Kernel Mode Decomposition and programmable/interpretable regression networks

Houman Owhadi

Oberwolfach, July 2019

AFOSR. Grant number FA9550-18-1-0271.
Games for Computation and Learning, 2018-2021.

Collaborators

Clint Scovel

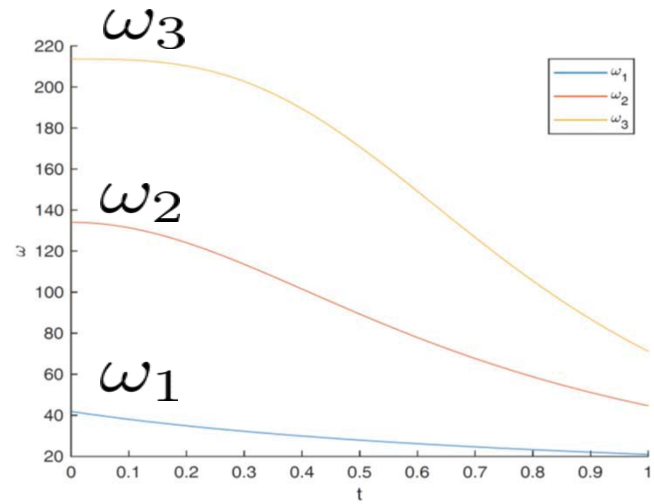
Gene Ryan Yoo

Kernel Mode Decomposition and programmable/interpretable regression networks, O., Scovel, Yoo, 2019
arXiv:1907.08592

Mode decomposition

Let $v_1, \ldots, v_m$ be unknown s.t.

m: unknown

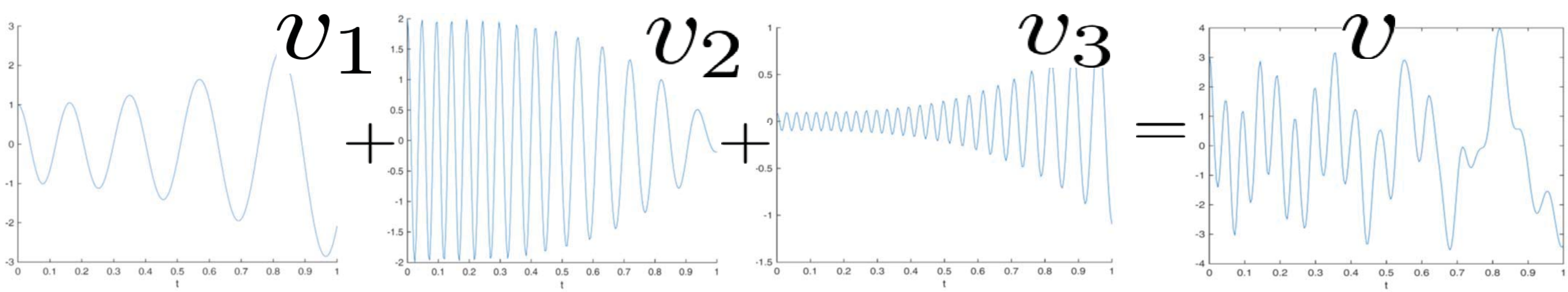


$$v_i(t) = a_i(t) \cos\big(\theta_i(t)\big)$$

a_i: unknown, slowly varying

$\omega_i := \dot{\theta}_i$: unknown, slowly varying, positive, well separated

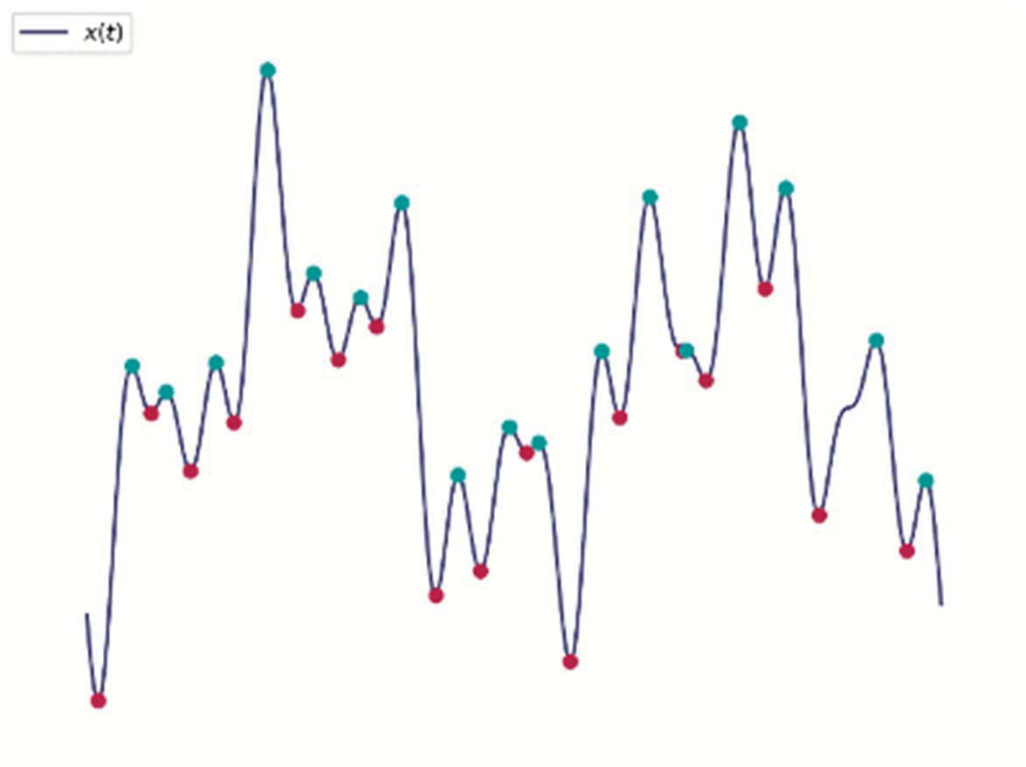
$$v_i(t) \approx a_i(\tau) \cos\big(\omega_i(\tau)(t-\tau) + \theta_i(\tau)\big) \text{ for } t \approx \tau$$



Problem Given $v = \sum_{i=1}^{m} v_i$ recover $v_1, \ldots, v_m$

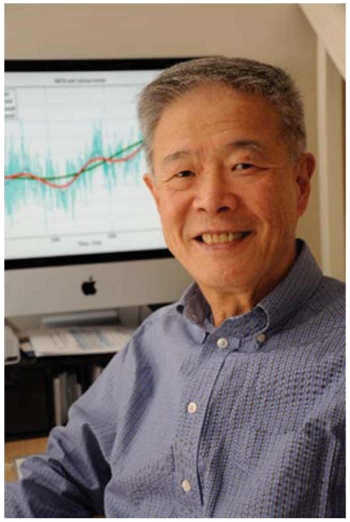
Empirical mode decomposition. Huang et Al, 1998

- $\approx 20,000$ citations
- Wide range of applications (meteorology, structural stability analysis, medical studies, etc. . .)
- Based on a heuristic "sifting process" (hard to analyze)



EMD's sifting process

Gif created by Geir Kulia and modified by Matt Hall, CC BY-SA 4.0, https://en.wikipedia.org/w/index.php?curid=57233178



N. E. Huang

A pattern recognition problem

Amenable to analysis

Numerical approximation

- Synchrosqueezed wavelet transform. Daubichies, Lu, Wu, 2011
- Variational mode decomposition. Dragomiretskiy, Zosso, 2014
- De-shape synchrosqueezing transform. Lin, Su, Wu, 2017

Statistical Inference

- Sparse time frequency representations. Hou, Shi, Tavallali 2011, 2014

Alchemy

Machine Learning

- Deep learning approach. Qiu, Ren, Suganthan, Amaratunga, 2017

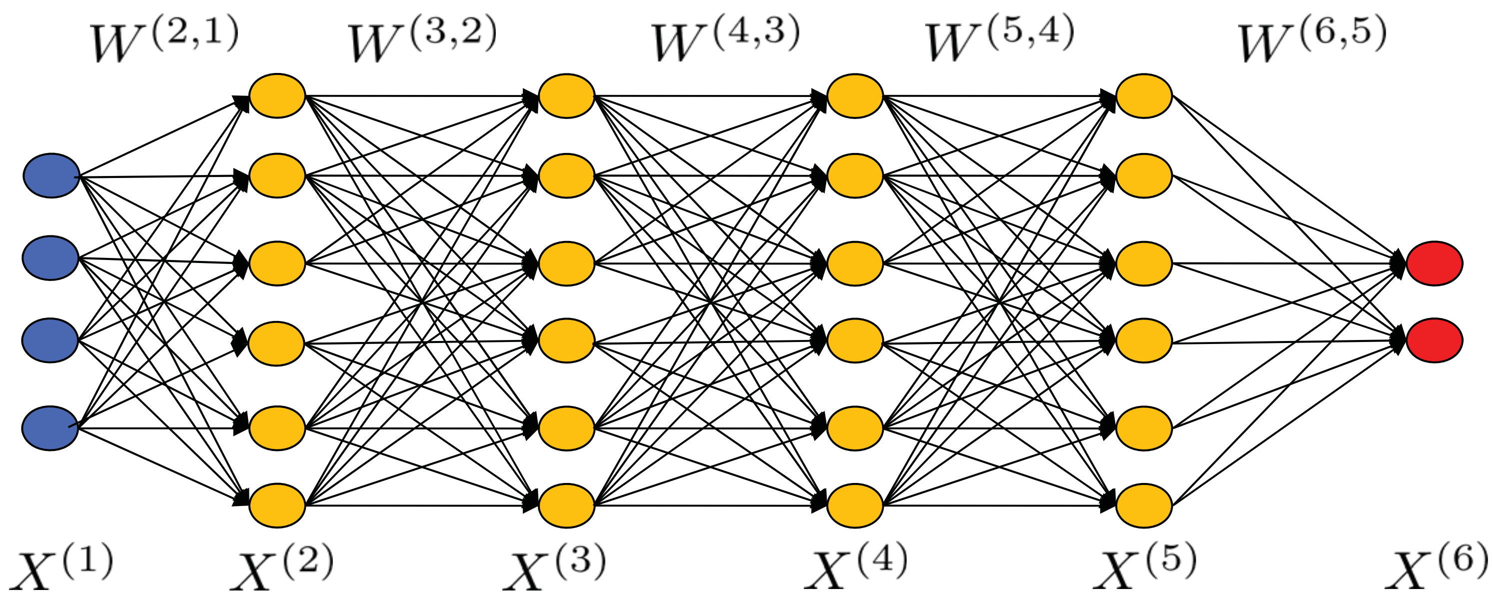
AI researchers allege that machine learning is alchemy

By Matthew Hutson | May. 3, 2018 , 11:15 AM

Ali Rahimi, a researcher in artificial intelligence (AI) at Google in San Francisco, California, took a swipe at his field last December—and received a 40-second ovation for it. Speaking at an AI conference, Rahimi charged that machine learning algorithms, in which computers learn through trial and error, have become a form of "alchemy." Researchers, he said, do not know why some algorithms work and others don't, nor do they have rigorous criteria for choosing one AI architecture over another. Now, in a paper presented on 30 April at the International Conference on Learning Representations in Vancouver, Canada, Rahimi and his collaborators document examples of what they see as the alchemy problem and offer prescriptions for bolstering AI's rigor.

"There's an anguish in the field," Rahimi says. "Many of us feel like we're operating on an alien technology."

"Machine learning has become alchemy"
Ali Rahimi
NIPS 2017 Test of Time Award



Nonlinear — Linear

$$X^{(n+1)} = s(W^{(n+1,n)} X^{(n)})$$

“Many of us feel like we’re operating on an alien technology”
Ali Rahimi, NIPS 2017 Test of Time Award

- We don’t know why ANNs work or why they don’t (no theory)
- developed through trial and error
- Some results are hard to replicate (many
- hyperparameters)
- Finding good architectures relies on guesswork

Question

Can the analysis of mode decomposition from the combined perspectives of numerical approximation and statistical inference be used as a Rosetta stone for deciphering mechanisms at play in ANNs?

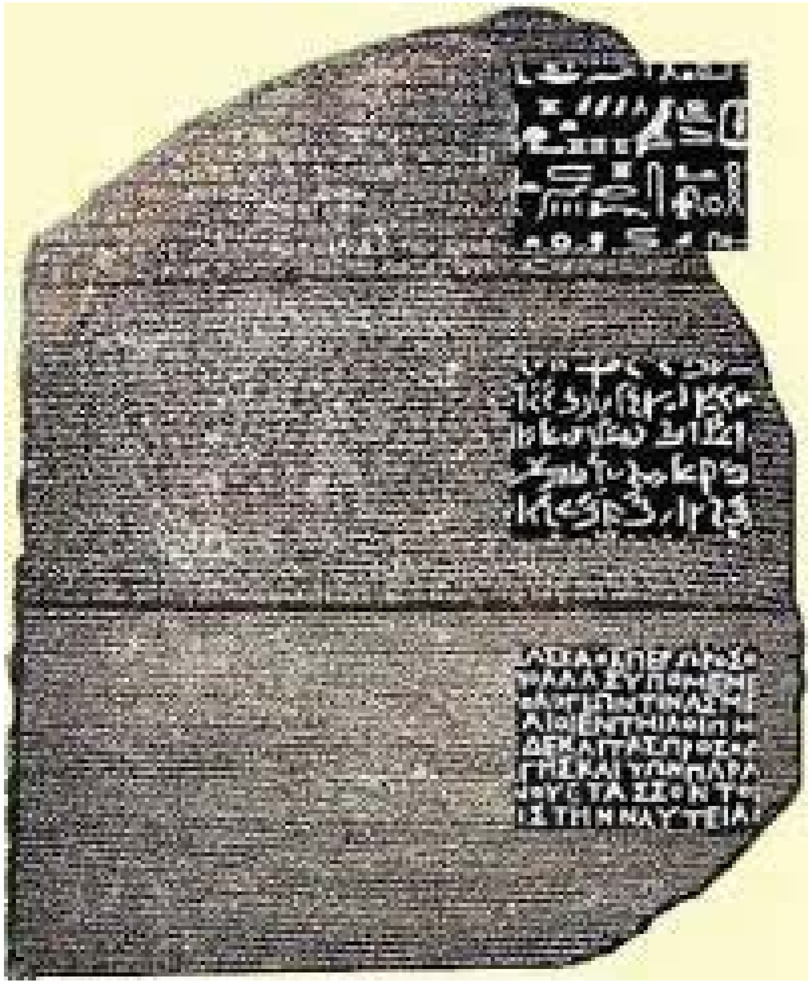
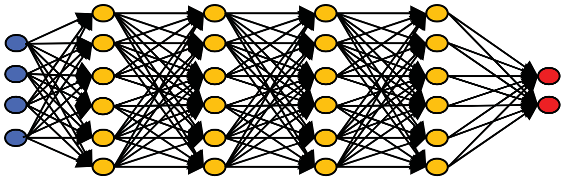


Image from Pinterest.com



$$w_i = K_i\big(\sum_j K_j\big)^{-1} v$$

$$w_i = \mathbb{E}\big[\xi_i \big| \sum_j \xi_j = v\big]$$

Mode decomposition with unknown waveforms

Let $v_1, \ldots, v_m$ be unknown s.t.

m: unknown

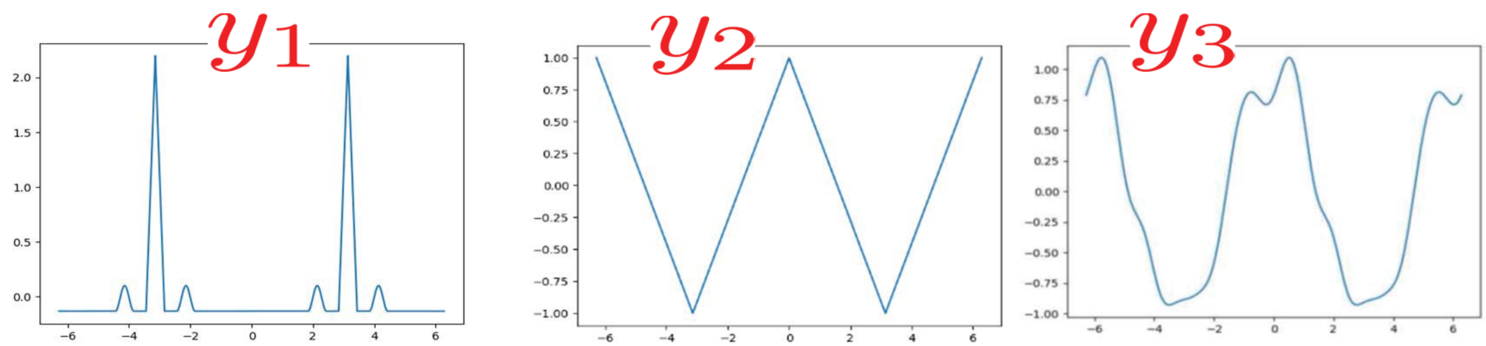
$$v_i(t) = a_i(t) y_i\big(\theta_i(t)\big)$$

a_i: unknown, slowly varying

$\omega_i := \dot{\theta}_i$: unknown, slowly varying, positive, well separated

(and $(k\dot{\omega}_j)_{t\in[-,1,1]} \not\equiv (k'\dot{\omega}_{j'})_{t\in[-1,1]}$ for $k, k' \in \mathbb{N}$)

y_i: unknown 2π-periodic



Problem Given $v = \sum_{i=1}^m v_i$ recover $v_1, \ldots, v_m$

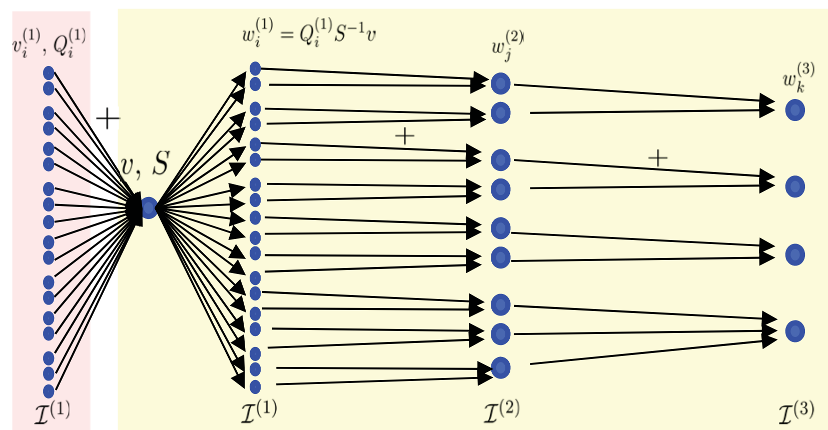
Solution

Kernel Mode Decomposition and programmable/interpretable regression networks, O., Scovel, Yoo, 2019

Algorithm achieving near machine precision recoveries.

Mode	$\frac{\lVert v_{i,e}-v_i\rVert_{L^2}}{\lVert v_i\rVert_{L^2}}$	$\frac{\lVert v_{i,e}-v_i\rVert_{L^\infty}}{\lVert v_i\rVert_{L^\infty}}$	$\frac{\lVert a_{i,e}-a_i\rVert_{L^2}}{\lVert a_i\rVert_{L^2}}$	$\lVert\theta_{i,e}-\theta_i\rVert_{L^2}$	$\frac{\lVert y_{i,e}-y_i\rVert_{L^2}}{\lVert y_i\rVert_{L^2}}$
$i=1$	6.59×10^{-3}	2.65×10^{-2}	1.52×10^{-5}	1.75×10^{-5}	6.65×10^{-3}
$i=2$	2.62×10^{-4}	5.61×10^{-4}	8.12×10^{-5}	1.25×10^{-4}	2.15×10^{-4}
$i=3$	6.55×10^{-4}	9.76×10^{-4}	3.99×10^{-4}	3.67×10^{-4}	3.43×10^{-4}

Looks like a Neural Network,
quacks like a Neural Network
but it is not a Neural Network!

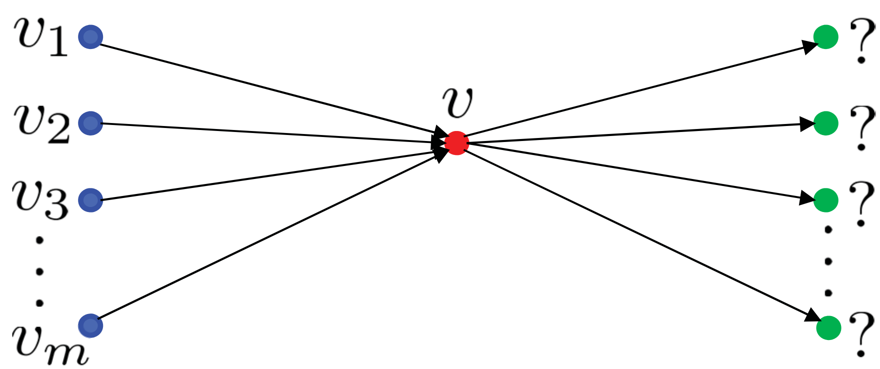


It is a regression network
Interpretable and amenable to analysis
Not trained but programed by composing two simple modules over a hierarchy

$V_1, \ldots, V_m$: Closed linear sub-spaces of V (Hilbert)

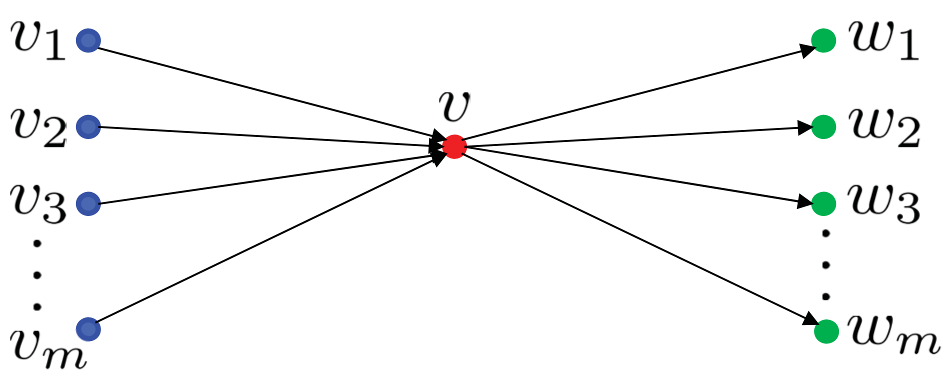
$$V_1 + \cdots + V_m = V$$

Let $v_1, \cdots, v_m$ be unknown elements of $V_1, V_2, \ldots, V_m$



Problem

Given $v = v_1 + \cdots + v_m$ recover $v_1, \ldots, v_m$



Player I

Chooses

$(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$

Player II

Sees $v = \sum_{i=1}^m v_i$

Chooses

$(w_1, \ldots, w_m) \in V_1 \times \cdots \times V_m$

Max → $\dfrac{\sum_{i=1}^m \|v_i - w_i\|_{V_i}^2}{\sum_{i=1}^m \|v_i\|_{V_i}^2}$ ← *Min*

$\| \cdot \|_{V_i}$: Quadratic norm on V_i

Theorem The optimal strategy of Player I is

$$(v_1, \ldots, v_m) \sim (\xi_1, \ldots, \xi_m)$$

$\xi_1, \ldots, \xi_m$: indep. centered GP with cov. op. $Q_1, \ldots, Q_m$

$$\xi_i \sim \mathcal{N}(0, Q_i)$$

$\|\cdot\|_{V_i} \Leftrightarrow Q_i \Leftrightarrow \xi_i \sim \mathcal{N}(0, Q_i)$

$$\|v_i\|_{V_i}^2 := [Q_i^{-1} v_i, v_i]$$

$Q_i : V_i^* \to V_i$ Symmetric $[\phi, Q_i \varphi] = [\varphi, Q_i \phi]$ for $\phi, \varphi \in V_i^*$
Postive: $[\phi, Q_i \phi] \geq 0$ for $\phi \in V_i^*$.

$$\xi_i : V_i^* \to \text{Gaussian space}$$

$$\phi \to [\phi, \xi_i] \sim \mathcal{N}(0, [\phi, Q_i \phi])$$

$$[\phi, Q_i \phi] = \|\phi\|_{V_i^*}^2 = \sup_{v \in V_i} \frac{[\phi, v]^2}{\|v\|_{V_i}^2}$$

Theorem The optimal strategy of Player II is

$$w_i = \mathbb{E}\Big[\xi_i \Big| \sum_j \xi_j = v\Big]$$

$\xi_1, \ldots, \xi_m$: indep. centered GP with cov. op. $Q_1, \ldots, Q_m$

$$\xi_i \sim \mathcal{N}(0, Q_i)$$

Variational Formulation

$$\underset{(w_1,\ldots,w_m)\in V_1\times\cdots\times V_m}{\operatorname{argmin}} \begin{cases} \text{Minimize } \sum_{i=1}^m \|w_i\|_{V_i}^2 \\ \text{Subject to } \sum_{i=1}^m w_i = v \end{cases}$$

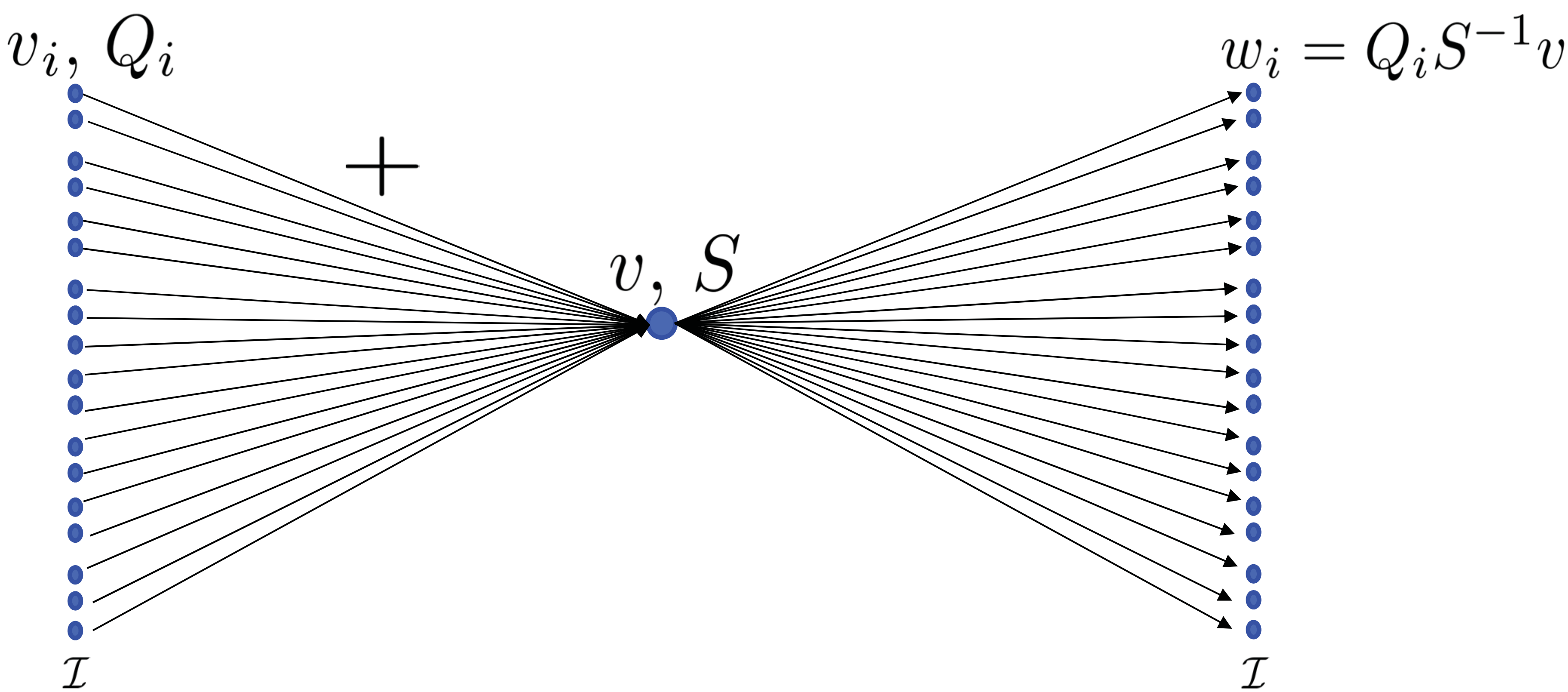
Galerkin optimality of the accuracy of the recovery

$$\sum_{i=1}^m \|v_i - w_i\|_{V_i}^2 = \inf_{\phi\in V^*} \sum_{i=1}^m \|v_i - Q_i\phi\|_{V_i}^2$$

