# 振幅を柔軟に調節できる時間周波数表現の凸最適化による実現*

◎新井慶大, 山田宏樹, 矢田部浩平 (農工大)

## 1 まえがき

音響信号処理において時間周波数 (T-F) 解析は欠かせない．T-F 解析では離散ガボール変換 (DGT) が広く用いられるが，これにより得られた T-F 表現は窓の影響で広がりを持ち，解析に影響を及ぼす．そこで，より明瞭な T-F 表現を得るためのスパース T-F 解析が長い間研究されてきた．その中でも，スペクトログラムの冗長性を利用し，信号の完全再構成制約のもとで T-F 表現を調節する，最適化に基づく手法が多数提案されている．$\ell_1$ ノルム最小化 [1] は最も標準的に用いられる手法であり，各 T-F ビンを独立に圧縮することでスパースな表現を得る．

一方，現実の音響信号の T-F 表現では複数のビンが関係を持つ．例えば，ピアノのような調波楽器音の振幅スペクトログラムは低ランク行列で近似しやすい．また，音響信号は連続的に変化するため振幅スペクトログラムは滑らかである．さらに，多くの音響信号は調波構造を持つ．これらの構造は音源分離や調波打撃音分離などの応用でしばしば利用される [2, 3]．そのため，低ランク性や滑らかさなどの特定の構造を振幅に残しつつ，スパースな T-F 表現を得ることが望ましい．これらの構造を振幅に反映するには，一般に非凸最適化問題を解く必要がある．しかし，非凸最適化は初期値が適切でないと大域的最適解を得られず，解析の意味では好ましくない．そこで本稿では，perspective 関数 [4] を用いて振幅にペナルティを課すことのできる凸最適化手法を提案し，振幅の柔軟な調節が可能な T-F 解析を実現する．

## 2 最適化に基づく T-F 解析

T-F 解析で用いる変換にはさまざまな種類があるが，本稿では DGT による T-F 解析のみを扱う．

### 2.1 $\ell_1$ ノルム最小化によるスパース T-F 表現

音響信号 $\boldsymbol{d} \in \mathbb{R}^L$ の窓 $\boldsymbol{w} \in \mathbb{R}^L$ を用いた DGT は

$$x_{m,n} = \sum_{l=0}^{L-1} d_l \, \overline{w_{l-an}} \, \mathrm{e}^{-2\pi \mathrm{i} ml/M} \tag{1}$$

で定義される．ただし，$\mathrm{i} = \sqrt{-1}$，$a \in \mathbb{N}$ は窓のシフト幅，$N, M \in \mathbb{N}$ はそれぞれ時間フレーム数と周波数ビン数，$n = 0, \ldots, N-1$ と $m = 0, \ldots, M-1$ はそれぞれ時間と周波数のインデックスである．また，信号長 $L$ は $N = L/a$ と $MN > L$ を満たすものとする．式 (1) は行列 $\boldsymbol{G}_w \in \mathbb{C}^{MN \times L}$ を用いて $\boldsymbol{x} = \boldsymbol{G}_w \boldsymbol{d}$ と表せる．ただし，$(G_w)_{m+nM,l} = \overline{w_{l-an}} \, \mathrm{e}^{-2\pi \mathrm{i} ml/M}$ である．$\boldsymbol{G}_w^{\mathrm{H}} \boldsymbol{G}_w$ が正則であれば，$\boldsymbol{w}$ の標準双対窓 $\boldsymbol{\gamma}^\star = (\boldsymbol{G}_w^{\mathrm{H}} \boldsymbol{G}_w)^{-1} \boldsymbol{w}$ が存在し，

$$\boldsymbol{G}_{\gamma^\star}^{\mathrm{H}} \boldsymbol{G}_w = \boldsymbol{G}_w^{\mathrm{H}} \boldsymbol{G}_{\gamma^\star} = \boldsymbol{I} \tag{2}$$

を満たす．ただし，$\boldsymbol{I}$ は単位行列である．

DGT により得られた T-F 表現は窓の影響で広がりを持ち，解析に影響を及ぼすため，できる限り広がりのない表現を得るための手法が多数研究されている．その中に，スパース最適化に基づく手法がある．例として，信号の完全再構成性を制約とし，スペクトログラムの $\ell_1$ ノルムを最小化する手法がある [1]．この手法では凸最適化問題

$$\min_{\boldsymbol{x} \in \mathbb{C}^{MN}} \ \|\boldsymbol{x}\|_1 \ \ \text{s.t.} \ \ \boldsymbol{G}_{\gamma^\star}^{\mathrm{H}} \boldsymbol{x} = \boldsymbol{d} \tag{3}$$

