

EMULAZIONE E STATISTICHE DELLA PMD ATTRAVERSO ALGORITMI MULTICANONICI MULTIVARIATI

Armando Vannucci, Nicola Rossi, Alberto Bononi
Dipartimento di Ingegneria dell'Informazione - Università degli Studi di Parma
Viale G.P.Usberti 181/A, 43100 Parma
cognome@tlc.unipr.it

We derive two different algorithms that extend the modified multicanonical method to the estimation of bivariate distributions. Applied to PMD, such technique efficiently generates fiber samples with very low first- and second-order PMD probability.

1. Introduzione

Da alcuni anni, la letteratura sui sistemi di comunicazione ottica si è arricchita di una nuova famiglia di strumenti statistici di simulazione, noti col nome di metodi "multicanonici". L'idea, introdotta nell'ambito della fisica dei solidi [1], è stata adattata in [2] all'emulazione di fibre affette da Dispersione Modale di Polarizzazione (PMD) e alla determinazione della densità di probabilità (pdf) del ritardo differenziale di gruppo (DGD), mediante l'algoritmo iterativo *Multicanonical Monte Carlo (MMC) standard*. Lo stesso problema è stato poi affrontato in [3] attraverso l'algoritmo *MMC modificato*, il cui pregio è quello di limitare l'effetto delle oscillazioni statistiche della stima, accelerando la convergenza delle iterazioni. Al di là della loro formulazione specifica, questi algoritmi presentano alcune caratteristiche generali che ne determinano la potenza e la flessibilità e che proviamo a sintetizzare di seguito.

Gli algoritmi MMC si prestano in particolare ad essere applicati come una sorta di "contenitore" di uno specifico sistema di simulazione, che ne controlli gli ingressi (*stato S* del sistema) e ne osservi le uscite (variabili *obiettivo X*) al fine di stimarne la pdf $p(X)$; altre quantità statistiche di interesse (probabilità di eventi errore, *outage*, etc.) si possono determinare a valle della stima di $p(X)$. In maniera simile a quanto accade nelle tecniche di *Importance Sampling*, lo spazio probabilistico S viene campionato deformandone le probabilità a-priori, in modo da aumentare la probabilità di occorrenza dei campioni più rari (*warping*), raffinando iterativamente la stima della pdf $p(X)$ prodotta e impiegando tale stima per il *warping* del ciclo successivo. Al contrario dell'*Importance Sampling*, l'iteratività consente di applicare i metodi multicanonici senza che si abbia alcuna cognizione a-priori sulla sagoma della pdf da determinare, potendo partire da una pdf arbitraria (ad esempio uniforme) al primo ciclo e affidandosi all'autoadattatività del metodo. Le visite ai campioni sperimentali avvengono attraverso il cosiddetto *Algoritmo di Metropolis*, che genera nello spazio dei campioni un cammino stocastico a incrementi indipendenti (cioè una catena di Markov): il campione aggiornato dopo l'incremento viene accettato o rigettato in funzione del rapporto tra le probabilità (deformate) dei due campioni coinvolti nella transizione.

L'applicazione di interesse per questo lavoro è l'emulazione di fibre ottiche affette da PMD attraverso un modello a 100 *lamine*, con DGD locale costante, la cui birifrangenza (uniformemente distribuita sulla sfera di Poincaré) rappresenta lo *stato S* del sistema. Le quantità stocastiche a cui siamo interessati sono: $X=|\Omega|/\Delta\tau$ e $Y=|\Omega_\omega|/\Delta\tau^2$, ovvero il modulo del vettore di PMD della fibra Ω e quello della sua derivata rispetto alla pulsazione Ω_ω , normalizzati rispetto al DGD medio $\Delta\tau$ delle fibre considerate. In [4,5], è stata stimata la pdf congiunta $p(X,Y)$ mediante la diretta estensione dell'algoritmo *MMC standard* [2] al caso bidimensionale. Esamineremo qui l'estensione del metodo *MMC modificato* [3], che raggiunge un'efficienza decine di volte superiore a quello standard. Tuttavia, come vedremo, la sua estensione al caso multivariato non è univoca: introdurremo due dei possibili algoritmi derivabili. La pdf della PMD non è, di per sé, un obiettivo scientificamente interessante, essendo disponibile un algoritmo [6] per il suo calcolo numerico, che useremo come riferimento. Tuttavia, un fondamentale prodotto degli algoritmi discussi di seguito è la

generazione efficiente di campioni di fibra rari, con elevati valori di PMD, necessari per attestare i limiti di prestazione nella simulazione di sistemi di trasmissione ottica.

