

Kuantum Hesaplamanın Bioinformatik ve Kimyadaki Problemlere Uygulanması

Program Kodu: 1002

Proje No:115E747

Proje Yürütücüsü:

Ammar DAŞKIN

(Ağustos 2016)

ÖNSÖZ

Bu rapordaki çalışma TÜBİTAK tarafından “Kuantum Hesaplamanın Bioinformatik ve Kimyadaki Problemlere Uygulanması” başlıklı ve 115E747 numaralı proje kapsamında desteklenmiştir. Bu rapor bu proje kapsamında yazılan 3 makaleden (İkisi SCI-expanded indeksli dergilerde yayınlanmıştır.) faydalanılarak yazılmıştır. Bu makalelere için lütfen Daskin (2015, 2016a,b)’ye bakınız.

İÇİNDEKİLER

ÖNSÖZ	ii
ÖZET	1
ABSTRACT	3
1. GİRİŞ	5
2. KUANTUM HESAPLAMA	7
3. KUANTUM FAZ HESAPLAMANIN İNDİRGENEMEYEN POZİTİF MATRİSLERE TATBİKİ	9
4. KUANTUM ISORANK	10
5. KUANTUM ASLİ UNSUR İNCELEME	11
6. SONUÇ	12
7. KAYNAKLAR	13

ÖZET

Kuantum bilgisayarlar kuantum mekaniksel özelliklerini kullanarak daha verimli algoritmalar tasarlama imkanı sunarlar. Bu münasebetle, kuantum bilgisayarların hesaplama gücünün günümüzde kullanılan konvansiyonel bilgisayarlara kıyasla daha kuvvetli olduğuna inanılmaktadır. Verilen bir matrisin özdeğerini (eigenvalue) bulmakta yada diğer algoritmaların temel parçası olarak kullanılmakta olan kuantum faz hesaplama algoritması hesaplama yönünden verimli olan en meşhur kuantum algoritmalarından bir tanesidir. Bu algoritmanın tatbiki başarısının girişte verilen tahmini özvektöre(eigenvector) bağlı olması, algoritmanın kullanım alanlarını daraltmaktadır.

Desteklenen proje kapsamında, kuantum faz hesaplama algoritması modifiye edilerek bu algoritmadaki bilinen kısıtlamalara çözüm sunulması ve böylelikle kullanım alanlarının genişletilmesi ve buna ek olarak modifiye edilmiş bu algoritmanın gücünden faydalanılarak biyoinformatik ve kuantum kimyası alanlarındaki zor problemlere verimli çözümler sunulması amaçlanmıştır. Bu amaç doğrultusunda, üç adet çalışma yapılmıştır (bakınız Daskin (2015, 2016a,b)).

Faz hesaplama algoritmasında girişte verilen tahmini özvektörün gerçek özvektöre yakınlığı algoritmanın başarısını belirlemektedir. İlk çalışmada (Daskin 2016b), algoritmanın negatif olmayan matrislere tatbikatında giriş vektörünün rolü incelenerek; söz konusu matrislerin dominant özdeğer-özvektör çiftinin pozitif olması ve diğer özvektörlerinde de en az bir negatif katsayı barındırması özelliği kullanılmış ve girişte verilen tahmini özvektörün pozitif özvektöre daha yakın olması sağlanabileceği ve dolayısıyla algoritmanın indirgenemeyen ve negatif olmayan matrisler için verimli bir şekilde kullanılabileceği gösterilmiştir. Örnek olarak stokastik özellik gösteren k-lokal Hamiltonianlar ve rastgele sayılar kullanılmış ve özdeğerlerinin hesaplanmasında,

Farklı türden canlılar için protein protein etkileşim ağlarının mukayeseli analizleri biyolojik süreçlerin anlaşılmasında önemli katkı sağlamaktadır. Bu ağlar arasındaki benzerliklerin anlaşılması ayrıca farklı türler arasında fonksiyonel benzerliklerin bulunmasına yardımcı olmaktadır. Ancak devamlı büyüyen etkileşim verilerinin kapasitesi günümüzde hesaplama yönünden zorluklar ortaya çıkarmaktadır. Protein protein etkileşim ağlarının mukayesesinde yaygın bir şekilde kullanılan algoritmalarından IsoRank diye isimlendirilen bir tanesi, ağlar arasındaki etkileşim skorlarını bu ağların Kronecker (tensor) çarpımından elde edilmiş bir stokastik matrisin özvektöründen hesaplanmaktadır. Stokastik matrisler ayrıca pozitif olmaları sebebiyle negatif olmayan matrisler grubuna dahildirler. Proje kapsamındaki ikinci çalışmada (Daskin 2015), protein protein etkileşim ağlarının mukayese edilmesinde; kuantum faz hesaplama yöntemi kullanılarak, klasik IsoRank algoritmasının kuantum versiyonu önerilmiş ve bu ağların sıralanmasının üstsel olarak daha verimli bir şekilde

kuantum bilgisayarlarca yapılabileceği gösterilmiştir.

Son çalışmada (Daskin 2016a) ise kuantum faz hesaplama algoritması kuantum katsayı yükseltme algoritması ile birlikte kullanılarak kuantum asli unsur inceleme (quantum principal component analysis) usulü tanımlanmıştır. Yapılan çalışmalar, kuantum bilgisayarların biyoinformatik ve kuantum kimyasının yanı sıra benzeri alanlardaki problemlere tatbikatına da, imkan sağlamaktadır.

Anahtar Kelimeler: Kuantum Faz Hesaplama, Negatif Olmayan Matrisler, Kuantum IsoRank, Kuantum Asli Unsur İnceleme

ABSTRACT

Quantum computers are believed to be more powerful than the digital computers since they allow us to design more efficient algorithms to solve different problems. Quantum phase estimation algorithm is one of the efficient algorithms which finds the lowest eigenvalue of a matrix more efficiently than known classical algorithms. Furthermore, with different settings, it has been adapted as a sub-frame of many quantum algorithms applied to a wide variety of applications in different fields, e.g. Shor's factoring algorithm. However, in the implementation of the algorithm, the requirement of a good approximate eigenvector as an input to the algorithm hinders the application of the algorithm to the problems where we do not have any a priori knowledge about the eigenvectors.

