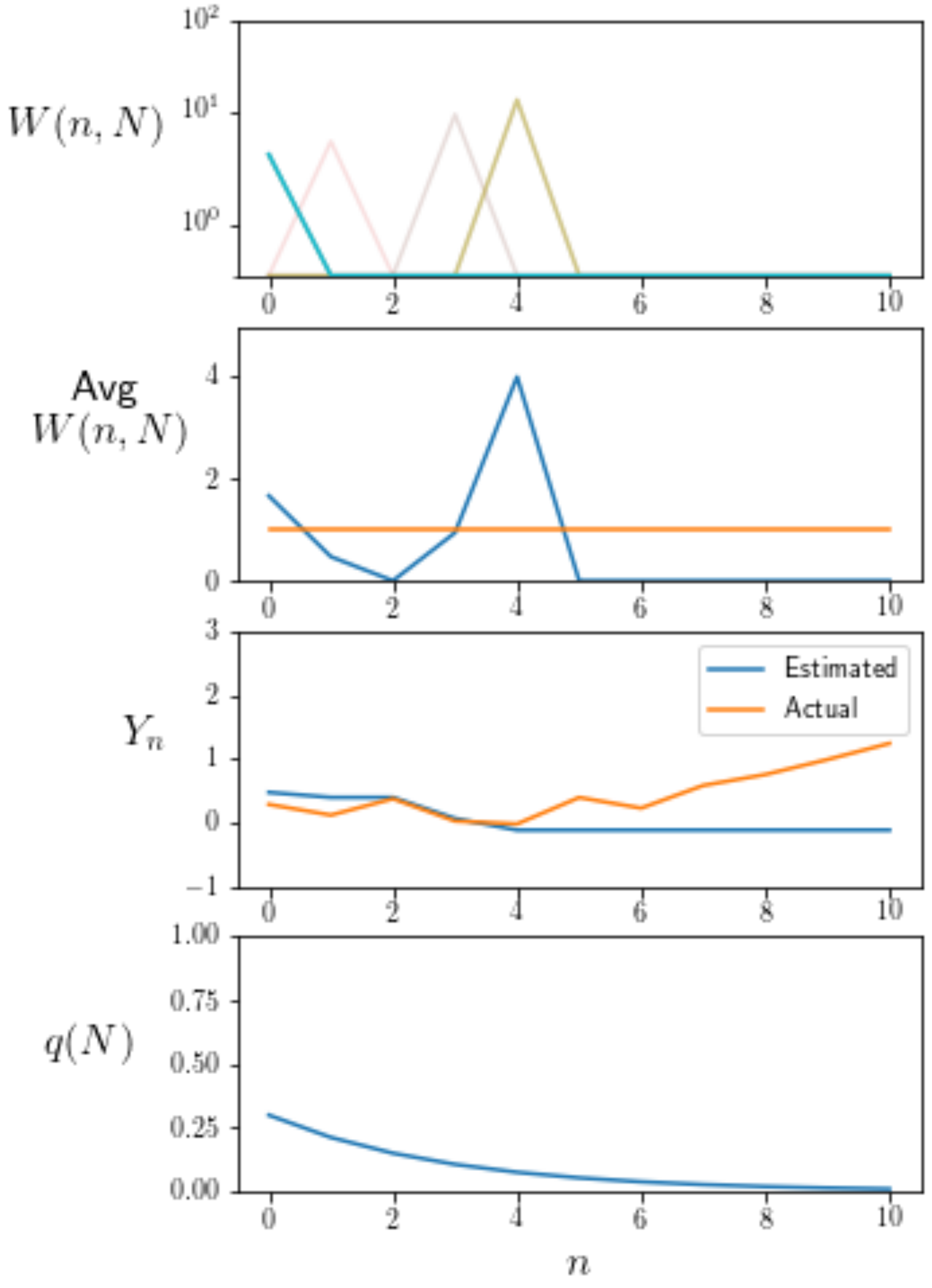
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



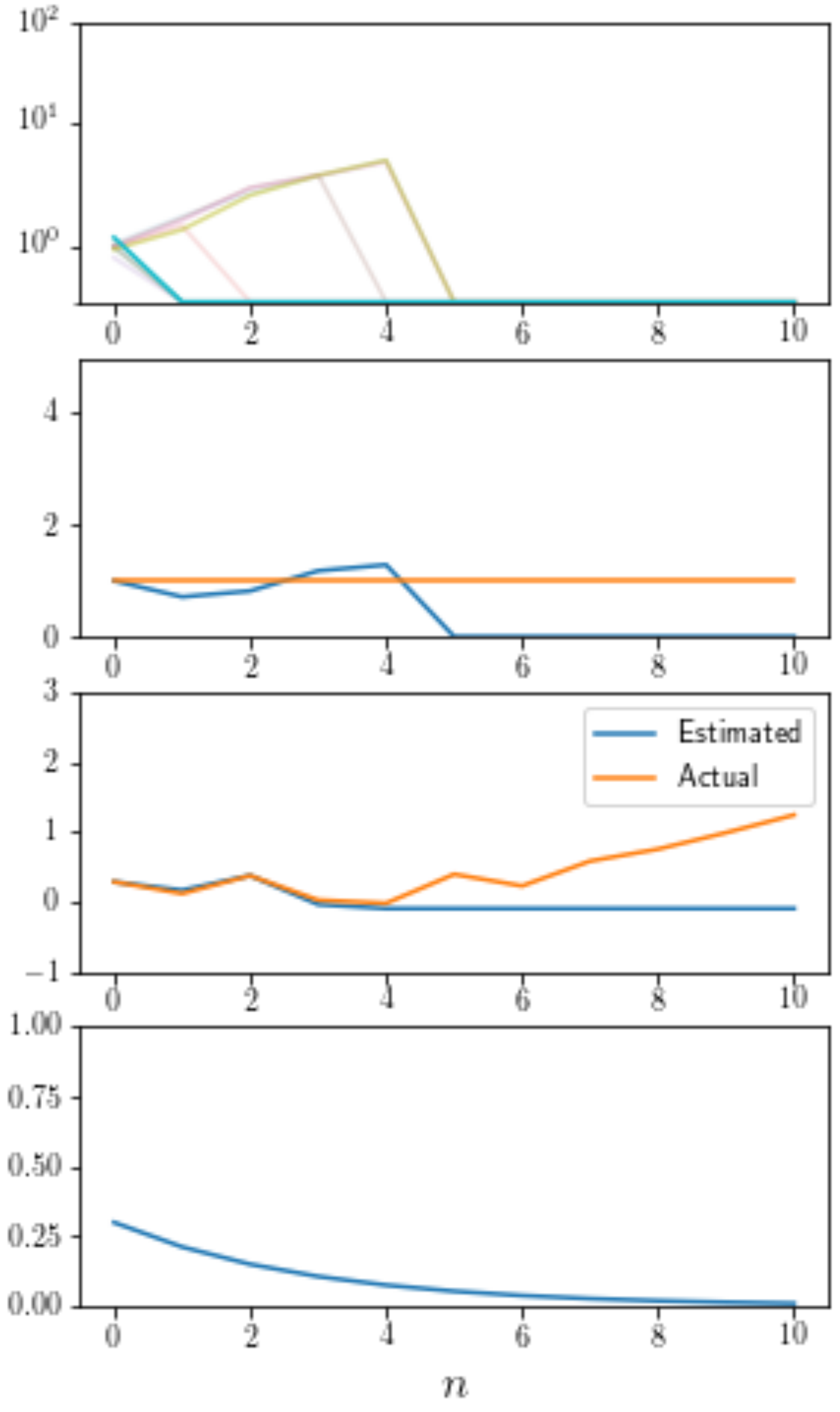
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



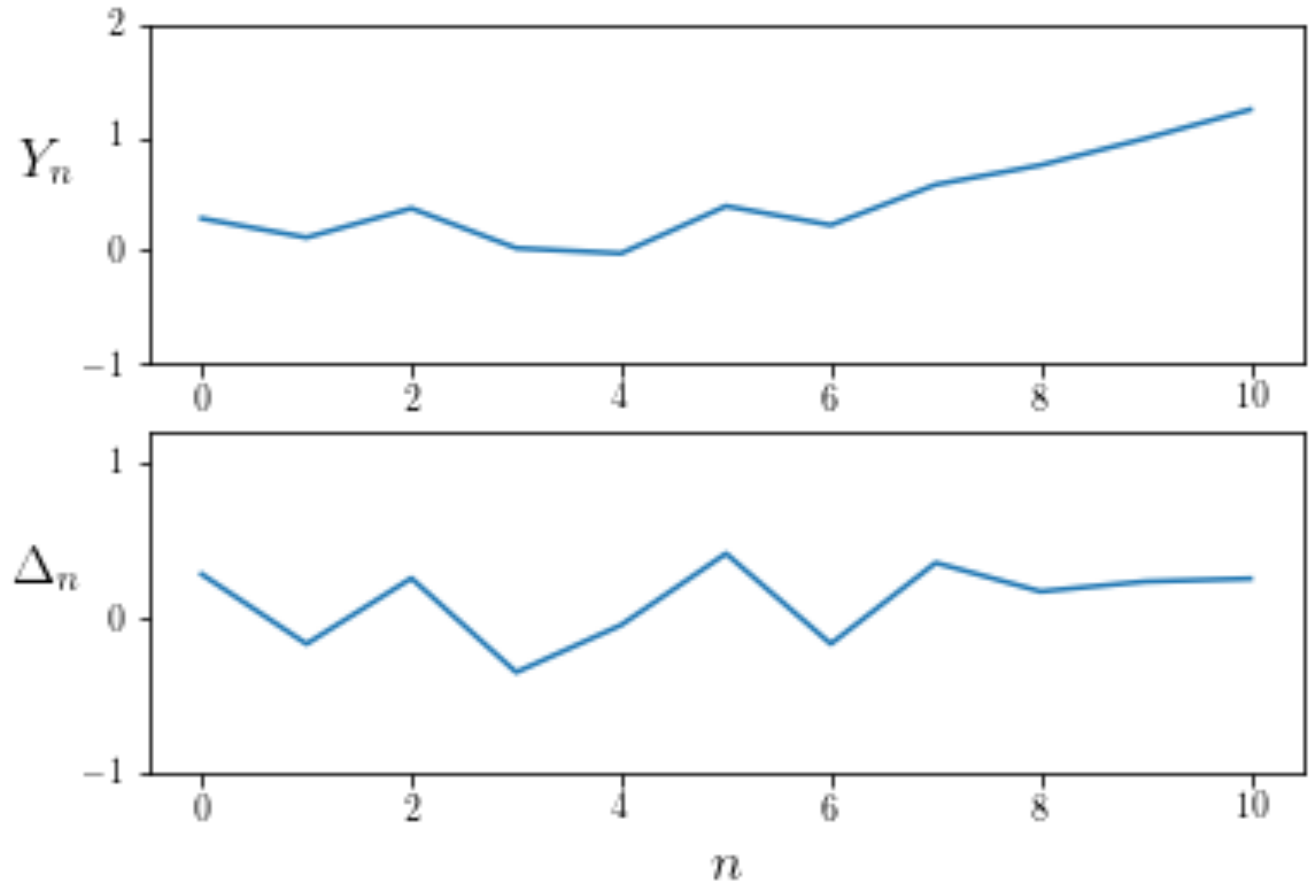
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

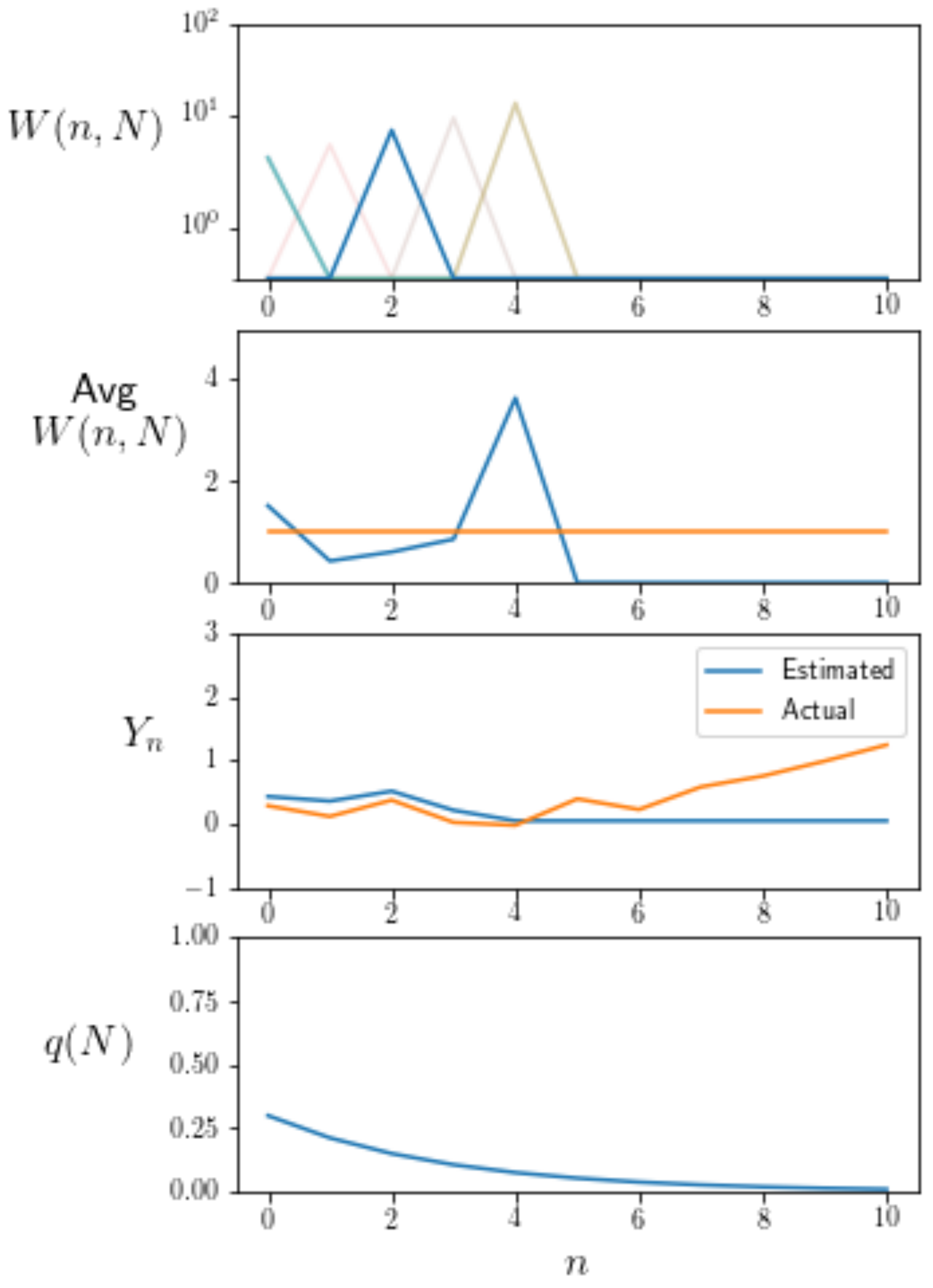
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



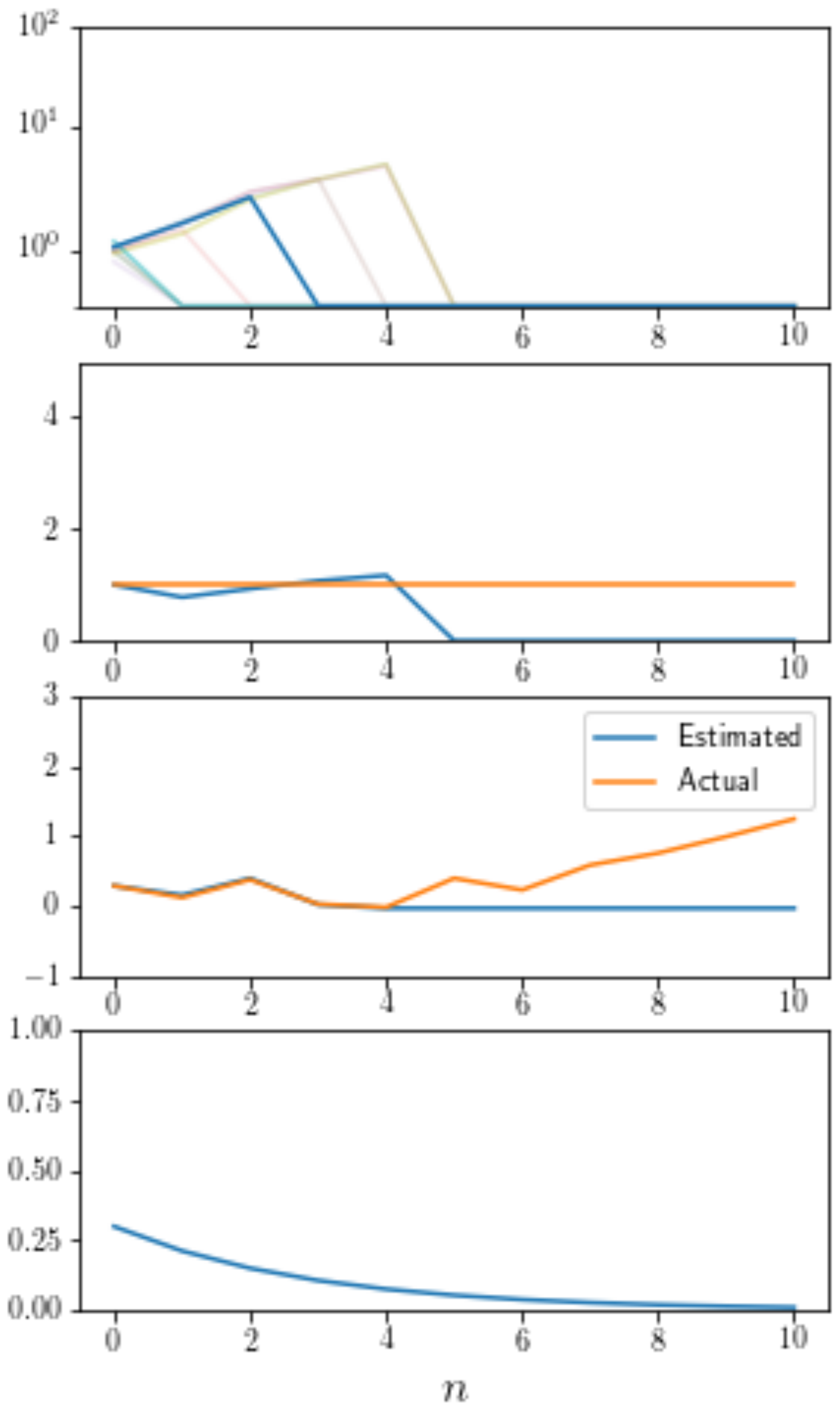
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



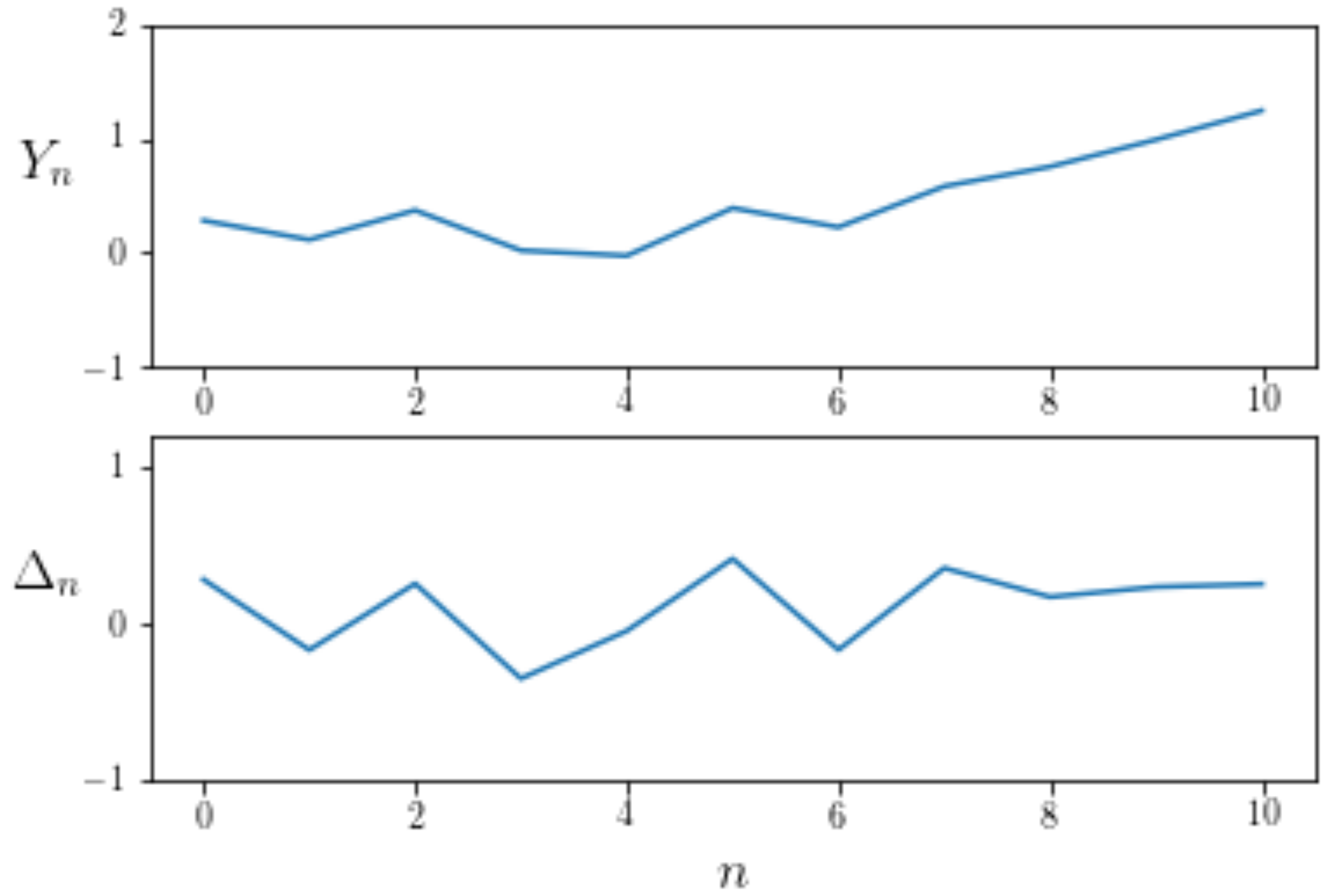
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

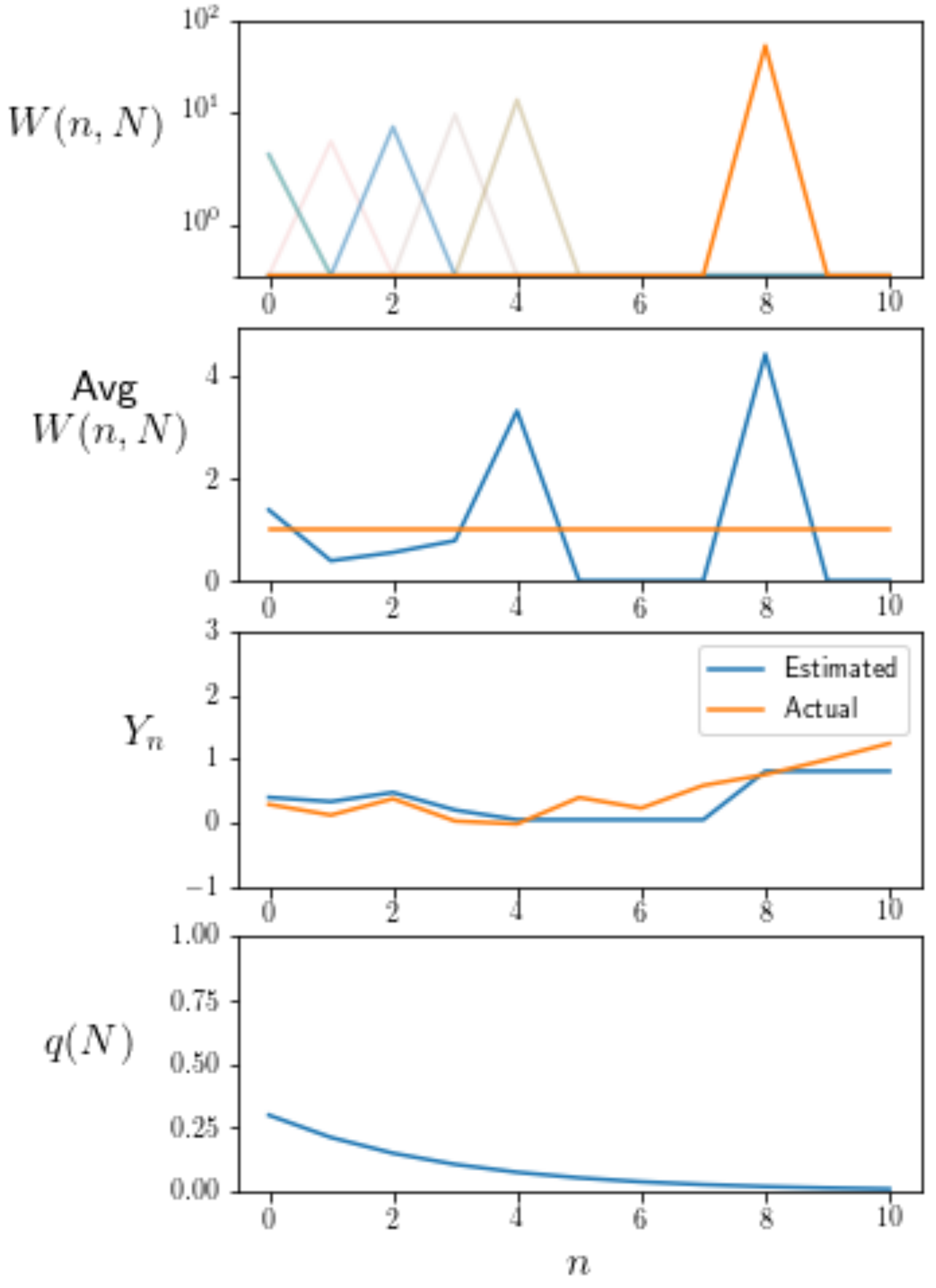
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



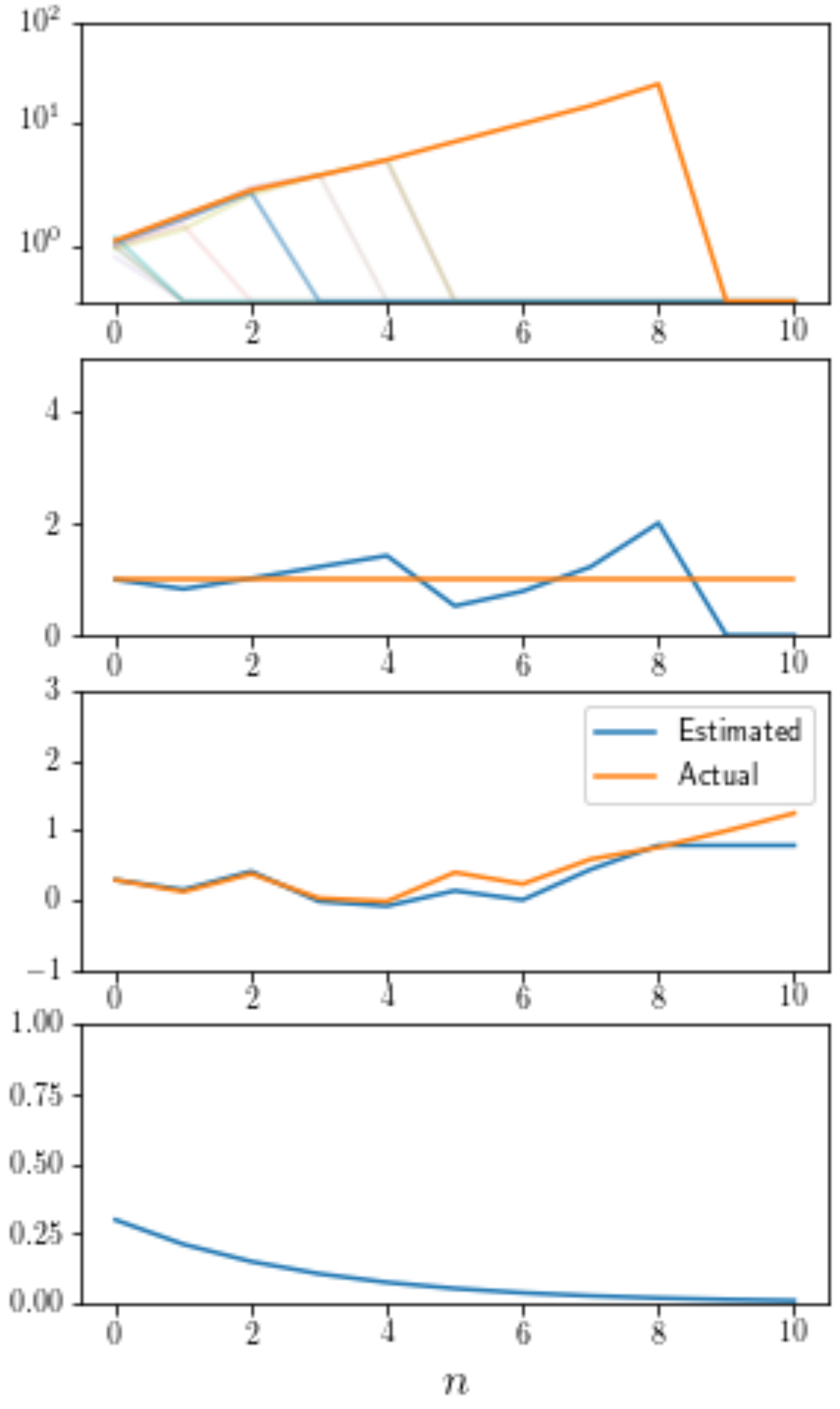
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



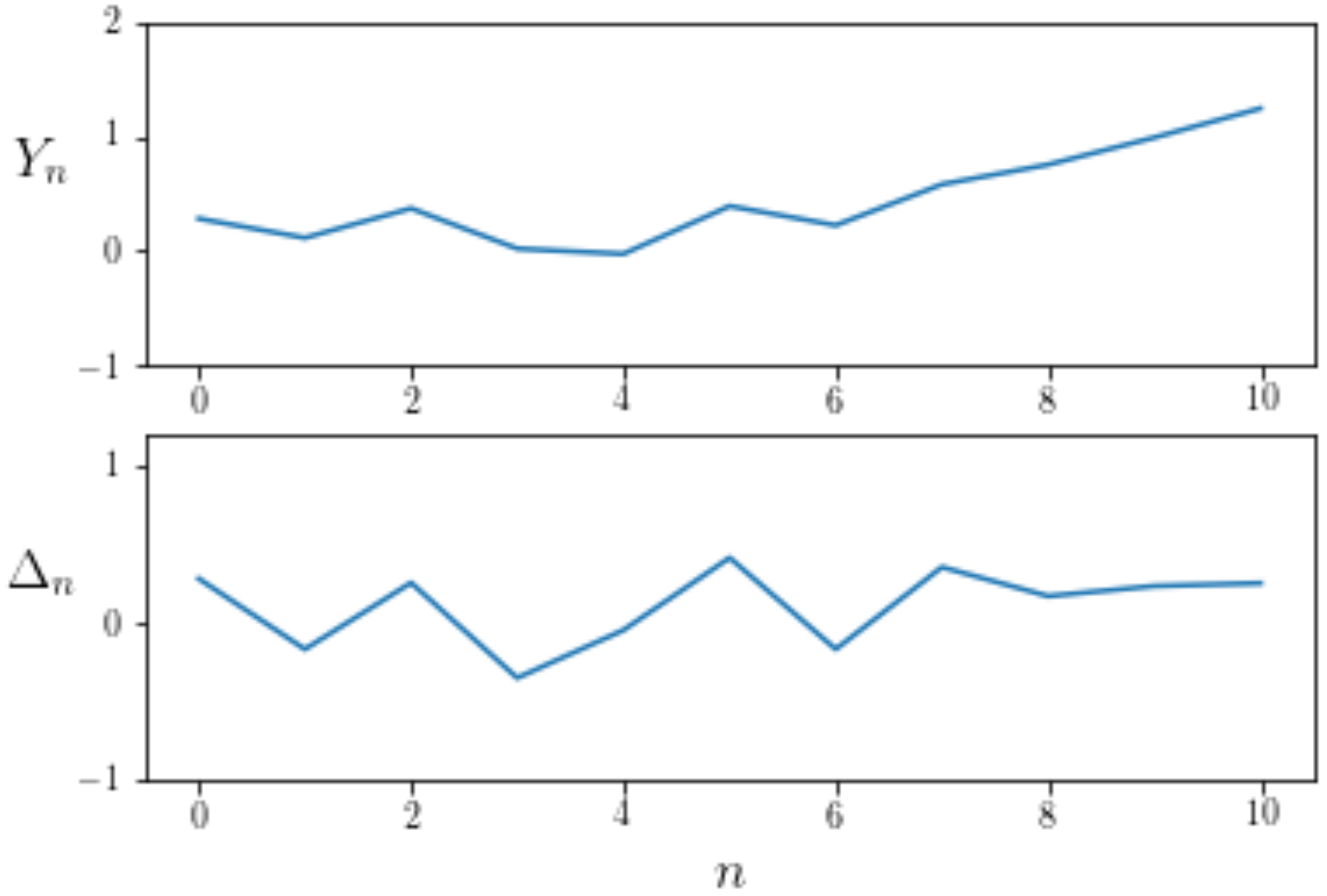
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

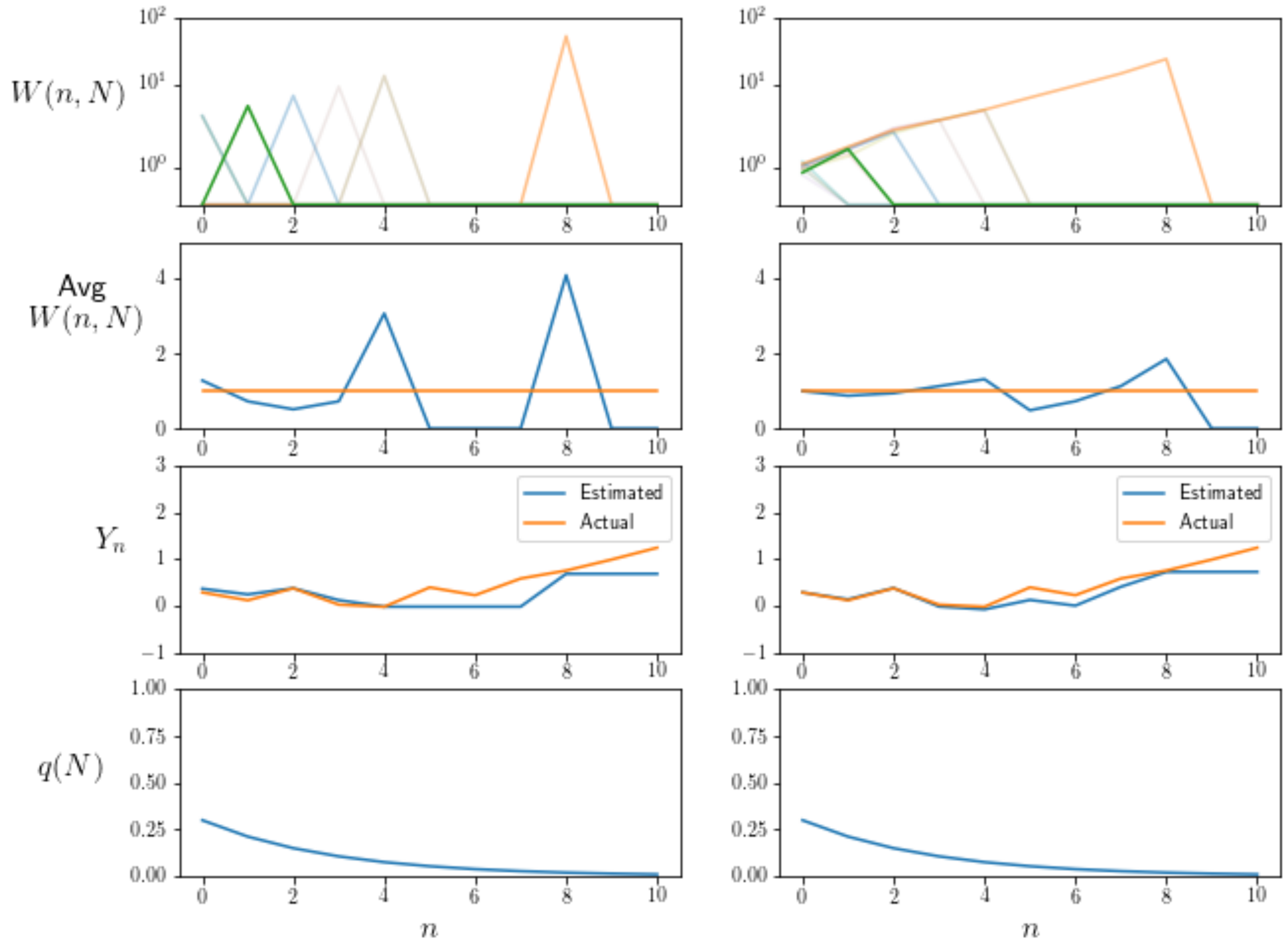


“Single sample”

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

“Russian roulette”

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



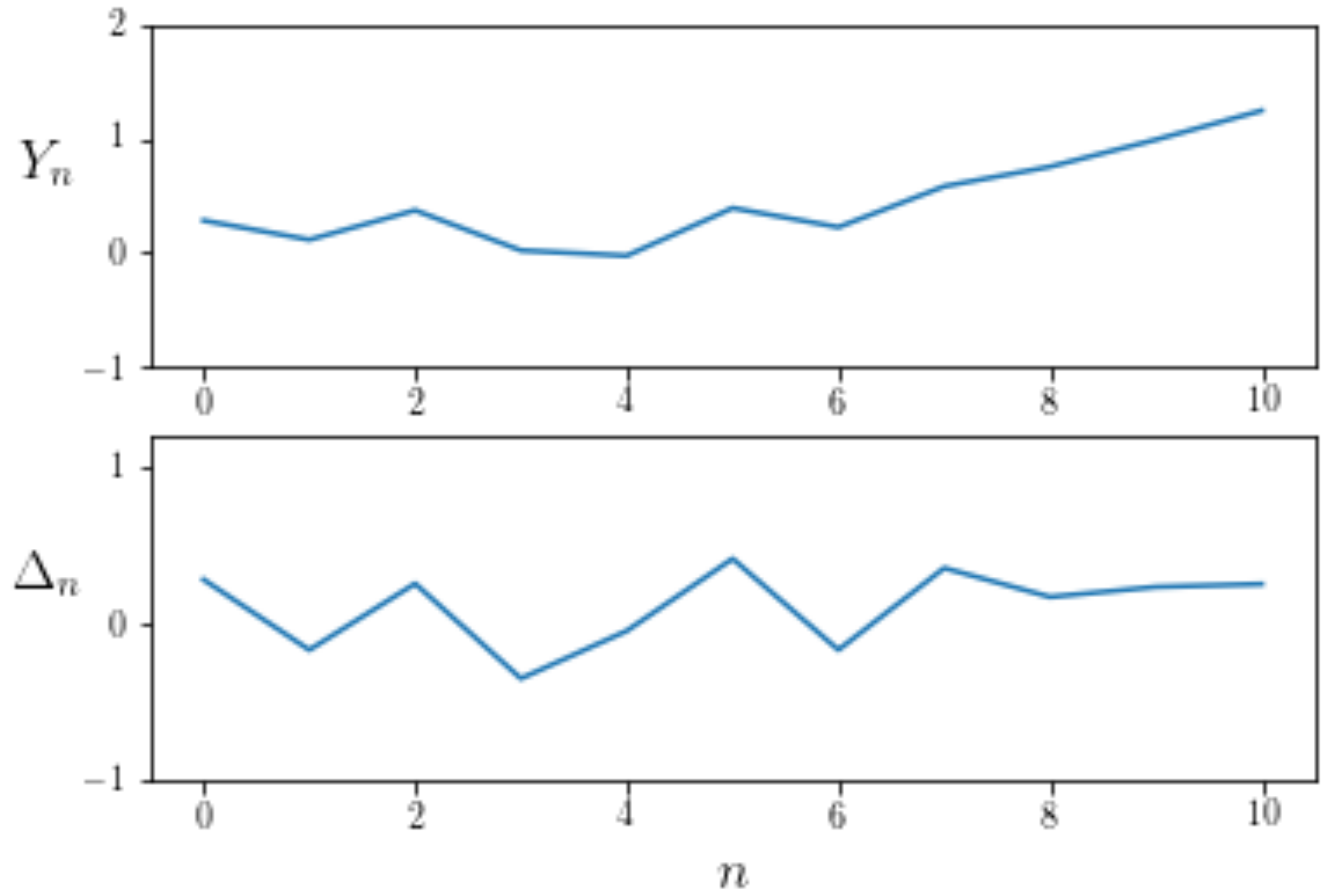
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

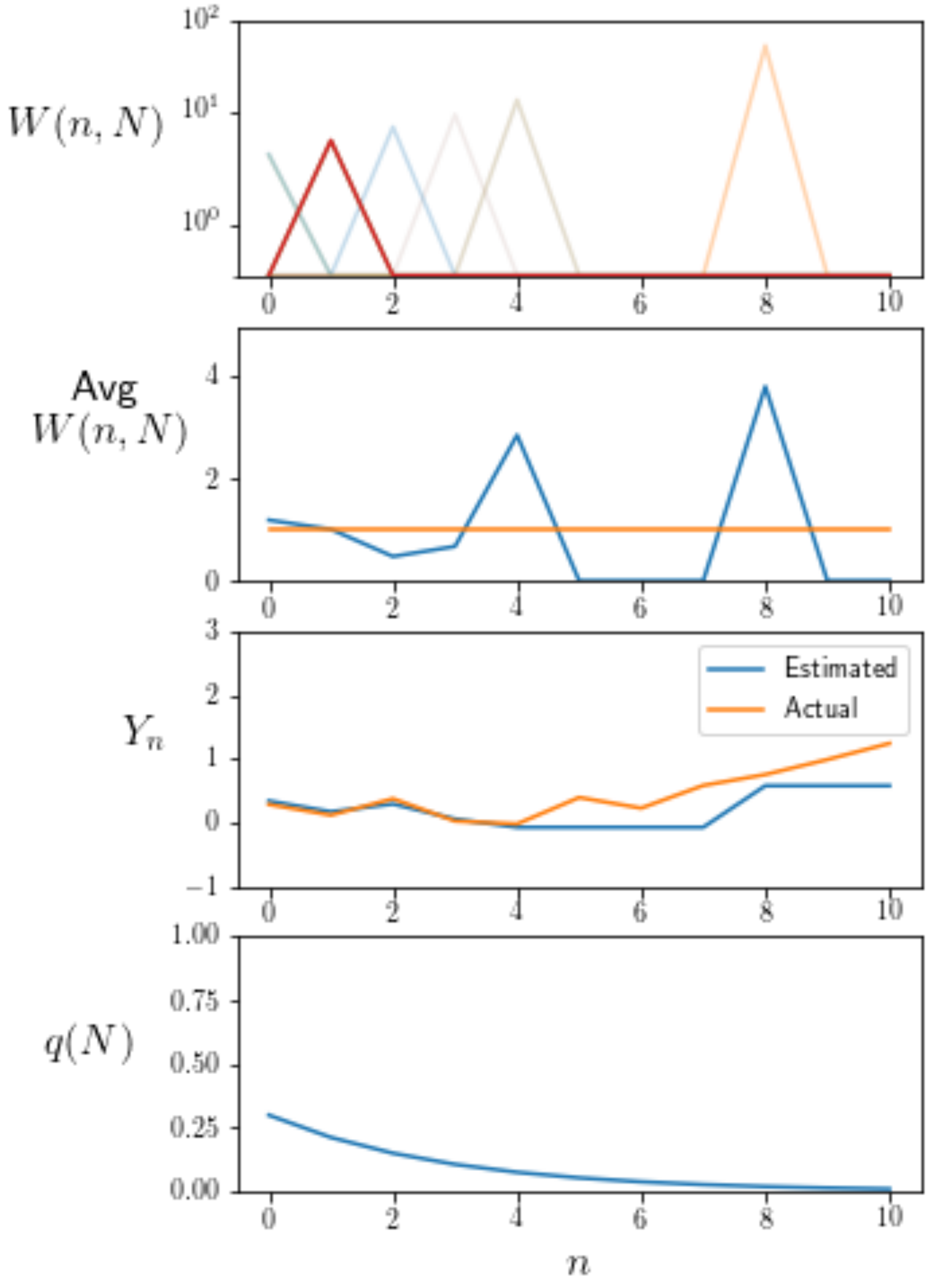
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



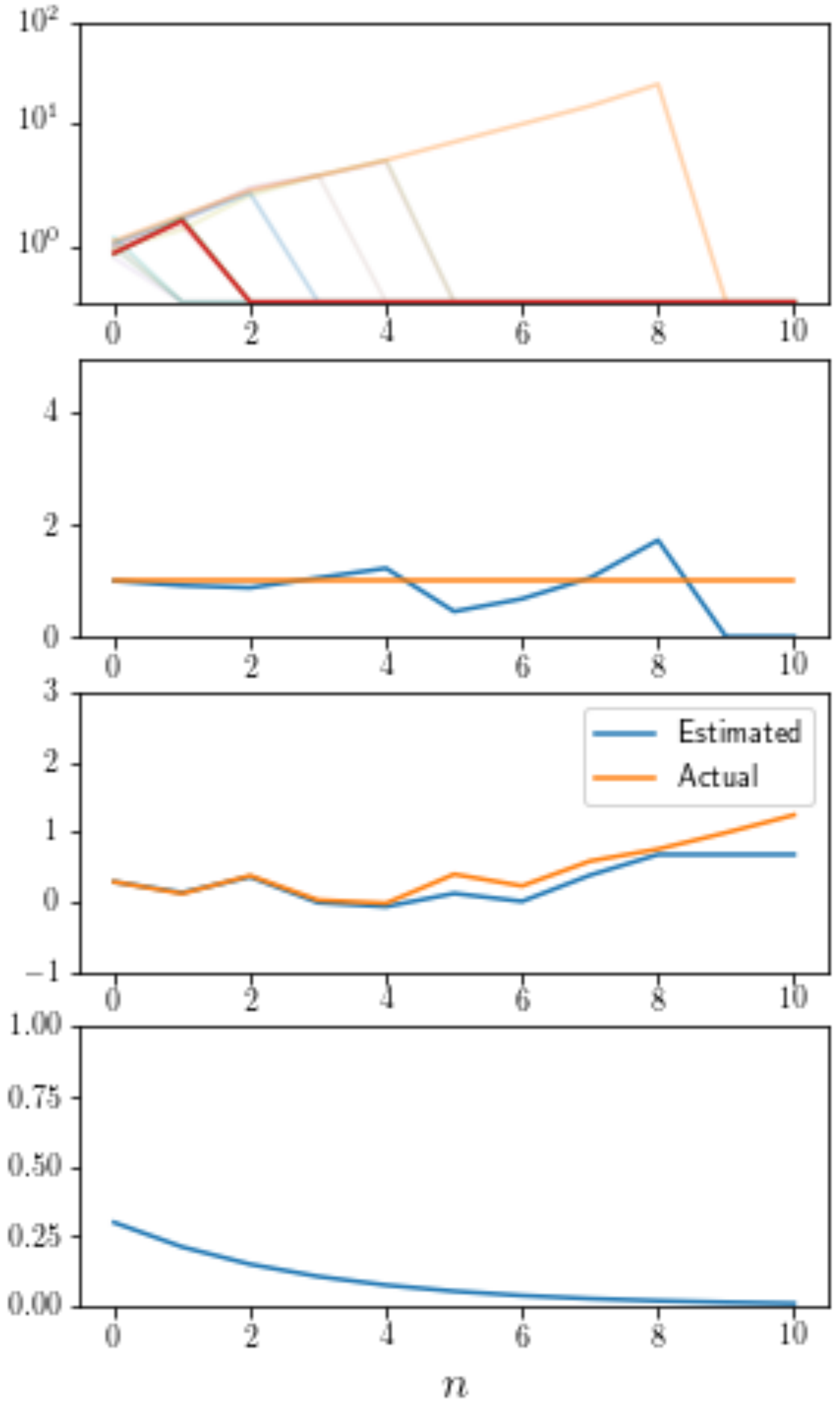
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



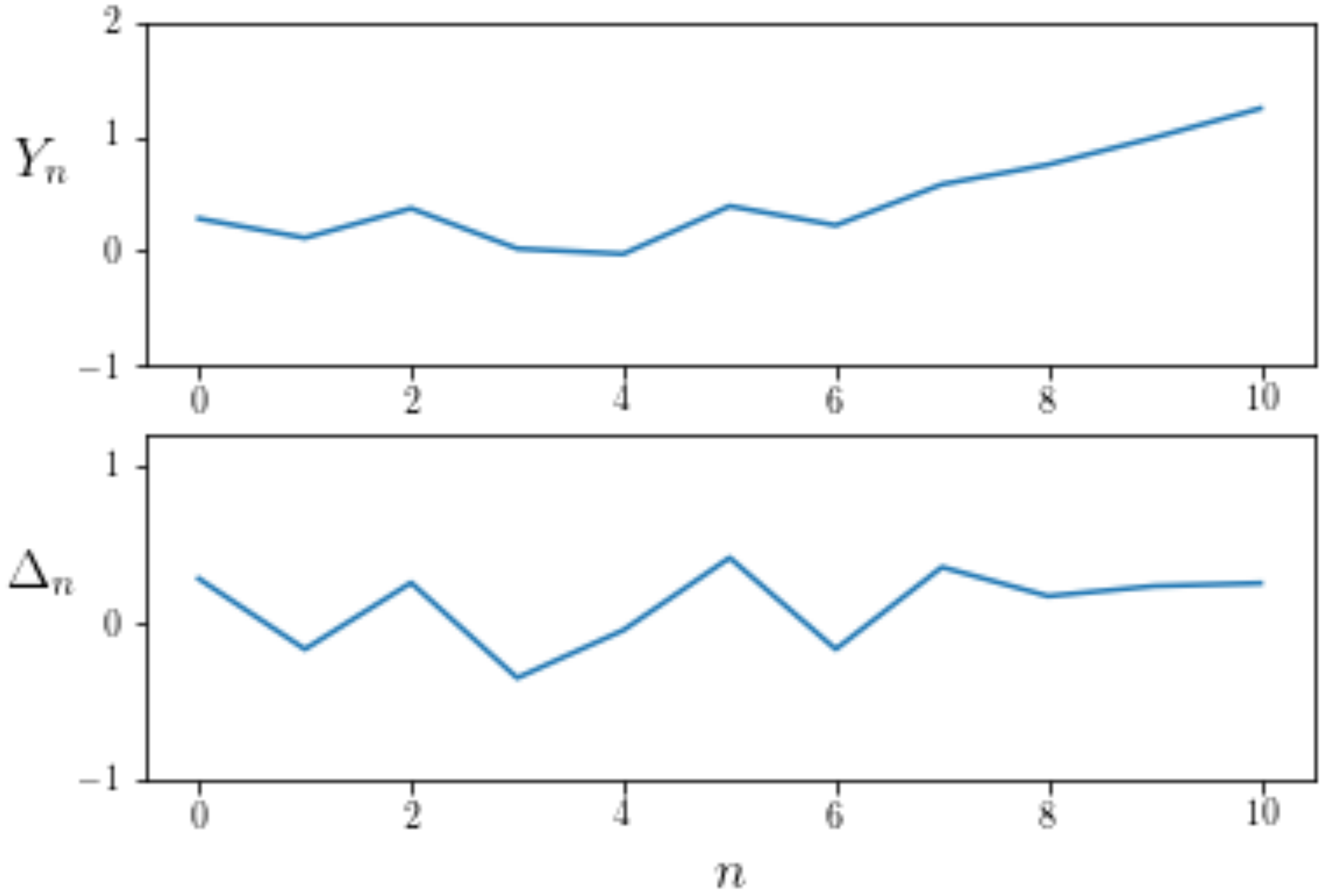
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

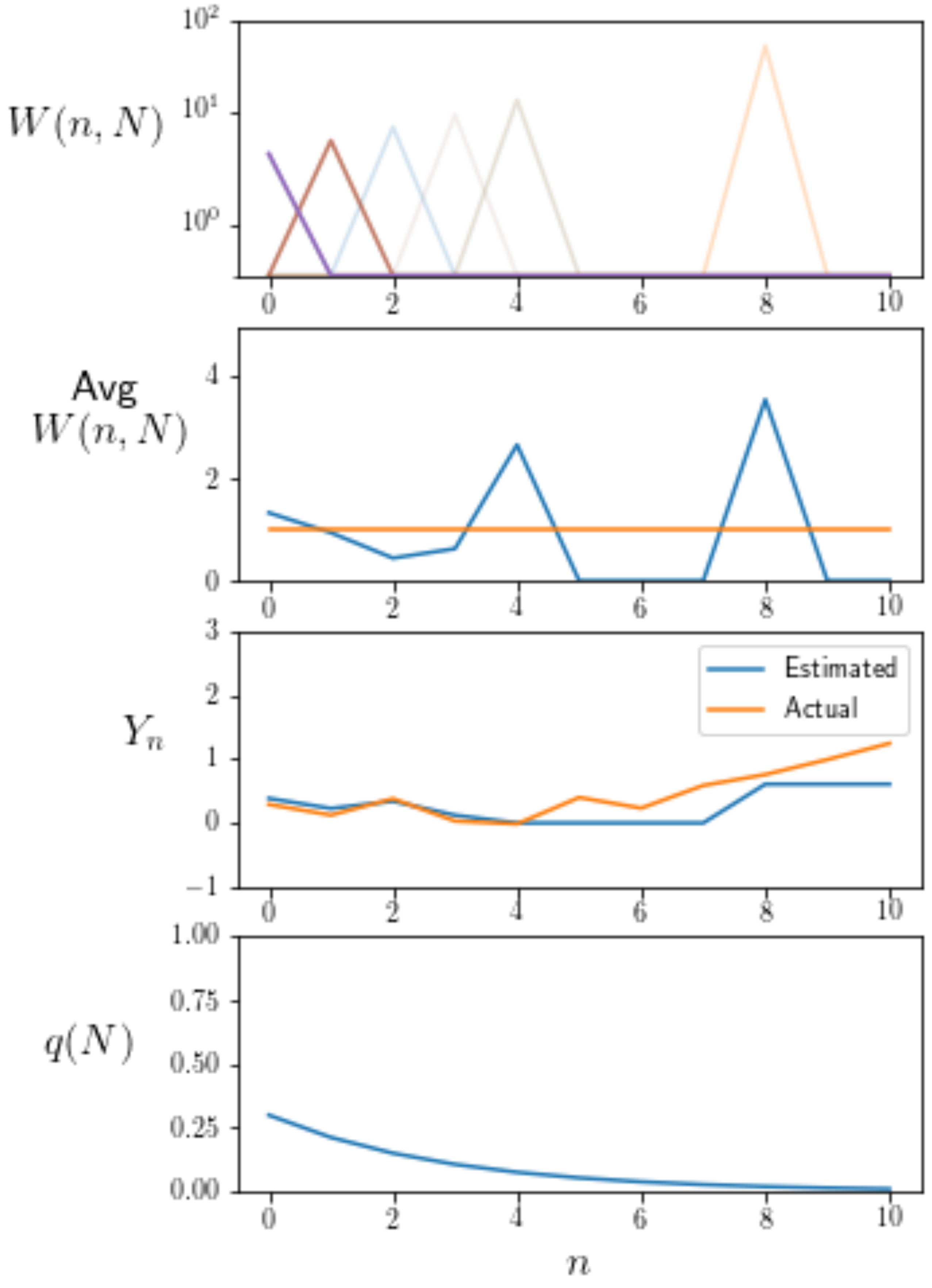
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



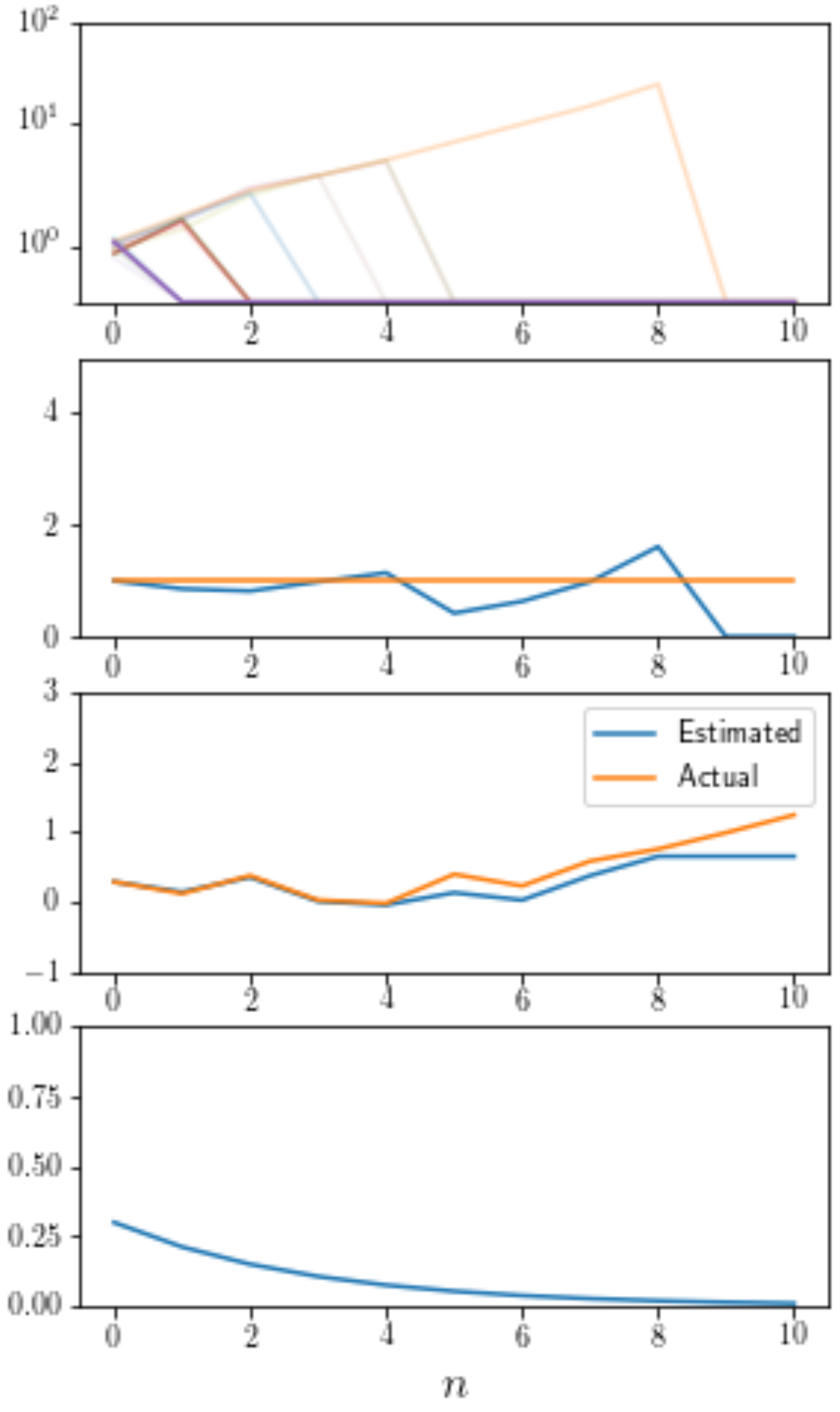
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



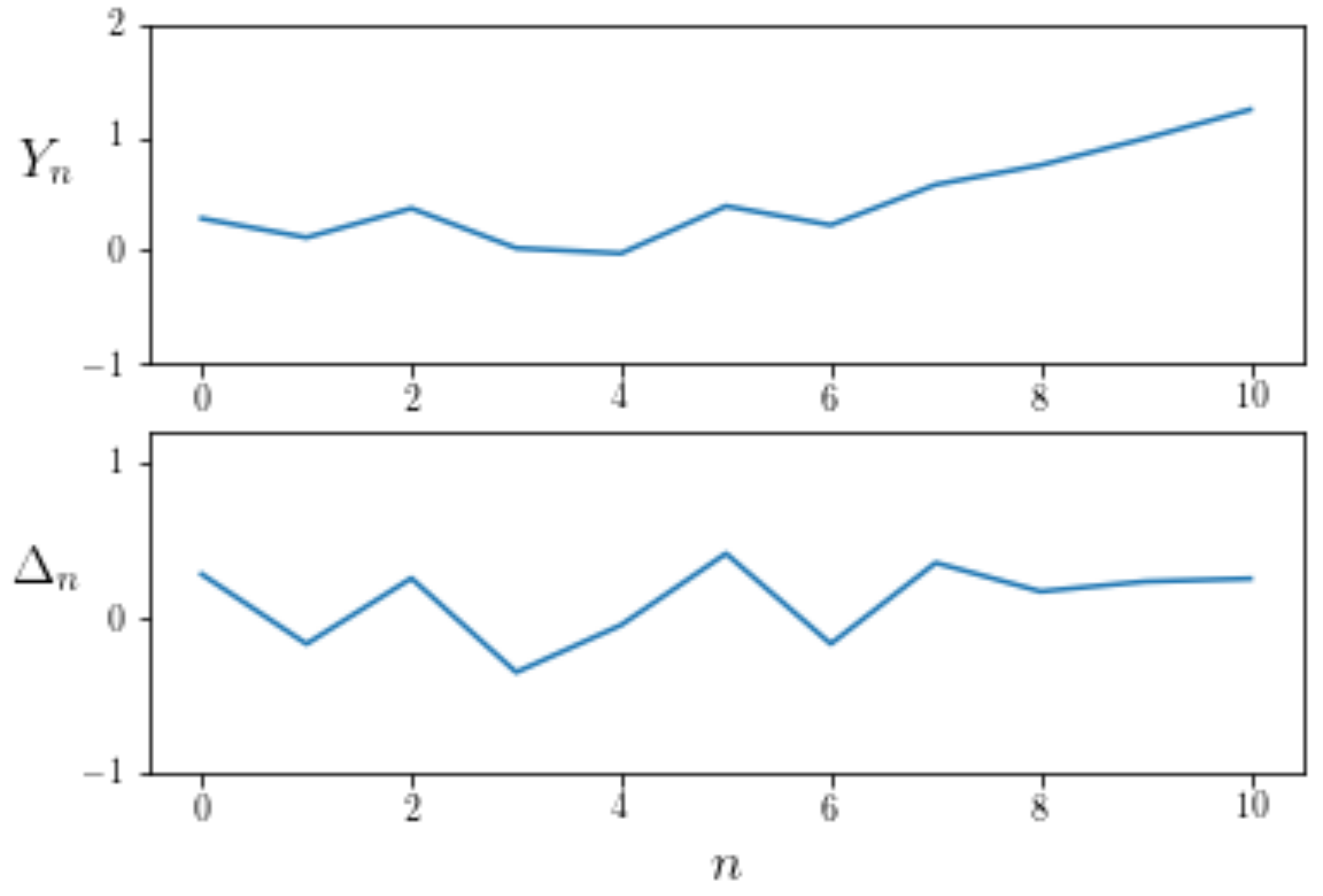
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

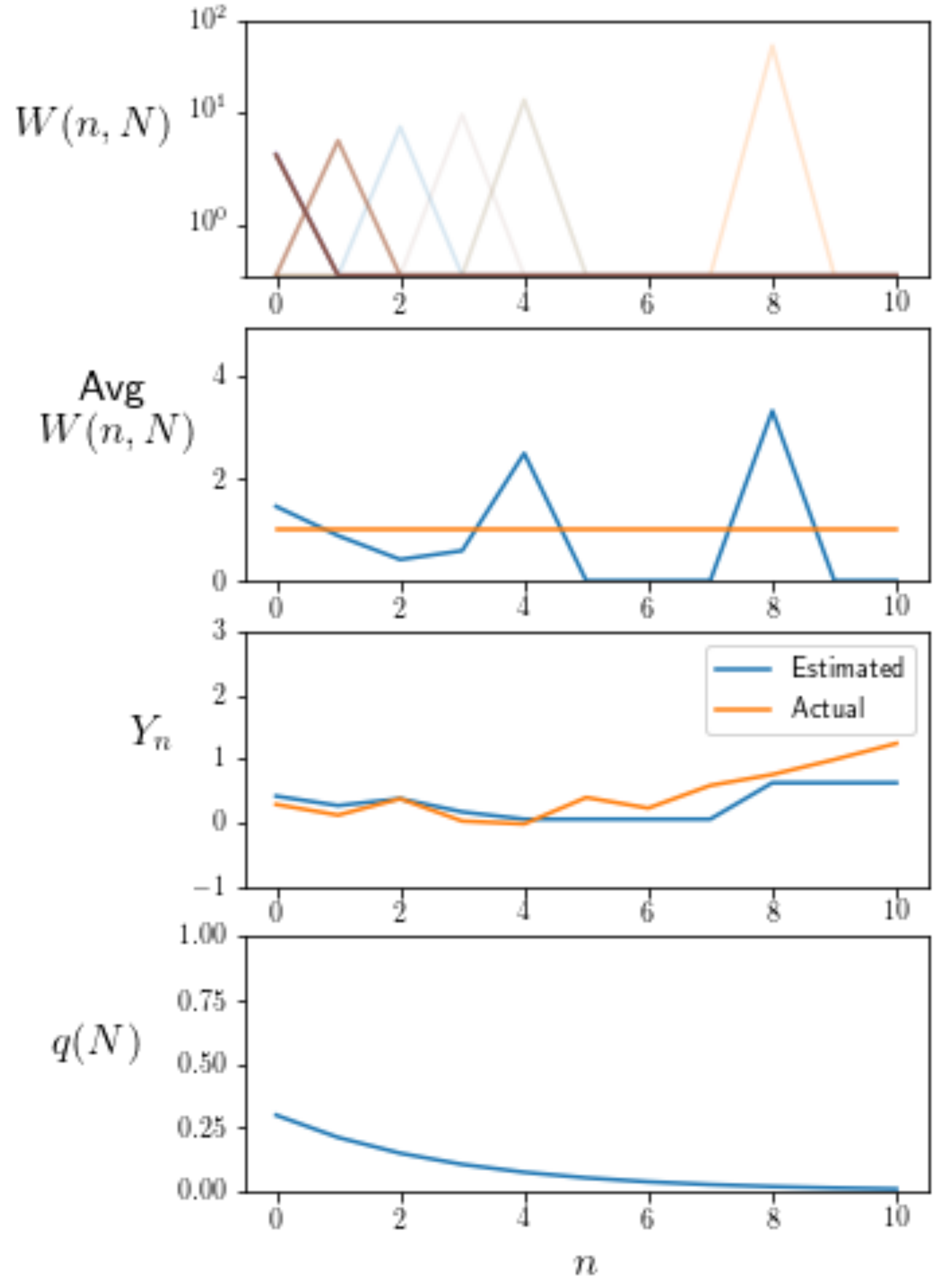
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