2. L'algoritmo multicanonico modificato per la stima di pdf bivariate

Nel metodo *MMC standard* [2,4,5], la regola di base per l'aggiornamento iterativo della stima di una pdf è: $p_{n+1}(X,Y,\dots)=p_n(X,Y,\dots)H_n(X,Y,\dots)/c_n$, dove $H_n(\cdot)$ rappresenta l'istogramma delle visite prodotte, all'iterazione n , dal cammino stocastico, mentre c_n è una costante di normalizzazione calcolabile a-posteriori. Seguendo l'approccio del metodo *MMC modificato*, esposto in [3] nel caso di una pdf monovariata, si assuma, come prima ipotesi, la seguente espressione per la pdf congiunta delle nostre due variabili:

$$p(X,Y)=\exp\{\beta^X(X,Y)X+\beta^Y(X,Y)Y-\alpha(X,Y)\} \quad (1)$$

Se le funzioni presenti all'esponente sono "lentamente variabili", ovvero se le derivate parziali di $\beta^X(\cdot)$, $\beta^Y(\cdot)$ e $\alpha(\cdot)$ rispetto a X o a Y sono sufficientemente piccole, allora è possibile esprimere, ad ogni $(n+1)$ -ma iterazione, $\beta^Z_{n+1}(X,Y)\approx\partial[\log p_{n+1}(X,Y)]/\partial Z$ con la derivata della pdf logaritmica rispetto alla variabile $Z=\{X,Y\}$ considerata. Impiegando per $p_{n+1}(X,Y)$ la regola di aggiornamento *MMC standard*, citata sopra, e discretizzando lo spazio delle variabili obiettivo (X,Y) in *bin* bidimensionali, con ampiezze $(\Delta X,\Delta Y)$ e valori centrali (X_k,Y_l) , le funzioni $\beta^Z(X,Y)$ andrebbero aggiornate come segue:

$$\begin{aligned} \beta^{X0}_{n+1}(X_k,Y_l) &= \beta^X_n(X_k,Y_l) + (\log H_n(X_{k+1},Y_l) - \log H_n(X_k,Y_l))/\Delta X \\ \beta^{Y0}_{n+1}(X_k,Y_l) &= \beta^Y_n(X_k,Y_l) + (\log H_n(X_k,Y_{l+1}) - \log H_n(X_k,Y_l))/\Delta Y \end{aligned} \quad (2)$$

Poiché l'istogramma $H_n(\cdot)$ in ciascun bin è una quantità aleatoria, i termini logaritmici in (2) sono soggetti a oscillazioni statistiche. Assumendo oscillazioni statisticamente indipendenti tra i bin, la varianza-campione di tali quantità è approssimabile con un'espansione in serie di Taylor del logaritmo, dunque la varianza nella stima di $\beta^{X0}(X,Y)$ e $\beta^{Y0}(X,Y)$ dopo il passo n e per il bin (X_k,Y_l) è

$$\begin{aligned} \sigma^2(\beta^{X0}_{n+1}(X_k,Y_l)) &\approx (H_n^{-1}(X_{k+1},Y_l) + H_n^{-1}(X_k,Y_l))/\Delta X^2 = (g_n^X(X_k,Y_l))^{-1} \\ \sigma^2(\beta^{Y0}_{n+1}(X_k,Y_l)) &\approx (H_n^{-1}(X_k,Y_{l+1}) + H_n^{-1}(X_k,Y_l))/\Delta Y^2 = (g_n^Y(X_k,Y_l))^{-1} \end{aligned} \quad (3)$$

