

Independent Component Analysis With Data-centric Contrast Functions For Separating Maternal And Twin Fetal ECG

Mallika Sridhar-Keralapura, Mehrdad Pourfathi and Birsen Sirkeci-Mergen *

Abstract—Intra-uterine fetal ECG monitoring is critical for identical twin gestations because of the increased risk for cardiac defects. Ultrasound echocardiography is important clinically but it does not provide entirely conclusive information pertaining to the fetal cardiac conduction system. On the other hand, fetal electrocardiography can be obtained indirectly by means of ordinary electrodes placed on the mother's abdomen, allowing potential separation of the fetal electrocardiogram (FECG) for determination of fetal ECG heart rate and morphology. Difficulties with the twin FECG separation is due to the large maternal ECG interference, surrounding noise, artifacts and the underlying similar twin ECG morphology, amplitudes and heart rates. The objective of this work is to investigate Fast-ICA, a signal separation technique, using standard contrast functions and a newly optimized data centric method in this context under different types of interference. We clearly show with a variety of simulations that the chosen polynomial based contrast functions ($3^{rd} - 6^{th}$ order) perform superior to the data based Pearson method. They work on par and in some cases better than standard ICA polynomial schemes like SKEW and POW3. Similar trends were seen with a sample of in-vivo data as well. Data-centric schemes used on the ICA have significant potential in true in-vivo situations where separation is based on the underlying data characteristics. *Keywords: ICA, Fast-ICA, ECG, twins, polynomial*

1 Introduction

Intra-uterine monitoring of fetal health seeks to assess several indicators like growth and maturation, oxygen availability, neurological function and cardiac function (to diagnose cardiac hypertrophy, arrhythmias and con-

genital heart defects - CHD) [1]. Fetal monitoring becomes more critical for identical twin gestations, since the risk for cardiac defects increase (9-15 fold) with such pregnancies [2]. With such conditions, both fetuses are at risk for heart failure, requiring simultaneous monitoring of the cardiac traces, to highlight signs of overload and dysfunction to optimize time of delivery [3]. Current trends in fetal cardiovascular diagnosis are generally based on ultrasound echocardiography. Ultrasound is clinically important and identifies 25-60% of major heart defects, between 16 and 22 weeks, but it does not provide entirely conclusive information pertaining to function of the fetal cardiac conduction system [4]. Furthermore echocardiography, requires a trained physician, frequent repositioning of the transducer and is not recommended for long-term recordings in the normal home environment [4]. Fetal Electrocardiography on the other hand is an attractive candidate for this measurement. Fetal electrocardiograms (FECG) can be obtained indirectly by means of ordinary electrodes placed on the mother's abdomen. Monitoring the FECG could potentially enable determination of the fetal heart rate signal with beat-to-beat accuracy, and allow analysis of morphological and temporal changes [1]. The FECG can also be used to measure RR intervals from which the fetal heart rate (FHR) and its variability can be determined, to diagnose conditions like tachycardia (FHR>180bpm) or bradycardia (FHR<110 bpm) commonly seen with cardiac defects [1]. Specifically, fetal hypoxia can be detected when the PR and RR intervals vary, whereas a depression of the ST segment is associated with acidosis [5].

Even though acquiring the FECG seems very attractive for fetal cardiac assessment, its use in clinics has been quite limited. FECG is a part of abdominal ECG waveforms and contains interferences that come from multiple sources. The omnipresent Maternal ECG (MECG), which can be 5-1000 times larger in its intensity, forms the largest interference. Electromyographic (EMG) activity, electrode drift, power-line coupling and thermal/electronic noise corrupt the recordings as well [6]. The problem becomes even more difficult in the case of twin fetuses that could have FECG signals of similar morphology, amplitudes and heart rates [3].

*Dr. Keralapura (email: mallika.keralapura@sjsu.edu) is the director of the Biomedical Systems Lab, EE department San Jose State University. Mehrdad Pourfathi (email: mehrdad.pourfathi@gmail.com) is a graduate student with interests in biological signal processing and Dr. Sirkeci-Mergen (email: Birsen.Sirkeci@sjsu.edu) is the director of the Wireless Communications Lab, EE department, San Jose State University. For correspondence, please contact Dr. Keralapura at Department of Electrical Engineering, One Washington Square, San Jose, CA 95192-0084. Tel/Fax: (408) 924-3915/3925. email: mallika.keralapura@sjsu.edu.