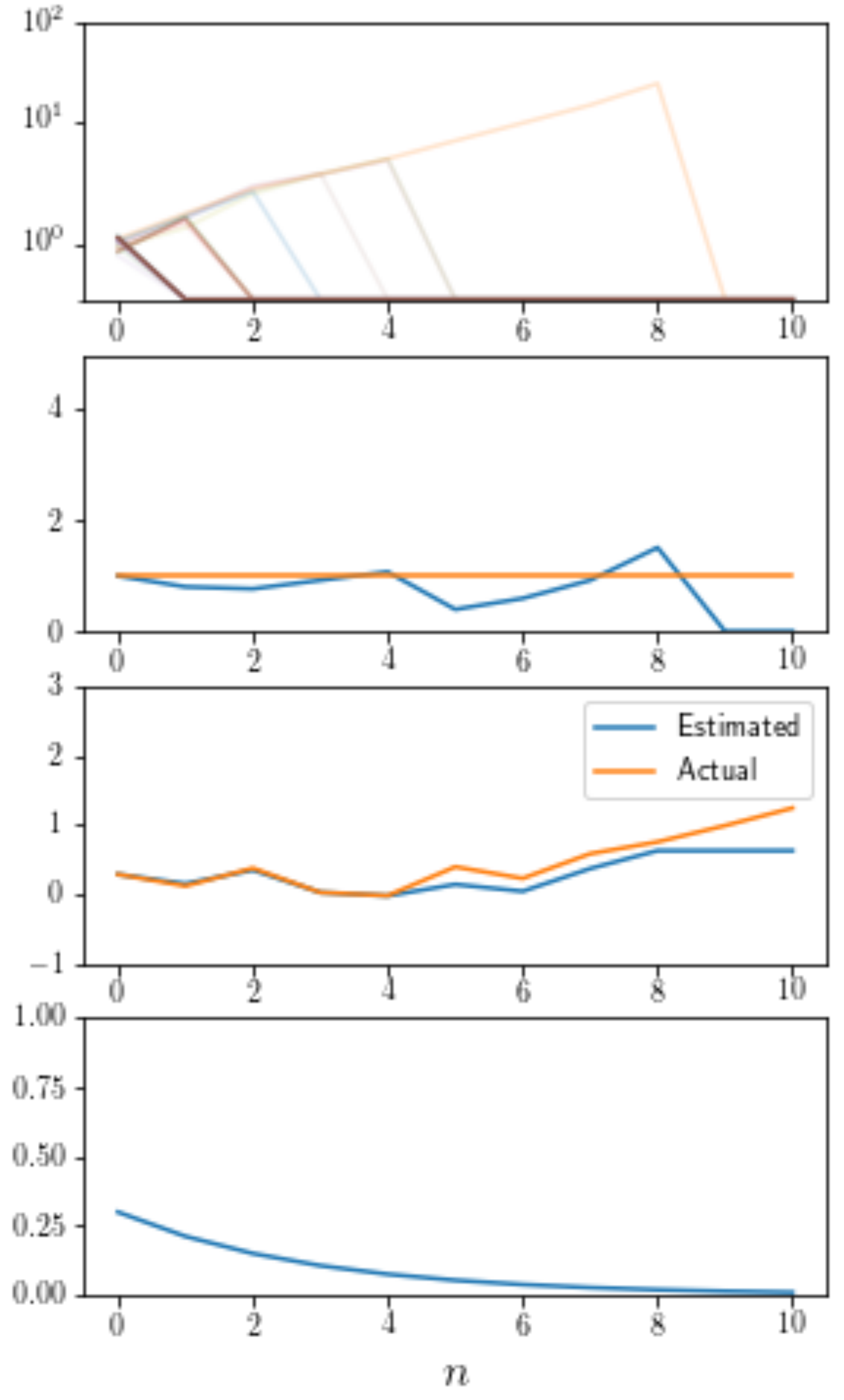
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



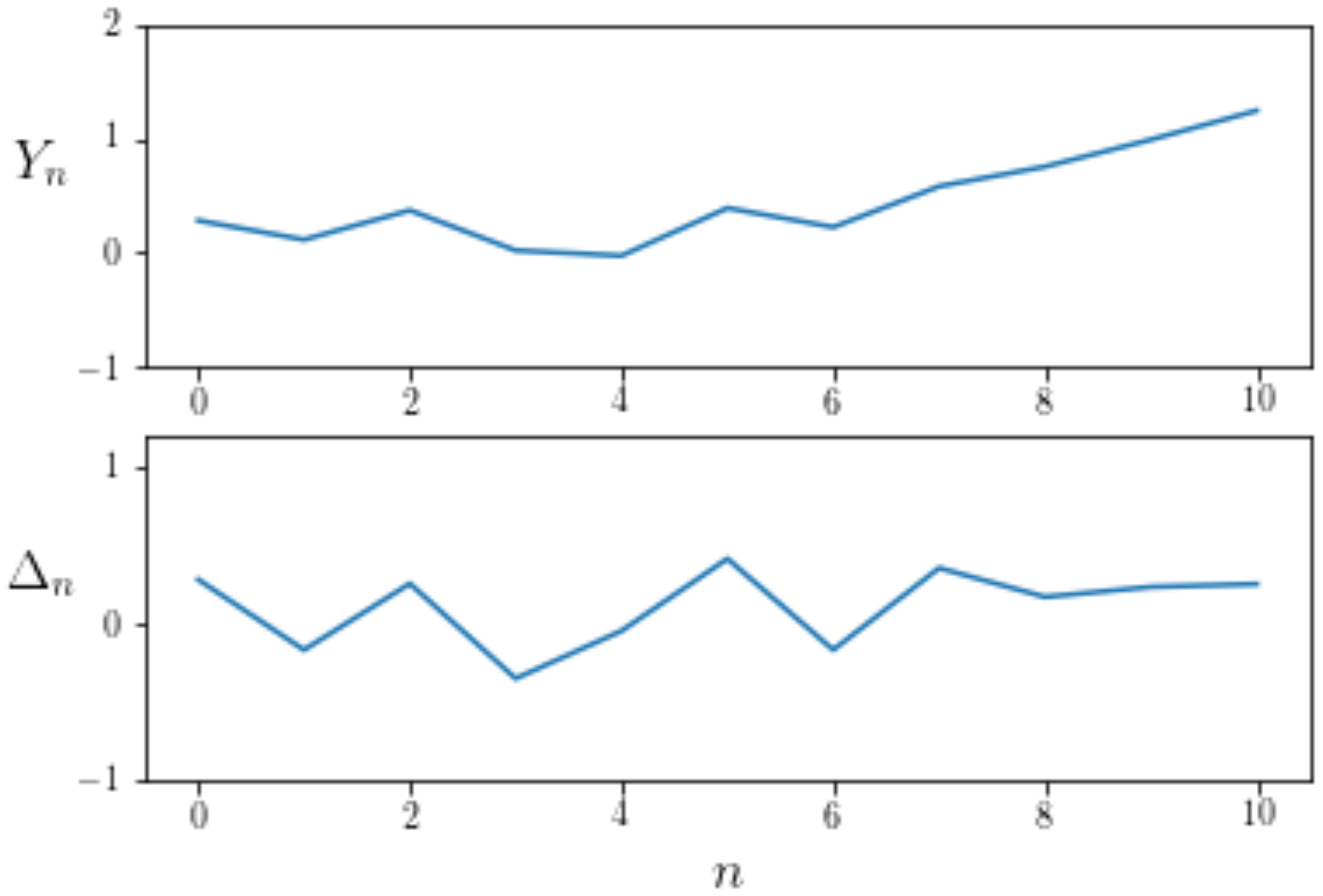
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

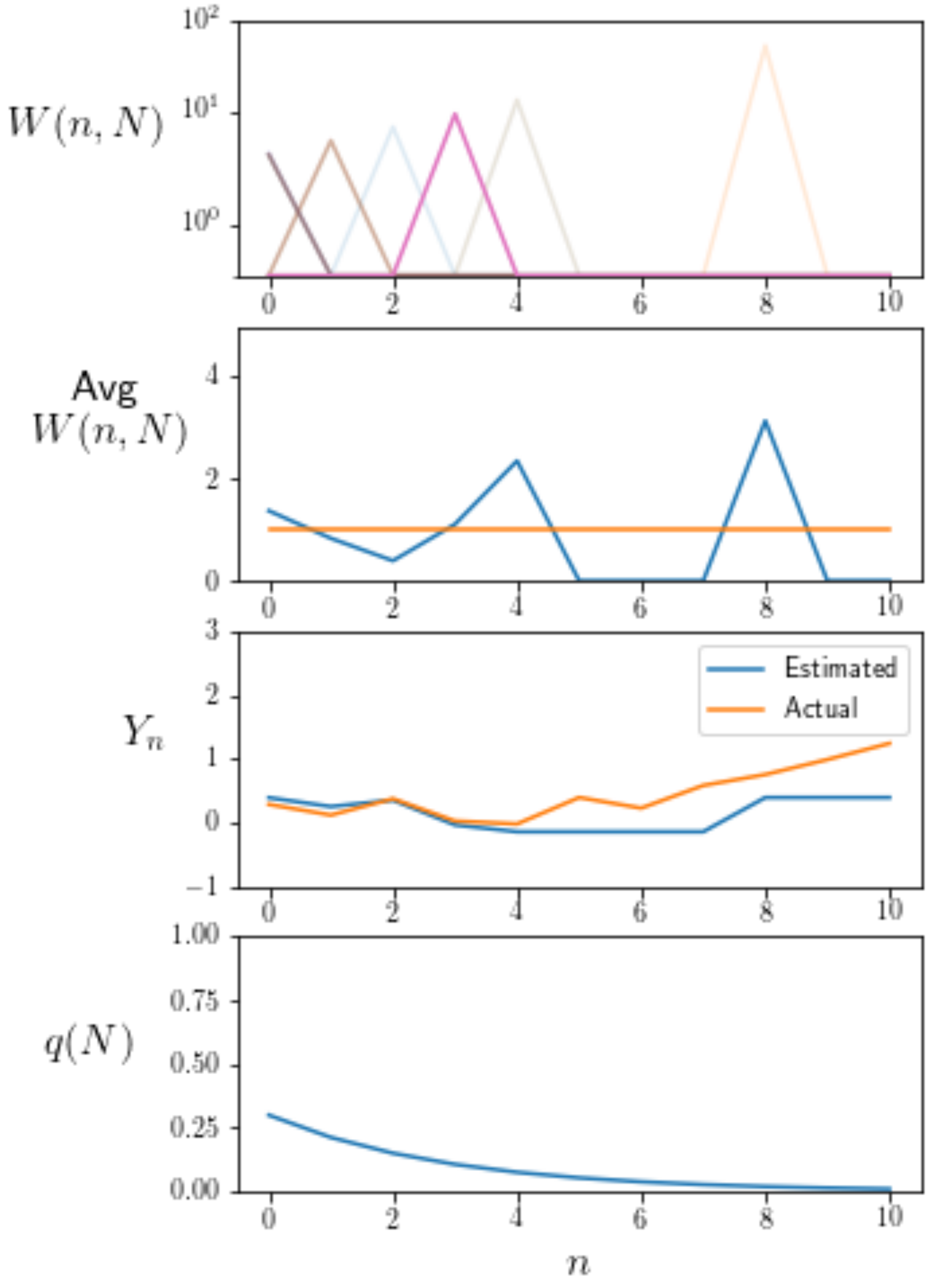
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



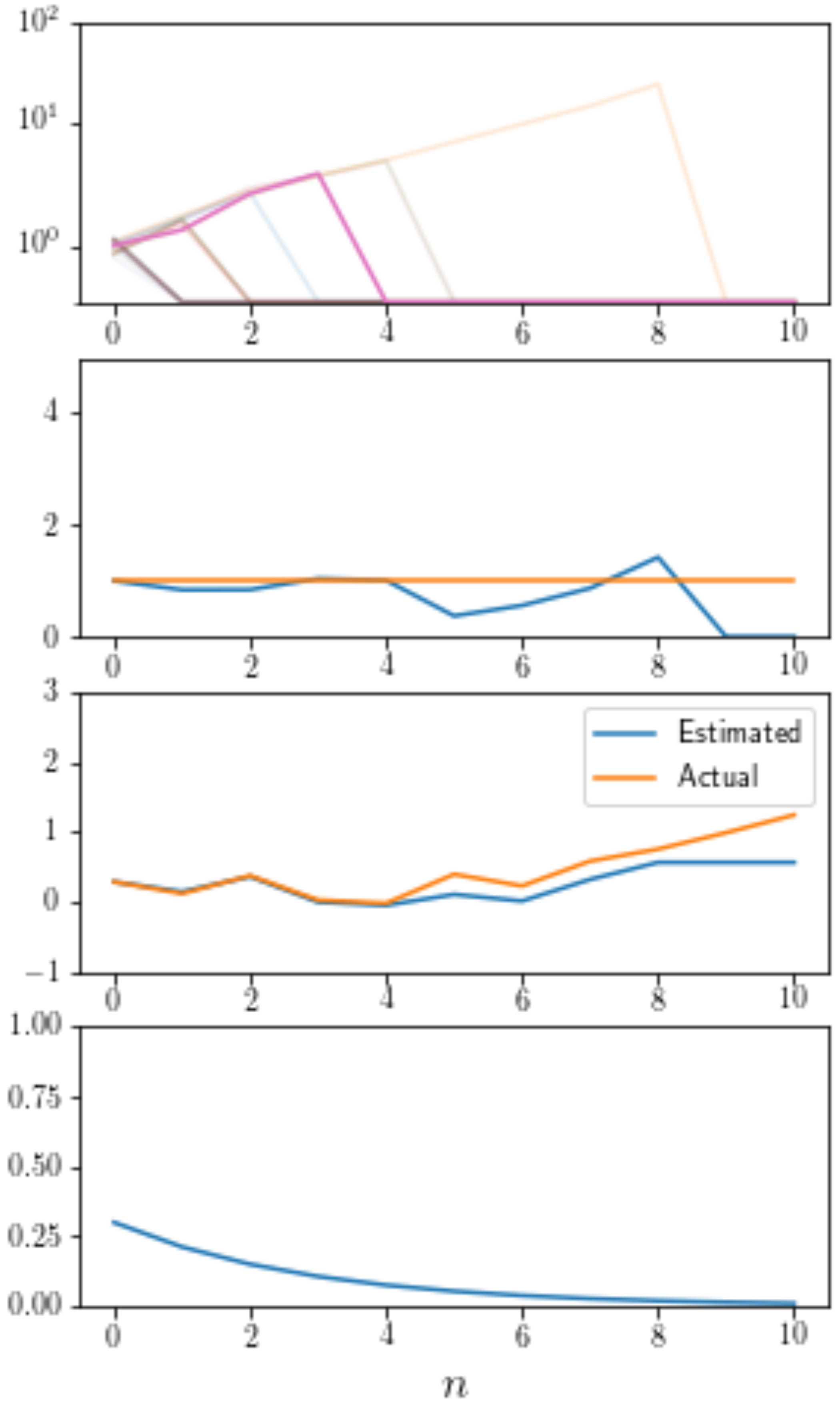
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



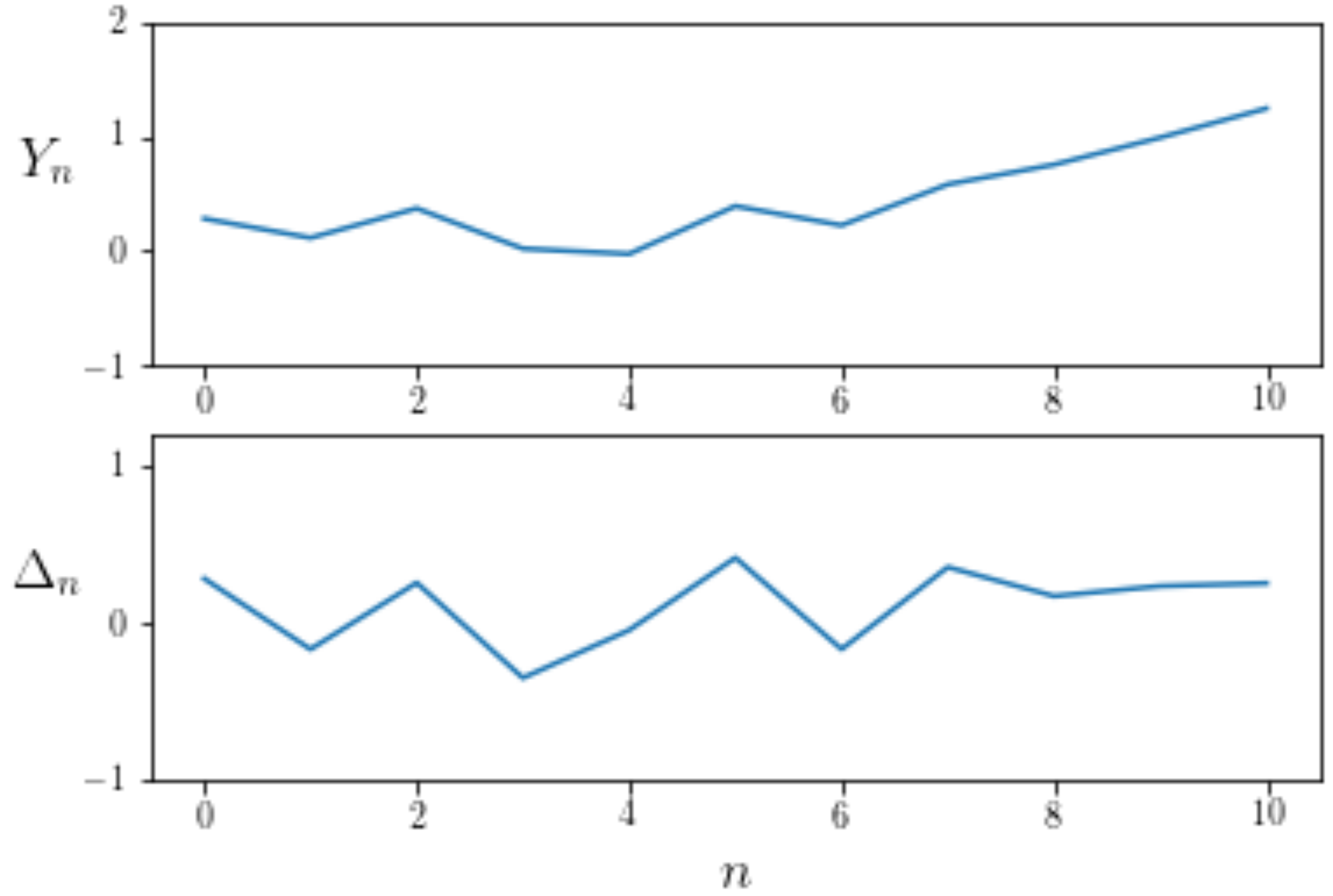
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

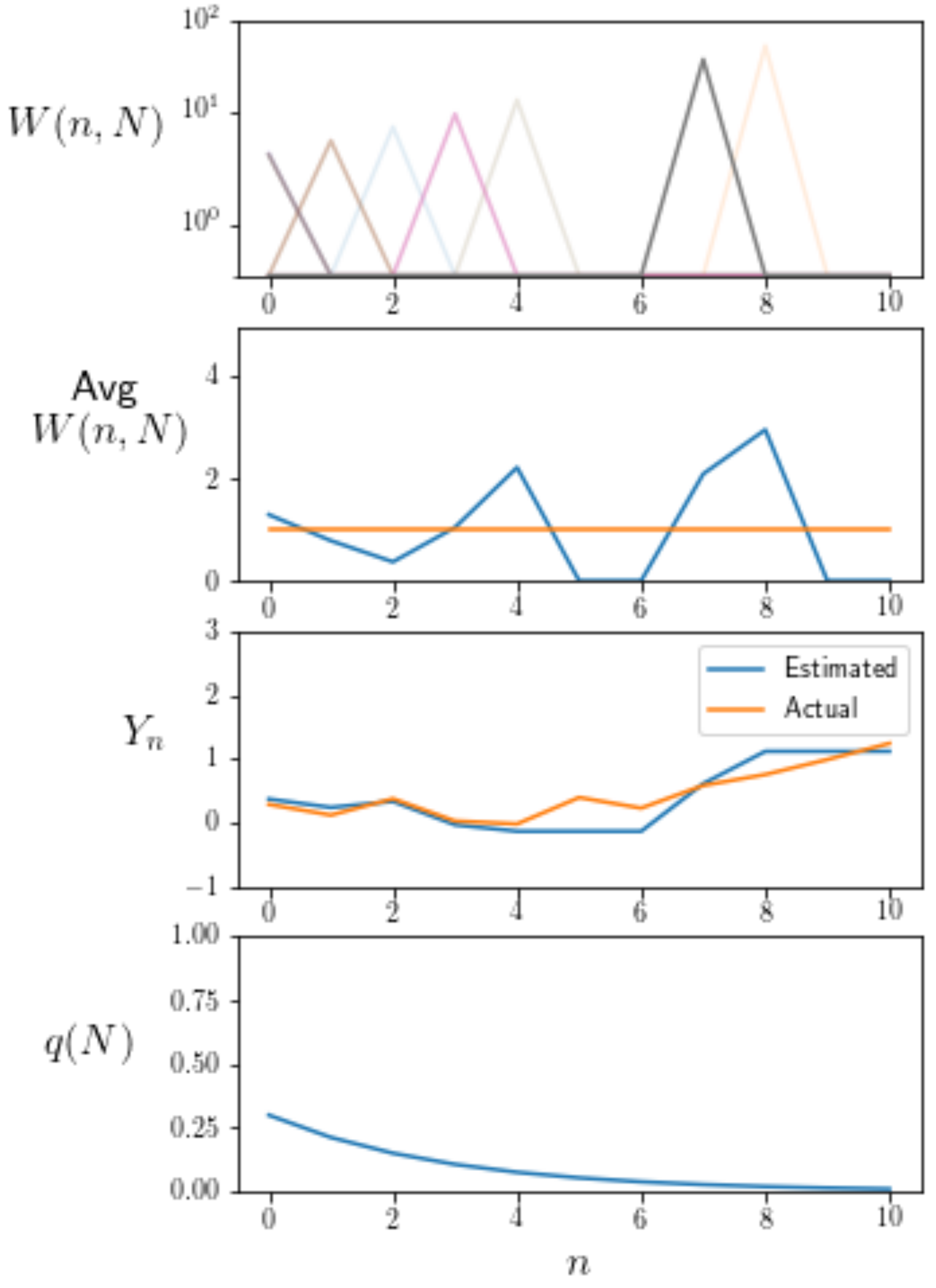
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



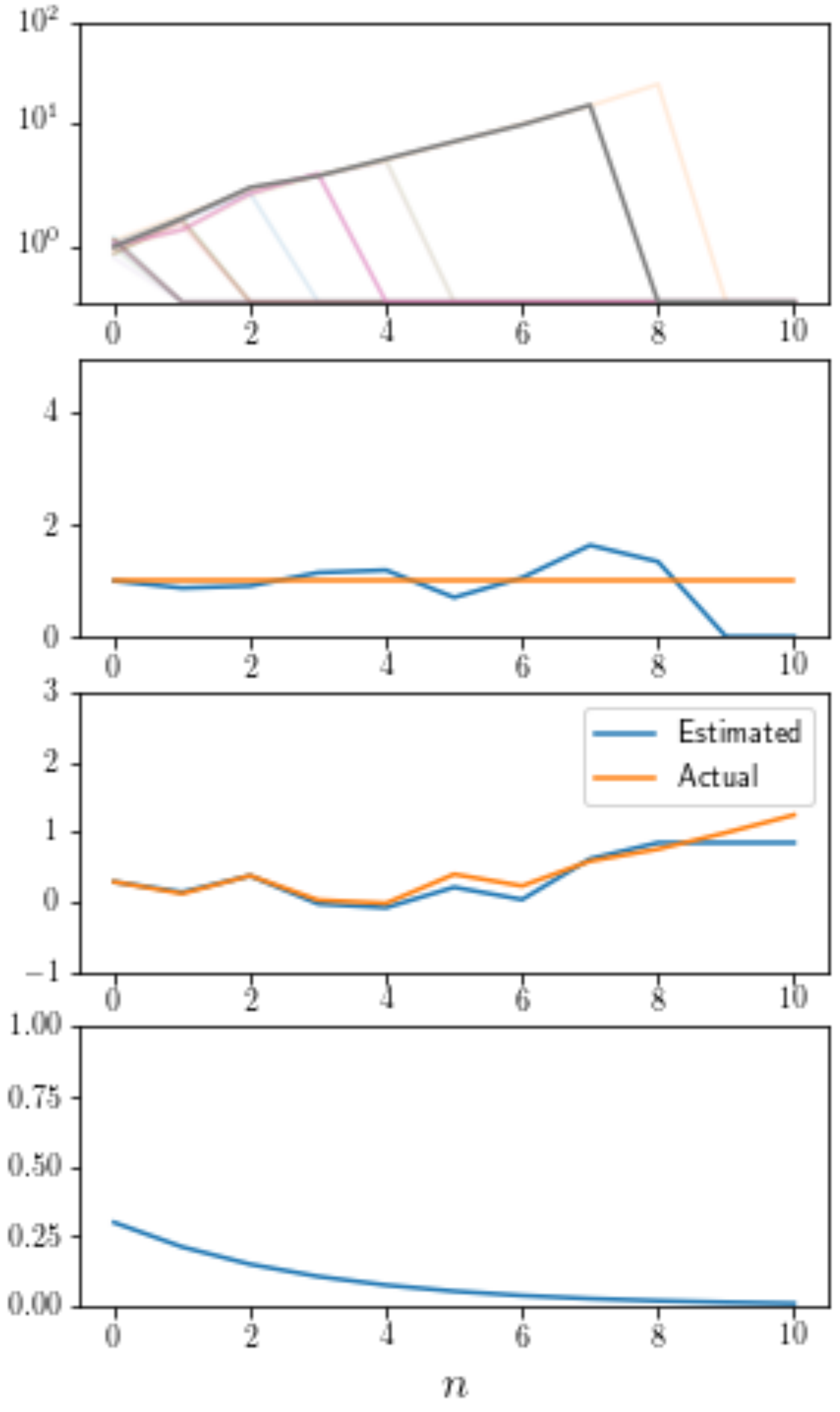
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



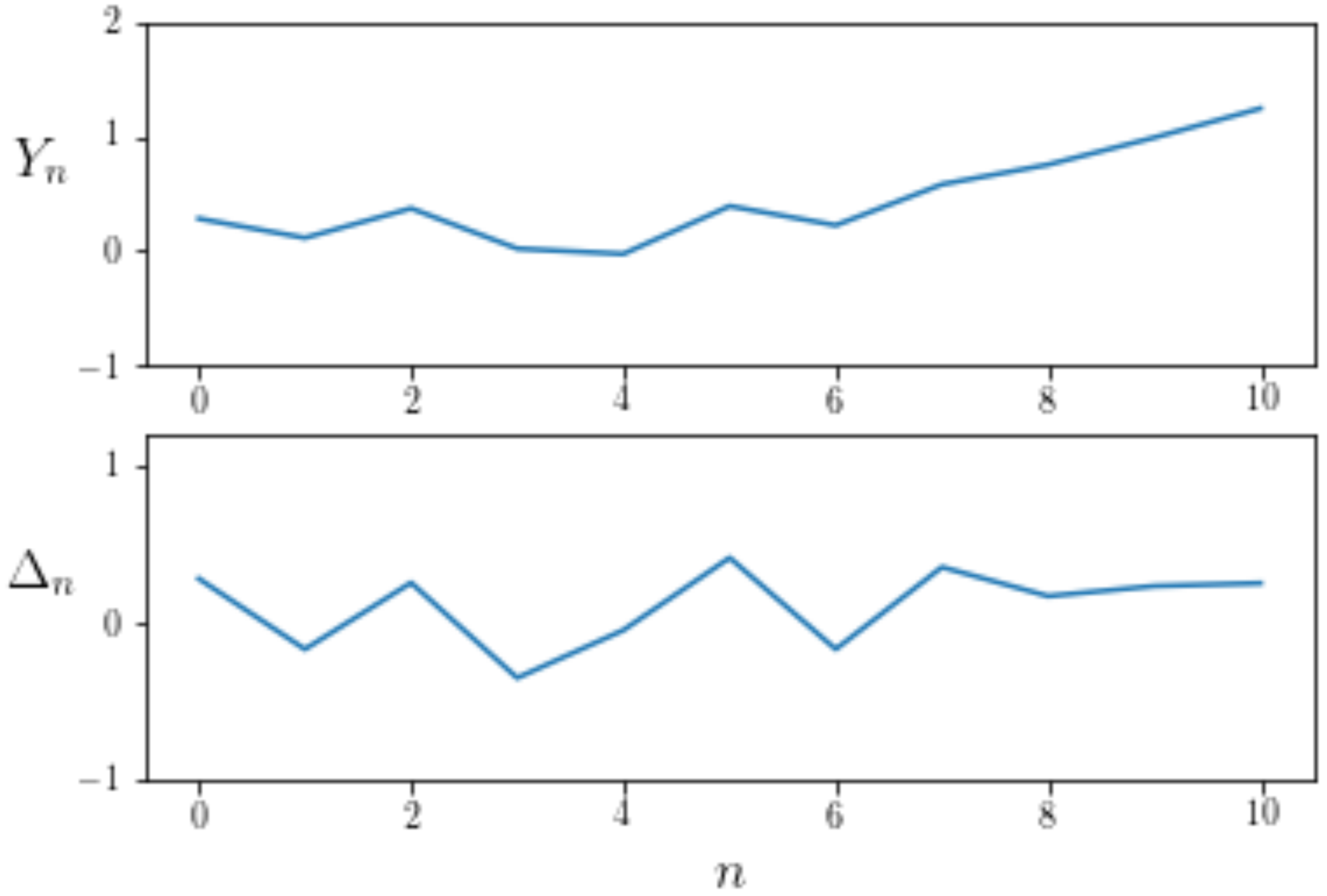
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

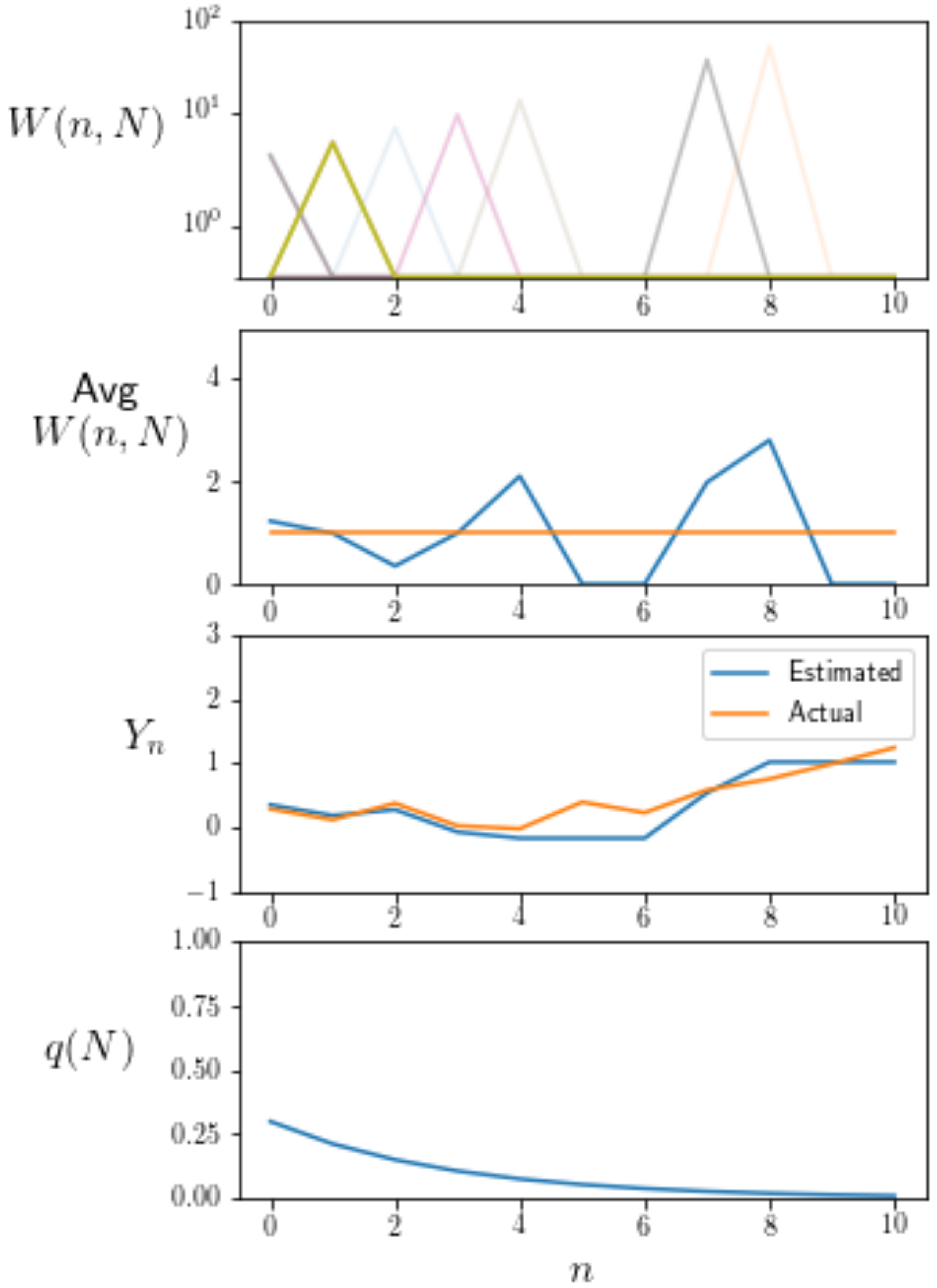
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



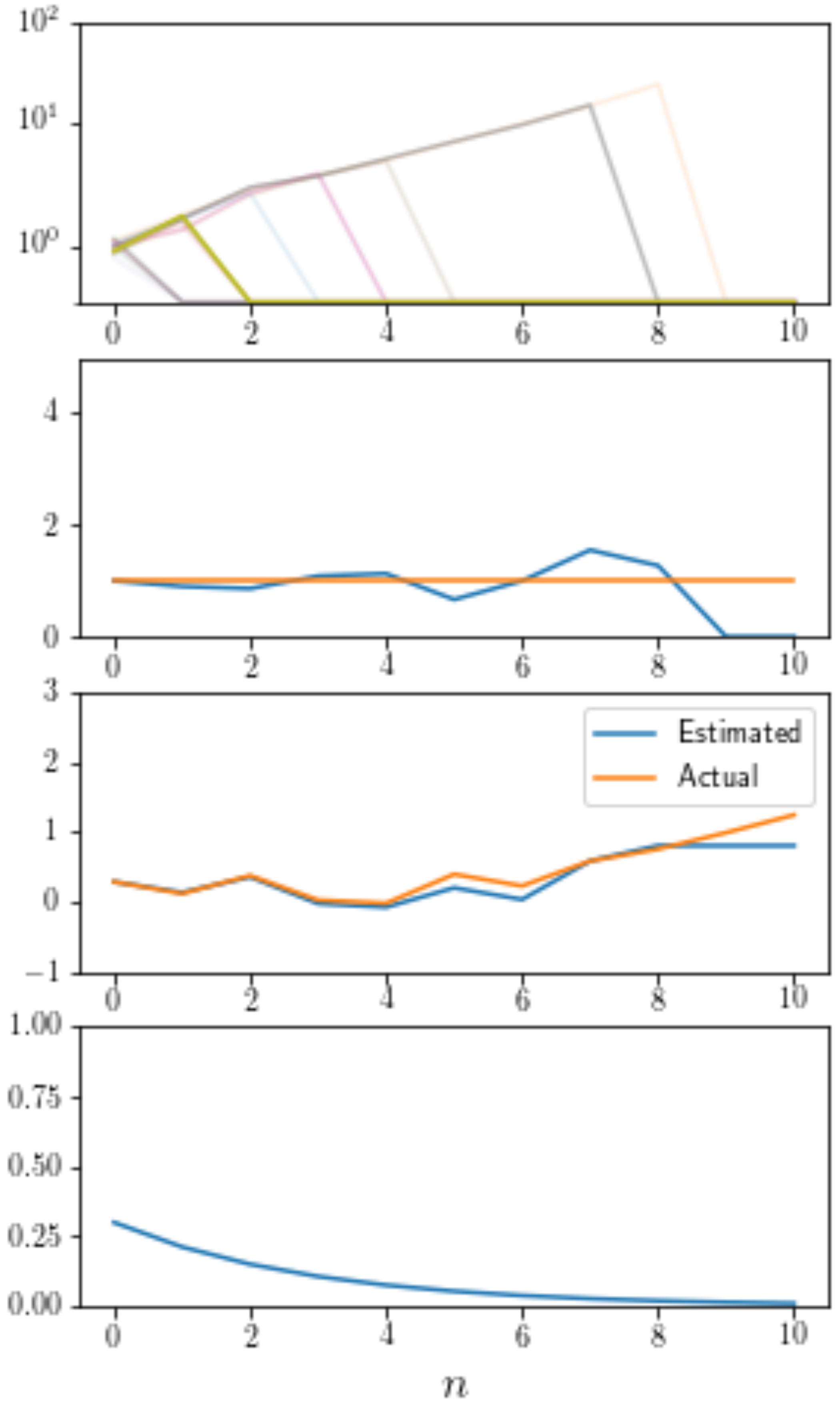
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



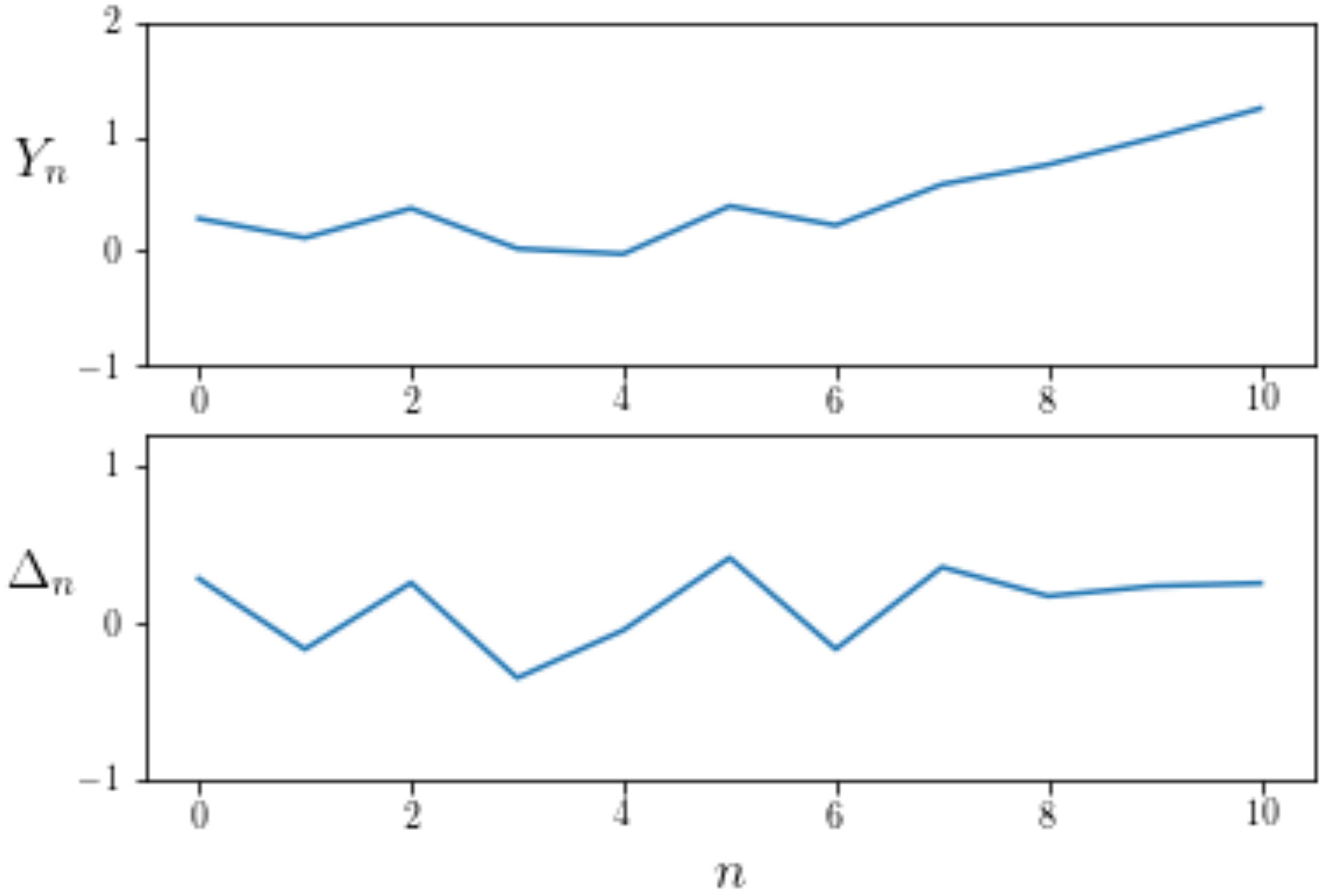
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n,N) \qquad N \in \{1, \ldots, H\} \sim q$$

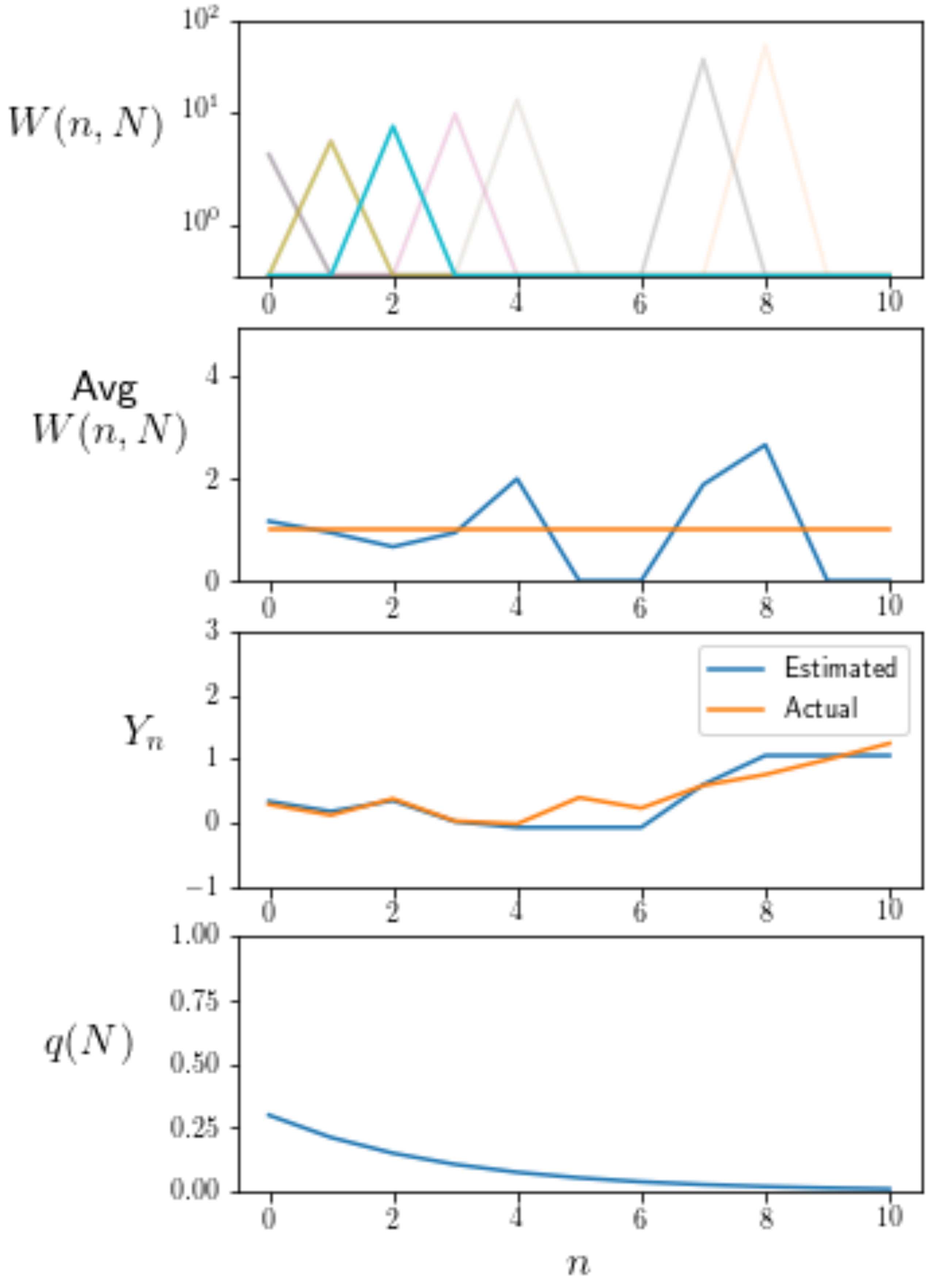
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



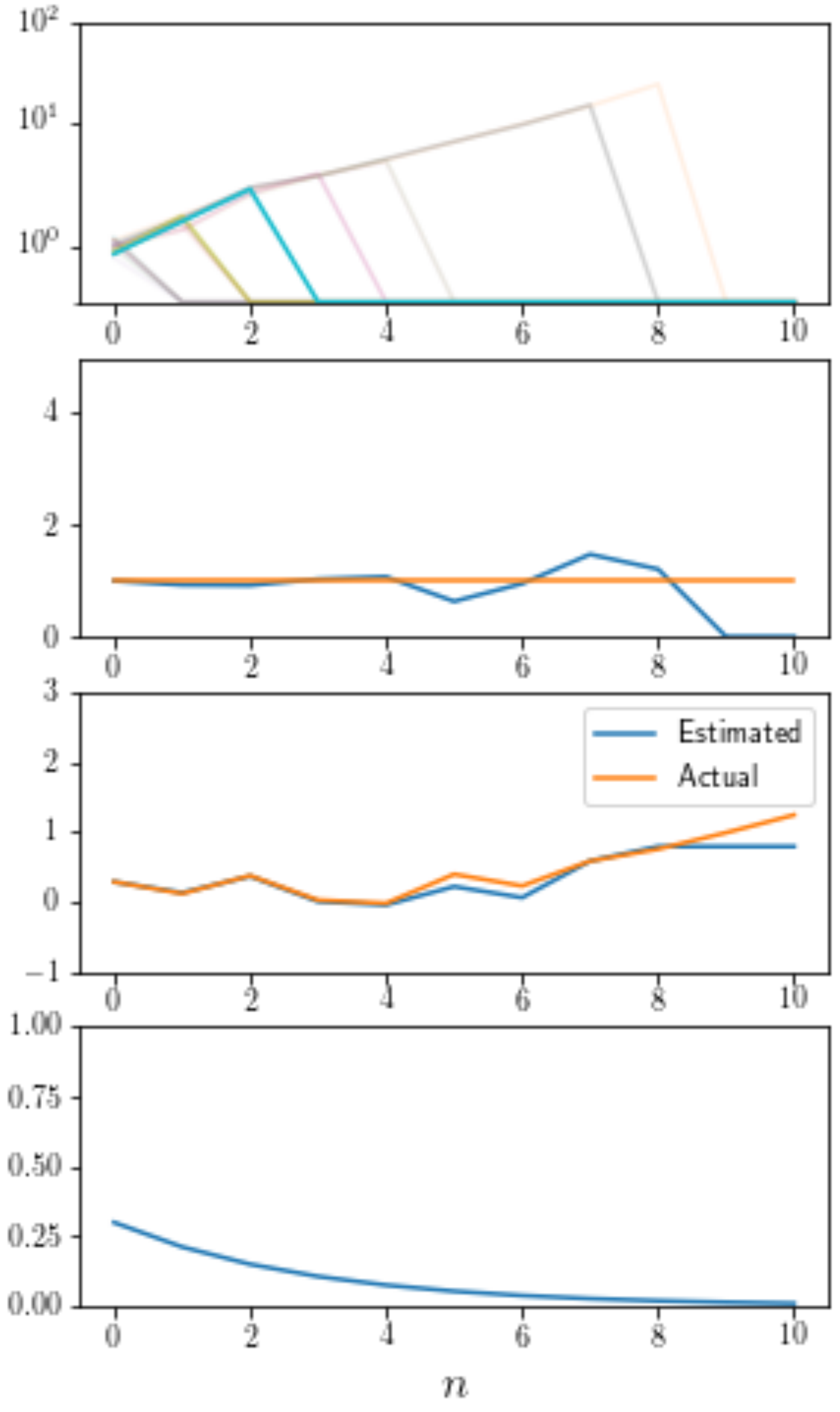
“Single sample”

$$W(n,N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



“Russian roulette”

$$W(n,N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



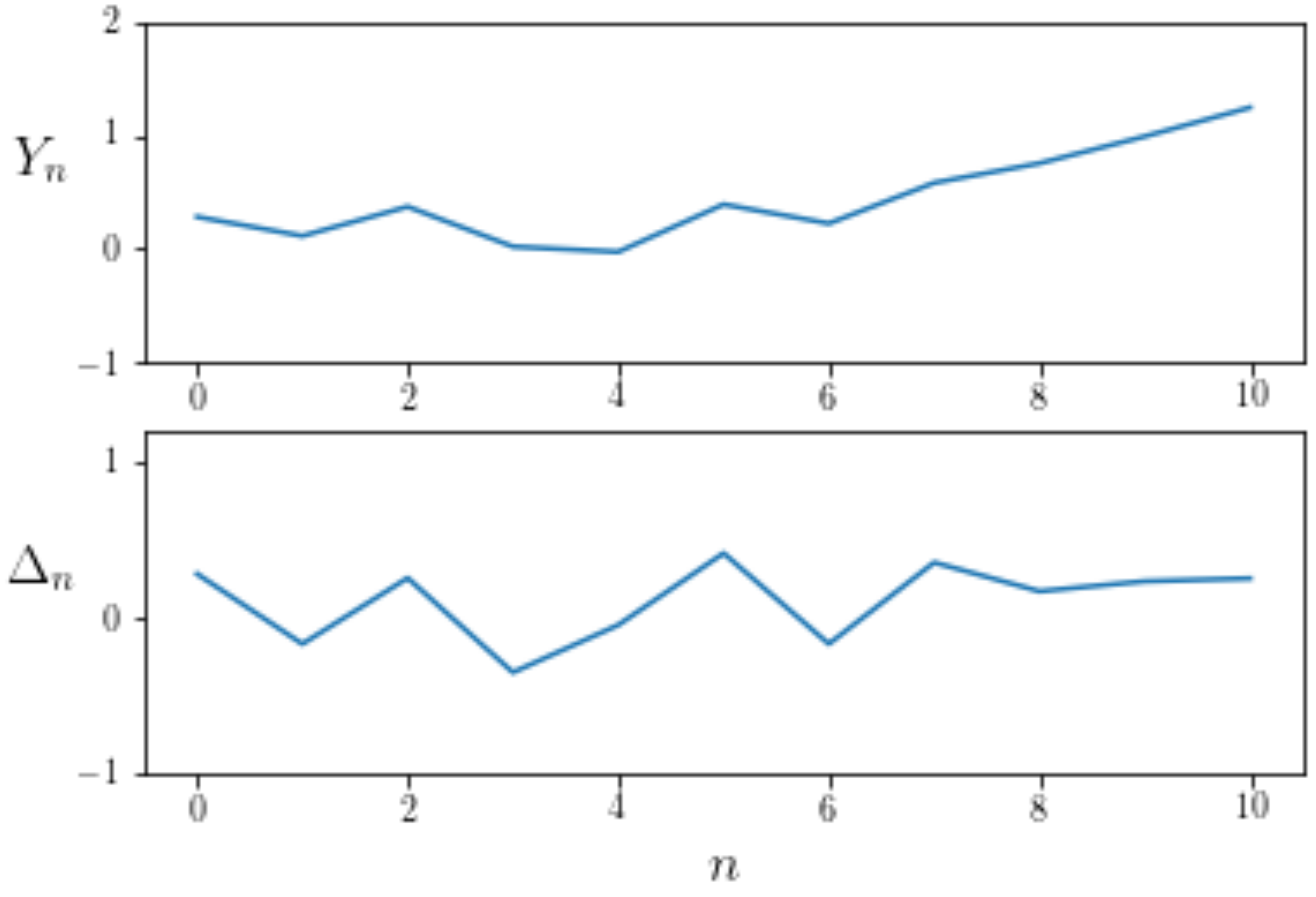
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

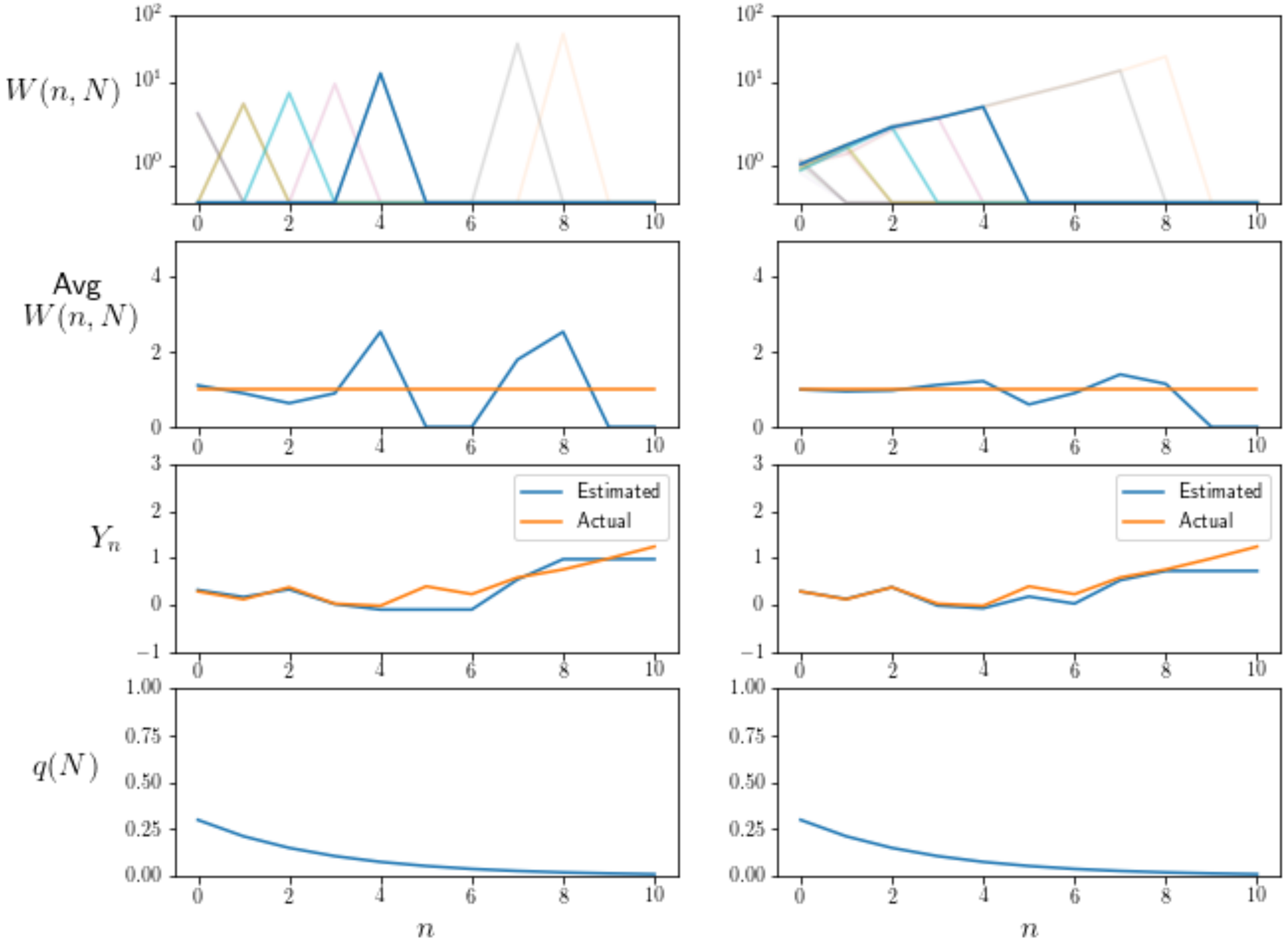


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



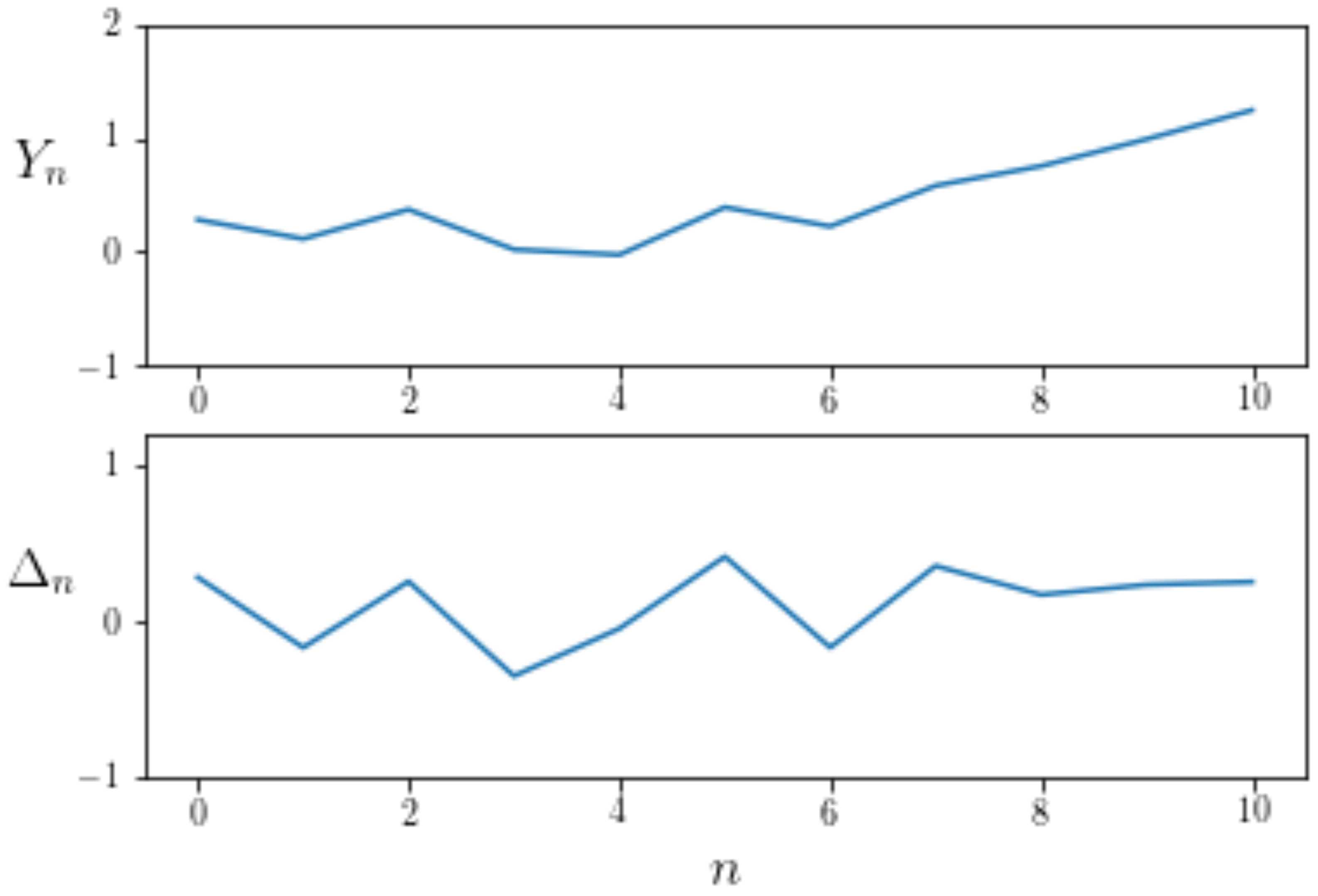
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

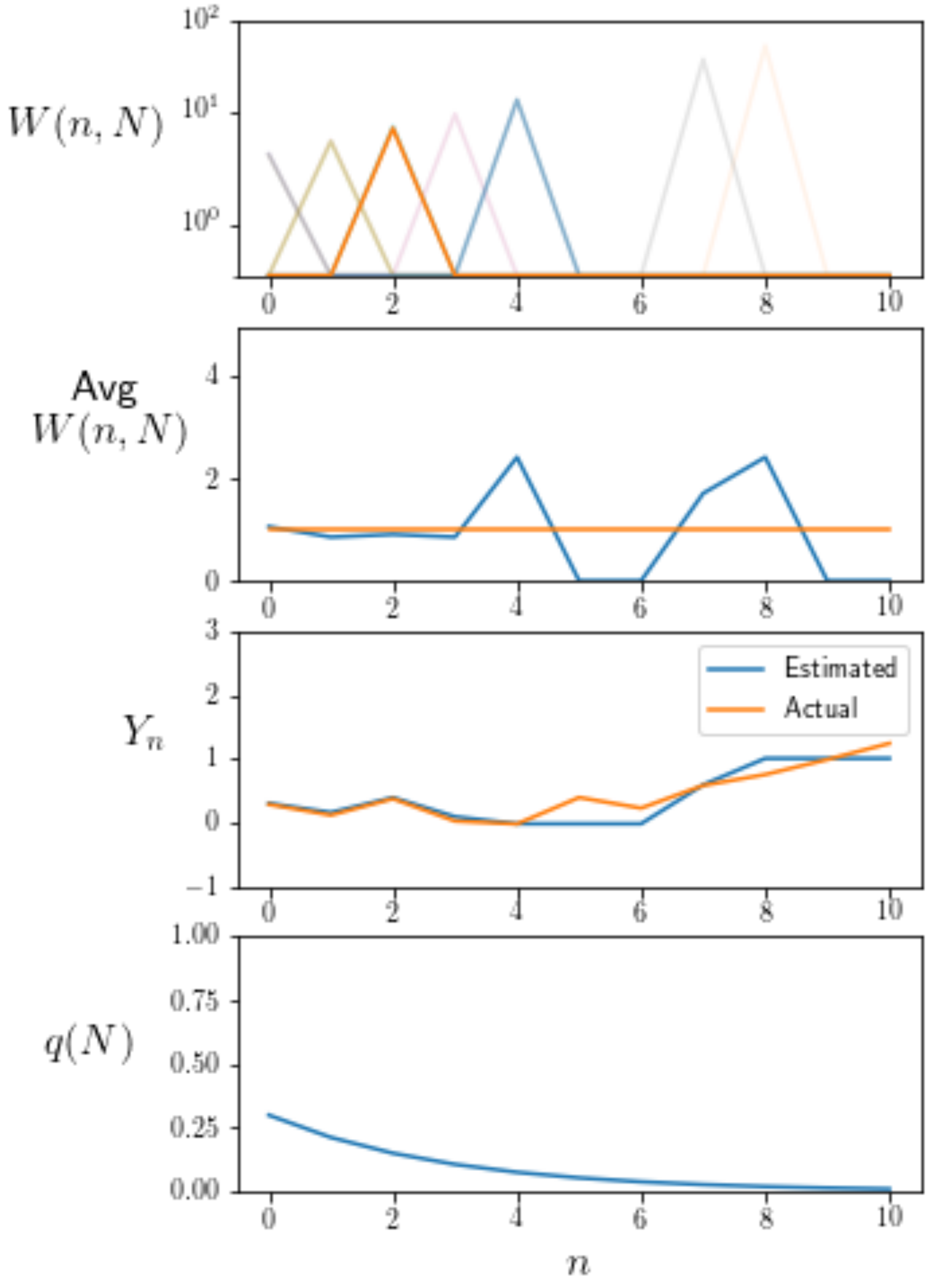
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



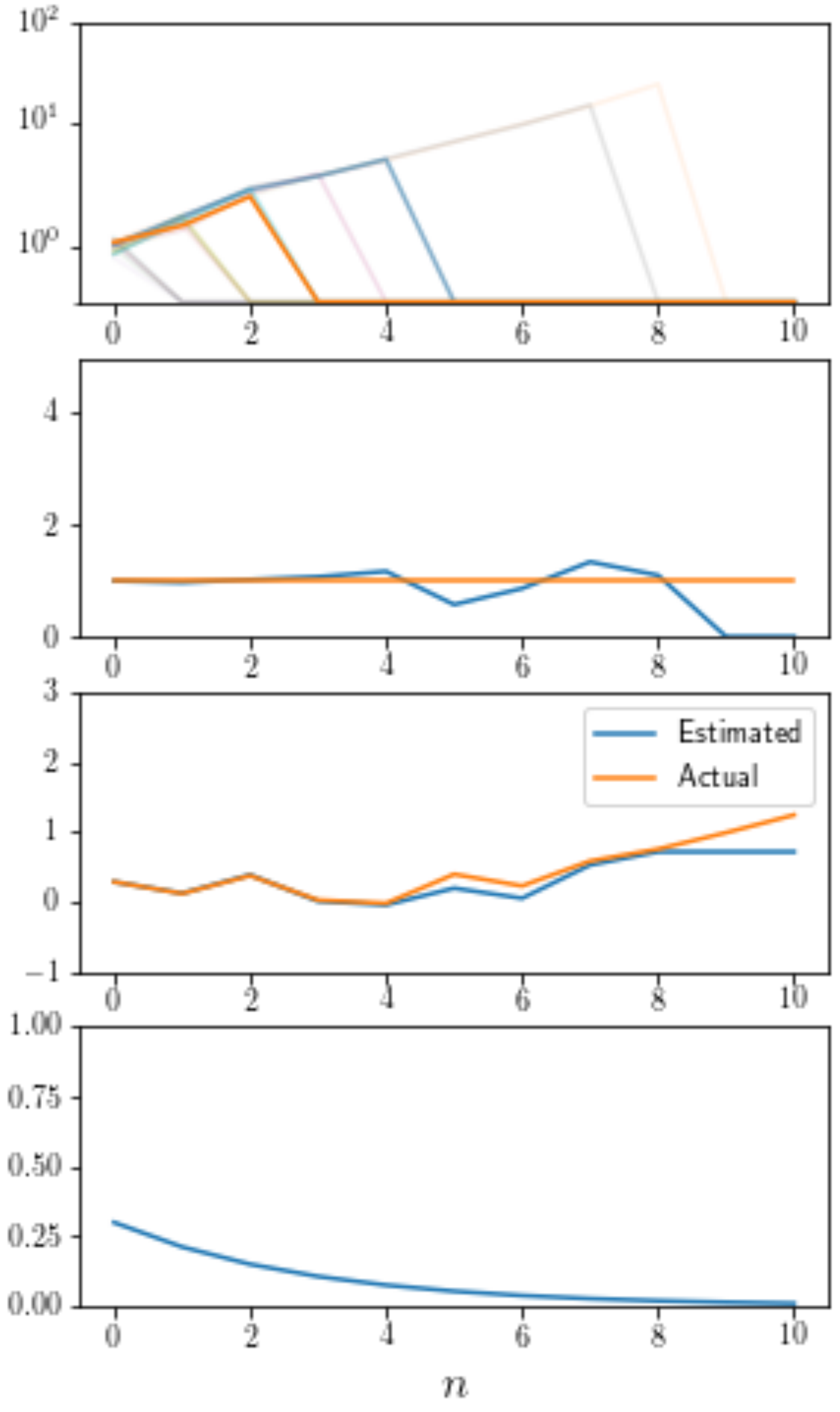
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



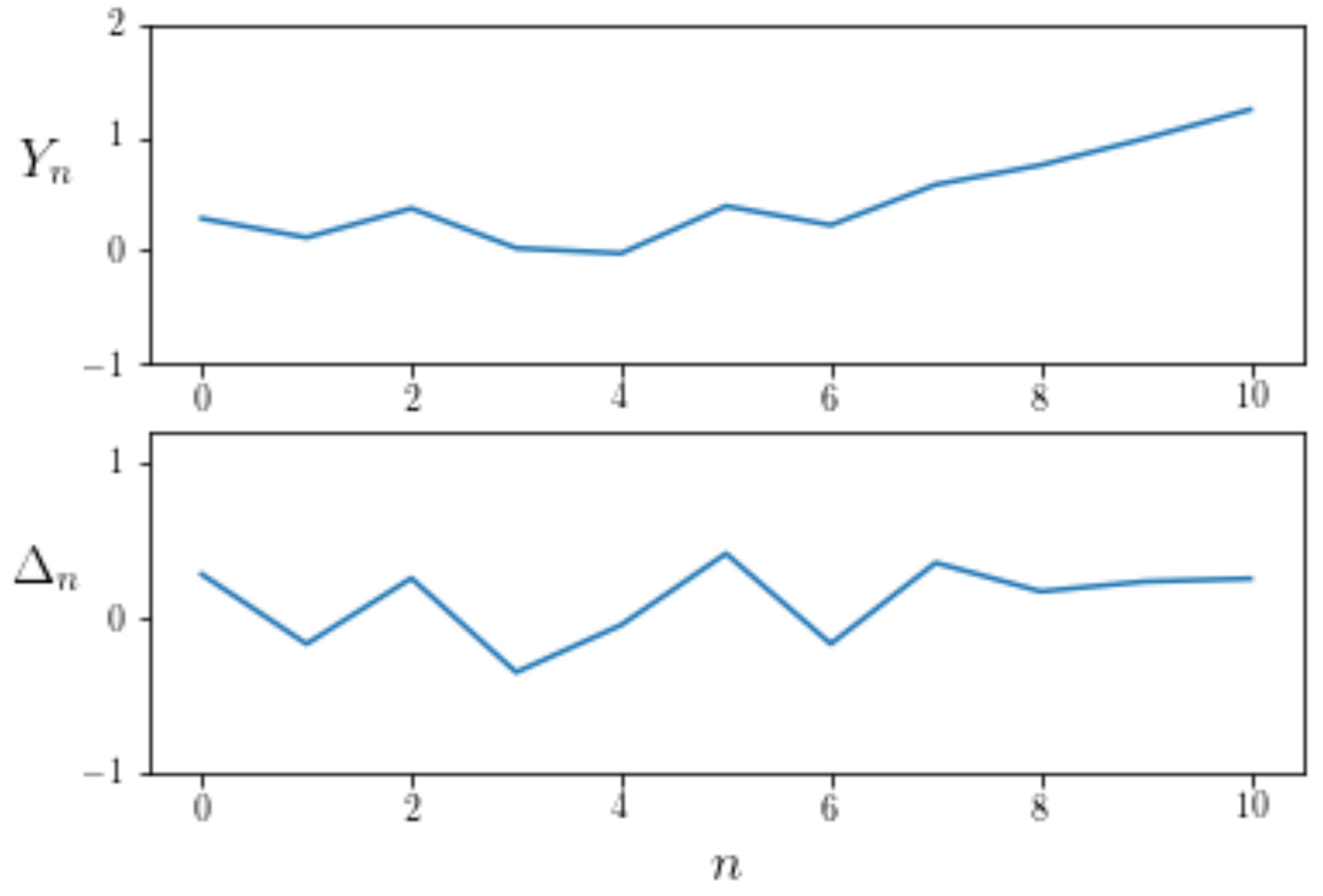
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

