33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Framework for Adaptive Multiscale Analysis of Nonhomogeneous Point Processes

Hannes Helgason, Jay Bartroff, and Patrice Abry.

***Abstract*— We develop the methodology for hypothesis testing and model selection in nonhomogeneous Poisson processes, with an eye toward the application of modeling and variability detection in heart beat data. Modeling the process' non-constant rate function using templates of simple basis functions, we develop the generalized likelihood ratio statistic for a given template and a multiple testing scheme to model-select from a family of templates. A dynamic programming algorithm inspired by network flows is used to compute the maximum likelihood template in a multiscale manner. In a numerical example, the proposed procedure is nearly as powerful as the super-optimal procedures that know the true template size and true partition, respectively. Extensions to general history-dependent point processes is discussed.**

***Index Terms*— point processes, nonhomogeneous Poisson processes, nonparametric detection, dynamic programming, multiple comparisons, heart rate analysis**

I. Introduction

A variety of quite different systems produce data consisting of occurrences in time (or space) of well identified events, although not taking the classical form of continuous-time signals $X(t)$ or time series X_n. In medicine and biology, this is notably the case for heart beat rhythm data [1], [2], [3] and neuron discharge spike trains [3], [4]. Other examples from science and engineering are numerous: IP packet flows in internet traffic [5], [6]; optics, with light photon counts; geophysics, with stick-slip mechanisms or earthquake epicenter locations (space) and occurrences (time). The characterization of the systems producing such data relies on the analysis of these lists of occurrences which are often modeled mathematically as *point processes*. Their theoretical properties, as well as corresponding signal processing analysis tools, have been studied thoroughly over the past decades (see, e.g., [7], [8] and references therein).

In point processes, *stationarity* corresponds to the constancy along time (or space) of the rate of occurrences of events. Here we address the non-stationary situation using a non-homogeneous Poisson process model [8] with methodology based on the Generalized Likelihood Ratio (GLR). Our original approach is based on modeling the non-constant rate with a set of elementary functions, and finding the best approximations within this class with an efficient network flow algorithm. The obtained solution is data-driven (i.e., adaptive) and multiscale, as well as weakly dependent on arbitrary parameters. This method is detailed in Section II. In Section III, this approach is tested via Monte-Carlo simulations of synthetic data consisting of nonhomogeneous Poisson process and is shown to provide practitioners with a methodology to test rate constancy using an efficient algorithm. Section IV discusses applications to the analysis of heart beat intervals.

H. Helgason (corresponding author) is with SIP Laboratory, KTH, Stockholm, Sweden, hannes.helgason@ee.kth.se; J. Bartroff is with Department of Mathematics, USC, Los Angeles, USA, bartroff@usc.edu; P. Abry is with Laboratoire de Physique, ENS Lyon/CNRS, Lyon, France, patrice.abry@ens-lyon.fr

II. Methodology

Assume we observe data from a counting process $N(t)$ in a fixed time interval $(0, T]$, where T is a constant chosen *a priori*. That is, $N(t)$ is the number of events in the time interval $(0, t]$, where $N(0) = 0$ by convention. Denote the series of event times by $\{\tau_i : i = 1, \ldots, N(T)\}$, where the times are in increasing order: $\tau_i < \tau_{i+1}$ for all i. Let $f_m(t \mid \mathcal{H}_{m-1})$ be the conditional distribution of the m-th event given τ_i, $i = 0, \ldots, m-1$. Here $\mathcal{H}_{m-1}$ denotes the history of the process for the time period $(0, \tau_{m-1}]$.

The methodology that follows (and is motivated by ideas in [9], [10]) is based on using simple models to fit the point process locally to capture its local behavior (or short term history dependence). The global behavior, or long term history dependence, is handled through global optimization by chaining together local fits in a meaningful way. Although we will describe the framework in the context of nonhomogeneous Poisson processes, the underlying idea can be extended and tailored to other types of point processes, such as RR interval data in heart rate analysis (see discussion in Section IV).

