Off-the-grid Blind Deconvolution and Demixing

Sasaced Razublikia, Ssajad Dacit, Mikael Skoglund, Gabor Fodoof, Carlolf Echionelet SchSchool Die Elektric Engineering rand Computer Science, KTH Royal Institute to for Choblogy g \$100 kholing Sweden Ericsson Research, Sweden

EEmail: { sraz, ssajado, skoglund, gaborff carlob | @kltlsese

Abstract+-We'e consider: thee problem of GB2D! (GB2D!) in scenarios where multiple users communicate messages through multiple unknownchannels|sandnal single BS! EBS!) Rollects libets contributionisu (This. sEenarioe arises ainsevarious recommunication fields a field ding windless communications at the Unternet of Things, oveT-thegair computation, and integrate discussing gardedomining gations...Imthis:setup, eachthiser/stmessagehisuconvolvedswith is multi-path channel formed by several scaled and delayed copies of Didac spikes of the BS Dieceives in dinear confidention of the convolved signals, and the goal is to recover the unknown amplitudes. continuotiseindekedydelays.pänddtransmittedowaveforms from/a compressed i vector of of casuréments at the BStd However, fin the absence tof anythor Rft! kHowledge of the transmitted arressages kndyddagnels, tGB2Dlaisaltighl yadsallgagi ngdarldaintrlactabli@lid generall/To address this issue, we assume that each user's message föllowissae distinctsmodulation schemeslivingninsagknövlovleve distensional subspaces: Bynexploiting these subspace-dissumptions ands the explansity; of ithe grounding at his channels sfor different rusers; syntriansform the nonlinear GB2Dhoroblem liftto an matrix tuple recovéryop roblem ofromo a (feW2 Dhearo bleasurem entso a Toi achieve this, vwey propose nafsemidefinite in programming noptimization that exploits the pspecific rhow differsional ratricture to fit he iomatrix tuple for recoveretificmlessages and continuous delays of different conjuntaries tion: paths drong a single received signal at diffe BSIt. Finally, nour ion medical respectingly rehowed that not not proposed Methiod effectivelyerrecovers call a transmitted timessages pando the continuous[cdelay] parametersalif the schittnells nwith gas sufficient numberoof samples arameters of the channels with a sufficient nuIndex Terms+pletomic norm minimization, blind channel estimaltiony, blind data Accovery, blind deconvolution, blind derhixing. estimation, blind data recovery, blind deconvolution, blind demixing.

I. Introduction

I. Introduction

In the near future, the Internet of Things (IoT) is expected in the near future, the Internet of Things (IoT) is to connect billions of wireless devices, surpassing the capacity expected to connect billions of wireless devices, surpassing of the current fitth-generation (5G) wireless system both the capacity of the current fitth-generation (5G) wireless technically and economically. One of the primary challenges system both technically and economically. One of the that of, the future wireless communication system, will face primary challenges that 6G, the future wireless communication system, will face is managing the massive number of IoT devices that generate cation system. Will face is managing the massive number of IoT devices that generate sporadic traffic. As the will significantly increase, and it is generally agreed among 6G market grows, this sporadic traffic along the massive number of market grows, this sporadic traffic access increase, and it is generally agreed among communications engineers that the current 5G channel access increase, and it is generally agreed among communications procedures cannot handle this volume of traffic.

Information and communication theory, require a large number

cannot handle this volume of traffic information and communication theory, require a large number

Saeed Razavikia and Carlo Fischione acknowledge the support of WSaSEd SSEaSiAliCant/Caffe, Hischfüne/Seknawjedge abel stippitet off/WASE, SSFaSACCOM avil, Elik Fel ASH ghroject ward a Divital features part by Digital Fußajnels Daßi gittel MikuteluSkogflund over ePstippforteel up pant byl DigitalvFuthrest Digital Fatdres Project PERCy supported the work of Gabor Fodor.

of Dilotstion training signals cosestimate dhe channel, reducing tdassignificafitrmestource no asten that i does not be scale retowards In Foreguinements, pThus, on injuicing the overhead passed the ekehanging-actitatintotypgs ifofatraining uinformation, hatchoes obárnolles timaribal offidredatáres lotatas sighmentai ás mízicos abye This his adspecially britical horgeommunications cave dynamic channelstisuch aschrillsmehrnwalzestiterahertz awherbachanhel coherence; tirises ares short. Talus the schanally state call formation changestrapidly. In these cases; the assumption of block fading nowlenger the idserOne suppreach to raddressing this issues is to shootporate chandelraging teffects fortoathorchangles timution brochsstochasintize spectrabelficieficlyl(sde, fedin §? j); but this headire.Oknowledge of the allahnesicorrelation structure at different tinhesowhielg imightf betschallengling doaobtel neitii gentiral channel environmentse Thereford, foliosituation sewhere; a large humbéri of devices transmites mall cambunts loft data sporadically overedynamic diffamels, tand-showliahnelightrdiationaltengturg ts) urktainwin, iteisecalcilalatorelveid itransmitting l'aleighale with sitichtlonger/byerheald/igformation than actual dataisThis saisels the question of twhether dhis live feasible manic channels, and theTo:hfaciletateorexplaination; uwterreoiss idekna vscentario cwhead multiple users stransmit meigagesythhough chultipler frequencyd sefectiveticharthels towards datcentralisBSIs(as tlescribed im [3f Edet[9]), this is BS a siece ives a combined signal comprising confloibutions afrom early basers in which vis them sprocessed sthrough a/sensingufiltele (seerFig:a22)mThe coages throughtanettisly estimatecythe lecansmitted needs agos and a channels of the received chreas(if erflents) 9t).thehBSRS which i vies a achallehging signishear problems contributions from all users, which is the the constitution of the control Tilgeoveralhesato computation (3); (3)ti super-resolutions singlemolecules imaging a [7] [c] [fr[7], [2], [7] einedti-uses weitipath chathrelBestimation [3]a [2]hablinginglibration in multi-channel saffibling csystems appeals, irandom access [2] palidaintegrated (tadar)gsensinglandicommunications[[?],[[?]]stpler-resolution single-molecule imaging [?], [?], [?], [?], multi-user A. Related work in ultipath charmel estimation [?], [?], blind calibration in mull hie chrobiteths of indexogeringe mes@lge@] and inchannels do stille modeln described (above) falls sinto and class not obtain do dec (rivolution techniques used to solve inverse problems. These techniques have made notable progress in addressing blind deconvolution problems, with a focus on sparse signals consisting of cobking lef user of thing | m (2 kagek and Thenconvein tionalimothodeinvolves assureing Hatathe continuous of hatinel parameterstilien onchrippede fineed domainvo fingridse whicherean

