Supplementary Material to
Dropout as a Low-Rank Regularizer for Matrix Factorization

Jacopo Cavazza[1], Pietro Morerio[1], Benjamin D. Haeffele[2], Connor Lane[2], Vittorio Murino[1], René Vidal[2]

[1] Pattern Analysis and Computer Vision, Istituto Italiano di Tecnologia, Genova, 16163, Italy
[2] Johns Hopkins University, Center for Imaging Science, Baltimore, MD 21218, USA

jacopo.cavazza@iit.it, pietro.morerio@iit.it, bhaeffele@jhu.edu,
connor.lane@jhu.edu, vittorio.murino@iit.it, rvidal@cis.jhu.edu

Contents

1 Proof of the equality between the dropout and its deterministic reformulation

Proposition 1. *For a fixed $m \times n$ matrix $\mathbf{X}$, consider the problem of factorizing $\mathbf{X}$ into the product $\mathbf{U}\mathbf{V}^\top$ where $\mathbf{U}$ is $m \times d$ and $\mathbf{V}$ is $n \times d$, for some $d \geq \rho(\mathbf{X}) := \operatorname{rank}(\mathbf{X})$. Define $\mathbf{r} = [r_1, \ldots, r_d]$, whose elements are* Bernoulli(θ) *i.i.d. where $0 < \theta < 1$. Furthermore, denote $\mathbf{u}_k \in \mathbb{R}^m$ and $\mathbf{v}_k \in \mathbb{R}^n$ the k-th column in $\mathbf{U}$ and $\mathbf{V}$, respectively, $k = 1, \ldots, d$. Then,*

$$\mathbb{E}_{\mathbf{r}} \left\| \mathbf{X} - \frac{1}{\theta}\mathbf{U}\operatorname{diag}(\mathbf{r})\mathbf{V}^\top \right\|_F^2 = \|\mathbf{X} - \mathbf{U}\mathbf{V}^\top\|_F^2 + \frac{1-\theta}{\theta}\sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2. \tag{1}$$

Proof. Equivalently, we will demonstrate that

$$\mathbb{E}_{\mathbf{r}} \|\theta\mathbf{X} - \mathbf{U}\operatorname{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = \theta^2 \|\mathbf{X} - \mathbf{U}\mathbf{V}^\top\|_F^2 + \theta(1-\theta)\sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2.$$

Since

$$\begin{aligned}
&\mathbb{E}_{\mathbf{r}} \|\theta\mathbf{X} - \mathbf{U}\operatorname{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = \\
&= \mathbb{E}_{\mathbf{r}} \left\| \begin{bmatrix} \theta X_{11} - \sum_{k=1}^d U_{1k} r_k V_{1k}, & \ldots, & \theta X_{1n} - \sum_{k=1}^d U_{1k} r_k V_{nk} \\ \vdots & \ddots & \vdots \\ \theta X_{m1} - \sum_{k=1}^d U_{mk} r_k V_{1k}, & \ldots, & \theta X_{mn} - \sum_{k=1}^d U_{mk} r_k V_{nk} \end{bmatrix} \right\|_F^2,
\end{aligned} \tag{2}$$

by definition of Frobenius norm and linearity of $\mathbb{E}_{\mathbf{r}}$, we elicit

$$\mathbb{E}_{\mathbf{r}} \|\theta\mathbf{X} - \mathbf{U}\operatorname{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} \mathbb{E}_{\mathbf{r}} \left[\left(\theta X_{ij} - \sum_{k=1}^{d} U_{ik} r_k V_{jk} \right)^2 \right]. \tag{3}$$

Use the bias-variance decomposition $\mathbb{E}[r^2] = \mathbb{V}[r] + \mathbb{E}[r]^2$, holding for a scalar random variable r.

$$\begin{aligned}
\mathbb{E}_{\mathbf{r}} \|\theta\mathbf{X} - \mathbf{U}\operatorname{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = &\sum_{i=1}^{m}\sum_{j=1}^{n} \mathbb{V}_{\mathbf{r}} \left[\theta X_{ij} - \sum_{k=1}^{d} U_{ik} r_k V_{jk} \right] + \\
&+ \sum_{i=1}^{m}\sum_{j=1}^{n} \left(\mathbb{E}_{\mathbf{r}} \left[\theta X_{ij} - \sum_{k=1}^{d} U_{ik} r_k V_{jk} \right] \right)^2.
\end{aligned} \tag{4}$$

Since $r_1, \ldots, r_d$ are i.i.d., use the properties of expectation $\mathbb{E}_{\mathbf{r}}$ and variance $\mathbb{V}_{\mathbf{r}}$ with respect to linear combinations of independent random variables.

$$\begin{aligned}
\mathbb{E}_{\mathbf{r}} \|\theta\mathbf{X} - \mathbf{U}\operatorname{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = &\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{d} U_{ik}^2 V_{jk}^2 \mathbb{V}_{\mathbf{r}}\left[r_k\right] + \\
&+ \sum_{i=1}^{m}\sum_{j=1}^{n} \left(\theta X_{ij} - \sum_{k=1}^{d} U_{ik} \mathbb{E}_{\mathbf{r}}\left[r_k\right] V_{jk} \right)^2.
\end{aligned} \tag{5}$$

Exploit the analytical formulas for expected value and variance of a Bernoulli(θ) distribution.

$$\begin{aligned}
\mathbb{E}_{\mathbf{r}} \|\theta\mathbf{X} - \mathbf{U}\operatorname{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = &\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{d} U_{ik}^2 V_{jk}^2 \cdot \theta(1-\theta) + \\
&+ \sum_{i=1}^{m}\sum_{j=1}^{n} \left(\theta X_{ij} - \sum_{k=1}^{d} U_{ik} \cdot \theta \cdot V_{jk} \right)^2.
\end{aligned} \tag{6}$$

Rearrange the terms.

$$\mathbb{E}_{\mathbf{r}}\|\theta\mathbf{X}-\mathbf{U}\mathrm{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = \theta(1-\theta)\sum_{k=1}^{d}\left(\sum_{i=1}^{m}U_{ik}^2\right)\left(\sum_{j=1}^{n}V_{jk}^2\right)+ \\ +\theta^2\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij}-\sum_{k=1}^{d}U_{ik}V_{jk}\right)^2. \tag{7}$$

Use the definition of row-by-column product of matrices

$$\mathbb{E}_{\mathbf{r}}\|\theta\mathbf{X}-\mathbf{U}\mathrm{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = \theta(1-\theta)\sum_{k=1}^{d}\left(\sum_{i=1}^{m}U_{ik}^2\right)\left(\sum_{j=1}^{n}V_{jk}^2\right)+ \\ +\theta^2\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij}-\left[\mathbf{U}\mathbf{V}^\top\right]_{ij}\right)^2. \tag{8}$$

Apply the definitions of squared Euclidean norm $\|\cdot\|_2^2$ and Frobenius norm $\|\cdot\|_F$

$$\mathbb{E}_{\mathbf{r}}\|\theta\mathbf{X}-\mathbf{U}\mathrm{diag}(\mathbf{r})\mathbf{V}^\top\|_F^2 = \theta(1-\theta)\sum_{k=1}^{d}\|\mathbf{u}_k\|_2^2\|\mathbf{v}_k\|_2^2+\theta^2\|\mathbf{X}-\mathbf{U}\mathbf{V}^\top\|_F.$$

This concludes the proof. □

Proposition 2.

$$0 = \inf_{d,\mathbf{U},\mathbf{V}} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2 \quad s.t. \quad d \geq \rho(\mathbf{X}), \mathbf{U} \in \mathbb{R}^{m \times d}, \mathbf{V} \in \mathbb{R}^{n \times d} \text{ and } \mathbf{U}\mathbf{V}^\top = \mathbf{X}. \tag{9}$$

Proof. Let $\mathbf{U}$ and $\mathbf{V}$ such that $\mathbf{U}\mathbf{V}^\top = \mathbf{X}$ for a particular choice of d. Denote

$$\Omega(\mathbf{U}, \mathbf{V}) = \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2 \tag{10}$$

and define

$$\mathbf{A} = \frac{\sqrt{2}}{2}[\mathbf{U}, \mathbf{U}] \in \mathbb{R}^{m \times 2d} \tag{11}$$

$$\mathbf{B} = \frac{\sqrt{2}}{2}[\mathbf{V}, \mathbf{V}] \in \mathbb{R}^{n \times 2d}. \tag{12}$$

Then

$$\mathbf{A}\mathbf{B}^\top = \left(\frac{\sqrt{2}}{2}\right)^2 \mathbf{U}\mathbf{V}^\top + \left(\frac{\sqrt{2}}{2}\right)^2 \mathbf{U}\mathbf{V}^\top = \frac{1}{2}\mathbf{X} + \frac{1}{2}\mathbf{X} = \mathbf{X} \tag{13}$$

and

$$\Omega(\mathbf{A}, \mathbf{B}) = \sum_{k=1}^{2d} \|\mathbf{a}_k\|_2^2 \|\mathbf{b}_k\|_2^2 \tag{14}$$

$$= \frac{1}{4} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2 + \frac{1}{4} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2 = \frac{1}{2}\Omega(\mathbf{U}, \mathbf{V}). \tag{15}$$

In light of this observation, suppose, by absurd that $\varepsilon > 0$ is the minimum of (9), being such value realizes for some matrix $\mathbf{U}$ and $\mathbf{V}$. Then, we can repeat the same construction and produce a pairs of matrix $\mathbf{A}$ and $\mathbf{B}$ such that $\Omega(\mathbf{A}, \mathbf{B}) = \frac{\varepsilon}{2}$. Thus, necessarily, (9) holds being the objective non-negative. □

2 Matrix factorization with variable size

Proposition 3. *For every* $0 < p < 1$*, define*

$$\theta(d) = \frac{p}{d-(d-1)p}. \tag{16}$$

Then, the following properties hold.

1. $0 < \theta(d) < 1$ *for all* $d \in \mathbb{N} \setminus \{0\}$.
2. $\dfrac{1-\theta(kd)}{\theta(kd)} = k\dfrac{1-\theta(d)}{\theta(d)}$ *for all* $k \in \mathbb{N} \setminus \{0\}$.

Proof. 1. We will prove $\theta(d) > 0$ and $\theta(d) < 1$ separately. Since $p > 0$, then $\theta(d) > 0$ if and only if $m-(m-1)p > 0$. But this is true since

$$m-(m-1)p = m - mp + p \geq m(1-p) > 0. \tag{17}$$

On the other hand, since the fraction $\theta(d)$ is positive, $\theta(d) < 1$ is verified if and only if

$$p < m-(m-1)p \tag{18}$$

if and only if

$$0 < m - mp \tag{19}$$

if and only if

$$p < 1 \tag{20}$$

which is actually true by assumption.

2. The property can also be verified analytically by noticing that

$$\frac{1-\theta(d)}{\theta(d)} = \frac{1-\dfrac{p}{d-(d-1)p}}{\dfrac{p}{d-(d-1)p}} = \frac{d-(d-1)p-p}{p} = d\frac{1-p}{p}. \tag{21}$$

This concludes the proof ∎

Proposition 4. *For any $m \times n$ matrix $\mathbf{X}$, consider the expression*

$$\|\mathbf{X}\|_{\triangle} = \min_{d,\mathbf{U},\mathbf{V}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2} \quad s.t. \quad d \geq \rho(\mathbf{X}), \mathbf{U} \in \mathbb{R}^{m\times d}, \mathbf{V} \in \mathbb{R}^{n\times d} \text{ and } \mathbf{U}\mathbf{V}^\top = \mathbf{X}. \tag{22}$$

where $\lambda_d = d^{\frac{1-p}{p}}$, for any $0 < p < 1$, $\mathbf{u}_k \in \mathbb{R}^m$ and $\mathbf{v}_k \in \mathbb{R}^n$ define the k-th column in $\mathbf{U}$ and $\mathbf{V}$, respectively, $k = 1, \ldots, d$. Then, equation (22) defines a quasi-norm over $m \times n$ matrices.