Several signal separation algorithms have been proposed in the literature for separating FECG from abdominal ECG primarily for singleton pregnancies [5]. Classical filtering methods and adaptive techniques, singular value decomposition techniques and wavelet transform based methods have been proposed. Only a few studies have been reported in the context of twin gestations. For example, Lathauwer *et. al.* [7] used abdominal data of twins to illustrate a signal separation concept. Comani *et. al.*, [3], measured and separated twin magnetocardiography data (an alternative approach to FECG) for diagnosis, whereas, Taylor *et. al* [8] was one of the only studies that has demonstrated single fetus and twin FECG detection on a large database of clinical records. Separation was done using blind source separation (BSS) techniques. The most common BSS technique is the independent component analysis (ICA) with different implementations like INFOMAX, the JADE, the FastICA and the MERMAID [9]. ICA research has been very vast with several estimation methods that have been proposed.

The objective of this work is to investigate Fast-ICA using standard contrast functions and comparing them with our newly optimized data centric contrast functions in the context of twin FECG extraction from abdominal ECG channel data simulated for normal and abnormal cardiac conditions under different types of interference.

2 Independent Component Analysis

The goal of the ICA (Independent Component Analysis) is to unmix sources that are assumed to be linearly mixed with unknown coefficients. In this context, it is common to model the abdominal ECG recordings of a pregnant woman carrying twins as a linear combination of twin fetal and maternal signals [5]. The linearly mixed sources can be represented in the vector form as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad (1)$$

where $\mathbf{s}(t) = [s_1(t)s_2(t) \dots s_n(t)]^T \in \mathcal{R}^N$ denotes the N source signals, $\mathbf{x}(t) = [x_1(t)x_2(t) \dots x_M(t)]^T \in \mathcal{R}^M$ denotes the observation vector, and $\mathbf{A} \in \mathcal{R}^{M \times N}$ is the unknown mixing matrix. It is assumed that $M \geq N$. ICA uses the assumption of statistical independence among source signals (which is true in this context) to estimate the source signals $\mathbf{s}(t)$ and the mixing matrix \mathbf{A} . In other words, ICA methods try to determine the demixing matrix \mathbf{W} , which is the inverse of \mathbf{A} , so that the rows of $\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t)$ become statistically independent. In the ICA context, independence is equivalent to non-gaussianity and we can measure this using kurtosis and negentropy [9]. For example, for a gaussian random variable kurtosis is zero. The negentropy for a random variable y is defined as $J(y) = H(y_{gauss}) - H(y)$, where $H(y) = \int f(y) \log f(y) dy$ denotes the information theoretic differential entropy, and y_{gauss} is a Gaussian random variable with the same covariance as y . Since Gaussian

random variables have the largest entropy among all the random variables with the same variance, the negentropy is always nonnegative. The larger the negentropy, the closer the random variable gets to be non-Gaussian. The main issue with negentropy calculations is the fact that the distribution of the random variable y is needed in calculation. Instead, we can use approximation of negentropy [9], eq. 2, for a non-quadratic function G , and zero-mean, unit variance Gaussian random variable ν . The nonlinear function G is known as *contrast function*.

$$J(y) \approx [E\{G(y)\} - E\{G(\nu)\}]^2 \quad (2)$$

Ideally, if the source distribution $f(y)$ is known, the function $G(y) = -\log f(y) = -\int \frac{f'(y)}{f(y)} dy$ would be the optimal choice for the contrast function. In general, the selection of the contrast function depends on the data and the application. Fast-ICA is a fixed point algorithm that maximizes an approximation of negentropy for non-gaussianity. The main iteration step of the Fast-ICA algorithm is

$$\mathbf{w} \leftarrow E\{\mathbf{z}g(\mathbf{w}^T \mathbf{z})\} - E\{g'(\mathbf{w}^T \mathbf{z})\}\mathbf{w} \quad (3)$$