を解くことでスパースな T-F 表現を得るが，過剰な局在化により元信号の構造がわかりにくくなるという課題がある．

### 2.2 グループ構造を反映したスパース T-F 解析

上記の課題に対する解決策の 1 つとして，信号の構造を考慮した上でスパース性を誘導する最適化手法がある．先行研究では，主変数に加えて信号の構造を反映する補助変数を導入し，信号のグループ構造を失わずにスパース性の誘導を可能にしている [5]．この手法では，グループ構造を考慮するために perspective 関数を補助変数への全変動制約のもとで最小化する．

我々は以前，上記の先行研究に基づき，音響信号のグループ構造を反映したスパース T-F 解析手法を提案した [6]．これは，スペクトログラム $\boldsymbol{x} \in \mathbb{C}^{MN}$ と補助変数 $\boldsymbol{\sigma} \in \mathbb{R}^{MN}$ に対し，以下の凸最適化問題を解くことで実現される．

$$\min_{(\boldsymbol{x},\boldsymbol{\sigma}) \in \mathbb{C}^{MN} \times \mathbb{R}^{MN}} \ \varphi(\boldsymbol{x},\boldsymbol{\sigma}) + \lambda \|\boldsymbol{D}\boldsymbol{\sigma}\|_{2,1} \ \ \text{s.t.} \ \ \boldsymbol{G}_{\gamma^\star}^{\mathrm{H}} \boldsymbol{x} = \boldsymbol{d} \tag{4}$$

ここで，$\varphi$ は 2 変数の組 $(\boldsymbol{x},\boldsymbol{\sigma})$ に関する凸関数

$$\varphi(\boldsymbol{x},\boldsymbol{\sigma}) = \sum_{k=1}^{MN} \phi(x_k, \sigma_k) \tag{5}$$

である．ただし，$\phi$ は $|\cdot|^2/2 + 1/2$ の perspective

*Realizing time-frequency representations with flexible adjustment of magnitude via convex optimization. By Keidai ARAI, Koki YAMADA, and Kohei YATABE (Tokyo University of Agriculture and Technology).

関数 [4] であり，

$$\phi(x_k,\sigma_k)=\begin{cases}\frac{|x_k|^2}{2\sigma_k}+\frac{\sigma_k}{2} & (\sigma_k>0)\\ 0 & (x_k=0 \text{ and } \sigma_k=0)\\ \infty & (\text{otherwise})\end{cases} \tag{6}$$

で定義される．また，$\boldsymbol{D}$ は $\boldsymbol{\sigma}$ を $M\times N$ の配列とみたときに，各要素における勾配の近似を返す差分作用素，$\|\cdot\|_{2,1}$ は $\ell_{2,1}$ ノルムを表す．式 (4) の凸最適化問題を解くと，図–1 右のように音響信号のグループ構造を反映したスパースな T-F 表現が得られる．

## 3 提案手法

信号のグループ構造のみならず，さまざまな構造を T-F 表現に反映させるために，[6] と同様の考えを利用し，振幅を柔軟に調節できる T-F 表現を凸最適化により実現する手法を提案する [7]．まず，提案手法において重要な役割を担う perspective 関数の役割について述べる．次に，提案手法が振幅への凸ペナルティを実現することを示し，その凸最適化問題を解くためのアルゴリズムを導出する．

### 3.1 perspective 関数の役割

式 (5), (6) で定義される perspective 関数の和 $\varphi$ が $\boldsymbol{x}$ に与える影響を考える．$\boldsymbol{\sigma}>\boldsymbol{0}$ のとき $\varphi$ は

$$\varphi(\boldsymbol{x},\boldsymbol{\sigma})=\frac{1}{2}(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}+\|\boldsymbol{\sigma}\|_1) \tag{7}$$

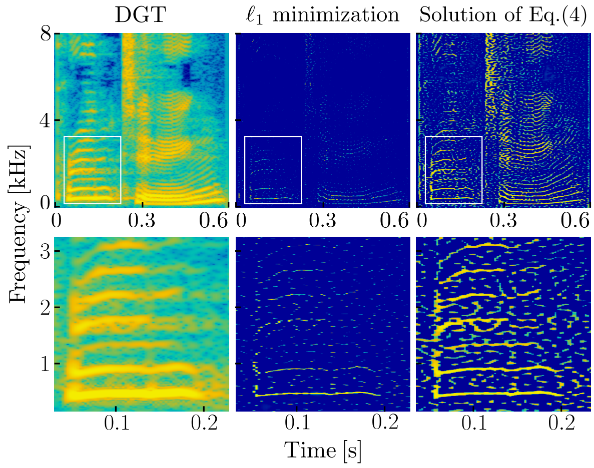
と表すことができる．ただし，$\boldsymbol{\Sigma}=\mathrm{diag}\,(\boldsymbol{\sigma})$ である．第 2 項から $\varphi$ は $\boldsymbol{\sigma}$ のスパース性を誘導する．さらに，第 1 項は重み $\boldsymbol{1}/\boldsymbol{\sigma}$ を持つ $\boldsymbol{x}$ の重み付き $\ell_2$ ノルムの 2 乗と解釈できるため，$\varphi$ を最小化すると $\boldsymbol{x}$ は $\boldsymbol{\sigma}$ の構造を持つように誘導される．したがって，$\boldsymbol{\sigma}$ に適切な正則化を施すことで，スパース性の誘導に加えて $\boldsymbol{x}$ に所望の構造を反映することができると考えられる．