In this project, the main goal has been to widen the use of quantum phase estimation algorithm in bioinformatics and chemistry by providing solutions to some of its well-known constraints. To achieve this goal, three main works given in Ref.(Daskin 2015, 2016a,b) have been completed. In the first work, a modification is introduced to make the phase estimation algorithm be used with ease for non-negative matrices. It is shown that the algorithm can be used efficiently to obtain the eigenvector of a non-negative matrix without a-priori approximate eigenvector which is one of the constraints in the conventional algorithm. In the numerical examples, the performance of the algorithm has been investigated for different instances of k -local Hamiltonians and random matrices.

Comparative analyses of protein-protein interaction networks play important roles in the understanding of biological processes. The alignment of multiple protein-protein interaction networks is necessary to understand the functional similarities across different species in bioinformatics. However, the growing enormity of available data on the networks becomes a computational challenge for the conventional alignment algorithms. The quantum phase estimation algorithm can generate the principal eigenvector of a stochastic matrix, which is also positive, with probability one. Using this property, in the second work (Daskin 2015), a quantum computing approach for the alignment of protein-protein interaction networks is introduced by following the classical algorithm IsoRank which uses the principal eigenvector of the stochastic matrix representing the Kronecker product of the normalized adjacency matrices of networks for the pairwise alignment. It is shown that this approach is able to provide exponential computational efficiency over the classical algorithm.

Finally in the third work(Daskin 2016a), a quantum version of the principal component analysis-a multivariate statistical method frequently used in science and engineering-is introduced. We believe this project will pave the way for the advancement of new quantum algorithms and quantum bioinformatics and ease simulation difficulties of quantum chemistry on quantum computers.

Keywords: Quantum Phase Estimation; Non-negative Matrices, Quantum IsoRank, Quantum Principal Component Analysis

1. GİRİŞ

Kuantum bilgisayarlar kuantum mekaniksel sistem tabanlıdır ve sistemin Hamiltonian operatörüne bağlı Schrödinger denkleminde göre davranırlar. Bu bilgisayarlar üzerinde, *entanglement* ve *superposition* gibi kuantum fenomeni kullanılarak yapılan hesaplama-kuantum hesaplama-yeni bir bilgi işleme usulü tanımlar. Temel olarak bilimin ilerlemesi ve bunun tabii sonucu olarak tabiatın daha iyi idrak edebilmesi, bilgisayarların hesaplama kapasitesiyle doğrudan ilişkilidir. Günümüzde kullanmış olduğumuz dijital bilgisayarlar, fiziksel limitlerine ulaşmaktadır. Intel'in kurucularından Gordon Moore'ın şimdiye kadar geçerli olan, her 18 ayda bir devre birimindeki transistör sayısının ikiye katlanır şeklinde (Moore 1965) ifade edilen tahmini artık sürdürülebilir olmaktan çıkmıştır (Powell 2008). Bu, kuantum hesaplama gibi yeni tür hesaplama usullerine ihtiyaç teşkil etmektedir.

Daha da önemlisi çevremizdeki küçük parçacıkların davranışları kuantum mekaniğiyle açıklanmaktadır. Ancak kuantum sistemlerin benzetiminin yapılması, günümüz bilgisayarlarında zordur. Mesela, moleküler simülasyonlarda temel olarak iki zorlukla karşılaşmaktadır (Van Gunsteren ve Berendsen 1990):

- Moleküler sistemce ulaşılabilir konfigürasyon uzayının büyüklüğü,
- Kullanılan moleküler modelin doğruluğu.

Kuantum simülasyonlarında problem büyüklüğü ve zorluğu simülasyona dahil edilen parçacık sayısına bağlı olarak ekponansiyel büyüdüğünden, bu tür simülasyonların günümüzde kullanılan bilgisayarlarla yapılması muhaldir (Lanyon vd. 2010).

Kuantum bilgisayarların kendilerinin de kuantum tabanlı olması bu tür benzetimlerin kullanımında verimlilik göstermektedir (Feynman 1982). Ancak yapılan hesaplamaların sonucunun ikili bitle ifade edilen dijital veriye haritalanması, kuantum bilgisayarların da dijital bilgisayarların limitleriyle sınırlanmaktadır (Bennett vd. 1997; Deutsch 1985; Martin-Delgado 2012): Mesela, dijital bilgisayarlarda karar verilemeyen problemler sınıfındaki bir problem yine kuantum bilgisayarlarda da karar verilemez sınıfına dahildir. Buna misal olarak bir programın ne zaman duracağı problemi (halting problem) verilebilir. Diğer taraftan, kuantum algoritmaların hesaplama karmaşıklığı klasik algoritmalarından daha azdır ve dolayısıyla daha verimli bir şekilde bir çok probleme çözüm sağlamaktadırlar. Shor'un bir tam sayıyı çarpanlarına ayırma algoritması (Shor 1994) ve kuantum faz hesaplama algoritması(Kitaev 1996) bilinen klasik algoritmalarından ekponansiyel daha hızlıdır-

lar. Faz hesaplama algoritması her ne kadar bir çok probleme verimli bir şekilde uygulanmasına rağmen; bu algoritmanın başarısının tahmini giriş özvektöre bağlı olması, tahmini özvektörün erişilemez olduğu yada kötü olduğu durumlarda, bu algoritmanın uygulanmasını oldukça zorlaştırmaktadır.

