

# Fast Dictionary Learning with a Smoothed Wasserstein Loss

## Supplementary Material

### 1 Proof of Theorem 3

**Theorem.** *The solution  $\Lambda^*$  of*

$$\min_{\Lambda \in \Sigma_k^m} \sum_{i=1}^m H_{X_i}(D\Lambda_i) - \rho_1 E(\Lambda_i) \text{ s.t. } D\Lambda \in \Sigma_s^m \quad (1)$$

is such that  $\forall i = 1 \dots m$ ,  $\Lambda_i^* = \frac{e^{-D^T G_i^* / \rho_1}}{\langle e^{-D^T G_i^* / \rho_1}, \mathbf{1} \rangle}$ , with

$$G_i^* \in \underset{g}{\operatorname{argmin}} H_{X_i}^*(G_i) - \rho_1 E_*(-D^T G_i / \rho_1). \quad (2)$$

*Proof.* Similarly to the proof of Theorem 1, the dual is

$$\begin{aligned} & \max_{G \in \mathbb{R}^{s \times m}} \min_{\Lambda \in \Sigma_k^m, Q \in \Sigma_s^m} \sum_{i=1}^m W_\gamma(X_i, Q_i) - \rho_1 E(\Lambda_i) + \langle D\Lambda - Q, G \rangle \\ &= \max_{G \in \mathbb{R}^{s \times m}} \min_{\Lambda \in \Sigma_k^m} \sum_{i=1}^m -\rho_1 E(\Lambda_i) + \langle \Lambda_i, D^T G_i \rangle - H_{X_i}^*(G_i) \\ &= \max_{G \in \mathbb{R}^{s \times m}} \min_{\Lambda \in \Sigma_k^m} \sum_{i=1}^m -\rho_1 E(\Lambda_i) - \rho \langle \Lambda_i, -D^T G_i / \rho_1 \rangle - H_{X_i}^*(G_i) \\ &= \max_{G \in \mathbb{R}^{s \times m}} \sum_{i=1}^m \rho_1 E_*(-D^T G_i / \rho_1) - H_{X_i}^*(G_i). \end{aligned} \quad (3)$$

Which is the same optimisation problem as Problem (2). The primal-dual relationship is derived from Equation (3). Note that since the derivative of  $x \rightarrow x \log x$  goes to infinity at 0,  $\Lambda^*$  is in the interior of  $\mathbb{R}_+^{k \times m}$ . So the gradient of the objective with respect to  $\Lambda$  is orthogonal to the simplex on  $\Lambda^*$ , *i.e.* is constant:  $\rho_1 \log \Lambda_i^* + D^T G_i^* = c \mathbf{1}$ . Thus  $\Lambda_i^* = e^{(c \mathbf{1} - D^T G_i^*) / \rho_1}$ . Since  $\Lambda_i^* \in \Sigma_k$ ,  $\langle \Lambda^*, \mathbf{1} \rangle = 1$ ,  $c = -\log \langle e^{-D^T G_i^* / \rho_1}, \mathbf{1} \rangle$  and  $\Lambda_i^* = \frac{e^{-D^T G_i^* / \rho_1}}{\langle e^{-D^T G_i^* / \rho_1}, \mathbf{1} \rangle}$ .

Note that the first order condition of Equation (3) with respect to  $G$  gives the same solution.  $\square$

### 2 Topics

Table 1: 24 topics learned on Reuters with Wasserstein NMF, the target words are a set of 3,000 common English words