Gli inversi delle varianze definiscono dei *parametri di affidabilità* $g_n^Z(\cdot)$, per ciascuna delle eq. (2), che vengono poi normalizzati alla somma dei valori assunti fino all'iterazione attuale,

ottenendo dei *coefficienti di affidabilità storica* $\hat{g}_n^Z(X_k,Y_l) = g_n^Z(X_k,Y_l) / \sum_{i=1}^n g_i^Z(X_k,Y_l)$ che

costituiscono il cuore del metodo *MMC modificato*. Difatti, le espressioni che vengono effettivamente impiegate per l'aggiornamento delle due funzioni $\beta^Z_{n+1}(X_k,Y_l)$, per Z coincidente con X o con Y , sono la combinazione lineare di due termini: uno è il secondo membro dell'eq.(2) (che equivale all'algoritmo standard), l'altro è il valore all'iterazione precedente $\beta^Z_n(X_k,Y_l)$, pesati rispettivamente da $\hat{g}_n^Z(X_k,Y_l)$ e dal suo complemento a uno. Esprimendo tali leggi di aggiornamento direttamente per la pdf da stimare, si ottengono nel nostro caso le seguenti regole fondamentali di iterazione:

$$\begin{aligned} p_{n+1}(X_{k+1},Y_{l+1}) &= p_n(X_{k+1},Y_{l+1}) \frac{p_{n+1}(X_k,Y_{l+1})}{p_n(X_k,Y_{l+1})} \left[\frac{H_n(X_{k+1},Y_{l+1})}{H_n(X_k,Y_{l+1})} \right]^{\hat{g}_n^X(X_k,Y_l)} \\ p_{n+1}(X_{k+1},Y_{l+1}) &= p_n(X_{k+1},Y_{l+1}) \frac{p_{n+1}(X_{k+1},Y_l)}{p_n(X_{k+1},Y_l)} \left[\frac{H_n(X_{k+1},Y_{l+1})}{H_n(X_{k+1},Y_l)} \right]^{\hat{g}_n^Y(X_k,Y_l)} \end{aligned} \quad (4)$$

A seconda di quale asse, X o Y , si consideri per l'aggiornamento iterativo della pdf, si può selezionare di volta in volta quale delle due regole impiegare, sulla base di qualche criterio di affidabilità. Tuttavia, come è intuitivo, la miglior performance si ottiene sfruttando tutta l'informazione disponibile: abbiamo verificato che la migliore evoluzione dell'algoritmo si ottiene combinando le due eq.(4) attraverso una media logaritmica pesata (che corrisponde ad una media aritmetica pesata per le funzioni $\beta^X(X,Y)$ e $\beta^Y(X,Y)$). I pesi sono scelti pari a $h^X / (h^X + h^Y)$ e $h^Y / (h^X + h^Y)$, dove, per ciascun *bin*, $h^Z(X_k,Y_l) = g_n^Z(X_k,Y_l) / \hat{g}_n^Z(X_k,Y_l)$

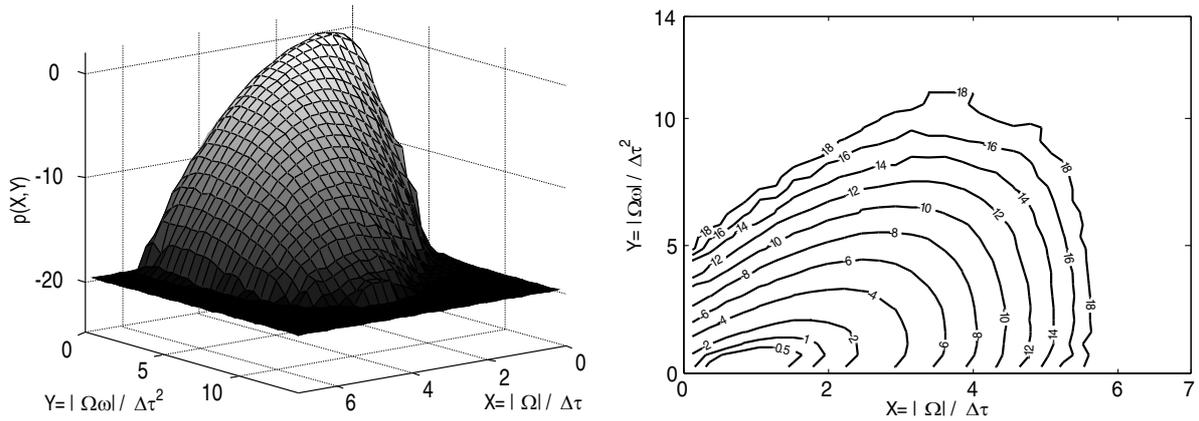


Fig.1: (sinistra) Stima della ottenuta con l'algorithm *MMC ad aggiornamento combinato*, dopo 6 cicli multicanonici ($5 \cdot 10^6$ fibre emulate). (destra) Curve di livello della pdf stimata.

