IMAGE: A MAP OF THE STARS OF THE ORION CONSTELLATION

Print ISSN: 2631-8490 Online ISSN: 2631-8504

JournalPreview

London Journal of Research in Science: Natural and Formal
Volume 23 | Issue 12 | Compilation 1.0

Great Britain Journals Press

WWW.JOURNALSPRESS.COM

Journal Content
In this Issue

Great Britain Journals Press

The Effect of Gravity on the Human Spirit

Ardeshir Irani

ABSTRACT

Time is a relative phenomenon between two regions of space with different values of gravity because the rate of passage of time depends on the strength of the gravity field. In high gravity fields time slows down compared to low gravity fields where time speeds up. This implies that if one goes from living in a high gravity region to a low gravity region then one sees the future. On the other hand, going from a low gravity region to a high gravity region one sees the past. Gravity has the same effect on the human spirit as it has on the human body and on non-living things, all of which depend on the rate of passage of time. On earth the spirit in the human body is affected by the earth's stronger gravity while a spirit without a body residing in the vacuum space of the quantum region where gravity has a negligible effect that depends only on the gravity due to the spirit's mass-energy. Hence spirits in the human body on earth will age slower than spirits in the quantum world. Based on the calculated spirit's mass-energy we prove that the spirit is quantum mechanical in nature.

Keywords: gravity, time dilation, human spirit, twin paradox, causality, exorcism.

Classification: LCC Code: QB1-QB991

Language: English

Great Britain Journals Press

LJP Copyright ID: 925611
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural and Formal

Volume 23 | Issue 12 | Compilation 1.0

The Effect of Gravity on the Human Spirit

Ardeshir Irani

ABSTRACT

Time is a relative phenomenon between two regions of space with different values of gravity because the rate of passage of time depends on the strength of the gravity field. In high gravity fields time slows down compared to low gravity fields where time speeds up. This implies that if one goes from living in a high gravity region to a low gravity region then one sees the future. On the other hand, going from a low gravity region to a high gravity region one sees the past. Gravity has the same effect on the human spirit as it has on the human body and on non-living things, all of which depend on the rate of passage of time. On earth the spirit in the human body is affected by the earth's stronger gravity while a spirit without a body residing in the vacuum space of the quantum region where gravity has a negligible effect that depends only on the gravity due to the spirit's mass-energy. Hence spirits in the human body on earth will age slower than spirits in the quantum world. Based on the calculated spirit's mass-energy we prove that the spirit is quantum mechanical in nature.

Keywords: gravity, time dilation, human spirit, twin paradox, causality, exorcism.

Author: The Dark Energy Research Institute, 12343 Chavers Avenue, Downey, Ca. 90242, USA.

I. MAIN TEXT

All spirits in the Multiverse have a limited life span depending on the life span of the Multiverse (Reference1), after which their energy is sent back into the infinite void as pure massless energy in the form of photons. Since different spirits age differently, depending on the time they reside in the classical material world or in the quantum spiritual world with its minimal gravity, implying that spirits can be recycled because of the Conservation of Energy Principle since the Energy of the spirit cannot be destroyed but it can be transformed from one form to another.

According to the Twin Paradox, one twin leaves earth in a spaceship destined for a planet near a Black Hole with very high gravity. He stays there for a few years and then returns home to earth to discover that while he is aged 5 years his twin on earth is aged 25 years. Higher gravity fields slow down time and the aging process. The traveling twin now sees the future because everything around him, including the place he was born in, has grown older. From the point of view of the stay-at-home twin he sees the past because his twin is younger, and his spaceship is a younger model of newly built spaceships.

Let us next look at the future of two twin spirits. One twin spirit decides to spend all its time in the quantum region while the other twin spirit decides to spend most of its time on earth by being born-again every time it leaves its material body for good. The born-again twin's spirit will age much slower and hence it will remain much younger due to the earth's greater gravity, and that is the benefit it gets from enduring the harsh realities of earthly existence. Its spirit will last much longer than that of its twin whose older spirit will be recycled sooner since it has aged faster. It should be noted that the life span of the spirit is much longer than the life span of the human body, and hence the same spirit can be born-again and again in different bodies, both male and female.

London Journal of Research in Science: Natural and Formal

While the human body is classical matter, the human spirit is quantum mechanical energy. The spirit remains in the classical world while in the active human body but can enter the quantum mechanical world instantaneously when it leaves the body. This can happen while the person is sleeping or has died. According to Quantum Mechanics, quantum objects can communicate and travel over huge spatial distances instantaneously as opposed to classical objects that cannot exceed the speed of light (Reference 2). As an estimate let us calculate how much mass a human spirit carries with it when it leaves the body. An 80 kg person falling to earth from 20 m would die instantly. This would indicate that E = 15,680 Joules of energy that killed him from the fall has increased his energy since Energy must be Conserved, and the spirit leaving his body would carry a mass equivalent according to E = m c^2 of m = $1.74x10^{-13}$ kg which is an upper limit since some of the energy from the fall would go into broken bones and dissipated heat energy. A more accurate estimate of spirit mass-energy based on levitation will be discussed at the end of the conclusion section. This small amount of mass that the spirit carries with it to the quantum region is responsible for aging the spirit since it has a gravity field of its own which can be referred to as "quantum gravity" since the spirit is quantum energy.

According to gravitational time dilation (Reference 3), $(t_1 - t_2)/\, t_2 = (\phi_1 - \phi_2)/c^2$. A clock at higher gravitational potential $\phi_1 > \phi_2$ will run faster $t_1 > t_2$ which means time will slow down because time between two ticks of the clock becomes quicker. The time between two ticks determines the passage of time and the aging process. When time between two ticks becomes zero as in the case of photons then time stands still or does not change and hence photons live for an infinite amount of time unless they interact with a material object to transfer their energy to that object. Time slows down on earth due to its larger gravity than time in the spirit quantum world. $\phi_1 = 6.25x10^7 m^2/s^2$ for the surface of the earth and $\phi_2 \approx 0$ for the vacuum quantum region. For significant time dilation the time for light to travel from earth to the quantum region must be large. As an example, for a time dilation $(t_1 - t_2)$ of 1 second since $(t_1 - t_2)/\, t_2 = \phi_1/c^2$, the distance $d_2 = t_2 c$ given by $d_2 = (t_1 - t_2)c^3/\, \phi_1 = 4.32x10^{17}$ meters or 45.6 light years. Once the light signal gets to the quantum region after 45.6 years the spirit can respond immediately but with a time dilation of 1 second which means the spirit will arrive 1 second in the past of current earth time. When grandfather died and went to the spirit world, he could see 45.6 years of the past earth years to relive his own youth visually. Waiting for 45.6 years he would see his own death when his granddaughter was 6 years old. He observed his granddaughter from the spirit world and noticed that she was about to step off the curb into the path of a speeding car. Grandfather's spirit instantaneously returned to earth entering the body of a person standing next to his granddaughter to pull her back to safety. Since Causality cannot be violated, whether he was able to save his granddaughter's life would depend on one of two factors: (1) His granddaughter was able to grow up in the interim 45.6 years to become a woman aged 51.6 years or (2) There exists a child's grave with her name on it. What has already happened cannot be changed or more specifically the past cannot change the future. The fact that Causality cannot be violated is true in all the physical Universe. Grandfather's example indicates that the spirits in the quantum world are not devoid of Feelings. This implies that the Energy of the spirit can carry intangibles such as Feelings, Intelligence, and Consciousness with it, qualities which all human beings have been endowed with, which implies that the spirit within the human being controls all these aspects of the mind and body. These intangibles remain locked in the spirit and begin to show up again in the born-again person to help the spirit evolve further. The further out in space you go the greater the time dilation which occurs not only due to gravity but also due to the speed of light because velocity also slows down time. For time dilation of 1 hour the distance between earth and the spirit world would be 164,160 light years. If the spirit located further out in space arrives

instantaneously to earth, after observing it for 164,160 years, it will arrive 1 hour in the past of current earth time. Whatever occurred either in the past or in that period of 1 hour due to time dilation cannot be changed by the spirit because of Causality.

II. CONCLUSION

There exist different types of spirits hovering all around us in the physical world. All spirits are ruled by the laws of nature, both classical and quantum mechanical. However, each spirit is different, just as each person has a different personality. Extreme variations in behavior, way beyond normal, for a brief period indicates that a different spirit, that may be good or bad, entering the human body has taken control of it and is responsible for causing these changes. While the person's own spirit is forced to stay dormant or leave the body, one or more spirits entering the person can control the mind and body for a brief period or much longer. As an extreme example, the movie Exorcist depicts the physical damage a troubled spirit can do by controlling the body and mind over a long period and refusing to leave on its own until an exorcism has been performed by a priest trained to perform the ritual to drive the evil spirit out. The Catholic Church refers to this control as the person being possessed by a demon. The possessed person has no memory of the cursing, trashing, and levitation that occurred during the exorcism. To levitate an 80 kg person 1 m above ground level for several minutes which has been stated to be an observed phenomenon during exorcism, the demonic spirit would have to possess an energy of E = 784 Joules to cancel the effect of gravity or more exactly a mass equivalent of 8.71×10^{-15} kg. This mass is smaller than Planck's mass of 2.18×10^{-8} kg (Reference 4) and therefore qualifies the spirit as being quantum mechanical in nature.

REFERENCES

1. Irani, A. (2021) Dark Energy, Dark Matter, and the Multiverse. Journal of High Energy Physics, Gravitation and Cosmology, 7, 172-190.
2. Irani, A. (2022) Faster than the speed of light is a Quantum Phenomenon. Journal of High Energy Physics, Gravitation and Cosmology, Vol.9, No.2, Apr. 2023.
3. Ta-Pei-Cheng (2015) A College Course on Relativity and Cosmology. Oxford University Press, Oxford, Pgs.64-69.
4. Ryden, Barbara (2018) Introduction to Cosmology. Cambridge University Press, Cambridge, Pg.3.

