

Data-Derived Wavelets for Nonlinear Wave Packet Detection

C.W. Curtis, J. Gilles

Problem Motivation

So, sparing one the usual details, at the end of the day we wish to solve the following nonlinear system of partial differential equations

$$\eta_t = (\mathcal{G}_0 + \epsilon \mathcal{G}_1 + \epsilon^2 \mathcal{G}_2) q, \quad (1)$$

$$q_t = -\eta + \tilde{\sigma} \eta_{xx} - \frac{\epsilon}{2} \left(-(\mathcal{G}_0 q)^2 - 2\epsilon \mathcal{G}_0 q \mathcal{G}_1 q + q_x^2 \right) - \epsilon^2 \left(\frac{3}{2} \tilde{\sigma} \eta_x^2 \eta_{xx} - \eta_x q_x \mathcal{G}_0 q \right), \quad (2)$$

where

$$\begin{aligned} \mathcal{G}_0 q &= -\mathcal{H} \partial_x q \\ \mathcal{G}_1 q &= -\mathcal{G}_0 (\eta \mathcal{G}_0 q) - \partial_x (\eta \partial_x q) \\ \mathcal{G}_2 q &= -\mathcal{G}_0 (\eta \mathcal{G}_1 q) + \frac{1}{2} \partial_x^2 (\eta^2 \mathcal{G}_0 q) + \frac{1}{2} \partial_x^2 \mathcal{H} (\eta^2 q_x). \end{aligned}$$

The nonlocality in the full equations comes from the Hilbert transform, \mathcal{H} , defined by

$$\mathcal{H}f = \frac{1}{2\pi} \int_{\mathbb{R}} dk e^{ikx} i \operatorname{sgn}(k) \hat{f}(k),$$

where $\hat{f}(k)$ is the Fourier transform of $f(x)$ defined by

$$\hat{f}(k) = \int_{\mathbb{R}} dx e^{-ikx} f(x).$$

In order to reduce the complexity of this model, we traditionally assume that

$$\begin{aligned} \eta(x, t) &= \eta_1(\xi, \tau) e^{i\theta} + \eta_1^*(\xi, \tau) e^{-i\theta} + \epsilon \left(\eta_2(\xi, \tau) e^{2i\theta} + \eta_2^*(\xi, \tau) e^{-2i\theta} \right) + \dots \\ q(x, t) &= q_1(\xi, \tau) e^{i\theta} + q_1^*(\xi, \tau) e^{-i\theta} + \epsilon \left(q_2(\xi, \tau) e^{2i\theta} + q_2^*(\xi, \tau) e^{-2i\theta} \right) + \dots \end{aligned}$$

where

$$\theta = k_0 x + \Omega t, \quad \xi = \epsilon(x + c_g t), \quad \tau = \epsilon^2 t.$$

Following the usual approximation arguments, we then get

$$\Omega_{\pm}(k_0) = \pm \sqrt{|k_0|(1 + \tilde{\sigma}k_0^2)},$$

and at $\mathcal{O}(\epsilon)$ we find that

$$c_g = \frac{1 + 3\tilde{\sigma}k_0^2}{2s\Omega},$$

where $s = \text{sgn}(k_0)$. Ultimately, we are able to describe the slow evolution of the envelope $\eta_1(\xi, \tau)$ via the Nonlinear-Schrödinger Equation (NLSE) given by

$$i\partial_{\tau}\eta_1 + \alpha_{nl}|\eta_1|^2\eta_1 + \alpha_d\partial_{\xi}^2\eta_1 = 0,$$

where

$$\begin{aligned}\alpha_d(k_0) &= \frac{(c_g^2 - 3|k_0|\tilde{\sigma})}{2\Omega}, \\ \alpha_{nl}(k_0) &= \frac{k_0 (sk_0^3 (8 + \tilde{\sigma}k_0^2 + 2(\tilde{\sigma}k_0^2)^2))}{(2s\Omega) (4\Omega^2 - s(2k_0(1 + 4\tilde{\sigma}k_0^2)))},\end{aligned}$$

One of the more important features of the NLS equation is that in the focusing case, i.e. $\alpha_d/\alpha_{nl} > 0$, it supports solitary traveling wave solutions, i.e. solitons, given by the formula

$$\eta_1(\xi, \tau) = \beta \sqrt{\frac{2\alpha_d}{\alpha_{nl}}} \text{sech}(\beta(\xi + c\tau)) e^{i(\varphi\tau - c(\xi + c\tau)/2\alpha_d)}, \quad c^2 = 4\alpha_d(\varphi - \alpha_d\beta^2).$$

As we know from Inverse-Scattering Theory (IST), the exponentially localized behavior of solitons typifies the response of the focusing NLS equation to rapidly decaying initial conditions.

