18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Detecting Trend in Randomly Switched Measurements

Dayu Huang, Marco Guerriero, and Xing Wang
General Electric Global Research Center
Niskayuna, NY 12309
Email: huang, marco.guerriero, x.wang@ge.com

***Abstract*—We introduce a new trend detection problem inspired by real-time monitoring applications where the origin of the measurements is uncertain: The observed sequence under the alternative hypothesis is the result of a random switching between two sequences, each with a trend. The association between each measurement sample and the two sequences is unknown to the detector.**

We propose a Generalized Mann-Kendall trend detection algorithm, and show via simulation that it achieves better performance than the Mann-Kendall algorithm for problems with randomly switched measurements.

We show that the test statistic can be calculated using an Mixed Integer Linear Programming (MILP) solver. We also show that computing the Generalized Mann-Kendall test statistic can be cast as a Max-Bisection problem, connecting the computation of test statistics to graph optimization.

***Index Terms*—Trend detection, Mann-Kendall test, randomly switched measurement, graph theory, switched system, Max-Bisection, mixed integer programming.**

I. Introduction

The goal of detecting a trend is to determine whether the values of a random variable generally increase (or decrease) over some period of time in statistical terms [1]. One of the earliest and most widely-used trend detection methods is the Mann-Kendall test [2], [3]. It is the omnipotent non-parametric trend test of choice [4]. As a non-parametric test, it makes no assumption on the probability distribution of each data point. Compared to the parametric tests which typically assumes a known distribution and homoscedasticity (homogeneous finite variance), Mann Kendall test's performance is significantly better when the data is not normally distributed. Yet for normally distributed data, Mann Kendall test's performance is still very close to that of parametric tests [5]. Moreover, it is robust against outliers because its test statistic is based on the sign of differences. These advantages make it a universally applicable test. Mann-Kendall test has received a lot of research attention, as evidenced by more than 2000 citations to Mann and Kendall's work. Many extensions have been proposed to accommodate other practical issues such as seasonality and autocorrelation (see, e.g., [6], [7]). It is a basic component of many analytic tools such as R, and commonly used in many applications such as ecology studies, hydrology, physiology, market analysis, social media and industrial diagnostic applications [8], [9], [10], [11]. For the new trend detection problem with random switched measurement introduced in this paper, we restrict our attention to extensions of Mann-Kendall test.

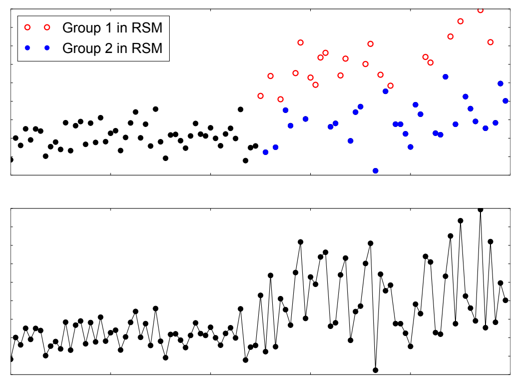

Fig. 1. Notional sketch (top panel) of the addressed problem with the RSMs marked by red circles and blue dots. However, with no access to the state information, one is left with the data points as in the bottom panel.

Trend detection in time-series signals will be automated in future real-time monitoring applications. In many such applications, signals might be affected or corrupted by other unmeasured events. The trend detection implementation should be made robust to these corruptions, eliminating the need for human intervention. In this paper, we focus on a particular issue which we call Randomly Switched Measurements (RSMs). As an example of RSMs, in the event of a failure, a machine could be randomly switching between alternative discrete operating modes. In another example, a machine could be operated by different operators and exhibit varying behavior between shifts. In the top panel of Fig. 1 we depict a notional example where red (circle) and blue (filled-in dot) samples are measurements corresponding to two different states. Both red and blue sequences exhibit an increasing trend starting at time $t = 50$. As shown in the bottom panel of Fig. 1, when the state information is not available, the trend might not be as visually significant as it was with the state information. As we will show in this paper, instead of applying the traditional Mann-Kendall test on the entire time series, one could achieve better detection performance by simutaneously estimating the state and computing Mann-Kendall test statistic on the sequence for each state.

A. Related work

The RSM trend detection framework in this paper is related to the areas of switched system [12], target tracking [13] and correlation clustering [14].