This page is intentionally left blank

Green Synthesis of Bimetallic Nanoparticle using Algae Extracts with the Evaluation of its Potential Applications

Preeti Maurya, Khushaboo Soni & Sanjay Singh

University of Allahabad

ABSTRACT

Wide-ranging usages of nanomaterials have become physiologically safe as a result of the global focus on studying green nanotechnology. The creation of nanoparticles with a regular shape and set of characteristics is the focus of nanotechnology. According to recent research in the field of nanotechnology, the size, shape, composition, level of crystallinity, and stability of metal nanoparticles are all factors that affect how they behave. Due to their great biocompatibility, stability, and relatively low toxicity, bimetallic nanoparticles are essential. Bimetallic nanostructures on graphene, zeolites, clays, fibers, and polymers, as well as non-supported bimetallic nanoparticles, are reviewed. Their production processes, resultant characteristics, and antibacterial activity are all presented. The ratio of the two metals in a bimetallic nanoparticle population as well as the internal distribution of the elements inside specific nanoparticles (such as the presence of homogeneous alloys, core-shell systems, and potential intermediary phases) are the two main elements that make up this population. Biological applications (in medicine and agriculture), environmental applications (in water treatment and removal of toxic contaminants), engineering applications (in nanosensors, nanochips, and nano-semiconductors), and chemical and physical applications (in optics, catalysis, and paints) are just a few of the many uses they have thanks to their synergistic properties.

Keywords: bimetallic nanoparticles, antidiabetic, green synthesis, nanoscale, biocompatibility.

Classification: LCC Code: T174.7

Language: English

Great Britain Journals Press

LJP Copyright ID: 925612
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural and Formal

Volume 23 | Issue 12 | Compilation 1.0

Green Synthesis of Bimetallic Nanoparticle using Algae Extracts with the Evaluation of its Potential Applications

Preeti Maurya[α], Khushaboo Soni[σ], & Sanjay Singh[ρ]

ABSTRACT

Wide-ranging usages of nanomaterials have become physiologically safe as a result of the global focus on studying green nanotechnology. The creation of nanoparticles with a regular shape and set of characteristics is the focus of nanotechnology. According to recent research in the field of nanotechnology, the size, shape, composition, level of crystallinity, and stability of metal nanoparticles are all factors that affect how they behave. Due to their great biocompatibility, stability, and relatively low toxicity, bimetallic nanoparticles are essential. Bimetallic nanostructures on graphene, zeolites, clays, fibers, and polymers, as well as non-supported bimetallic nanoparticles, are reviewed. Their production processes, resultant characteristics, and antibacterial activity are all presented. The ratio of the two metals in a bimetallic nanoparticle population as well as the internal distribution of the elements inside specific nanoparticles (such as the presence of homogeneous alloys, core-shell systems, and potential intermediary phases) are the two main elements that make up this population. Biological applications (in medicine and agriculture), environmental applications (in water treatment and removal of toxic contaminants), engineering applications (in nanosensors, nanochips, and nano-semiconductors), and chemical and physical applications (in optics, catalysis, and paints) are just a few of the many uses they have thanks to their synergistic properties.

Keywords: bimetallic nanoparticles, antidiabetic, green synthesis, nanoscale, biocompatibility.

*Author*α σ*:* Department of Botany, CMP Degree College, University of Allahabad, Prayagraj 211002 Uttar Pradesh india.

I. INTRODUCTION

The area of study of nanoparticles is known as nanotechnology. Thus, Nobel Prize winner Richard P. Feynman introduced nanotechnology in his 1959 speech "There's Plenty of Room at the Bottom" [1]. In the year 2008, the International Organization for Standardization (ISO) described a discrete nanoparticle as a nano-object with all three Cartesian dimensions less than 100nm. A more technical but broader definition was adopted by the European Union's commission in 2011: a natural, incidental, or manufactured material containing particles, in an unbound state or as an aggregate, where 50% or more of the particles in the number size distribution have one or more external dimensions that fall within the size range of 1 nm to 100 nm [2, 3]. Nanotechnology, as its name suggests, is a field of technology that works with matter at the nanoscale. It is a term used to describe the area of science and engineering that deals with manipulating atoms or molecules at the nanoscale to create systems and devices for diverse uses. The nanoparticles are the fundamental component of nanotechnology. These nanoparticles are small in size, have a large surface area, and show a characteristic known as quantum effects, which denotes erratic or unpredictable behavior [4,5].

Nanotechnology is an enabling technology that works with items as small as a nanometer. Three levels of development are anticipated for nanotechnology: materials, devices, and systems. The creation of

nanoparticles using biological entities has caught the interest of many researchers [7] due to their distinctive shape-dependent optical, electrical, and chemical [6] properties that have potential applications in nanobiotechnology [8].

The Environmental Protection Agency (EPA) asserts that NMs *"may display distinct features differing from the identical chemical molecule in a greater dimension*." The US Food and Drug Administration (USFDA) defines NMs as *"materials that have at least one dimension in the range of about 1 to 100 nm and show dimension-dependent behaviours*." The International Organization for Standardization (ISO), in a similar vein, has referred to NMs as *"materials with any outward nanoscale dimension or possessing internal nanoscale surface structure*." The definitions of nanofibers, nanoplates, nanowires, quantum dots, and other related words are based on this ISO standard. Similar to how the word "nanomaterial" is defined as *"a manmade or natural material that comprises unbound, aggregated, or agglomerated particles with exterior diameters in the range of 1-100 nm",* according to the EU Commission.

The British Standards Institution recently offered the following definitions for the terminology used in science:

- Nanoscale: a size range of around 1 to 1000 nm.
- Nanoscience: The study of matter at the nanoscale that focuses on comprehending its size- and structure-dependent features and examines the emergence of individual atoms or molecules or changes in bulk materials.
- Nanotechnology: Using scientific knowledge for a variety of industrial and healthcare purposes, it is the manipulation and control of matter on the nanoscale dimension.
- Materials with any internal or exterior structures on the nanoscale dimension are considered nanomaterials.

A material with one or more peripheral nanoscale dimensions is referred to as a nano-object.

Three exterior nanoscale dimensions are present in a nanoparticle, a nano-object. When nano-longest objects and shortest axes are different lengths, the words "nanorod" or "nanoplate" is used in place of "nanoparticle" (NP).

- Nanofiber: A nanomaterial is referred to as nanofiber if it has three dimensions total - two outside nanoscale dimensions that are comparable and one larger dimension.
- A multiphase structure containing at least one nanoscale phase is referred to as a nanocomposite.
- Nanostructure: a structure made up of interconnecting nanoscale building blocks.
- Nanostructured materials: Substances with either internal or external nanostructure [9,3]

According to their structure (metal base, carbon base, dendrimers, or liposome), dimension (zero, one, two, or three dimensions), or origin (natural or manmade), nanoparticles can be categorized into various types [10]. Metal-based nanoparticles, particularly those made of noble metals, provide greater benefits than other forms of nanoparticles. This is due to the very stable, biocompatible, and potential for large-scale manufacture of metal-based nanoparticles for use in biomedical and environmental applications [11]. The use of metal-based nanoparticles needs to be enhanced in several fields of study despite their intriguing capabilities and characteristics due to their toxicity, big size, cellular absorption, and chemical stability [12, 13]. Therefore, those restrictions must be overcome. Any two metals can be combined to create a bimetallic nanoparticle, which has intriguing and synergistic features that result in new, improved structural and physical properties, boosting its functionality and applicability [10]. The experimental results have also revealed a surprising enhancing capability and a potential workaround for the monometallic nanoparticles' drawbacks [4].

Due to their distinctive optical, electrical, magnetic, and catalytic properties—which, in most cases, are markedly different from those of their monometallic counterparts—bimetallic NPs have attracted considerable attention in the academic and technological fields over the past ten years. Two different kinds of metal nanoparticles are combined to create bimetallic NPs, which can have a wide range of morphologies and architectures [14]. They frequently display more fascinating features than the corresponding monometallic NPs, a phenomenon that is explained by the synergistic properties between the two separate metal components. By choosing the right metal combination and support as well as enhancing the composition of each metal type, characteristics, and performance may be tuned. Bimetallic nanostructures can be divided into two groups, namely mixed and segregated ones, which can then be further divided into alloy, intermetallic, subcluster, and core-shell types based on their atom configuration [15].

1.1 Green synthesis of bimetallic nanoparticles

The bottom-up approach needs a good soluble source of metals, often metal cations in the form of soluble salts or coordinated by appropriate ligands. Such a solution has a reducing agent added to it whose type has a significant impact on the particle characteristics. Many techniques use reducing agents such as sodium borohydride, glucose, or citrate. The so-called polyol approach, which uses high-boiling alcohol as the solvent and a reducing agent at the same time, is an exception. Surfactants, polymers, and polyelectrolytes, which are ideal capping agents, are frequently used to regulate particle development and colloidal stability. The type of nanoparticles that are created relies on several factors, including temperature, duration, and reagent concentrations. These are frequently difficult to properly manage to produce monodisperse nanoparticle populations. After all, the basic processes that control crystal growth and nucleation are complicated and yet poorly understood [16]. *Plectonema boryanum* (Cyanobacteria) produces intracellular silver nanoparticles [17, 18]. Algae such as *Sargassum wightii* [19] and *Chlorella vulgaris* can produce Au nanoparticles [20].

