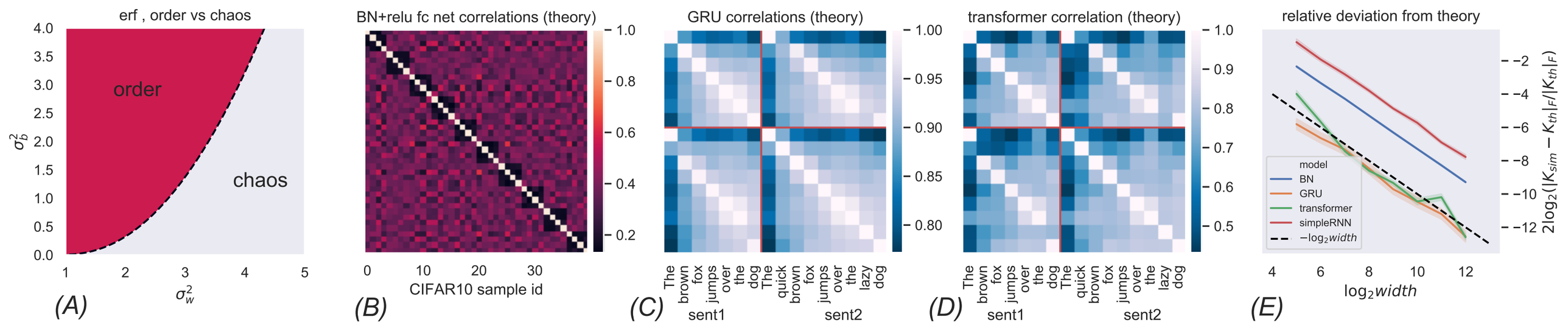
erf , order vs chaos
order
chaos
σ_b^2
σ_w^2
(A)
BN+relu fc net correlations (theory)
CIFAR10 sample id
(B)
GRU correlations (theory)
The brown fox jumps over the dog
The quick brown fox jumps over the lazy dog
sent1
sent2
(C)
transformer correlation (theory)
The brown fox jumps over the dog
The quick brown fox jumps over the lazy dog
sent1
sent2
(D)
relative deviation from theory
model
BN
GRU
transformer
simpleRNN
$-\log_2 width$
$\log_2 width$
$2\log_2(|K_{sim} - K_{th}|_F / |K_{th}|_F)$
(E)