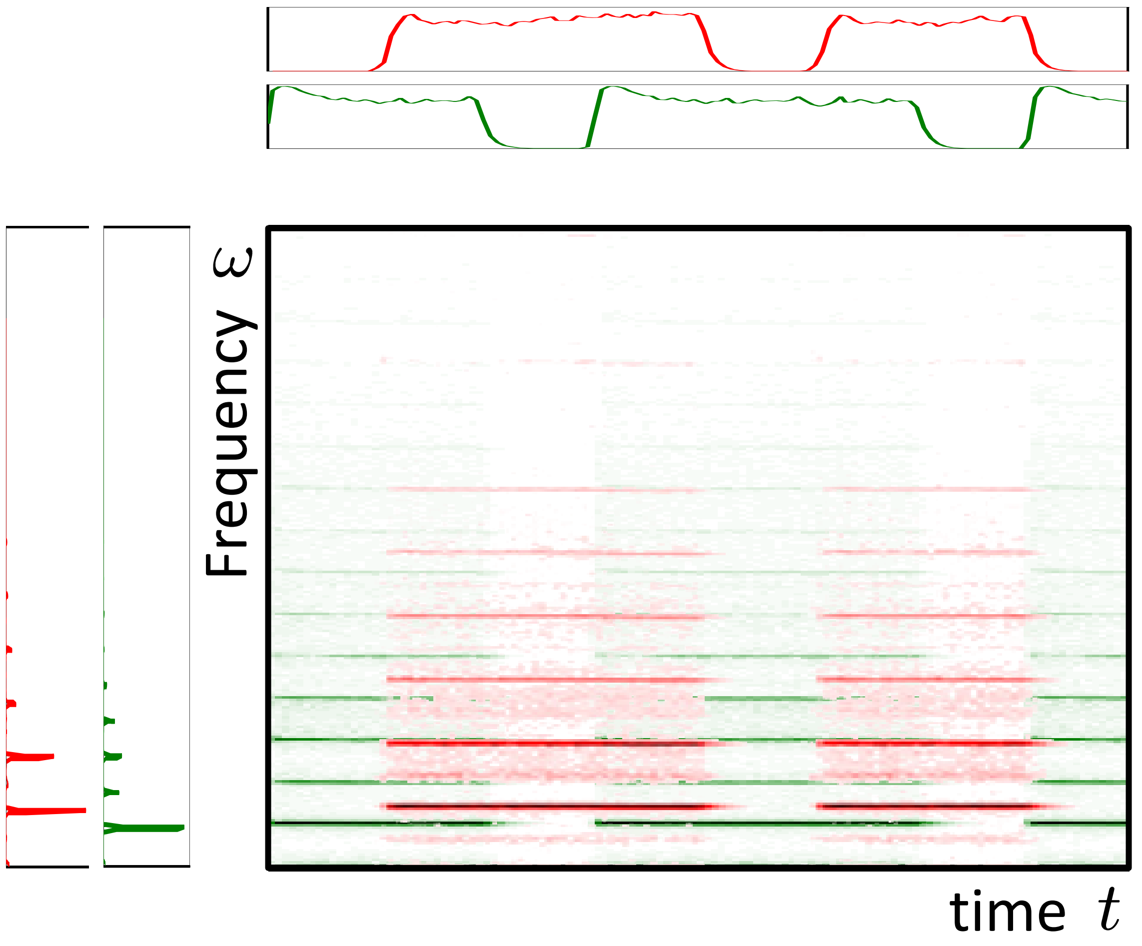
「第3回若手研究者フォーラム：音声言語処理について語ろう」
2011年10月8日

非負値行列因子分解とその発展形

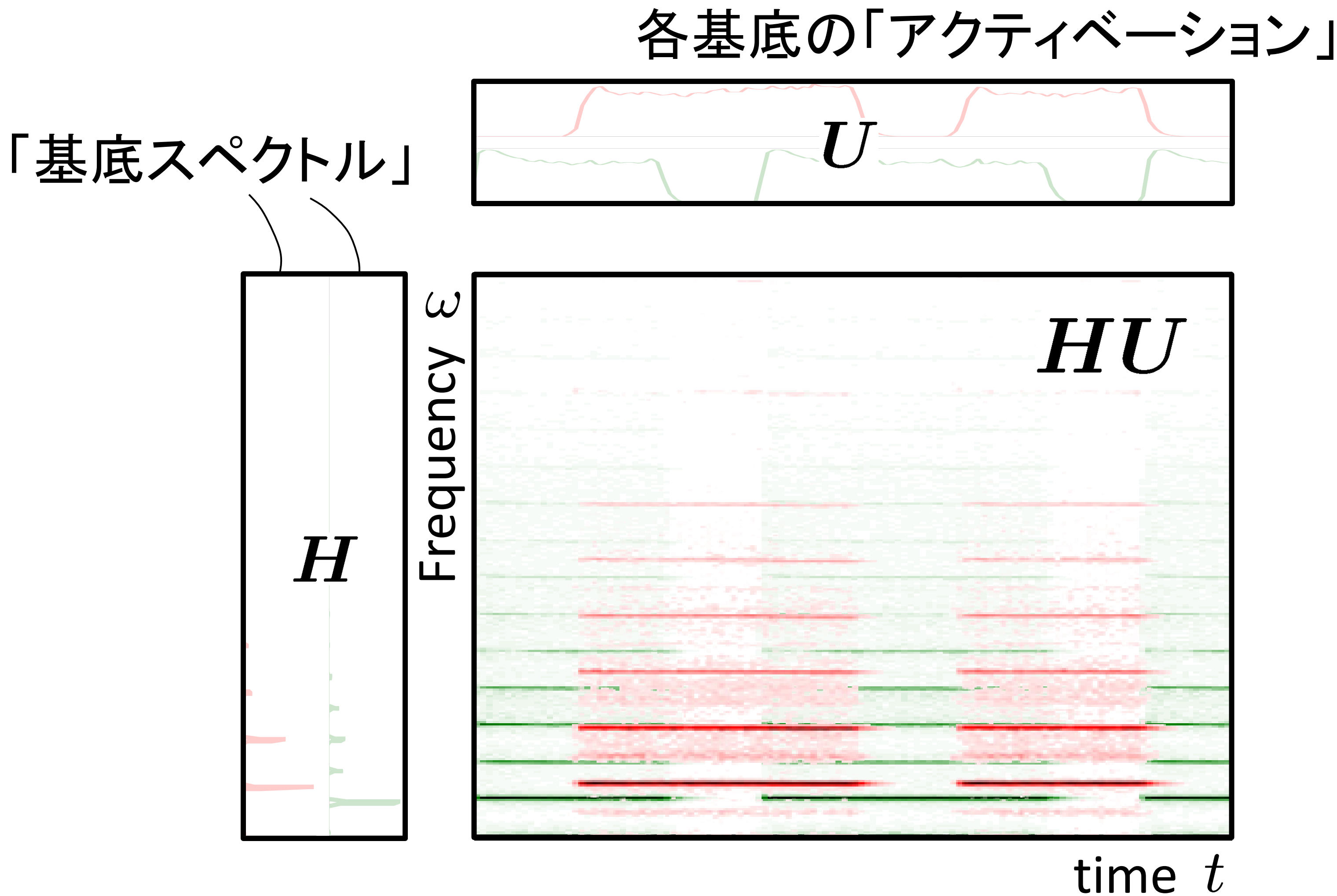
亀岡弘和

東京大学 大学院情報理工学系研究科
kameoka@hil.t.u-tokyo.ac.jp
NTTコミュニケーション科学基礎研究所
kameoka.hirokazu@lab.ntt.co.jp

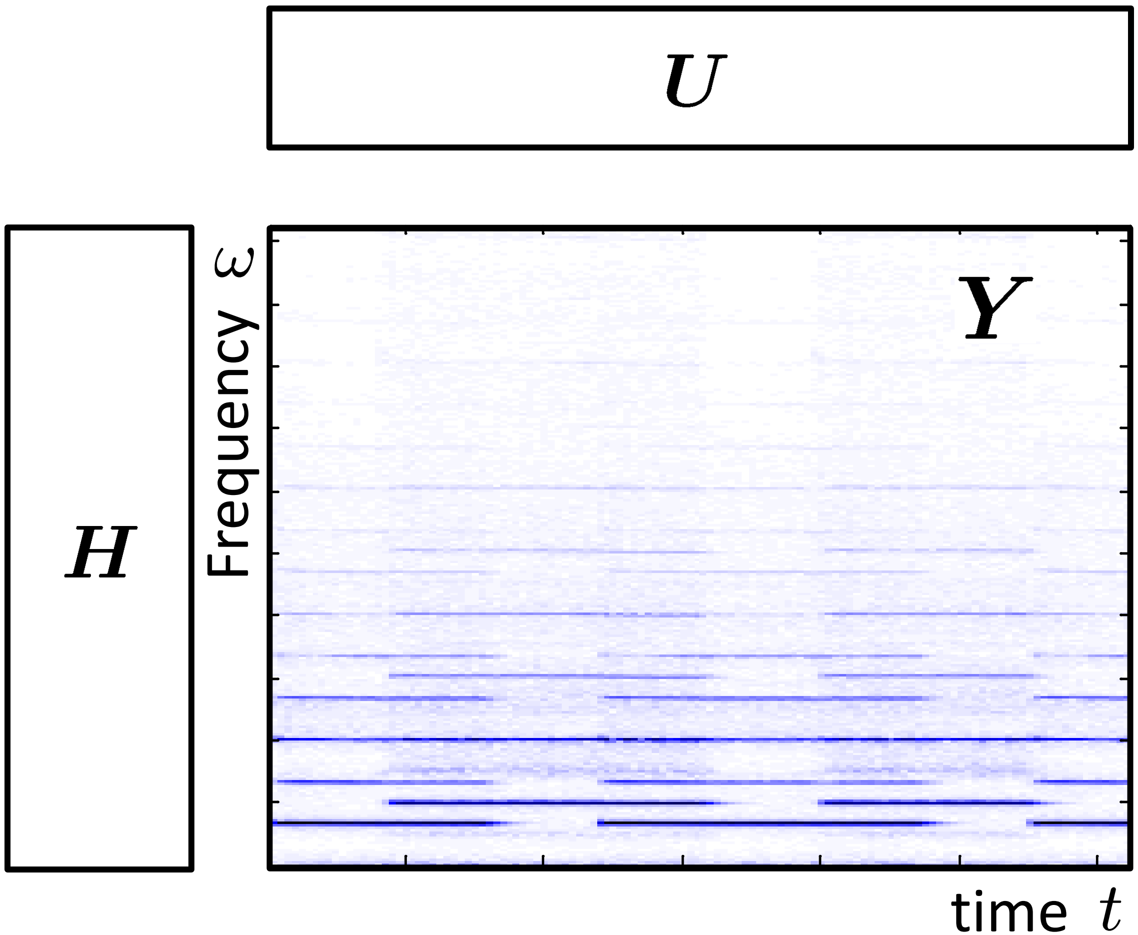
「行列の積」としてのスペクトログラム



「行列の積」としてのスペクトログラム



「行列の積」としてのスペクトログラム



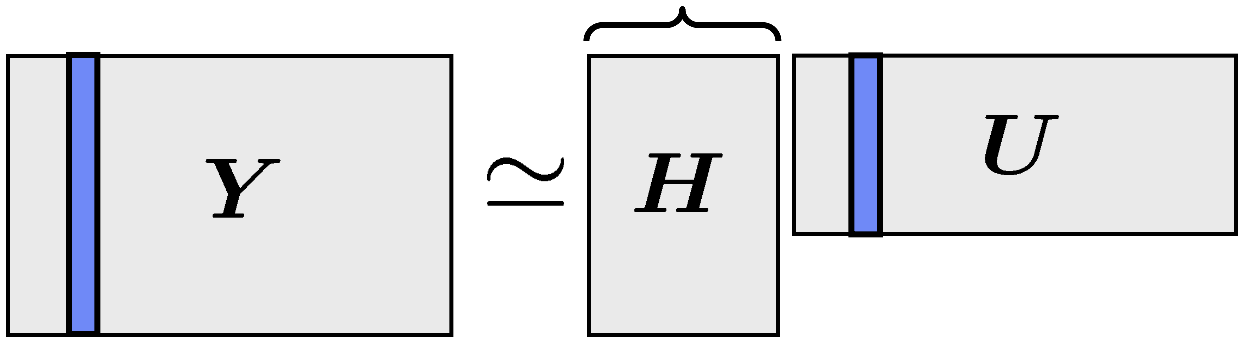
目次

1. 非負値行列因子分解（NMF）とは
   - 何に使えるのか（音響信号処理を題材として）
   - どのような性質があるのか
   - どのように求めるのか
   - 統計モデルとしての解釈

2. NMFの改良・拡張モデル
   - スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーション連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

非負値行列因子分解 (Nonnegative Matrix Factorization) とは

- 非負値行列を2つの非負値行列の積で表現



- 行列因子分解の応用場面
  - ブラインド信号分離(Blind Signal Separation)

    $\boldsymbol{H}$: 混合行列　$\boldsymbol{U}$ の各行: 源信号
  - 次元圧縮

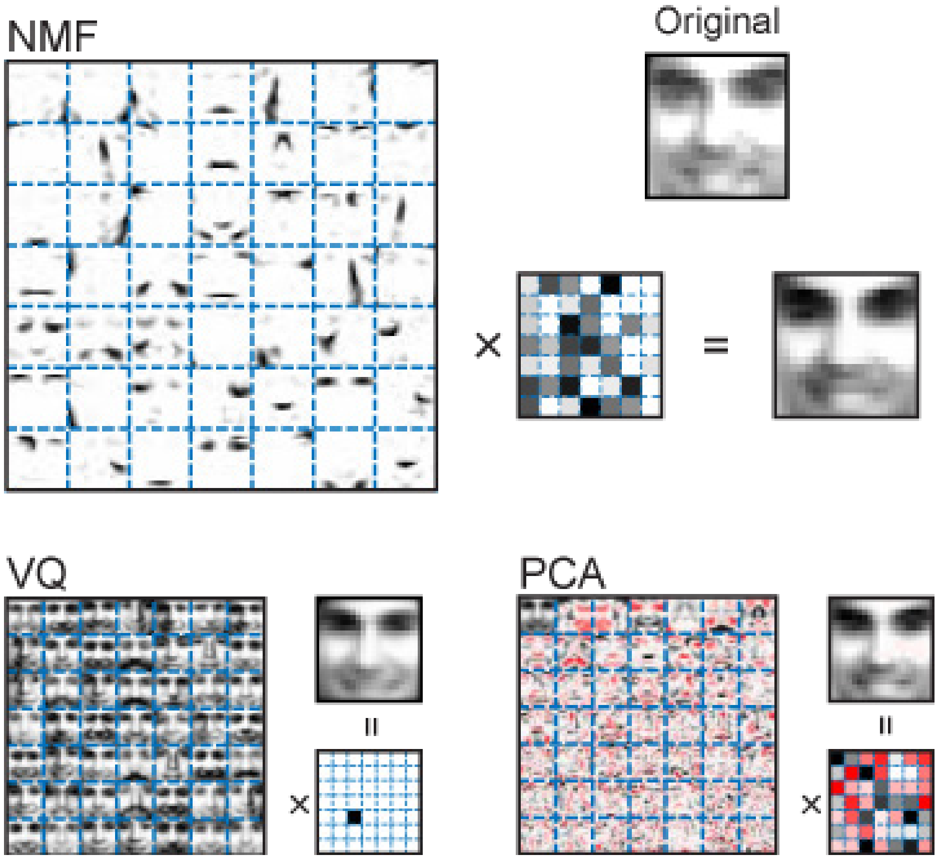
    $\boldsymbol{H}$ の各列: 基底ベクトル　$\boldsymbol{U}$ の各列: 結合係数

NMFが生まれた背景

- 画像処理分野で生まれた技術 [Lee1999]
  - 顔画像から顔パーツを抽出するのが目的
  - 概念自体は90年代前半に登場 [Paatero1994]
- 音のスペクトルを画像と見なして適用（後述）
  →音声分離・自動採譜等
  [Smaragdis2003]以降極めて多数…
- 効率的な反復アルゴリズム
  [Lee2000]

$$H_{\omega,k} \leftarrow H_{\omega,k} \frac{\left[\boldsymbol{Y}\boldsymbol{U}^{\mathrm{T}}\right]_{\omega,k}}{\left[\boldsymbol{H}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\right]_{\omega,k}}$$

$$U_{k,t} \leftarrow U_{k,t} \frac{\left[\boldsymbol{H}^{\mathrm{T}}\boldsymbol{Y}\right]_{k,t}}{\left[\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{U}\right]_{k,t}}$$



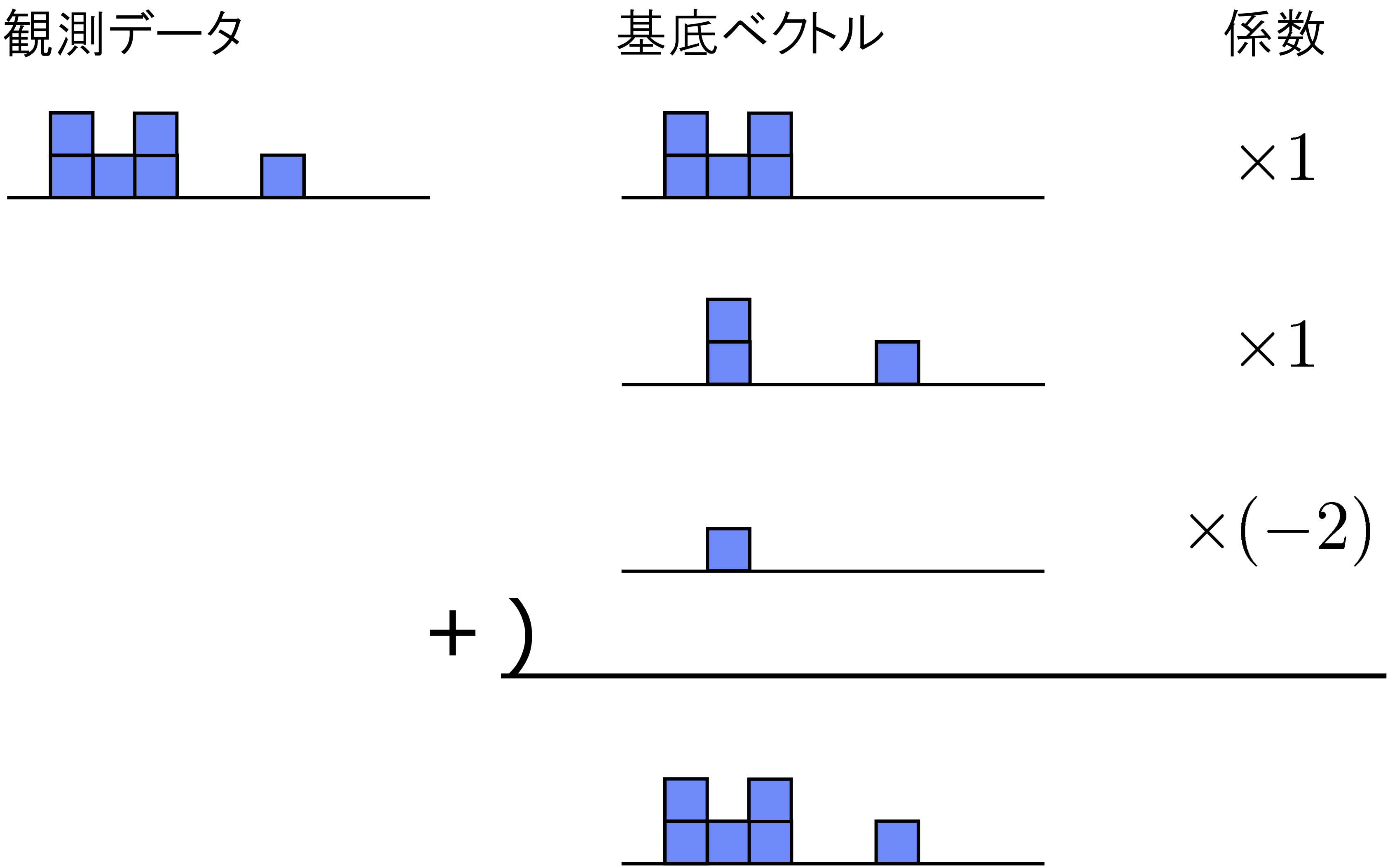
なぜ非負値なのか？その意図は？

- データ行列の非負性
  - 実世界には非負値データが多い
    （例）パワースペクトル，画素値，度数，
- 基底行列の非負性
  - 「非負値データの構成要素もまた非負値データであるべき(でないと物理的に意味をなさない！)」という考え方

  （例）負の値をもったパワースペクトルなんて解釈のしようがない
- 係数行列の非負性
  - 構成要素の混ざり方は「足し算」のみ
  - 係数行列をスパースに誘導 → 基底の情報量をアップ

