Electronic Supplementary Materials

For https://doi.org/10.1631/jzus.A2300251

Time-synchronous-averaging-spectrum based on super-resolution analysis and application in bearing fault signal identification

Zengle REN[1], Yuan WANG[2], Huiyue TANG[3], Xin'an CHEN[4], Wei FENG[1,5,6]

[1]*Guangdong Provincial Key Laboratory of Construction Robotics and Intelligent Construction, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China*

[2]*Shenzhen EBORAIL Technology Co., Ltd., Shenzhen 518038, China*

[3]*National Supercomputing Center in Shenzhen, Shenzhen 518055, China*

[4]*State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China*

[5]*University of Chinese Academy of Sciences, Beijing 100190, China*

[6]*Shenzhen University of Advanced Technology, Shenzhen 518107, China*

Section 1 Time Synchronous Averaging Spectrum (TSA-spectrum)

Definitions of TSA

TSA is defined as Definition 1. The only parameter of TSA is the operation cycle T, which is presented as Definition 2. By using TSA, a discrete sequence with length $L(L \gg T)$ is transformed into a sequence with the length of T, as illustrated in Fig. S1.

Definition #1: Time Synchronous Averaging (TSA). For a continuous signal $y = y(t)$, given a sampling frequency $F_s = 1/\mathrm{dt}$ and a sampling duration t, we can obtain a discrete sequence $\boldsymbol{Y} = \{y_i = y(i \cdot \mathrm{dt}) | i = 0,1,2,\ldots,L-1\}$, where $L = \lfloor t \cdot F_s \rfloor$. With a given positive integer T, we define the TSA of $\boldsymbol{Y}$ as

$$\mathbf{TSA}(\boldsymbol{Y}, T) = \left\{ \frac{1}{k} \sum_{j=0}^{k-1} y_{i+jT} \,\middle|\, i = 0,1,2,\ldots,T-1 \right\}, \tag{S1}$$

where $k = \lfloor L/T \rfloor$ is the total number of valid samples.

Particularly, when the sequence $\boldsymbol{Y}$ has an infinite length of L, the TSA of $\boldsymbol{Y}_\infty$ given parameter T is given as:

$$\mathbf{TSA}(\boldsymbol{Y}_\infty, T) = \left\{\lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} y_{i+jT} \,\middle|\, i = 0,1,2,\ldots,T-1\right\} \tag{S2}$$

where, $\boldsymbol{Y}_\infty$ represents a discrete sequence and $\boldsymbol{Y}$ has infinite number of elements.

Definition #2: Operation Cycle (OC) of TSA. The positive integer T of $\mathbf{TSA}(\boldsymbol{Y}, T)$ is called the operation cycle.

It should be emphasized that throughout this paper, we assume that $k = \lfloor L/T \rfloor$ is sufficiently large. Otherwise, $\mathbf{TSA}(\boldsymbol{Y}, T)$ may become meaningless due to a low signal to noise ratio (SNR). The value of $\lfloor L/T \rfloor$ is the key factor that contributes to the denoise characteristics of TSA. TSA provides an easy, concise, and special method to process a discrete sequence. TSA is a powerful tool for extracting the periodic components of a given sequence, especially for time series with a low SNR.

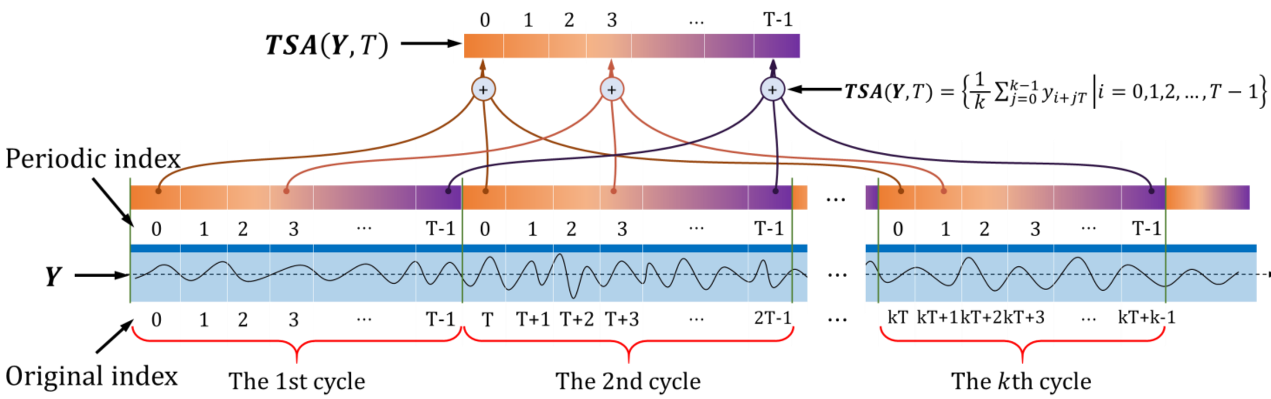

Fig. S1 Illustration of TSA

Basic theory of TSA

This section provides some of the basic mathematics of TSA. TSA is particularly useful when dealing with a quasiperiodic signal. Let N be the basic period of a sequence $\boldsymbol{Y} = \{y_i | i = 0,1,2,\ldots,L-1\}$, namely, $y_i = y_{i+N}$. All the following questions are about what we can obtain by applying TSA on $\boldsymbol{Y}$ with the operation cycle T. To answer this question, the following content is presented based on the congruence theory. We introduce two lemmas given as follows:

Lemma #1: Given $M, N \in \mathbb{N}^+$ and $(M, N) = 1$, the remainder set of $i \cdot M$ divided by N, denoted as $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N) = \{l_i |\, i \cdot M \equiv l_i \pmod N, i \in \mathbb{N}^+\}$, is given as $\{0,1,\ldots,N-1\}$.