where \mathbf{z} is the whitened and centered data, and $g()$ is the derivative of the contrast function $G()$, and $g'()$ is its second derivative. The vector \mathbf{w} denotes one column of the estimated inverse matrix \mathbf{W} . Commonly used contrast functions with Fast-ICA algorithm are listed as: (i) *Pow3*: $g(y) = y^3$; (ii) *Gauss*: $g(y) = y \exp(-y^2/2)$; (iii) *Tanh*: $g(y) = \tanh(y)$; (iv) *Skew*: $g(y) = y^2$; (v) *Pearson*: $g(y) = -\frac{y-a}{b_0+b_1y+b_2y^2}$ where a, b_0, b_1, b_2 are found using the method of moments. Among all the contrast functions above, Pearson system uses some statistical properties of the data to calculate constants a, b_0, b_1, b_2 in the contrast function. The rest do not use any of the properties of data for selection of the contrast function. In the next section, we describe a new method that uses certain characteristics of the ECG signal to derive a data-centric contrast function.

2.1 Empirical density based Fast-ICA

In this section, we propose a new way of obtaining the contrast function using an estimate of the underlying probability density function (pdf) of the sources (maternal and twin fetal). This way, certain properties of the source signals can be incorporated into the Fast-ICA algorithm. It is important to note that the mother ECG, the twin fetal ECG, and the noise signals have different morphological and temporal characteristics and utilizing the same contrast function could result in suboptimal solutions. Our goal is to incorporate the differences in the source signals into the Fast-ICA by utilizing different contrast functions that are functions of the empirical pdf.

We follow the following steps in order to obtain contrast functions that are data centric. (1) *Generate template*

source signal: More details are provided in the next section, (2) *Obtain the empirical pdf*: Among the various ways of obtaining pdf estimates, we choose 2 techniques - scaling the histogram of the observed data and kernel density estimation (kde), which is a non-parametric method, (3) *Obtain the contrast function*: we set the contrast function as the function: $G(y) = -\log(f_e(y))$, where $f_e(y)$ is the empirical pdf for the selected source. Note that if the pdf estimation was perfect, then this would be the optimal choice for the contrast function.

In the Fast-ICA algorithm, the first and the second derivatives of the contrast function $G(y)$ are needed. To obtain these from the pdf, we propose to fit a polynomial to $G(y)$ and then take the first and second derivatives analytically. In this case of polynomial fit with a fixed order L , the contrast function would be

$$G(y) = -\log(f_e(y)) \approx \sum_{i=0}^L a_i y^i.$$

The coefficient a_0 could be selected to be zero for simplicity. The polynomial order could also be optimized for different sources. Note that the standard Fast-ICA offers polynomial based contrast functions *skew* and *pow3* but with unoptimized coefficients.

2.2 Simulation of Data for Twins

We generated all the ECG source signals (maternal and twin fetal) based on a dynamical model presented in [11]. The model generates a trajectory in a three-dimensional (3-D) statespace with coordinates (x,y,z) . Each revolution on this circle corresponds to one RR-interval or heartbeat. Distinct points on the ECG, such as the P, Q, R, S, and T are described by events corresponding to negative and positive attractors/repellers in that direction. These events are placed at fixed angles along the unit circle given by θ_i . The resultant ECG signal is synthesized by solving a set of differential equations shown below:

$$\begin{aligned} \dot{x} &= \alpha x - \omega y; & \dot{y} &= \alpha y + \omega x \\ \dot{z} &= - \sum_{i \in \{P,Q,R,S,T\}} a_i \Delta\theta_i \exp\left(\frac{-\Delta\theta_i^2}{2b_i^2}\right) - (z - z_0)\lambda \end{aligned}$$

where $\alpha = 1 - \sqrt{x^2 + y^2}$, $\Delta\theta_i = (\theta - \theta_i) \bmod 2\pi$, $\theta = \text{atan2}(x, y)$, $z_0 = A \sin(2\pi f_2 t)$ and f_2 is the respiratory frequency. With some trial and error, we found suitable values for a_i , b_i and $\Delta\theta_i$, for each segment, resulting in more realistic maternal and fetal ECG signals with amplitudes, timing and frequency content close to clinically reported signals. Table 1 lists the values used into the ECG model and Table 2 shows the range of ECG amplitudes and timing values used in this paper. These values fall within the range reported by Shepovnikov *et al.* [12] for in-vivo maternal and fetal ECGs.