The RSMs could be generated by a system that exhibits switching between several subsystems [15]. Estimation of the state could be achieved if one uses a detailed control model under each different discrete state [16]. However, such a model might not be available in real-time monitoring applications considered in this paper, especially in the event of an anomaly. This is the reason why the use of non-parametric tests in this scenario is probably more effective.

In the context of target tracking, the two sequences with a trend could be recast as two target states where it is unknown to the tracker which measurement comes from which target. This is the data association problem [17]. However, for this data association approach to work, a probabilistic *parametric* measurement model is needed.

The proposed RSM framework is also related to the correlation clustering literature in which the data points are clustered based on their correlation [14]: The proposed framework groups the data points based on their alignment on a potential trend, introducing a new type of "correlation" between data points.

In this paper, the computation of the proposed test statistic is mapped to a graph optimization problem, making it possible to leverage a large body of research and software. A different way to represent Kendall's tau statistic using a graph is applied in [18], where Kendall's tau is treated as the cost of the associated permutation and the cost is defined based on the functional digraph of permutations.

B. Summary of results and contributions

The main contributions of this work are threefold:

1) A new mathematical formulation of the trend detection problem with RSM is proposed.
2) The Generalized Mann-Kendall test is proposed to address the RSM trend detection problem, and shown to perform better than the Mann Kendall test.
3) A connection is made between test statistic computation and graph optimization. It is shown that Generalized Mann-Kendall test can be cast as a graph Max-Bisection problem.

The paper is organized as follows: The mathematical formulation of the RSM trend detection problem is introduced in Section II. The Mann-Kendall test is presented in Section III. The Generalized Mann-Kendall test to address RSM is proposed in Section IV. We describe the algorithm to compute the test statistic using Mixed Integer Linear Programming (MILP) in Section V. The numerical results on the test performance are given in Section VI. The connection to Max-Bisection problem is given in Section VII. A few directions to extend this work are discussed in Section VIII and the paper is concluded in Section IX.

II. Problem Statement

The RSM trend detection problem in the case of increasing trend can be described using the following hypothesis testing formulation: Consider a real-valued random vector $\mathbf{Y}^n = \{Y_1, \ldots, Y_n\}$ of length n with independent elements. There are two composite hypotheses:

a) The null hypothesis H_0: The distribution of the random vector satisfies:

$$\text{Prob}\{Y_j \geq Y_i\} = 0.5, \text{ where time } j > i. \quad (1)$$

b) The alternative hypothesis H_1: The observation $\mathbf{Y}^n$ is associated with an i.i.d. Bernoulli random vector $\mathbf{U}^n = \{U_1, \ldots, U_n\}$, with parameter $p := \text{Prob}\{U_i = 1\} = 0.5$, which takes value in $\{0, 1\}^n$. The vector $\mathbf{U}^n$ divde the data into two sub-sequences: $\mathbf{Y}^{(0)} = \{Y_i : U_i = 0\}$ and $\mathbf{Y}^{(1)} = \{Y_i : U_i = 1\}$. Under H_1, for the data points in the same group has an increasing trend.

$$\text{Prob}\{Y_j \geq Y_i\} > 0.5, \text{ when time } j > i \text{ and } U_i = U_j. \quad (2)$$

The relationship between Y_j and Y_i when $U_i \neq U_j$ is arbitrary.

The value of $\mathbf{U}^n$ is *unknown* to the detection algorithm: A detection algorithm is a binary-valued function ϕ of the sequence $\mathbf{Y}^n$, where $\phi(\mathbf{Y}^n) = 1$ indicates that the algorithm decides in favor of H_1, i.e., there is a trend.

In other words, under the alternative hypothesis, the measurement $\mathbf{Y}^n$ is a sequence obtained from randomly switching between two vectors: $\mathbf{Y}^{(0)}$ and $\mathbf{Y}^{(1)}$, of which each has an upward trend. But the state of the switch at each time is unknown.

III. Mann-Kendall test

The Mann-Kendall test statistic [19] is defined as:

$$T^{\mathsf{MK}} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathsf{sign}(Y_j - Y_i), \quad (3)$$

where the value of $\mathsf{sign}(x)$ is 1 when $x > 0$, 0 when $x = 0$ and -1 when $x < 0$. The Mann-Kendall test is given by $\phi^{\mathsf{MK}}(\mathbf{Y}^n) = \mathbf{I}(T^{\mathsf{MK}} \geq \eta)$ where $\mathbf{I}$ is the indicator function.