Lemma #2: Given $M, N \in \mathbb{N}^+$ and $(M, N) = d > 1$, the remainder set of $i \cdot M$ divided by N, denoted as $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N) = \{l_i |\, i \cdot M \equiv l_i \pmod N, i \in \mathbb{N}^+\}$, is given as $\{0, d, 2d, \ldots, N-d\}$.

These foundational lemmas provide essential insights into the mathematical principles of TSA, paving the way for Proposition #1. The proof of Lemmas #1 and #2 and Proposition #1 are provided in the following.

Proposition #1: Let $\boldsymbol{Y} = \{y_i | i = 0,1,2,\dots,L-1\}$ be a discrete sequence sampled from a continuous periodic signal $y = y(t) = y(t + t_p)$ with the sampling frequency $F_s = 1/\mathrm{dt}$, where $t_p \in \mathbb{R}^+$ is the period of $y(t)$. If $t_p \cdot F_s = N \in \mathbb{N}^+$ and $L \gg N$ and given the operation cycle $T \in \mathbb{N}^+$ and $T \ll L$, with $(T, N) = d$, we have the following result

$$\mathbf{TSA}(\boldsymbol{Y}, T) = \left[\left\{\tfrac{1}{N'}\textstyle\sum_{j=0}^{N'-1} y_{i+jd} \,\middle|\, i = 0,1,\dots,d-1\right\}\right]_{T'} \tag{S3}$$

where $T' = T/d, N' = N/d$. $[\{\cdot\}]_k$ represents a copy of vector $\{\cdot\}$ with k times.

Proposition #1 underscores the essence of TSA in capturing periodic components within a signal. By exploiting the congruence theory, TSA isolates the cyclic variations, enhancing the extraction of periodic signals. Subsequently, we derive two significant corollaries from Proposition #1:

Corollary #1: Letting $\boldsymbol{Y} = \{y_i | i = 0,1,2,\dots,L-1\}$ be a discrete sequence with $y_i = y_{i+N}$ and taking operation cycle T satisfying $(T, N) = N$, we have $\mathbf{TSA}(\boldsymbol{Y}, T) = \{y_i | i = 0,1,2,\dots,T-1\}$.

Corollary #2: Letting $\boldsymbol{Y} = \{y_i | i = 0,1,2,\dots,L-1\}$ be a discrete sequence with $y_i = y_{i+N}$ and taking operation cycle T satisfying $(T, N) = 1$, we have $\mathbf{TSA}(\boldsymbol{Y}, T) = \{\bar{y} | i = 0,1,2,\dots,T-1\}$, where $\bar{y} = \frac{1}{N}\sum_{i=0}^{N-1} y_i$ is the average value of $\boldsymbol{Y}$ within in one period.

Corollary #1 shows us that TSA will produce the signal within the whole cycle if we only use the operation cycle T equal to the period N of the given signal. Corollary #2 shows that if we use an operation cycle T that shares no common divider with N, the TSA will produce a constant vector with each element equal to the average value of $\boldsymbol{Y}$ within one period.

These two corollaries have profound implications:

(1) **Precise Period Extraction:** Corollary #1 showcases the capability of TSA to faithfully reproduce the original signal when the operation cycle matches the signal's period. This property facilitates the extraction of periodic signals when the exact period is known. If we know the precise period N of sequence $\boldsymbol{Y}$, we can easily extract the periodic signal by applying TSA with the operation cycle of $T = N$.

(2) **Period Estimation in Unknown Cases:** Corollary #2 highlights the adaptability of TSA in cases where the exact period is unknown. TSA can be applied with various operation cycles, and the results from cycles significantly different from the signal's period can be statistically averaged, aiding in period estimation. If we don't know the precise period N of sequence $\boldsymbol{Y}$, we can search from a wide range of possible periods, and the results of the operation cycles that are far different from N can be averaged, which can be detected easily using a statistical method such as the standard deviation.

It should be noted that we present Proposition #1 in a relatively rigorous form by assuming that the sequence $\boldsymbol{Y}$ is sampled from a continuous signal $y(t)$ with a period $t_p \in \mathbb{R}^+$. Note that t_p is a real number, but the period of a discrete sequence is always an integer. In most cases, we expect that t_p is an integer multiple of the sampling interval $\mathrm{dt} = 1/F_s$, so the nature of TSA can be presented in an elegant form simply based on number theory. However, in practice, we can always observe a nonzero decimal part of t_p to be divided by dt. In this case, we provide Proposition #2 as a supplement to show how this nonzero decimal part influences the performance of TSA.

The situation becomes much more complicated when t_p, the period of $y(t)$, cannot be evenly divided by the sampling interval $\mathrm{dt} = 1/F_s$, that is, when $t_p \cdot F_s = N + \Delta$, where $N \in \mathbb{N}^+$ and $0 < \Delta < 1$. Particularly, we focus on the TSA with the operation cycle $T = N$. The proof of Proposition #2 is provided in the section of .

