Proof of Lemma 3.1. (a) $\nabla J(L_A(W,S)) = (ES^TA^{-T}, A^TW^TE) = L_{A^{-T}}\nabla J(W,S)$.
(b)

$$\begin{aligned} DJ(L_A(W,S))[L_A(G,H)] &= \langle \nabla J(L_A(W,S)), L_A(G,H)\rangle = \langle L_{A^{-T}}\nabla_J(W,S), L_A(G,H)\rangle \\ &= \langle(\nabla J(W,S)), (G,H)\rangle = DJ(W,s)[(G,H)]. \end{aligned}$$

(c)

$$\begin{aligned} \nabla^2 J(L_A(W,S))[L_A(G,H)] &= \Big((GA)(A^{-1}S)(A^{-1}S)^T + (WA)(A^{-1}H)(A^{-1}S)^T + E(A^{-1}H)^T, \\ &\qquad (WA)^T(WA)(A^{-1}H) + (WA)^T(GA)(A^{-1}S) + (GA)^TE\Big) \\ &= \Big((GSS^T + WHS^T + EH^T)A^{-T}, (A^T(W^TWH + W^TGS + G^TE)\Big) \\ &= L_{A^{-T}}\left(\nabla^2 J(W,S)[(G,H)]\right). \end{aligned}$$

(d) By definition, $D^2J(L_A(W,S))[L_A(G,H)] = \left\langle \nabla^2(L_A(W,S))[L_A(G,H)], L_A(G,H)\right\rangle$. Using part (c), the above expression is equal to $\left\langle L_{A^{-T}}(\nabla^2 J(W,S)[(G,H)]), L_A(G,H)\right\rangle = \left\langle \nabla^2 J(W,S)[(G,H)], (G,H)\right\rangle = D^2J(W,S)[(G,H)]$. □