A Data-Driven Approach to Wavelet Design

Given the form of our asymptotic ansatz, we are obliged to specify initial conditions $\eta_0(x)$ and $q_0(x)$ for Equations (1) and (2) consistent with using NLS as an approximating model. Thus, we in fact specify initial conditions for the NLS equation so that

$$\eta_0(x) = \eta_1(\epsilon x, 0)e^{ik_0x} + \eta_1^*(\epsilon x, 0)e^{-ik_0x}$$

and likewise, using the lead-order approximation

$$\partial_x q_0(x) = -s\Omega\eta_0(x),$$

we also see we should choose $\eta_1(x, 0)$ so that

$$\int_{\mathbb{R}} dx \eta_1(x, 0)e^{ik_0x/\epsilon} = 0.$$

Note, while these assumptions on initial conditions are somewhat restrictive, they reflect the particulars of the approximation which led us to the NLS equation. Other approximation schemes certainly exist which relax these conditions.

Starting from this initial data then, we are able either numerically, or even analytically in some cases, to generate an affiliated NLS flow given by $\eta_1(\xi, \tau)$, $0 \leq \tau \leq T$. We then discretely sample this signal in time, say at time steps $\tau_j = j\delta\tau$, $j = 0, \dots, N_T - 1$, $T = N_T\delta t$, thereby generating the N_T spatial functions $\eta_{1j}(\xi) = \eta_1(\xi, \tau_j)$. This motivates defining the *empirical measure* $\mu(A)$ where $A \subset L_2(\mathbb{R})$ is a Borel set (defined using the norm $\|\cdot\|_2$ on $L_2(\mathbb{R})$) and

$$\mu(A) = \lim_{N_T \rightarrow \infty} \frac{1}{N_T} \sum_{j=0}^{N_T-1} \tilde{I}_A(\eta_{1j}(\cdot)),$$

where $\tilde{I}_A(\cdot)$ is the indicator function. Letting the temporal flow operator S_τ be defined such that

$$\eta_1(\xi, \tau) = S_\tau \eta_1(\xi, 0),$$

with semigroup property $S_{\tau_1} S_{\tau_2} = S_{\tau_1 + \tau_2}$, and assuming $S_{-\tau}$ is well defined, we see that

$$\tilde{I}_A(\eta_{1j}(\cdot)) = \tilde{I}_A(S_{\tau_j} \eta_1(\cdot, 0)) = \tilde{I}_{S_{-\tau_j} A}(\eta_1(\cdot, 0)).$$

Therefore, we have that

$$\mu(S_{-\tau_k} A) = \lim_{N_T \rightarrow \infty} \frac{1}{N_T} \sum_{j=0}^{N_T-1} \tilde{I}_A(\eta_{1(j+k)}(\cdot)) = \mu(A),$$

so that the measure is invariant with respect to the discretized sampling times. Thus, we have a discretely invariant probability measure.

In turn then, we can define *empirical averages* of measureable functions

$$F : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$$

to be

$$\begin{aligned} \langle F \rangle &= \int_{L_2(\mathbb{R})} F(f) \mu(df) \\ &= \frac{1}{T} \int_0^T F(\eta_1(\xi, t)) dt \end{aligned}$$

Note, one should probably restrict to dense subspaces of $L_2(\mathbb{R})$, such as the Sobolev space $H^s(\mathbb{R})$, $s > 1/2$, which are algebras and thus supportive of a relatively wide class of maps and thus measurable functions in the present

context. From a practical point-of-view, or as practical as “measures over Hilbert spaces” can ever be, we can approximate the above integral with the usual ‘simple function’ formalism so that

$$\begin{aligned}\langle F \rangle &\sim \sum_{l=1}^{N_s} F_l \mu(A_l), \quad F_l(\xi) \in L_2(\mathbb{R}) \\ &\sim \frac{1}{N_T} \sum_{j=0}^{N_T-1} \sum_{l=1}^{N_s} F_l \tilde{I}_{A_l}(\eta_{1j}) \\ &\sim \frac{1}{N_T} \sum_{j=0}^{N_T-1} F(\eta_{1j}(\xi))\end{aligned}$$

which shows us that our formalism aligns with the usual frequency based determination of an “average”.