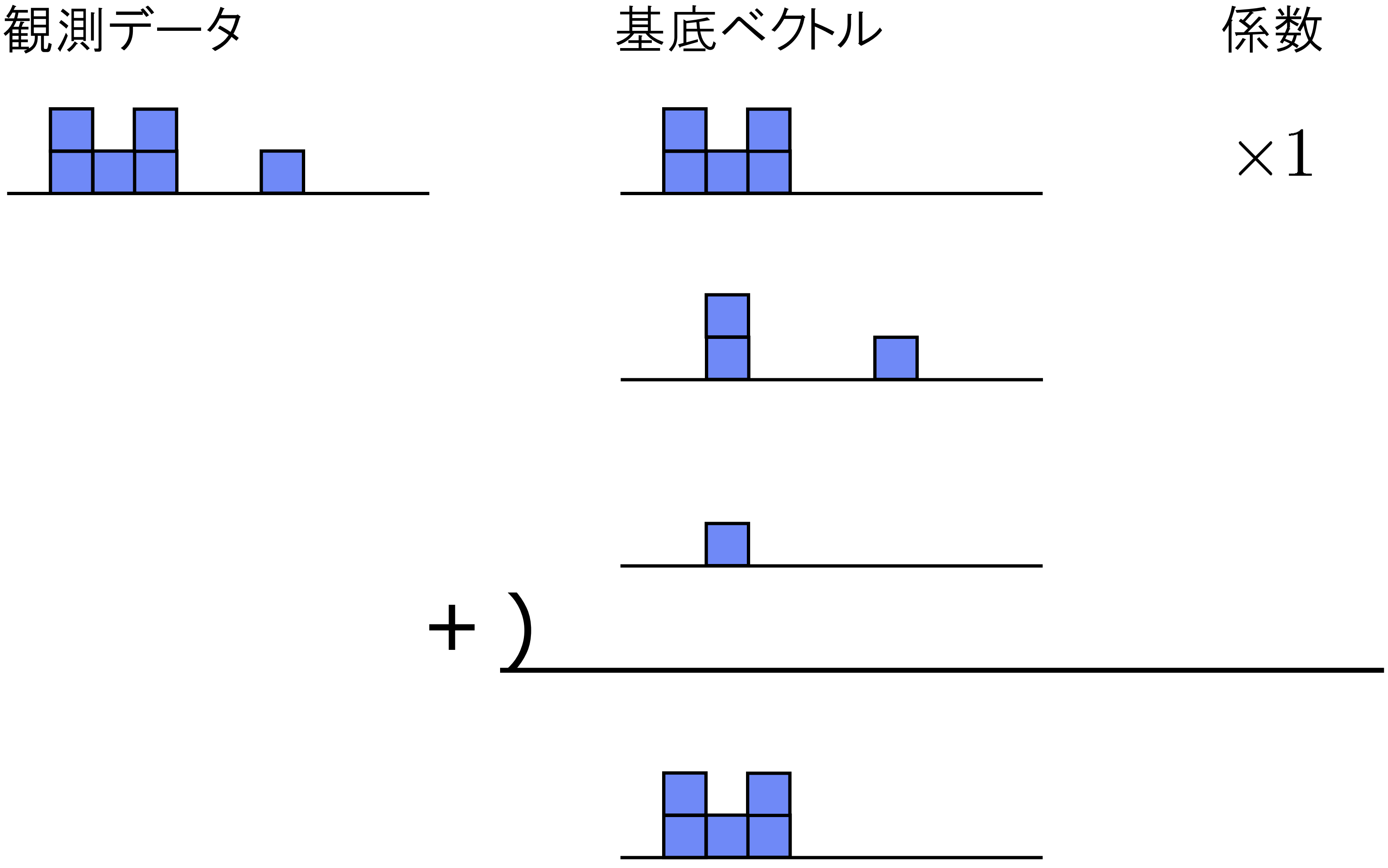
係数行列の非負性について

- 係数が負の値を取っても良い場合



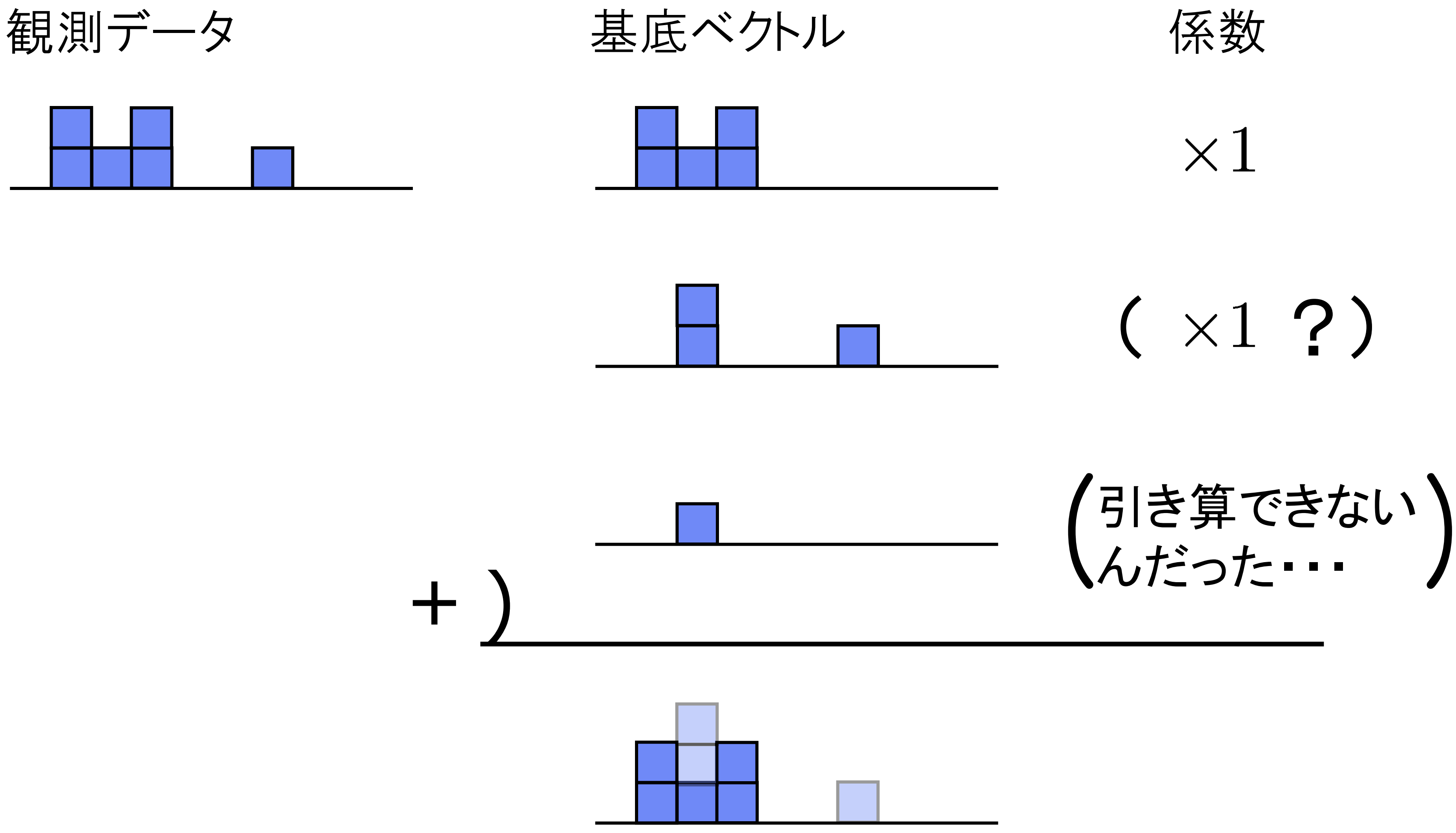
係数行列の非負性について

- 係数が負の値を取ってはならない場合



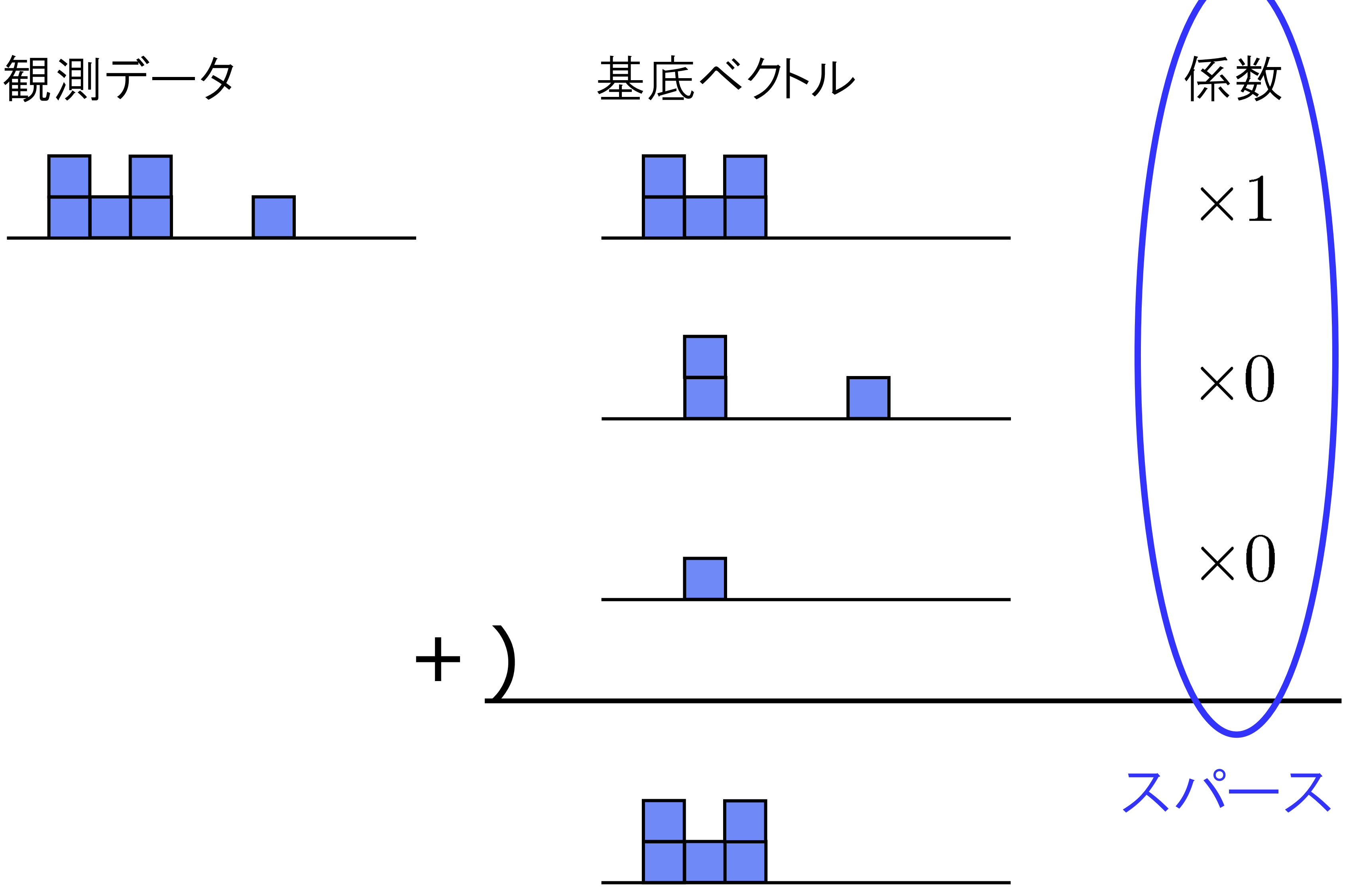
係数行列の非負性について

- 係数が負の値を取ってはならない場合



係数行列の非負性について

- 係数が負の値を取ってはならない場合

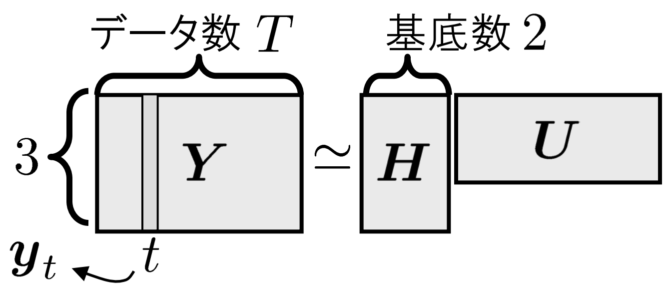


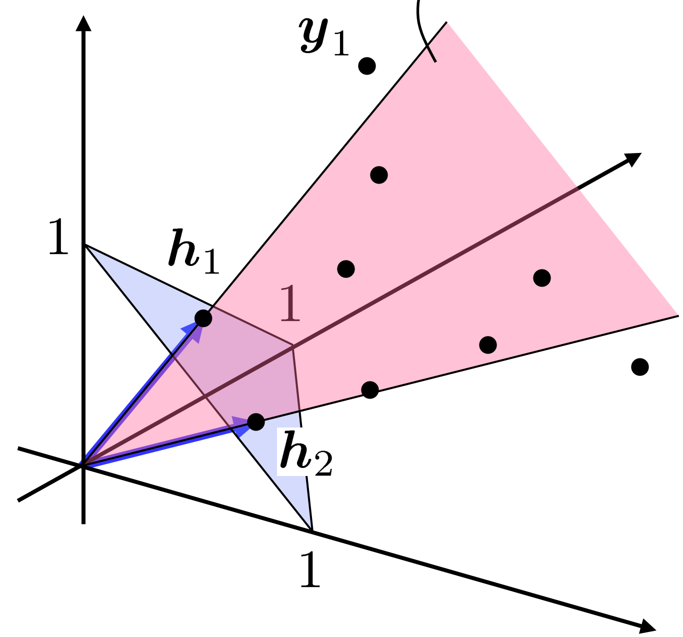
別の見方から…

- Frobeniusノルム規準のNMF

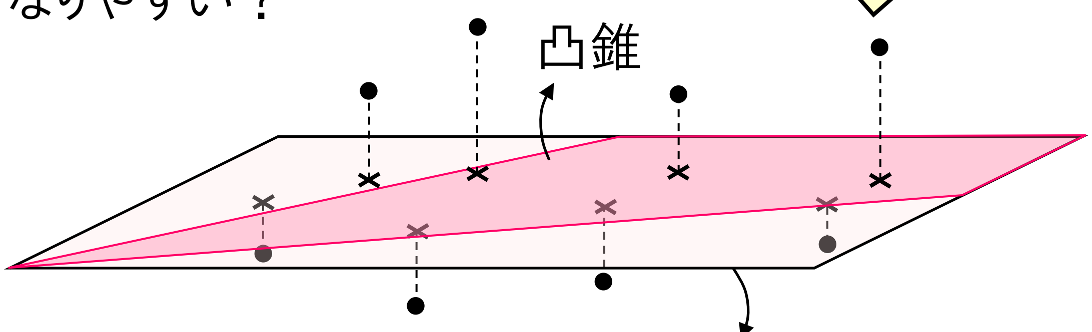
$$\text{minimize} \quad \|\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{U}\|_F^2 = \sum_t \left\| \boldsymbol{y}_t - \sum_k \boldsymbol{h}_k U_{k,t} \right\|_2^2$$

$$\text{subject to} \quad H_{\omega,k} \geq 0,\ U_{k,t} \geq 0$$





斜交基底だから係数はスパースになりやすい？



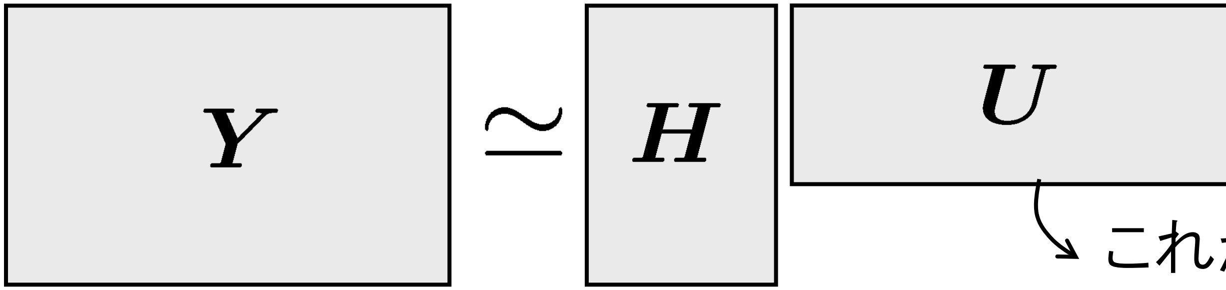
$\boldsymbol{h}_1$ と $\boldsymbol{h}_2$ が張る部分空間

$\boldsymbol{h}_1$ と $\boldsymbol{h}_2$ のなす角を大きくした方がお得！

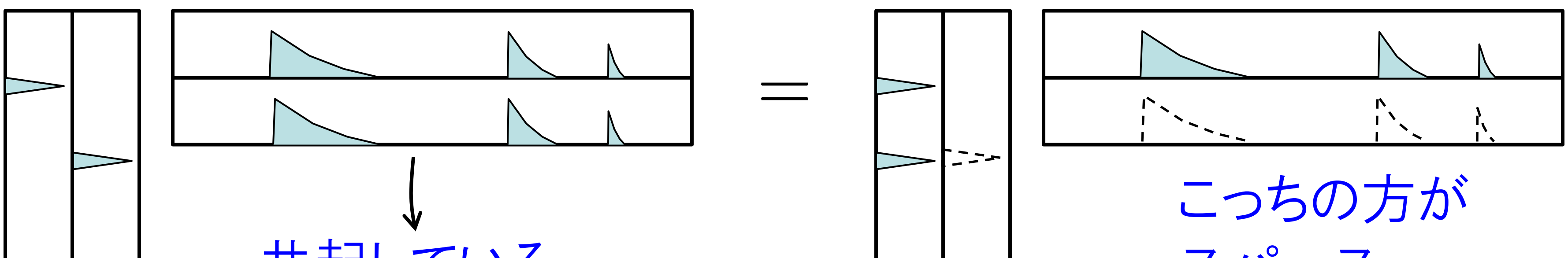
↓

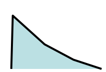
陽に制約を入れていないのに基底が直交化される傾向に

スパース性がもたらす効果



- Hがどうなれば Uはスパースになる？
  - 共起するパーツがあればひとまとめにした方がスパースに!



  - 共起頻度が高いパーツのペアをひとまとめにした方がスパースに!
    （上のひとまとめ操作では  を3つ分減らせている）

頻出するひとまとまりのパーツが
各基底ベクトルとして得られる傾向になる

NMFで音声スペクトログラムを分解してみる

- 音響信号 $y(u) \in \mathbb{R}$

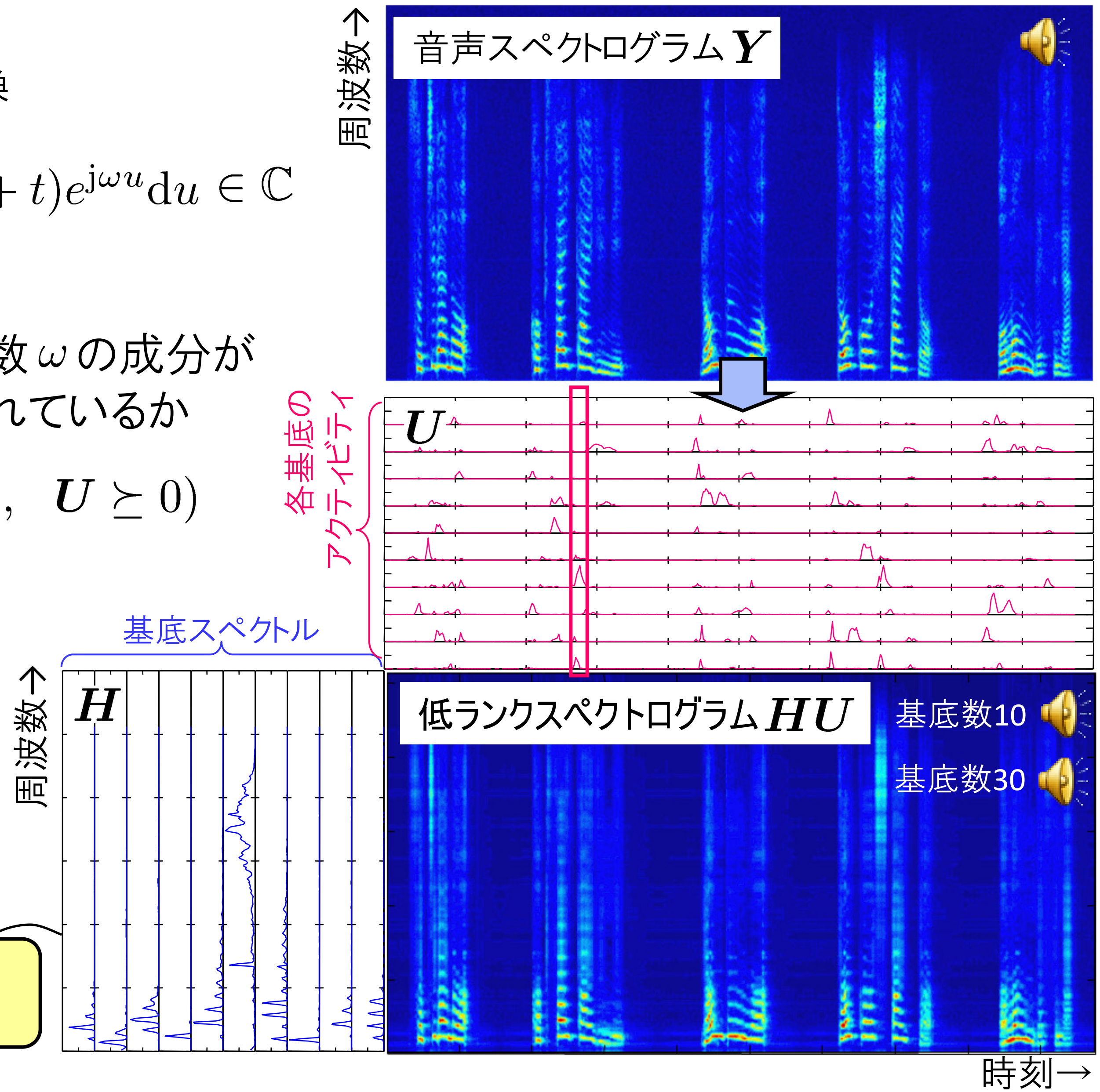
短時間フーリエ変換
（時間周波数分解）

$$Y_{\omega,t} = \int W(u) y(u+t) e^{\mathrm{j}\omega u} \mathrm{d}u \in \mathbb{C}$$

絶対値をとる

$|Y_{\omega,t}|$：時刻 t に周波数 ω の成分がどれほど含まれているか

$$\boldsymbol{Y} \simeq \boldsymbol{H}\boldsymbol{U} \quad (\boldsymbol{H} \succeq 0, \quad \boldsymbol{U} \succeq 0)$$

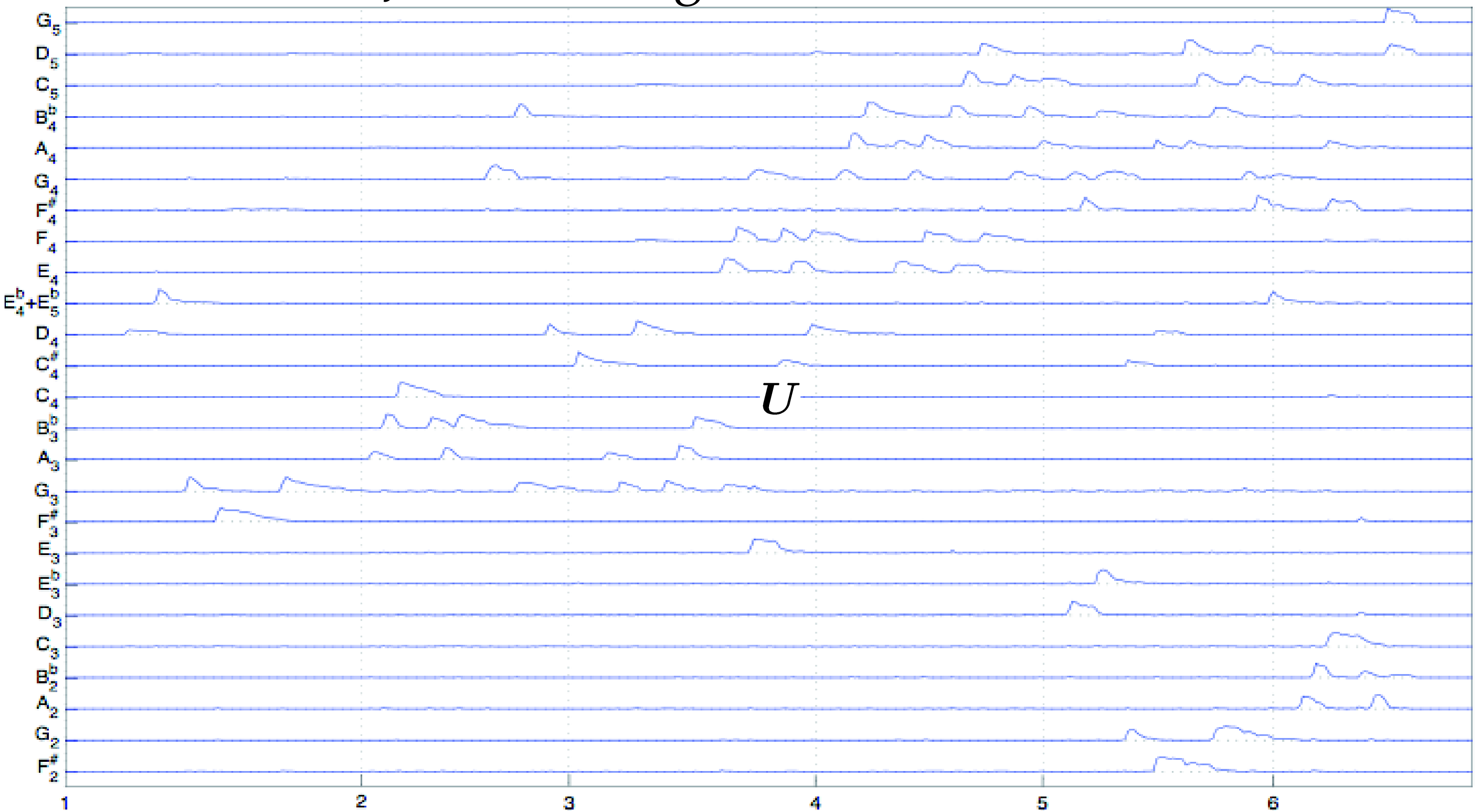


頻出するスペクトルパーツが獲得される

何に使えるのか？（1/3）

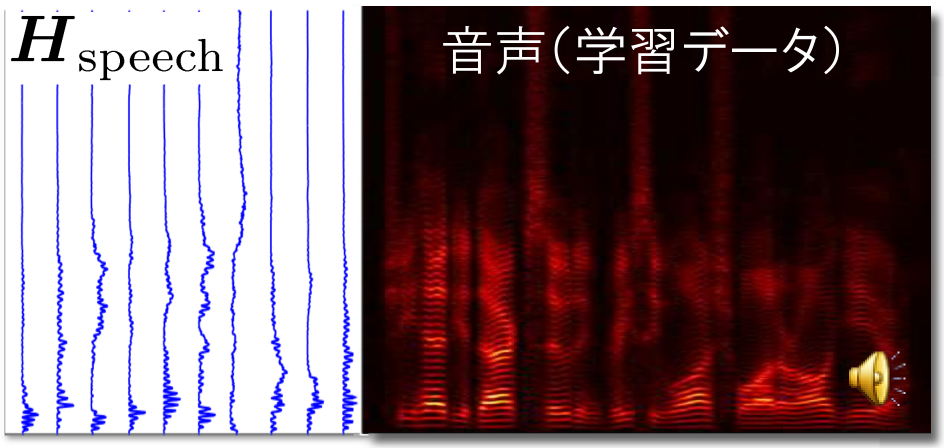
- 自動採譜 [P. Smaragdis et al., 2003]

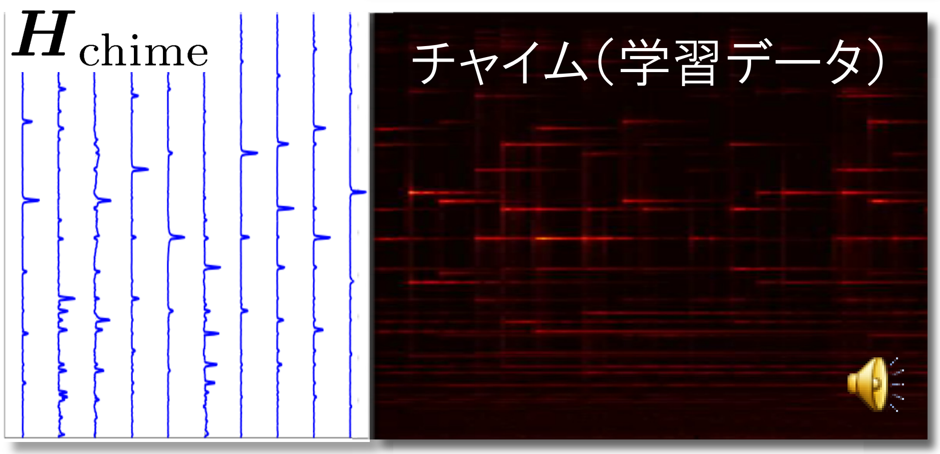
J.S. Bach: Fuge #16 in G minor



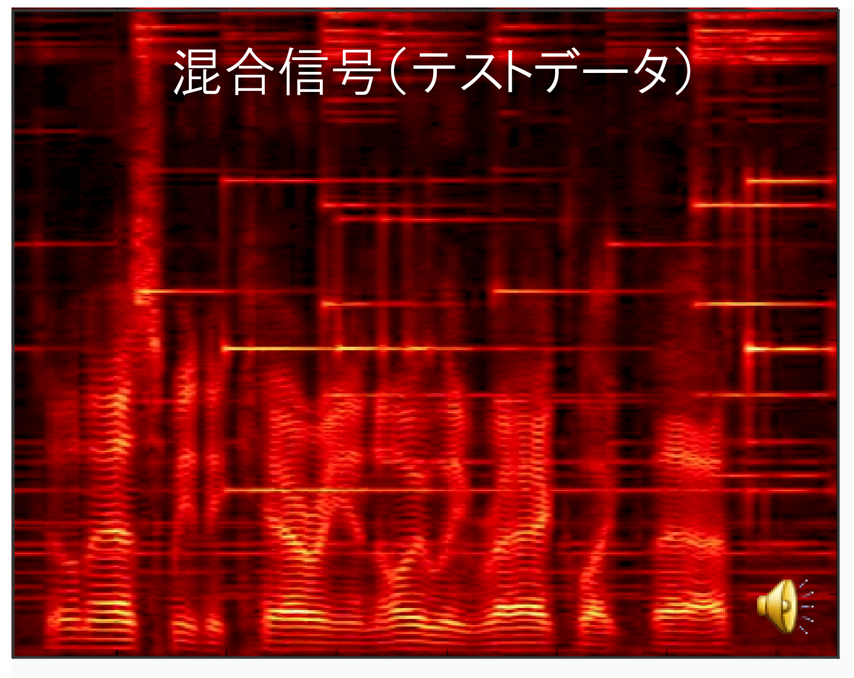
何に使えるのか？（2/3）

- 教師ありモノラル音源分離 [P. Smaragdis et al., 2007]





Y

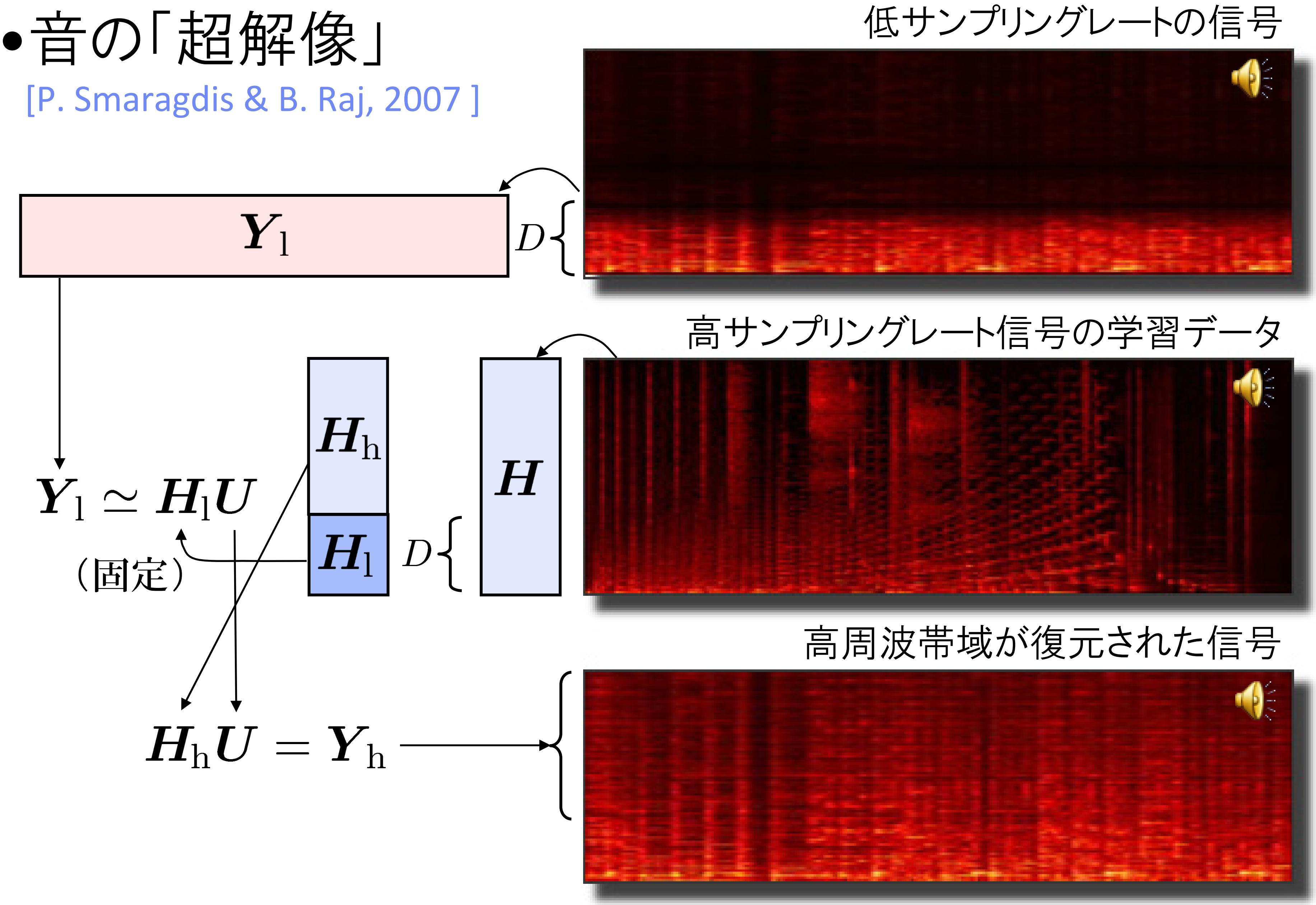


$$Y \simeq \underbrace{\begin{bmatrix} H_{\text{speech}} & H_{\text{chime}} \end{bmatrix}}_{\text{固定}} \begin{bmatrix} U_{\text{speech}} \\ U_{\text{chime}} \end{bmatrix}$$

何に使えるのか？（3/3）

- 音の「超解像」

[P. Smaragdis & B. Raj, 2007]



NMFの基本問題

- T 個の観測データ(非負値ベクトル) $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_T$
- K 個の非負値基底ベクトル$\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K$ の非負結合でどの観測データも良く表現できる基底セットを求めたい

$$\boldsymbol{y}_t \simeq \sum_{k=1}^{K} \boldsymbol{h}_k U_{k,t}, \ U_{k,t} \geq 0 \quad (t = 1, \ldots, T)$$

$$\left.\begin{aligned} \boldsymbol{Y} &:= (\boldsymbol{y}_1, \cdots, \boldsymbol{y}_T) \\ \boldsymbol{H} &:= (\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K) \\ \boldsymbol{U} &:= (U_{k,t})_{K \times T} \end{aligned}\right\} \quad \boldsymbol{Y} \simeq \boldsymbol{H}\boldsymbol{U} \qquad H_{\omega,k} \geq 0, \ U_{k,t} \geq 0$$

$\boldsymbol{Y}$ と $\boldsymbol{HU}$ の乖離度を表す規準

$$\begin{array}{ll} \text{minimize} & \mathcal{D}(\boldsymbol{Y}|\boldsymbol{HU}) \\ \text{subject to} & H_{\omega,k} \geq 0, \ U_{k,t} \geq 0 \end{array}$$