Theorem The optimal strategy of Player II is

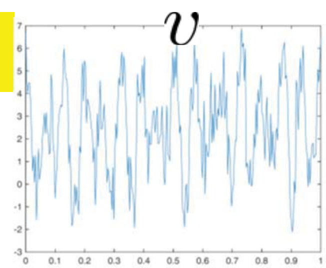
$$w_i = Q_i S^{-1} v$$

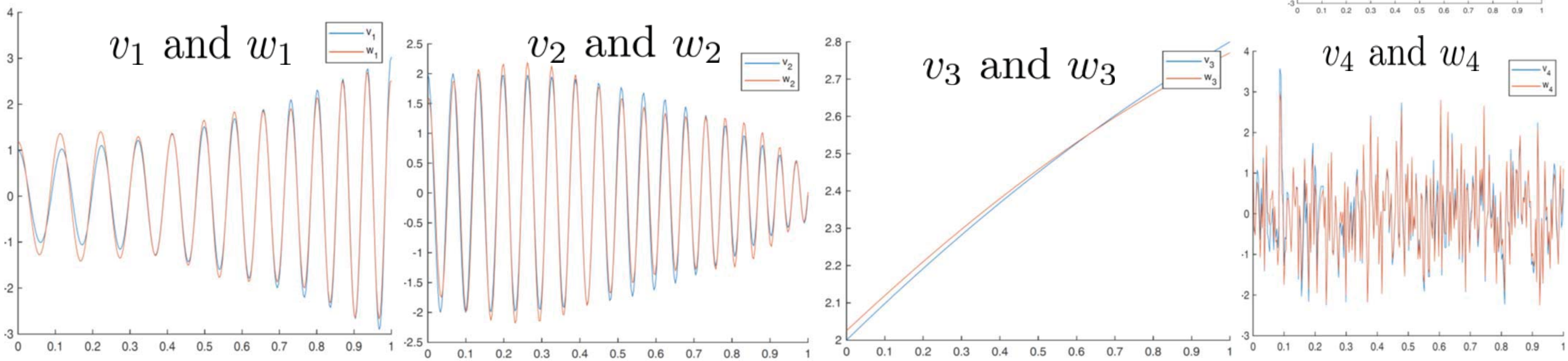
$S := \sum_j Q_j$: $V^* \to V$ always invertible



Example: Mode decomposition when only amplitudes are unknown

Given $v = v_1 + v_2 + v_3 + v_4$, recover v_1, v_2, v_3 and v_4





$a_1(t)\cos(\theta_1(t))$	$a_2(t)\cos(\theta_2(t))$	$a_3(t)$	White noise
$\xi_1(t) = \zeta_1(t)\cos(\theta_1(t))$	$\xi_2(t) = \zeta_2(t)\cos(\theta_2(t))$	$\xi_3(t) = \zeta_3(t)$	ξ_4
$\mathbb{E}[\zeta_i(s)\zeta_i(t)] = e^{-\frac{(t-s)^2}{\gamma^2}}$			$\mathbb{E}[\xi_4(s)\xi_4(t)] = \sigma^2\delta(t-s)$
$\mathbb{E}[\xi_i(t)\xi_i(s)] = K_i(s,t) = e^{-\frac{(t-s)^2}{\gamma^2}}\cos(\theta_i(s))\cos(\theta_i(t))$		$K_3(s,t) = e^{-\frac{(t-s)^2}{\gamma^2}}$	$K_4(s,t) = \delta(t-s)$

Solution

$$w_i = \mathbb{E}\left[\xi_i \,\middle|\, \sum_j \xi_j = v\right] \Longleftrightarrow w_i = K_i\left(\sum_j K_j\right)^{-1} v$$

The GP approach to linear mode decomposition is not new

- Thikonov regularization. Thikonov, 1943.
- Additive regression. Stone, 1985.
- Generalized additive models. Hastie and Tibshirani, 1990.
- Additive Gaussian processes. Duvenaud, Nickisch, and Rasmussen, 2011.
- Additive Covariance kernels. Durrande, Ginsbourger, Roustant, 2012. Durrande, Hensman, Rattray, and Lawrence, 2016.

Example: Mode decomposition when both amplitudes and phases are unknown

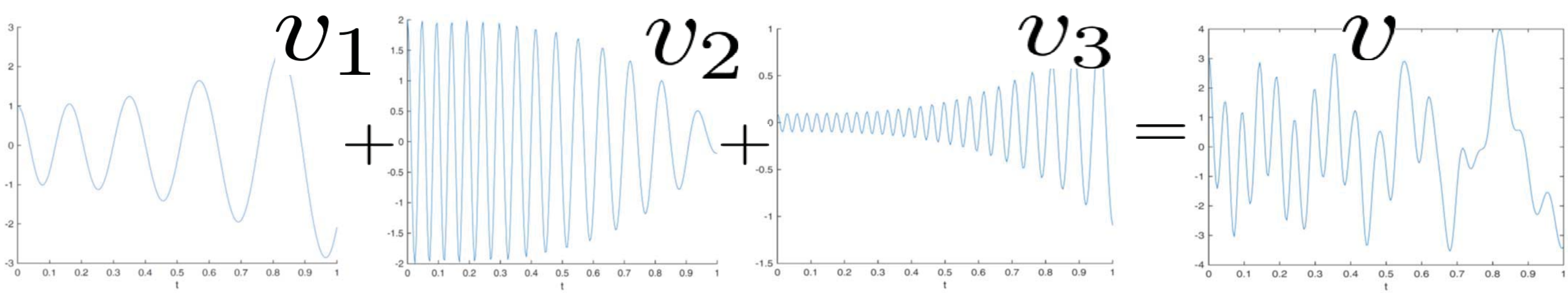
Let $v_1, \ldots, v_m$ be unknown s.t.

m: unknown

$$v_j(t) = a_j(t) \cos\left(\theta_j(t)\right)$$

a_j: unknown, slowly varying

$\omega_j := \dot{\theta}_j$: unknown, slowly varying, positive, well separated

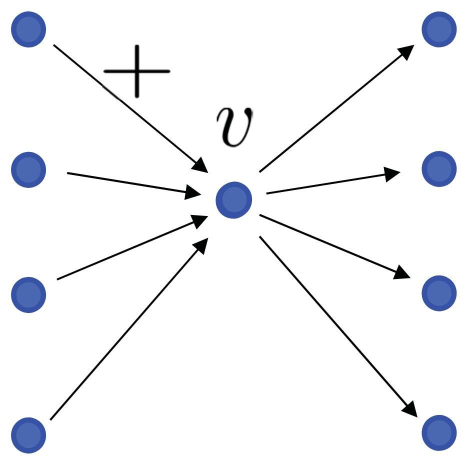


Problem Given $v = \sum_{i=1}^{m} v_i$ recover $v_1, \ldots, v_m$

Mode decomposition problem: non-linear

$$v_j(t) = a_j(t)\cos\big(\theta_j(t)\big)$$

$v_j \in V_j$ $\qquad$ $w_j \in V_j$

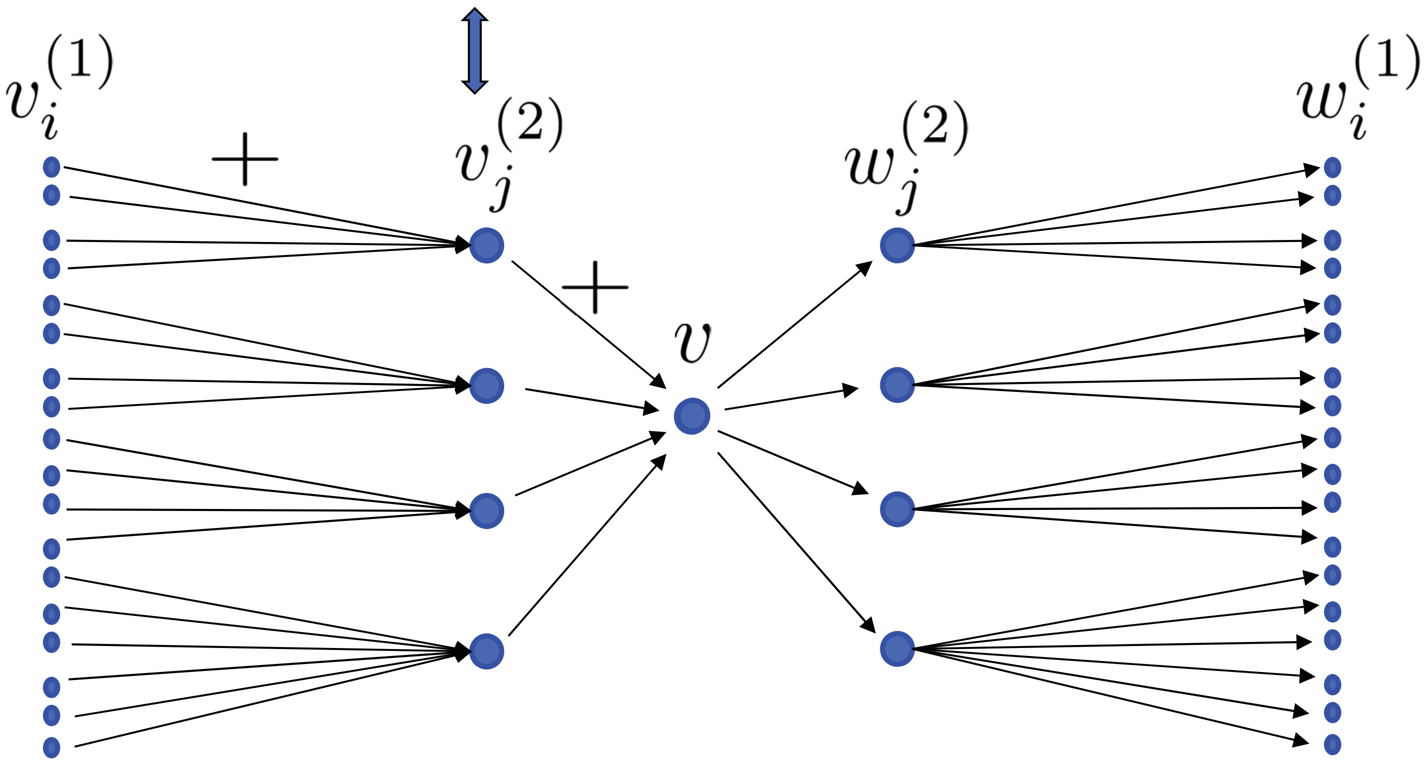


$$V_j := \Big\{ a(t)\cos\big(\theta(t)\big) \mid a, \theta \in \text{ linear spaces} \Big\} \text{ non-linear}$$

$$\zeta_a(t)\cos\big(\zeta_\theta(t)\big) \qquad \text{Not a GP}$$

GP $\qquad$ GP

$$v_j(t) = a_j(t)\cos\big(\theta_j(t)\big)$$

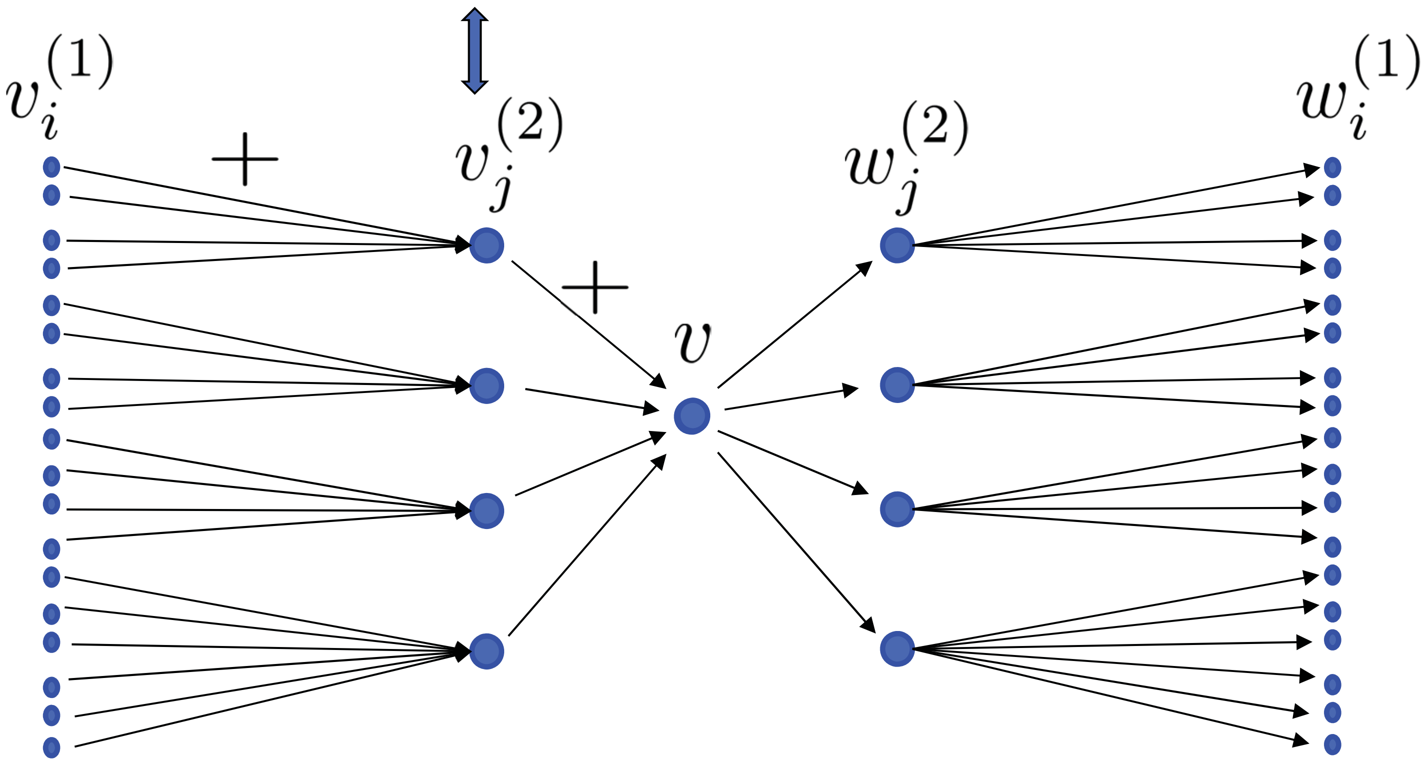


$$v_j^{(2)} = \sum_{i \rightsquigarrow j} v_i^{(1)}$$

$$v_i^{(1)} \in V_i^{(1)} \text{: Linear spaces}$$

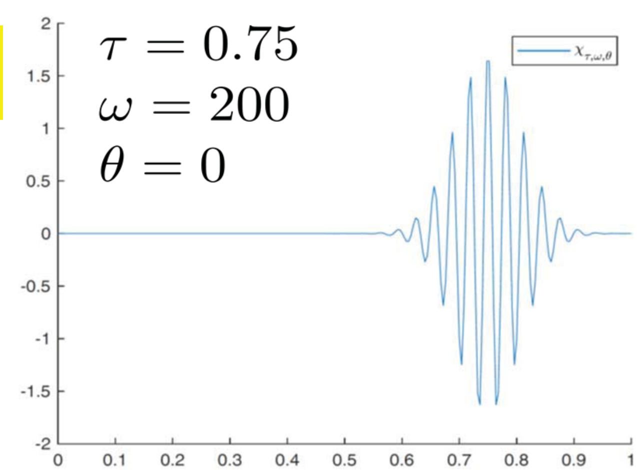
Idea: Recover the $v_j = a_j(t)\cos(\theta_j(t))$ as aggregates of finer modes living in linear spaces.

$$v_j(t) \approx a_j(\tau)\cos\big(\omega_j(\tau)(t-\tau)+\theta_j(\tau)\big) \text{ for } t \approx \tau$$



$$v_i^{(1)} \in \mathrm{Span}\{\chi_{\tau,\omega,\theta}\}$$

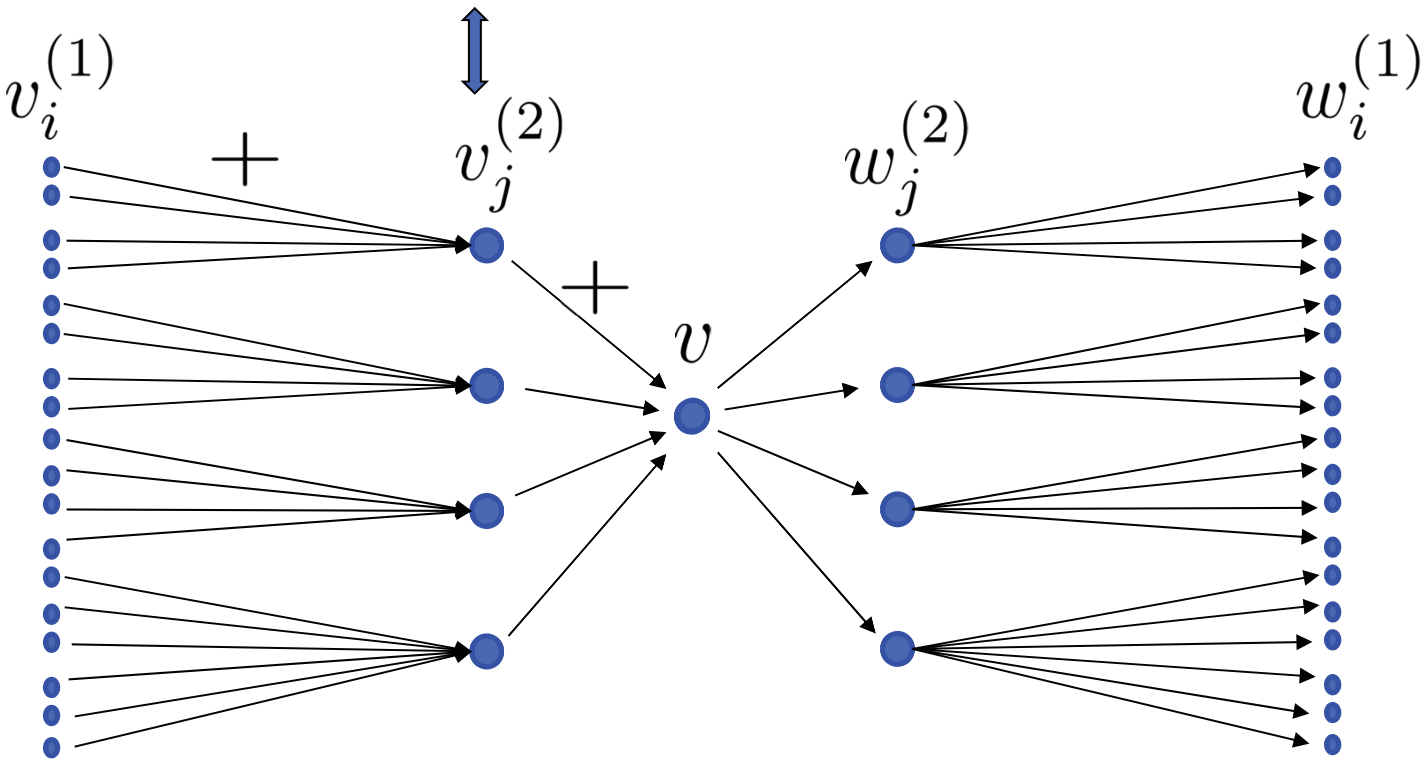
Gabor wavelets



$$\chi_{\tau,\omega,\theta}(t) := \left(\frac{2}{\pi^3}\right)^{\frac{1}{4}}\sqrt{\frac{\omega}{\alpha}}\cos\big(\omega(t-\tau)+\theta\big)e^{-\frac{\omega^2(t-\tau)^2}{\alpha^2}}$$

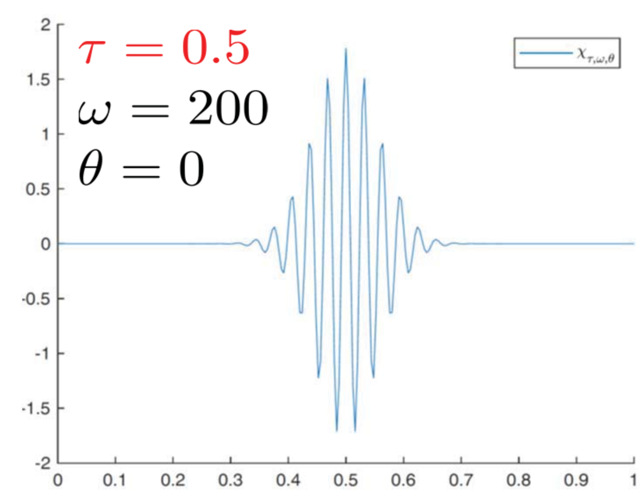
$$i \Longleftrightarrow (\tau,\omega,\theta)$$

$$v_j(t) \approx a_j(\tau)\cos\big(\omega_j(\tau)(t-\tau)+\theta_j(\tau)\big) \text{ for } t \approx \tau$$



$$v_i^{(1)} \in \mathrm{Span}\{\chi_{\tau,\omega,\theta}\}$$

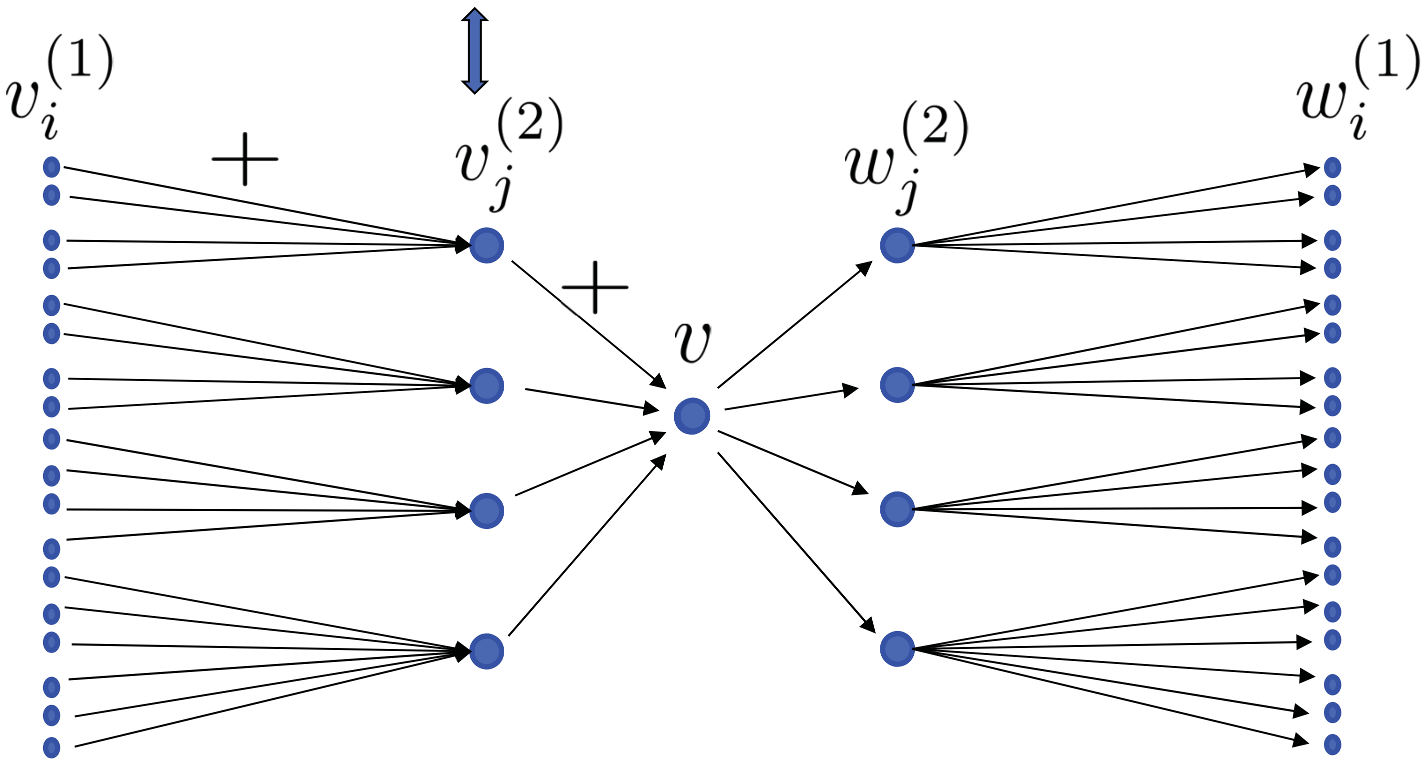
Gabor wavelets



$$\chi_{\tau,\omega,\theta}(t) := \left(\frac{2}{\pi^3}\right)^{\frac{1}{4}} \sqrt{\frac{\omega}{\alpha}}\cos\big(\omega(t-\tau)+\theta\big)e^{-\frac{\omega^2(t-\tau)^2}{\alpha^2}}$$

$$i \Longleftrightarrow (\tau, \omega, \theta)$$

$$v_j(t) \approx a_j(\tau)\cos\big(\omega_j(\tau)(t-\tau)+\theta_j(\tau)\big) \text{ for } t \approx \tau$$

