An adaptive phase-locking insight to unravel mRNAs synchrony from microarray experiments

Mariusz, Leslaw Zoltowski

***Abstract*—A bulk mRNA expression time course can slow or advance modulating in phase the averaged expression of each gene within the cell-cycle. This is also a synchronism induced at the cycle start to recognize periodic genes. Consistent with the above, a novel adaptive phase-locking insight into microarray experiment is provided. The budding-yeast expression of the major role in the cell-division-cycle could be clustered in phase by a new amplitude noise immune approach which compares well to meta-analysis of the web Cyclebase and to previous results in the field. The new approach to timing elucidation seems to be well matched to cellular phenomena and might be promising as the most accurate ever applied in Fourier context .**

I. Introduction

Biological cell life progresses a cooperation of concurrent biological threads; the each one occurs in series of discrete steps in a specific order while all of them could complete the subsequent stages: G0/G1, S, G2, M of a dividing and growing cell [1-3]. This is either extensive or intensive research that has been aimed to unravel different cellular events including the ones involved in cellular control; these all contributing to the recurrence of metabolism, protein synthesis, secretion, DNA replication, organelle biogenesis, chromosome segregation and cytoskeletal dynamics [4]. How bio-molecularly underpinned cellular threads nest the latter fundamental and observable processes is of great importance in combating disease or improving cognition.

Early methods by Spellman et al. [1] for identifying periodically expressed genes including the time-peak of expression proved superior than the other ones [5]. However their approach by the Fourier transform amplitude was unstable for small variations of ω radian frequency and averaging procedure over a range of 40 values was implemented [1, p 3276].

During a cell-cycle there are cellular events which can slow or advance its progression. These include nutrients availability or well-known check points to assess either the DNA-replication at S-synthesis or M-mitotic spindle status. Hence the bulk mRNA averaged expression synchronized at the start appears a waveform which is modulated in phase in the periodic i.e. the cell-cycle regulated genes case. This is the hypothesis that motivated the insight of this paper by an adaptive phase-locking approach to unravel mRNAs timing from a microarray experiment. An amplitude noise immune solution ever applied is to suit the experimental microarray context well.

M. Zoltowski is with Collegium Medicum Pharmacy Dept. of Nicolaus Copernicus University of Torun, phone: 48-691-833-860; (e-mail: mzoltowski@cm.umk.pl).

The next two sections are either on explanation or result presentation purpose whereas the last one serves conclusion and discussion.

II. Material and Methods

A. Expression data

The EMBL data of the budding-yeasts from high-density oligonucleotide tiling Affymetrix arrays were used; namely the alpha-factor mRNA transcripts from web repository [6]. Generally, so called microarray profiling results in signals of fluorescence relevant to the number of mRNA copies. Specifically, these mRNAs come from a population of synchronized cells and their fluorescently labeled cDNA counterparts are differentially (versus background) matched at each probe of the microarray by hybridization. They are samples of the cell-cycle time-course whose start is dictated at G1/S moment due to the population cells arrest by alpha pheromone.

B. Expression desynchronization

The synchronized expression of genes gets out of synchrony so fast as a few cell-cycle period due to the cell-to-cell variability [7] . This is a kind of variability termed noise by some biologists. It is reflected in the width of population distribution related to a particular event of the cell cycle. Variability can be also assessed by the ratio between the standard deviation and the mean expression level of the population. So the cell-cycle expressions, these by the log base two between the studied and the background ones, can be approximated statistically, e.g. by the periodic normal mixture (PNM, [8]). Accordingly:

$$P_{\omega}(t) = a_0 + \sum_{n=1}^{N} c_n \sin(n\omega t + \varphi_n) + n(t) \quad \text{(1a)}, \quad \text{where}$$

$\omega \sim N(\omega_0, \sigma_\omega)$ - the radian freq. relevant to a single cell expression period is distributed normally across the population, with ω_o mean value and σ_ω standard deviation. The other symbols of (1a) are a Fourier case familiar .