The Mann-Kendall statistic is calculated through pair-wise comparisons of each data point with all preceding data points, and determining the number of increases, decreases, and ties. A positive value for T^{MK} implies an upward or increasing temporal trend, whereas a negative value implies a downward or decreasing trend. A value of T^{MK} near zero suggests there is no significant upward or downward trend.

IV. Generalized Mann-Kendall Test

Under the alternative hypothesis, each data point comes from one of two possible states, and the data sequence for each state has an increasing trend. If the state U_i of each data Y_i is known (we refer to this case as "Oracle"), then the Mann-Kendall test should be applied to each sequence separately and then combined as in [20], [21]. Mathematically, we would have:

$$T^{\mathsf{Oracle}} = \sum_{i:U_i=0} \sum_{j:U_j=0,j>i} \mathsf{sign}(Y_j - Y_i) + \sum_{i:U_i=1} \sum_{j:U_j=1,j>i} \mathsf{sign}(Y_j - Y_i) \quad (4)$$

When the state is unknown we need to estimate it. Following the Generalized Likelihood Ratio Test principle [22], we extend the Mann-Kendall test to jointly detect the trend and estimate the state for each data point: We call the new test the Generalized Mann-Kendall (GMK) test.

In the GMK test, the test statistic T^{GMK} is given by the optimal value of the objective function related to the following optimization problem:

$$T^{\mathsf{GMK}} = \max_{\mathbf{U}^n:|\{i:U_i=1\}|=\lfloor n/2 \rfloor} \Big\{ \sum_{i:U_i=0} \sum_{j:U_j=0,j>i} \mathsf{sign}(Y_j - Y_i) + \sum_{i:U_i=1} \sum_{j:U_j=1,j>i} \mathsf{sign}(Y_j - Y_i) \Big\}, \quad (5)$$

where $\lfloor x \rfloor$ is the largest integer not greater than x and $|\{i : U_i = 1\}|$ is the number of U_is equal to one. Note that the vector $\mathbf{U}^n$ is the optimization variable to be solved. The GMK test is given by

$$\phi^{\mathsf{GMK}}(\mathbf{Y}^n) = \mathbf{I}(T^{\mathsf{GMK}} \geq \eta). \quad (6)$$

V. Algorithm

In this section, we show that GMK test statistic has an equivalent Mixed Integer Linear Programming (MILP) formulation [23], and thus can be solved using an MILP solver. Many commercial and open source solvers are available for MILP. In our numerical experiment, we use the open source package PULP in the COIN-OR project [24].

The following notations are used when defining the equivalent MILP problem:

1) $w_{i,j} = \mathsf{sign}(Y_j - Y_i)$ is defined over all i, j for $1 \leq i < j \leq n$.
2) $u \in \Re^n$ and will be restricted to $\{0, 1\}^n$ in the MILP formulation.

The equivalent MILP problem is given by:

$$\begin{array}{ll} \max & \sum_{i,j:w_{i,j}>0} w_{i,j}(1 - |u_i - u_j|) \\ & + \sum_{i,j:w_{i,j}<0} w_{i,j}(|u_i + u_j - 1|), \\ \text{s.t.} & \sum_i u_i = \lfloor n/2 \rfloor, \\ & u \in \{0,1\}^n. \end{array} \quad (7)$$

where we economized the notation by using a single summation sign $\sum_{i,j}$ in the objective function. The sign $\sum_{i,j:w_{i,j}>0}$ means that the summation is over all i, j satisfying $1 \leq i < j \leq n$ and $w_{i,j} > 0$.

Proposition 5.1: The MILP problem (7) is equivalent to the GMK optimization (5).

The proof is given in the Appendix.

A solver might not return the optimal solution to (7). In fact, as we will show in Section VII, the optimization in (5) can be cast into a Max-Bisection problem which is in general NP-hard. Although it remains to be seen whether the structures imposed by the GMK formulation can be leveraged to change the computaional complexity of the problem. For these problems, there are *approximation algorithms* with performance guarantees. An algorithm is a $\alpha-$approximation algorithm for a problem if and only if for every instance of the problem it can find a solution within a factor α of the optimal solution [25].