Bu raporda, faz hesaplama algoritmasının modifiye edilerek farklı uygulamalarını gösteren üç adet çalışmayı(Daskin 2015, 2016a,b)). kapsamaktadır. Yukarıda belirtildiği gibi, faz hesaplama algoritmasında girişte verilen tahmini özvektörün gerçek özvektöre yakınlığı algoritmanın başarısını belirlemektedir. Yapılan ilk çalışmada (Daskin 2016b), algoritmanın negatif olmayan matrislere tatbikatında giriş vektörünün rolü incelenerek; söz konusu matrislerin dominant özdeğer-özvektör çiftinin pozitif olması ve diğer özvektörlerinde de en az bir negatif katsayı barındırması özelliği kullanılmış ve girişte verilen tahmini özvektörün pozitif özvektöre daha yakın olması sağlanabileceği ve dolayısıyla algoritmanın indirgenemeyen ve negatif olmayan matrisler için verimli bir şekilde kullanılabileceği gösterilmiştir. Örnek olarak stokastik özellik gösteren k-lokal Hamiltonianlar ve rastgele sayılar kullanılmış ve özdeğerlerinin hesaplanmasında,

Farklı türden canlılar için protein protein etkileşim ağlarının mukayeseli analizleri biyolojik süreçlerin anlaşılmasında önemli katkı sağlamaktadır. Bu ağlar arasındaki benzerliklerin anlaşılması ayrıca farklı türler arasında fonksiyonel benzerliklerin bulunmasına yardımcı olmaktadır. Ancak devamlı büyüyen etkileşim verilerinin kapasitesi günümüzde hesaplama yönünden zorluklar ortaya çıkarmaktadır. Protein protein etkileşim ağlarının mukayesesinde yaygın bir şekilde kullanılan algoritmalarından IsoRank diye isimlendirilen bir tanesi, ağlar arasındaki etkileşim skorlarını bu ağların Kronecker (tensor) çarpımından elde edilmiş bir stokastik matrisin özvektöründen hesaplanmaktadır. Stokastik matrisler ayrıca pozitif olmaları sebebiyle negatif olmayan matrisler grubuna dahildirler. Proje kapsamındaki ikinci çalışmada (Daskin 2015), protein protein etkileşim ağlarının mukayese edilmesinde; kuantum faz hesaplama yöntemi kullanılarak, klasik IsoRank algoritmasının kuantum versiyonu önerilmiş ve bu ağların sıralanmasının üstsel olarak daha verimli bir şekilde kuantum bilgisayarlarca yapılabileceği gösterilmiştir.

Son çalışmada (Daskin 2016a) ise kuantum faz hesaplama algoritması kuantum katsayı yükseltme algoritması ile birlikte kullanılarak kuantum asli unsur inceleme (quantum principal component analysis) usulü tanımlanmıştır. Yapılan çalışmalar, kuantum bilgisayarların biyoinformatik ve kuantum kimyasının yanı sıra benzeri alanlardaki problemlere tatbikatına da, imkan sağlamaktadır.

Aşağıdaki kısımda ilk olarak kuantum hesaplama giriş niteliğinde kısa bilgiler verildikten sonra

yukarıda adı geçen proje katmanlarındaki yapılan çalışmalar özetlenmiştir.

2. KUANTUM HESAPLAMA

Projede temel olarak kuantum faz hesaplama algoritmasının modifikasyonunu, bu algoritmanın bahsi geçen problemlere uygulanabilir hale getirilmesini, ve kuantum katsayı yükseltme algoritması ile birlikte kullanılmasını incelemektedir. Bu münasebetle, aşağıda öncelikli olarak kuantum faz hesaplama algoritmasının tanımı yapılmış ve temelde yapılan modifikasyon anlatılmıştır.

a. Kuantum Hesaplama (Nielsen ve Chuang 2010): Kuantum bilgisayarlar, Dirac notasyonunda gösterilen $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ vektörlerini kullanarak bilgiyi temsil ederler. Bu iki vektör, *qubit* olarak adlandırılan bir kuantum sisteminin, ayırt edilebilir iki durumuyla ilişkilendirilmiştir. Genel olarak herhangi bir kuantum sisteminin durumu bu iki vektör kullanılarak gösterilebilir:

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle. \quad (1)$$

$|\psi\rangle$ bir qubitin herhangi bir zamandaki durumunu tanımlar. α_0 ve α_1 ise sistemin herhangi bir ölçüm sonucunda hangi durumda bulunacağını ihtimalini veren kompleks katsayılardır. Daha açık bir ifade ile, sistem bir ölçüm sonucunda $|\alpha_0|^2$ ihtimalle $|0\rangle$ durumunda ve $|\alpha_1|^2$ ihtimalle de $|1\rangle$ durumunda bulunur. Sistemdeki muhtemel durumlar, sistemde düşünülen qubit sayısı ile eksponansiyel orantılı bir şekilde büyür. İki qubitten oluşan farazi bir sistem düşünelim; bu sistemdeki ayırt edilebilir ölçülebilen durum sayısı dört olacaktır. Dolayısıyla, sistemin herhangi bir durumu bu dört durum kullanılarak ifade edilebilir:

$$|\psi\rangle = \alpha_0 |00\rangle + \alpha_1 |01\rangle + \alpha_2 |10\rangle + \alpha_3 |11\rangle. \quad (2)$$

Bir qubitte olduğu gibi, burada da $|\alpha_i|^2$, ölçüm sonucunda sistemin $|i\rangle$ durumunda bulunma ihtimalini verir. Aynı metodoloji ve tanımlama kullanılarak daha büyük sistemler bu şekilde tanımlanabilir.

Sistemin herhangi bir anda durumunu değiştirmek için kuantum operasyonları uygulanır. Bunlar matematiksel düzlemde üniter matrisler ile tanımlanabilirler. Kuantum sistemin durumuna orantılı bir şekilde bu matrislerinde boyutu da benzer şekilde büyümektedir. Farklı bir ifadeyle, bu operasyonların sistem üzerindeki etkileri, matris-vektör transformasyonu olarak da tanımlanabilirler.