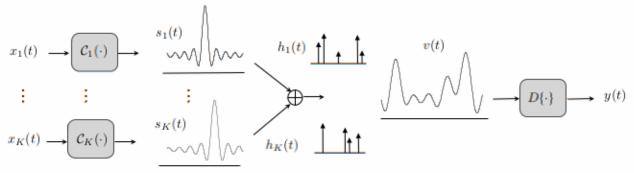


Fig. 1:: An illustration of old treather heiral timed of of B219 B25N cm blvery. Except ansmitt a wave fittings of the non-significant heiral timed by (tP_i) in which heir words the sum of the su

Ethestintatednusing well-vestablished-onethods sugheas &n midilnization.bHowever.otheobreidefinedolghidssmaythota accurately. snatch the true scontinuous andex wakies of paraméters, fleading [18] bášis This matche issues. I that the degrade sthe sperformable is off-chlindtdeconvolutiona/Foraddresst-theseicissuesa recente/work has focused oniciontinuous parameter estimation, within bliefid deconvolution metabol developed in fifth to receiver, obtainmens parameters in the ciase in a single cuserately match the true codn ithis workd we tackle achallenging and more generalized modelathat issuelyesha tmixtureleef: blindtldeconvolution problouns! with transmitted Tsignals desing especial as with edifferent bodebooks-dnd arbitráryuzhannels-thattdo not ifolitiwra spedifia bhanhellemodeil uFurthenetwodcolasidlemethe i avail a bil i tyecofea general reasing afriteriat ithehBS a shatfallows lusuter observe compressed time are or abilitions of data; samples insteach of the izhdle sadopleka as in dl? lesto, ensute pradtichity dodogenerality. Itrisbdetesyorthylithataounptopos edgmelhodeisigle terministic vand is findependente of obither the whatmey on message (distribution), follow setspitcifparthfromelexisting, statistical techniquese(stub) assaippbilkimaté megsageabassingirfe Ifilhat can tordy B&orkhon specific udistributions re Finally reweighting at remarkable onesulf thatawhen pills userseade of the learnet obdebook lee, transmit? their messagescontroploposeddogtineizalion problem tisvindependent ofithertotalseulmbethofluserketerministic and is independent either the channel or message distribution, which sets apart from existing statistical techniques (such as approx propose message bashimization haramework named GBSPEcithat deverages specific features of the channels and transmitted signals. Specifically, each user's channel has few result that when all users use the same codebook to transdominant scattering paths pand each user iemploys a distinct channel coding scheme GB2D! utilizes a lifting technique to convert the demixing of nonlinear problems into highdimensional matrices containing continuous channel parametelye whichestransforms, multiplea features roewhek, channels and 2 transtricted evigages intecificing to uspecific theaten a rofela matrix-auple itving signabig berediffied sionea Altra etable honyest bptifization problemattepinggseth to anclover the continuous chdisteh grachmeters doylipromoting ethe B@@lfictifeature bftithe trealrixi quplet, of obligwed that lae least-squages of robleshinten estimate theistrantschitteld dinessages all matddition on this niwerkom fewides conditions and act which the solution to GB2D is finiture and obtimal rands simulation results deinon strate of a effectiveness fin

recover the continuous channel parameters by proposed to recover the continuous channel parameters by promoting the specific feature of the matrix tuple, followed by a least-squared limited message recovery and channel estimation with additing their winder and sensing intermediately included by a least-squared limited message recovery and channel estimation with additing their winder and sensing intermediately included the continuous spending mulation results demonstrate its demonstrate of the channel communication frame channels by many provides a general communication frame channels by many provides a general communication frame channels by many provides a supply of the channels code book, and sufficients.

Specifically buckoptailutioniniaanoninymarized as follows Independent of the number of users: In a case wherein

- Blinders esseighersamer codebookaton transmitation with sage sinther proposed optimization problem B2 independent of the testal gramber of assers without spending train-
- Tractable complexity: www.sproposer.al tractable iconvex optimization problem do use over pontinuous schannels parameters by phonloting the lapecial cleature of the ich delicature of th
- Optimality condition: We provide a theoretical guaran-
- Icoldpandlentsofution ton GB2D his ranfque cand veptimal undersome sminimum reparation conditions condition the tipathghantel dplaysosed optimization problem is in-
- Messagenand (channel) distribution nares, arbitrary: In
- Contrastitocome ktatistic Wechannelses timation lenethods;
 ogticovariance based methods and approximate message
 passing tGB2D! doos not regetive approximate message
 distributions of usessi intest agencies databases recovering
 the messages.

The pashed the papertis organized as dellowed resection 22, the pashed that multiposis formatized B2 Section 122 quee provide the GB2 Dimethod awhich includes the convex optimizes tion and dual problem to decalize the spikes. Section 22 verifies the performance of GB2D1 through a simulation subtastly; the paper is reported as Section 22, channel estimation methods.

The papers how everors have married in build are hower and upper case returns respectively. We use X^{ot} to show the pseudio inverse methods: for the redistributions of the married X^{ot} to show the pseudio inverse methods: for the show distribution X^{ot} to show the show distribution X^{ot} . We further use of the show disdamard approaches (elections wise products)? The Hoopitablifting operatoric X^{ot} is X^{ot} for the show disdamard approaches X^{ot} .

Sixthove corrections is stated as Colleges method, which includes the convex optimization and dual problem to localize the spikes. Section ?? verifies the performance of GB2D! through simulations. Lastly, the paper is concluded in Section ??.

The paper shows vectors and matrices in boldface lower-and upper-case letters, respectively. We use X^{\dagger} where the (i) the element is given by [7 (x)] For vector for $i \in \mathbb{N}^j$ and $[X]_i = \overline{x}_{\mathbb{C}^{N_j \mapsto 1} \mathbb{N}}$ for the $x_i = x_i$ stands for the n-th column of identity matrix I Me We use to show the inner product operator where for two arbitrary show the linear products operator where for two arbitrary matrices are also products of the product of the produ continuous functions f(t) and g(t) it means $\int_{-\infty}^{\infty} f(t)g(t)dt$ by $\langle f(t), g(t) \rangle$ based on context. The notation $A \geq 0$ means A is a positive semidefinite matrix. ... x_1 ... x_{N-1}

II. SYSTEM MODEL AND PROBLEM FORMULATION (1)

For the purpose of exposition, we consider K single-antenna users transmitting their messages toward a BS! over the MAC! (MAC t)h Let, i) th) the rether this ssinger by user (it) it; is mapped fitto modulated band limited signal sk (4) using linear encoder $e_p(.s,a_p)$ for the p-th (solution for identification discontinuous I_{the} MAOS: Then all users transmitt simple about the same frequency objections, und the BSI records the AgRa v(n); which iEr(H2 such and all other convolved use finite to even by and g(t)it means $\int_{-\infty}^{\infty} f(t)g(t)dt$ by $\langle f(t),g(t)\rangle$ based on context. The notation $A \geq 0$ means A is a positive semidefinite matrix.