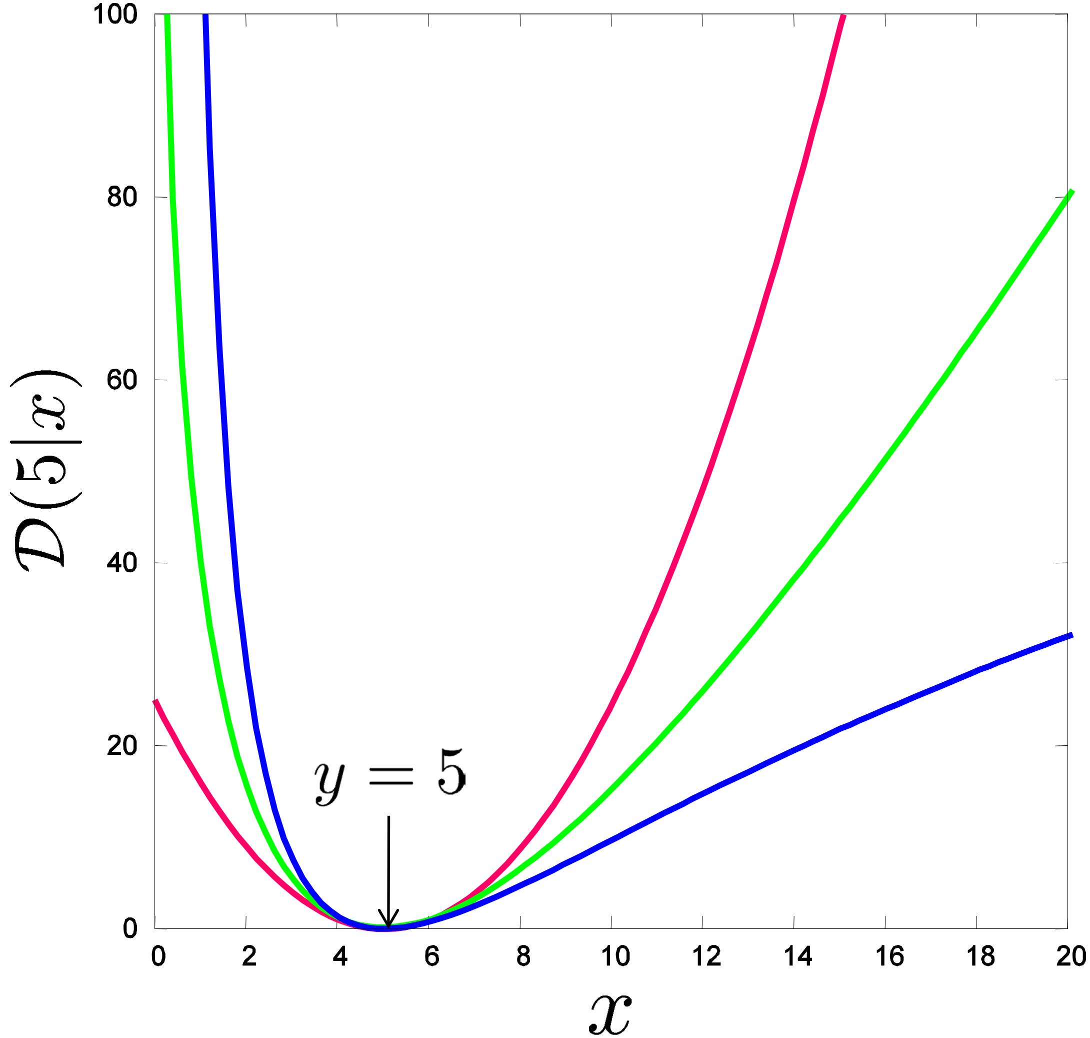
どうやって求めるのか

- NMFにおける代表的な最適化規準
  - Frobeniusノルム（二乗誤差）
  - Iダイバージェンス（一般化KLダイバージェンス）

- 押さえておくべき基本原理
  - 補助関数法
  - 凸不等式（Jensenの不等式）

二乗誤差とＩダイバージェンス

- $\mathcal{D}(y|x)$： x の y からの近さの度合い



二乗誤差

$$\mathcal{D}(x|y) = (y-x)^2$$

Ｉダイバージェンス

$$\mathcal{D}(x|y) = y\log\frac{y}{x} - (y-x)$$

板倉齋藤距離

$$\mathcal{D}(x|y) = \frac{y}{x} - \log\frac{y}{x} - 1$$

NMFにおける代表的な最適化規準

- Frobeniusノルム規準

なんとかしたい部分

$$\begin{aligned}&\text{minimize} \quad \|\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{U}\|_F^2 = \sum_{\omega,t}\left|Y_{\omega,t}-\sum_k H_{\omega,k}U_{k,t}\right|^2 \\ &\text{subject to} \quad H_{\omega,k}\geq 0,\ U_{k,t}\geq 0\end{aligned}$$

- Iダイバージェンス（一般化KLダイバージェンス）

$$\begin{aligned}&\text{minimize} \quad \sum_{\omega,t}\left(Y_{\omega,t}\log\frac{Y_{\omega,t}}{\sum_k H_{\omega,k}U_{k,t}}-Y_{\omega,t}+\sum_k H_{\omega,k}U_{k,t}\right) \\ &\text{subject to} \quad H_{\omega,k}\geq 0,\ U_{k,t}\geq 0\end{aligned}$$

なんとかしたい部分

いずれも $\boldsymbol{Y}=\boldsymbol{H}\boldsymbol{U}$ のとき 0 になる

押さえておくべき基本原理(1/2)

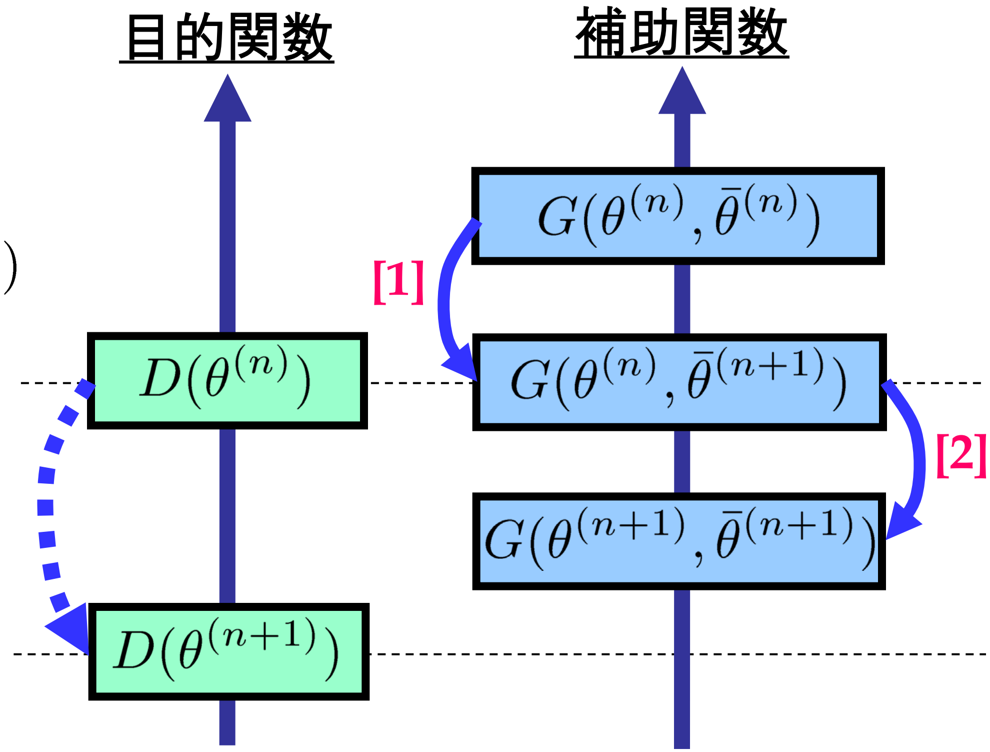
- 補助関数法
  - $D(\theta) = \min_{\bar{\theta}} G(\theta, \bar{\theta})$ を満たす $G(\theta, \bar{\theta})$を補助関数と定義
  - 反復アルゴリズム

[1] $\bar{\theta}^{(n+1)} = \underset{\bar{\theta}}{\operatorname{argmin}}\, G(\theta^{(n)}, \bar{\theta})$

[2] $\theta^{(n+1)} = \underset{\theta}{\operatorname{argmin}}\, G(\theta, \bar{\theta}^{(n+1)})$

  - 収束性

$$
\begin{aligned}
D(\theta^{(n)}) &= G(\theta^{(n)}, \bar{\theta}^{(n+1)}) \\
&\geq G(\theta^{(n+1)}, \bar{\theta}^{(n+1)}) \\
&\geq D(\theta^{(n+1)})
\end{aligned}
$$



目的関数を直接最小化するのが難しいなら、
とりあえずその上限関数を作ってみよう！

押さえておくべき基本原理（2/2）

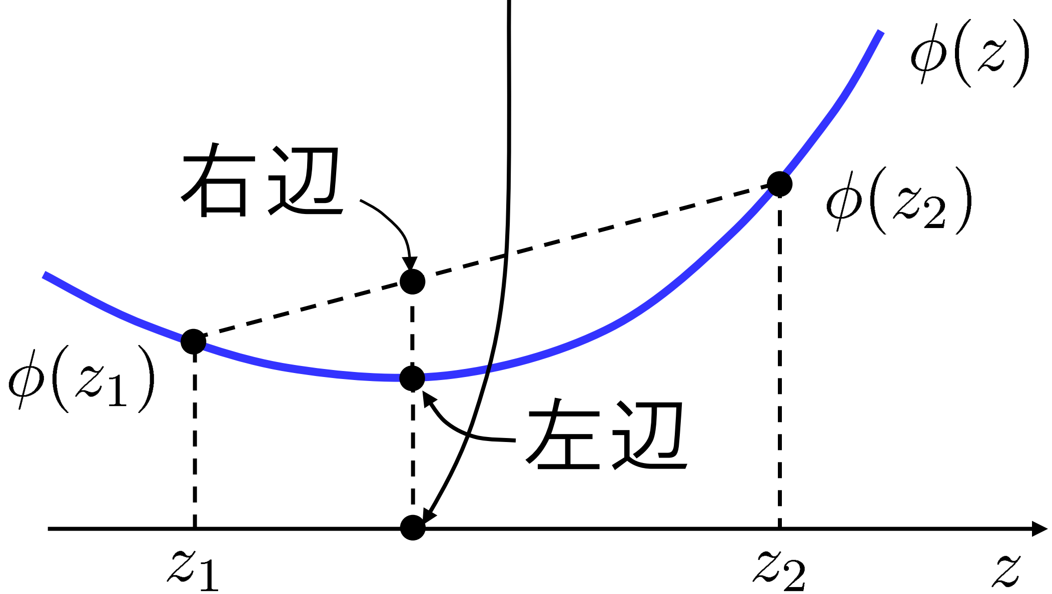
- Jensenの不等式
  - $\phi(\cdot)$：凸関数
  - $\lambda_i \geq 0,\ \sum_i \lambda_i = 1$

$$\phi\left(\sum_i \lambda_i z_i\right) \leq \sum_i \lambda_i \phi\left(z_i\right)$$

例えば,

$\phi(z) = -\log z$ の場合：

$$-\log\left(\sum_i \lambda_i z_i\right) \leq -\sum_i \lambda_i \log z_i$$



FrobeniusノルムFrobenius規準のNMFアルゴリズム

- 目的関数

$$\|\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{U}\|_F^2=\sum_{\omega,t}\left|Y_{\omega,t}-\sum_i H_{\omega,i}U_{i,t}\right|^2$$
$$=\sum_{\omega,t}\left(|Y_{\omega,t}|^2-2Y_{\omega,t}\sum_k H_{\omega,k}U_{k,t}+\underline{\left|\sum_k H_{\omega,k}U_{k,t}\right|^2}\right)$$

- 下線部に対してJensenの不等式を立ててみる

$$\left(\sum_i z_i\right)^2=\left(\sum_i \lambda_i\frac{z_i}{\lambda_i}\right)^2\le\sum_i\lambda_i\left(\frac{z_i}{\lambda_i}\right)^2=\sum_i\frac{{z_i}^2}{\lambda_i}$$

適用

$$\left|\sum_k H_{\omega,k}U_{k,t}\right|^2\le\sum_k\frac{H_{\omega,k}^2U_{k,t}^2}{\lambda_{k,\omega,t}}$$

- 補助関数が完成

$H_{\omega,k}$または$U_{k,t}$ごとの二次関数の和になっている

$$\|\boldsymbol{Y}-\boldsymbol{H}\boldsymbol{U}\|_F^2\le\sum_{\omega,t}\left(|Y_{\omega,t}|^2-2Y_{\omega,t}\sum_k H_{\omega,k}U_{k,t}+\sum_k\frac{H_{\omega,k}^2U_{k,t}^2}{\lambda_{k,\omega,t}}\right)$$

FrobeniusノルムＡ規準のNMFアルゴリズム

- 補助関数が完成したらあとはステップ1とステップ2を導出すればOK!

$$\mathcal{G} := \sum_{\omega,t}\left(|Y_{\omega,t}|^2 - 2Y_{\omega,t}\sum_k H_{\omega,k}U_{k,t} + \sum_k \frac{H_{\omega,k}^2 U_{k,t}^2}{\lambda_{k,\omega,t}}\right)$$

[1] $$\lambda_{k,\omega,t} \leftarrow \underset{\lambda_{k,\omega,t}}{\operatorname{argmin}}\, G = \frac{H_{\omega,k}U_{k,t}}{\sum_{k'} H_{\omega,k'}U_{k',t}}$$

代入

[2] $$H_{\omega,k} \leftarrow \underset{H_{\omega,k}}{\operatorname{argmin}}\, G = \frac{\sum_t Y_{\omega,t}U_{k,t}}{\sum_t \frac{U_{k,t}^2}{\lambda_{k,\omega,t}}}$$

$$U_{k,t} \leftarrow \underset{U_{k,t}}{\operatorname{argmin}}\, G = \frac{\sum_\omega Y_{\omega,t}H_{\omega,k}}{\sum_\omega \frac{H_{\omega,k}^2}{\lambda_{k,\omega,t}}}$$

$$H_{\omega,k} \leftarrow H_{\omega,k}\frac{\sum_t Y_{\omega,t}U_{k,t}}{\sum_t U_{k,t}\sum_{k'} H_{\omega,k'}U_{k',t}}$$

$$U_{k,t} \leftarrow U_{k,t}\frac{\sum_\omega Y_{\omega,t}H_{\omega,k}}{\sum_\omega H_{\omega,k}\sum_{k'} H_{\omega,k'}U_{k',t}}$$

Iダイバージェンス規準のNMFアルゴリズム

- 目的関数

$$
\begin{aligned}
\mathcal{D} &:= \sum_{\omega,t}\left(Y_{\omega,t}\log\frac{Y_{\omega,t}}{\sum H_{\omega,k}U_{k,t}} - Y_{\omega,t} + \sum_k H_{\omega,k}U_{k,t}\right)\\
&= \sum_{\omega,t}\left(Y_{\omega,t}\log Y_{\omega,t} \underline{- Y_{\omega,t}\log\sum_k H_{\omega,k}U_{k,t}} - Y_{\omega,t} + \sum_k H_{\omega,k}U_{k,t}\right)
\end{aligned}
$$

- 下線部に対してJensenの不等式を立ててみる

$$
-\log\left(\sum_i z_i\right) = -\log\left(\sum_i \lambda_i \frac{z_i}{\lambda_i}\right) \leq -\sum_i \lambda_i \log\left(\frac{z_i}{\lambda_i}\right)
$$

適用

$$
-\log\sum_k H_{\omega,k}U_{k,t} \leq -\sum_k \lambda_{k,\omega,t}\log\frac{H_{\omega,k}U_{k,t}}{\lambda_{k,\omega,t}}
$$

- 補助関数が完成

$$
\mathcal{D} \leq \sum_{\omega,t}\left(Y_{\omega,t}\log Y_{\omega,t} - Y_{\omega,t}\sum_k \lambda_{k,\omega,t}\log\frac{H_{\omega,k}U_{k,t}}{\lambda_{k,\omega,t}} - Y_{\omega,t} + \sum_k H_{\omega,k}U_{k,t}\right)
$$

Iダイバージェンス規準のNMFアルゴリズム

- 補助関数が完成したらあとはステップ1とステップ2を導出すればOK!

$$\mathcal{G} := \sum_{\omega,t}\left(Y_{\omega,t}\log Y_{\omega,t} - Y_{\omega,t}\sum_k \lambda_{k,\omega,t}\log\frac{H_{\omega,k}U_{k,t}}{\lambda_{k,\omega,t}} - Y_{\omega,t} + \sum_k H_{\omega,k}U_{k,t}\right)$$

[1] $$\lambda_{k,\omega,t} \leftarrow \underset{\lambda_{k,\omega,t}}{\operatorname{argmin}}\, G = \frac{H_{\omega,k}U_{k,t}}{\sum_{k'} H_{\omega,k'}U_{k',t}}$$

代入

[2] $$H_{\omega,k} \leftarrow \underset{H_{\omega,k}}{\operatorname{argmin}}\, G = \frac{\sum_t Y_{\omega,t}\lambda_{k,\omega,t}}{\sum_t U_{k,t}}$$

$$H_{\omega,k} \leftarrow H_{\omega,k}\frac{\sum_t \frac{Y_{\omega,t}U_{k,t}}{\sum_{k'} H_{\omega,k}U_{k',t}}}{\sum_t U_{k,t}}$$

$$U_{k,t} \leftarrow \underset{U_{k,t}}{\operatorname{argmin}}\, G = \frac{\sum_\omega Y_{\omega,t}\lambda_{k,\omega,t}}{\sum_\omega H_{\omega,k}}$$

$$U_{k,t} \leftarrow U_{k,t}\frac{\sum_\omega \frac{Y_{\omega,t}H_{\omega,k}}{\sum_{k'} H_{\omega,k'}U_{k',t}}}{\sum_\omega H_{\omega,k}}$$

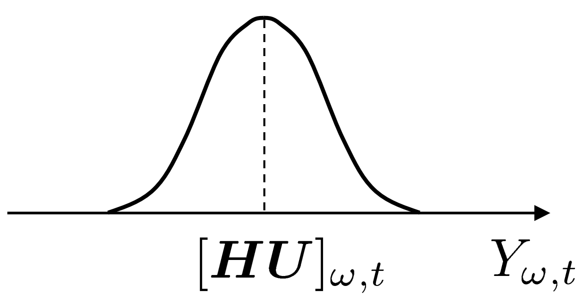
統計モデルとしての解釈

- NMFは以下を仮定した最尤推定問題と等価
  - Frobeniusノルム規準 $\boxed{Y_{\omega,t} \sim \mathcal{N}([\boldsymbol{HU}]_{\omega,t}, \sigma^2)}$

$$P(\boldsymbol{Y}|\boldsymbol{H},\boldsymbol{U}) = \prod_{\omega,t} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(Y_{\omega,t} - [\boldsymbol{HU}]_{\omega,t})^2}$$

$$\rightarrow -\log P(\boldsymbol{Y}|\boldsymbol{HU}) \stackrel{c}{=} \frac{1}{2\sigma^2} \sum_{\omega,t} \left(Y_{\omega,t} - [\boldsymbol{HU}]_{\omega,t}\right)^2$$

正規分布

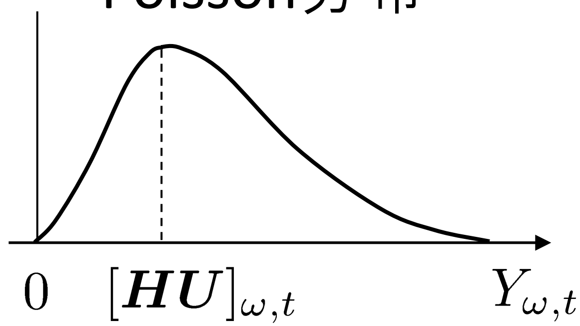


  - Iダイバージェンス規準 $\boxed{Y_{\omega,t} \sim \mathrm{Poisson}([\boldsymbol{HU}]_{\omega,t})}$

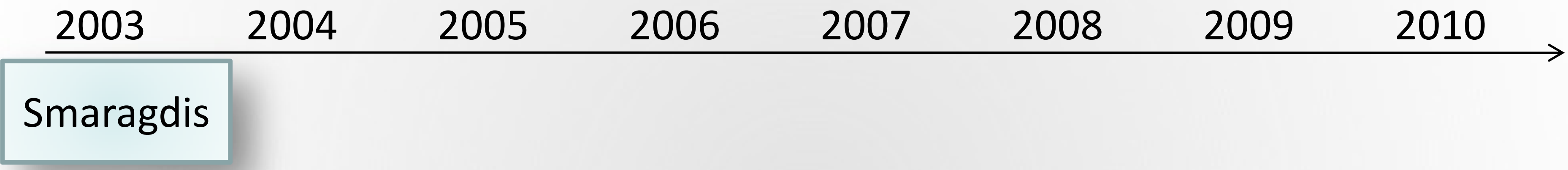
$$P(\boldsymbol{Y}|\boldsymbol{H},\boldsymbol{U}) = \prod_{\omega,t} \frac{{[\boldsymbol{HU}]_{\omega,t}}^{Y_{\omega,t}} e^{-[\boldsymbol{HU}]_{\omega,t}}}{Y_{\omega,t}!}$$

$$\rightarrow -\log P(\boldsymbol{Y}|\boldsymbol{HU}) \stackrel{c}{=} \sum_{\omega,t} \left(-Y_{\omega,t} \log[\boldsymbol{HU}]_{\omega,t} + [\boldsymbol{HU}]_{\omega,t}\right)$$

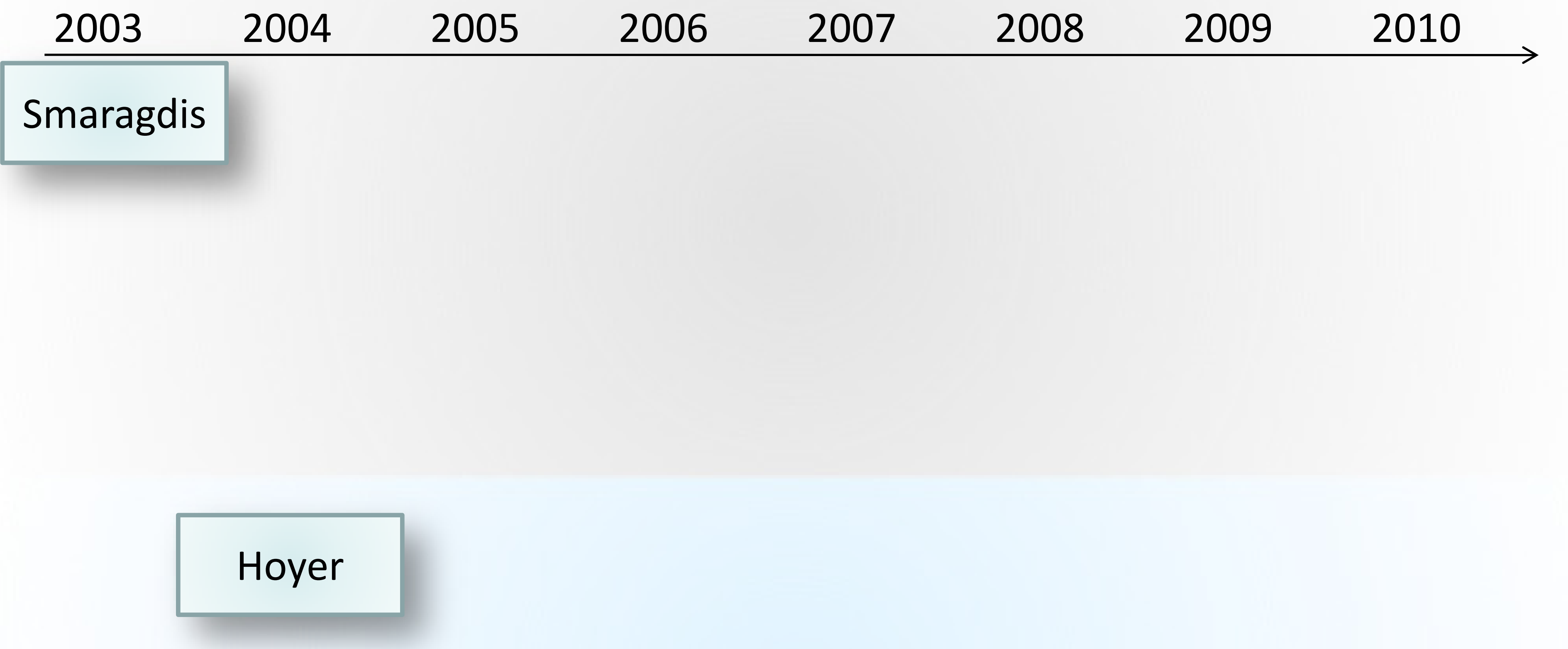
Poisson分布



NMFの改良･拡張モデル



NMFの改良･拡張モデル



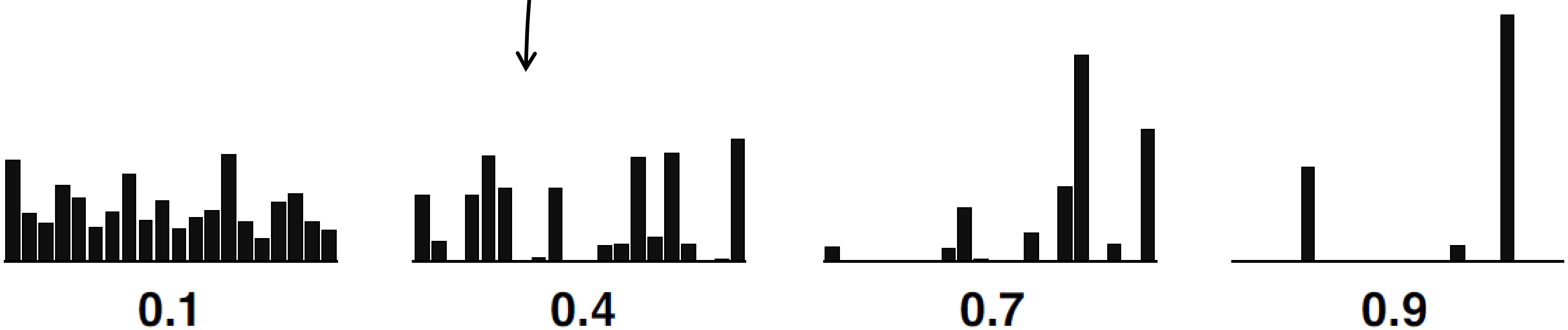
スパースNMF [Hoyer2004]

- アクティベーションを強制的にスパース化

$$\text{minimize} \quad \sum_{\omega,t} \mathcal{D}\left(Y_{\omega,t} \,\middle|\, \sum_i H_{\omega,i} U_{i,t}\right)$$

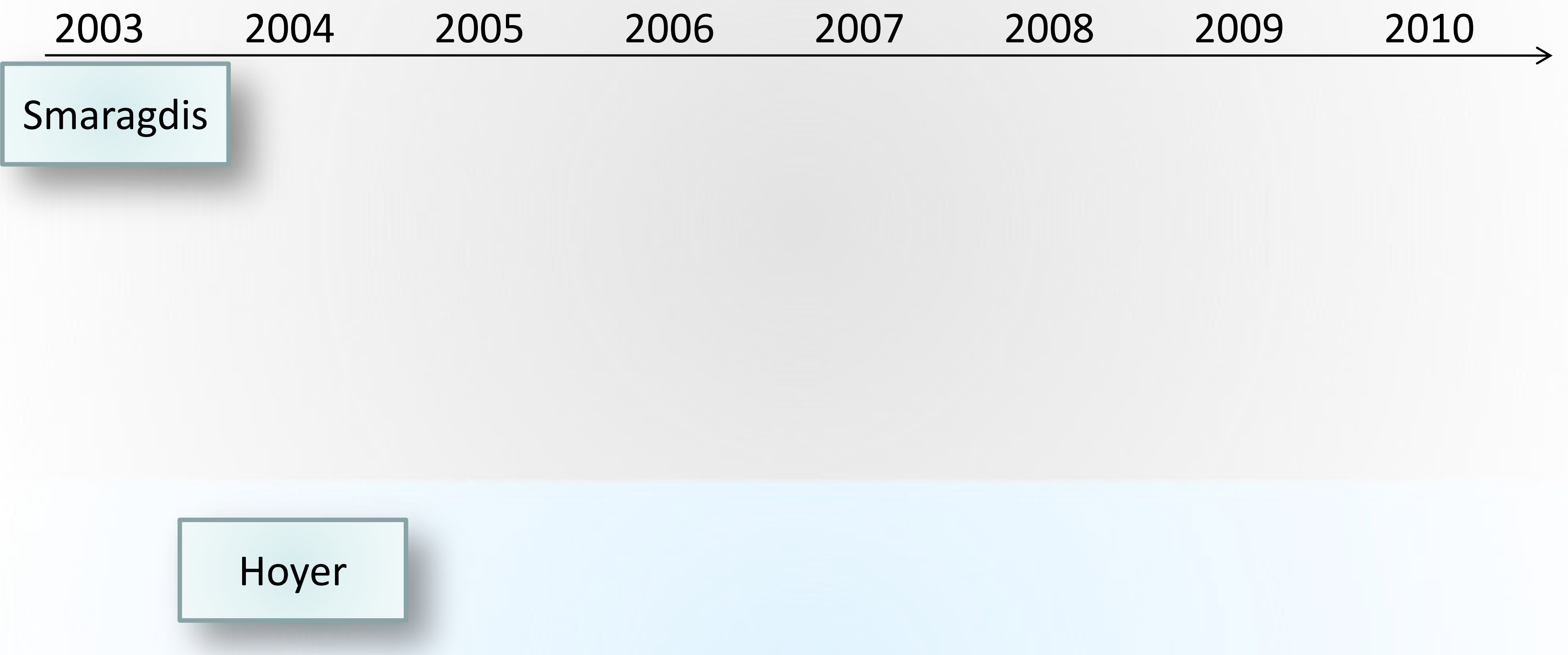
$$\text{subject to} \quad \frac{\sqrt{T} - \sum_{k,t} |U_{k,t}| / \sqrt{\sum_{k,t} U_{k,t}^2}}{\sqrt{T} - 1} = S_U$$