Proposition #2: Let $\boldsymbol{Y} = \{y_i | i = 0,1,2, \dots, L-1\}$ be a discrete sequence sampled from a continuous periodic signal $y = y(t) = y(t + t_p)$ with the sampling frequency $F_s = 1/\mathrm{dt}$, where t_p is the period of $y(t)$. Assuming that $t_p \cdot F_s = N + \Delta$, where $N \in \mathbb{N}^+$ and $0 < \Delta < 1$, and given the operation cycle $T = N$, we have the following result

$$\mathbf{TSA}(\boldsymbol{Y}, N) = \{\tilde{y}_i | i = 0,1, \dots, N-1\} \tag{S4}$$

where $\widetilde{\boldsymbol{Y}} = \{\tilde{y}_i | i = 0,1,2, \dots, L-1\}$ is the discrete sequence sampled from a filtered result of $y(t)$ by a moving average filter with a window length of $\lfloor L/N \rfloor \Delta \mathrm{dt}$.

By comparing Corollary #1 and Proposition #2, the influence of the nonzero decimal part of the period $t_p \cdot F_s$ on the result of $\mathbf{TSA}(\boldsymbol{Y}, N)$ can be understood as a smoothing effect. We introduce a new definition, called the smoothing effect length (SEL), given as Definition #3. The SEL is a quantitative description of the degree of the smoothing effect when the decimal part of $\boldsymbol{Y}$'s period t_p is not equal to zero. The SEL is determined by the ratio between L and N, the magnitude of Δ, and the sampling interval $\mathrm{dt} = 1/F_s$. Furthermore, the smoothing effect can only be observed when the operation cycle T is close to the cycle N of sequence $\boldsymbol{Y}$.

Definition #3: The smoothing effect length (SEL). The SEL, denoted as $L_{SE} = \lfloor L/N \rfloor \Delta \mathrm{dt}$, is the window length of the equivalent moving average filter in a TSA of a periodical signal $\boldsymbol{Y}$, where L is the length of $\boldsymbol{Y}$, dt is the sampling interval and $\Delta = t_p/\mathrm{dt} - \lfloor t_p/\mathrm{dt} \rfloor$ is the decimal part of $\boldsymbol{Y}$'s period t_p divided by dt. SEL plays a vital role in balancing noise reduction and signal preservation in TSA. Customizing SEL allows us to adapt TSA to various signal characteristics and analysis goals. We explore practical methods, such as Super Resolution Analysis (SRA), to optimize TSA's performance by adjusting SEL based on specific signal properties and analysis requirements.

Denoise property

In the beginning, we propose TSA based on the consideration of denoise characteristics. This section is focused on the denoise property of TSA. TSA performs robustly against noise. This can be addressed in Proposition 3, which is proven in the section of Proof of the Propositions 1-3.

Proposition 3 shows that we can use $\mathbf{TSA}(\widehat{\boldsymbol{Y}}, T)$ to approximate $\mathbf{TSA}(\boldsymbol{Y}, T)$. If $\lfloor L/T \rfloor$ is sufficiently large, $\mathbf{TSA}(\widehat{\boldsymbol{Y}}, T)$ is approximately equal to $\mathbf{TSA}(\boldsymbol{Y}, T)$. More exactly, the variance of this approximation decreases at a speed of $\lfloor L/T \rfloor^{-1}$. Proposition 3 tells us that we can simply enhance the denoise performance of TSA by increasing the sampling time of $\hat{y}(t)$.

Proposition #3: Given $\widehat{\boldsymbol{Y}} = \{\hat{y}_i | i = 0,1,2, \dots, L-1\}$ as a finite-length discrete sequence sampled from a continuous quasiperiodic signal $\hat{y} = \hat{y}(t) = y(t) + e(t)$ with the sampling frequency $F_s = 1/\mathrm{dt}$, where $y(t)$ is a continuous periodic signal with the period t_p and $e(t)$ is a noise term satisfying $\mathbb{E}[e(t)] = 0$ and $\mathbb{E}[e^2(t)] = \sigma^2$. When given the operation cycle T, we have the following result

1. $\mathbb{E}\left[\mathbf{TSA}(\widehat{\boldsymbol{Y}}, T) - \mathbf{TSA}(\boldsymbol{Y}, T)\right] = \{0 | i = 0,1,2, \dots, T-1\}$ (S5)
2. $\mathbb{E}\left[\left(\mathbf{TSA}(\widehat{\boldsymbol{Y}}, T) - \mathbf{TSA}(\boldsymbol{Y}, T)\right)^{\mathbf{2}}\right] = \left\{\frac{1}{k}\sigma^2 \middle| i = 0,1,2, \dots, T-1\right\}$ (S6)

where $k = \lfloor L/T \rfloor$.

TSA-spectrum

According to Propositions 1-3, it is obvious that if we know the major cycle N of a measured time series exactly, we can use TSA by setting the operation cycle $T = N$, so that we can extract the denoised signal hidden inside the measured signal. All other components with periods different from N, such as those T satisfying $(T, N) = 1$, are averaged or eliminated. However, the following problem concerns how to determine the accurate operation cycle of our interest. With this aim, this section introduces the TSA-spectrum, a tool used to analyze the hidden periods in a time series.

The definition of the TSA-spectrum is given as follows:

Definition #4: TSA-spectrum. Given a series of operation cycles $\boldsymbol{T} = \{T_i | i = 0,1,2, \dots.\}$, the TSA-spectrum $\mathbf{SPE}(\boldsymbol{Y}, \boldsymbol{T})$ of a discrete sequence $\boldsymbol{Y} = \{y_i = y(i \cdot \mathrm{dt}) | i = 0,1,2, \dots\}$ is defined as:

$$\mathbf{SPE}(\boldsymbol{Y}, \boldsymbol{T}) = \left\{\mathrm{std}\left(\mathbf{TSA}(\boldsymbol{Y}, T_i)\right) \middle| i = 0,1,2, \dots.\right\} \tag{S7}$$

where the operator $\mathrm{std}(\cdot)$ calculates the standard deviation.

