

## A Illustrative empirical results

We applied a variant of Algorithm 2 to the UCI “Bag of Words” dataset comprised of New York Times articles. This data set has 300000 articles and a vocabulary of size  $d = 102660$ ; we set  $k = 50$  and  $\alpha_0 = 0$ . Following [37], instead of using a single random  $\theta$  and obtaining singular vectors of  $\hat{W}^\top \text{Triples}_{\alpha_0}(\hat{W}\theta)\hat{W}$ , we used the following power iteration to obtain the singular vectors  $\{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_k\}$ :

$\{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_k\} \leftarrow$  random orthonormal basis for  $\mathbb{R}^k$ .  
Repeat:  
1. For  $i = 1, 2, \dots, k$ :  

$$\hat{v}_i \leftarrow \hat{W}^\top \text{Triples}_{\alpha_0}(\hat{W}\hat{v}_i)\hat{W}\hat{v}_i.$$
  
2. Orthonormalize  $\{\hat{v}_1, \hat{v}_2, \dots, \hat{v}_k\}$ .

The top 25 words (ordered by estimated conditional probability value) from each topic are shown below.

zzz_held send advisory publication released guard zzz_attn_editor undatedlined night advance zzz_andrew_pollack zzz_douglas_frantz billion zzz_jennifer zzz_dirk_johnson zzz_leslie cell zzz_linda games zzz_lee zzz_james_brooke zzz_winnipeg deal husband zzz usc	premature guard zzz_held released publication advisory send undatedlined zzz_washington_datelined zzz_istanbul zzz_attn_editor zzz_seth_mydan nyt zzz_johannesburg zzz_afghanistan zzz_jane_perlez zzz_john_broder zzz_warren zzz_melbourne zzz_lexington zzz_erik_eckholm zzz_bernard_simon substitute close point	las como los zzz_latin_trade articulo telefono transmiten fax una del articulos espanol paises sobre financial zzz_america_latina notas prohibitivo con revista tiene economia costo otros zzz_paris	sales economic consumer major home indicator weekly order claim scheduled listed dates jobless prices price market leading retailer economy index retail spending product cost producer	million shares public offering source initial debt bond billion share quarter revenue market zzz_calif school zzz_new_york cash stock percent securities zzz_credit_suisse_first_boston deal contract president expected	com question information zzz_eastern sport daily commentary business newspaper separate spot marked today zzz_tom_oder holiday need staffed development toder client eta directed additional reach washington	run inning hit game season home right games zzz_dodger left team start yankees pitcher ball pitch manager lead night homer field play ranger win hitter	women team woman job sport cancer look company group percent girl study game games female american number season breast play zzz_taliban right part male high
---	---	--	---	--	---	---	---

drug patient million company doctor companies percent cost program health care billion plan medical treatment zzz.aid disease cancer hospital prescription federal government product zzz.medicare study	player zzz.tiger_wood won shot play round win tournament tour right par final playing major ball hit lead golf guy hole course game played night set	article zzz_new-york misstated zzz_boston_globe zzz_united_states company president campaign zzz_clinton surname player incorrectly point film director office school home misspelled died information misidentified referred zzz_washington son	palestinian zzz_israel zzz_israeli zzz_yasser_arafat peace israeli israelis leader official attack zzz_bush zzz_west_bank zzz_palestinian violence security killed talk military jewish zzz_jerusalem soldier zzz_clinton zzz_sharon minister fire	tax cut percent zzz_bush billion plan bill taxes million zzz_congress zzz_george_bush economy money income government spending federal pay republican zzz_white_house zzz_senate zzz_democrat sales zzz_social_security proposal	cup minutes oil water add tablespoon food teaspoon pepper sugar large fat butter sauce serving hour fresh pan taste bowl cream onion serve medium pound	point game team shot play zzz_laker season half lead games quarter minutes night left goal king final played scored zzz_kobe_bryant rebound right win percent ball	yard game play season team touchdown quarterback coach defense quarter ball field pass run offense line running defensive zzz_nfl football receiver left win player zzz_giant	percent stock market fund investor companies analyst money investment economy point company quarter price billion earning prices firm index growth zzz_nasdaq shares rates rate interest
--	--	--	--	--	---	--	---	--