Table 1: Parameters input into the ECG model

MECG	P	Q	R	S	T
time (s)	-0.202	-0.044	0	0.044	0.216
$a_i; b_i$	0.3;0.19	0.3;0.1	28;0.08	-7;0.1	0.5;0.29
θ_i (deg)	-80.47	-16.94	0	16.94	84.71
FECG1;2	P	Q	R	S	T
time (s)	-0.084;-0.088	-0.02;-0.024	0;0	0.02;0.024	0.056;0.076
a_i	0.04;0.04	-3;-2	3.5;3	-4;-5	0.1;0.2
b_i	0.5;0.4	0.3;0.32	0.3;0.33	0.2;0.22	0.5;0.4
θ_i (deg)	-103.5;-116.31	-31.5;-42.46	0;0	31.5;42.46	72;95.08

Table 2: Durations and amplitudes of MECG and FECG

MECG	P-Q	Q-T	R-R	S-T	QRS
(sec)	0.158	0.26	0.83	0.172	0.088
FECG1;FECG2	P-Q	Q-T	R-R	S-T	QRS
(sec)	0.064;0.064	0.076;0.1	0.4;0.39	0.036;0.09	0.04;0.048
MECG	P	Q	R	S	T
(mV)	0.521	0.15	2.102	-0.266	0.804
FECG1;FECG2	P	Q	R	S	T
(mV)	0.135;0.133	0.053;0.077	0.242;0.231	0.089;0.058	0.142;0.141

2.3 Simulation Studies

The maternal and the fetal ECG signals are uniquely different in their ECG durations, amplitudes and also spectral content [12]. In this paper, we assume maternal heart rates fixed at 72 bpm and vary fetal hearts around an average of 150 bpm. At first a noise free scenario was tested where maternal and 2 fetal ECG signals were mixed using a randomly selected 3-by-3 mixing matrix to form 3 channels of data that was input into the algorithms. We then quantitatively compared the reconstruction performance of Fast-ICA with different standard contrast functions and our newly proposed data centric functions using a metric called the performance index or the PI-metric. Since ICA methods cannot exactly determine the scaling (the energy) of the sources, and their orders, the Amari distance [9] is useful for performance comparison. The Amari distance (or the PI metric) is defined as follows: Let e_{ij} be the (i,j)th element of the matrix $\mathbf{E} = \mathbf{WA}$. Then the PI metric is -

$$PI = \frac{1}{n} \sum_{i=1}^n \left\{ \left(\sum_{k=1}^n \frac{|e_{ik}|}{\max_j |e_{ij}|} - 1 \right) + \left(\sum_{k=1}^n \frac{|e_{ki}|}{\max_j |e_{ji}|} - 1 \right) \right\}. \quad (5)$$

We then repeated the simulations for noisy channel data using zero-mean unit variance Gaussian noise and electromyographic (EMG) artifacts. Gaussian noise is chosen to simulate the collection of the numerous sources of noise all added up together including electrode noise, environmental noise, electronics, etc. The effect of its variance is also tested using the ICA schemes. The EMG signal arises from the mother's abdominal muscle movements. We used data from physionet online signal archives [13] for this source. Other simulations involved testing the ICA techniques over acquisition time, closely similar fetal signals (primarily heart rate) and with in-vivo maternal ECG (from physionet). In-vivo fetal ECG data is currently unavailable online and requires a separate clinical study.