Proof. Using the definition of quasi-norm, we have to prove the following

- $\|\mathbf{X}\|_{\triangle} \geq 0$ for every $\mathbf{X} \in \mathbb{R}^{m\times n}$
- $\|\mathbf{X}\|_{\triangle} = 0 \Longleftrightarrow \mathbf{X} = \mathbf{0}$
- $\|\alpha\mathbf{X}\|_{\triangle} = |\alpha| \|\mathbf{X}\|_{\triangle}$ for every $\alpha \in \mathbb{R}$ and $\mathbf{X} \in \mathbb{R}^{m\times n}$
- $\|\cdot\|_{\triangle}$ is unitarily invariant. In formulæ

$$\|\mathbf{P}\mathbf{X}\mathbf{Q}\|_{\triangle} = \|\mathbf{X}\|_{\triangle} \qquad \text{for every } \mathbf{P} \in O_m \text{ and } \mathbf{Q} \in O_n, \tag{23}$$

  where $O_p := \{\mathbf{A} \in \mathbb{R}^{p\times p} \colon \mathbf{A}^\top\mathbf{A} = \mathbf{A}\mathbf{A}^\top = \mathbf{I}\}$ is the set of all orthonormal matrices of size p.

- Fix $\mathbf{X} \in \mathbb{R}^{m\times n}$ and arbitrary choose a pair of matrices $\mathbf{U}$ and $\mathbf{V}$, of suitable dimensions, such that $\mathbf{U}\mathbf{V}^\top = \mathbf{X}$. We get

$$\sqrt{\lambda_d \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2} \geq 0.$$

  Since the very same holds when computing the minimum over $d, \mathbf{U}$ and $\mathbf{V}$, we obtain $\|\mathbf{X}\|_{\triangle} \geq 0$.

- "$\|\mathbf{X}\|_{\triangle} = 0 \Rightarrow \mathbf{X} = \mathbf{0}$" Let $\overline{\mathbf{U}} \in \mathbb{R}^{m\times \overline{d}}$ and $\overline{\mathbf{V}} \in \mathbb{R}^{n\times \overline{d}}$ such that

$$\|\mathbf{X}\|_{\triangle} = \sqrt{\lambda_d \sum_{k=1}^{d} \|\overline{\mathbf{u}}_k\|_2^2 \|\overline{\mathbf{v}}_k\|_2^2}$$

  and assume

$$\sqrt{\lambda_d \sum_{k=1}^{d} \|\overline{\mathbf{u}}_k\|_2^2 \|\overline{\mathbf{v}}_k\|_2^2} = 0.$$

  Then

$$\sum_{k=1}^{d} \|\overline{\mathbf{u}}_k\|_2^2 \|\overline{\mathbf{v}}_k\|_2^2 = 0$$

  since $\lambda_d > 0$ (due to $0 < \theta(d) < 1$) and, also,

$$\|\overline{\mathbf{u}}_k\|_2^2 \|\overline{\mathbf{v}}_k\|_2^2 = 0 \qquad \text{for every } k = 1, \ldots, d, \tag{24}$$

  since the summation is composed by non-negative terms. By using the zero-product property, we elicit

$$\text{for every } k = 1, \ldots, d \qquad \|\overline{\mathbf{u}}_k\|_2^2 = 0 \text{ or } \|\overline{\mathbf{v}}_k\|_2^2 = 0 \tag{25}$$

  and

$$\text{for every } k = 1, \ldots, d \qquad \|\overline{\mathbf{u}}_k\|_2 = 0 \text{ or } \|\overline{\mathbf{v}}_k\|_2 = 0. \tag{26}$$

  This implies that

$$\text{for every } k = 1, \ldots, d \qquad \overline{\mathbf{u}}_k = \mathbf{0} \text{ or } \overline{\mathbf{v}}_k = \mathbf{0} \tag{27}$$

since $\|\cdot\|_2$ is a norm. But then, for any $i = 1, \dots, m$ and $j = 1, \dots, n$, the combination of the relationship

$$X_{ij} = \sum_{k=1}^{d} \overline{U}_{ik} \overline{V}_{jk} \tag{28}$$

combined with (27) gives

$$X_{ij} = 0 \quad \text{for every } i, j \tag{29}$$

which is the first half of the thesis. The second one, "$\|\mathbf{X}\|_\triangle = 0 \Leftarrow \mathbf{X} = \mathbf{0}$", follows from the following arguments. Assume $\mathbf{X} = \mathbf{0}$. Then the optimal decomposition $\mathbf{U}\mathbf{V}^\top = \mathbf{X}$ in the sense of (22) will be $\mathbf{U} = \mathbf{0}$ and $\mathbf{V} = \mathbf{0}$. This implies $\|\mathbf{X}\|_\triangle = 0$.

- (*Absolute homogeneity.*) Since we already proved (141), we can skip the case $\alpha = 0$ because

$$\|0\mathbf{X}\|_\triangle = \|\mathbf{0}\|_\triangle \overset{(141)}{=} 0 = 0 \cdot \|\mathbf{X}\|_\triangle. \tag{30}$$

Hence, let assume $\alpha \neq 0$. In such a case, by definition,

$$\|\alpha\mathbf{X}\|_\triangle = \underset{\substack{d \geq \rho(\alpha\mathbf{X}) \\ \mathbf{U} \in \mathbb{R}^{m \times d} \\ \mathbf{V} \in \mathbb{R}^{n \times d} \\ \text{s.t. } \mathbf{U}\mathbf{V}^\top = \alpha\mathbf{X}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2}. \tag{31}$$

Since $\alpha \neq 0$,

$$\|\alpha\mathbf{X}\|_\triangle = \underset{\substack{d \geq \rho(\mathbf{X}) \\ \mathbf{U} \in \mathbb{R}^{m \times d} \\ \mathbf{V} \in \mathbb{R}^{n \times d} \\ \text{s.t. } \mathbf{U}\mathbf{V}^\top = \alpha\mathbf{X}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2}. \tag{32}$$

Equivalently,

$$\|\alpha\mathbf{X}\|_\triangle = |\alpha| \underset{\substack{d \geq \rho(\mathbf{X}) \\ \mathbf{U} \in \mathbb{R}^{m \times d} \\ \mathbf{V} \in \mathbb{R}^{n \times d} \\ \text{s.t. } (\frac{1}{\alpha}\mathbf{U})\mathbf{V}^\top = \mathbf{X}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \left\| \frac{1}{\alpha}\mathbf{u}_k \right\|_2^2 \|\mathbf{v}_k\|_2^2}. \tag{33}$$

Since the transformation $\mathbf{U} \mapsto \widetilde{\mathbf{U}} := \frac{1}{\alpha}\mathbf{U}$ is invertible, we get

$$\|\alpha\mathbf{X}\|_\triangle = |\alpha| \underset{\substack{d \geq \rho(\mathbf{X}) \\ \widetilde{\mathbf{U}} \in \mathbb{R}^{m \times d} \\ \mathbf{V} \in \mathbb{R}^{n \times d} \\ \text{s.t. } \widetilde{\mathbf{U}}\mathbf{V}^\top = \mathbf{X}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\widetilde{\mathbf{u}}_k\|_2^2 \|\mathbf{v}_k\|_2^2} = |\alpha| \|\mathbf{X}\|_\triangle. \tag{34}$$

- Fix $\mathbf{X} \in \mathbb{R}^{m \times n}, \mathbf{P} \in O_m$ and $\mathbf{Q} \in O_n$ arbitrary chosen. Then, by definition,

$$\|\mathbf{P}\mathbf{X}\mathbf{Q}\|_\triangle = \underset{\substack{d \geq \rho(\mathbf{P}\mathbf{X}\mathbf{Q}) \\ \mathbf{U} \in \mathbb{R}^{m \times d} \\ \mathbf{V} \in \mathbb{R}^{n \times d} \\ \text{s.t. } \mathbf{U}\mathbf{V}^\top = \mathbf{P}\mathbf{X}\mathbf{Q}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2}. \tag{35}$$

Since $\mathbf{P}$ and $\mathbf{Q}$ are invertible, $\rho(\mathbf{PXQ}) = \rho(\mathbf{X})$ and the equation $\mathbf{UV}^\top = \mathbf{PXQ}$ is equivalent to $\mathbf{P}^\top\mathbf{UV}^\top\mathbf{Q}^\top = \mathbf{X}$. Therefore,

$$\|\mathbf{PXQ}\|_\triangle = \underset{\substack{d \geq \rho(\mathbf{X}) \\ \mathbf{U} \in \mathbb{R}^{m\times d} \\ \mathbf{V} \in \mathbb{R}^{n\times d} \\ \text{s.t. } (\mathbf{P}^\top\mathbf{U})(\mathbf{QV})^\top = \mathbf{X}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2}. \tag{36}$$

Define $\widetilde{\mathbf{U}} = \mathbf{P}^\top\mathbf{U}$ and $\widetilde{\mathbf{V}} = \mathbf{QV}$. Since the maps $\mathbf{U} \mapsto \widetilde{\mathbf{U}}$ and $\mathbf{V} \mapsto \widetilde{\mathbf{V}}$, we obtain

$$\|\mathbf{PXQ}\|_\triangle = \underset{\substack{d \geq \rho(\mathbf{X}) \\ \widetilde{\mathbf{U}} \in \mathbb{R}^{m\times d} \\ \widetilde{\mathbf{V}} \in \mathbb{R}^{n\times d} \\ \text{s.t. } \widetilde{\mathbf{U}}\widetilde{\mathbf{V}}^\top = \mathbf{X}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2} \tag{37}$$

and also

$$\|\mathbf{PXQ}\|_\triangle = \underset{\substack{d \geq \rho(\mathbf{X}) \\ \widetilde{\mathbf{U}} \in \mathbb{R}^{m\times d} \\ \widetilde{\mathbf{V}} \in \mathbb{R}^{n\times d} \\ \text{s.t. } \widetilde{\mathbf{U}}\widetilde{\mathbf{V}}^\top = \mathbf{X}}}{\text{minimum}} \sqrt{\lambda_d \sum_{k=1}^{d} \|\widetilde{\mathbf{u}}_k\|_2^2 \|\widetilde{\mathbf{v}}_k\|_2^2} = \|\mathbf{X}\|_\triangle \tag{38}$$

since the Euclidean norm is unitarily invariant and therefore $\|\widetilde{\mathbf{u}}_k\|_2^2 = \|\mathbf{Pu}_k\|_2^2 = \|\mathbf{u}_k\|_2^2$ and $\|\widetilde{\mathbf{v}}_k\|_2^2 = \|\mathbf{Qv}_k\|_2^2 = \|\mathbf{v}_k\|_2^2$. □

3 Variatonal form of the nuclear norm

Let us introduce the following technical result.

Lemma 1 (The Frobenius norm of a matrix is rotational invariant)**.** *For any arbitrary $m \times n$ matrix $\mathbf{Y}$ and for any $\mathbf{R}_1 \in O(m)$ and $\mathbf{R}_2 \in O(n)$, we have*

$$\|\mathbf{Y}\|_F = \|\mathbf{R}_1\mathbf{Y}\mathbf{R}_2\|_F. \tag{39}$$

Proof. Fix $\mathbb{Y} \in \mathbb{R}^{m\times n}$, $\mathbf{R}_1 \in O(m)$ and $\mathbf{R}_2 \in O(n)$ arbitrary chosen. By observing that, for a rotational matrix, inverse and adjoint are equal, we get

$$\begin{aligned}\|\mathbf{Y}\|_F^2 &= \langle \mathbf{Y}, \mathbf{Y}\rangle_F \\ &= \langle \mathbf{R}_1^\top\mathbf{R}_1\mathbf{Y}, \mathbf{Y}\rangle_F \\ &= \langle \mathbf{R}_1^\top\mathbf{R}_1\mathbf{Y}, \mathbf{Y}\mathbf{R}_2^\top\mathbf{R}_2\rangle_F \\ &= \langle \mathbf{R}_1\mathbf{Y}\mathbf{R}_2^\top, \mathbf{R}_1\mathbf{Y}\mathbf{R}_2^\top\rangle_F \\ &= \|\mathbf{R}_1\mathbf{Y}\mathbf{R}_2^\top\|_F^2.\end{aligned}$$

The thesis follows by square-rooting the extremal members of the previous chain of inequalities.

□

In order to define the nuclear norm, let us consider the family of Schatten/von Neumann ν-norms, $\nu \geq 1$, since the nuclear norm can be retrieved as a particular case.