Proposition 5.2: Let F_n be the cumulative distribution function of the test statistic T^{GMK} under the null hypothesis H_0, assuming that optimal solution is used. Let η be the threshold to achieve a certain probability of detection P_D and probability of false alarm P_F. Suppose an approximation algorithm with relative performance ratio α for (5) is used. Then the test based on the approximation algorithm using the threshold $\alpha\eta$ has the following detection performances measured in terms of P'_D and P'_F:

$$P'_D \geq P_D, \quad (8)$$

and

$$P'_F \leq 1 - F_n\big(\alpha F_n^{-1}(1 - P_F))\big), \quad (9)$$

where the right-hand-side is greater than or equal to P_F provided that $\alpha \leq 1$ and it converges to P_F as $\alpha \to 1$.

The proof is given in the Appendix.

VI. Numerical Results

We perform the following numerical experiment to demonstrate the performance of the Generalized Mann-Kendall test: Under the null hypothesis, the simulated data is given by

$$Y_t = N_t,$$

where N_t is a sequence of independent, identically distributed (i.i.d.) Gaussian random variables $\mathcal{N}(0, \sigma^2)$. Under the alternative hypothesis, the data Y_t is generated according the following model:

$$Y_t = \beta_1 t + U_t \beta_2 + N_t,$$

where N_t is a sequence of i.i.d. Gaussian random variables $\mathcal{N}(0, \sigma^2)$, U_t is sequence of i.i.d. Bernoulli random variables with $p = 0.5$ and $t = 1, \ldots, 10$.

We compare the performances of the following tests:

1) Mann Kendall test;
2) GMK test, with the test statistic given by the optimal solution to (5) calculated using an exhaustive enumeration over all feasible solutions;
3) GMK test calculated using the solver PULP to the MILP formulation (7);
4) Oracle: Mann Kendall test when the state information $\mathbf{U}^n$ is known to the detector. The Oracle case is only presented as an upper-bound on the best possible performance since it requires information that is not available.

The performances,of these tests in terms of Receiver Operating Characteristic (ROC) curve are depicted in Figure 2. The proposed GMK test achieves larger probability detection than Mann Kendall test for the same probability of false alarm.

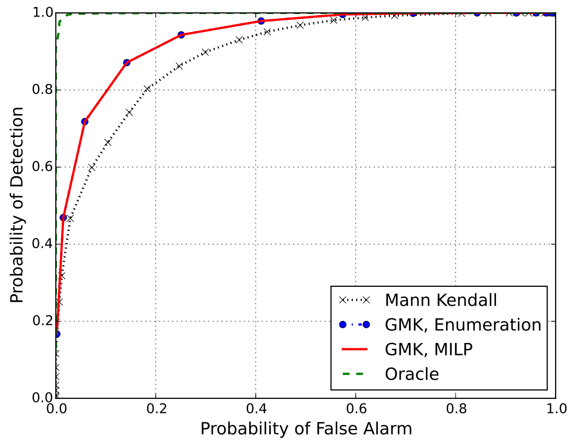

Fig. 2. Performance of Mann-Kendall test, GMK test with enumeration and MILP. $n = 10$, $\beta_1 = 0.08$, $\beta_2 = -1$, $\sigma = 0.1$. Note that the GMK with enumeration and MILP have the same performance

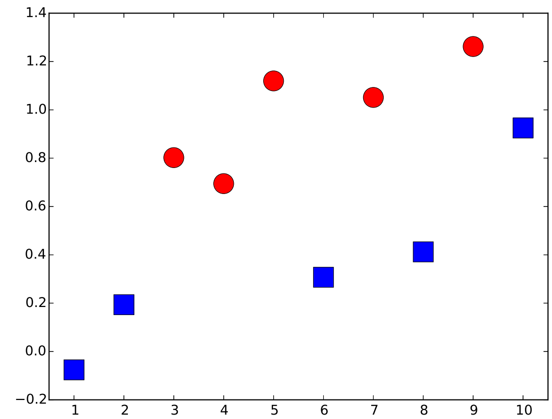

Fig. 4. Observations Y^n with $n = 10$. The state for a red round observation is 1 and the state for a blue square observation is 0.