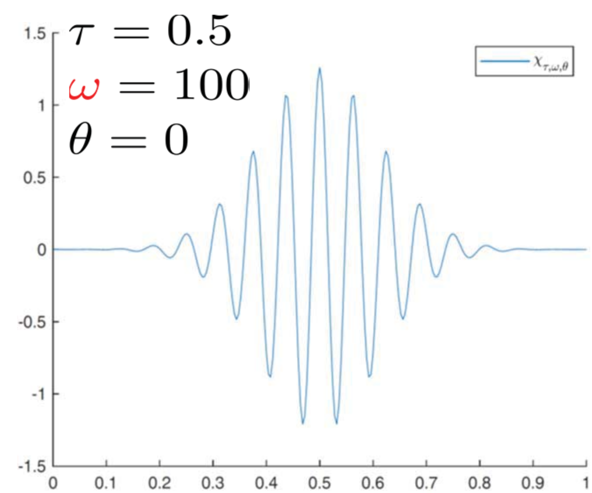


$$v_i^{(1)} \in \mathrm{Span}\{\chi_{\tau,\omega,\theta}\}$$

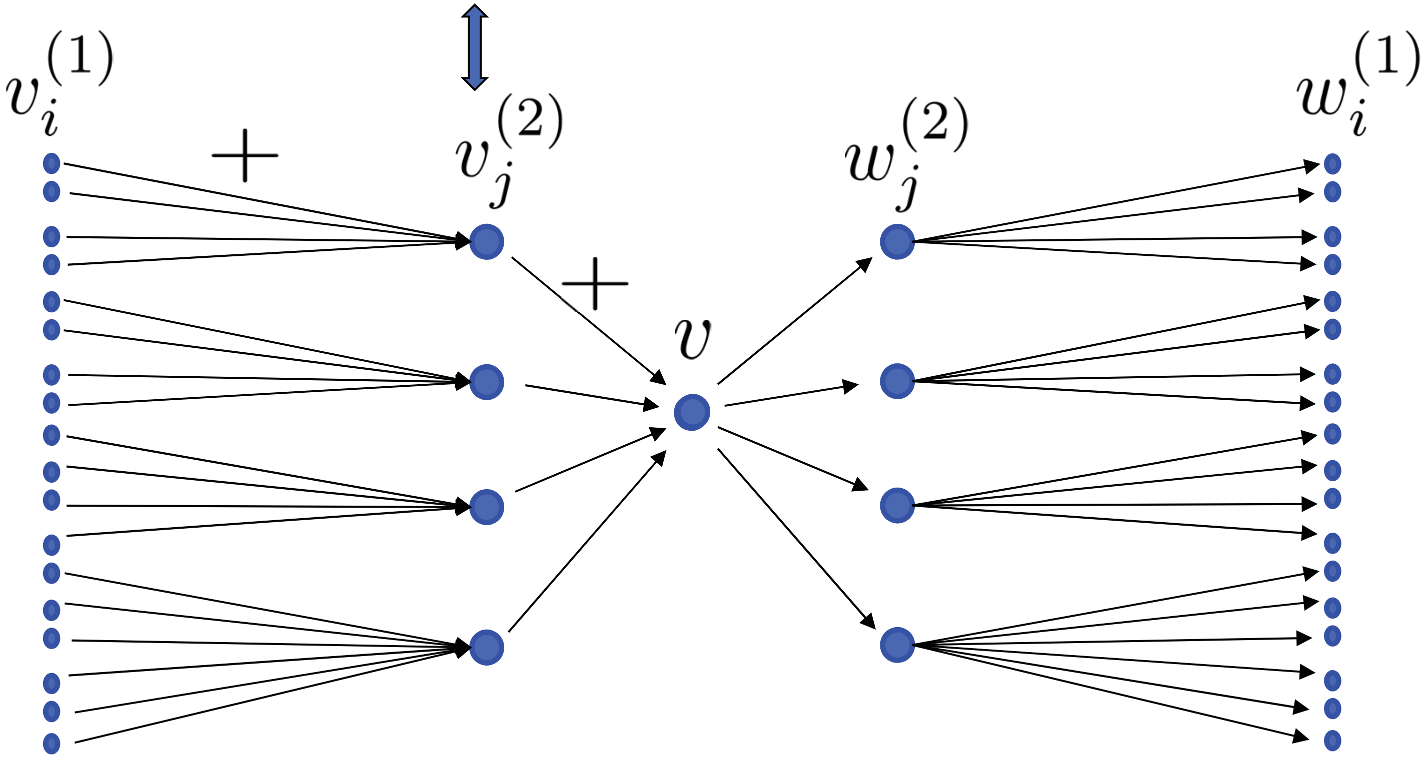
Gabor wavelets

$$\chi_{\tau,\omega,\theta}(t) := \left(\frac{2}{\pi^3}\right)^{\frac{1}{4}}\sqrt{\frac{\omega}{\alpha}}\cos\big(\omega(t-\tau)+\theta\big)e^{-\frac{\omega^2(t-\tau)^2}{\alpha^2}}$$

$$i \Longleftrightarrow (\tau,\omega,\theta)$$



$$v_j(t) \approx a_j(\tau)\cos\big(\omega_j(\tau)(t-\tau)+\theta_j(\tau)\big) \text{ for } t \approx \tau$$

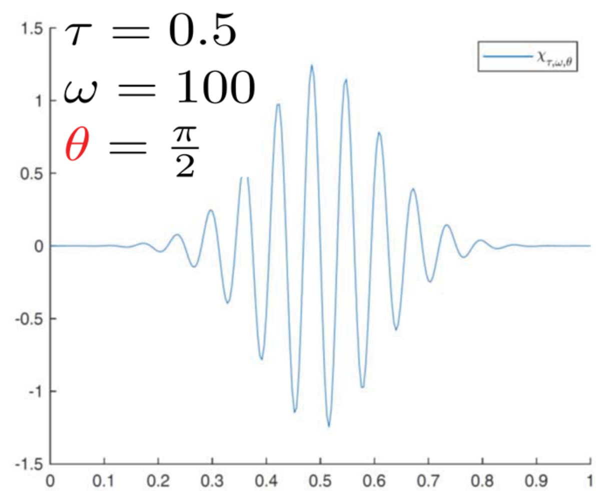


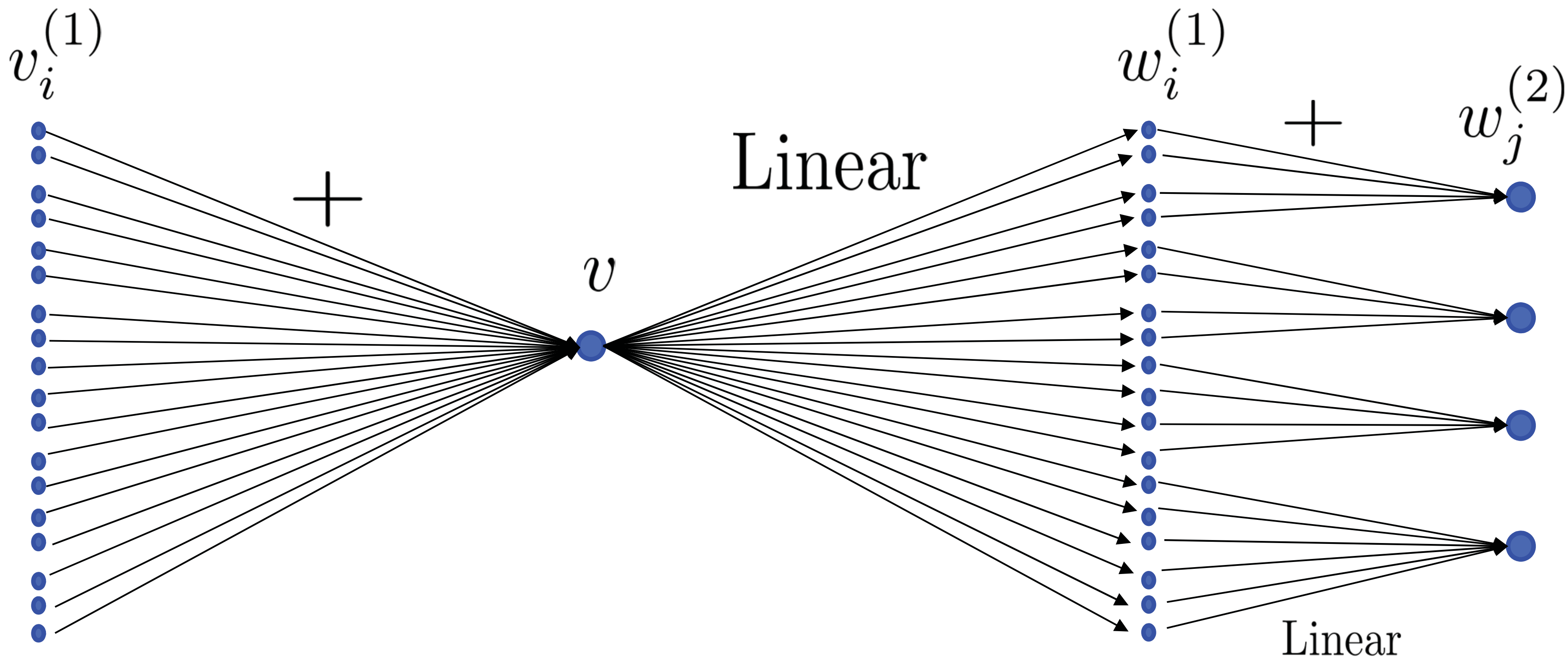
$$v_i^{(1)} \in \mathrm{Span}\{\chi_{\tau,\omega,\theta}\}$$

Gabor wavelets

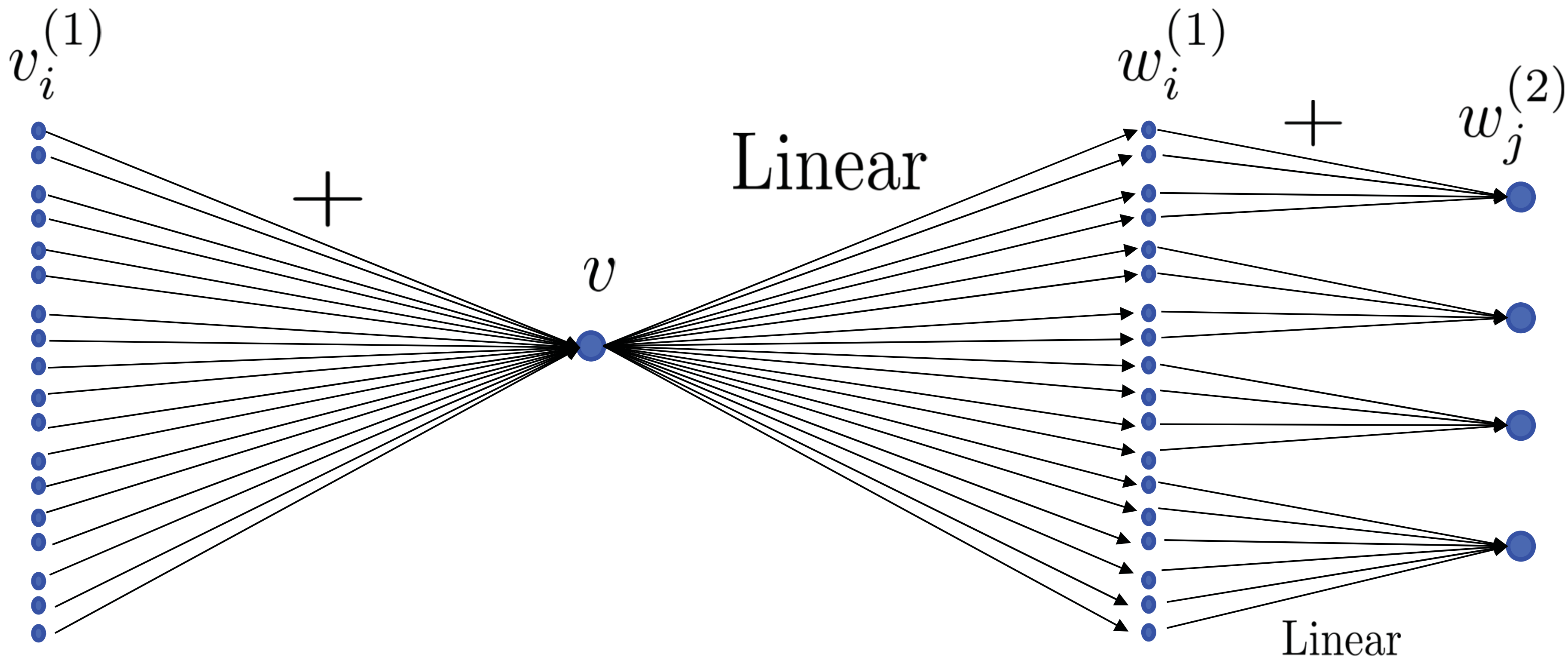
$$\chi_{\tau,\omega,\theta}(t) := \left(\frac{2}{\pi^3}\right)^{\frac{1}{4}}\sqrt{\frac{\omega}{\alpha}}\cos\big(\omega(t-\tau)+\theta\big)e^{-\frac{\omega^2(t-\tau)^2}{\alpha^2}}$$

$$i \Longleftrightarrow (\tau,\omega,\theta)$$





Where did the nonlinearity go?



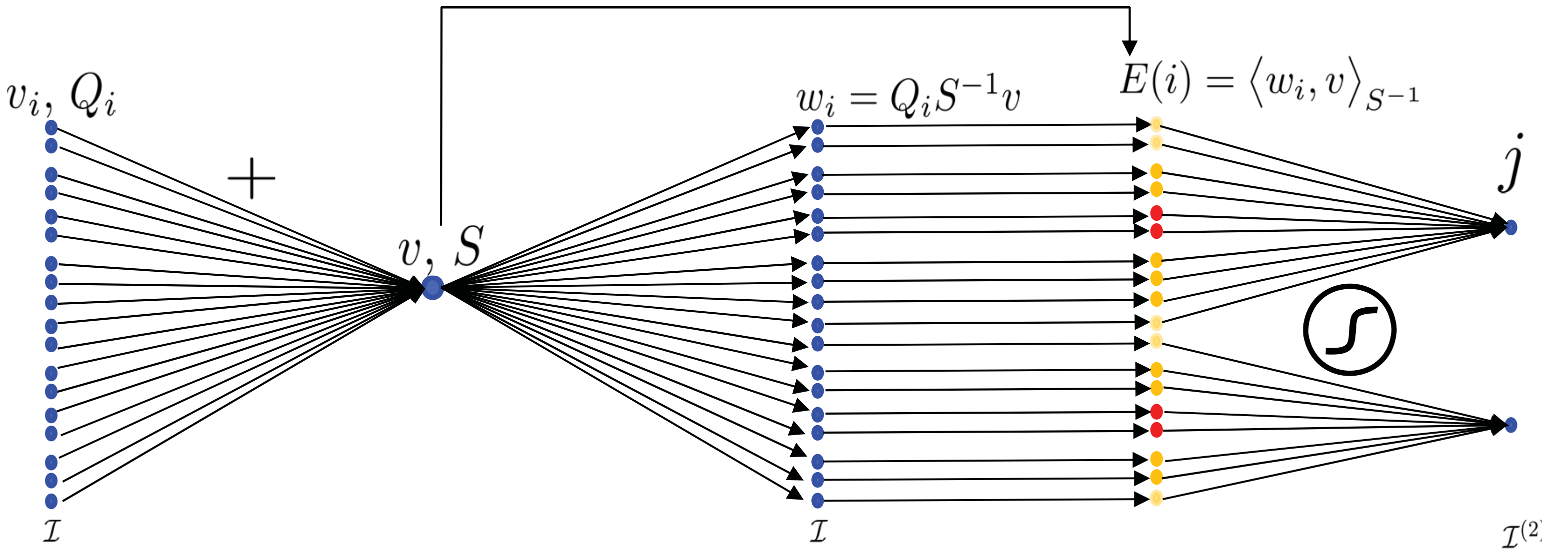
Where did the nonlinearity go?

It went in the identification of the ancestor/descendent relationship

$$i \rightsquigarrow j$$

Identification of ancestor/descendent relationships

- Compute energy of mode i, $E(i) := \|w_i\|^2_{V_i}$ (linear step)
- Combine $E(i)$ with nonlinear step (tresholding, graph-cut, argmax,...)

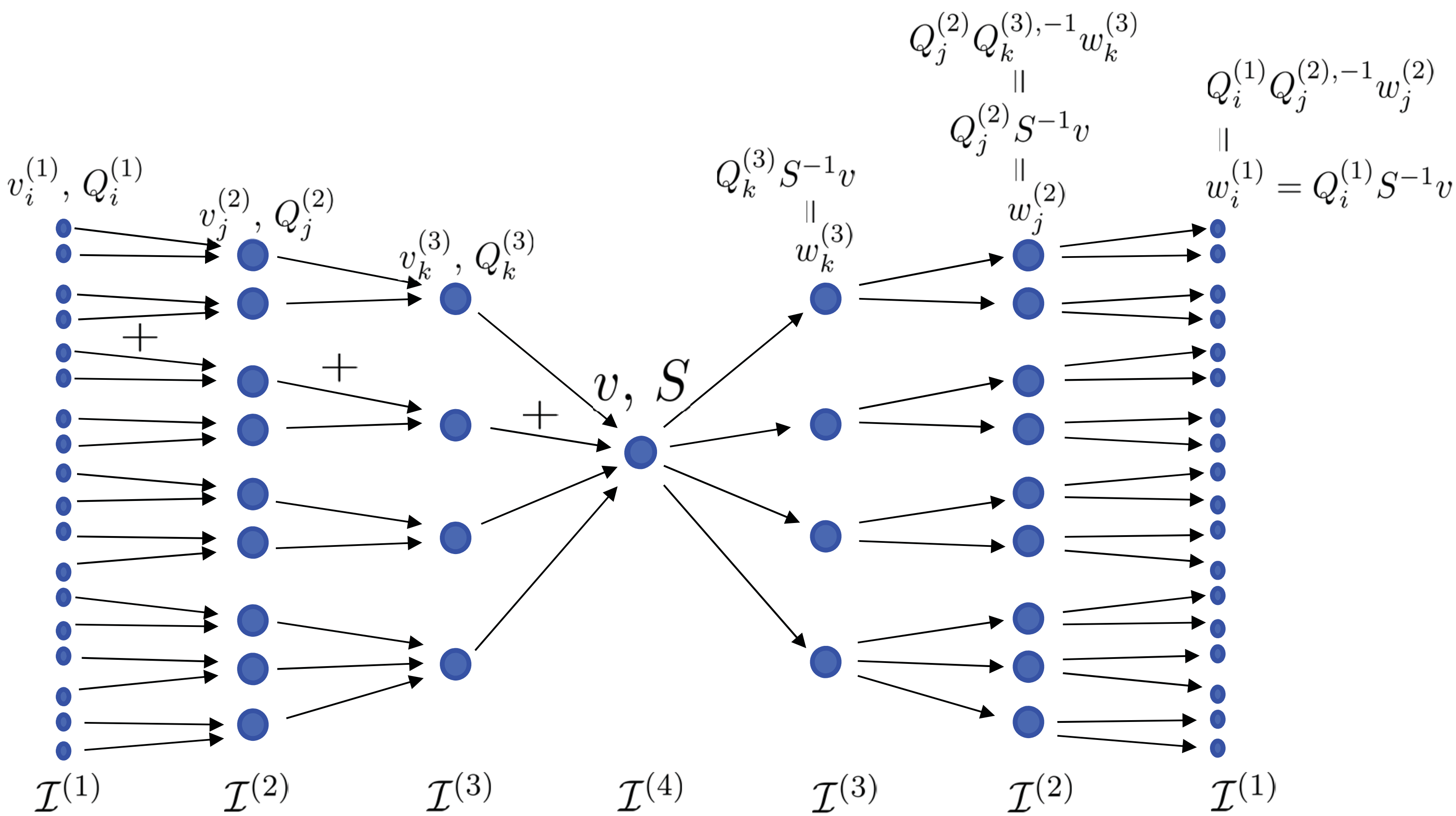


Theorem

$$E(i) = \langle w_i, v\rangle_{S^{-1}} = \mathrm{Var}\left(\langle \xi_i, v\rangle_{S^{-1}}\right)$$

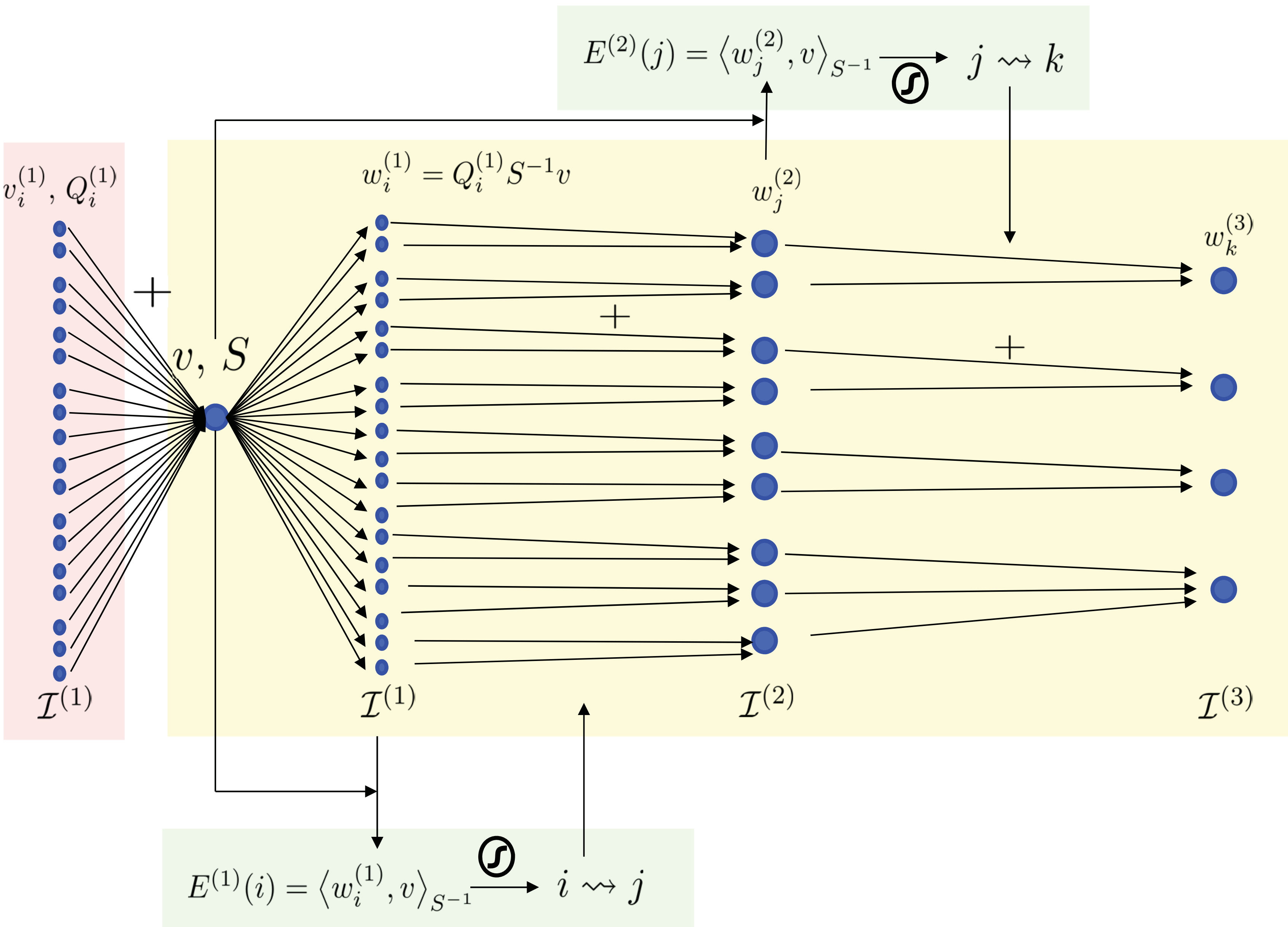
$$\sum_i E(i) = \|v\|^2_{S^{-1}} = \mathrm{Var}\left(\left\langle \sum_i \xi_i, v\right\rangle_{S^{-1}}\right)$$

Energy conservation ⟺ Variance decomposition



Strategy: repeated across levels of abstraction

Network decomposing and recomposing modes and kernels



The simplest network

Example: Mode decomposition when both amplitudes and phases are unknown

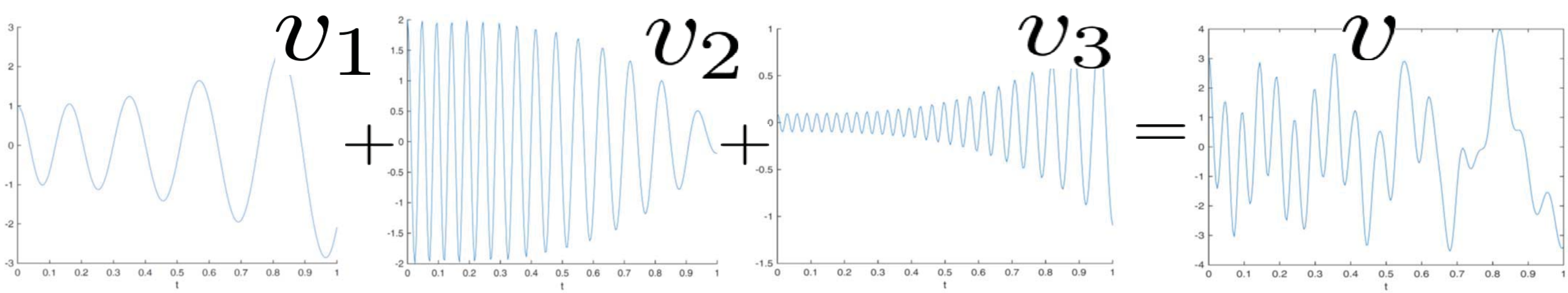
Let $v_1, \ldots, v_m$ be unknown s.t.

m: unknown

$$v_i(t) = a_i(t) \cos\big(\theta_i(t)\big)$$

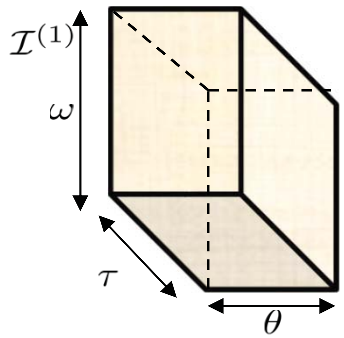
a_i: unknown, slowly varying

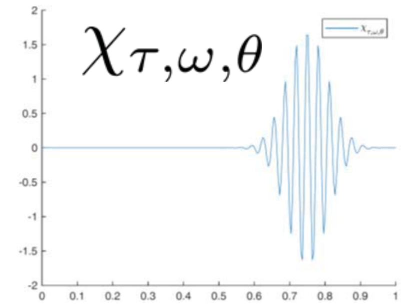
$\omega_i := \dot{\theta}_i$: unknown, slowly varying, positive, well separated



Problem Given $\sum_{i=1}^{m} v_i$ recover $v_1, \ldots, v_m$

Solution First recover fine modes $v_{\tau,\omega,\theta} \in \text{Span}\{\chi_{\tau,\omega,\theta}\}$





$$v_{\tau,\omega,\theta} \xrightarrow{\int_{-\pi}^{\pi}\int_{\omega_{\min}}^{\omega_{\max}}\int_0^1 \cdot\, d\tau\, d\omega\, d\theta} v$$