### 3.2 振幅への凸ペナルティの実現

$\varphi$ が $\boldsymbol{x}$ に反映する構造は $\boldsymbol{\sigma}$ で決まるため，$\boldsymbol{\sigma}$ に所望の構造を持たせることを考える．$\boldsymbol{\sigma}$ に与える構造を柔軟に変えられるように，$\boldsymbol{\sigma}$ に施す正則化を一般化したペナルティ関数として $\Psi\in\Gamma_0(\mathbb{R}^{MN})$ を導入する．ただし，$\Gamma_0(\mathcal{X})$ は実ヒルベルト空間 $\mathcal{X}$ から $\mathbb{R}\cup\{\infty\}$ への下半連続な真凸関数全体の集合である．式 (4) の目的関数の第 2 項を $\Psi$ に置き換え，

$$\min_{(\boldsymbol{x},\boldsymbol{\sigma})\in\mathbb{C}^{MN}\times\mathbb{R}^{MN}}\ \varphi(\boldsymbol{x},\boldsymbol{\sigma})+\Psi(\boldsymbol{\sigma})\ \text{ s.t. }\ \boldsymbol{G}^{\mathrm{H}}_{\gamma^\star}\boldsymbol{x}=\boldsymbol{d} \tag{8}$$

のように凸最適化問題を定式化する．$\Psi$ が $\boldsymbol{\sigma}$ に特定の構造を誘導すれば，$\varphi$ に含まれる重み付き $\ell_2$ ノルムを通して $\boldsymbol{x}$ にその構造が反映される．




図–1 従来手法により得られた音声信号の T-F 表現．各列は左から順に DGT，$\ell_1$ 最小化 (式 (3))，以前提案した手法 (式 (4)，$\lambda=1$) により得られたもので，下段は白枠の拡大図である．色の範囲は 100 dB とした．

ここで，式 (8) の問題を解くことで，凸関数 $\Psi$ が $\boldsymbol{x}$ の振幅 $|\boldsymbol{x}|$ に対して間接的にペナルティを課せることを以下に示す．

**定理 1.** 各 $\boldsymbol{x}$ に対し，$\varphi(\boldsymbol{x},\boldsymbol{\sigma})+\Psi(\boldsymbol{\sigma})$ の最小値を与える $\boldsymbol{\sigma}$ を $\boldsymbol{\sigma}^\star_{\boldsymbol{x}}$ とする．このとき $\Psi=0$ ならば $\boldsymbol{\sigma}^\star_{\boldsymbol{x}}=|\boldsymbol{x}|$ を満たす．$\Psi\neq 0$ かつ $\boldsymbol{\sigma}^\star_{\boldsymbol{x}}=|\boldsymbol{x}|$ ならば $\boldsymbol{\sigma}^\star_{\boldsymbol{x}}$ は $\Psi(\boldsymbol{\sigma})$ を最小化する．

**証明.** ある固定の $\boldsymbol{x}$ に対し，$\boldsymbol{\sigma}^\star_{\boldsymbol{x}}$ は $\varphi(\boldsymbol{x},\boldsymbol{\sigma})+\Psi(\boldsymbol{\sigma})$ を最小化することから，以下の最適性条件を満たす．

$$\boldsymbol{0}\in\frac{|\boldsymbol{x}|^2}{2}\odot\left(-\frac{1}{(\boldsymbol{\sigma}^\star_{\boldsymbol{x}})^2}\right)+\frac{1}{2}+\partial\Psi(\boldsymbol{\sigma}^\star_{\boldsymbol{x}}) \tag{9}$$

ただし，$\odot$ は要素ごとの乗算，$\partial\Psi$ は $\Psi$ の劣微分を表す．$\Psi=0$ ならば $\partial\Psi(\boldsymbol{\sigma}^\star_{\boldsymbol{x}})=\{\boldsymbol{0}\}$ より，式 (9) が成立するためには

$$\frac{|\boldsymbol{x}|^2}{2}\odot\left(-\frac{1}{(\boldsymbol{\sigma}^\star_{\boldsymbol{x}})^2}\right)+\frac{1}{2}=\boldsymbol{0} \tag{10}$$