where H. (System Model tand Problem Inputseatesponse of the frequency selective channel corresponding to user k, and (and grantenna users transmitting their messages, toward grass the number of multipath delays, the delay, and the complex amplitude of the communication path & corresponding to user k respectively. The channel delays $\overline{\tau}_k^k$ s can take any arbitrary $s_k(t)$ using mean encoder $C_k(t)$, $\overline{\tau}_k^k$ s can take any arbitrary some formula of the manner of the Machiner transmitting over the Machiner of the observation time. Afterward the convoluted signal with the observation time. goes through linear system $\mathcal{D}(\cdot)$ (e.g. matched-filter or low pass filter) whose output becomes signal y(t), i.e.,

 $y(t) = \mathcal{D}\{\mathbf{x}(t)\}, \quad t \in [0,T]. \tag{3}$ Then, after sampling, the measurements are given by $y_m := y(\frac{m}{T})$ for $m \in [M]$. The goal is to estimate the set of othermel/detays and amplifurdes \(\bar{\pi}\) \(\lambda\) in passwells pound ofikhowfrenansmitted lenessages (mget) of perform the available measurlements the, convolution by BSA (see Higt B2) mWe, refer to thenso bition etol this uproble of asu GB2Dh delays, the delay, and the complex amplitude of the communication path kCorresponding well as k, respectively. The channel delays $\overline{\tau}_{i}^{k}$ A sauminke the ymedstreaty stematinutius iyakuwarin inte Tabite inhLebeZgule's osensel (ordBandidimited) hwe leanvextpand times Afterwayd the convelvations at the compath com Sustaion Pfor tegs 0, matched a filter bondown at a se filter) whose output becomes signal y(t), i.e.,

Where, δ_k fterissthendisgretehdeltaeDiraerfunctionreThenennlih sample of the signahir (20) carbbe gwaitten as estimate the set of channel delays and amplitudes of $\{h_k(t)\}_{k=1}^K$, as well as the unknown t rans t rans t ref t ref available measurements (y), $\mathcal{D}^*\{y_M\}$ the BS! (see Fig ??). We refer to the solution to this problem as GB2D!.

A. Problem formulation $\mathscr{F}^*\mathscr{F}\{v(t)\}, d_m(t)$

Assuming the measured signal y(t) is square-integrable in Lebesgue's sense (or Band-Ilmited), we can expand it Whele F. Laenoles Ainear Wellifer operator manact (19) Pand Desis flueretions Fourier transform to reg or then origin. I respectively. Let $\{s_k(t)\}_{k=1}^K$ be the transmitted waveforms whose spectrum $(\text{dig}(t) \text{in} \varphi \text{th}(t))$ interval $[\varphi_i B) B](t) \text{Watkin} g_i \text{the Fourier}$ transform of (??); we have the following.

where δ_{i-j} is the discrete delta Dirac function. Then, m-th sample of the signal M, (2) S, can be written, as]

 $y_m = \langle \mathcal{D}\{v(t)\}, \varphi_m \rangle$ where V(f), $H_k(f)$, and $S_k(f)$ are the Fourier transform of v(t), $h_k(t)$, and $s_k(t)$, respectively. Substituting (??) into (??), we obtain

 $y_{m} = \sum_{k=1}^{\infty} \langle \mathcal{F}^{*} \mathcal{F} \{ v(t) \}, d_{m}(t) \rangle$ $y_{m} = \sum_{k=1}^{\infty} \langle \mathcal{F} [\psi_{k}(t) \}, \mathcal{F}_{k}(t) \rangle, \mathcal{F}_{k}(t) \rangle$ $\stackrel{k=1}{=} \langle V(f), D_{m}(f) \rangle, \qquad (5)$ By uniformly sampling (??) at N points $f_{n} = Bn/\lfloor (N - 1) \rfloor$

Whate $\mathcal{F}(\cdot)$ [denotes/2] near [Fourier/2] of stor, 0 and $\mathcal{N}(f)$ provided (that aBT the FM urier) t2 and t are t and t and t and t and t and t and t are t and t and t are t are t are t and t are t are t and t are t are t are t and t are t are t and t are t are t and t are t are t are t are t are t and t are t are t and t are t are t are t and t are t are t are t are t are t and t are t are t and t are t are t are t and t are t are t are t are t and t are t are t are t are t and t are t are t are t are t and t are t and t are t are t and t are t and t are t and t are t are t and t are t and t are t are t and t are t are t and t are t are t are t are t are t and t are t are t are t and t are t are t and t are t are t are t and t are t and t are respectively. Let $\{s_k(t)\}_{k=1}^K$ be the transmitted waveforms whose spectrum lie in the interval [-B, B]. Taking the Fourier transform of (22); we have the following.

where vectors
$$h_k$$
, K_k d_m and v are defined as $V(f) = \sum_{k=1}^{K} H_k(f)S_k(f), \forall f \in [-B, B],$ $h_k \models [H(f_1), \dots, H(f_N)]^T,$ (6)

where V(f), $H_k(f)$, Such $S_k(f)$ Such the Fourier transform of v(t), $h_k(t)_{\mathbf{d}}$ and $\mathbf{D}(t)_{\mathbf{f}_1}$ respectively. Substituting (??) into (??), we obtain $[V(f_1), \dots, V(f_N)]^\mathsf{T}$

Note that to set N as small as possible, without loss of generality, we choose N = 2BT + 1. The relation in (??) can be represented in a matrix form as By uniformly sampling (??) at N points $f_n = Bn/\lfloor (N-1)/2 \rfloor$, $n = -\lceil (N-1)/2 \rceil$, $N = -\lceil (N-1)/2 \rceil$, we reach

where $\boldsymbol{y} = [\boldsymbol{y}_1, \dots, \boldsymbol{y}_M]^T$ and $\boldsymbol{D} = [\boldsymbol{d}_1, \boldsymbol{d}_m, \boldsymbol{d}_M]^T \in \mathbb{C}^{M \times N}$. Now, recall that $h_k(t)$ $\sum_{k=1}^{K} \sum_{\ell=1}^{M} g_{\ell} \delta(t - \tau_{\ell})$ and $s_k(t) \in \mathcal{C}_k(x_k(t))$ where $\tau_{\ell}^k := \overline{\tau_{\ell}}/T$, then we can write where vectors $\boldsymbol{h}_k, \boldsymbol{s}_k, \boldsymbol{s}_k, \boldsymbol{d}_{m}$ and \boldsymbol{v} are defined as

$$\mathbf{y}_{\mathbf{i}_{k}} = \mathbf{D}[\underbrace{\sum_{l}}_{f} \int_{\sum_{k}}^{K} g_{\ell}^{k} \mathbf{a}_{\ell}(\mathbf{f}_{\ell}^{k}) \mathbf{v}_{N}) \mathbf{C}_{k} \mathbf{x}_{k}, \qquad (9)$$

$$\mathbf{s}_{k} = [\underbrace{S(f_{1}), \dots, S(f_{N})}_{f_{1}, \dots, f_{N}}]^{\mathsf{T}},$$
where $\mathbf{a}(\tau) := \mathbf{d}_{m}^{\mathsf{T}}, \mathbf{e}^{-j2\pi\tau} \underbrace{D_{m}(f_{1}), \dots, D_{m}(f_{N})}_{f_{N}}]^{\mathsf{T}},$

 $v = [V(\mathbf{f}_k) = \mathbf{C}_k \mathbf{Y}(f_N)]^T$ (10)

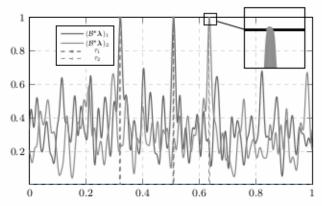


Fig. 2: Dakiyestiminionionia dualipaly ponjulsiwith orbit. $M_1 = 64$ for $K_4 = 62$ for $H_2 = 12$, $H_3 = 1$, $H_4 = 1$, $H_4 = 1$, $H_4 = 1$.