3 Results and Discussion

Fig 1 shows an example of the maternal and twin fetal ECG data used in the ICA simulations generated using the ECG dynamical model [11] and parameters in Table 1. Parameters were carefully chosen (see Table 2) so that, the ECG segments and amplitudes would be within the range of clinically reported results. It is clear from the figure that the two signals are not identical in terms of their morphological and temporal behavior as described in Taylor *et. al.* [8] and Shepoválnikov *et. al.* [12]. We also estimated the power spectral density using the standard Welch technique and calculated the center of gravity of the spectrums. Values obtained for MEGC, FECG1 and FECG2 were 9.27Hz, 20.22Hz and 17.59Hz. These values closely agree with those reported in [12]. The above 3 sources (for example: MEGC= 72, FECG1= 150, FECG2= 147bpm) were combined to form 3 channels of data using a 3-by-3 mixing matrix. With a starting random guess of the mixing matrix, the source data was estimated and the PI metric (eq. 5) was calculated. To assess the performance with noise (EMG and gaussian noise), the number of channels was increased to 5.

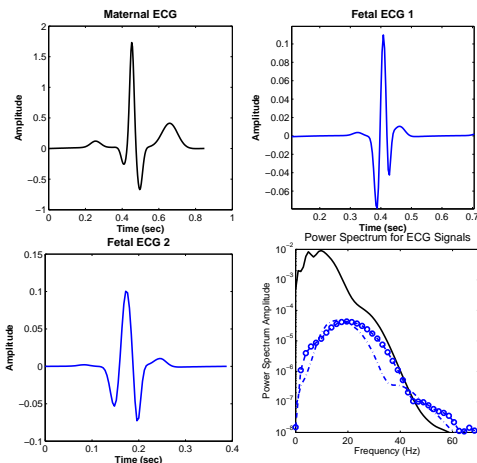


Figure 1: Simulated maternal and twin fetal ECG signals using parameters and values from Table 1 and 2. Power spectrum of the signals done using the Welch method.

For the data centric method, a robust estimate of the source probability density function (pdf) is needed. We do this by creating a normalized histogram of long length (10-15 minutes) ECG data and noise. The length of the data matrix is chosen such that the pdf shape converges. Fig 2 shows an example of the obtained density functions for sources and noises with both techniques. Note that the density functions look different across all sources motivating a data-centric method. We then fit a 3rd – 6th order polynomial to the contrast function and then use that information in the Fast-ICA routine (see section: ICA). Fig 3 shows the variation of the PI metric over the different schemes over data acquisition time (up to

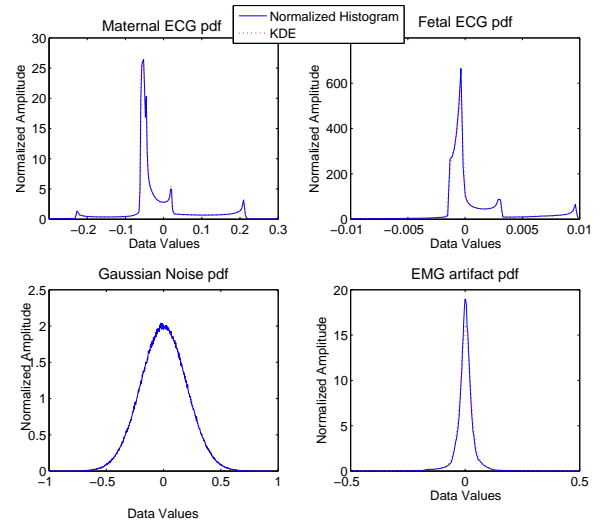


Figure 2: Probability density functions (pdfs) of the maternal, fetal signals, EMG artifact and gaussian noise, using normalized histogram and KDE methods.