The TSA-spectrum is an important tool for analyzing the hidden periods of different components in a time series. According to Propositions 1-3, if we set an operation cycle T_i close to the period of $\boldsymbol{Y}$, the value of the TSA-spectrum at that point is significantly large; otherwise, $\mathbf{TSA}(\boldsymbol{Y}, T_i)$ becomes an averaged value of $\boldsymbol{Y}$, and the standard deviation is small. Consequently, a larger spectrum value indicates that the operation cycle is more likely to be a hidden period in $\boldsymbol{Y}$. We can calculate the TSA-spectrum, analyze the key operation cycles, and finally extract the signal with those key operation cycles.

To illustrate the performance of the TSA-spectrum, we present two examples: (A) noisy sequence with single period and (B) noisy sequence with compound periods.

A. Noisy sequence with single period

Let $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}} = \{\hat{y}_i | i = 0,1,2,\dots,L-1\}$ be a finite-length discrete sequence sampled from a continuous noisy signal $\hat{y} = \hat{y}(t) = A \cdot \sin(2\pi t/N) + e(t)$ with the sampling frequency $F_s = 1$, where $e(t)$ is a white noise term satisfying $\mathbb{E}[e(t)] = 0$ and $\mathbb{E}[e^2(t)] = 1$. The TSA-spectrum of $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}}$ for two cases specified as (a) $A = 0.2,\ N = 200$, and (b) $A = 0.2,\ N = 200.3$ is illustrated in Fig. S2. On one hand, the TSA-spectrum of white noise with the finite length L will increase at the speed of $1/\sqrt{\lfloor L/T \rfloor}$. On the other hand, although the magnitude of the periodic term is much smaller than that of the noise term, a series of obvious peaks can be observed when the operation cycle is an integral multiple of N. The difference between Fig. S2(a) and Fig. S2(b) lies in whether or not the decimal part Δ of N is zero. According to Proposition 2, when $\Delta > 1$, $\mathbf{TSA}(\boldsymbol{Y}_{\boldsymbol{L}}, T)$ is equivalent to the smoothing-filtered result of $\boldsymbol{Y}_{\boldsymbol{L}}$ obtained from a moving average filter with the length of $k\Delta \mathrm{dt}$. As a result, the value of the TSA-spectrum at $T = N$ is smaller than that of $T = 2N$ and even smaller than that of $T = 3N$.

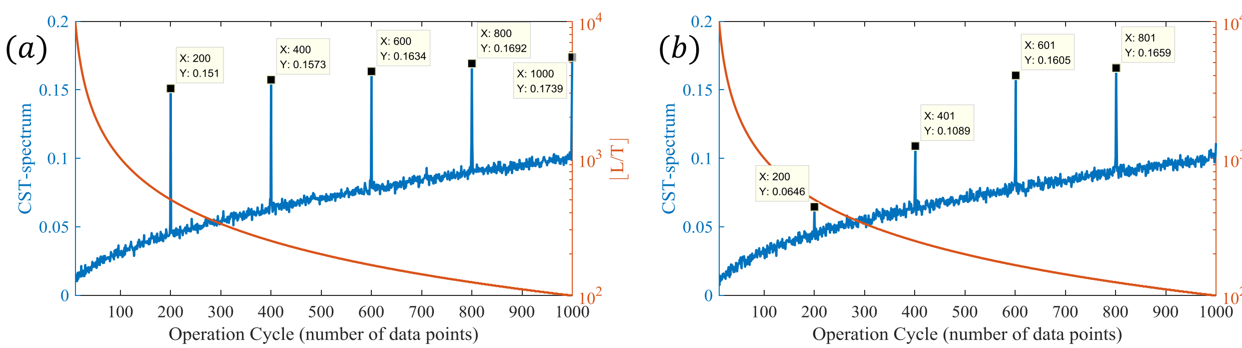

Fig. S2 The TSA-spectrum of $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}}$ with 100,000 data points. (a) $A = 0.2,\ N = 200$, and (b) $A = 0.2,\ N = 200.3$.

B. Noisy sequence with compound periods

Let $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}} = \{\hat{y}_i | i = 0,1,2,\dots,L-1\}$ be a finite-length discrete sequence sampled from a continuous noisy signal $\hat{y} = \hat{y}(t) = A_1 \cdot \sin(2\pi t/N_1) + A_2 \cdot \sin(2\pi t/N_2) + e(t)$ with the sampling frequency $F_s = 1$, were $e(t)$ is the white noise term. Given $A_1 = A_2 = 0.1$, $N_1 = 100$, and $N_2 = 150$, the TSA-spectrum of $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}}$ with 100,000 data points is illustrated in Fig. S3. Fig. S3 shows that the peaks are located at $(100,150,200,300,400,450,500,600,\dots.)$. Since there are two periodic components in $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}}$, there should be two groups of peaks in the TSA-spectrum. Let $\boldsymbol{P}_1 = (N_1, 2N_1, 3N_1, \dots)$ and

$\boldsymbol{P}_2 = (N_2, 2N_2, 3N_2, \dots)$ be the sets of operation cycles at the peaks caused by these two periodic components, respectively. Since $(N_1, N_2) = 50$ and $[N_1, N_2] = N_1 \cdot N_2/(N_1, N_2) = 300$. the TSA-spectrum shows a major period of 300, wherein the value of the TSA-spectrum becomes larger.