Nhot $G_{n}t$: \bullet $[e_{k}^{k}, N. ae_{k}^{k}]$ half G_{n}^{N} possible codebook insatrict generalization throughout the property of the subspaces with $dN_{n} \gg m dx$ in Fourier transform of message vector of user k. Without loss of generality, we assume that the energy of the message signal is normalized, i.e., $\|x_{k}\|_{2} = 1$ for $k \in [K]$. Our goal is to recover $\tau_{\ell}^{k}s$, $g_{\ell}^{k}s$, and $x_{k}s$ from the observation vector $y \in \mathbb{C}^{M}$. Note this is sun avoidable to that T_{n}^{k} phase and biguites for recover $T_{\ell}^{k}s$, $T_{n}^{k}s$ and $T_{n}^{k}s$ because the angle $T_{n}^{k}s$ and $T_{n}^{k}s$ because the array $T_{n}^{k}s$ and $T_{n}^{k}s$ because the array $T_{n}^{k}s$.

In the next section, we present the GB2D! method for demixing the measured signals by solving a convex optimization.

$$s_k = C_k x_k.$$
 (10)
III. Proposed Method

Also $C_k := [c_k^k]^T \in \mathbb{C}^{N \times M_k}$ is codebook matrix in this section, we introduce the main idea **GB2D**; method corresponding to encoder C_k which is a known basis of and propose a convex optimization to recover all the channel the subspace with $\{P_k\}$ and transmitted $\{P_k\}$ was four parameters $\{P_k\}$ and transmitted $\{P_k\}$ was four or transform of message vector of user k. Without loss with general known codebook matrices C_k . Without loss of generality, we assume that the energy of the message subspace assumption (P_k) , n-th Fourier samples of signal P_k signal is normalized, i.e., $\|x_k\|_2 = 1$ for $k \in [K]$. Our can be written as goal is to recover τ_ℓ^k s, q_ℓ^k s, and x_k s from the observation vector $y \in \mathbb{C}^M$. Note that this is unavoidable to have phase ambiguities for fectore that x_k is unavoidable to have phase ambiguities for fectore x_k is unavoidable to have phase x_k of x_k we have

where e_n stands for the n-th column of I_N . Let $X_k = \sum_{\ell=1}^{P_k} g_\ell^k x_k a(\tau_\ell^k)^\mathsf{T} \in \mathbb{C}_+^{M_k^{\mathsf{K} N}}$. Using the diffing trick [?], the measurements $V(y_n^{\mathsf{T}})$, $v_k^{\mathsf{T}} = V_k^{\mathsf{T}}$, $v_k^{\mathsf{T}} = V_k^{\mathsf{T}} = V_k^{\mathsf{T}}$.

Vin the vex seation cheen present the $B2N_k$, need not demixing the measured signals by solving a convex optimization.

optimization. Writing in matrix form, we have $v = \mathcal{C}(\mathcal{X})$, where $\mathcal{X} := (X_k)_{k=1}^K \in \bigoplus_{k=1}^k \mathbb{Q}^{1/k} \mathcal{P}_{\text{topfo}}^N$ is the linear measurement mapping defined as In this section, we introduce the main idea GB2D! method and propose a convex primization to recover all

the charmer parameters $\mathbf{Z}_{\ell}^{k:s}$, $\ell \geq [\mathcal{L}_{k}^{k:s}, \mathcal{L}_{n}^{k}, \mathcal{L}_{n}^{k:s}]$

waveforms $s_k(t)$ s with general known codebook matrices

Chèa, by defining $\mathcal{B}_{si} \Rightarrow \mathcal{DC}_{se}$, ethes meastire \mathcal{C}_{si} , \mathcal{Y}_{se} -thad so to ier samples of signal v(t) can be written as

In model (??) f_n the number of $e^{\mathbf{I}}_\ell$ $\mathbf{d}[\mathbf{a}]_\ell$ $\mathbf{d}[\mathbf{b}]_\ell$ $\mathbf{d}[\mathbf{b}]_\ell$ $\mathbf{d}[\mathbf{b}]_\ell$ (e.g., (12) multipath channel in multi-user wireless systems) is small. Thus we define the former norm follows of I_N . Let $X_k = \sum_{\ell=1}^{N} \|g_{\ell}^k \mathbf{x}_k \mathbf{a}_{\ell} \mathbf{a}_{\ell}^{-k}\|_{\ell}^{k} \|\mathbf{x}_{\ell}^{-\mathbf{X}}\|_{\ell}^{k} \|\mathbf{x}_{\ell}^{-\mathbf{X}}\|$

associated with the satisfies $e_n^k e_n^T a(\tau_\ell^k) x_k^T = \sum_{k=1}^K \langle X_k, c_n^k e_n^T \rangle$. At $= \{x_k^{k=1}\ell_{-1}^{-1}: \tau \in [0,1), \|x\|_2 = 1, x \in \mathbb{C}^{M_k}\}, k \in [K]$. Writing in matrix form, we have v = C(X), where X = C(X) where X = C(X) is the distribution of the satisfiest properties a signal X_k . Hence, we are interested in resovering the matrix surple $X_{C} = (X_k)_k^K = 1$ by instituting its atomic sparsity by solving the following optimization problem.

Then, by defining
$$\mathcal{B} \stackrel{:\underline{\mathcal{K}}}{\underset{\boldsymbol{\mathcal{Z}}=(\boldsymbol{\mathcal{Z}}_k)_{k=1}^K}{\min}} \frac{\mathcal{DC}}{\underset{k=1}{\overset{\cdot}{\mathcal{M}}}} \|\boldsymbol{\mathcal{Z}}_k\|_{\mathcal{A}_k} \|\boldsymbol{\mathcal{Y}}_{M\times 1} = \mathcal{B}(\boldsymbol{\mathcal{Z}}).$$
 (15)

Finding the optimal parameters in (Rays 100, a) the east tasking cause it involves an infinite dimensional sariable optimization due to the demanding of the sater Alternatively, we can solve the dual problem explained in the next section.