$v_{\tau,\omega,\theta} \in \text{Span}\{\chi_{\tau,\omega,\theta}\}$	$v = \int_{-\pi}^{\pi}\int_{\omega_{\min}}^{\omega_{\max}}\int_0^1 v_{\tau,\omega,\theta}\, d\tau\, d\omega\, d\theta$
$\xi_{\tau,\omega,\theta}(t) = \zeta(\tau,\omega,\theta)\chi_{\tau,\omega,\theta}(t)$ ↑ white noise	$\xi(t) = \int_{-\pi}^{\pi}\int_{\omega_{\min}}^{\omega_{\max}}\int_0^1$ $\zeta(\tau,\omega,\theta)\chi_{\tau,\omega,\theta}(t)\, d\tau\, d\omega\, d\theta$
$K_{\tau,\omega,\theta}(s,t) = \chi_{\tau,\omega,\theta}(s)\chi_{\tau,\omega,\theta}(t)$	$K(s,t) = \int_{-\pi}^{\pi}\int_{\omega_{\min}}^{\omega_{\max}}\int_0^1$ $K_{\tau,\omega,\theta}(s,t)\, d\tau\, d\omega\, d\theta$

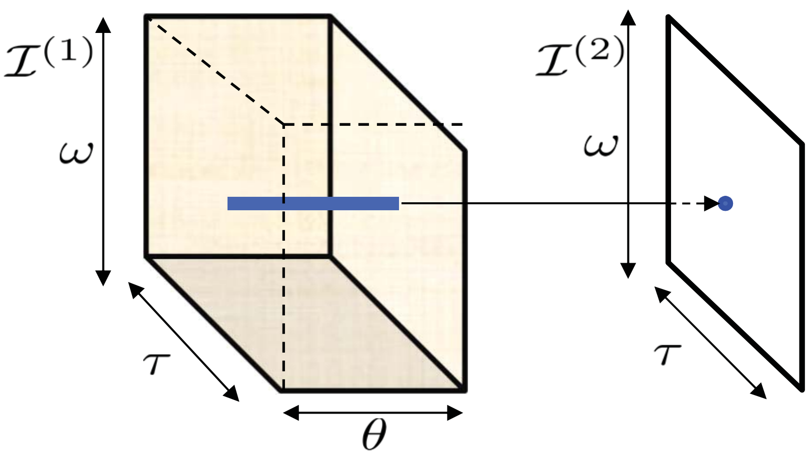
$$w_{\tau,\omega,\theta} \xleftarrow{K_{\tau,\omega,\theta}K^{-1}} v$$

Aggregate to level 2 modes $w_{\tau,\omega}$

$$w_{\tau,\omega,\theta} \xrightarrow{\quad \int_{-\pi}^{\pi} \cdot \, d\theta \quad} w_{\tau,\omega}$$

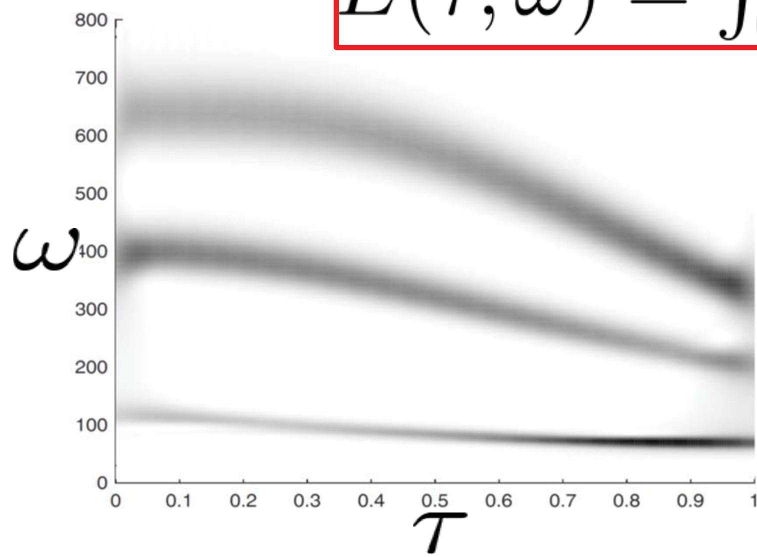
$$w_{\tau,\omega} = \int_{-\pi}^{\pi} w_{\tau,\omega,\theta} \, d\theta$$

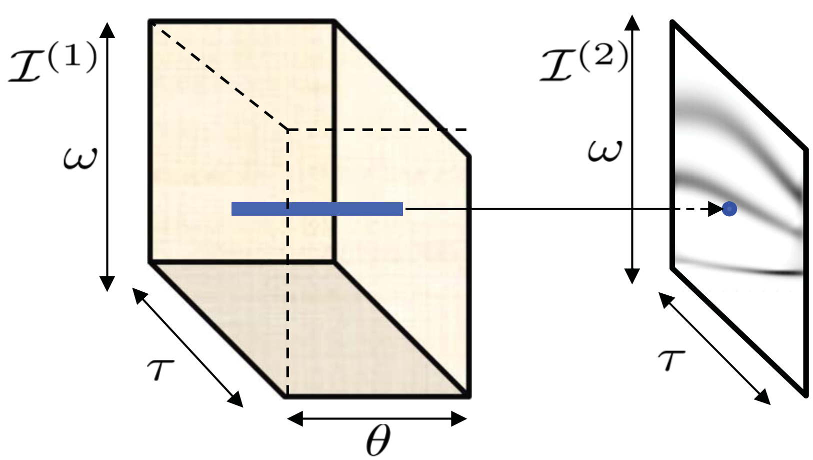
$(\tau, \omega, \theta) \to (\tau, \omega)$



Compute alignment energy of level 2 modes

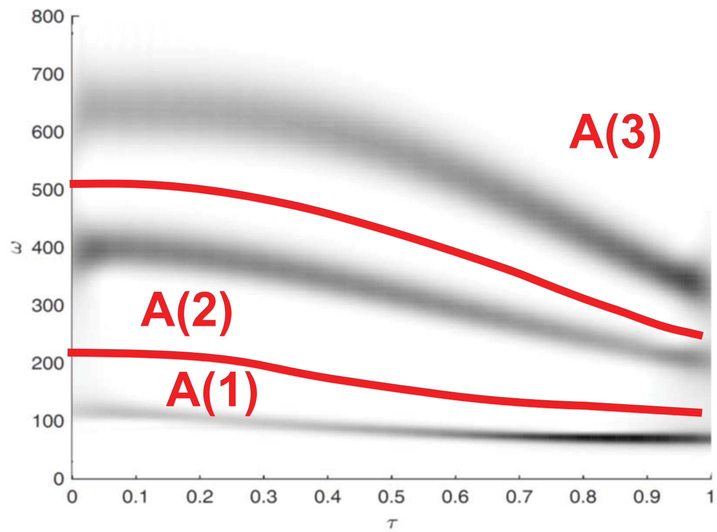
$$E(\tau, \omega) = \int_0^1 w_{\tau,\omega} K^{-1} v$$

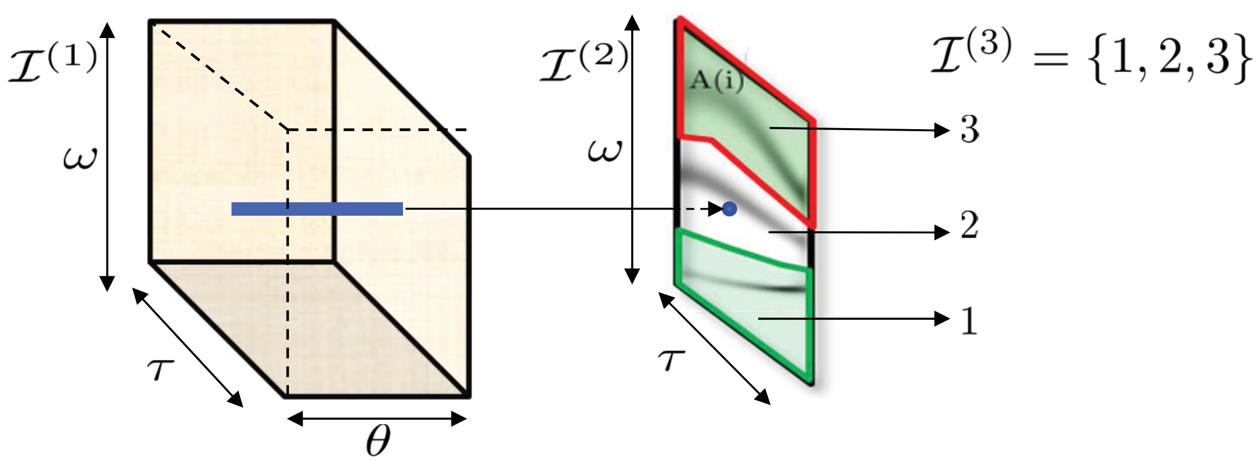




Partition the time-frequency domain to identify level 3 modes

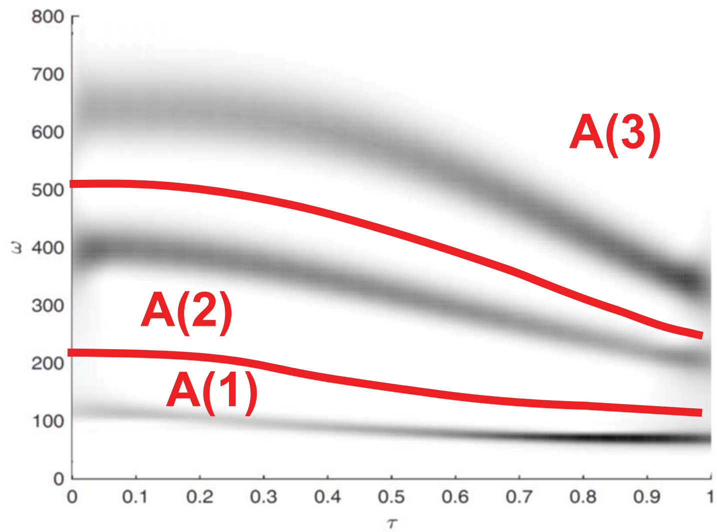
$$(\tau, \omega) \to i \iff (\tau, \omega) \in A(i)$$



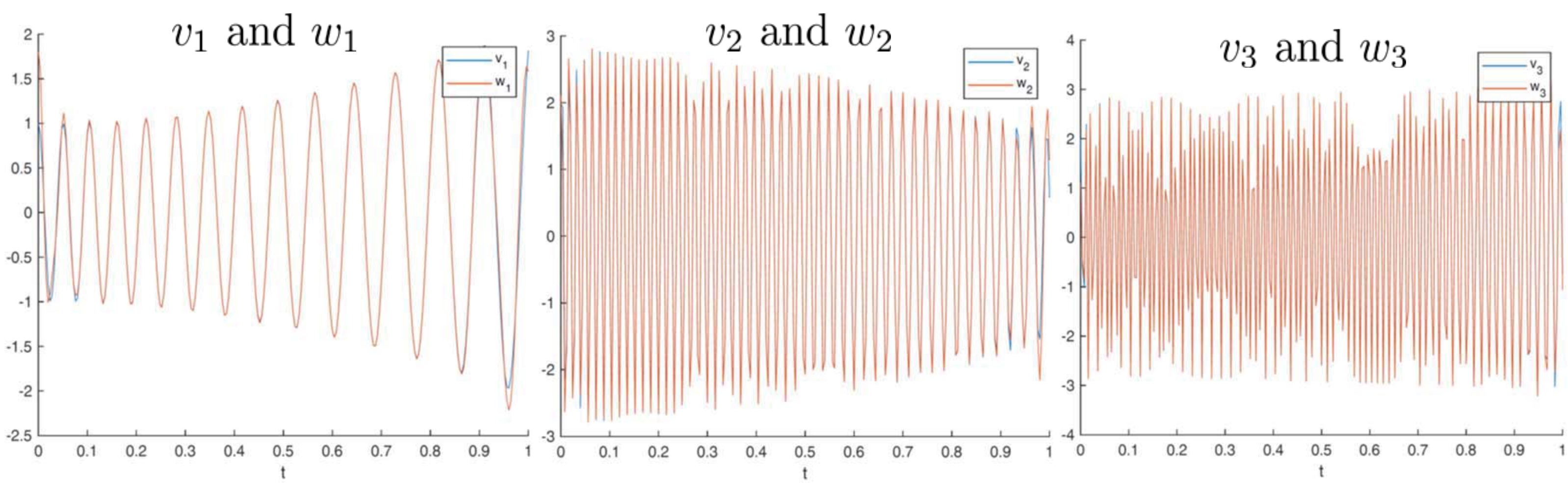


Aggregate to level 3 modes

$$w_{\tau,\omega} \xrightarrow{\int_{A(i)} \cdot \, d\tau \, d\omega} w_i$$

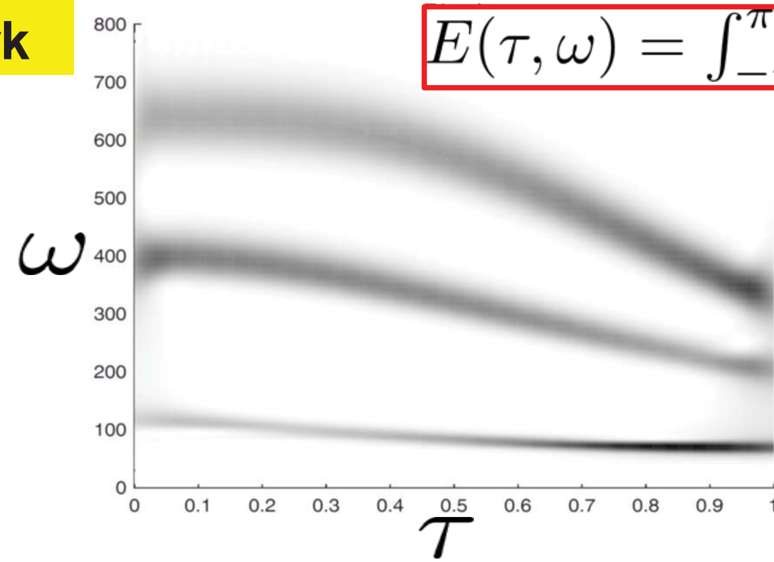


$$w_i = \int_{A(i)} w_{\tau,\omega} d\tau \, d\omega$$



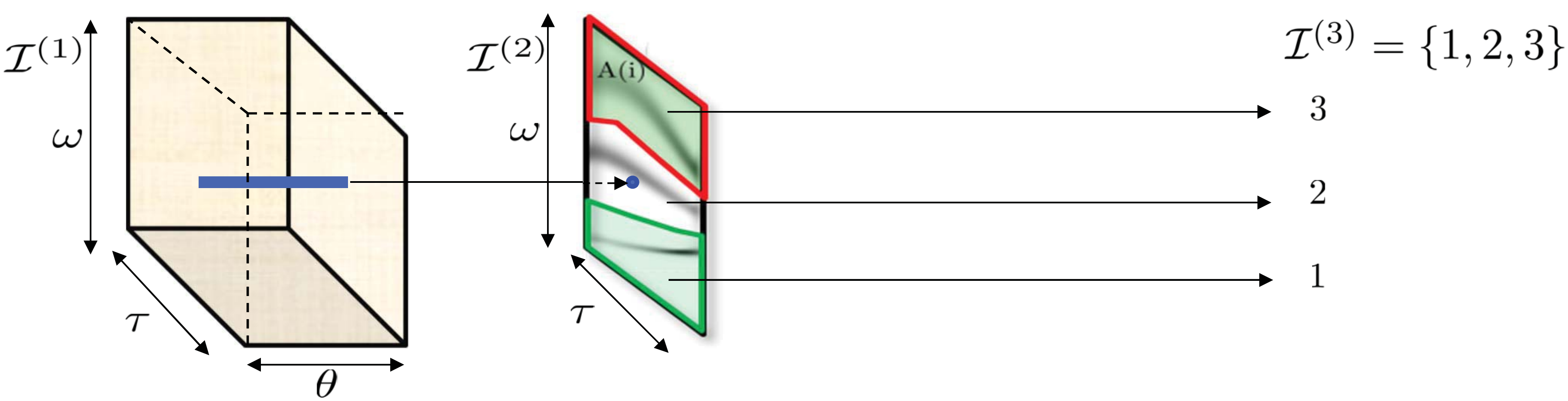
The simplest network

$$E(\tau,\omega) = \int_{-\pi}^{\pi} E(\tau,\omega,\theta)\, d\theta$$

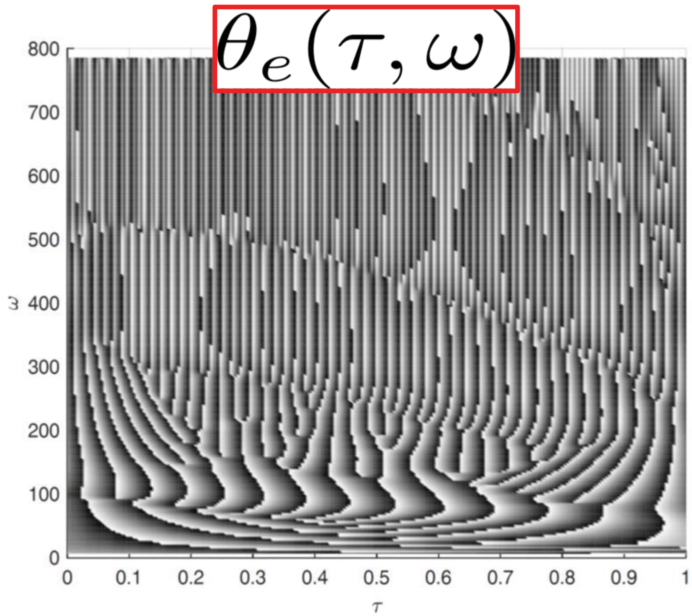


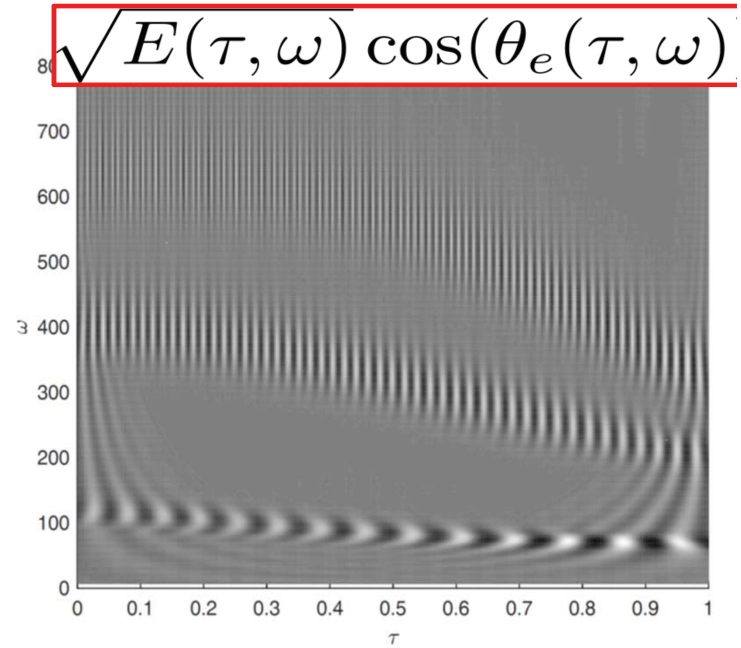
$$\int_0^1 w_{\tau,\omega,\theta} K^{-1} v \qquad \int_0^1 w_{\tau,\omega} K^{-1} v$$

$$E(\tau,\omega,\theta) \xrightarrow{\int_{-\pi}^{\pi} \cdot\, d\theta} E(\tau,\omega)$$

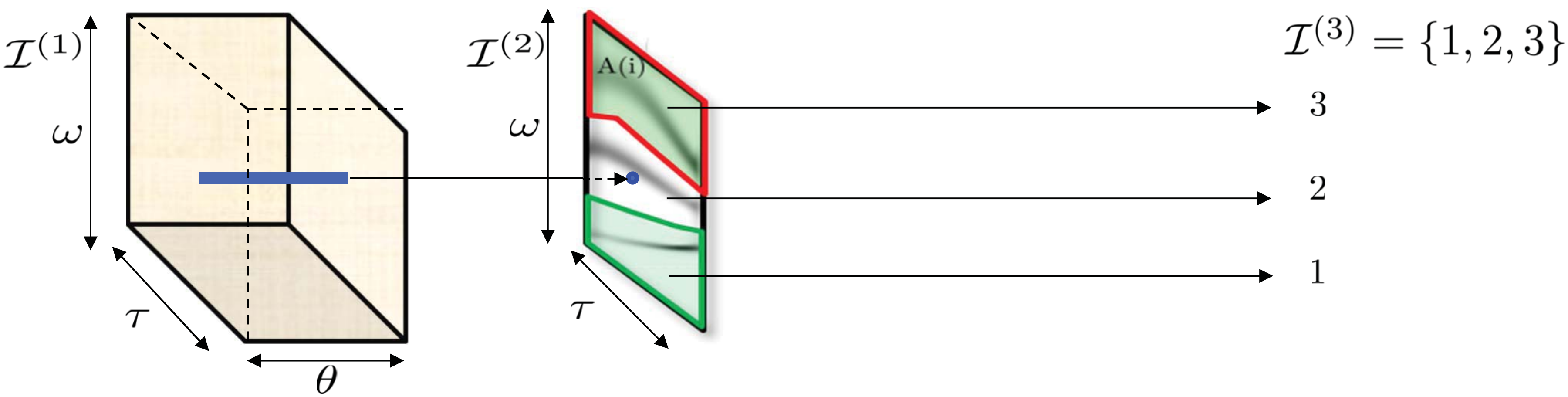


Estimated phase $\theta_e(\tau,\omega) := \mathrm{argmax}_\theta\, E(\tau,\omega,\theta)$



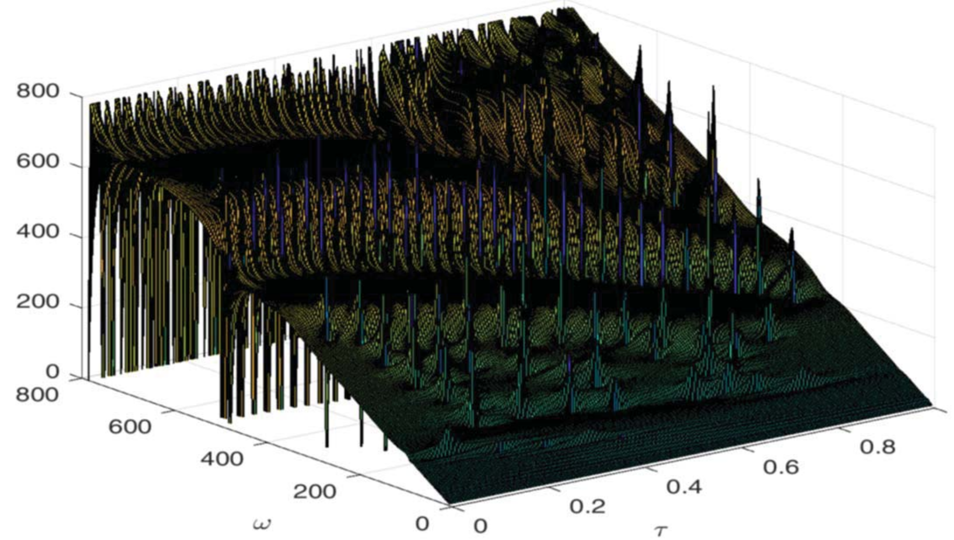


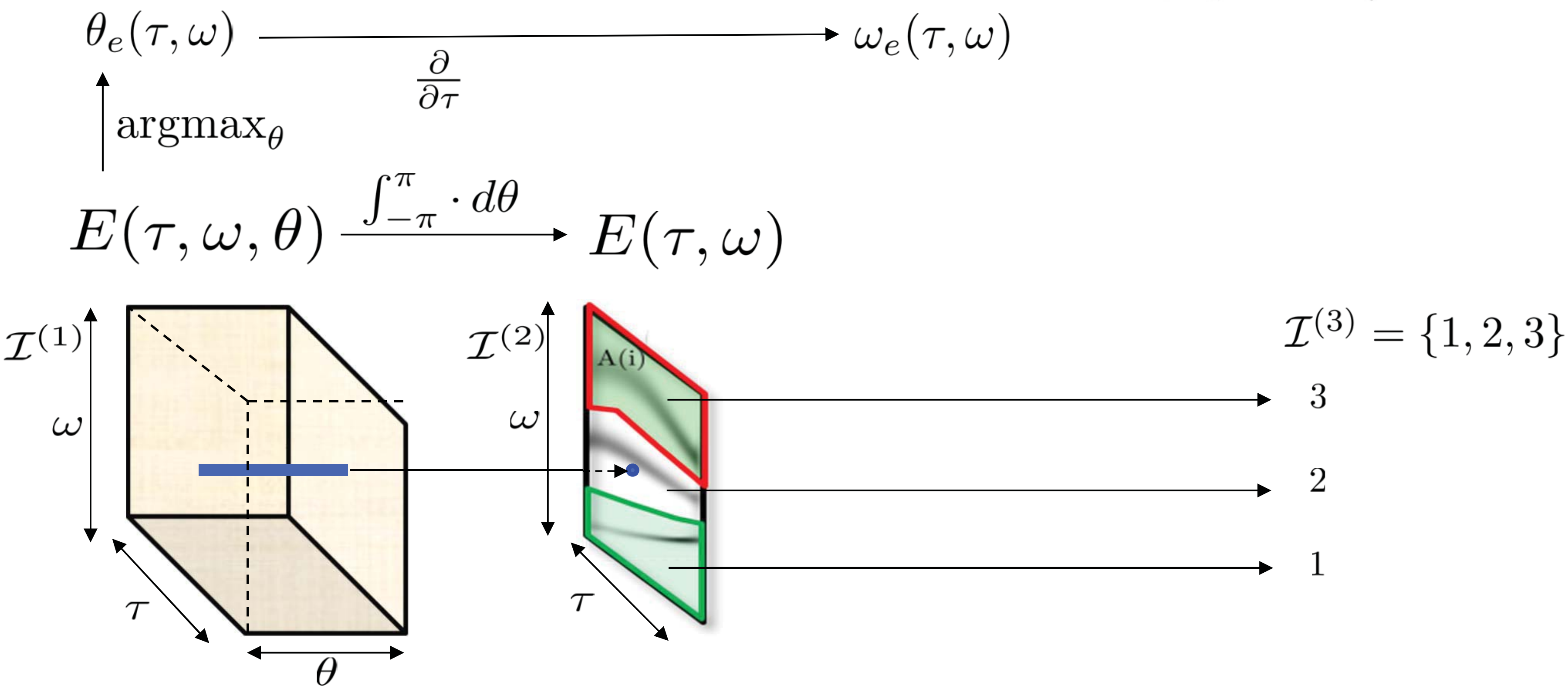
$$\theta_e(\tau,\omega) \xleftarrow{\mathrm{argmax}_\theta} E(\tau,\omega,\theta) \xrightarrow{\int_{-\pi}^{\pi}\cdot\, d\theta} E(\tau,\omega)$$