となればよい．したがって $\boldsymbol{\sigma}^\star_{\boldsymbol{x}}=|\boldsymbol{x}|$ が成立する．一方，$\Psi\neq 0$ ならば式 (9) に $\boldsymbol{\sigma}^\star_{\boldsymbol{x}}=|\boldsymbol{x}|$ を代入して $\boldsymbol{0}\in\partial\Psi(|\boldsymbol{x}|)$ を得る．したがって $\boldsymbol{\sigma}^\star_{\boldsymbol{x}}=|\boldsymbol{x}|$ は $\Psi(\boldsymbol{\sigma})$ を最小化する．■

定理 1 より $\Psi$ は $|\boldsymbol{x}|$ への間接的な凸ペナルティとみなすことができる．よって，凸性を保持したまま $\Psi$ の誘導する構造を振幅に反映することが可能となる．

### 3.3 凸最適化アルゴリズムの導出

式 (8) の具体的な形式として以下の式を考える．

$$\min_{(\boldsymbol{x},\boldsymbol{\sigma})\in\mathbb{C}^{MN}\times\mathbb{R}^{MN}}\ \varphi(\boldsymbol{x},\boldsymbol{\sigma})+\lambda\psi(\boldsymbol{B}\boldsymbol{\sigma})+\iota_{\mathcal{C}}(\boldsymbol{x}) \tag{11}$$

ただし，$\boldsymbol{B}\in\mathbb{C}^{J\times MN},\psi\in\Gamma_0(\mathbb{R}^J),\lambda>0$ であり，$\mathcal{C}=\{\boldsymbol{x}\in\mathbb{C}^{MN}\mid\boldsymbol{G}^{\mathrm{H}}_{\gamma^\star}\boldsymbol{x}=\boldsymbol{d}\}$ は信号の完全再構成条

件を満たすスペクトログラム全体の集合，$\iota_{\mathcal{C}}(\boldsymbol{x})$ は以下で定義される $\mathcal{C}$ の指示関数である．

$$\iota_{\mathcal{C}}(\boldsymbol{x}) = \begin{cases} 0 & (\boldsymbol{x} \in \mathcal{C}) \\ \infty & (\boldsymbol{x} \notin \mathcal{C}) \end{cases} \tag{12}$$

まず，ペナルティ関数 $\psi \circ \boldsymbol{B}$ の例をいくつか挙げる．

**例 1.** 式 (11) における $\psi \circ \boldsymbol{B}$ の誘導する構造

(i) スパース性
$\psi = \|\cdot\|_1, \boldsymbol{B} = \boldsymbol{I}$ とすると，$\psi \circ \boldsymbol{B}$ は振幅 $|\boldsymbol{x}|$ をスパースに誘導する．なお，式 (7) より $\varphi$ には $\|\boldsymbol{\sigma}\|_1$ が含まれるため，$\psi = 0$ としてもスパース性を誘導する．

(ii) 低ランク性
$\psi = \|\cdot\|_*, \boldsymbol{B} = \boldsymbol{I}$ とすると，$\psi \circ \boldsymbol{B}$ は $|\boldsymbol{x}|$ を低ランクに誘導する．ただし，$\|\cdot\|_*$ は核ノルムであり，$\boldsymbol{\sigma} \in \mathbb{R}^{MN}$ を $M \times N$ の行列に変形してから適用する．

(iii) 全変動
$\psi = \|\cdot\|_{2,1}, \boldsymbol{B} = \boldsymbol{D}$ とすると，$\psi \circ \boldsymbol{B}$ は $|\boldsymbol{x}|$ の滑らかさを誘導する．ただし，$\boldsymbol{D}$ は $M \times N$ 個の各要素における勾配の近似を返す差分作用素である．このとき，式 (8) は式 (4) に一致する．

(iv) 調波強調
$\psi = \|\cdot\|_{2,1}, \boldsymbol{B} = \boldsymbol{CD}$ とすると，$\psi \circ \boldsymbol{B}$ は $|\boldsymbol{x}|$ の調波構造を強調する．ただし，$\boldsymbol{C}$ は周波数方向の離散コサイン変換 (DCT) を表す．DCT 係数がスパースになると $|\boldsymbol{x}|$ の周期的な成分，すなわち調波成分が強調される．