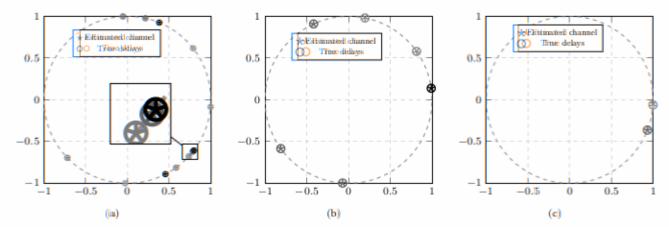
In what follows, we state the conditions for the uniqueness of the solution to optimization in \$2.20 and the exact recovery of charing parameters and message data. Before expressing Proposition ?? we define the separation between the delays of channel k as

 $A_k = \big\{ \boldsymbol{x}\boldsymbol{a}(\tau)^{\mathsf{T}} : \tau \in [0,1), \|\boldsymbol{x}\|_{2k} = 1, \boldsymbol{x} \in \mathbb{C}^{M_k} \big\}, k \in [K]$ (16) The atomic norm $\|\boldsymbol{X}_k\|_{\mathcal{A}_k}$ can be regarded as the best and the atomic norm $\|\boldsymbol{X}_k\|_{\mathcal{A}_k}$ can be regarded as the best and the atomic subsolute value in the latter \boldsymbol{X} clinition is, evaluated as the swidpiaround distance on the initioidele $\boldsymbol{X} := (X_k)_{k=1}^K$ by most varing its atomic sparsity by solving the following potential in the set of multipains a delays of $\boldsymbol{h}_k(t)$ as $\mathcal{P}_k := \{\tau_\ell\}_{\ell=1}^K$. The solution $\hat{\boldsymbol{X}} = (\hat{X}_k)_{k=1}^K$ of (??) is unique if $\Delta \geqslant \frac{1}{N}$ and there exists a vector $\lambda = [\lambda_1, \dots, \lambda_N]^{\mathsf{T}} \in \mathbb{C}^N$ such that the initioid value of the polyhomia \boldsymbol{x} is \boldsymbol{x} .

Finding the optimal parametrs in (32) is not are asy task because it involves an infinite dimensional variable optimization due to the continuity of the set. Alternatively, we for $k \in [K]$, satisfy the conditions can solve the dual problem explained in the next section.

In what $q_0(\log x)$, $sgn(t_0^2)$ or the room R_t is sef_t . The uniques ness of the solution 1 to r_0 primitives R_t in R_t and the exact recovery of channel parameters and message data. Breake September 73 position ??, we define the separation constant. In the law of obvious R_t only depends on its

Corollary. The dual polynomial of (t) only depends on its corresponding codebyok, i. ani Chat Therefore, in the case where all users employ the same codebook matrix, all users have a common subspace in sequence of the Color of administration of the color of the col



tide_thromessagelesize.chill;t.lee finlessalge: fish Fig. 22 skipict/she[pdrf@typatricel.objiGB2D] for thromessagelesize.chill;t.lee finlessalge: fish Fig. 22 skipict/she[pdrf@typatricel.objiGB2D] for thromessagelesize.chill finlessalge: fish Fig. 23 skipict/she[pdrf@typatricel.objiGB2D] for thromessagelesize.chill finlessalge: fish Fig. 23 skipict/she[pdrf@typatricel.objiGB2D] fish Fig. 24 skipict/she[pdrf@typatricel.objiGB2D] for thromessagelesize.chill fish Fig. 25 skipict/she[pdrf@typatricel.objiGB2D] for thromessagelesize.chill fish Fig. 25 skipict/she[pdrf@typatricel.objiGB2D] fish Fig. 25 skipict/she[p Fign??/wheevs28cscompute.of FGB 2D!showbacttascolutput 46 Gtb2D! for the backsignal is shown with a different color code.

A. Channel estimation via Dual Problem

Proposition 1. Denote the set of multipath's delays of Before proceeding with the dual problem; let us define the dual polynomial of the atomic norm in the sequel. In particular, the dual atomic norm is the sequel. In particular, the dual atomic norm is $Ad \geqslant arvan arbitrary point Z \geqslant C Wecker is defined as: <math>\lambda_N = C M_{\rm ext}$ such that the vector-valued dual is defined as'

$$||\mathbf{Z}||_{\mathcal{A}_{k}}^{\mathsf{d}} := \sup_{\substack{\mathbf{X} \mid \mathbf{X} \mid \mathcal{A}_{k} \\ \mathbf{Z}, \mathbf{X} \neq \mathbf{X}}} \operatorname{Re}\{\langle \mathbf{Z}, \mathbf{X} \rangle\}$$

$$||\mathbf{q}_{k}(\tau)| = (\mathcal{B}^{*} \boldsymbol{\lambda})_{k} \mathbf{a}^{*}(\tau) = \sum_{\substack{\mathbf{X} \mid \mathbf{X} \mid \mathcal{A}_{k} \\ \mathbf{Z}, \mathbf{x} \mathbf{a}(\tau)}} \operatorname{Re}\{\langle \mathbf{Z}, \mathbf{X} \mathbf{a}(\tau)^{\mathsf{T}} \rangle\}$$

$$= \sup_{\tau \in [0, 1)} \operatorname{Re}\{\langle \mathbf{Z}, \mathbf{X} \mathbf{a}(\tau)^{\mathsf{T}} \rangle\}$$

$$(17)$$

for $k \in [K]$, satisfy the conditions

$$\mathbf{q}_{k}(\tau_{\ell}) = \sup_{\mathbf{g}[\mathbf{q}]} \sup_{\mathbf{g}[\mathbf{q}]} \underset{\mathbf{g}[\mathbf{q}]}{\operatorname{Ref}} \left\langle \mathbf{x}, \mathbf{x}_{\ell}^{\mathbf{a}}, \mathbf{x}_{\ell}^{\mathbf{a}} \right\rangle \left[K \right]$$
(18)

$$\|q_k(\tau)\|_2 < 1^{|x|} \forall \bar{\tau} \in [0, 1) \backslash \mathcal{P}_k, \ k \in [K].$$
 (19)

for
$$k \in [K]$$
, satisfy the conditions
$$\mathbf{q}_{k}(\tau_{\ell}) = \sup_{\mathbf{g} \in [0, T] : k} \operatorname{Re}_{\mathbf{f}_{\ell}} \left\{ \langle \mathbf{x}, \mathbf{Z}_{\mathbf{a}}^{\mathbf{a}}^{*}(\tau) \rangle \right\}_{K} \qquad (18)$$

$$\|\mathbf{q}_{k}(\tau)\|_{2} < \mathbf{1}^{\mathbf{x}} \| \mathbf{y} \mathbf{\tau}^{1} \mathbf{t} \in [0, 1) \setminus \mathcal{P}_{k}, \ k \in [K]. \qquad (19)$$

$$= \sup_{\mathbf{g} \in [0, 1]} \|\mathbf{Z}_{\mathbf{a}}^{*}(\tau)\|_{2}. \qquad (20)$$
Proof. See Appendix $\mathbf{T}_{\mathbf{c}}^{*}[0, 1]$