Estimated frequency

$$\omega_e(\tau,\omega) = \frac{\partial}{\partial\tau}\theta_e(\tau,\omega)$$

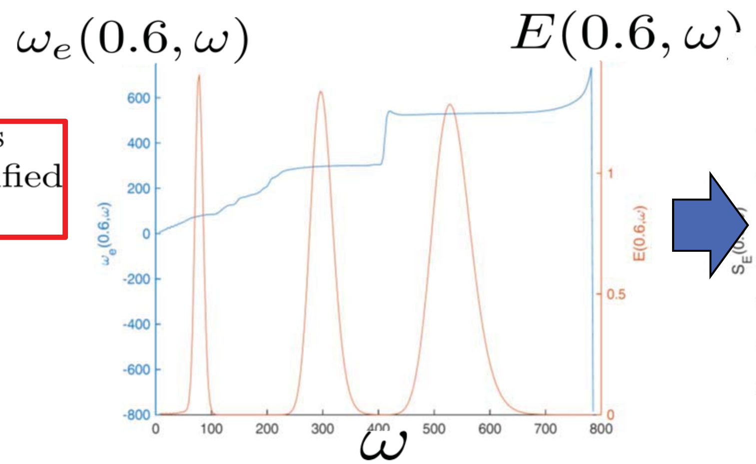


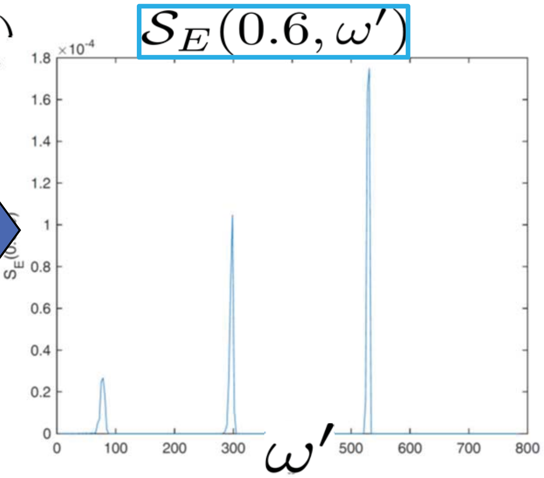


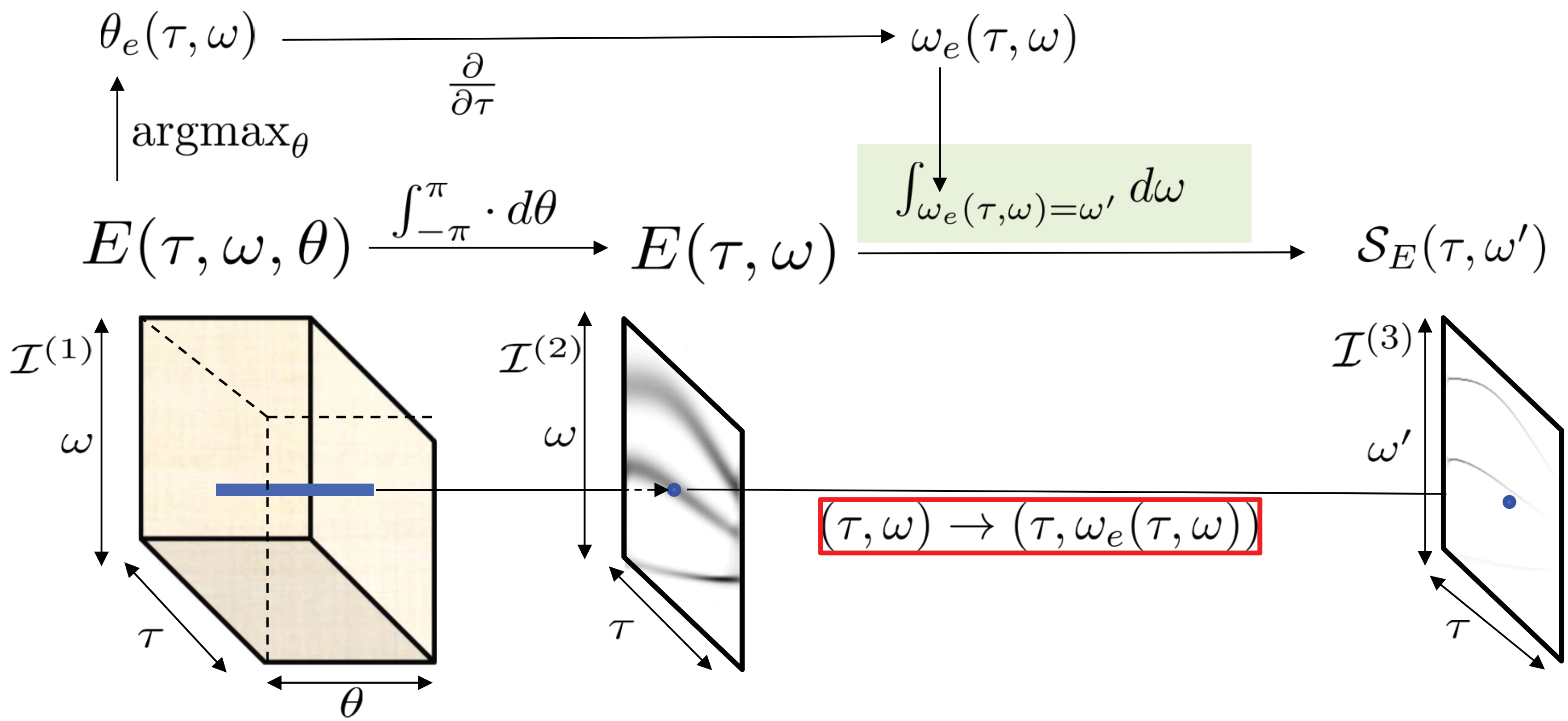
Synchro-squeezing

[Daubichies, Lu, Wu, 2011]

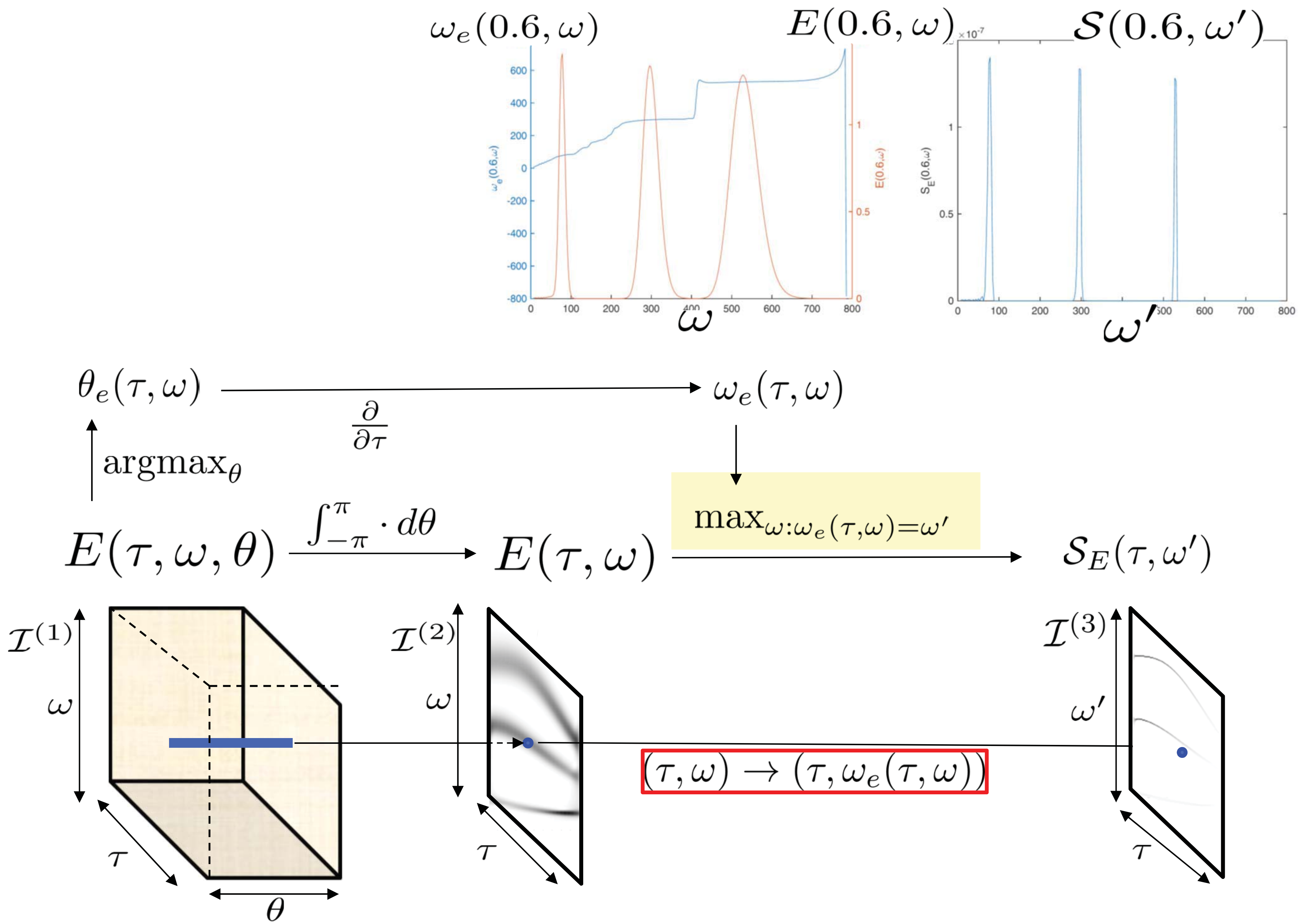
The synchrosqueezed transform is the energy of level 3 modes identified via the map $(\tau, \omega) \to (\tau, \omega_e(\tau, \omega))$







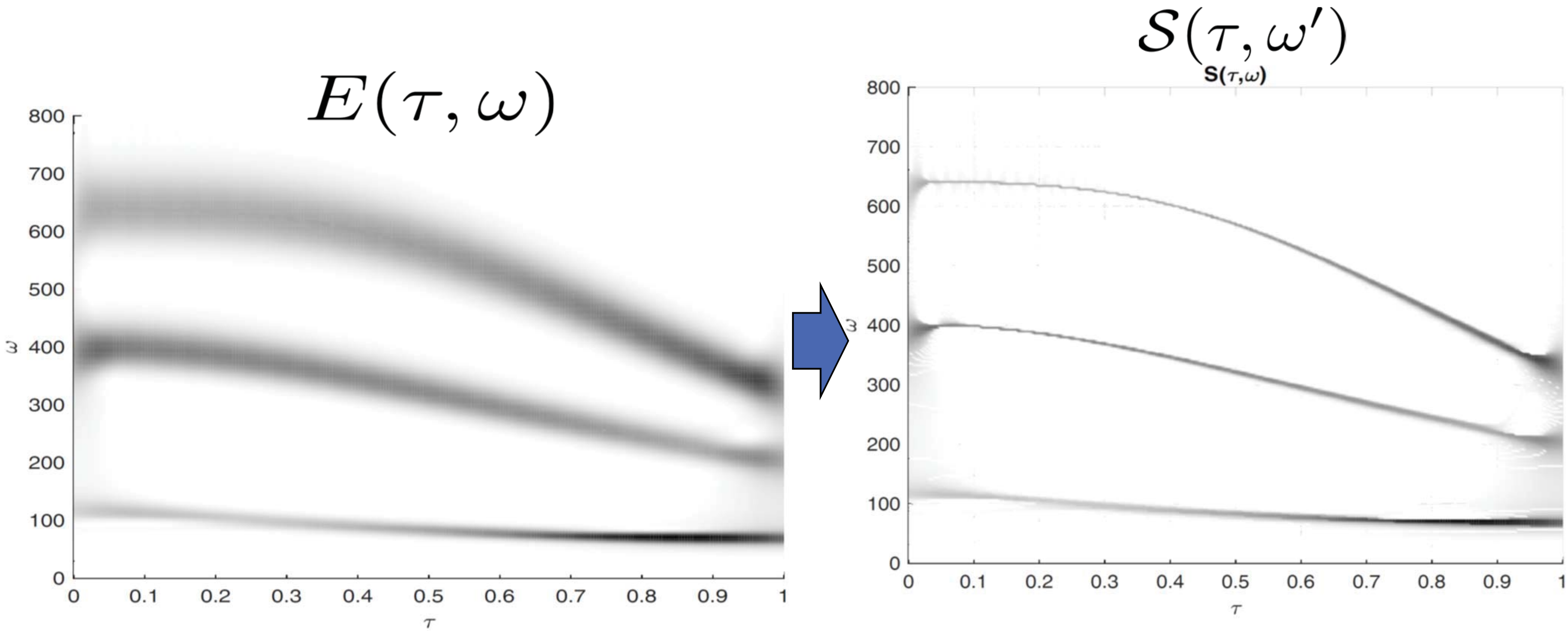
Max-squeezing



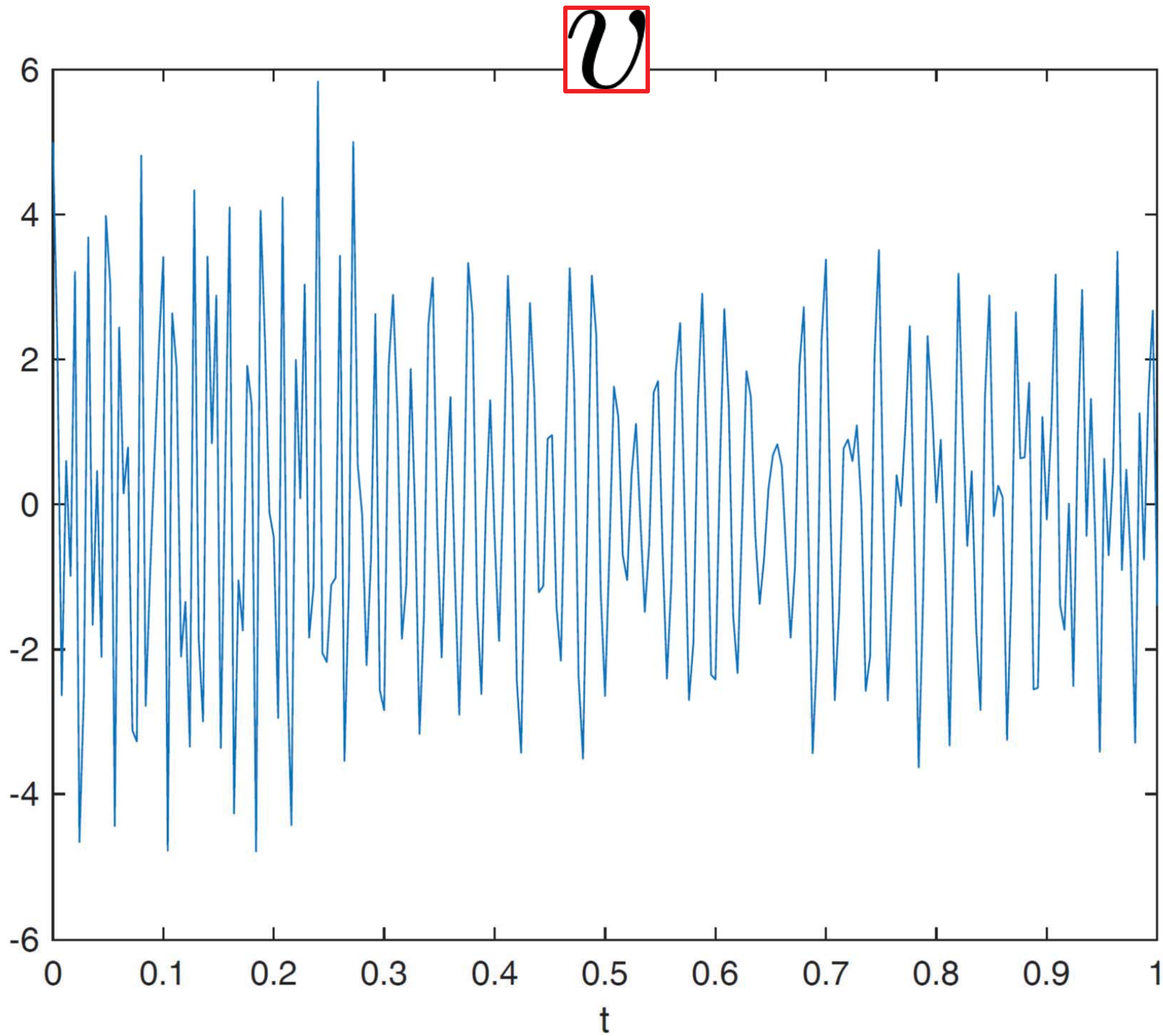
Max-squeezed energy

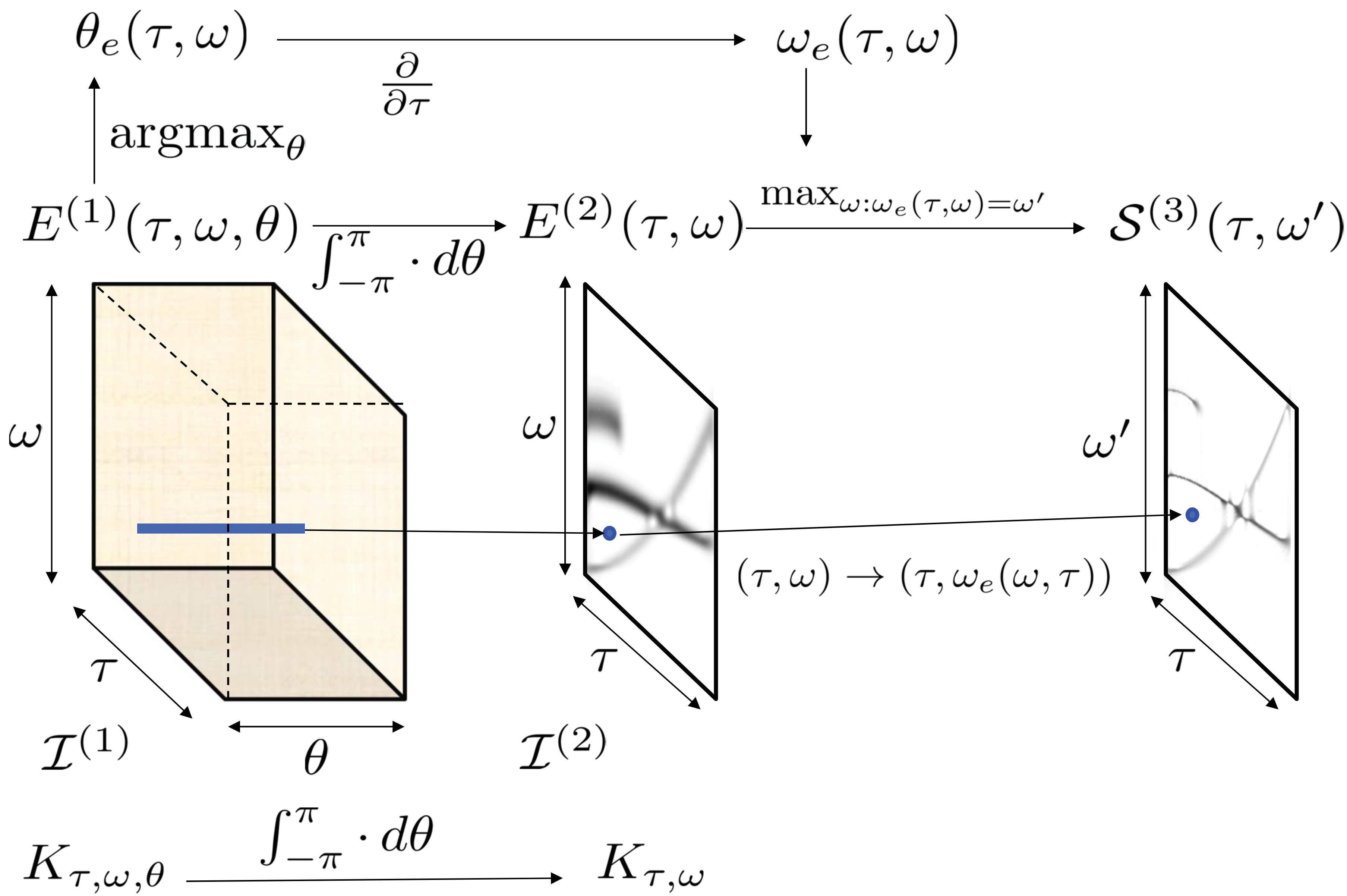
$$(\tau, \omega) \to (\tau, \omega_e(\tau, \omega))$$

$$\mathcal{S}(\tau, \omega') = \max_{\omega : \omega_e(\tau,\omega) = \omega'} E(\tau, \omega)$$