Mesela;

$$U \begin{vmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{vmatrix} = \begin{vmatrix} \tilde{\alpha}_0 \\ \tilde{\alpha}_1 \\ \tilde{\alpha}_2 \\ \tilde{\alpha}_3 \end{vmatrix}, \quad (3)$$

Yukarıda, U 4x4 bir üniter matris ve aynı zamanda bir kuantum operasyonudur. Bu türden matrislerin elemanları akıllı bir şekilde seçilerek farklı türden operasyonlar tanımlanabilir ve kuantum dünyasının sunduğu süperpozisyon ve entanglement (iç-içe-geçme) gibi fenomenlerle birlikte kullanılarak verimli algoritmalar geliştirilebilir.

b. Kuantum Fourier Transformasyonu (Nielsen ve Chuang 2010): Kuantum faz hesaplamasının temel parçalarından olan kuantum Fourier transformasyonunu (QFT) kısa bir şekilde açıklayacağız. QFT aslında ayrık Fourier transformasyonunun kuantum kompleks katsayılarına tatbikine denk gelmektedir. Bu, $|0\rangle, \dots, |N-1\rangle$ olarak verilen kuantum durum kümesini aşağıdaki şekilde dönüştürür:

$$|j\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi i jk/N} |k\rangle. \quad (4)$$

QFT, $O(\log^2 N)$ hesaplama adımı gerektirdiğinden üstel olarak DFT den daha hızlıdır.

c. Faz Hesaplama Algoritması (Kitaev 1996): Kuantum kompleks katsayılarıyla kodlanmış olarak verilen bir özvektör $|\mu_j\rangle$ için; kuantum faz hesaplama algoritmasının (phase estimation algorithm (PEA)) temel amacı $U|\mu_j\rangle = e^{i2\pi\phi_j}|\mu_j\rangle$ şeklinde tanımlanan özdeğer denkleminde, ϕ_j ile ifade edilmiş faz değerini hesaplamaktır. Bu algoritmanın icrasında iki register (kütük) kullanılır: $|reg_1\rangle$ and $|reg_2\rangle$. Bunların her birisi faz değerini ve özvektörü tutabilmeleri için yeterli sayıda qubitten müteşekkildir. Algoritmanın başlangıcında, $|reg_1\rangle$ 'da bütün qubitler sıfıra eşitlenirken, $|reg_2\rangle$ tahmini olarak verilmiş olan özvektörü tutar. QFT ve U^{2^j} şeklinde tanımlanmış kontrollü operasyonların uygulanması sonucu $|reg_1\rangle$ üstünde faz değerinin Fourier transformasyonu görülür. Binaenaleyh, QFT'nin tersi $|reg_1\rangle$ 'e uygulanarak, $|reg_1\rangle$ değeri faz değerine eşitlenir ve özdeğer denklemindeki faz değeri hesaplanmış olur. Burada eğer matris U , bir Hamiltonian H matrisinin zaman tekâmülü, yani $U = e^{i2\pi H}$ ise, o zaman H matrisinin özdeğeride hesaplanmış olur.

3. KUANTUM FAZ HESAPLAMANNIN İNDİRGENEMEYEN POZİTİF MATRİSLERE TATBİKİ

Elemanlarının her biri sıfır veya müspet olan simetrik bir matris için negatif değildir denir. Ve yine eğer böyle bir matrisle ilişkilendirilen graf kuvvetlice bağlı ise veya bu matris blok-diyagonal şekle indirgenemiyorsa, bu matrise için ayrıca indirgenemez denir. Perron-Frobenius teoremine göre indirgenemeyen ve negatif olmayan bir matris yine negatif olmayan bir özvektöre ve bu vektörle ilintilenmiş ve mutlak değeri diğer özdeğerlerden büyük olan bir özdeğere sahiptir (Meyer 2000). Bu türden matrisler geniş kapsamlı olarak çalışılmasının (bakınız Bapat ve Raghavan (1997); Berman ve Plemmons (1979)) yanı sıra; özvektör elemanlarının dağılımı ve toplamalarının, matris elemanlarıyla (Minc 1970) ve moleküler graflardaki rastgele patika sayısı (Gutman vd. 2001) ilişkili olduğu gösterildi.

Faz hesaplama algoritmasının temel başarısı, verilen giriş vektörünün özvektöre yakınlığına bağlıdır. Pozitif matrislerin sadece bir özvektörün pozitif olması özelliği kullanılarak bu algoritmanın başarısının giriş tahmini vektöre bağımlılığı azaltılabilir olacağını bu projenin ilk merhalesi olarak yapılan çalışmada (bakınız Daskin (2016b)) gösterilmiştir. Bu çalışmada, indirgenemeyen ve negatif olmayan matrisler için; faz hesaplama algoritmasındaki tahmini başlangıç özvektörü yerine eşit katsayılardan oluşan bir giriş değeri verilmiştir (Bu basit bir şekilde, başlangıç $|0\rangle$ vektörüne Hadamard kapıları uygulanarak başarılabılır.). Bu başlangıç değeri, algoritmanın sonunda birinci registerda (kütükte) özdeğerlerin süperpozisyonu üretmektedir. Süperpozisyonadaki her bir özdeğerin ihtimal katsayısı normalleştirilmiş özvektörlerinin-elemanları-topamları ile belirlenmektedir. Mevzubahis olan matrislerin asli özvektör ve özdeğeri diğerlerinden daha büyük olduğu ve negatif olmadıkları için ihtimal katsayısını belirleyen toplam çoğunlukla diğerlerinden çok daha büyük olmaktadır. Buda algoritmanın asli özvektör ve özdeğer çiftini hesaplamada kolaylıkla kullanılabileceğini göstermektedir. Ayrıca bu çalışmada, ikinci kütüğe Hadamard kapıları tatbik edildikten sonra ikinci kütüğün $|0\rangle$ değerinde ölçülmesine bağlı olarak birinci kütüğün ölçülmesinin, algoritmanın asli özdeğeri hesaplamadaki başarı ihtimali yükselttiği not edilmiştir. Dahası, başarı ihtimalinin önceden tahmin edilemeyeceği sorusu sorulmuş ve mevzubahis matrisin stokastik matrise yakınlığı kullanılarak bu soruya cevap aranmıştır.