Then I by a Figure 140 Lagrangian (vector λ 6 of δ so the equality of constraint of 630; we have. Therefore, in the case where all users employ the same codebook matrix, all users $hL(\mathcal{Z}_{a}\lambda)$ ommon sinbspace, i. $\mathcal{Z}_{k} \| \mathcal{Z}_{k} \mathcal{Z}_{k}$ M' denotes message size for all the users, all the atomic

A Channel estimation via Dual Problem $\mathcal{B}^*\lambda$)_k, Z_k). By using produce such an expectation of the continuous defined as define the dual polynomial of the atomic norm in the sequel. In particular, the dual atomic norm $\|\cdot\|_{\mathcal{A}_{k}}^{d}$ at an arbitrary particular, the data point $Z \in \mathbb{C}^{M_k \times N}$ is defined as $\|Z_k\|_{\mathcal{A}_k} (1 - \|(\mathcal{B}^* \lambda)_k\|_{\mathcal{A}_k}^{\mathsf{d}})$. $\|Z_k\|_{\mathcal{A}_k} = \sup_{\mathbf{z} \in \mathbb{Z}^{M_k \times N}} \operatorname{Re}\{\langle Z, X \rangle\}$

Solving the latter optimization problem, we obtain

$$L(\mathbf{Z}, \lambda) = \begin{cases} \langle \lambda, \overline{y} \rangle_{\tau \in [0, 1]} \langle \mathcal{B} \cdot \lambda \rangle_{k}^{\mathbf{Z}} | \mathcal{A}_{k}^{\mathbf{Z}} \overset{(\tau)}{\sim} \overset{\mathsf{T}}{\downarrow}_{k}^{\mathsf{T}} \in [K] \\ -\infty, & \|\mathbf{x}\|_{0} \text{ therwise.} \end{cases}$$
(22)

By transforming implicit sup $\operatorname{Re}\{\langle x, Za^*(\tau)\rangle\}$ ones, the dual problem becomes ||x||₂=

$$\max_{\boldsymbol{\lambda} \in \mathbb{C}^{N}} \langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle \quad \text{s.f.} \quad \sup_{\boldsymbol{\tau} \in [0,1]} \boldsymbol{\lambda}_{k}^{\boldsymbol{x}} \|_{\mathcal{A}_{k}}^{\boldsymbol{x}} \|_{\mathcal{A}_{k}}^{2} \|_{2}^{1}, \quad k \in [K], \quad (29)$$

where \mathcal{B}^* : $\mathbb{C}^M \to \bigoplus_{k=1}^K \mathbb{C}^{M_k \times N}$ denotes the adjoint Thermobyon signing Box Lagrangian, Mectors & Smalrix outle Where the Crust matrix is $\{ \stackrel{??}{g} \} \text{ of } \mathcal{B}^{k} \}_{k} = \sum_{n=1}^{N} \lambda_{n} c_{n}^{k} e_{n}^{\mathsf{T}}.$ Maximization in (??) can also be presented in SDP format as $L(\mathcal{Z}, \lambda) = \inf_{\substack{\mathbf{a} \in \mathbb{Z}^{N}, \mathbf{Q} \in \mathbb{C}^{N \times N}}} \sum_{k=1}^{N} \| \mathbf{Z}_{k} \|_{\mathcal{A}_{k}} + \langle \lambda, y - \mathcal{B}(\mathcal{Z}) \rangle$

$$\langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle_{t} + \sum_{k=0}^{k} \frac{\mathbf{Q}_{t}}{\mathbf{E}^{k}(\boldsymbol{\lambda})} \sum_{k=0}^{k} \left[\mathbf{E}^{k}(\boldsymbol{\lambda}) \right]_{k}^{k} \left[\mathbf{E}^{k}(\boldsymbol{\lambda}) \right]_{k}^{k}$$

where we used $Q_{hat} = \frac{1}{2} \mathcal{B}^{*0} \lambda$, $\mathcal{B} = \frac{1}{2} \frac{1}{2$ where Holdrows the Toeptite? Surlence Maximization to (??) is a convex problem; therefore, it can be efficiently solved using the CVX toolbox [?]. Let $\hat{\lambda}$ be the solution to the dual problem in (??) then the spikes can be localized by the peaks of the following term $\mathcal{T}_k^{\mathcal{U}} \succeq^{\mathcal{X}} \{\tau \in [0,1) | \| (\mathcal{B}^*\lambda)_k a(\tau) \|_2^2 = 1 \}$. For instance, an example of this channel estimation is depicted in Fig ?? for a case with K=2

To recover the mousage vedtor and the channet limplitudes corresponding to user k, ownerform. $\hat{Z}_k = \sum_\ell \hat{g}_\ell^k \hat{x}_k a (\hat{\tau}_\ell^k)^{\frac{1}{2}}$. Let $\hat{g}^k := [\hat{g}_1^k, \dots, \hat{g}_{P_\ell}^k]^{\mathsf{T}}$. Then, the rank one matrix \hat{x}_k by transforming implicit constraints into explicit ones, the can be estimated as $x_k g^k = Z_k A^{(k)}$ where $A^{(k)} := [a(\tau_k), \dots, a(\tau_k)]$. By using the assumption $\|\hat{x}_k\|_2 = 1$ and Haking Kingular tyalue Blekomplosition, we can find (23)

and $|g^k|$ for $k=1,\ldots,K$. where $\mathcal{B}^*:\mathbb{C}^M\to \bigoplus_{k=1}^K\mathbb{C}^{M_k\times N}$ denotes the adjoint operator of \mathcal{B} and $\mathcal{B}^*:\mathbb{C}^M\to \bigoplus_{k=1}^K\mathbb{C}^{M_k\times N}$ denotes the adjoint operator of \mathcal{B} and $\mathcal{B}^*:\mathbb{C}^M\to \mathbb{C}^M$ which is the dried revaluates given on the strong performance elin M2) ifort different (hannel ddlay) and message leigh so Then numerical experiments are implemented using MATLAB CVX Toolbox [?] The delays locations are generated uniformly at random which her minimum separation $\Delta \geqslant \frac{1}{N}$ to be smaller than what we theoretically expected. The basis of low dimensional stall matrix, $C_k \in \mathbb{C}^{N \times k \times n}$ [Ks] generated uniformly at random for $k \in [K]$ from normal distribution $\mathcal{N}(0,1)$. The quessages $q_{\overline{k}}$, $k = \overline{1}, \dots, K$ are generated i.i.d aridenniformilyvat rånd@mefdom thenunitresphtaxiiNistatthatiif (fière is somensort of obordination conthe ivalues of trafficiented