次に，式 (11) の問題を解くためのアルゴリズムを導出する．$\boldsymbol{L} = [[\boldsymbol{I}\ \boldsymbol{O}]^{\mathrm{T}}, [\boldsymbol{O}\ \boldsymbol{B}]^{\mathrm{T}}]$, $\boldsymbol{y} = [\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{\sigma}^{\mathrm{T}}]^{\mathrm{T}}$, $f(\boldsymbol{y}) = \varphi(\boldsymbol{x}, \boldsymbol{\sigma})$, $g(\boldsymbol{L}\boldsymbol{y}) = \iota_{\mathcal{C}}(\boldsymbol{x}) + \lambda\psi(\boldsymbol{B}\boldsymbol{\sigma})$ と定義すると，式 (11) は $\min_{\boldsymbol{y}} f(\boldsymbol{y}) + g(\boldsymbol{L}\boldsymbol{y})$ と変形できる．これを Chambolle–Pock の主双対分離法 [8] により解くと Algorithm 1 が得られる．ただし，Algorithm 1 中の近接作用素 $\mathrm{prox}_{\tau\varphi}$ と $P_{\mathcal{C}}$ はそれぞれ以下のように求められる．

式 (5) より $\varphi$ は要素ごとの計算であるため，$\tau\varphi$ の近接作用素は要素ごとに計算でき，

$$\mathrm{prox}_{\tau\varphi}(\boldsymbol{x}, \boldsymbol{\sigma}) = (\mathrm{prox}_{\tau\phi}(x_k, \sigma_k))_{k=1}^{MN}. \tag{13}$$

とかける．各要素の近接作用素 $\mathrm{prox}_{\tau\phi}$ は

$$\mathrm{prox}_{\tau\phi}(x_k, \sigma_k) = \begin{cases} (0, 0) & (2\tau\sigma_k + |x_k|^2 \leq \tau^2) \\ \left(0, \sigma_k - \frac{\tau}{2}\right) & (x_k = 0 \text{ and } 2\sigma_k > \tau) \\ \left(x_k - \tau s \frac{x_k}{|x_k|}, \sigma_k + \tau \frac{s^2 - 1}{2}\right) & (\text{otherwise}) \end{cases} \tag{14}$$

**Algorithm 1** 提案手法のアルゴリズム

**Input:** $\tau > 0,\ \mu > 0,\ \rho^{[i]} \in (0, 2)\ (i = 0, 1, 2, \ldots)$,
$\boldsymbol{x}^{[0]} \in \mathbb{C}^{MN}, \boldsymbol{\sigma}^{[0]} \in \mathbb{R}^{MN}, \boldsymbol{u}^{[0]} \in \mathbb{C}^{MN}, \boldsymbol{v}^{[0]} \in \mathbb{R}^{J}$

**for** $i = 0, 1, 2, \ldots$ **do**

$(\boldsymbol{x}^{[i+\frac{1}{2}]}, \boldsymbol{\sigma}^{[i+\frac{1}{2}]}) = \mathrm{prox}_{\tau\varphi}(\boldsymbol{x}^{[i]} - \tau\boldsymbol{u}^{[i]}, \boldsymbol{\sigma}^{[i]} - \tau\boldsymbol{B}^{\mathrm{H}}\boldsymbol{v}^{[i]})$

$\tilde{\boldsymbol{u}}^{[i+\frac{1}{2}]} = \boldsymbol{u}^{[i]} + \mu(2\boldsymbol{x}^{[i+\frac{1}{2}]} - \boldsymbol{x}^{[i]})$

$\boldsymbol{u}^{[i+\frac{1}{2}]} = \tilde{\boldsymbol{u}}^{[i+\frac{1}{2}]} - \mu P_{\mathcal{C}}(\tilde{\boldsymbol{u}}^{[i+\frac{1}{2}]}/\mu)$

$\tilde{\boldsymbol{v}}^{[i+\frac{1}{2}]} = \boldsymbol{v}^{[i]} + \mu\boldsymbol{B}(2\boldsymbol{\sigma}^{[i+\frac{1}{2}]} - \boldsymbol{\sigma}^{[i]})$

$\boldsymbol{v}^{[i+\frac{1}{2}]} = \tilde{\boldsymbol{v}}^{[i+\frac{1}{2}]} - \mu\,\mathrm{prox}_{(\lambda/\mu)\psi}(\tilde{\boldsymbol{v}}^{[i+\frac{1}{2}]}/\mu)$

$\boldsymbol{x}^{[i+1]} = \boldsymbol{x}^{[i]} + \rho^{[i]}(\boldsymbol{x}^{[i+\frac{1}{2}]} - \boldsymbol{x}^{[i]})$

$\boldsymbol{\sigma}^{[i+1]} = \boldsymbol{\sigma}^{[i]} + \rho^{[i]}(\boldsymbol{\sigma}^{[i+\frac{1}{2}]} - \boldsymbol{\sigma}^{[i]})$

