1. Parameterizations ('pauli' and 'upper') for Hermitian matrix of size N×N

- pauli
  - Generalized Pauli basis of length N^2
  - One real coefficient per basis matrix
  - Total real parameters: N^2
- upper
  - $N(N-1)/2$ complex upper-triangular entries
  - N real diagonal entries
  - Total real parameters: N^2
- Both parametrizations perform similarly

2. Matrix training methods

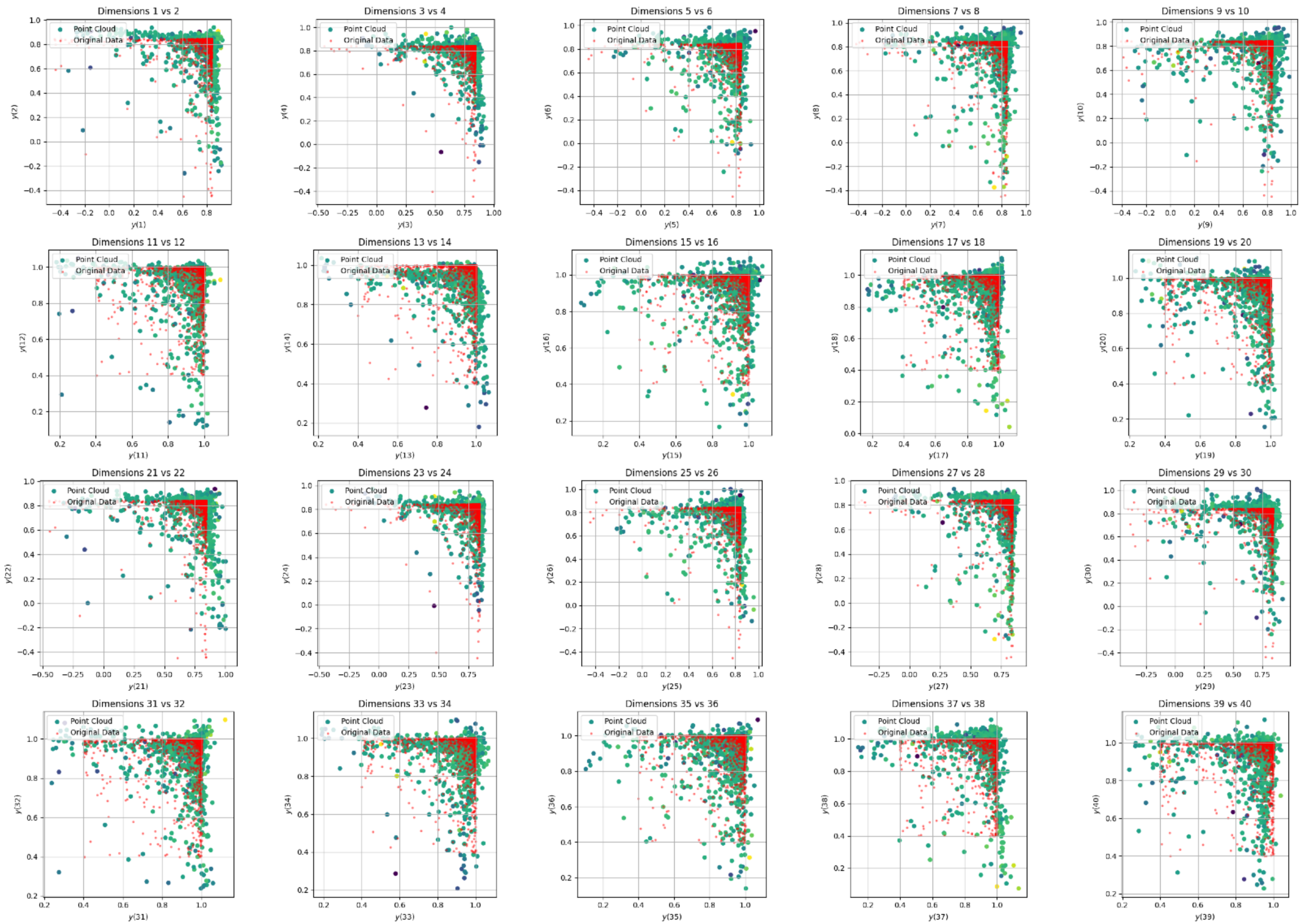
- Custom loop (Optax): L2 regularization + Gradient clipping + Learning-rate scheduler
- Line-search minimization (JAXOPT LBFGS)
- Analytic gradient computation + simple descent

The following plots use the LBFGS solver.