1.2 Green synthesis of ZnO NPs using microalgae and macroalgae

Algae are a class of photosynthetic organisms that can be either single-celled (like chlorella) or multicellular (like brown algae). Basic plant components like roots and leaves are absent in algae. The color of the marine algae—Rhodophyta has red pigment, Phaeophyta has brown pigment, and Chlorophytes has green pigment—is used to classify them. Algae have been widely used in the manufacture of Au and Ag nanoparticles, but their use in the synthesis of ZnO nanoparticles is more restricted and has been documented in fewer works [21]. Because of its capacity to break down hazardous metals and transform them into less toxic forms, microalgae have received particular attention [22]. *Sargassum muticum* and *S. myriocystum*, both members of the Sargassaceae family, were employed to synthesize ZnO NP. Using XRD and FE-SEM, *Sargassum muticum* researchers looked at the size of NPs, which revealed similar ranges and a hexagonal wurtzite structure with hydroxyl groups and sulfated polysaccharides present. *S. myriocystum* used DLS and AFM to compare sizes, revealing distinct size ranges with the presence of hydroxyl and carbonyl stretching in NPs with a wide range of shapes [24]. Some of the macro- and microalgae are used in the synthesis of ZnO NP [25].

1.3 Biological Method of bimetallic nanoparticles

Typically, creating nanomaterials using physical or chemical means is pricy, arduous, time-consuming, and harmful to the environment. In addition, they create highly poisonous by-products, demand a lot of energy, and could be dangerous to people's health. Therefore, a quicker, less expensive synthesis method that can get beyond these constraints is required. The biological method of making nanoparticles, often known as the "green synthesis method," is an additional and more biocompatible

way to make nanomaterials. With the green synthesis process, nanoparticles are created without the use of risky or expensive chemicals. Instead, natural resources are employed to create the nanoparticles, resulting in a more ecologically and biologically friendly final product. In general, there are two approaches to producing biological synthesis: either by employing microorganisms (bacteria, fungi, yeast, etc.) or plants (i.e., leaves, stems, fruits, seeds, bark, peels, shoots, roots, etc.) as the reducing and stabilizing agent. This method of creating nanoparticles is known as creating biogenic nanoparticles or biogenic nanomaterials [27, 4].

Table 1: Different types of bimetallic nanoparticles by green synthesis method

Forms of Bimetallic Nanoparticles	Type of Bimetallic Nanoparticle	Structure of the Bimetallic Nanoparticle	Method of Synthesis	References
Gold-based	Au-Pt	Cubic crystal structure	Green synthesis	[26]
	Au-Ag	Alloy & Core-shell	Green synthesis	[27]
Silver-based	Ag-Cu	Alloy	Green synthesis	[28]
	Ag-Au	Alloy	Green synthesis	[29]
	Ag-Fe	Spherical shape	Green synthesis	[30]
	Ag-Pd	Cubic Crystalline structure	Green synthesis	[31]
	Ag-Zn	Wurtzite hexagonal	Green synthesis	[32]
Copper based	Cu-Ag	Alloy	Green synthesis	[33]
Iron-based	Fe-Zn	Spherical	Green synthesis	[34]
	Fe-Cu	Crystalline	Green synthesis	[35]
Platinum-based	Pt-Pd	Crystalline	Green synthesis	[36]
Palladium based	Pd-Ag	Spherical	Green synthesis	[37]
	Pd-Cu	Crystalline alloy	Green synthesis	[38]

1.4 Potential Applications of Bimetallic Nanoparticles

Numerous medical fields, such as diagnostic (bio-imaging), therapeutic (cancer therapy), and preventative (antimicrobial, antioxidant, and antidiabetic drug delivery) use bimetallic nanoparticles (Table 2). Bimetallic nanoparticles made of Au-Fe and Ni-Co are strongly magnetic, making them appropriate for use as contrast agents in CT and MRI imaging for diagnosis and prognosis [39] as well as theranostic agents for malignancies [40]. Similar to Cu-Fe, Pd-Pt, Au-Co, Au-Co, Ag-Cu, and Au-Pt, these materials have been exploited for cancer therapies and anticancer action [42, 43–46] as well as Au-Bi for the suppression of tumor cells [47]. Bimetallic nanoparticles are widely employed in preventive medicine, and many of them serve as antibacterial, antioxidant, anti-diabetic, anti-Alzheimer, anti-inflammatory, and drug-delivery agents [4].

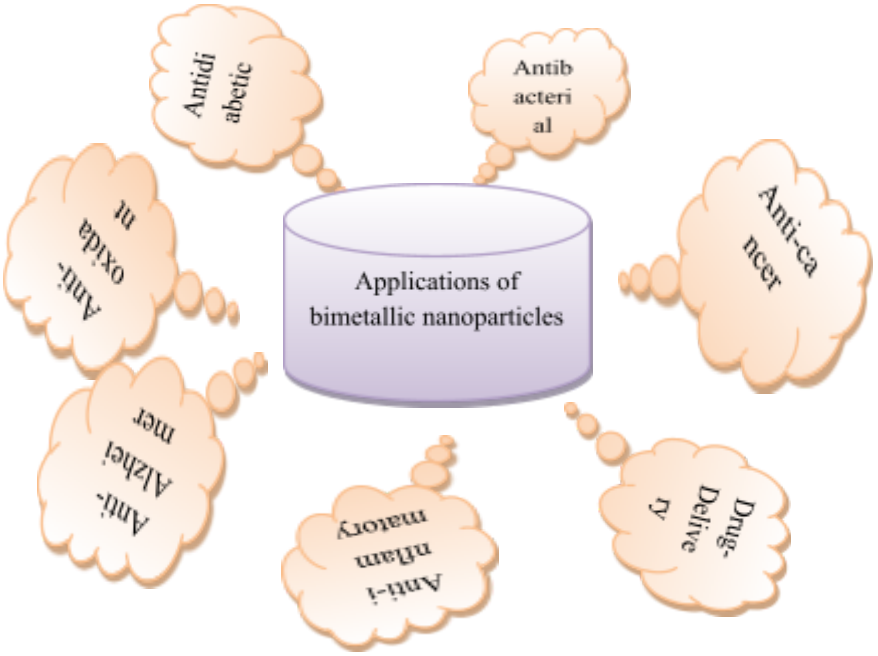

Fig. 1: Potential Applications of Bimetallic Nanoparticles

Table 2: Biological Applications of Bimetallic nanoparticles [4].

Applications	Bimetallic Nanoparticles	References
Antimicrobial agents	Pd-Pt	[48]
	Ag-Fe	[49]
	CuO-NiO	[50]
	Ag-Au	[51-54]
	Cu-Ag	[55]
	Ag-Cu	[56]
	Au-Pt	[57]
	Cu-Zn	[58,59]
	Cu-Ni	[59]
Antioxidant	CuO/NiO	[50]
	Ag-Cu	[60]
	Au-Ag	[61]
	Mn-Cu	[62]
	Pt-Pd	[63]
	Ag-Au	[64]
Antidiabetic	Au-Ag	[65]
	Ag-Au	[66]
	Ag/ ZnOVI	[67]
	Ag-ZnO	[68]
Anti-Alzheimer	Ag-Au	[66]
Anti-Inflammatory	Ag/ ZnOVI	[67]
	Zn-Fe2O4	[68]
	CuFeO2	[69]
	Au@Ag	[70]

Drug delivery	Au@Pd	[71]
	Au-Pt	[72]
	Pd-Pt	[73]

1.5 Prospects

A considerable amount of work has been reported for bimetallic nanoparticles. In summary, they extend the field of potential applications beyond monometallic nanoparticles. It will be interesting to blend three or more metals in one nanoparticle. This touches the field of high-entropy alloys in materials science. The synthesis and the structural characterization will be more complex, but a possible fine-tuning of the properties could give rise to new applications, for example, in heterogeneous catalysis. There is the perspective to make the bimetallic nanoparticles smaller until they reach the size of ultrasmall particles which meet the area of atom-sharp clusters. In this case, the particle diameters are about 2 nm and below. From such particles, novel applications (e.g., imaging in cell biology) and better cell wall permeation can be expected.

The current advancements in nanotechnology suggest a sustainable development in the green synthesis of bimetallic nanoparticles (BMNPs) through green approaches. Though challenging, nano phytotechnology has versatile methods to achieve desired unique properties like optic, electronic, magnetic, therapeutic, and catalytic efficiencies. The review also highlights the prospective future direction to improve the reliability, and reproducibility of biosynthesis methods, their actual mechanism in research works, and the extensive application of biogenic bimetallic NPs [74, 75].

II. CONCLUSION

The field of nanotechnology has witnessed considerable advancement over the last decade. It has been applied in many areas, including biology, medicine, engineering, environment, physics, and chemistry. This is because of the fascinating and synergistic effect between the two metals. This review comprehends the overview of bimetallic nanoparticles, types, synthesis, characterization, application, and toxicity. The toxicity of NPs of the eukaryotic cell is a legitimate concern and remains uncharacterized. In recent years, the vast spectrum of AgNPs used in medicine, cosmetics, biosensors, therapies, and other fields has prompted the development of innovative green production techniques. This review article emphasizes the numerous uses for silver nanoparticles, pointing out that the most widely used way of producing and extracting silver nanoparticles is through the synthesis of algae. A novel, developing method enables the controlled and high-quality production of algae-mediated nanoparticles (NPs), which enhances the properties and usefulness of these NPs for commercial use.

REFERENCES

1. Khan I, Saeed K, Khan I. Nanoparticles: Properties, applications, and toxicities. Arabian journal of chemistry. 2019 Nov 1;12(7):908-31, https://doi.org/10.1016/j.arabjc.2017.05.011.
2. King, S., Jarvie, H., and Dobson, P., (2019, May 14). nanoparticle. Encyclopaedia Britannica. https://www.britannica.com/science/nanoparticle.
3. Singh S., Maurya, P.,& Soni, K.(2023). Nanoparticles: Their Classification, Types and Properties. International Journal of Innovative Research in Technology,9(8),159-166.http://ijirt.org/Article?manuscript=157866
4. Idris, D. S., & Roy, A. (2023). Synthesis of Bimetallic Nanoparticles and Applications—An Updated Review. *Crystals*, *13*(4), 637. https://doi.org/10.3390/cryst13040637
5. Rotello, V. (Ed.). (2004). *Nanoparticles: building blocks for nanotechnology*. Springer Science & Business Media.ISBN 978-0-306-48287-8.