So the bulk i.e. averaged mRNA of the cell-cycle follows the weighted sum of the single-cell terms of (1a). That is, but additive noise term *n(t)*, given by the Fourier transform of Gaussian function:

$$P(t) = a_o + \sum_{n=1}^{N} c_n \int_{-\infty}^{\infty} \sin(n\omega t + \varphi_n) \frac{\exp[-(\frac{(\omega - \omega_o)}{2\sigma_\omega})^2]}{\sqrt{2\pi}\sigma_\omega} d\omega =$$

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

$$= a_o + \sum_{n=1}^{N} c_n \sin(n\omega_o t + \varphi_n) \exp(-\frac{1}{2} n^2 \sigma_\omega^2 t^2) \quad (1b).$$

The averaged expression of (1b) shows a fast decay of synchronism and supports further considerations.

C. Cell-cycle period estimation from expression

Firstly, to make a feasible digital signal processing (DSP) case, the sampling rate that is the screening frequency of mRNAs by microarrays was converted up 10 times with an aid of cubic splines. Secondly, the idea behind this insight was to phase-lock a first harmonic of each averaged expression alike in (1b) by a harmonic oscillator of ω_o free running radian frequency. This ω_o could be estimated by a maximum of the geometric mean of a set of N_i Fourier-spectra including these mRNAs that are cell-cycle regulated:

$$\left|H_{GM}(\omega)\right| = \left(\prod_{i=1}^{N_i} \left|H_i(\omega)\right|\right)^{\frac{1}{N_i}} \quad (2a).$$ Herein $\left|H_i(\omega)\right|$ is by

$$H(\omega) = \sum_{n=0}^{N_t-1} x(n)w(n)\exp(-jn\omega) \quad (2b)$$ as the Fourier transform of the expression $x(n)$ ($n=0,1,\ldots, N_t-1$). The weighting by $w(n)$ is to alleviate Gibbs's phenomenon.

In spectral calculations ω scales as $\sim 2\pi T_S / T_o$ with the sampling period T_s and $T_o=2\pi/\omega_o$.

D. Selecting first-harmonic of mRNA expression

Prior to phase-locked oscillator is a selective filter, the not delaying in phase one. It was implemented by a linear-phase-FIR-kernel selective around ω_o:

$$h_{\omega_o}^{*} = \{h_{\omega_o}^{*}(n) : -N_{FIR} \le n \le N_{FIR}\} = \\ = Z^{-1}\{H_{\omega_o}(z)H_{\omega_o}(z^{-1})\}. \quad (3a)$$

This is Z-inverse of a causal $H(z)$ and not causal $H(z^{-1})$ filtering cascade case that is in the above, with $N_{FIR}+1$ taps of the each FIR filter. Such filtering could be repeated several times to improve the spectral content attenuation around $\omega \approx 0$. This is also the expression power loss (EPL) due to the prior filtering that could be defined to account for the power change from initial P_o to the P_{FIR}: $\rho_{EPL} =$

$$= 10\log_{10}(P_{FIR} / P_o) \ . \quad (3b)$$

E. Adaptive phase-locking of an oscillator

The adaptive phase-locking of the oscillator stands for the control of its phase φ. This of $osc(n) = \exp[j\varphi_{osc}(n)] =$

$$= \exp(j[\omega_o n + \varphi(n-1)]) \quad (4a),$$ is controlled according to:

$$\varphi(n) = \varphi(n-1) + K_1\varphi_{err}(n) \quad n= 1\ldots N_t-1 \quad (4b).$$

Phase-locking is to agree the oscillator phase with the one of the expression waveform by $x_a(n)$ of (4d), given an error:

$$\varphi_{err}(n) = \varphi_a(n) - \varphi(n-1) = angle[x_a(n)\,osc(n)^*] \quad (4c).$$

$$x_a(n) = x_{FIR}(n) + jHT[x_{FIR}(n)] = \left|x_a(n)\right|\exp[j\,angle(x_a(n))] \\ = \left|x_a(n)\right|\exp[j\,\varphi_a(n)] \quad (4d).$$

By HT in (4d), this is Hilbert transform that acts upon the FIR filtered expression denoted by $x_{FIR}(n)$.

Finally, in view of (4)s a phase φ_{osc} of controlled oscillator mimics smoothly the one of x_a in (4d).