スパース性尺度（1のときスパース，0のとき非スパース）

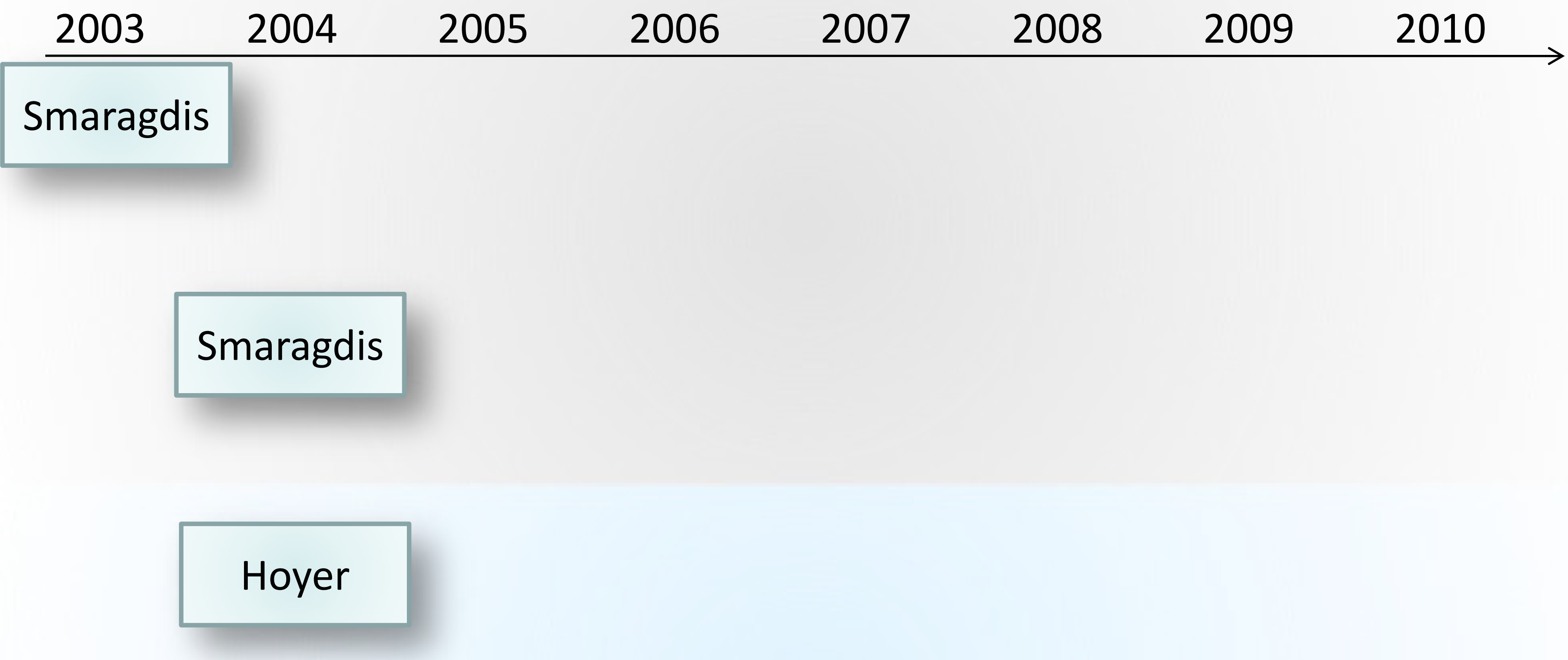


P. O. Hoyer, “Non-negative matrix factorization with sparseness constraints,” *J. Mach. Learning Res., vol. 5, pp. 1457–1469, 2004.*

NMFの改良･拡張モデル



NMFの改良·拡張モデル

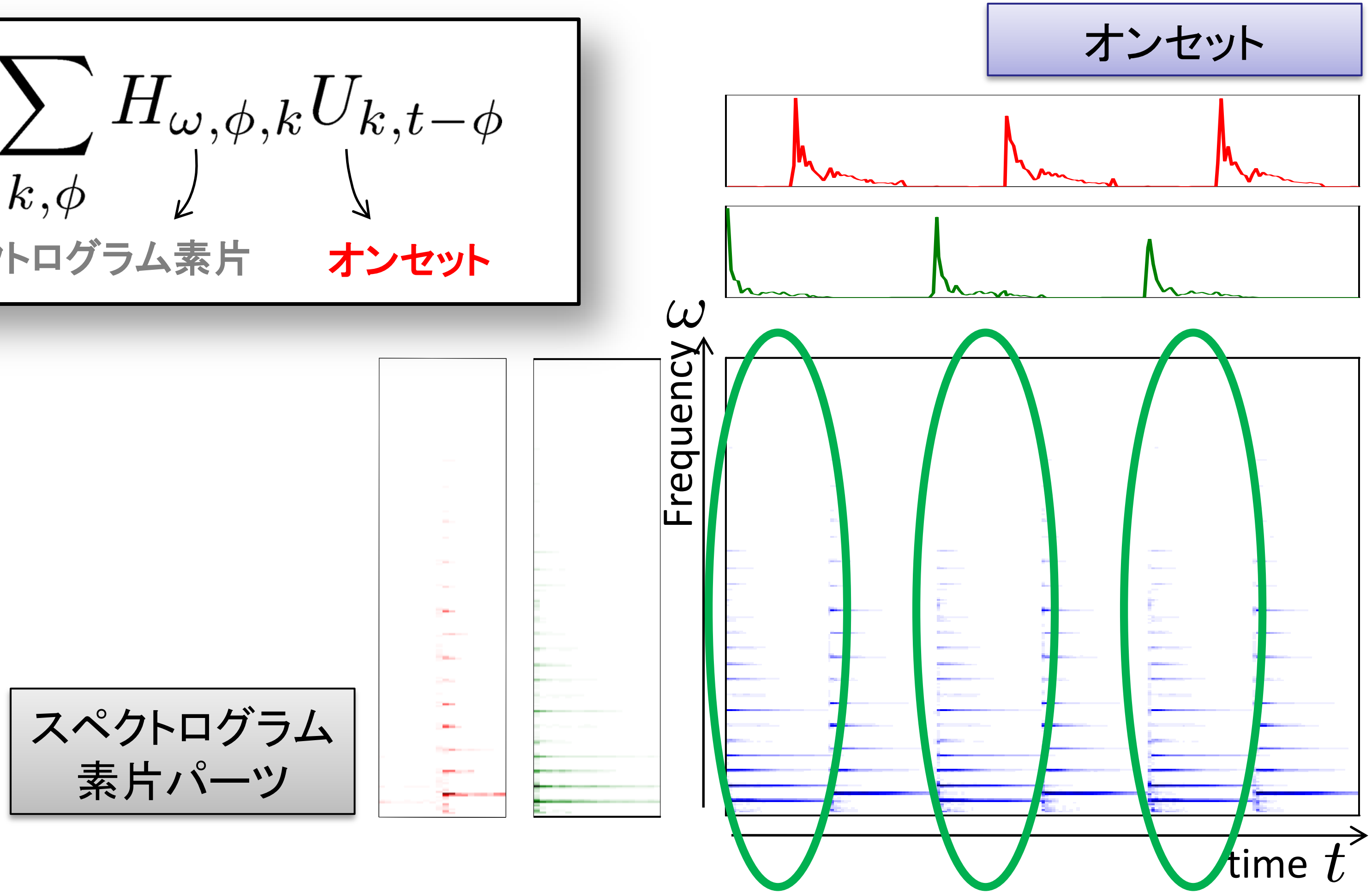


Nonnegative Matrix Factor Deconvolution [Smaragdis2004]

- スペクトログラム素片パーツを時間方向に畳み込んで楽音スペクトログラムを表現

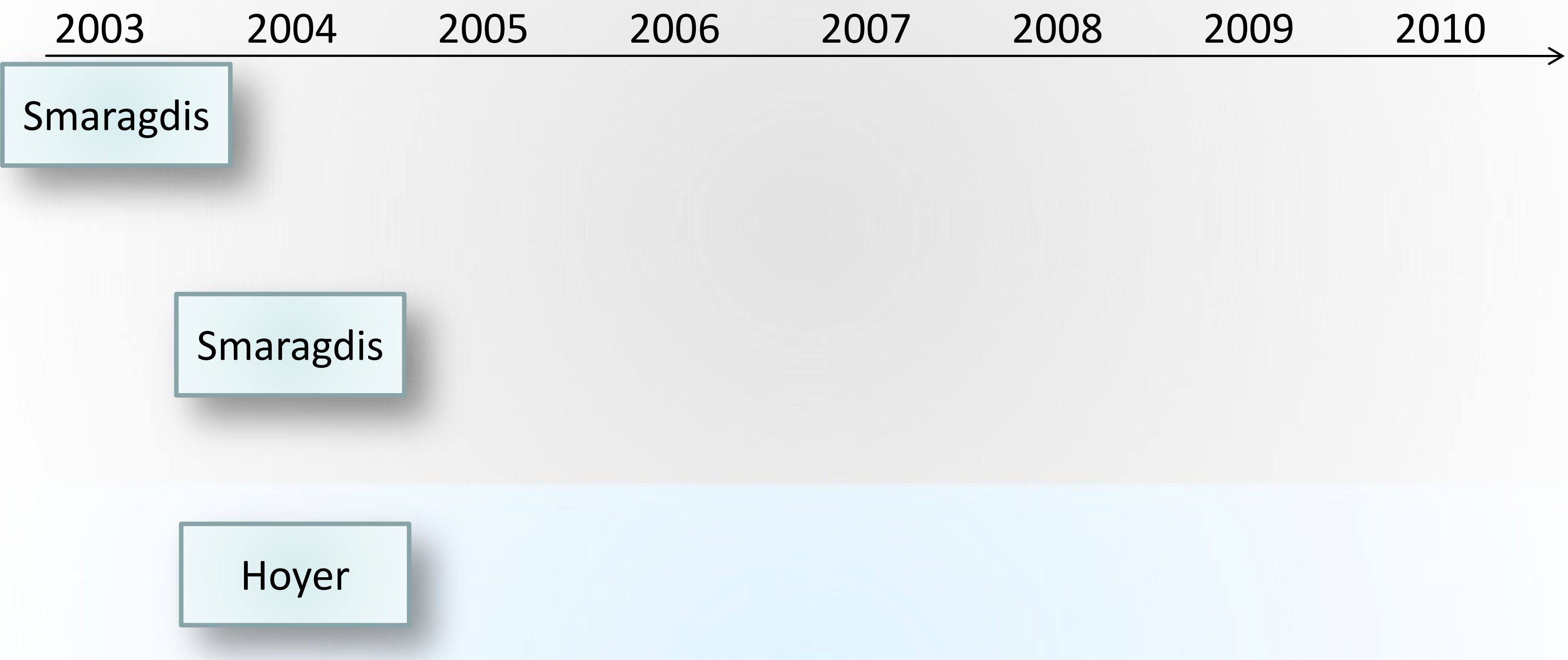
$$Y_{\omega,t} \simeq \sum_{k,\phi} H_{\omega,\phi,k} U_{k,t-\phi}$$

スペクトログラム素片　　オンセット



P. Smaragdis, "Non-negative matrix factor deconvolution; extraction of multiple sound sources from monophonic inputs," in *Proc. ICA2004, pp. 494–499, 2004.*

NMFの改良･拡張モデル



NMFの改良･拡張モデル

Bregmanダイバージェンス規準NMF [Dhillon2005]

- Bregmanダイバージェンス [L.M. Bregman, 1967]

$$\mathcal{D}_\varphi(y|x) := \varphi(y) - \varphi(x) - \varphi'(x)(y-x)$$（φ は任意の微分可能な凸関数）

⇒ y と x が近いほど小さい

- ◆ $\varphi(z) = z^2$ ：二乗誤差誤差
- ◆ $\varphi(z) = z \log z$ ：I ダイバージェンス
- ◆ $\varphi(z) = -\log z$ ：板倉齋藤距離

- $$\text{minimize} \sum_{\omega,t} \mathcal{D}_\varphi\left(\sum_k H_{\omega,k} U_{k,t} \middle| Y_{\omega,t}\right)$$

1. $\varphi'(xy) = \varphi'(x)\varphi'(y)/\varphi'(1)$ を満たすクラス（$\varphi(z) = z^2$ はこのタイプ）

$$H_{\omega,k} \leftarrow H_{\omega,k} \cdot \varphi'^{-1}\left(\frac{\varphi'(1)\sum_t \varphi'(Y_{\omega,t})U_{k,t}}{\sum_t \varphi'\left(\sum_k H_{\omega,k}U_{k,t}\right)U_{k,t}}\right)$$

Uの更新式も同様に解析的に求まる

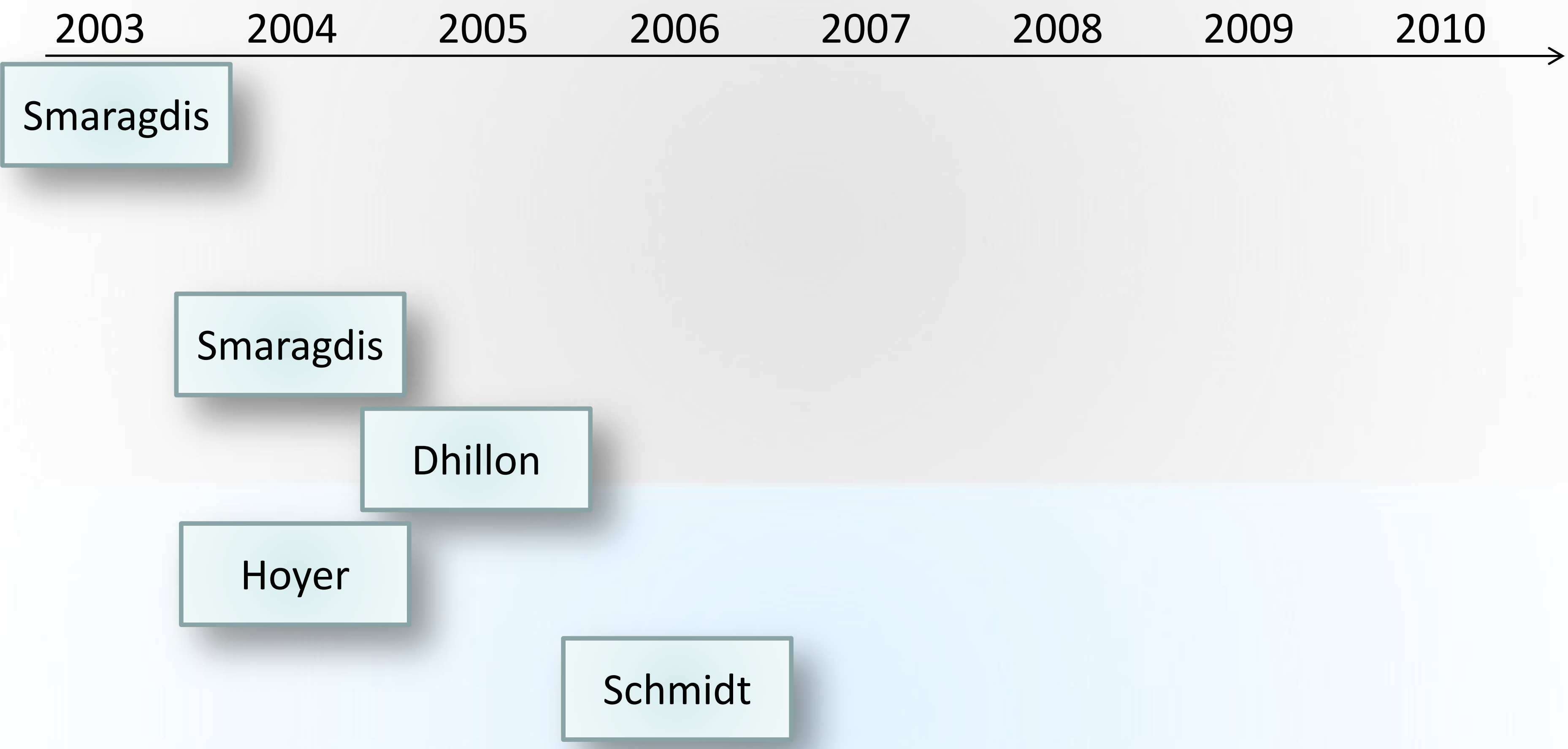
2. $\varphi'(xy) = \varphi'(x) + \varphi'(y) + \varphi'(1)$ を満たすクラス（$\varphi(z) = z\log z, -\log z$ はこのタイプ）

$$H_{\omega,k} \leftarrow H_{\omega,k} \cdot \varphi'^{-1}\left(\sum_t \left[\varphi'(Y_{\omega,t}) - \varphi'\left(\sum_k H_{\omega,k}U_{k,t}\right) - \varphi'(1)\right]U_{k,t} \middle/ \sum_t U_{k,t}\right)$$

I.S. Dhillon and S. Sra, “Generalized nonnegative matrix approximations with Bregman divergences ,” Proc. *NIPS 2005, pp. 283-290, 2005.*

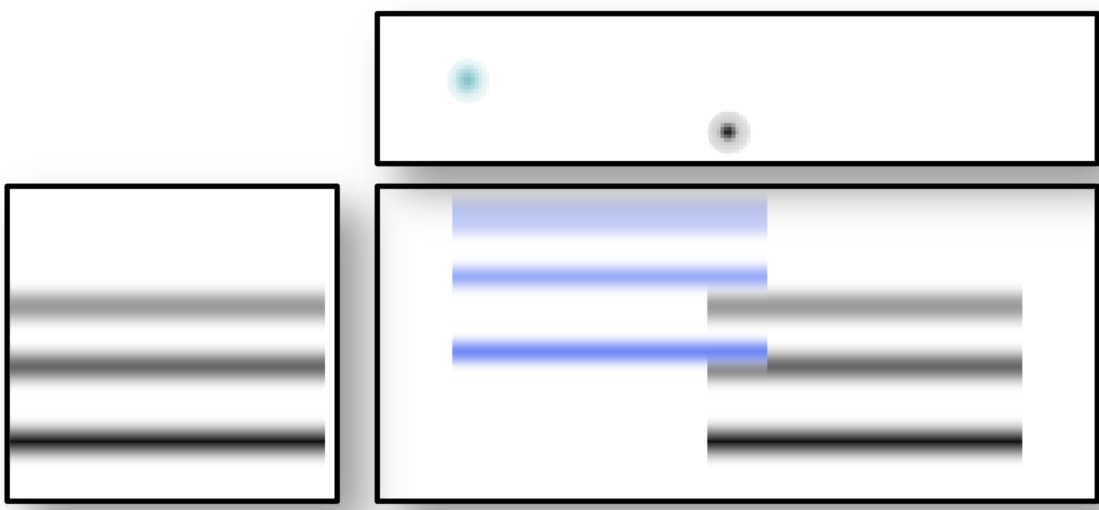
NMFの改良·拡張モデル

NMFの改良･拡張モデル



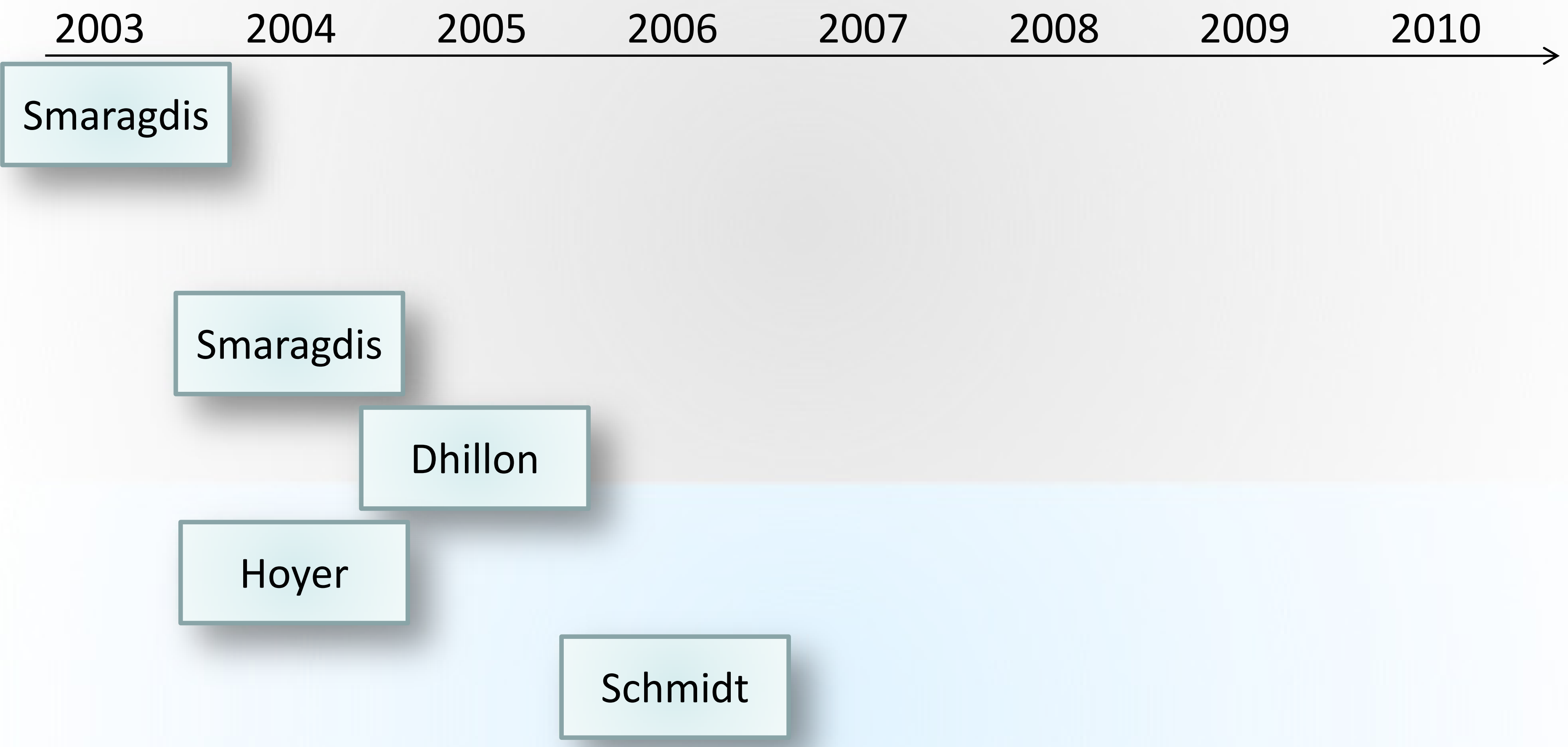
Nonnegative Matrix Factor 2D Deconvolution [Schmidt2006]

- 楽音の調波構造が対数周波数軸上でシフト不変であると仮定（ピッチが変わっても倍音パワー比は不変という仮定）
- スペクトログラムパーツの対数周波数-時間平面での2次元畳み込みで楽音スペクトログラムを表現

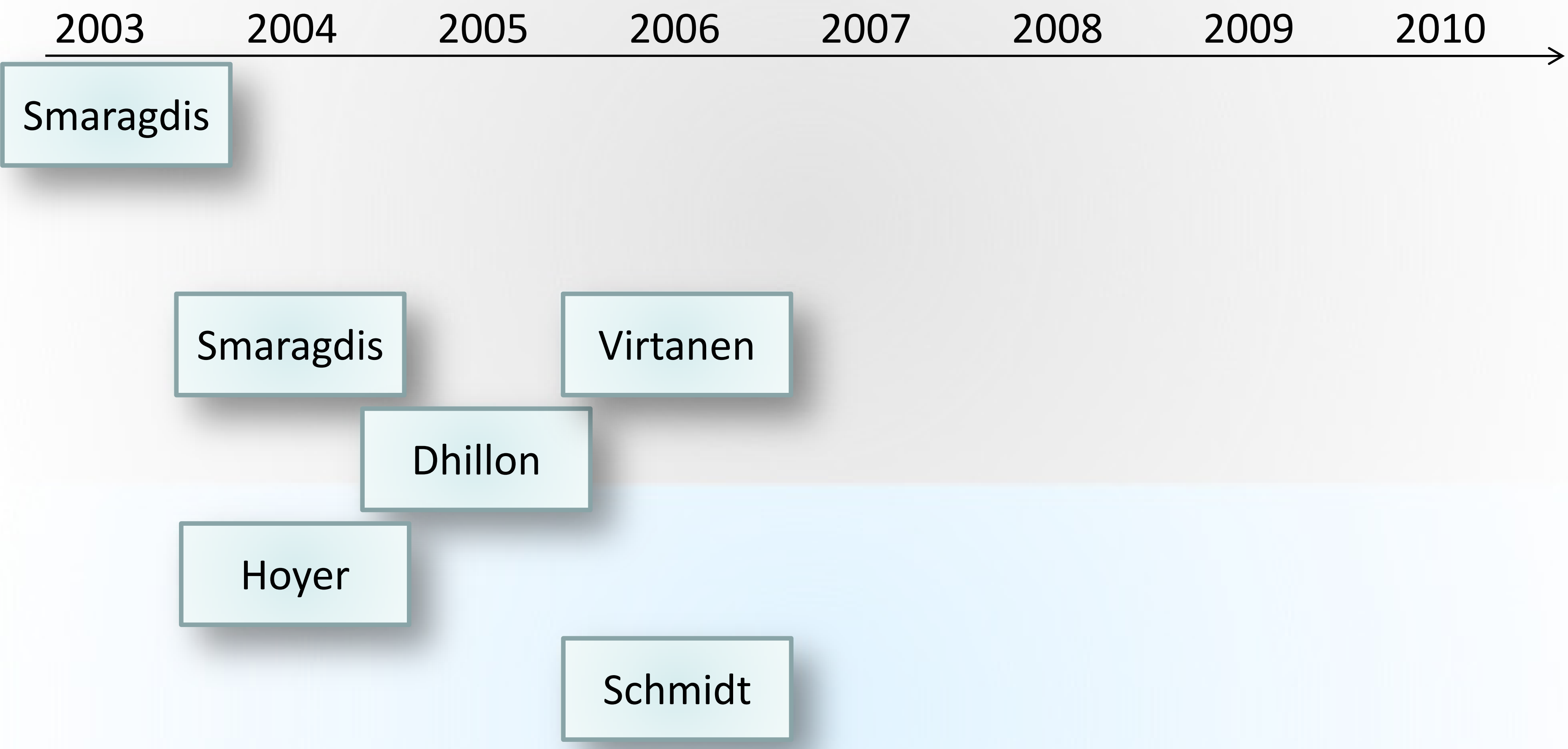


M. N. Schmidt andM.Mørup, “Nonnegative matrix factor 2-D deconvolution for blind single channel source separation,” in Proc. *ICA2006, pp. 700-707, 2006.*

NMFの改良･拡張モデル

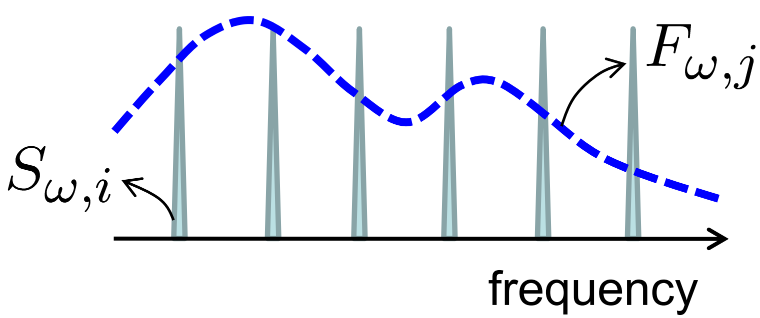


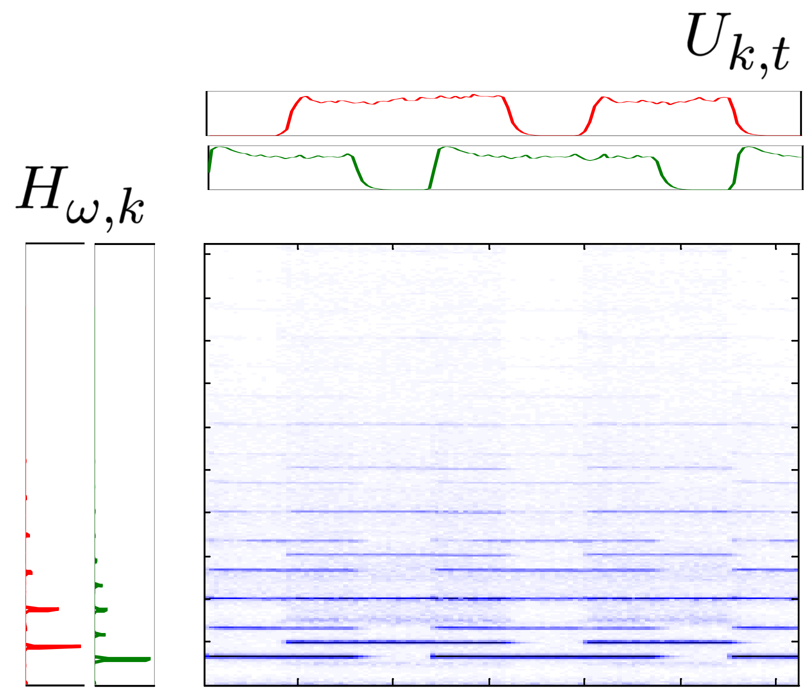
NMFの改良・拡張モデル



ソースフィルタモデルに基づくNMF [Virtanen2006]