| 1  | 2  | 3  | 4  | 5   | 6   |
|--|--|--|--|---|---|
| quarter<br>increase<br>result<br>share<br>report<br>expect<br>due<br>rise<br>cost<br>significant | offer<br>bid<br>tender<br>plan<br>request<br>approve<br>board<br>make<br>receive<br>seek             | million<br>income<br>tax<br>note<br>cash<br>pay<br>total<br>debt<br>net<br>amount            | split<br>board<br>stock<br>record<br>approve<br>annual<br>declare<br>company<br>set<br>april               | net<br>saint<br>note<br>company<br>profit<br>million<br>share<br>cheerio<br>refresh<br>situate      | rate<br>interest<br>market<br>bank<br>money<br>rise<br>today<br>forecast<br>percent<br>dollar |
| 7  | 8  | 9  | 10   | 11  | 12  |
| trade<br>japan<br>states<br>export<br>foreign<br>minister<br>policy<br>market<br>economy<br>deal | purchase<br>deal<br>sell<br>sale<br>buy<br>cash<br>subject<br>stake<br>plan<br>amount                | oil<br>gas<br>price<br>barrel<br>energy<br>increase<br>day<br>supply<br>raise<br>coast       | loss<br>net<br>profit<br>wagon<br>year<br>million<br>cheerio<br>note<br>suck<br>refresh                    | stock<br>share<br>common<br>stake<br>exchange<br>commission<br>total<br>company<br>percent<br>price | million<br>net<br>cheerio<br>profit<br>news<br>telecom<br>refresh<br>plonk<br>litre<br>cock   |
| 13   | 14   | 15   | 16   | 17  | 18  |
| include<br>tax<br>gain<br>sale<br>current<br>fund<br>income<br>result<br>million<br>such         | group<br>unit<br>stake<br>company<br>sell<br>telecom<br>part<br>news<br>plan<br>would                | net<br>refresh<br>cheerio<br>suck<br>plonk<br>situate<br>engine<br>news<br>cock<br>reproduce | technology<br>international<br>system<br>service<br>provide<br>data<br>make<br>telecom<br>research<br>news | chairman<br>comment<br>news<br>week<br>official<br>west<br>interview<br>today<br>monday<br>part     | bank<br>loan<br>credit<br>debt<br>mortgage<br>fund<br>state<br>public<br>money<br>part        |
| 19   | 20   | 21   | 22   | 23  | 24  |
| pay<br>record<br>set<br>wagon<br>april<br>previous<br>bracket<br>before<br>curl<br>news          | april<br>february<br>june<br>october<br>news<br>july<br>december<br>january<br>november<br>september | year<br>million<br>end<br>last<br>month<br>period<br>ago<br>week<br>after<br>time            | wagon<br>cheerio<br>profit<br>litre<br>net<br>news<br>plonk<br>cock<br>brake<br>lorry                      | company<br>business<br>firms<br>industry<br>firm<br>telecom<br>product<br>stake<br>news<br>same     | profit<br>company<br>telecom<br>percent<br>retail<br>pence<br>net<br>rise<br>tumble<br>cent   |

Table 2: 15 topics learned on BBCsport with Wasserstein NMF, the target words are a set of 3,000 common English words

| 1   | 2  | 3  | 4   | 5  |
|---|--|--|---|--|
| play<br>game<br>team<br>cup<br>world<br>good<br>top<br>great<br>coach<br>important    | wales<br>rugby<br>captain<br>bath<br>nick<br>sir<br>rob<br>centre<br>jack<br>brown                 | year<br>april<br>time<br>july<br>june<br>january<br>season<br>career<br>october<br>make        | scotland<br>france<br>irish<br>italy<br>game<br>coach<br>side<br>match<br>paris<br>win        | knee<br>back<br>return<br>left<br>week<br>fit<br>problem<br>hope<br>time<br>play             |
| 6   | 7  | 8  | 9   | 10   |
| league<br>club<br>boss<br>ham<br>cup<br>states<br>side<br>match<br>draw<br>coach      | make<br>people<br>interview<br>television<br>criticism<br>lot<br>thing<br>talk<br>money<br>referee | club<br>deal<br>coach<br>contract<br>move<br>offer<br>future<br>summer<br>chairman<br>chief    | win<br>good<br>victory<br>lose<br>feel<br>happy<br>hard<br>beat<br>final<br>game              | nought<br>ball<br>captain<br>total<br>run<br>butcher<br>percent<br>catch<br>mark<br>bat      |
| 11  | 12   | 13   | 14  | 15   |
| goal<br>minute<br>kick<br>ball<br>half<br>penalty<br>referee<br>lead<br>score<br>side | seed<br>tennis<br>open<br>set<br>april<br>match<br>beat<br>champion<br>july<br>round               | cricket<br>test<br>india<br>australia<br>africa<br>south<br>zealand<br>match<br>series<br>tour | champion<br>world<br>race<br>gold<br>event<br>record<br>britain<br>london<br>compete<br>title | test<br>drug<br>ban<br>guilty<br>decision<br>sport<br>court<br>evidence<br>committee<br>case |