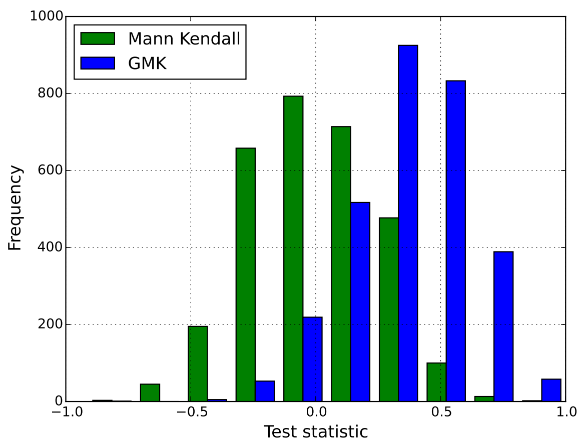

Fig. 3. Null distribution of Mann-Kendall and GMK test statistic for $n = 10$ samples.

There is a gap between the performance of the Oracle case and our proposed GMK test.

The threshold to achieve a certain probability of false alarm can be determined numerically via Monte-Carlo simulations. The distributions of the Mann-Kendall test and the GMK test statistics under the null hypothesis H_0, are shown in Figure 3. It is observed that the GMK test statistic has a nonzero mean under the null hypothesis while the Mann Kendall test statistic is known to have a zero mean under the null hypothesis [19]. The analytical characterization of this null distribution will be investigated in future work.

VII. Graph Representation and Max-Bisection

In this section, we show that the optimization problem in (5) can be cast into a well-known graph problem called Max-Bisection, which is a constrained version of the Max-Cut problem [26]. This connection leads to alternative computational algorithms to solve (5). In addition, other non-parametric test such as the Theil-Sen slope estimator [27], [28] can be defined on a graph by adjusting the weights of the graph edges.

Consider a weighted complete undirected graph $G = (V, E)$ with n vertices where $V = \{1, \ldots, n\}$ is the set of vertices and the edge set $E = \{(ij) : 1 \leq i, j \leq n\}$ is the set of *unordered* pairs of indices (ij). Let $w(ij)$ denote weight of the edge $(ij) \in E$. The assignment of edge weight for Mann-Kendall is given as follows:

$$w(ij) = \mathsf{sign}(Y_j - Y_i)\mathsf{sign}(j - i). \qquad (10)$$

The toy example shown in Figure 4 illustrates a 10-point sequence with two states. Its corresponding graph representation depicted in Figure 5.

The Mann Kendall test statistic can be written as the total weight of all edges over the graph:

$$T^{\mathsf{MK}} = \sum_{(ij) \in E} w(ij).$$

Similarly, the GMK test statistic can be represented as

$$T^{\mathsf{GMK}} = \max_{S \subseteq V : |S| = \lfloor \frac{n}{2} \rfloor} \sum_{(ij) : i,j \in S} w(ij) + \sum_{(ij) : i,j \in \bar{S}} w(ij), \quad (11)$$

where S is a subset of V and $\bar{S} := V \setminus S$. In other words, the GMK test statistic is the maximum of the sum of the total weight of edges in the subgraphs induced by S and $\bar{S}$. The corresponding graph with the optimal S^{opt} that solves (11) is depicted in Figure 6. Only edges within the two subgraphs are depicted. A suboptimal $S^{sub-opt}$, which might be obtained by randomly cutting the graph, would lead to more purple edges as depicted in Figure 7.

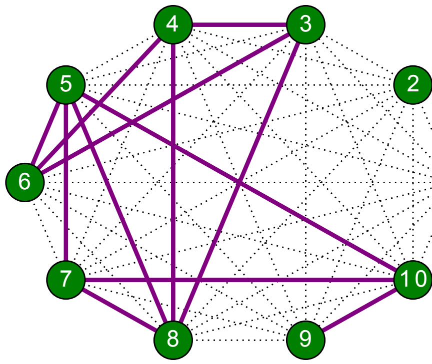

Fig. 5. Original graph for the Mann-Kendall test: A solid purple line is an edge with a *negative* weight $w(ij)$. A dotted black line is an edge with a *positive* weight $w(ij)$.

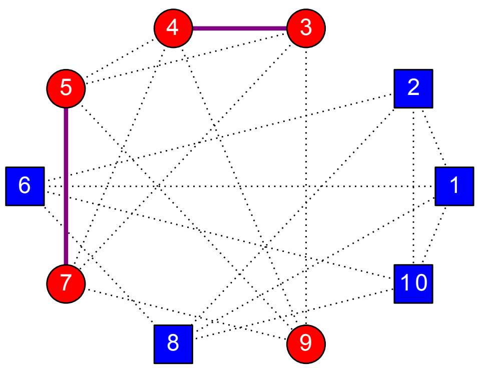

Fig. 6. Graph for GMK / Max-Bisection: The vertices are partitioned into two subsets represented by blue square nodes and red round nodes. Only edges of the two induced subgraphs are shown since those are the weights counted towards the summation in GMK test statistic. The partition is calculated from GMK, and it correctly recovers the actual states of the observations. Comparing to Figure 5, the number of solid purple lines is significantly reduced.