Eğer $(C^2)^{\otimes n}$ üzerinde davranan Hamiltonian operatör $\mathcal{H}$, k qubit arasındaki etkileşimi gösteren lokal Hamiltonianların toplamı şeklinde yazılabilirse, bu Hamiltonian k -local n -qubit Hamiltonian olarak adlandırılır (Bravyi vd. 2008). Matematiksel olarak şöyle ifade edilebilir: $\mathcal{H} = \sum_s H_s$. Lokal Hamiltonian problemi $\mathcal{H}$ 'nin en düşük özdeğerinin bulunması olarak da formüle edilebilir. Eğer

lokal Hamiltonianların diagonal olmayan elemanları pozitif ise bu operatörlerle eşleştirilen Gibbs yoğunluk matrisi, $\rho = e^{-\beta\mathcal{H}}/Tr(-\beta\mathcal{H})$, $\beta \geq 0$ olduğunda pozitif elemanlara sahiptir. Bu sebeple $\mathcal{H}$ 'nin en düşük özdeğeri pozitif olan bir özvektörle eşleşmiştir. Sayısal simülasyonlarda rastgele matrisler ve 3-lokal Hamiltonianlar kullanılmıştır. Ve başarı ihtimallerinin yüksek olduğu ve bu değerlerinin tahmini hesaplanan başarı ihtimalleriyle örtüştüğü gözlemlenmiştir.

4. KUANTUM ISORANK

Protein protein etkileşim (PPI) ağları genel olarak graflar kullanılarak gösterilir. Mesela $G(V, E)$ grafi V düğüm ve $(v_i, v_j) \in E$ kenar kümeleri ile bir PPI ağı tanımlamakta ve v_i ve v_j bağlayan kenar i ve j proteinleri arasındaki etkileşimi göstermektedir. PPI ağlarının karşılaştırmalı analizleri biyolojik proseslerin anlaşılmasında önemli bir yer teşkil etmektedir. Farklı türlerin PPI ağlarının sıralanması ortak yönlerinin belirlenmesi bu türler arasındaki korunan bölümlerin kataloglanmasını ve fonksiyonel benzerliklerin ortaya çıkarılmasını sağlar. Bu ağların sıralanmasının kesin olarak elde edilmesi NP-Hard sınıfında bir problem olmasından ötürü, bu ağların sıralanması için son yıllarda bir çok farklı algoritma önerildi: (Kelley vd. 2004; Klau 2009; Koyutürk vd. 2006; Li vd. 2007; Singh vd. 2007; Zaslavskiy vd. 2009) ve diğerleri (Clark ve Kalita 2014). Ancak devamlı olarak büyüyen ulaşılabilir protein etkileşim verisi, halen daha hesaplama yönünden yeni zorluklar ortaya çıkarmaktadır. Bu algoritmalarından IsoRank (Singh vd. 2007), iki nodun sıralanma skorunun komşularının skoruna bağlı olması fikrini kullanarak PPI ağlarının sıralanması için meşhur derecelendirme algoritmalarına (misal, Google arama motorunda da kullanılan PageRank (Page vd. 1999)) benzer bir formülasyon sunmaktadır. IsoRank PPI ağlarının sıralanmasını aşağıdaki şekilde formüle eder:

$$R_{ij} = \sum_{u \in N(i)} \sum_{v \in N(j)} \frac{1}{|N(u)||N(v)|} R_{uv}, \quad (5)$$

Yukarıda, $N(a)$, a 'nın komşu kümesini; $|N(a)|$ bu kümenin büyüklüğünü; V_1 ve V_2 , G_1 ve G_2 'nin düğüm kümelerini tanımlamakta ve $i \in V_1$ ve $j \in V_2$. Ayrıca R fonksiyonel benzerlik matrisi olmakla birlikte bu matrisin durgun durumu ağların sıralaması için kullanılır. Biz yukarıdaki denklemi şöyle de ifade edebiliriz:

$$R = AR. \quad (6)$$

Burada matris elemanları şöyle tanımlanmıştır:

$$A[i, j][u, v] = \begin{cases} \frac{1}{|N(u)||N(v)|} & \text{if } (i, u) \in E_1 \text{ ve } (j, v) \in E_2, \\ 0 & \text{diğer.} \end{cases} \quad (7)$$

Denklemde, A , $|V_1||V_2| \times |V_1||V_2|$ bir stokastik matristir ve giriş grafların komşu matrislerinin Kronecker çarpımı olarak tanımlanabilir: $A = A_1 \otimes A_2$, burada A_i normalizasyonu yapılmış graf G_i 'nin komşuluk matrisidir ve stokastiktir. Bu denklem bir özdeğer problemi ve bu problemde bulunan özvektör durgun durumu tanımlar. Özvektördeki katsayıların düğümler arasındaki benzerlik skorlarını vermektedir. Bu vektörden bir açgözlü algoritma (greedy algorithm) yardımıyla çözüm elde edilebilmektedir.

IsoRank algoritması çok başarılı bir şekilde büyük ağlara uygulanabilmesine rağmen (Bayati vd. 2009), algoritmanın hesaplama karmaşıklığının ağ sayısına bağlı olarak üstsel orantıda büyümesi (Çalışma zamanı, E kenar sayılı m tane ağ için $O(E^m)$), algoritmanın çok sayıda ağa tatbikatını engellemektedir (Nir vd. 2008).