- $\{i, j\}$ 番目の基底スペクトルを i 番目のソーススペクトルと j 番目のフィルタの積で表現





従来NMF：

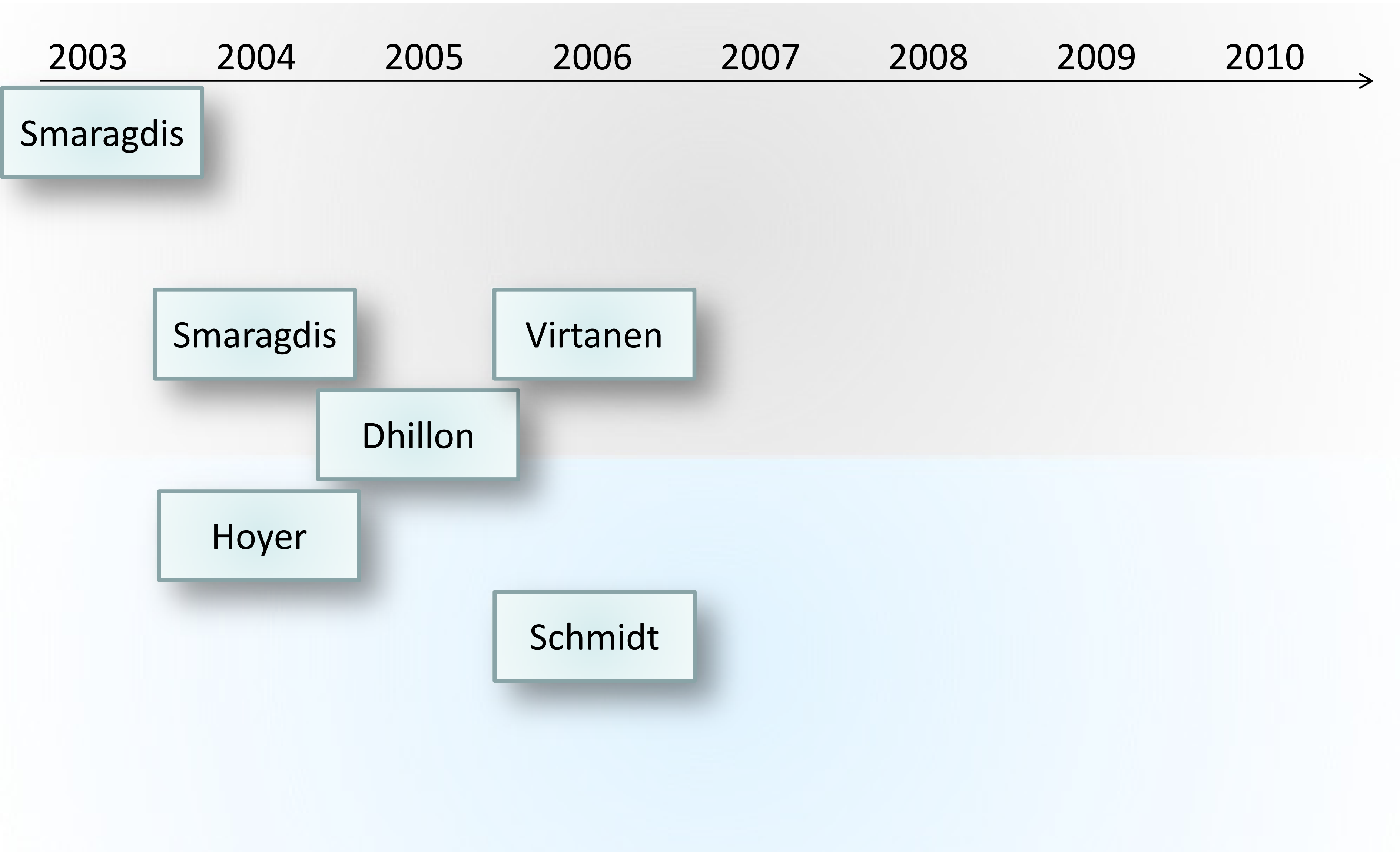
$$\sum_k H_{\omega,k} U_{k,t}$$

ソースフィルタNMF：

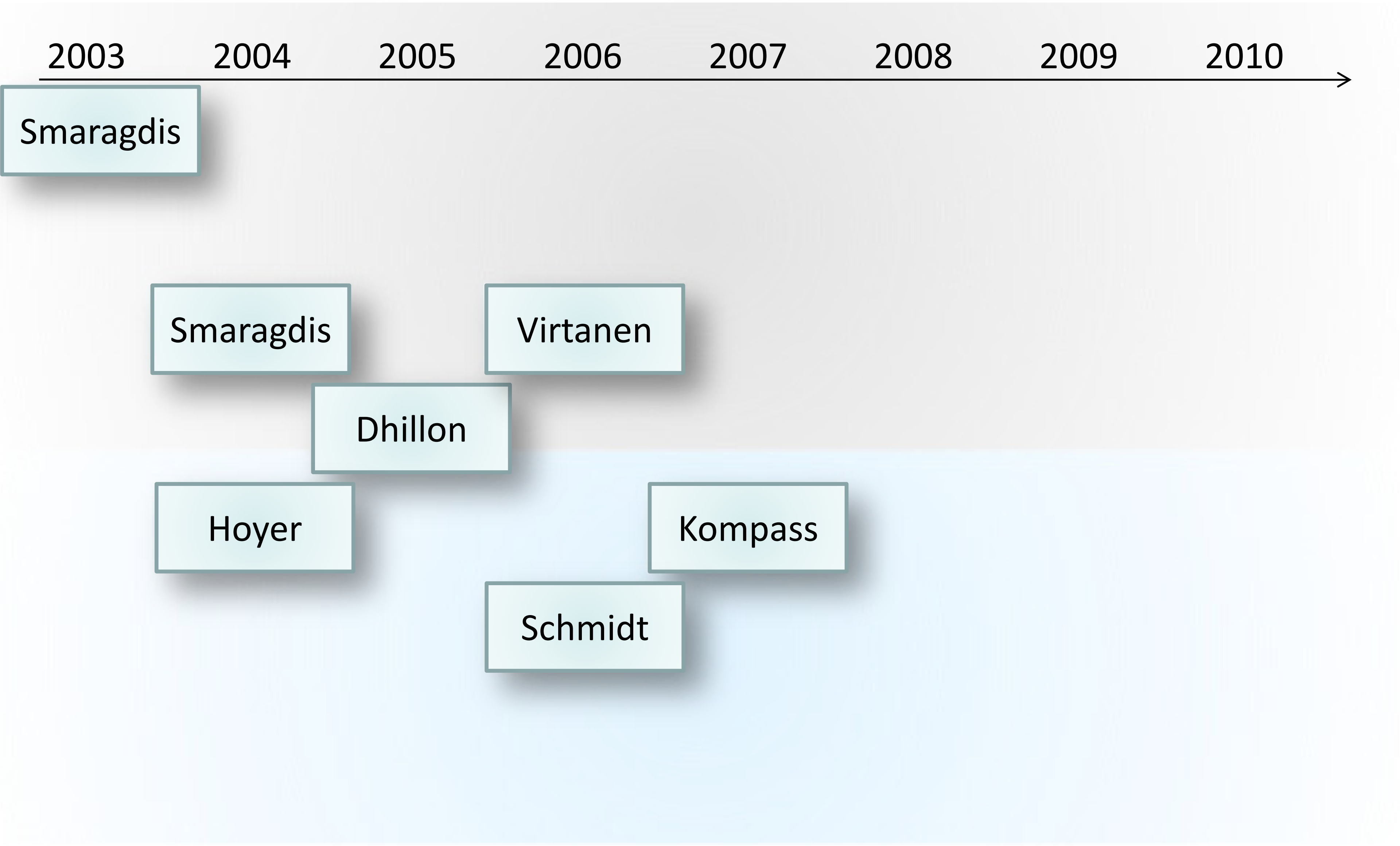
$$\sum_{i,j} \underline{S_{\omega,i} F_{\omega,j}} U_{i,j,t}$$

T. Virtanen and A. Klapuri, "Analysis of polyphonic audio using source-filter model and non-negative matrix factorization," in *Proceedings of the Advances in Models for Acoustic Processing, Neural Information Processing Systems Workshop*, 2006.

NMFの改良・拡張モデル



NMFの改良・拡張モデル

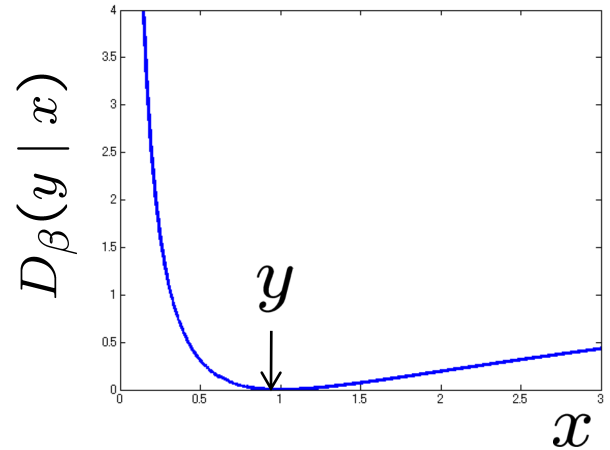


βダイバージェンス規準NMF [Kompass2007]

- βダイバージェンス[Eguchi2001]を局所最小化するアルゴリズム

$$\mathcal{D}_\beta\left(y|x\right) = \frac{1}{\beta(\beta-1)}\left(y^\beta + (\beta-1)x^\beta - \beta y x^{\beta-1}\right)$$

- $\beta=0$: 板倉齋藤距離
- $\beta=1$: I ダイバージェンス
- $\beta=2$: 二乗誤差

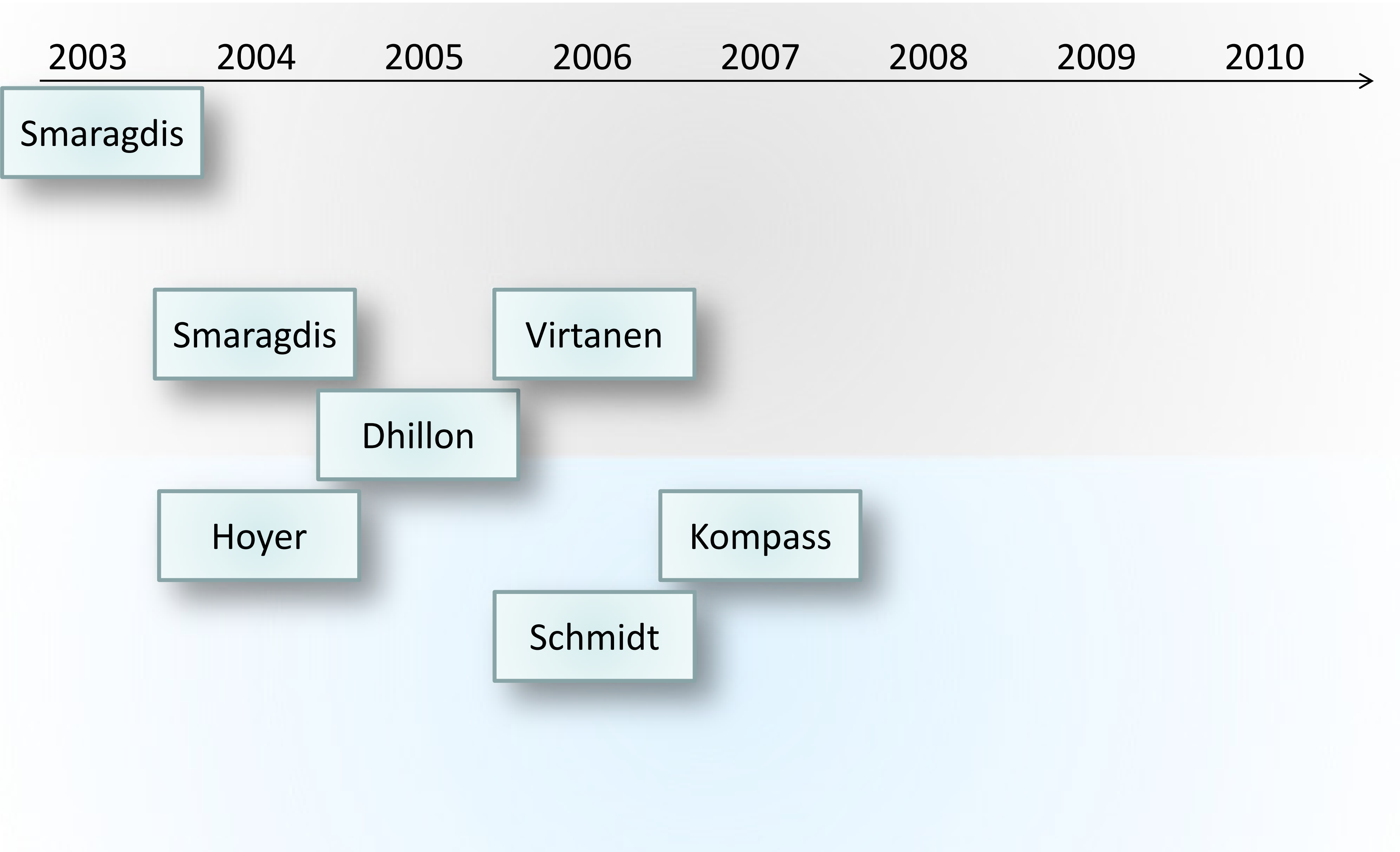


$$\boldsymbol{H} \leftarrow \boldsymbol{H} \cdot \frac{(\boldsymbol{Y} \cdot (\boldsymbol{HU})^{\beta-2})\boldsymbol{U}^{\mathrm{T}}}{(\boldsymbol{HU})^{\beta-1}\boldsymbol{U}^{\mathrm{T}}} \quad \boldsymbol{U} \leftarrow \boldsymbol{U} \cdot \frac{\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{Y} \cdot (\boldsymbol{HU})^{\beta-2})}{\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{HU})^{\beta-1}}$$

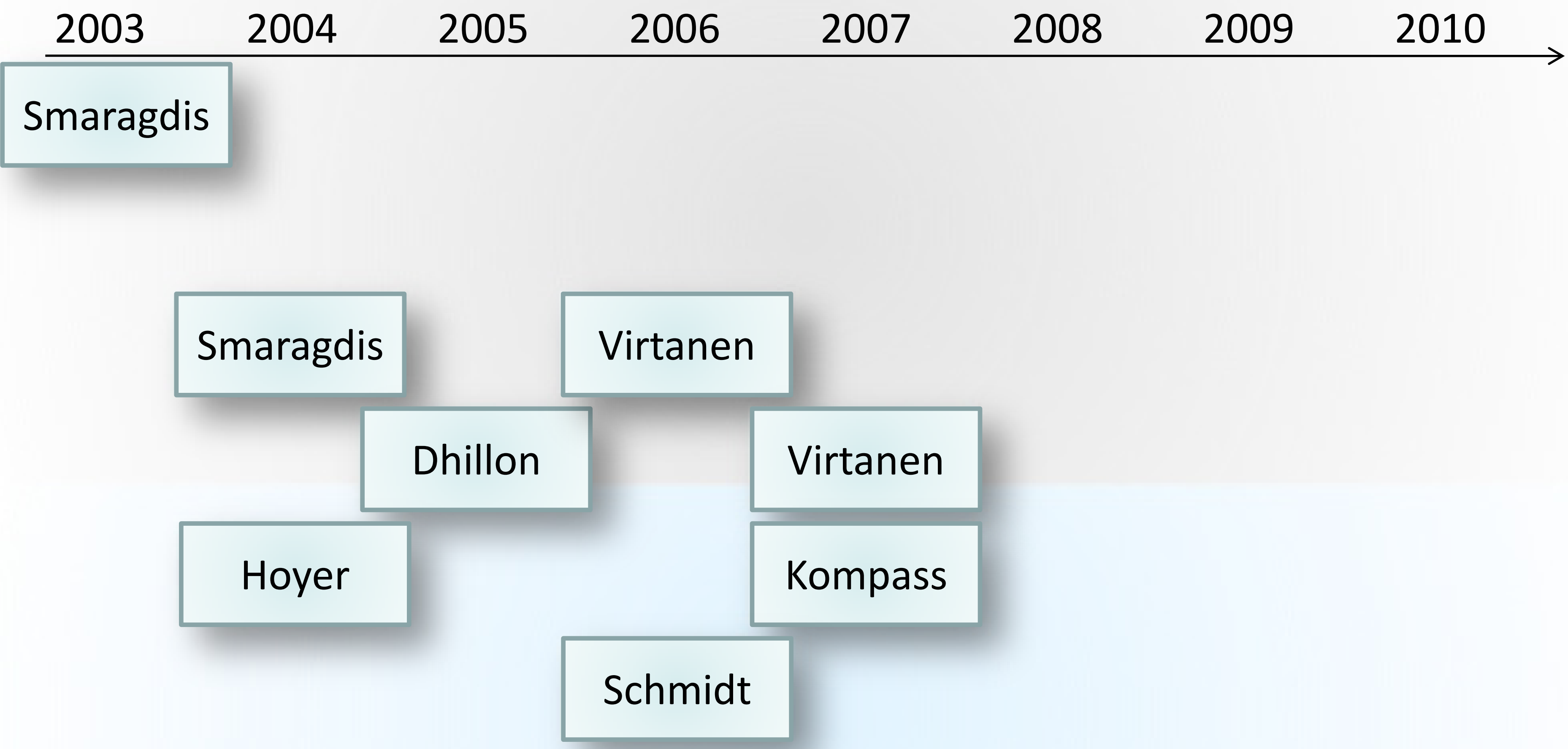
$1 \leq \beta \leq 2$ の場合しか収束性が保証されない

R. Kompass, “A generalized divergence measure for nonnegative matrix factorization,” Neural Computation, 19(3), pp. 780–791, 2007.

NMFの改良・拡張モデル

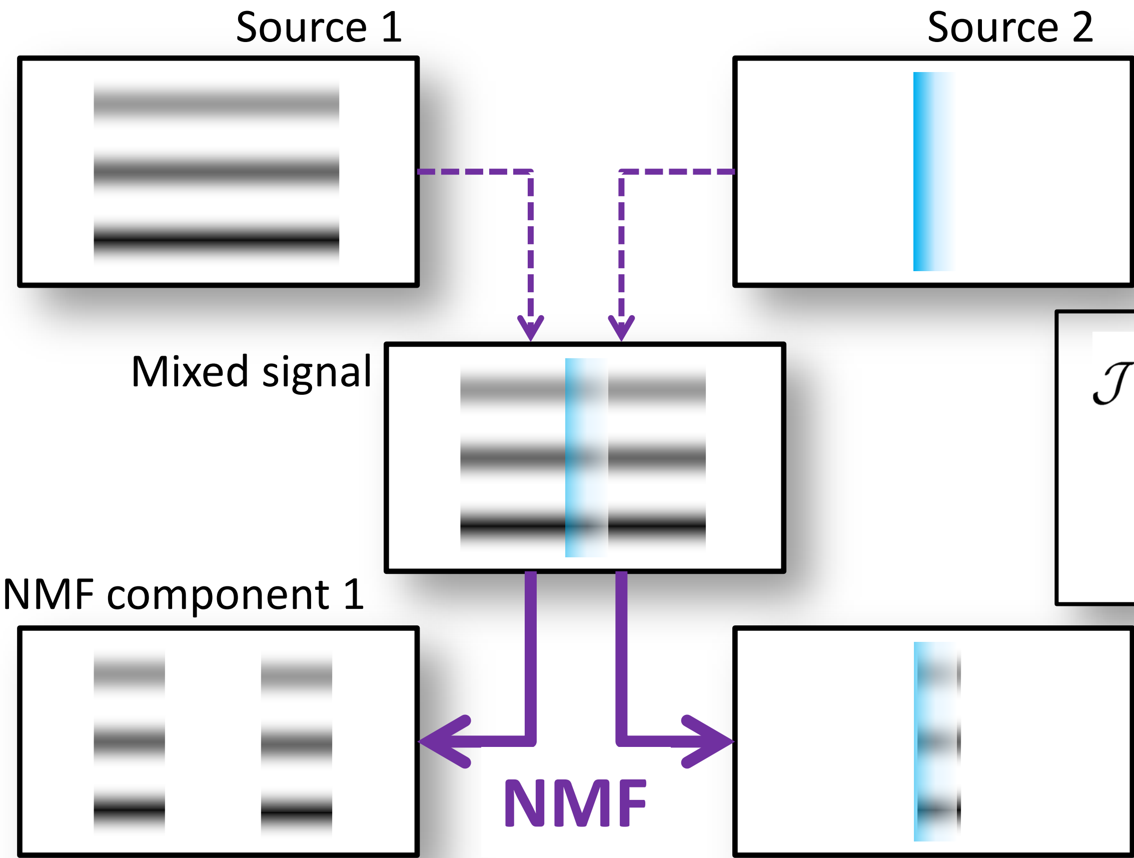


NMFの改良･拡張モデル



アクティベーションの連続性規準入りNMF [Virtanen2007]

- 楽器音のパワーエンベロープは滑らかという仮定



音源のストリームが
分断されてしまっている！

別音源のストリームが
混入してしまっている！

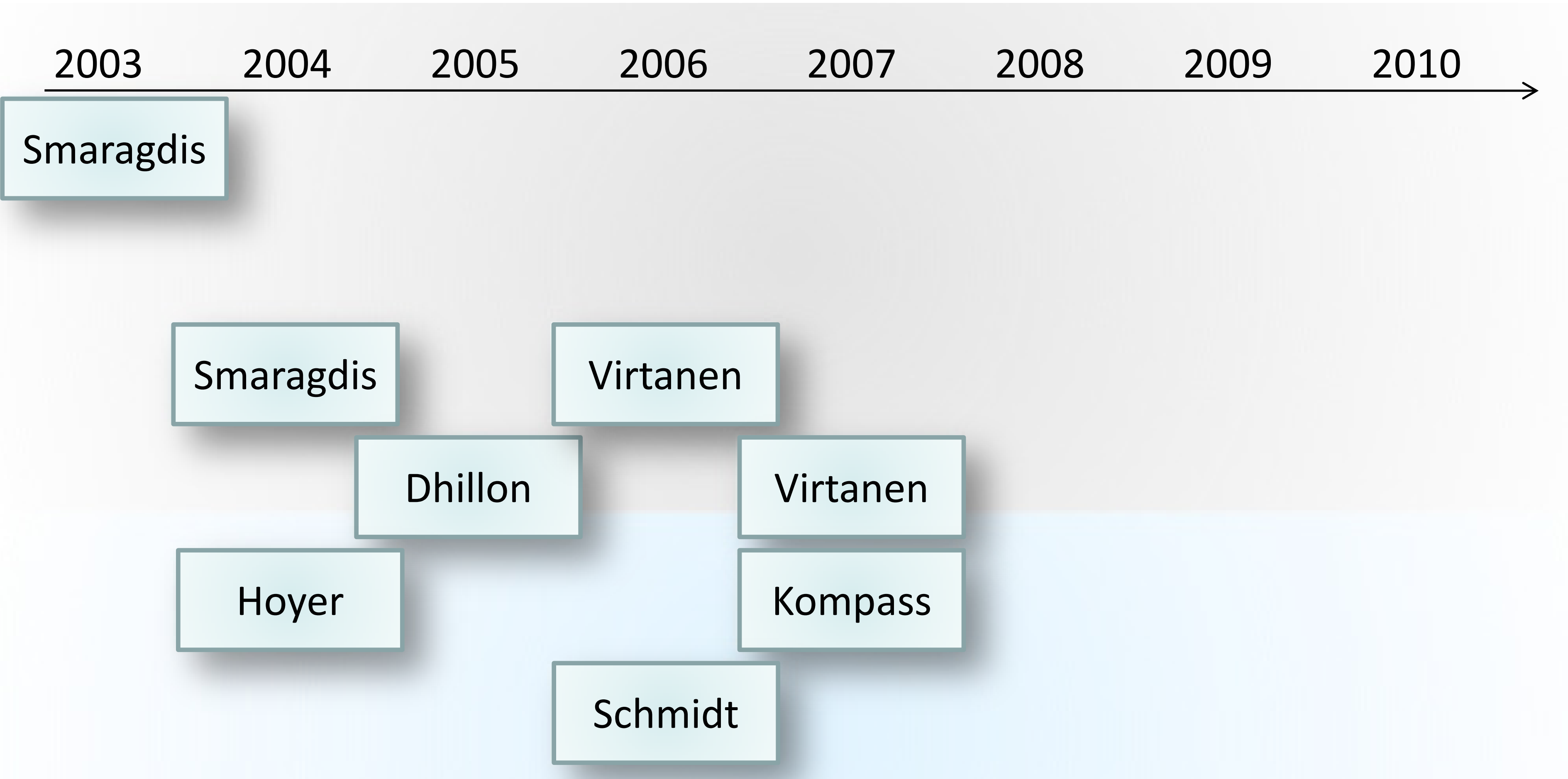
そこで，アクティベーション
の連続性規準を導入・・・

$$\mathcal{J} = \sum_{\omega,t} \mathcal{D}\left(Y_{\omega,t} \,\middle|\, \sum_{k} H_{\omega,k} U_{k,t}\right) + \underline{\sum_{k,t} (U_{k,t} - U_{k,t-1})^2}$$

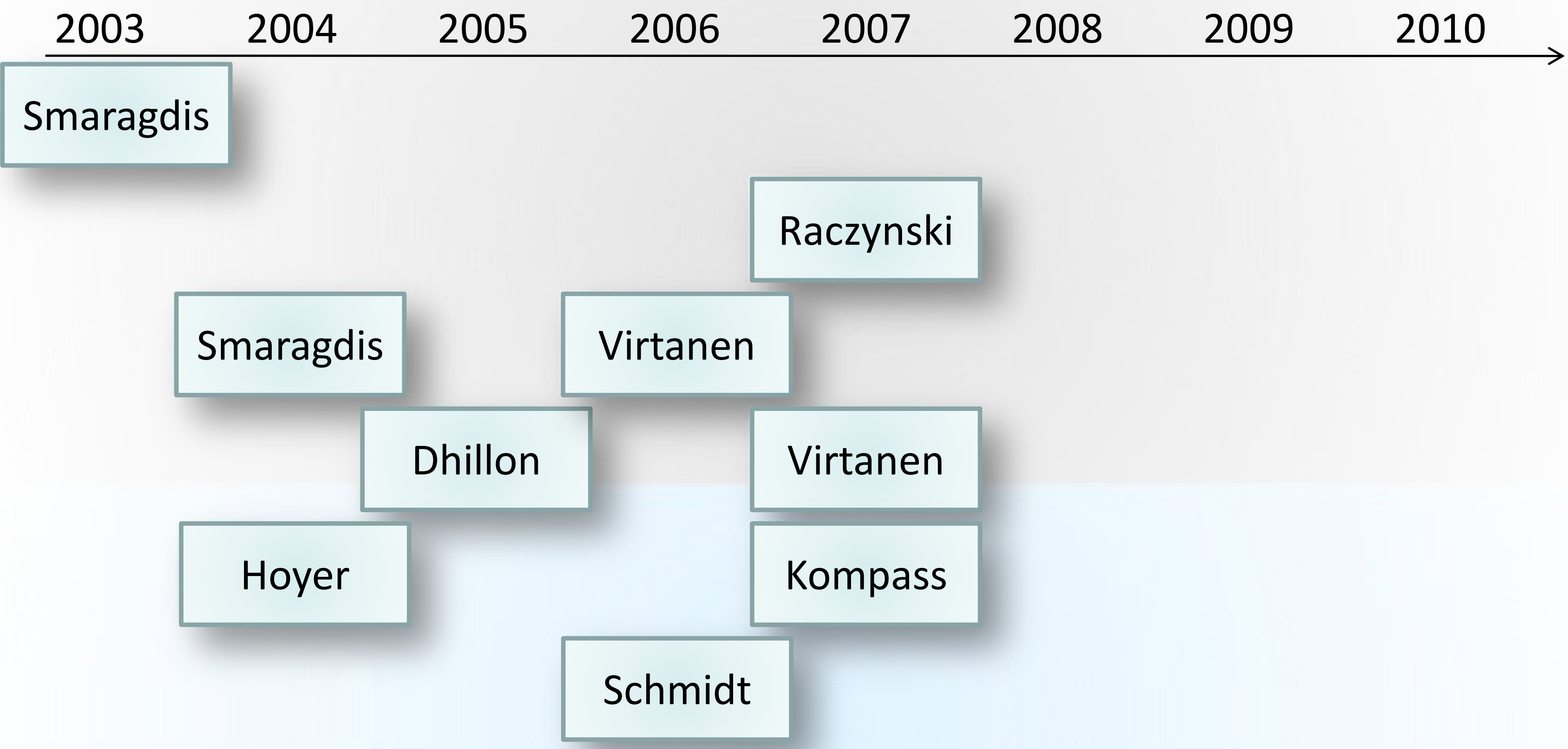
- アクティベーションが滑らかな
ほど小さい値をとるコスト関数
- この項を含めて最適化を行う
ことで，左図のようなことが
起こりにくくなる

T. Virtanen, “Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria,” *IEEE Trans. on Audio, Speech and Language Processing, vol. 15, no. 3, pp. 1066*–1074, 2007.