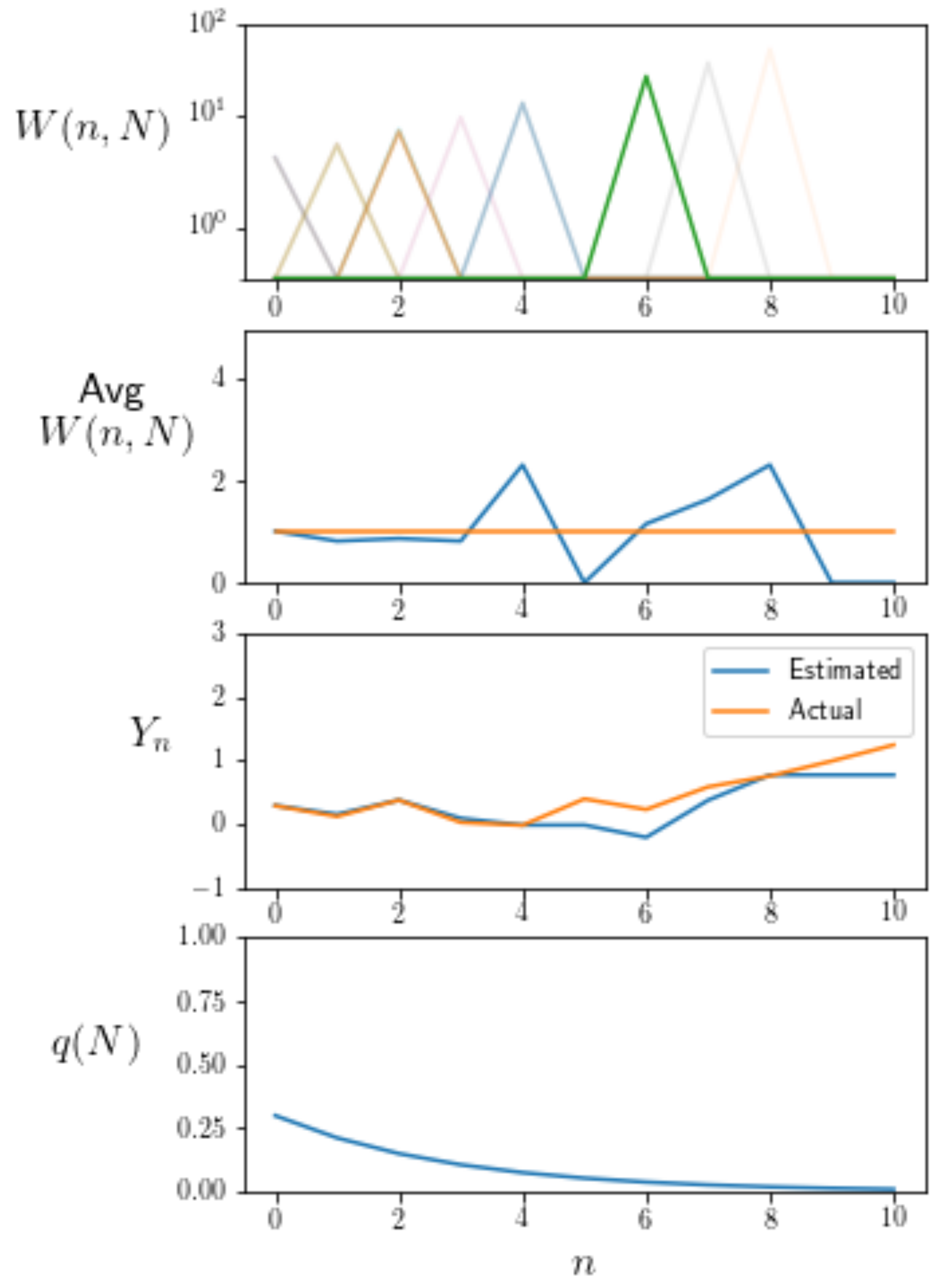
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



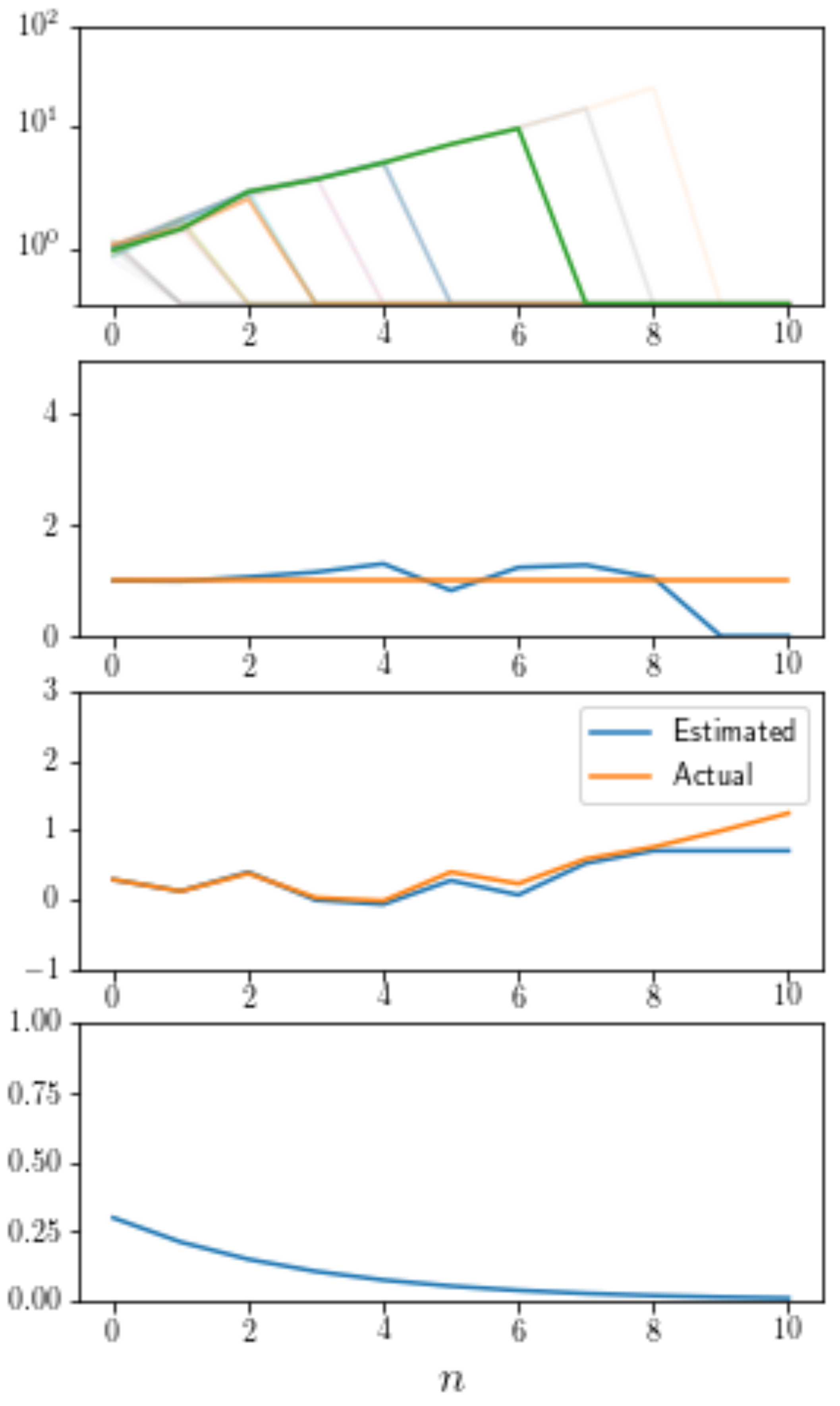
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



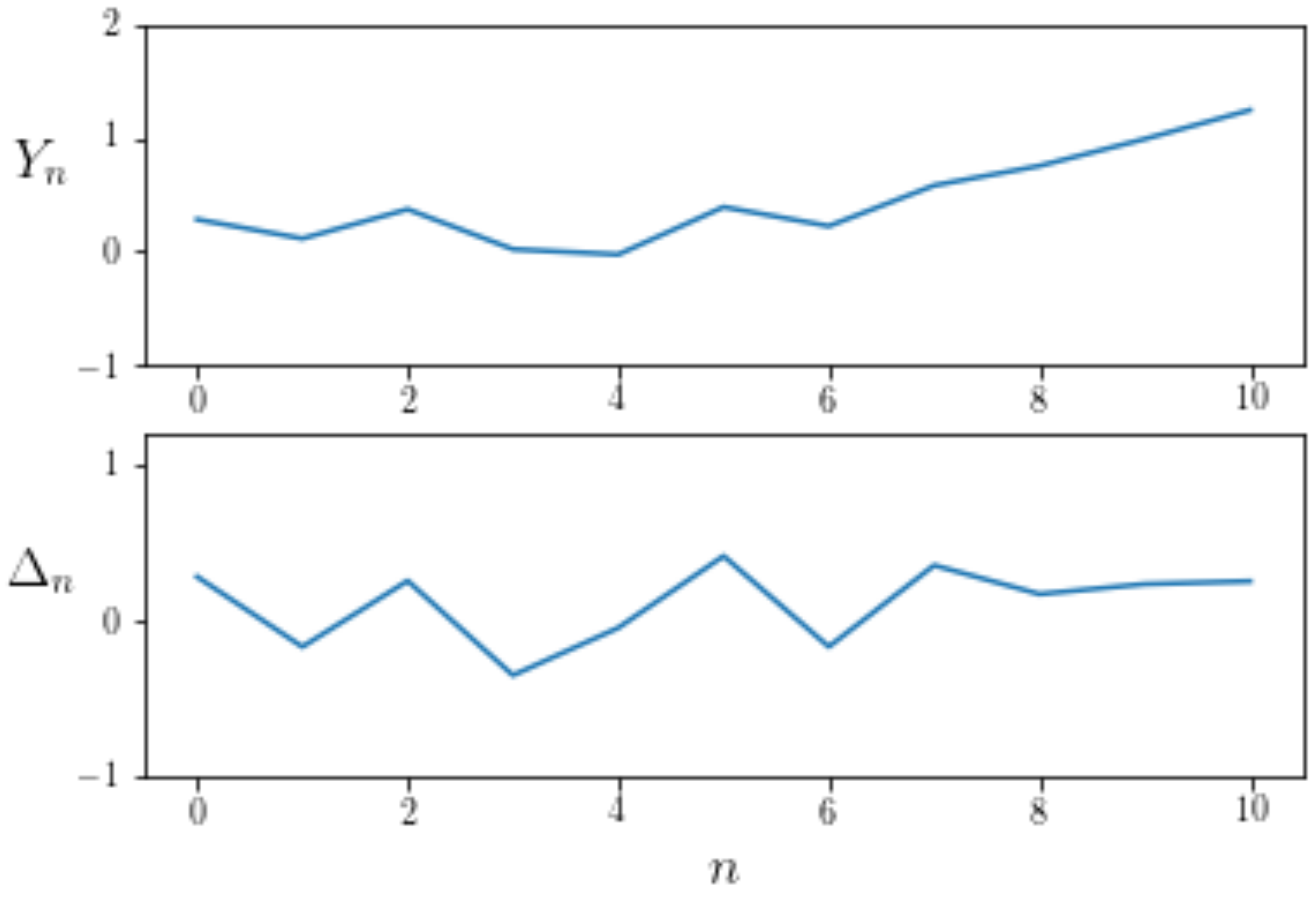
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

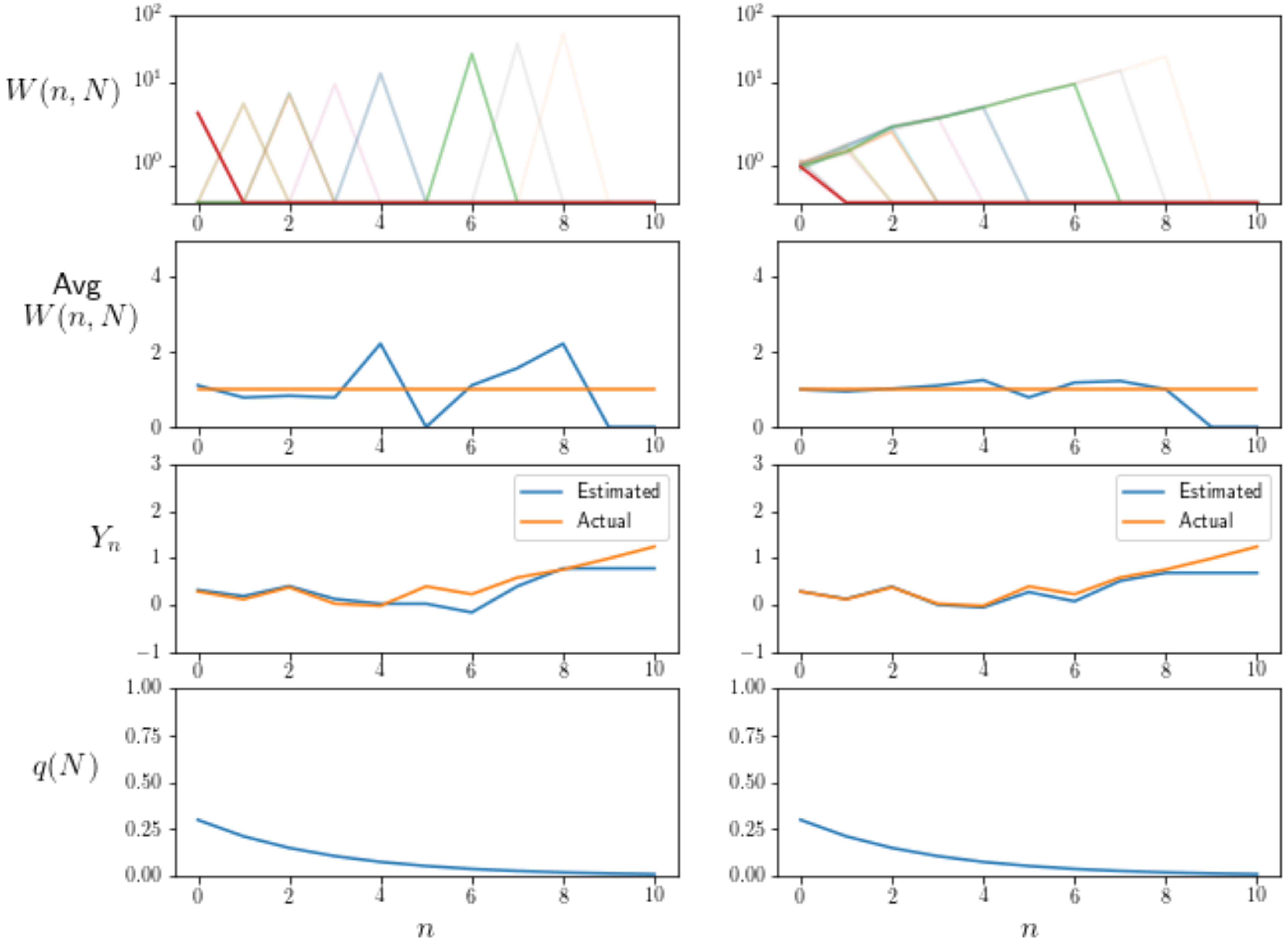


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



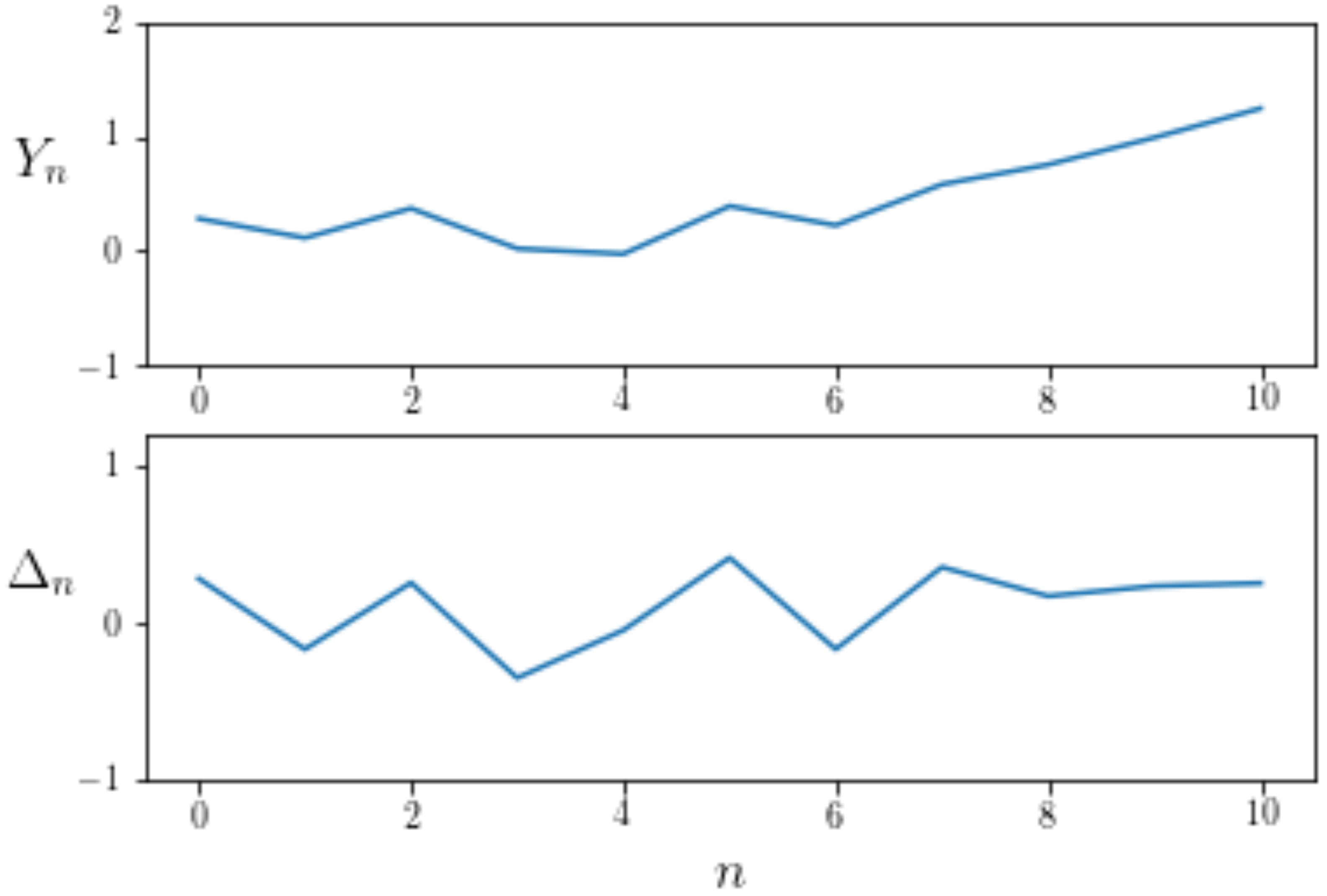
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

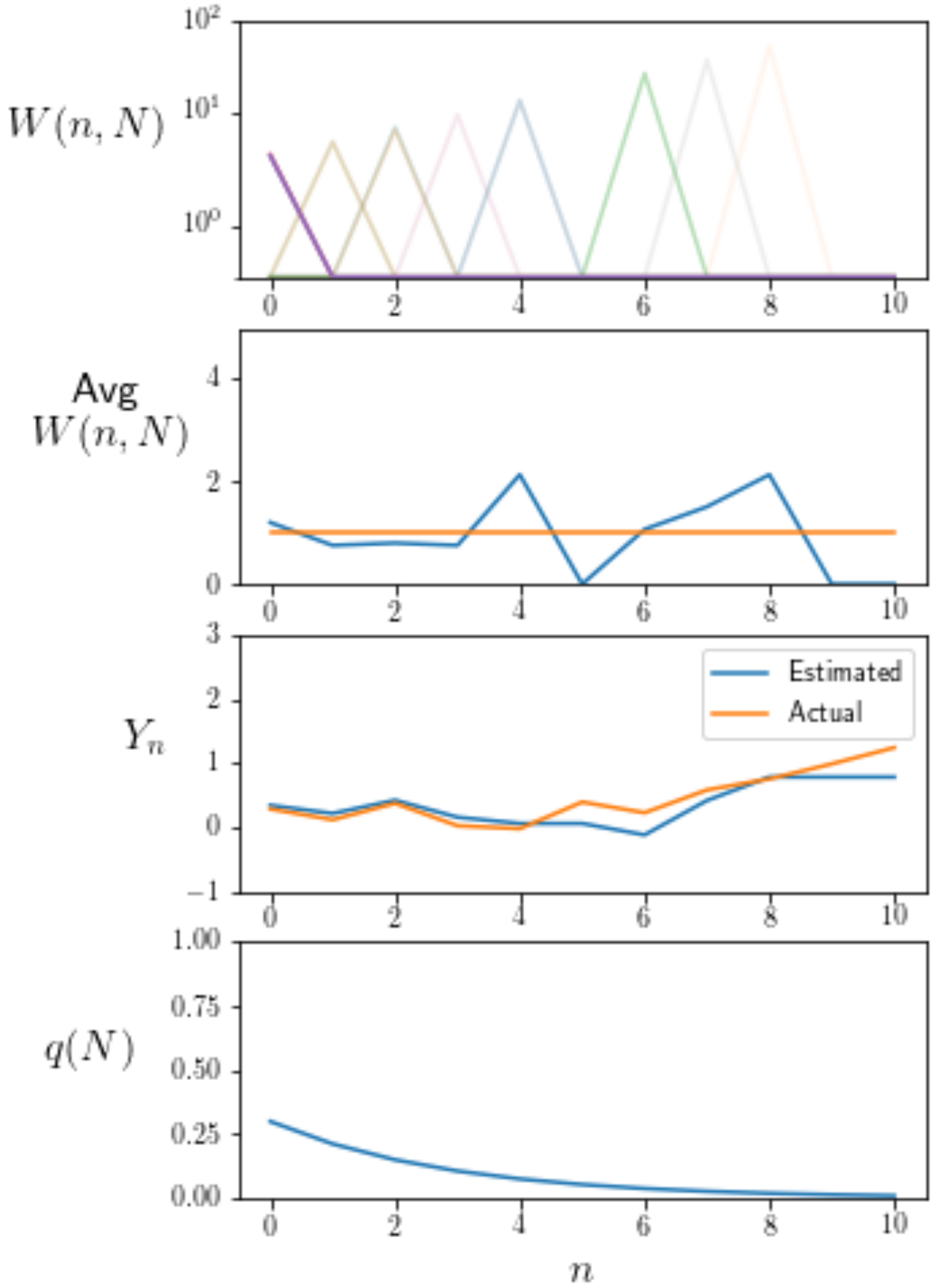
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



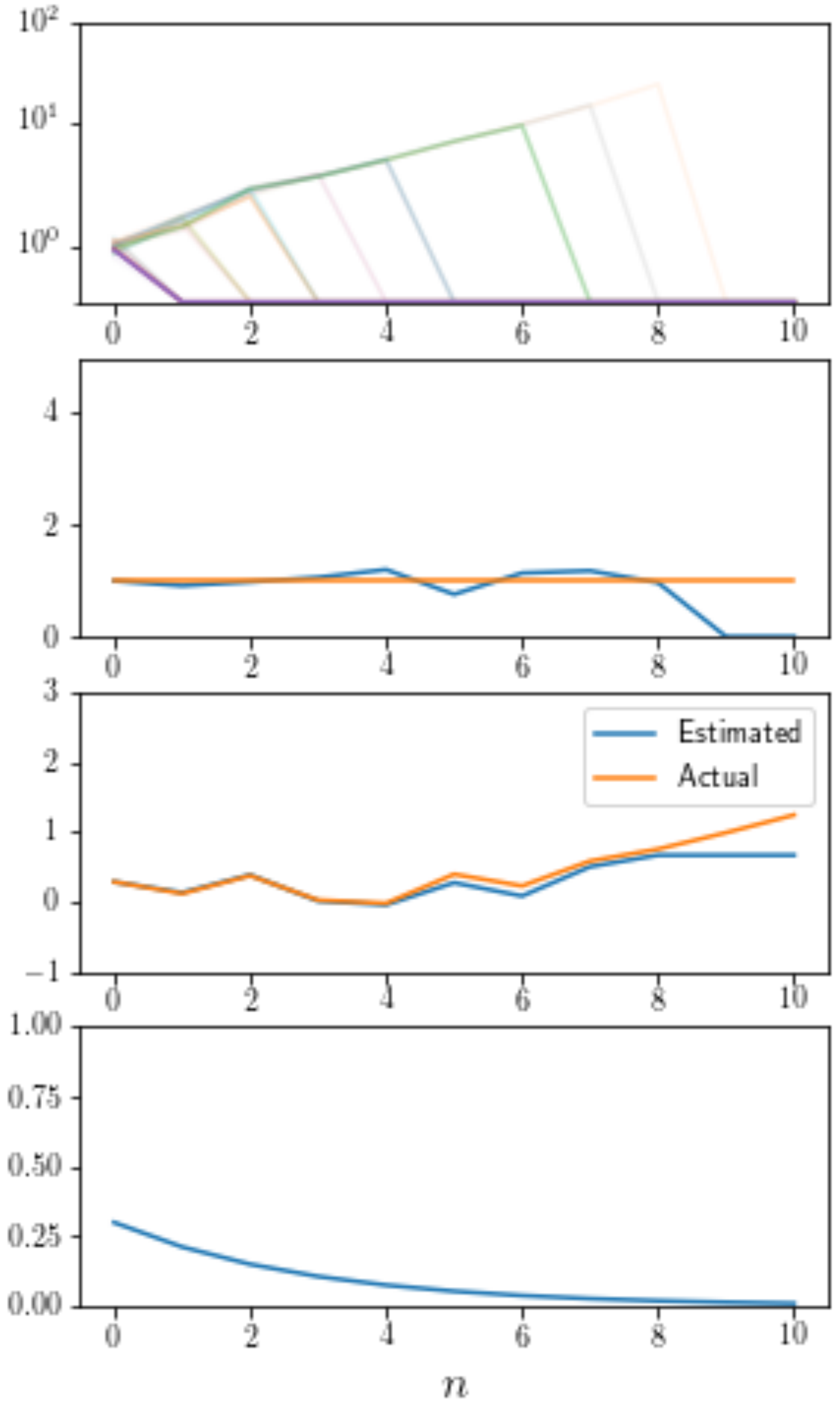
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



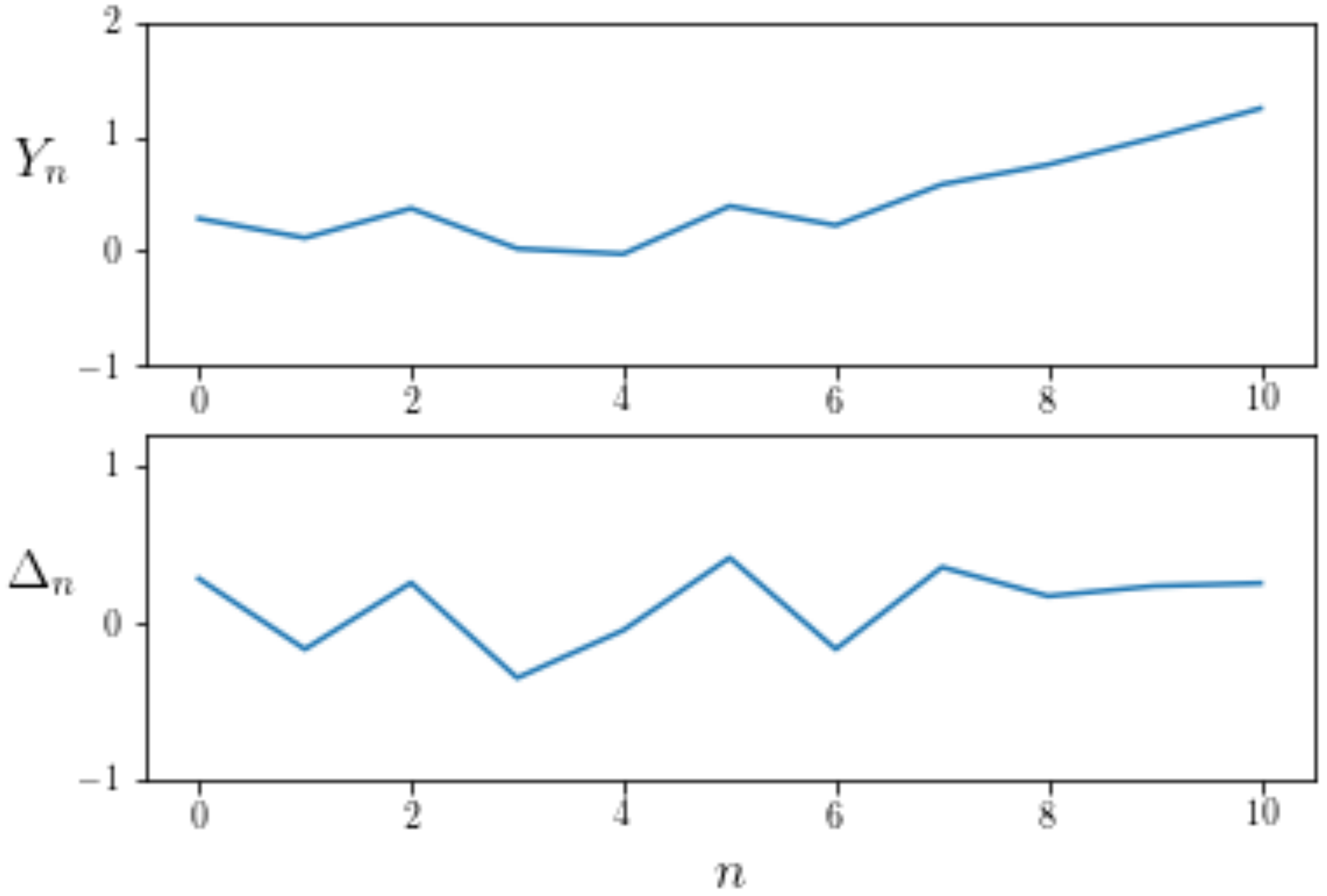
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

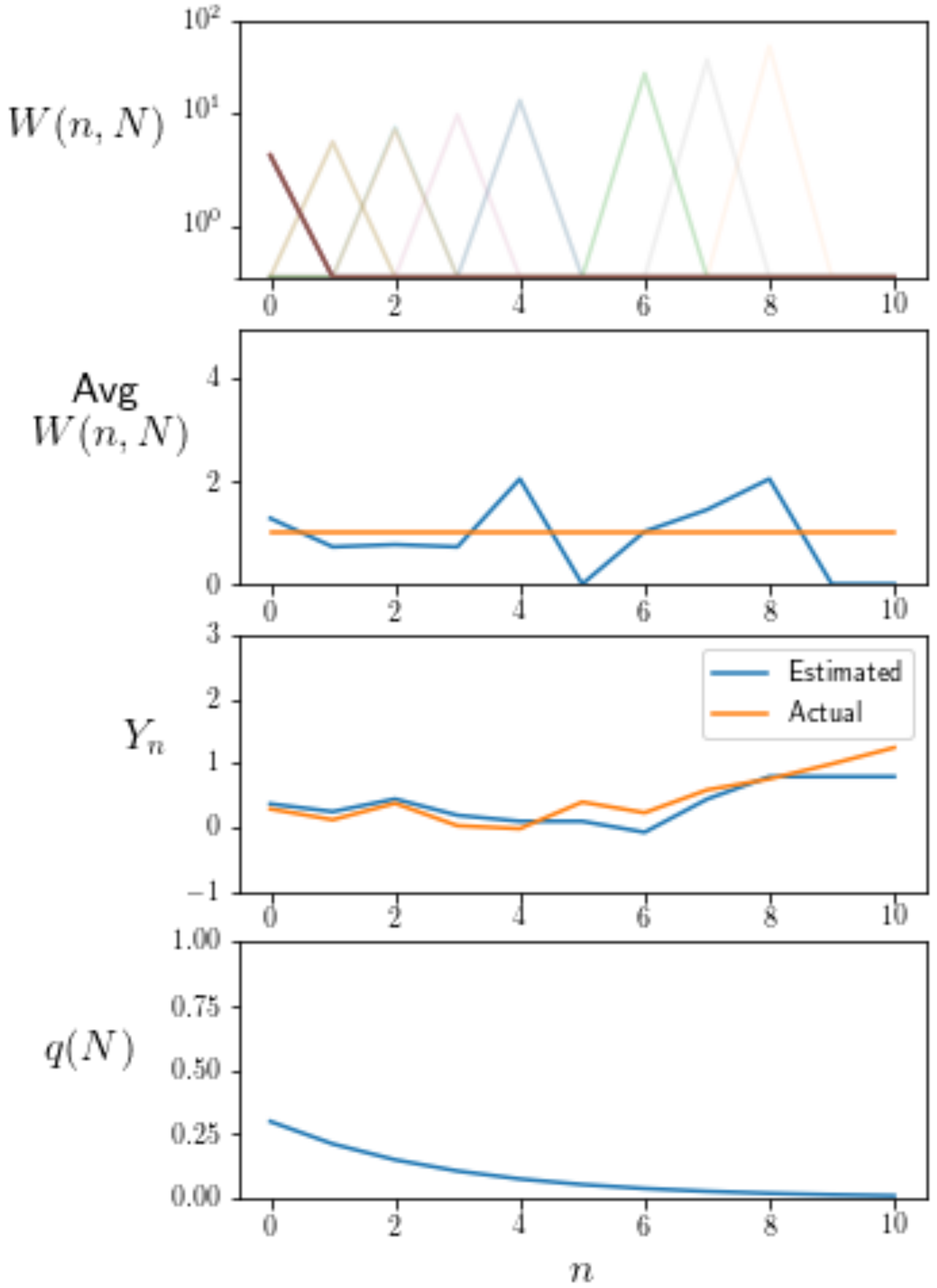
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



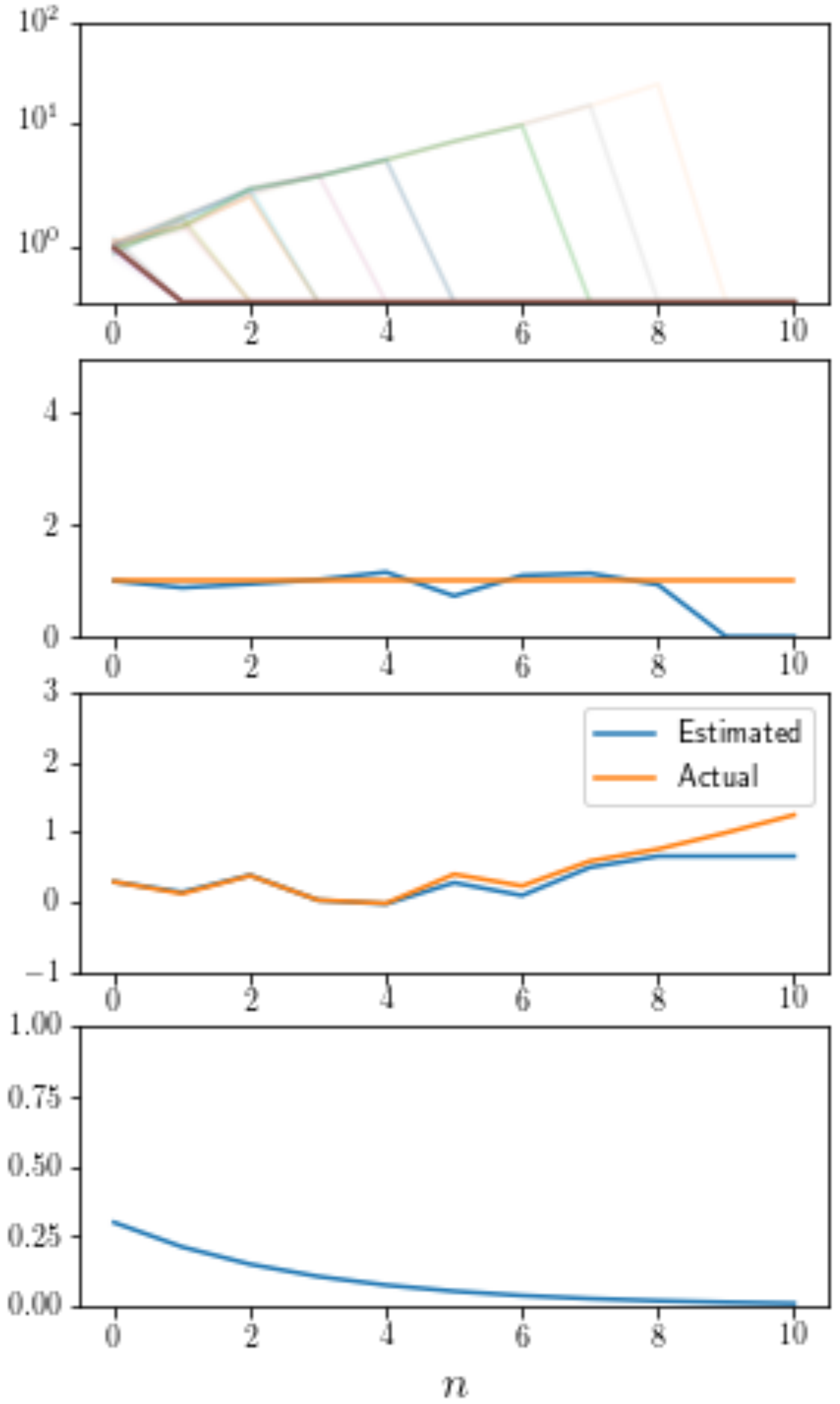
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



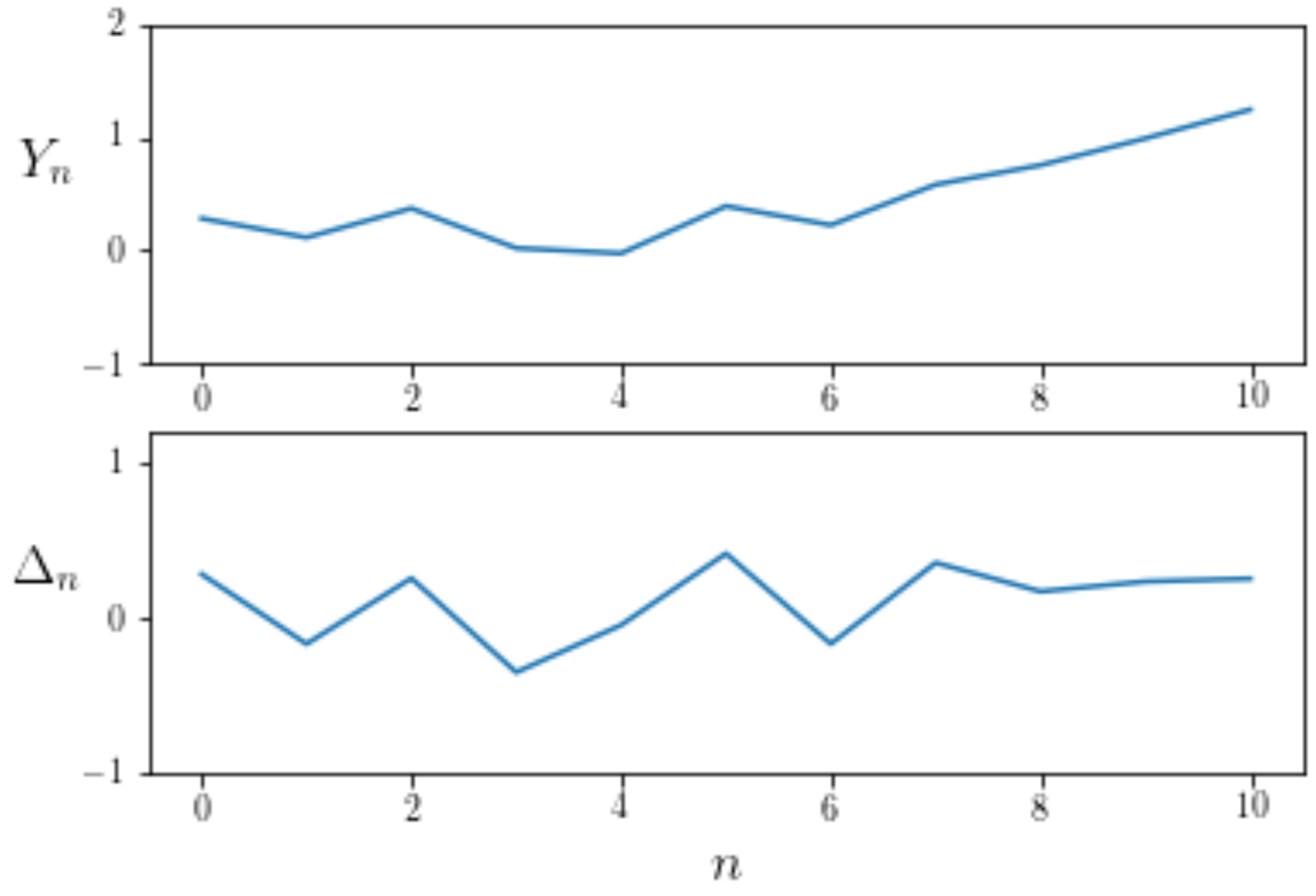
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

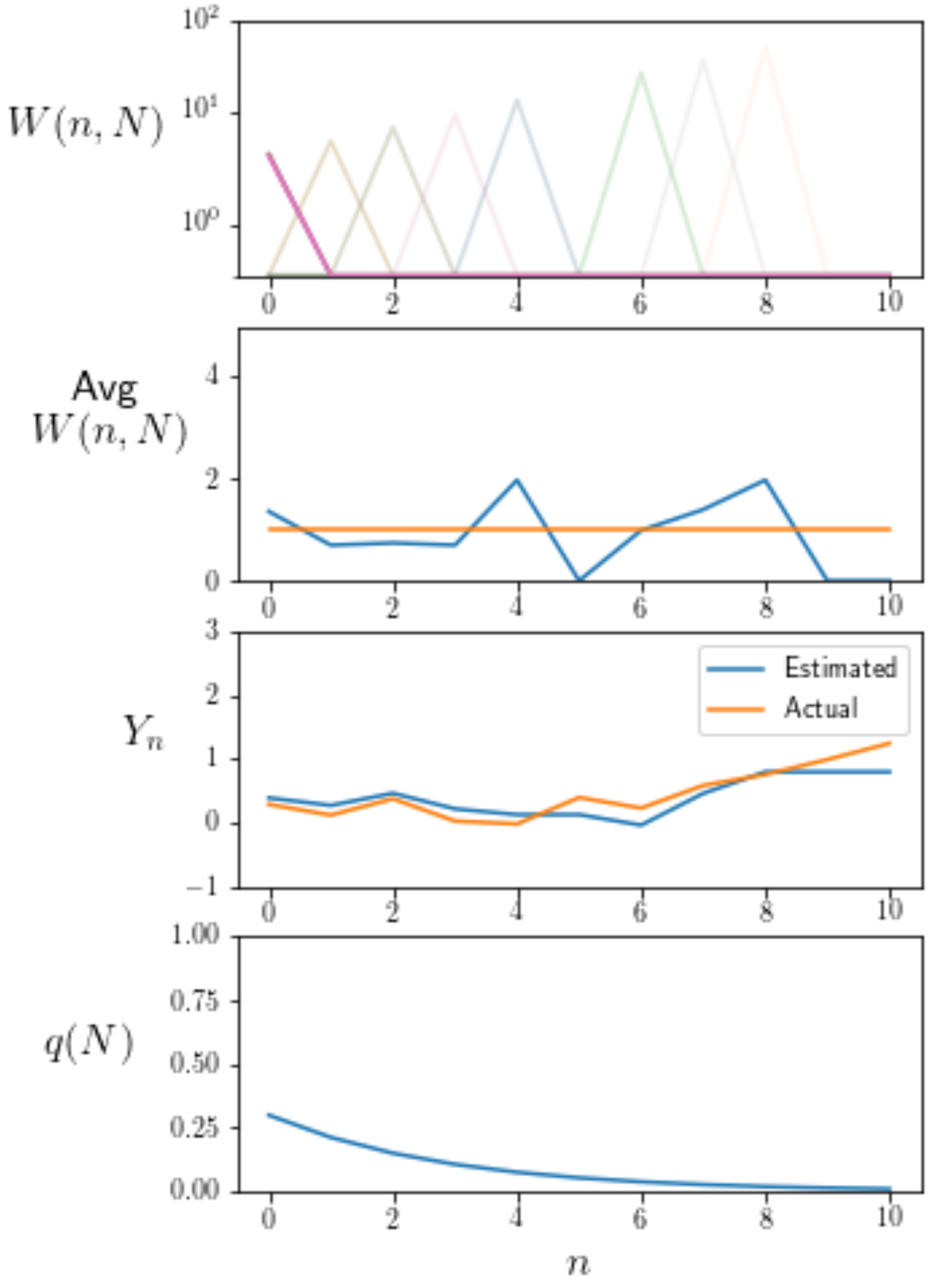
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



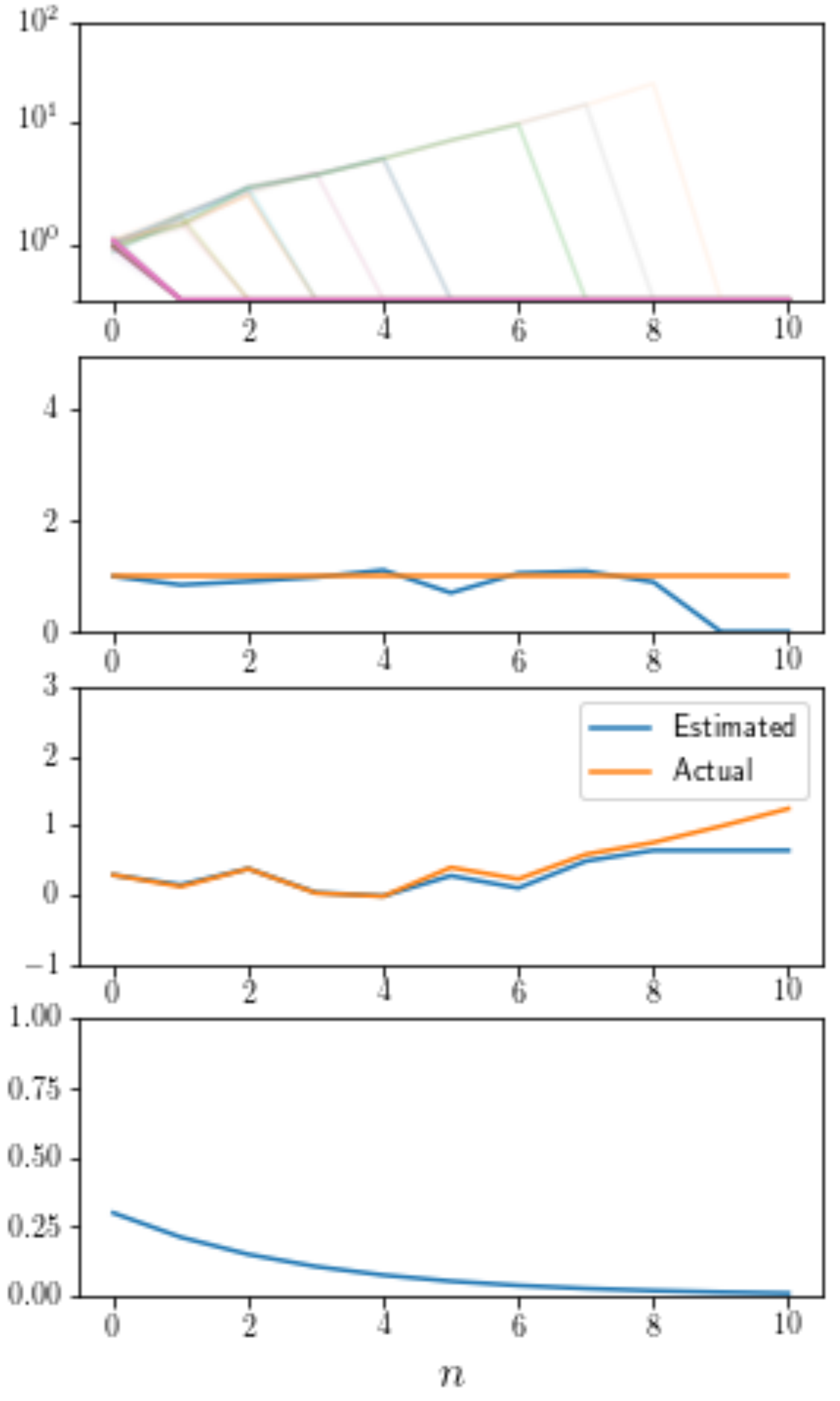
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



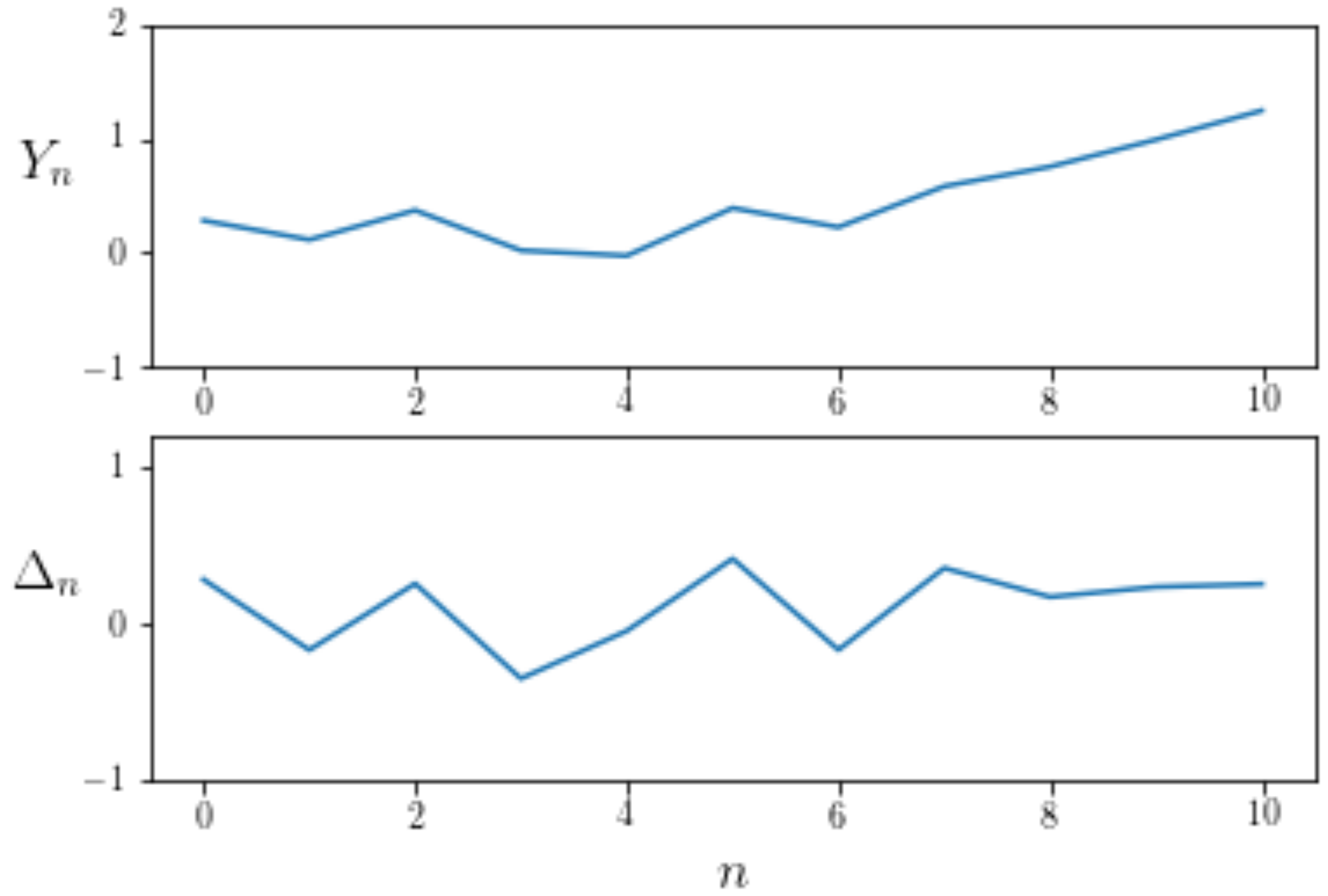
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

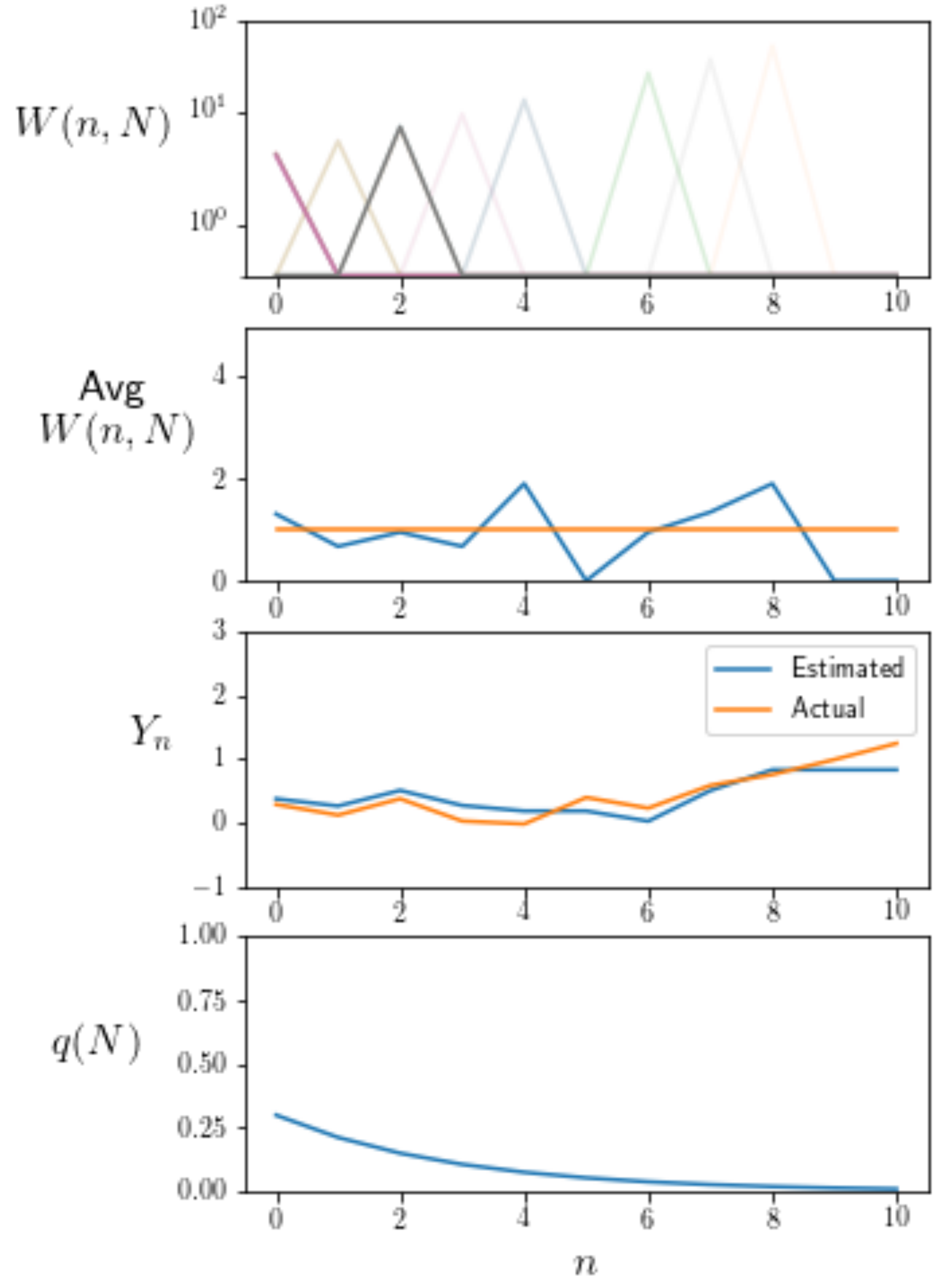
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



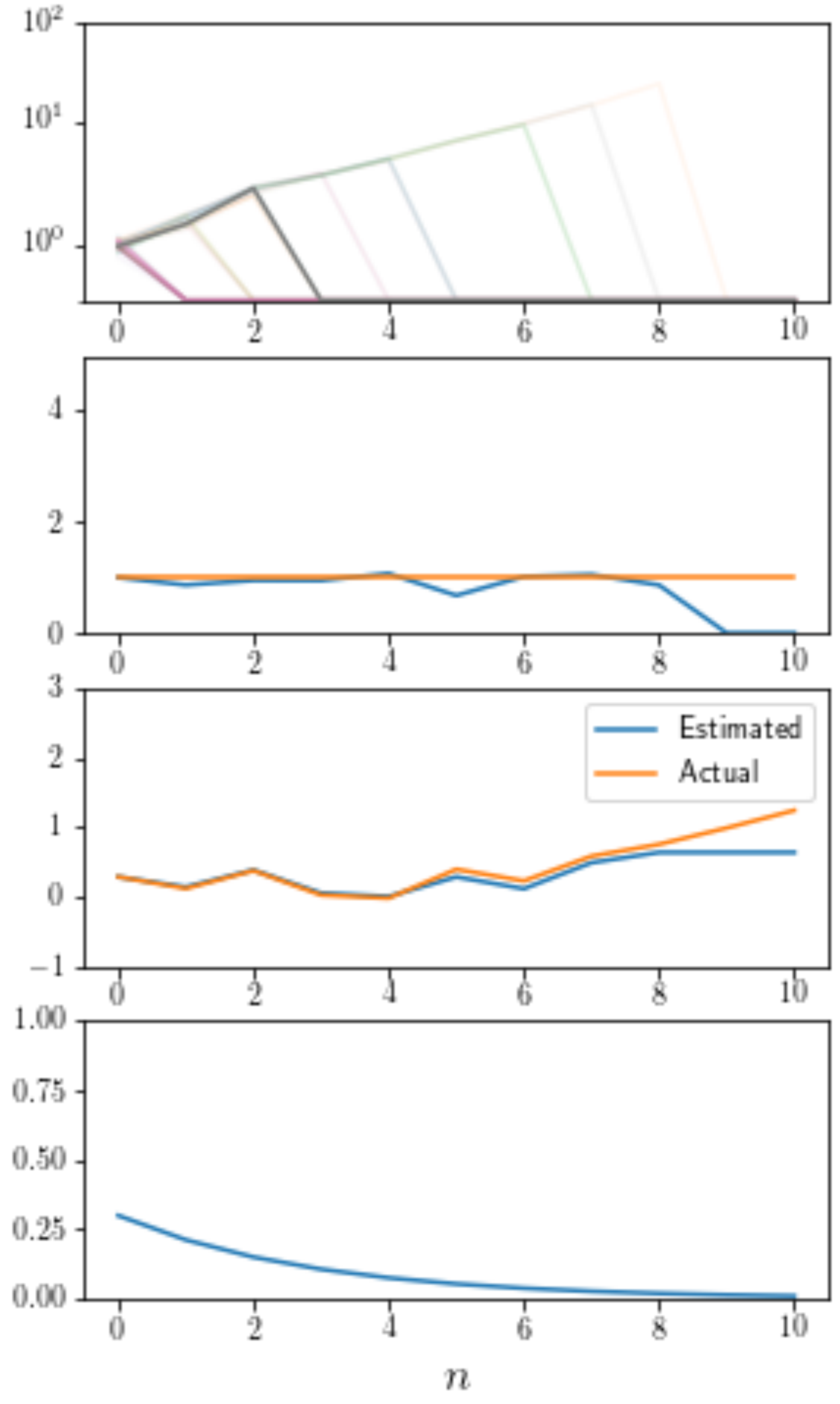
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



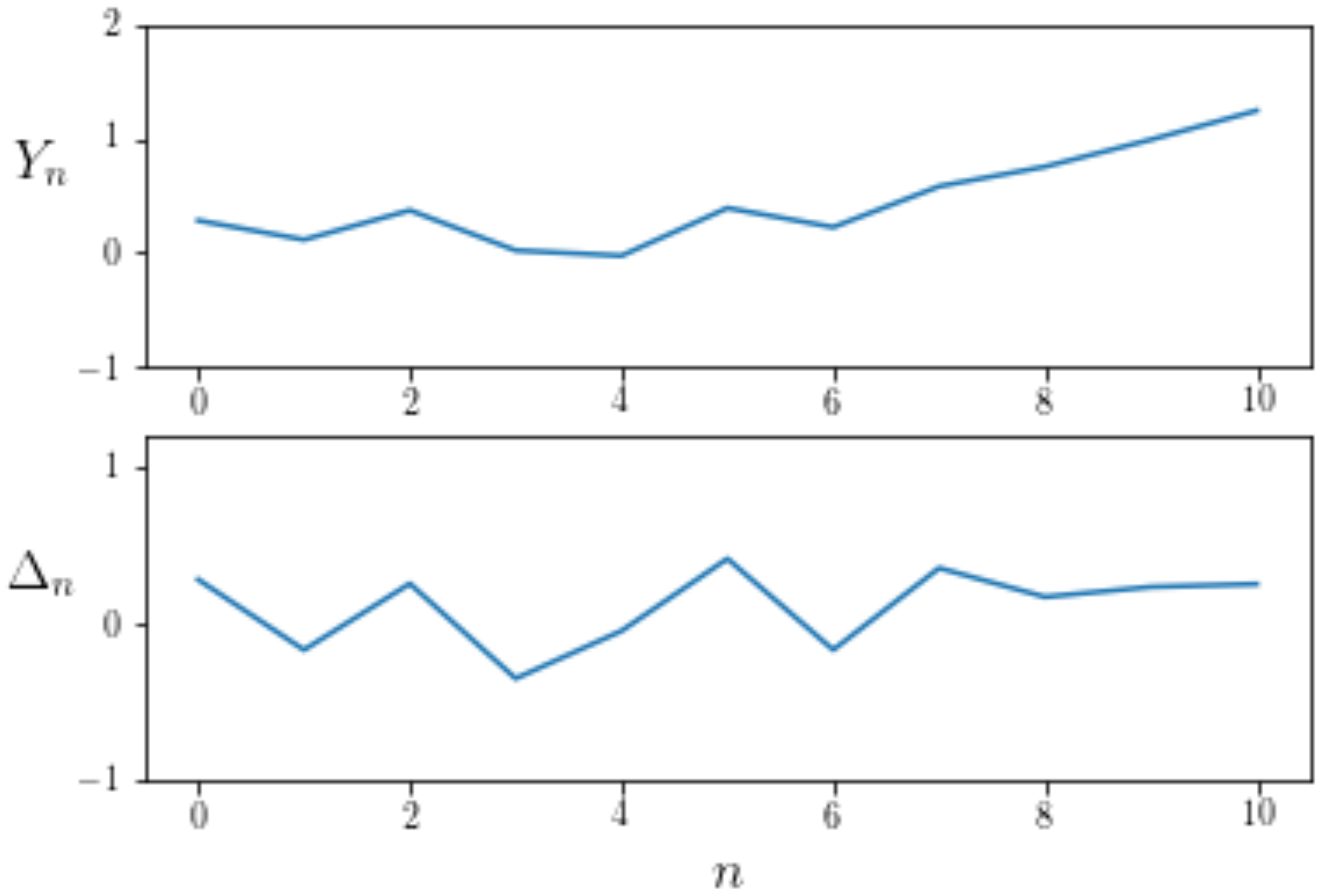
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

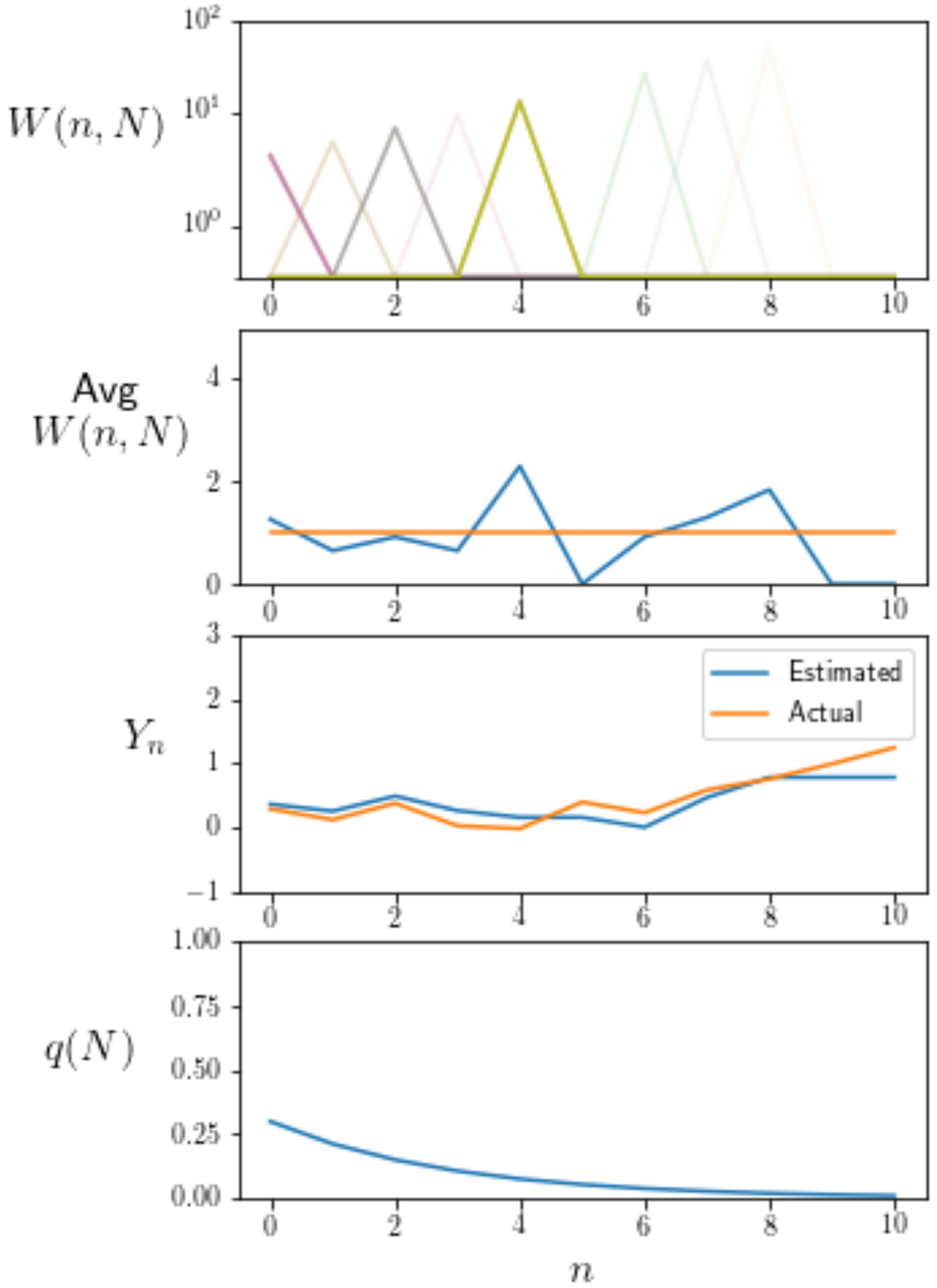
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