A cut of a graph is a partition of the vertices of a graph into two disjoint subsets S and $\bar{S}$. The cut set is the set of edges crossing the two sub-graphs:

$$\mathcal{E}(S, \bar{S}) = \{(ij) : i \in S, j \in \bar{S}\}.$$

The objective of Max-Bisection problem is to find a bi-partition which maximizes the number of crossing edges with the constraint that the cardinality of the two sub-graphs is the same. For a graph whose edge weight is denoted by $w'(ij)$,

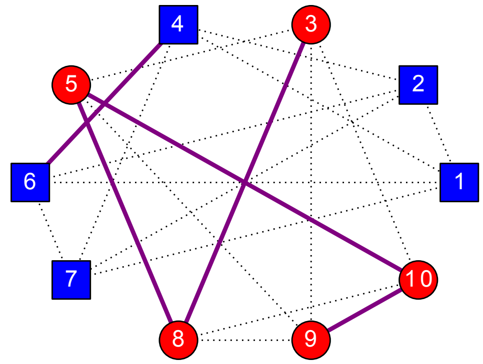

Fig. 7. Graph after a sub-optimal cut: Comparing to Figure 6, there are more solid purple lines left after the suboptimal cut.

the Max-Bisection is the following optimization problem:

$$T^{\text{Bisection}} = \max_{S \subseteq V : |S| = \lfloor \frac{n}{2} \rfloor} \sum_{(ij) \in \mathcal{E}(S, \bar{S})} w'(ij) \qquad (12)$$

We now show that our proposed problem (5) can be cast into the Max-Bisection problem (12) for a different but related graph.

Theorem 7.1: The GMK test statistic can be written as a linear function of the Max-Bisection problem's objective function value given in $T^{\text{Bisection}}$, where the weight $w'(ij)$ is defined as

$$w'(ij) = -0.5w(ij) + 0.5. \qquad (13)$$

The relationship is given by

$$T^{\text{GMK}} = T^{\text{MK}} + 2T^{\text{Bisection}} - \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil. \qquad (14)$$

The proof is given in the Appendix.

The general Max-Bisection problem is NP-hard [29]. There is a large body of work to find approximation algorithms to this problem (See [26] for references).

VIII. Extensions

In this section, we discuss several possible extensions of the proposal GMK test.

A. General switching probability p

The test described in this paper only addresses the case $p = 0.5$, i.e., under the alternative hypothesis the probability that the observation comes from either state is the same. This leads to the cardinality constraint in the Max-Bisection problem and the equality constraint in (7). For other known p, following the current framework, we can still formulate the problem as a max-cut problem with a different cardinality constraint or a different constant on the right-hand-side of the equality constraint in (7). The main challenge is to address how the test statistic from the two Mann-Kendall tests should be combined: The asymptotic distribution of the Mann-Kendall test under the null hypothesis is shown to be a Gaussian of which the variance is dependent on the number of samples in the test.

B. Mixed and Multiple Trends

We only address the case where there are two possible states, and the measurements in both states have an increasing trend under the alternative hypothesis. The test can be extended to the case where it is an increasing trend for one state and a decreasing trend for the other. The Generalized Mann-Kendall test statistic for this case is given by

$$\max_{S\subseteq V}\{|\sum_{(ij):i,j\in S} w(ij)| + |\sum_{(ij):i,j\in \bar{S}} w(ij)|.\} \tag{15}$$

It can also be extended to the cases where there are multiple states where the number of states is k:

$$\max_{\{S_1,S_2,..,S_k\} \text{ is a partition of V}} \sum_k |\sum_{(ij):i,j\in S_k} w(ij)|. \tag{16}$$

In many applications, it is unknown how many states the alternative hypothesis might have. While the objective function could be easily modified to be the average weight for a k-partition of the graphs for any k, it is unclear how to choose the optimal k. Inspirations could be drawn from the correlation clustering literature (See [14] and references therein) where the number of clusters is determined automatically .