F. In Conclusion:

- Timing information preserving, the first-harmonic x_1 of averaged mRNA transcript is available via the tracked phase of the phase-locked oscillator i.e.:

$$x_1(n) \cong \sqrt{2P_{FIR}}\cos\varphi_{osc}(n)\,,\ n = \ldots\ldots L \quad (5a).$$

- During the cell-cycle, both the timing and the synchrony status assessing synchrony index is defined as:

$$d_s^k = \frac{1}{M}\sum_{n=N_{start}}^{N_{end}} \exp(j[\varphi_{osc}^k(n) - \varphi_{osc}^1(n)]), \quad (5b).$$
$$M = N_{start} - N_{end} + 1$$

  This is φ_{osc}^k phase of the k^{th} mRNA minus the phase of the 1^{st} reference expression that is in (5b). (5b) stands an arithmetic mean from N_{start} to N_{end} of phasors whose phase at moment n is the phase difference between the k^{th} gene and the first one. The first gene starting the cell-cycle serves the periodic i.e. the cell-cycle regulated reference mRNA.

- For a k^{th} gene mRNA transcript in the perfect synchrony with reference: $\left|d_s^k\right| = 1$ (5c) and a constant $\Delta\varphi^k = angle(d_s^k)$ shows the common phase of phasors of (5b) in this case.

- The index of (5b) serves a synchrony insight into genes expression of the next section.

- Besides a proper value of synchrony index, the periodic cell-cycle regulated genes should show least ρ_{EPL}-s and reasonable $\sqrt{2P_{FIR}}$ amplitudes.

III. Results

The synchrony index d_s^k was calculated for multiple mRNAs from some important groups of cellular threads. One such a group related to transcriptional cell program is shown in Fig.2. The moments N_{start} and N_{end} were chosen to encompass internal samples of the expression alleviating boundary effects. On insight purpose, twenty genes of Fig.1 were of concern ($k=0,1,\ldots 19$). They contribute in three ways to the cell-cycle dependent transcriptional regulatory network responsible for expression of genes whose products

perform stage-specific functions [9-10]. Firstly, a direct transcriptive regulation is related to an expression of this gene that belongs to the specific stage cluster of co-regulated genes. Secondly, indirect enzymatic regulation of another

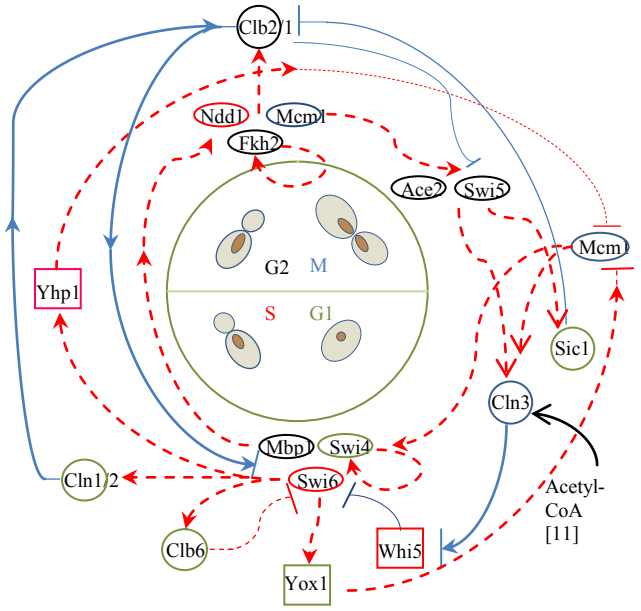

Figure 1. The Budding Yeast cell-cycle [9-10]. Transcriptional and posttranslational regulation by red and blue lines respectively. Activating transcription factors (TFs) by ovals whereas the repressors by rectangles. Cell-cycle regulatory proteins are denoted by circles. The cycle- phases: G1 in green; S in red; G2 in black and M in blue

