Problem 1.

Apply ADMM from page 11.16 with $x_1 = u, x_2 = x, z = w, t = 1$, the algorithm is

$$\begin{aligned}
u^+ &= \underset{v}{\arg\min}\, g(v) + w^T v + \frac{1}{2}\|x - v\|_2^2 \\
x^+ &= \underset{y}{\arg\min}\, f(y) + w^T y + \frac{1}{2}\|y - u^+\|_2^2 \\
w^+ &= w + (x^+ - u^+).
\end{aligned}$$

Complete the squares in line 1, we get

$$\begin{aligned}
\underset{v}{\arg\min}\, g(v) + w^T v + \frac{1}{2}\|x - v\|_2^2 &= \underset{v}{\arg\min}\, g(v) + \frac{1}{2}\|x - v - w\|_2^2 - \|w\|_2^2 + 2w^T x \\
&= \underset{v}{\arg\min}\, g(v) + \frac{1}{2}\|x - v - w\|_2^2 = \mathrm{prox}_g(x + w).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\underset{y}{\arg\min}\, f(y) + w^T y + \frac{1}{2}\|y - u^+\|_2^2 &= \underset{y}{\arg\min}\, f(y) + \frac{1}{2}\|y - u^+ + w\|_2^2 + w^T u^+ - \frac{1}{2}\|w\|_2^2 \\
&= \underset{y}{\arg\min}\, f(y) + \frac{1}{2}\|y - u^+ + w\|_2^2 = \mathrm{prox}_f(u^+ - w).
\end{aligned}$$

Concluding above, ADMM is indeed equivalent to Douglas-Rachford splitting.

Problem 2.

First we establish the useful operators:

$$\begin{aligned}
\mathrm{prox}_{\alpha\|\cdot\|_*}(X) &= \underset{Y}{\arg\min}\, \|Y\|_* + \frac{1}{2\alpha}\|Y - X\|_F^2 \\
&= U_X \left(\underset{\Sigma_Y}{\arg\min}\, \mathrm{tr}(\Sigma_Y) + \frac{1}{2\alpha}\|\Sigma_Y - \Sigma_X\|_F^2 \right) V_X^T \\
&= U_X\, \mathrm{shrink}_\alpha(\Sigma_X) V_X^T, \text{ where } X = U_X \Sigma_X V_X^T.
\end{aligned} \tag{1}$$

(a) Decompose the optimization problem with two variables $x \in \mathbb{R}^n, Y \in \mathbb{R}^{p\times q}$,

$$\begin{array}{ll}
\text{minimize} & \|Y\|_* + \frac{1}{2}\|x - a\|_2^2 \\
\text{subject to} & H(x) - Y = 0
\end{array}$$

Introducing the dual variable $Z \in \mathbb{R}^{p\times q}$, the three steps involved in ADMM are

$$\begin{aligned}
x^+ &= \underset{\tilde{x}}{\arg\min}\, \frac{1}{2}\|\tilde{x} - a\|_2^2 + Z : H(\tilde{x}) + \frac{t}{2}\|H(\tilde{x}) - Y\|_F^2 \\
Y^+ &= \underset{\tilde{Y}}{\arg\min}\, \|\tilde{Y}\|_* + Z : (-\tilde{Y}) + \frac{t}{2}\|H(x^+) - \tilde{Y}\|_F^2 \\
Z^+ &= Z + t(H(x^+) - Y^+).
\end{aligned}$$

Here $Z : X = \mathrm{tr}(Z^T X) = \sum_{i,j} Z_{ij}X_{ij}$ denotes the inner product corresponding to the Frobenius norm. First observe that

$$
\begin{aligned}
Z : H(x) &= Z_{11}x_1 + \left(\sum_{i+j=3} Z_{ij}\right) x_2 + \cdots + \left(\sum_{i+j=n+1} Z_{ij}\right) x_n \\
\|H(x) - Y\|_F^2 &= (x_1 - Y_11)^2 + \sum_{i+j=3}(x_2 - Y_{ij})^2 + \cdots + \sum_{i+j=n+1}(x_n - Y_{ij})^2.
\end{aligned}
$$

Denote the index set $S(k) := \{(i,j) : {}^{i=1,\cdots,p}_{j=1,\cdots,q}, i+j = k+1\}$ to simplify notations. Notice that the minimization involved in the first step is separable for each $\tilde{x}_k$, $(k = 1, \cdots, n)$; for fixed $k = 1, \cdots, n$, the minimizer x_k^+ satisfies

$$
\begin{aligned}
x_k^+ - a_k + \sum_{S(k)} Z_{ij} + t\sum_{S(k)}(x_k^+ - Y_{ij}) &= 0 \\
x_k^+ &= \frac{1}{1 + t\#S(k)}\left(a_k + \sum_{S(k)}(tY_{ij} - Z_{ij})\right).
\end{aligned}
$$

This gives a succinct implementation for the first step. For the second step, the minimization is

$$
\begin{aligned}
Y^+ &= \arg\min_{\tilde{Y}} \|\tilde{Y}\|_* + Z : (-\tilde{Y}) + \frac{t}{2}\|H(x^+) - \tilde{Y}\|_F^2, \\
&= \arg\min_{\tilde{Y}} \|\tilde{Y}\|_* + \frac{t}{2}\left\|\tilde{Y} - H(x^+) - \frac{1}{t}Z\right\|_F^2 + Z : H(x^+) - \frac{1}{2t}\|Z\|_F^2 \\
&= \mathrm{prox}_{\frac{1}{t}\|\cdot\|_*}\left(H(x^+) + \frac{1}{t}Z\right).
\end{aligned}
$$

Use the formula in (1) to evaluate the result. Note that one singular value decomposition will be computed during the evaluation.