Definition 1 (Schatten/von Neumann ν norms)**.** *Fix $\nu \in [1, +\infty[$ arbitrary chosen*[1]*. For any $m \times n$ matrix $\mathbf{Y}$ we define the Schatten/von Neumann ν-norm*

$$\|\mathbf{Y}\|_{\mathsf{S},\nu} = \left(\sum_{i=1}^{n} [\sigma_i(\mathbf{Y})]^\nu\right)^{1/\nu} \tag{40}$$

where $\sigma_i(\mathbf{Y})$ are the singular values of $\mathbf{Y}$ and we assume $m \geq n$, the latter hypothesis being non restrictive upon matrix inversion.

Lemma 2. *When $\nu = 2$, the Schatten/von Neumann norm $\|\cdot\|_{\mathsf{S},2}$ equals the Frobenius norm $\|\cdot\|_F$.*

Proof. Consider the singular value decomposition

$$\mathbf{Y} = \mathbf{L}\boldsymbol{\Sigma}\mathbf{R}^\top, \tag{41}$$

where Σ stacks $\sigma_1(\mathbf{Y}), \ldots, \sigma_n(\mathbf{Y})$ on the diagonal. By definition

$$\|\mathbf{Y}\|_{\mathsf{S},2}^2 = \sum_{i=1}^{n} [\sigma_i(\mathbf{Y})]^2 \tag{42}$$

and by definition of Frobenius norm

$$\|\mathbf{Y}\|_{\mathsf{S},2}^2 = \|\boldsymbol{\Sigma}\|_F^2. \tag{43}$$

By using (39) and (41),

$$\|\boldsymbol{\Sigma}\|_F^2 = \|\mathbf{L}\boldsymbol{\Sigma}\mathbf{R}^\top\|_F^2 = \|\boldsymbol{Y}\|_F^2, \tag{44}$$

[1]Although the case $nu = +\infty$ is allowed in the literature, we will skip it here for simplicity

which is the thesis.

□

Definition 2. *Nuclear norm For any $m \times n$ matrix $\mathbf{Y}$, let us define the nuclear norm as*

$$\|\mathbf{Y}\|_\star = \sum_{i=1}^{n} \sigma_i(\mathbf{Y}), \tag{45}$$

being $\sigma_i(\mathbf{Y})$ the singular values of $\mathbf{Y}$, $i = 1, \dots, n$.

Proposition 5. *For any $p > 0$,*

$$\|\mathbf{Y}\|_\star = \underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \|\mathbf{U}\|_F \|\mathbf{V}\|_F \tag{46}$$

Proof. We will demonstrate the thesis by proving that

$$\|\mathbf{Y}\|_\star \leq \underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \|\mathbf{U}\|_F \|\mathbf{V}\|_F \tag{47}$$

and

$$\underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \|\mathbf{U}\|_F \|\mathbf{V}\|_F \leq \|\mathbf{Y}\|_\star \tag{48}$$

hold at the same time

- *Proof of* (47). For a generic d, let $\mathbf{U}$ and $\mathbf{V}$ two $m \times d$ and $n \times d$ dimensional matrices, respectively, such that $\mathbf{Y} = \mathbf{U}\mathbf{V}^\top$. Then,

$$\|\mathbf{Y}\|_\star = \|\mathbf{U}\mathbf{V}^\top\|_\star \tag{49}$$

and

$$\|\mathbf{Y}\|_\star = \sum_{i=1}^{n} \sigma_i(\mathbf{U}\mathbf{V}^\top) \tag{50}$$

by definition of nuclear norm (45). Let

$$\rho = \text{rank}(\mathbf{Y}) = \text{rank}(\mathbf{U}\mathbf{V}^\top) \leq \min(\text{rank}(\mathbf{U}), \text{rank}(\mathbf{V})). \tag{51}$$

Then,

$$\|\mathbf{Y}\|_\star = \sum_{i=1}^{\rho} \sigma_i(\mathbf{U}\mathbf{V}^\top), \tag{52}$$

since $\sigma_{\rho+1}(\mathbf{Y}) = \cdots = \sigma_n(\mathbf{Y}) = 0$. Use Von Neumann inequality

$$\|\mathbf{Y}\|_\star \leq \sum_{i=1}^{\rho} \sigma_i(\mathbf{U})\sigma_i(\mathbf{V}) \tag{53}$$

Apply Cauchy-Schwartz inequality.

$$\|\mathbf{Y}\|_\star \leq \sqrt{\sum_{i=1}^{\rho} \sigma_i(\mathbf{U})^2} \sqrt{\sum_{j=1}^{\rho} \sigma_j(\mathbf{V})^2}. \tag{54}$$

The square-rooting is an increasing function. Therefore, through (51),

$$\|\mathbf{Y}\|_\star \leq \sqrt{\sum_{i=1}^{\rho} \sigma_i(\mathbf{U})^2} \sqrt{\sum_{j=1}^{\text{rank}(\mathbf{V})} \sigma_j(\mathbf{V})^2} \leq \sqrt{\sum_{i=1}^{\text{rank}(\mathbf{U})} \sigma_i(\mathbf{U})^2} \sqrt{\sum_{j=1}^{\text{rank}(\mathbf{V})} \sigma_j(\mathbf{V})^2}, \tag{55}$$

since adding non-negative addends to the summations. Hence, using (40) with $\nu = 2$,

$$\|\mathbf{Y}\|_\star \le \|\mathbf{U}\|_{\mathsf{S},2}\|\mathbf{V}\|_{\mathsf{S},2}. \tag{56}$$

Apply Lemma 2.

$$\|\mathbf{Y}\|_\star \le \|\mathbf{U}\|_F\|\mathbf{V}\|_F. \tag{57}$$

Consider an arbitrary $p > 0$, then

$$\|\mathbf{Y}\|_\star \le \|\mathbf{U}\|_F\|\mathbf{V}\|_F. \tag{58}$$

Since d, $\mathbf{U}$ and $\mathbf{V}$ are generic, we can minimize both terms with respect to them. In this case, the inequality is trivially preserved. Then,

$$\|\mathbf{Y}\|_\star \le \underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \|\mathbf{U}\|_F\|\mathbf{V}\|_F, \tag{59}$$

obtaining (47).

- *Proof of* (48). Consider *a* thin singular value decomposition $\mathbf{Y} = \mathbf{L}\boldsymbol{\Sigma}\mathbf{R}^\top$ for $\mathbf{Y}$ and let us choose

$$\overline{\mathbf{U}} = \mathbf{L}\boldsymbol{\Sigma}^{1/2} \quad \text{and} \quad \overline{\mathbf{V}} = \mathbf{R}\boldsymbol{\Sigma}^{1/2}, \tag{60}$$

being $\boldsymbol{\Sigma}^{1/2}$ the diagonal matrix obtaining from $\boldsymbol{\Sigma}$ entrywise square-rooting all its entries. Note that

$$\overline{\mathbf{U}}\,\overline{\mathbf{V}}^\top = \mathbf{L}\boldsymbol{\Sigma}^{1/2}\left(\mathbf{R}\boldsymbol{\Sigma}^{1/2}\right)^\top = \mathbf{L}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}\mathbf{R}^\top = \mathbf{L}\boldsymbol{\Sigma}\mathbf{R}^\top = \mathbf{Y}. \tag{61}$$

Hence,

$$\underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \|\mathbf{U}\|_F\|\mathbf{V}\|_F \le \|\overline{\mathbf{U}}\|_F\|\overline{\mathbf{V}}\|_F \tag{62}$$

and

$$\underset{\mathbf{U},\mathbf{V}}{\text{minimum}} \|\mathbf{U}\|_F\|\mathbf{V}\|_F \le \|\boldsymbol{\Sigma}^{1/2}\|_F\|\boldsymbol{\Sigma}^{1/2}\|_F \tag{63}$$

by using (39) in the right member. Then

$$\underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \|\mathbf{U}\|_F\|\mathbf{V}\|_F \le \left(\|\boldsymbol{\Sigma}^{1/2}\|_F\right)^{2p}. \tag{64}$$

Observe that

$$\left(\|\boldsymbol{\Sigma}^{1/2}\|_F\right)^{2p} = \left(\sum_{i=1}^{\text{rank}(\mathbf{Y})}\left(\sqrt{\sigma_i(\mathbf{Y})}\right)^2\right)^p = \left(\sum_{i=1}^{\text{rank}(\mathbf{Y})}\sigma_i(\mathbf{Y})\right)^p = \text{trace}(\boldsymbol{\Sigma}) \tag{65}$$

Thus, by means of the definition of nuclear norm (45),

$$\begin{aligned}\underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \|\mathbf{U}\|_F\|\mathbf{V}\|_F &\le \text{trace}(\boldsymbol{\Sigma}) \\ &= \|\mathbf{Y}\|_\star \end{aligned} \tag{66}$$

which gives (48).

□

We can achieve one other equivalent expression for the nuclear norm as in Prop. 5. In order to do this, let us consider the following necessary technical result.

Lemma 3. *For any $a, b \in \mathbb{R}$, and $p > 0$*

$$(2ab) = \underset{\eta > 0}{\text{minimum}} \left(\frac{a^2}{\eta} + \eta b^2\right) \tag{67}$$

Proof. Fix arbitrary $a, b \in \mathbb{R}$ and $\eta > 0$. Then

$$0 \leq (a - \eta b)^2,$$

$$0 \leq a^2 - 2\eta ab + \eta^2 b^2,$$

$$2\eta ab \leq a^2 + \eta^2 b^2.$$

Divide each term by $\eta > 0$.

$$2ab \leq \frac{a^2}{\eta} + \eta b^2.$$

The inequality is preserved after p-th order exponentiation, $p > 0$.

$$(2ab) \leq \left(\frac{a^2}{\eta} + \eta b^2\right).$$

Thus, the function $f_{a,b}(\eta) = (\frac{a^2}{\eta} + \eta b^2)$ upper bounds $(2ab)$ for any a, b, η. Then, since

$$f_{a,b}(a/b) = \left(\frac{a^2}{a} b + \frac{a}{b} b^2\right) = (ab + ab) = (2ab), \tag{68}$$

we get the thesis by using the definition of minimum. □

Proposition 6. *For any $p > 0$,*

$$\|\mathbf{Y}\|_\star = \underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \frac{1}{2} \left(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2\right) \tag{69}$$

Proof. It is enough to observe that, for any matrix $\mathbf{U}$ and $\mathbf{V}$ such that $\mathbf{U}\mathbf{V}^\top = \mathbf{Y}$, also $\mathbf{U}/\sqrt{\eta}$ and $\sqrt{\eta}\mathbf{V}$ still satisfy the same property, $\eta > 0$. Then, for any arbitrary $\eta > 0$,

$$\underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \frac{1}{2} \left(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2\right) = \underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \left(\frac{1}{2} \frac{\|\mathbf{U}\|_F^2}{\eta} + \frac{1}{2} \eta \|\mathbf{V}\|_F^2\right) \tag{70}$$

Minimize over η and apply (67).

$$\underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \frac{1}{2} \left(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2\right) = \underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \left(\|\mathbf{U}\|_F \|\mathbf{V}\|_F\right). \tag{71}$$

The thesis is direct corollary of Prop. 5. □

Proposition 7. *For any $p > 0$,*

$$\|\mathbf{Y}\|_\star = \underset{\mathbf{U}, \mathbf{V} \colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \frac{1}{2} \left(\sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 + \sum_{k=1}^{d} \|\mathbf{v}_k\|_2^2\right), \tag{72}$$

where $\mathbf{u}_k \in \mathbb{R}^m$ and $\mathbf{v}_k \in \mathbb{R}^n$ denote the k-th column of $\mathbf{U}$ and $\mathbf{V}$, respectively.