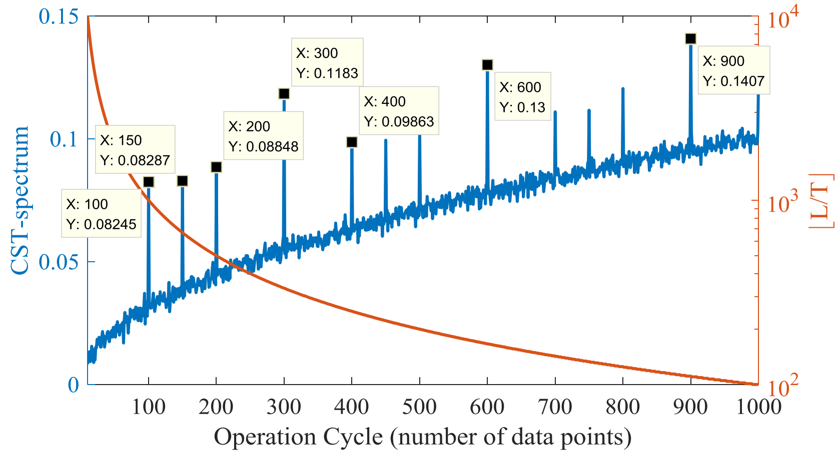

Fig. S3 The TSA-spectrum of $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}}$ with 100,000 data points, given $A_1 = A_2 = 0.1$, $N_1 = 100$, and $N_2 = 150$.

The TSA-spectrum is elaborated upon as a tool adept at discerning hidden periods within a time series. For signals that exhibit changing periodicities, the TSA-spectrum will manifest multiple distinct peaks, representing these dynamic periodic components. Specifically, in the context of feature extraction for such signals, the key parameters to consider would be the significant peaks in the TSA-spectrum, representing the varied periodicities inherent in the signal.

The incorporation of Super-Resolution Analysis is grounded in enhancing the resolution and precision of the TSA-spectrum. The core value of the TSA-spectrum lies in its ability to detect nuanced periodicities, even in noise-rich environments. Super-Resolution Analysis bolsters this capability by offering a more refined resolution, making it feasible to differentiate between closely-spaced periodicities and further augmenting the TSA's efficacy in unearthing concealed periodic patterns amidst noise.

Super resolution analysis (SRA)

According to Proposition 2, we can enhance the performance of the TSA-spectrum by minimizing the smoothing effect length (SEL). In this section, we propose a super-resolution analysis (SRA) achieved by increasing the sampling frequency (decreasing the sampling interval) through interpolation. SRA can be useful for reducing the SEL, so that the extracted signal is able to keep its original waveform. This is addressed by the following example. First, let $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}} = \{\hat{y}_i | i = 0,1,2, \dots, L-1\}$ be a finite-length discrete sequence sampled from a continuous noisy signal $\hat{y} = \hat{y}(t) = A \cdot \sin(2\pi t/N) + e(t)$ with the sampling frequency $F_s = 1$. $e(t)$ is a white noise term. Second, let $\widehat{\boldsymbol{Y}}_{\boldsymbol{sL}}$ be the linear interpolation of $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}}$ with the equivalent sampling frequency $F_s = s > 1$. Finally, according to Definition 3, the SEL of $\mathbf{TSA}(\widehat{\boldsymbol{Y}}_{sL}, \lfloor sN \rfloor)$ becomes 1/s of that of $\mathbf{TSA}(\widehat{\boldsymbol{Y}}_{L}, \lfloor N \rfloor)$.

Taking $N = 200.33$ as an example, the TSA-spectrum of $\widehat{\boldsymbol{Y}}_L$ with 100,000 data points is illustrated in Fig. S4(a). In contrast, the TSA-spectrum of $\widehat{\boldsymbol{Y}}_{sL}$ with $F_s = 10$ is presented in Fig. S4(b). For the original $\widehat{\boldsymbol{Y}}_L$ with $F_s = 1$, the first cycle T=200 is hard to observe since the decimal part of N is 0.33. When we apply SRA with $F_s = 10$, Fig. S4(b) shows that the first peak of the TSA-spectrum is located at 200.3, which is much closer than that of $\widehat{\boldsymbol{Y}}_L$ without SRA.

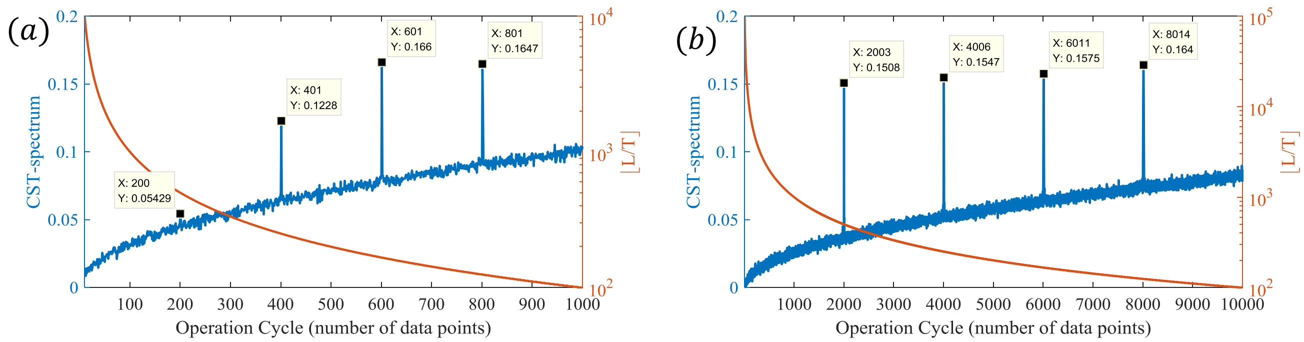

Fig. S4 The TSA-spectrum of $\widehat{\mathrm{Y}}_{\mathrm{L}}$ given $\mathrm{M} = 200.33$. (a) $\mathrm{F_s} = 1$ and (b) $\mathrm{F_s} = 10$

Proof of the Propositions 1-3

First, we prove Lemmas 1-2.