Nonhomogeneous Poisson Processes. A nonhomogeneous Poisson process with a positive *rate function* $r(t)$ is a counting process $\{N(t), t \geq 0\}$, where $N(t)$ is the number of events in the time interval $(0, t]$ and satisfying: (i) $N(0) = 0$; (ii) for disjoint intervals $[t_0, t_1]$ and $[s_0, s_1]$, $N(t_1) - N(t_0)$ and $N(s_1) - N(s_0)$ are independent; (iii) the probability of observing m events in the interval $(t_0, t_1]$ is $P(N(t_1) - N(t_0) = m) = (m!)^{-1} e^{-R(t_0,t_1)} R(t_0, t_1)^m$, where $R(t_0, t_1) := \int_{t_0}^{t_1} r(t)dt$. That is, $N(t_1) - N(t_0)$, $0 \leq t_0 < t_1$, is distributed as $Poi(R(t_0, t_1))$.

It is easy to show that the conditional distribution of the m-th event given the history $\mathcal{H}_{m-1}$ is given by $f_m(t \mid \mathcal{H}_{m-1}) = e^{-R(\tau_{m-1},t)} r(t)$. This equation holds for $m = 1$ with the convention $\tau_0 = 0$. The joint probability density of observing events at the times $\tau_1, \ldots, \tau_{N(T)}$ in the interval $(0, T]$ can be expressed as (see, e.g., [7])

$$f_{1,\ldots,N(T)}(\tau_1, \ldots, \tau_{N(T)}) = e^{-R(0,T)} \prod_{m=1}^{N(T)} r(\tau_m). \quad (1)$$

Given data $\{\tau_n : n = 1, \ldots, N(T)\}$, we can use (1) to calculate the likelihood of different models of the rate function $r(t)$ and motivate choices for goodness-of-fit measures. The log-likelihood function for this data is

$$l(r|\tau_1, \ldots, \tau_N) = \sum_{m=1}^{N} \log(r(\tau_m)) - \int_0^T r(t)dt. \quad (2)$$

Assume the rate function $r(t)$ belongs to a given functional class $\mathcal{F}$ (possibly nonparametric, e.g., a class of smooth curves, etc.). In practice, the class $\mathcal{F}$ would typically be based on the prior knowledge one has about the system under consideration; for example, physiological restrictions in the case of biomedical applications.

Maximum Likelihood Principle. Given data $\{\tau_n : n = 1, \ldots, N\}$, the *maximum likelihood estimate* (MLE) of the rate function is the $\hat{r} \in \mathcal{F}$ maximizing the log-likelihood function in (2), i.e.,

$$\hat{r} = \arg\max_{r \in \mathcal{F}} l(r|\tau_1, \ldots, \tau_N). \quad (3)$$

Unless the functions in $\mathcal{F}$ have similar *complexity* there is a need for adding some some regularization to the functional in the optimization problem (3); otherwise there is a risk of overfitting the data, resulting in an estimator with large variance. Letting $\Lambda(r)$ be a functional measuring the complexity of r (i.e., $\Lambda(r)$ is large for "complex" r, in some sense), one could consider the restricted log-likelihood problem

$$\hat{r}_{\Lambda,\Lambda_0} = \arg\max_{r \in \mathcal{F},\ \Lambda(r) \le \Lambda_0} l(r|\tau_1, \ldots, \tau_N), \quad \Lambda_0 > 0. \quad (4)$$

In general, optimization problems of the type (3) and (4) can be very difficult or impossible to solve. Here we take the simplifying approach of approximating $\mathcal{F}$ with simplified sets $\widetilde{\mathcal{F}}$ of models which, if chosen correctly, can lead to a computationally tractable optimization problem that approximates the general problem.

Linear Combinations of Templates. Consider models $\widetilde{r}$ of the form

$$\widetilde{r}(t) = \widetilde{r}_V(t) = \sum_{i \in V} \alpha_i g_i(t), \quad (5)$$

where V is a finite set of indices, (α_i) is a sequence of scalars, and g_i is a *template* supported on a given interval I_i. For example, the templates g_i in (5) could be members of classic function bases (e.g., wavelets, B-splines, etc.) or piecewise linear functions.