Projenin bu ayağında (bakınız Daskin (2015)), kuantum faz hesaplaması bir önceki kısımda tanımlanan ayarlarla, temel IsoRank (Singh vd. 2007) algoritmasındaki A matrisi için kullanılarak, IsoRank algoritmasının kuantum versiyonu tanımlanmıştır. A matrisi simetrik olmadığı için algoritmada bu matrisin nasıl kullanılabileceği ve bunun sonucu nasıl değiştirdiği incelenmiştir. Ayrıca bir ölçüm stratejisi tanımlanarak sonuç özvektöründen çözümün nasıl elde edilebileceği gösterilmiş ve sonuç olarak algoritmanın kuantum versiyonun klasik versiyona oranla, hesaplama yönünden, üstsel verimlilik sağladığı not edilmiştir. Çalışmanın bu kısmı örnek üç adet PPI ağlarının sıralanması gösterilerek sonuçlandırılmıştır.

5. KUANTUM ASLİ UNSUR İNCELEME

Asli unsur inceleme (Principal component analysis (PCA)), asli unsurları temsil eden elipsoidleri verilen veri kümesine uydurarak, bir problemin boyutunu düşürmek veya çok boyutlu veri kümesinin hulasasını elde etmek için kullanılan çok değişkenli istatistiksel bir tekniktir (F.R.S. 1901; Jolliffe 2002). Asli unsur inceleme, J adet değişkenle ifade edilen I adet farklı gözlemlerdeki değişimi açıklamaktadır. Eğer bu gözlemleri $I \times J$ boyutlu X matrisi ile ifade edersek; PCA, verinin lineer izdüşümünde varyantın (değişimin) maksimum olduğu q sayıda orthonormal eksen bulmaktadır. Verinin asli unsurları olarak isimlendirilen bu q adet eksen, X matrisinin en büyük q adet özdeğeriyle ilişkilendirilmiş özvektörleridir (Jolliffe 2002; Shlens 2014).

Evvelki kısımlarda tartışıldığı üzere kuantum faz hesaplama algoritmasına eşit bir süperpozisyon değeri verildiğinde çıkışta bütün özvektörlerin süperpozisyonu izlenmektedir. Bu proje kapsamında yapılan Daskin (2016a) daki çalışmada, bu özellik kuantum katsayı yükseltme algoritmasıyla birlikte kullanılarak belirli bir aralıktaki özdeğerlere karşılık gelen özvektör birleşiminin elde edilebileceği gösterilmiştir. Dolayısıyla herhangi bir X matrisinin asli unsurlarının hesaplanabileceği gösterilmiştir. Kuantum katsayı algoritması verilen bir kuantum vektöründeki istenilen bir değerin katsayısını işaretleyip onun mutlak değerini yükseltmede kullanılmaktadır. Bu algoritma temelde iki operatöre bağlı olarak çalışmaktadır. Birinci operatör istenilen kuantum durumlarının katsayılarının işaretini değiştirmekte diğeri ise bir nevi işaretlenen bu katsayıların değerlerini yükseltmede kullanılmaktadır. Yapılan çalışmada öncelikle yüksek özdeğerlere karşılık gelen durumlarının işaretlenmesi için birinci operatör tanımlanmış daha sonra ikinci operatörün tatbiki ile katsayılarının nasıl değiştiği analitik ve sayısal olarak gösterilmiştir. Ayrıca tanımlanan operatörler için gerekli kuantum devrelerin tasarımına nasıl gidileceği verilmiş ve bunların hesaplama zorluğu kısmi olarak analiz edilmiştir. Sonuç olarak, tanımlanan metodolojinin nerelerde kullanılabileceği, klasik asli unsur analizinden farkları ve avantajları tartışılmıştır.

6. SONUÇ

Projenin başlangıcında kuantum hesaplamasında kullanılan algoritmaların modifiye edilerek bu algoritmalarındaki bilinen kısıtlamalara çözüm sunulması ve böylelikle kullanım alanlarının genişletilmesi ve buna ek olarak modifiye edilmiş bu algoritmaların gücünden faydalanılarak bioinformatik ve kuantum kimyası alanlarındaki zor problemlere verimli çözümler sunulması amaçlanmıştır. Bu amaca yönelik bu proje kapsamında üç tane çalışma yapılmıştır (Daskin 2015, 2016a,b).

Bu amaçlar doğrultusunda kuantum hesaplamasında kullanılan algoritmaların kısıtlamaları belirlenip bu kısıtlamalara ne türden çözümler sunulabileceği araştırılmıştır. Kuantum algoritmalarından faz hesaplama algoritmasının bilinen kısıtlamalarından bir tanesi, başlangıçta iyi bir tahmini özvektöre ihtiyaç duymasıdır. Projenin ilk aşamasında (Daskin 2016b), geniş kullanım alanlarına sahip olan negatif olmayan matrisler için bu algoritmanın girişteki tahmini özvektörden bağımsız olarak doğru sonuç üretilebileceği gösterilmiştir. Mevzubahis matrislerin çok farklı bilimsel alanlarda kullanımı mevcuttur. Kuantum kimyası ve fizik bunlar arasındadır. Kuantum bilgisayarların önerilmesinin sebeplerinden bir tanesi; özelliklerinin elde edilmesi hesaplama yönünden çok zor olan kuantum sistemlerinin bu bilgisayarlar kullanıldığında daha kolay hale geleceğinin beklenmesidir. Kuantum sistemlerinin enerji ve özelliklerini gösteren kuantum Hamiltonianlarının bir kısmı

stokastik özellikler göstermektedir. Modifiye edilmiş faz algoritmasının bu türden sistemlere uygulanması ve bu sistemlerin anlaşılmasında kullanımı gösterilmiştir.

Negatif matrislerin alt kategorisi niteliğindeki stokastik matrisler biyoinformatikte farklı tür- lere ait protein etkileşim ağlarının mukayesesinde önemli bir kullanım alanına sahiptirler. Protein etkileşim ağlarının kapasitesinin büyüklüğü hesaplama yönünden zorluklar ortaya çıkarmakta ve bu ağların günümüz bilgisayar teknolojisiyle karşılaştırılmasını imkansız hale getirmektedir. Pro- jede modifiye edilen faz hesaplama algoritmasının protein etkileşim ağlarının mukayesesinde etkin bir şekilde kullanılması amaçlanmış ve klasik algoritmalarından IsoRank algoritmasının kullanımda karşılaşılan hesaplama zorluklarının kuantum bilgisayarlar tarafından çözüme kavuşturulabileceği gösterilmiştir (Daskin 2015).