Lemma #1: Given $M, N \in \mathbb{N}^+$ and $(M, N) = 1$, the remainder set of $i \cdot M$ divided by N, denoted as $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N) = \{l_i |\, i \cdot M \equiv l_i \ (\mathrm{mod}\ N), i \in \mathbb{N}^+\}$, is given as $\{0,1, \dots, N-1\}$.

Proof:

(1) We know that for all $x \in \mathbb{N}^+$, the remainder set of x divided by N contains N elements, given as $\boldsymbol{L}(\mathbb{N}^+, N) = \{0,1, \dots, N-1\}$.

(2) There is a basic period for $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N)$, such that

$i \cdot M \equiv (i + N) \cdot M \equiv l_i \ (\mathrm{mod} N)$,

so $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N)$ is determined only by the remainder set of $\{iM |\, i = 1,2, \dots, N\}$.

(3) We can prove that when $(M, N) = 1$, we cannot find any two different integers $s, t \in [1, N]$, and $s > t$ satisfying

$s \cdot M \equiv t \cdot M \ (\mathrm{mod} N)$.

Otherwise, let us suppose that there exist two different integers s and $t \in [1, N]$ satisfying $s \cdot M \equiv t \cdot M \ (\mathrm{mod} N)$. Then, we get $(s - t) \cdot M \equiv 0 \ (\mathrm{mod} N)$. This indicates that $(s - t) \cdot M$ is divisible by N. However, the considering $(M, N) = 1$ and that M and N do not share any common divisor larger than 1, we have $s - t \equiv 0 \ (\mathrm{mod} N)$. This is impossible since $s, t \in [1, N]$ and $s > t$, $s - t \equiv s - t > 0 \ (\mathrm{mod} N)$.

(4) By combining the previous results, we know that for every integer $i \in [1, N]$, we will have a different remainder l_i with $l_i \equiv iM (\text{mod}\, N)$, so there are N different remainders for $\{iM | i = 1,2, \dots, N\}$ to be divided by N. Considering that there are, at most, N elements in the remainder set $\boldsymbol{L}(\mathbb{N}^+, N)$, we have

$$\boldsymbol{L}(\mathbb{N}^+ \cdot M, N) = \boldsymbol{L}(\mathbb{N}^+, N) = \{0,1, \dots, N-1\}$$

Done.

Lemma #2: Given $M, N \in \mathbb{N}^+$ and $(M, N) = d > 1$, the remainder set of $i \cdot M$ divided by N, denoted as $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N) = \{l_i | \, i \cdot M \equiv l_i \,(\text{mod}\, N), i \in \mathbb{N}^+\}$, is given as $\{0, d, 2d, \dots, N-d\}$.

Proof:

(1) We know that for all $x \in \mathbb{N}^+$, the remainder set of x divided by N contains N elements, given as $\boldsymbol{L}(\mathbb{N}^+, N) = \{0,1, \dots, N-1\}$.

(2) There is a basic period for $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N)$, such that

$i \cdot M \equiv (i + N) \cdot M \equiv l_i \,(\text{mod} N)$,

so $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N)$ is determined only by the remainder set of $\{iM | \, i = 1,2, \dots, N\}$.

(3) Since $(M, N) = d > 1$, we can rewrite M and N as $M = dM'$ and $N = dN'$, where $M', N' \in \mathbb{N}^+$ and $(M', N') = 1$. Then, we get

$$\boldsymbol{L}(\mathbb{N}^+ \cdot M, N) = \boldsymbol{L}(\mathbb{N}^+ \cdot dM', dN') = \boldsymbol{L}(\mathbb{N}^+ \cdot M', N') \times d$$

It can be interpreted that $\boldsymbol{L}(\mathbb{N}^+ \cdot M, N)$ can be given as $\boldsymbol{L}(\mathbb{N}^+ \cdot M', N')$ with each element multiplied by d. By using Lemma #1, we know that $\boldsymbol{L}(\mathbb{N}^+ \cdot M', N') = \{0,1,2, \dots, N'-1\}$. Last, we have

$\boldsymbol{L}(\mathbb{N}^+ \cdot M, N) = \boldsymbol{L}(\mathbb{N}^+ \cdot M', N') \times d = \{0,1,2, \dots, N'-1\} \times d = \{0, d, 2d, \dots, (N'-1)d\}$.

Done.

Second, we can prove Propositions 1-3 using Lemmas 1-2.

Proposition #1: Let $\boldsymbol{Y} = \{y_i | i = 0,1,2, \dots, L-1\}$ be a discrete sequence sampled from a continuous periodic signal $y = y(t) = y(t + t_p)$ with the sampling frequency $F_s = 1/\text{dt}$, where t_p is the period of $y(t)$. If $t_p \cdot F_s = N \in \mathbb{N}^+$ and $L \gg N$ and given the operation cycle $T \in \mathbb{N}^+$ and $T \ll L$, with $(T, N) = d$, we have the following result

$$\textbf{TSA}(\boldsymbol{Y}, T) = \left[\left\{ \frac{1}{N'} \sum_{j=0}^{N'-1} y_{i+jd} \,\middle|\, i = 0,1, \dots, d-1 \right\} \right]_{T'}$$

where $T' = T/d, N' = N/d$. $[\{\cdot\}]_k$ represents a copy of vector $\{\cdot\}$ with k times.