è pari al rapporto tra l'affidabilità al ciclo attuale e l'affidabilità storica delle due regole di aggiornamento lungo $Z=X$ o $Z=Y$ in (4).

Chiamiamo questo algoritmo *MMC ad aggiornamento combinato*. Come nel caso monodimensionale, $p_n(X_1, Y_1)$ viene fissato ad un valore arbitrario (risolto poi attraverso la condizione di normalizzazione) e i valori di $p_n(X_k, Y_1)$ e $p_n(X_1, Y_l)$ vengono determinati applicando l'algoritmo monodimensionale [3] lungo i due assi. La Fig.1 riporta la stima della pdf ottenuta con l'algorithm *MMC ad aggiornamento combinato*, dopo 6 iterazioni, con un totale di $5 \cdot 10^6$ campioni di fibra emulate. Così come nel caso monodimensionale si ha un effetto di propagazione lungo X dell'ultimo valore "affidabile" calcolato per $p(X)$, generando intervalli con $p(X)$ costante laddove questa è incognita, qui le zone non esplorate dal cammino stocastico presentano una probabilità molto simile tra loro (nell'ordine di 10^{-18}), il che garantisce, come abbiamo verificato, una corretta evoluzione dell'algoritmo nei cicli successivi. L'andamento della pdf segue quello calcolabile dalla teoria [6], con un'evidente sovrastima nella parte piatta (mai visitata dai campioni) e, per contro, una sottostima che si enfatizza in prossimità del dominio visitato. Una fondamentale causa per la sottostima dei valori di probabilità per elevata PMD si può attribuire al modello a lamine con DGD locale costante, che, come noto, presenta un limite superiore teorico alla PMD di primo e secondo ordine.

L'estensione del metodo *MMC modificato* al caso bivariato, tuttavia, non è univoca. Per dimostrare un secondo possibile approccio, si assuma la seguente espressione, alternativa alla (1), per la pdf:

$$p(X, Y) = \exp\{\beta(X, Y)XY - \alpha(X, Y)\} \quad (5)$$

Ipotizzando che le funzioni all'esponente siano lentamente variabili in X e Y , possiamo valutare la derivata seconda mista della pdf logaritmica e ottenere una singola regola di aggiornamento per $\beta(X, Y)$. Procedendo in maniera analoga a quanto esposto sopra, si perviene infine ad una singola regola di aggiornamento della pdf: all'iterazione n e per ciascun bin, risulta

$$p_{n+1}(X_{k+1}, Y_{l+1}) = p_n(X_{k+1}, Y_{l+1}) \frac{p_{n+1}(X_{k+1}, Y_l) p_{n+1}(X_k, Y_{l+1})}{p_n(X_{k+1}, Y_l) p_n(X_k, Y_{l+1})} \frac{p_n(X_k, Y_l)}{p_{n+1}(X_k, Y_l)} \left[\frac{H_n(X_{k+1}, Y_{l+1}) H_n(X_k, Y_l)}{H_n(X_{k+1}, Y_l) H_n(X_k, Y_{l+1})} \right]^{\hat{g}_n(X_k, Y_l)} \quad (6)$$

dove il singolo *coefficiente di affidabilità storica*, $\hat{g}_n(X_k, Y_l)$ è definito, analogamente al caso precedente, in base all'inverso della varianza-campione del termine di aggiornamento di $\beta(X, Y)$, e dipende a sua volta dai valori dell'istogramma delle visite:

$$\sigma^2(\beta_{n+1}^0(X_k, Y_l)) \approx (H_n^{-1}(X_{k+1}, Y_{l+1}) + H_n^{-1}(X_k, Y_{l+1}) + H_n^{-1}(X_{k+1}, Y_l) + H_n^{-1}(X_k, Y_l)) / (\Delta X^2 \Delta Y^2) = g_n^{-1}(X_k, Y_l) \quad (7)$$