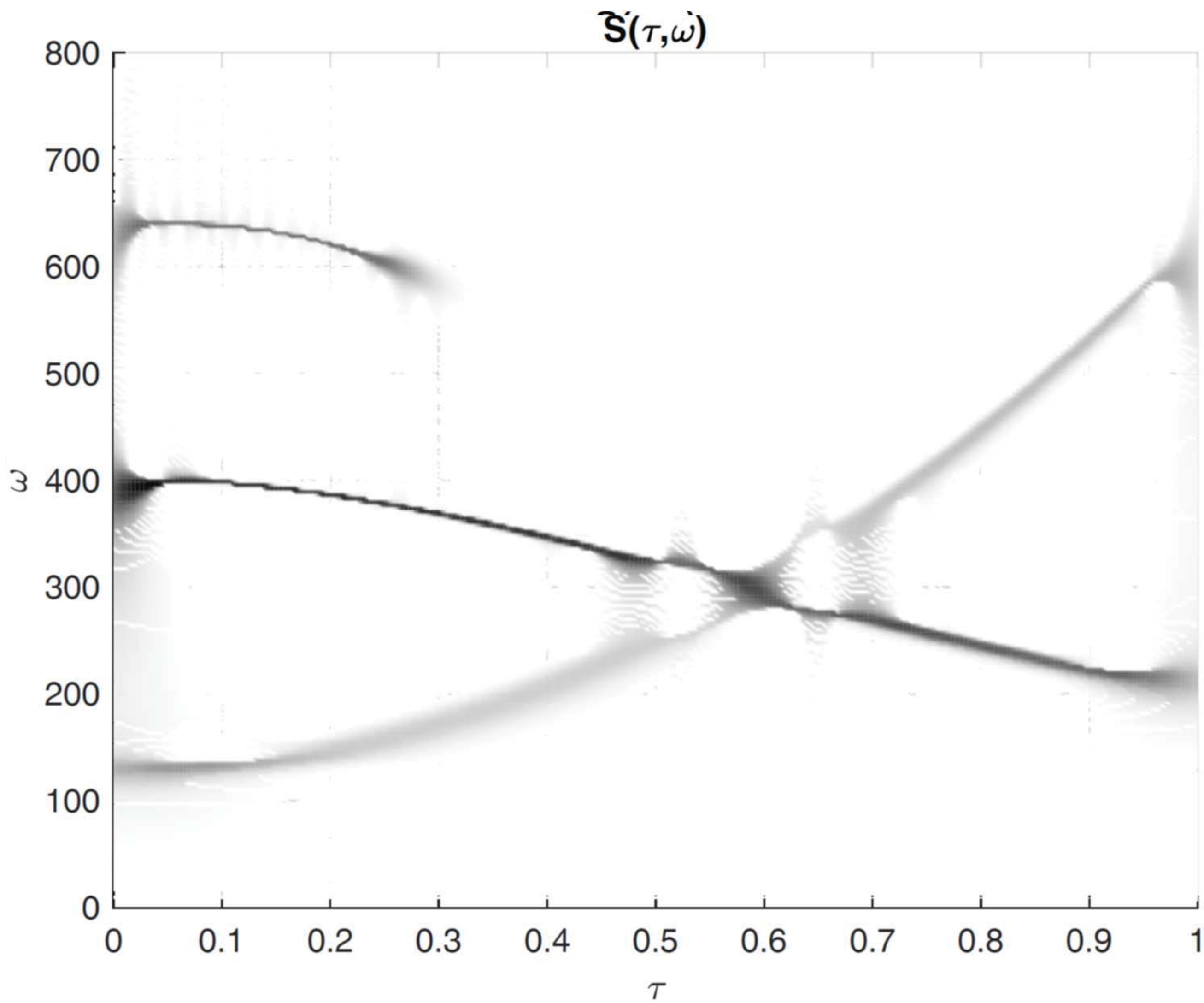
Example

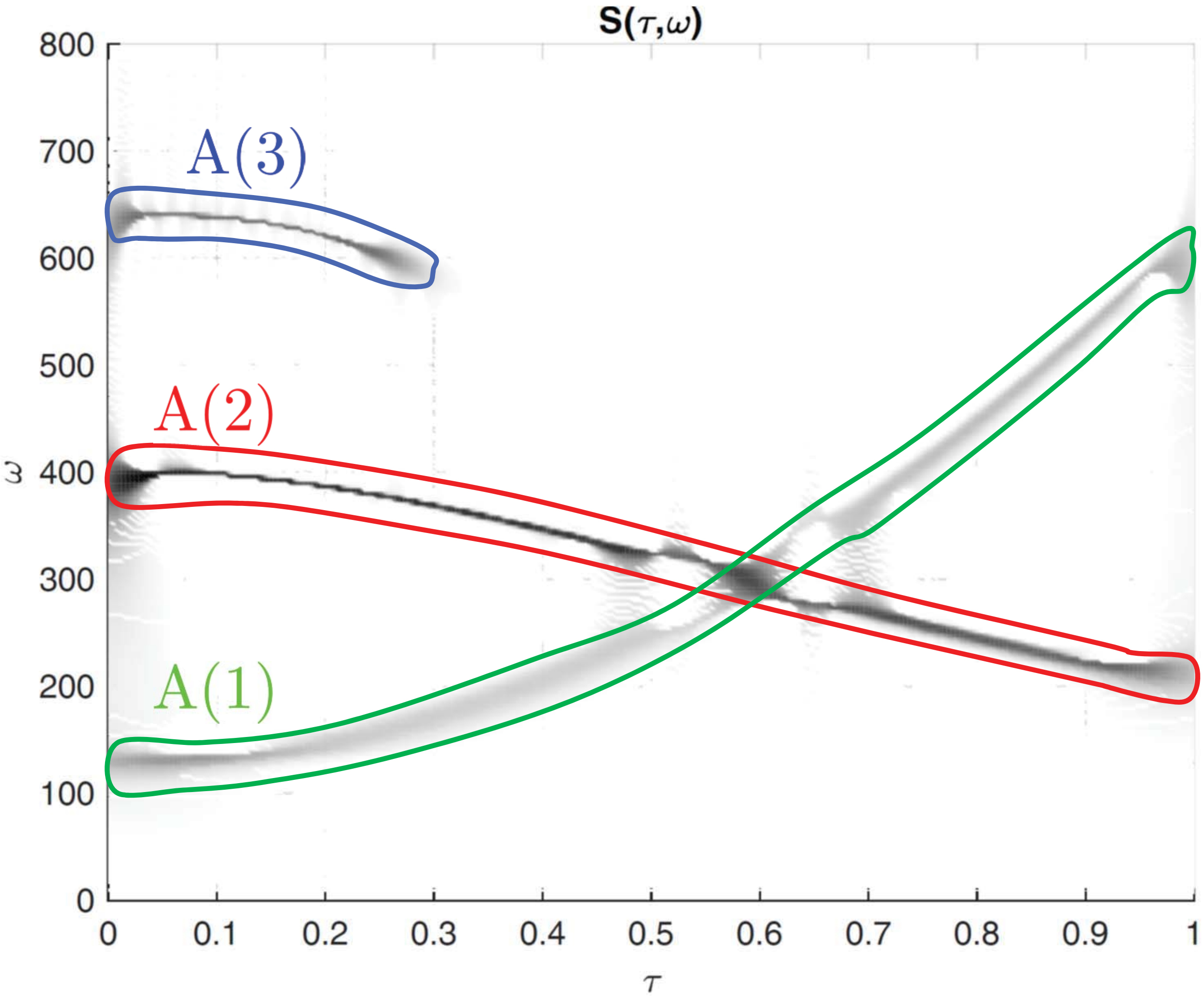




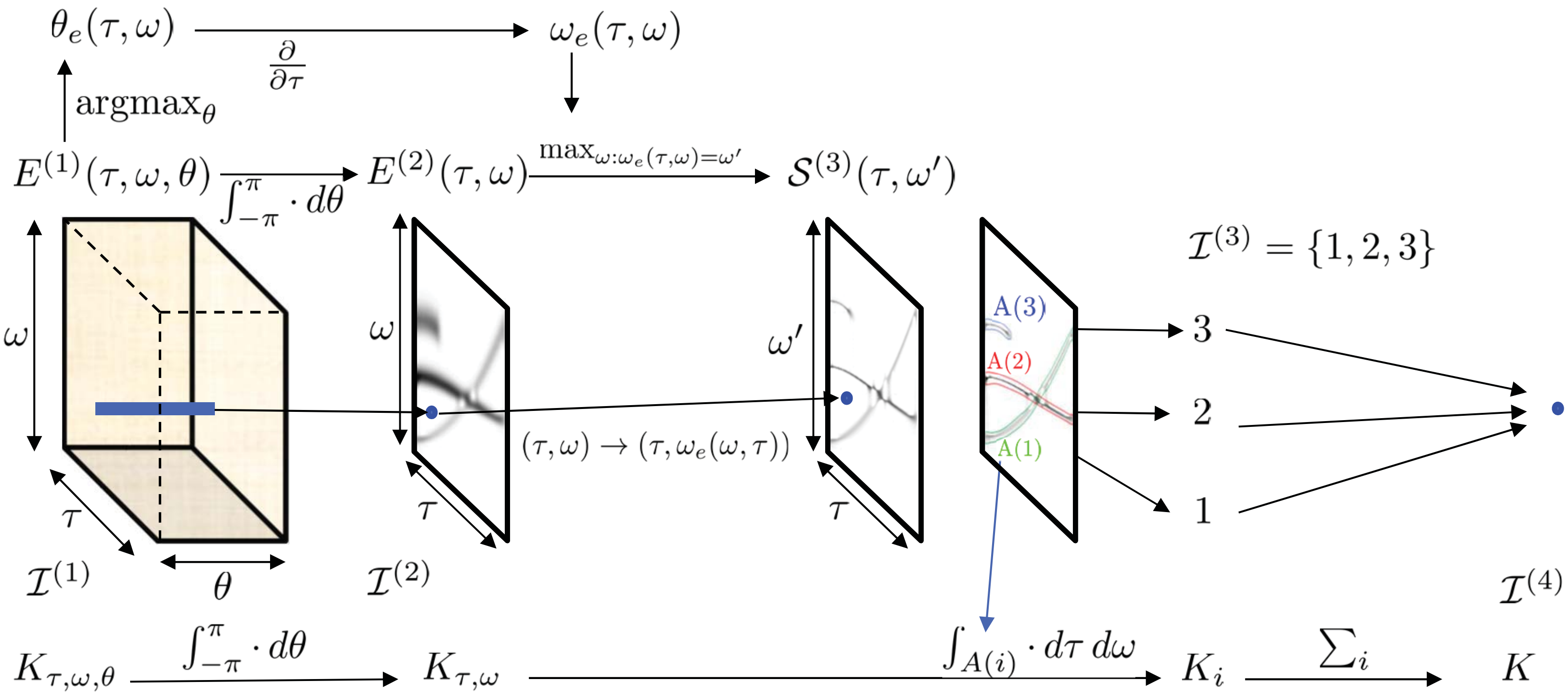
$\theta_e(\tau,\omega)$
$\frac{\partial}{\partial\tau}$
$\omega_e(\tau,\omega)$
argmax_θ
$E^{(1)}(\tau,\omega,\theta)$
$\int_{-\pi}^{\pi}\cdot d\theta$
$E^{(2)}(\tau,\omega)$
$\max_{\omega:\omega_e(\tau,\omega)=\omega'}$
$\mathcal{S}^{(3)}(\tau,\omega')$
ω
ω
ω'
$(\tau,\omega)\to(\tau,\omega_e(\omega,\tau))$
τ
τ
τ
$\mathcal{I}^{(1)}$
θ
$\mathcal{I}^{(2)}$
$K_{\tau,\omega,\theta}$
$\int_{-\pi}^{\pi}\cdot d\theta$
$K_{\tau,\omega}$

Max-squeezed Energy

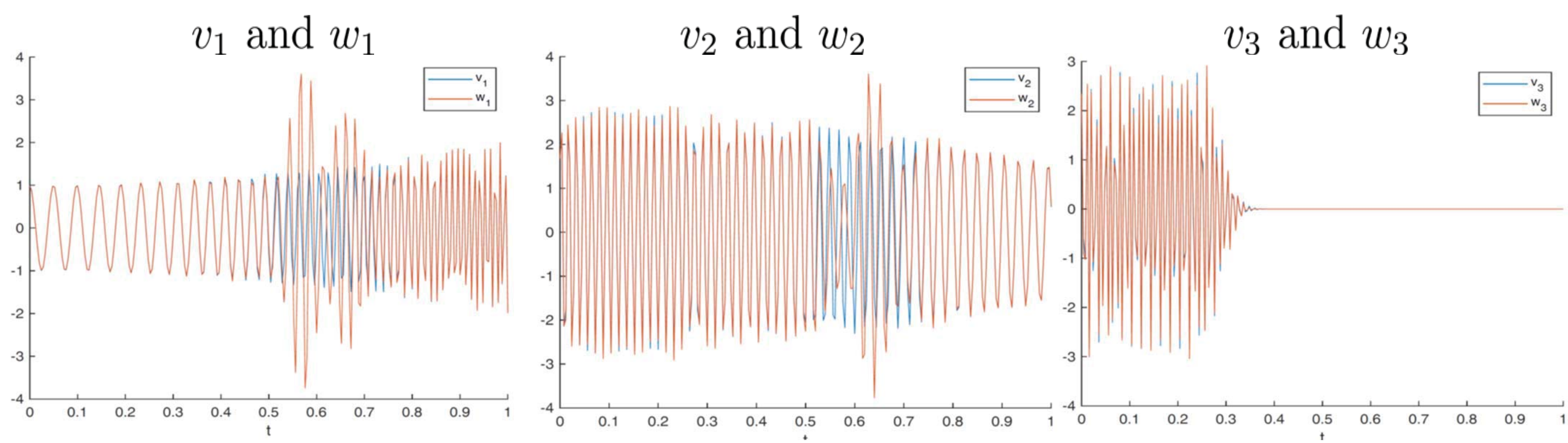


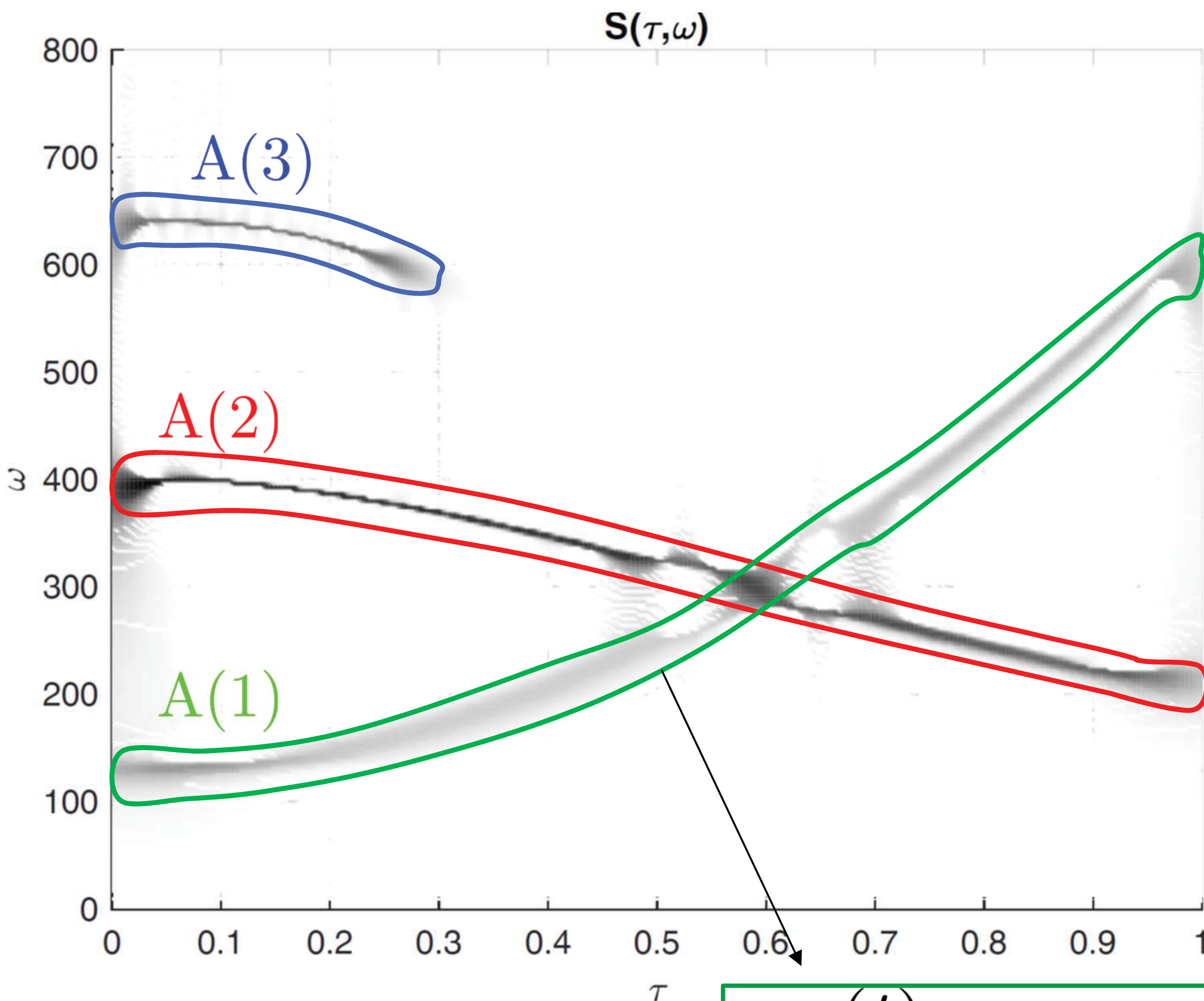


$$K_i = \int_{A(i)} K_{\tau,\omega}\, d\tau\, d\omega$$



$$w_i = K_i\left(\sum_j K_j\right)^{-1} v$$

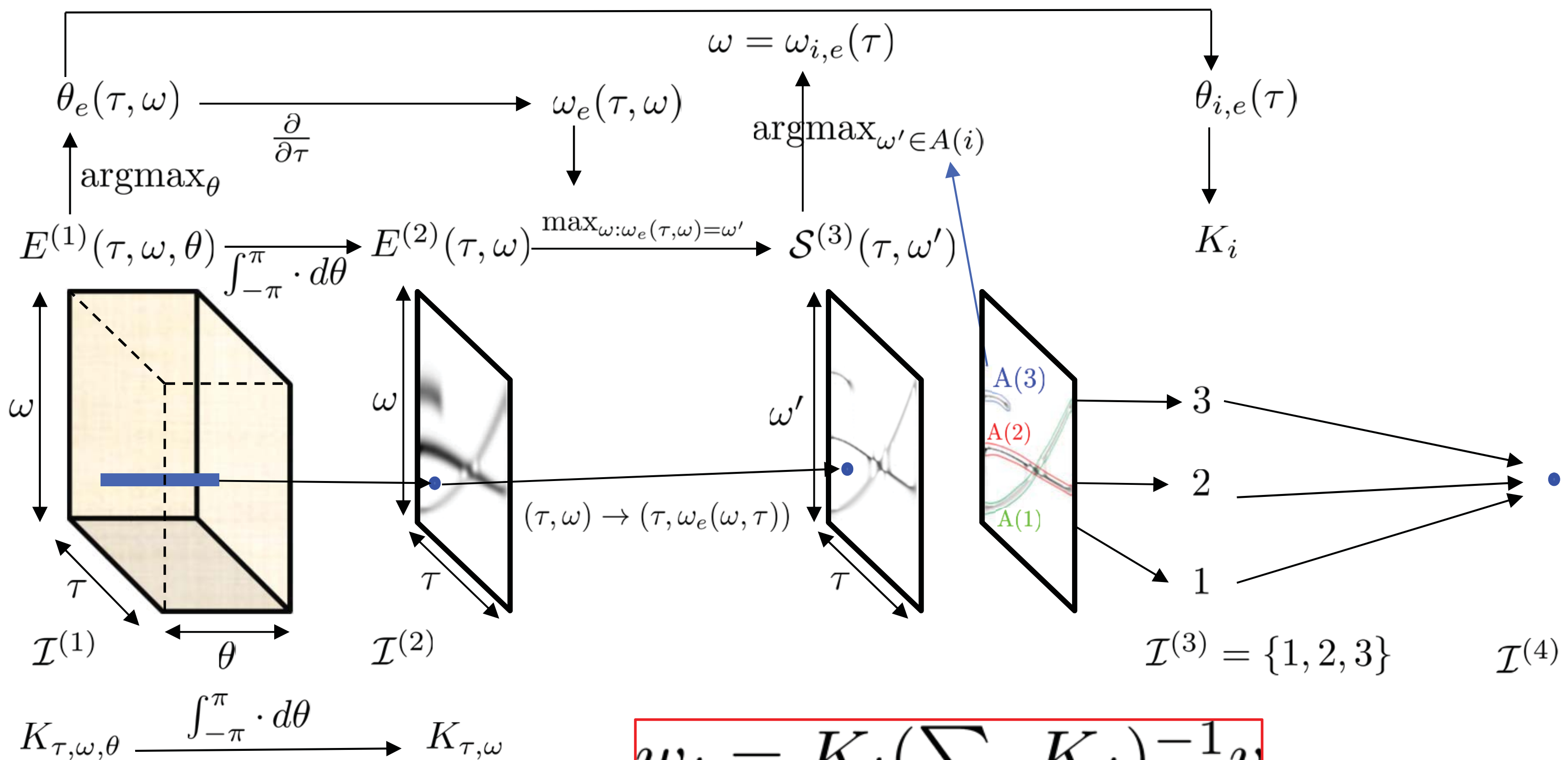




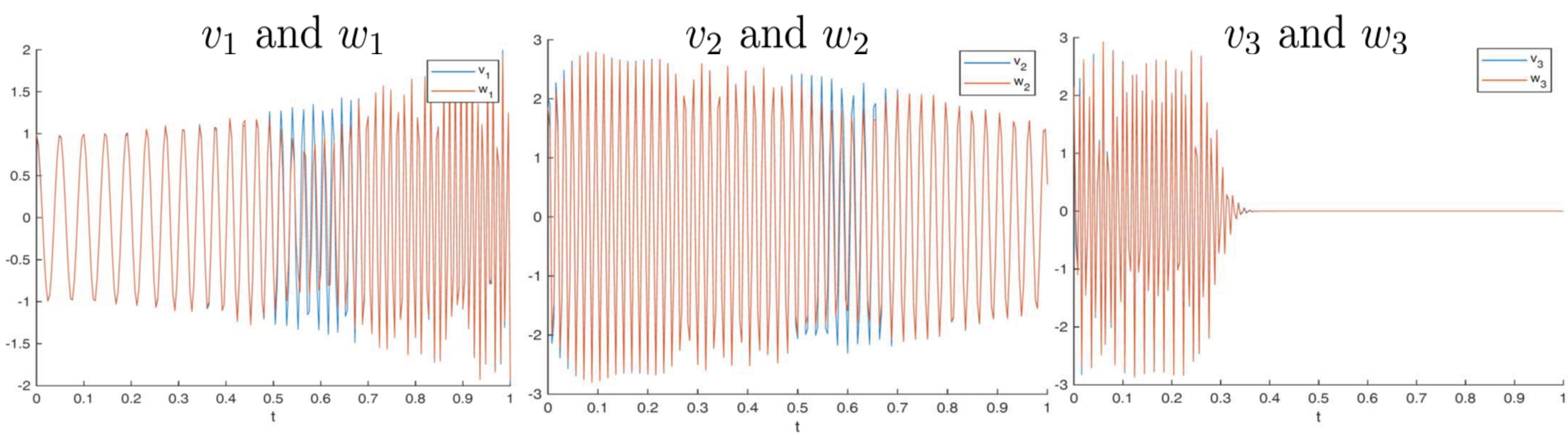
$$\omega_{i,e}(t) = \mathrm{argmax}_{(t,\omega)\in A(i)}\, \mathcal{S}(t,\omega)$$

$$\theta_{i,e}(t) = \theta_e(t, \omega_{i,e}(t))$$

$$K_i(s,t) = e^{-\frac{(t-s)^2}{\gamma^2}} \left(\cos(\theta_{i,e}(t)) \cos(\theta_{i,e}(s)) + \sin(\theta_{i,e}(t)) \sin(\theta_{i,e}(s)) \right)$$



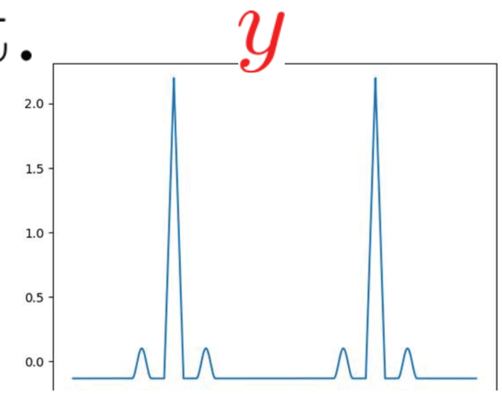
$$w_i = K_i\left(\sum_j K_j\right)^{-1} v$$



Ex: Amplitudes and phases unknown, waveforms known but non-trigonometric

Let $v_1, \ldots, v_m$ be unknown s.t.

m: unknown

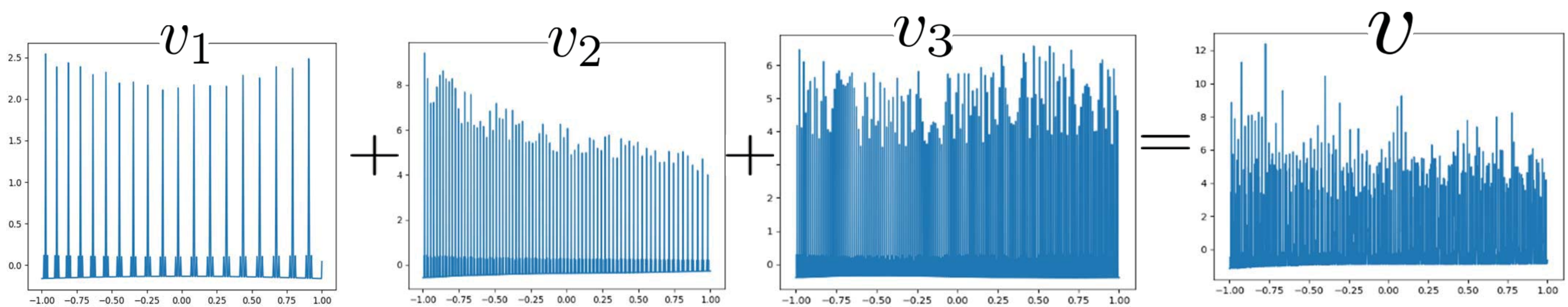


$$v_i(t) = a_i(t) y(\theta_i(t))$$

a_i: unknown, slowly varying

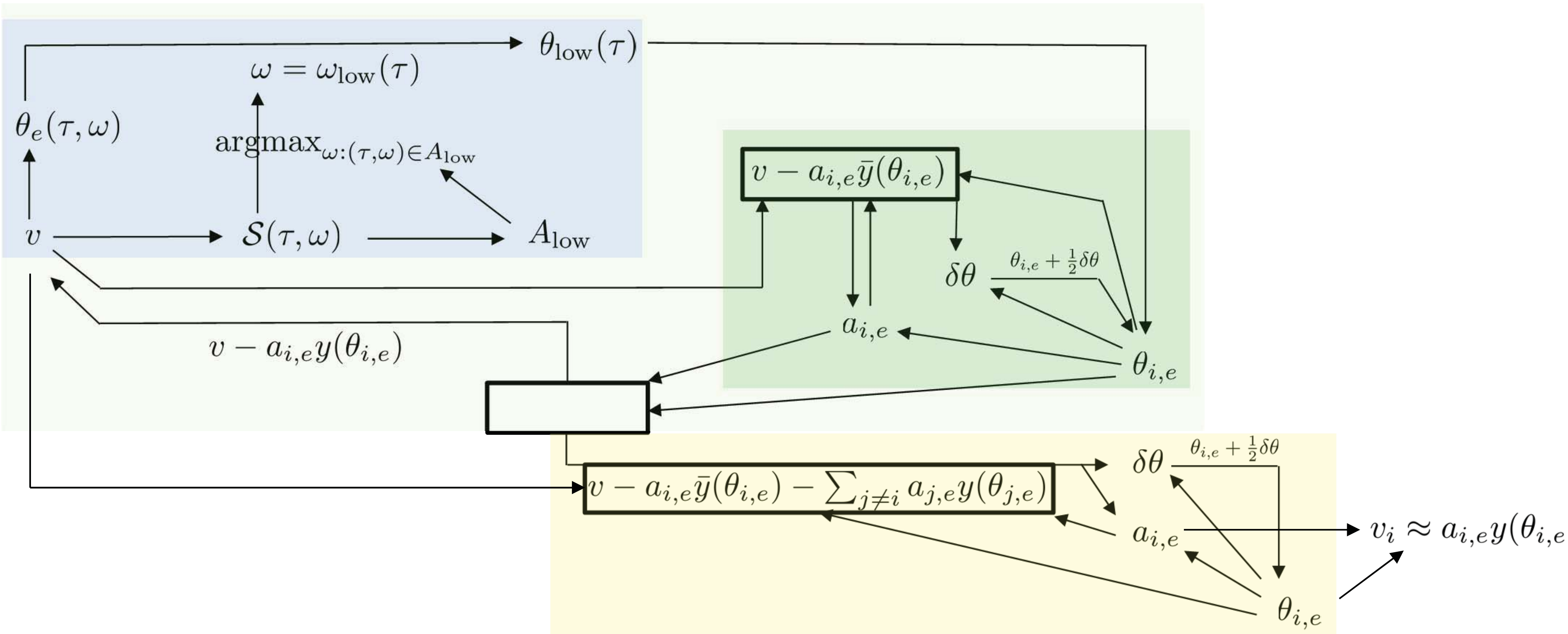
$\omega_i := \dot{\theta}_i$: unknown, slowly varying, positive, well separated

