Fast Dictionary Learning with a Smoothed Wasserstein Loss
Supplementary Material

1 Proof of Theorem 3

Theorem. *The solution Λ^* of*

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

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

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

Proof. Similarily to the proof of Theorem 1, the dual is

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

Which is the same optimisation problem as Problem (2). The primal-dual relationship is derived from Equation (3). Note that since the derivative of $x \to x \log x$ goes to infinity at 0, Λ^* is in the interior of $\mathbb{R}_+^{k \times m}$. So the gradient of the objective with respect to Λ is orthogonal to the simplex on Λ^*, *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	offer	million	split	net	rate
increase	bid	income	board	saint	interest
result	tender	tax	stock	note	market
share	plan	note	record	company	bank
report	request	cash	approve	profit	money
expect	approve	pay	annual	million	rise
due	board	total	declare	share	today
rise	make	debt	company	cheerio	forecast
cost	receive	net	set	refresh	percent
significant	seek	amount	april	situate	dollar

7	8	9	10	11	12
trade	purchase	oil	loss	stock	million
japan	deal	gas	net	share	net
states	sell	price	profit	common	cheerio
export	sale	barrel	wagon	stake	profit
foreign	buy	energy	year	exchange	news
minister	cash	increase	million	commission	telecom
policy	subject	day	cheerio	total	refresh
market	stake	supply	note	company	plonk
economy	plan	raise	suck	percent	litre
deal	amount	coast	refresh	price	cock

13	14	15	16	17	18
include	group	net	technology	chairman	bank
tax	unit	refresh	international	comment	loan
gain	stake	cheerio	system	news	credit
sale	company	suck	service	week	debt
current	sell	plonk	provide	official	mortgage
fund	telecom	situate	data	west	fund
income	part	engine	make	interview	state
result	news	news	telecom	today	public
million	plan	cock	research	monday	money
such	would	reproduce	news	part	part

19	20	21	22	23	24
pay	april	year	wagon	company	profit
record	february	million	cheerio	business	company
set	june	end	profit	firms	telecom
wagon	october	last	litre	industry	percent
april	news	month	net	firm	retail
previous	july	period	news	telecom	pence
bracket	december	ago	plonk	product	net
before	january	week	cock	stake	rise
curl	november	after	brake	news	tumble
news	september	time	lorry	same	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	wales	year	scotland	knee
game	rugby	april	france	back
team	captain	time	irish	return
cup	bath	july	italy	left
world	nick	june	game	week
good	sir	january	coach	fit
top	rob	season	side	problem
great	centre	career	match	hope
coach	jack	october	paris	time
important	brown	make	win	play

6	7	8	9	10
league	make	club	win	nought
club	people	deal	good	ball
boss	interview	coach	victory	captain
ham	television	contract	lose	total
cup	criticism	move	feel	run
states	lot	offer	happy	butcher
side	thing	future	hard	percent
match	talk	summer	beat	catch
draw	money	chairman	final	mark
coach	referee	chief	game	bat