C. Slope

In many practical trend detection applications, a trend is important only if it has a significant slope. The Mann-Kendall test is only concerned with the *persistency* of the trend, and it is ignorant of the difference in values between samples. The Theil-Sen estimator of the slope [27], [28], on the other hand, calculates the median between slope between samples. Since the Theil-Sen slope estimator can be interpreted as the median weight of the graph where each weight is given by the slope between two observations, the proposed generalization can be also applied to the Theil-Sen slope estimator.

IX. Conclusions

In this paper we introduced the problem of detecting a trend in a sequence obtained from randomly interleaving two sequences, each exhibiting an upward trend. The problem is complicated the unknown *data association*: it is unknown to the detector which measurement sample belongs to which sequence. We proposed a GMK trend detection algorithm, which is based on the solution of a MILP problem, and we showed via simulation that it out-performs the Mann-Kendall algorithm. We also provided a *graph* formulation of the GMK problem and demonstrated the equivalence between the GMK test statistic computation and the Max-Bisection problem.

Acknowledgement

The authors would like to thank Jayakrishnan Unnikrishnan, Michael Lexa, Tao Yang, Frederick Wheeler and Steven Tyber for some inspiring discussions.

Appendix

A. Proof of Proposition 5.1

Proof: On letting $U_i = u_i$, it is easy to see that the sets of feasible solutions to (5) and (7) are the same. It remains to show that the objective function is equivalent: It is easy to verify that:

$$1 - |u_i - u_j| = \begin{cases} 0, & \text{if } u_i = 0, u_j = 1 \text{ or } u_i = 1, u_j = 0, \\ 1, & \text{if } u_i = u_j = 0, \text{ or } u_i = u_j = 1. \end{cases}$$

The same right-hand-side applies to $|u_i + u_j - 1|$. Therefore, the objective function in (7) is equal to

$$\sum_{i,j:u_i=u_j,w_{i,j}>0} w_{i,j} + \sum_{i,j:u_i=u_j,w_{i,j}<0} w_{i,j} = \sum_{i,j:u_i=u_j} w_{i,j},$$

of which the right-hand-side is equivalent to the objective function in (5) by the definition of $w_{i,j}$. ■

B. Proof of Proposition 5.2

Proof: Consider any sequence of observations $\mathbf{Y}^n$, if T^{GMK}, using the optimal solution, satisfies $T^{\mathsf{GMK}} \geq \eta$, then its value $T^{\mathsf{GMK}'}$ based on the approximation algorithms, satisfies $T^{\mathsf{GMK}'} \geq \alpha\eta$. Consequently, $P_D' \geq P_D$.

For the probability of false alarm, it follows from $T^{\mathsf{GMK}'} \leq T^{\mathsf{GMK}}$ that $P_F' \leq 1 - F_n(\alpha\eta)$. This is equivalent to the conclusion in (9) since $\eta = F_n^{-1}(1 - P_F)$. ■

C. Proof of Theorem 7.1

Proof: It follows from the definition of w and w' in (10) and (13) that

$$w(ij) + 2w'(ij) = 1.$$

Summing this over all edges in $\mathcal{E}(S, \bar{S})$ and applying the fact that there are $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2}$ edges in $\mathcal{E}(S, \bar{S})$ leads to,

$$\sum_{(ij)\in\mathcal{E}(S,\bar{S})} w(ij) + 2 \sum_{(ij)\in\mathcal{E}(S,\bar{S})} w'(ij) = \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil.$$

On the other hand,

$$\begin{aligned} &\sum_{(ij):i,j\in S} w(ij) + \sum_{(ij):i,j\in \bar{S}} w(ij) \\ &= \sum_{(ij)\in E} w(ij) - \sum_{(ij)\in\mathcal{E}(S,\bar{S})} w(ij) \\ &= \sum_{(ij)\in E} w(ij) + 2 \sum_{(ij)\in\mathcal{E}(S,\bar{S})} w'(ij) - \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil, \end{aligned}$$

where the first term on the right-hand side of the last equation is equal to T^{MK}, and the second term is twice the objective function in the definition of $T^{\mathsf{Bisection}}$ in (12). Substituting this into (11) leads to the conclusion in (14). ■

References

[1] D. Helsel and R. Hirsch, *Statistical Methods in Water Resources*. Elsevier, Amsterdam, 1992.

[2] H. Mann, “Nonparametric test against trend,” *Econometrica*, vol. 13, pp. 245–259, 1945.