6. Lengke MF, Fleet ME, Southam G (2007) Biosynthesis of silver nanoparticles by filamentous cyanobacteria from a silver (I) nitrate complex. Langmuir 23 (5): 2694–2699. https://doi.org/10.1021/la0613124.
7. Sharma, A., Sharma, S., Sharma, K., Chetri, S. P., Vashishtha, A., Singh, P. & Agrawal, V. (2016). Algae as crucial organisms in advancing nanotechnology: a systematic review. Journal of applied phycology, 28 (3), 1759-1774.
8. Sanjay Singh, Preeti Maurya, Khushaboo Soni. Utilization of Algae for the Green Synthesis of Silver Nanoparticles and Their Applications. *American Journal of Nano Research and Applications*. Vol. 11, No. 1, 2023, pp. 1-9. doi: 10.11648/j.nano.20231101.11
9. Jeevanandam, J., Barhoum, A., Chan, Y. S., Dufresne, A., & Danquah, M. K. (2018). Review on nanoparticles and nanostructured materials: history, sources, toxicity, and regulations. Beilstein journal of nanotechnology, 9, 1050–1074. https://doi.org/10.3762/bjnano.9.98.
10. Mazhar, T., Shrivastava, V., & Tomar, R. S. (2017). Green synthesis of bimetallic nanoparticles and its applications: a review. Journal of Pharmaceutical Sciences and Research, 9(2), 102.
11. Klębowski, B., Depciuch, J., Parlińska-Wojtan, M., & Baran, J. (2018). Applications of noble metal-based nanoparticles in medicine. International journal of molecular sciences, 19(12), 4031.
12. Kumar, H., Venkatesh, N., Bhowmik, H., & Kuila, A. (2018). Metallic nanoparticle: a review. Biomed. J. Sci. Tech. Res, 4(2), 3765-3775.
13. Auffan, M., Rose, J., Wiesner, M. R., & Bottero, J. Y. (2009). Chemical stability of metallic nanoparticles: a parameter controlling their potential cellular toxicity in vitro. Environmental Pollution, 157(4), 1127-1133.
14. Belenov, S. V., Volochaev, V. A., Pryadchenko, V. V., Srabionyan, V. V., Shemet, D. B., Tabachkova, N. Y., et al. (2017). Phase behavior of Pt-Cu nanoparticles with different architecture upon their thermal treatment. Nanotechnol. Russia 12, 147–155. doi:10.1134/S1995078017020033
15. Arora, N., Thangavelu, K., & Karanikolos, G. N. (2020). Bimetallic nanoparticles for antimicrobial applications. *Frontiers in Chemistry*, *8*, 412.doi.org/10.3389/fchem.2020.00412
16. Loza, K., Heggen, M., & Epple, M. (2020). Synthesis, structure, properties, and applications of bimetallic nanoparticles of noble metals. *Advanced functional materials*, *30*(21), DOI: 10.1002/adfm.201909260
17. Ramezani, F., Jebali, A., &Kazemi, B. (2012). A green approach for the synthesis of gold and silver nanoparticles by Leishmania sp. *Applied biochemistry and biotechnology*, *168*, 1549-1555.DOI 10.1007/s12010-012-9877-3
18. Klaus, T., Joerger, R., Olsson, E., & Granqvist, C. G. (1999). Silver-based crystalline nanoparticles, microbially fabricated. Proceedings of the National Academy of Science, 96, 13611–13614.
19. Lengke, M., Fleet, M., & Southam, G. (2007). Biosynthesis of silver nanoparticles by filamentous cyanobacteria from a silver (I) nitrate complex. Langmuir, 23(5), 2694–2699.
20. Xie, J., Lee, J. Y., Wang, D. I., & Ting, Y. P. (2007). Identification of active biomolecules in the high-yield synthesis of single-crystalline gold nanoplates in algal solutions. Small, 3(4), 672–682.
21. Agarwal, H., Kumar, S. V., & Rajeshkumar, S. (2017). A review on green synthesis of zinc oxide nanoparticles–An eco-friendly approach. *Resource-Efficient Technologies*, *3*(4), 406-413.https://doi.org/ 10.1016/j.reffit.2017.03.002
22. Thema, F. T., Manikandan, E., Dhlamini, M. S., & Maaza, M. J. M. L. (2015). Green synthesis of ZnO nanoparticles via Agathosma betulina natural extract. Materials Letters, 161, 124-127.doi: 10.1016/j.matlet.2015.08.052.
23. Bird, S. M., El-Zubir, O., Rawlings, A. E., Leggett, G. J., & Staniland, S. S. (2016). A novel design strategy for nanoparticles on nanopatterns: interferometric lithographic patterning of Mms6 bio templated magnetic nanoparticles. Journal of Materials Chemistry C, 4(18), 3948-3955., doi: 10.1039/C5TC03895B.
24. Rajiv, P., Rajeshwari, S., & Venkatesh, R. (2013). Bio-Fabrication of zinc oxide nanoparticles using leaf extract of Parthenium hysterophorus L. and its size-dependent antifungal activity against plant fungal pathogens. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 112, 384-387. doi 10.1016/j.saa.2013.04.072.

25. Behera, A., Mittu, B., Padhi, S., Patra, N., & Singh, J. (2020). Bimetallic nanoparticles: Green synthesis, applications, and future perspectives. In Multifunctional hybrid nanomaterials for sustainable agri-food and ecosystems (pp. 639-682). Elsevier. ISBN 978-0-12-821354-4.
26. Chaturvedi, V. K., Yadav, N., Rai, N. K., Bohara, R. A., Rai, S. N., Aleya, L., & Singh, M. P. (2021). Two birds with one stone: oyster mushroom mediated bimetallic Au-Pt nanoparticles for agro-waste management and anticancer activity. Environmental Science and Pollution Research, 28, 13761-13775.
27. Kumari, M. M., Jacob, J., & Philip, D. (2015). Green synthesis and applications of Au–Ag bimetallic nanoparticles. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 137, 185-192.
28. Ansari, Z., Saha, A., Singha, S. S., & Sen, K. (2018). Phytomediated generation of Ag, CuO, and Ag-Cu nanoparticles for dimethoate sensing. Journal of Photochemistry and Photobiology A: Chemistry, 367, 200-211.
29. Elemike, E. E., Onwudiwe, D. C., Fayemi, O. E., & Botha, T. L. (2019). Green synthesis and electrochemistry of Ag, Au, and Ag–Au bimetallic nanoparticles using golden rod (Solidago canadensis) leaf extract. Applied Physics A, 125, 1-12.
30. Wu, M. L., Chen, D. H., & Huang, T. C. (2001). Synthesis of Au/Pd bimetallic nanoparticles in reverse micelles. Langmuir, 17(13), 3877-3883.
31. Sivamaruthi, B. S., Ramkumar, V. S., Archunan, G., Chaiyasut, C., & Suganthy, N. (2019). Biogenic synthesis of silver palladium bimetallic nanoparticles from fruit extract of Terminalia chebula–In vitro evaluation of anticancer and antimicrobial activity. Journal of Drug Delivery Science and Technology, 51, 139-151.
32. Singh, I., Mazhar, T., Shrivastava, V., & Tomar, R. S. (2022). Bio-assisted synthesis of bi-metallic (Ag-Zn) nanoparticles by leaf extract of Azadirachta indica and its antimicrobial properties. International Journal of Nano Dimension, 13(2), 168-178.
33. Valodkar, M., Modi, S., Pal, A., & Thakore, S. (2011). Synthesis and antibacterial activity of Cu, Ag, and Cu–Ag alloy nanoparticles: A green approach. Materials Research Bulletin, 46(3), 384-389.
34. Oruç, Z., Ergüt, M., Uzunoğlu, D., & Özer, A. (2019). Green synthesis of biomass-derived activated carbon/Fe-Zn bimetallic nanoparticles from lemon (Citrus limon (L.) Burm. f.) wastes for heterogeneous Fenton-like decolorization of Reactive Red 2. Journal of Environmental Chemical Engineering, 7(4), 103231.
35. Younas, U., Hassan, S. T., Ali, F., Hassan, F., Saeed, Z., Pervaiz, M., ... & Ali, S. (2021). Radical scavenging and catalytic activity of Fe-Cu bimetallic nanoparticles synthesized from Ixora finlaysoniana extract. Coatings, 11(7), 813.
36. Liao, M., Li, W., Peng, J., Zhang, F., Xu, W., & Huang, Z. (2021). Enhancement of anodic oxidation of formic acid on Pd–Fe bimetallic nanoparticles by thermal treatment. International Journal of Hydrogen Energy, 46(17), 10239-10246.
37. Gulbagça, F., Aygun, A., Altuner, E. E., Bekmezci, M., Gur, T., Sen, F., ... & Vasseghian, Y. (2022). Facile bio-fabrication of Pd-Ag bimetallic nanoparticles and its performance in catalytic and pharmaceutical applications: Hydrogen production and in-vitro antibacterial, anticancer activities, and model development. Chemical Engineering Research and Design, 180, 254-264.
38. Sultana, S., Mech, S. D., Hussain, F. L., Pahari, P., Borah, G., & Gogoi, P. K. (2020). Green synthesis of graphene oxide (GO)-anchored Pd/Cu bimetallic nanoparticles using Ocimum sanctum as bio-reductant: an efficient heterogeneous catalyst for the Sonogashira cross-coupling reaction. RSC advances, 10(39), 23108-23120.
39. Amendola, V., Scaramuzza, S., Litti, L., Meneghetti, M., Zuccolotto, G., Rosato, A., ... & Colombatti, M. (2014). Magneto-plasmonic Au-Fe alloy nanoparticles designed for multimodal SERS-MRI-CT imaging. Small, 10(12), 2476-2486.
40. Lu, Y., Zhang, P., Lin, L., Gao, X., Zhou, Y., Feng, J., & Zhang, H. (2022). Ultra-small bimetallic phosphides for dual-modal MRI imaging-guided photothermal ablation of tumors. Dalton Transactions, 51(11), 4423-4428.