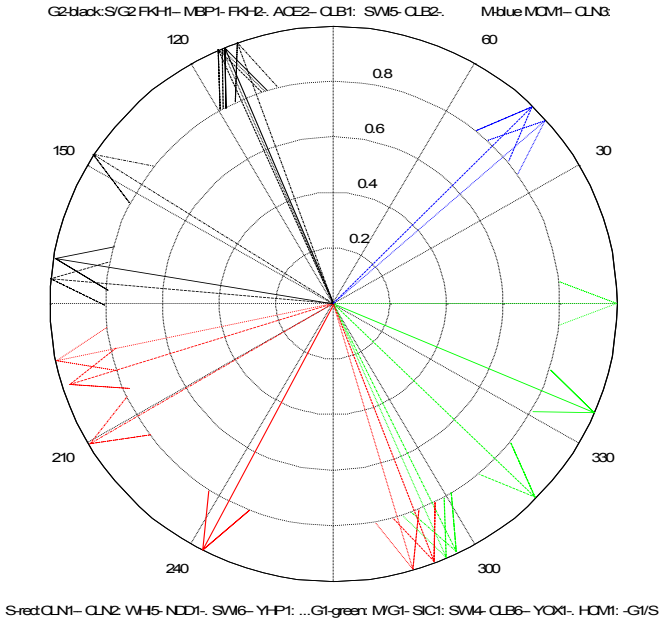

Figure 2. The clock-wise compass map of mRNA expression peak-time within cycle, related to the transcriptional regulation of Fig.1. The clock-wise: In green (G1): SIC1, SWI4, CLB6, YOX1, HCM1; In red (S): CLN1, CLN2, WHI5, NDD1, SWI6, YHP1; In black (G2): FKH1, MBP1, FKH2, ACE2, CLB1, SWI5, CLB2; In blue (M): MCM1, CLN3.

cluster of genes can occur e.g. by feedback repressing action to switch off genes of previous stage e.g. Cln3 in Fig.1. Finally, one cluster of genes is repressed or switched off by a transcriptional repressor, a member of another gene cluster e.g. Yox1 in Fig1.

This is the expression of SIC1 inhibitor of cdc28-Clb kinase-cyclin complexes controlling G1/S stage transition to prevent premature S stage that was used in (5b) index calculations as the reference. SIC1's expression peak-time that is almost at start of stage G1 is by the first horizontal clock-wise green arrow in the compass map of Fig.2. This map is a polar representation of the index. Hence the next expressions shown in the clock-wise order in Fig.2 can also be viewed by their peak- times. Their description follows in the caption below Fig.2 and they can be confronted with the transcriptional and posttranslational regulation of Fig.1.

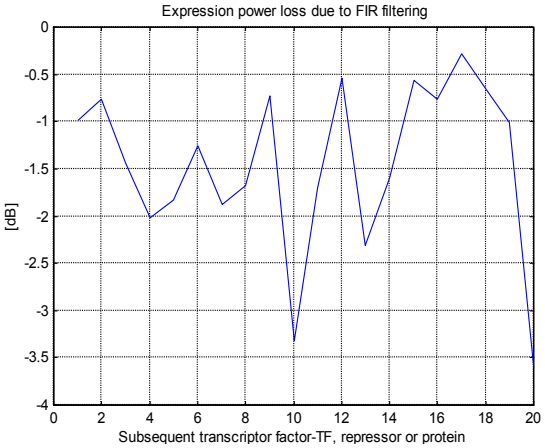

Figure 3. Expression power loss in [dB] due to Sec.2 FIR filtering for the subsequent genes given in the caption of Fig.2.

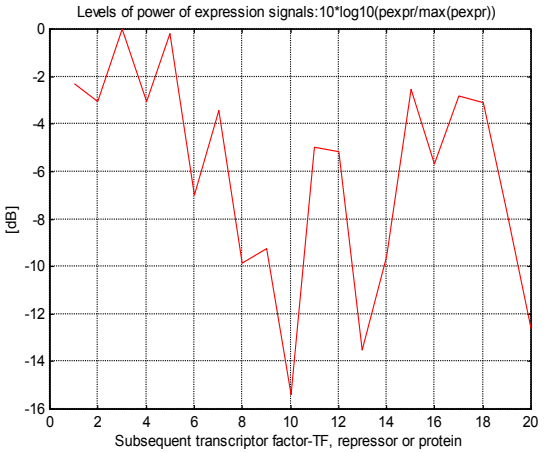

Figure 4. Expression power in [dB] of the subsequent genes given in the caption of Fig.2 versus the maximal expression power

The next Figs 3-4 are related to the power of the subsequent expressions of Fig.2. From Fig.3, one can learn the power loss of about 3.5 dB maximally. This is due to the selective filtering of Sec.2, as well-fitted to the cell-cycle. With an aid of Fig.4, the power extent of the considered mRNAs can be rated from 0 to -15 dB. This is versus the maximal power. The Figs 2-4 are of assistance when assessing periodicity of mRNA. They are different in a not cell-cycle regulated genes case .