[3] M. G. Kendall, “A new measure of rank correlation,” *Biometrika*, pp. 81–93, 1938.

[4] G. Hofmann and N. Balakrishnan, “A nonparametric test for trend based on initial ranks,” *Journal of Statistical Computation and Simulation*, vol. 76, no. 9, pp. 829–837, 2006.

[5] S. Yue and P. Pilon, “A comparison of the power of the t test, Mann-Kendall and bootstrap tests for trend detection,” *Hydrological Sciences Journal*, vol. 49, no. 1, pp. 21–37, 2004.

[6] R. M. Hirsch and J. R. Slack, “A nonparametric trend test for seasonal data with serial dependence,” *Water Resources Research*, vol. 20, no. 6, pp. 727–732, 1984.

[7] K. H. Hamed and A. R. Rao, “A modified mann-kendall trend test for autocorrelated data,” *Journal of Hydrology*, vol. 204, no. 1, pp. 182–196, 1998.

[8] G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*. Holden Day, 1976.

[9] W. Melek, Z. Lu, A. Kapps, and W. Fraser, “Comparison of trend detection algorithms in the analysis of physiological time-series data,” *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 4, pp. 639–651, April 2005.

[10] R. Hirsch, Alexander, and R. Smith, “Selection of methods for the detection and estimation of trends in water quality,” *Water Resources Research*, vol. 27, pp. 803–813, 1991.

[11] S. Fatichi, S. Barbosa, E. Caporali, and E. Silva, “Deterministic versus stochastic trends: Detection and challenges,” *Journal of Geophysical Research*, vol. 114, 2009.

[12] X. Xu and P. Antsaklis, “Optimal control of switched systems based on parameterization of the switching instants,” *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 2–16, 2004.

[13] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. Storrs, CT: YBS Publishing, 2011.

[14] N. Bansal, A. Blum, and S. Chawla, “Correlation clustering,” in *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, Vancouver, Canada, Nov 1992, pp. 238–250.

[15] D. Liberzon and A. Morse, “Basic problems in stability and design of switched systems,” *IEEE Control Systems Magazine*, vol. 19, pp. 59–70, 1999.

[16] A. Tanwani, H. Shim, and D. Liberzon, “Observability for switched linear systems: characterization and observer design,” *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 891–904, 2013.

[17] Y. Bar-Shalom, F. Daum, and J. Huang, “The probabilistic data association filter,” *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 82–100, December 2009.

[18] F. Farnoud and O. Milenkovic, “An axiomatic approach to constructing distances for rank comparison and aggregation,” *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6417–6439, 2014.

[19] H. Daniels, “Rank correlation and population models,” *Journal of Royal Statistical Society Series B*, vol. 12, pp. 171–181, 1950.

[20] E. J. Dietz and T. J. Killeen, “A nonparametric multivariate test for monotone trend with pharmaceutical applications,” *Journal of the American Statistical Association*, vol. 76, no. 373, pp. 169–174, 1981.

[21] R. Hirshc, J. Slack, and R. Smith, “Techniques of trend analysis for monthly water quality data,” *Water Resources Research*, vol. 18, no. 1, pp. 107–121, 1982.

[22] G. Casella and R. Berger, *Statistical Inference*. Duxbury Press, 2001.

[23] D. Bertsimas and J. N. Tsitsiklis, *Introduction to linear optimization*. Athena Scientific Belmont, MA, 1997, vol. 6.

[24] R. Lougee-Heimer, “The common optimization interface for operations research: Promoting open-source software in the operations research community,” *IBM Journal of Research and Development*, vol. 47, no. 1, pp. 57–66, 2003.

[25] V. V. Vazirani, *Approximation algorithms*. Springer, 2001.

[26] A. Frieze and M. Jerrum, “Improved approximation algorithms for max k-cut and max bisection,” *Algorithmica*, vol. 18, pp. 67–81, 1997.

[27] H. Theil, “A rank-invariant method of linear and polynomial regression analysis,” in *Henri Theils Contributions to Economics and Econometrics*. Springer, 1992, pp. 345–381.

[28] P. Sen, “Estimates of the regression coefficient based on Kendalls tau,” *Journal of the American Statistical Association*, vol. 63, pp. 1379–1389, 1968.

[29] D. Johnson, “The NP-completeness column: an ongoing guide,” *Journal of Algorithms*, pp. 434–451, 1985.