Assume that the templates $(g_i)_{i \in V}$ in (5) have disjoint time supports I_i. Inserting $\widetilde{r}$ into (2) gives

$$l(\widetilde{r}|\tau_1, \ldots, \tau_N) = \sum_{m=1}^{N} \log(\widetilde{r}(\tau_m)) - \int_0^T \widetilde{r}(t)dt$$

$$= \sum_{i \in V} \left[n_i \log(\alpha_i) + \sum_{m:\ \tau_m \in I_i} \log(g_i(\tau_m)) - \alpha_i \int_{I_i} g_i(t)dt \right]$$

where $n_i = \#\{\tau_m \in I_i\}$ is the number of events occurring in the time interval I_i. As a measure of the goodness-of-fit of $\widetilde{r}$ to the data, we use the generalized likelihood principle, and maximize $l(\widetilde{r}|\tau_1, \ldots, \tau_N)$ with respect to the coefficients $(\alpha_i)_{i \in V}$. This leads to the unrestricted MLE $\widehat{\alpha}_i = n_i(\int_{I_i} g_i(t)dt)^{-1}$ and the generalized log-likelihood function

$$G(\widetilde{r}|\tau_1, \ldots, \tau_N) := \max_{(\alpha_i)_{i \in V}} l(\widetilde{r} \mid \tau_1, \ldots, \tau_N) = \sum_{i \in V} c(i)$$

$$\text{where} \quad c(i) := n_i \log(n_i) - n_i - n_i \log\left(\int_{I_i} g_i(t)dt \right) + \sum_{\tau_m \in I_i} \log(g_i(\tau_m)).$$

We interpret $c(i)$ as the *local fit* of the template g_i to the data.

III. Hypothesis Testing with Piecewise Constant Templates

As above, assume that we observe a nonhomogeneous Poisson process on a fixed, known interval $(0, T]$ with unknown rate function $r(t)$. We wish to test

$$H_0 : r(t) = \alpha, \qquad \text{versus} \qquad H_1 : r(t) = \alpha + \mu(t), \quad (6)$$

where $\alpha > 0$ is an unknown constant and $\mu(t)$ is a piecewise constant function with mean zero over $(0, T]$, not identically zero. Note that we assume nothing further about $\mu(t)$, e.g., the number and positions of the points where $\mu(t)$ changes levels and the levels themselves are all unknown.

GLR for Fixed Partition. Fix a partion $P = \{I_i\}_{i=1}^L$ of $(0, T]$ composed of left-open, right-closed intervals and suppose the rate function is of the form $r(t) = \sum_{i=1}^L \alpha_i 1_{I_i}(t)$, where α_i are unknown. Abusing our notation in (2), we write the log-likelihood function as $l(\alpha_1, \ldots, \alpha_L) = \sum_{i=1}^L (n_i \log \alpha_i - \alpha_i |I_i|)$, and note that $n_i \sim Poi(\alpha_i |I_i|)$ are independent. Then the unrestricted MLE is $\widehat{\alpha}_i = n_i/|I_i|$ and the H_0-restricted MLE is $\widetilde{\alpha} = N/T$, where $N = \sum_{i=1}^L n_i$. This gives the log-GLR statistic $S(P) := l(\widehat{\alpha}_1, \ldots, \widehat{\alpha}_L) - l(\widetilde{\alpha}, \ldots, \widetilde{\alpha})$ as

$$S(P) = \sum_{i=1}^{L} n_i \log(n_i/|I_i|) - N \log(N/T). \quad (7)$$

Under H_0 and assuming $\alpha \cdot \min_i |I_i|$ to be large, $n_i \approx N(\alpha|I_i|, \alpha|I_i|)$, and this can be used to show that $2S(P)$ is approximately χ^2_{L-1} under H_0, which is also observed in Monte Carlo simulations.

Model Selection and Algorithms. In practice one would typically not know the size L of the partition or the size $|I_i|$ of the intervals. Here we give an adaptive method and algorithm for estimating these.

Assume that $\Delta t > 0$ is the length of the smallest possible interval size in any partition under consideration (i.e., Δt sets the finest analysis scale) and assume that $T = M\Delta t$,

for some integer M. The partitions considered will be of the form $\{I_i\}_{i=1}^{L}$, $L \le M$, where

$$I_i = (k_i\,\Delta t, k_{i+1}\,\Delta t],$$

with $0 \le k_i < k_{i+1} \le M$, $0 \le i \le L-1$. Let $\mathcal{P}$ denote the finite set of partitions of $(0,T]$ constructed this way. First assume we have assigned a local cost $c(I)$ to each interval $I = (k'\,\Delta t, k\,\Delta t]$, $0 \le k' < k \le M$. The goal is to solve the optimization problems

$$C(L) = \max_{P\in\mathcal{P},\#P=L} \sum_{I\in P} c(I) \tag{8}$$

where the maximization is done over all partitions $P \in \mathcal{P}$ of size L, $1 \le L \le L_{\max} \le M$.