500s) for noiseless (left) and noisy (right) data. The new data-centric schemes well surpasses the data based Pearson method (both noiseless and noisy) and performs on par and sometimes better than the standard polynomial schemes (such as POW3 and SKEW). Beyond 200s of data for noiseless and 400s for noisy channels, the PI metric stabilized indicating that a relatively long segment of data (maybe 10 mins) should be taken before ICA estimation for fetal reconstruction. PI metrics under 0.1 in this case offered a clear visual morphological reconstruction (example shown later). For increased heart rates, around 180 bpm, (see Fig 4), the PI metric increased in value due to errors in maternal ECG reconstruction. FECGs were however estimated properly (result not shown). Fig 5 shows the variation in mean PI metric over the different ICA schemes with closely similar twin fetal signals and heart rates for noisy channel data centered around 150 bpm and 180 bpm (both types common with cardiac defects). Closely similar fetal signal is a common occurrence with identical twins. Here we investigated the separation performance when fetal signals had a heart rate difference between 0 and 5 bpm. Note that the input signals have no rate variability. The data-centric method with 4th order performs the best for FECG1=150bpm but the POW3 method surpasses this slightly for a higher mean FECG1 of 180 bpm. We also noticed an overall shift in the metric with increase in heart rate indicating ICA reconstruction difficulties with the maternal ECG at these rates. Fig 6 is an example in-vivo data set experimented on. The MEGC was obtained from physionet from real ECG databases and the fetal signals were simulated (see section: simulation of twin data). EMG gaussian noise were also mixed into the channel. Real fetal signals are unavailable online and hence were simulated. The figure shows clean reconstruction with the polynomial method.

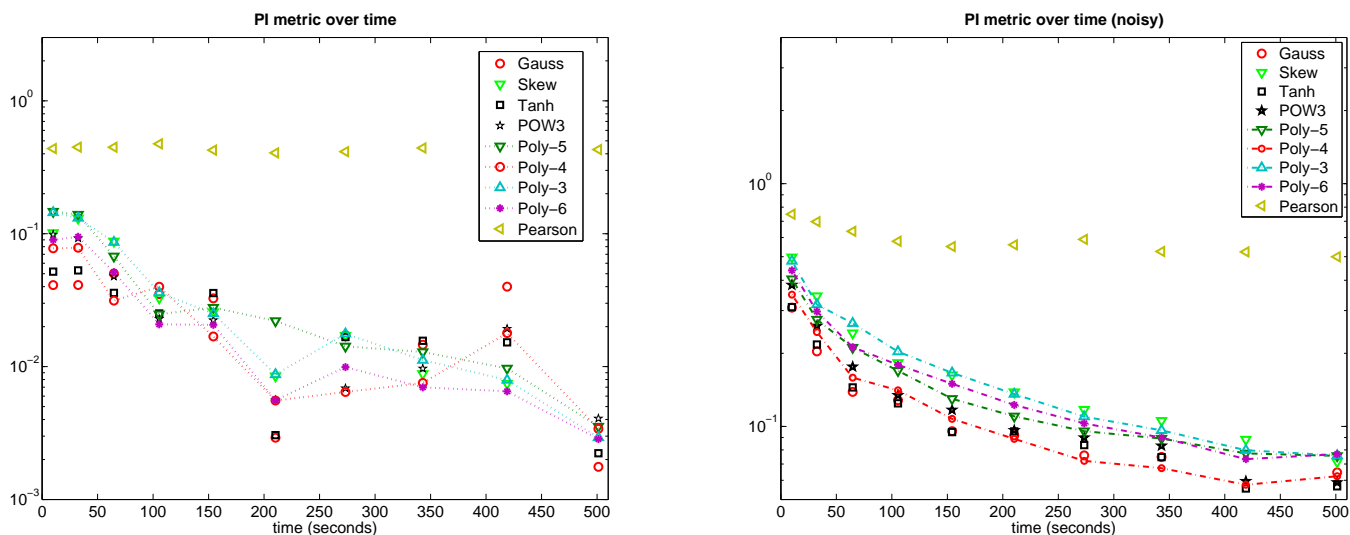


Figure 3: Variation in mean PI metric over time (10-500 sec) for the different Fast-ICA schemes and our proposed data-centric implementations using polynomial ($3^{rd} - 6^{th}$ order) for data with no noise (left) and with noise (right).

Further the figure also shows the PI metrics for all the techniques with the poly-4 performing as expected.