Proof:

We prove this proposition within two parts: (1) $(T, N) = d = 1$; (2) $(T, N) = d > 1$.

Part 1:

For the case when $(T, N) = 1$, we need to prove that

$$\textbf{TSA}(\boldsymbol{Y}, T) = \left\{ \frac{1}{N} \sum_{i=0}^{N-1} y_i \,\middle|\, i = 0,1,2, \dots, N-1 \right\}$$

Since $y(t) = y(t + t_p)$ and $t_p \cdot F_s = N \in \mathbb{N}^+$, we know that $y_{i+jN} = y_i$ for all $j \in \mathbb{N}^+$.
By using Lemma #1, we know that when given T with $(T, N) = 1$, the remainder set of $\{j \cdot T | j \in \mathbb{N}^+\}$ divided by N is given as $\{0,1, \dots, N-1\}$, which is also the remainder set of $\{i + j \cdot T | j \in \mathbb{N}^+\}$ for $i = 0,1,2, \dots, N-1$. Since $L \gg T$, we know that $k = \lfloor L/T \rfloor$ is sufficiently large. As a result, the occurrence frequency of each remainder in $\{0,1, \dots, N-1\}$ is the same. We have

$$\frac{1}{k}\sum_{j=0}^{k-1} y_{i+jT} = \frac{1}{k}\left(\frac{k}{N}\sum_{i=0}^{N-1} y_i\right) = \frac{1}{N}\sum_{i=0}^{N-1} y_i$$

Part 2:
When $(T, N) = d > 1$, we can rewrite N as $N = dN'$, where $N' \in \mathbb{N}^+$ and $(T, N') = 1$.
Then, we can then divide $\boldsymbol{L}(\mathbb{N}^+, N) = \{0,1,2, \dots, N-1\}$ into d groups $\{\boldsymbol{R_i} | i = 0,1,2, \dots, d-1\}$, with the ith group given as:

$$\boldsymbol{R_i} = \{i, d+i, 2d+i, \dots, N-d+i\}$$

By using Lemma #2, we know that when given T with $(T, N) = d > 1$, the remainder set of $\{j \cdot T | j \in \mathbb{N}^+\}$ divided by N is $\boldsymbol{R_0}$. Generally, for $i = 0,1,2, \dots, d-1$, the remainder set of $\{i + j \cdot T | j \in \mathbb{N}^+\}$ divided by N is $\boldsymbol{R_i}$. Now, we can also divide $\mathbb{N}^+$ into d groups according to its remainder divided by N. The ith group related to $\boldsymbol{R_i}$ is given as $\boldsymbol{G_i}$:

$$\boldsymbol{G_i} = \{i + j \cdot T | j \in \mathbb{N}^+\}; i = 0,1, \dots, d-1$$

Since $L \gg T$, this indicates that $k = \lfloor L/T \rfloor$ is sufficiently large. As a result, the occurrence frequency of each R_i is the same, and all elements in $\boldsymbol{R_i}$ share the same occurrence frequency.
Now, let us focus on the first $i < d$ elements of $\mathbf{TSA}(\boldsymbol{Y}, T)$. We have

$$\frac{1}{k}\sum_{j=0}^{k-1} y_{i+jT} = \frac{1}{k}\left(\frac{k}{N'}\sum_{j=0}^{N'-1} y_{i+jd}\right) = \frac{1}{N'}\sum_{j=0}^{N'-1} y_{i+jd}\,; i = 0,1, \dots, d-1$$

The first d elements of $\mathbf{TSA}(\boldsymbol{Y}, T)$ can be written as

$$\left\{\frac{1}{N'}\sum_{j=1}^{N'} y_{i+jd} \,\middle|\, i = 0,1, \dots, T'-1\right\}$$

Furthermore, according to the congruence properties stating that for $A, b \in \mathbb{N}^+$, if $A \equiv b \pmod{N}$, we have $A + c \equiv b + c \pmod{N}$ satisfied for all $c \in \mathbb{Z}$. It can be proven that the remainder set of $\boldsymbol{G_i} + kd$ for $k \in \mathbb{Z}$ divided by N also equals $\boldsymbol{R_i}$. As a result, it can be verified that there is a period of d in $\mathbf{TSA}(\boldsymbol{Y}, T)$.

$$\frac{1}{k}\sum_{j=0}^{k-1} y_{i+jT+kd} = \frac{1}{k}\left(\frac{k}{N'}\sum_{j=0}^{N'-1} y_{i+jd+kd}\right) = \frac{1}{N'}\sum_{j=0}^{N'-1} y_{i+jd}\,; k = 0,1,2, \dots, T'$$

This indicates that $\mathbf{TSA}(\boldsymbol{Y}, T)$ consists of a number of T' copies of the first d elements of $\mathbf{TSA}(\boldsymbol{Y}, T)$, which can be written as

$$\mathbf{TSA}(\boldsymbol{Y},T) = \left[\left\{\frac{1}{N'}\sum_{j=0}^{N'-1} y_{i+jd}\,\middle|\, i = 0,1,\dots,d-1\right\}\right]_{T'}$$

Done.