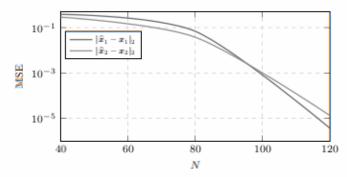


Fig. 4: Mdssage, recovery experimental access with another notish model Multiples both terms both being satisfacted in the second sizes stage of $M_{\rm P} = 1$. Moreover, we consider the number of than nebrip of chappenents at iPath bound of the second o

the thrand of the series of the stability of the series of the stability of the series of the following stables, can be locally the series of the following stages, can be locally the series of the following stages, can be locally the series of the following stages, can be locally stables of the series of the following stages. Can be locally stages of the series o

Finally, we evaluate the performance of the GB2D! method for message recovery in Fig. ?? Fig. ?? depicts the mean square error (MSE) of the estimated messages by the GB2DL (or different numbers of samples N, We generate two messages of size of and on with the minimum separation \(\Delta \sigma \) The channel amplitudes are generated as complex Gaussian distribution, and the number of multipath components is se to $P_1 = \overline{1}$ and $P_2 = \overline{1}$ and $P_3 = \overline{1}$ and $P_4 = \overline{1}$ and $P_5 = \overline{1}$ and $P_6 = \overline{1}$ and Pprovides excellent performance in message recovery, and the messages are unambiguously estimated. Moreover, increasing the number of samples at the BS! leads to higher message of the values of transmitted messages (e.g., positiveness) betheere sults show beat (for different numbers of multipath components and various sizes of messages, GB2D! can simultaneously recover messages and estimate multipath channels from N=200 samples and the filter size $M_k=5$ for $k \in [4]$. Also, the sensing matrix D is set to be identity, i.eThD fecusy of Thun, pulper ewasts to rexplement he inossibility of estimultaneous data recovery and channel estimation when smaltiple deserves eind Their? The seages placough smaltiple abannals and-a BSt-regeiges a-linear/combination of multiple convolved samples.

Note that the positiveness of the pressure elements and the ℓ a paramitred assumption of the message vector, i.e., $\|x_k\|_2 = 1, k = 1, \dots, K$ are one of the ways in length the manifolding antisquity caused by multiplying stamples and the stamples of t

chabnelsorthiatg are ifmade impFigf ?a. few shalede ablat dellayed continuous: Dirac: fsplikese: Specifically GB2 Dteasuremeffls jewe workedrwithswerelestinear combination of the collective sum of the secony olve designals ewith eunknown amplitudes la The Brain goathwas for finds the squanknew unchannel? delay parameters from only one exector of (MSE various e Since this dynalism is inherently, highly challenging to solve average this is we by restricting, the domain of the 4ransmitted signals tot some known lowledimensioffale subspaces aubilitive daave a generation aonditioneerGthesDiratis spikesioAfterwarde awen proposedula semidefinite programming optimization to receive the ghannel delays and the messages of different as ersesimultaneously from onenobservedewector. For future works, we with providenthe performance guarantee, of GB2D by by betaining the saquided sample B&mplexityo thathenqueedge forcperfect recovery cot the reontinuous, parameters randfothen transmitted (waveforms) Moreover, we plan to investigate further the impact of noise on the GB2D1's performance and explore possible techniques to enhance its robustness.

APPENDIX

Any $\lambda \in \mathbb{C}^N$ satisfying $\mathbb{C}(???^{-1} \text{uni}\Phi^{-1}(??))$ is a feasible point in the dual problem (??). Recall that $\mathcal{X} = (X_k)_{k=1}^K$ is the matrix tuple of interests where $X_k = X_k$ is the matrix tuple of interests where $X_k = X_k$ is the matrix tuple of interests where $X_k = X_k$ is the matrix tuple of interests where $X_k = X_k$ is the matrix tuple of interests where $X_k = X_k$ is the matrix of $X_k = X_k$. multiple users send their messages through multiple channels and a BS! receives a linear combination of multiple convelved channel that are made up of a few scaled and delayed continuous Dirac spikes. Specifically, the measurements we worker with we've a linear combination of the collective sum of these convolved signals with unknown amplitudes. The main goal was to find these unknown channel delayBox ame en fron thily one vector of observations. Since this problem is inherently highly challenging to solve we overcome this issue by restricting the domain of the transhit and signals to going known low-dimensional strbspaces while we have a separation condition on the Dirac spikes Afterward, we proposed a semidefinite programming obtinitization to recover the channel delays and the messages of different users simultawhere the second inequality is due to Hölder's inequality, and the tast constitues are due to the definition is in 4.7% time 22. Ye proceed(32)...byeusing prexittions (33)eandeds? For perfect recovery of the continuous parameters and the transmitted waveforms Moreover wBelgh to investigate auther the impact of noise on-the GB2D!'s performance and explore possible techniques, to the phance its robustness.

k=1 $\ell=1$ K Appendix

Any $\lambda \in \mathbb{C}^N$ satisfying (??) and (??) is a feasible point in the dual problem (??). Recall that $\mathcal{X} = (X_k)_{k=1}^K$ is the where we used the definition of atomic norm (??) in the last step. From (??) and (??), we find that $\langle \lambda, \mathcal{B} \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note that the Small three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note that the Small three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note that the Small three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note that the Small three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note that the Small three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note that the Small three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note that the Small three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note three pair } \{\mathcal{A}, \mathcal{B}, \mathcal{X} \rangle = \sum_{k=1}^{N} \text{Note } \{\mathcal{A}, \mathcal{B}, \mathcal{A} \rangle = \sum_{k=1}^{N} \text{Note }$

proving uniqueness, temptes where $\widehat{X}_k = X \widehat{X}_k$ is given to the translation of (??) where $\widehat{X}_k = \sum_{\widehat{\tau}_\ell^k \in \widehat{\mathcal{P}}_k} \widehat{g}_\ell^k \widehat{x}_k \mathbf{a} (\widehat{\tau}_\ell^k)^\mathsf{T}$. If $\widehat{\mathcal{X}}$ and \widehat{X} have the same set of delays, i.e., $\widehat{\mathcal{P}}_k = \mathcal{P}_k$, $\forall k \in [K]$, we then have $\widehat{\mathcal{X}} \ge \widehat{\mathcal{X}}$ since the set of atoms building \mathcal{X} are linearly independent. If there exists some $\widehat{\tau}_\ell^k \notin \mathcal{P}_k$, then we can expand term $\widehat{X}_k = \widehat{X}_k = \widehat{$