(b) Consider the following decomposition:

$$
\begin{aligned}
\text{minimize} \quad & \|Y\|_* + \frac{1}{2}\|x - a\|_2^2 + \delta_{\|u\|_2 \le \gamma}(u) \\
\text{subject to} \quad & H(x) - Y = 0 \\
& Dx - u = 0
\end{aligned}
$$

Introduce dual variable $Z \in \mathbb{R}^{p\times q}, z \in \mathbb{R}^{n-1}$; the three steps involved in ADMM can be written as the following five steps (thanks to the separability of the minimization):

$$\begin{aligned}
x^+ &= \arg\min_{\tilde{x}} \frac{1}{2}\|\tilde{x}-a\|_2^2 + Z : H(\tilde{x}) + z^T D\tilde{x} + \frac{t}{2}\left(\|H(\tilde{x}) - Y\|_F^2 + \|D\tilde{x} - u\|_2^2\right) \\
Y^+ &= \arg\min_{\tilde{Y}} \|\tilde{Y}\|_* + Z : (-\tilde{Y}) + \frac{t}{2}\left(\|H(x^+) - \tilde{Y}\|_F^2 + \|Dx^+ - u\|_2^2\right) \\
u^+ &= \arg\min_{\tilde{u}} \delta_{\|u\|_2\le\gamma}(\tilde{u}) + z^T(-\tilde{u}) + \frac{t}{2}\left(\|H(x^+) - Y^+\|_F^2 + \|Dx^+ - \tilde{u}\|_2^2\right) \\
Z^+ &= Z + t(H(x^+) - Y^+) \\
z^+ &= z + t(Dx^+ - u^+).
\end{aligned}$$

The first step follows a similar derivation as in part (a), for each index $k = 1, \cdots, n$, the minimizer x_k^+ satisfies

$$\begin{aligned}
x_k^+ - a_k + \sum_{S(k)} Z_{ij} + (D^T z)_k + t\sum_{S(k)}(x_k^+ - Y_{ij}) + t(D^T(Dx-u))_k &= 0 \\
x_k^+ &= \frac{1}{1+t\#S(k)}\left(a_k + \sum_{S(k)}(tY_{ij} - Z_{ij}) - (D^T z)_k - t(D^T(Dx-u))_k\right)
\end{aligned}$$

One not completely unrelated observation is that the matrix $D^T D$ is the discrete Laplacian,

$$(D^T Dx)_k = \begin{cases} x_1 - x_2, & k = 1 \\ x_n - x_{n-1}, & k = n \\ 2x_k - x_{k-1} - x_{k+1}, & \text{else} \end{cases}.$$

The second step is exactly identical to that in part (a) and will cost one singular value decomposition of a $p \times q$ matrix. The third step can be computed by a simple projection,

$$\begin{aligned}
u^+ &= \arg\min_{\tilde{u}} \delta_{\|u\|_2\le\gamma}(\tilde{u}) + z^T(-\tilde{u}) + \frac{t}{2}\|Dx^+ - \tilde{u}\|_2^2 \\
&= \arg\min_{\tilde{u}} \delta_{\|u\|_2\le\gamma}(\tilde{u}) + \frac{t}{2}\left\|Dx^+ - \tilde{u} + \frac{1}{t}z\right\|_2^2 - z^T Dx^+ - \frac{1}{2t}\|z\|_2^2 \\
&= \operatorname{proj}_{B_\gamma(0)}\left(Dx^+ + \frac{1}{t}z\right).
\end{aligned}$$

Here $\operatorname{proj}_{B_\gamma(0)}$ denotes the Euclidean projection on the ball (w.r.t. Euclidean distance) of radius γ.

(c) Replace the derivation in part (a) with 1-norm in the first step.

$$sgn(x_k - a_k) + t\#S(k)x_k \ni \sum_{S(k)}(tY_{ij} - Z_{ij}).$$

There are three cases,

$$x_k - a_k > 0 \Leftrightarrow t\#S(k)x_k = \sum_{S(k)}(tY_{ij} - Z_{ij}) - 1$$

$$x_k - a_k < 0 \Leftrightarrow t\#S(k)x_k = \sum_{S(k)}(tY_{ij} - Z_{ij}) + 1$$

$$x_k = a_k \Leftrightarrow \sum_{S(k)}(tY_{ij} - Z_{ij}) - t\#S(k)x_k \in [-1, 1]$$

With further analysis, we have

$$\sum_{S(k)}(tY_{ij} - Z_{ij}) - t\#S(k)a_k \in [-1, 1] \Rightarrow x_k = a_k$$

$$\sum_{S(k)}(tY_{ij} - Z_{ij}) - t\#S(k)a_k > 1 \Rightarrow x_k = \frac{1}{t\#S(k)}\left(\sum_{S(k)}(tY_{ij} - Z_{ij}) - 1\right)$$

$$\sum_{S(k)}(tY_{ij} - Z_{ij}) - t\#S(k)a_k < 1 \Rightarrow x_k = \frac{1}{t\#S(k)}\left(\sum_{S(k)}(tY_{ij} - Z_{ij}) + 1\right)$$

(d) Replace the derivation in part (b) with 1-norm in the first step.

$$sgn(x_k - a_k) + t\#S(k)x_k + t(D^T Dx)_k \ni \sum_{S(k)}(tY_{ij} - Z_{ij}) - (D^T z)_k + t(D^T u)_k.$$

I spent quite a while pondering for a closed form for this step, but haven't have much luck.