With this formalism in place, we define the autocorrelation operator K so that for $\varphi \in L_2(\mathbb{R})$ we have

$$\begin{aligned}K\varphi &= \left\langle \eta_1(\xi, \cdot), \left(\varphi(\xi'), \eta_1(\xi', \cdot) \right) \right\rangle \\ &= \int_{\mathbb{R}} \left\langle \eta_1(\xi, \cdot) \eta_1^*(\xi', \cdot) \right\rangle \varphi(\xi') d\xi',\end{aligned}$$

where (\cdot, \cdot) denotes the usual L_2 inner product and tacit Fubini-Tonelli argument has been used to exchange the order of averaging and integration. This move is justified if

$$\left\langle |\eta_1(\xi, \cdot)|^2 \right\rangle < \infty, \quad \left\langle |\eta_1(\xi, \cdot)|^2 \right\rangle \in L_1(\mathbb{R}).$$

Letting the kernel of K , say $k(\xi, \xi')$ be given by

$$k(\xi, \xi') = \left\langle \eta_1(\xi, \cdot) \eta_1^*(\xi', \cdot) \right\rangle,$$

existing results show that if $k(\xi, \xi) \in L_1(\mathbb{R})$, then this essentially ensures that K is a Hilbert-Schmidt operator, and therefore has a norm-convergent representation in terms of its eigenfunctions in the form

$$K = \sum_{l=1}^{\infty} \mu_l \mathbb{P}_l, \quad \mathbb{P}_l f = (f, \varphi_l) \varphi_l, \quad \mu_1 \geq \mu_2 \geq \dots \geq 0.$$

The orthonormal modes φ_l are called the *principal components* of the affiliated data set $\{\eta_{1j}(\xi)\}_{j=0}^{N_T-1}$.

For our future purposes, we need some estimates on the far-field behavior and regularity of $\varphi_1(\xi)$ based on the far-field behavior of the functions $\eta_{1j}(\xi)$.

Letting $\hat{\varphi}(k)$ denote the Fourier transform, we formally see that the principal components must satisfy the integral equation

$$\lambda \hat{\varphi}(k) = \frac{1}{2\pi} \int_{\mathbb{R}} d\tilde{k} \left\langle \hat{\eta}_1(k, \cdot) \hat{\eta}_1^*(\tilde{k}, \cdot) \right\rangle \hat{\varphi}(\tilde{k})$$

If we treat the infinite limit of the averaging operation as an integral in time, we are then motivated to write

$$\frac{1}{2\pi} \int_{\mathbb{R}} \left\langle \hat{\eta}_1(k, \cdot) \hat{\eta}_1^*(\tilde{k}, \cdot) \right\rangle \hat{\varphi}(\tilde{k}) d\tilde{k} = \frac{1}{T} \int_0^T \hat{\eta}_1(k, t) m(t) dt,$$

where

$$m(t) = (\varphi(\cdot), \eta_1(\cdot, t)).$$

When we then compute the s -Sobolev norm of $K\varphi$, say $\|K\varphi\|_s$, we find using Minkowski's inequality, the Cauchy-Schwarz inequality, and basic results about Sobolev spaces that

$$\begin{aligned} \|K\varphi\|_s &= \frac{1}{T\sqrt{2\pi}} \left(\int_{\mathbb{R}} (1+k^2)^s \left| \int_0^T \hat{\eta}_1(k, t) m(t) dt \right|^2 dk \right)^{1/2} \\ &\leq \frac{1}{T} \int_0^T |m(t)| \|\eta_1(\cdot, t)\|_s dt \\ &\leq \frac{\|\varphi\|_2}{T} \int_0^T \|\eta_1(\cdot, t)\|_s^2 dt. \end{aligned}$$

