Logistic regression with non-convex regularizer; mushrooms EF21; k=2; 1x 10<sup>2</sup> EF21; k=2; 4xEF21; k=2; 8x EF21; k=2; 16x 10<sup>1</sup> EF21; k=2; 32x EF21; k=2; 64xEF21; k=2; 128x EF21; k=2; 256x 10<sup>0</sup>  $||\nabla f(x)||^2$ EF21; k=2; 512x EF21; k=2; 1024x $10^{-1}$  $10^{-2}$ 10<sup>-3</sup> 0.5 1.0 2.0 3.0 0.0 1.5 2.5 1e6 bits/n