Revisiting *spatiotemporal imaging with partially separable functions*

Kaibo Tang

January 8, 2026

1 Theory

Instead of proving the more general Theorem 2.1 in [Zhi07], we present a more practical case by considering the measure spaces to be $\mathbb{R}^n$ and $\mathbb{R}^m$, which admits a more elementary and easier to follow proof.

Theorem 1.1. *Let $n, m \in \mathbb{N}$. The set S of all finite sums of functions of the form $f(x,y) = g(x)h(y)$, where $g \in L^2(\mathbb{R}^n), h \in L^2(\mathbb{R}^m)$, is dense in $L^2(\mathbb{R}^{n+m})$.*

Proof.

1. (Reduction to approximation of simple functions). It is standard[1] that the set of simple functions is dense in $L^2(\mathbb{R}^{n+m})$.[2] Hence, we only need to consider $\varphi \in L^2(\mathbb{R}^{n+m})$ simple.

2. (Reduction to approximation of characteristic functions). Let

$$\varphi = \sum_{j=1}^{N} c_j \chi_{A_j} \tag{1}$$

be the canonical representation of φ. I claim that it suffices to show that we can approximate characteristic functions of finite measurable

[1] Recall that the Lebesgue integral is defined using the integral of simple functions. A reference for this can be found in almost any text in real analysis.

[2] In fact, this holds for any L^p space where $1 \leq p \leq \infty$ and of course with respect to the corresponding L^p-norm.

sets by functions in S. Indeed, let $c = \max_j\{|c_j|\}$. For any fixed $\varepsilon > 0$, we can always find $f_j \in S$ such that[3]

$$\|f_j - \chi_{A_j}\|_2 < \frac{\varepsilon}{cN}. \tag{2}$$

Let

$$f = \sum_{j=1}^{N} c_j f_j. \tag{3}$$

Then,

$$\|f - \varphi\|_2 \leq \sum_{j=1}^{N} |c_j| \|f_j - \chi_{A_j}\|_2 < \varepsilon. \tag{4}$$

3. Let μ denote the Lebesgue measure on $\mathbb{R}^{n+m}$. Fix $E \subset \mathbb{R}^{n+m}$ measurable with $\mu(E) < \infty$. In addition, fix $\varepsilon > 0$.

   (a) Recall[4] that μ and the Lebesgue outer measure μ^* coincide for every measurable set. Hence, by the definition of μ^*, there exists bounded open intervals $I_j \subset \mathbb{R}^{n+m}$ such that

   $$E \subset U := \bigcup_{j=1}^{\infty} I_j \tag{5}$$

   and that

   $$\mu(U) < \mu(E) + \frac{\varepsilon^2}{2}, \quad \text{i.e., } \mu(U \setminus E) < \frac{\varepsilon^2}{2}. \tag{6}$$

   (b) Define

   $$U_N := \bigcup_{j=1}^{N} I_j. \tag{7}$$

   Continuity of μ from below implies that there exists $N \in \mathbb{N}$ such that

   $$\mu(U) - \mu(U_N) < \frac{\varepsilon^2}{2}, \quad \text{i.e., } \mu(U \setminus U_N) < \frac{\varepsilon^2}{2}. \tag{8}$$

[3] Here, $\| \cdot \|_2$ denotes the L^2-norm.

[4] This is in fact the most efficient treatment of the Lebesgue measure that I know of. One would first define measures and outer measures. Then, one would prove Caratheodory's theorem and apply it to the special case of Lebesgue outer measure, which yields a definition of the Lebesgue measure. A reference for this can be found in the unpublished graduate text by Mark Williams titled *Multivariable real analysis*.

(c) Finally, since $U_N \subset U$ and $E \subset U$, we have [5]

$$\mu(U_N \triangle E) \leq \mu(U \setminus E) + \mu(U \setminus U_N) < \varepsilon^2. \tag{9}$$

Hence,

$$\|\chi_{U_N} - \chi_E\|_2^2 = \int_{\mathbb{R}^{n+m}} |\chi_{U_N} - \chi_E|^2 \, d\mu = \mu(U_N \triangle E) < \varepsilon^2. \tag{10}$$

Therefore,

$$\|\chi_{U_N} - \chi_E\|_2 < \varepsilon. \tag{11}$$

We finish the proof by checking that $\chi_{U_N} \in S$, which is trivial.