M_beta : E_dim = 40, I_dim = 10

A_init = 1

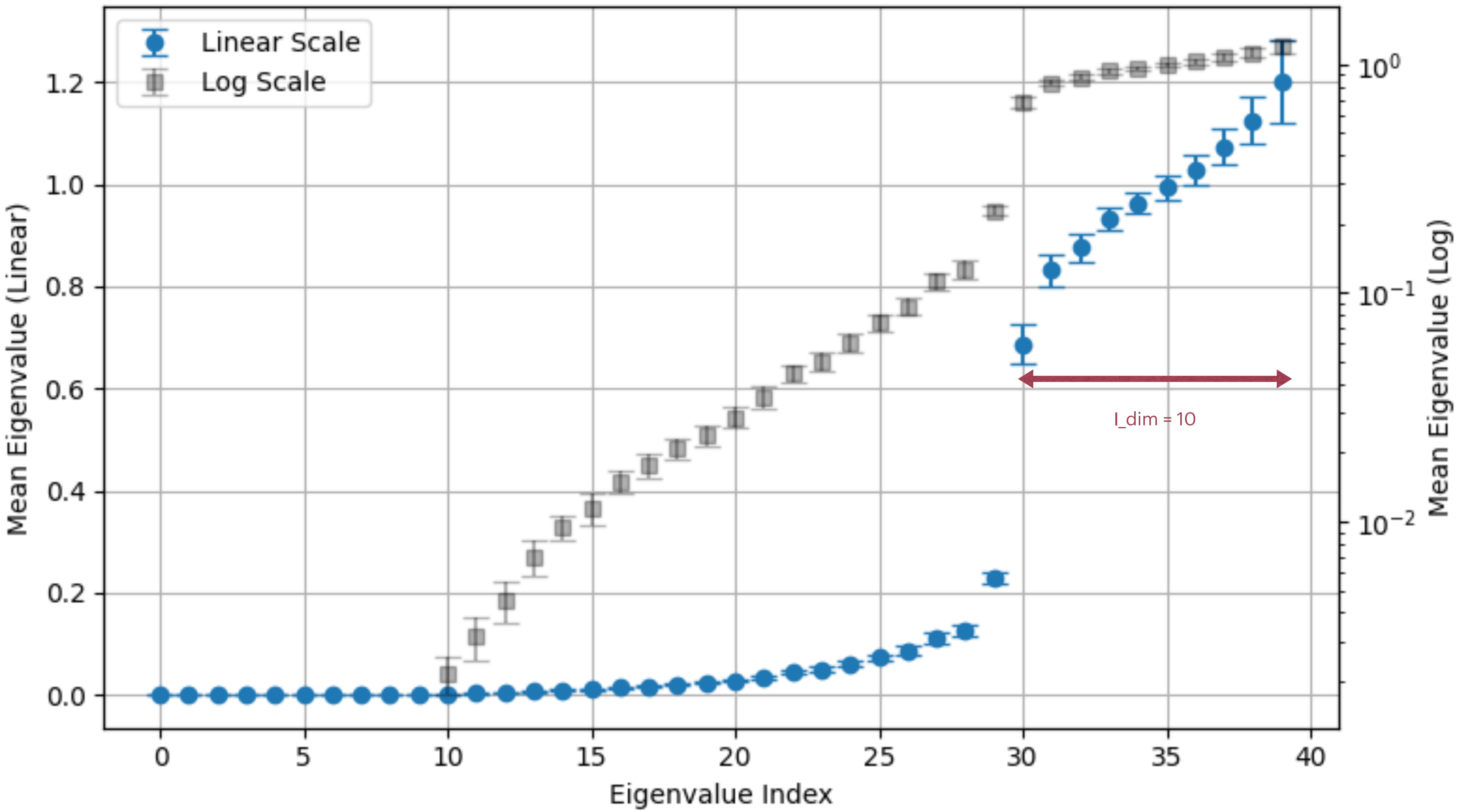
noise = 0.0



M_beta : E_dim = 40, I_dim = 10

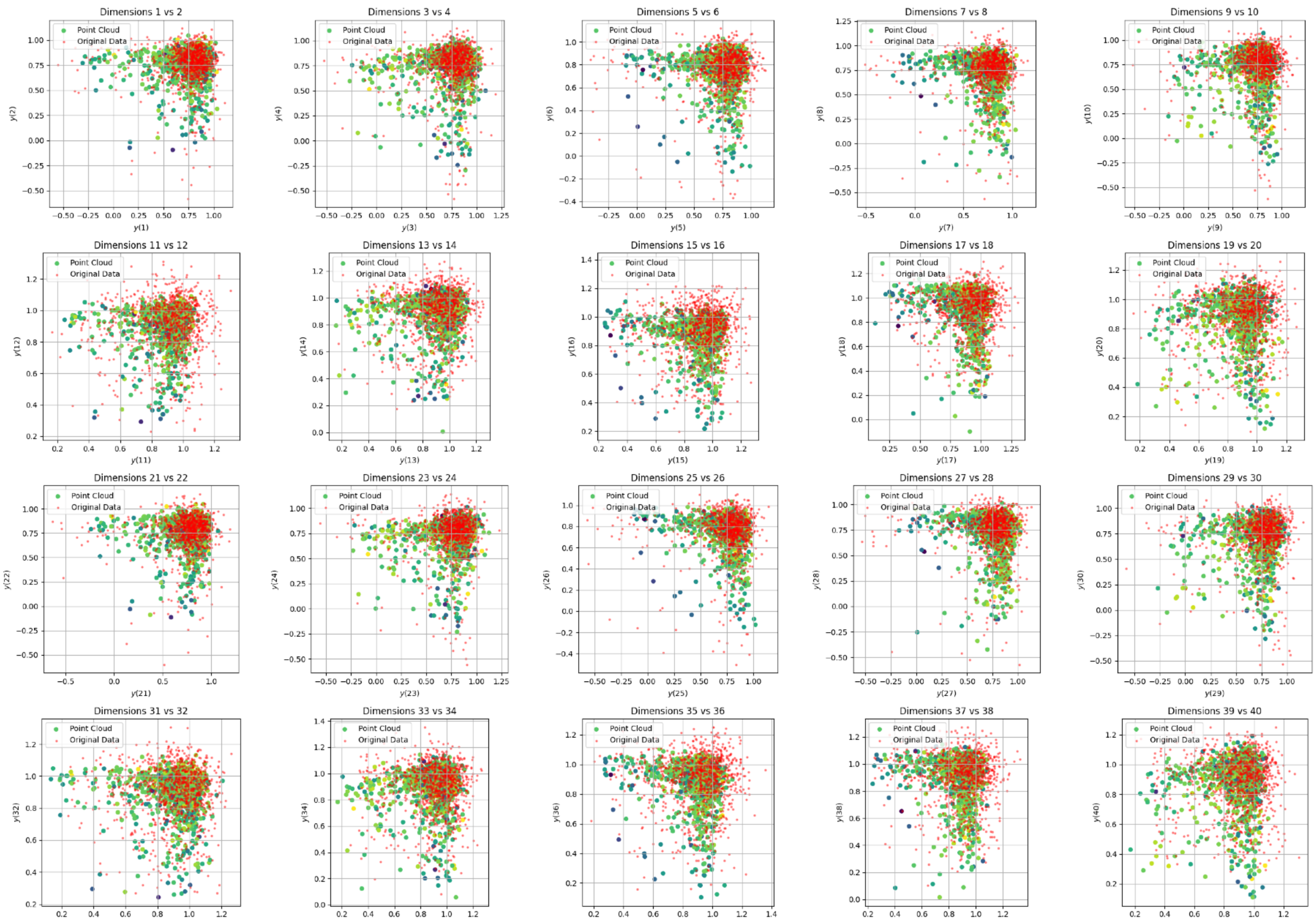
A_init = 1

noise = 0.0



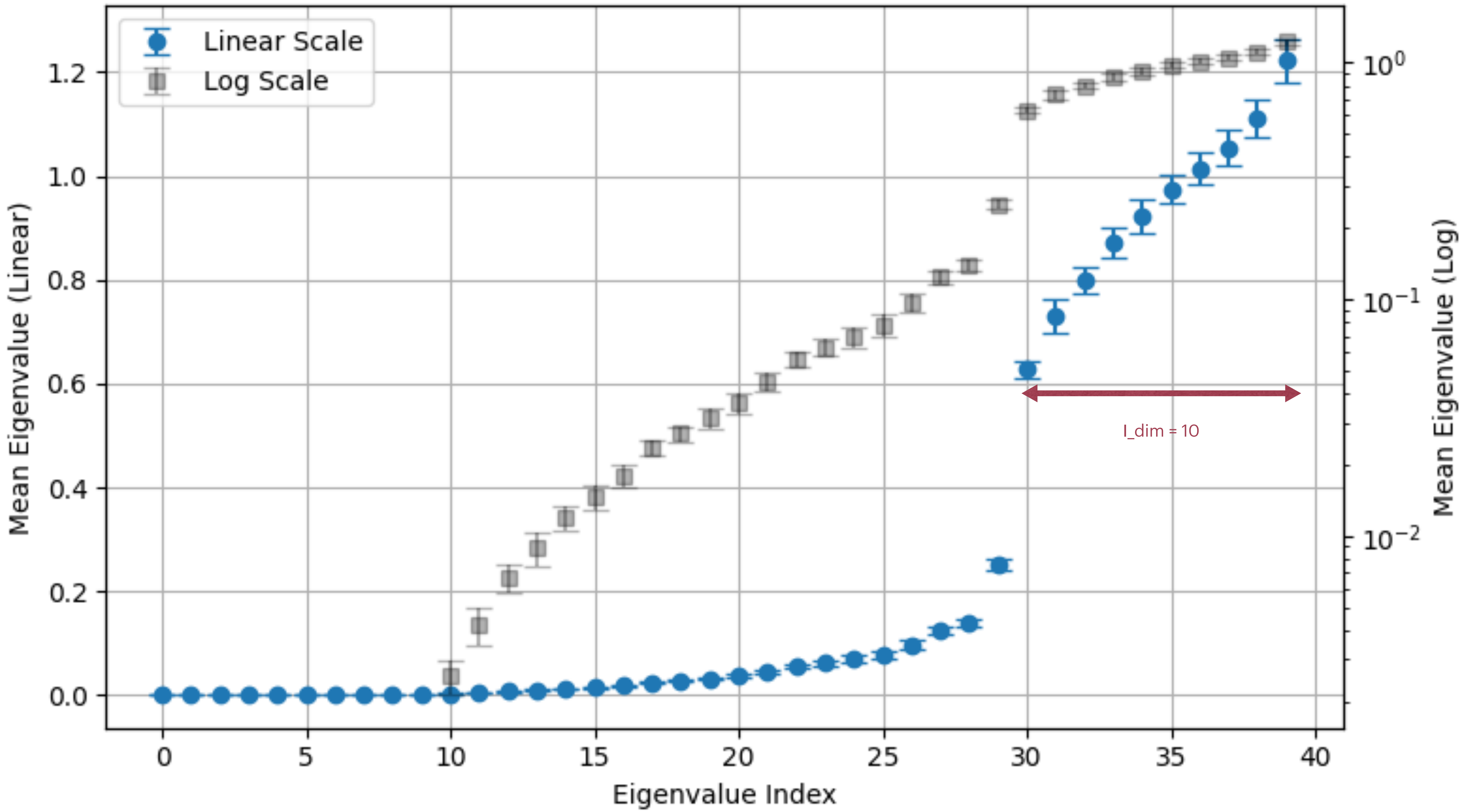
M_beta : E_dim = 40, I_dim = 10
A_init = 1