zzz_al_gore campaign president zzz_george_bush zzz_bush zzz_clinton vice presidential million democratic night voter election vote plan zzz_bill_bradley ballot zzz_governor_bush republican zzz_florida right votes poll court candidates	zzz_george_bush president campaign republican zzz_john_mccain election zzz_texas presidential political zzz_enron governor administration democratic zzz_white_house voter nation public zzz_clinton republican candidate point question percent zzz_party	car race driver team won win racing track season lap point sport seat races road run look right zzz_nascar drive zzz_winston_cup owner start big ago	book children ages author read newspaper web writer written sales find history list word published school zzz_new_york right boy writing american reading game reader won	zzz_taliban attack zzz_afghanistan official military zzz_u_s zzz_united_states terrorist war bin laden zzz_american zzz_bush government group forces zzz_pakistan country leader american afghan troop terrorism nation zzz_pentagon	com www site web sites information online mail internet telegram visit find zzz_internet computer org newspaper offer free services company official list user companies customer	zzz_bush percent campaign zzz_enron administration president zzz_white_house money plan republican company million zzz_republican official zzz_texas election show political zzz_mccain energy zzz_washington zzz_united_states voter fund zzz_al_gore	court case law lawyer federal government decision trial zzz_microsoft right judge legal ruling attorney death system company zzz_supreme_court election cases prosecutor public zzz_florida ballot states	percent number group rate million sales survey according study quarter average economy american increase rose black student level school season poll newspaper job consumer government
--	---	--	---	--	---	--	---	--

company	show	game	computer	film	team	bill	cell	election
percent	network	games	system	movie	player	zzz_senate	patent	ballot
million	season	season	program	director	season	law	human	vote
business	zzz_nbc	play	zzz_microsoft	play	game	right	research	voter
companies	zzz_cb	goal	mail	character	coach	zzz.white.house	group	campaign
billion	program	king	software	actor	play	zzz.congress	scientist	political
analyst	television	team	window	show	games	vote	zzz_enron	votes
stock	series	won	web	movies	right	member	study	official
quarter	night	player	company	million	league	president	disease	zzz.florida
executive	zzz_new.york	coach	million	part	million	legislation	information	democratic
deal	zzz_abc	played	information	zzz.hollywood	deal	zzz.clinton	found	race
sales	tonight	period	need	look	manager	group	team	zzz_republican
share	hour	left	technology	big	need	zzz.house	public	recount
zzz_enron	look	playing	user	young	contract	republican	doctor	republican
chief	zzz.fox	night	security	music	guy	campaign	government	won
market	air	win	zzz.internet	set	point	federal	death	leader
employees	viewer	right	problem	screen	played	money	cancer	candidate
customer	rating	com	internet	writer	baseball	election	researcher	zzz.al.gore
president	game	playoff	money	television	agent	support	stem	zzz-party
product	early	power	home	making	fan	zzz_republican	official	poll
executives	big	guy	network	love	playing	measure	problem	candidates
financial	talk	zzz_new.york	product	played	job	issue	called	party
earning	event	record	called	producer	free	passed	medical	presidential
operation	hit	shot	help	guy	sport	percent	director	win
cent	award	minutes	number	kind	basketball	billion	question	result

money million fund zzz_enron campaign program group plan government firm company pay worker help job political lawyer member account effort billion employees financial question need	police officer official president government attack case told office member public death group zzz_new_york chief black lawyer prosecutor security building campaign night hour home found	team game win won zzz_u_s play games official point run home zzz_united_states sport zzz_new_york attack tournament american percent minutes zzz_olympic final player company lead zzz_washington	air water million high building power plant plan cost hour system wind part weather area home rain shower front program billion night feet low miles	family children home father mother son parent child friend school boy wife house told daughter kid night help care left official room money hour job	music song group part zzz_new_york company million band show album companies record play right business look artist home industry member black sound night called fan	official government zzz_united_states zzz_china zzz_u_s zzz_american country administration zzz_clinton million nation countries president economic foreign power chinese zzz_russia political plan meeting leader trade percent right	companies job worker company business firm zzz_new_york attack president employees plan need law percent customer industry number cost terrorist security market information help official economy	president program zzz_bush group game member zzz_clinton care leader health zzz_white_house vice plan job children patient executive worker doctor school decision director zzz_congress administration chief
---	--	---	--	--	---	--	--	---

government companies political country president campaign leader business election zzz.bush win war company zzz.internet billion race power support market team democratic won public web industry	season team won race win attack home record games zzz.u.s final zzz.clinton night million zzz.olympic winning coach championship patient playoff victory american trial medal series	right zzz.united-states american war student look need show home question black military left country com women word put zzz.american help room zzz.u.s zzz.america percent job	test zzz.seattle_post_intelligencer zzz.hearst_news_service zzz.kansas.city look testing houston ellipses anthrax student glories mark night rare zzz.texas result risk exam system scores missile zzz.washington body according scientist	file onlytest sport notebook zzz.los_angeles onlyendpar zzz.joe_haakenson_san_gabriel_valley_tribune zzz.anaheim_angel frontend zzz.seattle_pi zzz.seattle_post_intelligencer zzz.chuck zzz.abcdcfg_test added zzz.los_angeles_dodger read zzz.calif output email internet zzz.brian_dohn files zzz.scott_wolf wrote consumer
--	--	---	--	---