Supplemental Information for "Microresonator Soliton Crystal Pattern Fitting Algorithm"

Van Thuy Hoang[1] and Jonathan M. Silver[1,*]

[1] *National Physical Laboratory, Hampton Road, Teddington TW11 0LW, UK*

(Dated: November 27, 2023)

I. THEORETICAL BACKGROUND

In the time domain, the complex electric field amplitude of a microcomb with angular pump frequency ω_{pump} and spacing ω_{rep} takes the form

$$E(t) = \sum_n \tilde{F}_n \exp(i(\omega_{\mathrm{pump}} + n\omega_{\mathrm{rep}})t) \tag{1}$$

where $\tilde{F}_n$ is the complex amplitude of comb line $n \in \mathbb{Z}$ relative to the pump. This can be expressed as

$$E(t) = \exp(i\omega_{\mathrm{pump}}t)F(\tau) \tag{2}$$

where the dimensionless time $\tau = \omega_{\mathrm{rep}}t/(2\pi)$ and

$$F(\tau) = \sum_{n=-\infty}^{\infty} \tilde{F}_n \exp(2i\pi n\tau). \tag{3}$$

$F(\tau)$ is periodic with period 1, and

$$\tilde{F}_n = \int_0^1 F(\tau)\exp(-2i\pi n\tau)d\tau. \tag{4}$$

A soliton crystal comb consists of a repeating series of hyperbolic-secant-shaped solitons with the same duration, amplitude and phase with respect to the pump, plus a monochromatic pump background. Mathematically, this can be expressed as

$$F(\tau) = \frac{A\beta}{\pi}\sum_{k=-\infty}^{\infty}\sum_{r=0}^{n_{\mathrm{s}}-1} \operatorname{sech}(\beta(\tau - \tau_r - k)) + B \tag{5}$$

where n_{s} is the number of solitons in each repeat, A and B are in general complex amplitudes, β is the inverse dimensionless pulse duration and $\tau_r \in [0, 1)$ is the dimensionless time of the centre of the r^{th} soliton within each repeat. We shall choose the carrier phase such that A is real and positive. The reason for the prefactor β/π will become clear when we take the Fourier transform. Eq. (5) is equivalent to

$$F(\tau) = A(P * S)(\tau) + B \tag{6}$$

where the pattern function

$$P(\tau) = \sum_{k=-\infty}^{\infty}\sum_{r=0}^{n_{\mathrm{s}}-1} \delta(\tau - \tau_r - k), \tag{7}$$

the soliton function

$$S(\tau) = \frac{\beta}{\pi}\sum_{k=-\infty}^{\infty} \operatorname{sech}(\beta(\tau - k)), \tag{8}$$

and the convolution operator $*$ is defined by

$$(F * G)(\tau) := \int_0^1 F(\tau')G((\tau - \tau') \bmod 1)d\tau'. \tag{9}$$

Importantly, the solitons are well separated from each other with little or no overlap. In practical terms, the patterns of the combs we fitted had minimum spacings in τ of around $8/\beta$.

Using the tilde notation to denote the Fourier series of any periodic function with period 1 as defined for F in Eq. (4), we have

$$\tilde{F}_n = A\tilde{P}_n\tilde{S}_n + B\delta_{n0} \tag{10}$$

with

$$\tilde{P}_n = \sum_{r=0}^{n_{\mathrm{s}}-1} \exp(-2i\pi n\tau_r) \tag{11}$$

and

$$\tilde{S}_n = \operatorname{sech}\left(\frac{n\pi^2}{\beta}\right). \tag{12}$$

Note that $\tilde{P}_0 = n_{\mathrm{s}}$ and $\tilde{S}_0 = 1$. We thus have

$$|\tilde{F}_n| = A|\tilde{P}_n||\tilde{S}_n|(1 + (A_0 - 1)\delta_{n0}) \tag{13}$$

where

$$A_0 = \frac{|An_{\mathrm{s}} + B|}{An_{\mathrm{s}}} \tag{14}$$

is a scaling factor for the pump comb line. It will only be possible to determine A_0 using this algorithm and not B, which would require additional information about the in-coupling coefficient of the resonator or the pump detuning from resonance.