NMFの改良・拡張モデル



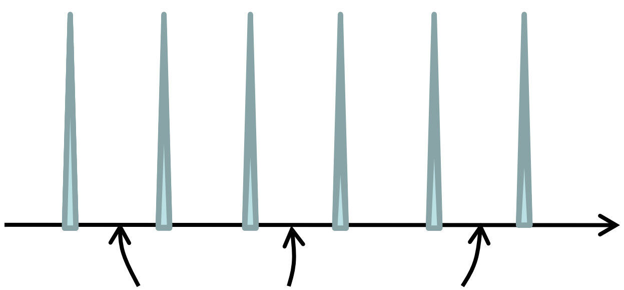
NMFの改良･拡張モデル



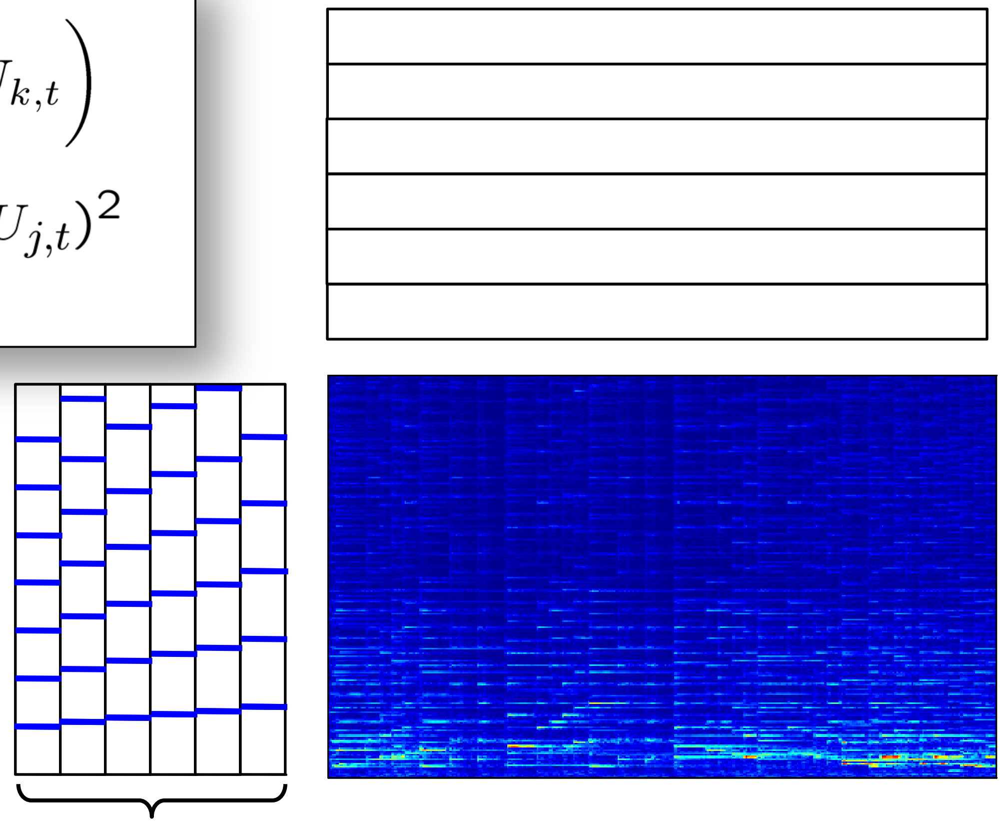
ハーモニックNMF [Raczynski2007]

- 調波構造形の基底スペクトル
- アクティベーションの無相関規準

$$
\mathcal{J} = \sum_{\omega,t} \mathcal{D}\left(Y_{\omega,t} \middle| \sum_{k} H_{\omega,k} U_{k,t} \right) + \sum_{k \neq j} \sum_{t} (U_{k,t} - U_{j,t})^2
$$



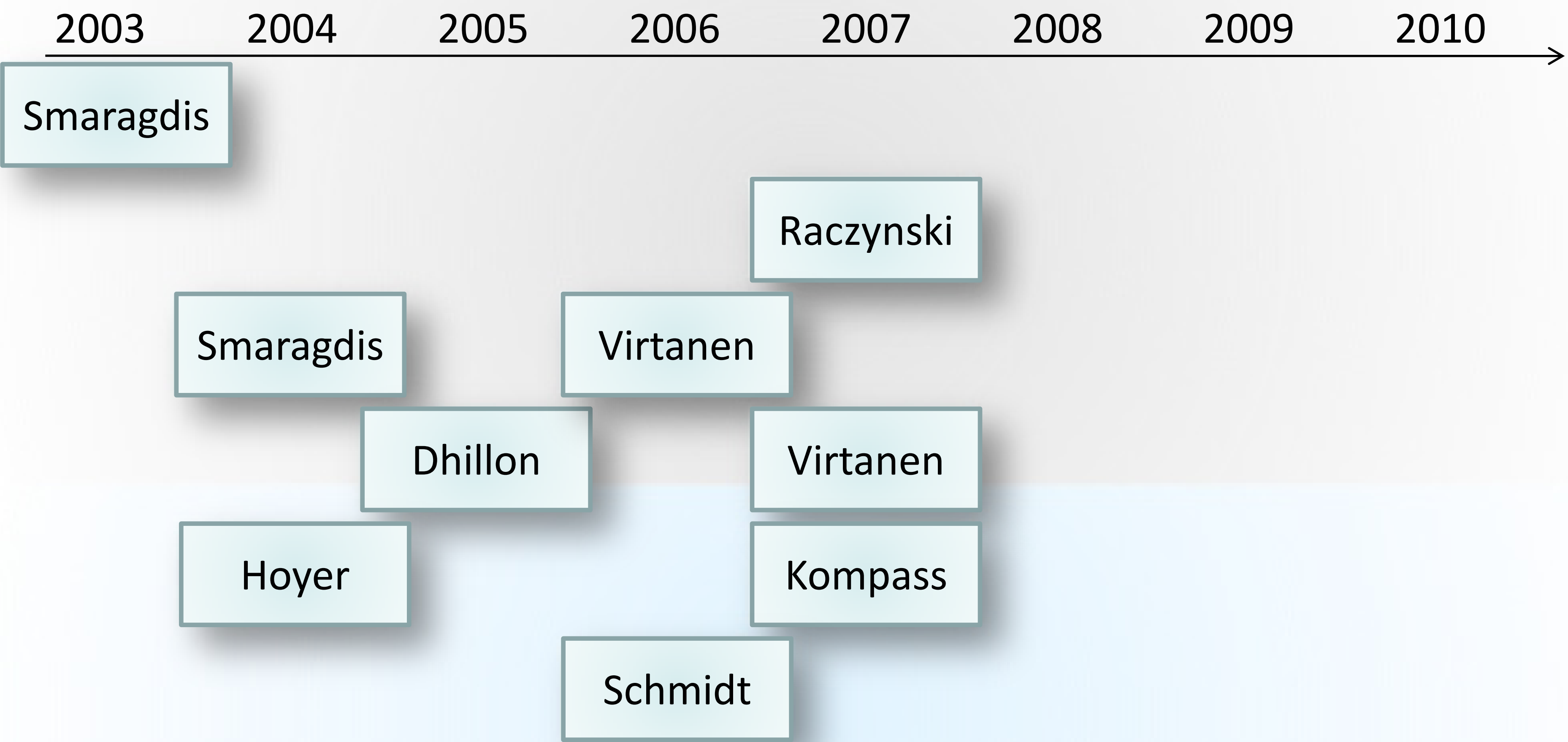
基本周波数･高調波周波数
以外の成分を0に初期設定



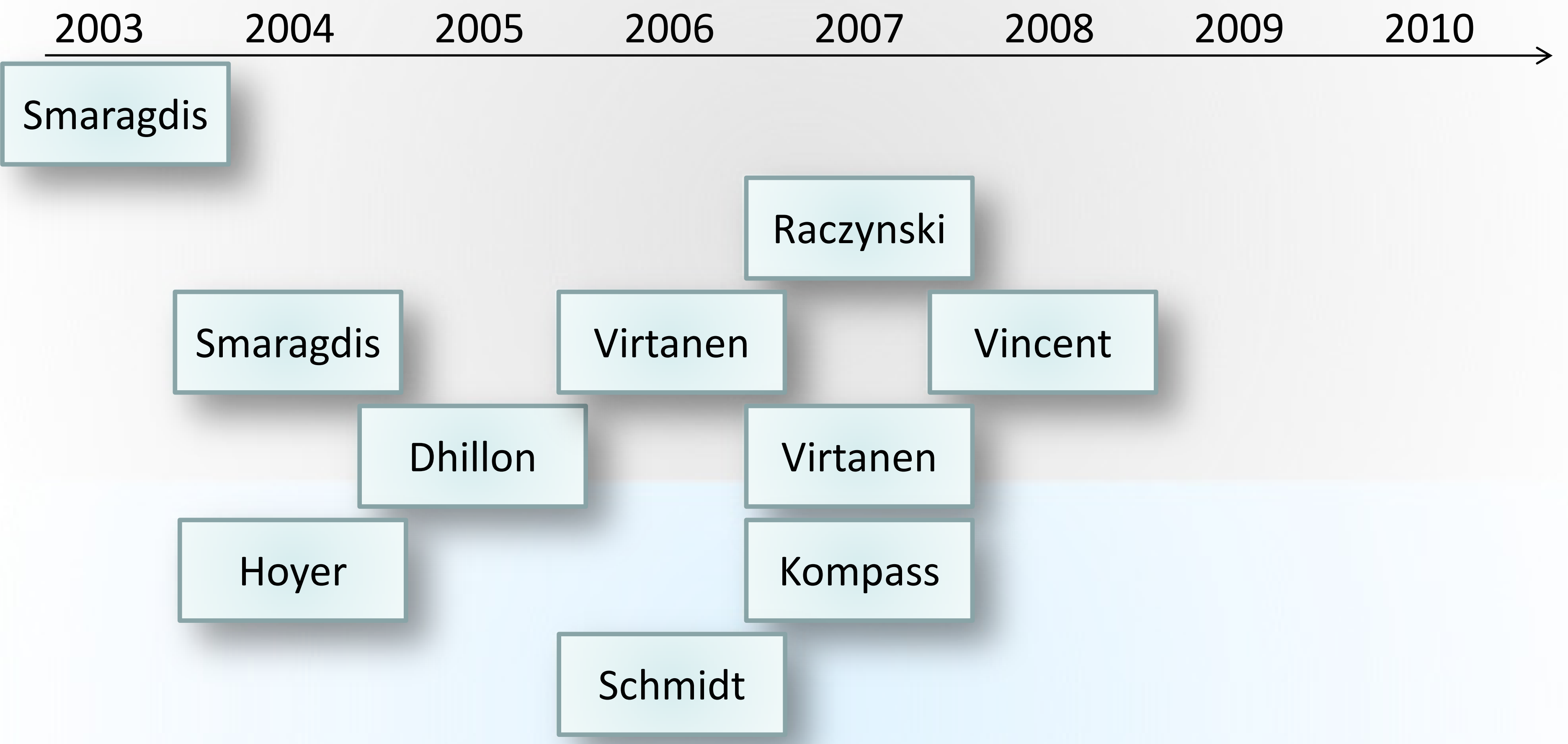
音名に対応する分だけ基底スペクトルを用意

Stanislaw Andrzej Raczynski, Nobutaka Ono, Shigeki Sagayama, "Multipitch analisys with Harmonic Nonnegative Matrix Approximation," *Proc. of ISMIR*, pp.381-386, 2007.

NMFの改良･拡張モデル

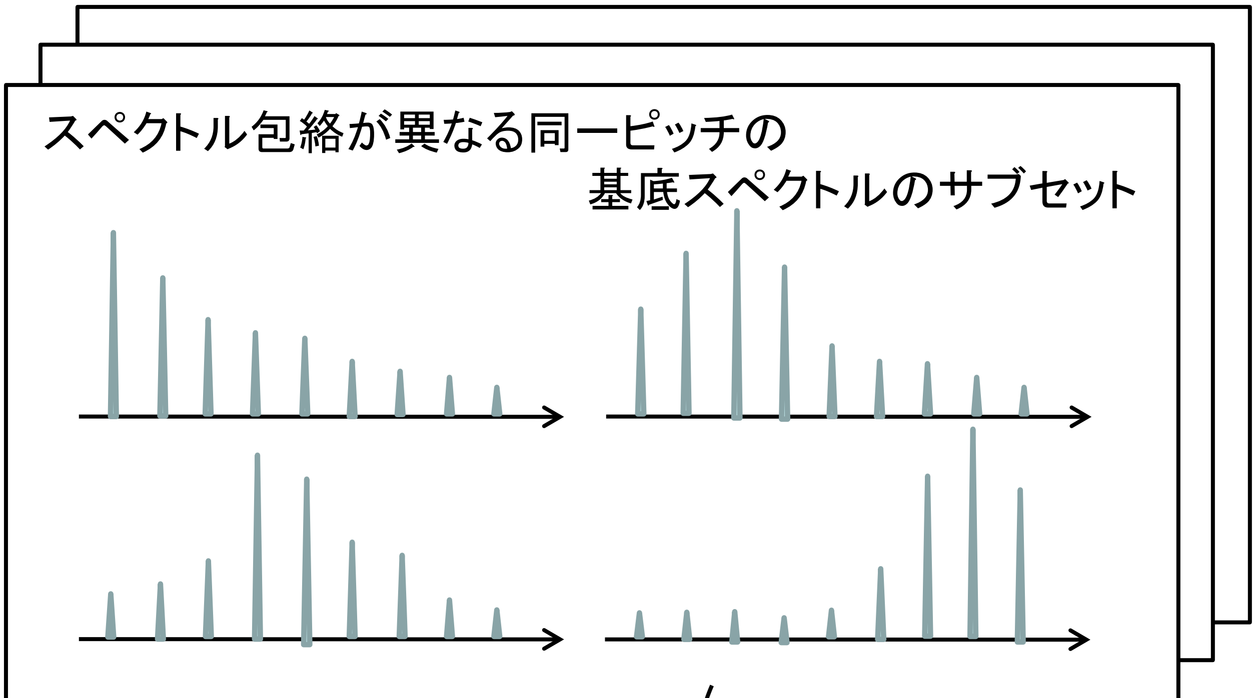


NMFの改良･拡張モデル

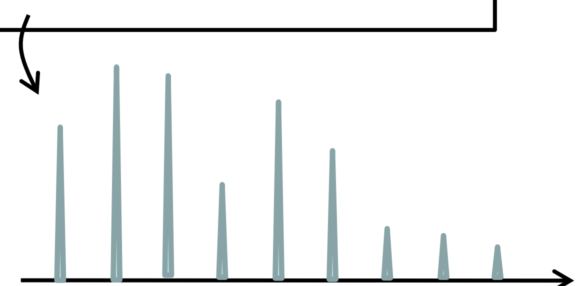


ハーモニックNMF [Vincent2008]

- 調波構造形の基底スペクトル

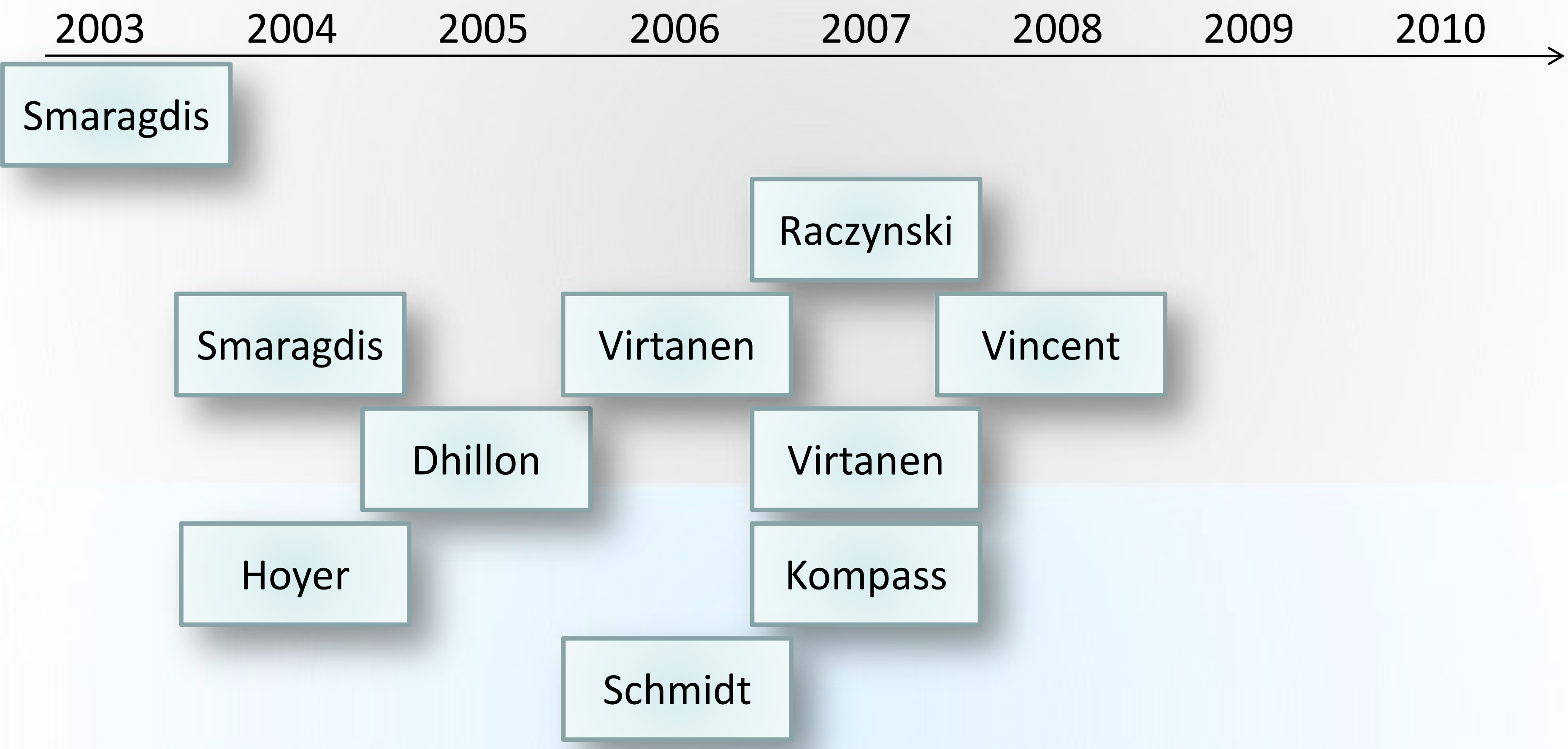


これらの重ね合わせでさまざま
なスペクトル包絡形状を表現

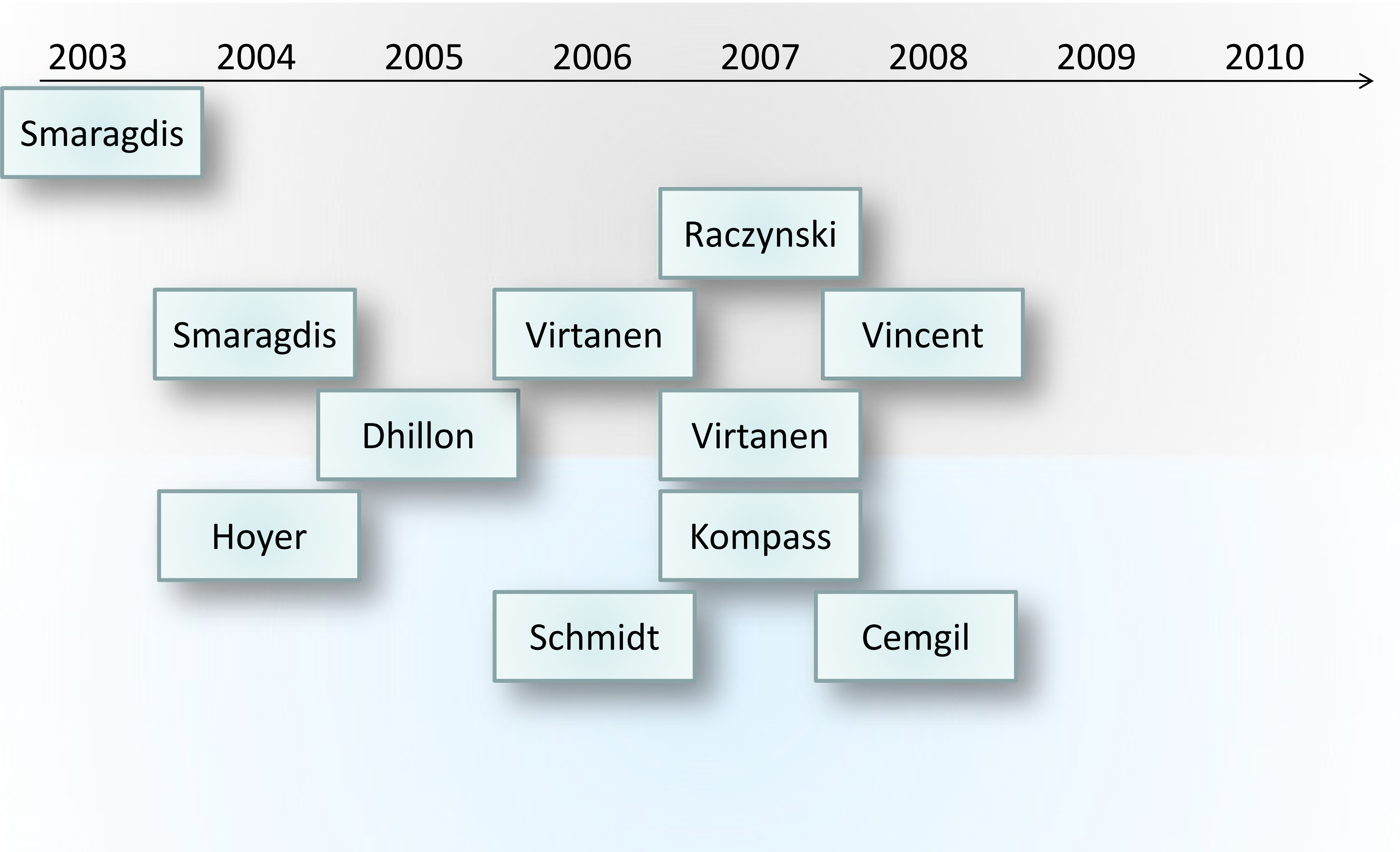


E. Vincent, N. Bertin, and R. Badeau, “Harmonic and inharmonic nonnegative matrix factorization for polyphonic pitch transcription,” in P*roc. ICASSP’08, pp. 109-112* , 2008

NMFの改良・拡張モデル



NMFの改良･拡張モデル



ベイジアンNMF [Cemgil2008]

- IダイバージェンスNMFを生成モデルの観点から解釈

$$\mathcal{J} = \sum_{\omega,t} \mathcal{D}\Big(Y_{\omega,t} \Big| \sum_k H_{\omega,k} U_{k,t}\Big)$$
$$= \sum_{\omega,t} \left(Y_{\omega,t} \log \frac{Y_{\omega,t}}{\sum_k H_{\omega,k} U_{k,t}} - Y_{\omega,t} + \sum_k H_{\omega,k} U_{k,t} \right)$$

$$\Updownarrow$$

$$Y_{\omega,t} \sim \mathrm{Poisson}\Big(\sum_k H_{\omega,k} U_{k,t}\Big)$$

$$\mathrm{Poisson}(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

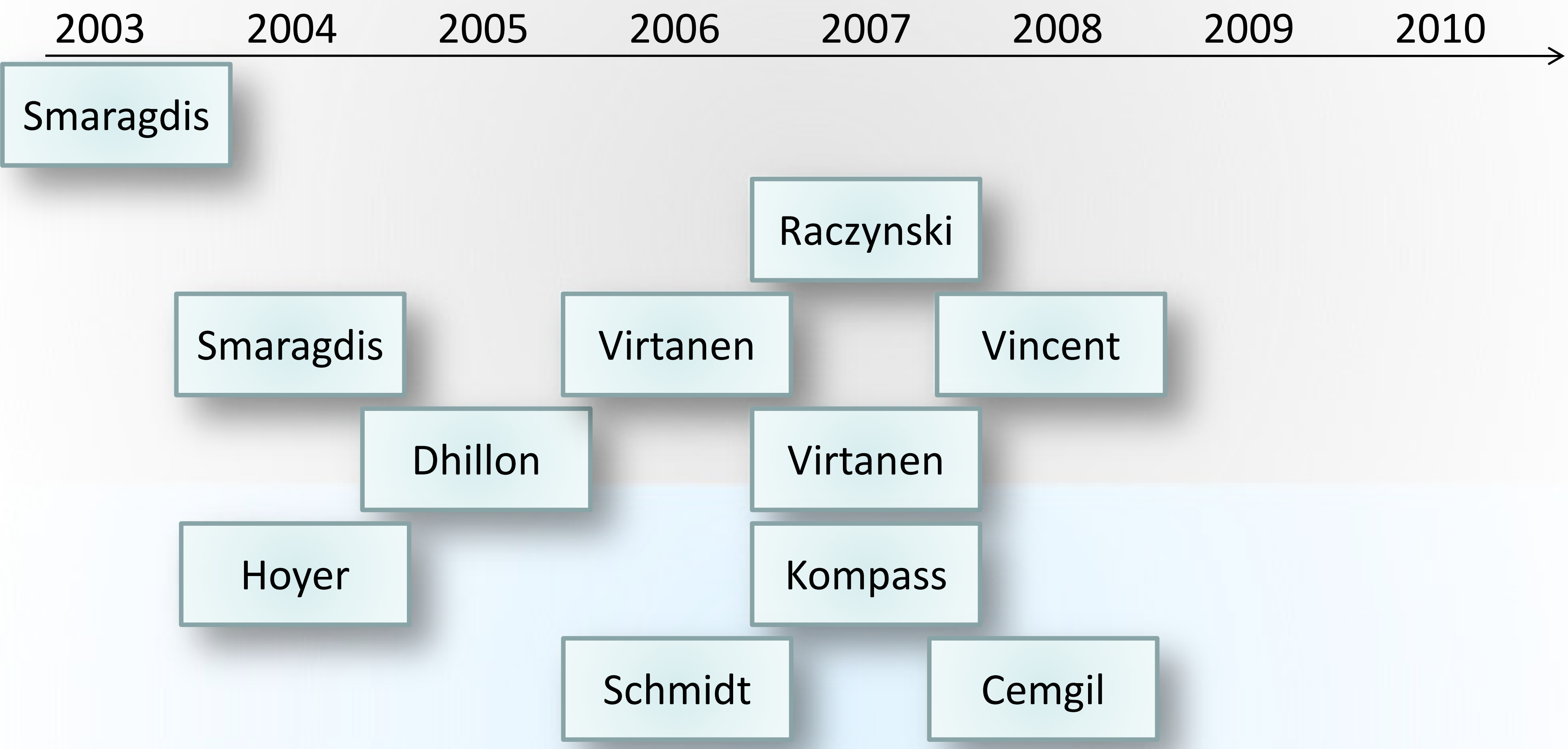
$$\Updownarrow$$

$$C_{\omega,t,k} \sim \mathrm{Poisson}(H_{\omega,k} U_{k,t})$$
$$Y_{\omega,t} = \sum_k C_{\omega,t,k}$$

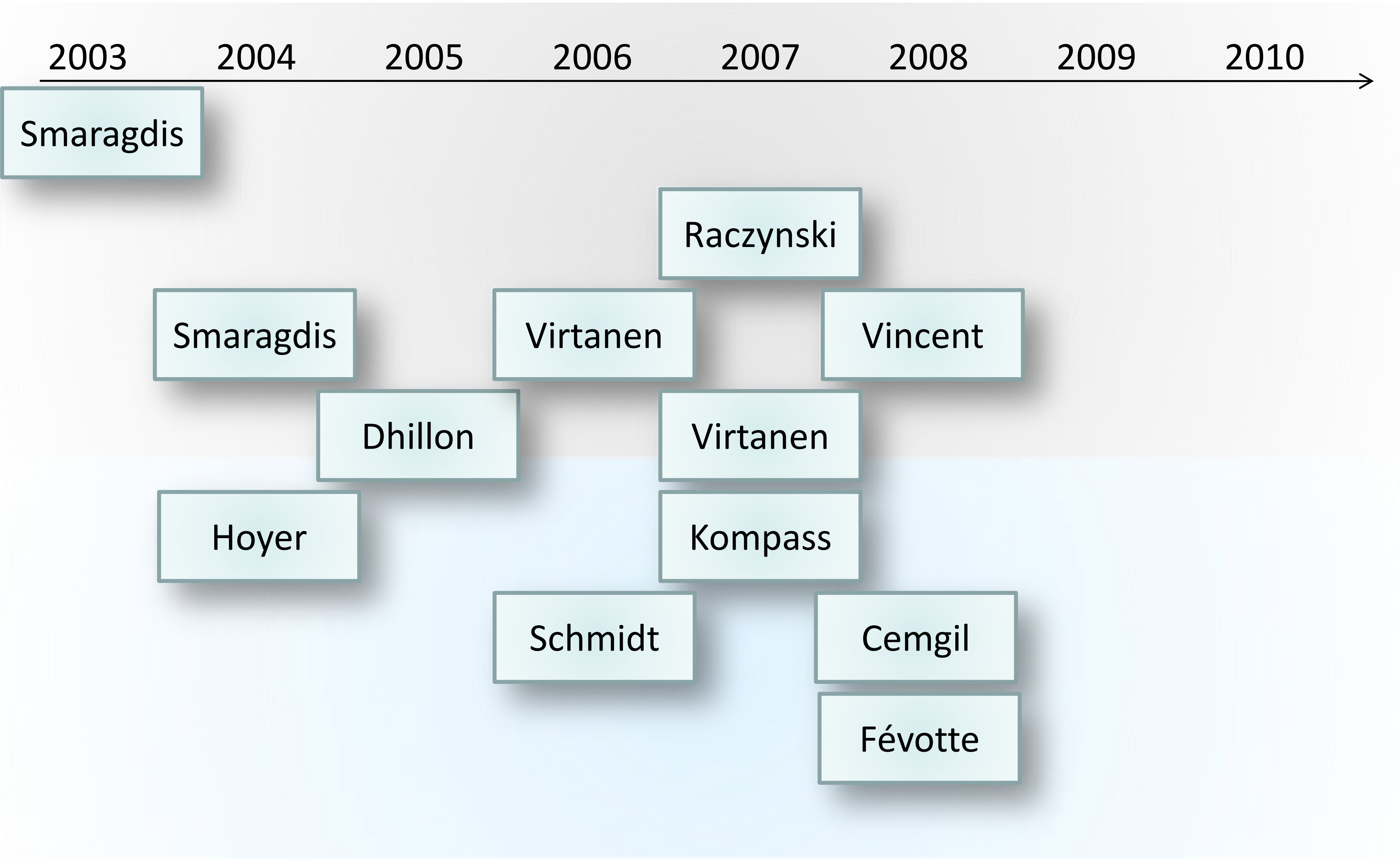
$$H_{\omega,k} \sim \mathrm{Gamma}(a,b)$$
$$U_{k,t} \sim \mathrm{Gamma}(c,d)$$

A.T. Cemgil, “Bayesian inference for nonnegative matrix factorization models,” Technical Report CUED/F-INFENG/TR.609, University of Cambridge, 2008.