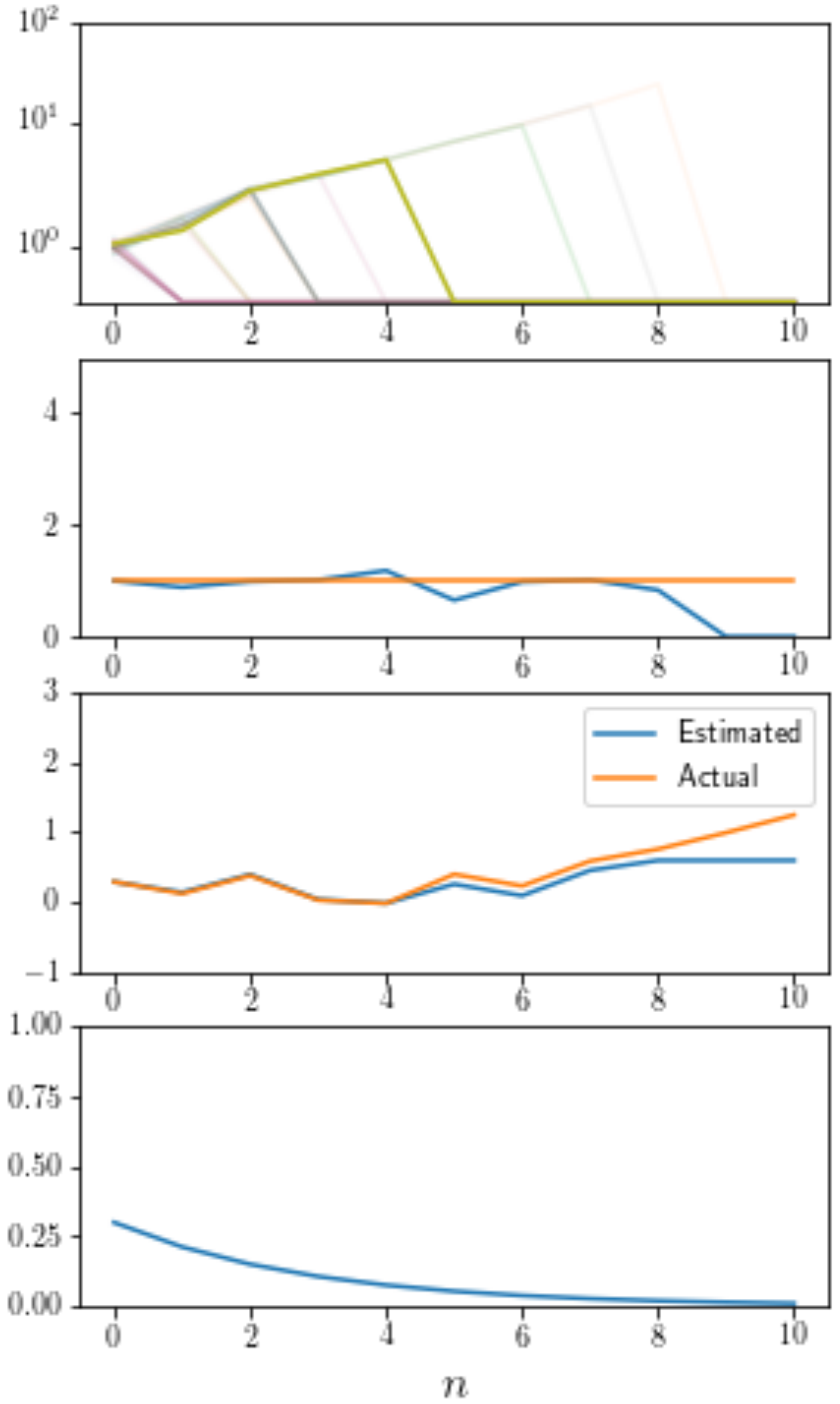
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



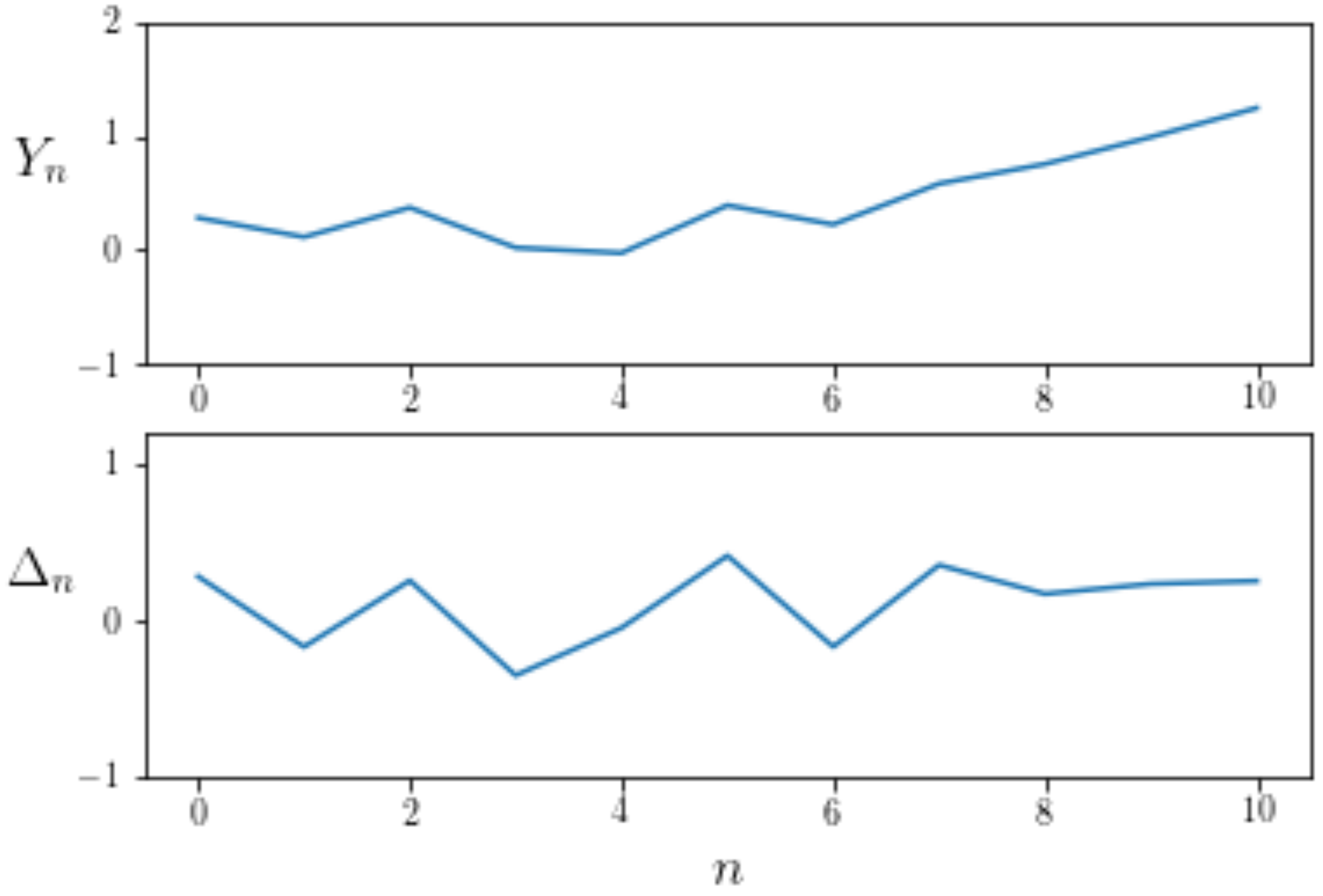
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

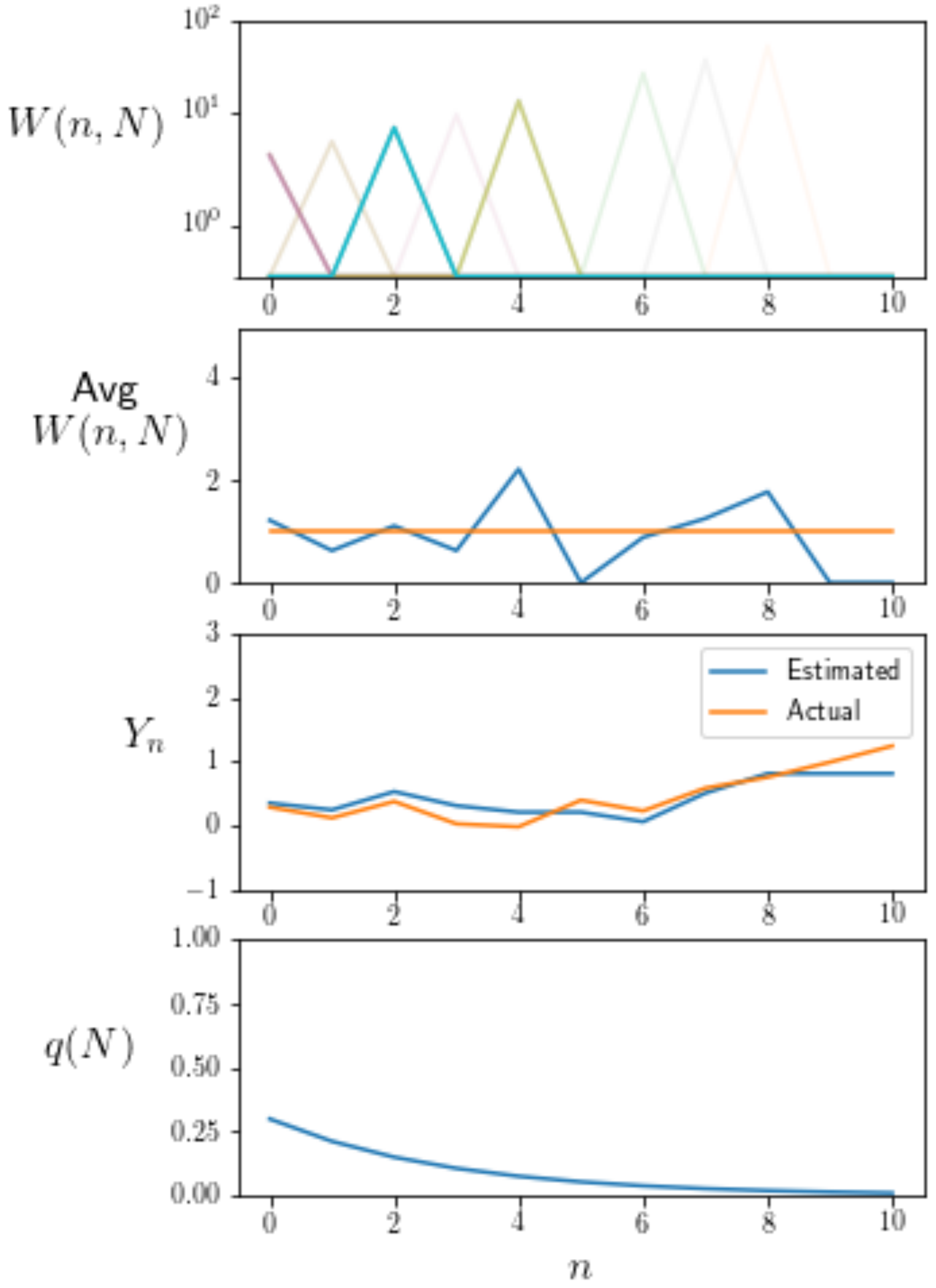
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



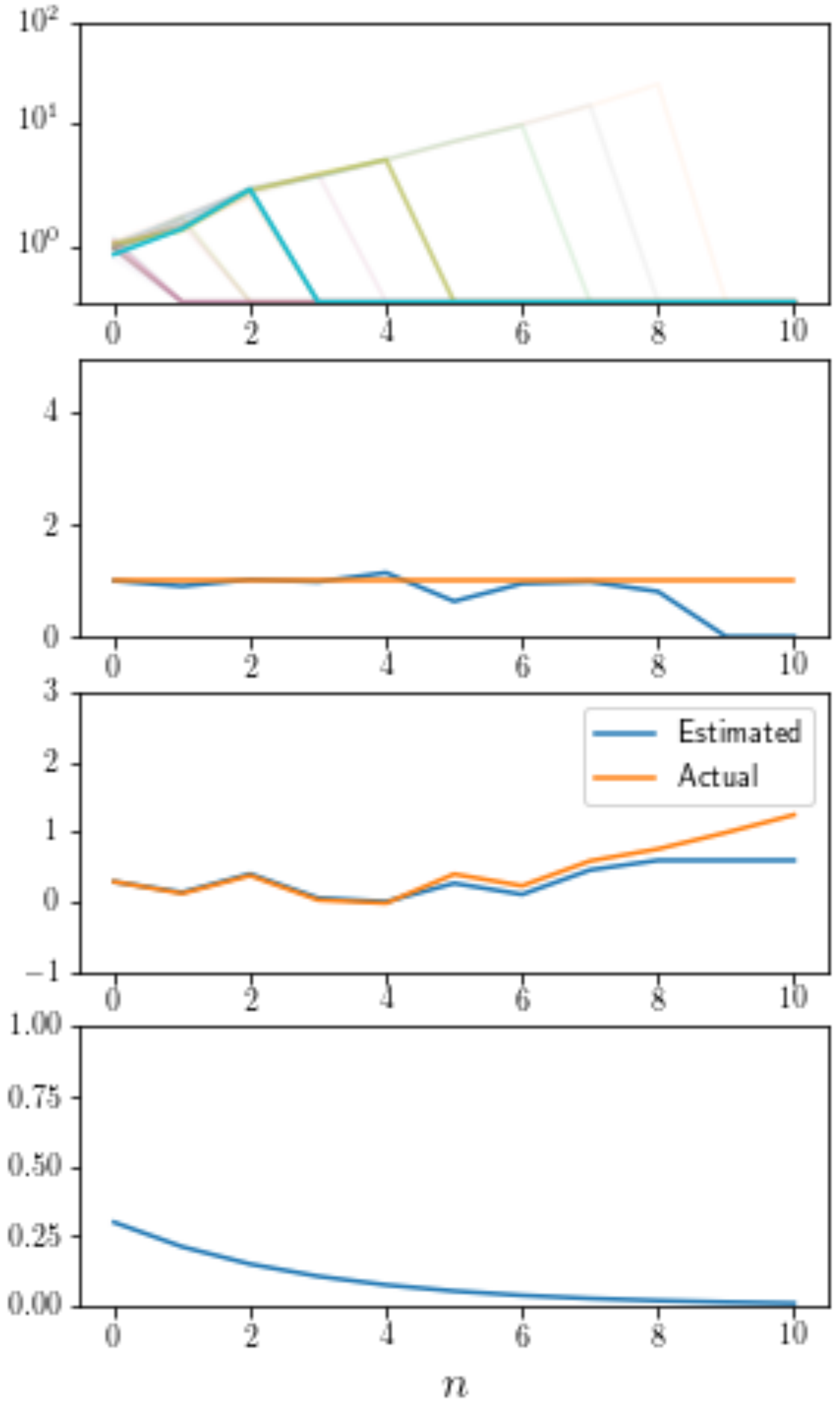
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



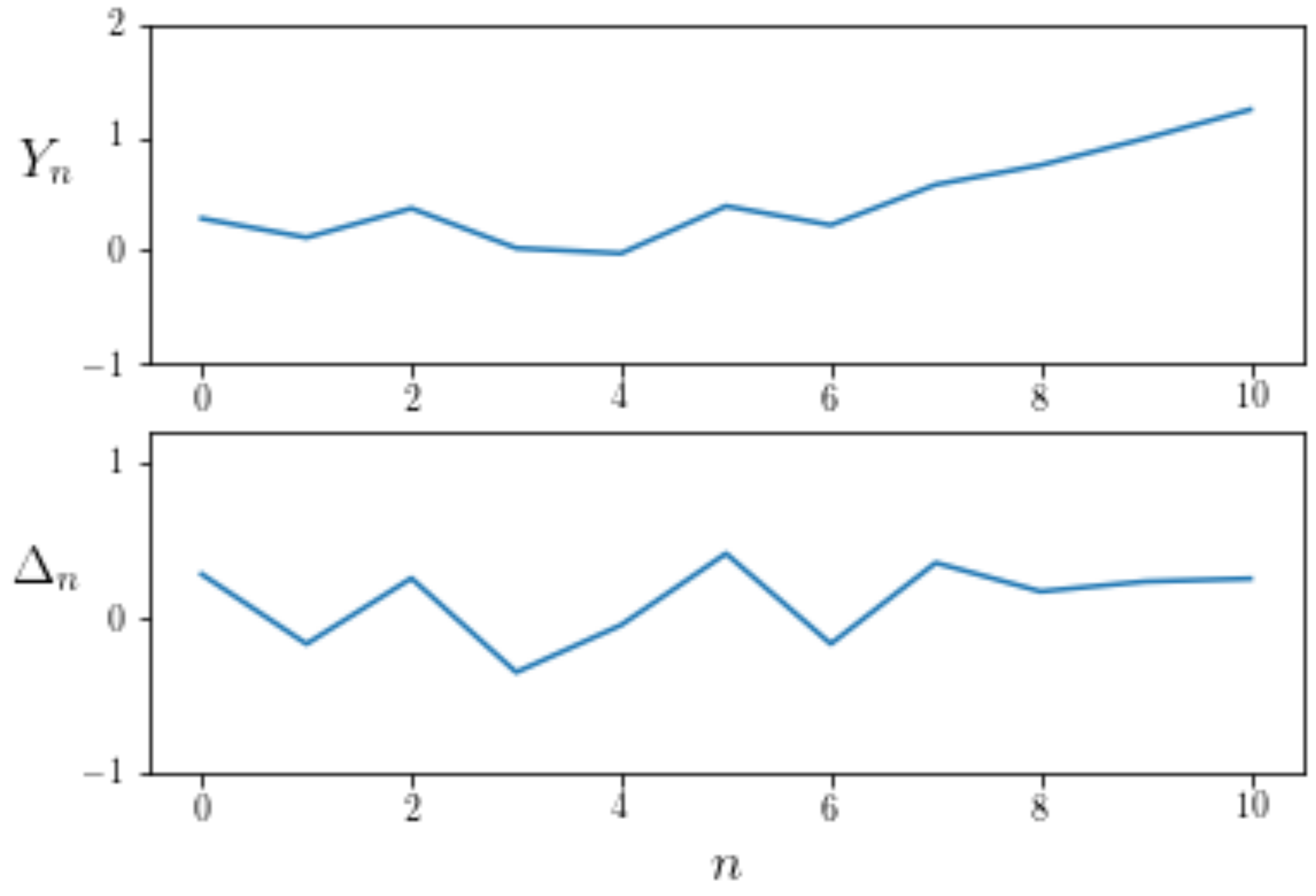
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

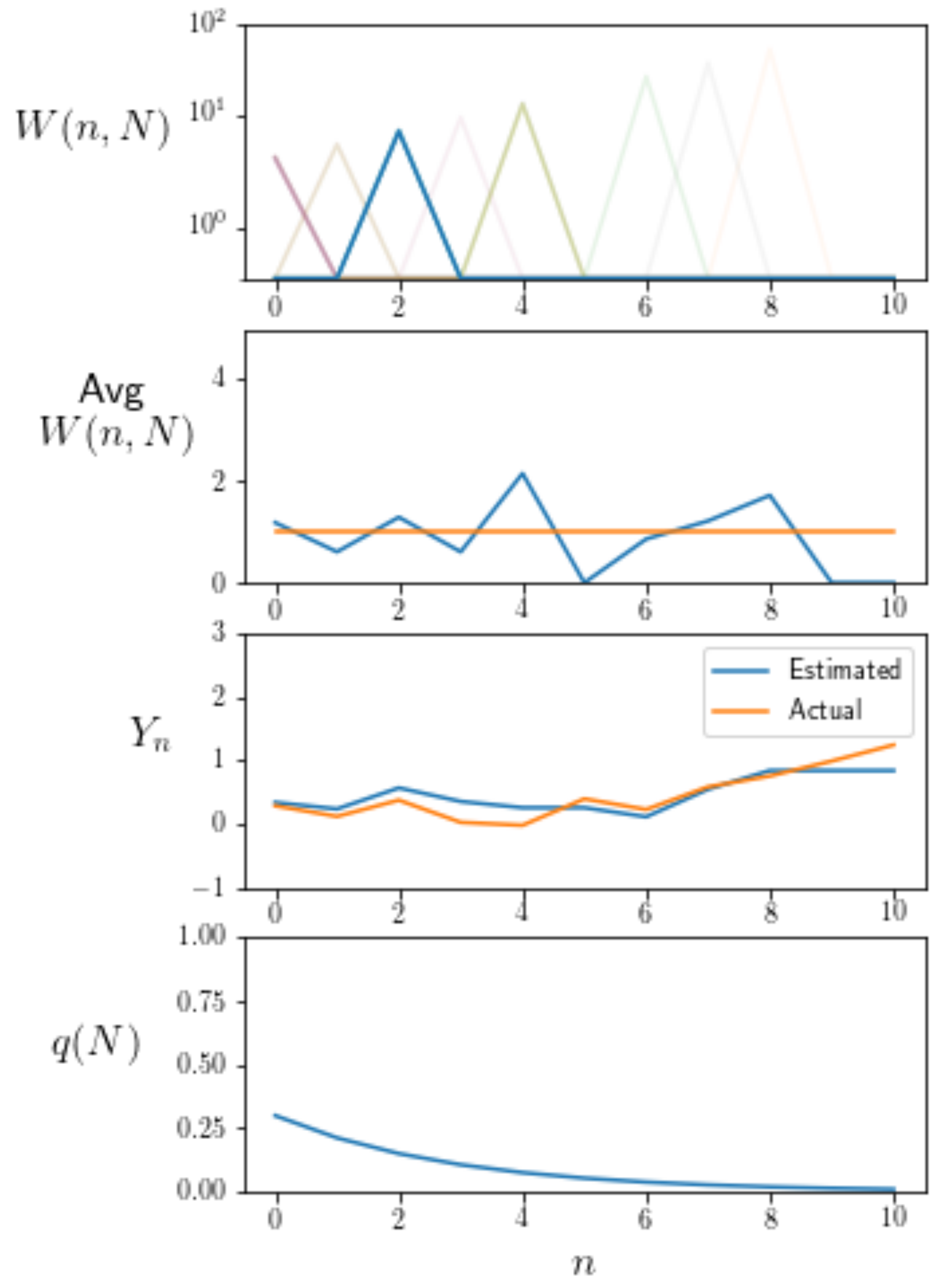
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



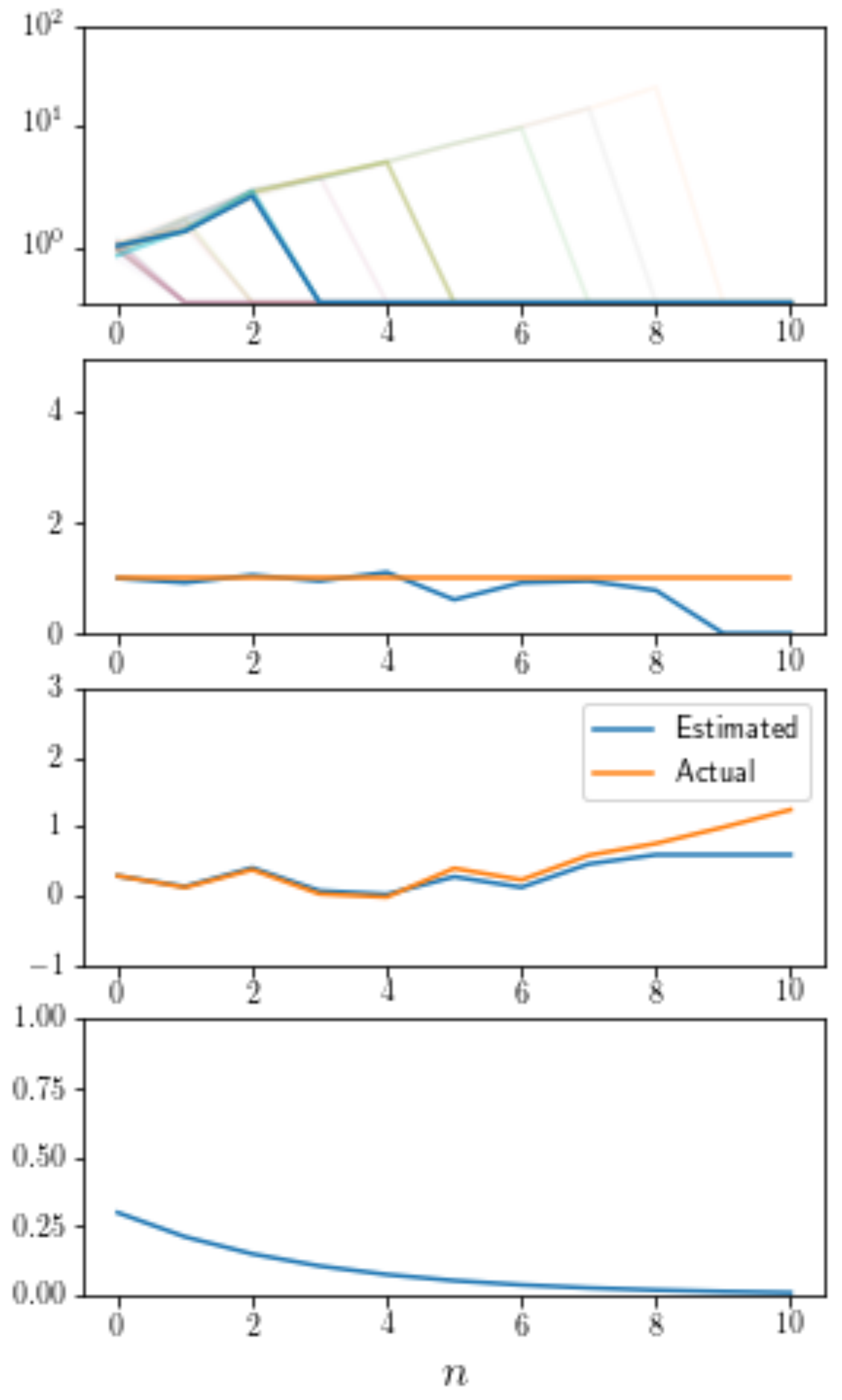
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



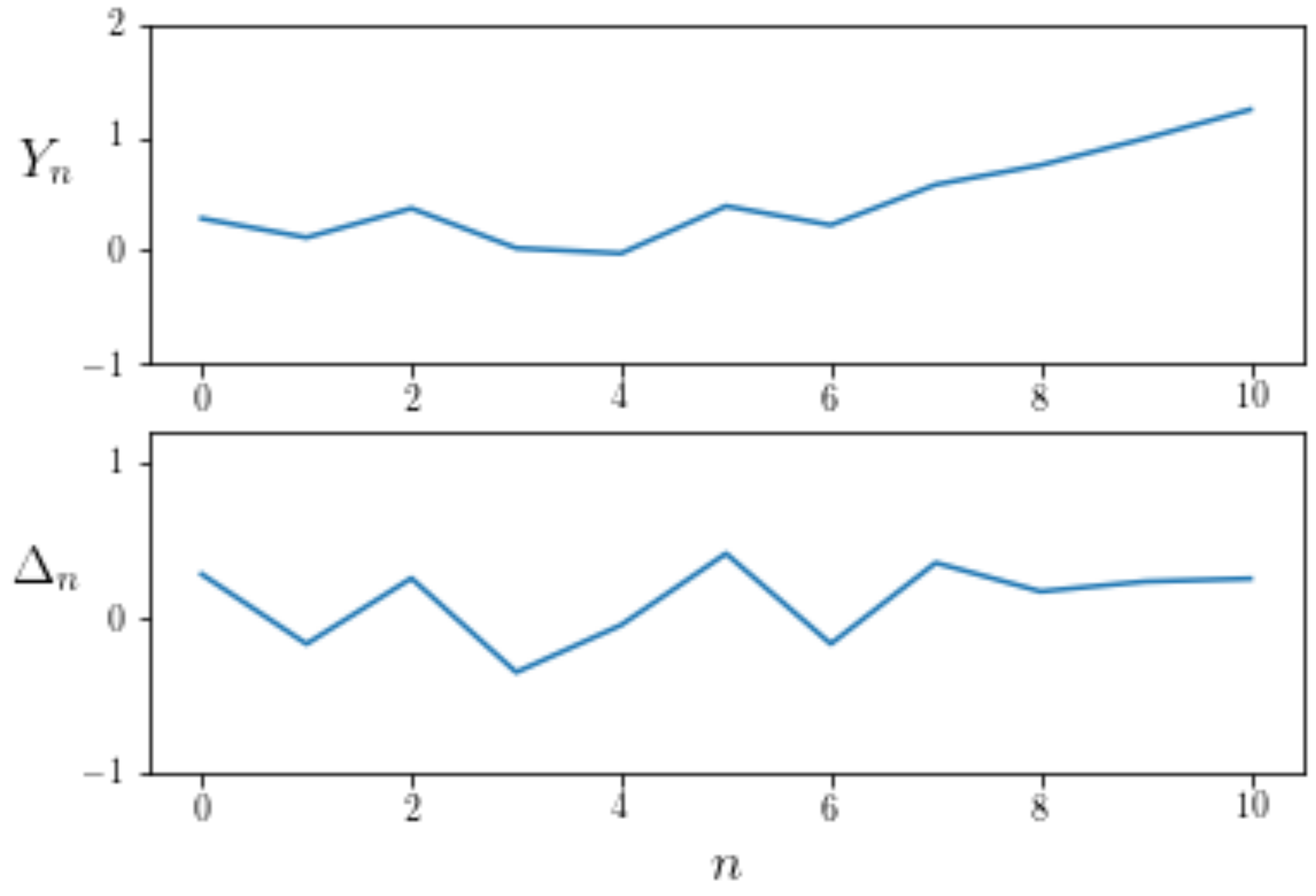
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

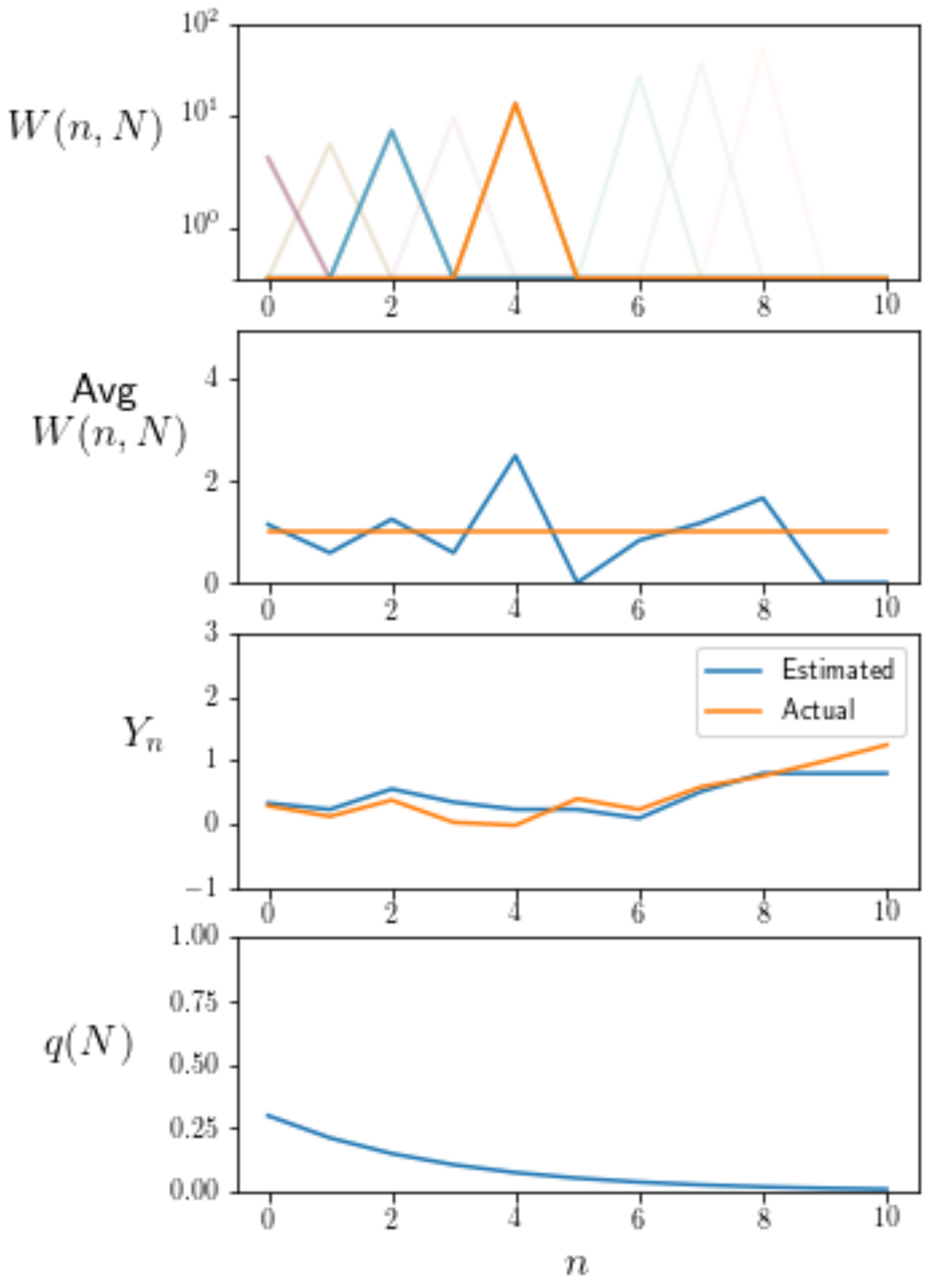


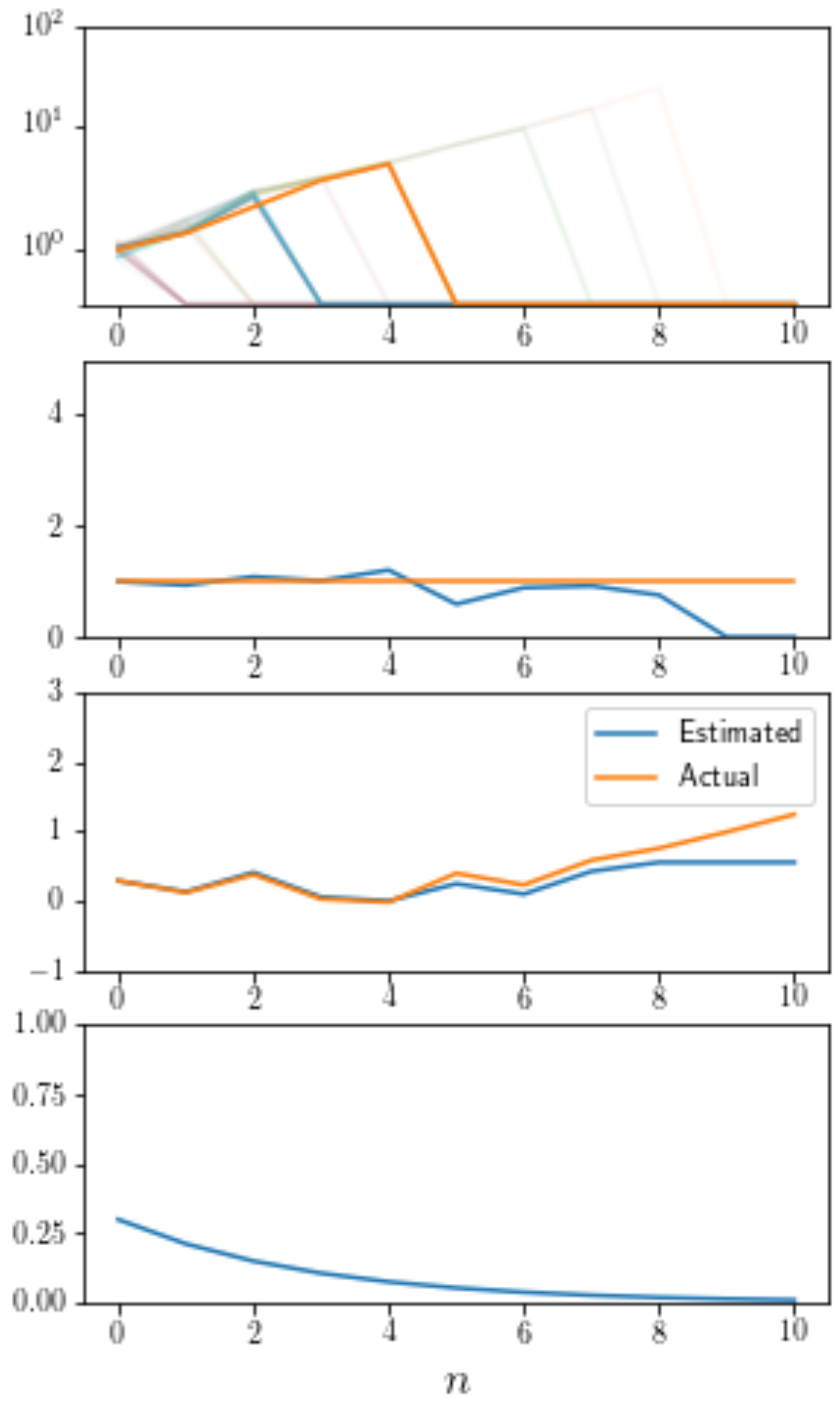
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$





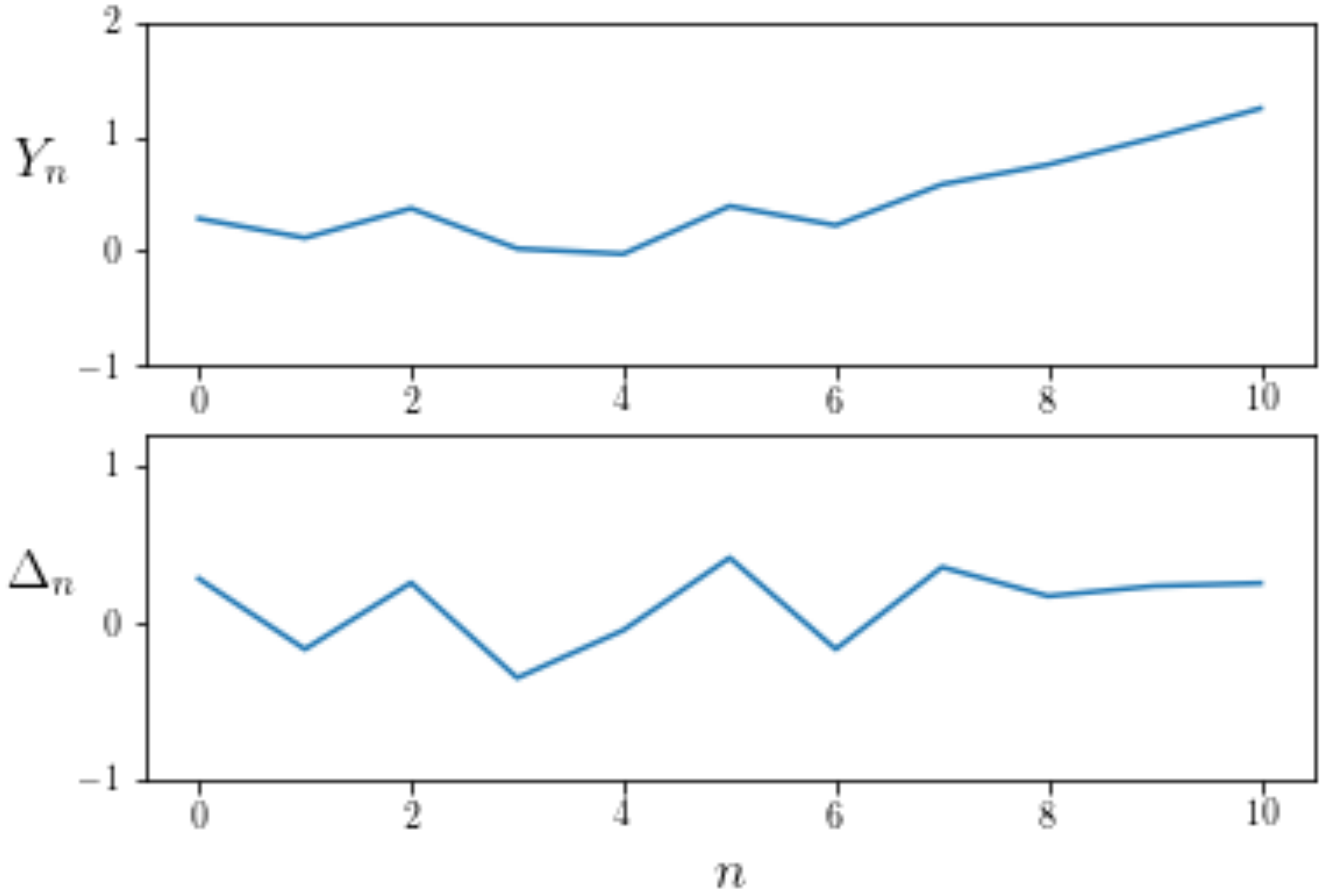
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

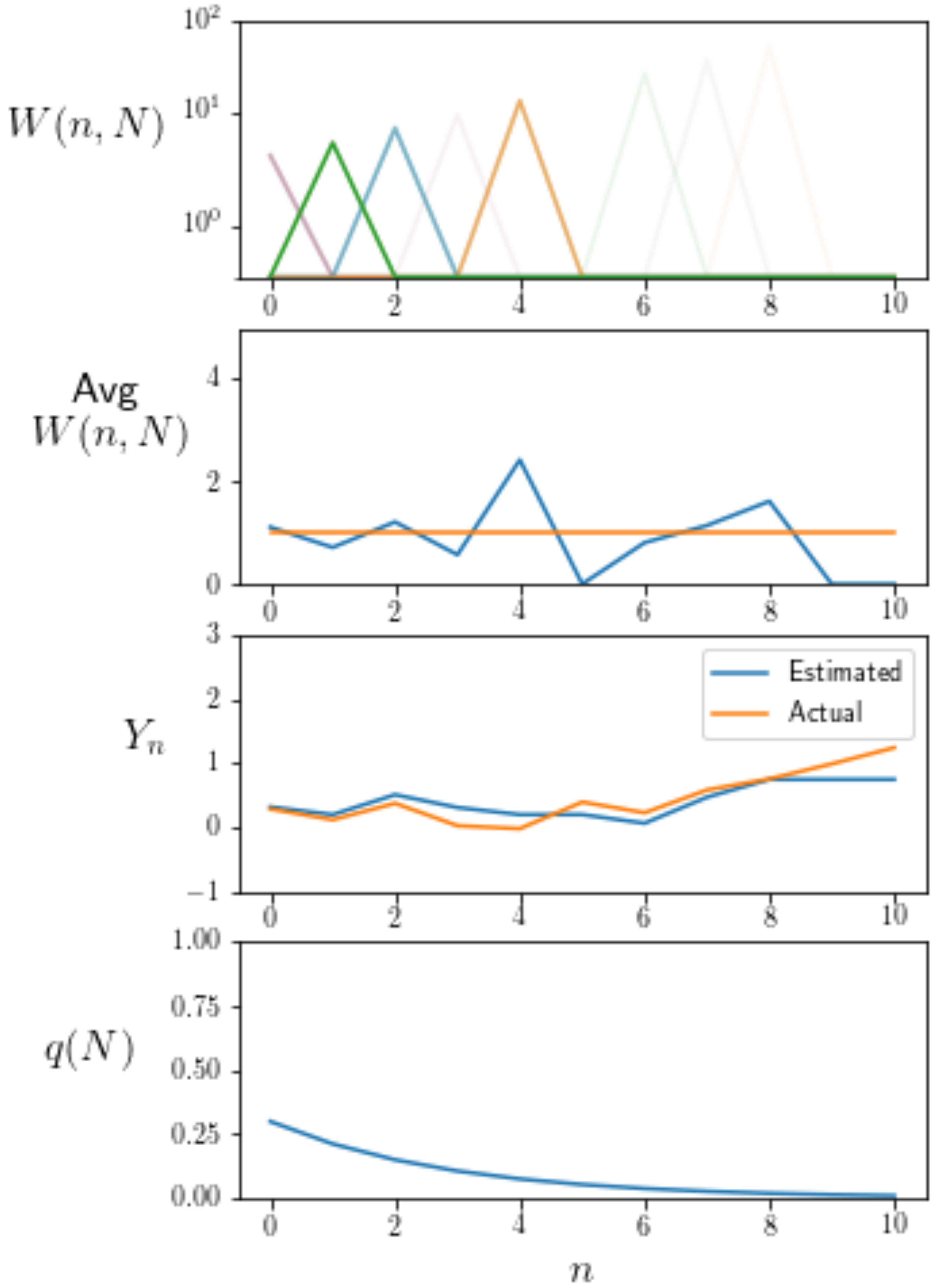
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



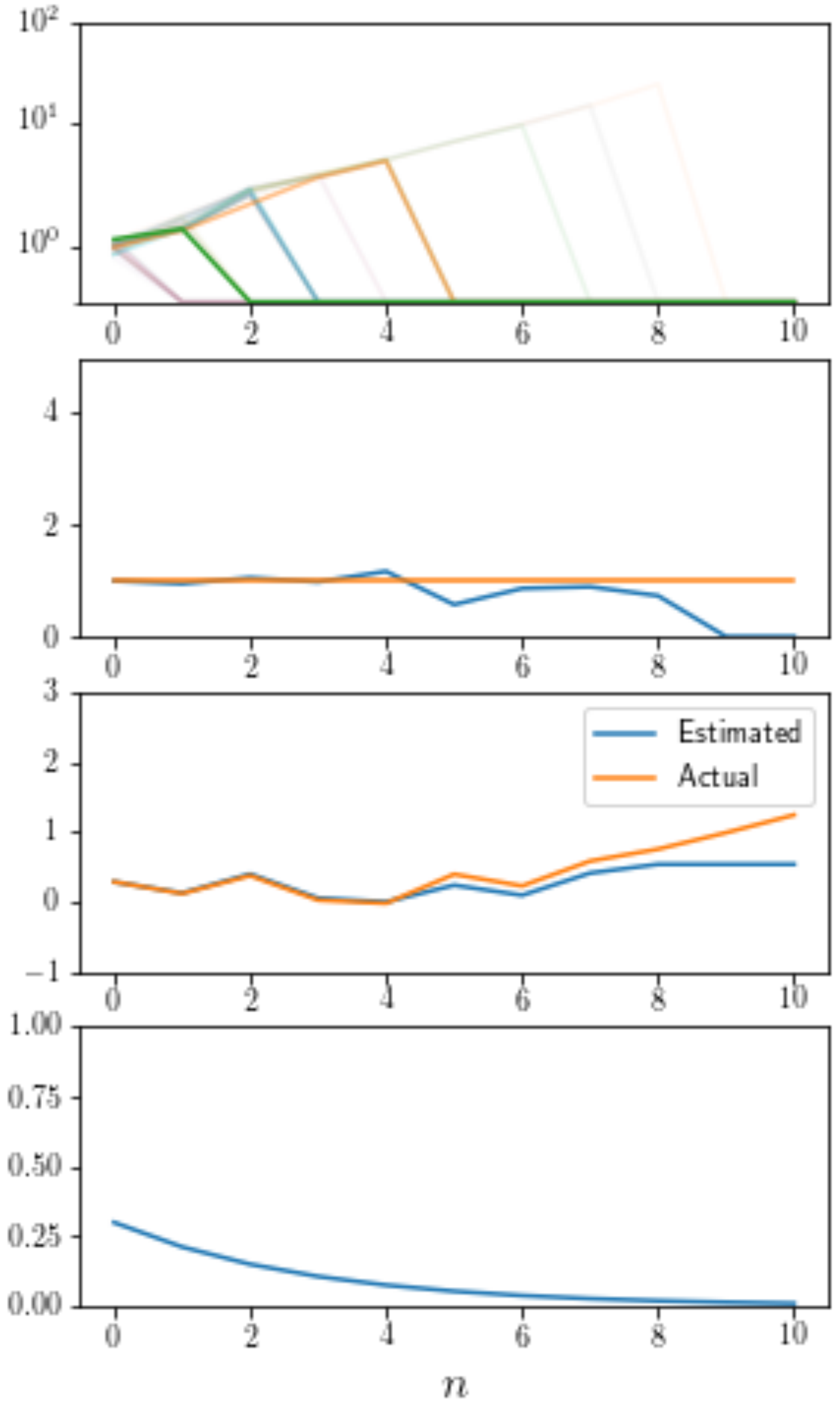
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



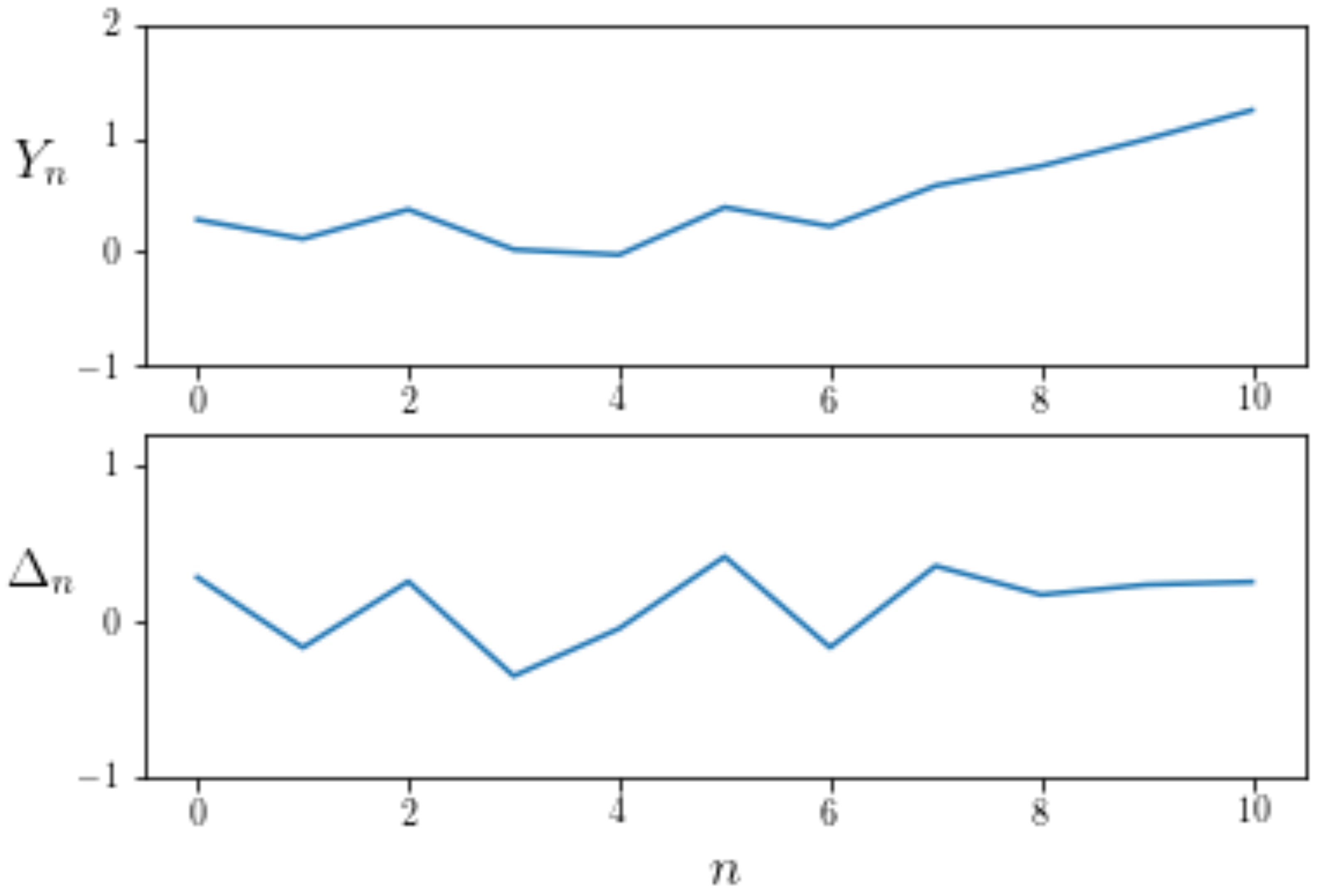
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

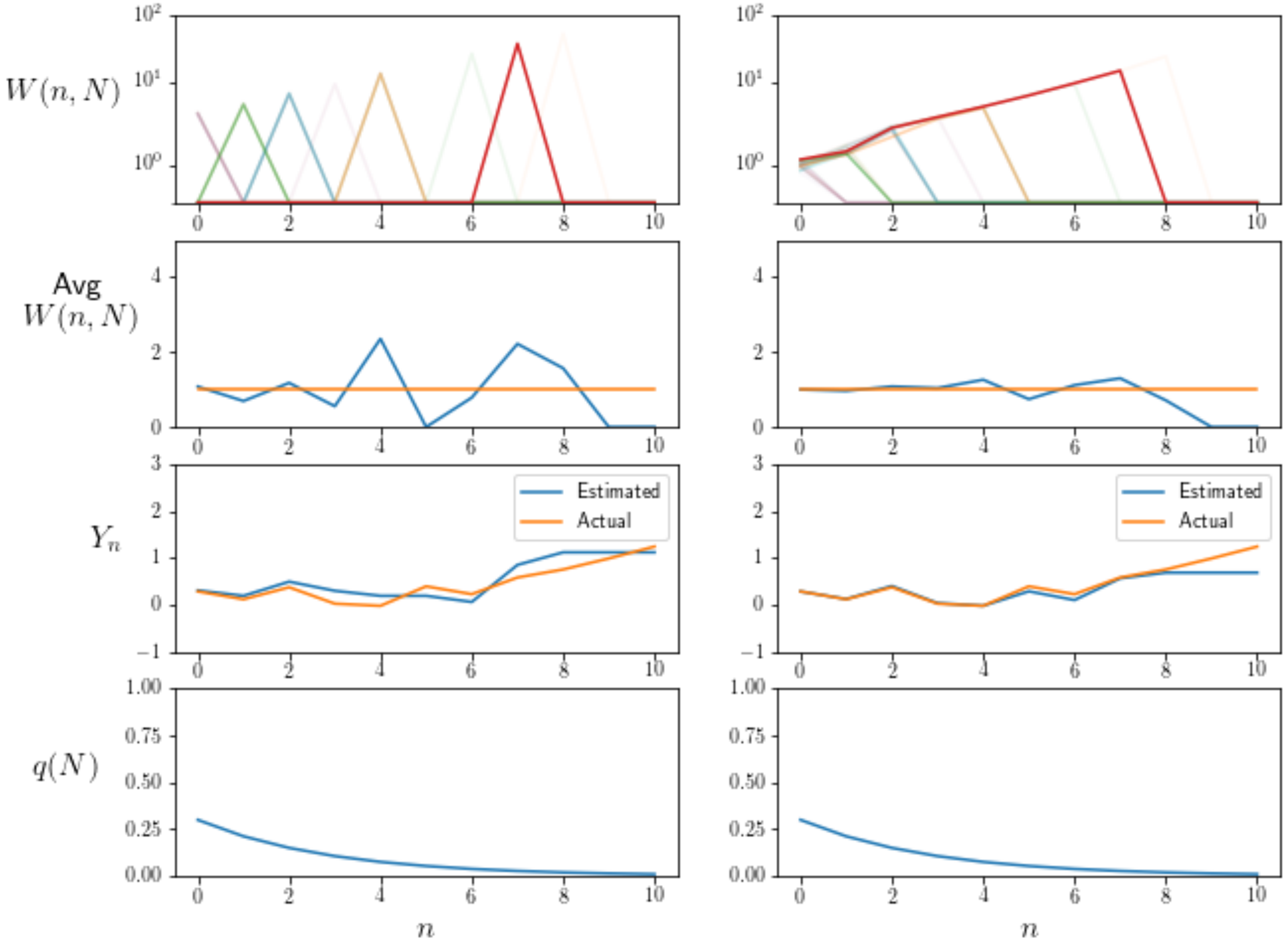


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



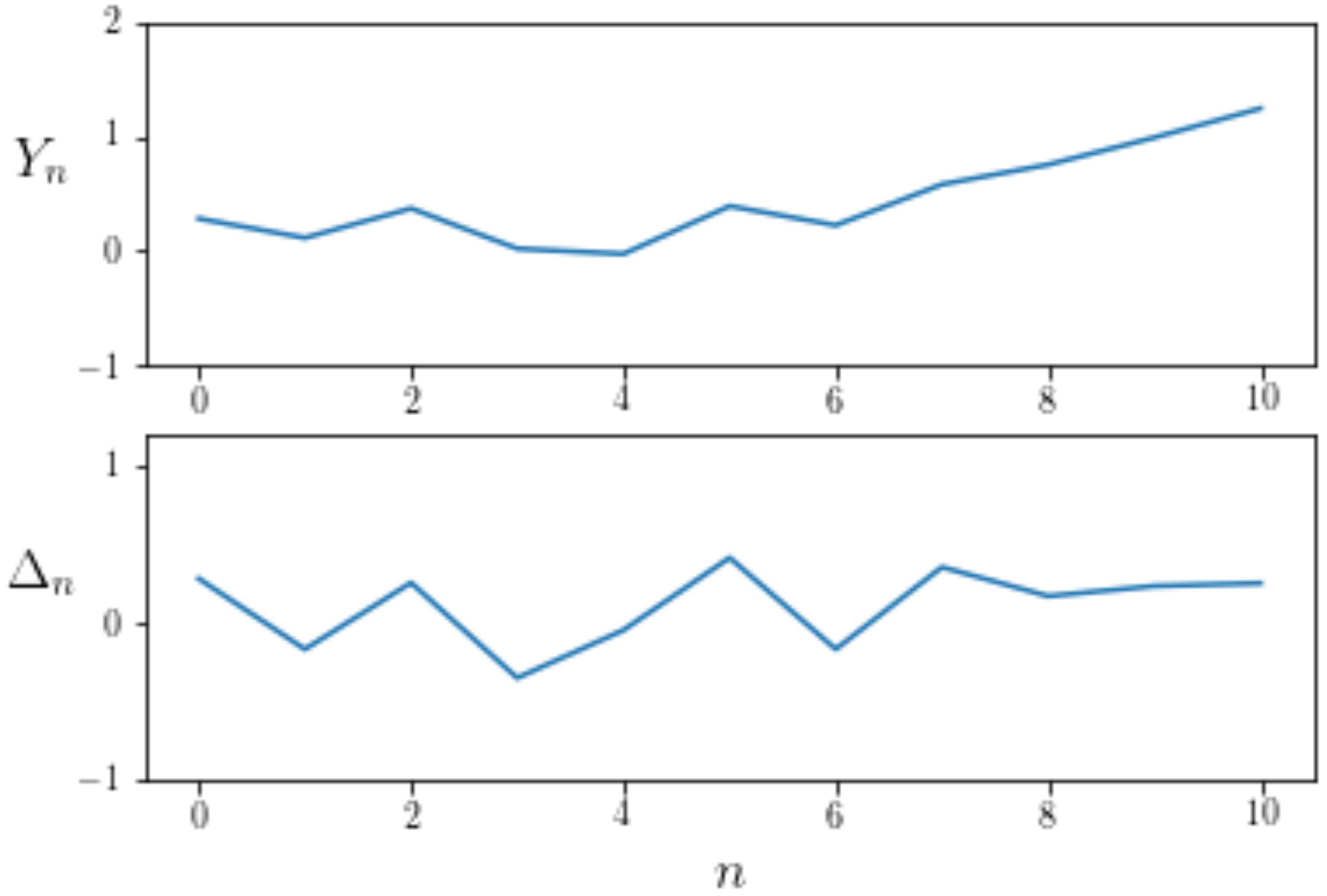
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

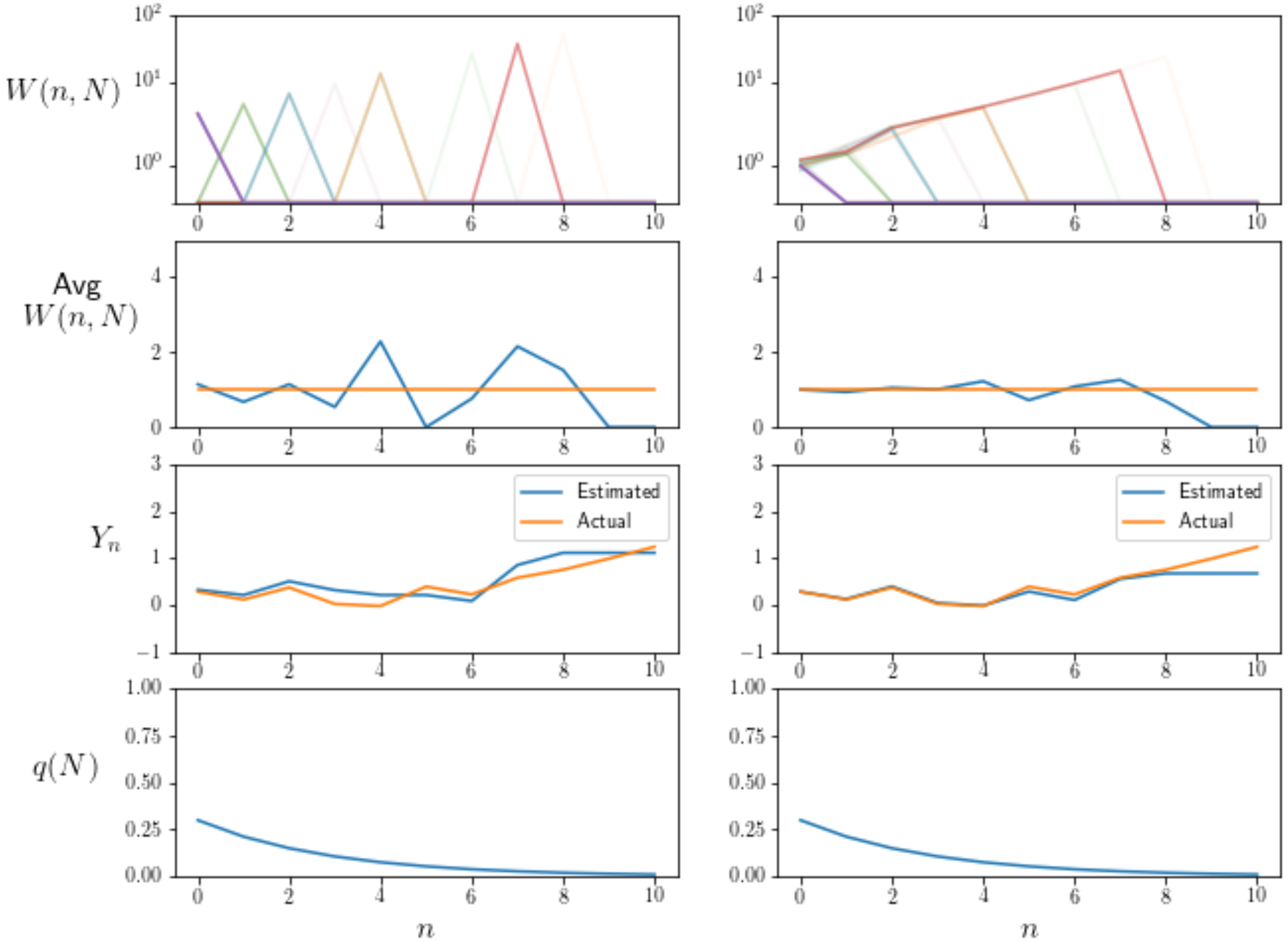


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



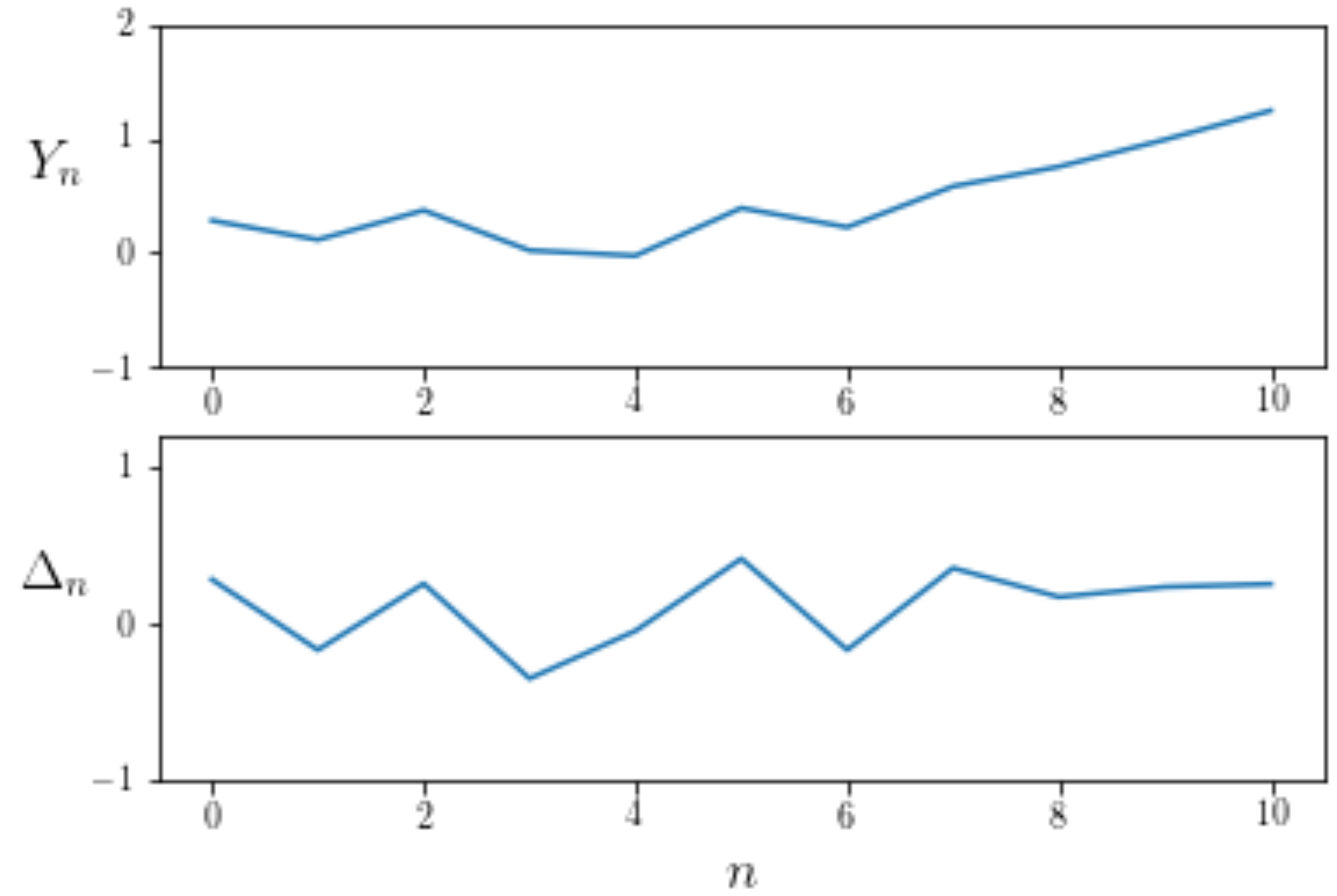
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