noise = 0.1



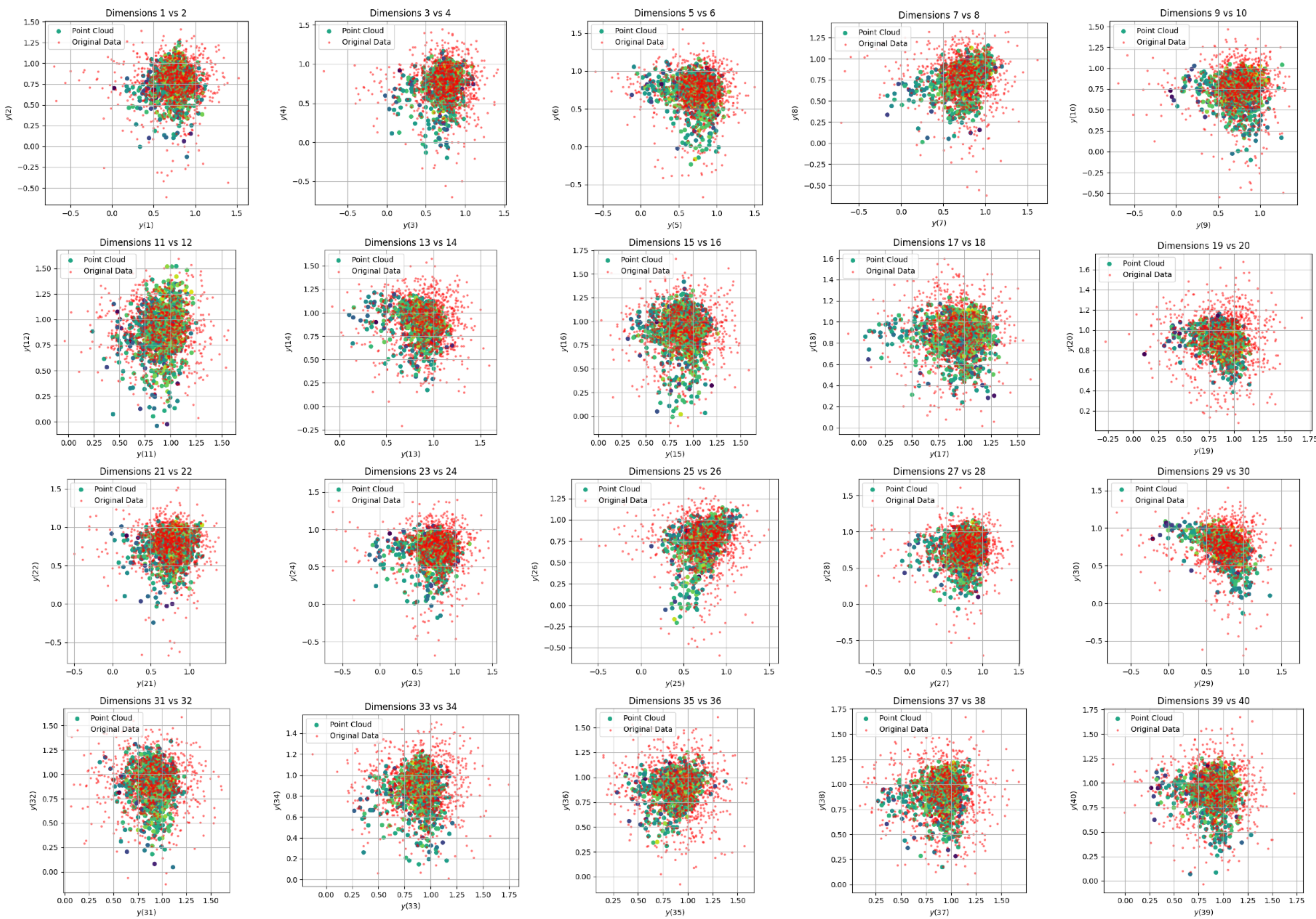
M_beta : E_dim = 40, I_dim = 10
A_init = 1

noise = 0.1



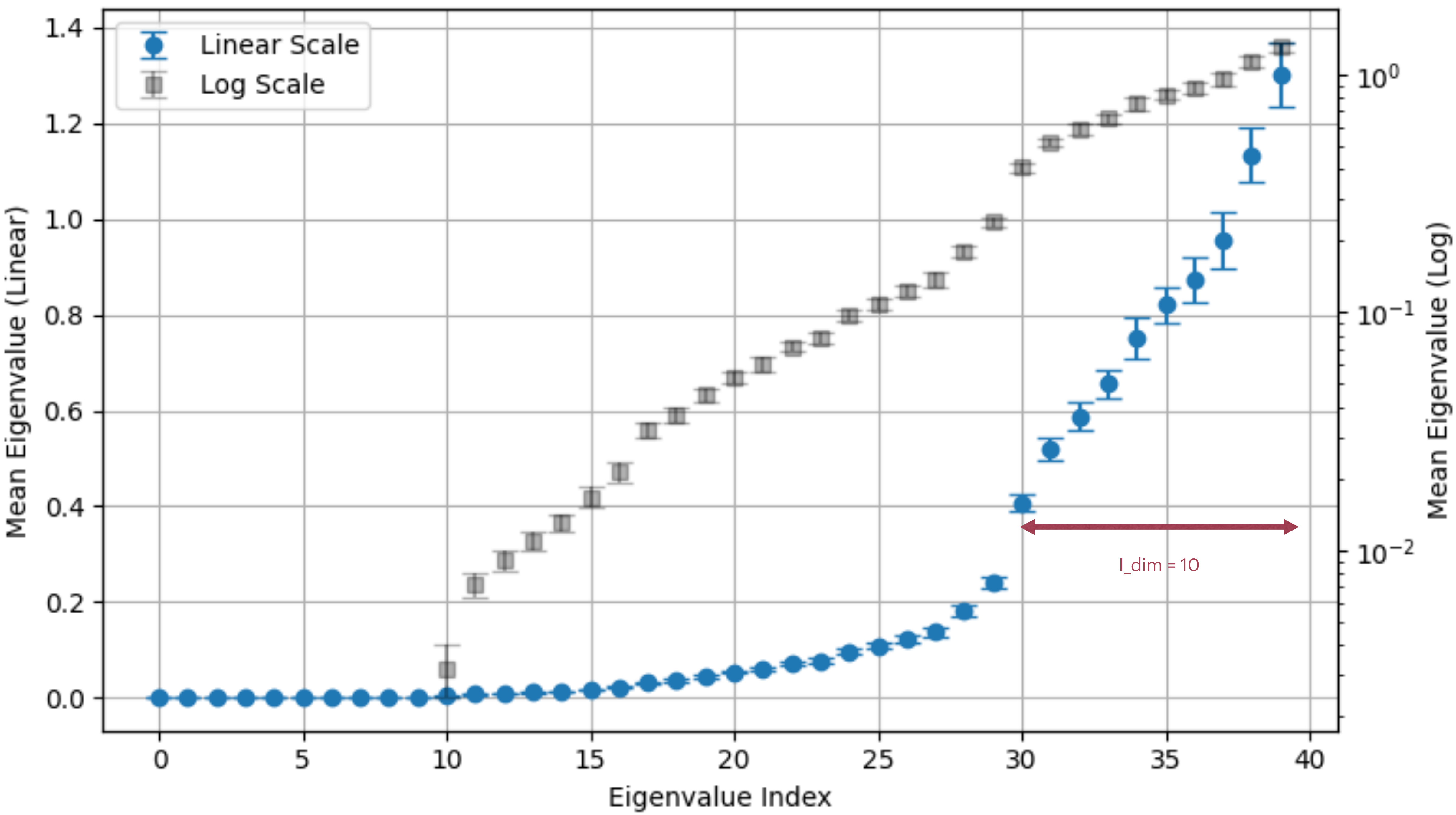
M_beta : E_dim = 40, I_dim = 10
A_init = 2