Proof. It follows by applying (69) and observing that

$$\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 = \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 + \sum_{k=1}^{d} \|\mathbf{v}_k\|_2^2 \tag{73}$$

□

Proposition 8.

$$\|\mathbf{Y}\|_\star = \underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2, \tag{74}$$

where $\mathbf{u}_k \in \mathbb{R}^m$ *and* $\mathbf{v}_k \in \mathbb{R}^n$ *denote the* k*-th column of* $\mathbf{U}$ *and* $\mathbf{V}$*, respectively.*

Proof. Apply (67) for any $k = 1, \ldots, d$. Then,

$$\|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2 = \underset{\eta_k > 0}{\text{minimum}} \left(\frac{1}{2} \frac{\|\mathbf{u}_k\|_2^2}{\eta_k} + \eta_k \|\mathbf{v}_k\|_2^2 \right). \tag{75}$$

Since we have a decoupled minimization problem, we can commute summation and minimum, rewriting

$$\sum_{k=1}^{d} \|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2 = \sum_{k=1}^{d} \underset{\eta_k > 0}{\text{minimum}} \left(\frac{1}{2} \frac{\|\mathbf{u}_k\|_2^2}{\eta_k} + \eta_k \|\mathbf{v}_k\|_2^2 \right) \tag{76}$$

into

$$\sum_{k=1}^{d} \|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2 = \underset{\eta_1, \ldots, \eta_d > 0}{\text{minimum}} \left(\frac{1}{2} \sum_{k=1}^{d} \frac{\|\mathbf{u}_k\|_2^2}{\eta_k} + \sum_{k=1}^{d} \eta_k \|\mathbf{v}_k\|_2^2 \right). \tag{77}$$

Minimize both terms with respect to d, $\mathbf{U}$ and $\mathbf{V}$.

$$\begin{aligned} &\underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2 \\ &= \underset{\eta_1, \ldots, \eta_d > 0, d, \mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \left(\frac{1}{2} \sum_{k=1}^{d} \frac{\|\mathbf{u}_k\|_2^2}{\eta} + \eta \|\mathbf{v}_k\|_2^2 \right). \end{aligned} \tag{78}$$

Equivalently,

$$\begin{aligned} &\underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2 \\ &= \underset{\eta_1, \ldots, \eta_d > 0, d, \mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \left(\frac{1}{2} \sum_{k=1}^{d} \|\mathbf{u}_k / \sqrt{\eta_k}\|_2^2 + \|\sqrt{\eta_k} \mathbf{v}_k\|_2^2 \right). \end{aligned} \tag{79}$$

We can discard the factors η_k exploiting the condition $\mathbf{U}\mathbf{V}^\top = \mathbf{Y}$.

$$\underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2 = \underset{d, \mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \left(\frac{1}{2} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 + \|\mathbf{v}_k\|_2^2 \right).$$

We get

$$\underset{\mathbf{U},\mathbf{V}\colon \mathbf{U}\mathbf{V}^\top = \mathbf{Y}}{\text{minimum}} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2 \|\mathbf{v}_k\|_2 = \|\mathbf{Y}\|_\star$$

by means of (72). □

Observation. Due to the fact that, in all previous formulas, d was fixed and generic, the variational forms still apply if we select d to be the arg minimum of the objective fuctions whose minimization over $\mathbf{U}$ and $\mathbf{V}$ gives the variational forms itself. Therefore,

$$\|\mathbf{X}\|_\star = \min_{d\geq\rho(\mathbf{X}),\mathbf{U}\in\mathbb{R}^{m\times d},\mathbf{V}\in\mathbb{R}^{n\times d}:\ \mathbf{U}\mathbf{V}^\top=\mathbf{X}} \frac{1}{2}\left(\|\mathbf{U}\|_F^2+\|\mathbf{V}\|_F^2\right) \tag{80}$$

$$= \min_{d\geq\rho(\mathbf{X}),\mathbf{U}\in\mathbb{R}^{m\times d},\mathbf{V}\in\mathbb{R}^{n\times d}:\ \mathbf{U}\mathbf{V}^\top=\mathbf{X}} \|\mathbf{U}\|_F\|\mathbf{V}\|_F \tag{81}$$

$$= \min_{d\geq\rho(\mathbf{X}),\mathbf{U}\in\mathbb{R}^{m\times d},\mathbf{V}\in\mathbb{R}^{n\times d}:\ \mathbf{U}\mathbf{V}^\top=\mathbf{X}} \frac{1}{2}\left(\sum_{k=1}^d\|\mathbf{u}_k\|_2^2+\sum_{k=1}^d\|\mathbf{v}_k\|_2^2\right) \tag{82}$$

$$= \min_{d\geq\rho(\mathbf{X}),\mathbf{U}\in\mathbb{R}^{m\times d},\mathbf{V}\in\mathbb{R}^{n\times d}:\ \mathbf{U}\mathbf{V}^\top=\mathbf{X}} \sum_{k=1}^d\|\mathbf{u}_k\|_2\|\mathbf{v}_k\|_2. \tag{83}$$

4 (Lower) convex envelope of the regularizer

Proposition 9. *The convex envelope of* $\frac{1}{2}\|\mathbf{X}\|_\triangle^2$ *is* $\frac{1-p}{2p}\|\mathbf{X}\|_\star^2$.

Proof. First, recall that the convex envelope of a function f is the largest closed, convex function g such that $g(x)\leq f(x)$ for all x and is given by $g=(f^*)^*$, where f^* denotes the Fenchel dual of f, defined as $f^*(q)\equiv\sup_x\langle q,x\rangle-f(x)$. Let $\Theta(\mathbf{X})=\frac{1}{2}\|\mathbf{X}\|_\triangle^2$, given by

$$\Theta(\mathbf{X}) = \inf_{\substack{d\geq\rho(\mathbf{X})\\ \mathbf{U}\in\mathbb{R}^{m\times d}\\ \mathbf{V}\in\mathbb{R}^{n\times d}\\ \text{s.t. } \mathbf{U}\mathbf{V}^\top=\mathbf{X}}} \frac{\lambda_d}{2}\sum_{k=1}^d\|\mathbf{u}_k\|_2^2\|\mathbf{v}_k\|_2^2. \tag{84}$$

and note that this can be equivalently written by the equation

$$\Theta(\mathbf{X}) = \inf_{\substack{d\geq\rho(\mathbf{X})\\ \mathbf{U}\in\mathbb{R}^{m\times d}\\ \mathbf{V}\in\mathbb{R}^{n\times d}\\ \Lambda\in\mathbb{R}^d}} \frac{\lambda_d}{2}\|\Lambda\|_2^2 \ \text{ s.t. } \sum_{k=1}^d\Lambda_k\mathbf{u}_k\mathbf{v}_k^T=\mathbf{X} \ \text{ and } \ (\|\mathbf{u}_k\|_2,\|\mathbf{v}_k\|_2)\leq(1,1) \ \ \forall k. \tag{85}$$

This gives the Fenchel dual of Θ as

$$\Theta^*(\mathbf{Q}) = \sup_d \sup_{\substack{\mathbf{U}\in\mathbb{R}^{m\times d}\\ \mathbf{V}\in\mathbb{R}^{n\times d}\\ \Lambda\in\mathbb{R}^d}} \sum_{k=1}^d\Lambda_k\left\langle\mathbf{Q},\mathbf{u}_k\mathbf{v}_k^T\right\rangle-\frac{\lambda_d}{2}\|\Lambda\|_2^2 \ \text{ s.t. } \ (\|\mathbf{u}_k\|_2,\|\mathbf{v}_k\|_2)\leq(1,1) \ \ \forall k. \tag{86}$$

Now, note that if we define the vector $\mathbf{B}_d(\mathbf{U},\mathbf{V})\in\mathbb{R}^d$ as

$$\mathbf{B}_d(\mathbf{U},\mathbf{V}) = \begin{bmatrix}\left\langle\mathbf{Q},\mathbf{u}_1\mathbf{v}_1^T\right\rangle\\ \left\langle\mathbf{Q},\mathbf{u}_2\mathbf{v}_2^T\right\rangle\\ \vdots\\ \left\langle\mathbf{Q},\mathbf{u}_d\mathbf{v}_d^T\right\rangle\end{bmatrix}, \tag{87}$$

then from (86) we have that

$$\Theta^*(\mathbf{Q}) = \sup_d \sup_{\substack{\mathbf{U}\in\mathbb{R}^{m\times d}\\ \mathbf{V}\in\mathbb{R}^{n\times d}}} \sup_{\Lambda\in\mathbb{R}^d} \langle\mathbf{B}_d(\mathbf{U},\mathbf{V}),\Lambda\rangle-\frac{\lambda_d}{2}\|\Lambda\|_2^2 \ \text{ s.t. } \ (\|\mathbf{u}_k\|_2,\|\mathbf{v}_k\|_2)\leq(1,1) \ \ \forall k \tag{88}$$

$$= \sup_d \sup_{\substack{\mathbf{U}\in\mathbb{R}^{m\times d}\\ \mathbf{V}\in\mathbb{R}^{n\times d}}} \frac{1}{2\lambda_d}\|\mathbf{B}_d(\mathbf{U},\mathbf{V})\|_2^2 \ \text{ s.t. } \ (\|\mathbf{u}_k\|_2,\|\mathbf{v}_k\|_2)\leq(1,1) \ \ \forall k. \tag{89}$$

where the final equality comes from noting that the supremum w.r.t. Λ is the definition of the Fenchel dual of the squared ℓ_2 norm evaluated at $\mathbf{B}_d(\mathbf{U}, \mathbf{V})$.

Now, from(89) and the definition of $\mathbf{B}_d(\mathbf{U}, \mathbf{V})$ note that for a fixed value of d, (89) is optimized w.r.t. $(\mathbf{U}, \mathbf{V})$ by choosing all the columns of $(\mathbf{U}, \mathbf{V})$ to be equal to the maximum singular vector pair, given by

$$\sup_{\mathbf{u} \in \mathbb{R}^m, \mathbf{v} \in \mathbb{R}^n} \left\langle \mathbf{Q}, \mathbf{u}\mathbf{v}^T \right\rangle \quad \text{s.t.} \quad (\|\mathbf{u}\|_2, \|\mathbf{v}\|_2) \leq (1, 1). \tag{90}$$

Note also that for this optimal choice of $(\mathbf{U}, \mathbf{V})$ we have that $\mathbf{B}_d(\mathbf{U}, \mathbf{V}) = \sigma(\mathbf{Q})\mathbf{1}_d$ where $\sigma(\mathbf{Q})$ denotes the largest singular value of $\mathbf{Q}$ and $\mathbf{1}_d$ is a vector of all ones of size d. Plugging this in (89) gives

$$\Theta^*(\mathbf{Q}) = \sup_d \frac{1}{2\lambda_d} \|\sigma(\mathbf{Q})\mathbf{1}_d\|_2^2 = \sup_d \frac{\sigma^2(\mathbf{Q})d}{2\lambda_d} = \left(\frac{p}{1-p}\right) \frac{\sigma^2(\mathbf{Q})}{2}, \tag{91}$$

where recall $\lambda_d = d(1-p)/p$. The result then follows by noting the well-known duality between the spectral norm (largest singular value) and the nuclear norm and basic properties of the Fenchel dual. □

5 Optimal dropout for MF satisfies the optimality conditions for the convex lower bound

For a fixed $\mathbf{Y} \in \mathbb{R}^{m\times n}$, find $\mathbf{U} \in \mathbb{R}^{m\times d}$, and $\mathbf{V} \in \mathbb{R}^{n\times d}$, $d \geq \text{rank}(\mathbf{Y})$, consider the dropout problem

$$\min_{\mathbf{U},\mathbf{V},d} \mathbb{E}_{\mathbf{r}} \left\| \mathbf{Y} - \frac{1}{\theta}\mathbf{U}\text{diag}(\mathbf{r})\mathbf{V}^\top \right\|_F^2, \tag{92}$$

where $\mathbf{r}$ is a d-dimensional random vector with i.i.d. Bernoulli(θ) entries, $0 < \theta < 1$. We know that

$$\mathbb{E}_{\mathbf{r}} \left\| \mathbf{Y} - \frac{1}{\theta}\mathbf{U}\text{diag}(\mathbf{r})\mathbf{V}^\top \right\|_F^2 = \|\mathbf{Y} - \mathbf{U}\mathbf{V}^\top\|_F^2 + \frac{1-\theta}{\theta}\sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2\|\mathbf{v}_k\|_2^2 \tag{93}$$

and for a particular choice $\theta = \theta(d)$, we get

$$\mathbb{E}_{\mathbf{r}} \left\| \mathbf{Y} - \frac{1}{\theta(d)}\mathbf{U}\text{diag}(\mathbf{r})\mathbf{V}^\top \right\|_F^2 = \|\mathbf{Y} - \mathbf{U}\mathbf{V}^\top\|_F^2 + d\frac{1-p}{p}\sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2\|\mathbf{v}_k\|_2^2 \tag{94}$$

Definition 3 (Convex envelope)**.** *The convex envelope of a real-valued function f is the largest closed, convex function g such that $g \leq f$.*

Definition 4. *Fix $\mathbf{Y} \in \mathbb{R}^{m\times n}$ arbitrary. Then, for any integer $d \geq \text{rank}(\mathbf{Y})$ and $\mathbf{U} \in \mathbb{R}^{m\times d}$, $\mathbf{V} \in \mathbb{R}^{n\times d}$, define*

$$\Omega(\mathbf{Y}) = \inf_{d,\mathbf{U},\mathbf{V}} d\tfrac{1-p}{p}\sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2\|\mathbf{v}_k\|_2^2 \quad s.t. \quad \mathbf{U}\mathbf{V}^\top = \mathbf{Y}. \tag{95}$$

Proposition 10. *The convex envelope of $\frac{1}{2}\Omega(\mathbf{Y})$ is $\frac{1-p}{2p}\|\mathbf{Y}\|_\star^2$.*

• If we can show that a local minimum of

$$f(\mathbf{U},\mathbf{V},d) = \left[\|\mathbf{Y} - \mathbf{U}\mathbf{V}^\top\|_F^2 + d\tfrac{1-p}{p}\sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2\|\mathbf{v}_k\|_2^2 \right] \tag{96}$$

satisfies the optimality conditions for its lower bound

$$f(\mathbf{U},\mathbf{V},d) \geq F(\mathbf{X}) = \|\mathbf{X} - \mathbf{Y}\|_F^2 + \tfrac{1-p}{p}\|\mathbf{X}\|_\star^2, \tag{97}$$

then we know we've also found a global minimum of (96).

