Best performances on logistic loss real-sim w8a rcv1 0.7 optimizer optimizer optimizer - SGD SGD - SGD 0.6 0.6 0.6 SARAH SARAH SARAH L-SVRG L-SVRG L-SVRG 0.5 Adam Adam Adam precond $P(w_t)$ precond $P(w_t)$ precond $\frac{N_{t}}{N_{t}}$ 0.4 none none - none -- hutchinson hutchinson – hutchinson 0.3 0.2 0.2 0.2 0.1 0.0 0.0 40 60 80 20 60 80 80 0 20 100 0 40 100 0 20 40 60 100 **Effective Passes Effective Passes Effective Passes** w8a real-sim rcv1 10^{-2} optimizer optimizer optimizer - SGD SGD - SGD 10^{-4} 10^{-4} SARAH SARAH SARAH 10^{-4} L-SVRG L-SVRG L-SVRG Adam Adam $||\nabla P(w_t)||^2$ Adam $||\nabla P(w_t)||^2$ $||\nabla P(w_t)||^2$ precond precond precond 10^{-6} 10^{-6} 10^{-6} none none hutchinson hutchinson hutchinson 10^{-8} 10^{-8} 10^{-8} 10^{-10} 0 20 40 60 80 100 20 40 60 80 100 0 20 40 60 80 100 0 **Effective Passes Effective Passes Effective Passes** w8a real-sim rcv1 optimizer optimizer optimizer - SGD SGD - SGD SARAH SARAH SARAH 10^{-1} L-SVRG L-SVRG L-SVRG 10^{-1} Adam Adam Adam precond precond precond Error Error Error — none none none – hutchinson — hutchinson hutchinson 10^{-2} 10^{-1} 10^{-2} 10^{-3} 60 20 40 80 100 20 40 60 80 100 20 40 60 80 100 0 0 0 **Effective Passes Effective Passes Effective Passes**