noise = 0.2



M_beta : E_dim = 40, I_dim = 10
A_init = 2

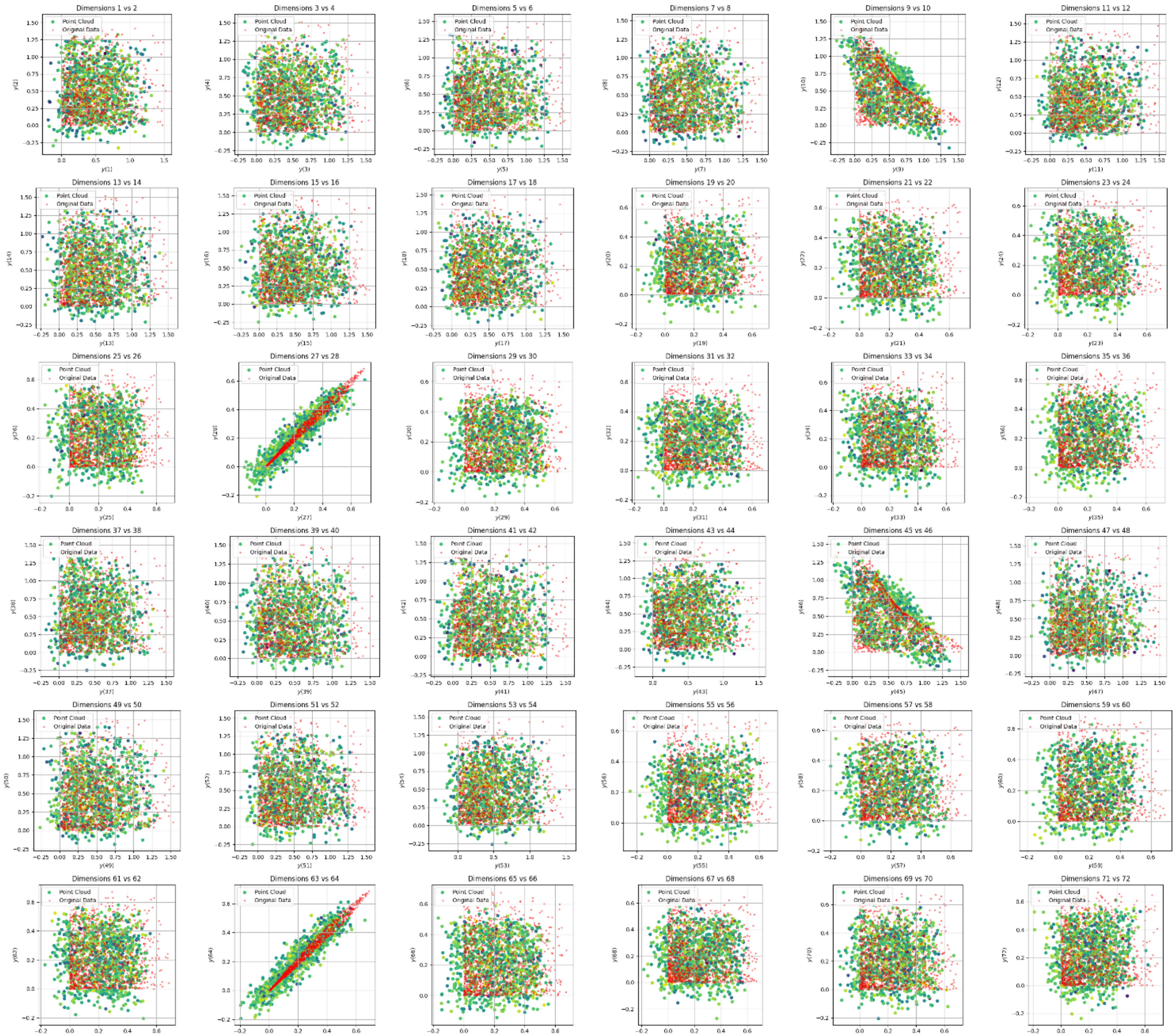
noise = 0.2



MN_1 : E_dim = 72, I_dim = 1

A_init = 0.5

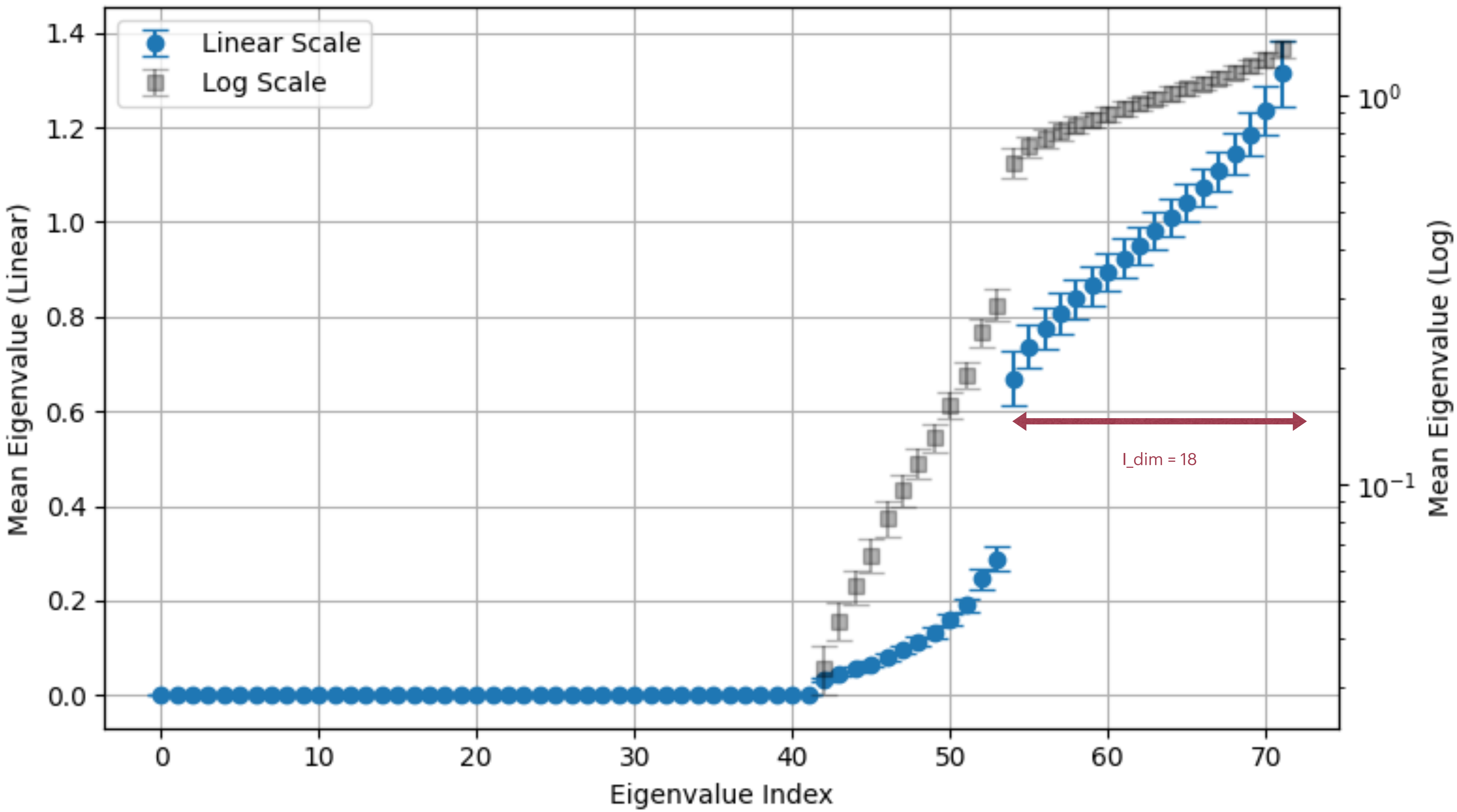
noise = 0.0



MN_1 : E_dim = 72, I_dim = 18