Lemma 4. *Let $(\bar{\mathbf{U}}, \bar{\mathbf{V}}, \bar{d})$ a local minimizer of (96). Then,*

$$\left\langle \mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top \right\rangle = \bar{d}\tfrac{1-p}{p}\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 \tag{98}$$

Proof. Since $(\bar{\mathbf{U}}, \bar{\mathbf{V}})$ is a local minimizer of $f(\mathbf{U},\mathbf{V},\bar{d})$, then there exists $\delta > 0$ such that, for any $\epsilon > 0$, $\epsilon < \delta$, we must have

$$f(\bar{\mathbf{U}}, \bar{\mathbf{V}}, \bar{d}) \leq f(\bar{\mathbf{U}} + \epsilon\bar{\mathbf{U}}, \bar{\mathbf{V}} + \epsilon\bar{\mathbf{V}}, \bar{d}) = f((1+\epsilon)\bar{\mathbf{U}}, (1+\epsilon)\bar{\mathbf{V}}, \bar{d}). \tag{99}$$

That is

$$\|\mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 + \bar{d}\tfrac{1-p}{p}\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 \leq \|\mathbf{Y} - (1+\epsilon)^2\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 + \bar{d}\tfrac{1-p}{p}(1+\epsilon)^4\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2. \tag{100}$$

Exploit the first order approximations $(1+\epsilon)^2 = 1+2\epsilon+O(\epsilon^2)$ and $(1+\epsilon)^4 = 1+4\epsilon+O(\epsilon^2)$. Since $\|\mathbf{A}+(1+\epsilon)^2\mathbf{B}\|_F^2 = \|\mathbf{A}+(1+2\epsilon)\mathbf{B}\|_F^2 + O(\epsilon^2)$,

$$\|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 + \bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 \leq \tag{101}$$

$$\|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top - 2\epsilon\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 + \bar{d}\frac{1-p}{p}(1+4\epsilon)\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 + O(\epsilon^2) \tag{102}$$

and also, by deleting $\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2$ and rearranging terms,

$$0 \leq \|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top - 2\epsilon\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 - \|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 + 4\epsilon\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 + O(\epsilon^2). \tag{103}$$

Divide by $\varepsilon = 2\epsilon$,

$$0 \leq \tfrac{1}{\varepsilon}\left(\|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top - \varepsilon\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 - \|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2\right) + 2\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 + O(\varepsilon). \tag{104}$$

Take the limit as $\varepsilon \to 0$ and use the definition of one-sided directional derivative for a differentiable function h

$$\langle \nabla h(\mathbf{x}), \mathbf{d}\rangle = \nabla_{\mathbf{d}} h(\mathbf{x}) = \lim_{\varepsilon\to 0}\frac{h(\mathbf{x}+\varepsilon\mathbf{d}) - h(\mathbf{x})}{\varepsilon} \tag{105}$$

for $h(\cdot) = \|\mathbf{Y}-\cdot\|_F^2$, $\mathbf{x} = \bar{\mathbf{U}}\bar{\mathbf{V}}^\top$ and $\mathbf{d} = \bar{\mathbf{U}}\bar{\mathbf{V}}^\top$. Then,

$$0 \leq \left\langle \left[\nabla_{\mathbf{X}}\|\mathbf{X}-\mathbf{Y}\|_F^2\right](\bar{\mathbf{U}}\bar{\mathbf{V}}^\top), \bar{\mathbf{U}}\bar{\mathbf{V}}^\top\right\rangle + 2\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 \tag{106}$$

$$= 2\left\langle \bar{\mathbf{U}}\bar{\mathbf{V}}^\top - \mathbf{Y}, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top\right\rangle + 2\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 \tag{107}$$

and, once the factor 2 is simplified,

$$0 \leq -\left\langle \mathbf{Y}-\mathbf{U}\mathbf{V}^\top, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top\right\rangle + \bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2. \tag{108}$$

Note that by $\epsilon > 0$ and sufficiently small, we must also have,

$$f(\bar{\mathbf{U}}, \bar{\mathbf{V}}, \bar{d}) \leq f(\bar{\mathbf{U}} - \epsilon\bar{\mathbf{U}}, \bar{\mathbf{V}} - \epsilon\bar{\mathbf{V}}, \bar{d}) = f((1-\epsilon)\bar{\mathbf{U}}, (1-\epsilon)\bar{\mathbf{V}}, \bar{d}). \tag{109}$$

By applying the same steps as before, we get

$$0 \leq \tfrac{1}{\varepsilon}\left(\|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top + \varepsilon\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2 - \|\mathbf{Y}-\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_F^2\right) - 2\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 + O(\varepsilon). \tag{110}$$

Take the limit as $\varepsilon \to 0$ and use the definition of directional derivative (105).

$$0 \leq \left\langle \left[\nabla_{\mathbf{X}}\|\mathbf{X}-\mathbf{Y}\|_F^2\right](\bar{\mathbf{U}}\bar{\mathbf{V}}^\top), -\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\right\rangle - 2\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 \tag{111}$$

$$= 2\left\langle \bar{\mathbf{U}}\bar{\mathbf{V}}^\top - \mathbf{Y}, -\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\right\rangle - 2\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2 \tag{112}$$

$$= 2\left\langle \mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top\right\rangle - 2\bar{d}\frac{1-p}{p}\sum_{k=1}^{\bar{d}}\|\bar{\mathbf{u}}_k\|_2^2\|\bar{\mathbf{v}}_k\|_2^2. \tag{113}$$

Simplify the common factor 2, after changing the signs,

$$0 \geq -\left\langle \mathbf{Y} - \mathbf{U}\mathbf{V}^\top, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top \right\rangle + \bar{d}\frac{1-p}{p} \sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2 \tag{114}$$

The thesis comes by combining (108) and (114).

Theorem 1 (Necessary and sufficient conditions for local minima of (96) to be global for (97))**.** *Let* $(\bar{\mathbf{U}}, \bar{\mathbf{V}}, \bar{d})$ *be a local minimizer of problem* (96)*:*

$$\min_{\mathbf{U},\mathbf{V},d} \left[\|\mathbf{Y} - \mathbf{U}\mathbf{V}^\top\|_F^2 + d\frac{1-p}{p} \sum_{k=1}^{d} \|\mathbf{u}_k\|_2^2 \|\mathbf{v}_k\|_2^2 \right] \tag{115}$$

and consider its convex lower bound (97)*:*

$$\min_{\mathbf{X}} \|\mathbf{X} - \mathbf{Y}\|_F^2 + \tfrac{1-p}{p} \|\mathbf{X}\|_\star^2, \tag{116}$$

- **Sufficient conditions.** *Assume that*

$$\|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star = \sqrt{\bar{d} \sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2} \tag{117}$$

  and

$$\mathbf{p}^\top (\mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top)\mathbf{q} \leq \tfrac{1-p}{p} \|\mathbf{p}\|_2 \|\mathbf{q}\|_2 \|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star \tag{118}$$

  for any column vector $\mathbf{p} \in \mathbb{R}^m$ *and* $\mathbf{q} \in \mathbb{R}^n$*. Then* $\bar{\mathbf{U}}\bar{\mathbf{V}}^\top$ *is a global minimizer for* (97)*.*

- **Necessary condition.** *Assume that* $\bar{\mathbf{U}}\bar{\mathbf{V}}^\top$ *is a global minimizer for* (97)*. Then,* (117) *is satisfied.*

Proof. • The first order optimality condition for

$$\min_{\mathbf{X}} \|\mathbf{X} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{X}\|_\star^2 \tag{119}$$

is

$$0 \in \nabla_{\mathbf{X}} \|\mathbf{X} - \mathbf{Y}\|_F^2 + \partial\left(\tfrac{1-p}{p} \|\mathbf{X}\|_\star^2\right) \tag{120}$$

$$\Leftrightarrow 0 \in 2(\mathbf{X} - \mathbf{Y}) + 2\tfrac{1-p}{p} \|\mathbf{X}\|_\star \partial \|\mathbf{X}\|_\star \tag{121}$$

$$\Leftrightarrow \mathbf{Y} - \mathbf{X} \in \tfrac{1-p}{p} \|\mathbf{X}\|_\star \partial \|\mathbf{X}\|_\star \tag{122}$$

$$\Leftrightarrow \frac{\mathbf{Y} - \mathbf{X}}{\frac{1-p}{p}\|\mathbf{X}\|_\star} \in \partial \|\mathbf{X}\|_\star. \tag{123}$$

Recall that

$$\partial \|\mathbf{X}\|_\star = \left\{ \mathbf{W} \in \mathbb{R}^{m \times n} \colon \langle \mathbf{X}, \mathbf{W} \rangle = \|\mathbf{X}\|_\star, \quad \mathbf{p}^\top \mathbf{W} \mathbf{q} \leq \|\mathbf{p}\|_2 \|\mathbf{q}\|_2 \,\forall\, (\mathbf{p}, \mathbf{q}) \right\}. \tag{124}$$

Therefore, since (118) easily reads as the inequality in (124) for

$$\mathbf{W} = \frac{\mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top}{\frac{1-p}{p}\|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star}, \tag{125}$$

the thesis will follow if we show

$$\left\langle \bar{\mathbf{U}}\bar{\mathbf{V}}^\top, \frac{\mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top}{\frac{1-p}{p}\|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star} \right\rangle = \|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star \tag{126}$$

or, equivalently,

$$\left\langle \mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top \right\rangle = \tfrac{1-p}{p} \|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star^2. \tag{127}$$

We can observe that (127) easily follows from (117) if considering equation (98) of Lemma 4.

- Let $(\bar{\mathbf{U}}, \bar{\mathbf{V}}, \bar{d})$ be a global minimizer of (96) and assume that $\bar{\mathbf{U}}\bar{\mathbf{V}}^\top$ satisfies the optimality conditions for (97).