Chiamiamo questo algoritmo - per il quale si impiegano gli stessi passi di inizializzazione

descritti in precedenza - *MMC ad aggiornamento correlato*, poiché la definizione iniziale (5) implica correlazione statistica tra le variabili X e Y. Applicando la (6) al caso della PMD, tuttavia, i risultati che si ottengono sono coerenti con i valori effettivi di $p(X,Y)$ [6] solo in una zona limitata, mentre sono affetti da pesanti sovrastime della pdf nelle zone del piano (X,Y) con elevati valori di Y, non esplorate dall'algoritmo. Difatti, a differenza dell'algoritmo *MMC ad aggiornamento combinato*, l'eq.(6) implica qui una dominanza nella propagazione del più elevato tra i valori stimati nei bin adiacenti lungo i due assi, generando una sorta di "altopiano" che risulta particolarmente nocivo per la stima della pdf. Tale effetto compromette, già dalla prima iterazione, la probabilità di effettuare visite ai bin coinvolti e di migliorare la stima nei cicli successivi.

3. Conclusioni

Abbiamo mostrato come l'estensione dell'algoritmo *MMC modificato* [3] al caso di pdf bivariate non sia univoca. Partendo da diverse possibili assunzioni sull'espressione della pdf $p(X,Y)$, abbiamo derivato e implementato diversi algoritmi multicanonici bivariati. In questo lavoro abbiamo illustrato due di tali algoritmi, chiamati *MMC ad aggiornamento combinato* e *MMC ad aggiornamento correlato*, applicandoli alla stima della pdf congiunta della PMD di 1° e 2° ordine di fibre emulate con il *modello a lamine*. Benché il primo sia risultato il migliore per il problema di nostro interesse, non è possibile sancire la superiorità di tale approccio in generale. Difatti, come si intuisce dal confronto degli assunti di partenza (1) e (5), la prestazione dei vari algoritmi risulta fortemente dipendente dallo specifico problema in oggetto, ovvero dalla natura della pdf da stimare. Come termine di confronto, in [4] la pdf in oggetto viene stimata, mediante l'algoritmo *MMC standard*, fino a valori nell'ordine di 10^{-18} impiegando $120 \cdot 10^6$ campioni di fibra emulata. Con l'algoritmo *MMC ad aggiornamento combinato* proposto qui, un'analoga stima della pdf ha richiesto un numero di campioni di fibra emulata 24 volte inferiore ($5 \cdot 10^6$ campioni), consentendo un corrispondente risparmio sui tempi di calcolo.

Ringraziamenti

Gli autori desiderano ringraziare il Prof. Enrico Forestieri per le interessanti discussioni e per l'aiuto fornito nell'implementazione del suo metodo analitico di calcolo.

Bibliografia

- [1] B. A. Berg, T. Neuhaus, "Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions", *Physical Review Letters*, Vol. 68, N. 1, *gennaio 1992*, pp.9-12.
- [2] D. Yevick, "Multicanonical Communication System Modeling – Application to PMD Statistics", *IEEE Photonics Technology Letters*, Vol. 14, N. 11, *novembre 2002*, pp.1512-1514.
- [3] D. Yevick, "The Accuracy of Multicanonical System Models", *IEEE Photonics Technology Letters*, Vol. 15, N. 2, *febbraio 2003*, pp.224-226.
- [4] S. L. Fogal, W. L. Kath, G. Biondini, "Multicanonical Monte Carlo simulations of first- and second-order PMD", in *IEEE LEOS Topical Meeting Polarization Mode Dispersion*, Vancouver (Canada), 14-16 *luglio 2003*, pp.33-34.
- [5] D. Yevick, "Multicanonical Evaluation of Joint Probability Density Functions in Communication System Modeling", *IEEE Photonics Technology Letters*, Vol. 15, N. 14, *novembre 2003*, pp.1540-1542.
- [6] E. Forestieri, "A Fast and Accurate Method for Evaluating Joint Second-Order PMD Statistics", *IEEE Journal of Lightwave Technology*, Vol. 11, N. 11, *novembre 2003*, pp.2942-2952.