The following Algorithm 1 for finding the optimal partition is based on dynamic programming and ideas from the network flows [11] literature. Let $d_k(\ell)$ be a *distance label* which stores the cost of the tentative optimal partition of size ℓ, $0 \le \ell \le k$, for the interval $(0, k\Delta t]$. We also let $\text{pred}_k(\ell) = k'$ be the time index which precedes the index k in the tentative optimal partition of size ℓ of the interval $(0, k\Delta t]$. For each left-endpoint $t = k'\,\Delta t$, let $A_{k'}$ be the set of right-endpoint indices for intervals $(k'\,\Delta t, k\,\Delta t]$ under consideration. When it terminates, $d_k(\ell)$ consists of the optimal partition cost of $(0,,k\Delta t]$ using ℓ intervals and $\text{pred}_k(\ell)$ is the time index which precedes time index k for this partion. Thus, Algorithm 1 finds the best partitions of sizes $\ell = 1,\ldots,L_{\max}$ in just one sweep over the set of time indices $k' = 0,\ldots,M-1$.

Algorithm 1 Calculation of Optimal Partition

Initialize distance labels:
Set $d_0(\cdot) = 0$ and $d_k(\cdot) = -\infty$ for $k = 1,\ldots,M$
Update distance labels by marching forward in time:
for $k' = 0,\ldots,M-1$
  for all $k \in A_{k'}$ (the set of ks in intervals $(k'\,\Delta t, k\,\Delta t]$)
    set $I = (k'\,\Delta t, k\,\Delta t]$ and calculate $c(I)$
    for $\ell = 1,\ldots,L_{\max}$
      if $d_k(\ell) > d_{k'}(\ell-1) + c(I)$
        $d_k(\ell) = d_{k'}(\ell-1) + c(I)$
        $\text{pred}_k(\ell) = k'$.

The computational complexity is directly proportional to the number of subintervals of $(0,T]$ we consider. For example, if we include all intervals $(k'\,\Delta t, k\,\Delta t]$, $0 \le k' < k \le M$, we have $O(M^2) = O(T^2(\Delta t)^{-2})$ intervals. Note that since the algorithm marches forward in time, it has the potential to being adapted to online processing of data; this property is useful, for example, in heart rate monitoring. Also, since no type of "orthogonality" exists it is not possible to solve (8) using a greedy tree algorithm by subsequent splitting of branches. That is, if P_L^* is the optimal partition using L segments, then P_{L-1}^* is not necessarily retrieved by joining two intervals in P_L^*, i.e., pruning the tree.

Multiple Comparisons and Proposed Hypothesis Test. Given data $(\tau_1,\ldots,\tau_N)$, we calculate $C(L)$ for a number of different values of $L = L_1,\ldots,L_K$ and estimate the P-value P_k for $C(L_k)$, $k = 1,\ldots,K$, by Monte Carlo simulation under H_0. Since the intervals corresponding to different L_k overlap, $P_1,\ldots,P_K$ are dependent and, under H_0, not simply K independent Unif$[0,1]$ variables as in the usual univariate case. Under H_1, we would expect to observe low P-values for the coordinates of C corresponding to models whose complexity is close to the one of the true $\mu(t)$. Therefore it seems reasonable to look for evidence against H_0 by considering the minimum observed P-value $P^* := \min(P_1,...,P_K)$, whose distribution under H_0 – and hence critical value – can be estimated by Monte Carlo.

Numerical Experiment. As a numerical exampe, we consider the testing problem (6) and use Monte Carlo to evaluate our proposed test. We use a standard synthesis method based on time rescaling to generate point processes on the time interval $(0,T]$ (see, e.g., [7]). In all the simulations we take $T = 1024$ and $\alpha = 1$. For our statistic C we take L_k for $k = 1,\ldots,7 = K$ as $2,3,4,8,16,32,64$, respectively; this choice covers a large range of segmentation sizes Ls using few values (considering the full range $L = 1,\ldots,2^6$ imposes difficulties in using Monte Carlo simulations for simulating the joint distribution of C). The smallest time interval considered is taken to be $\Delta t = T/2^9 = 2$.

For comparison we consider two other statistics: (i) $C(L)$ in (8), where the true number of partitions L under H_1 is assumed to be known; (ii) $S(P)$ in (7) which knows the true partition P under H_1 exactly. For simulations under H_1, we consider two examples with $\beta > 0$:

(a) $\mu_1(t) = \beta(1_{[0,T/2]}(t) - 1_{(T/2,T]}(t))$;
(b) $\mu_2(t) = \beta(1_{[0,T/2]}(t) - 3\cdot 1_{(T/2,3T/4]}(t) + 1_{(3T/4,T]}(t))$.