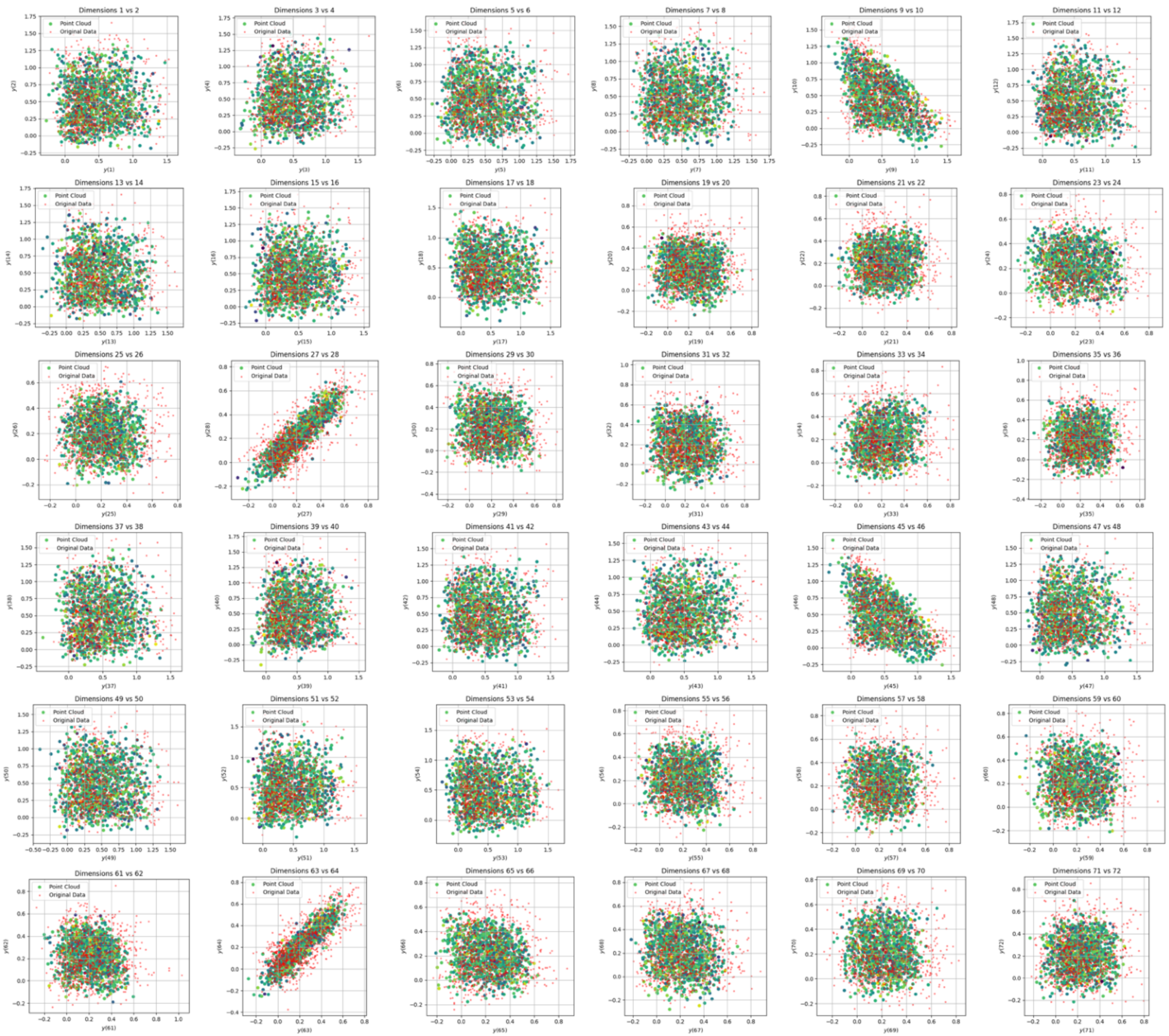
A_init = 0.5

noise = 0.0

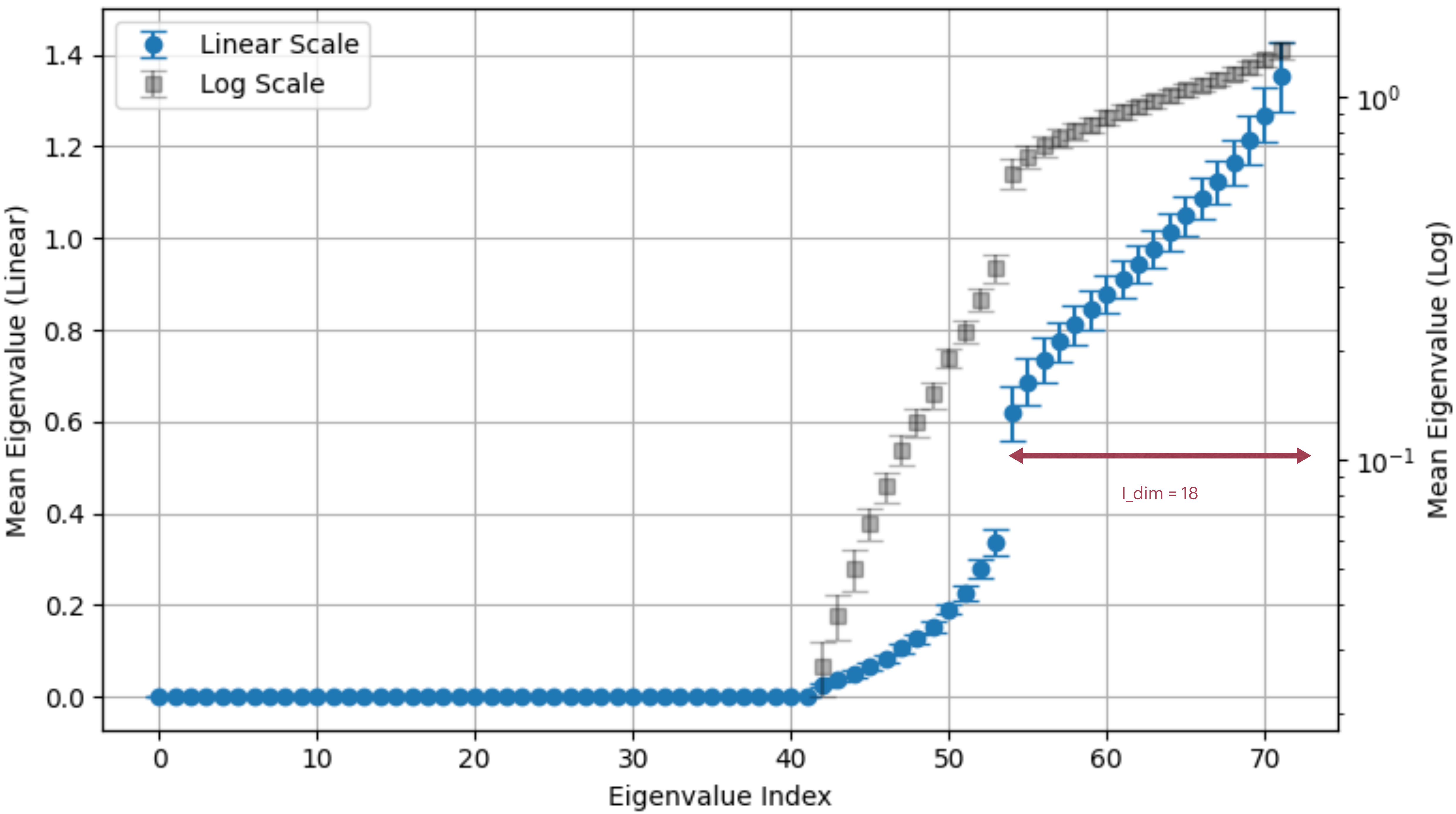


MN_1 : E_dim = 72, I_dim = 1
A_init = 0.5

noise = 0.1

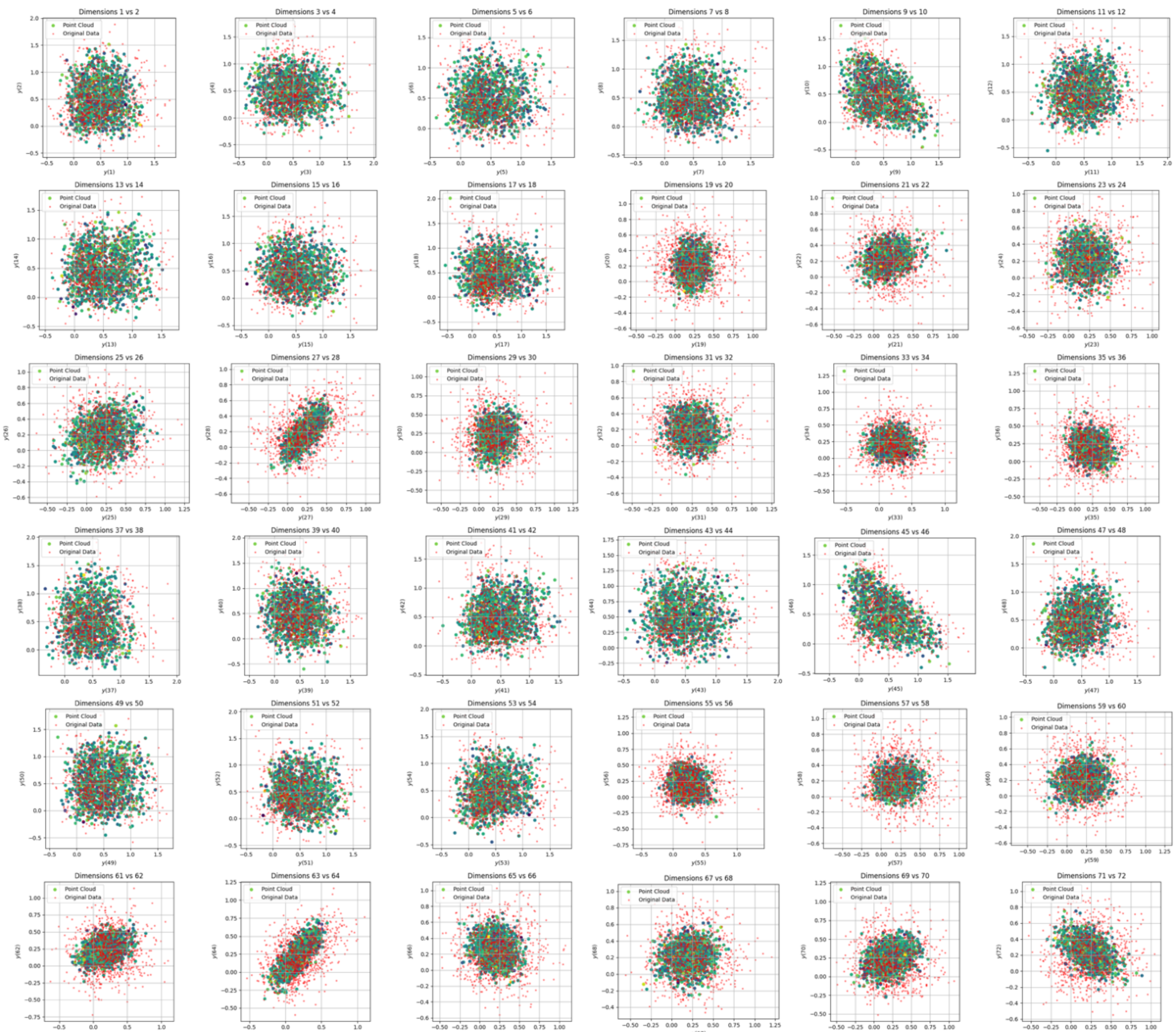