$\boldsymbol{u}^{[i+1]} = \boldsymbol{u}^{[i]} + \rho^{[i]}(\boldsymbol{u}^{[i+\frac{1}{2}]} - \boldsymbol{u}^{[i]})$

$\boldsymbol{v}^{[i+1]} = \boldsymbol{v}^{[i]} + \rho^{[i]}(\boldsymbol{v}^{[i+\frac{1}{2}]} - \boldsymbol{v}^{[i]})$

**end for**

で求められる [4]．ただし，$s > 0$ は 3 次方程式

$$s^3 + \left(\frac{2}{\tau}\sigma_k + 1\right)s - \frac{2}{\tau}|x_k| = 0 \tag{15}$$

の唯一の正の根であり，$p = \frac{2}{\tau}\sigma_k + 1,\ q = -\frac{2}{\tau}|x_k|$, $r = -\frac{p^3}{27} - \frac{q^2}{4}$ とするとカルダノの公式から

$$s = \begin{cases} \sqrt[3]{-\frac{q}{2} + \sqrt{-r}} + \sqrt[3]{-\frac{q}{2} - \sqrt{-r}} & (r < 0) \\ 2\sqrt[3]{-\frac{q}{2}} & (r = 0) \\ 2\sqrt[3]{\sqrt{\frac{q^2}{4} + r}}\cos\left(\frac{\arctan(-2\sqrt{r}/q)}{3}\right) & (r > 0) \end{cases} \tag{16}$$

で得られる [5]．また，$\mathcal{C} = \{\boldsymbol{x} \in \mathbb{C}^{MN} \mid \boldsymbol{G}_{\gamma^\star}^{\mathrm{H}}\boldsymbol{x} = \boldsymbol{d}\}$ への射影 $P_{\mathcal{C}}$ は，DGT のフレーム作用素の性質より

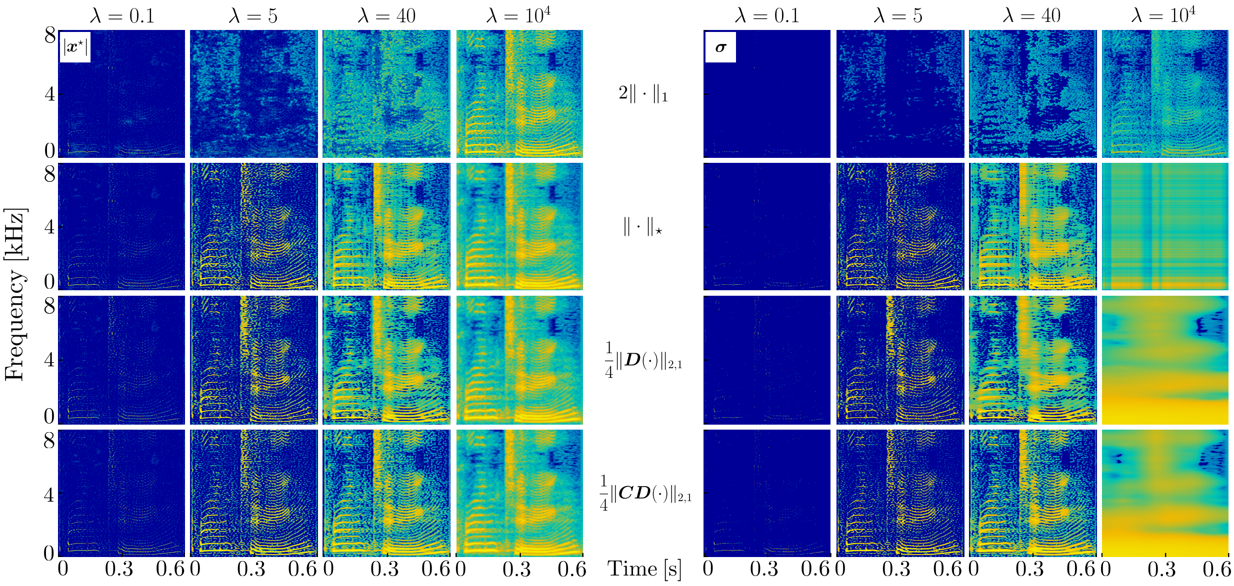
$$\begin{aligned} P_{\mathcal{C}}(\boldsymbol{x}) &= \boldsymbol{x} - \boldsymbol{G}_{\gamma^\star}(\boldsymbol{G}_{\gamma^\star}^{\mathrm{H}}\boldsymbol{G}_{\gamma^\star})^{-1}(\boldsymbol{G}_{\gamma^\star}^{\mathrm{H}}\boldsymbol{x} - \boldsymbol{d}) \\ &= \boldsymbol{x} - \boldsymbol{G}_w(\boldsymbol{G}_{\gamma^\star}^{\mathrm{H}}\boldsymbol{x} - \boldsymbol{d}) \end{aligned} \tag{17}$$