4 Summary and Conclusion

In this study, we investigated Fast-ICA standard contrast functions and formulated new data-centric polynomial based contrast functions to extract simulated twin FECG from abdominal ECG channel data, under different types of interference for purposes of twin fetal health assessment. We clearly show that the polynomial based contrast functions perform superior to the data based pearson method and works on par and in some cases better than standard ICA polynomial schemes like SKEW and POW3. We show a variety of simulations with and without noise and quantify reconstruction performance using a PI metric that is scale and order independent. PI metrics for Poly-4 were in many cases the lowest offering best performance. When estimation was done over data acquisition time, it was clear that adequate data was needed for good performance. In general a good rule of thumb may be acquiring 600s (or 10 min) of data before ICA estimation. The performance of the ICA schemes, however, degraded slightly with increased heart rates (180 bpm) possibly due to increase information content in the signal. With this scenario MECG reconstruction imperfections caused the increase in PI values. It was also clear from the simulations that equal fetal heart rates could not be separated, but a difference of 1 bpm was sufficient to significantly improve the performance for most techniques. Again poly-4 performed the best at FECG1=150bpm whereas POW3 was slightly better at FECG1=180bpm. While testing on the in-vivo data, similar trends were seen with the Poly-4 being on par

than the other ICA schemes. We believe that these data-centric schemes with ICA will be more powerful in true in-vivo situations when access to abdominal ECG data is available with twin fetal gestations.

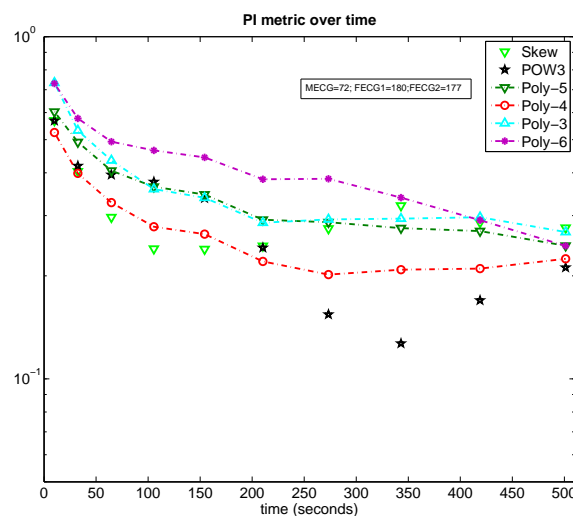


Figure 4: Variation in mean PI metric (eq. 5) over time (10-500 sec) for the different Fast-ICA schemes and our proposed data-centric implementations using polynomial ($3^{rd} - 6^{th}$ order) for data with increased heart rates (FECG1=180bpm; FECG2=177bpm).

References

[1] Lewis, M. J., "Review of electromagnetic source investigations of the fetal heart," *Medical Engineering and Physics*, V25, pp. 801-810, 2003.

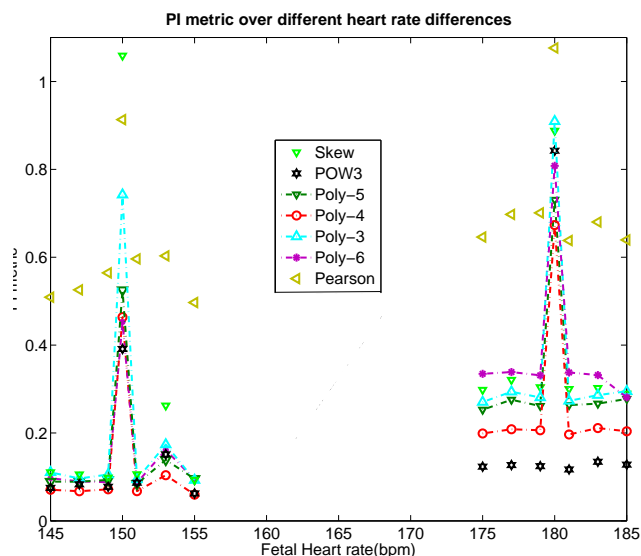


Figure 5: Variation in mean PI metric (eq. 5) over the different ICA schemes (standard and newly proposed) with closely similar twin fetal signals and heart rates for noisy channel data. The figure shows results using $FECG1=150\text{bpm}$ and $145 \leq FECG2 \leq 155$ and $FECG2=180\text{bpm}$ and $175 \leq FECG2 \leq 185$.