MN_1 : E_dim = 72, I_dim = 1
A_init = 0.5
noise = 0.1



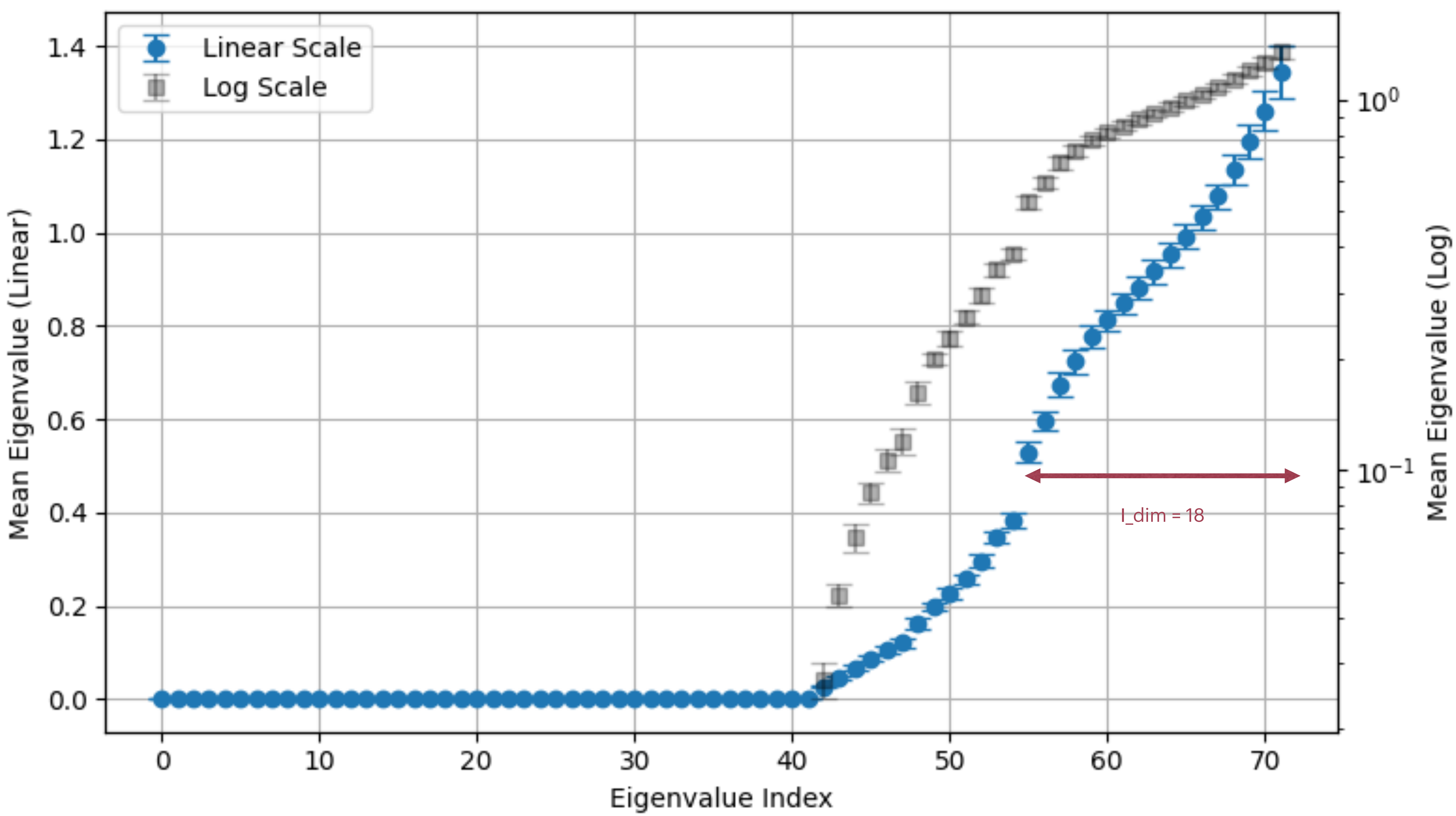
MN_1 : E_dim = 72, I_dim = 1
A_init = 0.5

noise = 0.2



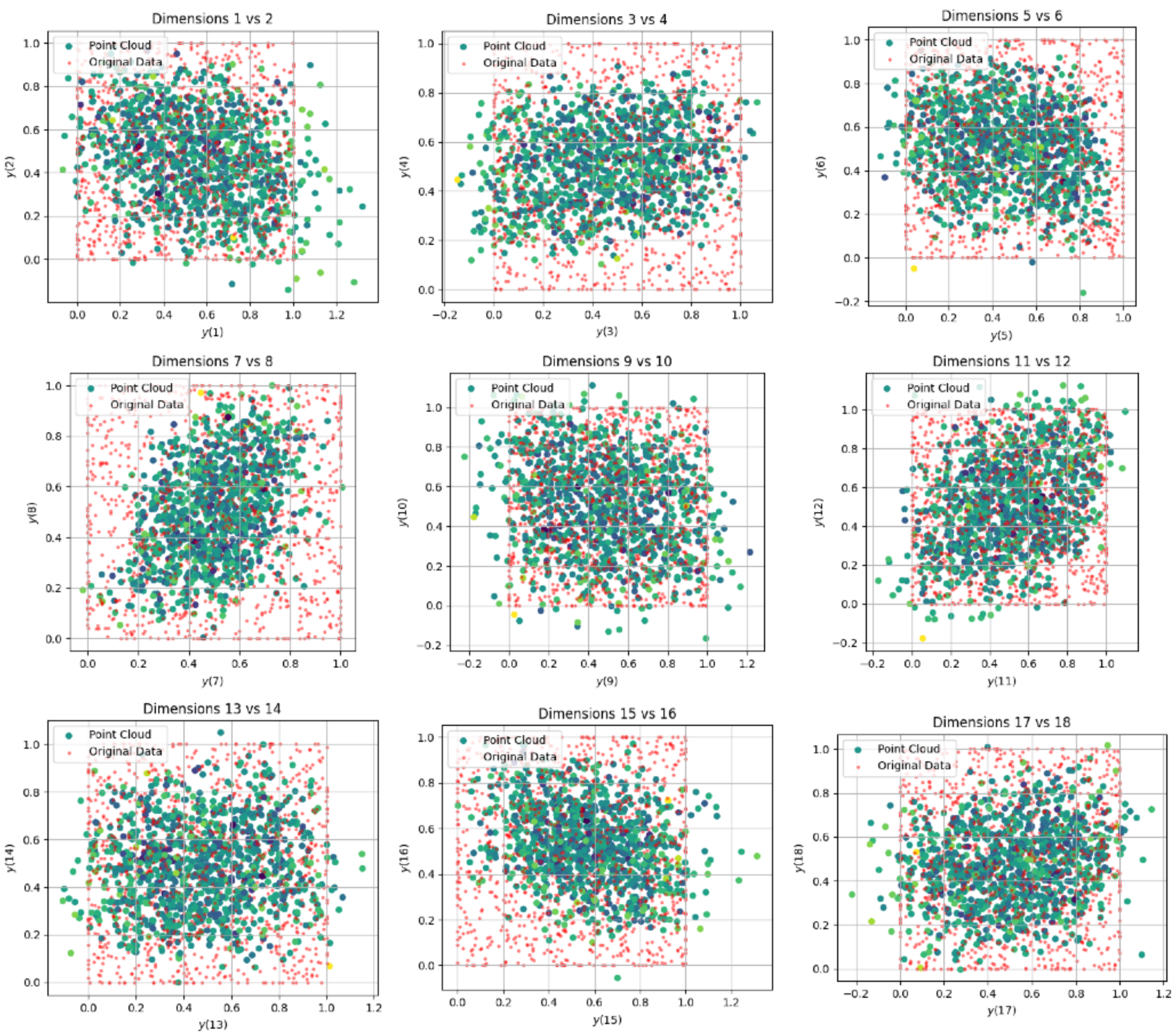
MN_1 : E_dim = 72, I_dim = 1
A_init = 0.5

