

Adding small random noise to  $x'$  recovers the correct label prediction

