Did you forget the constraint that the probability distribution adds to 1?

You're correct. In deriving the probability distribution, I neglected to explicitly enforce the normalization constraint that the probability distribution $p(\epsilon)$ must integrate to 1. Let's correct this.

Step 1: Revisiting the Maximum Entropy Framework

The goal is to find the probability distribution $p(\epsilon)$ that maximizes entropy under the noisy constraints and the normalization constraint.

The entropy $H(p)$ is:

$$H(p) = -\int p(\epsilon) \log p(\epsilon) \, d\epsilon$$

The constraints are:

1. $E[\epsilon_i] = \mu_\epsilon$
2. $E[\epsilon_i^2] = \sigma^2 + \delta_\epsilon^2$
3. $E[x_i \cdot \epsilon_i] = \gamma$
4. Normalization: $\int p(\epsilon) \, d\epsilon = 1$

Step 2: Lagrange Function with Normalization

We introduce Lagrange multipliers λ_1, λ_2, λ_3, and λ_0 (for the normalization constraint). The Lagrange function is:

$$\mathcal{L} = -\int p(\epsilon) \log p(\epsilon) \, d\epsilon + \lambda_0 \left(\int p(\epsilon) \, d\epsilon - 1 \right) + \lambda_1 \left(\int p(\epsilon)\epsilon \, d\epsilon - \mu_\epsilon \right) + \lambda_2 \Big($$

Step 3: Functional Derivative

To find the probability distribution $p(\epsilon)$, take the functional derivative of $\mathcal{L}$ with respect to $p(\epsilon)$ and set it to zero:

$$\frac{\delta \mathcal{L}}{\delta p(\epsilon)} = -\log p(\epsilon) - 1 + \lambda_0 + \lambda_1 \epsilon + \lambda_2 \epsilon^2 + \lambda_3 \epsilon x = 0$$