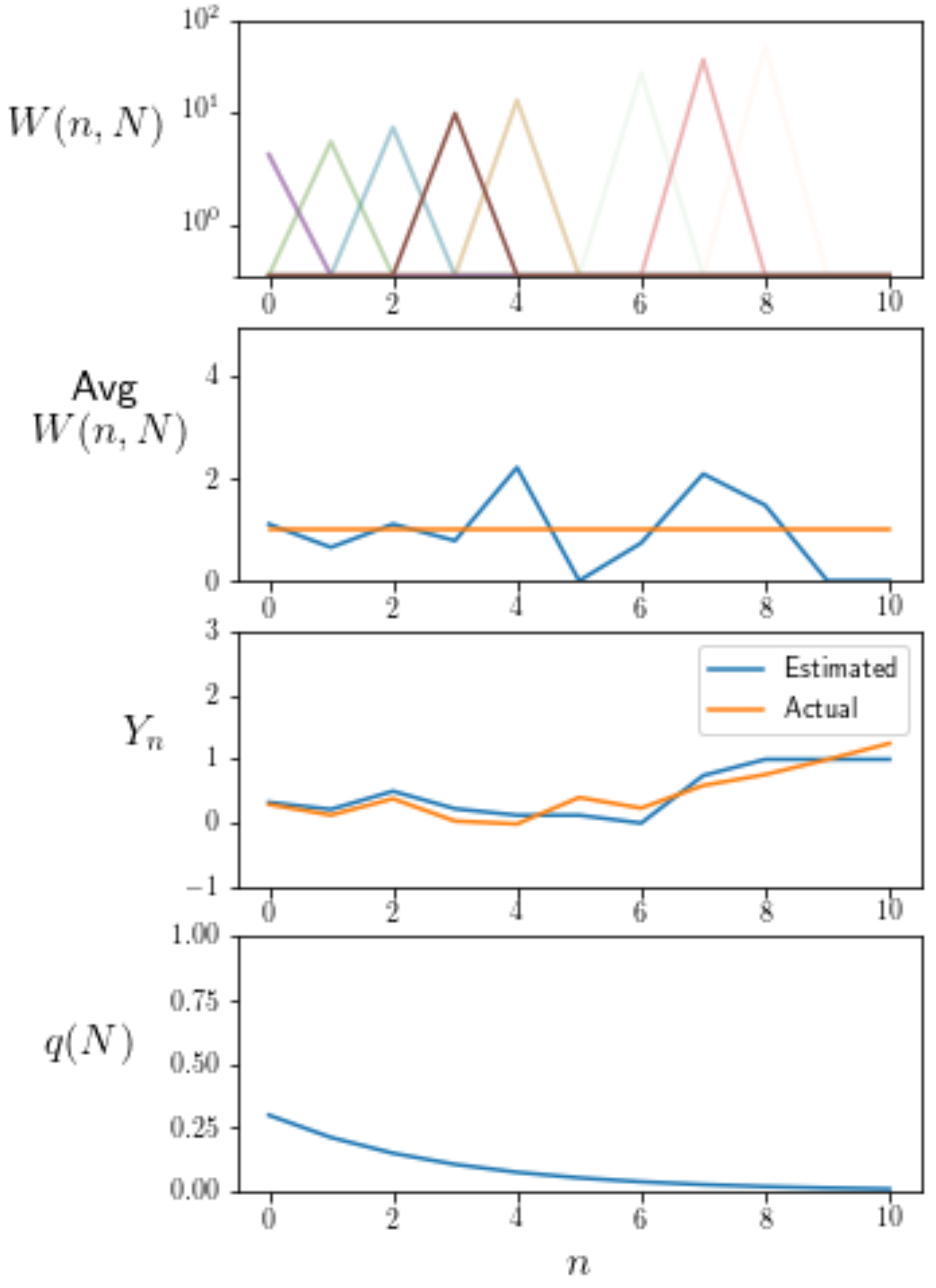
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



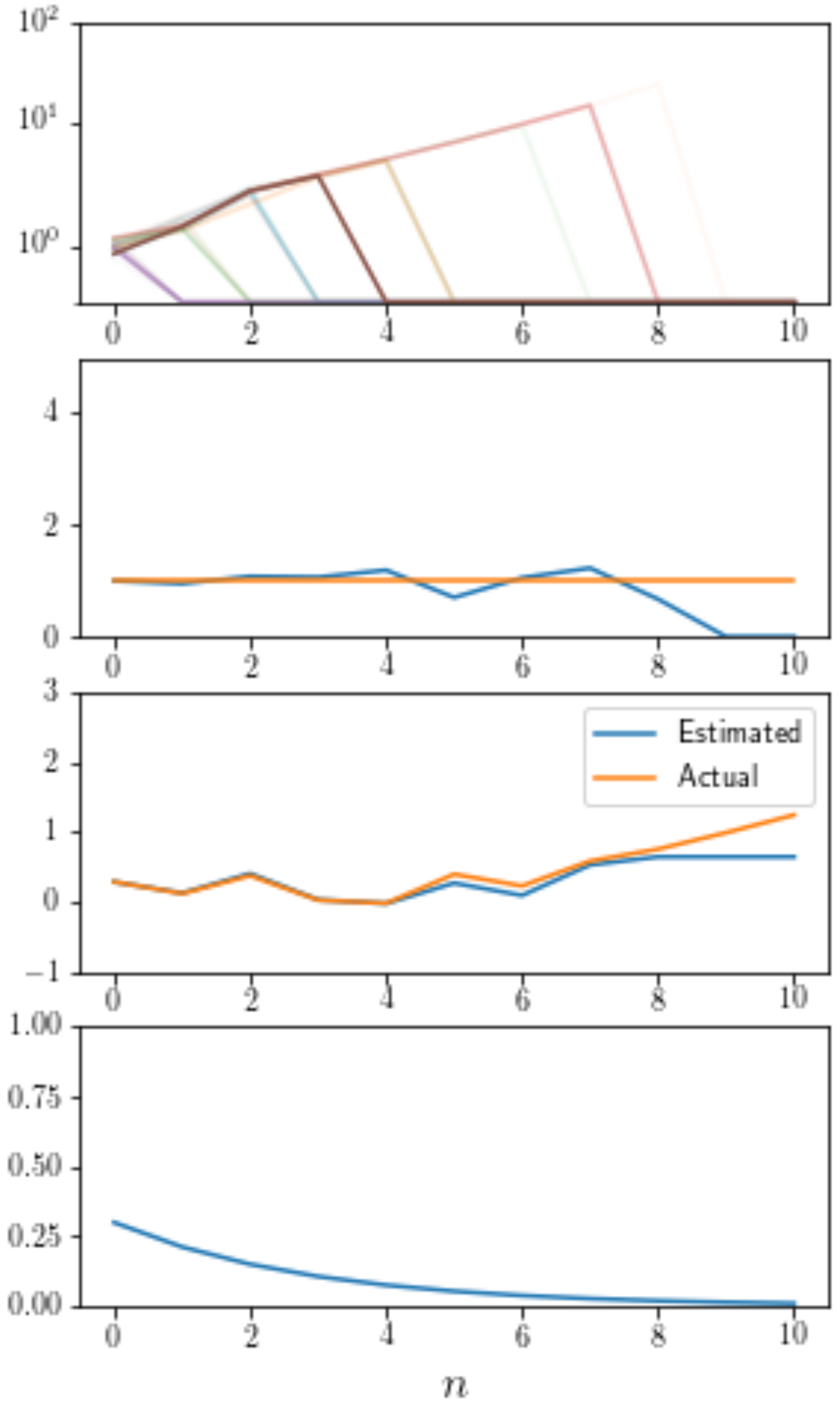
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



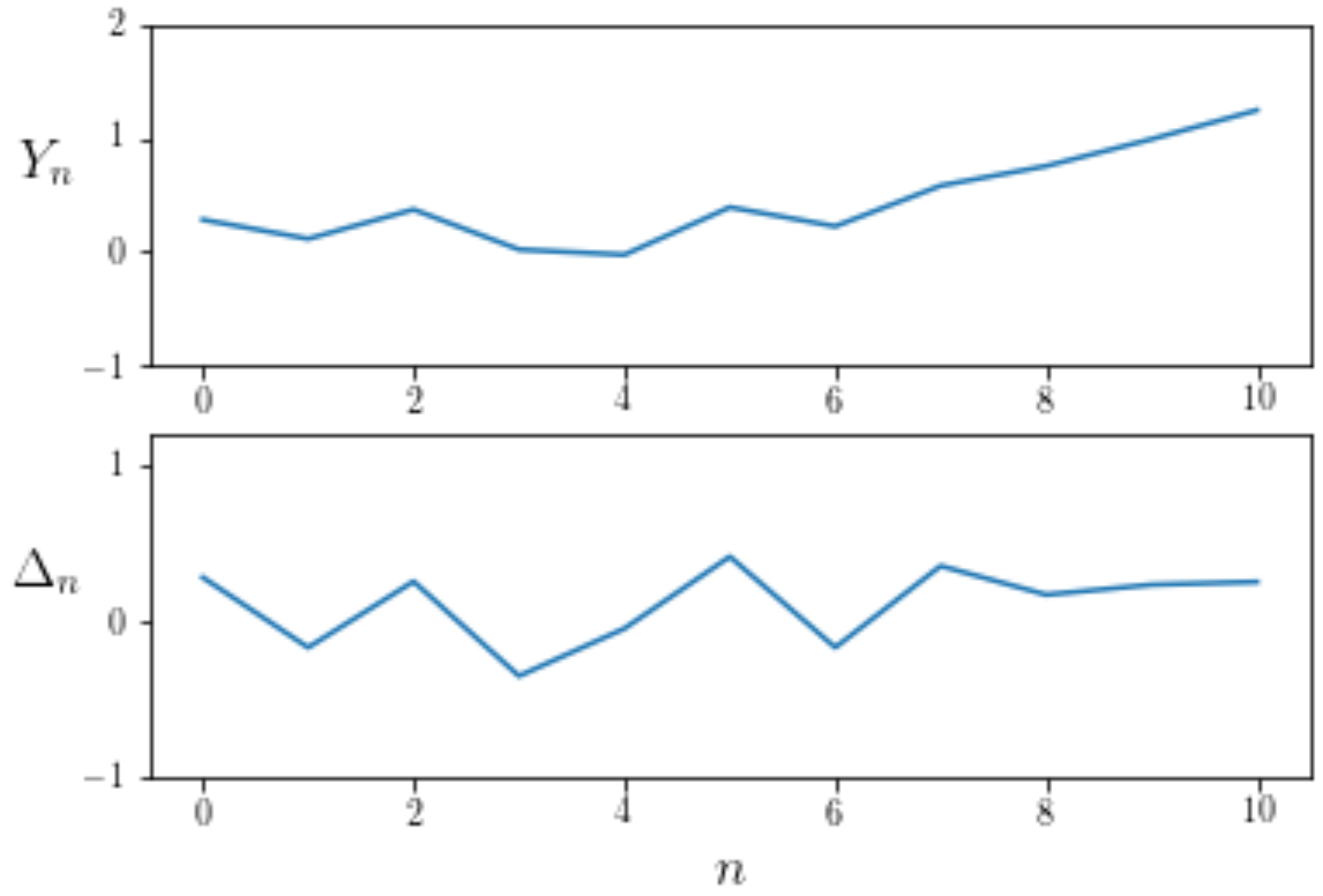
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

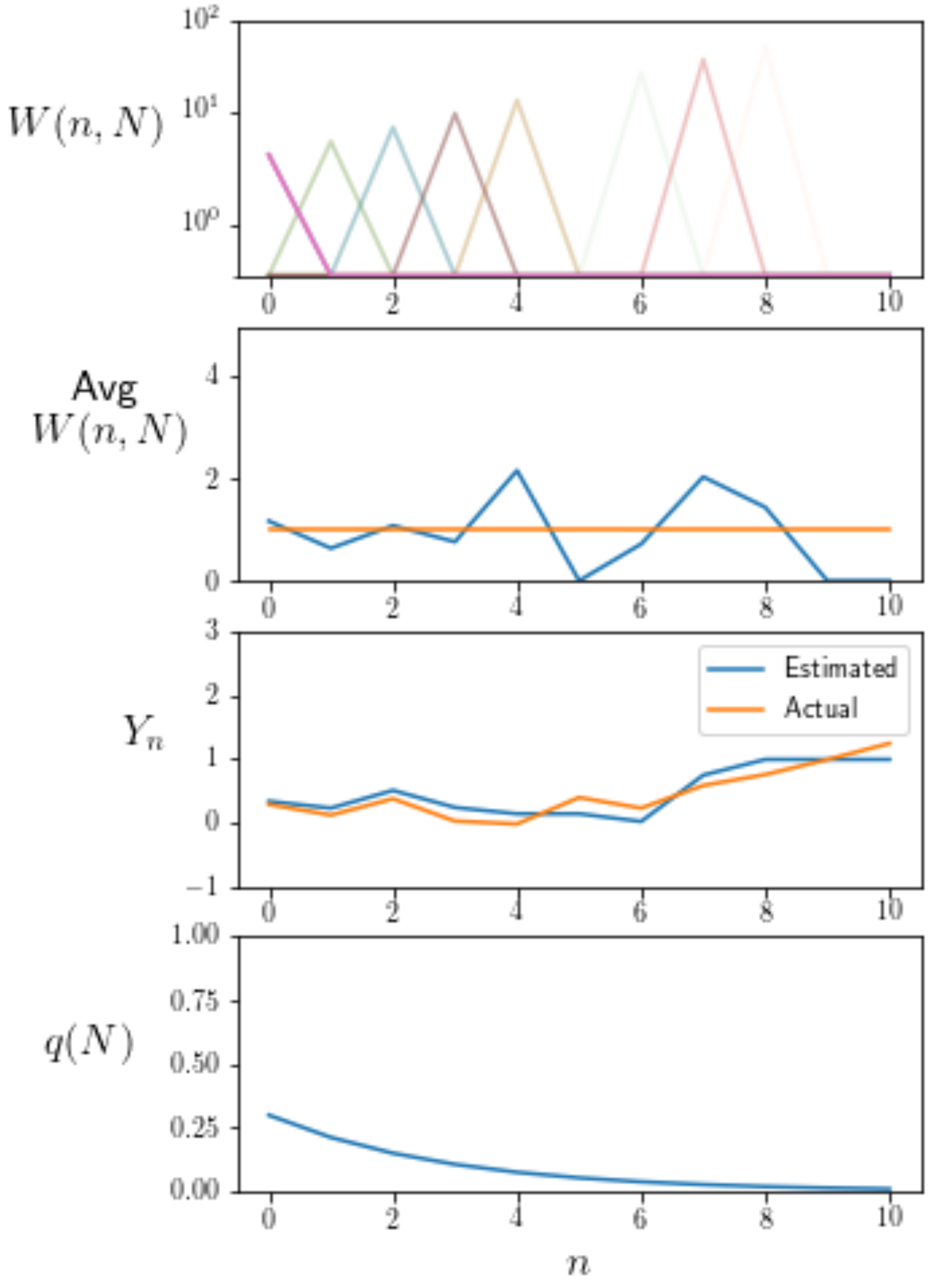
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



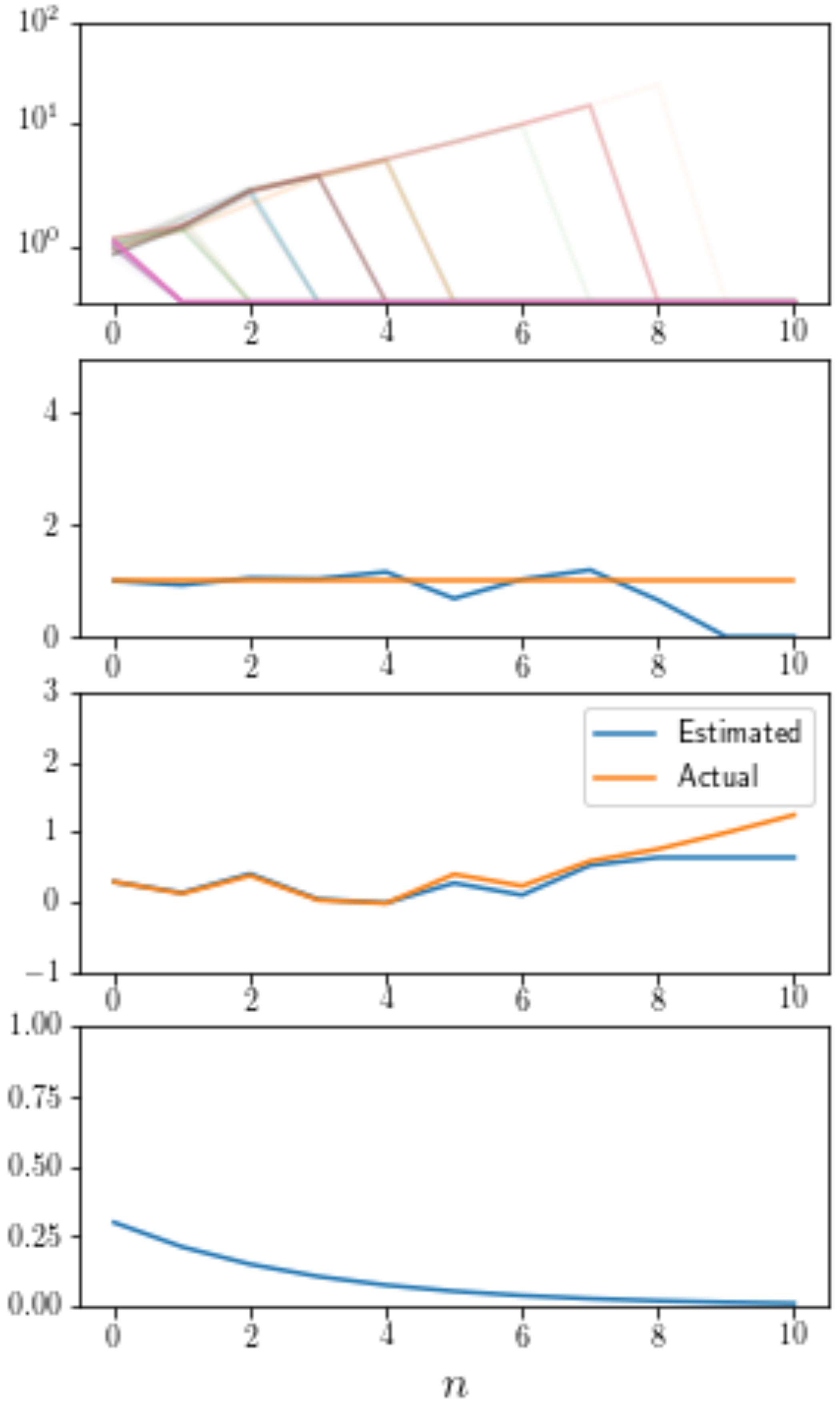
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



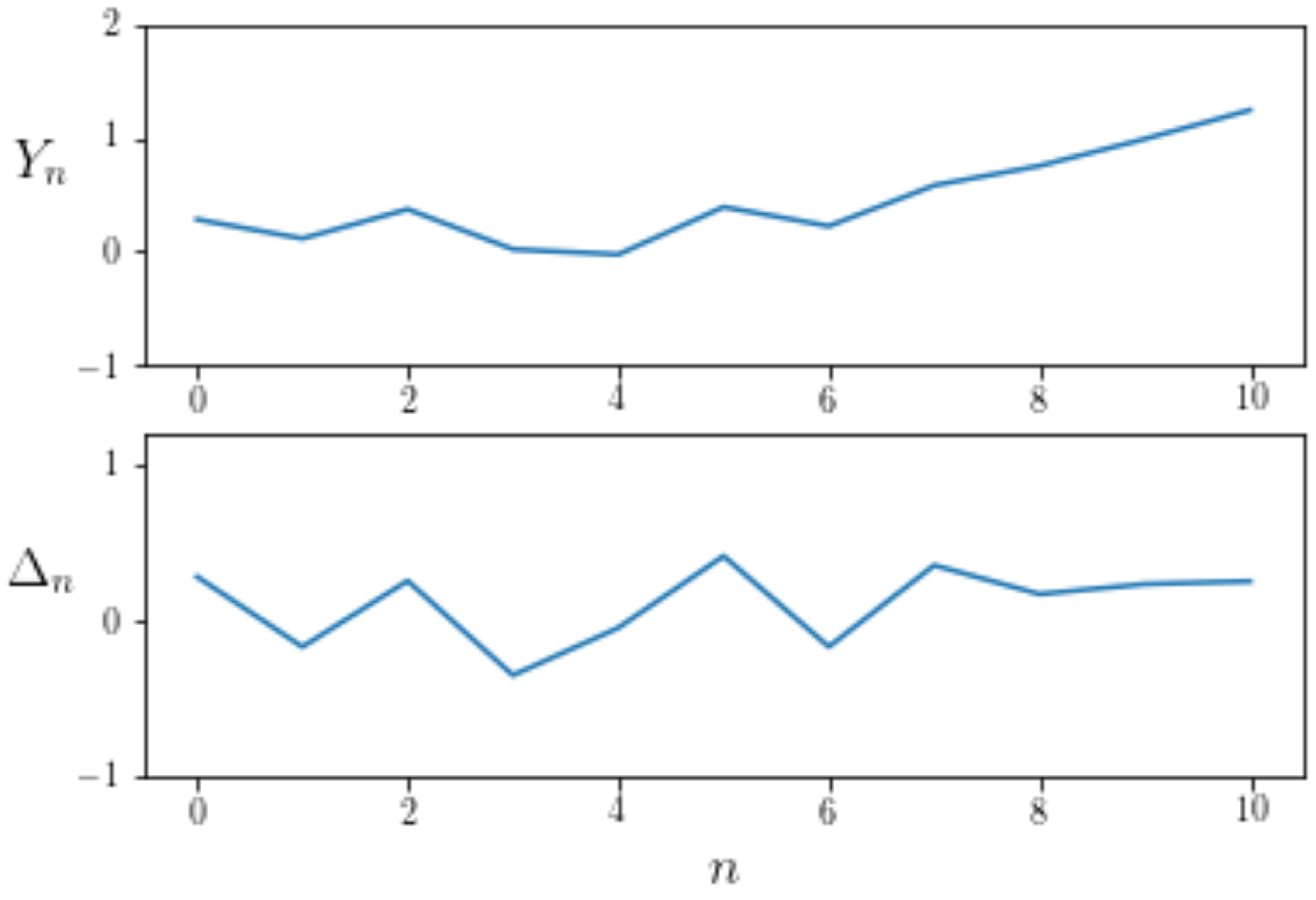
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

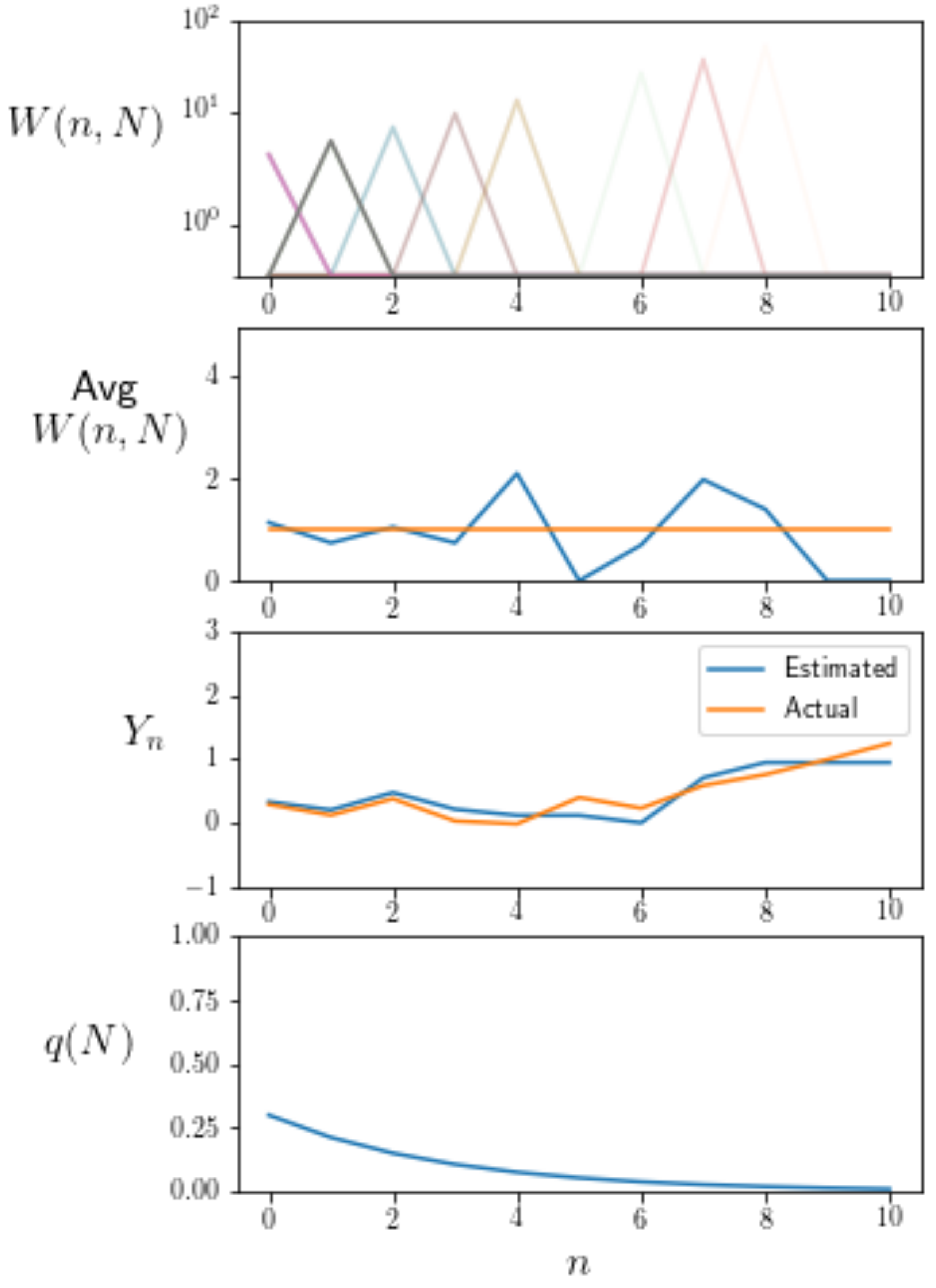
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



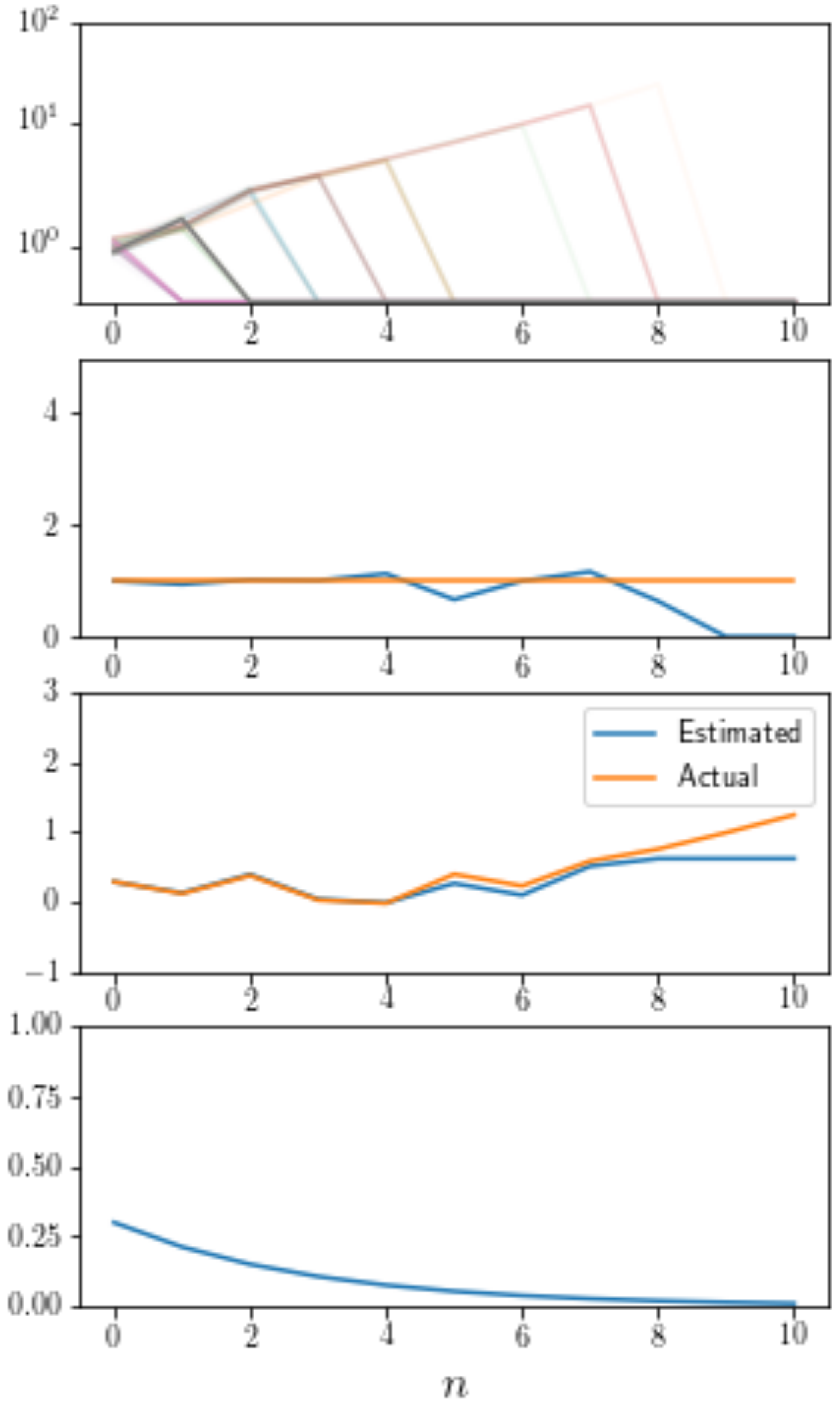
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



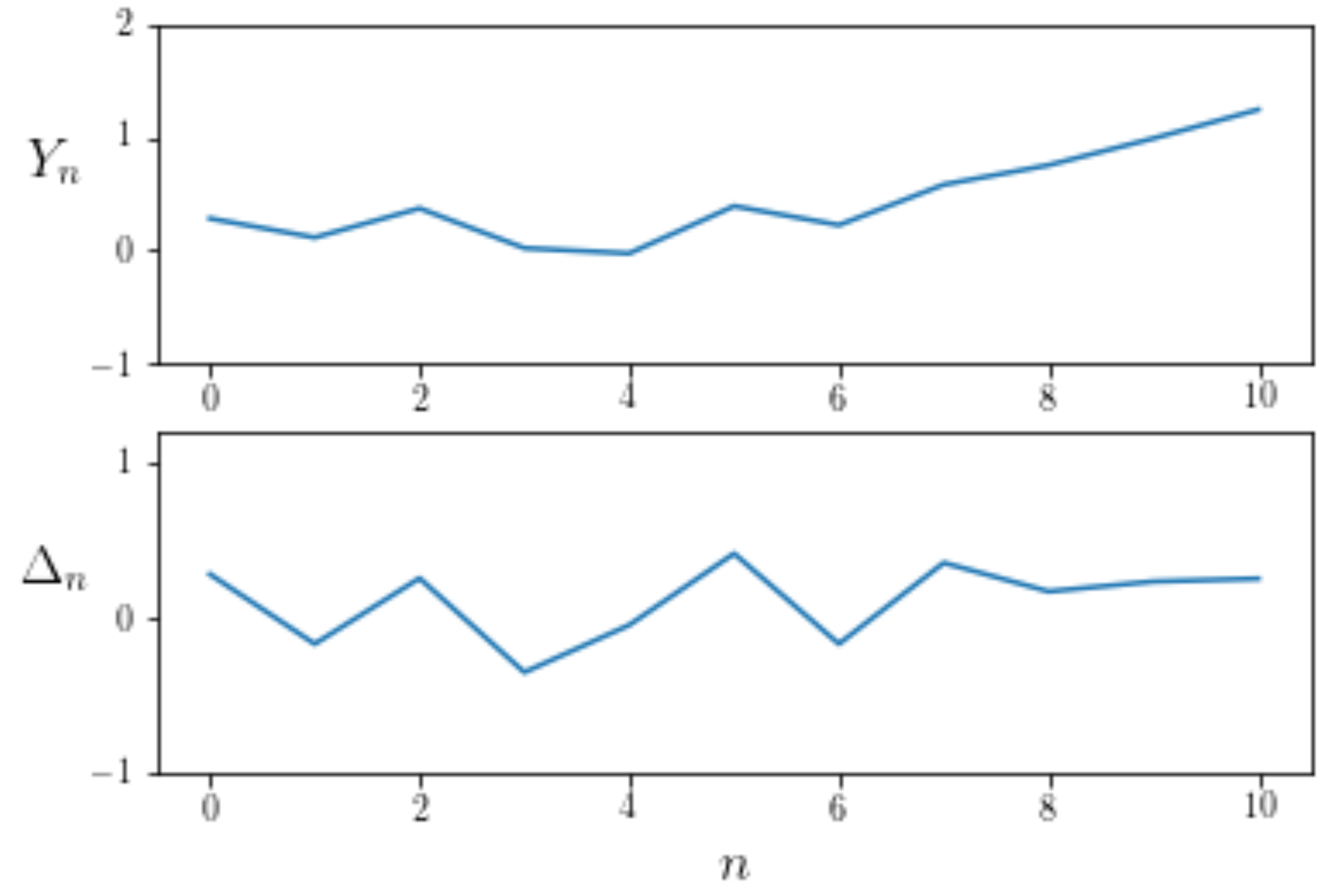
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

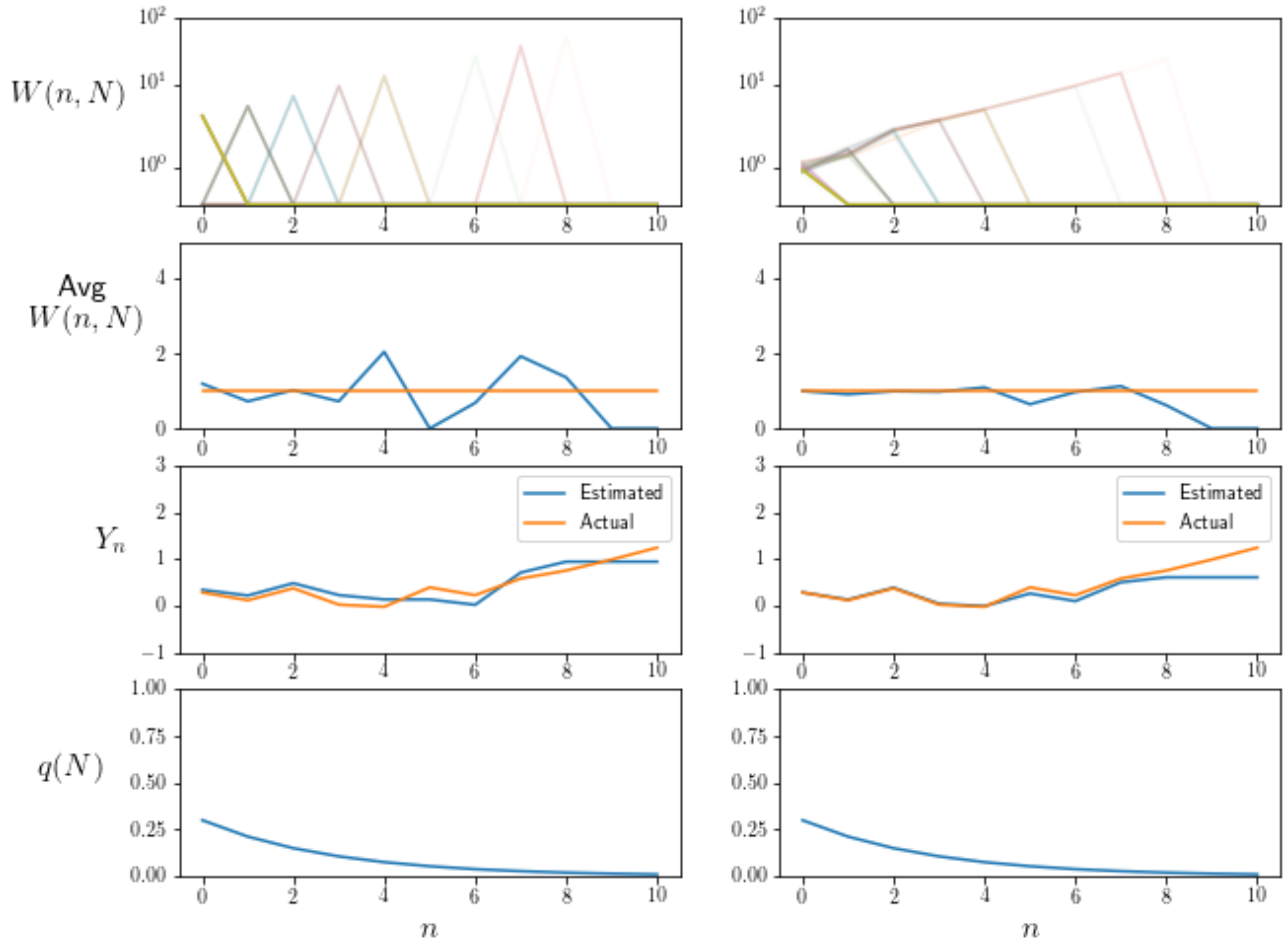


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



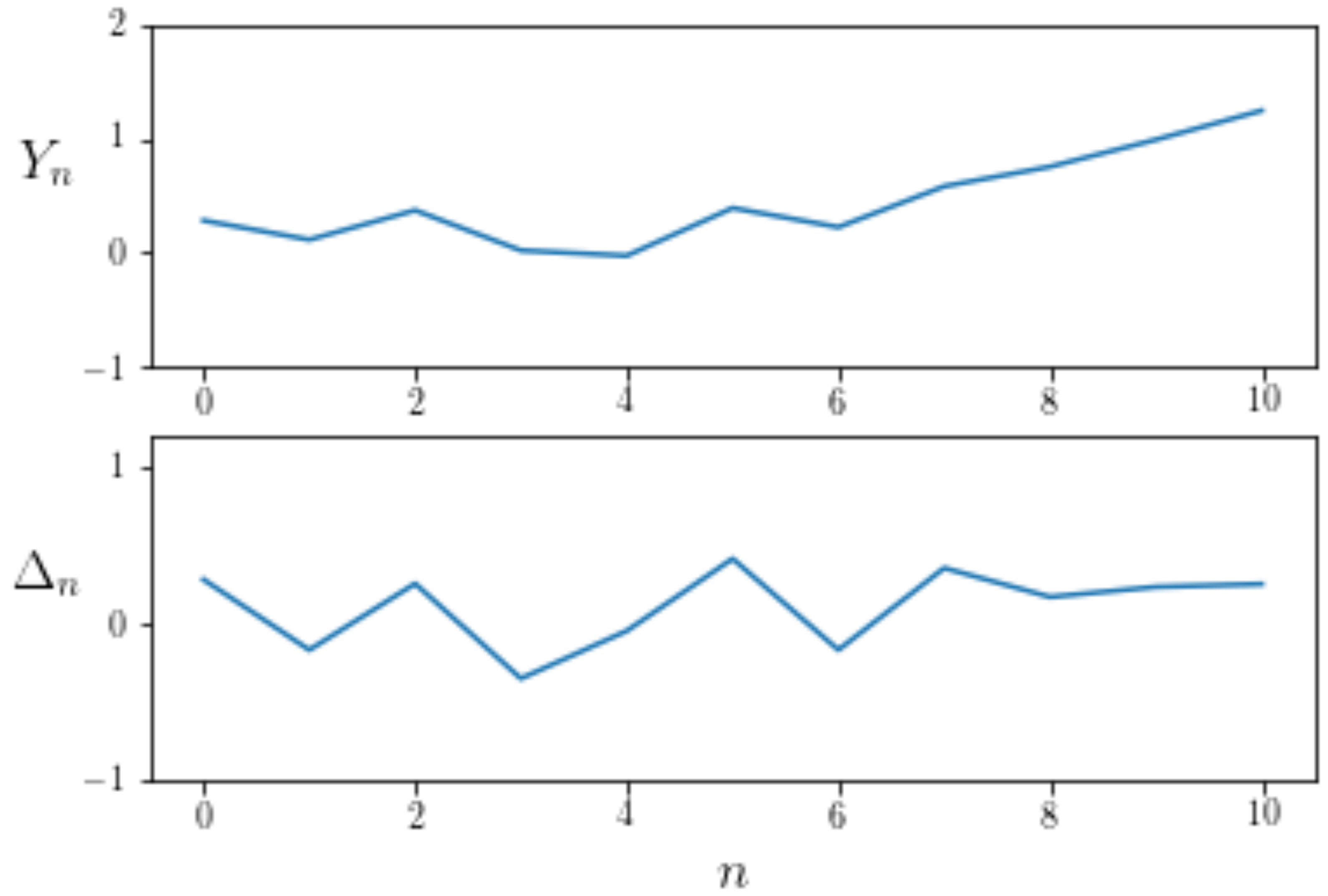
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

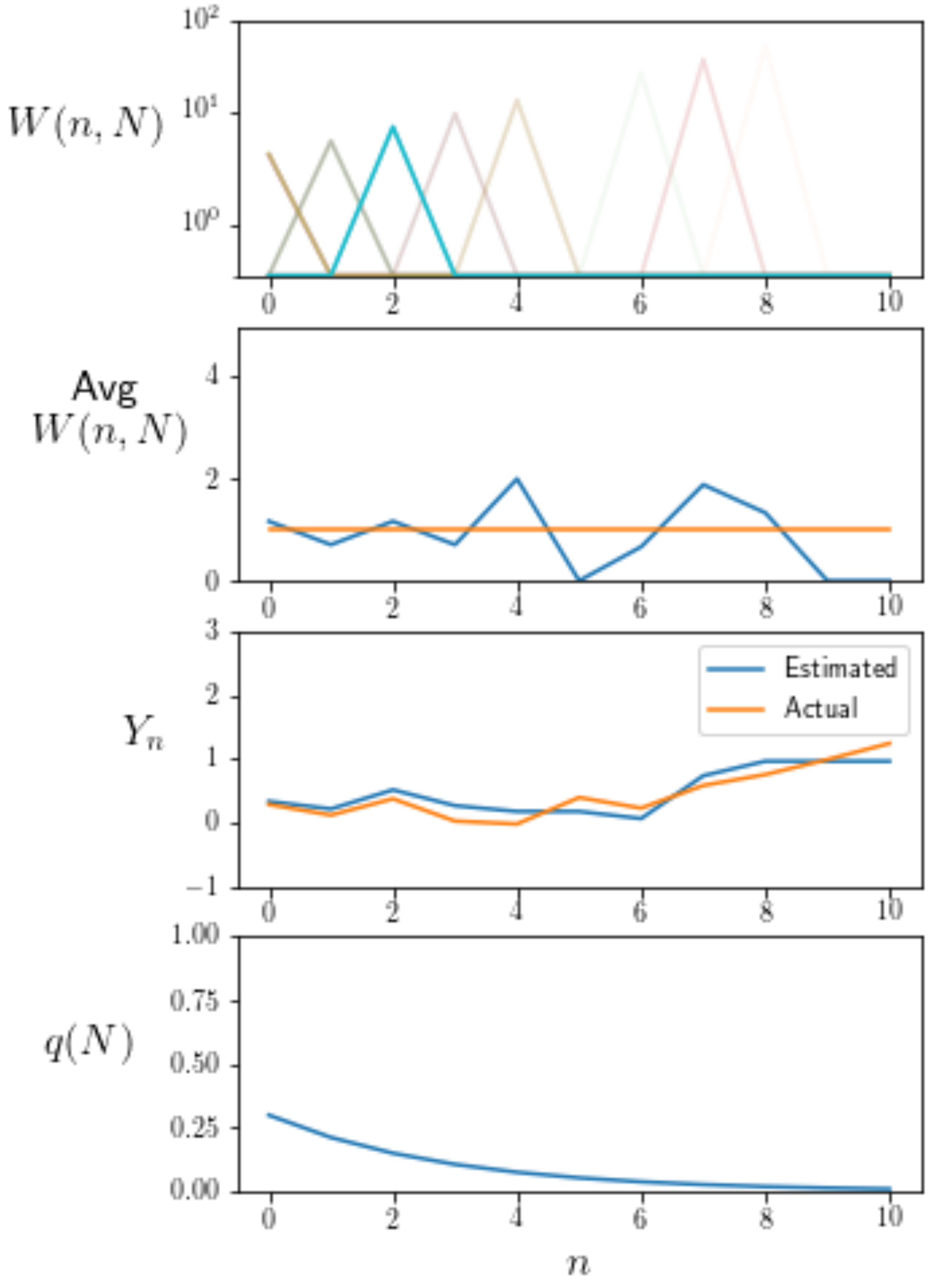
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



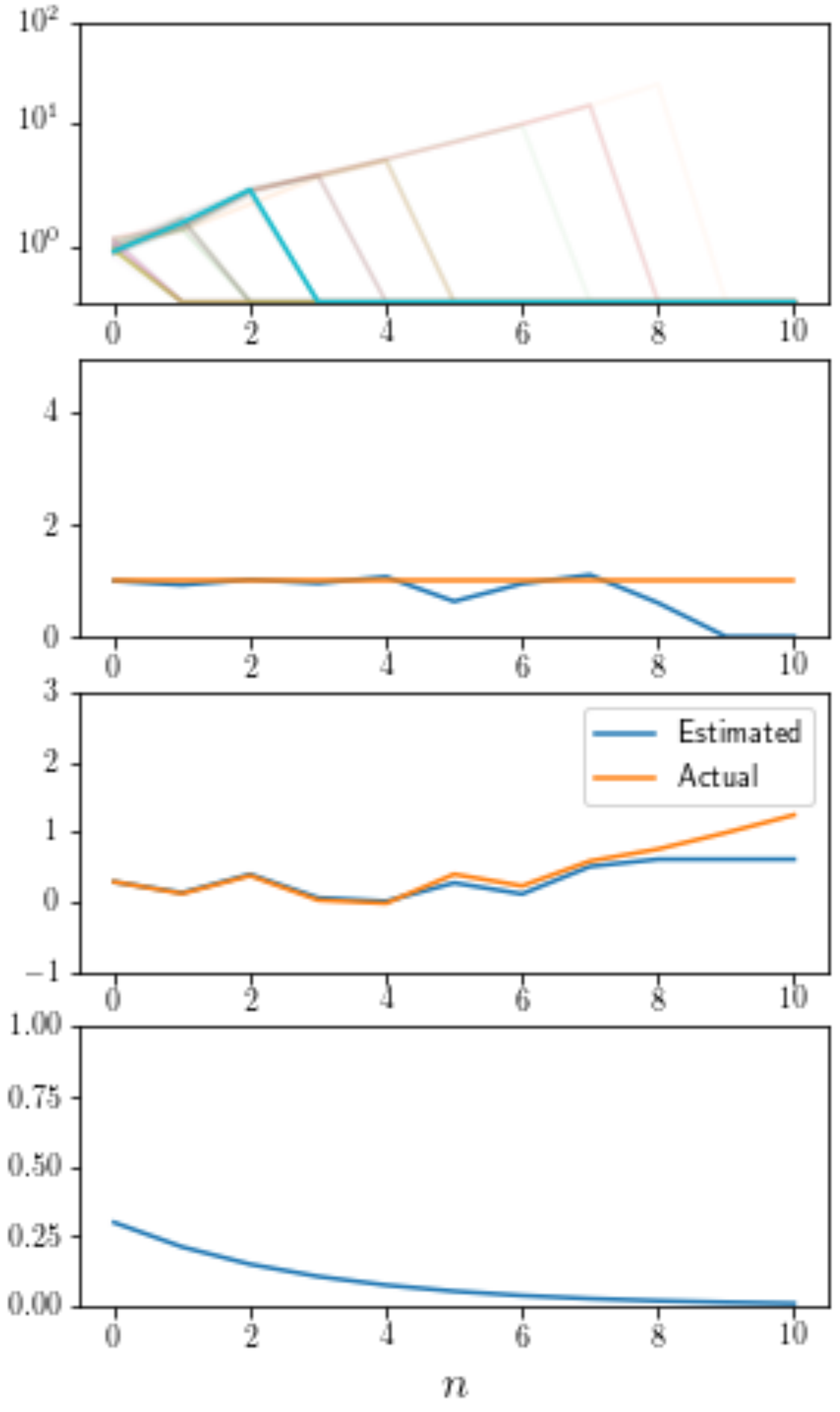
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



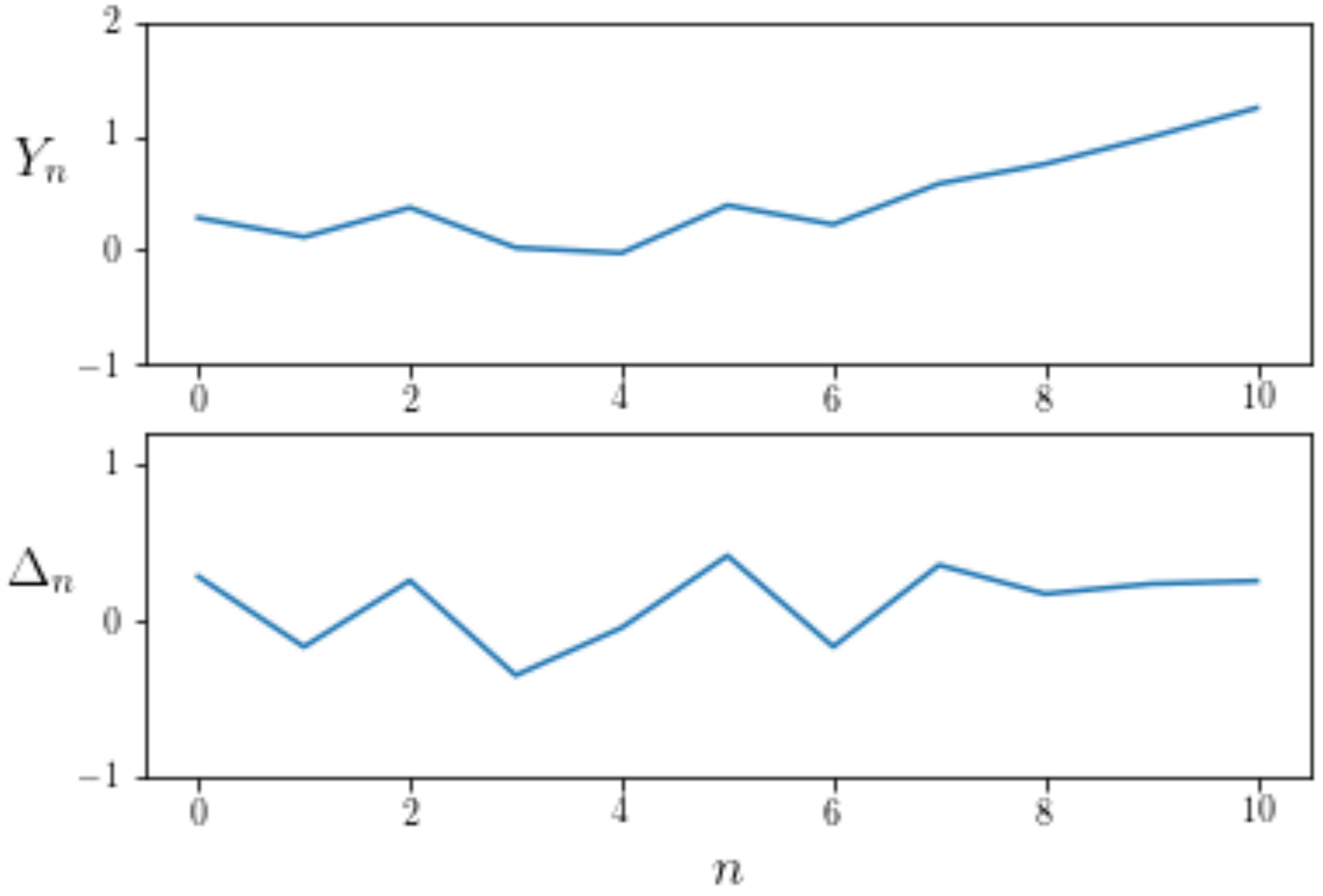
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

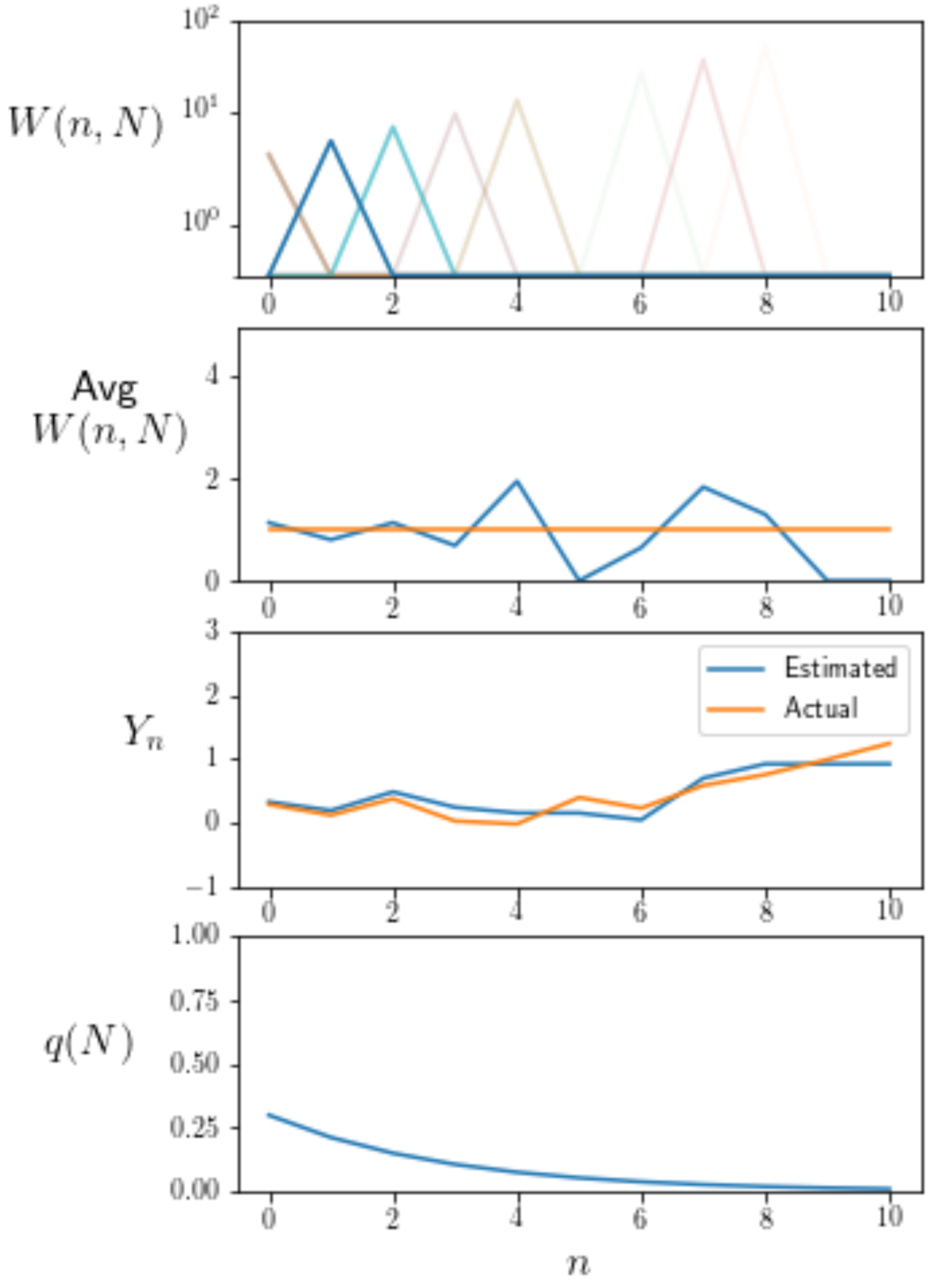
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



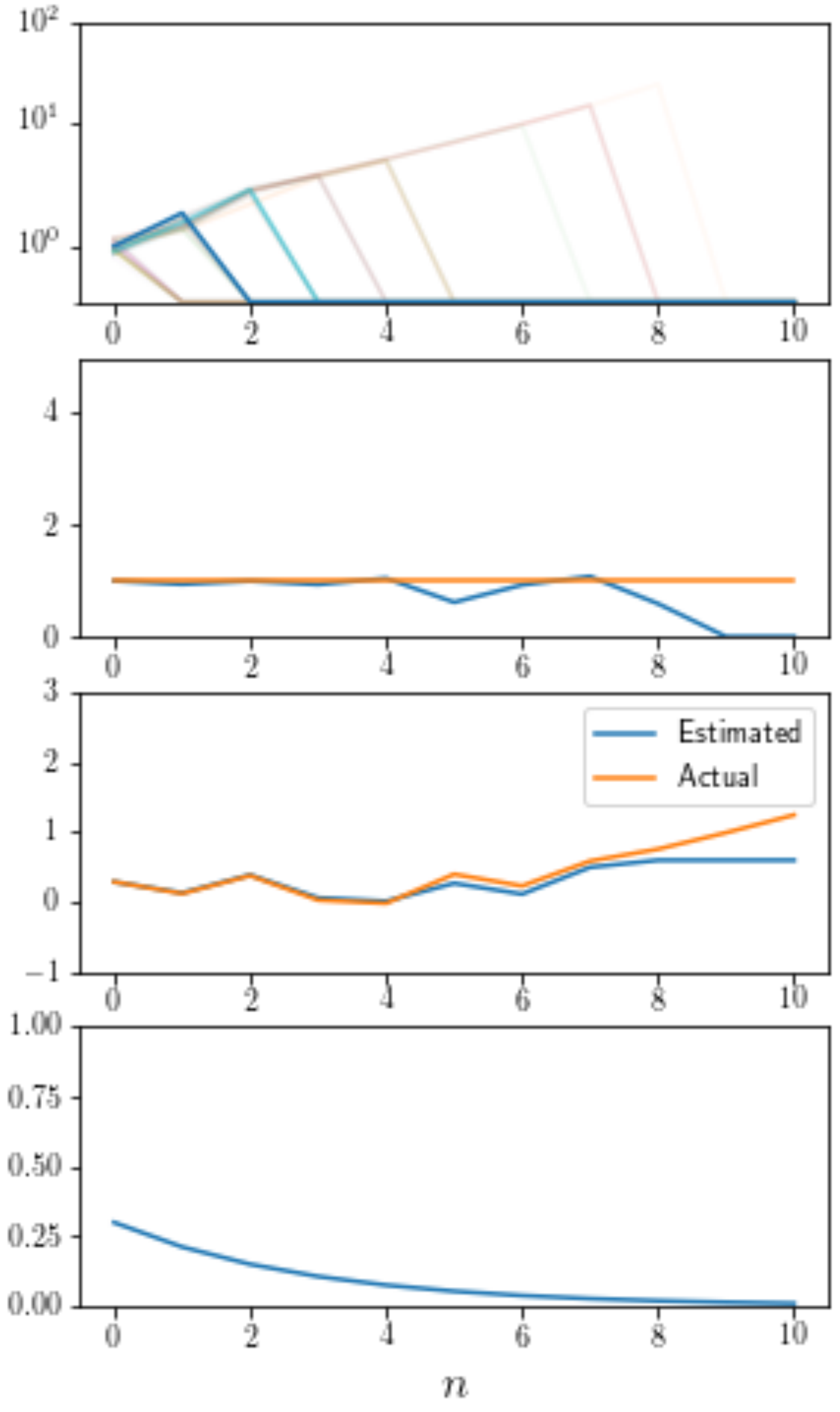
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



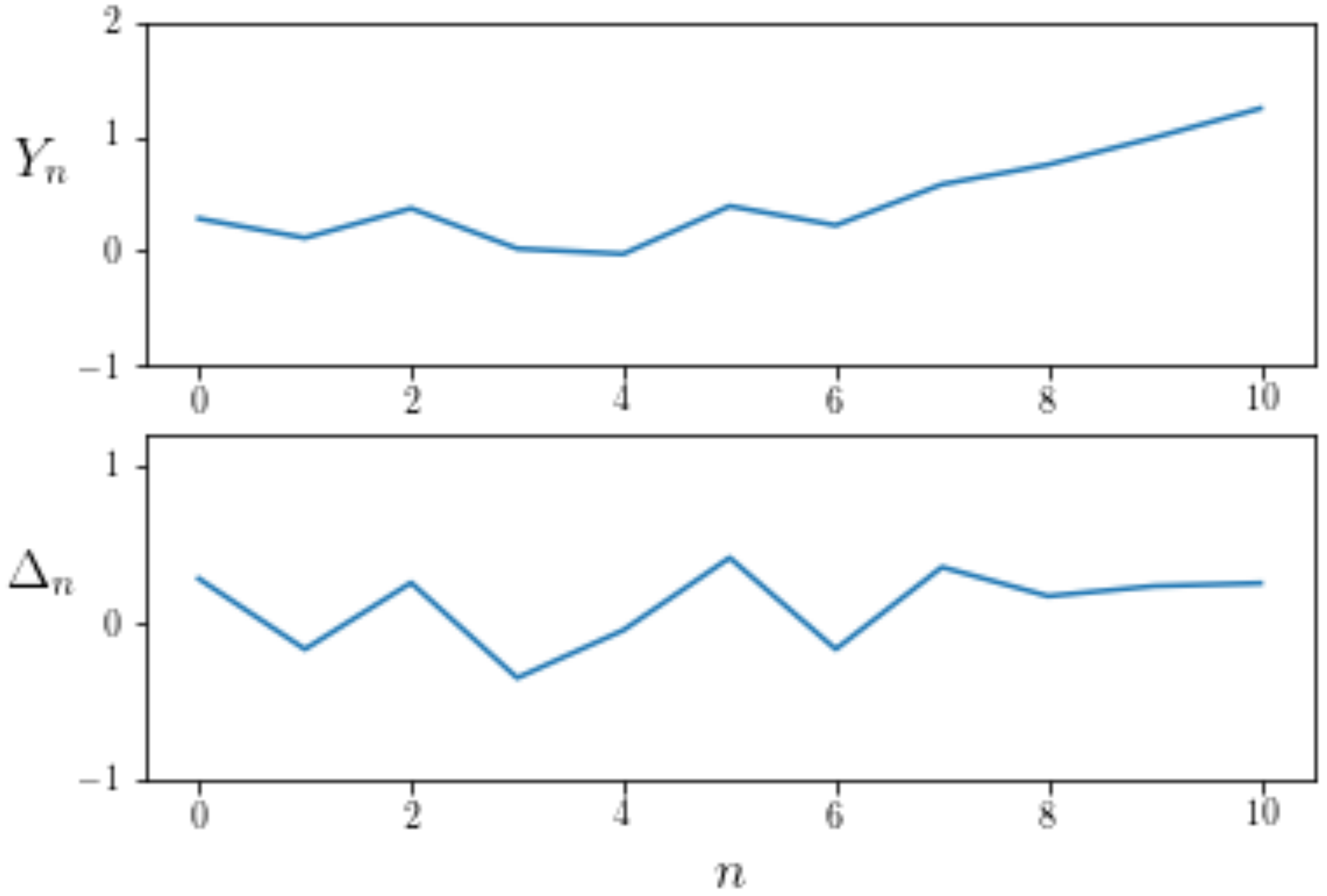
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

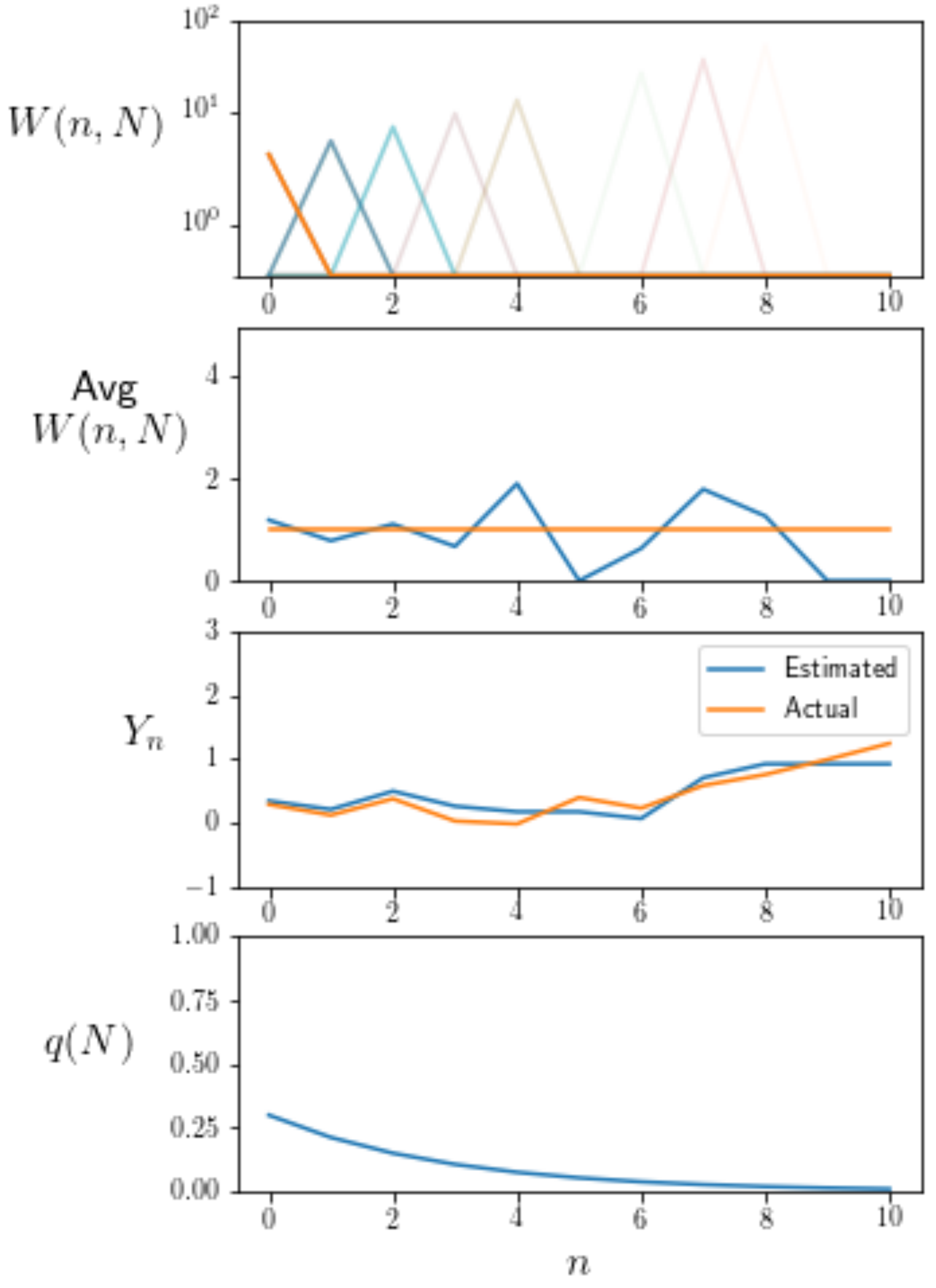


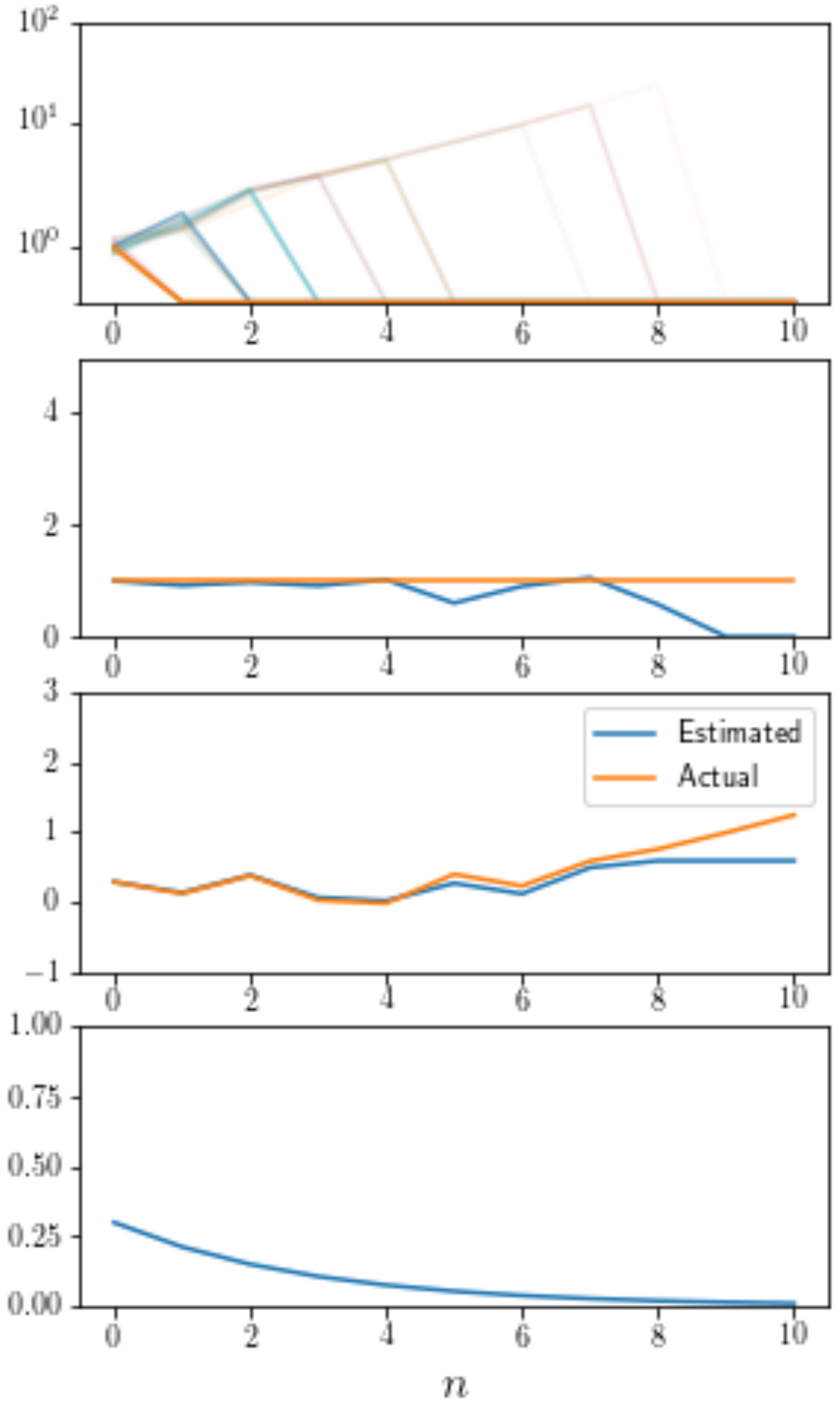
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$