Table 3: 24 topics learned on Reuters with Wasserstein NMF, the target words are a set of 6,000 words in French

| 1   | 2   | 3   | 4  | 5  | 6  |
|---|---|---|--|--|--|
| note<br>aex<br>comprend<br>corp<br>conclut<br>electrabel<br>claes<br>motorola<br>sambre<br>cools                            | march<br>conomique<br>nationales<br>conomiques<br>dveloppement<br>international<br>national<br>ncessit<br>niveau<br>internationaux    | perte<br>profits<br>pertes<br>bnfices<br>profit<br>rentabilit<br>corp<br>aex<br>motorola<br>anhyp                         | mois<br>niveau<br>surtout<br>actuellement<br>base<br>mise<br>permettre<br>notamment<br>avant<br>permet       | aex<br>waterloo<br>corp<br>invest<br>motorola<br>cera<br>nikkei<br>abn<br>anhyp<br>warrant                                   | moi<br>mes<br>ma<br>rapport<br>m<br>convaincu<br>proposition<br>mon<br>demand<br>prfr  |
| 7   | 8   | 9   | 10   | 11   | 12   |
| dividende<br>tv<br>serveurs<br>tlphonique<br>telecom<br>supports<br>lecteurs<br>lectronique<br>multimedia<br>communications | avril<br>prvue<br>prvues<br>aex<br>dfinitif<br>charge<br>corp<br>invest<br>dfinitive<br>pb  | suivi<br>engag<br>manquer<br>livrer<br>lance<br>poursuit<br>charge<br>dtermine<br>gestes<br>afficher                      | fixes<br>aex<br>corp<br>electrabel<br>motorola<br>anhyp<br>cera<br>claes<br>warrant<br>abn                   | electrabel<br>claes<br>gib<br>cools<br>warrants<br>sambre<br>kredietbank<br>bbl<br>tractebel<br>cera                         | pb<br>corp<br>etienne<br>paribas<br>inc<br>invest<br>bell<br>nikkei<br>rupo<br>legrand   |
| 13  | 14  | 15  | 16   | 17   | 18   |
| electrabel<br>claes<br>gib<br>warrants<br>sambre<br>cera<br>motorola<br>kredietbank<br>anhyp<br>hanart                      | milliards<br>milliard<br>aex<br>chiffre<br>estimations<br>total<br>sek<br>grimp<br>corp<br>electrabel                                 | core<br>commerce<br>japonaises<br>commerciaux<br>amricain<br>commercial<br>usa<br>amricaine<br>asiatiques<br>exportations | banque<br>banques<br>crdit<br>emprunts<br>bancaires<br>bancaire<br>change<br>liquidits<br>cotation<br>cote   | ptrole<br>ptroliers<br>ptrolier<br>stock<br>lectrique<br>produite<br>lectriques<br>rendements<br>constructions<br>mtalliques | gestion<br>financement<br>oprations<br>financier<br>financiers<br>utilisant<br>effectues<br>concours<br>surveillance<br>office |
| 19  | 20  | 21  | 22   | 23   | 24   |
| vente<br>fabrication<br>vendre<br>acheter<br>industriels<br>vendus<br>achetant<br>acheteurs<br>vendu<br>industries          | entreprises<br>actionnaires<br>crdit<br>bancaires<br>cotation<br>acquisitions<br>actionnaire<br>capital<br>transaction<br>hypothcaire | recettes<br>taux<br>rduits<br>revenus<br>diminution<br>baisse<br>infrieur<br>diminuer<br>compenser<br>augmentations       | occidentale<br>situe<br>situ<br>village<br>roland<br>angleterre<br>cinaste<br>sud-est<br>allemandes<br>villa | mars<br>octobre<br>runion<br>semaines<br>tard<br>jours<br>dbut<br>dclaration<br>tenue<br>premiers                            | offrent<br>bas<br>offre<br>porte<br>offertes<br>facilite<br>accder<br>acquirit<br>lie<br>gamme                                 |