Thus, we expect that the operator K is regularizing if

$$\frac{1}{T} \int_0^T \|\eta_1(\cdot, t)\|_s^2 dt < \infty, \quad (3)$$

which is to say this condition guarantees that the regularity of the data is passed on to the principal components. Moreover, letting φ be an eigenvector of K leads us to the eigenvalue bound

$$|\lambda| \leq \frac{1}{T} \int_0^T \|\eta_1(\cdot, t)\|_2^2 dt.$$

We might need an $L_1(\mathbb{R})$ estimate on our eigenfunctions. Thus, we estimate

$$\begin{aligned} \|K\varphi\|_1 &\leq \|\varphi\|_2 \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \left| \left\langle \eta_1(\xi, \cdot) \eta^*(\xi', \cdot) \right\rangle \right|^2 d\xi' \right)^{1/2} d\xi \\ &\leq \frac{\|\varphi\|_2}{T} \int_{\mathbb{R}} \int_0^T |\eta_1(\xi, t)| \|\eta_1(\cdot, t)\|_2 dt d\xi \\ &\leq \frac{\|\varphi\|_2}{T} \int_0^T \|\eta_1(\cdot, t)\|_1 \|\eta_1(\cdot, t)\|_2 dt, \end{aligned}$$

which shows us that data in $L_1(\mathbb{R})$ gives us eigenfunctions in the same space.

Building Frames from Principal Components

Assuming our data regularity requirement, i.e. Equation (3) holds, we now determine conditions which ensures that an eigenfunction of K can be used to build a frame for a mutli-resolution analysis (MRA). We remind ourselves though that the primary principal component vector $\varphi(\xi)$ in fact corresponds to a slowly varying function $\varphi(\epsilon x)$. Letting $\epsilon = 2^{-J}$ for the ease of computation and exposition, we see by choosing the foundation for the scaling function of an MRA to be $\varphi(x)$ that we are building a discrete wavelet transformation (DWT) whose coarsest scales correspond to the leading-order approximations used to derive the NLS approximation. Thus we are building an MRA specifically designed to identify fast scales relative to the slow amplitude approximation and all intermediary scales in between. This makes a potential DWT into a highly effective band-pass filter.

Thus, we need to show there exist positive constants A and B such that

$$A \leq \sum_m |\hat{\varphi}(k + 2\pi m)|^2 \leq B.$$

Taking then the definition of the s^{th} -Sobolev norm so that

$$\begin{aligned} \|\varphi\|_s^2 &= \frac{1}{2\pi} \int_{\mathbb{R}} (1 + k^2)^s |\hat{\varphi}(k)|^2 dk \\ &= \frac{1}{2\pi} \sum_m \int_0^{2\pi} (1 + (k + 2\pi m)^2)^s |\hat{\varphi}(k + 2\pi m)|^2 dk, \end{aligned}$$

which given the positivity of all terms involved gives us that the function

$$\Gamma_s(k) = \sum_m (1 + (k + 2\pi m)^2)^s |\hat{\varphi}(k + 2\pi m)|^2$$

is in $L_1([0, 2\pi])$ if $\varphi \in H^s(\mathbb{R})$. Moreover, we see that the Fourier-average of the 2π -periodic function $\Gamma_s(k)$ is $\|\varphi\|_s^2$. The L_1 convergence ensures that $\Gamma_s(k)$ exists except possibly on a set of Lebesgue-measure zero. Thus, for almost all k , we have that for a chosen value of $\tilde{\epsilon} > 0$ there is some integer $M > 0$ such that

$$|\hat{\varphi}(k + 2\pi m)|^2 \leq \frac{\tilde{\epsilon}}{(1 + (k + 2\pi m)^2)^s}, \quad |m| > M.$$

We therefore have that so long as $s > 1/2$, then the function $\Gamma(k) = \Gamma_0(k)$ is defined by a sum which converges uniformly on $[0, 2\pi]$, and thus we have the necessary frame bound.