  Since $(\bar{\mathbf{U}}, \bar{\mathbf{V}}, \bar{d})$ is a local minimizer of (96), then, thanks to Lemma 4, we have

$$\left\langle \mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top \right\rangle = \bar{d}\tfrac{1-p}{p} \sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2 \tag{128}$$

  At the same time, since $\bar{\mathbf{U}}\bar{\mathbf{V}}^\top$ satisfies the optimality conditions for (97), we have

$$\mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top \in \tfrac{1-p}{p} \|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star \partial \|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star. \tag{129}$$

  Considering that

$$\partial \|\mathbf{X}\|_\star = \left\{ \mathbf{W} \colon \langle \mathbf{X}, \mathbf{W} \rangle = \|\mathbf{X}\|_\star, \quad \mathbf{p}^\top \mathbf{W} \mathbf{q} \leq \|\mathbf{p}\|_2 \|\mathbf{q}\|_2 \; \forall \, (\mathbf{p}, \mathbf{q}) \right\}, \tag{130}$$

  in particular,

$$\langle \mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top, \bar{\mathbf{U}}\bar{\mathbf{V}}^\top \rangle = \tfrac{1-p}{p} \|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star^2. \tag{131}$$

  Combine (98) and (131), we get

$$\|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star^2 = \bar{d} \sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2 \tag{132}$$

  and (117) follows by square-rooting each member of (132). □

Lemma 5. *For any $n \in \mathbb{N}_+$ and $a_1, \ldots, a_n \in \mathbb{R}$, we get*

$$\left(\sum_{k=1}^{n} a_k \right)^2 \leq n \sum_{k=1}^{n} a_k^2 \tag{133}$$

Proof. For any n-dimensional column vectors $\mathbf{x}$ and $\mathbf{y}$, Cauchy-Schwartz inequality is

$$\mathbf{x}^\top \mathbf{y} \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \tag{134}$$

and therefore

$$(\mathbf{x}^\top \mathbf{y})^2 \leq \|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2. \tag{135}$$

The thesis follows by applying the previous inequality to the case $\mathbf{x} = [a_1, \ldots, a_n]^\top$ and $\mathbf{y} = [1, \ldots, 1]$. Indeed, in such a case,

$$(\mathbf{x}^\top \mathbf{y})^2 = \sum_{k=1}^{n} a_k, \quad \|\mathbf{x}\|_2^2 = \sum_{k=1}^{n} a_k^2 \quad \text{and} \quad \|\mathbf{y}\|_2^2 = n. \tag{136}$$

□

Observation: a consequence of (117)**, i.e. another necessary condition.**

Assume that $(\bar{\mathbf{U}}, \bar{\mathbf{V}}, \bar{d})$ is a local minimizer of (96), being at the same time $\bar{\mathbf{U}}\bar{\mathbf{V}}^\top$ the global minimizer of (97). Then, for what we said, equation (117) is satisfied. That is,

$$\|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star = \sqrt{\bar{d} \sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2}. \tag{137}$$

Exploit the variational form of the nuclear norm to obtain

$$\sqrt{\bar{d} \sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2} \leq \sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2 \|\bar{\mathbf{v}}_k\|_2 \tag{138}$$

and, equivalently,

$$\bar{d}\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2 \leq \left(\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2 \|\bar{\mathbf{v}}_k\|_2\right)^2. \tag{139}$$

Apply (133) for $n = \bar{d}$ and $a_k = \|\bar{\mathbf{u}}_k\|_2 \|\bar{\mathbf{v}}_k\|_2$. Then

$$\left(\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2 \|\bar{\mathbf{v}}_k\|_2\right)^2 \leq \bar{d}\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2. \tag{140}$$

By combining the last two inequalities, we get

$$\boxed{\bar{d}\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2^2 \|\bar{\mathbf{v}}_k\|_2^2 = \left(\sum_{k=1}^{\bar{d}} \|\bar{\mathbf{u}}_k\|_2 \|\bar{\mathbf{v}}_k\|_2\right)^2} \tag{141}$$

Analyzing the polar

We are interested in a better understanding of equation (118):

$$\mathbf{p}^\top(\mathbf{Y} - \bar{\mathbf{U}}\bar{\mathbf{V}}^\top)\mathbf{q} \leq \tfrac{1-p}{p}\|\mathbf{p}\|_2\|\mathbf{q}\|_2\|\bar{\mathbf{U}}\bar{\mathbf{V}}^\top\|_\star \qquad \text{for all } \mathbf{p} \in \mathbb{R}^m \text{ and } \mathbf{q} \in \mathbb{R}^n. \tag{142}$$

Therefore, we will provide a few sufficient conditions built on top of the following technical result.

Lemma 6. *For any* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *and* $\mathbf{x} \in \mathbb{R}^n$*,*

$$\|\mathbf{A}\mathbf{x}\|_2 \leq \|\mathbf{A}\|_2\|\mathbf{x}\|_2 \tag{143}$$

$$\|\mathbf{A}\mathbf{x}\|_2 \leq \|\mathbf{A}\|_F\|\mathbf{x}\|_2 \tag{144}$$

$$\|\mathbf{A}\mathbf{x}\|_2 \leq \sqrt{mn}\|\mathbf{A}\|_{\max}\|\mathbf{x}\|_2 \tag{145}$$

Proof of (143). The thesis comes after the definition of operator norm,

$$\|\mathbf{A}\|_2 = \sup\left\{\tfrac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} : \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}\right\}. \tag{146}$$

Proof of (144). We get (144) from (143) by using the fact that $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F$ through the following chain of inequalities.

$$\|\mathbf{A}\|_2^2 = \sigma_1(\mathbf{A})^2 = \max_i[\sigma_i(\mathbf{A})^2] = \max_i \lambda_i(\mathbf{A}^\top\mathbf{A}) \leq \sum_{i=1}^m \lambda_i(\mathbf{A}^\top\mathbf{A}) = \operatorname{trace}(\mathbf{A}^\top\mathbf{A}) = \|\mathbf{A}\|_F^2. \tag{147}$$

Proof of (145). By definition of Euclidean norm,

$$\|\mathbf{A}\mathbf{x}\|_2^2 = \sum_{i=1}^m\left(\sum_{j=1}^n A_{ij}x_j\right)^2. \tag{148}$$

For any $i = 1, \ldots, m$ apply (133).

$$\begin{aligned}\|\mathbf{A}\mathbf{x}\|_2^2 &\leq n\sum_{i=1}^m\sum_{j=1}^n A_{ij}^2x_j^2 \leq n\sum_{i=1}^m\sum_{j=1}^n \max_{p,q} A_{pq}^2x_j^2\\ &= mn\max_{p,q} A_{pq}^2\sum_{j=1}^n x_j^2 = mn\max_{p,q} A_{pq}^2\|\mathbf{x}\|_2^2.\end{aligned} \tag{149}$$

The thesis comes by square-rooting the extremal terms in the chain of equalities (149) and noticing that

$$\sqrt{\max_{p,q} A_{pq}^2} = \max_{p,q}|A_{pq}| = \|\mathbf{A}\|_{\max}. \tag{150}$$

□

Proposition 11. *Equation* (118)*:*

$$\mathbf{p}^\top(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q} \le \tfrac{1-p}{p}\|\mathbf{p}\|_2\|\mathbf{q}\|_2\|\mathbf{U}\mathbf{V}^\top\|_\star \tag{151}$$

is satisfied for any column vectors $\mathbf{p}\in\mathbb{R}^m$ *and* $\mathbf{q}\in\mathbb{R}^n$ *if any of the following conditions are matched.*

1. $$\sigma_1(\mathbf{Y}) \le \tfrac{1-p}{p}\|\mathbf{U}\mathbf{V}^\top\|_\star - \sigma_1(\mathbf{U}\mathbf{V}^\top) = \tfrac{1-p}{p}\sum_{i=1}^{\min(m,n)}\sigma_i(\mathbf{U}\mathbf{V}^\top) - \sigma_1(\mathbf{U}\mathbf{V}^\top) \tag{152}$$
2. (117) *and* (98) *hold for* $\mathbf{U},\mathbf{V}$*, each having* d *columns, and*
$$\|\mathbf{Y}\|_F^2 \le \|\mathbf{U}\mathbf{V}^\top\|_F^2 - \tfrac{1-p}{p}\|\mathbf{U}\mathbf{V}^\top\|_\star\left(2\|\mathbf{U}\mathbf{V}^\top\|_\star - 1\right) \tag{153}$$
3. $$\|\mathbf{Y}-\mathbf{U}\mathbf{V}^\top\|_{\max} \le \frac{1}{\sqrt{mn}}\tfrac{1-p}{p}\|\mathbf{U}\mathbf{V}^\top\|_\star \tag{154}$$

Proof. As the main tool to prove (152), (153) and (154), we notice that, thanks to Cauchy-Schwartz inequality, we get

$$\mathbf{p}^\top(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q} \le \|\mathbf{p}\|_2\|(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \tag{155}$$

for any column vector $\mathbf{p}\in\mathbb{R}^m$ and $\mathbf{q}\in\mathbb{R}^n$.

1. We will obtain (118) if we are able to show that (152) implies

$$\|(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \le \tfrac{1-p}{p}\|\mathbf{q}\|_2\|\mathbf{U}\mathbf{V}^\top\|_\star. \tag{156}$$

By means of (143), we get

$$\|(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \le \|\mathbf{Y}-\mathbf{U}\mathbf{V}^\top\|_2\|\mathbf{q}\|_2 = \sigma_1(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\|\mathbf{q}\|_2. \tag{157}$$

Since, for any pairs of matrices $\mathbf{A},\mathbf{B}\in\mathbb{R}^{m\times n}$ one gets

$$\sigma_1(\mathbf{A}+\mathbf{B}) \le \sigma_1(\mathbf{A}) + \sigma_1(\mathbf{B}), \tag{158}$$

then

$$\|(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \le \left(\sigma_1(\mathbf{Y}) + \sigma_1(\mathbf{U}\mathbf{V}^\top)\right)\|\mathbf{q}\|_2, \tag{159}$$

since the singular values of $-\mathbf{U}\mathbf{V}^\top$ are the same ones of $\mathbf{U}\mathbf{V}^\top$. It's easy to see that (159) combined with the assumption (152) gives (156).

2. We will obtain (118) if we are able to show that (153) implies

$$\|(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \le \tfrac{1-p}{p}\|\mathbf{q}\|_2\|\mathbf{U}\mathbf{V}^\top\|_\star \tag{160}$$

Since, by using (144), we get

$$\|(\mathbf{Y}-\mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \le \|\mathbf{Y}-\mathbf{U}\mathbf{V}^\top\|_F\|\mathbf{q}\|_2. \tag{161}$$

Now,

$$\|\mathbf{Y}-\mathbf{U}\mathbf{V}^\top\|_F^2 = \|\mathbf{Y}\|_F^2 + \|\mathbf{U}\mathbf{V}^\top\|_F^2 - 2\langle\mathbf{Y},\mathbf{U}\mathbf{V}^\top\rangle \tag{162}$$

$$= \|\mathbf{Y}\|_F^2 + \|\mathbf{U}\mathbf{V}^\top\|_F^2 - 2\langle\mathbf{Y}-\mathbf{U}\mathbf{V}^\top+\mathbf{U}\mathbf{V}^\top,\mathbf{U}\mathbf{V}^\top\rangle \tag{163}$$

$$= \|\mathbf{Y}\|_F^2 - \|\mathbf{U}\mathbf{V}^\top\|_F^2 - 2\langle\mathbf{Y}-\mathbf{U}\mathbf{V}^\top,\mathbf{U}\mathbf{V}^\top\rangle \tag{164}$$

and, by combining (98) with (117), we get

$$\|\mathbf{Y}-\mathbf{U}\mathbf{V}^\top\|_F^2 = \|\mathbf{Y}\|_F^2 - \|\mathbf{U}\mathbf{V}^\top\|_F^2 - 2\tfrac{1-p}{p}\|\mathbf{U}\mathbf{V}^\top\|_\star^2. \tag{165}$$

Therefore, $\|\mathbf{Y}-\mathbf{U}\mathbf{V}^\top\|_F^2 \le \tfrac{1-p}{p}\|\mathbf{U}\mathbf{V}^\top\|_\star$ rewrites

$$\|\mathbf{Y}\|_F^2 - \|\mathbf{U}\mathbf{V}^\top\|_F^2 - 2\tfrac{1-p}{p}\|\mathbf{U}\mathbf{V}^\top\|_\star^2 - \tfrac{1-p}{p}\|\mathbf{U}\mathbf{V}^\top\|_\star \le 0 \tag{166}$$

and therefore we get the thesis.