Table 4: 15 topics learned on BBCsports with Wasserstein NMF, the target words are a set of 6,000 words in French

| 1  | 2   | 3   | 4   | 5  |
|--|---|---|---|--|
| mm<br>commentaires<br>demand<br>minute<br>volontiers<br>m<br>propos<br>mettre<br>prsident<br>monsieur            | warrants<br>hanart<br>tessengerlo<br>claes<br>akzo<br>roland<br>anhyp<br>bekaert<br>gevaert<br>rupo       | jouer<br>acteurs<br>jou<br>quipe<br>acteur<br>gagner<br>interlocuteurs<br>jeu<br>puissante<br>cder                        | victoire<br>gagner<br>esprer<br>cder<br>impression<br>bataille<br>prtendre<br>fin<br>moiti<br>attend  | drogue<br>commis<br>tests<br>tribunal<br>interdit<br>preuves<br>test<br>comit<br>condamn<br>organisme                              |
| 6  | 7   | 8   | 9   | 10   |
| suivants<br>articles<br>partiellement<br>accepte<br>victoire<br>vingt<br>douze<br>gagner<br>amricain<br>champion | victime<br>morts<br>mort<br>dommages<br>condamn<br>victimes<br>provoqu<br>commis<br>perte<br>touch        | rduits<br>stock<br>diminuer<br>levs<br>montants<br>porte<br>homme<br>double<br>leve<br>jeunes                             | club<br>dclar<br>saison<br>fin<br>ajoute<br>acteur<br>angleterre<br>contrat<br>immobilier<br>football | angleterre<br>royaume-uni<br>espagne<br>france<br>grande-bretagne<br>britanniques<br>britannique<br>voyageurs<br>victoire<br>moiti |
| 11   | 12  | 13  | 14  | 15   |
| jours<br>semaines<br>semaine<br>samedi<br>vendredi<br>avril<br>mardi<br>week-end<br>octobre<br>heures            | mois<br>niveau<br>international<br>depuis<br>prsent<br>pays<br>crer<br>ncessit<br>avant<br>internationaux | nationales<br>national<br>nationale<br>dveloppement<br>concurrence<br>nationaux<br>priorit<br>plans<br>march<br>dvelopper | sud<br>core<br>mexique<br>canada<br>afrique<br>test<br>asiatiques<br>asie<br>jours<br>amrique         | jeux<br>football<br>championnat<br>sportif<br>olympique<br>sport<br>sportifs<br>joueurs<br>coupe<br>avril                          |

Table 5: 24 topics learned on Reuters with Euclidean NMF

|            |            |           |         |          |             |
|------------|------------|-----------|---------|----------|-------------|
| 1          | 2          | 3         | 4       | 5        | 6           |
| mln        | net        | loss      | cts     | cts      | dlrs        |
| shrs       | cts        | cts       | div     | net      | loss        |
| avg        | shr        | shr       | qtly    | sales    | mln         |
| revs       | dlrs       | year      | record  | shr      | mths        |
| shares     | mln        | revs      | pay     | reuter   | oil         |
| cts        | year       | qtr       | prior   | mln      | split       |
| mths       | revs       | corp      | reuter  | qtr      | stock       |
| stg        | reuter     | shares    | march   | mths     | net         |
| net        | qtr        | trade     | april   | st       | billion     |
| corp       | oper       | group     | payout  | note     | fed         |
| 7          | 8          | 9         | 10      | 11       | 12          |
| profit     | dlrs       | cts       | loss    | billion  | pct         |
| loss       | mln        | qtr       | shrs    | bank     | rate        |
| year       | share      | shr       | avg     | corp     | group       |
| stg        | loss       | net       | oper    | assets   | days        |
| pretax     | year       | revs      | net     | deposits | stake       |
| sales      | group      | share     | st      | loans    | prime       |
| tax        | note       | mths      | reuter  | yen      | shares      |
| billion    | gain       | st        | bank    | group    | discount    |
| nil        | oper       | profit    | profit  | fed      | rates       |
| includes   | includes   | company   | sales   | dividend | tender      |
| 13         | 14         | 15        | 16      | 17       | 18          |
| dividend   | corp       | net       | cts     | company  | company     |
| stock      | dlrs       | revs      | dlrs    | shares   | reuter      |
| april      | shares     | bank      | shr     | stock    | corp        |
| split      | qtr        | rate      | crude   | march    | oil         |
| cts        | stake      | reuter    | avg     | common   | unit        |
| record     | st         | shr       | shrs    | reuter   | acquisition |
| payable    | exchange   | prime     | raises  | split    | terms       |
| corp       | rate       | year      | oil     | group    | sell        |
| share      | american   | effective | bbl     | billion  | sale        |
| declared   | investment | note      | wti     | trade    | acquire     |
| 19         | 20         | 21        | 22      | 23       | 24          |
| stg        | april      | oper      | st      | quarter  | offer       |
| bank       | sets       | cts       | trade   | year     | acquisition |
| mln        | fed        | march     | oil     | earnings | tender      |
| dlrs       | shares     | fed       | mln     | bank     | fed         |
| market     | year       | mln       | qtr     | expects  | dlr         |
| money      | reuter     | rate      | crude   | sees     | stock       |
| england    | net        | sets      | deficit | share    | rejects     |
| company    | repurchase | excludes  | stg     | japan    | shares      |
| today      | dlr        | tax       | raises  | american | split       |
| assistance | cts        | share     | bbl     | rates    | buy         |