41. Koo, S., Park, O. K., Kim, J., Han, S. I., Yoo, T. Y., Lee, N., ... & Hyeon, T. (2022). Enhanced chemodynamic therapy by Cu–Fe peroxide nanoparticles: tumor microenvironment-mediated synergistic Fenton reaction. ACS nano, 16(2), 2535-2545.
42. Aygun, A., Gulbagca, F., Altuner, E. E., Bekmezci, M., Gur, T., Karimi-Maleh, H., ... & Sen, F. (2023). Highly active PdPt bimetallic nanoparticles synthesized by one-step bioreduction method: Characterizations, anticancer, antibacterial activities, and evaluation of their catalytic effect for hydrogen generation. International Journal of Hydrogen Energy, 48(17), 6666-6679.
43. Garfinkel, D. A., Tang, N., Pakeltis, G., Emery, R., Ivanov, I. N., Gilbert, D. A., & Rack, P. D. (2022). Magnetic and Optical Properties of Au–Co Solid Solution and Phase-Separated Thin Films and Nanoparticles. ACS Applied Materials & Interfaces, 14(13), 15047-15058.
44. Yang, H., Zhang, X., Velu, P., Liu, X., & Vijayalakshmi, A. (2022). Enhanced green mediated synthesis of optimized Ag-Cu bimetallic nanoparticles using Leucas aspera and its application in Anti-cancer activity against alveolar cancer. Materials Letters, 313, 131645.
45. Koyyati, R., Kudle, K. R., Nagati, V., Merugu, R., & Padigya, P. R. M. (2022, April). Extracellular Synthesis of Mono and Bimetallic Nanocomposites from Novel Strains of Rhodopseudomonas palustris and Evaluation of Their Biomedical Applications. In Macromolecular Symposia (Vol. 402, No. 1, p. 2100378).
46. Oladipo, A. O., Iku, S. I., Ntwasa, M., Nkambule, T. T., Mamba, B. B., & Msagati, T. A. (2020). Doxorubicin conjugated hydrophilic AuPt bimetallic nanoparticles fabricated from Phragmites australis: Characterization and cytotoxic activity against human cancer cells. Journal of Drug Delivery Science and Technology, 57, 101749.
47. He, F., Ji, H., Feng, L., Wang, Z., Sun, Q., Zhong, C., ... & Lin, J. (2021). Construction of thiol-capped ultrasmall Au–Bi bimetallic nanoparticles for X-ray CT imaging and enhanced antitumor therapy efficiency. Biomaterials, 264, 120453.
48. Aygun, A., Gulbagca, F., Altuner, E. E., Bekmezci, M., Gur, T., Karimi-Maleh, H., ... & Sen, F. (2023). Highly active PdPt bimetallic nanoparticles synthesized by one-step bioreduction method: Characterizations, anticancer, antibacterial activities and evaluation of their catalytic effect for hydrogen generation. International Journal of Hydrogen Energy, 48(17), 6666-6679.
49. Padilla-Cruz, A. L., Garza-Cervantes, J. A., Vasto-Anzaldo, X. G., García-Rivas, G., León-Buitimea, A., & Morones-Ramírez, J. R. (2021). Synthesis and design of Ag–Fe bimetallic nanoparticles as antimicrobial synergistic combination therapies against clinically relevant pathogens. Scientific Reports, 11(1), 5351.
50. Krishnan, S. G., Nand, D., Sivadas, S., Alfarhan, A., & Muthusamy, K. (2022). CuO/NiO bimetallic nanocomposite: A facile DNA assisted synthetic approach and evaluation of bio efficacy. Journal of King Saud University-Science, 34(1), 101718.
51. Velmurugan, P., Mohanavel, V., Sekar, P., Vijayanand, S., Chinnathambi, A., Govindasamy, C., & Sivakumar, S. (2022). Influence of dissolved oxygen on the synthesis of Ag-Au mono and bimetallic nanostructure using Cudrania tricuspidata leaf extract and its broad-spectrum antibacterial activity. Materials Letters, 310, 131471.
52. Ramakritinan, C. M., Kaarunya, E., Shankar, S., & Kumaraguru, A. K. (2013). Antibacterial effects of Ag, Au, and bimetallic (Ag-Au) nanoparticles synthesized from red algae. Solid State Phenomena, 201, 211-230.
53. Lateef, A., Ojo, S. A., Folarin, B. I., Gueguim-Kana, E. B., & Beukes, L. S. (2016). Kolanut (Cola nitida) mediated synthesis of silver–gold alloy nanoparticles: antifungal, catalytic, larvicidal, and thrombolytic applications. Journal of Cluster Science, 27, 1561-1577.
54. Ghosh, S., Rana, D., Sarkar, P., Roy, S., Kumar, A., Naskar, J., & Kole, R. K. (2022). Ecological safety with multifunctional applications of biogenic mono and bimetallic (Au–Ag) alloy nanoparticles. Chemosphere, 288, 132585.
55. Mohanlall, V., & Biyela, B. (2022). Biocatalytic and biological activities of Kigelia africana mediated silver monometallic and copper-silver bimetallic nanoparticles. Indian Journal of Biochemistry and Biophysics (IJBB), 59(1), 94-102.

56. Zhou, F., Zhu, Y., Yang, L., Yang, D. Q., & Sacher, E. (2022). Ag NP catalysis of Cu ions in the preparation of AgCu NPs and the mechanism of their enhanced antibacterial efficacy. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 632, 127831.
57. Formaggio, D. M. D., de Oliveira Neto, X. A., Rodrigues, L. D. A., de Andrade, V. M., Nunes, B. C., Lopes-Ferreira, M., ... & Tada, D. B. (2019). In vivo toxicity and antimicrobial activity of AuPt bimetallic nanoparticles. Journal of Nanoparticle Research, 21, 1-16.
58. Singh, A. K. (2022). Flower extract-mediated green synthesis of bimetallic CuZn oxide nanoparticles and its antimicrobial efficacy in hydrocolloid films. Bioresource Technology Reports, 18, 101034.
59. Zhang, S., Dong, H., He, R., Wang, N., Zhao, Q., Yang, L., ... & Li, J. (2022). Hydro electroactive Cu/Zn coated cotton fiber nonwovens for antibacterial and antiviral applications. International Journal of Biological Macromolecules, 207, 100-109.
60. Yang, H., Zhang, X., Velu, P., Liu, X., & Vijayalakshmi, A. (2022). Enhanced green mediated synthesis of optimized Ag-Cu bimetallic nanoparticles using Leucas aspera and its application in Anti-cancer activity against alveolar cancer. Materials Letters, 313, 131645.
61. Ahmad, M. M., Kotb, H. M., Mushtaq, S., Waheed-Ur-Rehman, M., Maghanga, C. M., & Alam, M. W. (2022). Green synthesis of Mn+ Cu bimetallic nanoparticles using Vinca rosea extract and their antioxidant, antibacterial, and catalytic activities. Crystals, 12(1), 72.
62. Ghosh, S., Nitnavare, R., Dewle, A., Tomar, G. B., Chippalkatti, R., More, P., ... & Chopade, B. A. (2015). Novel platinum–palladium bimetallic nanoparticles synthesized by Dioscorea bulbifera: anticancer and antioxidant activities. International Journal of Nanomedicine, 7477-7490.
63. Bhanja, S. K., Samanta, S. K., Mondal, B., Jana, S., Ray, J., Pandey, A., & Tripathy, T. (2020). Green synthesis of Ag@ Au bimetallic composite nanoparticles using a polysaccharide extracted from Ramaria botrytis mushroom and performance in catalytic reduction of 4-nitrophenol and antioxidant, antibacterial activity. Environmental Nanotechnology, Monitoring & Management, 14, 100341.
64. Malapermal, V., Mbatha, J. N., Gengan, R. M., & Anand, K. (2015). Biosynthesis of bimetallic Au-Ag nanoparticles using Ocimum basilicum (L.) with antidiabetic and antimicrobial properties. Advanced materials letters (Online) 6, 1050–1057.
65. Sher, N., Ahmed, M., Mushtaq, N., & Khan, R. A. (2022). Enhancing the antioxidant, antidiabetic, and anti-Alzheimer performance of Hippeastrum hybridum (L.) using silver nanoparticles. Applied Organometallic Chemistry, 36(7), e6724.
66. D'Souza, J. N., Nagaraja, G. K., Prabhu, A., Navada, K. M., Kouser, S., & Manasa, D. J. (2022). AgVI and Ag/ZnOVI nanostructures from Vateria indica (L.) exert antioxidant, antidiabetic, anti-inflammatory, and cytotoxic efficacy on triple-negative breast cancer cells in vitro. International Journal of Pharmaceutics, 615, 121450.
67. Bakur, A., Elshaarani, T., Niu, Y., & Chen, Q. (2019). Comparative study of antidiabetic, bactericidal, and antitumor activities of MEL@ AgNPs, MEL@ ZnONPs, and Ag–ZnO/MEL/GA nanocomposites prepared by using MEL and gum arabic. RSC advances, 9(17), 9745-9754.
68. Imraish, A., Abu Thiab, T., Al-Awaida, W., Al-Ameer, H. J., Bustanji, Y., Hammad, H., ... & Al-Hunaiti, A. (2021). In vitro anti-inflammatory and antioxidant activities of ZnFe2O4 and CrFe2O4 nanoparticles synthesized using Boswellia carteri resin. Journal of Food Biochemistry, 45(6), e13730.
69. Antonoglou, O., Lafazanis, K., Mourdikoudis, S., Vourlias, G., Lialiaris, T., Pantazaki, A., & Dendrinou-Samara, C. (2019). Biological relevance of CuFeO2 nanoparticles: Antibacterial and anti-inflammatory activity, genotoxicity, DNA and protein interactions. Materials Science and Engineering: C, 99, 264-274.
70. Orlowski, P., Zmigrodzka, M., Tomaszewska, E., Ranoszek-Soliwoda, K., Pajak, B., Slonska, A., ... & Krzyzowska, M. (2020). Polyphenol-conjugated bimetallic Au@ AgNPs for improved wound healing. International journal of nanomedicine, 4969-4990.
71. Oladipo, A. O., Nkambule, T. T., Mamba, B. B., & Msagati, T. A. (2020). The stimuli-responsive properties of doxorubicin adsorbed onto bimetallic Au@ Pd nano dendrites and its potential application as drug delivery platform. Materials Science and Engineering: C, 110, 110696.