Proposition #2: Let $\boldsymbol{Y} = \{y_i | i = 0,1,2,\dots,L-1\}$ be a discrete sequence sampled from a continuous periodic signal $y = y(t) = y(t+t_p)$ with the sampling frequency $F_s = 1/\mathrm{dt}$, where t_p is the period of $y(t)$. Assuming that $t_p \cdot F_s = N + \Delta$, where $N \in \mathbb{N}^+$ and $0 < \Delta < 1$, and given the operation cycle $T = N$, we have the following result

$$\mathbf{TSA}(\boldsymbol{Y},N) = \{\tilde{y}_i | i = 0,1,\dots,N-1\}$$

where $\widetilde{\boldsymbol{Y}} = \{\tilde{y}_i | i = 0,1,2,\dots,L-1\}$ is the discrete sequence sampled from a filtered result of $y(t)$ by a moving average filter with the window length of $\lfloor L/N \rfloor \Delta\mathrm{dt}$.

Proof:

First, let $\widetilde{\boldsymbol{Y}} = \{\tilde{y}_i | i = 0,1,2,\dots,L-1\}$ be the discrete sequence sampled from a filtered result of $y(t)$ by a moving average filter with the window length $\lfloor L/N \rfloor \Delta\mathrm{dt}$, given the sampling frequency $F_s = 1/\mathrm{dt}$, over a sampling duration of $L\mathrm{dt}$ seconds. We have

$\tilde{y}_i = \frac{1}{k}\sum_{j=0}^{k-1} y\big((i - j\Delta)\mathrm{dt}\big)$

where $k = \lfloor L/T \rfloor$, $\Delta = t_p/\mathrm{dt} - \lfloor t_p/\mathrm{dt} \rfloor$ is the decimal part of t_p divided by dt.

Second, since $t_p \cdot F_s = N + \Delta$, we have $y(t) = y(t+t_p) = \cdots = y(t + jt_p); j \in \mathbb{N}^+$, which yields

$y(i\cdot\mathrm{dt}) = y\big(i\cdot\mathrm{dt} + t_pF_s\cdot\mathrm{dt}\big) = y\big((i+N+\Delta)\cdot\mathrm{dt}\big) = \cdots = y\big((i+jN+j\Delta)\cdot\mathrm{dt}\big)$

$$y_{i+jN} = y\big((i+jN)\cdot\mathrm{dt}\big) = y\big((i-j\Delta)\cdot\mathrm{dt}\big)$$

Finally, taking $T = N$, the ith element of $\mathbf{TSA}(\boldsymbol{Y},N)$ is given as

$\frac{1}{k}\sum_{j=0}^{k-1} y_{i+jN} = \frac{1}{k}\sum_{j=0}^{k-1} y\big((i-j\Delta)\cdot\mathrm{dt}\big) = \tilde{y}_i$

Done.

Proposition #3: Given $\widehat{\boldsymbol{Y}}_{\boldsymbol{L}} = \{\hat{y}_i | i = 0,1,2,\dots,L-1\}$ as a finite-length discrete sequence sampled from a continuous noisy signal $\hat{y} = \hat{y}(t) = y(t) + e(t)$ with the sampling frequency $F_s = 1/\mathrm{dt}$, where $y(t)$ is a continuous periodic signal with period t_p and $e(t)$ is a noise term satisfying $\mathbb{E}[e(t)] = 0$ and $\mathbb{E}[e^2(t)] = \sigma^2$, when given the operation cycle T, we have the following result

1. $\mathbb{E}\big[\mathbf{TSA}(\widehat{\boldsymbol{Y}},T) - \mathbf{TSA}(\boldsymbol{Y},T)\big] = \{0 | i = 0,1,2,\dots,T-1\}$
2. $\mathbb{E}\left[\left(\mathbf{TSA}(\widehat{\boldsymbol{Y}},T) - \mathbf{TSA}(\boldsymbol{Y},T)\right)^{\mathbf{2}}\right] = \left\{\frac{1}{k}\sigma^2 \middle| i = 0,1,2,\dots,T-1\right\}$

where $k = \lfloor L/T \rfloor$.

Proof:

According to Definition 1, we have

$$\mathbf{TSA}(\widehat{\boldsymbol{Y}}, T) - \mathbf{TSA}(\boldsymbol{Y}, T) = \left\{ \frac{1}{k} \sum_{j=0}^{k-1} (\hat{y}_{i+jT} - y_{i+jT}) \middle| i = 0,1,2, \dots, T-1 \right\}$$

$$= \left\{ \frac{1}{k} \sum_{j=0}^{k-1} e_{i+jT} \middle| i = 0,1,2, \dots, T-1 \right\}$$

Considering that $\mathbb{E}[e(t)] = 0$ and $\mathbb{E}[e^2(t)] = \sigma^2$, we have $\mathbb{E}[e_i] = 0$, $\mathbb{E}[e_i e_j] = 0$ and $\mathbb{E}[e_i^2] = 0$. Then, we have

$$\mathbb{E}\left[\frac{1}{k} \sum_{j=0}^{k-1} e_{i+jT}\right] = \frac{1}{k} \sum_{j=0}^{k-1} \mathbb{E}[e_{i+jT}] = 0$$

$$\mathbb{E}\left[\left(\frac{1}{k} \sum_{j=0}^{k-1} e_{i+jT}\right)^2\right] = \frac{1}{k^2} \mathrm{E}\left[\sum_{j=0}^{k-1} e_{i+jT}^2\right] = \frac{1}{k} \sigma^2$$

Done.