IV. CONCLUSION

By their averaged expressions, the most important transcriptional regulation "players" [1-3,9-10] appear in almost perfect synchrony each versus the other. This is so since their arrows have lengths almost equal to 1 in the compass map of Fig.2. This is also the genes expression clustering case in this figure that is based on transcript's phase of a peak-time during the cell-cycle.

By averaging in (5b), the synchrony status index provides the mean phase differences for the interval from N_{start} to N_{end}. These phase differences are between the subsequent mRNA-expression-waveforms, so quite similar differences can be viewed between peak-times of these waveforms.

Adaptive phase-locking resynchronization of genes expression is robust against *n(t)* of (1a) noise contributing to the amplitude distortion. This is so since when restoring a phase of an incoming waveform, the adaptive phase-locking acts as a tracking filter of a very narrow bandwidth for the noise. The less K_1 of (4b) either the narrower this bandwidth or the less the capability of tracking of phase-modulation is [12].

Provided in Fig.2 data from the expression phase-clustering are in accord with the meta-analysis of Cyclebase [13]. Also this is a plausible specific finding that MCM1- TF is herein the cell-cycle regulated; the last but one blue arrow clock-wise in Fig.2. So MCM1 peaks in mitosis. Accordingly, this feature is confirmed by researchers [14] .

This is also the phase-locking that might be a cooperation principle of different cellular threads [15]. Hence applied mechanism to the timing elucidation seems to be well matched to cellular phenomena and might be promising to the field.

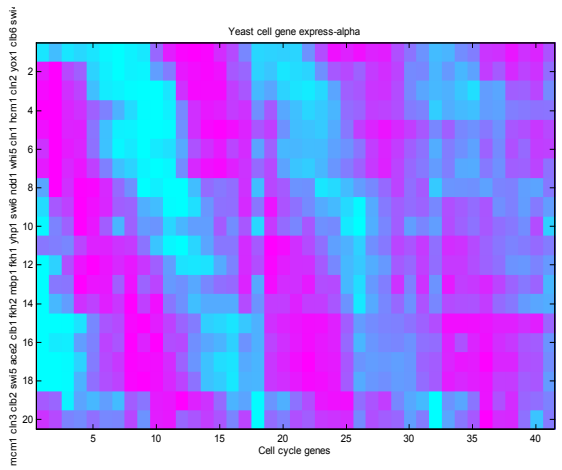

Figure 5. Cell-cycle regulated genes of Fig.2 by their crude mRNA expression from web repository (Sec.2) versus subsequent stages of screening.

While Fig.2 provides a sophisticated image of expression from a phase-locking approach, a straightforward amplitude visualization of the original data [6] from web repository is given in Fig.5 on comparison purpose. This is a time course of the same genes that are shown in Fig.2 . Also the Cyclebase information on these genes is provided in Table 1. The method described depends on how well the each RNA expression profile is represented by its first harmonic. A (1b) constrained extension to higher harmonics is possible.

TABLE 1. CONSIDERED CELL-CYCLE EXPRESSIONS ACCORDING TO CYCLEBASE [13]