11	12	13	14	15
goal	seed	cricket	champion	test
minute	tennis	test	world	drug
kick	open	india	race	ban
ball	set	australia	gold	guilty
half	april	africa	event	decision
penalty	match	south	record	sport
referee	beat	zealand	britain	court
lead	champion	match	london	evidence
score	july	series	compete	committee
side	round	tour	title	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	marché	perte	mois	aex	moi
aex	économique	profits	niveau	waterloo	mes
comprend	nationales	pertes	surtout	corp	ma
corp	économiques	bénéfices	actuellement	invest	rapport
conclut	développement	profit	base	motorola	m
electrabel	international	rentabilité	mise	cera	convaincu
claes	national	corp	permettre	nikkei	proposition
motorola	nécessité	aex	notamment	abn	mon
sambre	niveau	motorola	avant	anhyp	demandé
cools	internationaux	anhyp	permet	warrant	préféré
7	**8**	**9**	**10**	**11**	**12**
dividende	avril	suivi	fixes	electrabel	pb
tv	prévue	engagé	aex	claes	corp
serveurs	prévues	manquer	corp	gib	etienne
téléphonique	aex	livrer	electrabel	cools	paribas
telecom	définitif	lancée	motorola	warrants	inc
supports	charge	poursuit	anhyp	sambre	invest
lecteurs	corp	chargée	cera	kredietbank	bell
électronique	invest	déterminée	claes	bbl	nikkei
multimédia	définitive	gestes	warrant	tractebel	rupo
communications	pb	afficher	abn	cera	legrand
13	**14**	**15**	**16**	**17**	**18**
electrabel	milliards	corée	banque	pétrole	gestion
claes	milliard	commerce	banques	pétroliers	financement
gib	aex	japonaises	crédit	pétrolier	opérations
warrants	chiffre	commerciaux	emprunts	stock	financier
sambre	estimations	américain	bancaires	électrique	financiers
cera	total	commercial	bancaire	produite	utilisant
motorola	sek	usa	change	électriques	effectuées
kredietbank	grimpé	américaine	liquidités	rendements	concours
anhyp	corp	asiatiques	cotation	constructions	surveillance
hanart	electrabel	exportations	cote	métalliques	office
19	**20**	**21**	**22**	**23**	**24**
vente	entreprises	recettes	occidentale	mars	offrent
fabrication	actionnaires	taux	située	octobre	basé
vendre	crédit	réduits	situé	réunion	offre
acheter	bancaires	revenus	village	semaines	portée
industriels	cotation	diminution	roland	tard	offertes
vendus	acquisitions	baisse	angleterre	jours	facilite
achetant	actionnaire	inférieur	cinéaste	début	accéder
acheteurs	capital	diminuer	sud-est	déclaration	acquérir
vendu	transaction	compenser	allemandes	tenue	liée
industries	hypothécaire	augmentations	villa	premiers	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	warrants	jouer	victoire	drogue
commentaires	hanart	acteurs	gagner	commis
demandé	tessenderlo	joué	espérer	tests
minute	claes	équipe	céder	tribunal
volontiers	akzo	acteur	impression	interdit
m	roland	gagner	bataille	preuves
propos	anhyp	interlocuteurs	prétendre	test
émettre	bekaert	jeu	fin	comité
président	gevaert	puissante	moitié	condamné
monsieur	rupo	céder	attend	organisme

6	7	8	9	10
suivants	victime	réduits	club	angleterre
articles	morts	stock	déclaré	royaume-uni
partiellement	mort	diminuer	saison	espagne
accepte	dommages	élevés	fin	france
victoire	condamné	montants	ajoutée	grande-bretagne
vingt	victimes	portée	acteur	britanniques
douze	provoqué	homme	angleterre	britannique
gagner	commis	double	contrat	voyageurs
américain	perte	élevée	immobilier	victoire
champion	touché	jeunes	football	moitié

11	12	13	14	15
jours	mois	nationales	sud	jeux
semaines	niveau	national	corée	football
semaine	international	nationale	mexique	championnat
samedi	depuis	développement	canada	sportif
vendredi	présent	concurrence	afrique	olympique
avril	pays	nationaux	test	sport
mardi	créer	priorité	asiatiques	sportifs
week-end	nécessité	plans	asie	joueurs
octobre	avant	marché	jours	coupe
heures	internationaux	développer	amérique	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	england	year	6	league
6	test	world	year	cup
play	team	chelsea	england	club
set	series	england	world	real
final	cricket	indoor	3	arsenal
good	game	wales	7	united
match	pakistan	mourinho	champion	chelsea
7	tour	time	time	win
title	australia	european	game	england
cup	india	record	open	season
6	7	8	9	10
6	england	half	united	open
7	players	ball	kenteris	club
seed	injury	goal	greek	australian
4	back	liverpool	thanou	year
club	half	side	iaaf	capriati
3	wales	time	club	play
chelsea	ireland	benitez	drugs	arsenal
england	jones	minutes	liverpool	football
2	nations	ireland	athens	injury
agassi	game	game	tests	women
11	12	13	14	15
holmes	williams	match	sri	souness
country	injury	rugby	cricket	cup
athletes	australian	chelsea	day	johnson
cross	women	fa	lanka	play
season	final	united	match	year
drugs	set	goal	play	newcastle
indoor	scotland	referee	played	move
athletics	australia	final	tour	uefa
race	open	mutu	team	bellamy
doping	davenport	lions	jayasuriya	goal

Table 7: 24 topics learned on Reuters with KL NMF

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