NMFの改良･拡張モデル



NMFの改良・拡張モデル



板倉齋藤距離規準NMF [Févotte2008]

- 観測信号をガウス性信号の重ね合わせとしてモデル化
- パワースペクトル密度がNMF型の構造をもつと仮定

$$C_{\omega,t,k} \sim \mathcal{N}_{\mathbb{C}}(0, H_{\omega,k}U_{k,t})$$
$$Y_{\omega,t} = \sum_k C_{\omega,t,k}$$

→ ここでの $Y_{\omega,t}$ と $C_{\omega,t,k}$ は複素スペクトル成分（複素数値）であることに注意。

⇕

$$Y_{\omega,t} \sim \mathcal{N}_{\mathbb{C}}\Big(0, \sum_k H_{\omega,k}U_{k,t}\Big)$$

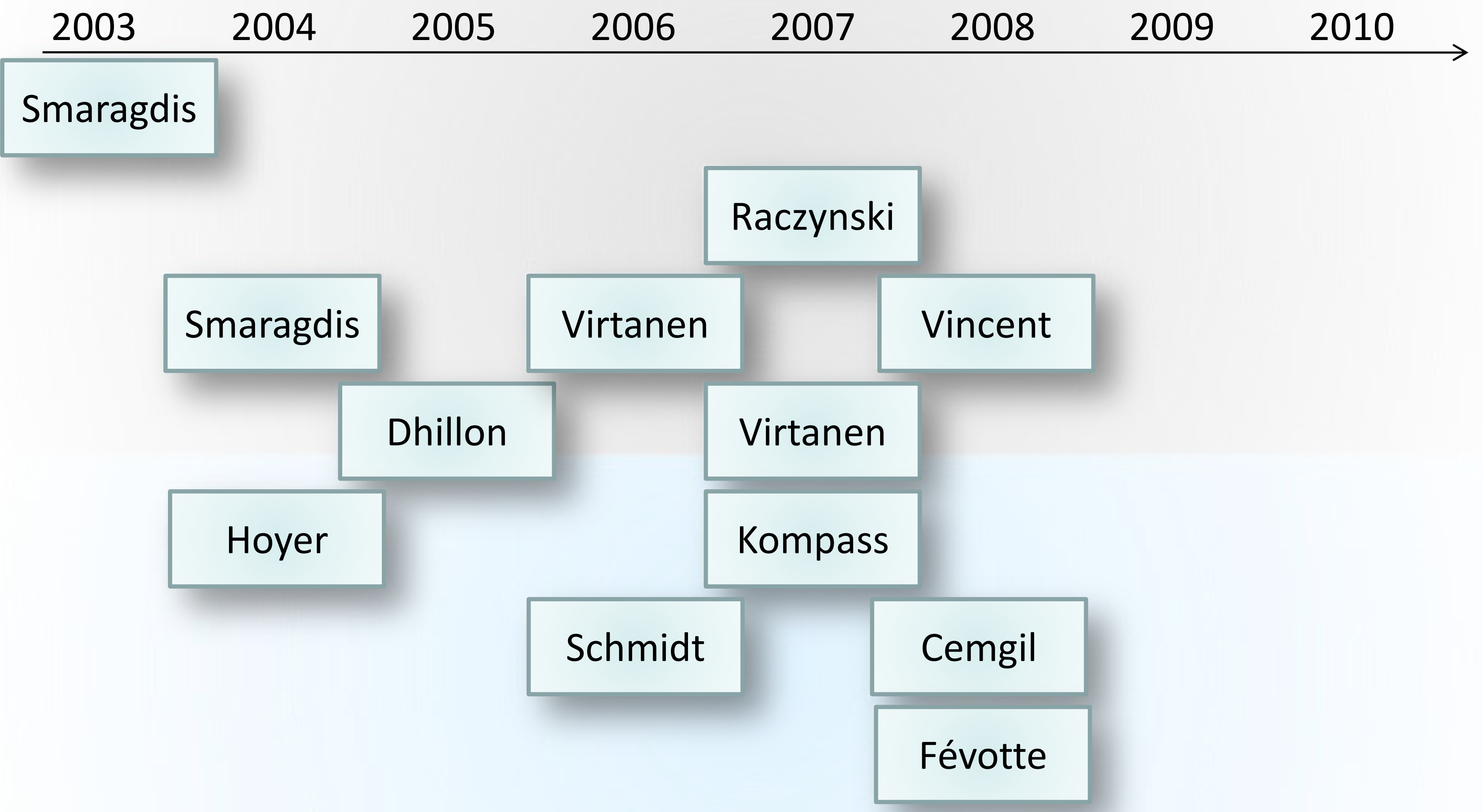
→ $C_{\omega,t,k}$ と $C_{\omega,t,k'}$ の独立性を仮定

⇕

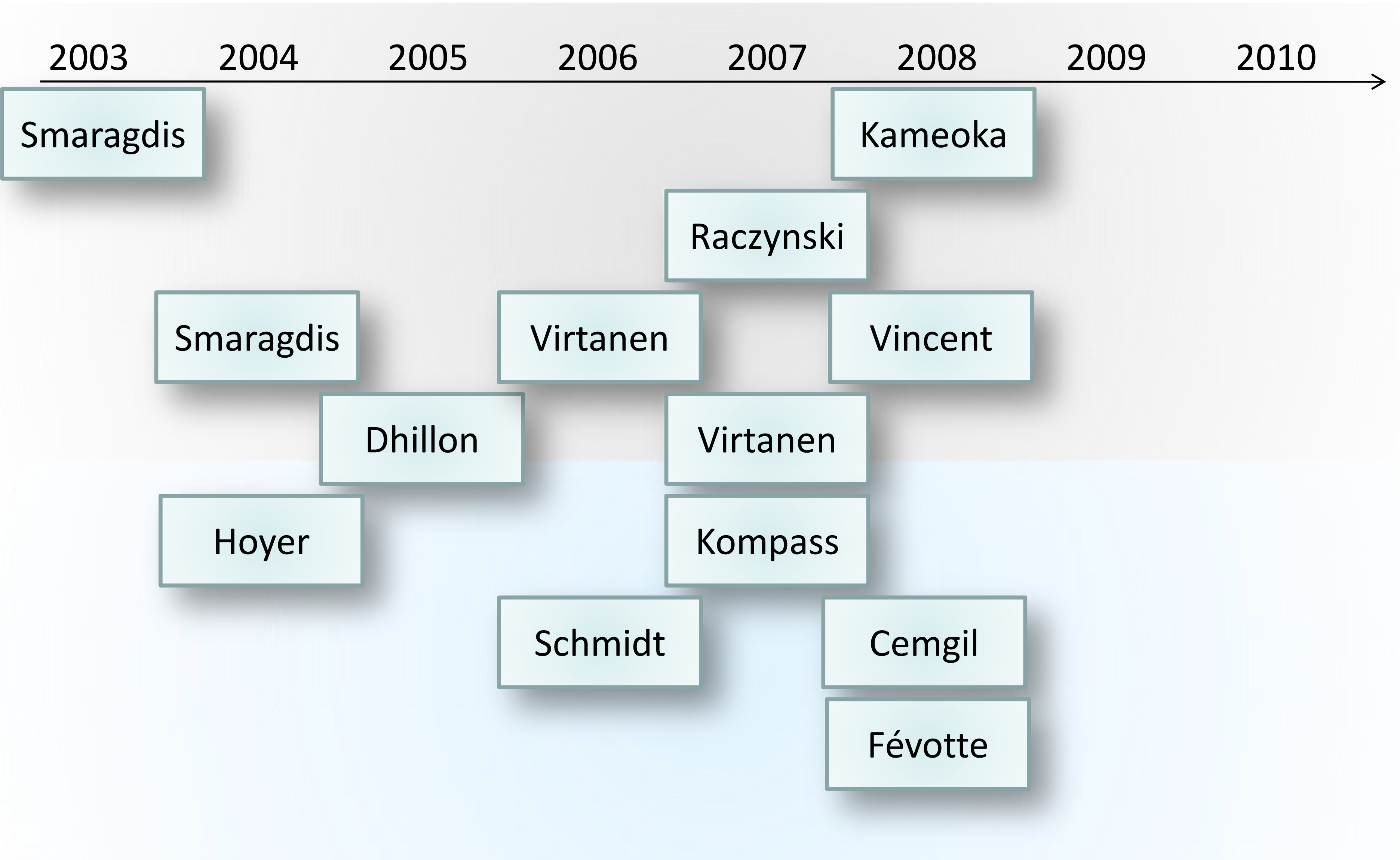
$$\mathcal{J} = \sum_{\omega,t} \mathcal{D}\Big(Y_{\omega,t}\Big|\sum_k H_{\omega,k}U_{k,t}\Big) = \sum_{\omega,t}\left(\frac{Y_{\omega,t}}{\sum_k H_{\omega,k}U_{k,t}} - \log\frac{Y_{\omega,t}}{\sum_k H_{\omega,k}U_{k,t}} - 1\right)$$

C. Févotte, N. Bertin, and J.-L. Durrieu, “Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis,” *Neural Computation, vol. 21, no. 3,* Mar. 2009,

NMFの改良･拡張モデル



NMFの改良･拡張モデル



複素NMF [Kameoka2008]

- パワースペクトルも振幅スペクトルも本来は非加法的
- 複素スペクトル領域モデルでNMFと同様のスペクトルパーツ獲得機能が実現できないか？

$$x(u) = \sum_k c^{(k)}(u)$$

⇩ **(STFT)**

$$X_{\omega,t} = \sum_k C_{\omega,t,k}$$

⇩ **(magnitude)**

$$|X_{\omega,t}| = \left|\sum_k C_{\omega,t,k}\right|$$

$$\neq \sum_k |C_{\omega,t,k}|$$

NMFにおけるモデルが厳密には不適切であることを示している

要素信号の複素スペクトログラム $C_{\omega,t,k}$ の振幅成分が“rank1”構造をもつように混合信号をモデル化

$$X_{\omega,t} = \sum_k |C_{\omega,t,k}| e^{j\phi_{\omega,t,k}}$$

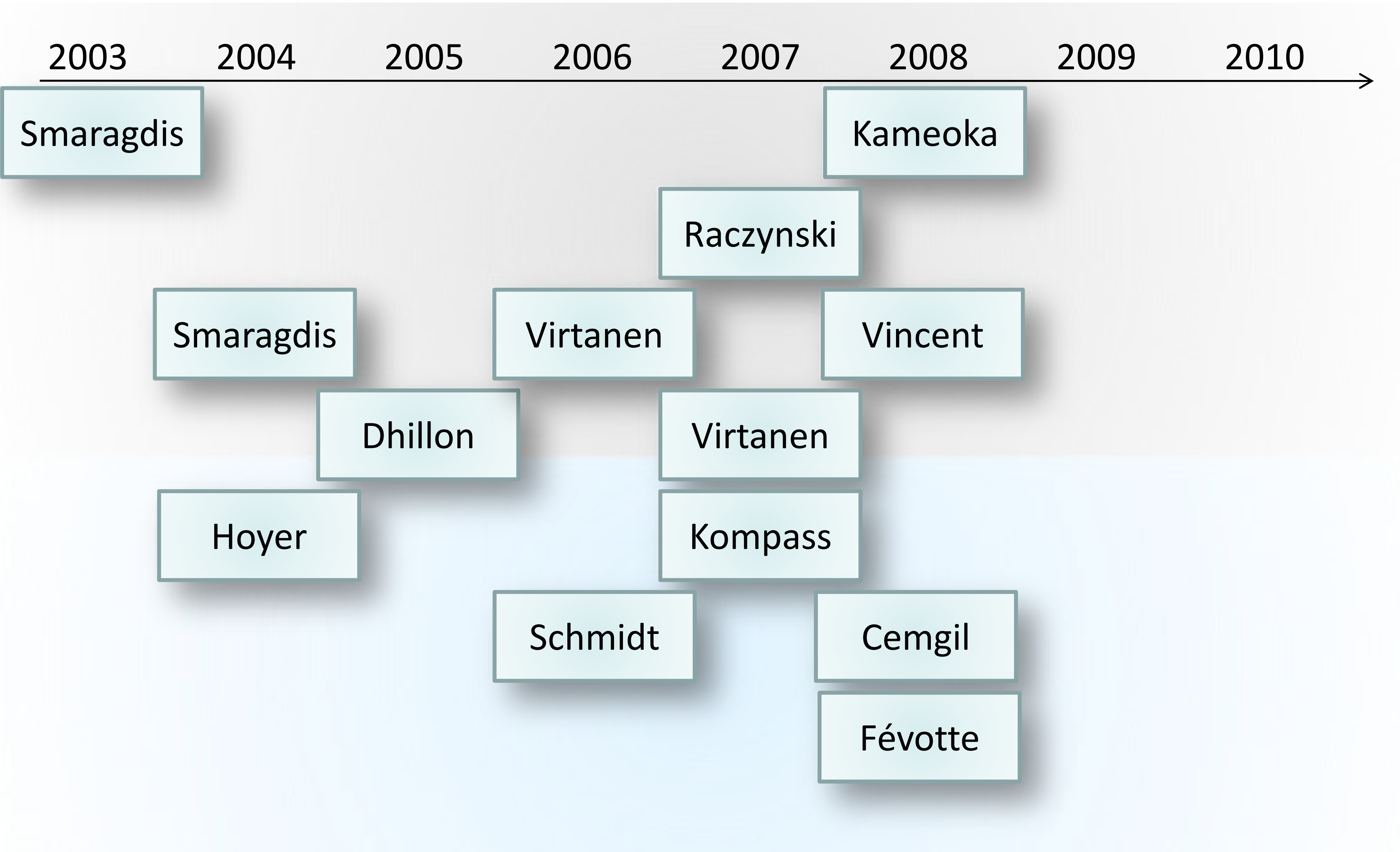
$$= \sum_k H_{\omega,k} U_{k,t} e^{j\phi_{\omega,t,k}}$$

$$\sum_{\omega,t} |Y_{\omega,t} - X_{\omega,t}|^2 \to \text{minimum}$$

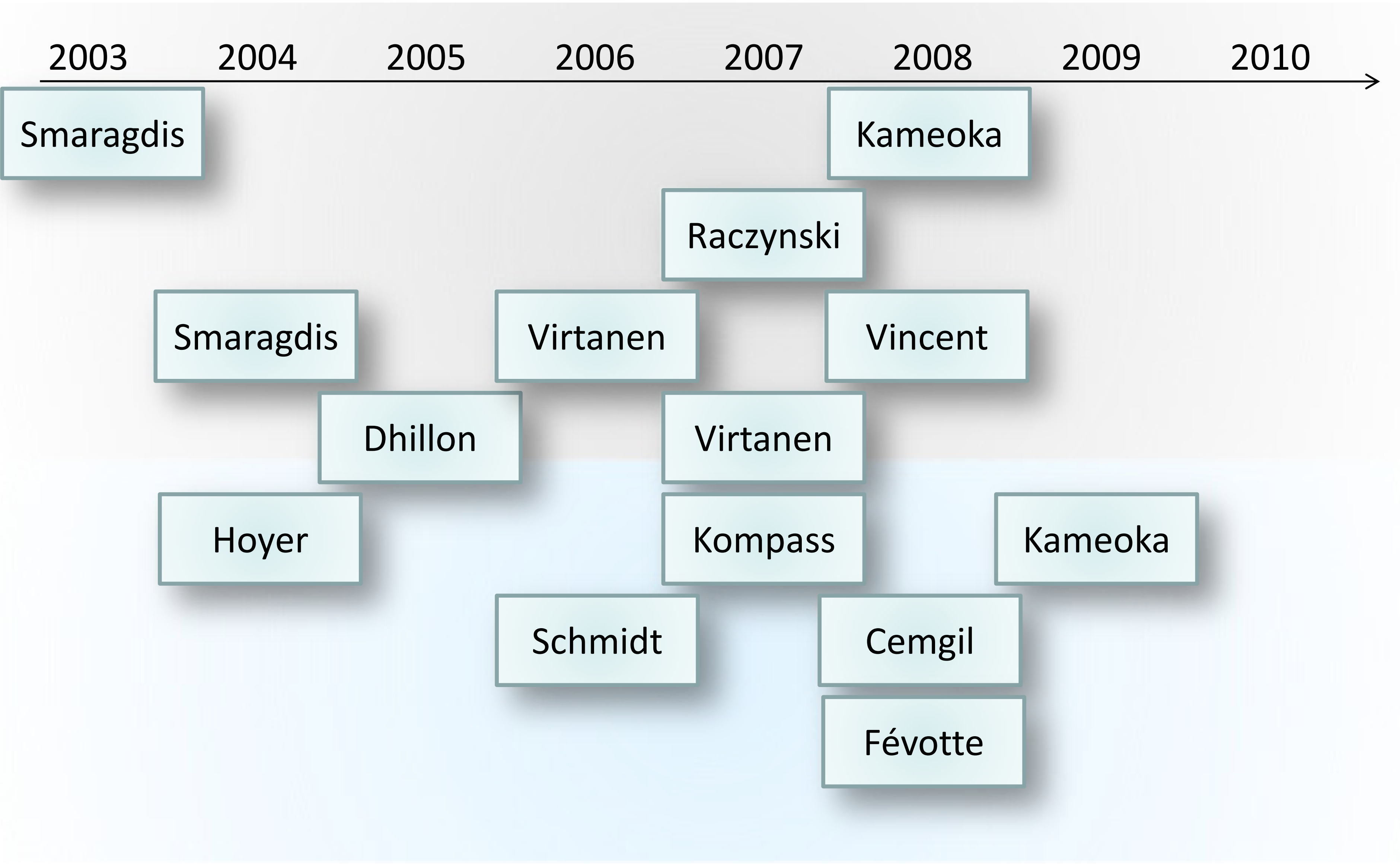
ここでの $Y_{\omega,t}$ と $C_{\omega,t,k}$ も複素スペクトル成分（複素数値）であることに注意。

亀岡, 小野, 柏野, 嵯峨山, “複素NMF: 新しいスパース信号分解表現と基底系学習アルゴリズム,” 日本音響学会2008年秋季研究発表会講演論文集, 2-8-13, pp. 657-660, 2008.

NMFの改良・拡張モデル



NMFの改良･拡張モデル



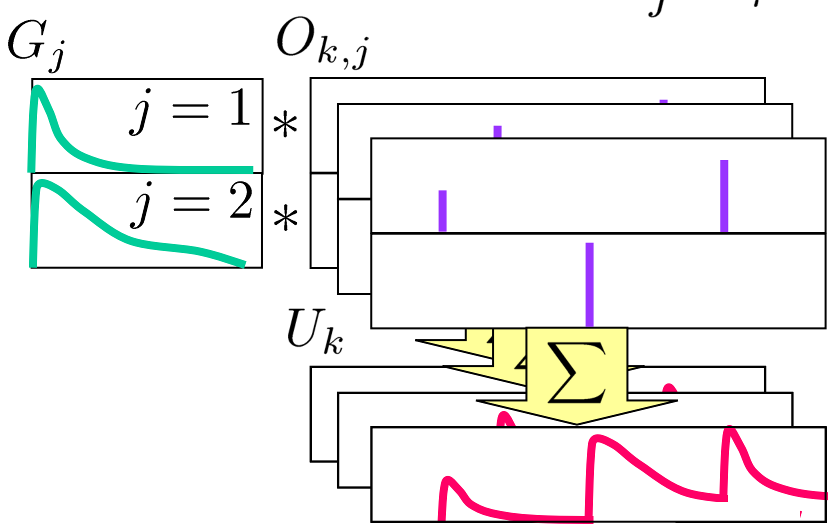
MusicFactorizer [Kameoka2009]

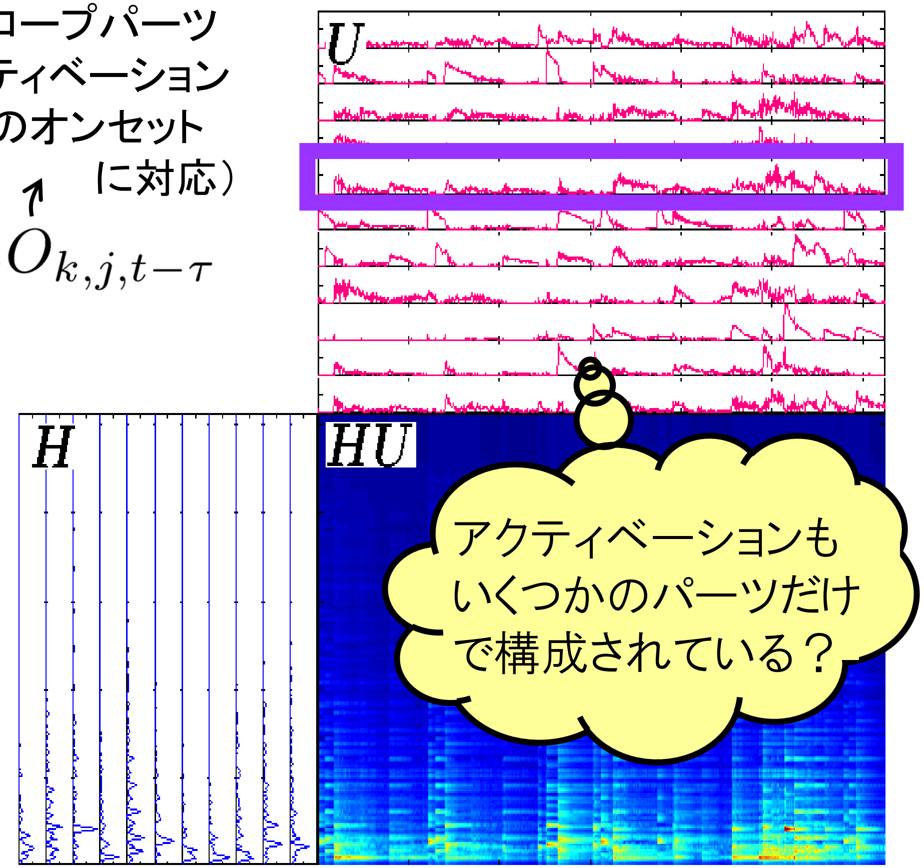
- 音楽信号は限られた種類の音価の音だけで構成される
  →NMFのアクティベーションもまたパーツベースド表現ができる？
- アクティベーションをエンベロープパーツの畳み込み混合でモデル化

$$Y_{\omega,t} \simeq \sum_k H_{\omega,k} \underline{U_{k,t}}$$

エンベロープパーツ
のアクティベーション
（ノートのオンセット
に対応）

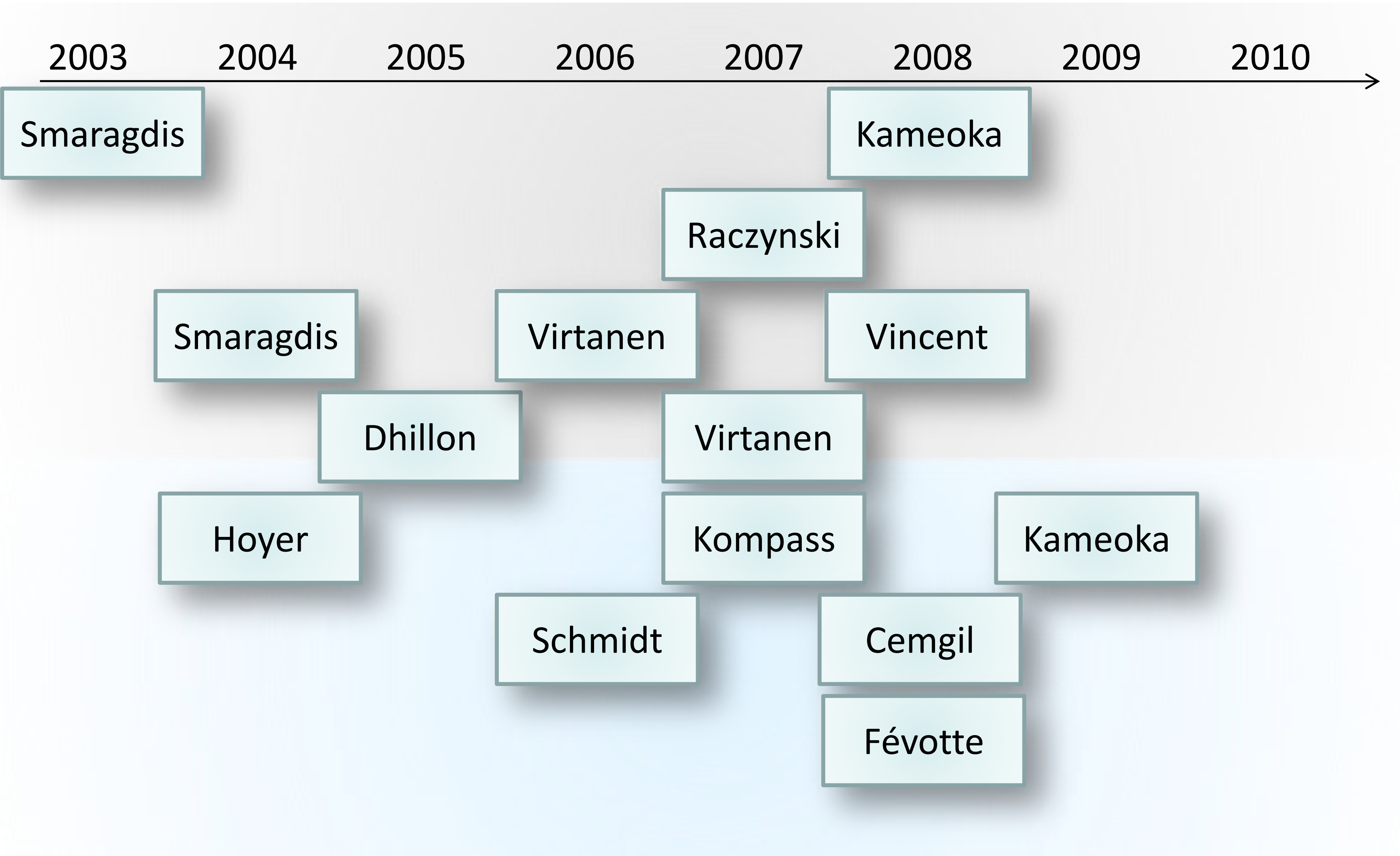
「畳み込み混合」 $U_{k,t} = \sum_j \sum_\tau G_{j,\tau} O_{k,j,t-\tau}$



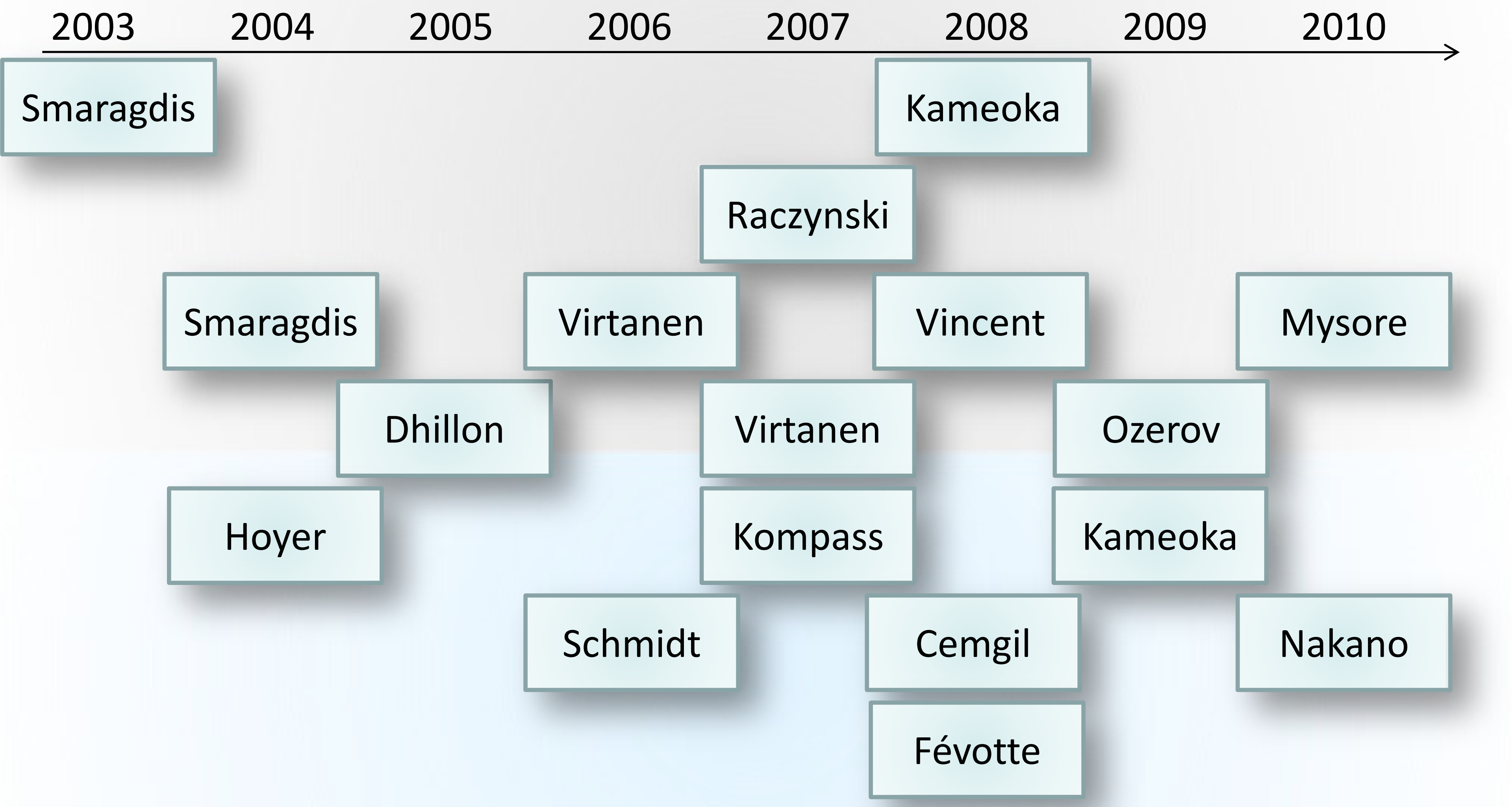


亀岡, ルルー, 大石, 柏野, "Music Factorizer: 音楽音響信号をノート単位で編集できるインタフェース," 情報処理学会研究報告, 2009-MUS-81-9, Jul. 2009.