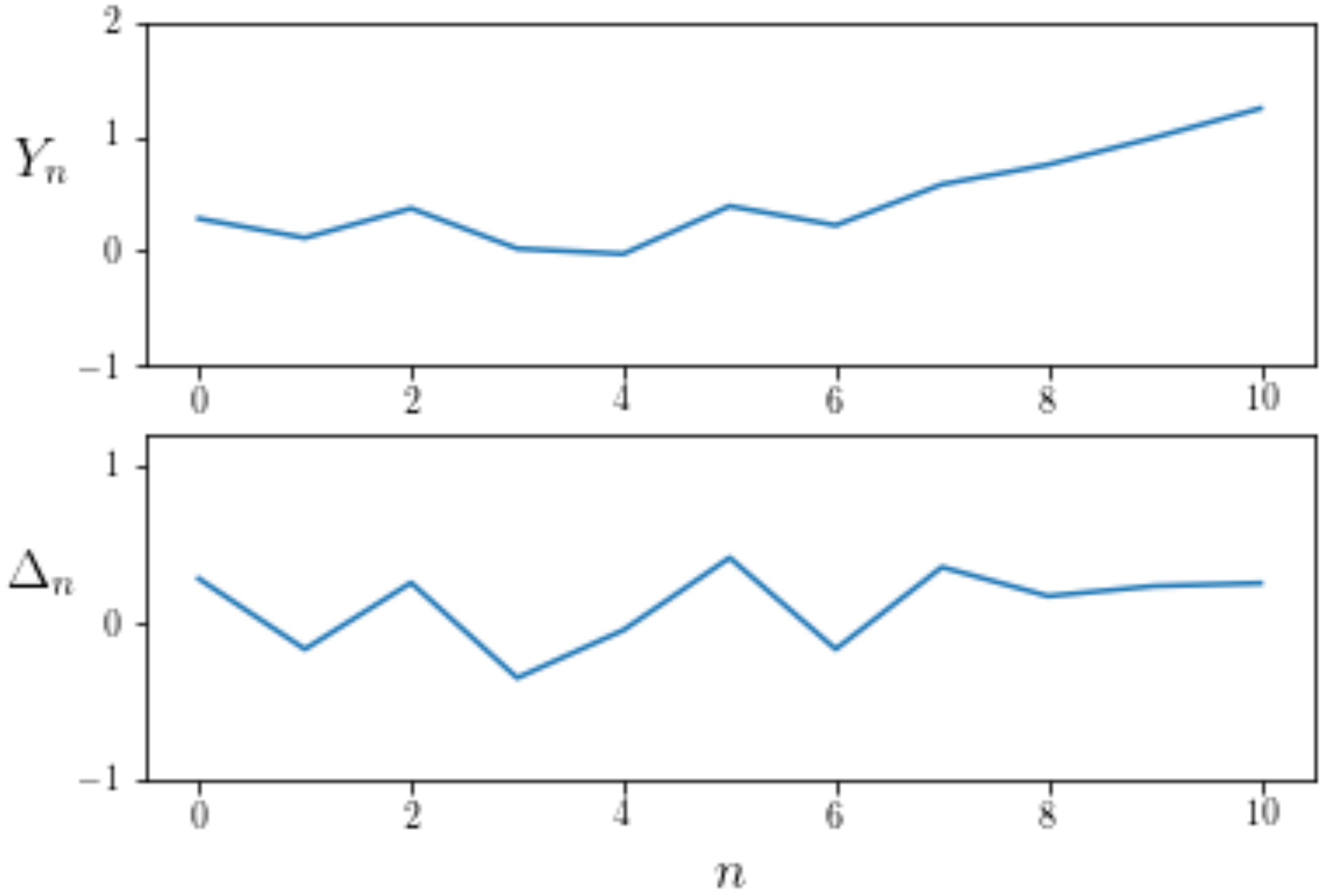
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

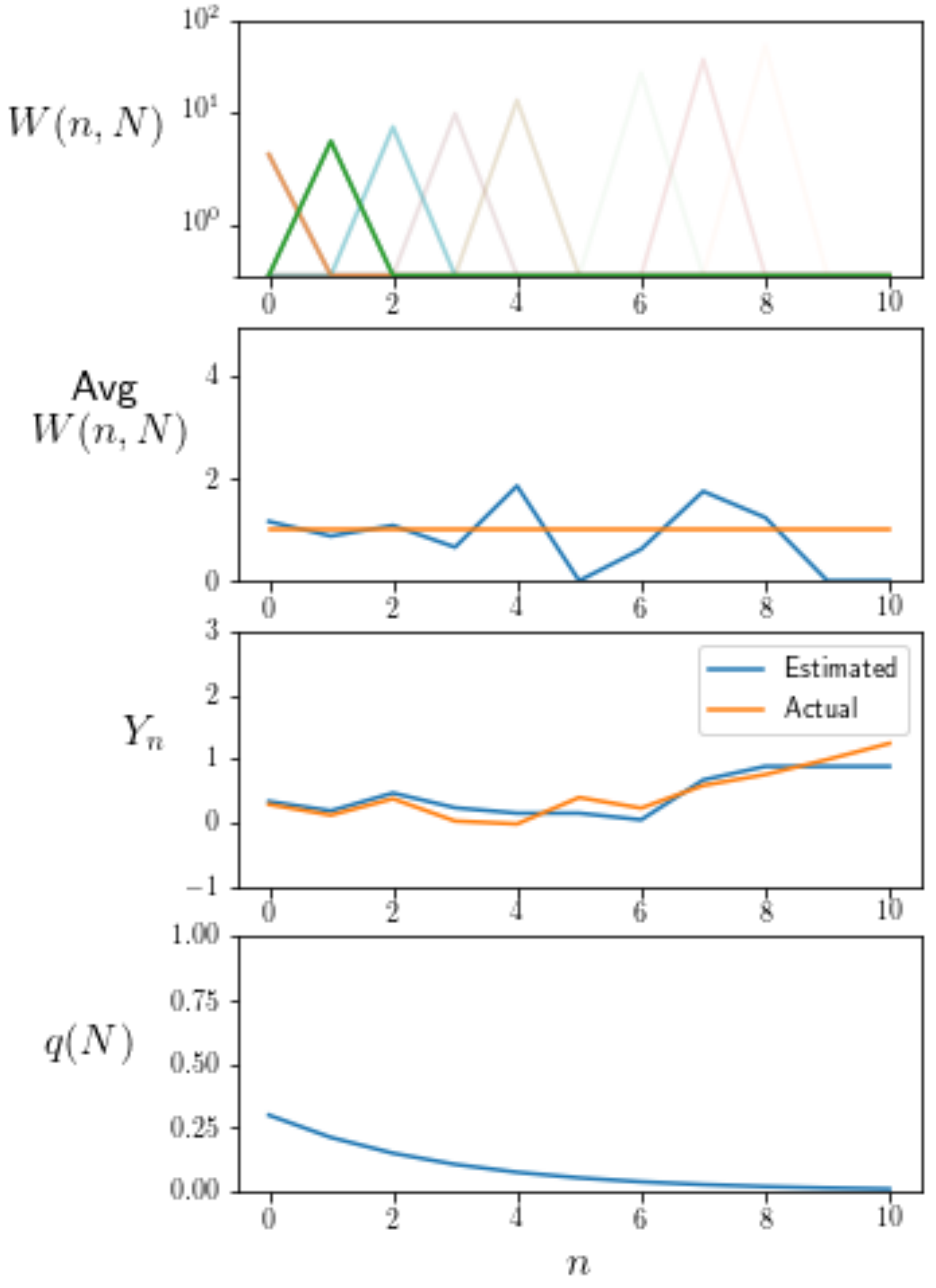
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



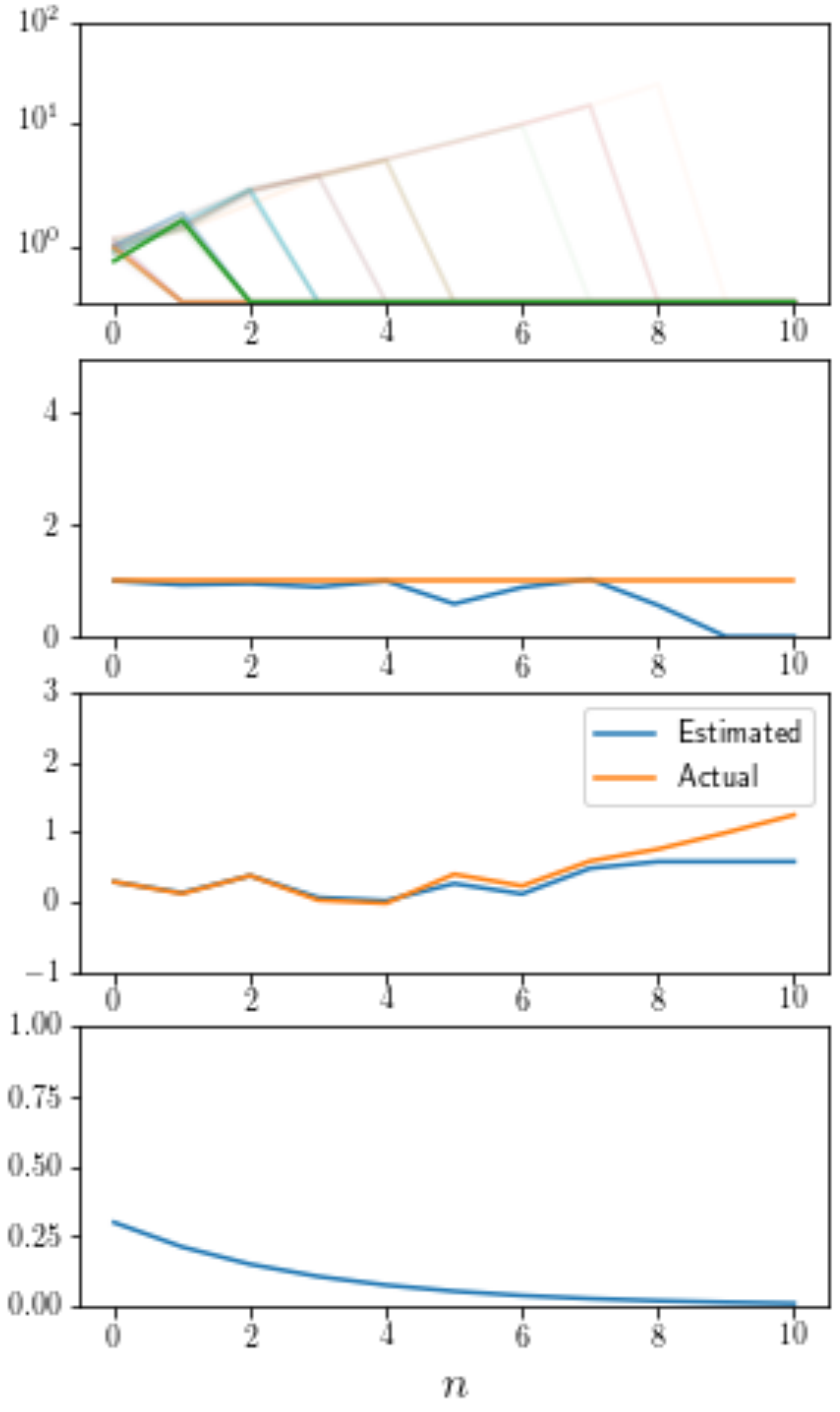
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



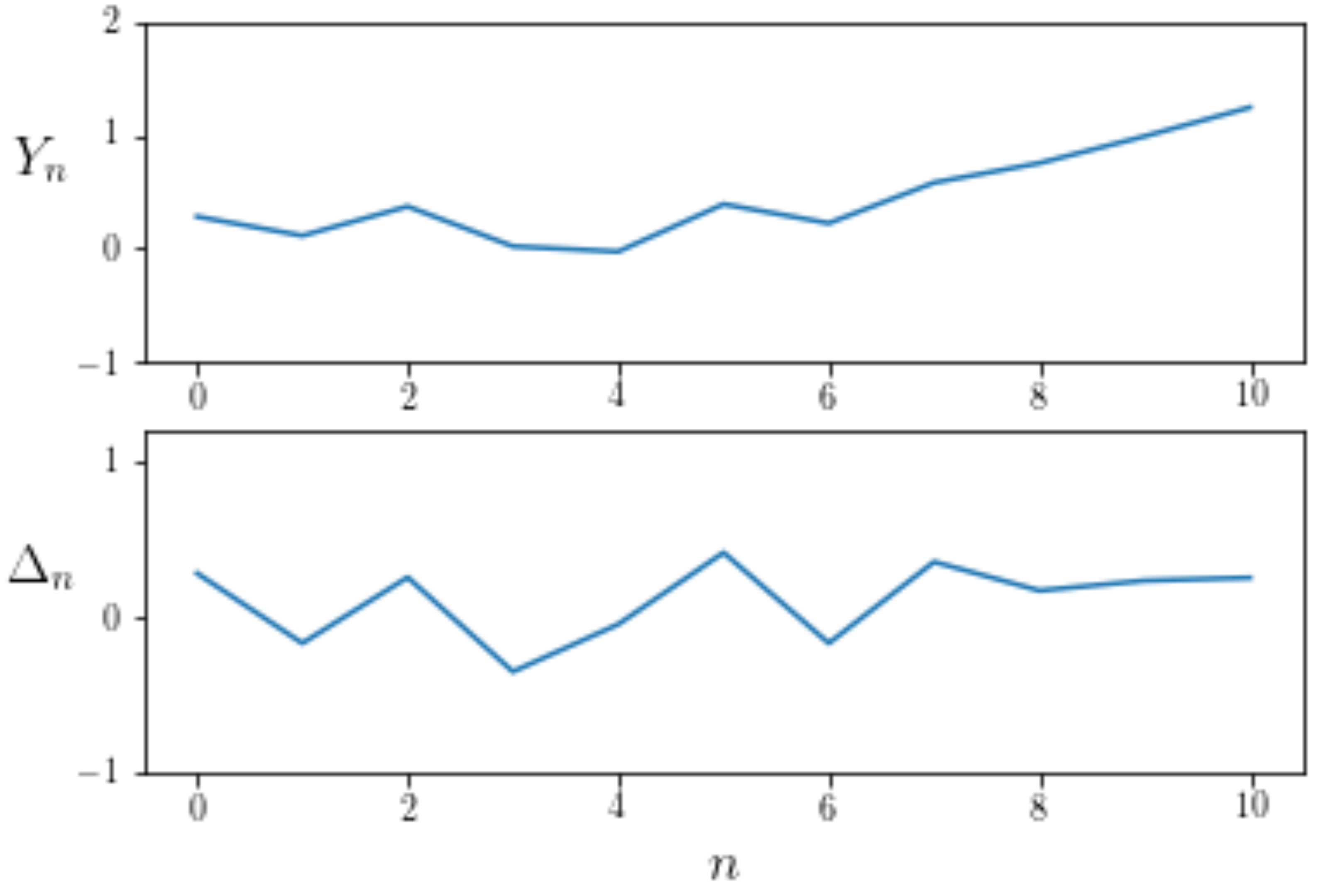
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

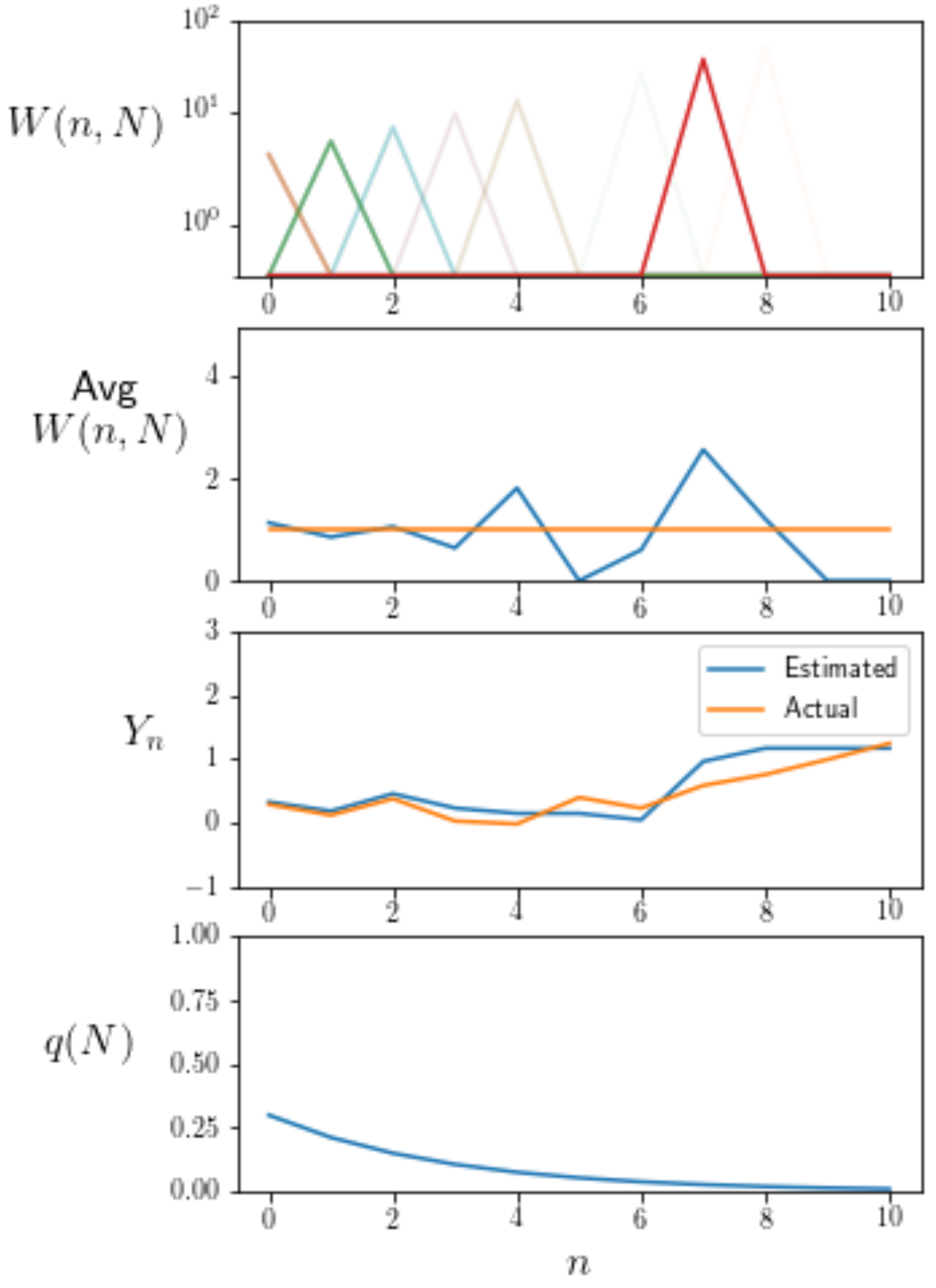
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



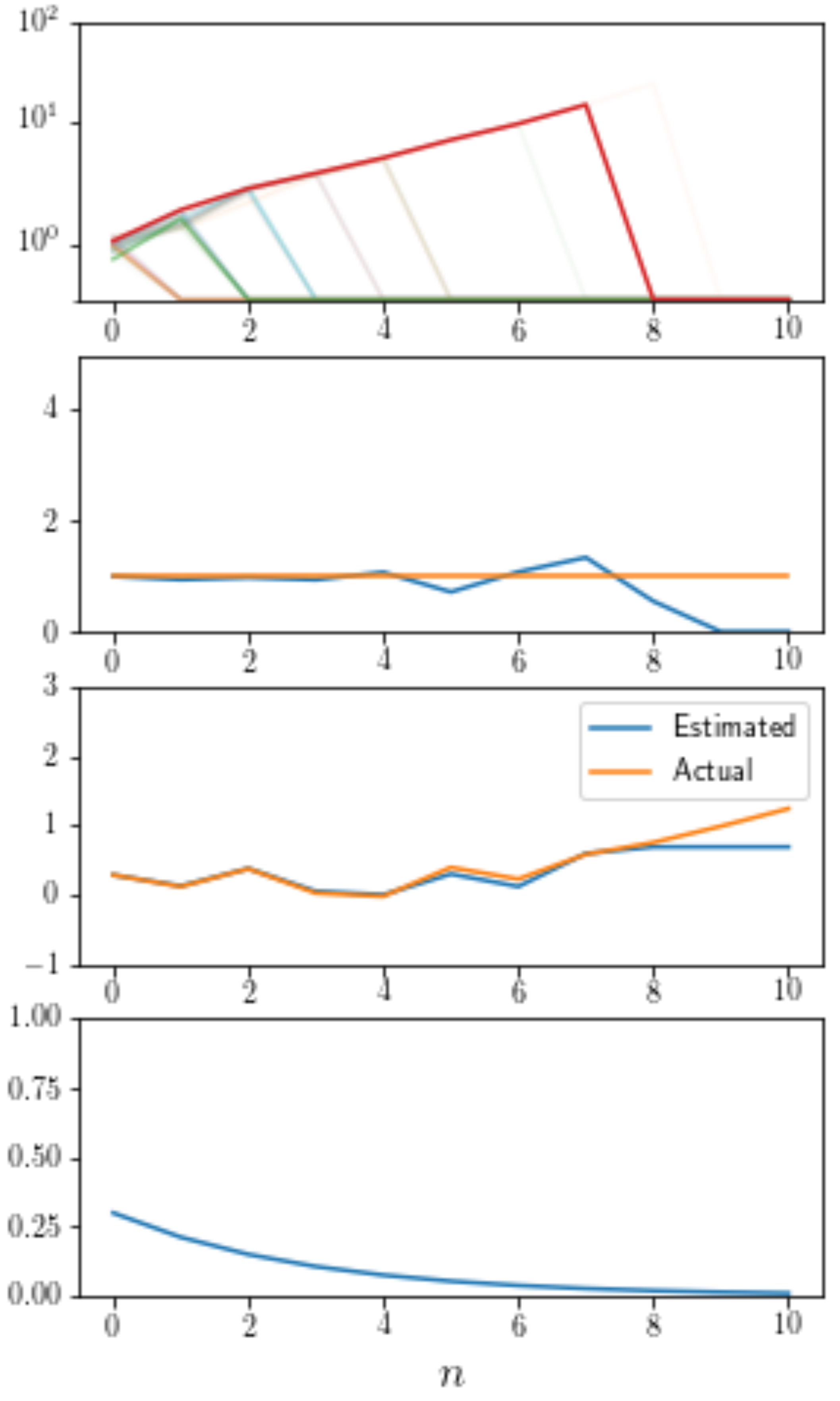
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



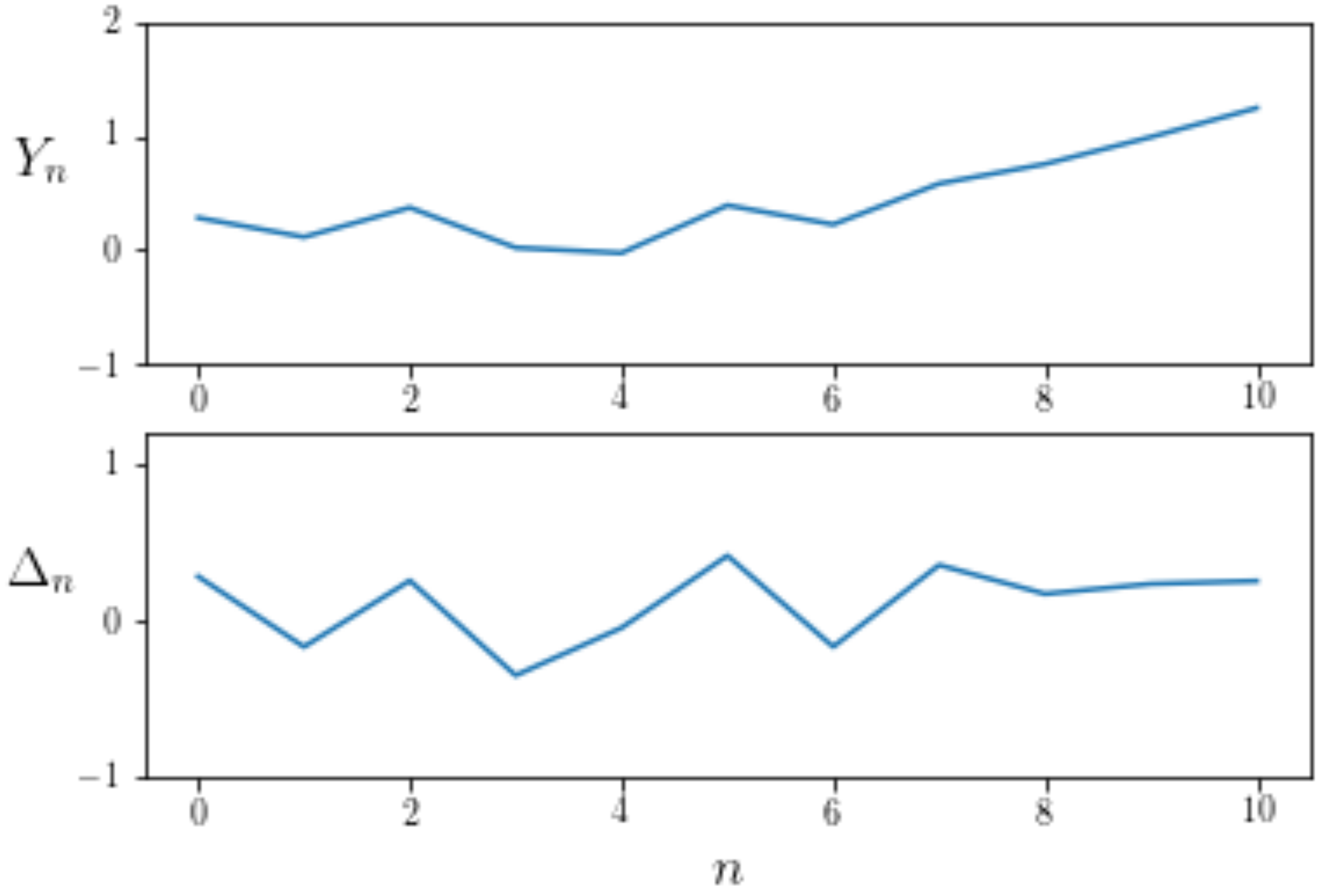
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

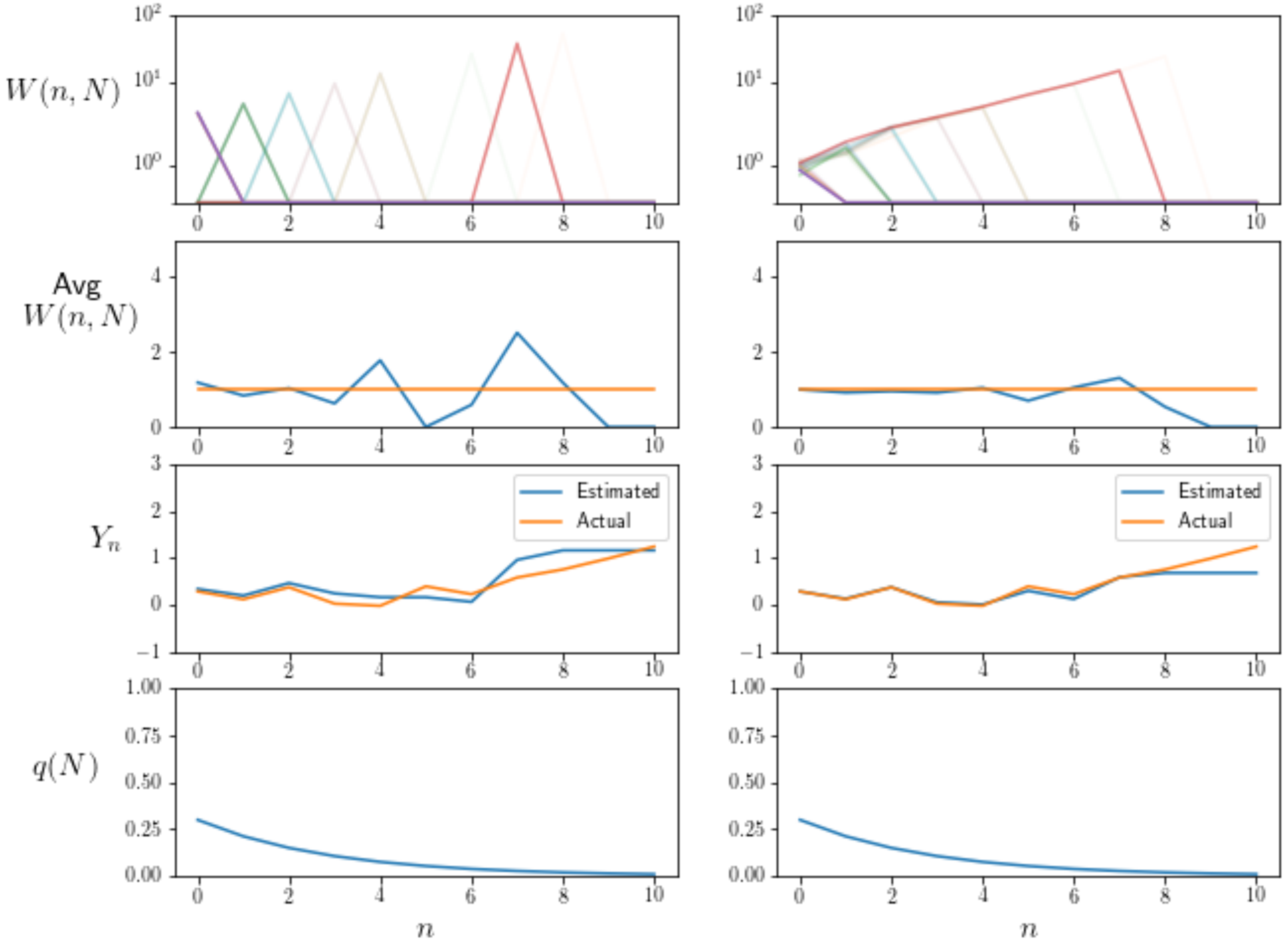


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



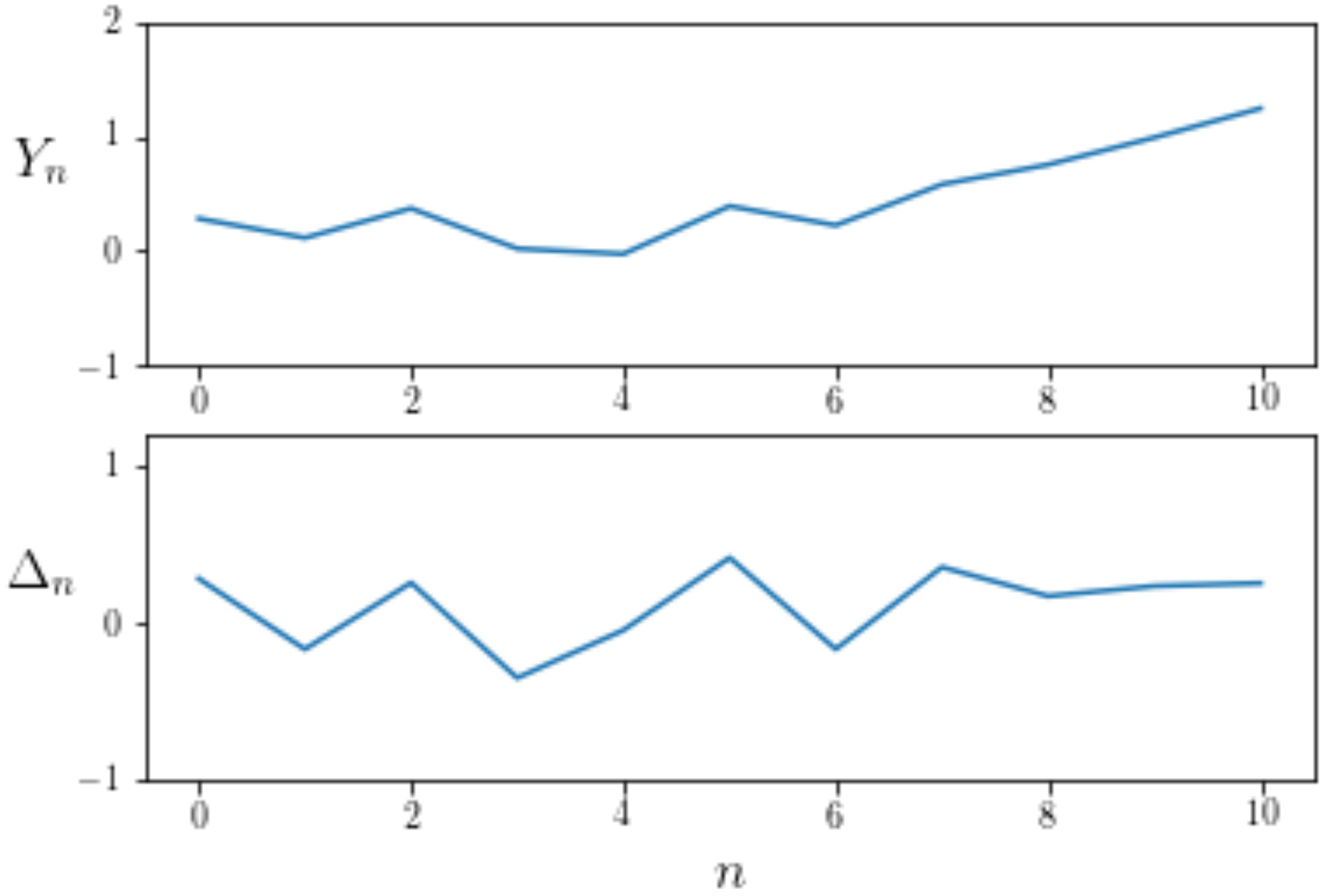
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

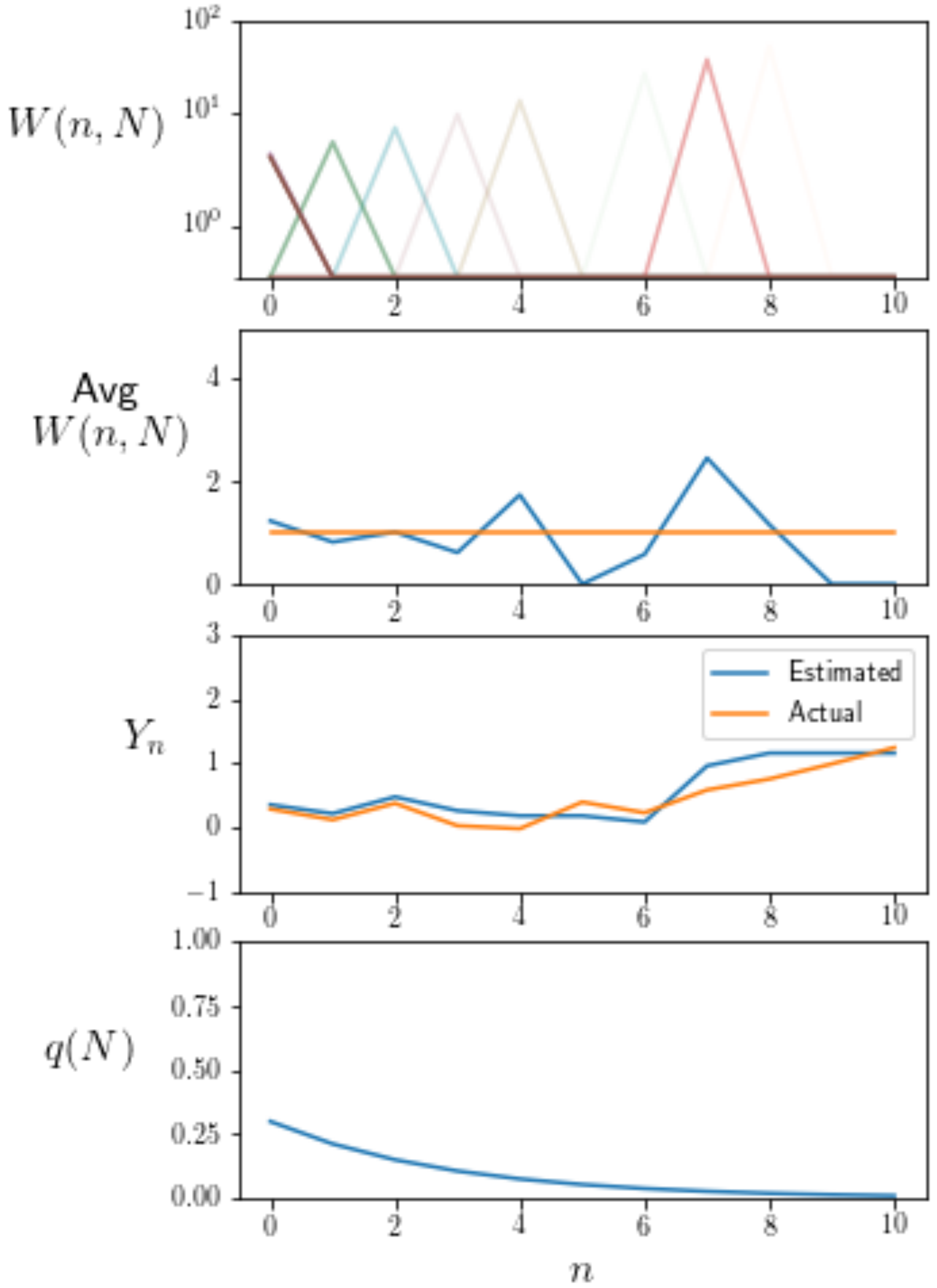
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



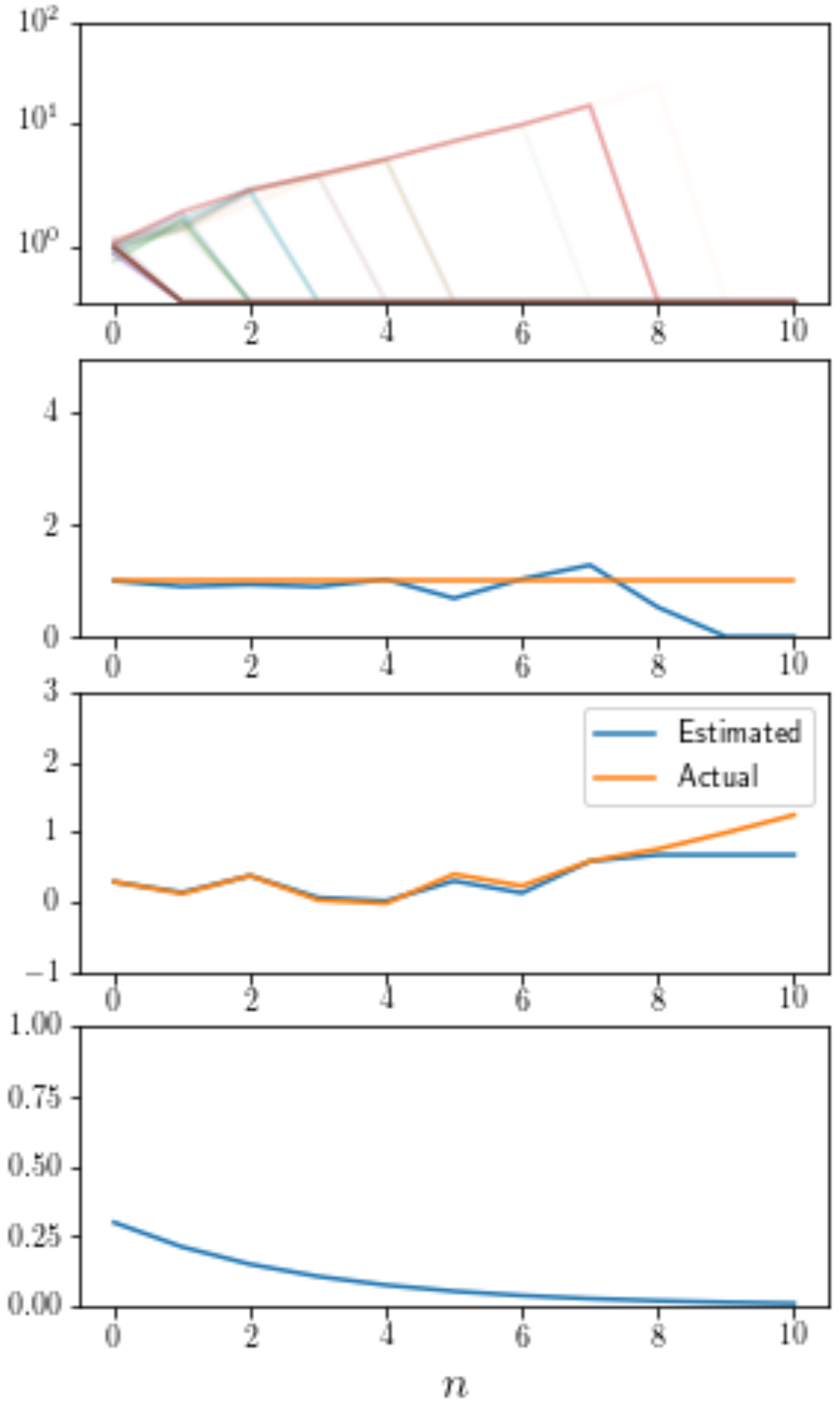
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



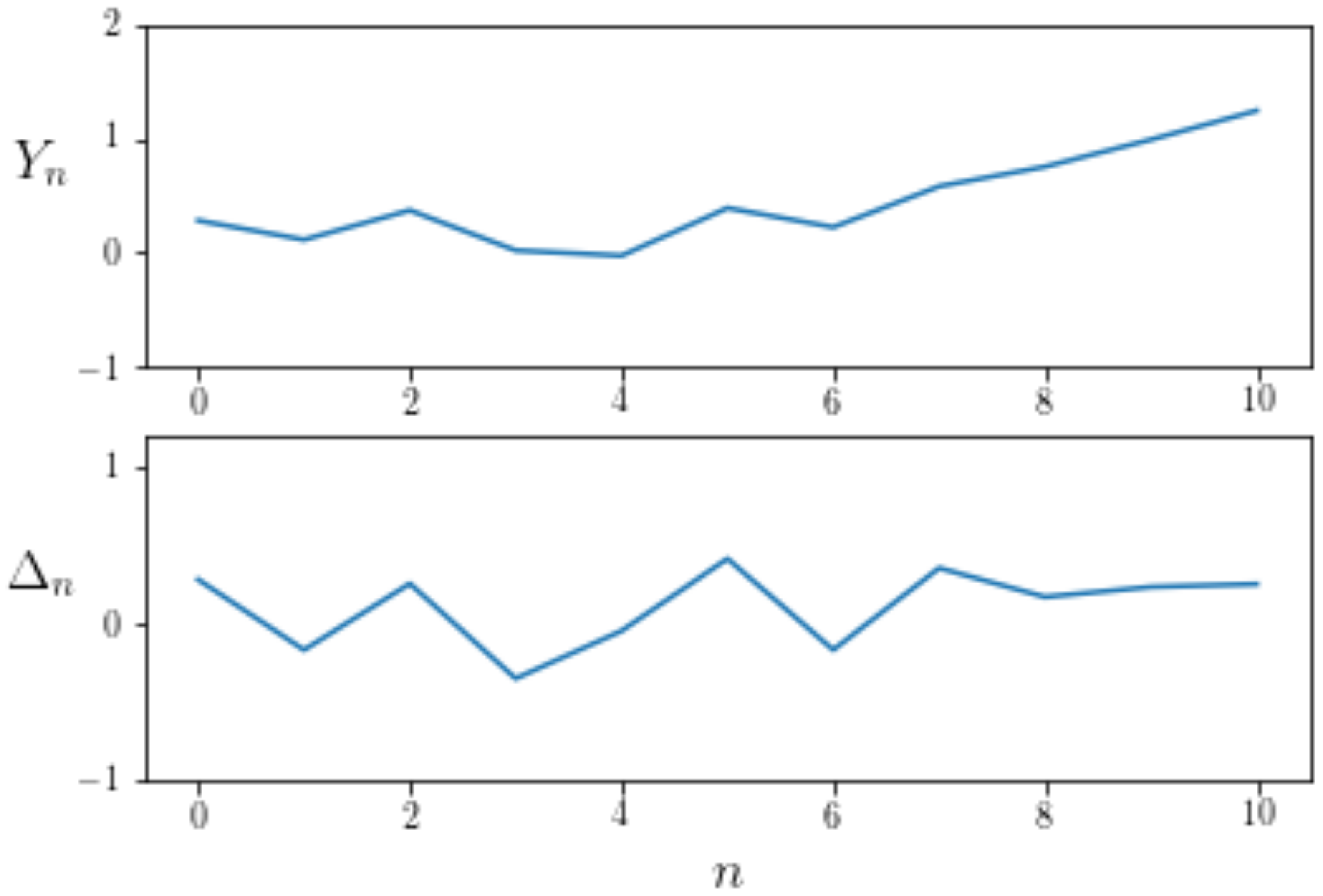
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

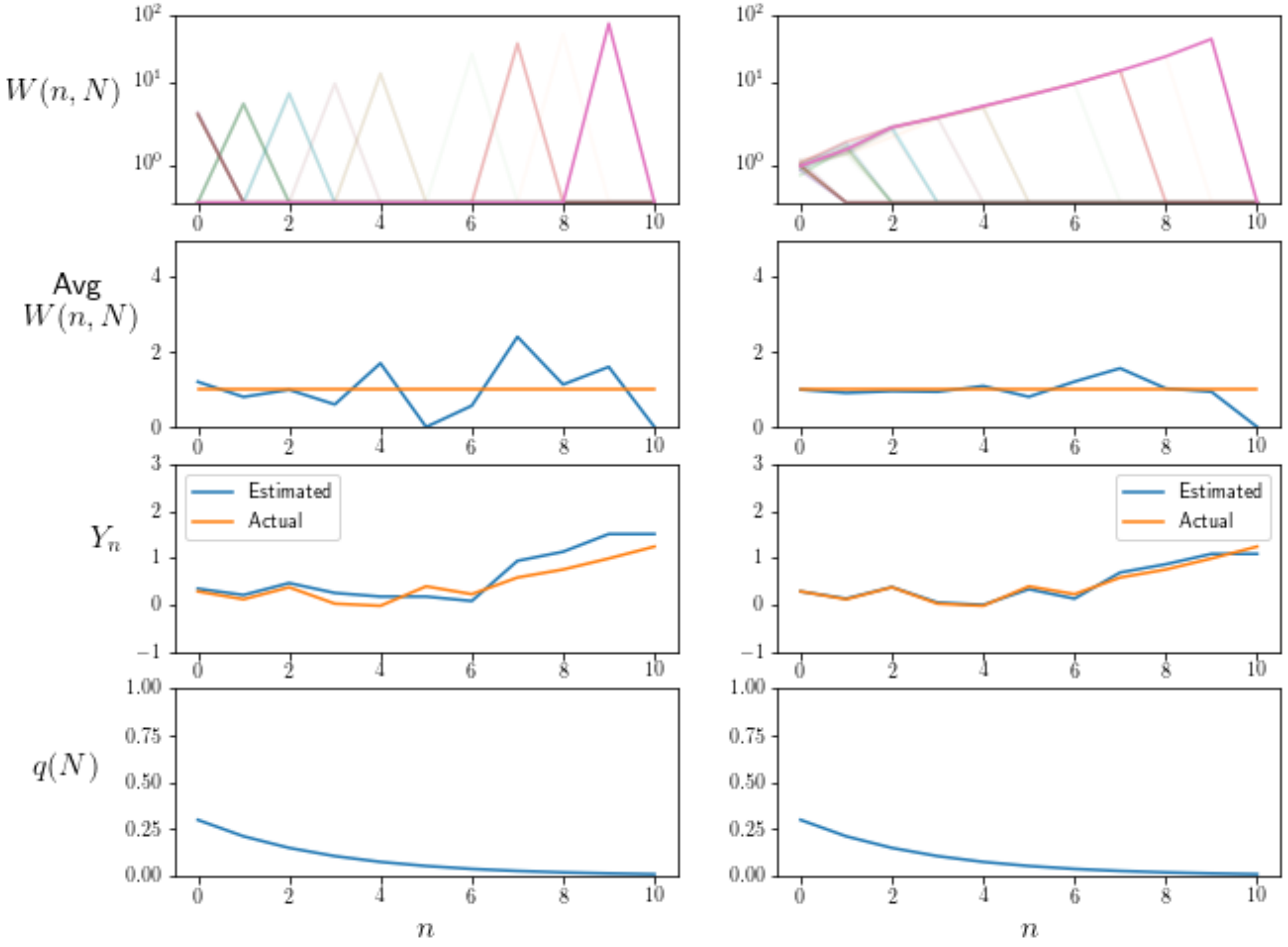


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



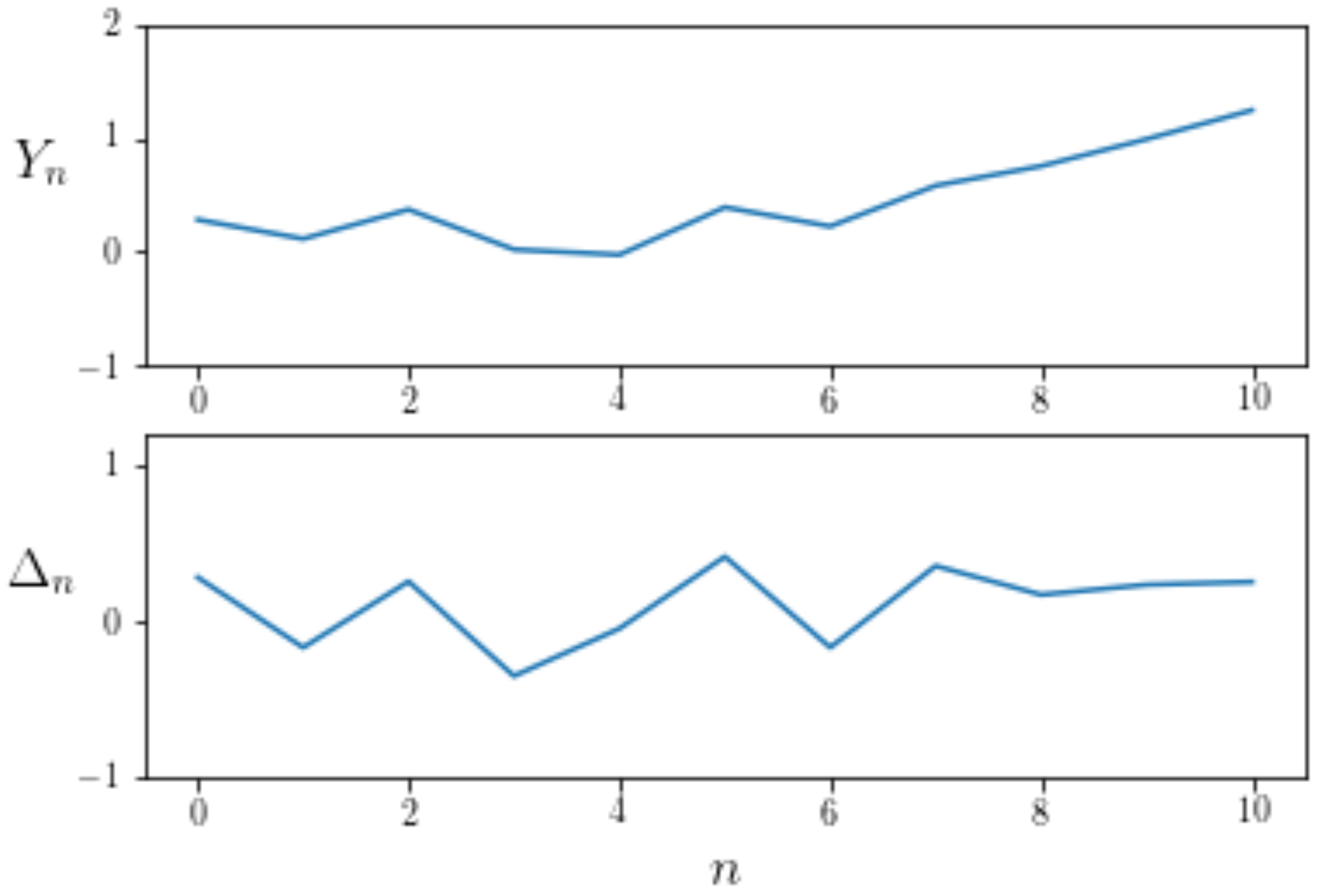
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

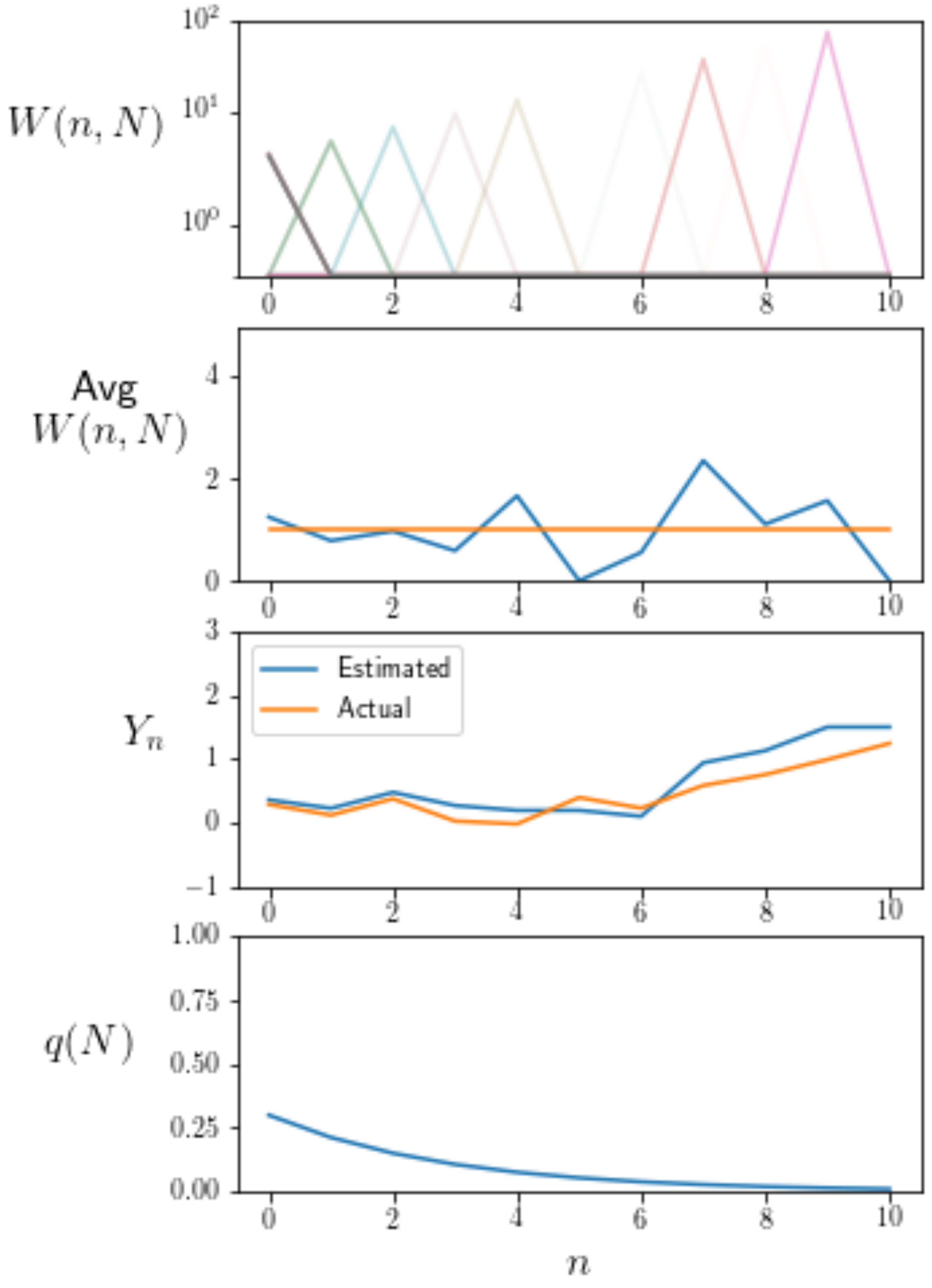
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



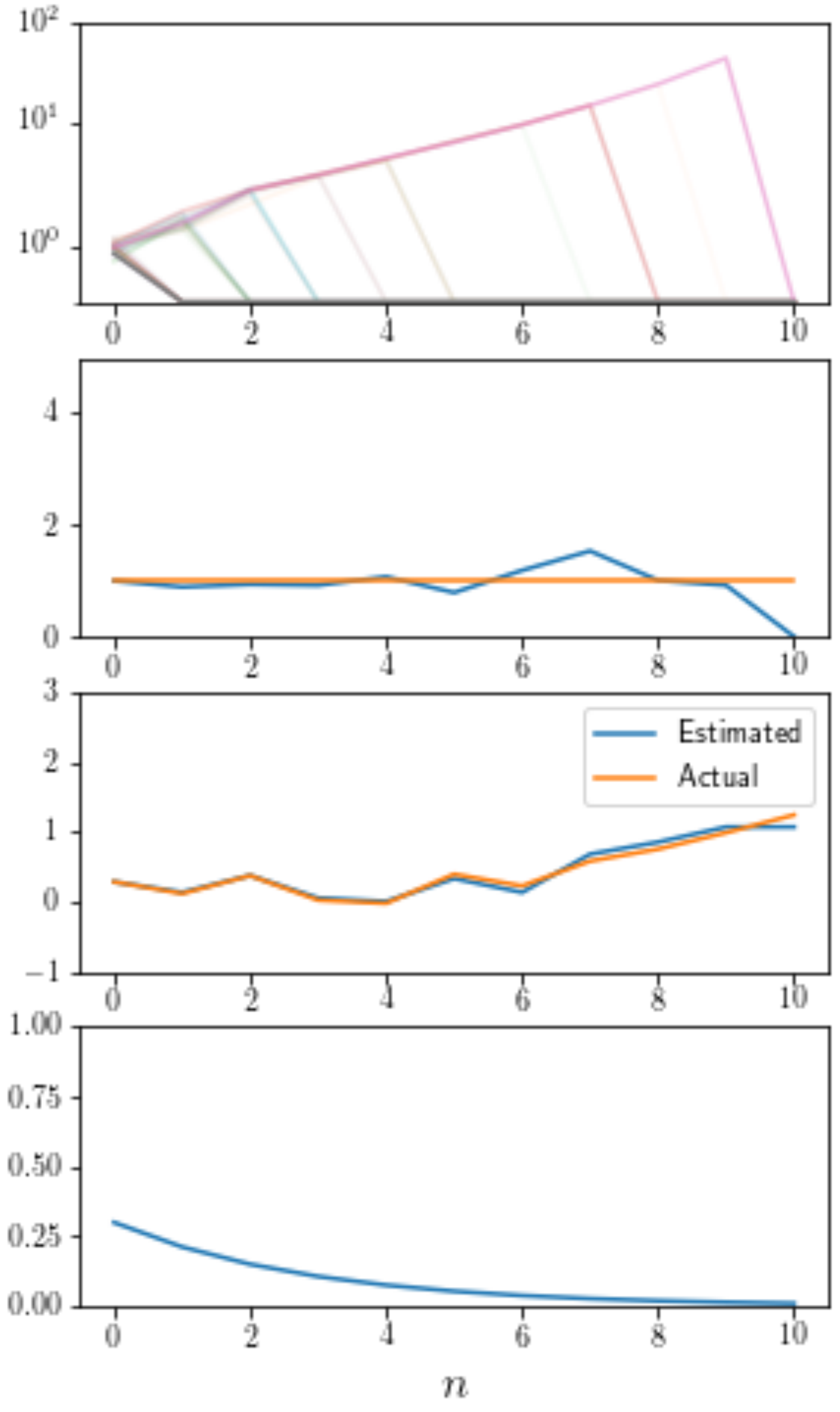
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



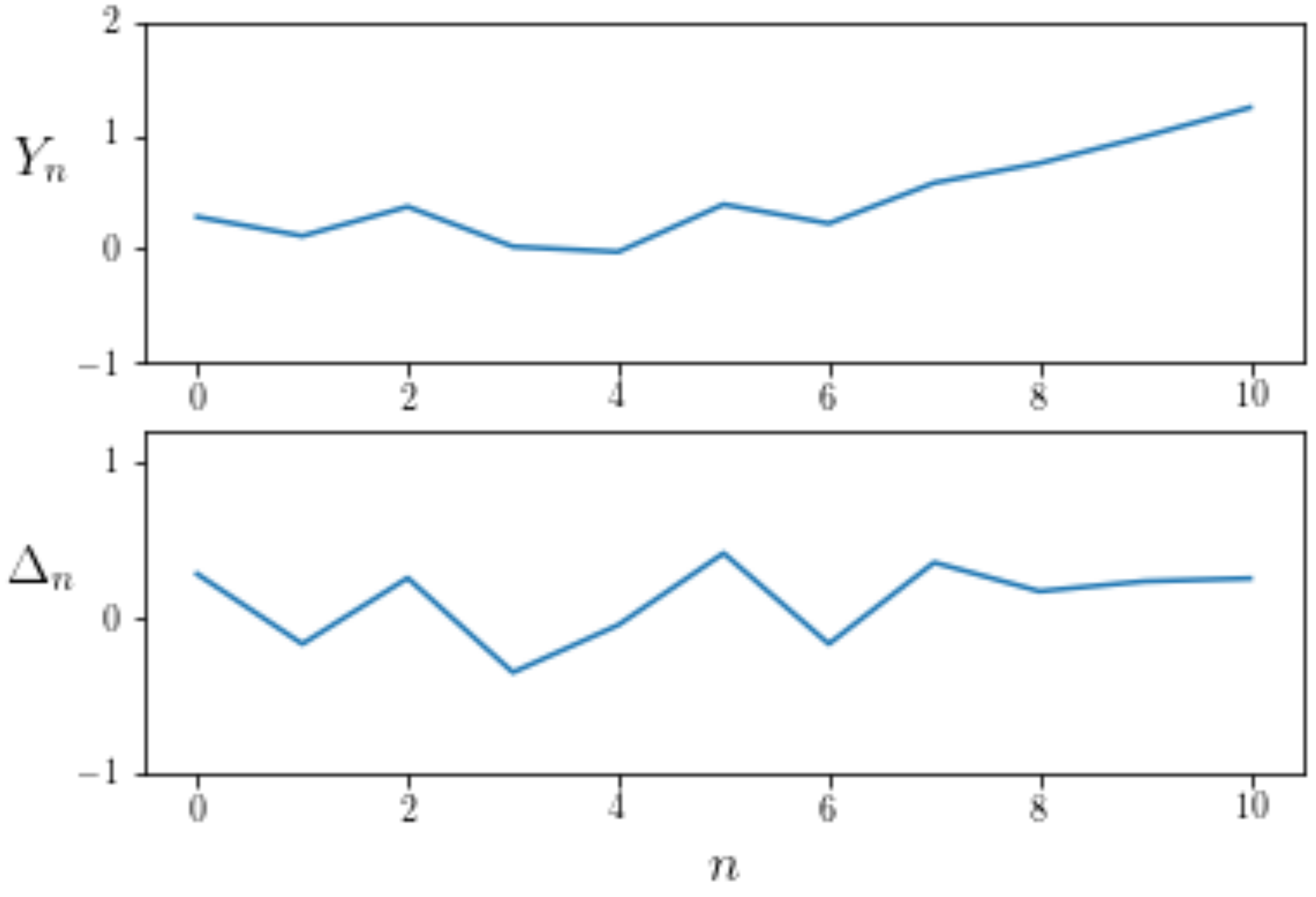
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

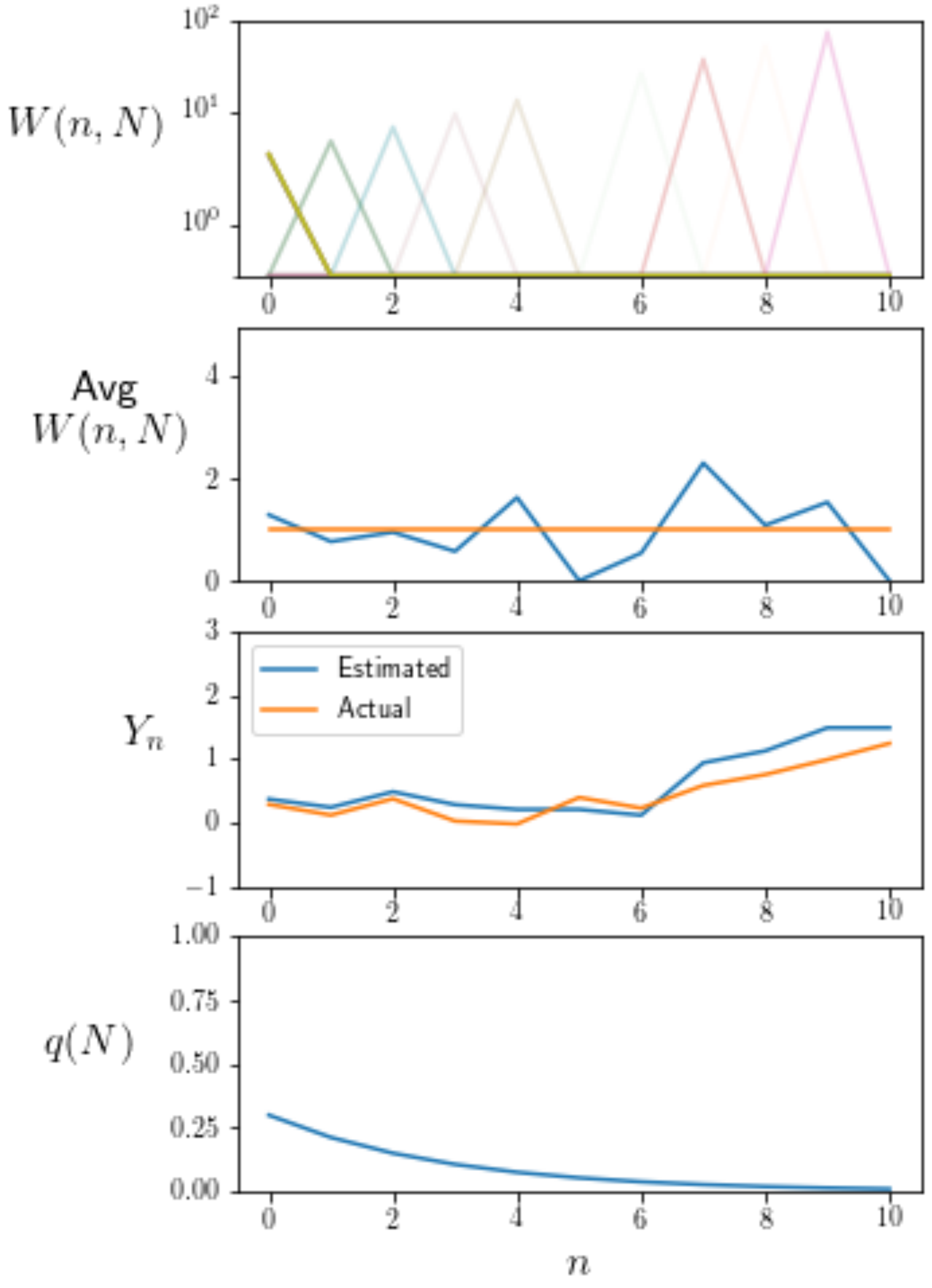
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