noise = 0.2

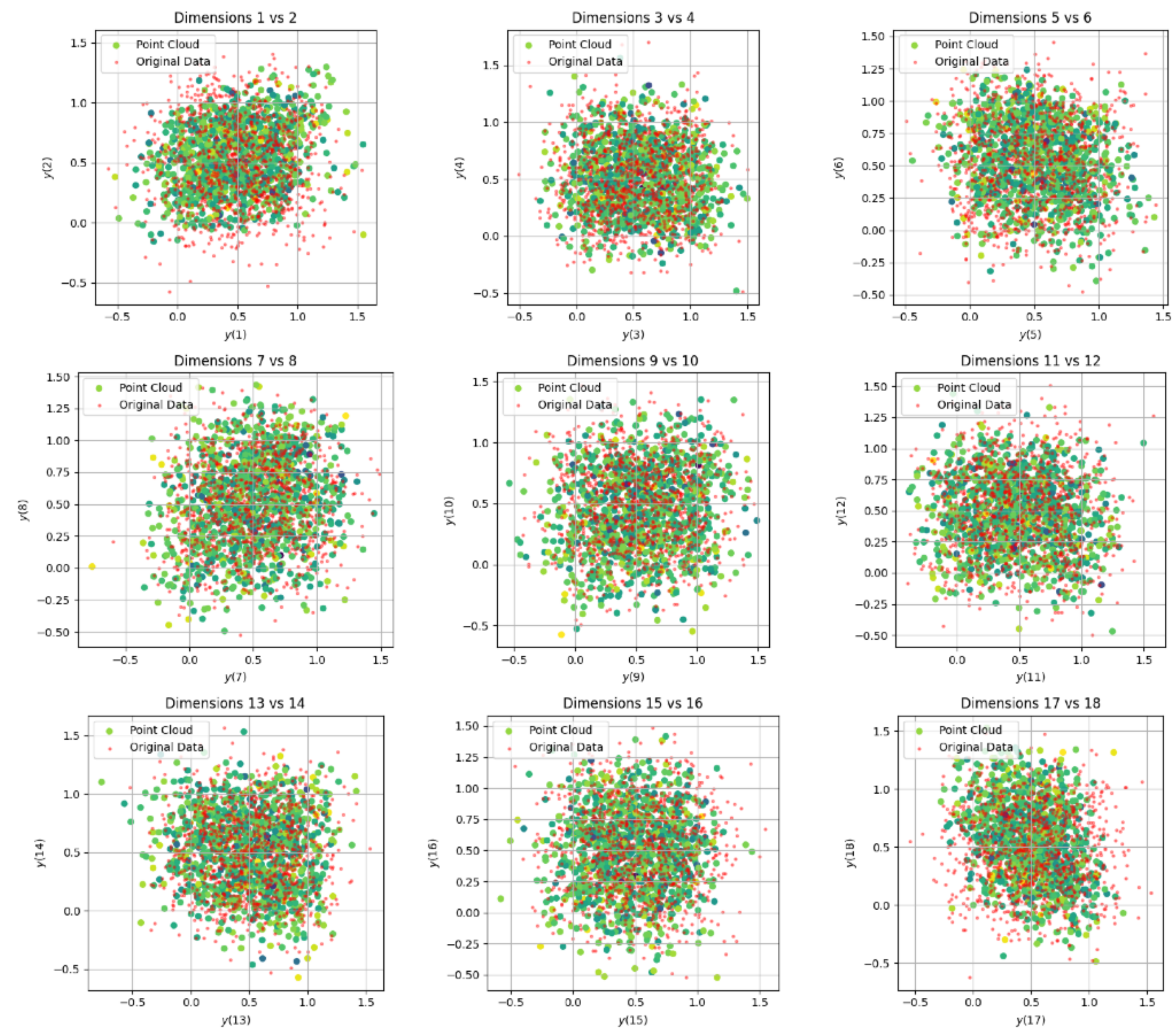


M10b : E_dim = 18, I_dim = 17
A_init = 2

noise = 0.0

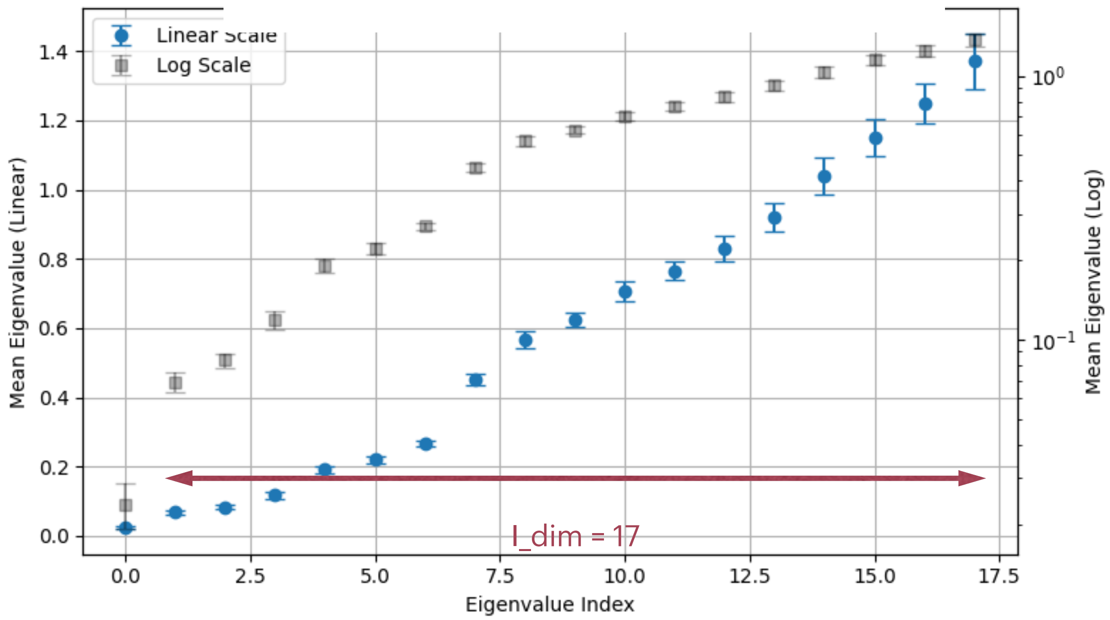


noise = 0.2

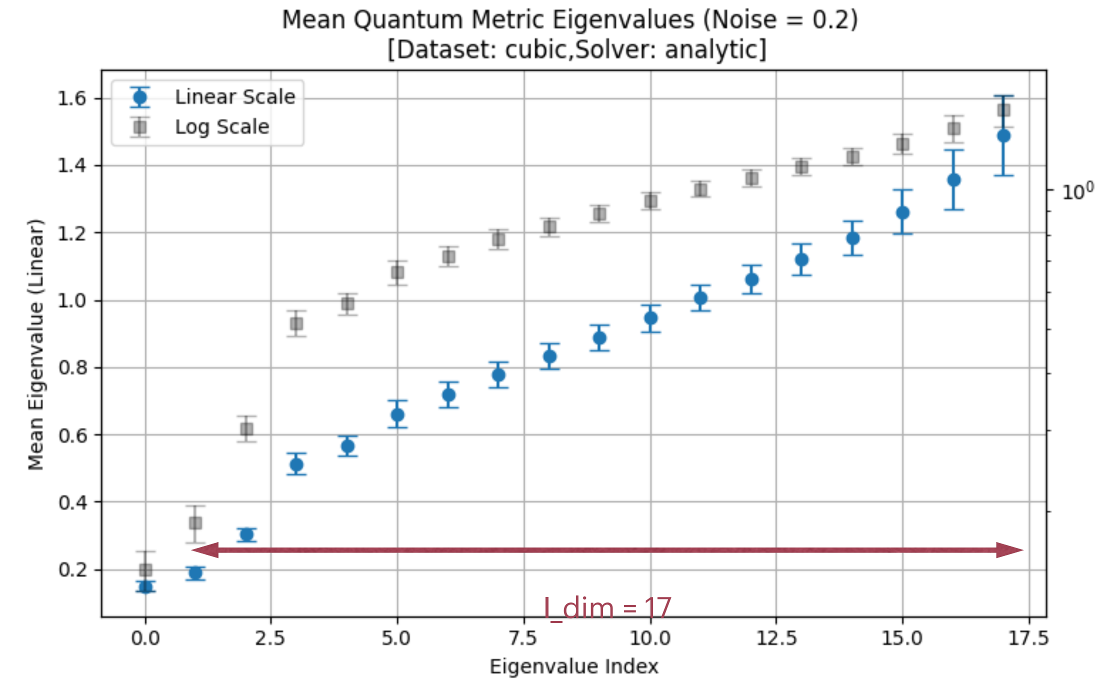


M10b : E_dim = 18, I_dim = 17
A_init = 2

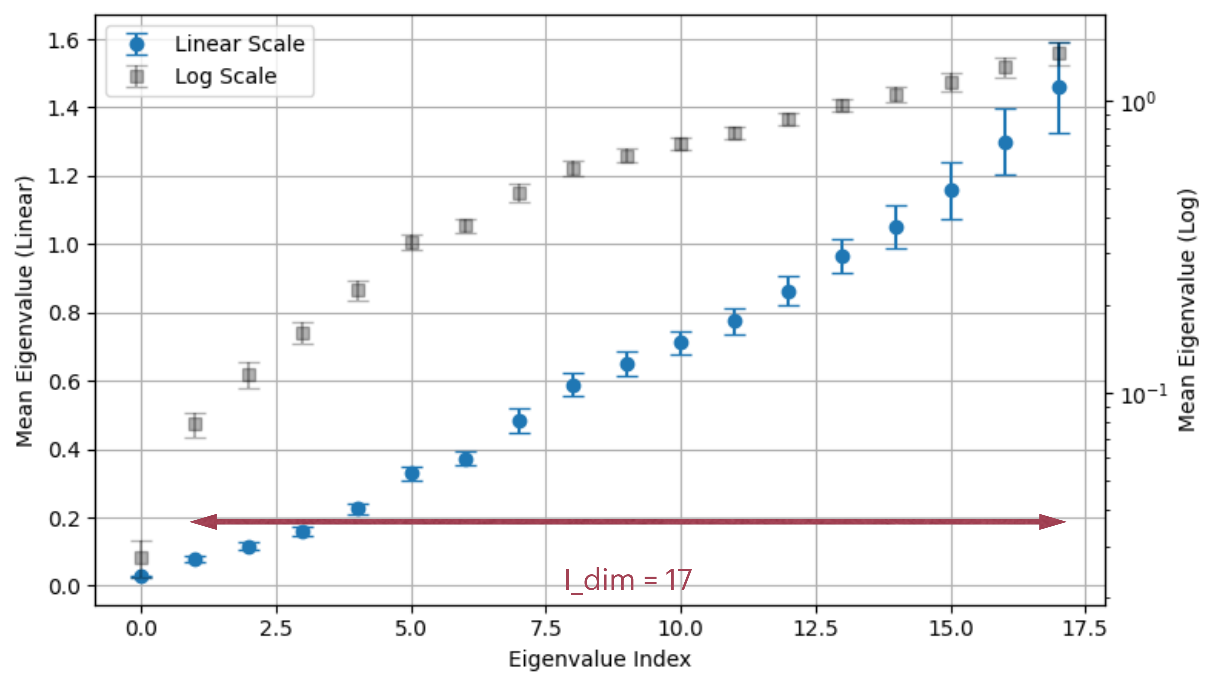