NMFの改良・拡張モデル

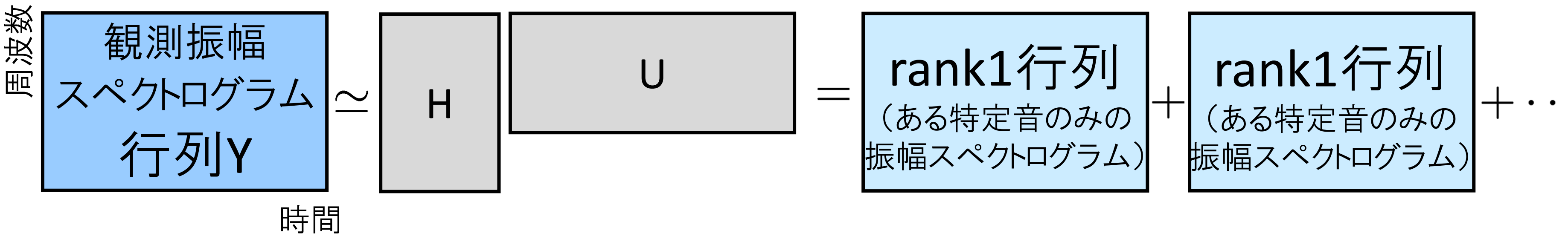


NMFの改良・拡張モデル



状態遷移NMF [Ozerov2009,Nakano2010,Mysore2010]

- NMF：スペクトログラムをrank1行列の和に分解していることに相当



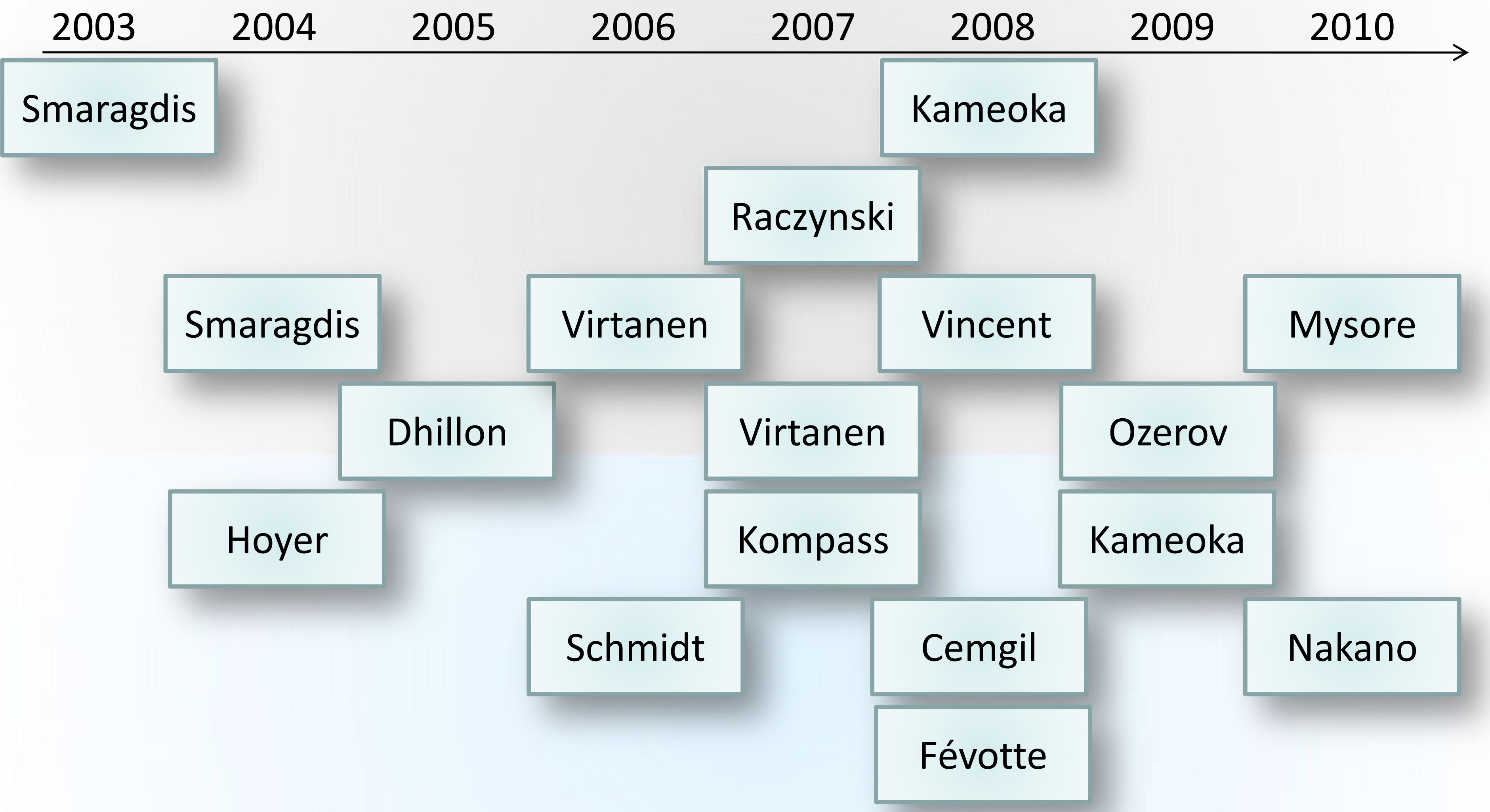
  – 楽音のスペクトログラムは実際にはrank1にならない
  （ピアノ音にattack/decay/sustainなどの複数状態があるように，
  通常，楽音スペクトルの形は発音中に多様に時間変化する）

- 各楽音スペクトログラムをrank1行列でモデル化する代わりに
  非負値を出力する隠れマルコフモデル（HMM）でモデル化

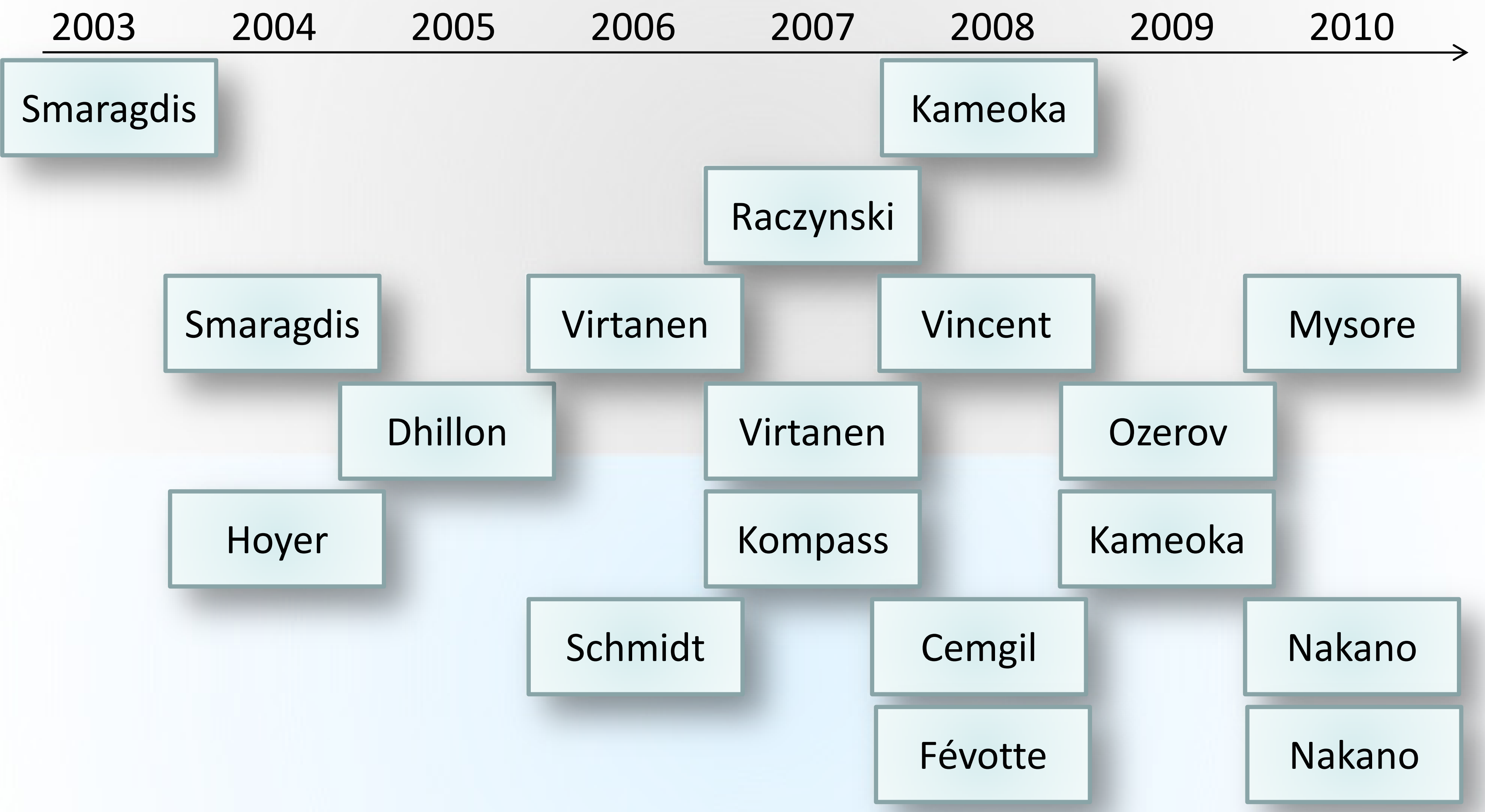
→Factorial HMM [Ghahramani1997]と同形のモデルになる

A. Ozerov, C. Févotte, M. Charbit, “Factorial scaled hidden Markov model for polyphonic audio representation and source separation,” in *Proc. WASPAA'09*, 2009.
中野, 北野, ルルー, 亀岡, 小野, 嵯峨山, "可変基底NMFに基づく音楽音響信号の解析," 情報処理学会研究報告, 2010-MUS-84-10, Feb. 2010.
G. J. Mysore, P. Smaragdis, B. Raj, “Non-negative hidden Markov modeling of audio with application to source separation,” in *Proc. LVA/ICA'2010*. pp.140-148, 2010.

NMFの改良･拡張モデル



NMFの改良·拡張モデル



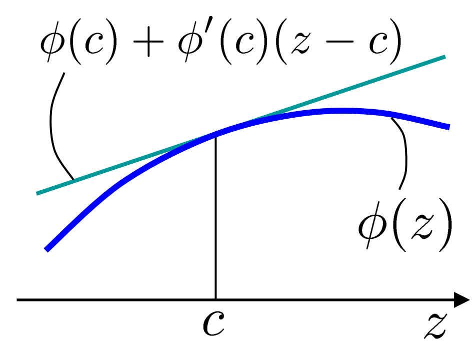
βダイバージェンス規準NMF [Nakano2010]

- すべてのβで収束性が保証された乗法更新アルゴリズム

$$\mathcal{J} = \sum_{\omega,t} \mathcal{D}_\beta \left(Y_{\omega,t} \middle| \sum_k H_{\omega,k} U_{k,t} \right)$$

$$= \sum_{\omega,t} \left[\frac{Y_{\omega,t}^{\beta}}{\beta(\beta-1)} + \frac{1}{\beta} \underbrace{\left(\sum_k H_{\omega,k} U_{k,t} \right)^{\beta}}_{\text{凸/凹}} - \frac{Y_{\omega,t}}{\beta-1} \underbrace{\left(\sum_k H_{\omega,k} U_{k,t} \right)^{\beta-1}}_{\text{凸/凹}} \right]$$

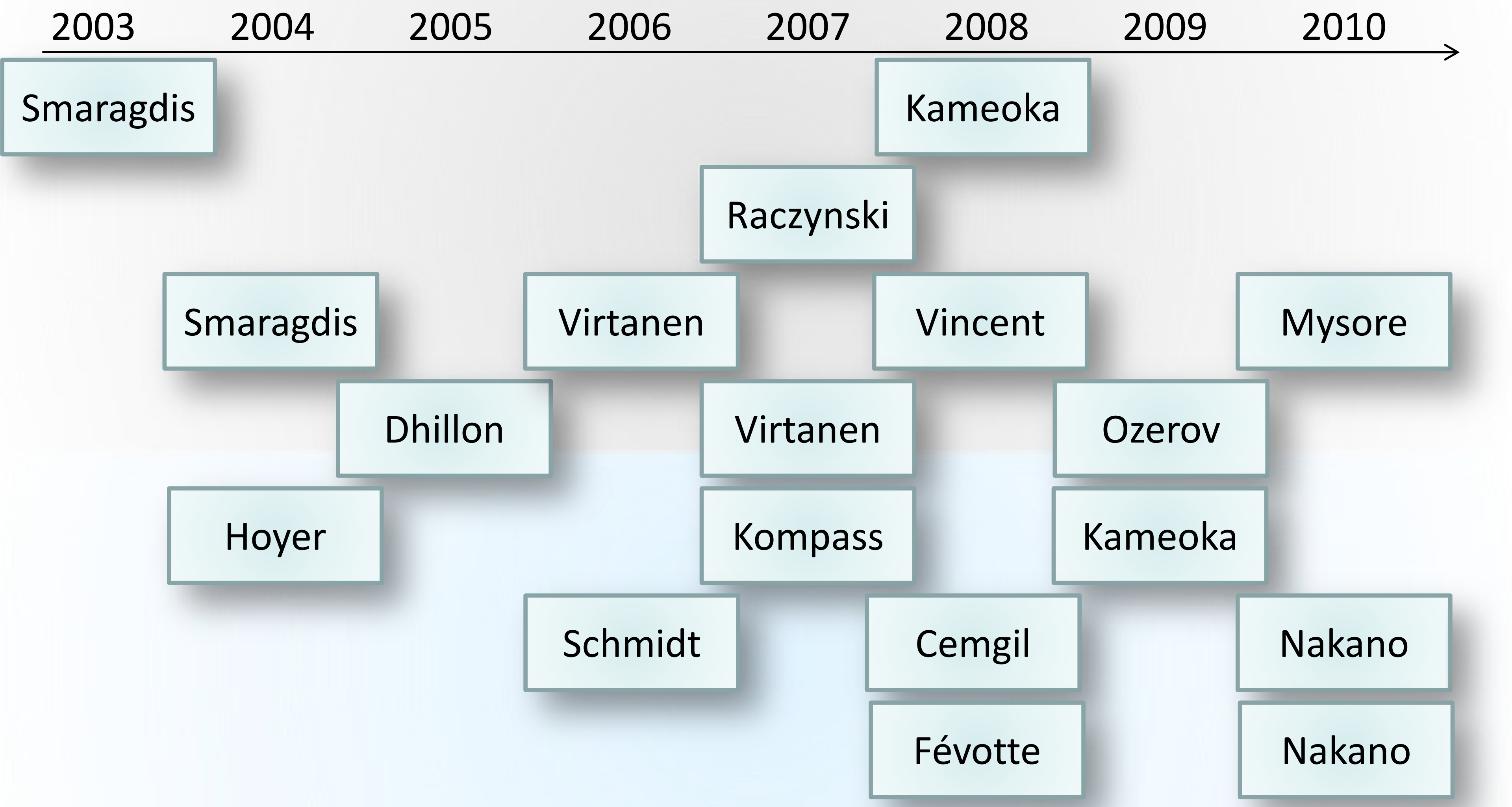
- 補助関数の設計方針
  - Jensen不等式 $\phi\left(\sum_i \lambda_i z_i\right) \leq \sum_i \lambda_i \phi(z_i)$ $\left\{ \begin{array}{l} \phi : \text{凸関数} \\ \sum_i \lambda_i = 1 \end{array} \right.$
  - 接線不等式 $\phi(z) \leq \phi(c) + \phi'(c)(z-c)$ ϕ : 凹関数



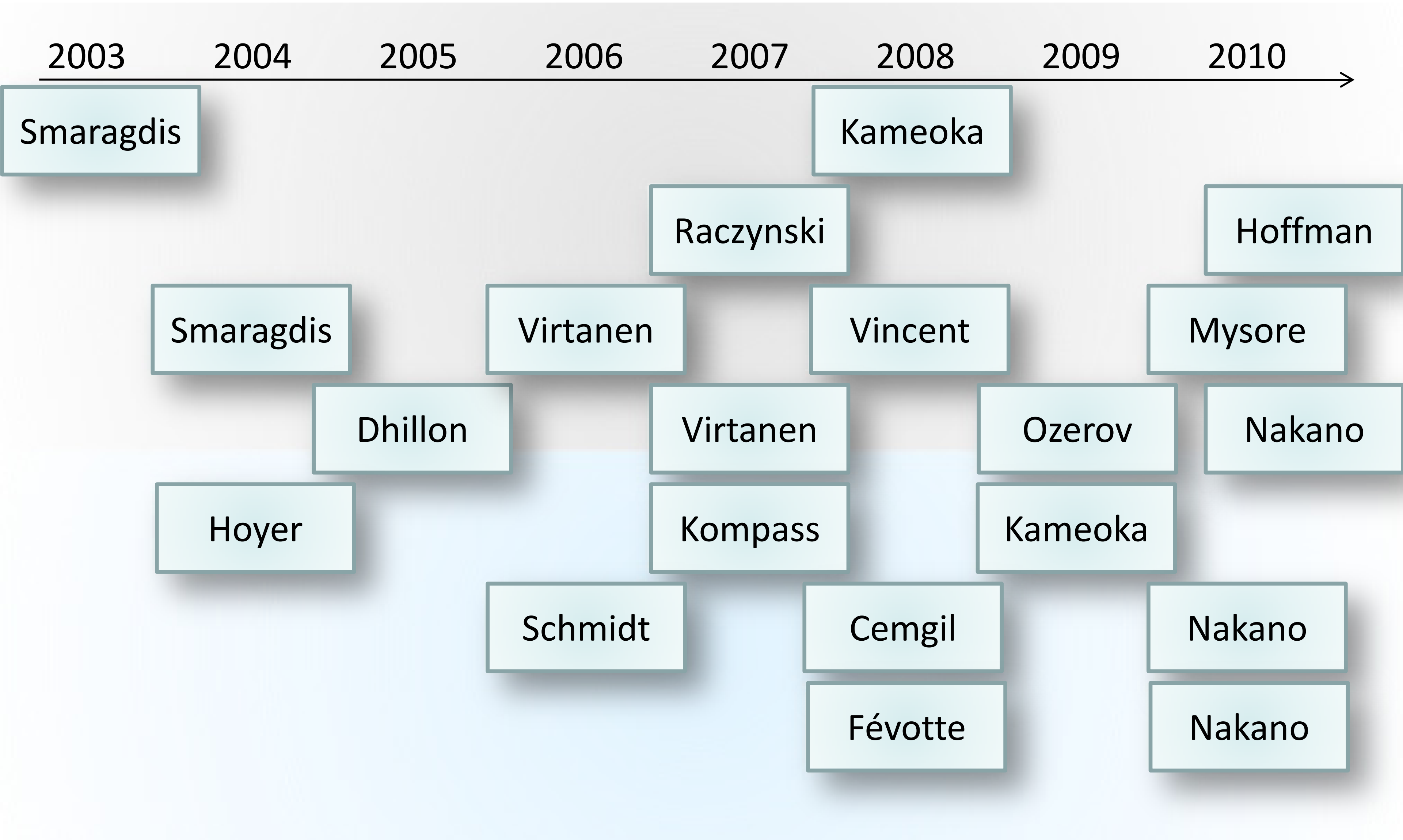
凸関数項の上限を凸不等式，凹関数項の上限を接線不等式を使って設計し，補助関数を設計[Kameoka2006]

M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono, S. Sagayama, “Convergence-guaranteed multiplicative algorithms for non-negative matrix factorization with beta-divergence,” In *Proc. MLSP 2010*, in CD-ROM, 2010.

NMFの改良･拡張モデル



NMFの改良･拡張モデル

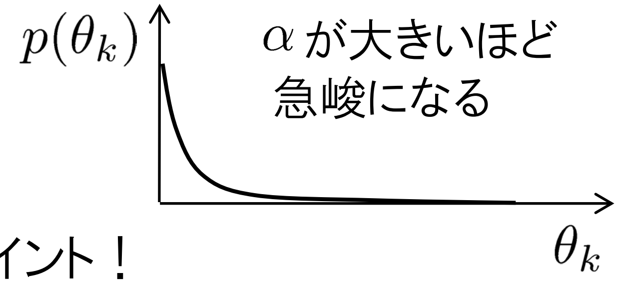


ノンパラメトリックベイズNMF [Hoffman2011, Nakano2011]

- NMFにおける基底数をデータから推論できるようにしたい

→ Gamma Process (GaP) NMF [Hoffman2011]

  - スペクトログラムの生成モデル



$$H_{\omega,k} \sim \mathrm{Gamma}(a, a)$$
$$U_{k,t} \sim \mathrm{Gamma}(b, b)$$
$$\theta_k \sim \mathrm{Gamma}(\alpha/K, \alpha c)$$
$$Y_{\omega,t} \sim \mathrm{Exponential}\Big(\sum_k \theta_k H_{\omega,k} U_{k,t}\Big)$$

→ ここがポイント！

→FévotteのIS-NMF方式

  - $\theta_l \sim \mathrm{Gamma}(\alpha/K, \alpha c)$の意味
    - $\theta_1, \theta_2, \ldots$をスパースにする

      ＝不要なコンポーネントはできるだけ使わない

- 状態遷移NMFにおけるコンポーネント数と状態数をデータから推論できるようにしたい

→ Infinite Factorial infinite Hidden Markov Model (iFiHMM) [Nakano2011]

まとめ

1. 非負値行列因子分解（NMF）とは
   – 何に使えるのか（音響信号処理を題材として）
   – どのような性質があるのか
   – どのように求めるのか
   – 統計モデルとしての解釈

2. NMFの改良・拡張モデル
   – スパースNMF, NMFD, NMF2D, ソースフィルタNMF, アクティベーション連続性規準入りNMF, ハーモニックNMF, 板倉齋藤距離規準NMF, 複素NMF, 状態遷移NMF, ノンパラメトリックベイズNMF, etc...

参考文献

[1] D. D. Lee and H. S. Seung, “Learning the parts of objects with nonnegative matrix factorization,” *Nature*, vol. 401, pp. 788–791, 1999.

[2] P. Paatero and U. Tapper, “Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values,” *Environmetrics* **5**: 111–126, 1994.

[3] P. Smaragdis and J. C. Brown, “Non-negative matrix factorization for music transcription,” Proc. *WASPAA 2003*, pp. 177–180, 2003.

[4] D. D. Lee and H. S. Seung, “Algorithms for nonnegative matrix factorization,” in Adv. *NIPS 2000*, pp. 556–562, 2000.

[5] P. Smaragdis, B. Raj and M.V. Shashanka, “Supervised and semi-supervised separation of sounds from single-channel mixtures," in Proc. *ICA 2007*, pp. 414-421, 2007.

[6] P. Smaragdis and B. Raj, ”Example-driven bandwidth expansion,” in Proc. *WASPAA 2007*, pp. 135-138, 2007.

[7] P. O. Hoyer, “Non-negative matrix factorization with sparseness constraints,” *J. Mach. Learning Res., vol. 5, pp. 1457–1469, 2004.*

[8] P. Smaragdis, “Non-negative matrix factor deconvolution; extraction of multiple sound sources from monophonic inputs,” in *Proc. ICA 2004, pp. 494–499, 2004.*

[9] L. M. Bregman, “The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming,” *USSR Computational Mathematics and Mathematical Physics* **7**: pp. 200–217, 1967.

参考文献

[10] M. N. Schmidt and M. Mørup, "Nonnegative matrix factor 2-D deconvolution for blind single channel source separation," in Proc. *ICA 2006, pp. 700-707, 2006.*

[11] T. Virtanen and A. Klapuri, "Analysis of polyphonic audio using source-filter model and non-negative matrix factorization," in *Adv. NIPS 2006,* 2006.

[12] S. Eguchi and Y. Kano, "Robustifying maximum likelihood estimation," Technical report, Institute of Statistical Mathematics, Research Memo. 802, 2001.

[13] R. Kompass, "A generalized divergence measure for nonnegative matrix factorization," *Neural Computation*, vol. 19, no. 3, pp. 780–791, 2007.

[14] T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," *IEEE Trans. on Audio, Speech and Language Processing, vol. 15, no. 3, pp. 1066*–1074, 2007.

[15] S. A. Raczynski, N. Ono and S. Sagayama, "Multipitch analisys with harmonic nonnegative matrix approximation," in *Proc. of ISMIR 2007*, pp.381-386, 2007.

[16] E. Vincent, N. Bertin and R. Badeau, "Harmonic and inharmonic nonnegative matrix factorization for polyphonic pitch transcription," in P*roc. ICASSP 2008, pp. 109-112 ,* 2008.

[17] A. T. Cemgil, "Bayesian inference for nonnegative matrix factorization models," Technical Report CUED/F-INFENG/TR.609, University of Cambridge, 2008.

[18] C. Févotte, N. Bertin and J.-L. Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis," *Neural Computation,* vol. 21, no. 3, pp. 793-830, 2009.

[19] 亀岡, 小野, 柏野, 嵯峨山, "複素NMF: 新しいスパース信号分解表現と基底系学習アルゴリズム," 日本音響学会2008年秋季研究発表会講演論文集, 2-8-13, pp. 657-660, 2008.

参考文献

[20] 亀岡, ルルー, 大石, 柏野, "Music Factorizer: 音楽音響信号をノート単位で編集できるインタフェース," 情報処理学会研究報告, 2009-MUS-81-9, 2009.

[21] Z. Ghahramani and M. I. Jordan, “Factorial hidden Markov models,” Machine Learning, vol. 29, pp. 245-273, 1997.

[22] A. Ozerov, C. Févotte and M. Charbit, “Factorial scaled hidden Markov model for polyphonic audio representation and source separation,” in *Proc. WASPAA 2009*, pp. 121-124, 2009.

[23] 中野, 北野, ルルー, 亀岡, 小野, 嵯峨山, "可変基底NMFに基づく音楽音響信号の解析," 情報処理学会研究報告, 2010-MUS-84-10, 2010.

[24] G. J. Mysore, P. Smaragdis and B. Raj, “Non-negative hidden Markov modeling of audio with application to source separation,” in *Proc. LVA/ICA 2010*. pp.140-148, 2010.

[25] 亀岡, 後藤, 嵯峨山, “スペクトル制御エンベロープによる混合音中の周期および非周期成分の選択的イコライザ,” 情報処理学会研究報告, 2006-MUS-66-13, pp. 77-84, 2006.

[26] M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono and S. Sagayama, “Convergence-guaranteed multiplicative algorithms for non-negative matrix factorization with beta-divergence,” In *Proc. MLSP 2010*, pp. 283-288, 2010.

[27] M. Hoffman, D. Blei and P. Cook, “Bayesian nonparametric matrix factorization for recorded music,” in *Proc. ICML,* pp.439-446, 2010.

[28] 中野, ルルー, 亀岡, 中村, 小野, 嵯峨山, “スペクトログラムのベイジアンノンパラメトリックモデリングに基づく音楽信号の解析,” 情報処理学会研究報告, 2011-MUS-91-6, 2011.