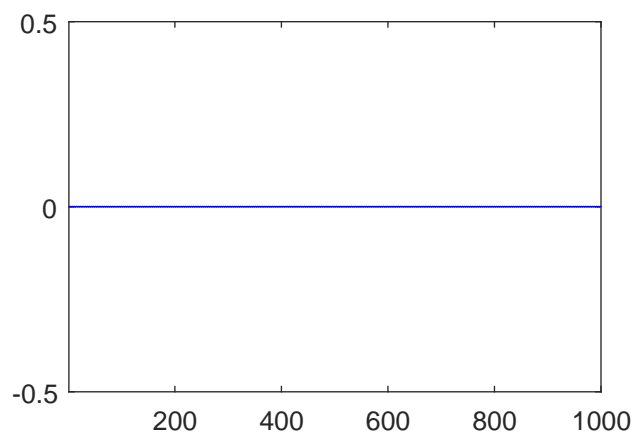
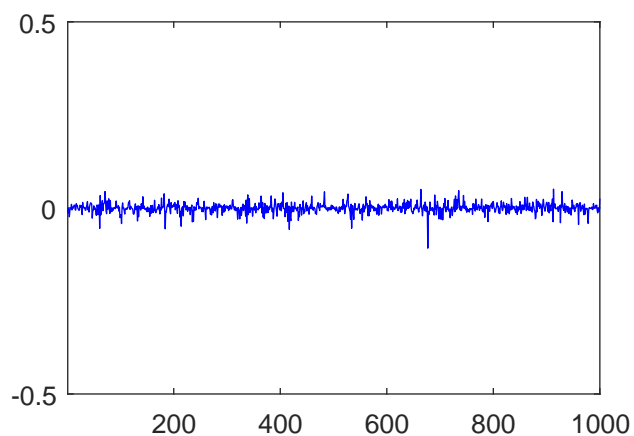


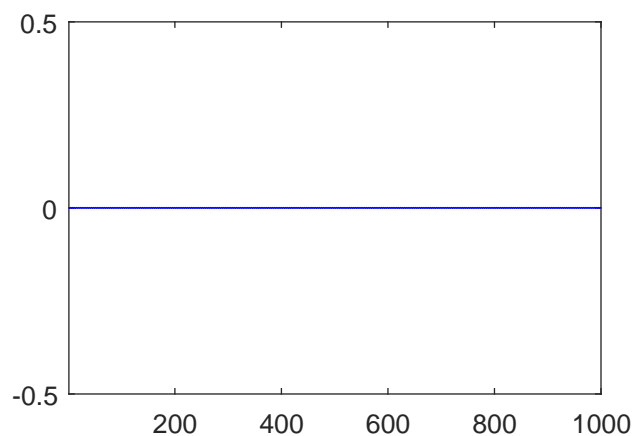
(a)gradient of a non-zero channel
scaling factor at Convergence state of
traditional sparsity training



(b)gradient of a zerochannel
scaling factor at Convergence state of
traditional sparsity training



(c)gradient of a non-zero channel
scaling factor at Convergence state of
mask sparsity training



(d)gradient of a zero channel
scaling factor at Convergence state of
mask sparsity training