The example $\mu_1(t)$ describes a change-point problem where the system being monitored changes state in the middle of the observation, while $\mu_2(t)$ models a sudden "dip" in rate and then returns to the previous rate. In the context of fetal heart rate monitoring, these examples can be thought of as ideal models of tachycardia and a deceleration [12], respectively.

The *mean variance* $\overline{\mu^2} := T^{-1}\int_0^T \mu^2(t)dt$ gives a measure of $\mu(t)$ from constancy: the smaller $\overline{\mu^2}$, the more difficult H_1 will be to detect. In the examples above, the mean variance is $\overline{\mu_1^2} = \beta^2$ for μ_1, and $\overline{\mu_2^2} = 3\beta^2$ for μ_2, both proportional to β^2.

The number of Monte Carlo realizations under H_0 were 10,000 for each statistic. For the alternatives, the number of realizations were 1,000 for each different value of $\overline{\mu^2}$. Monte Carlo was performed under H_0 to choose thresholds so that the Type I error probability was 5%. Fig. 1 shows the detection rates, i.e., the power, as the mean variance $\overline{\mu^2}$ is varied. We see that the price paid for being adaptive is low: to reach detection rate of 95%, the proposed multivariate statistic needs the mean variance $\overline{\mu^2}$ to be about 10% greater than for $C(L)$ where L is taken to be known; the ratio is about 30% when compared to the statistic which knows the

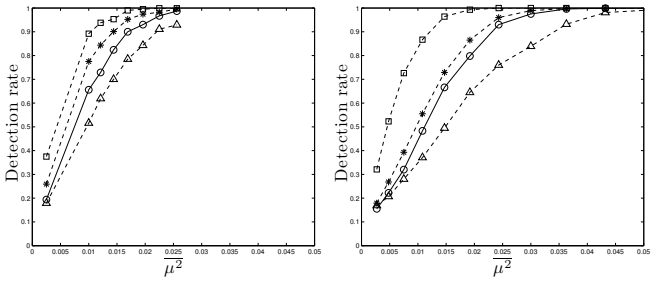

Fig. 1. Detection rates vs. $\overline{\mu^2}$ for P(Type I error) $= 5\%$. Left: $r(t) = \alpha + \mu_1(t)$. Right: $r(t) = \alpha + \mu_2(t)$. Labels: '$\square$' for $S(P)$ from (7) with partition P known; '$*$' for $C(L)$ in (8) with $L = 2$; '$\circ$' for P^* statistic based on $C = (C(L_1), \ldots, C(L_K)$; '$\triangle$' corresponds to $C(8)$ for μ_1 and to $C(2)$ for μ_2.

segmentation exactly. Fig. 1 also shows that the proposed model selection is helpful: the detection rates using $C(8)$ – an overly complicated model for the one-step alternative μ_1 – are worse than for the proposed test; using $C(2)$ when the truth is μ_2 (so that "correct" model corresponds to $L = 3$) is also not as good as the more adaptive test.

We observed that in the critical range $\overline{\mu_2^2} \in [0.025, 0.045]$ (break-down of power for $\mu_2(t)$), the minimum P-value in our proposed statistic was achieved at the "correct" coordinate $C(L_k = 3)$ in around 60–70% of the realizations. For $\mu_1(t)$, the proportion of realizations where the minimum P-value was achieved at $C(L_k = 2)$ rose steadily from 70% to almost 100% as $\overline{\mu_1^2}$ was increased from 0.01 to 0.025.

IV. Conclusions and Perspectives

Point processes that are history dependent make modeling using Poisson processes inappropriate. Much real data of interest violates the independence property of Poisson processes. Nevertheless, the above framework can be adapted to this setting by using the same guiding philosophy: (i) analyze local behavior of the point process through *local fits* $c(i)$ of a multiscale set of templates g_i; (ii) analyze global behavior by "chaining" together local fits.