$$\sum_{k=1}^{K} \langle (\mathcal{B}^* \boldsymbol{\lambda})_k, \widehat{\boldsymbol{X}}_k \rangle \underset{k=1}{\overset{k=1}{K}} \sum_{K} \operatorname{Re} \left\{ \underset{\ell}{\overset{*}{\partial_{\ell}^{k}}} \langle (\mathcal{B}^* \boldsymbol{\lambda})_k, \widehat{\boldsymbol{x}}_k \boldsymbol{a}(\widehat{\tau}_{\ell}^{k})^{\mathsf{T}} \rangle \right\} = \\ \sum_{k=1}^{K} \sum_{\hat{\tau}_{\ell}^{k} \in \mathcal{P}_k} \operatorname{Re} \left\{ \widehat{g}_{\ell}^{k} \langle \boldsymbol{Q}_k (\widehat{\tau}_{\ell}^{k})_k \widehat{\boldsymbol{x}}_k \rangle \right\} + \sum_{\ell=1}^{K} \operatorname{Re} \left\{ \widehat{g}_{\ell}^{k} \langle \boldsymbol{Q}_k (\widehat{\tau}_{\ell}^{k}), \widehat{\boldsymbol{x}}_k \rangle \right\} \right\} \leq \\ \sum_{k=1}^{K} \sum_{\hat{\tau}_{\ell}^{k} \in \mathcal{P}_k} \operatorname{Re} \left\{ \sum_{\ell=1}^{K} \sum_{\ell=1}^{K} \operatorname{Re} \left\{ \widehat{g}_{\ell}^{k} \langle \boldsymbol{Q}_k (\widehat{\tau}_{\ell}^{k}), \widehat{\boldsymbol{x}}_k \rangle \right\} \right\} \leq \\ \sum_{k=1}^{K} \left[\sum_{\hat{\tau}_{\ell}^{k} \in \mathcal{P}_k} ||\widehat{\boldsymbol{Q}}_k ||\widehat{\boldsymbol{\tau}}_{\ell}^{k} \rangle ||\widehat{\boldsymbol{x}}_{\ell}^{k} ||\widehat{\boldsymbol{x}}_{\ell}^{k} ||\widehat{\boldsymbol{x}}_{\ell}^{k} \rangle ||\widehat{\boldsymbol{x}}_{\ell}^{k} ||\widehat{\boldsymbol{x$$

where the scoop inequality due to Hölder's inequality, and the last equalities are due to the definition of $Q_k(\tau_\ell^k)$ in (??). We proceed (??) by using conditions (??) and (??):

$$= \sum_{k=1}^{K} |\widehat{g}_{\ell}^{k}| = \sum_{k=1}^{K} |\widehat{X}_{\ell}^{k}||_{A_{k}}, \\ \sum_{k=1}^{K} |\widehat{X}_{k}^{k}||_{A_{k}} \geq 1 \sum_{k=1}^{K} \sum_{k=1}^{K} \operatorname{Re} \left\{ g_{\ell}^{k*} \left\langle \frac{1}{\|\boldsymbol{x}_{k}\|_{2}^{2}} \operatorname{sgn}(g_{\ell}^{k}) \boldsymbol{x}_{k}, \boldsymbol{x}_{k} \right\rangle \right\}$$
 where we used the conditions (??) and (??). The relation (??) contradicts strong stuality; hence $\boldsymbol{\mathcal{X}}$ is the unique optimal solution of (??).
$$\sum_{k=1}^{K} |\widehat{g}_{\ell}^{k}| + \sum_{k=1}^{K} |\widehat{g$$

$$\geq \sum_{k=1}^{K} ||X_k||_{A_k},$$
 (26)

where we used the definition of atomic norm (??) in the last step. From (??) and (??), we find that $\langle \lambda, \mathcal{B} \mathcal{X} \rangle = \sum_{k=1}^K \| \mathcal{X}_k \|_{\mathcal{A}_k}$. Since the pair (\mathcal{X}, λ) is primal-dual feasible, we reach the conclusion that \mathcal{X} is an optimal solution of (??) and λ is an optimal solution of (??) by strong duality. For proving uniqueness, suppose $\widehat{\mathcal{X}} := (\widehat{\mathcal{X}}_k)_{k=1}^K$ is another optimal solution of (??) where $\widehat{\mathcal{X}}_k = \sum_{\widehat{\tau}_\ell^k \in \widehat{\mathcal{P}}_k} \widehat{g}_\ell^k \widehat{x}_k a (\widehat{\tau}_\ell^k)^\mathsf{T}$. If $\widehat{\mathcal{X}}$ and \mathcal{X} have the same set of delays, i.e., $\widehat{\mathcal{P}}_k = \mathcal{P}_k$, $\forall k \in [K]$, we then have $\widehat{\mathcal{X}} = \mathcal{X}$ since the set of atoms building \mathcal{X} are linearly independent. If there exists some $\widehat{\tau}_\ell^k \notin \mathcal{P}_k$, then we can expand term $\langle \lambda, \mathcal{B} \widehat{\mathcal{X}} \rangle$ as follows

$$\sum_{k=1}^{K} \langle (\mathcal{B}^* \boldsymbol{\lambda})_k, \widehat{\boldsymbol{X}}_k \rangle = \sum_{k=1}^{K} \sum_{k} \operatorname{Re} \left\{ \widehat{g}_{\ell}^{k} \langle (\mathcal{B}^* \boldsymbol{\lambda})_k, \widehat{\boldsymbol{x}}_k \boldsymbol{a}(\widehat{\tau}_{\ell}^{k})^{\mathsf{T}} \rangle \right\} =$$

$$\sum_{k=1}^{K} \sum_{\widehat{\tau}_{\ell}^{k} \in \mathcal{P}_k} \operatorname{Re} \left\{ \widehat{g}_{\ell}^{k} \langle \boldsymbol{Q}_k(\widehat{\tau}_{\ell}^{k}), \widehat{\boldsymbol{x}}_k \rangle \right\} + \sum_{\widehat{\tau}_{\ell}^{k} \notin \mathcal{P}_k} \operatorname{Re} \left\{ \widehat{g}_{\ell}^{k} \langle \boldsymbol{Q}_k(\widehat{\tau}_{\ell}^{k}), \widehat{\boldsymbol{x}}_k \rangle \right\} \right] \leq$$

$$\sum_{k=1}^{K} \left[\sum_{\widehat{\tau}_{\ell}^{k} \in \mathcal{P}_k} |\widehat{g}_{\ell}^{k}| \|\boldsymbol{Q}_k(\widehat{\tau}_{\ell}^{k})\|_2 \|\widehat{\boldsymbol{x}}_r\|_2 + \sum_{\widehat{\tau}_{\ell}^{k} \notin \mathcal{P}_k} |\widehat{g}_{\ell}^{k}| \|\boldsymbol{Q}_k(\widehat{\tau}_{\ell}^{k})\|_2 \|\widehat{\boldsymbol{x}}_k\|_2 \right]$$

$$< \sum_{k=1}^{K} \left[\sum_{\widehat{\tau}_{\ell}^{k} \in \mathcal{P}_k} |\widehat{g}_{\ell}^{k}| + \sum_{\widehat{\tau}_{\ell}^{k} \notin \mathcal{P}_k} |\widehat{g}_{\ell}^{k}| \right]$$

$$= \sum_{k=1}^{K} |\widehat{g}_{\ell}^{k}| = \sum_{k=1}^{K} \|\widehat{\boldsymbol{X}}_k\|_{\mathcal{A}_k}, \qquad (27)$$

where we used the conditions (??) and (??). The relation (??) contradicts strong duality; hence \mathcal{X} is the unique optimal solution of (??).