72. Stavropoulou, A. P., Theodosiou, M., Sakellis, E., Boukos, N., Papanastasiou, G., Wang, C., ... & Efthimiadou, E. K. (2022). Bimetallic gold-platinum nanoparticles as a drug delivery system coated with a new drug to target glioblastoma. Colloids and Surfaces B: Biointerfaces, 214, 112463.
73. Chu, C., Bao, Z., Sun, M., Wang, X., Zhang, H., Chen, W., ... & Wang, D. (2020). NIR Stimulus-responsive PdPt bimetallic nanoparticles for drug delivery and chemo-photothermal therapy. Pharmaceutics, 12 (7), 675.
74. Sasireka, K. & Lalitha, P. (2021). Biogenic synthesis of bimetallic nanoparticles and their applications. Reviews in Inorganic Chemistry, 41 (4), 223-244. https://doi.org/10.1515/revic-2020-0024
75. Chopra H, Bibi S, Singh I, Hasan MM, Khan MS, Yousafi Q, Baig AA, Rahman MM, Islam F, Emran TB and Cavalu S (2022). Green Metallic Nanoparticles: Biosynthesis to Applications. Front. Bioeng. Biotechnol. 10:874742. doi 10.3389/fbioe.2022.874742

This page is intentionally left blank

The Quantile Method for Symbolic Principal Component Analysis

Manabu Ichino

ABSTRACT

In this article, we present a new quantification method to realize the principal component analysis (PCA) for symbolic data tables. We first describe the nesting property for the monotone point sequences and the correlation matrix by the rank correlation coef icient. Then, we present the object splitting method by which interval valued data table can be transformed to a usual numerical data table. We are able to apply the traditional PCA to this transformed data table. The quantile method is a generalization of the object splitting method, and can manipulate histograms, nominal multi-value types, and other types simultaneously. We present several experimental results in order to illustrate the usefulness of the quantile method. © 2011 Wiley Periodicals, Inc. Statistical Analysis and Data Mining 4: 184–198, 2011

Keywords: PCA; monotone structure; rank correlation; histogram; quantile; sub-object.

Classification: LCC Code: QA76.9.D343, G.3, 62H25

Language: English

Great Britain Journals Press

LJP Copyright ID: 925613
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural and Formal

Volume 23 | Issue 12 | Compilation 1.0

The Quantile Method for Symbolic Principal Component Analysis

Manabu Ichino*

ABSTRACT

In this article, we present a new quantification method to realize the principal component analysis (PCA) for symbolic data tables. We first describe the nesting property for the monotone point sequences and the correlation matrix by the rank correlation coefficient. Then, we present the object splitting method by which interval valued data table can be transformed to a usual numerical data table. We are able to apply the traditional PCA to this transformed data table. The quantile method is a generalization of the object splitting method, and can manipulate histograms, nominal multi-value types, and other types simultaneously. We present several experimental results in order to illustrate the usefulness of the quantile method. © 2011 Wiley Periodicals, Inc. Statistical Analysis and Data Mining 4: 184–198, 2011.

Keywords: PCA; monotone structure; rank correlation; histogram; quantile; sub-object.

I. INTRODUCTION

The generalization of the principal component analysis (PCA) is an important research theme in the symbolic data analysis [1–4]. The main purpose of the traditional PCA is to transform a number of possibly correlated variables into a small number of uncorrelated variables called principal components. Chouakria [5] proposed the extension of the PCA to interval data as *vertices principal component analysis* (V-PCA). Chouakria *et al.* [6] proposed also the *centers method* of PCA (C-PCA) for interval data, and they presented a comparative example for the V-PCA and the C-PCA. Lauro and Palumbo [7] proposed *symbolic object principal component analysis* (SO-PCA) as an extended PCA to any numerical data structure. Lauro *et al.* [8] summarize various methods of SO-PCA for interval data. The author also proposed a general "Symbolic PCA" (S-PCA) based on the quantification method by using the generalized Minkowski metrics [9,10]. In this approach, we first transform the given symbolic data table to a usual numerical data table, and then we execute the traditional PCA on the transformed data table.

In this article, another quantification method for symbolic data tables based on the monotone structures of objects is presented. In Section 2, first we describe the case of point sequences in a d-dimensional Euclidean space. The monotone structures are characterized by the nesting of the Cartesian join regions associated with pairs of objects. If the given point sequence is monotone in the Euclidean d space, the property is also satisfied in any feature axis. In other words, a nesting structure of the given point sequence in the d space confines the orders of points in each feature axis to be similar. Therefore, we can evaluate the degree of similarity between features based on the Kendall or the Spearman's rank correlation coefficients. Then, we can execute a traditional PCA based on the correlation matrix by the selected rank correlation coefficient. Secondly, we describe the "object splitting method" for SO-PCA for interval-valued data [11]. This method splits each of N symbolic objects described by d interval-valued features into the two d-dimensional vertices called the "minimum sub-object" and the "maximum sub object". We should point out the fact that any interval object can be reproduced from the minimum and the maximum sub-objects. Moreover, the nesting structure of interval objects in the d space confines the orders of the minimum and the

London Journal of Research in Science: Natural and Formal

maximum sub-objects in each feature axis to be similar. Therefore, we can evaluate again the degree of similarity between features based on the Kendall or the Spearman's rank correlation coefficients on the $(2 \times N) \times d$ standard numerical data table. We can execute a traditional PCA based on the correlation matrix by the selected rank correlation coefficient. As a further extension to manipulate histogram data, nominal multi-valued data, and others, we describe the "quantile method" for S-PCA [12] in Section 4.

The problem is how to obtain a common numerical rep resentation of objects described by mixed types of features. For example, in histogram data, the numbers of subinter vals (bins) of the given histograms are mutually different in general. Therefore, we first define the cumulative dis tribution function for each histogram. Then, we select a common integer number m to generate the "quantiles" for all histograms. As the result, for each histogram, we have an $(m + 1)$-tuple composed of $(m - 1)$ quantiles and the *minimum* and the *maximum* values of the whole interval of the histogram. Then, we split each object into $(m + 1)$ sub-objects: the *minimum* sub-object, $(m - 1)$ *quantile* sub objects and the *maximum* sub-object. By virtue of the monotonic property of the distribution function, $(m + 1)$ sub-objects of an object satisfy automatically a nesting structure. Therefore, the nesting of N objects described by the minimum and the maximum sub-objects in the d space confines the orders of $N \times (m + 1)$ sub-objects in each feature axis to be similar. Again, we can evaluate the degree of similarity between features by the Kendall or the Spearman's rank correlation coefficient, and then execute a traditional PCA.

Interval-valued data may be regarded as a special histogram-valued data, where only *one* bin organizes the histogram. Furthermore, we can also split nominal multi-valued data into $(m + 1)$ sub-objects based on the distribution function associated with rank values attached to categorical values of an object. Therefore, by the quantile method we can transform a given general $N \times d$ symbolic data table to an $\{N \times (m + 1)\} \times d$ standard numerical data table, and then we can execute a traditional PCA on the transformed data table. In Section 5, we describe several experimental results in order to show the effectiveness of the quantile method. Section 6 is a summary.

II. MONOTONE STRUCTURES AND OBJECT SPLITTING METHOD

In this section, we describe some properties of monotone structures for point sequence and for interval objects. Then, we describe the object splitting method for S-PCA.

2.1. Monotone Structures for Point Sequence

Let a set of N objects $\boldsymbol{U}$ be represented by $\boldsymbol{U} = \{\omega_1, \omega_2, \ldots, \omega_N\}$. Let each object ω_i be described by d numerical features, i.e. a vector $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{id})$ in a d-dimensional Euclidean space $\boldsymbol{R}^d$.

DEFINITION 1: Rectangular region spanned by $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$.

Let $\boldsymbol{J}(\omega_i, \omega_j)$ be a rectangular region in $\boldsymbol{R}^d$ spanned by the vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, and be defined by the following Cartesian product of d closed intervals.

$$\begin{aligned} \boldsymbol{J}(\omega_i, \omega_j) &= [\min(x_{i1}, x_{j1}), \max(x_{i1}, x_{j1})] \\ &\times [\min(x_{i2}, x_{j2}), \max(x_{i2}, x_{j2})] \\ &\times \cdots \times [\min(x_{id}, x_{jd}), \max(x_{id}, x_{jd})], \end{aligned} \tag{1}$$

where $\min(a, b)$ and $\max(a, b)$ are the operators to take the minimum value and the maximum value from a and b, respectively.

In the following, we call $\boldsymbol{J}(\omega_i, \omega_j)$ as the Cartesian join (region) of objects ω_i and ω_j [9,10,13].

DEFINITION 2: Nesting structure

If a series of objects $\omega_1, \omega_2, \ldots, \omega_N$ satisifies the nesting property

$$\boldsymbol{J}(\omega_1, \omega_k) \subseteq \boldsymbol{J}(\omega_1, \omega_{k+1}), k = 1, 2, \ldots, N-1, \tag{2}$$

the series is called a "nesting structure with the starting point ω_1 and the ending point ω_N".