y: Known non-trigonometric



Problem Given $v = \sum_{i=1}^{m} v_i$ recover $v_1, \ldots, v_m$

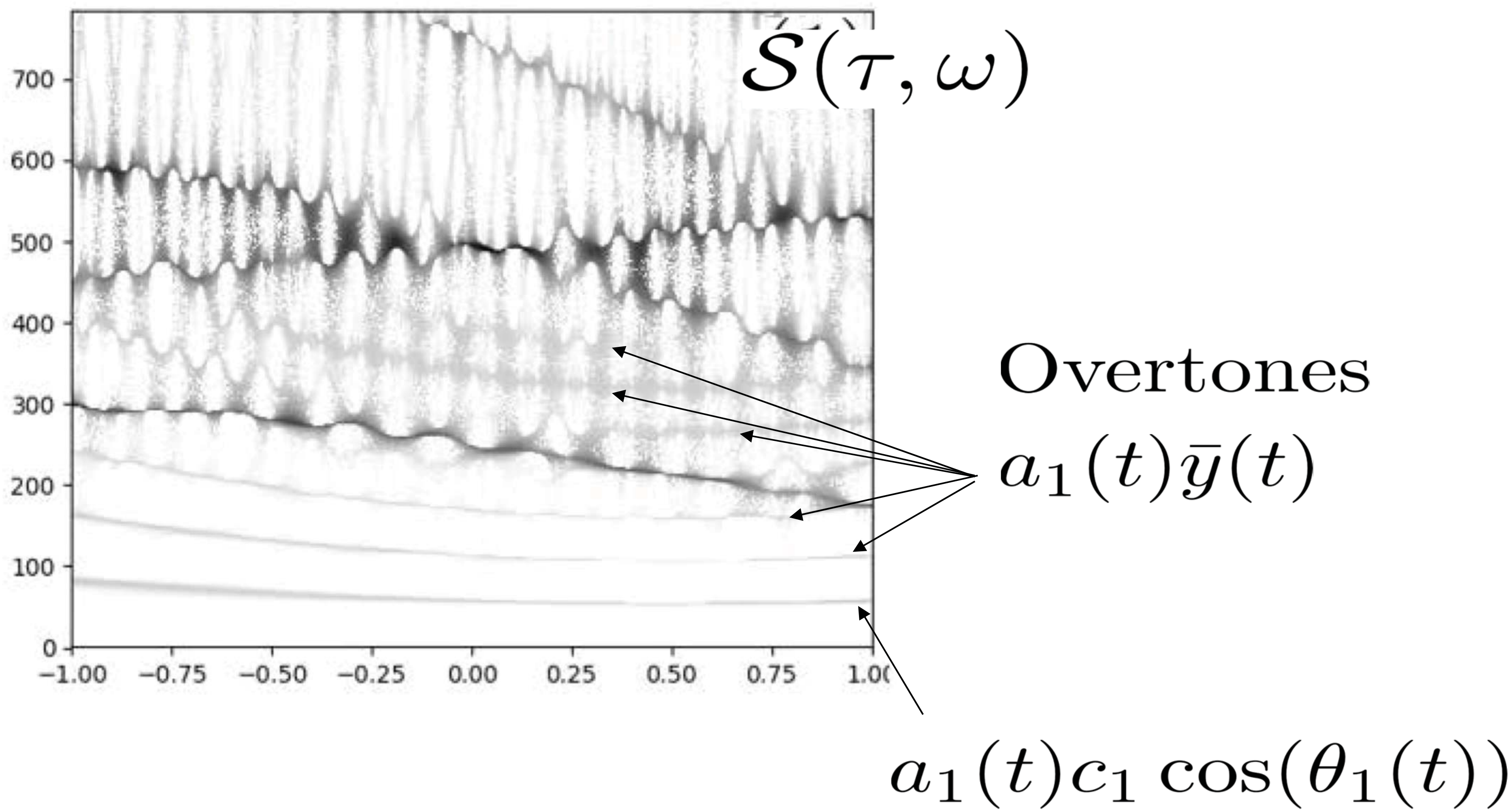
The network

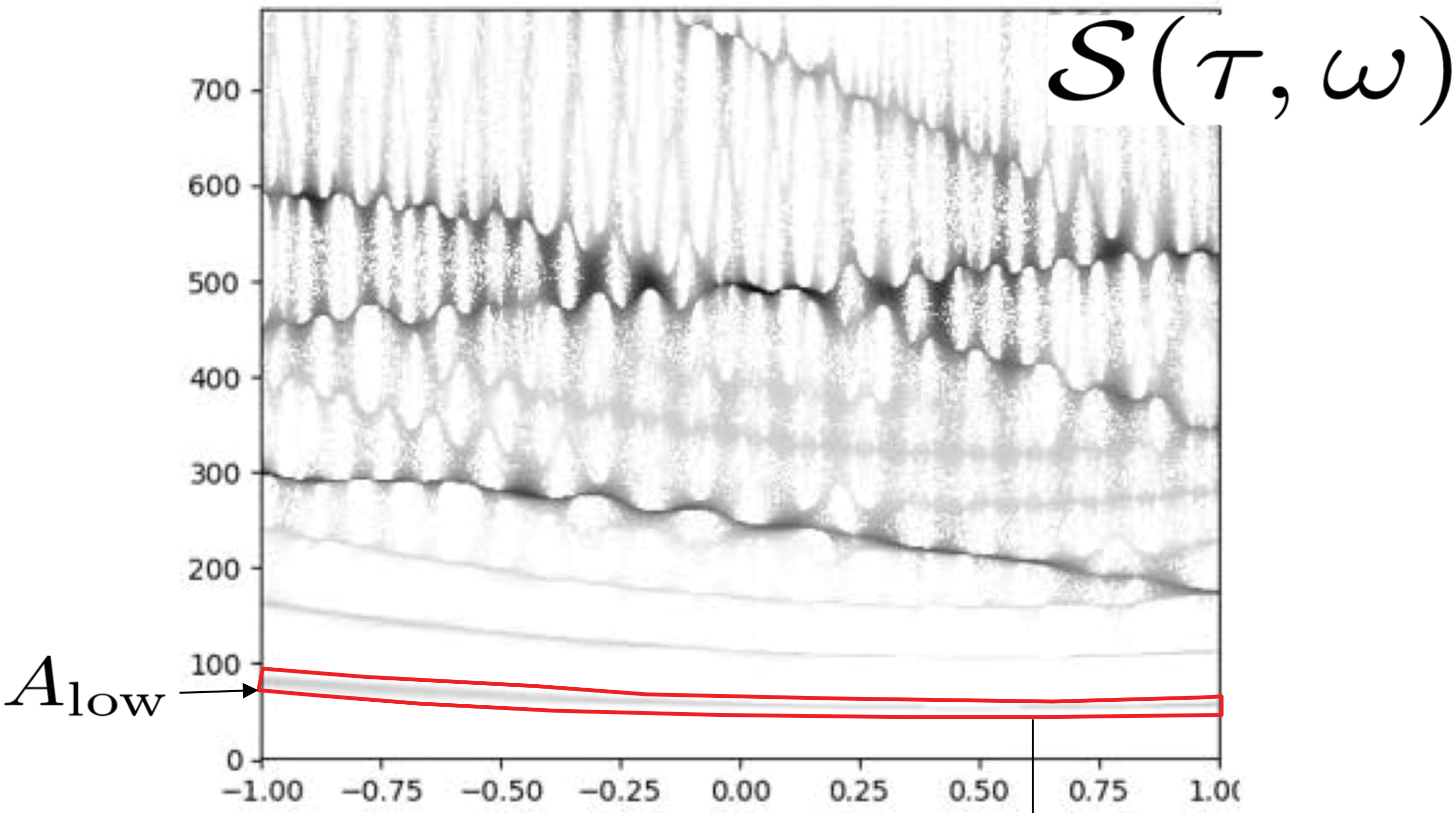


$\theta_{\text{low}}(\tau)$
$\omega = \omega_{\text{low}}(\tau)$
$\theta_e(\tau, \omega)$
$\text{argmax}_{\omega:(\tau,\omega)\in A_{\text{low}}}$
v
$\mathcal{S}(\tau, \omega)$
A_{low}
$v - a_{i,e}\bar{y}(\theta_{i,e})$
$\delta\theta$
$\theta_{i,e} + \frac{1}{2}\delta\theta$
$a_{i,e}$
$\theta_{i,e}$
$v - a_{i,e}y(\theta_{i,e})$
$v - a_{i,e}\bar{y}(\theta_{i,e}) - \sum_{j\neq i} a_{j,e}y(\theta_{j,e})$
$\delta\theta$
$\theta_{i,e} + \frac{1}{2}\delta\theta$
$a_{i,e}$
$v_i \approx a_{i,e}y(\theta_{i,e})$
$\theta_{i,e}$

$$y(t) = c_1 \cos(t) + \underbrace{\sum_{n=2}^{\infty} c_n \cos(nt + d_n)}_{\bar{y}(t)}$$

Max squeezed energy



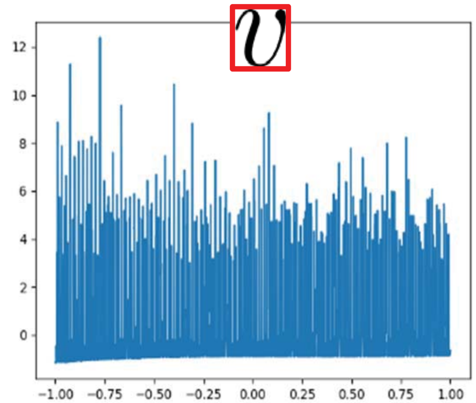


$$\omega_{1,e}(t) = \mathrm{argmax}_{(t,\omega)\in A_{\mathrm{low}}} \mathcal{S}(t, \omega)$$

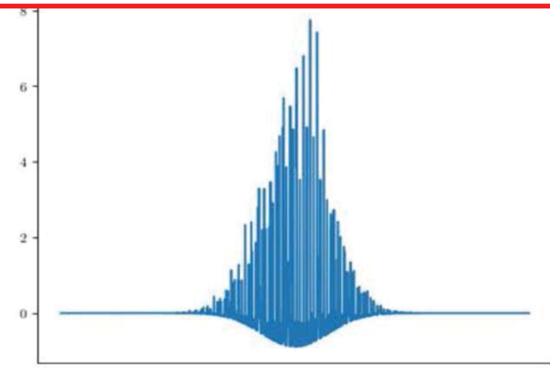
$$\theta_{1,e}(t) = \theta_e(t, \omega_{1,e}(t))$$

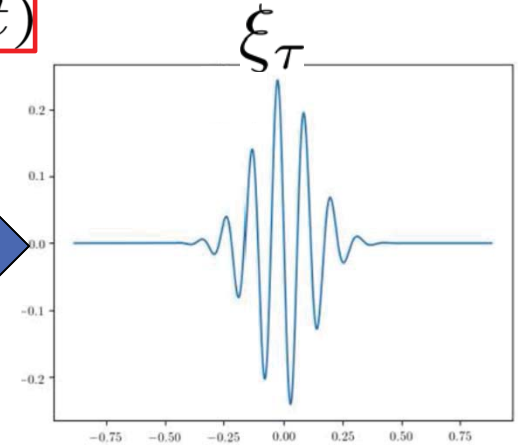
Estimated phase of mode 1

Next step: refine the estimated phase of mode 1



$$v_\tau(t) = e^{-\left(\frac{\dot\theta_{1,e}(\tau)(t-\tau)}{\alpha}\right)^2} v(t)$$





$$Z_c := \lim_{\sigma\downarrow 0} \mathbb{E}\left[X_c \middle| \xi_\tau + \xi_\sigma = v_\tau\right]$$

$$\xi_\tau(t) := \left(X_c\cos(\theta_{1,e}(t)) + X_s\sin(\theta_{1,e}(t))\right)e^{-\left(\frac{\dot\theta_{1,e}(\tau)(t-\tau)}{\alpha}\right)^2}$$

$$X_c, X_s \sim \mathcal{N}(0,1) \qquad \xi_\sigma \sim \mathcal{N}\left(0, \sigma^2\delta(t-s)\right)$$

For $t \approx \tau$

$$a_1(t)\cos(\theta_1(t)) \approx Z_c\cos(\theta_{1,e}(t)) + Z_s\sin(\theta_{1,e}(t))$$

$$a_1(\tau) \approx \sqrt{Z_c{}^2 + Z_s{}^2} \qquad (\theta_1 - \theta_{1,e})(\tau) \approx \operatorname{atan2}(-Z_s, Z_c)$$

Estimate of mode 1

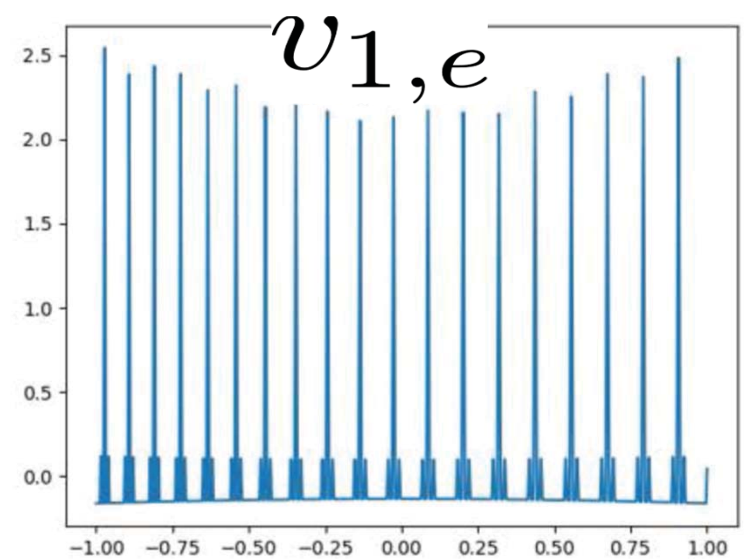
$$\theta_{1,e}(t)$$

Micro-local KMD

$+$

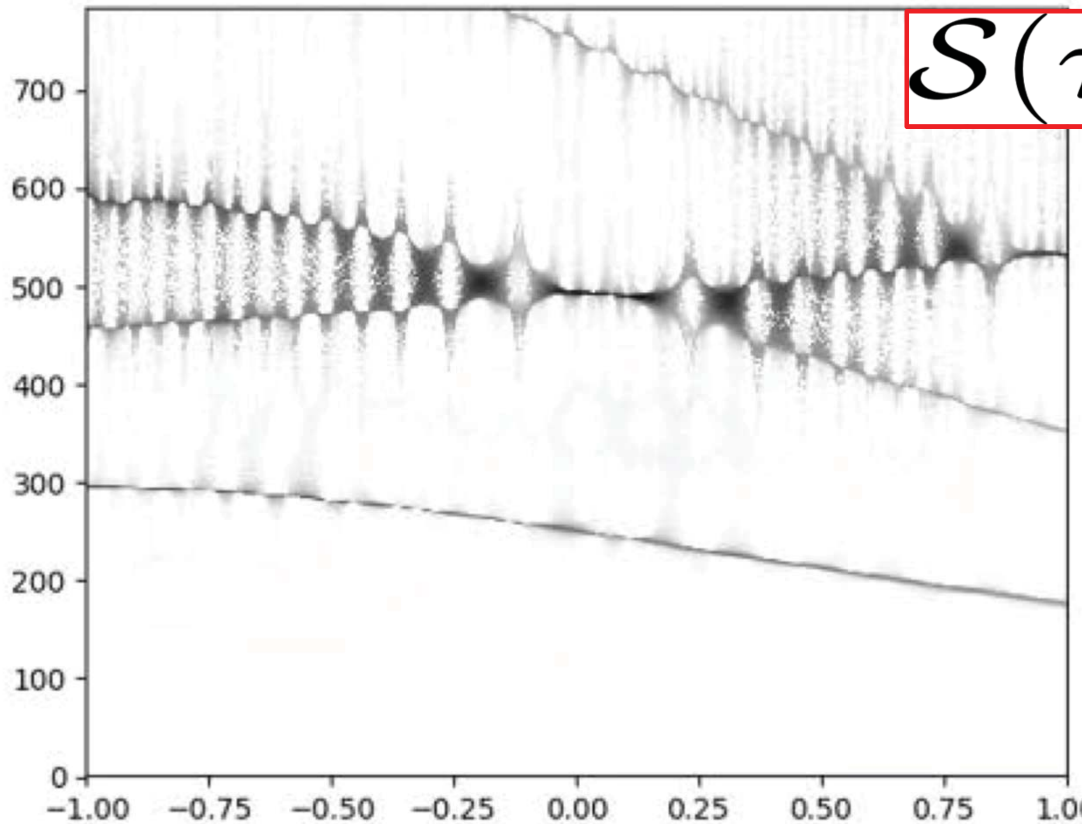
$$a_{1,e}(t),\ \delta\theta_{1,e} \approx \theta_1 - \theta_{1,e}$$

$$v_{1,e} = a_{1,e}(t)y(\theta_{1,e}(t))$$



Max-squeezed Energy

$$v - v_{1,e}$$

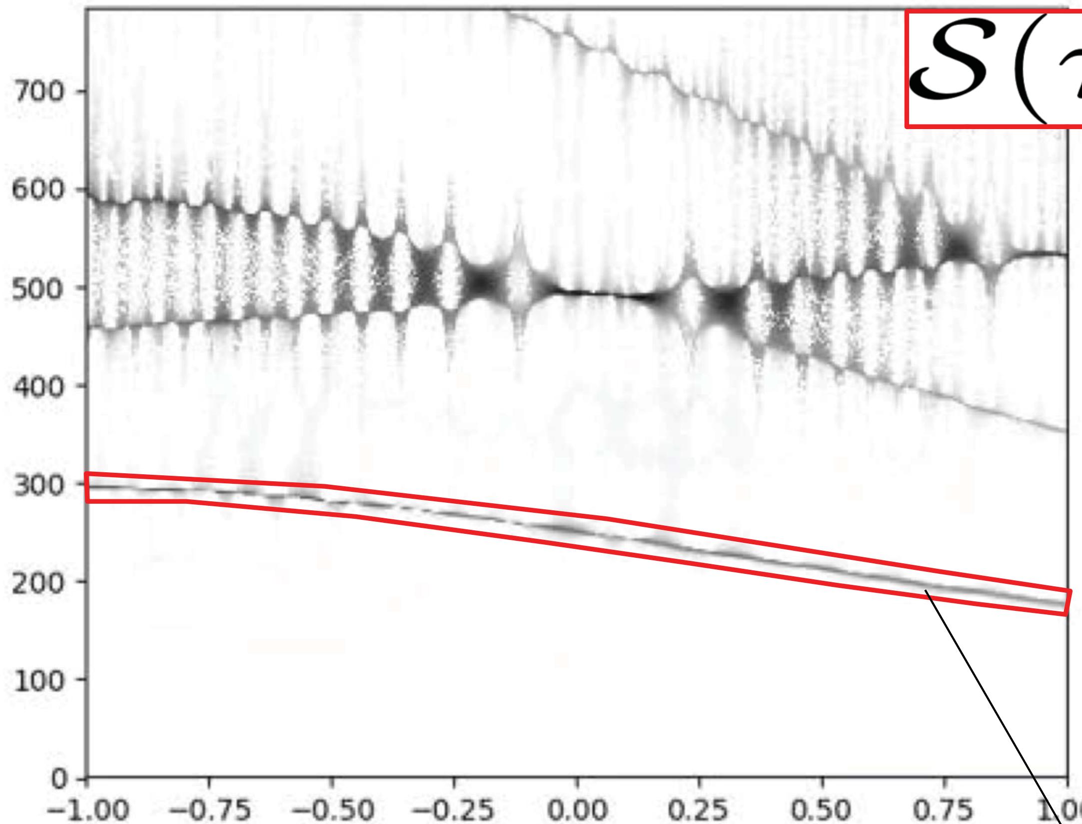


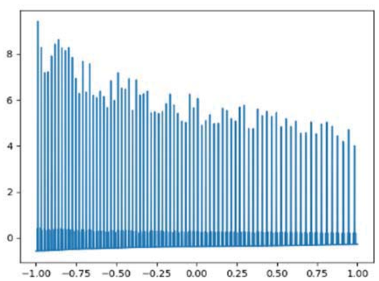
$$S(\tau, \omega)$$

Max-squeezed Energy

$$v - v_{1,e}$$

$$\mathcal{S}(\tau, \omega)$$





$$v_{2,e}$$

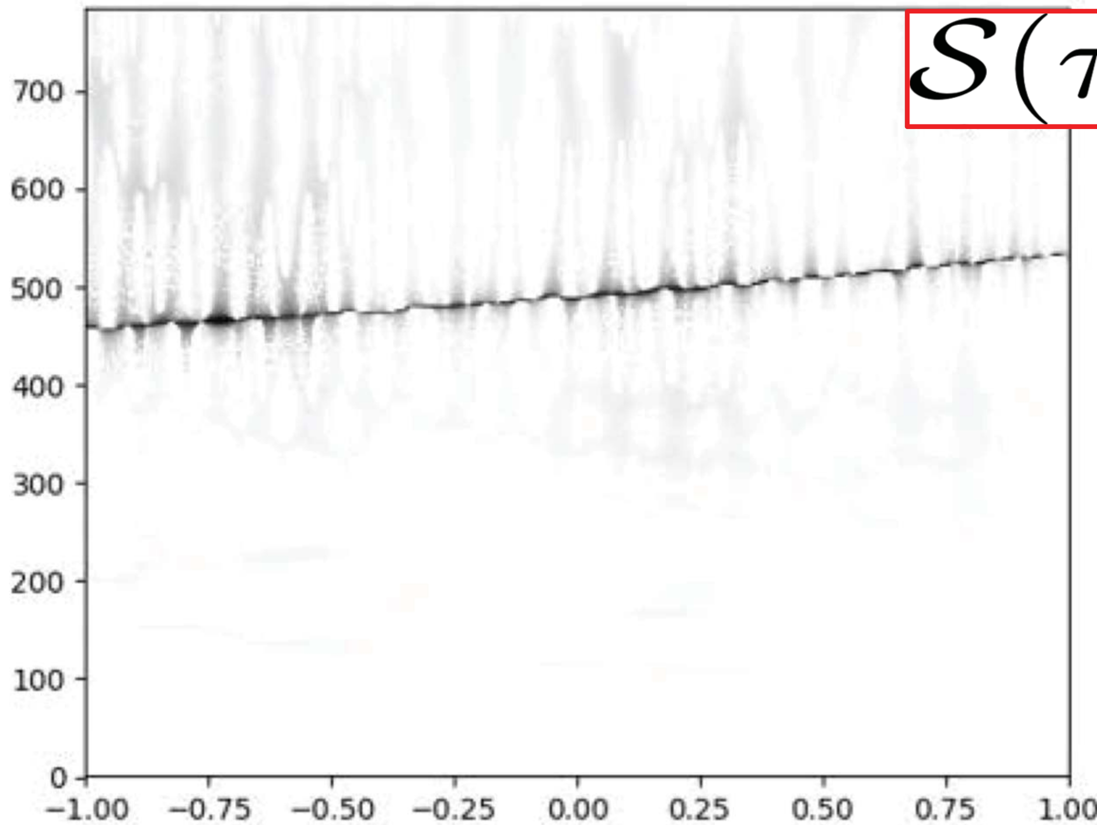
$$\omega_{2,e}(t) = \mathrm{argmax}_{(t,\omega)\in A_{\mathrm{low}}} \mathcal{S}(t, \omega)$$

$$\theta_{2,e}(t) = \theta_e(t, \omega_{2,e}(t))$$

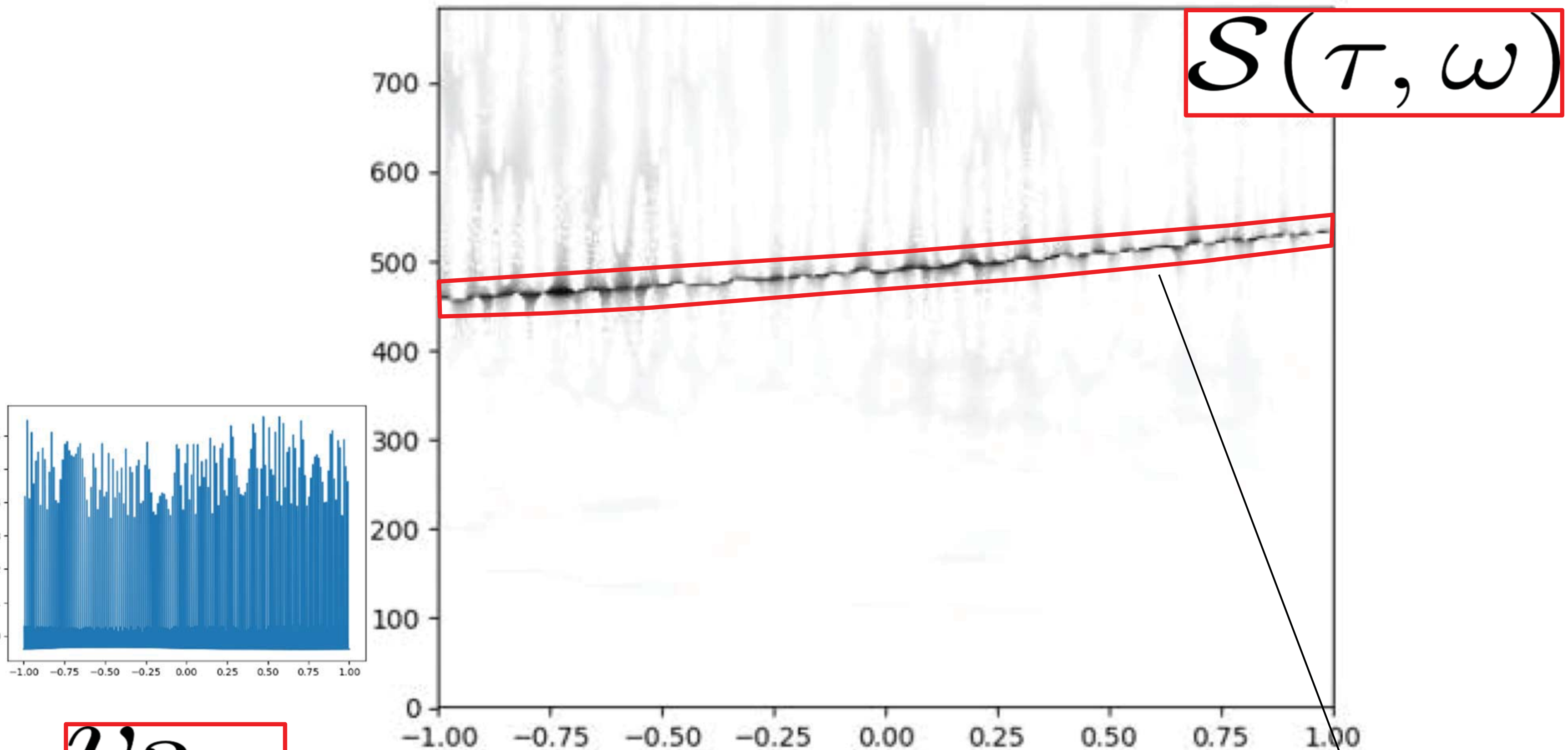
Max-squeezed Energy

$$v - v_{1,e} - v_{2,e}$$

$$S(\tau, \omega)$$



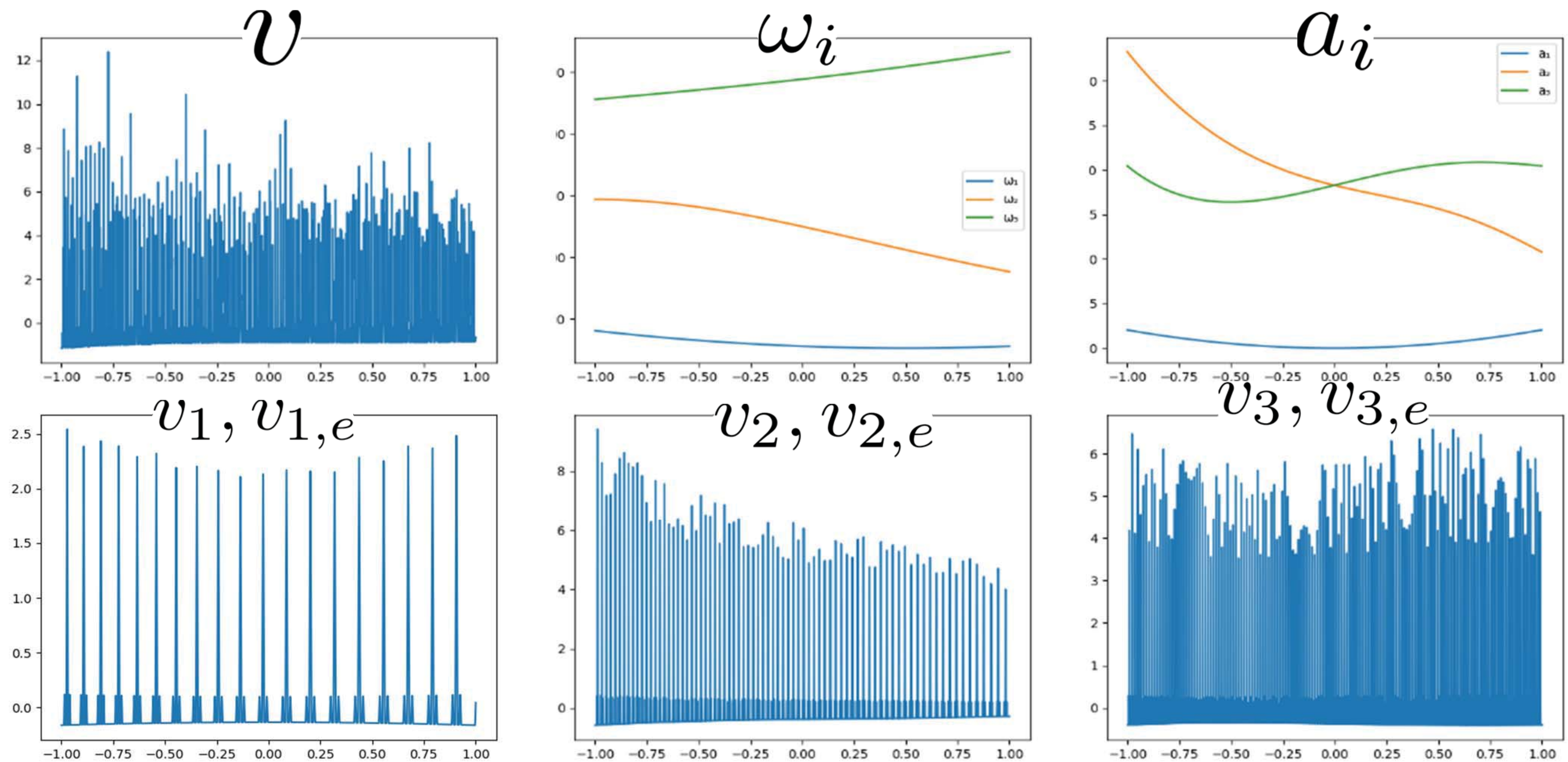
Max-squeezed Energy



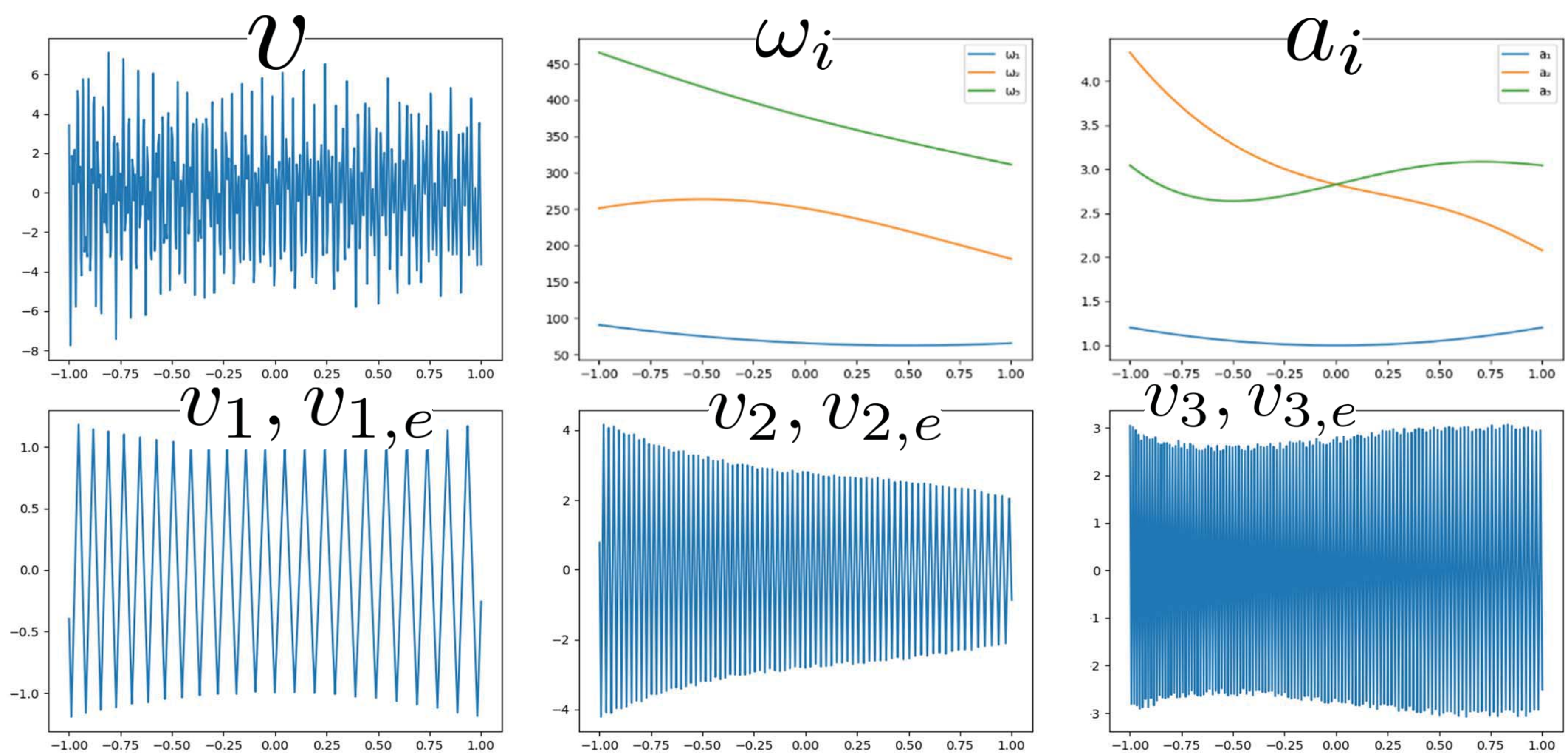
$$v_{3,e}$$

$$\omega_{3,e}(t) = \operatorname{argmax}_{(t,\omega)\in A_{\text{low}}} \mathcal{S}(t,\omega)$$