□

Remark 1.1.1. *We consider the $(k,t)-$space signal given by $s : K \times T \to \mathbb{C}$ where $K \subset \mathbb{R}^d$ for $d-$dimensional imaging and $T \subset \mathbb{R}$ so that Theorem 1.1 applies. Hence, we can approximate s arbitrarily well by finite sums of functions of the form $f(k,t) = g(k)h(t)$, i.e.,*

$$s(k,t) = \sum_{\ell=1}^{L} c_\ell(k)\varphi_\ell(t). \tag{12}$$

This is equation (5) in [Zhi07].

As a rather surprising result, the following theorem (Theorem 2.4 in [Zhi07], stated below without proof) provides a necessary and sufficient condition for (12) to be *exact*.

Theorem 1.2. *Let $s(k,t)$ be defined over $K \times T$ where $K = \{k_1, \dots, k_n\}$ and $T = \{t_1, \dots, t_m\}$. Let $C \in \mathbb{C}^{n \times m}$ be given by $C_{ij} = s(k_i, t_j)$. The following are equivalent:*

1. $\operatorname{rank}(C) = L$,

2. *There exists linearly independent sets $\{c_1, \dots, c_L\} \subset \mathbb{C}^{|K|}$ and $\{\varphi_1, \dots, \varphi_L\} \subset \mathbb{C}^{|T|}$ such that*

$$s(k,t) = \sum_{\ell=1}^{L} c_\ell(k)\varphi_\ell(t). \tag{13}$$

[5]Here, $\triangle$ denotes the symmetric difference of two sets, i.e., $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

The theorem above considers the discrete case where s is only sampled at discrete locations in the $(k,t)-$space. Here, we present and prove a generalization of the theorem where we consider $s \in L^2(K \times T)$ where $K \subset \mathbb{R}^d$ and $T \subset \mathbb{R}$ are measure spaces.

Theorem 1.3. *Let $K \subset \mathbb{R}^d$ and $T \subset \mathbb{R}$. Let μ, ν be the Lebesgue measure on K, T, respectively. Define $S : L^2(T) \to L^2(K)$ by*

$$(Sf)(k) := \int_T s(k,t) f(t)\, d\nu(t). \tag{14}$$

The following are equivalent:

1. $\operatorname{rank}(S) = L$,
2. *There exists linearly independent sets $\{c_1, \dots, c_L\} \subset L^2(K)$ and $\{\varphi_1, \dots, \varphi_L\} \subset L^2(T)$ such that*

$$s(k,t) = \sum_{\ell=1}^{L} c_\ell(k) \varphi_\ell(t) \tag{15}$$

for almost every (k,t).

Proof.

1. $(\Longrightarrow)$. We identify the left- and right-hand sides of (15) via the Riesz representation theorem.
   (a) On the one hand, it is clear that $\Lambda_k : L^2(T) \to \mathbb{C}$ given by[6]

$$\Lambda_k(f) := (Sf)(k) = (f, \overline{s(k,\cdot)}) \tag{16}$$

   is a bounded linear functional for almost every k.
   (b) On the other hand, fix $f \in L^2(T)$. Let $\{c_1, \dots, c_L\} \subset L^2(K)$ be

[6]Here, we use $(\cdot,\cdot)$ to denote the inner product on the implied Hilbert space.

an orthonormal basis of range(S).

$$\Lambda_k(f) = (Sf)(k) = \sum_{\ell=1}^{L}(Sf, c_\ell)c_\ell(k) = \sum_{\ell=1}^{L}(f, S^* c_\ell)c_\ell(k) \tag{17}$$

$$= \int_T \left(\sum_{\ell=1}^{L} c_\ell(k)\overline{(S^* c_\ell)(t)} \right) f(t)\ d\nu(t) \tag{18}$$

$$= \left(f, \sum_{\ell=1}^{L} \overline{c_\ell(k)}(S^* c_\ell)(\cdot) \right). \tag{19}$$

(c) By the Riesz representation theorem, we know that for almost every k, the representation of Λ_k is unique. Taking $\varphi_\ell = S^* c_\ell$,[7] we see that

$$s(k,t) = \sum_{\ell=1}^{L} c_\ell(k)\varphi_\ell(t) \tag{20}$$

for almost every (k,t).

2. ($\Longleftarrow$).

(a) (rank(S) $\leq L$). Fix $f \in L^2(T)$. Notice that

$$(Sf)(k) = \sum_{\ell=1}^{L} c_\ell(k) \left(\int_T \varphi_\ell(t) f(t)\ d\nu(t) \right) = \sum_{\ell=1}^{L} c_\ell(k)(f, \overline{\varphi_\ell}). \tag{21}$$

Hence,

$$\text{range}(S) \subset \text{span}\{c_1, \ldots, c_L\} \tag{22}$$

and we have rank(S) $\leq L$.