3. We will obtain (118) if we are able to show that (154) implies

$$\|(\mathbf{Y} - \mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \le \tfrac{1-p}{p}\|\mathbf{q}\|_2\|\mathbf{U}\mathbf{V}^\top\|_\star. \tag{167}$$

   But, this is straightforward as consequence of (145) which states

$$\|(\mathbf{Y} - \mathbf{U}\mathbf{V}^\top)\mathbf{q}\|_2 \le \sqrt{mn}\|\mathbf{Y} - \mathbf{U}\mathbf{V}^\top\|_{\max}\|\mathbf{q}\|_2. \tag{168}$$

□

6 Closed-form solution for the convex lower bound of dropout for MF

Proposition 12. *Let* $\mathbf{X} = \mathbf{L}\mathbf{\Sigma}\mathbf{R}^\top$ *be the singular value decomposition of* $\mathbf{X}$. *The optimal solution to*

$$\min_{\mathbf{Y}} \quad \|\mathbf{X} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{Y}\|_\star^2 \tag{169}$$

is given by $\mathbf{Y} = \mathbf{L}\mathcal{S}_\mu(\mathbf{\Sigma})\mathbf{R}^\top$, *where* $\lambda > 0$, $\mu = \frac{\lambda d}{1+\lambda d}\bar{\sigma}_d(\mathbf{X})$, $\bar{\sigma}_d(\mathbf{X})$ *is the average of the top* d *singular values of* $\mathbf{X}$, d *represents the largest integer such that* $\sigma_d(\mathbf{X}) > \frac{\lambda d}{1+\lambda d}\bar{\sigma}_d(\mathbf{X})$, *and* $\mathcal{S}_\mu$ *is defined as the shrinkage thresholding operator which set to zero all singular values of* $\mathbf{X}$ *which are less or equal to* μ.

Proof. Since both the nuclear norm $\|\cdot\|_\star$ and the Frobenius norm $\|\cdot\|_F$ are rotationally invariant, up to non-restrictive rotations applied to the data matrix $\mathbf{X}$, the thesis can be equivalently proved by considering the following result.

Let $\mathbf{x} = [x_1, \ldots, x_r]$ a fixed vector with $x_i \geq x_{i+1} > 0$. Define μ_d as the average of the first d entries of $\mathbf{x}$ Then, the optimal solution to the optimization problem

$$\min_{\mathbf{a}\in\mathbb{R}^r} \|\mathbf{a} - \mathbf{x}\|_2^2 + \lambda\|\mathbf{a}\|_1^2 \tag{170}$$

is given by $\mathbf{a} = [a_1, \ldots, a_r]$ where

$$a_i = \begin{cases} x_i - \dfrac{\lambda d}{1+\lambda d}\mu_d & i = 1, \ldots, d \\ 0 & i = d+1, \ldots, r \end{cases} \tag{171}$$

where d is the largest positive integer less or equal to r such that all ai given in (171) are positive.

In order to prove this claim, first note that the objective function is strictly convex and, hence, there is a unique global minimum. If $\lambda = 0$ the global minimizer is precisely $\mathbf{x}$, which is consistent with the formula given in the statement of the proposition. So, suppose that $\lambda > 0$. Next, notice that if $\mathbf{a} = [a_1, a_2, \ldots, a_r]$ is an optimal solution, then all a_i must be non-negative. Indeed, if say $a_1 < 0$, then the vector $[-a_1, a_2, \ldots, a_r]$ already gives a smaller objective value. Now, the first order optimality condition of our problem rewrites

$$\mathbf{0} \in (\mathbf{a} - \mathbf{x}) + \lambda\|\mathbf{a}\|_1 \partial\|\mathbf{a}\|_1. \tag{172}$$

There are two cases for each coordinate i of (172).

$$a_i = x_i - \lambda\|\mathbf{a}\|_1, \text{ if } a_i > 0, \text{ and } x_i = \lambda\|\mathbf{a}\|_1\xi_i, \text{ if } a_i = 0. \tag{173}$$

where ξ_i in (173) is some number in the interval $[0, 1]$. Notice that since $x_i > 0$ for every i, the second condition in (173) guarantees that the global solution can not be the zero vector, otherwise $\|\mathbf{a}\|_1 = 0$ and so $x_i = 0$ for every i. Thus, suppose that exactly the first $k \geq 1$ coordinates of $\mathbf{a}$ are non-zero. Then sum the equations $a_i = x_i - \lambda k\|\mathbf{a}\|_1$ for $i = 1, \ldots, k$. We get

$$\|\mathbf{a}\|_1 = k\mu_k - \lambda k\|\mathbf{a}\|_1 \tag{174}$$

which gives

$$\|\mathbf{a}\|_1 = \frac{k}{1+\lambda k}\mu_k. \tag{175}$$

Then (173) and (175) give

$$a_i = x_i - \frac{\lambda k}{1+\lambda k}\mu_k > 0 \text{ for } i = 1, \ldots, k \text{ and } a_i = 0 \text{ for } i = k+1, \ldots, r. \tag{176}$$

Now, let d be the largest integer such that $a_i = x_i - \dfrac{\lambda d}{1+\lambda d}\mu_d > 0$ and define the vector

$$\mathbf{v} = \left[x_1 - \frac{\lambda d}{1+\lambda d}\mu_d, \ldots, x_d - \frac{\lambda d}{1+\lambda d}\mu_d, \underbrace{0, \ldots, 0}_{r-d \text{ times}}\right]. \tag{177}$$

If $d = r$, then $\mathbf{v}$ satisfies the optimality condition (173) and so it is the global minimizer. So suppose that $d < r$. In that case, to show that $\mathbf{v}$ is the global minimizer it suffices to show that

$$x_{d+1} - \frac{\lambda d}{1 + \lambda d}\mu_d \leq 0. \tag{178}$$

since this is equivalent to saying that for any $i >$d there exists $\xi_i \in [0, 1]$ such that $x_i = \lambda \|\mathbf{v}\|_1 \xi_i$ in which case $\mathbf{v}$ satisfies the optimality condition (173). Now by the maximality of d, we have that

$$x_{d+1} - \frac{\lambda(d+1)}{1 + \lambda(d+1)}\mu_{d+1} \leq 0. \tag{179}$$

Equivalently, we get the following chain of inequalities

$$\left(1 - \frac{\lambda}{1 + \lambda(d+1)}\right) x_{d+1} - \frac{\lambda}{1 + \lambda(d+1)} \sum_{k=1}^{d} x_k \leq 0 \tag{180}$$

$$\frac{1 + \lambda d}{1 + \lambda(d+1)} x_{d+1} - \frac{\lambda}{1 + \lambda(d+1)} \sum_{k=1}^{d} x_k \leq 0 \tag{181}$$

$$x_{d+1} - \frac{\lambda d}{1 + \lambda d}\mu_d \leq 0 \tag{182}$$

from which we obtain the desired condition. □

7 Complementary results for the numerical simulation

In this Section we will present the complete version of Figure 1 in the paper. Also, we provide more examples of digits' reconstruction on MNIST dataset.

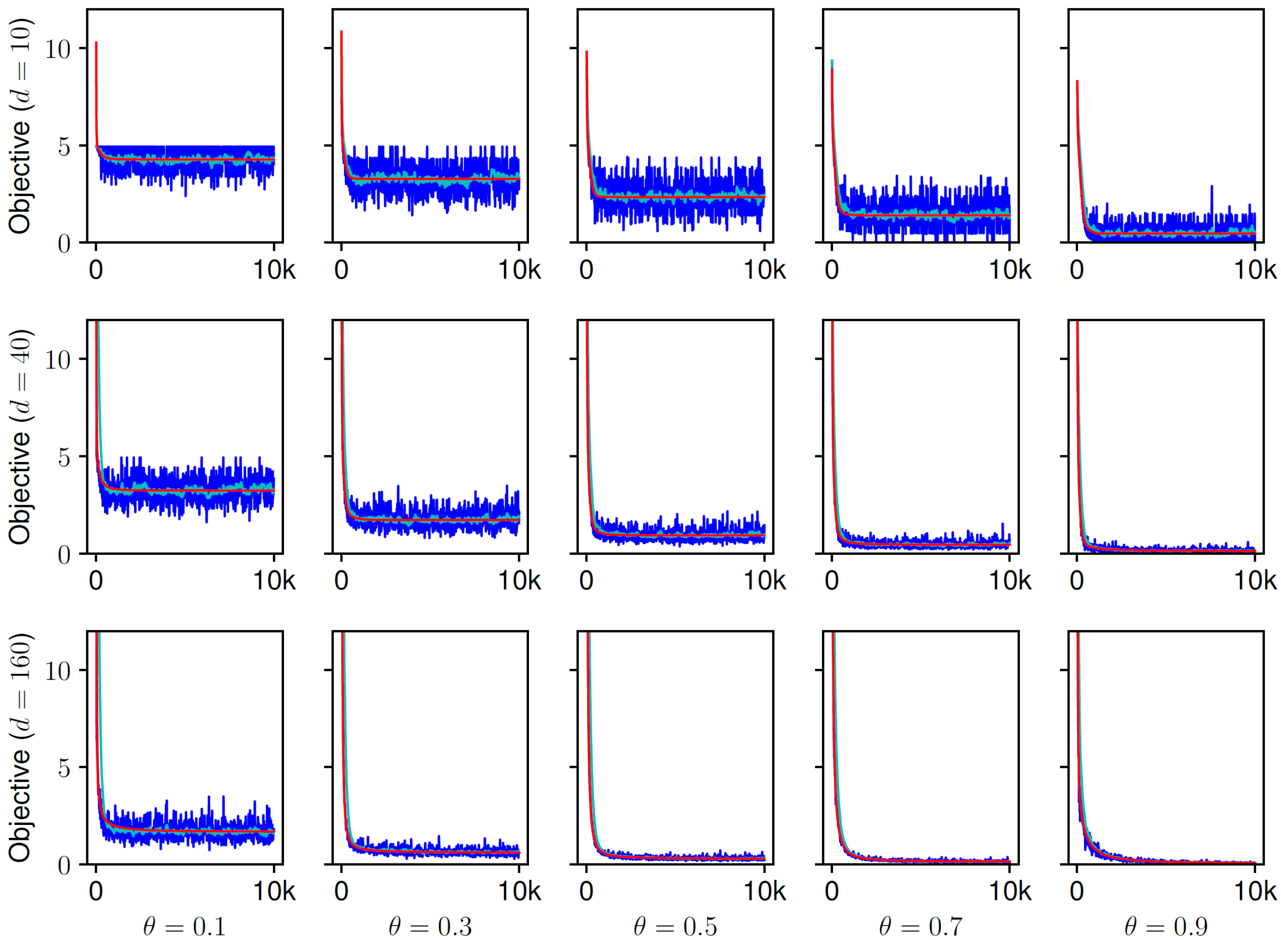

Figure 1: For $\theta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and $d \in \{10, 40, 160\}$ we compare the deterministic problem (red) with its stochastic counterpart (blue). The exponential moving average of the stochastic objective is shown in cyan. Best viewed in color.

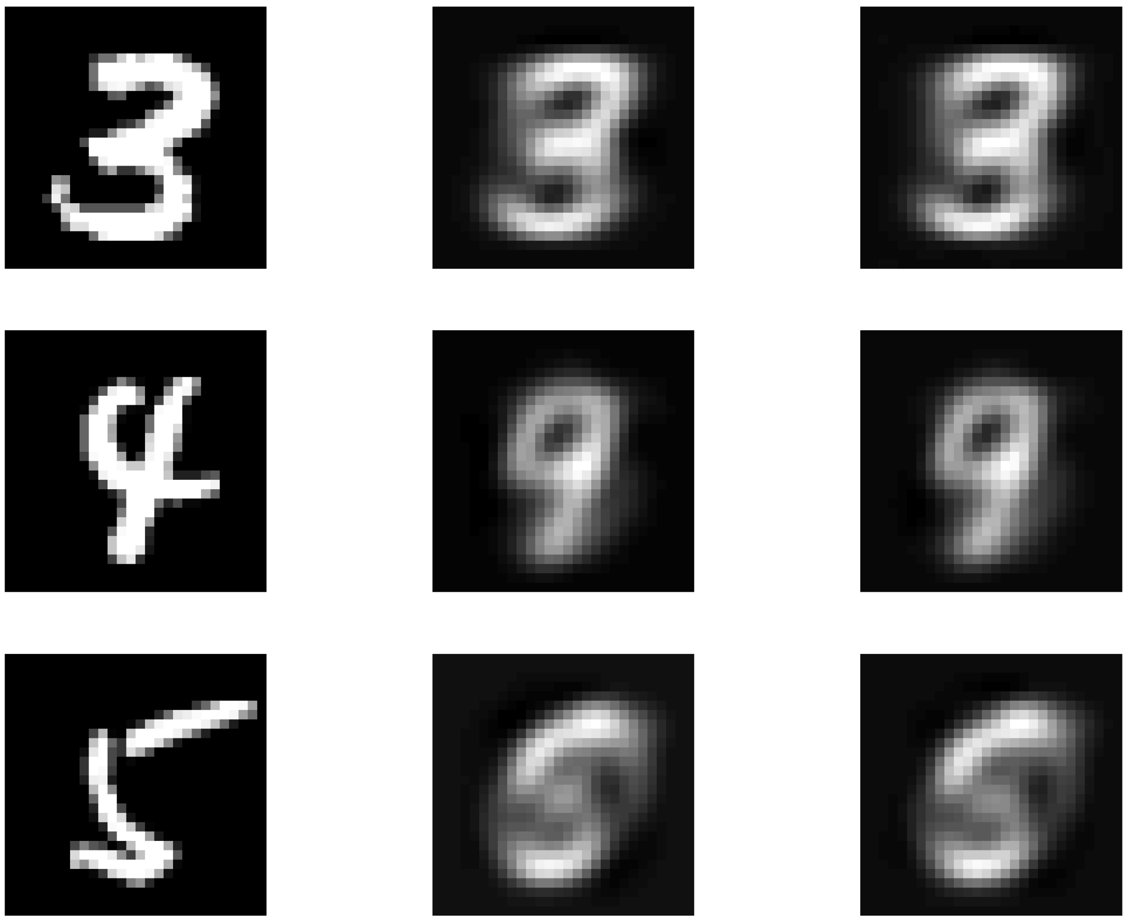

Figure 2: Experiments on MNIST dataset, whose original images are reported in the first column. For each of those, we compute dropout for MF with $\theta = 0.5$ (second column) and the corresponding closed-form solution (third column).