For computational purposes, we must discretise time, allowing us to use the Fast Fourier Transform (FFT) algorithm. To do this, we restrict n to the interval $[-N/2, N/2)$ for some positive integer N that is a few times the number of observable comb lines. N is normally chosen to be a power of 2 to maximise the efficiency of the FFT. For around 300 observable comb lines

* jonathan.silver@npl.co.uk

covering a dynamic range in power of around 70 dB, we find that the minimum value of N that works well with our algorithm is $2^{10} = 1024$.

To use F as an example, we define

$$F_m = \frac{1}{N} F\left(\frac{m}{N}\right) \tag{15}$$

for integer $m \in [0, N)$. We thus have

$$F_m = \frac{1}{N} \sum_{n=-\lfloor \frac{N}{2} \rfloor}^{\lceil \frac{N}{2} \rceil - 1} \tilde{F}_n \exp\left(\frac{2i\pi nm}{N}\right) \tag{16}$$

and

$$\tilde{F}_n = \sum_{m=0}^{N-1} F_m \exp\left(\frac{-2i\pi nm}{N}\right), \tag{17}$$

in other words $\tilde{F}$ is the discrete Fourier transform (DFT) of F. The convolution operator in the time domain takes the form

$$(F * G)_m := \sum_{m'} F_{m'} G_{(m-m') \bmod N}. \tag{18}$$

In the case that $\tau_r = m_r/N,\ m_r \in \mathbb{Z}$ for all r, Eq. (7) translates unambiguously to

$$P_m = \sum_{r=0}^{n_\mathrm{s}} \delta_{m-m_r}, \tag{19}$$

which exactly reproduces Eq. (11). However, we will need to represent continuous values of τ_r as the algorithm involves minimising a quantity with respect to continuous variations of the soliton positions. The way we do this is to define

$$\delta_m(\tau) := \begin{cases} 1 - \{N\tau\}, & m = \lfloor N\tau \rfloor \\ \{N\tau\}, & m = \lfloor N\tau + 1 \rfloor \bmod N \\ 0, & \text{otherwise} \end{cases} \tag{20}$$

where $\{x\}$ denotes the fractional part $x - \lfloor x \rfloor = x \bmod 1$, and let

$$P_m = \sum_{r=0}^{n_\mathrm{s}-1} \delta_m(\tau_r). \tag{21}$$

This gives

$$\tilde{P}_n = \sum_{r=0}^{n_\mathrm{s}-1} \exp\left[\frac{-2i\pi n \lfloor N\tau_r \rfloor}{N}\right]\left(1 + \{N\tau_r\}\left(\exp\left[\frac{-2i\pi n}{N}\right] - 1\right)\right), \tag{22}$$

which for $|n| \ll N$ yields

$$\begin{aligned} \tilde{P}_n &\simeq \sum_{r=0}^{n_\mathrm{s}-1} \exp\left(\frac{-2i\pi n \lfloor N\tau_r \rfloor}{N}\right)\left(1 - \{N\tau_r\}\frac{2i\pi n}{N}\right) && (23) \\ &\simeq \sum_{r=0}^{n_\mathrm{s}-1} \exp\left(\frac{-2i\pi n \lfloor N\tau_r \rfloor}{N}\right) \exp\left(\frac{-2i\pi n \{N\tau_r\}}{N}\right), && (24) \end{aligned}$$

which is equivalent to Eq. (11). For $|n|$ of order N the error becomes large, however if $\tilde{S}_n$ falls away to zero for $|n|$ much less than N, then these errors are insignificant for $\tilde{F}_n$ and this method provides a fast and accurate way of calculating $\tilde{F}_n$ that allows continuous variation of τ_r.

II. DETERMINING COMB PARAMETERS AND SOLITON NUMBER

Having generated $|\tilde{F}_n|$ from an experimental spectrum as described in the main text, the first step of the algorithm is to estimate A, A_0, β and n_s. This is done via two successive methods, the first being a direct calculation and the second involving a curve fit using the results of the first as an initial guess.