(b) (rank(S) $\geq L$). We show that $c_\ell \in$ range(S) for every ℓ. Since

$$(Sf_\ell)(k) = \sum_{j=1}^{L} c_j(k)(f_\ell, \overline{\varphi_j}), \tag{23}$$

[7] Notice that we have not checked that $\{\varphi_\ell\}$ is a linearly independent set. However, doing so is straightforward and is left as an exercise.

it suffices to find $f_\ell \in L^2(T)$ such that $(f_\ell, \overline{\varphi_j}) = \delta_{\ell j}$. Consider the Gram matrix $G \in \mathbb{C}^{L\times L}$ where

$$G_{ij} = (\overline{\varphi_i}, \overline{\varphi_j}). \tag{24}$$

Since $\{\overline{\varphi_\ell}\}$ is linearly independent set, G is invertible. We take

$$f_\ell = \sum_{i=1}^{L} (G^{-1})_{\ell i} \overline{\varphi_i}, \tag{25}$$

in which case,

$$(f_\ell, \overline{\varphi_j}) = \sum_{i=1}^{L} (G^{-1})_{\ell i} G_{ij} = \delta_{\ell j}. \tag{26}$$

□

2 Acquisition and reconstruction

In this section, we assume the $(k,t)-$space data $s(k,t)$ is $L-$th order partially separable. In practice, it is rarely the case that L is known *a priori*. We will start by assuming L is known and fixed, and describe an analytic solution to the reconstruction problem. We will also discuss two ways to determine L from data.

2.1 Acquisition

We follow the data acquisition scheme 1(b) in [Zhi07], which is reproduced below in Figure 1. The acquired data is grouped into two overlapping sets. One set (circles) is sampled with high temporal resolution, i.e.,

$$s_1(k,t) \quad \forall (k,t) \in K_1 \times T_1 = \{k_1^1, \ldots, k_n^1\} \times \{t_1^1, \ldots, t_m^1\}; \tag{27}$$

the other set (dots) is acquired extensively in the $k-$space to achieve high spatial resolution, i.e.,

$$s_2(k,t) \quad \forall (k,t) \in K_2 \times T_2 = \{k_1^2, \ldots, k_p^2\} \times \{t_1^2, \ldots, t_q^2\}. \tag{28}$$

In particular, note that we have $T_1 \supset T_2$ and $K_1 \subset K_2$.

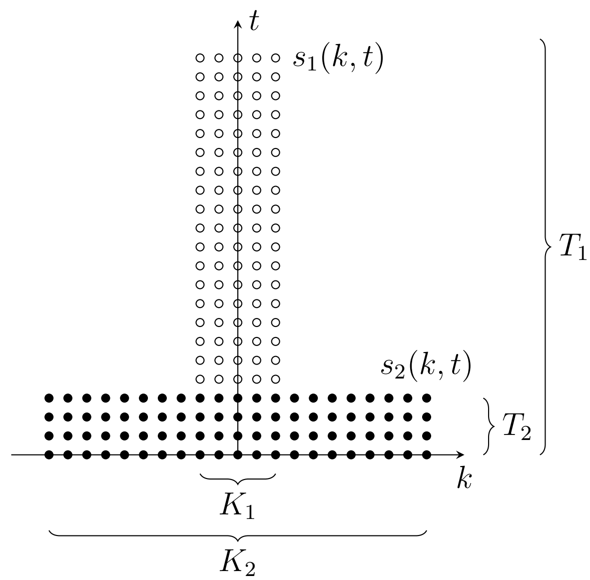

Figure 1: Data acquisition scheme for $L-$th order partially separable model.

2.2 Reconstruction

We first use $s_1(k,t)$ to estimate the temporal basis functions $\varphi_\ell(t)$. Then, the spatial basis functions $c_\ell(k)$ can be estimate from $s_2(k,t)$ and $\varphi_\ell(t)$. Once $c_\ell(k)$ is determined for all $k \in K_2$ and $\varphi_\ell(t)$ is determined for all $t \in T_1$, we can recover $s(k,t)$ for all $(k,t) \in K_2 \times T_1$ using Equation (12). Finally, the underlying image $\rho(r,t)$ can be reconstructed from $s(k,t)$ at high spatiotemporal resolution.

2.2.1 Estimating temporal basis functions

Let $C := C^* + E \in \mathbb{C}^{n\times m}$ denote the possibly noisy observation of $s_1(k,t)$ at discrete locations $(k,t) \in K_1 \times T_1$. Specifically, we let

$$C_{ij} := C^*_{ij} + E_{ij} \tag{29}$$

where $C^*_{ij} = s_1(k^1_i, t^1_j)$ and $E_{ij} \sim N(0, \sigma^2)$.