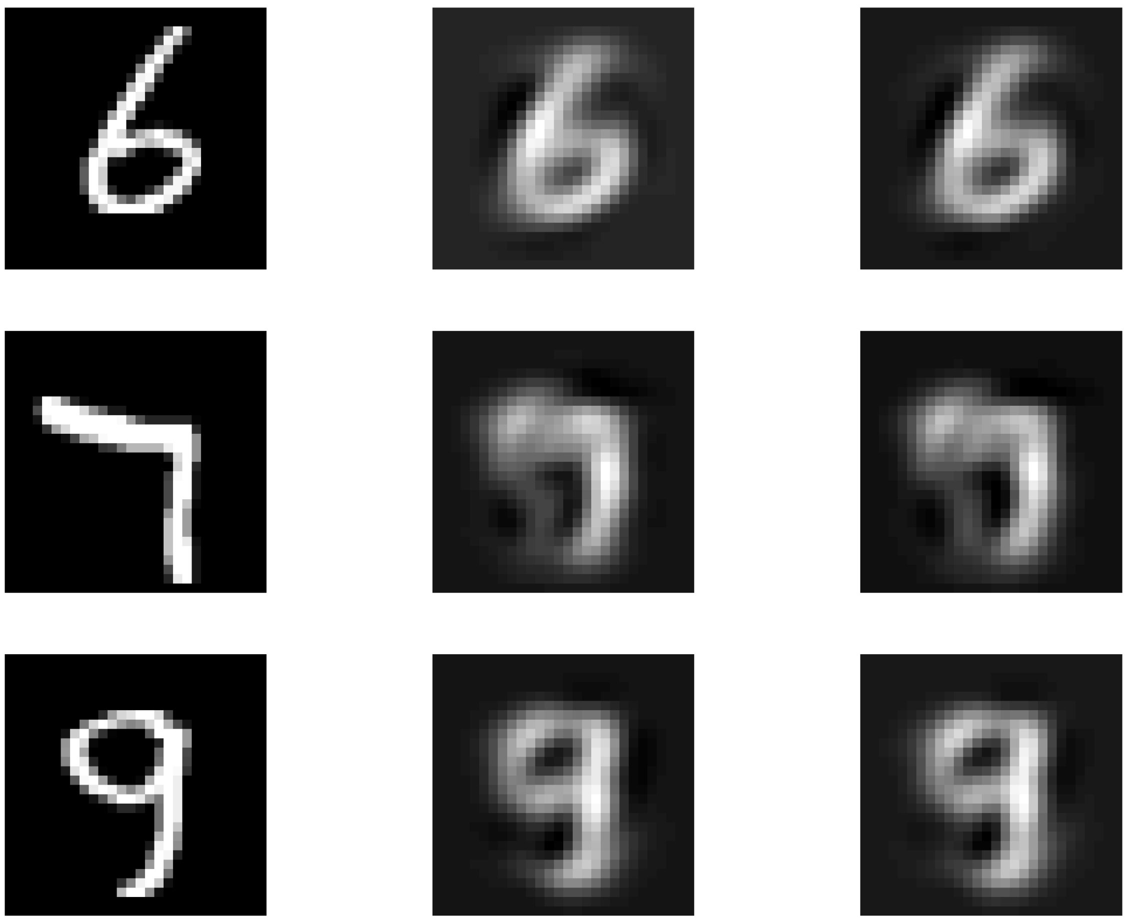

Figure 3: Experiments on MNIST dataset, whose original images are reported in the first column. For each of those, we compute dropout for MF with $\theta = 0.5$ (second column) and the corresponding closed-form solution (third column).

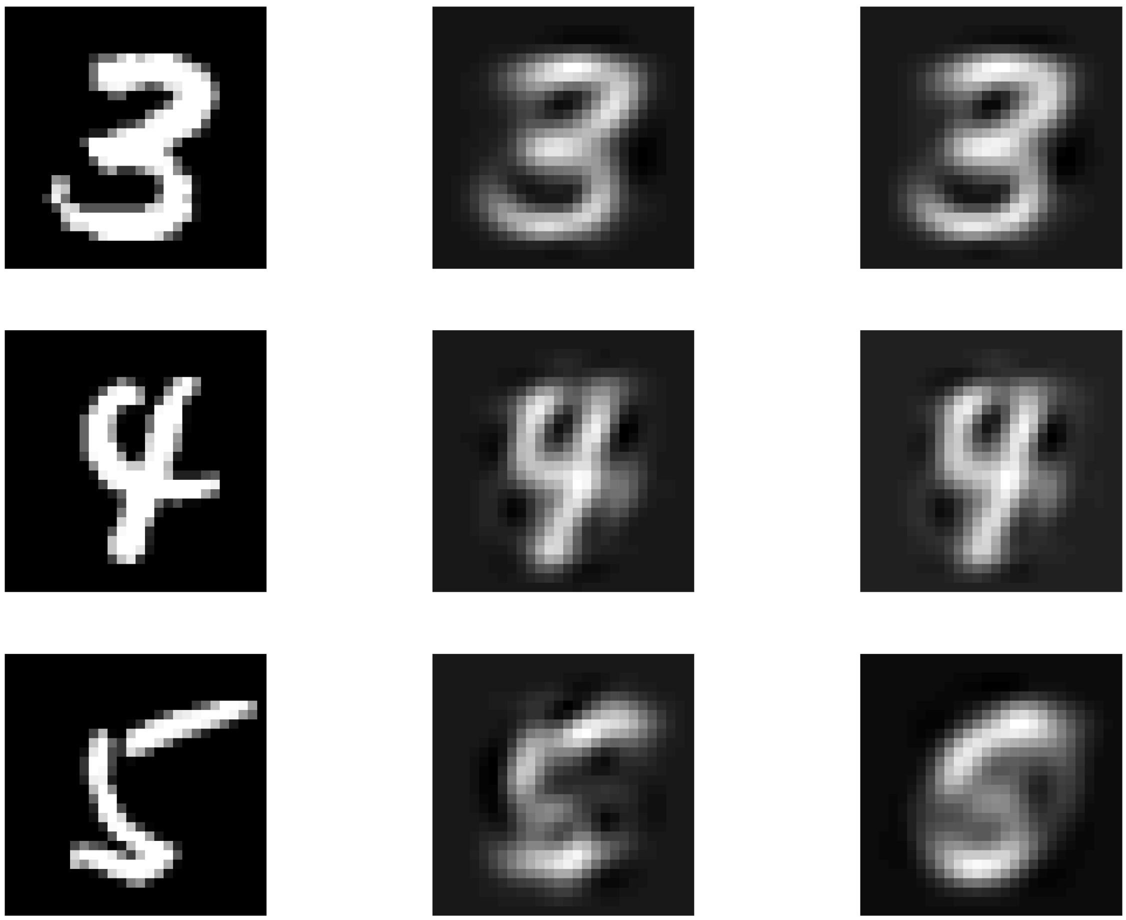

Figure 4: Experiments on MNIST dataset, whose original images are reported in the first column. For each of those, we compute dropout for MF with $\theta = 0.8$ (second column) and the corresponding closed-form solution (third column).

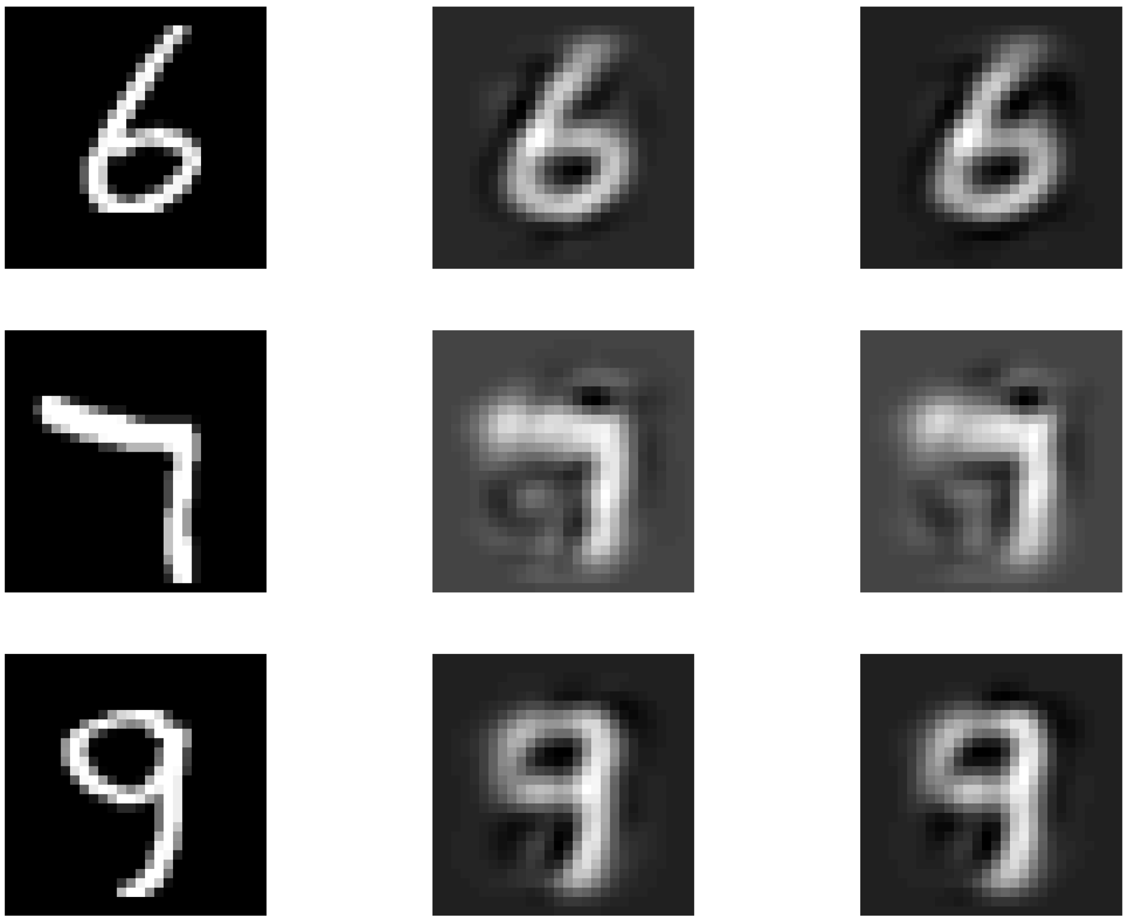

Figure 5: Experiments on MNIST dataset, whose original images are reported in the first column. For each of those, we compute dropout for MF with $\theta = 0.8$ (second column) and the corresponding closed-form solution (third column).