In Fig. 1, (a) is a monotone increasing series, and (b) is a monotone decreasing series of objects. It should be noted that the two series of objects show the same nesting structures with starting point ω_1 and ending point ω_5.

PROPOSITION 1: If a series of objects $\omega_1, \omega_2, \ldots, \omega_N$ is *a nesting structure* with the starting point ω_1 and the ending point ω_N in the space $\boldsymbol{R}^d$, the series satisfies the same structure in each feature (axis) of the space $\boldsymbol{R}^d$.

Proof: From the definition of rectangular region as in Eq. (1), we have

$$\begin{aligned}\boldsymbol{J}(\omega_1, \omega_k) = &[\min(x_{11}, x_{k1}), \max(x_{11}, x_{k1})] \\ &\times [\min(x_{12}, x_{k2}), \max(x_{12}, x_{k2})] \\ &\times \cdots \times [\min(x_{1d}, x_{kd}), \\ &\max(x_{1d}, x_{kd})],\end{aligned} \tag{3}$$

and

$$\begin{aligned}\boldsymbol{J}(\omega_1, \omega_{k+1}) = &[\min(x_{11}, x_{k+1,1}), \max(x_{11}, x_{k+1,1})] \times [\min(x_{12}, x_{k+1,2}), \max(x_{12}, x_{k+1,2})] \\ &\times \cdots \times [\min(x_{1d}, x_{k+1,d}), \max(x_{1d}, x_{k+1,d})].\end{aligned} \tag{4}$$

Statistical Analysis and Data Mining DOI:10.1002/sam

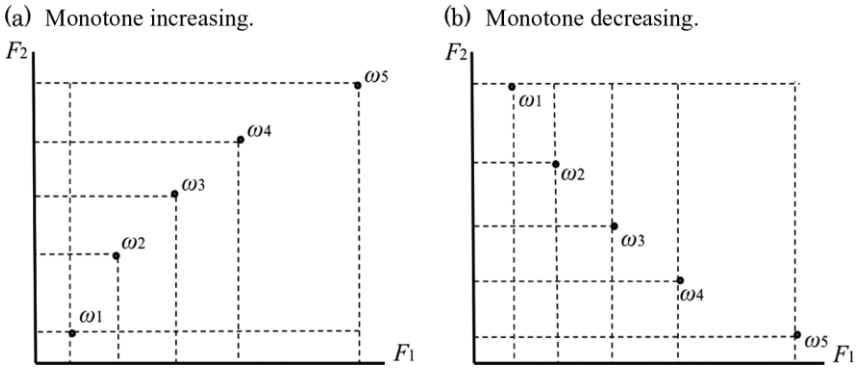

Fig. 1: Monotone structures of point sequence.

Therefore, the relations of the Cartesian join regions $\boldsymbol{J}(\omega_1, \omega_k) \subseteq \boldsymbol{J}(\omega_1, \omega_{k+1})$, $k = 1, 2, \ldots, N-1$, in Definition 2, require the following relations for each feature, i.e. for each j $(= 1, 2, \ldots, d)$,

$$[\min(x_{1j}, x_{kj}), \max(x_{1j}, x_{kj})] \subseteq [\min(x_{1j}, x_{k+1,j}), \max(x_{1j}, x_{k+1,j}], \; k = 1, 2, \ldots, N-1. \tag{5}$$

Although, there exist several ways to define the mono tone sequences of objects, i.e. monotone structures, we use the following definition.

DEFINITION 3: Monotone structure of a series of points.

A series of objects $\omega_1, \omega_2, \ldots, \omega_N$ is called a monotone structure, if the series satisfies the nesting structure of Definition 2.

Since, for a pair of features, we can evaluate the degree of similarity between two sets of orders of objects for the same object set $\boldsymbol{U}$ by using the Kendall or the Spearman's rank correlation coefficient, we have Proposition 2.

PROPOSITION 2: Correlation matrix $\boldsymbol{S}$.

If a series of objects $\omega_1, \omega_2, \ldots, \omega_N$ is a monotone structure in the space $\boldsymbol{R}^d$, the absolute value of each off diagonal element of the $d \times d$ correlation matrix $\boldsymbol{S}$ takes the maximum value *one* in the sense of the Kendall or the Spearman's rank correlation coefficient.

Proof: From Definition 3, any monotone structure must satisfy the nesting property of Definition 2. Then, from Proposition 1, the given series of objects has the identical nesting structure for each feature. This property exactly restricts the order of objects for each feature to be the same way or the reverse way according to the series of objects is monotone increasing or monotone decreasing. Therefore, if a series of objects is a monotone structure in $\boldsymbol{R}^d$, the absolute value of the correlation coefficient for each pair of features takes the maximum value *one* in the sense of the Kendall or the Spearman's rank correlation coefficient.

From Proposition 2, if many off-diagonal elements of $\boldsymbol{S}$ take highly correlated values, we can expect the existence of a large eigenvalue of $\boldsymbol{S}$, and that the corresponding eigenvector reproduces well the original nesting property of the set of objects in the space $\boldsymbol{R}^d$.

EXAMPLE 1: As an intuitive example, suppose that the given set of objects in $\boldsymbol{R}^d$ organizes an approximate monotone structure which is monotone increasing along each of d features, and the degrees of similarity between two features are the same for all possible pairs. Therefore, all off-diagonal elements of $\boldsymbol{S}$ take an identical value ρ, $0 < \rho < 1$. Then, it is known [14] that d eigenvalues of $\boldsymbol{S}$ become

$$\lambda_1 = 1 + (d-1)\rho \text{ and } \lambda_2 = \lambda_3 = \cdots = \lambda_d = 0, \tag{6}$$

and the eigenvector for λ_1 is

$$\boldsymbol{a}_1 = (1/\sqrt{d}, 1/\sqrt{d}, \ldots, 1/\sqrt{d}). \tag{7}$$

Therefore, the given monotone structure of objects in $\boldsymbol{R}^d$ is approximately reproduced around the eigenvector $\boldsymbol{a}_1$. As a particular case, when $\rho = 1$, the given set of objects organizes a complete monotone structure in the space $\boldsymbol{R}^d$. Then, the eigenvalue λ_1 becomes d, i.e. its contribution ratio is 100%, and the order of the given object sequence in the space $\boldsymbol{R}^d$ is exactly reproduced on the eigenvector $\boldsymbol{a}_1$.

In the above, we characterized monotone structures by the nesting property, and obtain the correlation matrix $\boldsymbol{S}$. The monotone structures include any linear structure as a special case. On the other hand, a monotone structure may be approximated well by an appropriately selected linear structure. This suggests that we can use also the Pearson correlation coefficient to evaluate the degree of similarity between two features instead of the Kendall and the Spearman's rank correlation coefficients.

2.2. Monotone Structures for Interval Objects

Let each object be described by d interval-valued features. Then, an object $\omega_k \in \boldsymbol{U}$ becomes a hyper rectangle in $\boldsymbol{R}^d$, i.e. the Cartesian product of d closed intervals:

$$\boldsymbol{I}_k = I_{k1} \times I_{k2} \times \cdots \times I_{kd}, \tag{8}$$

where each interval I_{kp} is given by

$$I_{kp} = [x_{kp(\min)}, x_{kp(\max)}], \; p = 1, 2, \ldots, d. \tag{9}$$

Then, we can define the minimum vertex $\boldsymbol{x}_{k(\min)}$ and the maximum vertex $\boldsymbol{x}_{k(\max)}$ by

$$\boldsymbol{x}_{k(\min)} = (x_{k1(\min)}, x_{k2(\min)}, \ldots, x_{kd(\min)}) \text{ and } \boldsymbol{x}_{k(\max)} = (x_{k1(\max)}, x_{k2(\max)}, \ldots, x_{kd(\max)}). \tag{10}$$

DEFINITION 4: The minimum sub-object and the max imum sub-object

Let the minimum vertex $\boldsymbol{x}_{k(\min)}$ and the maximum vertex $\boldsymbol{x}_{k(\max)}$ for each object $\omega_k \in \boldsymbol{U}$ be called the minimum sub object and the maximum sub-object, and be denoted by $\omega_{k(\min)}$ and $\omega_{k(\max)}$, respectively.

EXAMPLE 2: In Table 1, the minimum and the maxi mum sub-objects of *Linseed oil* under the first four inter val features are represented by the vertices $\boldsymbol{x}_{\text{Linseed}(\min)}$ = (0.930, −27, 170, 118) and $\boldsymbol{x}_{\text{Linseed}(\max)}$ = (0.935, −18, 204, 196), respectively.

PROPOSITION 3: From Definition 1, any interval object $\omega_k \in \boldsymbol{U}$ is represented in the space $\boldsymbol{R}^d$ by the Cartesian join region $\boldsymbol{J}(\omega_{k(\min)}, \omega_{k(\max)})$.