Now, Theorem 1.2 suggests that $\text{rank}(C^*) = L$. Hence, we can recover the true underlying data C^* by solving the following low-rank matrix approximation problem, which we denote as $(*)$,[8]

$$\begin{aligned} &\text{minimize} \quad \|C - B\| && (30) \\ &\text{subject to} \quad \operatorname{rank}(B) \le L. && (31) \end{aligned}$$

Here, we may take the norm $\|\cdot\|$ to be either the Frobenius or the spectral norm. The Eckart-Young-Mirsky theorem[9] guarantees that $(*)$ has an optimal solution with a closed-form expression, i.e.,

$$\hat{C} = \sum_{\ell=1}^{L} \lambda_\ell u_\ell v_\ell^H, \tag{32}$$

where $\{\lambda_\ell\}$ are the singular values of C in non-increasing order, and $\{u_\ell\}$ and $\{v_\ell\}$ are the left and right singular vectors, respectively.[10] The solution $\hat{C}$ is unique if and only if $\lambda_L > \lambda_{L+1}$.

In the more practical situation where L is not known *a priori*, there are at least two ways to choose a reasonable L that balances model complexity (in this case, $\operatorname{rank}(B)$) and goodness-of-fit. From a statistical point of view, one can pick the L that minimizes the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC). Alternatively, one might plot the residual norm $\|C - \hat{C}_L\|$ against L and pick the L that corresponds to the "elbow" of the $L-$shaped curve.

Finally, we set the $\ell-$th temporal basis function $\varphi_\ell(t)$, which is defined over T_1, to be the $\ell-$th left singular vector, i.e.,

$$\begin{bmatrix} \varphi_\ell(t_1^1) & \cdots & \varphi_\ell(t_m^1) \end{bmatrix}^T := u_\ell. \tag{33}$$

2.2.2 Estimating spatial basis functions

We estimate the spatial basis functions $c_\ell(k)$ by fitting Equation (12) to $s_2(k,t)$. Specifically, for fixed i, which is used to index $k-$space location $k_i^2 \in K_2$, we have

$$s_2(k_i^2, t_j^2) = \sum_{\ell=1}^{L} c_\ell(k_i^2)\varphi_\ell(t_j^2), \quad j = 1, \ldots, q, \tag{34}$$

[8] We take the feasible set to be the set of all complex-valued matrices of rank at most L to account for poor and/or degenerate sampling of the $(k,t)-$space that results in lower rank of C^*.

[9] The statement and proof of the theorem are well-known and thus omitted here.

[10] If we replace C in $(*)$ with S as in Equation (14), we would arrive at an infinite-dimensional analog of $(*)$. As a matter of fact, this problem was solved in [Sch07].

where $\varphi_\ell(t_j^2)$ are known for every j, which is used to index temporal location $t_j^2 \in T^2$. In matrix form, this is given by

$$\begin{bmatrix} \varphi_1(t_1^2) & \cdots & \varphi_L(t_1^2) \\ \varphi_1(t_2^2) & \cdots & \varphi_L(t_2^2) \\ \vdots & \ddots & \vdots \\ \varphi_1(t_q^2) & \cdots & \varphi_L(t_q^2) \end{bmatrix} \begin{bmatrix} c_1(k_i^2) \\ c_2(k_i^2) \\ \vdots \\ c_L(k_i^2) \end{bmatrix} = \begin{bmatrix} s_2(k_i^2, t_1^2) \\ s_2(k_i^2, t_2^2) \\ \vdots \\ s_2(k_i^2, t_q^2) \end{bmatrix}. \tag{35}$$

We can solve the system of equations above in the least-squares sense for $c_\ell(k_i^2)$ for every ℓ. Repeating the same routine for every i, we obtain the spatial basis function $c_\ell(k)$ for every $k \in K_2$.

References

[Sch07] Erhard Schmidt. “Zur Theorie der linearen und nichtlinearen Integralgleichungen: I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener”. In: *Mathematische Annalen* 63.4 (Dec. 1907), pp. 433–476. ISSN: 0025-5831, 1432-1807. DOI: 10.1007/BF01449770.

[Zhi07] Zhi-Pei Liang. “Spatiotemporal Imaging with Partially Separable Functions”. In: *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Arlington, VA, USA: IEEE, 2007, pp. 988–991. ISBN: 978-1-4244-0671-5. DOI: 10.1109/ISBI.2007.357020.