From Frames to MRAs

In order to generate a fast discrete wavelet decomposition of a signal, we let the primary principal component vector $\varphi(x)$ play the role of scaling

function, thereby defining the low-pass filter $\hat{h}(k)$, which is a 2π periodic function satisfying the equation

$$\hat{h}(k) = \sqrt{2} \frac{\hat{\varphi}(2k)}{\hat{\varphi}(k)} \frac{\Gamma(k)}{\Gamma(2k)}$$

In turn, the essential feature we need to store associated with filter is its corresponding l_2 representation, i.e. we need the coefficients h_n such that

$$\hat{h}(k) = \sum_n h_n e^{-ink},$$

or

$$h_n = \frac{1}{2\pi} \int_0^{2\pi} e^{ink} \hat{h}(k) dk.$$

This likewise specifies the high-pass filter $\hat{g}(k)$ whose l_2 representation is given by

$$g_n = (-1)^{n-1} h_{-n+1}^*.$$

Since our data on a given signal, say $f(x)$, and the principal component $\varphi(x)$ come from data sampled on an interval $[-L, L]$, we work with interpolated approximations to both of the form

$$f(x) = \sum_{j=0}^{N_s-1} f_j S_\delta(x - x_j), \quad \varphi(x) = \sum_{j=0}^{N_s-1} \varphi_j S_\delta(x - x_j),$$

where we sample at the the equispaced points $x_j = -L + j\delta$, $\delta = 2L/N_s$, and $S_\delta(x)$ is given by

$$S_\delta(x) = \text{sinc}\left(\frac{\pi}{\delta}x\right).$$

This in turn gives us the approximations

$$\hat{f}(k) = \delta B\left(\frac{\delta}{\pi}k\right) \sum_{j=0}^{N_s-1} f_j e^{-ikx_j}, \quad \hat{\varphi}(k) = \delta B\left(\frac{\delta}{\pi}k\right) \sum_{j=0}^{N_s-1} \varphi_j e^{-ikx_j},$$

where $B(k)$ is given by

$$B(k) = \begin{cases} 1 & |k| \leq 1, \\ 0 & |k| > 1. \end{cases}$$

If we fix $-\pi \leq k < \pi$, then we can readily show that,

$$\sum_{m=-M_2}^{M_1} e^{-im\theta_n} = e^{i(M_2-M_1)\theta_n/2} \frac{\sin((M_1+M_2+1)\theta_n/2)}{\sin(\theta_n/2)}, \quad \theta_n = 2\pi\delta n$$

then letting $1/(2\delta) = M + \tilde{\theta}$, $M \in \mathbb{Z}^+$, we see

$$\Gamma^2(k) = \sum_{j,l} \varphi_j \varphi_l^* S(j-l; k)$$

where for $\tilde{k} = k/(2\pi)$, and letting $\tilde{s} = \text{sgn}(\tilde{k})$, we have

$$S(n; k) = \frac{\delta^2 e^{i\delta kn}}{\sin(\pi\delta n)} \begin{cases} e^{-i\pi\tilde{s}\delta n} \sin(2\pi\delta Mn) & \tilde{\theta} < |\tilde{k}|, \tilde{\theta} + |\tilde{k}| < 1 \\ \sin(2\pi\delta(M+1/2)n) & \tilde{\theta} \geq |\tilde{k}|, \tilde{\theta} + |\tilde{k}| < 1 \\ e^{-2\pi i\tilde{s}\delta n} \sin(2\pi\delta(M+1/2)n) & \tilde{\theta} < |\tilde{k}|, \tilde{\theta} + |\tilde{k}| \geq 1 \\ e^{-i\pi\tilde{s}\delta n} \sin(2\pi\delta(M+1)n) & \tilde{\theta} \geq |\tilde{k}|, \tilde{\theta} + |\tilde{k}| \geq 1 \end{cases}$$

With $\Gamma(k)$ computed, we can find the rescaled wavelet function at discretized nodes x_m is given by

$$\tilde{\varphi}_m = \sum_{l=0}^{N_s-1} \varphi_l K(m-l), \quad K(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\Gamma\left(\frac{k}{\delta}\right)} e^{ink} dk.$$

We can then find the initial approximation coefficients via the formula

$$\begin{aligned} a_n^{(0)} &= \int_{\mathbb{R}} f(x) \tilde{\varphi}^*(x-n) dx, \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(k) \hat{\tilde{\varphi}}^*(k) e^{ikn} dk, \\ &= \delta \sum_{j,l} f_j \tilde{\varphi}_l^* \tilde{K}(j-l; n), \quad \tilde{K}(m; n) = \frac{(-1)^m}{1 - \delta m/n} \text{sinc}\left(\frac{\pi n}{\delta}\right). \end{aligned}$$

Results