A. Initial estimate of A, A_0, β and n_s

For the first method, we define the background-free comb amplitudes

$$\tilde{F}'_n := A\tilde{P}_n\tilde{S}_n = \tilde{F}_n - B\delta_{n0}, \quad F'_m := A(P * S)_m \tag{25}$$

Because the solitons are distinct with very little overlap, we have

$$\sum_m |F'_m|^2 \simeq n_\mathrm{s} A^2 \sum_m |S_m|^2, \tag{26}$$

which implies that

$$\sum_n |\tilde{F}'_n|^2 \simeq n_\mathrm{s} A^2 \sum_n |\tilde{S}_n|^2. \tag{27}$$

Now for $1 \ll \beta/\pi^2 \ll N$,

$$\sum_n |\tilde{S}_n|^2 \simeq \int_{-\infty}^{\infty} \mathrm{sech}^2\left(\frac{\pi^2 x}{\beta}\right) dx = \frac{2\beta}{\pi^2} \tag{28}$$

and so

$$\sum_n |\tilde{F}'_n|^2 \simeq \frac{2n_\mathrm{s} A^2 \beta}{\pi^2}. \tag{29}$$

We can use similar arguments to show that

$$\begin{aligned} \sum_n n^2 |\tilde{F}_n|^2 \simeq \sum_n n^2 |\tilde{F}'_n|^2 &\simeq n_\mathrm{s} A^2 \int_{-\infty}^{\infty} x^2 \mathrm{sech}^2\left(\frac{\pi^2 x}{\beta}\right) dx \\ &= \frac{n_\mathrm{s} A^2 \beta^3}{6\pi^4} \end{aligned} \tag{30}$$

and that

$$\begin{aligned} \sum_n n^4 |\tilde{F}_n|^2 \simeq \sum_n n^4 |\tilde{F}'_n|^2 &\simeq n_\mathrm{s} A^2 \int_{-\infty}^{\infty} x^4 \mathrm{sech}^2\left(\frac{\pi^2 x}{\beta}\right) dx \\ &= \frac{7 n_\mathrm{s} A^2 \beta^5}{120\pi^6}. \end{aligned} \tag{31}$$

Combining Eq. (30) and Eq. (31) gives

$$\beta \simeq \sqrt{\frac{20\pi^2 \sum_n n^4|\tilde{F}_n|^2}{7\sum_n n^2|\tilde{F}_n|^2}}, \tag{32}$$

which we can calculate from the experimental values of $|\tilde{F}_n|$. We now use this value of β combined with Eqs. (29) and (30) to estimate

$$\sum_n |\tilde{F}'_n|^2 \simeq \frac{12\pi^2}{\beta^2}\sum_n n^2|\tilde{F}_n|^2, \tag{33}$$

from which we obtain

$$|\tilde{F}'_0|^2 = \sum_n |\tilde{F}'_n|^2 - \sum_{n\neq 0}|\tilde{F}_n|^2 \tag{34}$$

and thus

$$A_0 = \frac{|\tilde{F}_0|}{|\tilde{F}'_0|}. \tag{35}$$

Since $\tilde{P}_0 = n_\mathrm{s}$ and $\tilde{S}_0 = 1$, we also know that

$$\tilde{F}'_0 = |\tilde{F}'_0| = n_\mathrm{s}A, \tag{36}$$

and using this and Eq. (29) we can find n_s and A from $|\tilde{F}'_0|$, $\sum_n |\tilde{F}'_n|^2$ and β.

B. Better estimate of A, A_0, β and n_s via fit

The continuous-time inverse Fourier transform (as in Eq. (3)) of $\tilde{C}_n = |\tilde{F}_n|^2$ is given by the autocorrelation function

$$C(\tau) = \int_0^1 F(\tau')F^*((\tau+\tau') \bmod 1)d\tau'. \tag{37}$$

This is analogous to the Patterson function in X-ray crystallography. Since

$$|\tilde{F}_n|^2 = |\tilde{F}'_n|^2 + A^2n_\mathrm{s}^2(A_0^2-1)\delta_{m0}, \tag{38}$$