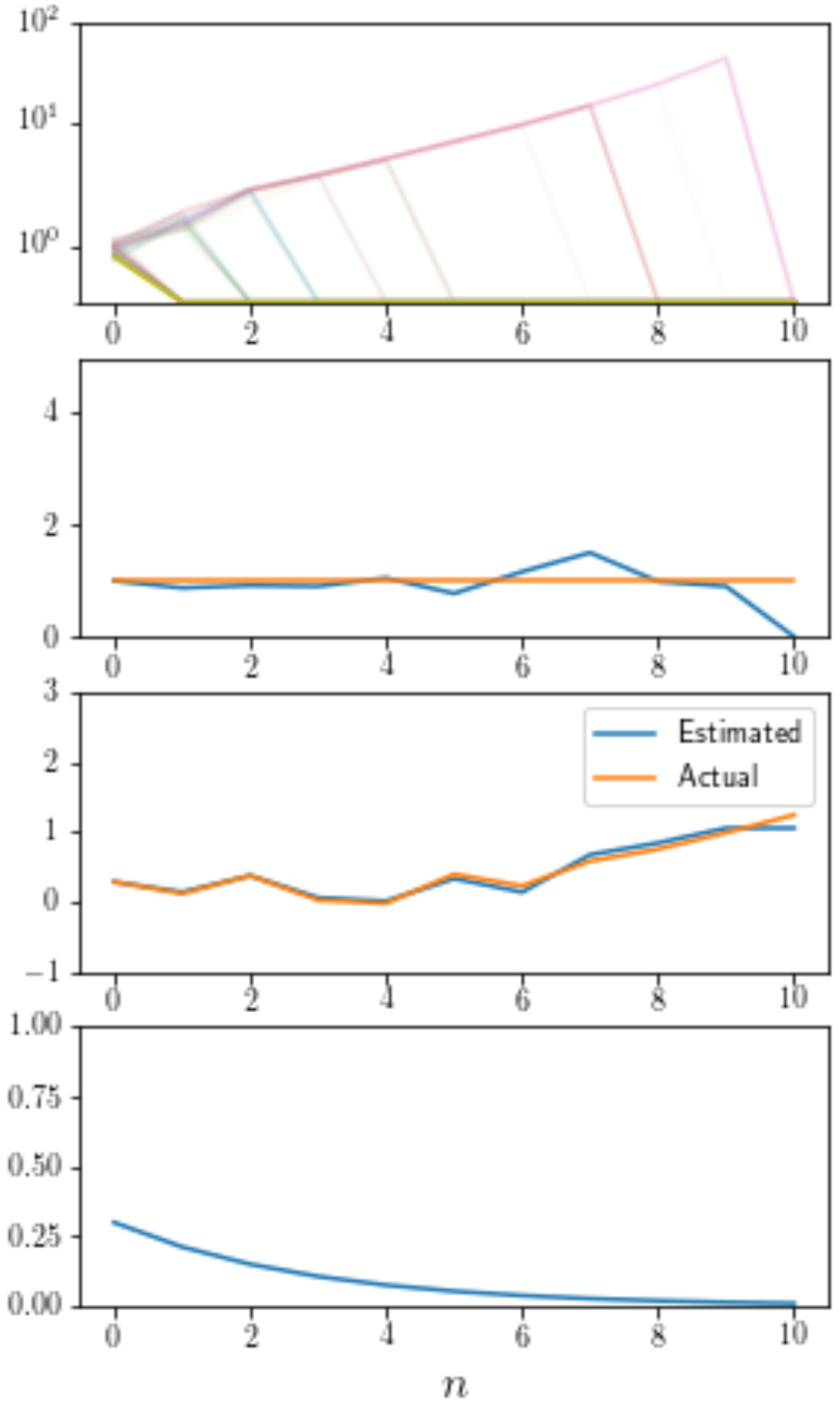
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



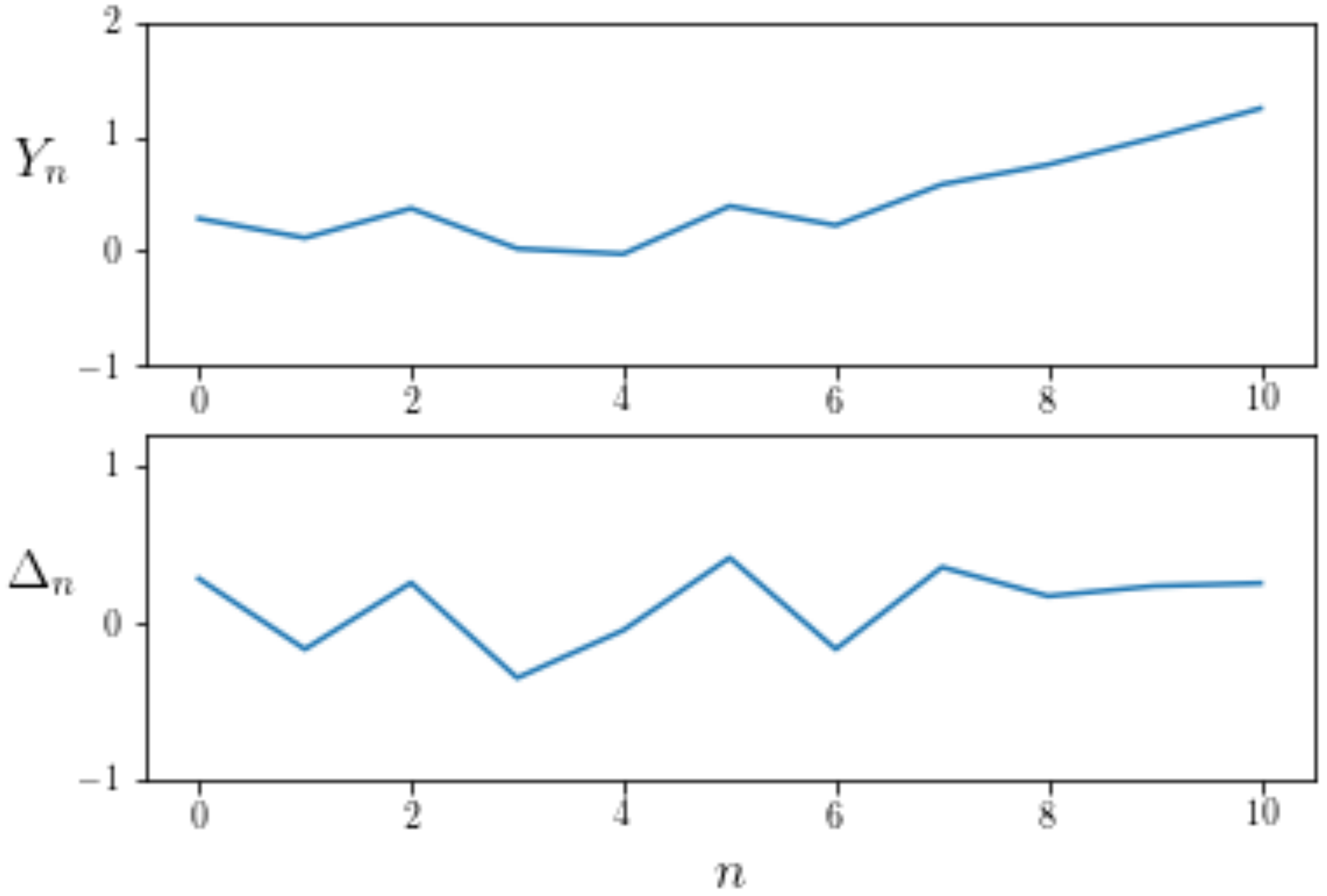
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n,N) \qquad N \in \{1,\ldots,H\} \sim q$$

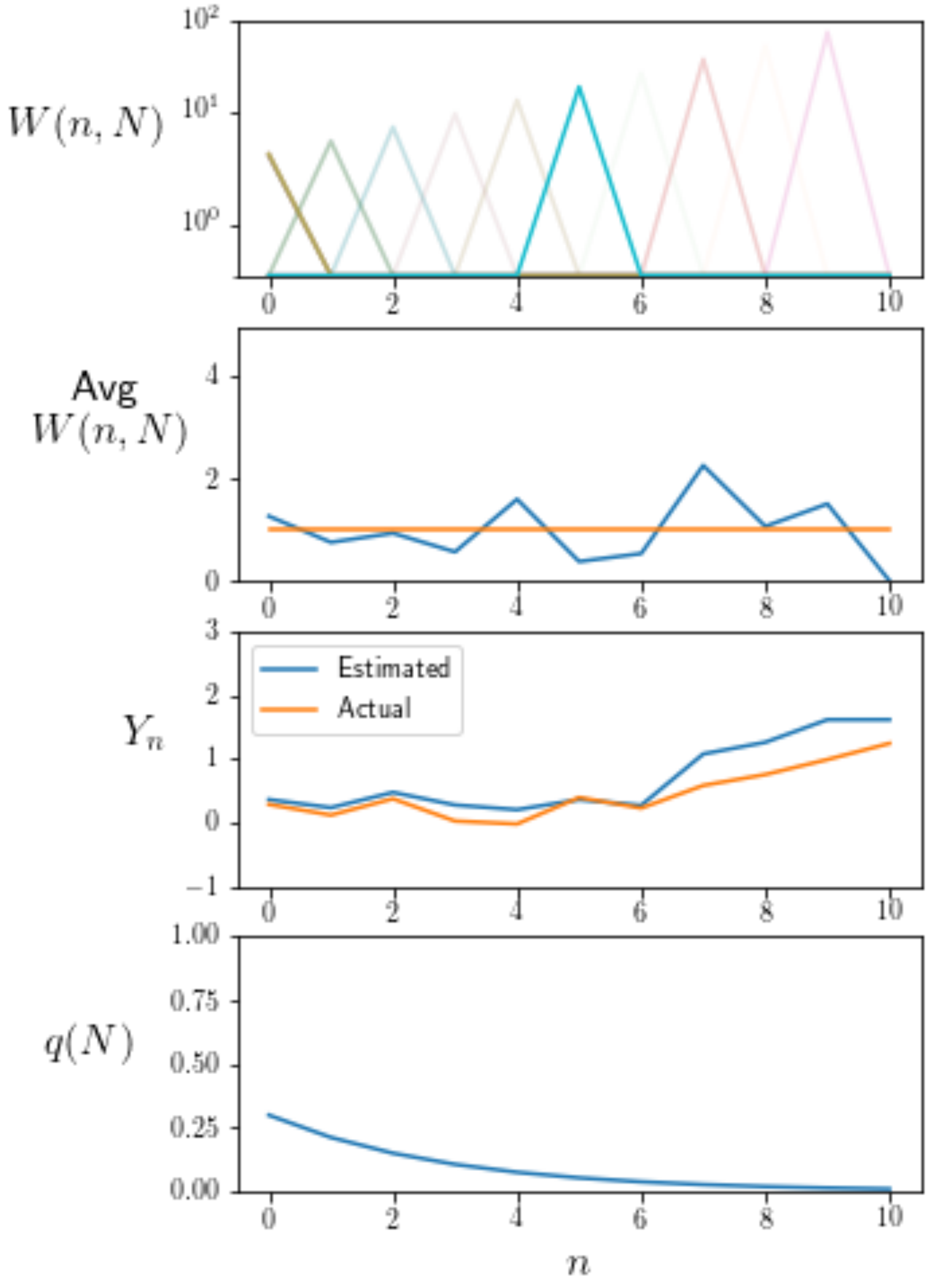
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



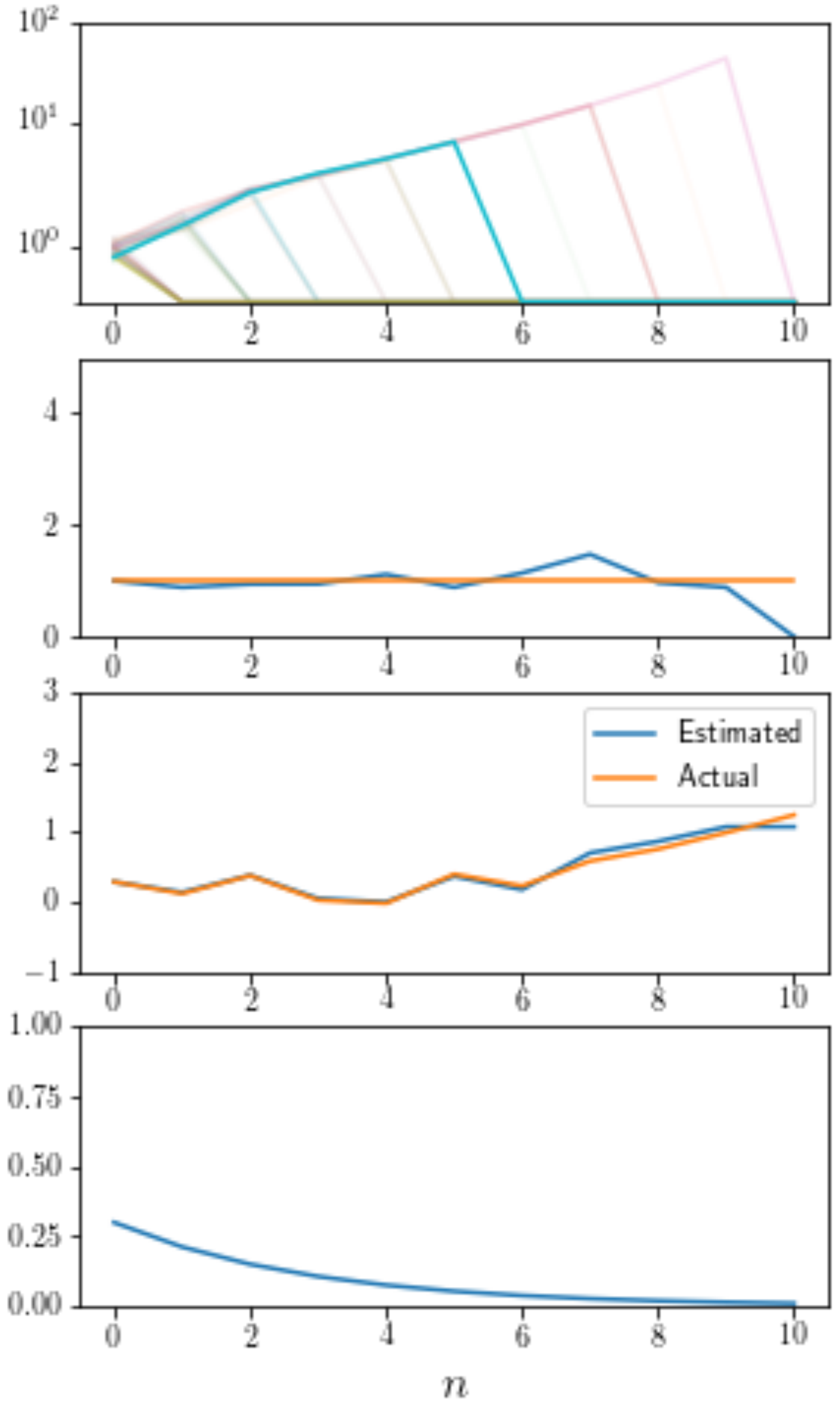
"Single sample"

$$W(n,N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n,N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



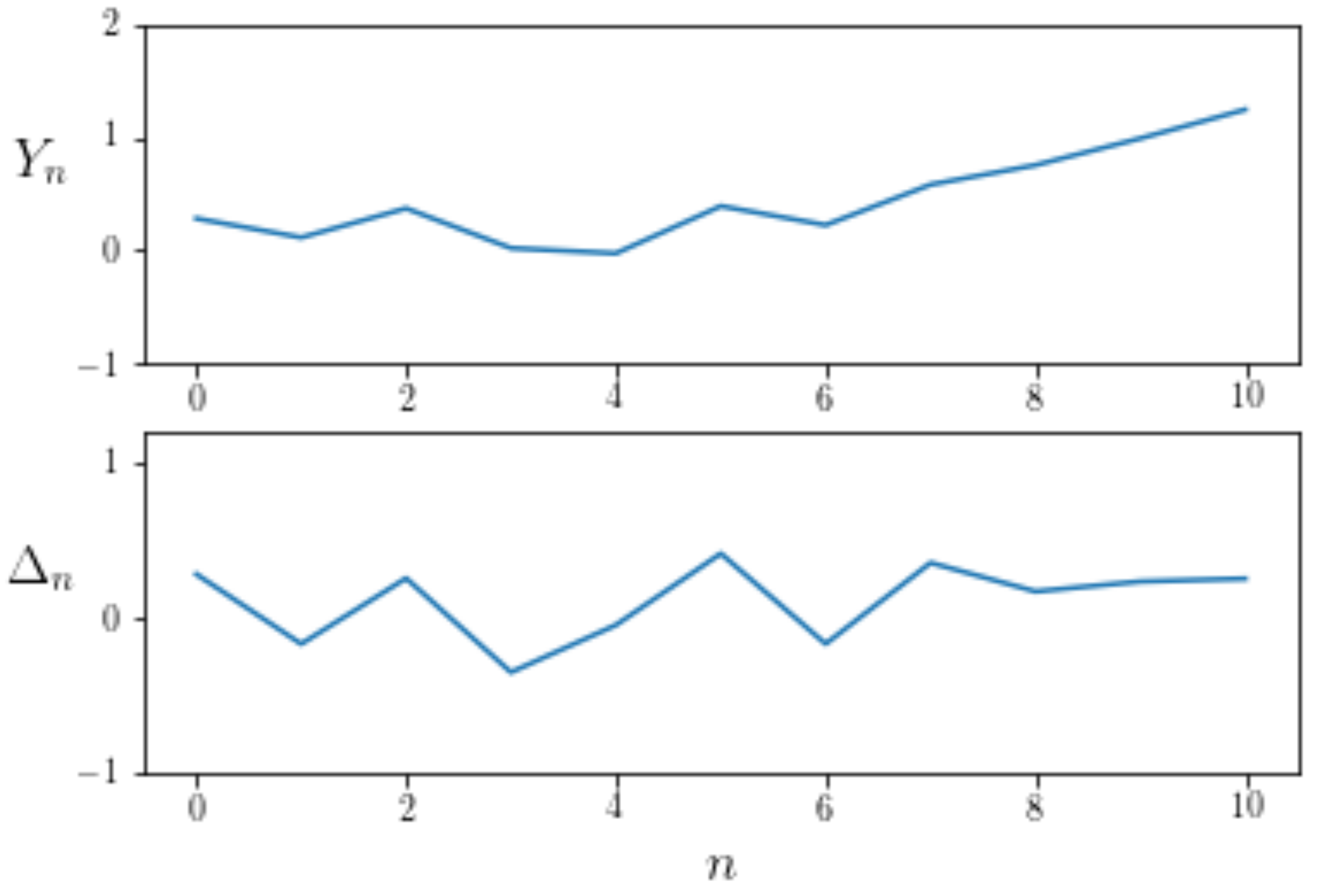
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

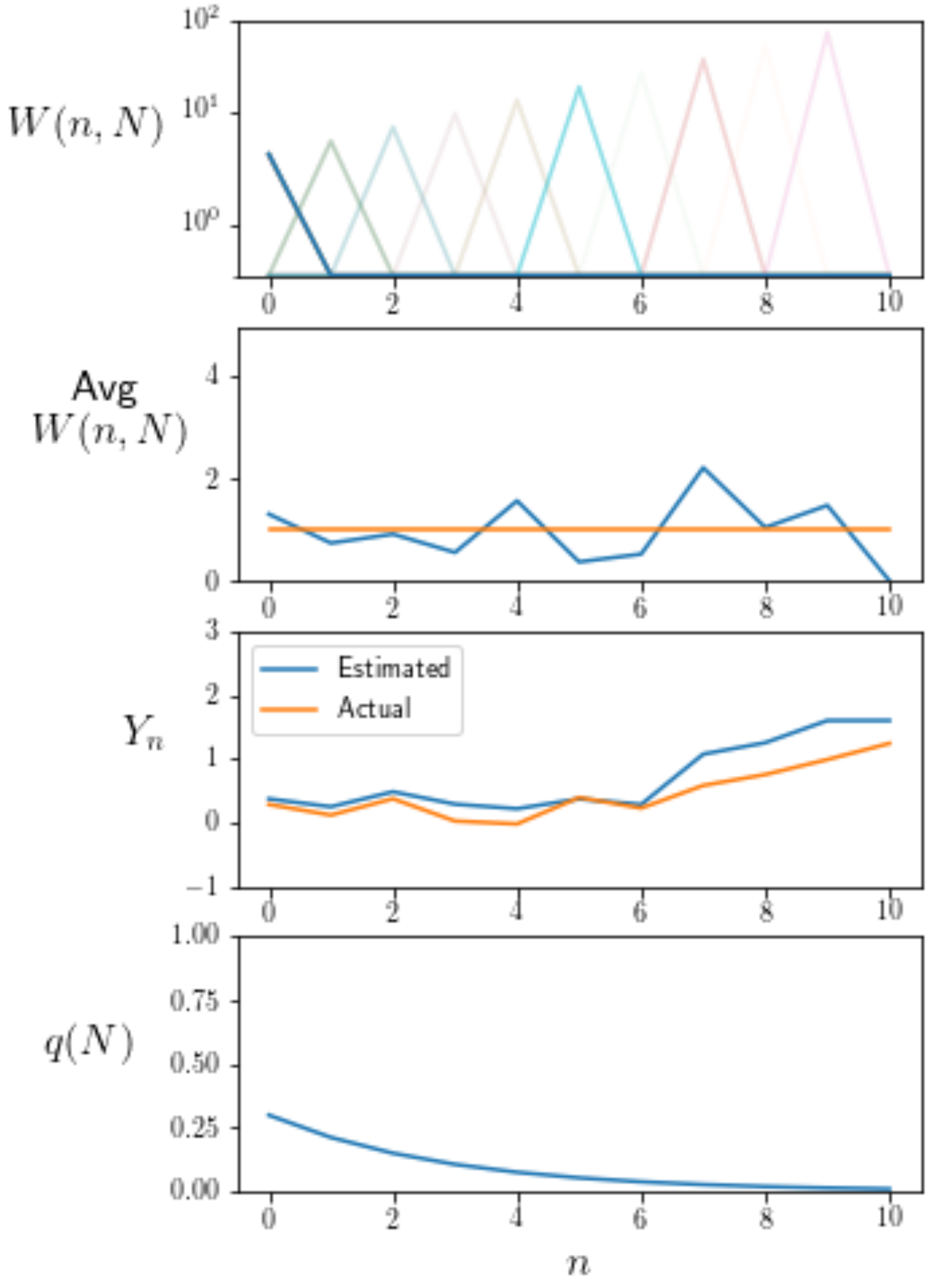
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



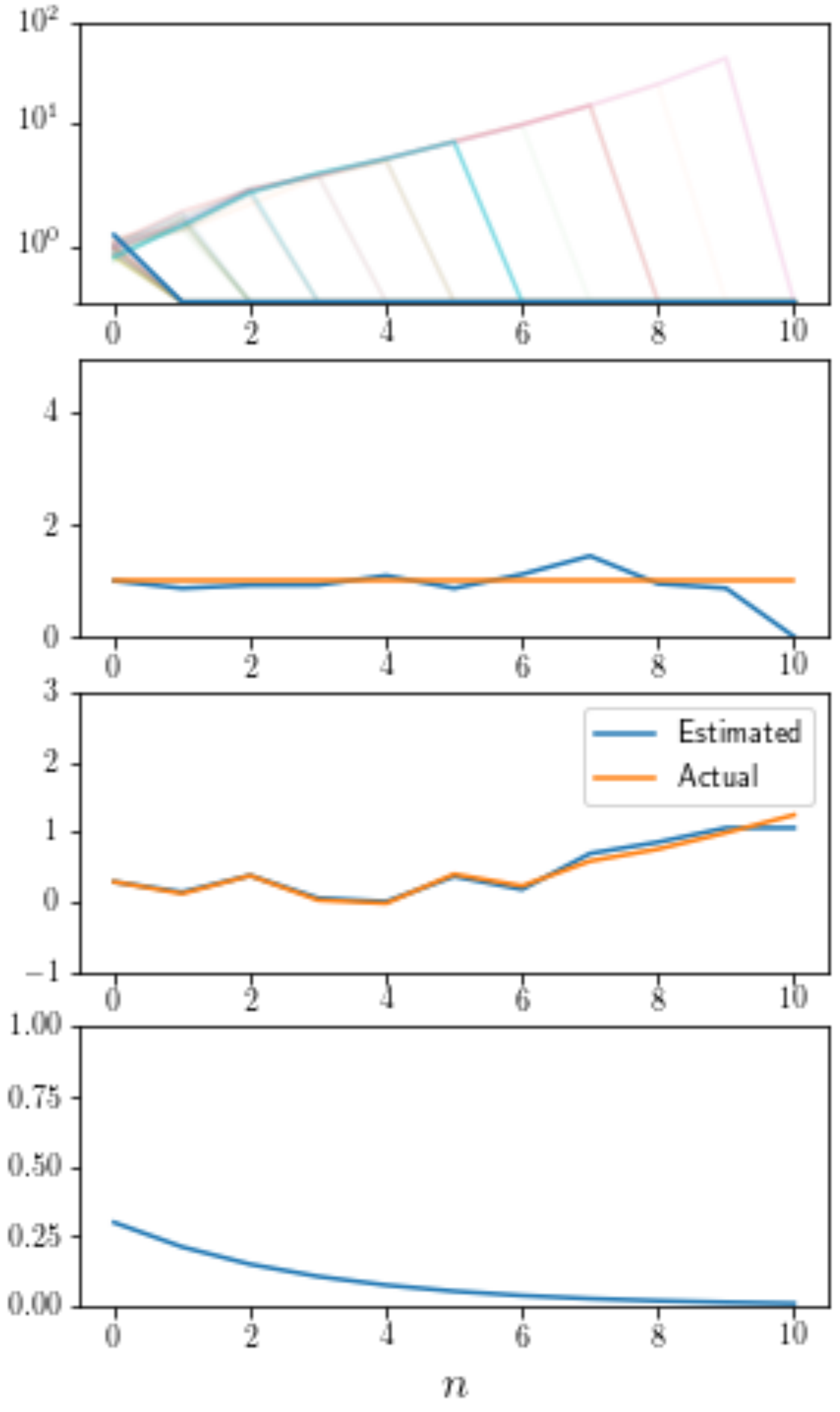
“Single sample”

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



“Russian roulette”

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



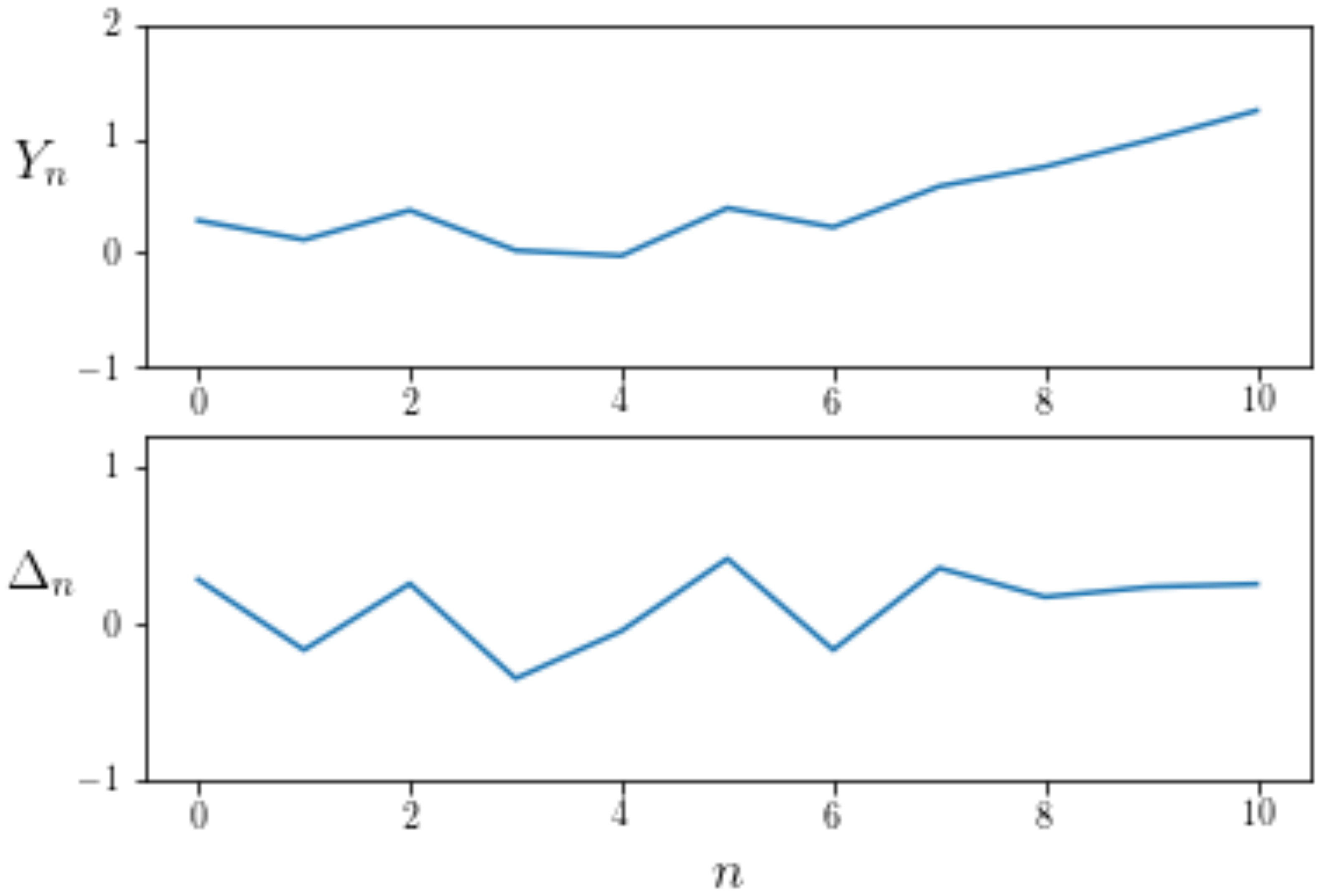
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

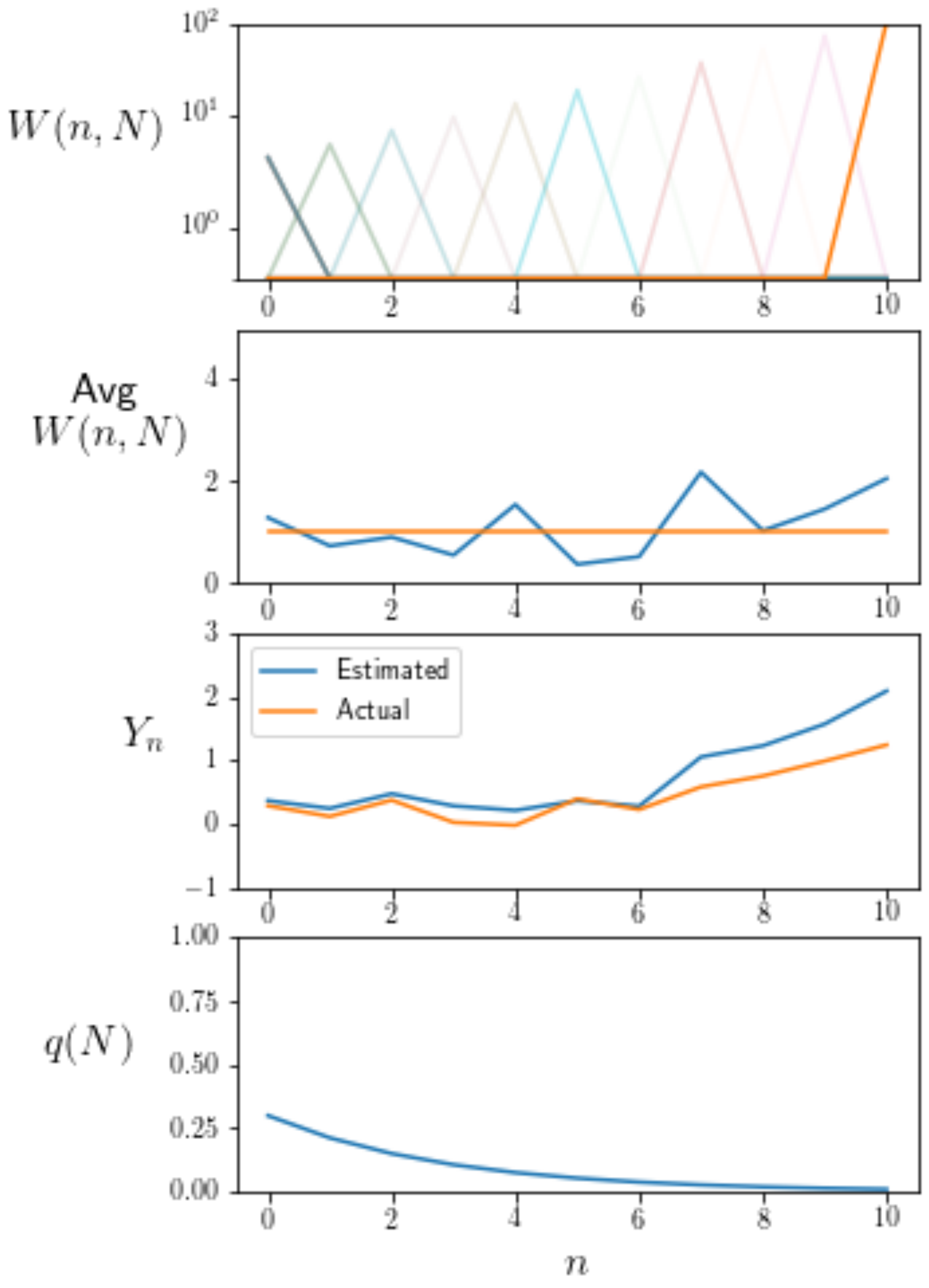
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



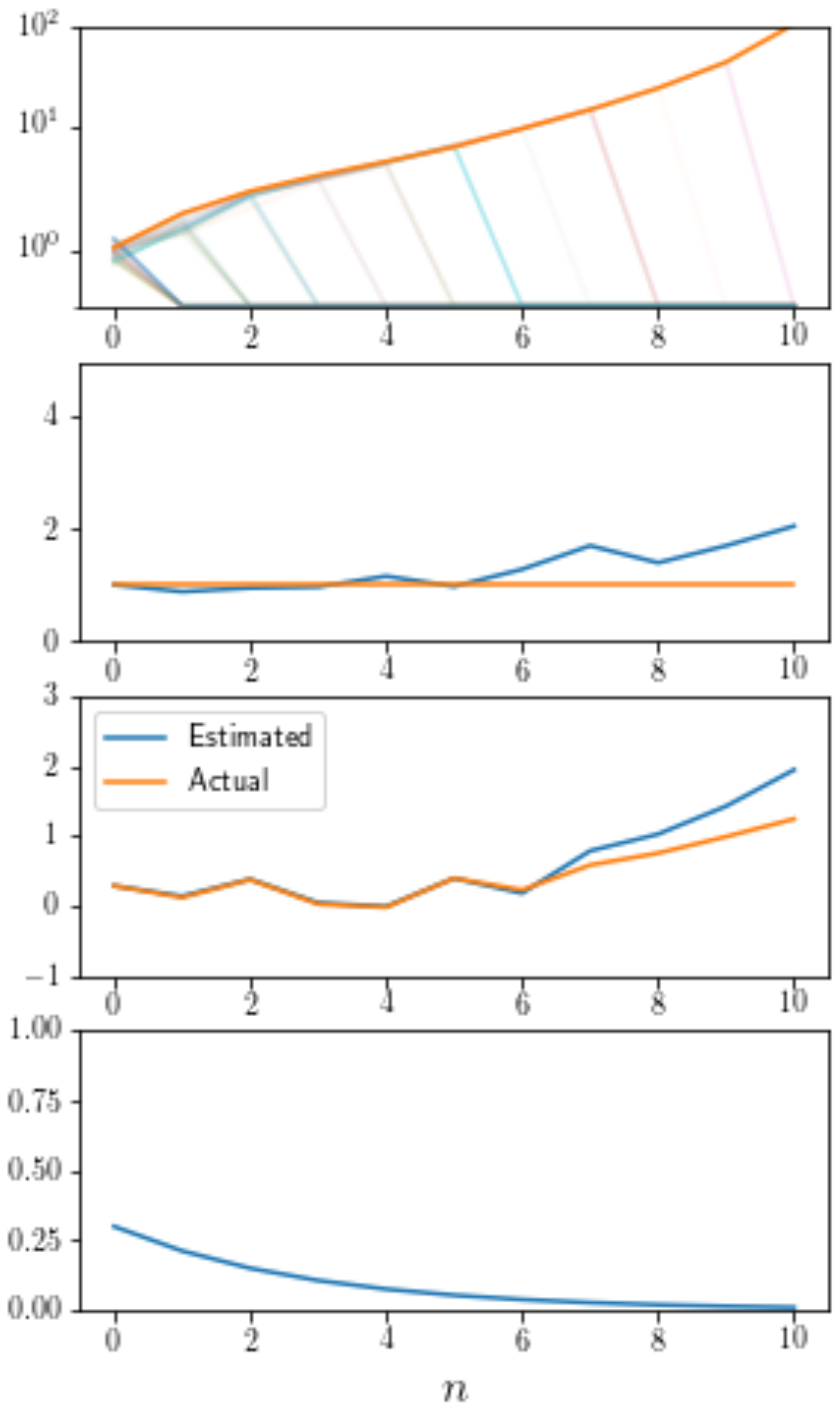
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



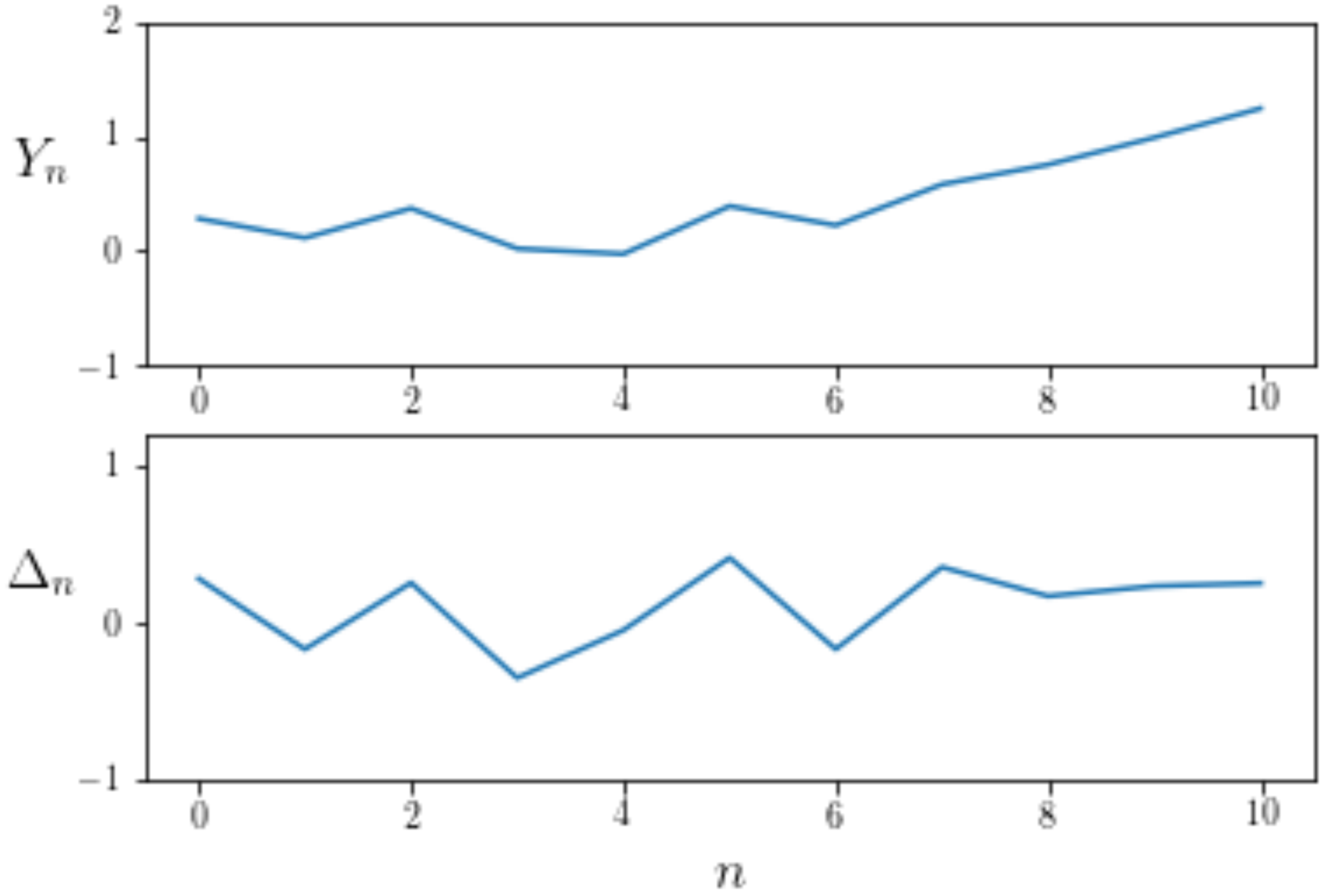
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

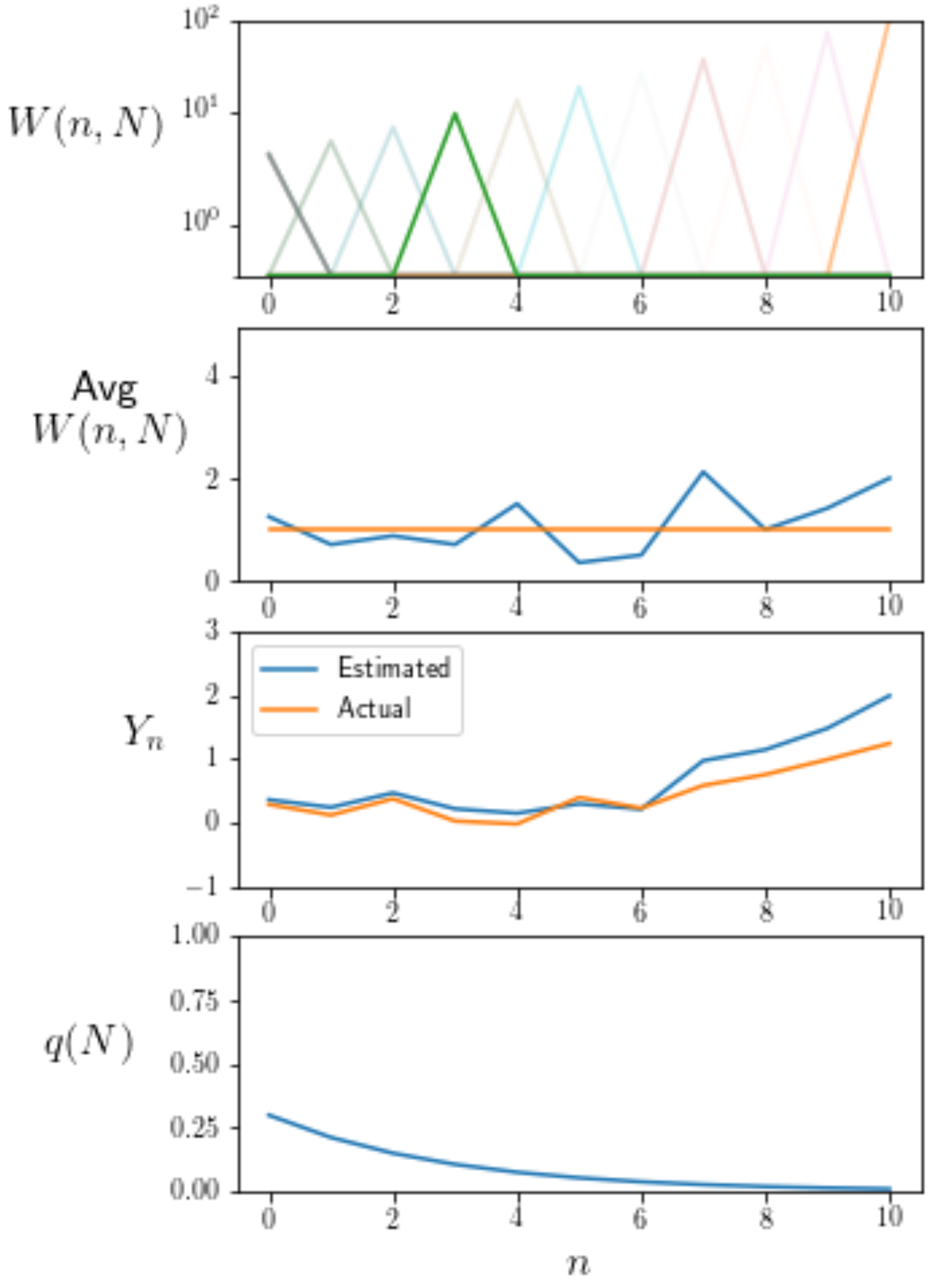
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



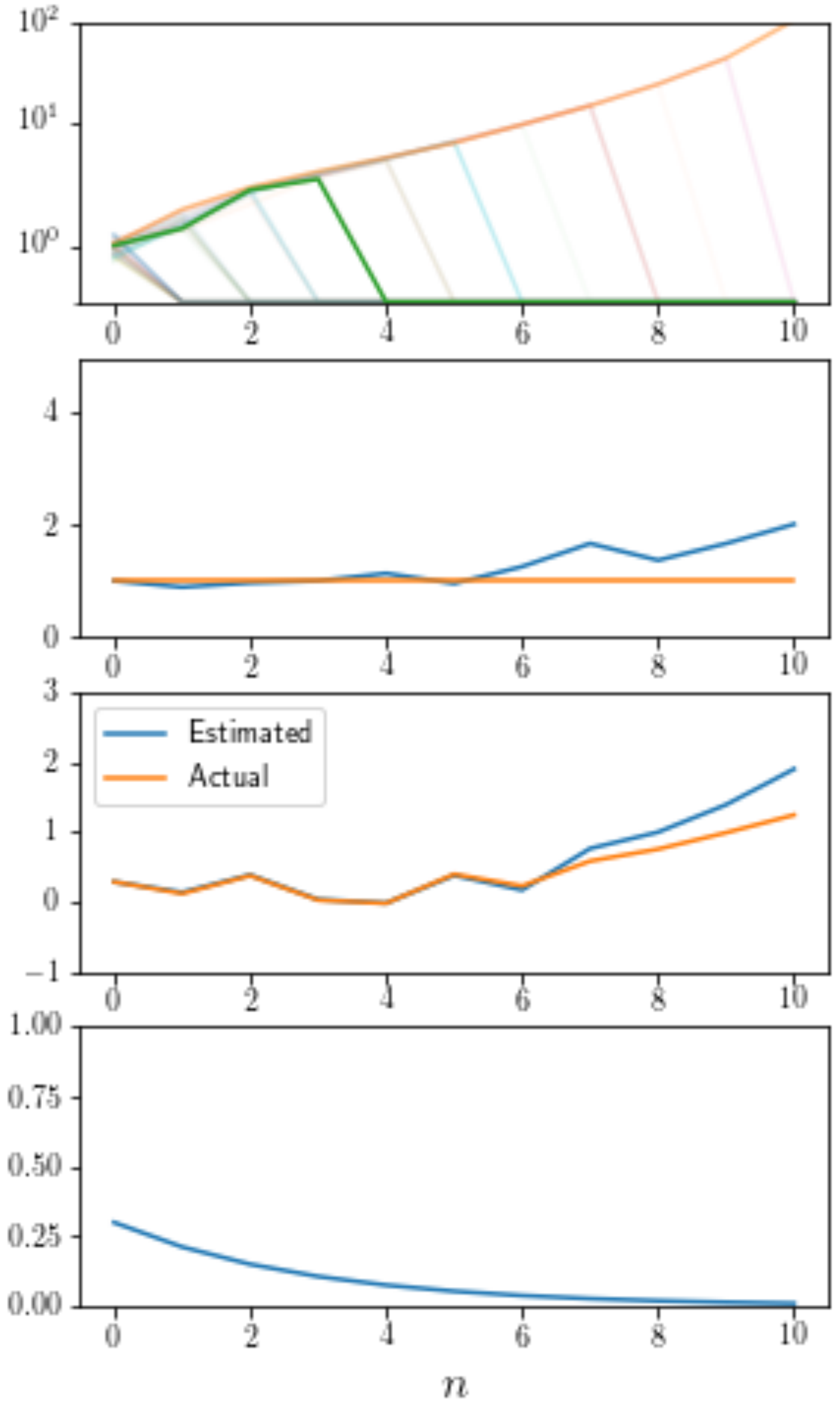
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



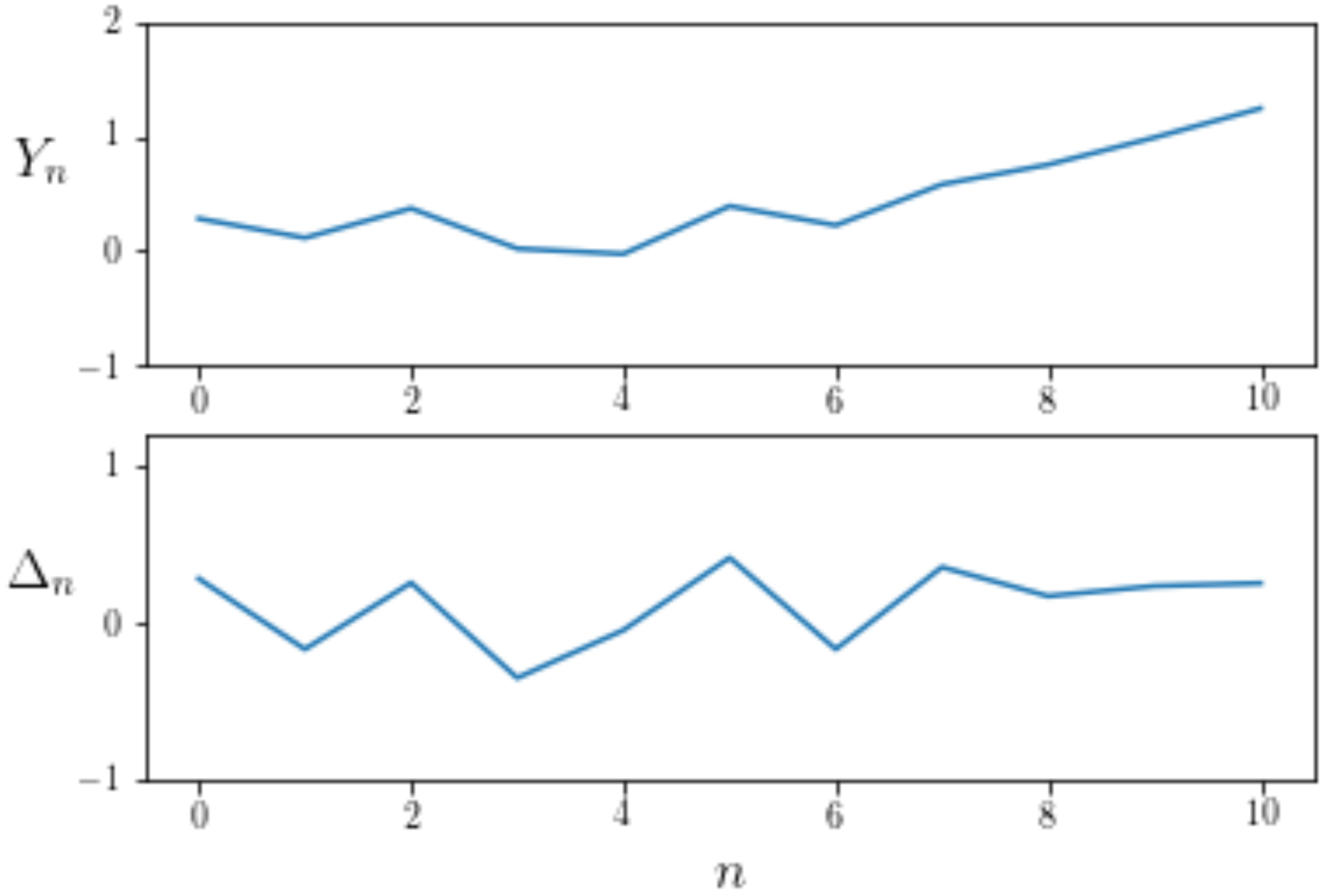
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

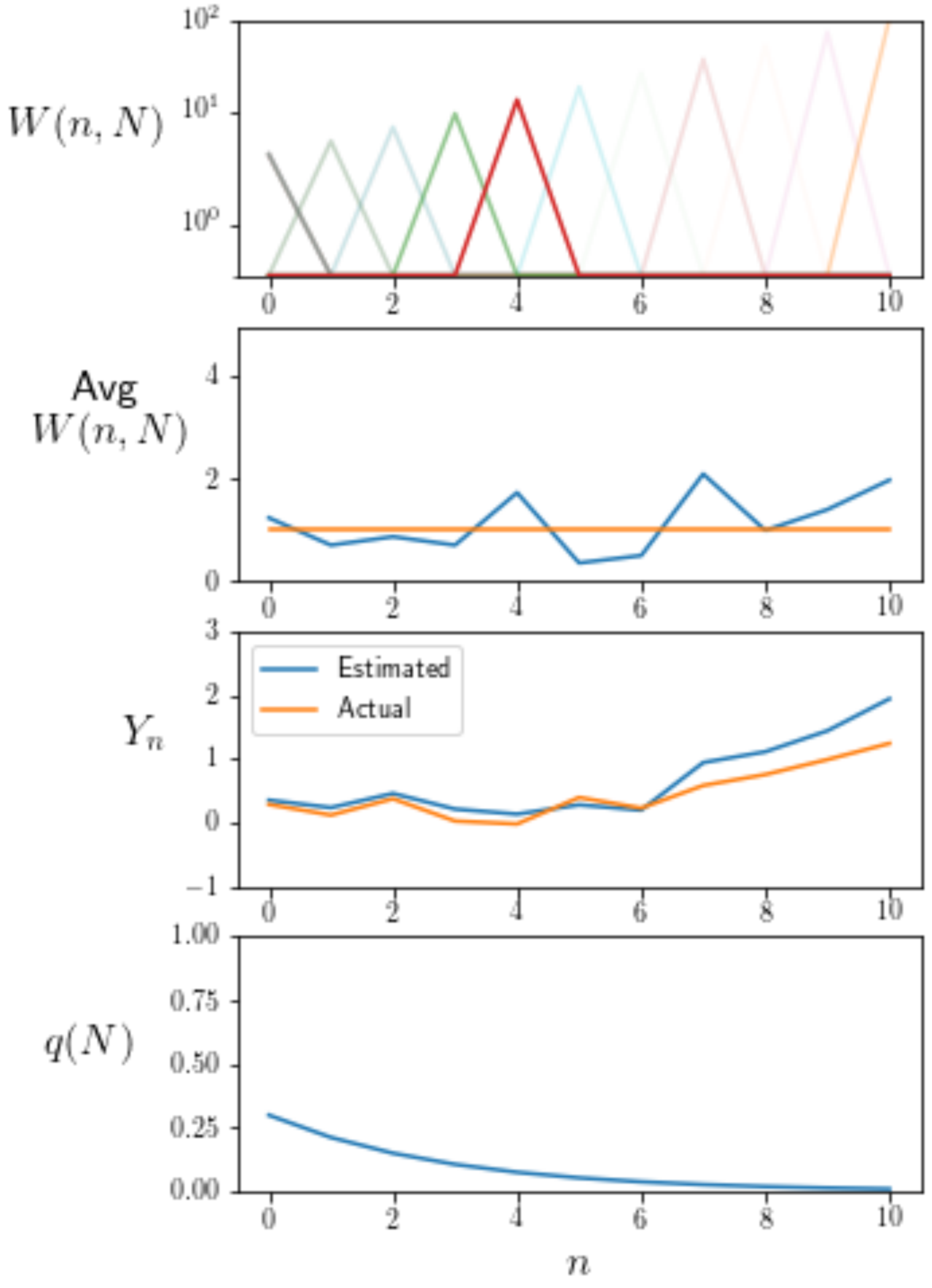
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



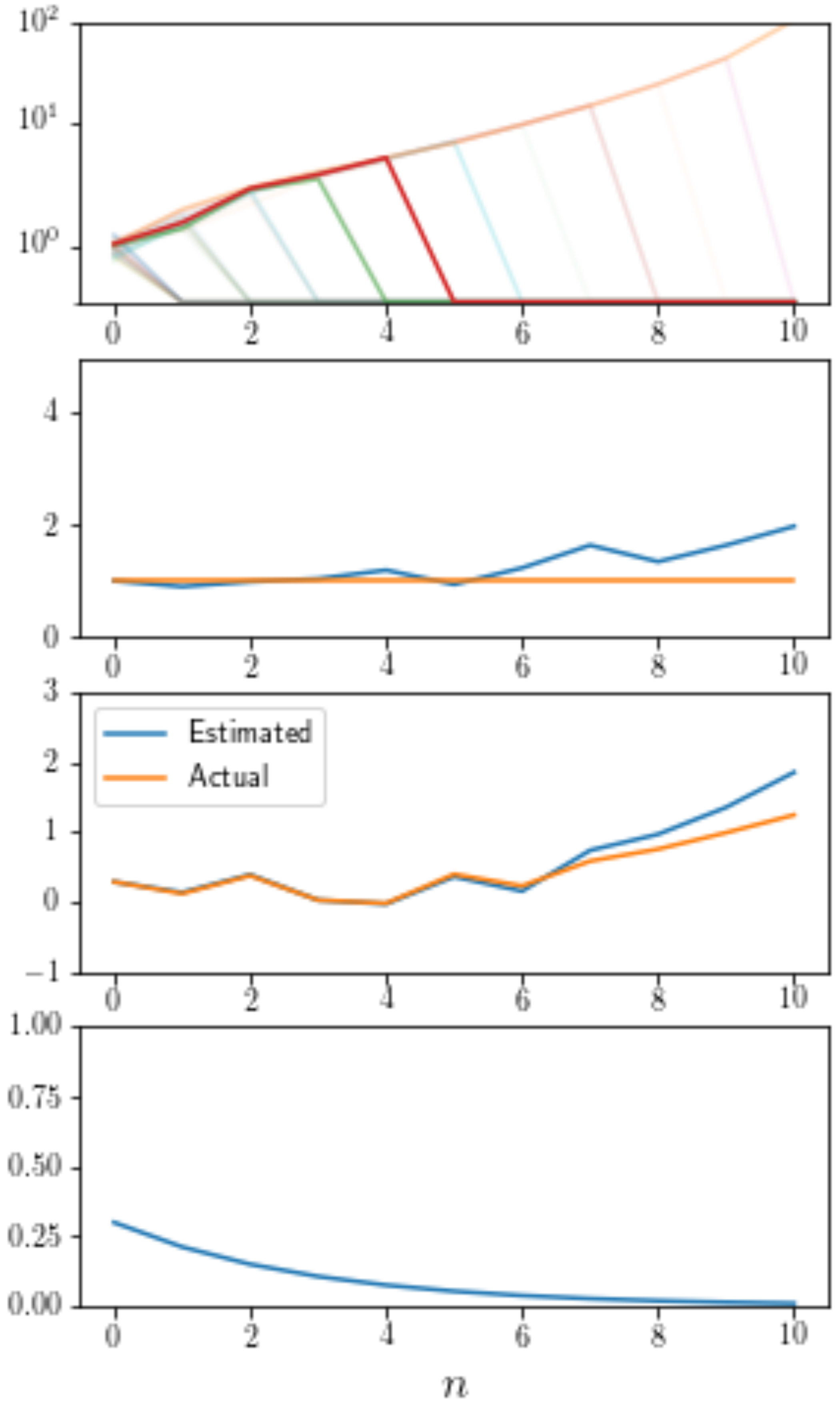
“Single sample”

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



“Russian roulette”

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



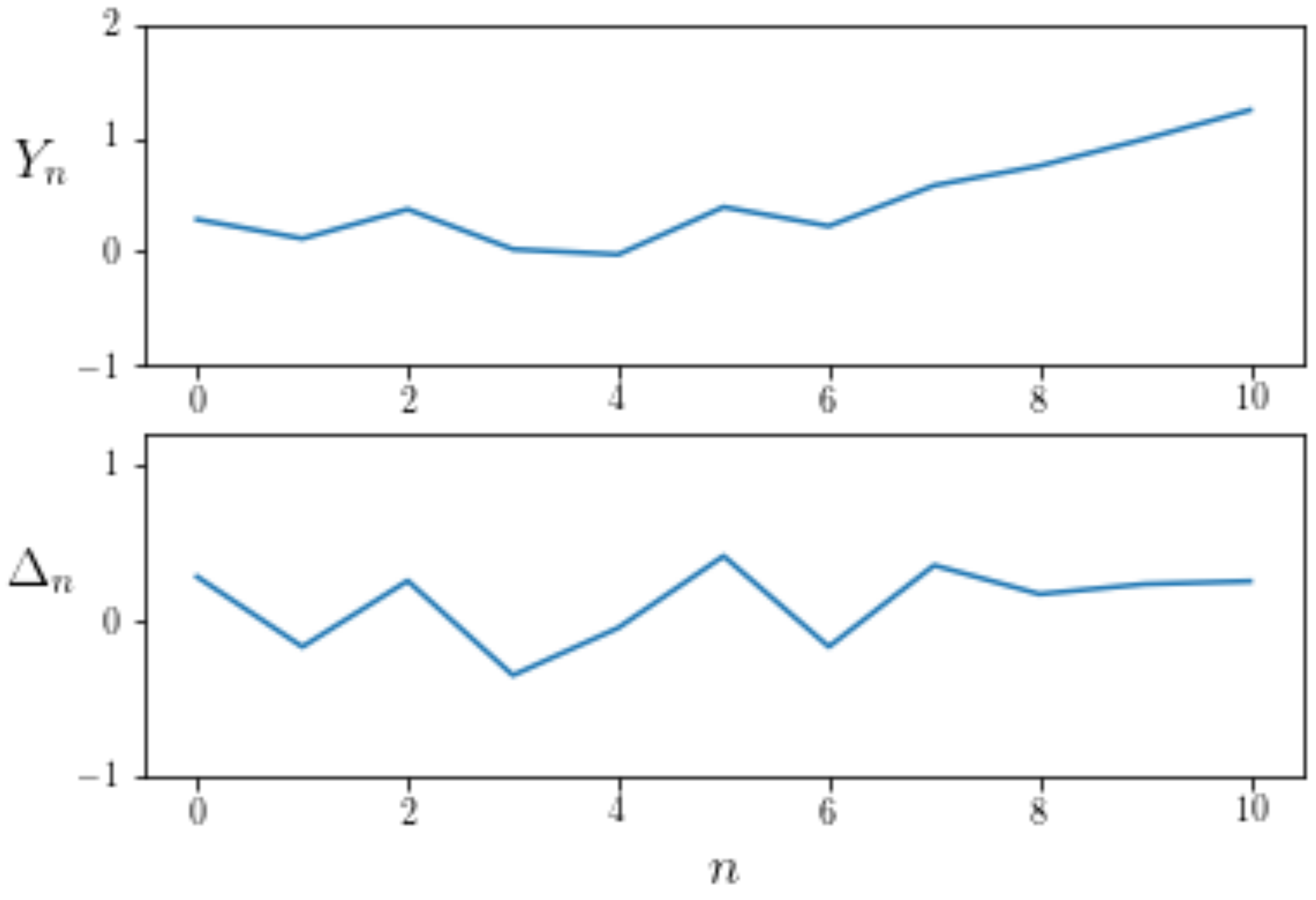
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

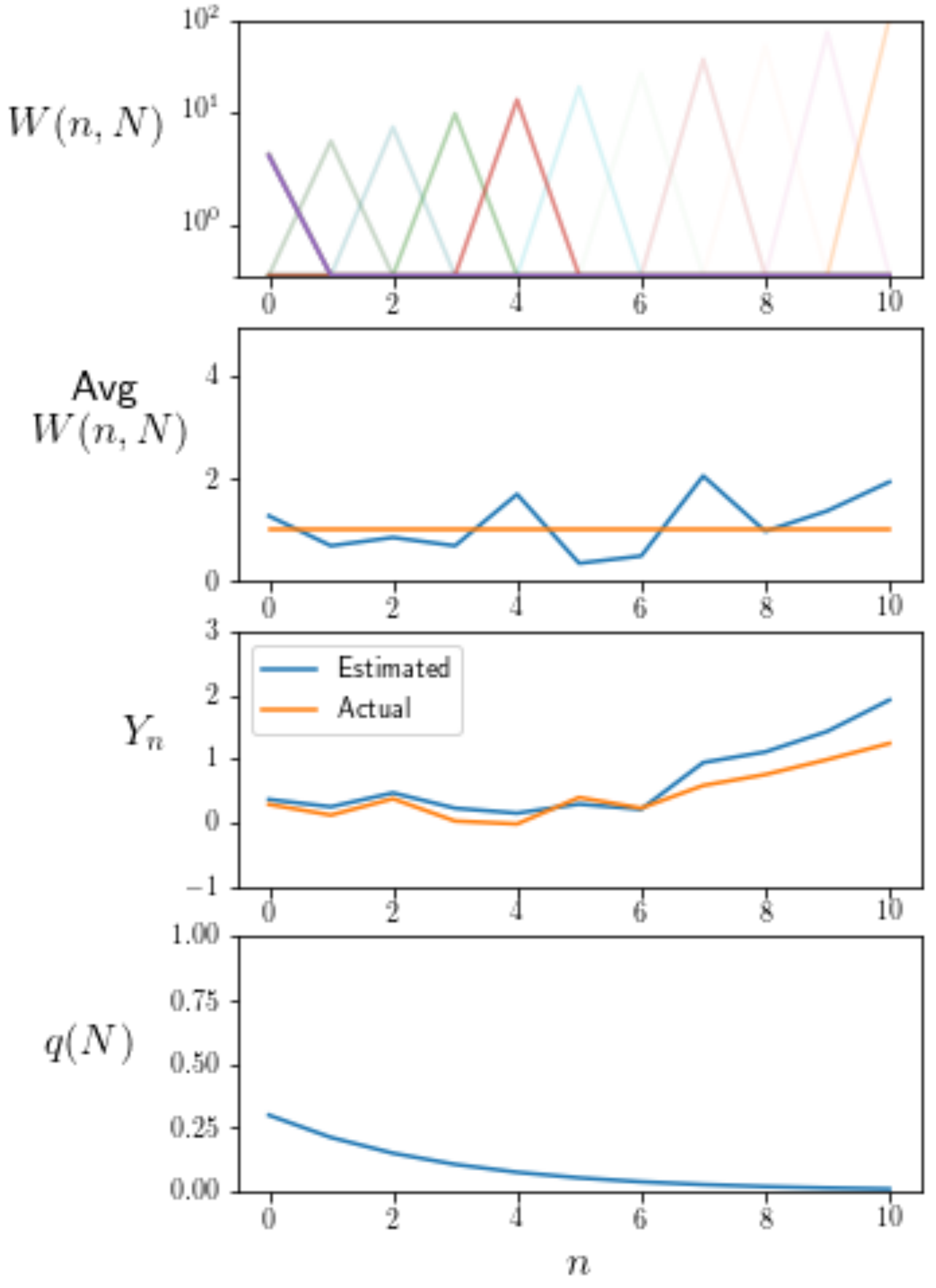
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