$$\theta_{3,e}(t) = \theta_e(t, \omega_{3,e}(t))$$



Mode	$\frac{\|v_{i,e}-v_i\|_{L^2}}{\|v_i\|_{L^2}}$	$\frac{\|v_{i,e}-v_i\|_{L^\infty}}{\|v_i\|_{L^\infty}}$	$\frac{\|a_{i,e}-a_i\|_{L^2}}{\|a_i\|_{L^2}}$	$\|\theta_{i,e}-\theta_i\|_{L^2}$
$i=1$	2.72×10^{-3}	1.04×10^{-2}	9.11×10^{-4}	3.25×10^{-4}
$i=2$	1.72×10^{-3}	4.30×10^{-3}	3.04×10^{-4}	2.31×10^{-4}
$i=3$	5.28×10^{-3}	2.89×10^{-2}	1.02×10^{-3}	8.91×10^{-4}



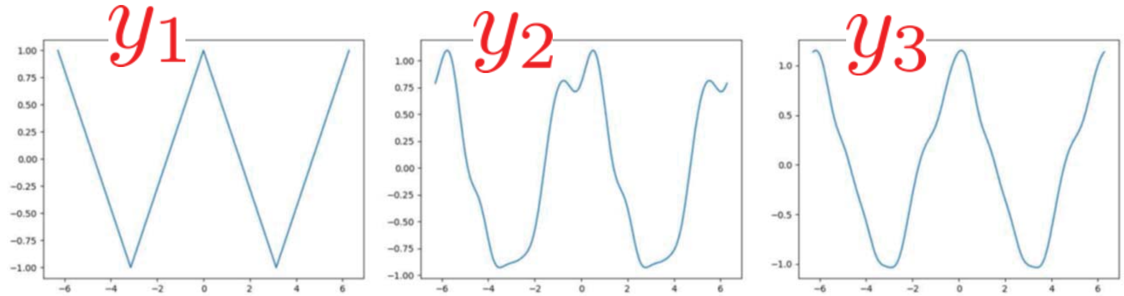
Mode	$\frac{\|v_{i,e}-v_i\|_{L^2}}{\|v_i\|_{L^2}}$	$\frac{\|v_{i,e}-v_i\|_{L^\infty}}{\|v_i\|_{L^\infty}}$	$\frac{\|a_{i,e}-a_i\|_{L^2}}{\|a_i\|_{L^2}}$	$\|\theta_{i,e}-\theta_i\|_{L^2}$
$i=1$	1.94×10^{-7}	3.65×10^{-7}	9.99×10^{-8}	1.50×10^{-7}
$i=2$	3.71×10^{-6}	2.52×10^{-6}	5.68×10^{-7}	3.56×10^{-6}
$i=3$	2.19×10^{-6}	3.89×10^{-6}	1.40×10^{-6}	1.51×10^{-6}

Ex: Amplitudes, phases waveforms are all unknown

Let $v_1, \dots, v_m$ be unknown s.t.

m: unknown

$$v_i(t) = a_i(t)\, y_i(\theta_i(t))$$

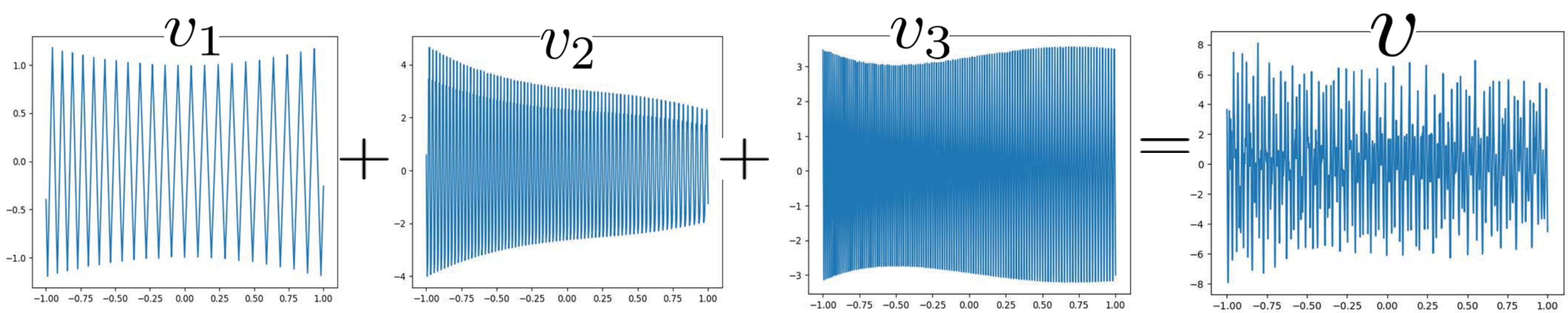


a_i: unknown, slowly varying

$\omega_i := \dot{\theta}_i$: unknown, slowly varying, positive, well separated

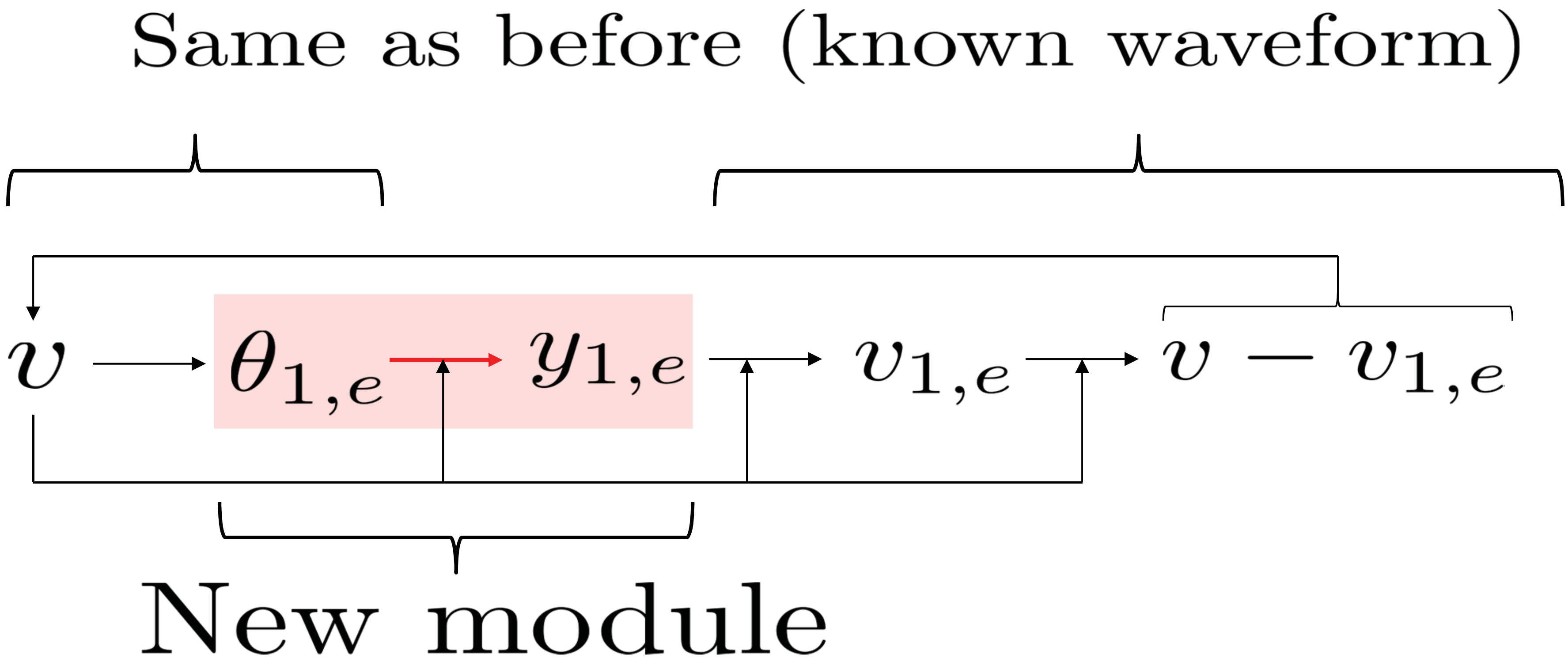
(and $(k\dot{\omega}_j)_{t\in[-,1,1]} \not\equiv (k'\dot{\omega}_{j'})_{t\in[-1,1]}$ for $k, k' \in \mathbb{N}$)

y_i: Unknown



Problem Given $v = \sum_{i=1}^{m} v_i$ recover $v_1, \dots, v_m$

Network


Same as before (known waveform)
v
$\theta_{1,e}$
$y_{1,e}$
$v_{1,e}$
$v - v_{1,e}$
New module

New module

$$\theta_{1,e} \longrightarrow y_{1,e}$$

$$v_\tau(t) := e^{-\left(\frac{\dot\theta_{1,e}(\tau)(t-\tau)}{\alpha}\right)^2} v(t)$$

$$\xi_\tau(t) = e^{-\left(\frac{\dot\theta_{1,e}(\tau)(t-\tau)}{\alpha}\right)^2} \zeta(t)$$

$$\zeta_\tau(t) = X_{1,c} \cos\left(\theta_{1,e}(t)\right) + \sum_{k=2}^{k_{\max}} \left(X_{k,c} \cos\left(k\theta_{1,e}(t)\right) + X_{k,s} \sin\left(k\theta_{1,e}(t)\right)\right)$$

$$Z_{k,j}(\tau, \theta_{1,e}, v) := \lim_{\sigma \downarrow 0} \mathbb{E}\left[X_{k,j} \middle| \xi_\tau + \xi_\sigma = v_\tau\right]$$

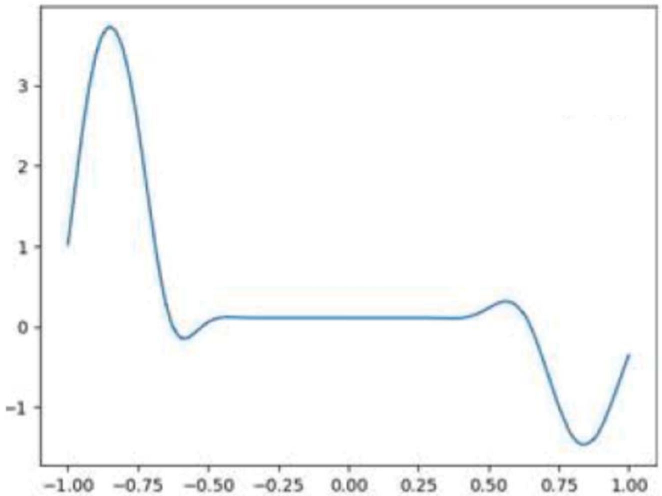
Estimator for the Fourier coefficients of y_1

$$c_{k,j}(\tau, \theta_{1,e}, v) := \frac{Z^y_{k,j}(\tau,\theta_{1,e},v)}{Z^y_{1,c}(\tau,\theta_{1,e},v)}$$

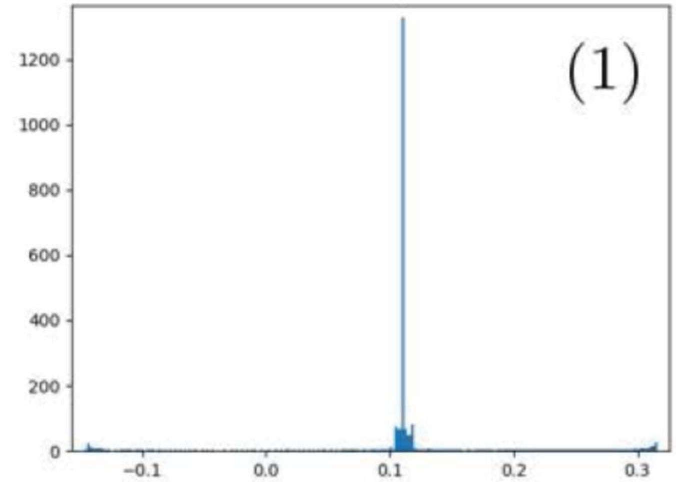
$$y_1(t) = \cos(t) + \sum_{k=2}^{k_{\max}} c_{1,(k,c)} \cos(kt) + c_{1,(k,s)} \sin(kt)$$

$\tau \rightarrow c_{3,c}(\tau, \theta_{1,e}, v)$

Histogram of the values of $\left(c_{3,c}(\tau, \theta_{1,e}, v)\right)_{\tau \in T}$



Non-robust

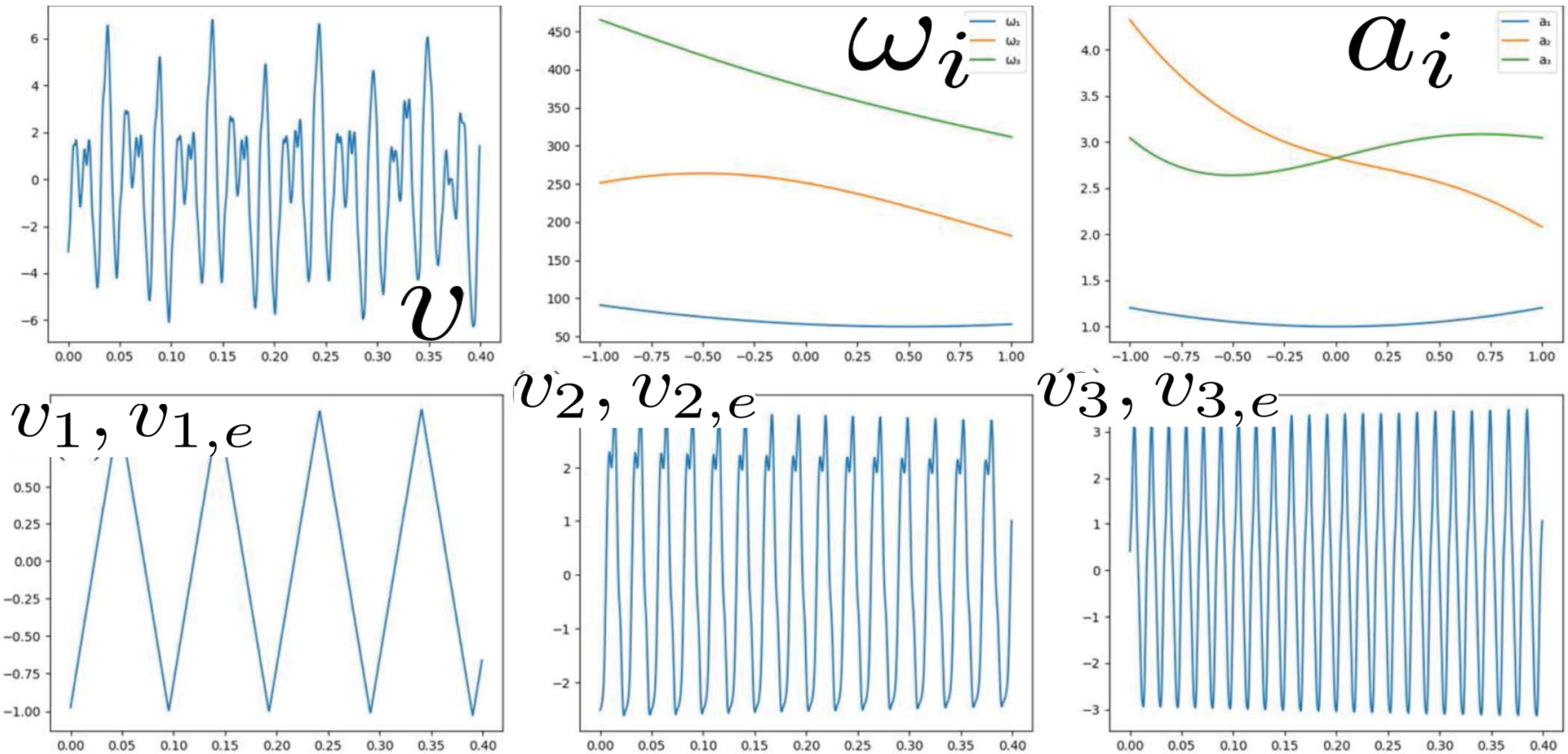


Discrete MLE

$c_{3,c}(\theta_{1,e}, v)$

Robust estimator of the Fourier coefficients of the waveform y_1

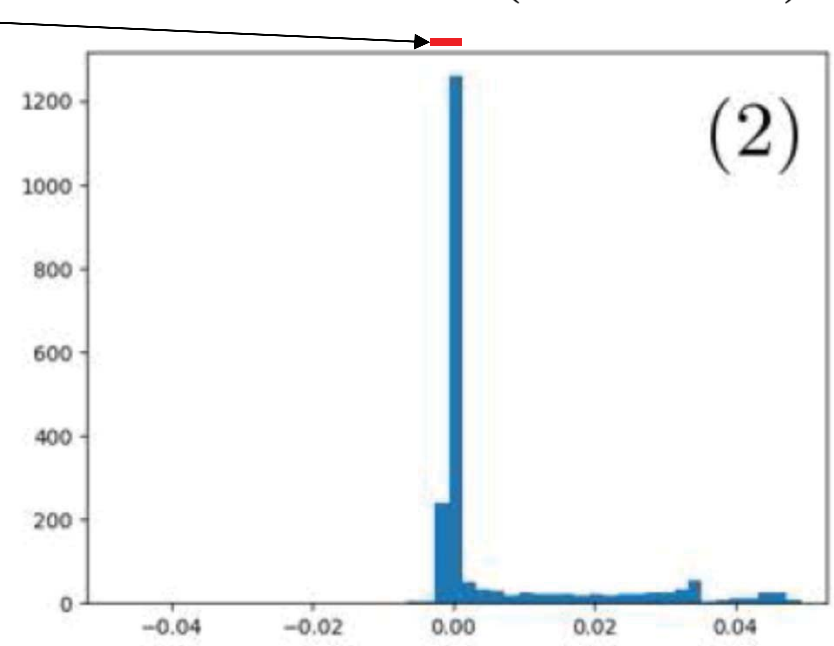
Results



Mode	$\frac{\Vert v_{i,e}-v_i\Vert_{L^2}}{\Vert v_i\Vert_{L^2}}$	$\frac{\Vert v_{i,e}-v_i\Vert_{L^\infty}}{\Vert v_i\Vert_{L^\infty}}$	$\frac{\Vert a_{i,e}-a_i\Vert_{L^2}}{\Vert a_i\Vert_{L^2}}$	$\Vert \theta_{i,e}-\theta_i\Vert_{L^2}$	$\frac{\Vert y_{i,e}-y_i\Vert_{L^2}}{\Vert y_i\Vert_{L^2}}$
$i=1$	6.59×10^{-3}	2.65×10^{-2}	1.52×10^{-5}	1.75×10^{-5}	6.65×10^{-3}
$i=2$	2.62×10^{-4}	5.61×10^{-4}	8.12×10^{-5}	1.25×10^{-4}	2.15×10^{-4}
$i=3$	6.55×10^{-4}	9.76×10^{-4}	3.99×10^{-4}	3.67×10^{-4}	3.43×10^{-4}

Discrete MLE estimator

N: Number of values of τ
I: Interval of $\mathbb{R}$
N_I: Number of values of τ contained in I.
$I_{\max}$: Interval fixed size δ $(= 0.002)$ maximizing N_I



$$T_I := \{\tau \in T | c_{k,j}(\tau, \theta, v) \in I\}$$

$$c_{k,j}(\theta, v) := \begin{cases} \frac{1}{N_{I_{\max}}} \sum_{\tau \in T_{I_{\max}}} c_{k,j}(\tau, \theta, v) & \text{for } \frac{N_{I_{\max}}}{N} \geq 0.05 \\ 0 & \text{for } \frac{N_{I_{\max}}}{N} < 0.05 \end{cases}$$

$$\chi_{\tau,\omega,\theta}(t) := \left(\frac{2}{\alpha^2\pi^3}\right)^{\frac{1}{4}} \omega^{\frac{1-\beta}{2}} \cos(\omega(t-\tau)+\theta) e^{-\frac{\omega^2(t-\tau)^2}{\alpha^2}}, \qquad t \in \mathbb{R},$$

$$K_{\tau,\omega,\theta}(s,t) := \chi_{\tau,\omega,\theta}(s)\chi_{\tau,\omega,\theta}(t), \quad s,t \in \mathbb{R},$$

$$K_\beta(s,t) = \int_{-\pi}^{\pi}\int_{\mathbb{R}_+}\int_{\mathbb{R}} K_{\tau,\omega,\theta}(s,t)\,d\tau\,d\omega\,d\theta, \quad s,t \in \mathbb{R},$$

Theorem

Let

$$(\mathcal{K}_\beta f)(s) := \frac{1}{H(\beta)}\int_{\mathbb{R}} K_\beta(s,t) f(t)\,dt$$

$$H(\beta) := 2^{\beta-1}\sqrt{\pi}(\sqrt{2}\alpha)^{1-\beta}\Gamma(\frac{\beta}{2}))e^{-\frac{\alpha^2}{2}}\,{}_1F_1\left(\frac{\beta}{2},\frac{1}{2};\frac{\alpha^2}{2}\right),$$

$${}_1F_1(\alpha,\gamma;z) = 1 + \frac{\alpha}{\gamma}\frac{z}{1!} + \frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{z^2}{2!} + \frac{\alpha(\alpha+1)(\alpha+2)}{\gamma(\gamma+1)(\gamma+2)}\frac{z^3}{3!} + \dots,$$

Confluent hypergeometric function

Then

$$\mathcal{K}_{\beta_1}\mathcal{K}_{\beta_2} f = \mathcal{K}_{\beta_1+\beta_2} f, \quad f \in \mathcal{S}, \quad \beta_1,\beta_2 > 0, \beta_1+\beta_2 < 1,$$

$$\lim_{\beta\to 0}(\mathcal{K}_\beta f)(x) = f(x), \quad x \in \mathbb{R}, \quad f \in \mathcal{S}$$

$\mathcal{S} := \{f \in C^\infty(\mathbb{R}) : \sup_{x\in\mathbb{R}} |x^{m_1} D^{m_2} f(x)| < \infty, m_1, m_2 \in \mathbf{N}\}$

$$y(t) := \sum_{-N}^{N} c_n e^{int}$$

$$\chi_{\tau,\omega,\theta}(t) := \omega^{\frac{1-\beta}{2}} y\big(\omega(t-\tau)+\theta\big) e^{-\frac{\omega^2}{\alpha^2}|t-\tau|^2}$$

$$K_\beta(s,t) := \Re \int_{-\pi}^{\pi} \int_{\mathbb{R}_+} \int_{\mathbb{R}} \chi_{\tau,\omega,\theta}(s)\chi^*_{\tau,\omega,\theta}(t) d\tau d\omega d\theta \,.$$

Theorem

$$K_\beta(s,t) = 2\pi|s-t|^{\beta-1} \sum_{n=-N}^{N} a_n(s,t)|c_n|^2$$

$$a_n(s,t) = \frac{\alpha\sqrt{\pi}}{2\sqrt{2}}(\sqrt{2}\alpha)^{1-\beta}\Gamma(\frac{1-\beta}{2})e^{-\frac{|n|\alpha^2}{2}} {}_1F_1\Big(\frac{\beta}{2};\frac{1}{2};\frac{|n|\alpha^2}{2}\Big)\,.$$

$$K_0(s,t) = \alpha^2\pi^2|s-t|^{-1}\|y\|^2 \,.$$

$$\|y\|^2 := \sum_{n=-N}^{N} e^{-\frac{|n|\alpha^2}{2}}|c_n|^2$$

Thank you