Projenin son kısmında faz hesaplama algoritması katsayı yükseltme algoritması ile birlikte kul- lanılarak bir kuantum asli unsur inceleme algoritması tanıtılmıştır (Daskin 2016a). Asli unsur inceleme algoritmasının makine öğrenmesi ve örüntü tanıma alanlarında sıklıkla kullanılmakta olduğundan, bu çalışmanın ileride kuantum bilgisayarların bu alanlarda da etkin bir biçimde kul- lanılmasına katkıda bulunması beklenmektedir.

7. KAYNAKLAR

- Bapat, R. B. ve Raghavan, T. E. (1997). *Nonnegative matrices and applications*, volume 64. Cambridge University Press.
- Bayati, M., Gerritsen, M., Gleich, D., Saberi, A., ve Wang, Y. (2009). Algorithms for large, sparse network alignment problems. In *Data Mining, 2009. ICDM '09. Ninth IEEE International Conference on* (pp. 705–710).
- Bennett, C. H., Bernstein, E., Brassard, G., ve Vazirani, U. (1997). Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5), 1510–1523.
- Berman, A. ve Plemmons, R. J. (1979). Nonnegative matrices. *The Mathematical Sciences, Classics in Applied Mathematics*, 9.
- Bravyi, S., Divincenzo, D. P., Oliveira, R., ve Terhal, B. M. (2008). The complexity of stoquastic local hamiltonian problems. *Quantum Info. Comput.*, 8(5), 361–385.
- Clark, C. ve Kalita, J. (2014). A comparison of algorithms for the pairwise alignment of biological networks. *Bioinformatics*, 30(16), 2351–2359.
- Daskin, A. (2015). Quantum isorank: Efficient alignment of multiple ppi networks. *arXiv preprint arXiv:1506.05905*.
- Daskin, A. (2016a). Obtaining a linear combination of the principal components of a matrix on quantum computers. *Quantum Information Processing*, (pp. 1–15).

- Daskin, A. (2016b). Quantum eigenvalue estimation for irreducible non-negative matrices. *International Journal of Quantum Information*, 14(01), 1650005.
- Deutsch, D. (1985). Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A: Mathematical and Physical Sciences*, 400(1818), 97–117.
- Feynman, R. (1982). Simulating physics with computers. *International Journal of Theoretical Physics*, 21, 467–488. 10.1007/BF02650179.
- F.R.S., K. P. (1901). Liii. on lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 6*, 2(11), 559–572.
- Gutman, I., Rucker, C., ve Rucker, G. (2001). On walks in molecular graphs. *Journal of Chemical Information and Computer Sciences*, 41(3), 739–745.
- Jolliffe, I. (2002). *Principal component analysis*. Wiley Online Library.
- Kelley, B. P., Yuan, B., Lewitter, F., Sharan, R., Stockwell, B. R., ve Ideker, T. (2004). Pathblast: a tool for alignment of protein interaction networks. *Nucleic Acids Research*, 32(suppl 2), W83–W88.
- Kitaev, A. (1996). Quantum measurements and the abelian stabilizer problem. *Electronic Colloquium on Computational Complexity (ECCC)*, 3(3).
- Klau, G. (2009). A new graph-based method for pairwise global network alignment. *BMC Bioinformatics*, 10(Suppl 1), S59.
- Koyutürk, M., Kim, Y., Topkara, U., Subramaniam, S., Szpankowski, W., ve Grama, A. (2006). Pairwise alignment of protein interaction networks. *Journal of Computational Biology*, 13(2), 182–199.
- Lanyon, B. P., Whitfield, J. D., Gillett, G. G., Goggin, M. E., Almeida, M. P., Kassal, I., Biamonte, J. D., Mohseni, M., Powell, B. J., Barbieri, M., Aspuru-Guzik, A., ve White, A. G. (2010). Towards quantum chemistry on a quantum computer. *Nature Chemistry*, 2(2), 106–111.
- Li, Z., Zhang, S., Wang, Y., Zhang, X.-S., ve Chen, L. (2007). Alignment of molecular networks by integer quadratic programming. *Bioinformatics*, 23(13), 1631–1639.
- Martin-Delgado, M. A. (2012). On quantum effects in a theory of biological evolution. *Scientific Reports*, 2(302).
- Meyer, C. (2000). *Matrix analysis and applied linear algebra book and solutions manual*, volume 2. Society for Industrial and Applied Mathematics.
- Minc, H. (1970). On the maximal eigenvector of a positive matrix. *SIAM Journal on Numerical Analysis*, 7(3), 424–427.
- Moore, G. E. (1965). Cramming more components onto integrated circuits. *G.E. Moore Electronics*, 38(8), 261–276.
- Nielsen, M. A. ve Chuang, I. L. (2010). *Quantum computation and quantum information*. Cambridge university press.
- Nir, Y., Roded, S., ve William Stafford, N. (2008). Improved network-based identification of protein orthologs. *Bioinformatics*, 24(16), i200.
- Page, L., Brin, S., Motwani, R., ve Winograd, T. (1999). *The PageRank Citation Ranking: Bringing Order*

- to the Web. Technical Report 1999-66, Stanford InfoLab. Previous number = SIDL-WP-1999-0120.
- Powell, J. (2008). The quantum limit to Moore’s law. *Proceedings of the IEEE*, 96(8), 1247–1248.
- Shlens, J. (2014). A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*.
- Shor, P. W. (1994). Algorithms for quantum computation: Discrete logarithms and factoring. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on* (pp. 124–134).: IEEE.
- Singh, R., Xu, J., ve Berger, B. (2007). Pairwise global alignment of protein interaction networks by matching neighborhood topology. In T. Speed ve H. Huang (Eds.), *Research in Computational Molecular Biology*, volume 4453 of *Lecture Notes in Computer Science* (pp. 16–31). Springer Berlin Heidelberg.
- van Gunsteren, W. F. ve Berendsen, H. J. C. (1990). Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. *Angewandte Chemie International Edition in English*, 29(9), 992–1023.
- Zaslavskiy, M., Bach, F., ve Vert, J.-P. (2009). Global alignment of protein-protein interaction networks by graph matching methods. *Bioinformatics*, 25(12), i259–1267.