we have

$$C(\tau) = C'(\tau) + A^2n_\mathrm{s}^2(A_0^2-1) \tag{39}$$

where

$$C'(\tau) = \int_0^1 F'(\tau')F'((\tau+\tau') \bmod 1)d\tau' \tag{40}$$

in which the real function

$$F'(\tau) = A\sum_{r=0}^{n_\mathrm{s}-1} S(\tau-\tau_r) \tag{41}$$

for $S(\tau)$ as defined in Eq. (8). We thus have

$$C(\tau) = A^2\left[\sum_{r=0}^{n_\mathrm{s}-1}\sum_{r'=0}^{n_\mathrm{s}-1} T(\tau-\tau_r+\tau_{r'}) \ + n_\mathrm{s}^2(A_0^2-1)\right] \tag{42}$$

for

$$T(\tau) := \int_0^1 S(\tau')S(\tau'+\tau)d\tau' = \frac{2\beta}{\pi^2}\sum_{k=-\infty}^{\infty}\mathrm{cschc}\left(\beta\left(\tau+k\right)\right) \tag{43}$$

where we define $\mathrm{cschc}(x) := 1/\mathrm{sinhc}(x) = x/\sinh(x)$. Since the solitons have a minimum separation that is a few times their width, $C(\tau)$ will have a high, isolated peak centred at $\tau = 0$ with height $2A^2\beta n_\mathrm{s}/\pi^2$, coming from the sum of the autcorrelations of each soliton, with just the tails of the nearest peaks superimposed on it. With a little reasoning, we can thus say that for $|\tau|$ significantly less than the minimum soliton separation,

$$\begin{aligned} C(\tau) \simeq \frac{2A^2\beta n_\mathrm{s}}{\pi^2}\,&\mathrm{cschc}(\beta\tau) + J\cosh(\beta\tau) \\ &+ Km\sinh(\beta\tau) + A^2(A_0^2-1)n_\mathrm{s}^2 \end{aligned} \tag{44}$$

where J and K are real constants, and therefore when $\beta/\pi^2 \ll N$, the discrete-time inverse Fourier transform (as in Eq. (16)) C_m of $\tilde{C}_n$ can be expressed as

$$C_m \simeq \frac{H\,\mathrm{cschc}(Gm) + J\cosh(Gm) + Km\sinh(Gm) + L}{N} \tag{45}$$

for $|m/N|$ significantly less than the minimum soliton separation in τ, where

$$G = \frac{\beta}{N}, \quad H = \frac{2A^2\beta n_\mathrm{s}}{\pi^2} \quad \text{and} \quad L = A^2(A_0^2-1)n_\mathrm{s}^2. \tag{46}$$

To find better estimates of A, A_0, β and n_s, we first calculate C_m as the inverse FFT of $|\tilde{F}_n|$, then fit Eq. (45) to this over the range $0 \leq m < m_\mathrm{c}$ for some cutoff m_c using G, H, J, K and L as free parameters. The initial guesses of G, H and L are found from the previous etimates of A, A_0, β and n_s via Eq. (46), while those of J and K are taken as 0. The new values of A, A_0, β and n_s are then calculated from the fitted values of G, H and L using Eq. (46) together with the relation $\tilde{F}_0 = AA_0n_\mathrm{s}$.

The cutoff m_c has to be large enough to ensure an accurate fit but not so large that it gets too close to N times the minimum soliton separation in τ, causing this approximation to break down. It was found that a good way to decide m_c was to let $m_\mathrm{c} = \mu_\mathrm{c}N/\beta$ for relative cutoff μ_c, in other words to set m_c to be a fixed ratio to the soliton width $1/\beta$ based on the initial estimate of β. A value of $\mu_\mathrm{c} = 7$ was found to work well for all the combs we tried, giving a value of n_s always within 0.4 of the correct value found by actually fitting the soliton pattern, and producing accurate values of A and β as judged by the excellent match of the fitted simulated spectra (produced without further ajdustment of A, A_0 or β) to the experimental spectra. We also plot a graph of the fitted value of n_s vs. μ_c for μ_c between 2 and 10 to give an idea of the uncertainty of n_s. While there is sometimes considerable variation in the fitted value of n_s over this entire range of μ_c, there tends to be a flatter region around $\mu_\mathrm{c} = 7$, giving extra confidence in the value of n_s obtained there.