Table 6: 15 topics learned on BBCSports with Euclidean NMF

| 1  | 2   | 3  | 4  | 5  |
|--|---|--|--|--|
| win<br>6<br>play<br>set<br>final<br>good<br>match<br>7<br>title<br>cup                             | england<br>test<br>team<br>series<br>cricket<br>game<br>pakistan<br>tour<br>australia<br>india          | year<br>world<br>chelsea<br>england<br>indoor<br>wales<br>mourinho<br>time<br>european<br>record | 6<br>year<br>england<br>world<br>3<br>7<br>champion<br>time<br>game<br>open                    | league<br>cup<br>club<br>real<br>arsenal<br>united<br>chelsea<br>win<br>england<br>season        |
| 6  | 7   | 8  | 9  | 10   |
| 6<br>7<br>seed<br>4<br>club<br>3<br>chelsea<br>england<br>2<br>agassi                              | england<br>players<br>injury<br>back<br>half<br>wales<br>ireland<br>jones<br>nations<br>game            | half<br>ball<br>goal<br>liverpool<br>side<br>time<br>benitez<br>minutes<br>ireland<br>game       | united<br>kenteris<br>greek<br>thanou<br>iaaf<br>club<br>drugs<br>liverpool<br>athens<br>tests | open<br>club<br>australian<br>year<br>capriati<br>play<br>arsenal<br>football<br>injury<br>women |
| 11   | 12  | 13   | 14   | 15   |
| holmes<br>country<br>athletes<br>cross<br>season<br>drugs<br>indoor<br>athletics<br>race<br>doping | williams<br>injury<br>australian<br>women<br>final<br>set<br>scotland<br>australia<br>open<br>davenport | match<br>rugby<br>chelsea<br>fa<br>united<br>goal<br>referee<br>final<br>mutu<br>lions           | sri<br>cricket<br>day<br>lanka<br>match<br>play<br>played<br>tour<br>team<br>jayasuriya        | souness<br>cup<br>johnson<br>play<br>year<br>newcastle<br>move<br>uefa<br>bellamy<br>goal        |