Define $y_m = (\Delta\tau_m)^{-1}$, $\Delta\tau_m =$, $m \geq 1$, $\tau_0 = 0$, and consider the time series (τ_m, y_m); this type of transformation of point process into time series is very popular in processing RR-intervals in heart rate analysis. Instead of modeling the inter-arrival times $\Delta\tau_m$ or the conditional distribution of the m-th event, one could consider modeling y_m. One possibility is to consider *additive noise models*

$$y_m = \mu_m + z_m, \qquad m \geq 1, \tag{9}$$

where $\mu_m = \mu(\tau_m)$ and μ is some unknown function; (z_m) is a sequence of possibly correlated random variables. The problem is now to learn about μ and z_m given the observations (τ_m, y_m), and this is a classical function estimation problem involving time series. Note that the time $\Delta\tau_m$ between samples is uneven, so many popular signal processing tools cannot be directly applied to the problem (e.g., Fourier analysis and wavelet based methods). In the time series model (9), the function μ could be interpreted as the *trend*, responsible for the underlying "large scale structure" of the process. The sequence (z_m) can be thought of as *variability*, or "small scale fluctuations," of the process.

The model (5) can be generalized by taking $\alpha_i = 1$ but allowing g_i to include unknown parameters to be estimated. The log-likelihood is still a sum $\sum_{i \in V} c(i)$ of local fits but where $c(i) = \sum_{\tau_m \in I_i} \log(g_i(\tau_m)) - \int_{I_i} g_i(t)dt$. Models of this form can be useful when one wants to impose some sort of regularity condition like *continuity*. For example, suppose the class of rate functions $\mathcal{F}$ can be well approximated by near-continuous piecewise linear functions. Then a discrete set of linear templates $g_i(t) = (at + b)1_I(t)$ would be appropriate. Models of this form have been considered in the context of time series data in [10] for oscillatory signals and [12] for fetal heart rate times series, and the framework and algorithms therein are therefore extendible to the analysis of point processes by replacing the local fits.

In future works we will investigate the proposed framework by adapting local fits to models of type (9) for applications to heart rate analysis. There we would like to compare the methodology with other appealing approaches, such as those developed in [13], which are based on time-varying autoregressive modeling.

References

[1] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. C. Berger, and R. J. Cohen, "Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control," *Science*, vol. 213, no. 4504, pp. 220–222, 1981.

[2] A. L. Goldberger, L. A. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and S. H. E., "Fractal dynamics in physiology: alterations with disease and aging," *Proc. Natl. Acad. Sci. USA*, vol. 99, no. Suppl 1, pp. 2466–2472, 2002.

[3] S. B. Lowen and M. C. Teich, *Fractal-based Point Processes*. Wiley Series in Probability and Statistics, 2005.

[4] P. Abry and P. Flandrin, "Point processes, long-range dependence and wavelets," *Wavelet Applications in Medecine and Biology, A. Aldroubi, M. Unser, eds.,*, vol. CRC Press, pp. 413–437, 1996.

[5] B. K. Ryu and S. B. Lowen, "Point process models for self-similar network traffic, with applications," *Stochastic models*, vol. 14, no. 3, pp. 735–761, 1998.

[6] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, "Multiscale network traffic analysis, modeling, and inference using wavelets, multifractals, and cascades," *IEEE Signal Processing Magazine*, vol. 3, no. 19, pp. 28–46, May 2002.

[7] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*, 2nd ed. Springer–Verlag, 2003.

[8] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*. Springer–Verlag, 1991.

[9] D. L. Donoho and X. Huo, "Beamlets and multiscale image analysis," in *Multiscale and multiresolution methods*, ser. Lect. Notes Comput. Sci. Eng. Berlin: Springer–Verlag, 2002, vol. 20, pp. 149–196.

[10] E. J. Candès, P. Charlton, and H. Helgason, "Detecting highly oscillatory signals by chirplet path pursuit," *Appl. Comput. Harmon. Anal.*, vol. 24, pp. 14–40, 2008.

[11] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network flows. Theory, algorithms and applications*. Prentice-Hall, New York, 1993.

[12] H. Helgason, P. Abry, P. Gonçalvès, C. Gharib, P. Gaucherand, and M. Doret, "Adaptive multiscale complexity analysis of fetal heart rate," *IEEE Transactions on Biomedical Engineering*, 2011, To appear.

[13] R. Barbieri, E. C. Matten, A. A. Alabi, and E. M. Brown, "A point-process model of human heartbeat intervals: new definitions of heart rate and heart rate variability," *Am. J. Physiol Heart Circ. Physiol.*, vol. 288, pp. H424–H435, 2005.