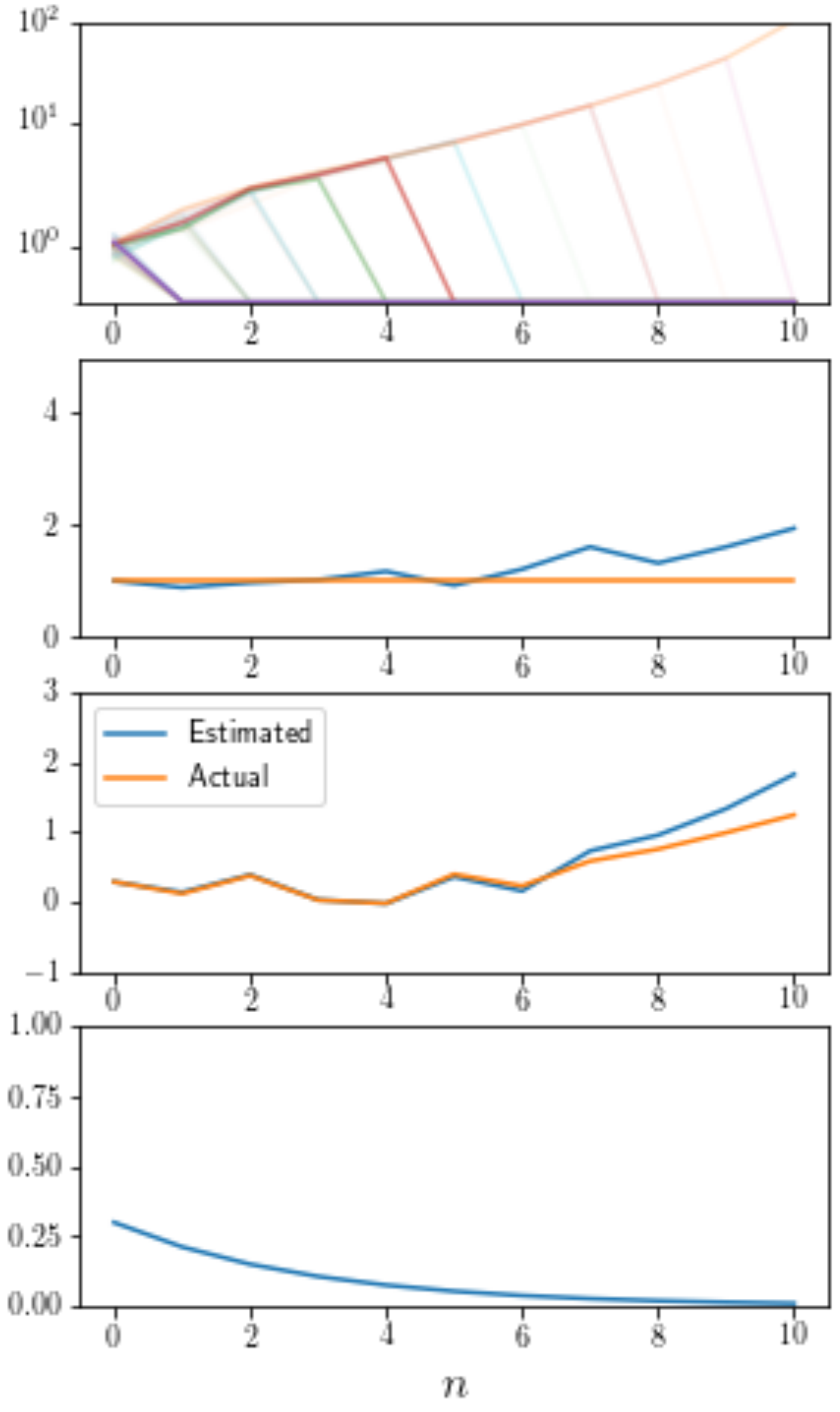
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



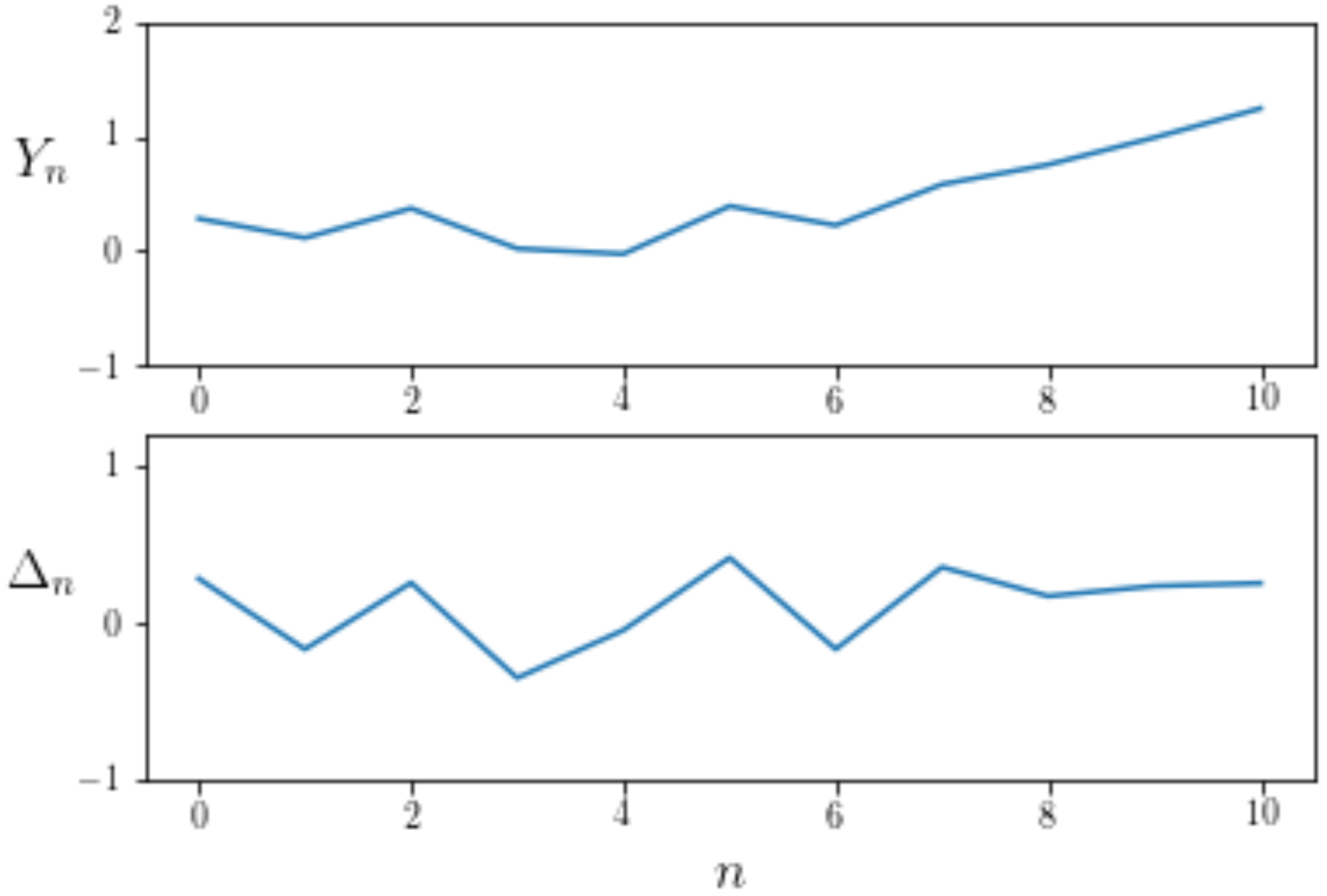
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth

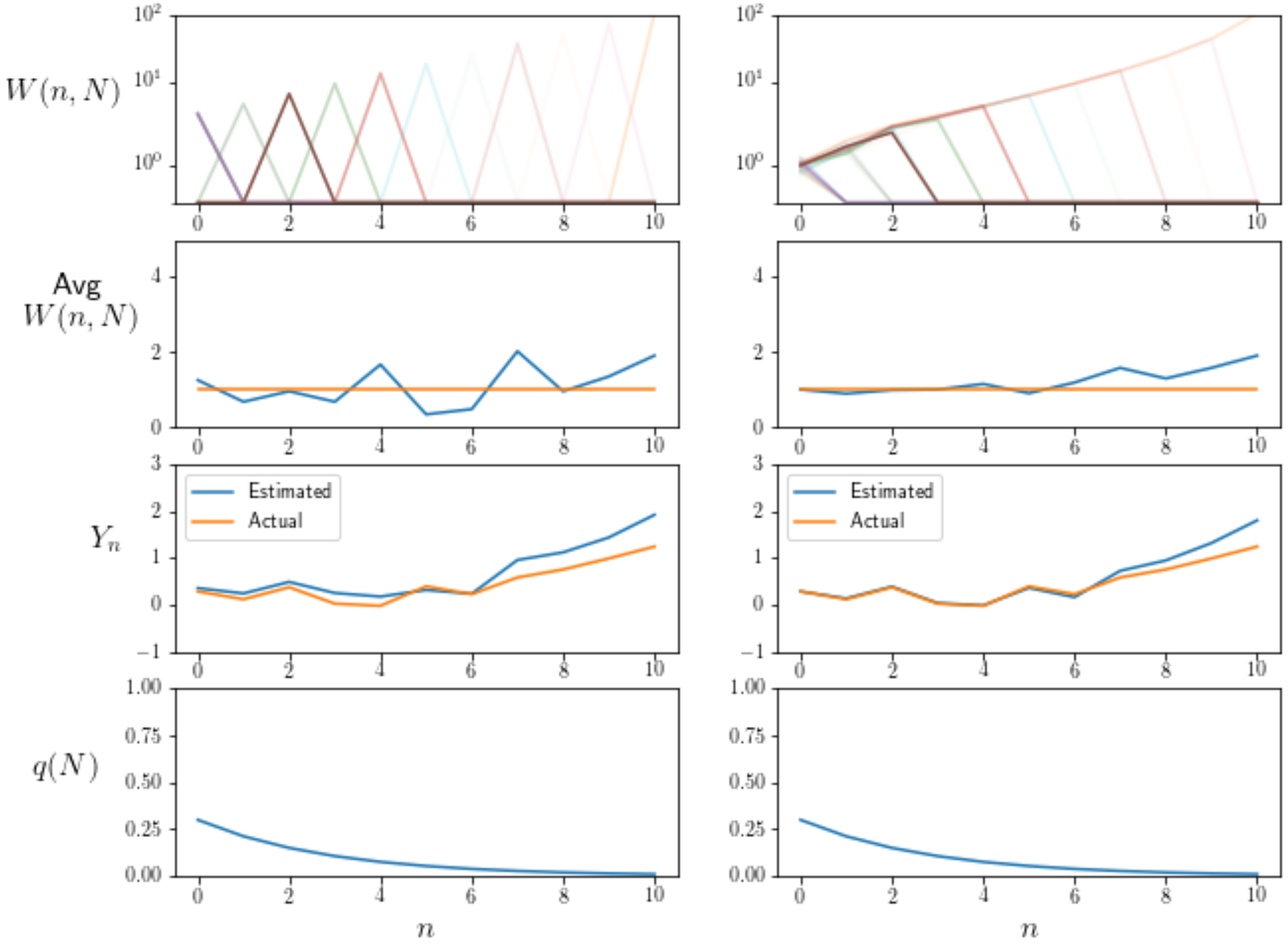


"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$

"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



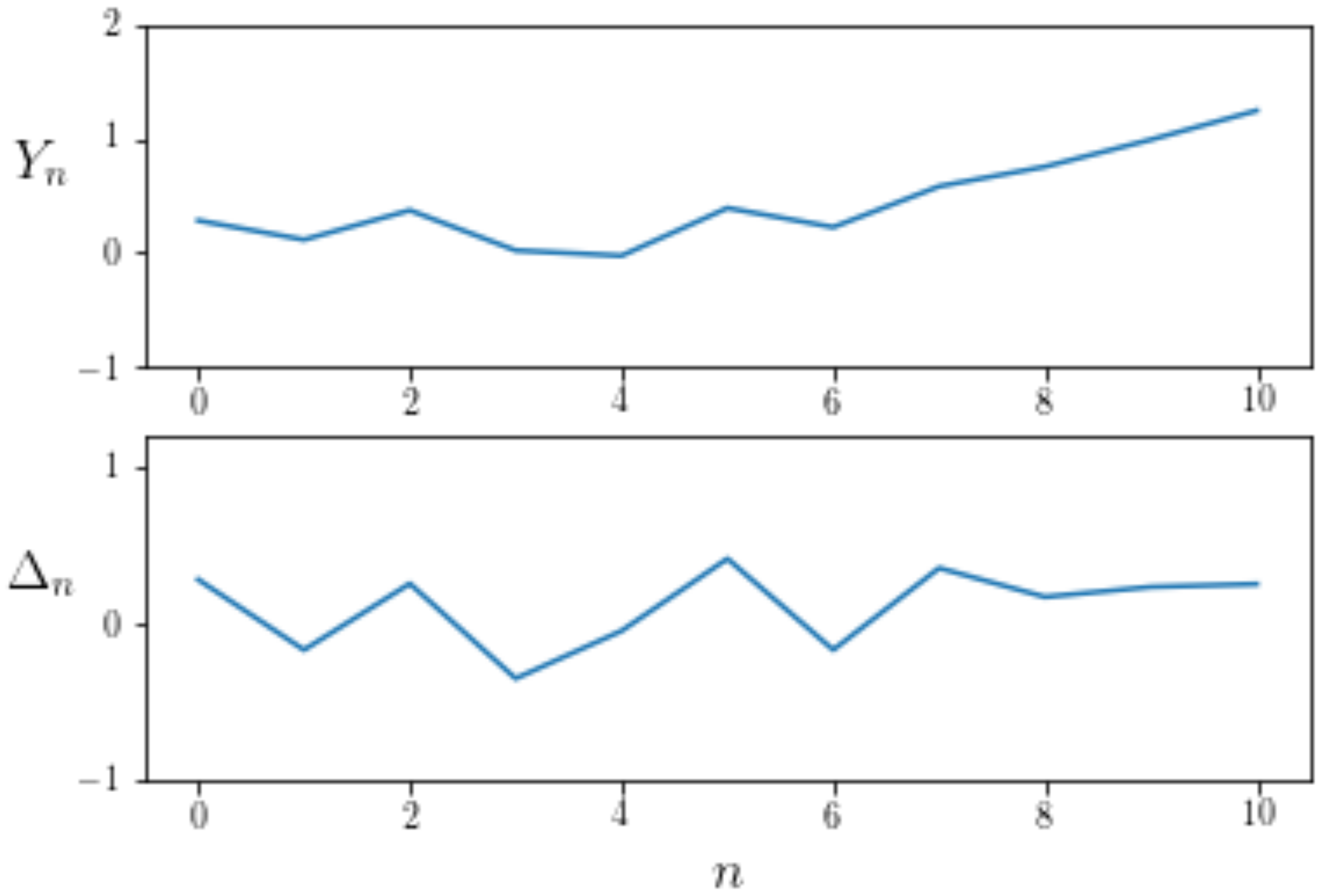
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

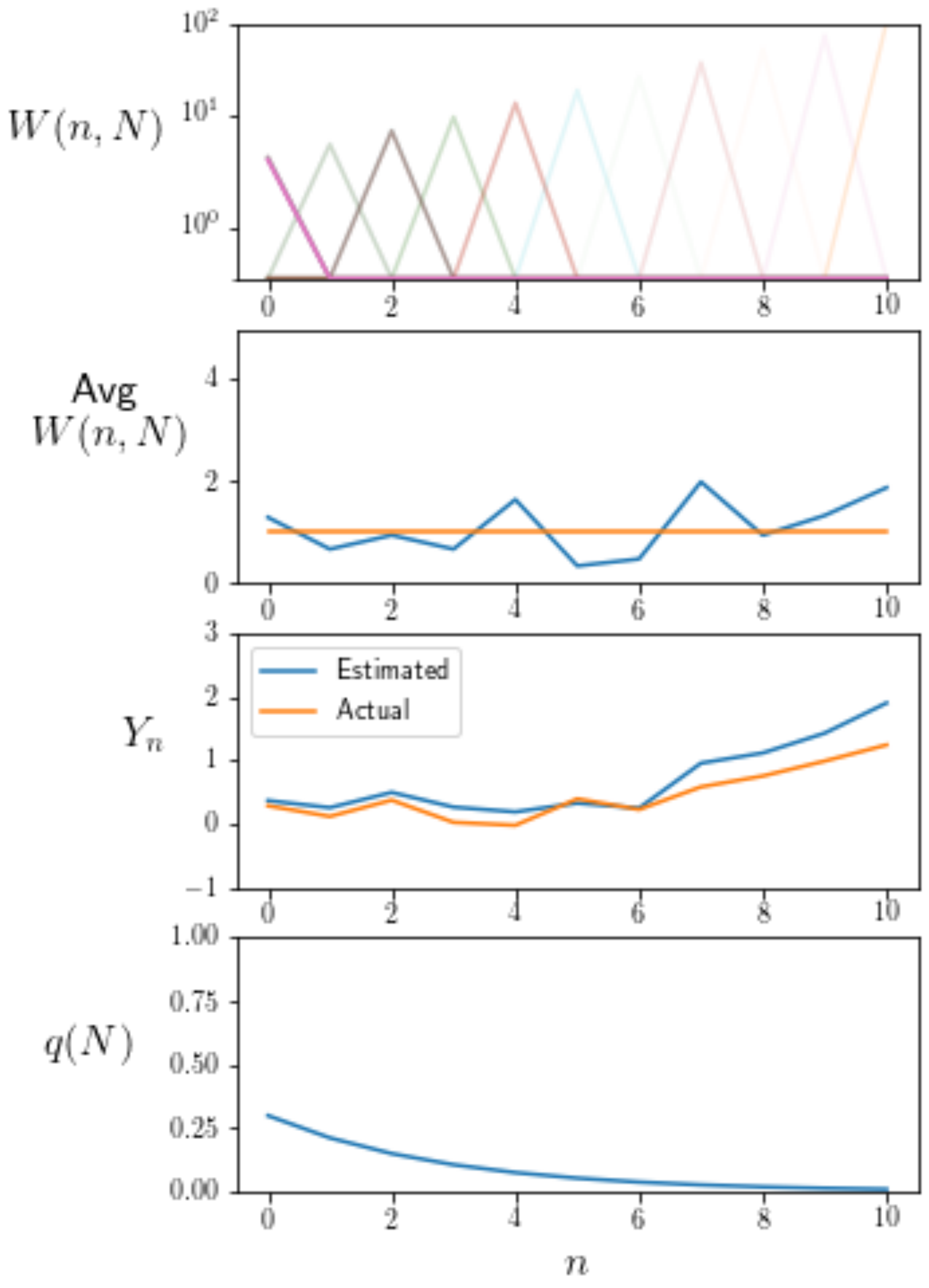
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



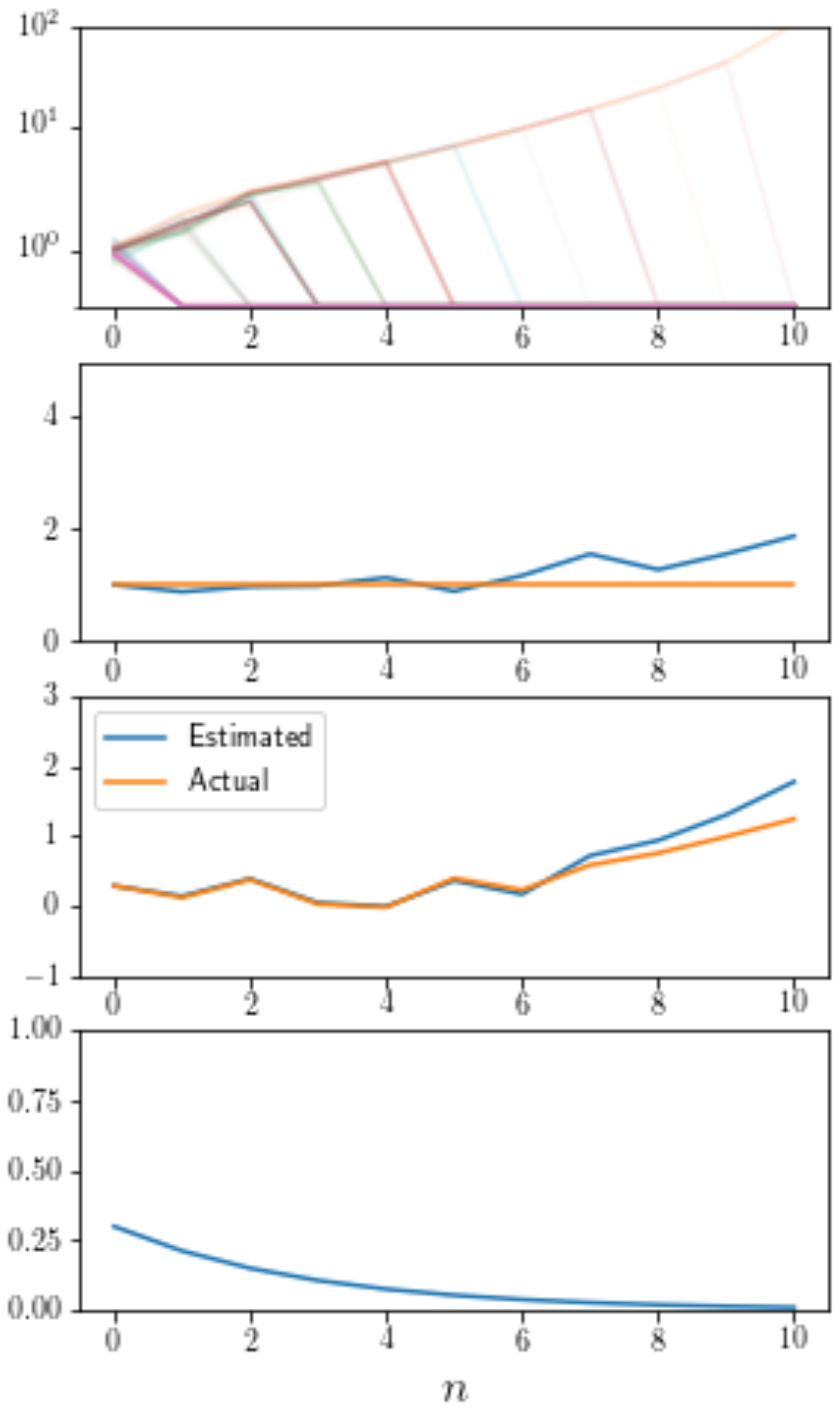
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



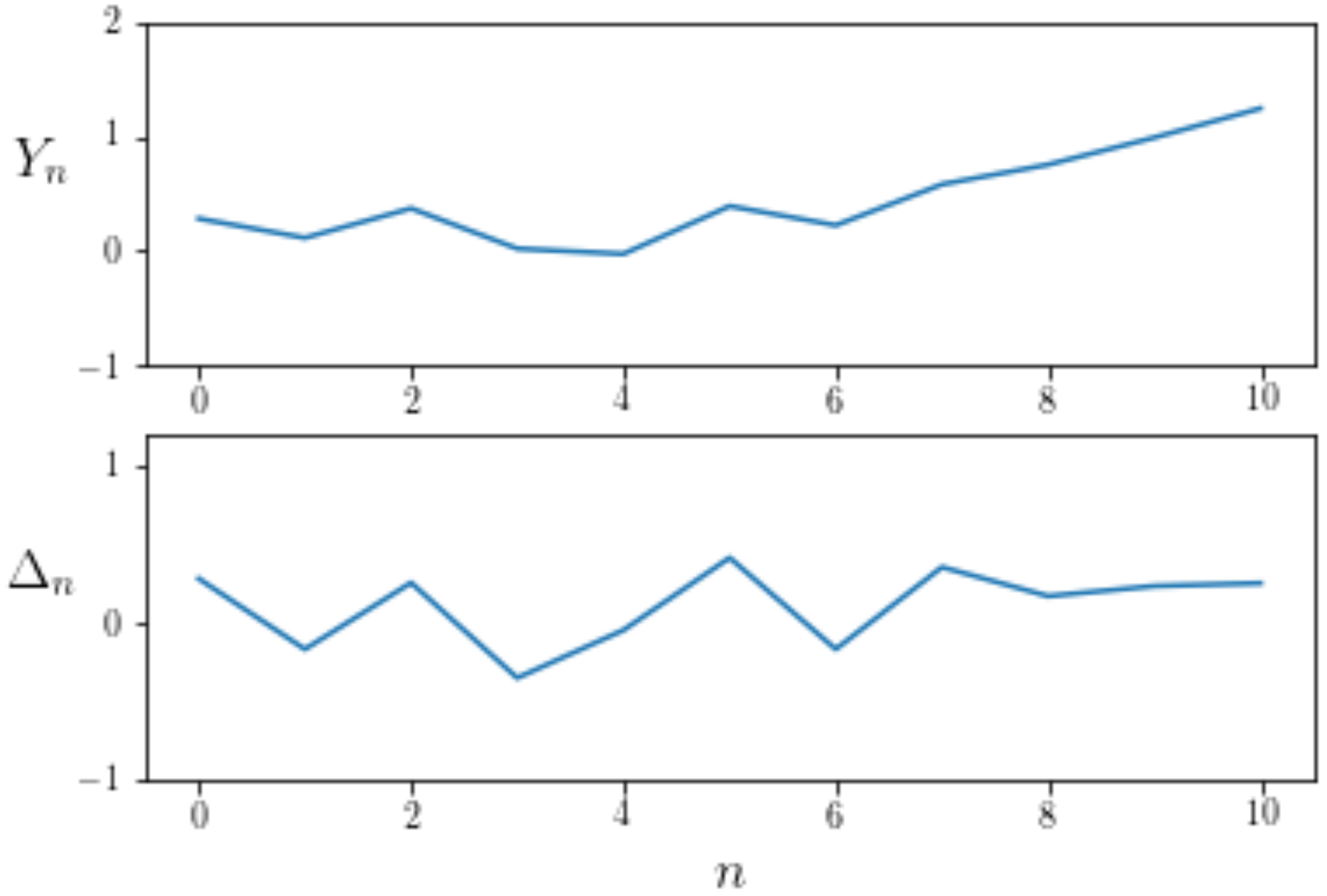
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

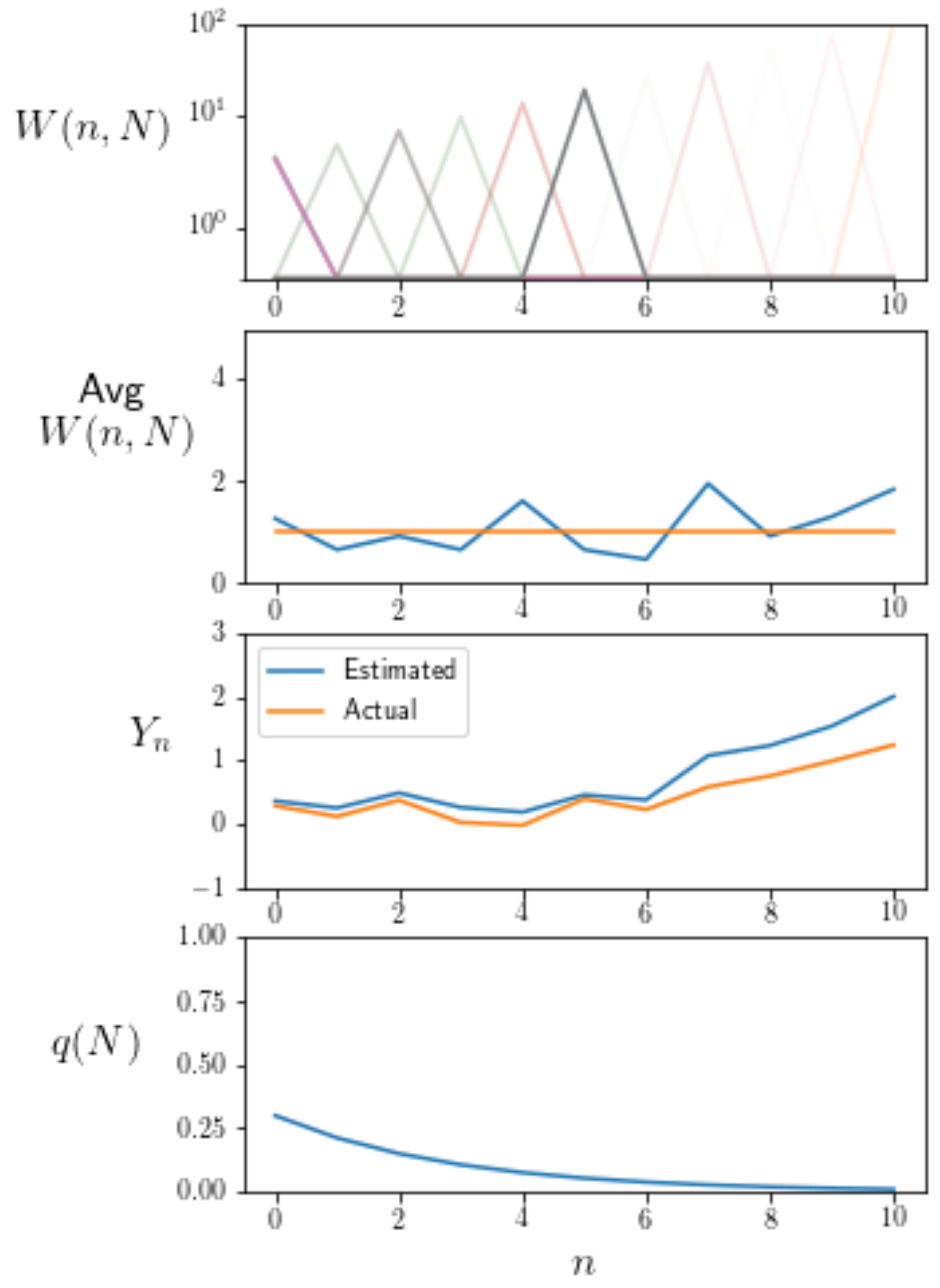
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



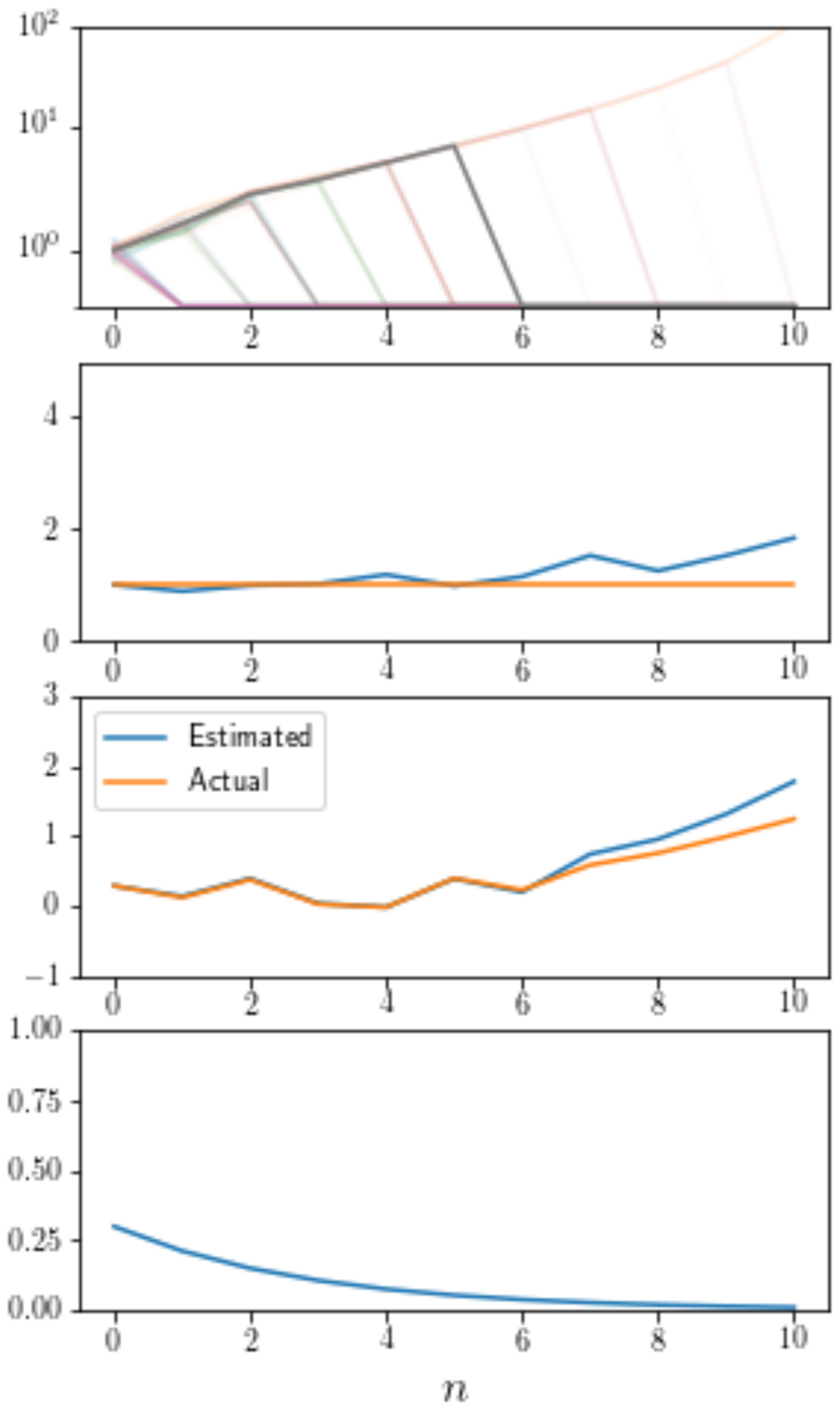
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



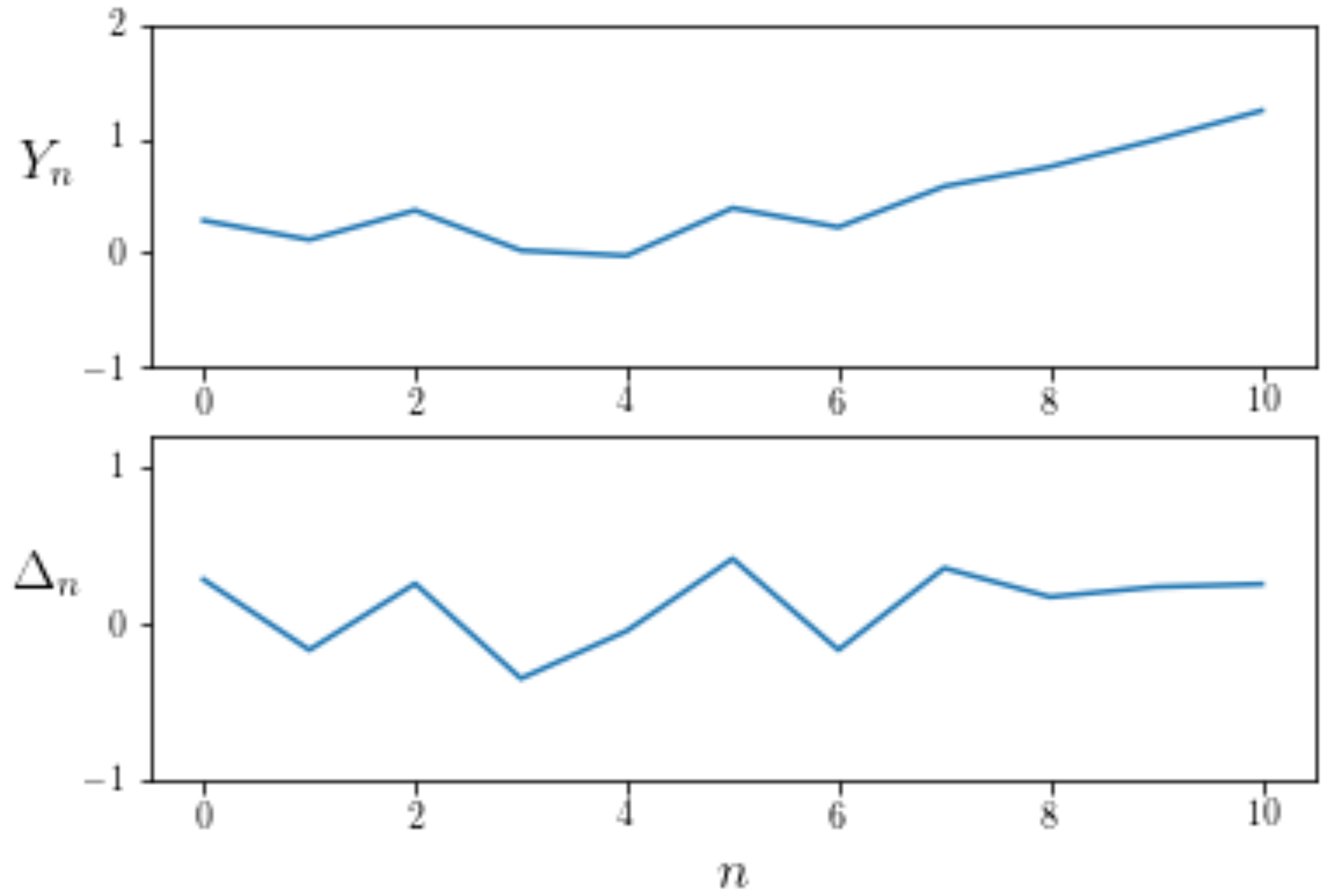
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \dots, H\} \sim q$$

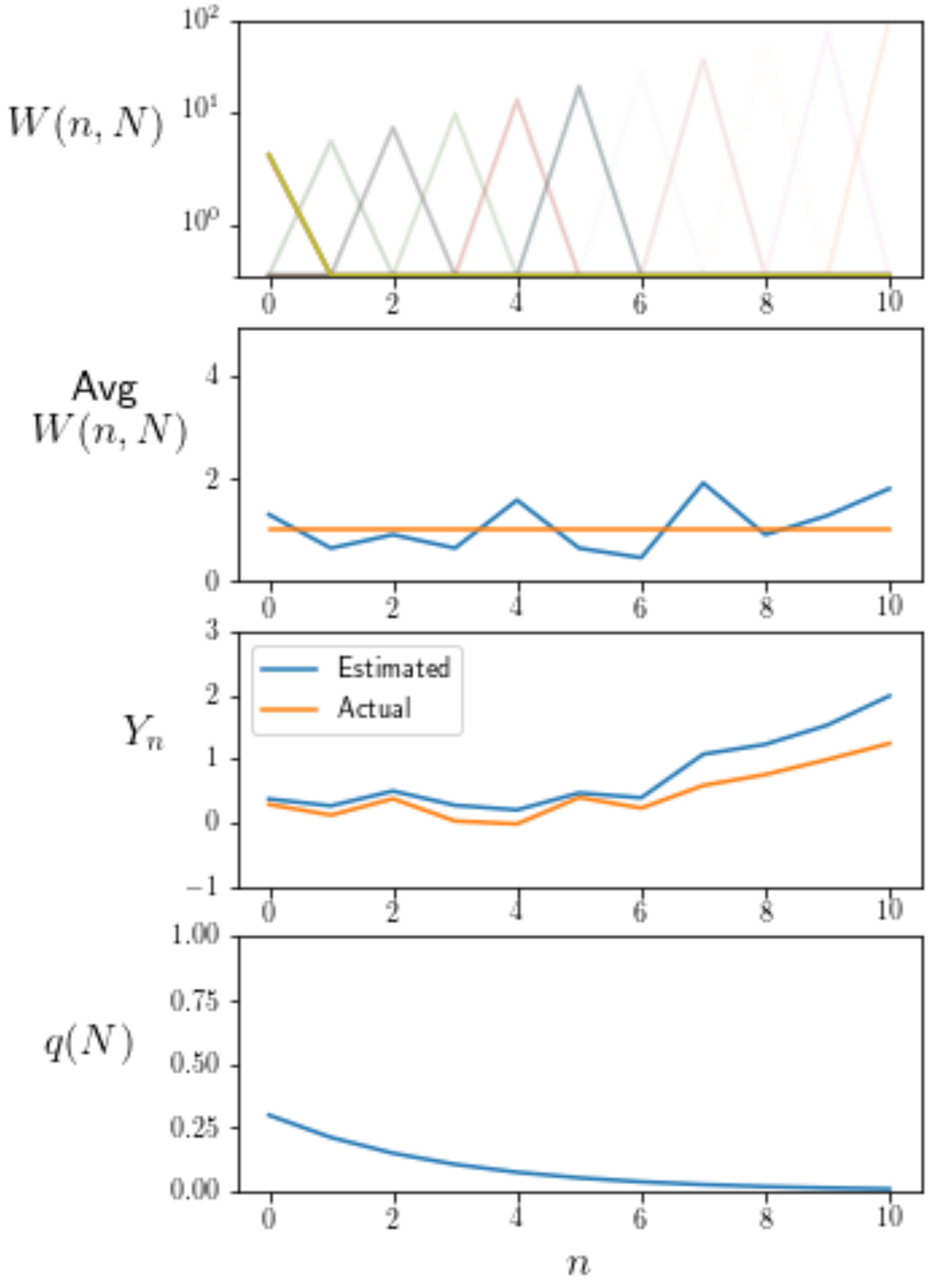
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



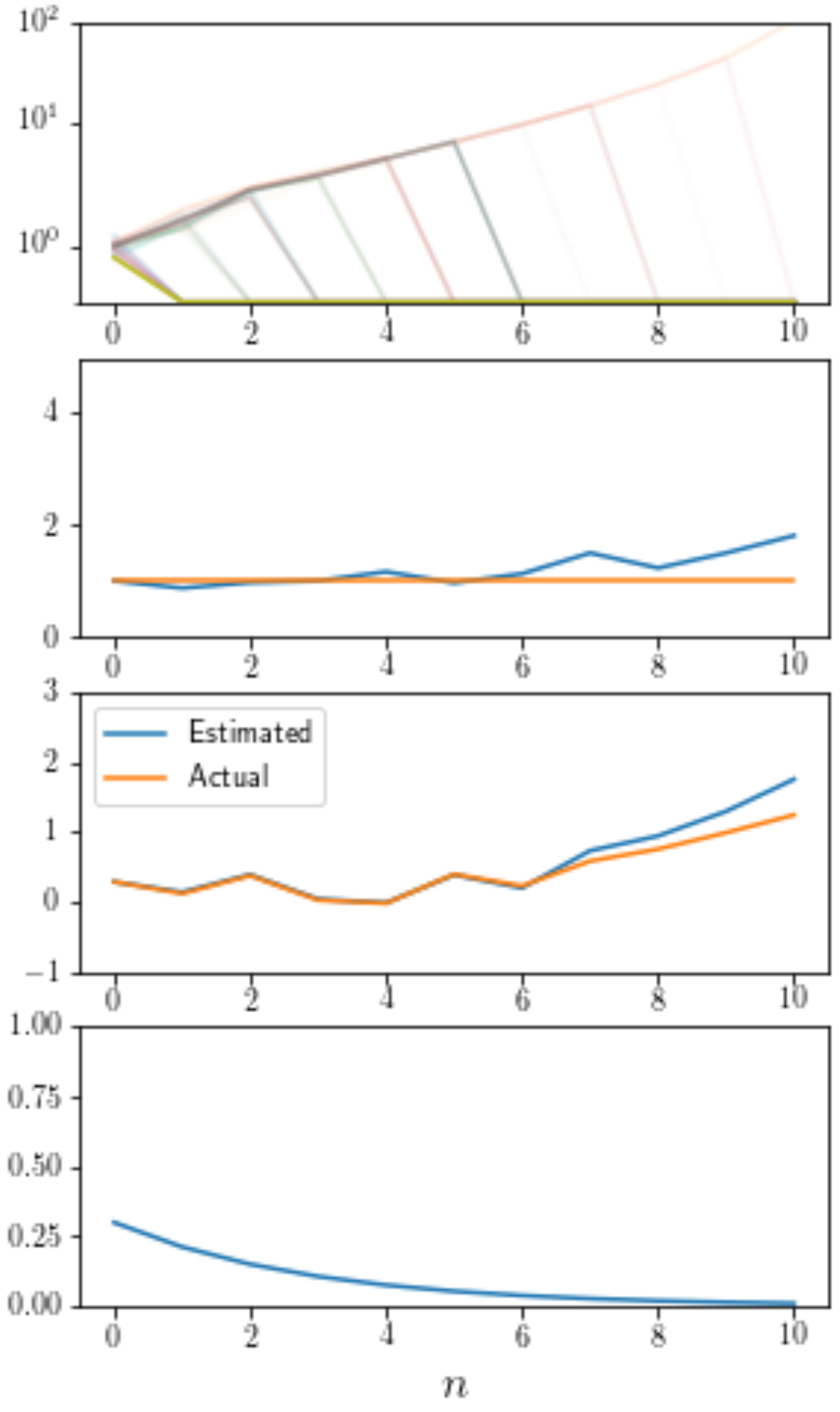
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



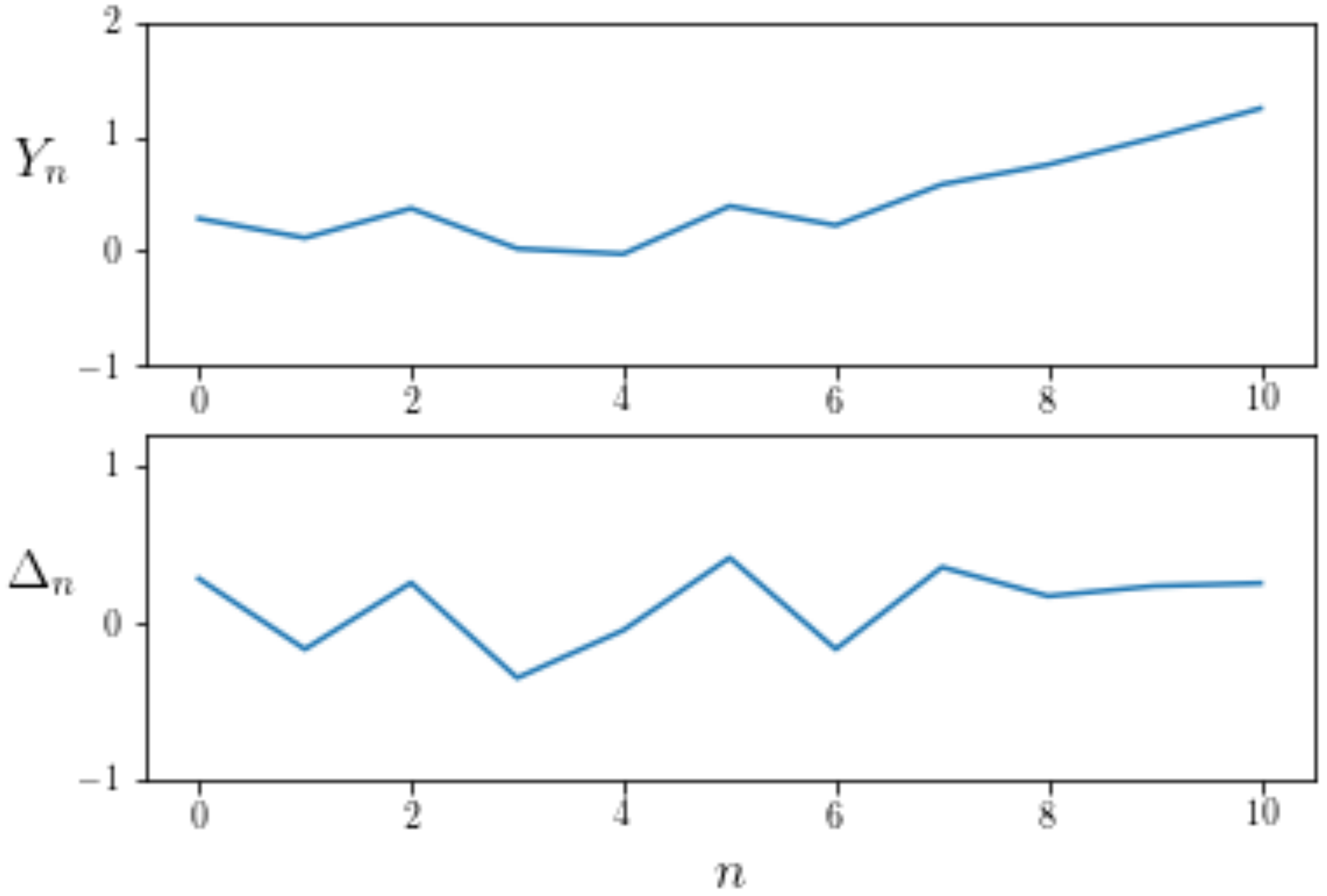
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

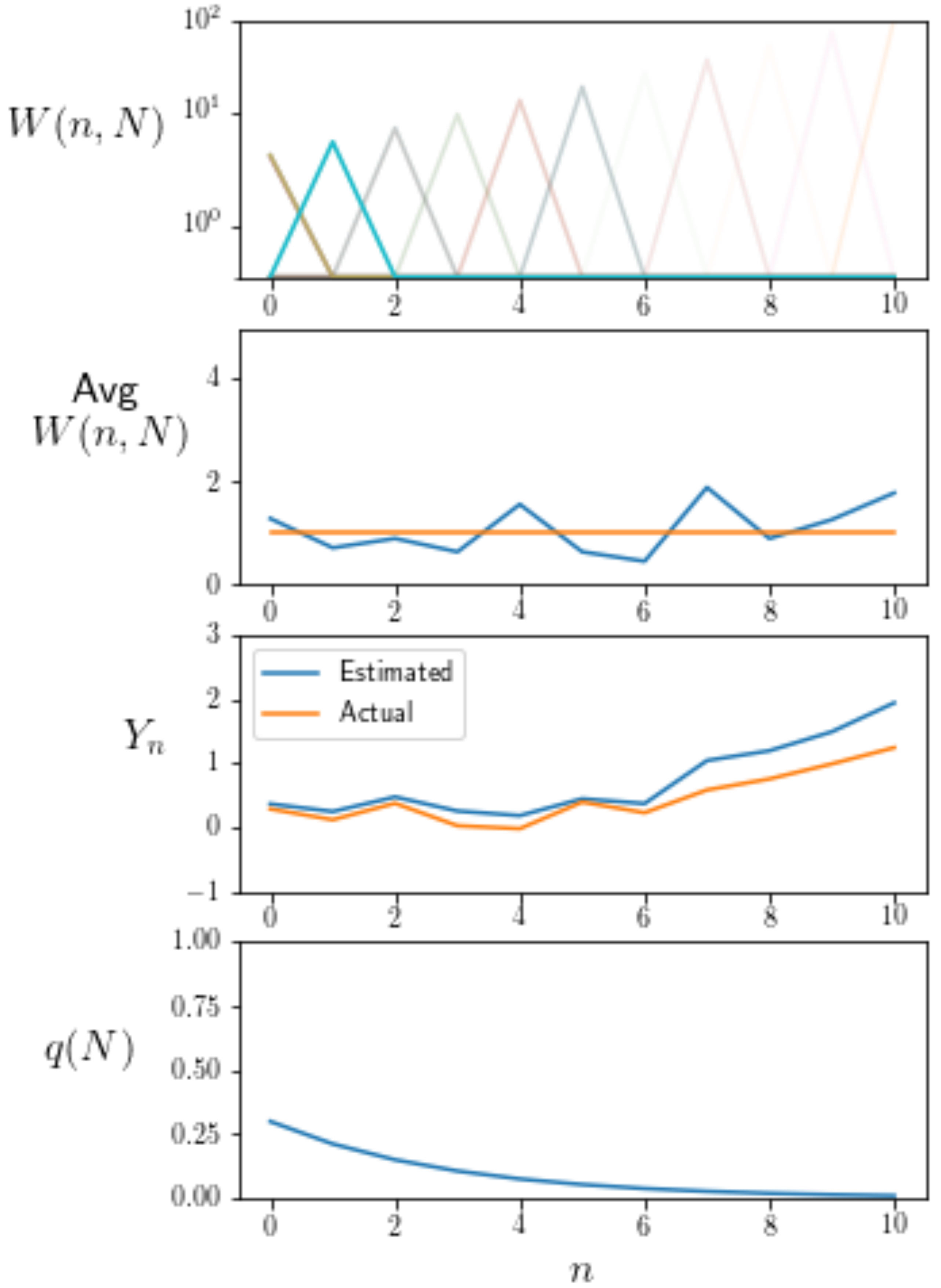
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



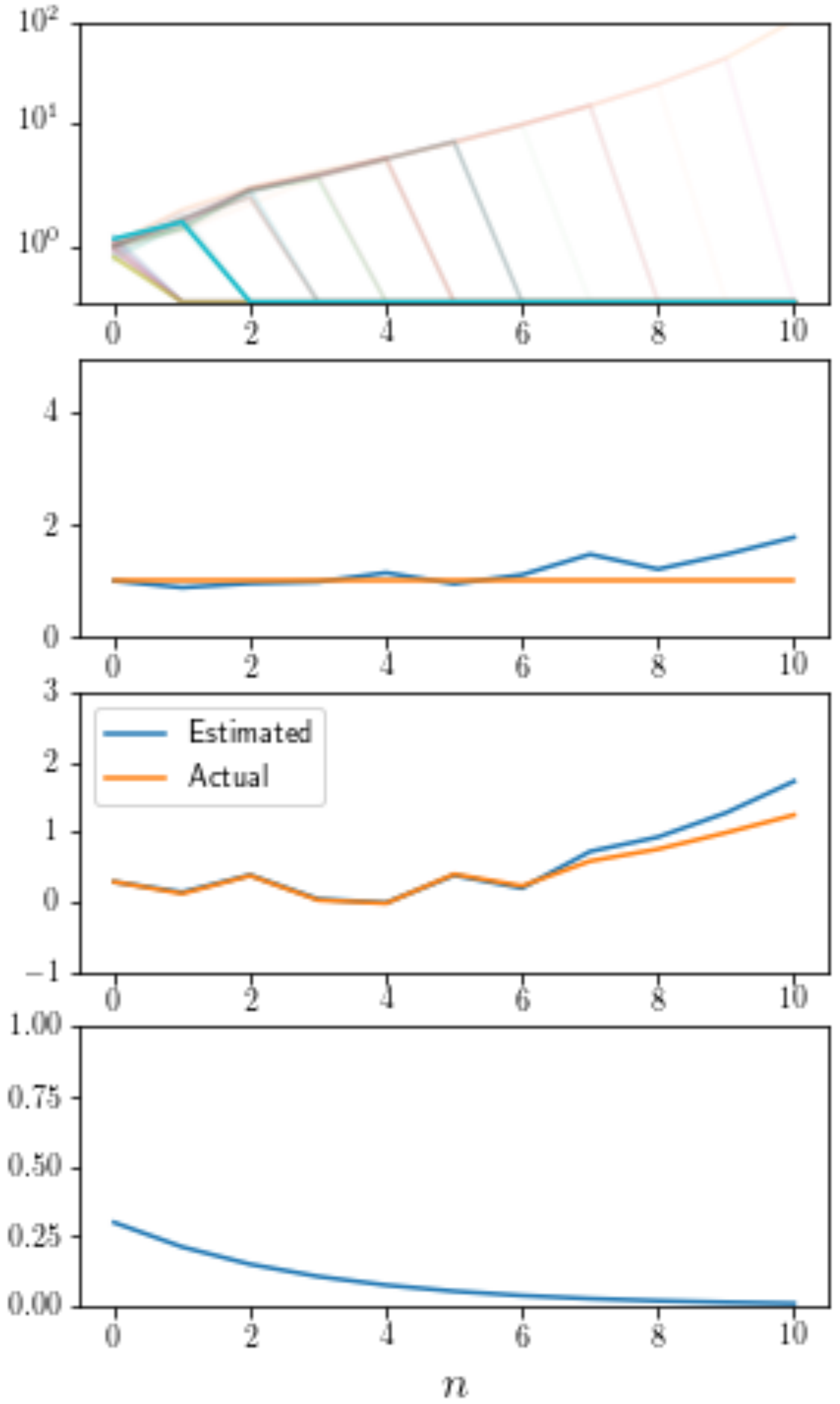
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



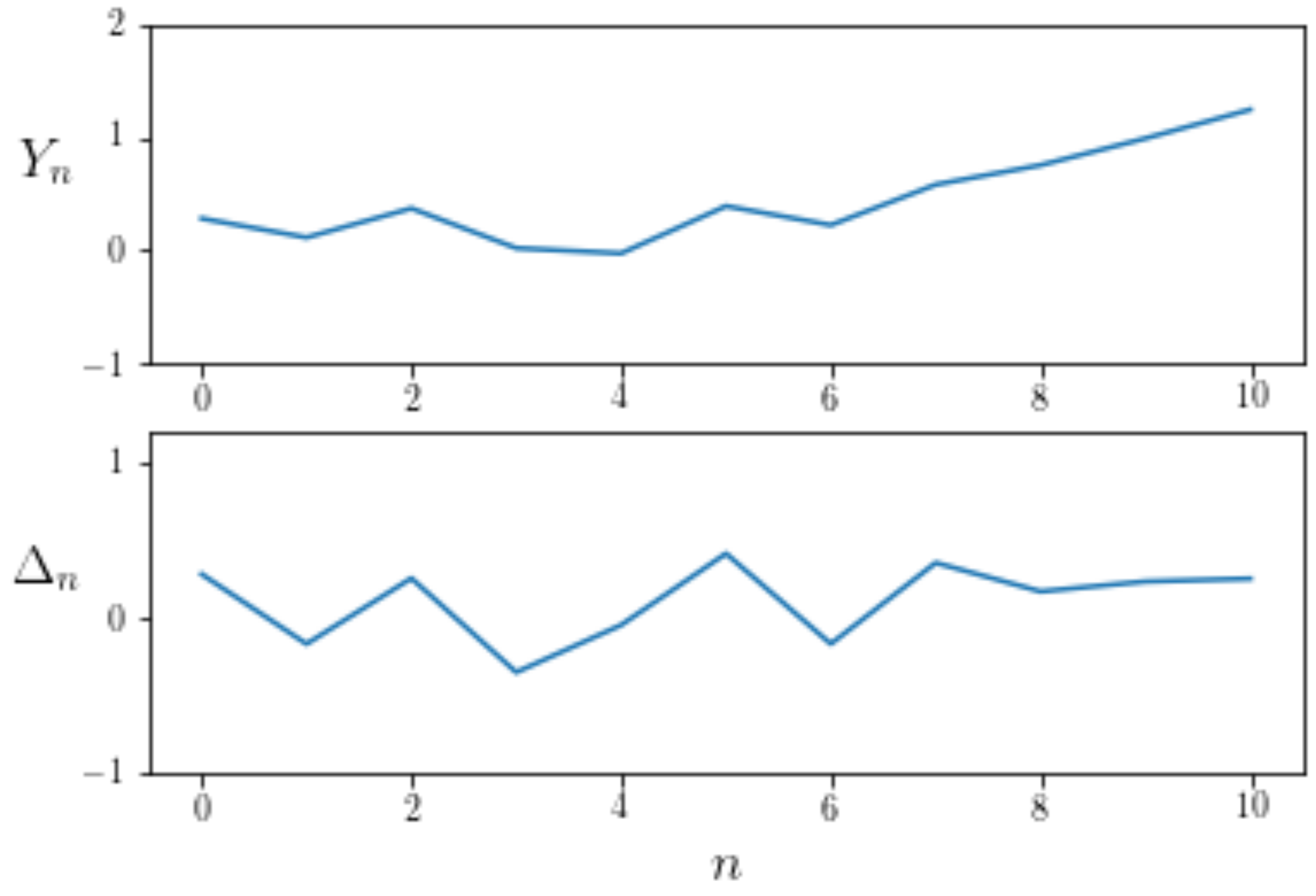
DEMONSTRATION

General form

$$\hat{Y}_H = \sum_{n=1}^{N} \Delta_n W(n, N) \qquad N \in \{1, \ldots, H\} \sim q$$

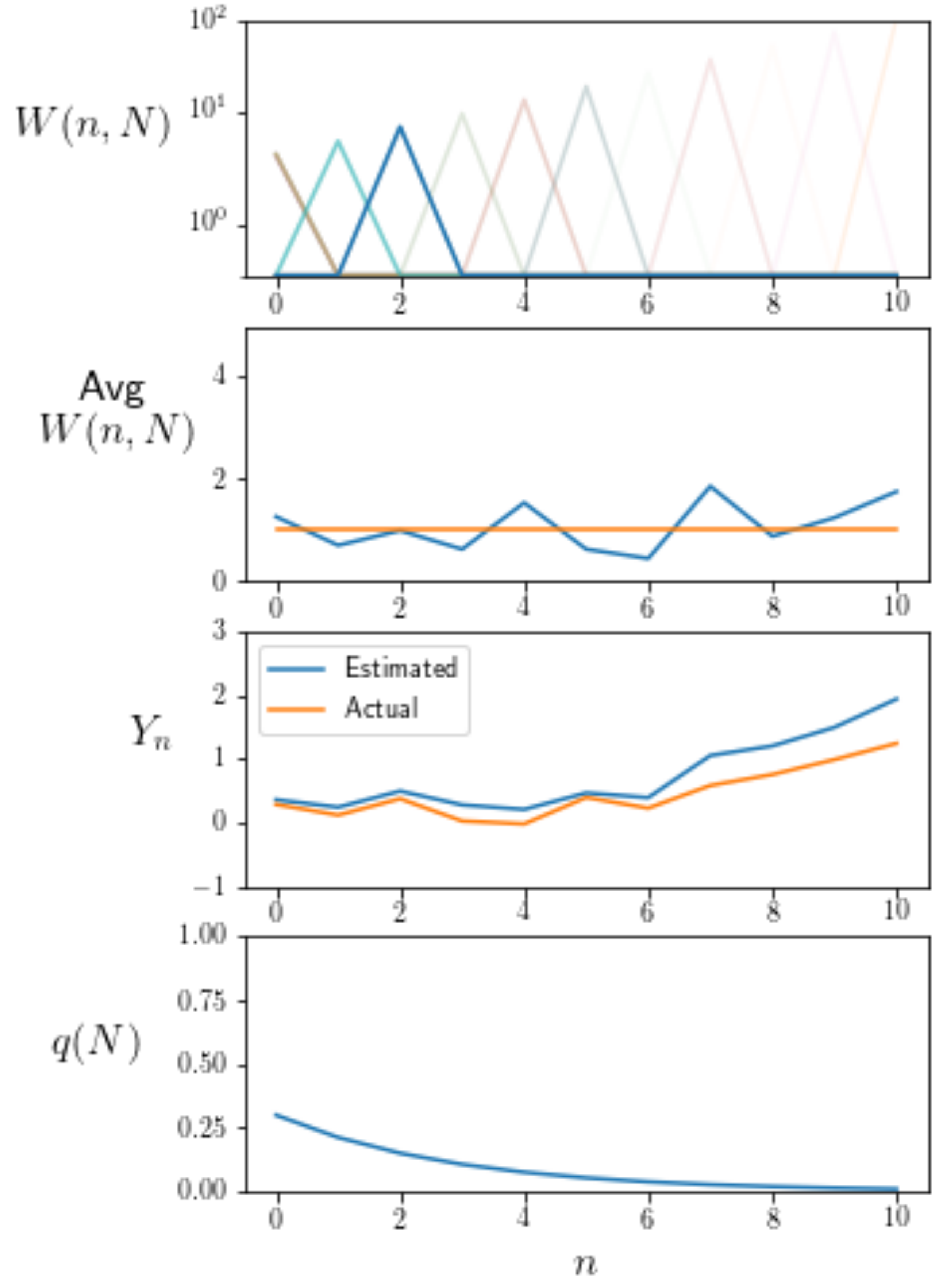
$$\Delta_n = \begin{cases} Y_n - Y_{n-1} & n > 1 \\ Y_1 & n = 1 \end{cases}$$

Ground truth



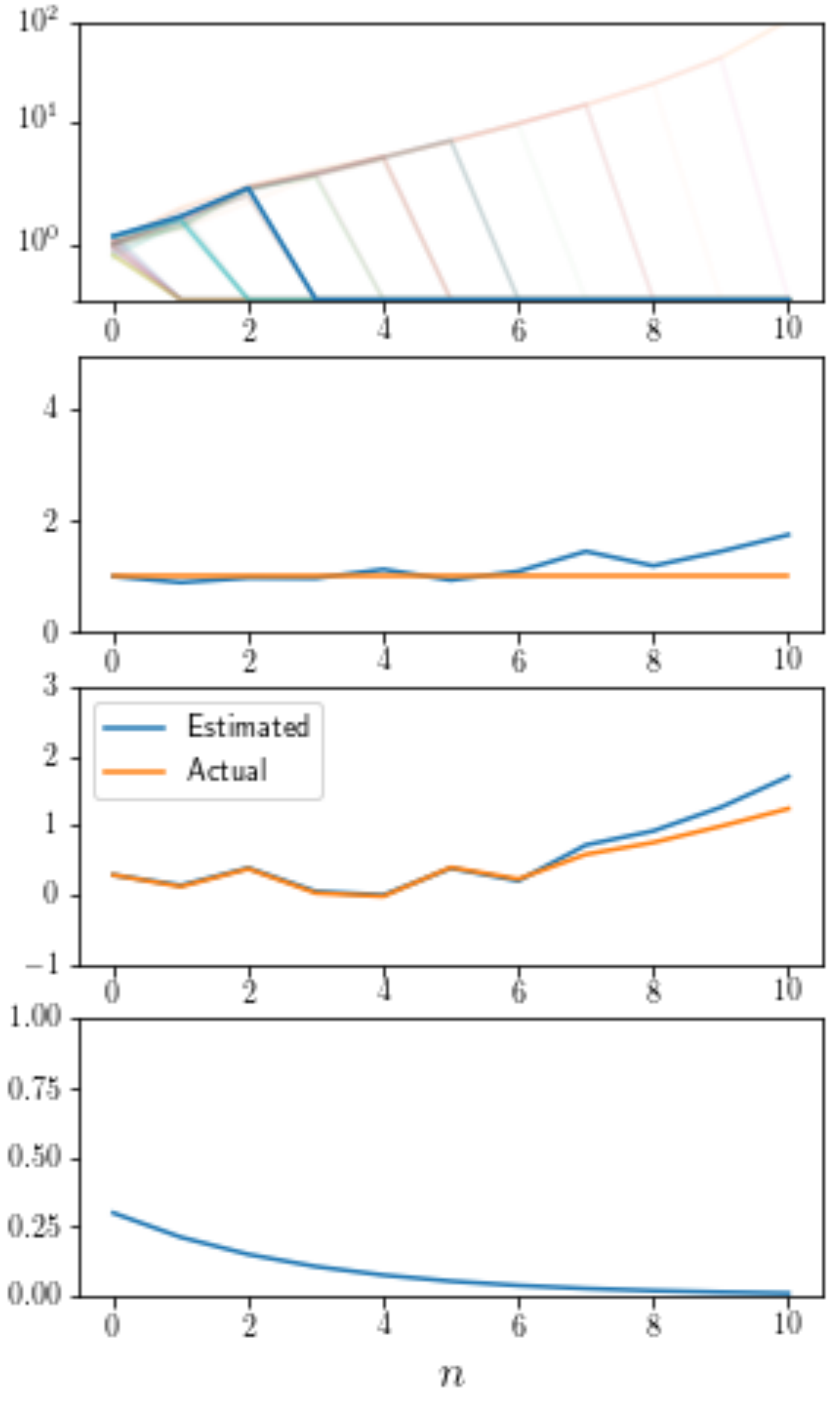
"Single sample"

$$W(n, N) = \frac{1}{q(N)} \mathbb{1}\{n = N\}$$



"Russian roulette"

$$W(n, N) = \frac{1}{1 - \sum_{n'=1}^{n-1} q(n')} \mathbb{1}\{N \geq n\}$$



VISIT THE POSTER FOR…

- Randomized Telescopes for optimization = SGD for limits
- Conditions for finite computation and variance
- Provable convergence for convex optimization of infinite loops and limits
- Adapting sampling and learning rate online to balance computation and variance, and maximize optimization efficiency
- Experiments: meta-optimization of learning rates, variational inference of ODE parameters, optimizing RNNs

THANKS!