TÜBİTAK
PROJE ÖZET BİLGİ FORMU

Proje Yürütücüsü:	Yrd. Doç. Dr. ANMER DAŞKIN
Proje No:	115E747
Proje Başlığı:	Kuantum Hesaplamanın Bioinformatik Ve Kimyadaki Problemlere Uygulanması
Proje Türü:	1002 - Hızlı Destek
Proje Süresi:	8
Araştırmacılar:	
Danışmanlar:	
Projenin Yürütüldüğü Kuruluş ve Adresi:	İSTANBUL MEDENİYET Ü. MÜHENDİSLİK MİMARLIK F. BİLGİSAYAR MÜHENDİSLİĞİ B.
Projenin Başlangıç ve Bitiş Tarihleri:	01/01/2016 - 01/09/2016
Onaylanan Bütçe:	24108.0
Harcanan Bütçe:	24100.0

TÜBİTAK

Öz:	Kuantum bilgisayarlar kuantum mekaniksel özelliklerini kullanarak daha verimli algoritmalar tasarlama imkanı sunarlar. Bu münasebetle, kuantum bilgisayarların hesaplama gücünün günümüzde kullanılan konvansiyonel bilgisayarlara kıyasla daha kuvvetli olduğuna inanılmaktadır. Verilen bir matrisin özdeğerini (eigenvalue) bulmakta yada diğer algoritmaların temel parçası olarak kullanılmakta olan kuantum faz hesaplama algoritması hesaplama yönünden verimli olan en meşhur kuantum algoritmalarından bir tanesidir. Bu algoritmanın tatbiki başarısının girişte verilen tahmini özvektöre(eigenvector) bağlı olması, algoritmanın kullanım alanlarını daraltmaktadır. Desteklenen proje kapsamında, kuantum faz hesaplama algoritması modifiye edilerek bu algoritmadaki bilinen kısıtlamalara çözüm sunulması ve böylelikle kullanım alanlarının genişletilmesi ve buna ek olarak modifiye edilmiş bu algoritmanın gücünden faydalanılarak biyoinformatik ve kuantum kimyası alanlarındaki zor problemlere verimli çözümler sunulması amaçlanmıştır. Bu amaç doğrultusunda, üç adet çalışma yapılmıştır. Faz hesaplama algoritmasında girişte verilen tahmini özvektörün gerçek özvektöre yakınlığı algoritmanın başarısını belirlemektedir. İlk çalışmada, algoritmanın negatif olmayan matrislere tatbikatında giriş vektörünün rolü incelenerek; söz konusu matrislerin dominant özdeğer-özvektör çiftinin pozitif olması ve diğer özvektörlerinde de en az bir negatif katsayı barındırması özelliği kullanılmış ve girişte verilen tahmini özvektörün pozitif özvektöre daha yakın olması sağlanabileceği ve dolayısıyla algoritmanın indirgenemeyen ve negatif olmayan matrisler için verimli bir şekilde kullanılabileceği gösterilmiştir. Örnek olarak stokastik özellik gösteren k-lokal Hamiltonianlar ve rastgele sayılar kullanılmış ve özdeğerlerinin hesaplanmasında, Farklı türden canlılar için protein protein etkileşim ağlarının mukayeseli analizleri biyolojik süreçlerin anlaşılmasında önemli katkı sağlamaktadır. Bu ağlar arasındaki benzerliklerin anlaşılması ayrıca farklı türler arasında fonksiyonel benzerliklerin bulunmasına yardımcı olmaktadır. Ancak devamlı büyüyen etkileşim verilerinin kapasitesi günümüzde hesaplama yönünden zorluklar ortaya çıkarmaktadır. Protein protein etkileşim ağlarının mukayesesinde yaygın bir şekilde kullanılan algoritmalarından IsoRank diye isimlendirilen bir tanesi, ağlar arasındaki etkileşim skorlarını bu ağların Kronecker (tensor) çarpımından elde edilmiş bir stokastik matrisin özvektöründen hesaplanmaktadır. Stokastik matrisler ayrıca pozitif olmaları sebebiyle negatif olmayan matrisler grubuna dahildirler. Proje kapsamındaki ikinci çalışmada \citep{Daskin2015IsoRank}, protein protein etkileşim ağlarının mukayese edilmesinde; kuantum faz hesaplama yöntemi kullanılarak, klasik IsoRank algoritmasının kuantum versiyonu önerilmiş ve bu ağların sıralanmasının üstsel olarak daha verimli bir şekilde kuantum bilgisayarlarca yapılabileceği gösterilmiştir. Son çalışmada ise kuantum faz hesaplama algoritması kuantum katsayı yükseltme algoritması ile birlikte kullanılarak kuantum asli unsur inceleme (quantum principal component analysis) usulü tanımlanmıştır. Yapılan çalışmalar, kuantum bilgisayarların biyoinformatik ve kuantum kimyasının yanı sıra benzeri alanlardaki problemlere tatbikatına da, imkan sağlamaktadır.
Anahtar Kelimeler:	Kuantum Faz Hesaplama, Negatif Olmayan Matrisler, Kuantum IsoRank, Kuantum Asli Unsur İnceleme
Fikri Ürün Bildirim Formu Sunuldu Mu?:	Hayır