のように求められる．なお，$(\lambda/\mu)\psi$ の近接作用素はペナルティ関数 $\psi$ に依存する．

Algorithm 1 により求まる点列 $(\boldsymbol{x}^{[i]}, \boldsymbol{\sigma}^{[i]})_{i\in\mathbb{N}}$ は，$\tau\mu\|\boldsymbol{L}\|_{\mathrm{op}}^2 \leq 1$ かつ $\sum_{i\in\mathbb{N}}\rho^{[i]}(2 - \rho^{[i]}) = \infty$ を満たすとき，式 (11) で与えられる問題の大域的最適解に収束する [8]．ただし，$\|\boldsymbol{L}\|_{\mathrm{op}} = \max_{\boldsymbol{x}\neq\boldsymbol{0}}[\|\boldsymbol{L}\boldsymbol{x}\|_2/\|\boldsymbol{x}\|_2]$ は $\boldsymbol{L}$ の作用素ノルムで，$\boldsymbol{L}$ の最大特異値に一致する．

### 3.4 提案手法の性質に関する補足

式 (11) 中の正則化パラメータ $\lambda$ は，$\boldsymbol{\sigma}$ への正則化の強さを変える役割を担っている．しかし，$\lambda$ を大きくしても，$\boldsymbol{\sigma}$ に対して $\psi \circ \boldsymbol{B}$ で決まる構造を必ずしも強く誘導するとは限らない．これは，$\psi \circ \boldsymbol{B}$ の誘導する構造が，式 (11) に含まれる等式制約を満たさない可能性があることに起因する．例えば，例 1 の (i) と (ii) では $\lambda$ が大きいとき $\boldsymbol{\sigma} = \boldsymbol{0}$ が誘導されるが，等式制約により $\boldsymbol{x}$ は $\boldsymbol{0}$ になりえない．この場合，$\boldsymbol{\sigma}$ はむしろ定数に近づくため，$\varphi$ に含まれる重み付きノルムを通しても $\boldsymbol{x}$ には構造が反映されず，$\psi \circ \boldsymbol{B}$ の誘導する構造は失われる可能性がある．




図–2　提案手法により得られた音声信号の T-F 表現 $|\boldsymbol{x}^{\star}|$ (左) とそれに対応する補助変数 $\boldsymbol{\sigma}$ (右)．各行は上から順に $\psi(\boldsymbol{B}(\cdot)) = 2\|\cdot\|_1, \|\cdot\|_*, \frac{1}{4}\|\boldsymbol{D}(\cdot)\|_{2,1}, \frac{1}{4}\|\boldsymbol{C}\boldsymbol{D}(\cdot)\|_{2,1}$ に対応しており，係数は図の視認性のために調整した．また，各列は左から順に $\lambda = 0.1, 5, 40, 10^4$ に対応する．色の範囲は 100 dB とした．

$\boldsymbol{\sigma}$ をある正の値 $\tilde{\boldsymbol{\sigma}} > \boldsymbol{0}$ に固定すると，式 (11) は $\boldsymbol{x}$ のみに関する凸最適化問題

$$\min_{\boldsymbol{x}\in\mathbb{C}^{MN}} \boldsymbol{x}^{\mathrm{H}}\tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{x} + \iota_{\mathcal{C}}(\boldsymbol{x}) \quad (\tilde{\boldsymbol{\Sigma}} = \mathrm{diag}(\tilde{\boldsymbol{\sigma}})) \tag{18}$$

に帰着し，その解は $\boldsymbol{x}^{\star}_{\tilde{\boldsymbol{\sigma}}} = \tilde{\boldsymbol{\Sigma}}\boldsymbol{G}_{\gamma^{\star}}(\boldsymbol{G}^{\mathrm{H}}_{\gamma^{\star}}\tilde{\boldsymbol{\Sigma}}\boldsymbol{G}_{\gamma^{\star}})^{-1}\boldsymbol{d}$ となる．$\boldsymbol{\sigma}$ が正の定数 $c\boldsymbol{1}$ $(c > 0)$ に固定されていれば，$\boldsymbol{x}^{\star}_{\tilde{\boldsymbol{\sigma}}}$ は最小ノルム解 $\boldsymbol{G}_w\boldsymbol{d}$ になるため，$\lambda$ を過剰に大きくすると $\boldsymbol{G}_w\boldsymbol{d}$ に近い解 $\boldsymbol{x}^{\star}$ が得られる．しかし，$\lambda$ と $\psi\circ\boldsymbol{B}$ の選び方によらず $\varphi$ の第 2 項 $(\|\boldsymbol{\sigma}\|_1/2)$ がスパース性を誘導するため，実際には $\boldsymbol{x}^{\star} = \boldsymbol{G}_w\boldsymbol{d}$ は成立しないことに注意する．