[2] Bahtiyar, M. O., Dulay, A. T., Weeks, B. P., Friedman, A. H., and Copel, J. A., "Prevalence of Congenital Heart Defects in Monochorionic/Diamniotic Twin Gestations," *Journal Ultrasound in Medicine*, V26, pp. 1491-1498, 2007.

[3] Comani, S., Mantini, D., Alleva, G., Gabriele, E., Liberati, M., and Romani, G. L., "Simultaneous monitoring of separate fetal magnetocardiographic signals in twin pregnancy," *Physiological Measurement*, V26, pp. 193-201, 2005.

[4] Martens, S. M., Rabotti, C., Misch, M., and Sluiter, R. J., "A robust fetal ECG detection method for abdominal recordings," *Physiological Measurement*, V28, pp. 373-388, 2007.

[5] Comon, P., and Jutten, C., *Handbook of Blind Source Separation, Independent Component Analysis and Applications*, 1st Edition, Academic Press, 2009.

[6] Gupta, A., Srivastava, M. C., Khandelwal, V., and Gupta, A., "A Novel approach to Fetal ECG Extraction and Enhancement Using Blind Source Separation (BSS-ICA) and Adaptive Fetal ECG Enhancer (AFE)," *IEEE ICICS 2007*, pp. 983-987 2007.

[7] Lathauwer, L. D., Moor, B., and Vandewalle, J., "Fetal Electrocardiogram Extraction by Blind Source Subspace Separation," *IEEE Transactions on Biomedical Engineering*, V47, N5, pp. 567-572, 2000.

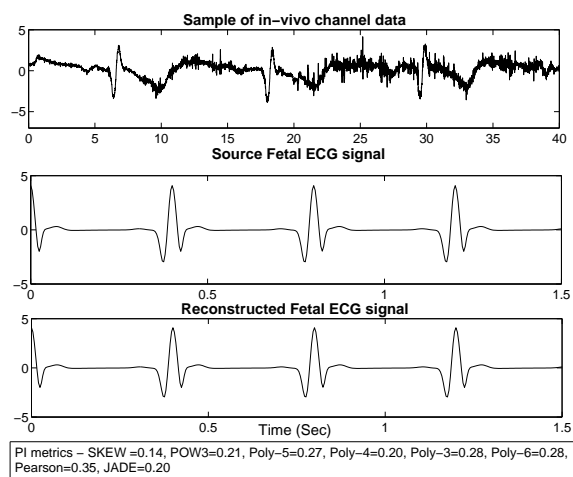


Figure 6: Top: Example of in-vivo based constructed channel data, Middle: source FECG1, Bottom: Reconstructed FECG1 using the polynomial method with order 4. The PI metrics for all techniques with this data set is also tabulated in the figure.

[8] Taylor, M. J., Smith, M. J., Thomas, M., Green, A. R., Cheng, F., Oseku-Afful, S., Wee, L. Y., Fisk, N. M. and Gardiner, H. M., "Non-invasive fetal electrocardiography in singleton and multiple pregnancies," *British Journal Obstetric Gynaecology*, V110, pp. 668-678 10/2003.

[9] Hyvarinen, A., Karhunen, J., and Oja, P., *Independent component analysis*, Wiley interscience, 2001.

[10] Karvanen, J., and Koivunen, V., "Blind separation methods based on pearson system and its extensions," *Signal Processing*, V82, pp. 663-673 2002.

[11] McSharry, P. E., Clifford, G. D., Tarassenko, J., and Smith, L. A., "Dynamical Model for Generating Synthetic Electrocardiogram Signals," *Signal Processing*, V50, N3, pp. 289-294 2003.

[12] Shepovalnikov, R. A., Nemirko, A. P., Kalinichenko, A. N., and Abramchenko, V. V., "Investigation of Time, Amplitude, and Frequency Parameters of a Direct Fetal ECG Signal during Labor and Delivery," *Pattern Recognition and Image Analysis*, V16, N1, pp. 74-76 2006.

[13] Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. Ch., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., Stanley, H. E., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation* V101, N23, pp. e215-e220 2000.