noise = 0.0



noise = 0.2

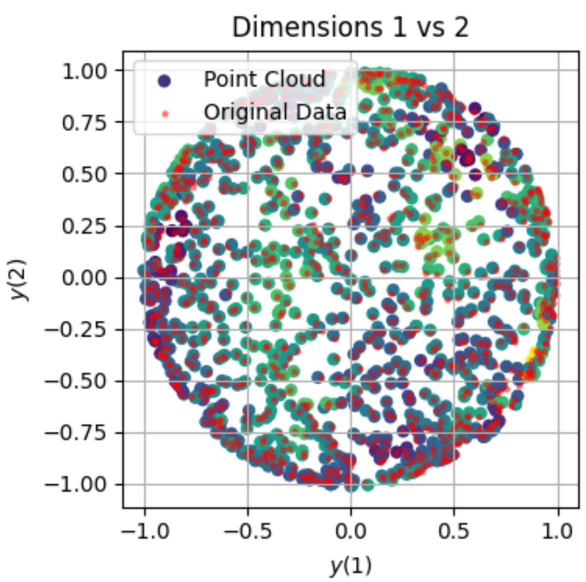


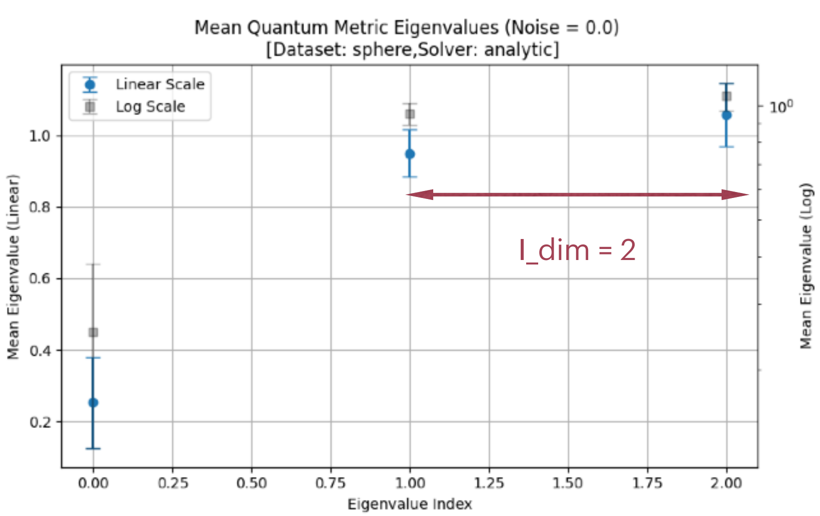
noise = 0.1



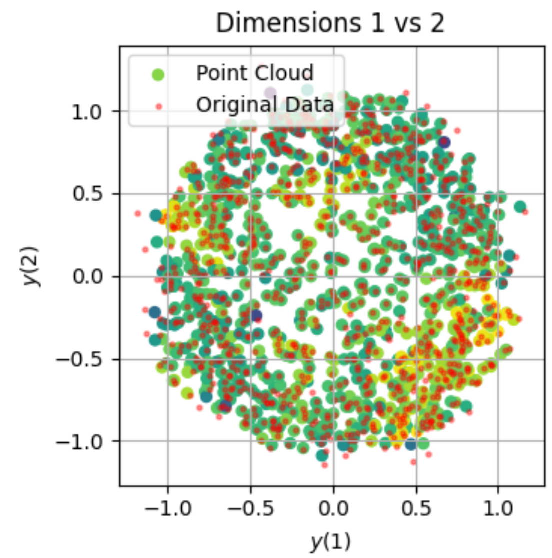
sphere : E_dim = 3, I_dim = 2
A_init = 0.5

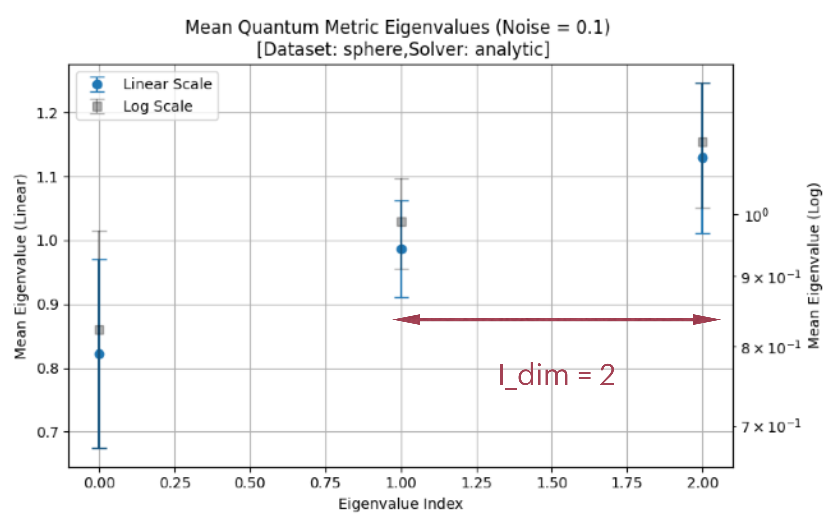
noise = 0.0





noise = 0.1





noise = 0.2