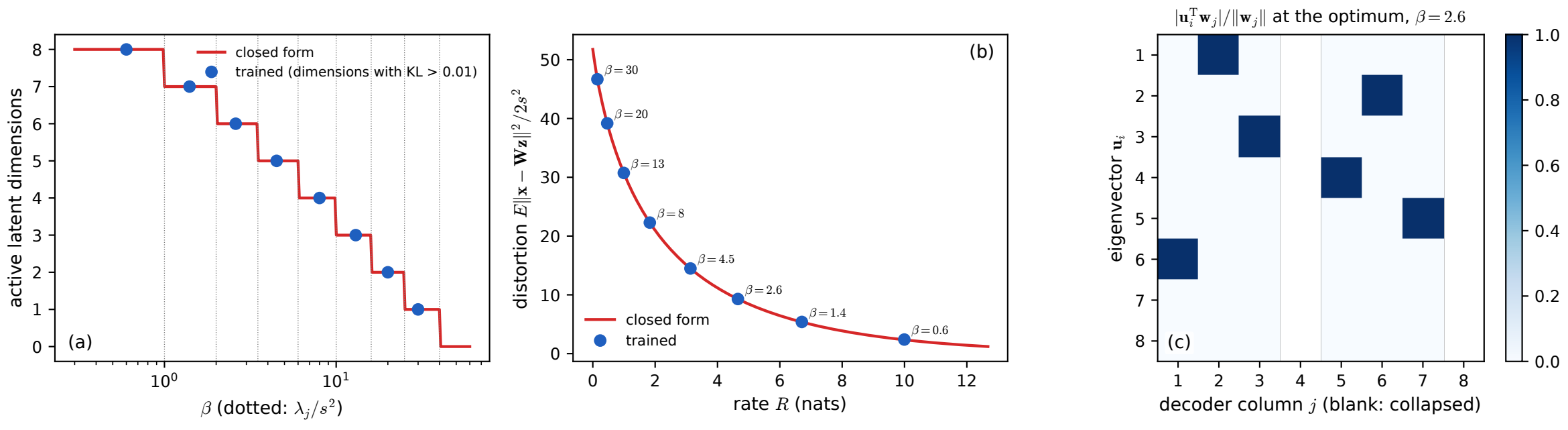

(a)
active latent dimensions
closed form
trained (dimensions with KL > 0.01)
$\beta$ (dotted: $\lambda_j/s^2$)
(b)
distortion $E\|\mathbf{x}-\mathbf{W}\mathbf{z}\|^2/2s^2$
$\beta=30$
$\beta=20$
$\beta=13$
$\beta=8$
$\beta=4.5$
$\beta=2.6$
$\beta=1.4$
$\beta=0.6$
closed form
trained
rate $R$ (nats)
$|\mathbf{u}_i^{\mathrm{T}}\mathbf{w}_j|/\|\mathbf{w}_j\|$ at the optimum, $\beta=2.6$
eigenvector $\mathbf{u}_i$
(c)
decoder column $j$ (blank: collapsed)