Expr.	Short description	Peak at:
SIC1	Inhibitor of cdc28-Clb kinase complexes	G1 (4%)
SWI4	SBF TF activator in concert with MBF	G1(13%)
CLB6	B-cyclin, activates cdc28 for DNA synthesis	G1(14%)
YOX1	Transcrip. repress., binds to MCM1p & ECBs	G1(20%)
HCM1	Forkhead TF, S-phase specific	G1/S(25%)
CLN1	Cyclin activ. cdc28 to promote G1-S transition	G1/S(25%)
CLN2	Cyclin activ. cdc28 to promote G1-S transition	G1(22%)
WHI5	Repress. of G1-transcrip.,binds to SCB of SBF	S(34%)
NDD1	Transcrip. activ. essential for nuclear division	S(41%)
SWI6	T co-F, binds: Swi4p, Mbp1p, at G1/S & M	S(38%)
YHP1	T. repress., binds to Mcm1p & ECB of SBF	S/G2(49%)
FKH1	Forkhead TF, minor role in G2/M expression	S/G2(50%)
MBP1	MluI box TF with Swi6p, regul. from G1 to S	S/G2(52%)
FKH2	Forkhead TF, major role in G2/M expression	G2(53%)
ACE2	TF that activates early G1-specific genes	G2(64%)
CLB1	B-cyclin, activates cdc28 for transition G2-M	G2(63%)
SWI5	TF for genes of G1/M and G1phases	G2(65%)
CLB2	B-cyclin, activates cdc28 for transition G2-M	G2(66%)
MCM1	TF for cell-type specific & pheromone respons.	?
CLN3	G1-cyclin, activates cdc28 for transition G1-S	M(86%)

REFERENCES

[1] P.T. Spellman, G. Sherlock, M. Q. Zhang et al., "Comprehensive Identification of Cell Cycle-regulated Genes", *Molecular Biology of the Cell*, vol.9, December 1998, pp 3273-97.

[2] I. Simon, J. Barnett, N. Hannett et al., "Serial Regulation of Transcriptional Regulators in the Yeast Cell Cycle", *Cell*, vol.106, September 21, pp 697-708.

[3] D.A. Orlando et al., "Global control of all transcription by coupled CDK network oscillators", *Nature*, 2008, June 12, pp 944-47.

[4] T. Hunt, K. Nasmyth and B. Novák, "The cell cycle", *Phil. Trans. R. Soc. B* , vol.366, 14 November 2011, pp 3494–97.

[5] Ulrik de Lichtenberg, L. J. Jensen et al.," Comparison of computational methods for the identification of cell cycle regulated genes", *Bioinformatics*, vol. 21, no. 7, 2005, pp 1164–71 .

[6] M.V. Granovskaia, L. J. Jensen, M. E. Ritchie et al.," High-resolution transcription atlas of the mitotic cell cycle in budding yeast", *Genome Biology*, 2010, vol.11, R24.

[7] D. Muzzey and A. van Oudenaarden, "Quantitative Time-Lapse Fluorescence Microscopy in Single Cells", *Annu. Rev. Cell Dev. Biol.*,vol.25, 2009, pp 301-27.

[8] X. Lu, W. Zhang, Z.S. Qin, K.E. Kwast and J.S. Liu, "Statistical resynchronization and Bayesian detection of periodically expressed genes", *Nucleic Acids Research*, 2004, vol.32, No 2, pp 447-55.

[9] C. Wittenberg and S.I. Reed," Cell cycle – dependent transcription in yeast: transcription factors and transcriptomes", *Oncogene*, vol.24, 2005, pp 2746-55.

[10] J. Bähler," Cell Cycle Control of Gene Expression in Budding and Fission Yeast", *Annu. Rev. Genet.*, 2005, vol.39, pp 69-94.

[11] L. Shi and B. P. Tu," Acetyl-CoA induces transcription of the key G1 cyclin CLN3 to promote entry into the cell division cycle in Sacchar. cerevisiae", *PNAS*, vol.110, no 18, April 30, 2013, pp 7318-23.

[12] M. Zoltowski, "Some advances and refinements in digital phase-locked loops (DPLLs)", *Signal Processing*, vol.81, 2001, pp 735-89, Elsevier Science B.V.

[13] http://www.cyclebase.org

[14] S. Pondugula, D. W. Neef et al., "Coupling Phosphate Homeostasis to Cell Cycle-Specific Transcription: Mitotic Activation of *Saccharom. cerevisiae PHO5* by Mcm1 and Forkhead Proteins", *Mol. & Cell. Biol.*, vol. 29, No 18, Sept. 2009, pp 4891–905.

[15] Y. Lu and F. R. Cross, "Periodic Cyclin-Cdk Activity Entrains an Autonomous Cdc14 Release Oscillator", *Cell*, vol.141, April 16, 2010, pp 268–79.