Table 7: 24 topics learned on Reuters with KL NMF

| 1   | 2   | 3  | 4   | 5  | 6  |
|---|---|--|---|--|--|
| acquire<br>agreement<br>reuter<br>buy<br>agreed<br>corp<br>company<br>acquisition<br>insurance<br>terms     | note<br>oper<br>year<br>includes<br>net<br>gain<br>tax<br>share<br>dlr<br>operations                  | offer<br>bid<br>tender<br>company<br>share<br>merger<br>corp<br>takeover<br>management<br>acquisition    | billion<br>profit<br>stg<br>assets<br>january<br>francs<br>year<br>pretax<br>plc<br>general     | trade<br>japan<br>japanese<br>foreign<br>reuter<br>house<br>states<br>government<br>told<br>reagan | stock<br>split<br>shares<br>common<br>share<br>shareholders<br>company<br>pct<br>reuter<br>merger    |
| 7   | 8   | 9  | 10  | 11   | 12   |
| note<br>ended<br>months<br>company<br>full<br>march<br>end<br>period<br>reuter<br>public                    | bank<br>pct<br>rate<br>national<br>stg<br>prime<br>market<br>reuter<br>money<br>today                 | pct<br>federal<br>savings<br>loan<br>fed<br>days<br>bank<br>banks<br>interest<br>rate                    | group<br>financial<br>united<br>reuter<br>west<br>crowns<br>allied<br>owned<br>company<br>foods | loss<br>profit<br>cts<br>shr<br>net<br>revs<br>reuter<br>qtr<br>year<br>ct                         | dividend<br>payable<br>record<br>reuter<br>april<br>trust<br>declared<br>cts<br>share<br>march       |
| 13  | 14  | 15   | 16  | 17   | 18   |
| stores<br>union<br>reuter<br>air<br>western<br>taft<br>port<br>warner<br>video<br>borg                      | pct<br>shares<br>stake<br>group<br>investment<br>securities<br>common<br>stock<br>exchange<br>reuter  | mln<br>net<br>avg<br>shrs<br>sales<br>year<br>reuter<br>industries<br>enterprises<br>international       | oil<br>crude<br>prices<br>gas<br>canada<br>energy<br>canadian<br>petroleum<br>barrel<br>bbl     | year<br>quarter<br>earnings<br>company<br>expects<br>share<br>reported<br>pct<br>sees<br>results   | exchange<br>dollar<br>rates<br>baker<br>treasury<br>rate<br>currency<br>interest<br>japan<br>finance |
| 19  | 20  | 21   | 22  | 23   | 24   |
| company<br>pct<br>president<br>reuter<br>executive<br>chairman<br>chief<br>international<br>venture<br>corp | corp<br>reuter<br>sale<br>acquisition<br>unit<br>products<br>completed<br>company<br>sell<br>business | american<br>resources<br>service<br>pacific<br>corp<br>america<br>southern<br>power<br>chrysler<br>north | dlrs<br>share<br>corp<br>cash<br>harper<br>offers<br>electric<br>row<br>ford<br>cos             | cts<br>april<br>record<br>div<br>pay<br>reuter<br>qtrly<br>prior<br>march<br>sets                  | cts<br>net<br>shr<br>qtr<br>revs<br>reuter<br>st<br>mths<br>corp<br>jan                              |



Table 8: 15 topics learned on BBCsports with KL NMF

| 1             | 2             | 3         | 4          | 5             |
|---------------|---------------|-----------|------------|---------------|
| 6             | wales         | liverpool | open       | chelsea       |
| 7             | england       | win       | year       | arsenal       |
| 3             | nations       | league    | australian | mourinho      |
| set           | game          | cup       | women      | league        |
| seed          | injury        | play      | world      | game          |
| 4             | france        | benitez   | cup        | football      |
| final         | rugby         | gerrard   | tennis     | players       |
| match         | scotland      | club      | injury     | team          |
| win           | coach         | newcastle | number     | boss          |
| 5             | robinson      | united    | grand      | wenger        |
| 6             | 7             | 8         | 9          | 10            |
| england       | minutes       | drugs     | lions      | pakistan      |
| south         | half          | iaaf      | rugby      | india         |
| africa        | goal          | kenteris  | england    | test          |
| test          | kick          | doping    | tour       | series        |
| day           | ball          | greek     | year       | australia     |
| series        | lead          | thanou    | zealand    | day           |
| vaughan       | time          | athens    | back       | cricket       |
| cricket       | penalty       | olympics  | world      | team          |
| play          | shot          | test      | cup        | tour          |
| tour          | arsenal       | athletes  | match      | khan          |
| 11            | 12            | 13        | 14         | 15            |
| world         | game          | england   | club       | cricket       |
| year          | players       | game      | real       | icc           |
| race          | united        | ireland   | year       | players       |
| olympic       | match         | half      | contract   | international |
| indoor        | time          | goal      | move       | sri           |
| european      | spain         | side      | player     | ireland       |
| champion      | england       | ball      | football   | club          |
| record        | international | france    | told       | day           |
| championships | vieira        | line      | madrid     | deal          |
| holmes        | coach         | back      | team       | england       |