## 4　数値例

提案手法 (式 (11)) においてペナルティ関数 $\psi\circ\boldsymbol{B}$ を変えたときの，得られた音声信号 $\boldsymbol{d}$ の T-F 表現を比較した．窓関数 $\boldsymbol{w}$ として窓長が $2^9$ のハン窓を用い，シフト幅 $a = 2^6$，周波数ビン数 $M = 2^{12}$ とした．Algorithm 1 の反復回数は 5000 回とし，$\tau = 1/2$, $\mu = 1/5$, $\rho^{[i]} = 1.99$, $\boldsymbol{x}^{[0]} = \boldsymbol{G}_w\boldsymbol{d}$, $\boldsymbol{\sigma}^{[0]} = |\boldsymbol{G}_w\boldsymbol{d}|$, $\boldsymbol{u}^{[0]} = \boldsymbol{0}$, $\boldsymbol{v}^{[0]} = \boldsymbol{0}$ とした．ペナルティ関数 $\psi\circ\boldsymbol{B}$ については例 1 に示した 4 種類を用いた．

得られた T-F 表現 $|\boldsymbol{x}|$ とそれに対応する $\boldsymbol{\sigma}$ を図–2 に示す．$\lambda = 0$ のとき，式 (7) より $\varphi$ に $\|\boldsymbol{\sigma}\|_1$ が含まれることから，提案手法は $\ell_1$ ノルム最小化 (式 (3)) と一致する．そのため，$\lambda$ が小さいときはペナルティ関数によらずスパースな T-F 表現が得られた．また，$\lambda$ を大きくしていくと，ペナルティ関数 $\psi\circ\boldsymbol{B}$ の誘導する構造が振幅に反映された T-F 表現が得られた．図–2 を見ると，$\lambda = 5, 40$ のときは $\boldsymbol{\sigma}$ が $|\boldsymbol{x}|$ と類似していることがわかる．これより，定理 1 で示したように $\psi\circ\boldsymbol{B}$ は振幅への間接的なペナルティを課しているといえる．一方，$\lambda$ を過剰に大きくすると，等式制約の影響で $\psi\circ\boldsymbol{B}$ が $|\boldsymbol{x}|$ に与える効果は歪められ，ペナルティ関数によらず同様の T-F 表現が得られた．

## 5　むすび

本稿では，振幅を柔軟に調節できる T-F 表現を凸最適化により実現する手法を提案した．今後は提案手法の性質のさらなる調査や実応用への適用に取り組む．

### 参考文献


[1] G. E. Pfander and H. Rauhut, “Sparsity in time-frequency representations,” *J. Fourier Anal. Appl.*, vol. 16, no. 2, pp. 233–260, 2010.

[2] H. Sawada, N. Ono, H. Kameoka, D. Kitamura, and H. Saruwatari, “A review of blind source separation methods: Two converging routes to ILRMA originating from ICA and NMF,” *APSIPA Trans. Signal Inf. Process.*, vol. 8, no. 1, e12, 2019.

[3] H. Tachibana, N. Ono, H. Kameoka, and S. Sagayama, “Harmonic/percussive sound separation based on anisotropic smoothness of spectrograms,” *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 22, no. 12, pp. 2059–2073, 2014.

[4] P. L. Combettes and C. L. Müller, “Perspective maximum likelihood type estimation via proximal decomposition,” *Electron. J. Stat.*, vol. 14, no. 1, pp. 207–238, 2020.

[5] H. Kuroda and D. Kitahara, “Block-sparse recovery with optimal block partition,” *IEEE Trans. Signal Process.*, vol. 70, pp. 1506–1520, 2022.

[6] 新井慶大，山田宏樹，矢田部浩平，“音響信号のグループ構造を自動適応させるスパース時間周波数解析,” 日本音響学会講演論文集, pp. 167–168, Sep. 2023.

[7] K. Arai, K. Yamada, and K. Yatabe, “Versatile Time-Frequency Representations Realized by Convex Penalty on Magnitude Spectrogram,” *IEEE Signal Process. Lett.*, vol. 30, pp. 1082–1086, 2023.

[8] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, “Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists,” *SIAM Rev.*, vol. 65, no. 2, pp. 375–435, 2023.