Approximate Unlearning

$\mathbb{E}[\nabla^2 f]^{-1}(\mathbb{E}[g])$

$[\nabla^2 f]^{-1}_{T_L}(g_{T_L})$

$[\nabla^2 f]^{-1}_{T_L}(g_{T_L})$

$[\nabla^2 f]^{-1}_{T_L}(g_{T_L})$

$[\nabla^2 f]^{-1}_{T_L}(g_{T_L})$

$\epsilon - \epsilon' > 0$