Proof: From Eq. (1) in Definition 1 and (8–10), we see that

$$\begin{aligned} \boldsymbol{J}(\omega_{k(\min)}, \omega_{k(\max)}) &= [x_{k1(\min)}, x_{k1(\max)}] \\ &\times [x_{k2(\min)}, x_{k2(\max)}] \\ &\times \ldots \times [x_{kd(\min)}, x_{kd(\max)}] \\ &= I_{k1} \times I_{k2} \times \cdots \times I_{kd} = \boldsymbol{I}_k. \end{aligned}$$

From Eq. (8), d respective intervals for ω_i and ω_j are

$$I_{ip} = [x_{ip(\min)}, x_{ip(\max)}], \; p = 1, 2, \ldots, d, \text{ and } I_{jp} = [x_{jp(\min)}, x_{jp(\max)}], \; p = 1, 2, \ldots, d. \tag{11}$$

Thus the closed interval I_{ijp} generated from two intervals I_{ip} and I_{jp} becomes

$$\begin{aligned} I_{ijp} = [&\min(x_{ip(\min)}, x_{jp(\min)}), \\ &\max(x_{ip(\max)}, x_{jp(\max)})], \; p = 1, 2, \ldots, d. \end{aligned} \tag{12}$$

London Journal of Research in Science: Natural and Formal

DEFINITION 5: We define the Cartesian join region $\boldsymbol{J}(\omega_i, \omega_j)$ based on Eq. (12) by

$$\begin{aligned}\boldsymbol{J}(\omega_i, \omega_j) &= I_{ij\,1} \times I_{ij\,2} \times \cdots \times I_{ijd} \\ &= [\min(x_{i1(\min)}, x_{j\,1(\min)}), \\ &\quad \max(x_{i1(\max)}, x_{j\,1(\max)})] \\ &\quad \times [\min(x_{i2(\min)}, x_{j\,2(\min)}), \\ &\quad \max(x_{i2(\max)}, x_{j\,2(\max)})] \\ &\quad \times \ldots \times [\min(x_{id(\min)}, x_{jd(\min)}), \\ &\quad \max(x_{id(\max)}, x_{jd(\max)})].\end{aligned} \tag{13}$$

In this definition, we should note that, for each k, $\boldsymbol{J}(\omega_k, \omega_k)$ is equivalent to $\boldsymbol{J}(\omega_{k(\min)}, \omega_{k(\max)})$. Furthermore,

Table 1: Fats' and oils' data [10].

Object	Specific gravity (g/cm^3), F_1	Freezing point (°C), F_2	Iodine value, F_3	Saponification value, F_4	Major acids, F_5
1. Linseed	0.930–0.935	−27 to −18	170–204	118–196	L, Ln, O, P, M
2. Perilla	0.930–0.937	−5 to −4	192–208	188–197	L, Ln, O, P, S
3. Cotton	0.916–0.918	−6 to −1	99–113	189–198	L, O, P, M, S
4. Sesame	0.920–0.926	−6 to −4	104–116	187–193	L, O, P, S, A
5. Camellia	0.916–0.917	−21 to −15	80–82	189–193	L, O
6. Olive	0.914–0.919	0–6	79–90	187–196	L, O, P, S
7. Beef	0.860–0.870	30–38	40–48	190–199	O, P, M, S, C
8. Hog	0.858–0.864	22–32	53–77	190–202	L, O, P, M, S, Lu

L = linoleic acid, Ln = linolenic acid, O = oleic acid, P = palmitic acid, M = myristic acid, S = searic acid, A = arachic acid, C = capric acid, Lu = lauric acid.

$\boldsymbol{J}(\omega_{k(\min)}, \omega_{k(\min)})$ and $\boldsymbol{J}(\omega_{k(\max)}, \omega_{k(\max)})$ are reduced to the minimum vertex $\boldsymbol{x}_{k(\min)}$ and the maximum vertex $\boldsymbol{x}_{k(\max)}$ in Eq. (10), respectively.

DEFINITION 6: Nesting structure for interval objects If a series of interval objects $\omega_1, \omega_2, \ldots, \omega_N$ satisifies the nesting property

$$\boldsymbol{J}(\omega_1, \omega_k) \subseteq \boldsymbol{J}(\omega_1, \omega_{k+1}), k = 1, 2, \ldots, N-1, \tag{14}$$

the series is called a "nesting structure with the starting object ω_1 and the ending object ω_N".

Fig. 2 shows a series of five interval objects. It should be noted that the nesting order of objects in each feature axis is the same as that in the two-dimensional space.

PROPOSITION 4: If a series of interval objects $\omega_1, \omega_2, \ldots, \omega_N$ is *a nesting structure* with the starting object ω_1 and the ending object ω_N in the space $\boldsymbol{R}^d$, the series has the same nesting in each feature (axis) of the space $\boldsymbol{R}^d$.

Proof: From the definition of the Cartesian join region as seen in Eq. (13), we have

$$\begin{aligned}\mathbf{J}(\omega_1, \omega_k) = &[\min(x_{11(\min)}, x_{k1(\min)}),\\ &\max(x_{11(\max)}, x_{k1(\max)})]\\ &\times [\min(x_{12(\min)}, x_{k2(\min)}),\\ &\max(x_{12(\max)}, x_{k2(\max)})]\\ &\times \ldots \times [\min(x_{1d(\min)}, x_{kd(\min)}),\\ &\max(x_{1d(\max)}, x_{kd(\max)})],\end{aligned} \tag{15}$$

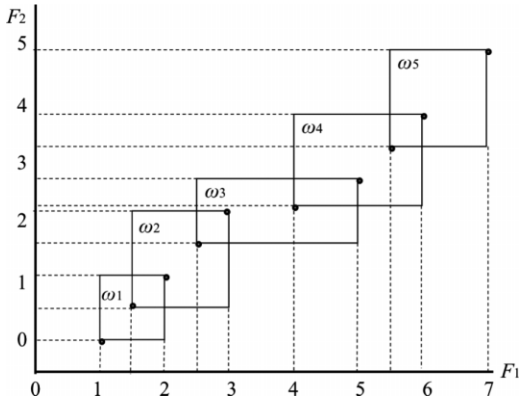

Fig. 2: A monotone structure of interval objects.

and

$$\begin{aligned}\boldsymbol{J}(\omega_1, \omega_{k+1}) = &[\min(x_{11(\min)}, x_{k+1,1(\min)}),\\ &\max(x_{11(\max)}, x_{k+1,1(\max)})]\\ &\times [\min(x_{12(\min)}, x_{k+1,2(\min)}),\\ &\max(x_{12(\max)}, x_{k+1,2(\max)})]\\ &\times\cdots\times [\min(x_{1d(\min)}, x_{k+1,d(\min)}),\\ &\max(x_{1d(\max)}, x_{k+1,d(\max)})].\end{aligned} \tag{16}$$

Therefore, the relations of the Cartesian join regions $\boldsymbol{J}(\omega_1, \omega_k) \subseteq \boldsymbol{J}(\omega_1, \omega_{k+1})$, $k = 1, 2, \ldots, N-1$, in Definition 5, require the following relations for each feature, i.e. for each j $(= 1, 2, \ldots, d)$,

$$[\min(x_{1j(\min)}, x_{kj(\min)}), \max(x_{1j(\max)}, x_{kj(\max)})] \subseteq [\min(x_{1j(\min)}, x_{k+1,j(\min)}), \max(x_{1j(\max)}, x_{k+1,j(\max)})],\ k = 1, 2, \ldots, N-1. \tag{17}$$

We define the monotone structure of interval objects by the same way in Definition 3.

DEFINITION 7: Monotone structure of a series of interval objects.
A series of interval objects $\omega_1, \omega_2, \ldots, \omega_N$ is called a monotone structure, if the series satisfies a nesting structure in Definition 6.

According to Definition 7, we assume a series of interval objects $\omega_1, \omega_2, \ldots, \omega_N$ is a monotone structure in the space $\boldsymbol{R}^d$. Then, from Proposition 4, the series of objects satisfies the same nesting in each feature axis. However, the nesting in (17) is based on the closed intervals generated from two objects. Therefore, we cannot evaluate the degree of similarity between two features by direct use of the Kendall or the Spearman's rank correlation coefficient. To remove this difficulty, we split each interval object into the minimum sub-object and the maximum sub-object.

PROPOSITION 5: Monotone conditions by sub-objects. Let a series of interval objects $\omega_1, \omega_2, \ldots, \omega_N$ be *monotone* in the space $\boldsymbol{R}^d$. Then, at least one condition of the following must be satisfied.

(1) The series of the minimum sub-objects, $\omega_{1(\min)}, \omega_{2(\min)}, \ldots, \omega_{N(\min)}$, is *monotone* in $\boldsymbol{R}^d$.

(2) The series of the maximum sub-objects, $\omega_{1(\max)}, \omega_{2(\max)}, \ldots, \omega_{N(\max)}$, is *monotone* in $\boldsymbol{R}^d$.

Proof: Assume that the conditions (1) and (2) are negated simultaneously. Then, there exists a nesting order k in which the object ω_k satisfies the nesting property in $\boldsymbol{R}^d$ but the corresponding minimum sub-object $\omega_{k(\min)}$ and the maximum sub-object $\omega_{k(\max)}$ breaks the nesting property in $\boldsymbol{R}^d$, simultaneously. This contradicts the fact given in Proposition 3.

DEFINITION 8: Strongly monotone structure and weak ly monotone structure.

If a series of interval objects $\omega_1, \omega_2, \ldots, \omega_N$ in $\boldsymbol{R}^d$ satisfies both conditions (1) and (2) in Proposition 5, we call the series of objects as *strongly monotone* in $\boldsymbol{R}^d$. On the other hand, if the series of objects satisfies only one condition, we call the series of objects as *weakly monotone* in $\boldsymbol{R}^d$.
Fig. 2 shows a case of a *strongly monotone* struc ture, whereas Fig. 3 illustrates a case of a *weakly mono tone* structure.

If a series of interval objects $\omega_1, \omega_2, \ldots, \omega_N$ in the space $\boldsymbol{R}^d$ is given, we can obtain the $d \times d$ correlation matrix $\boldsymbol{S}$ by splitting each object into the minimum and the maximum sub-objects and by using the Kendall or the Spearman's rank correlation coefficient.

PROPOSITION 6: Property of correlation matrix $\boldsymbol{S}$ by the object splitting.

(1) If the given series of objects is strongly monotone in a pair of features, the corresponding correlation coefficient shows a strictly high score for $2N$ sub objects by the object splitting.
(2) If the given series of interval objects is weakly monotone, the correlation coefficient shows a degraded score compared to the case (1).

Proof: From Proposition 5, if the given interval objects in $\boldsymbol{R}^d$ is monotone, the series of the minimum sub-objects in

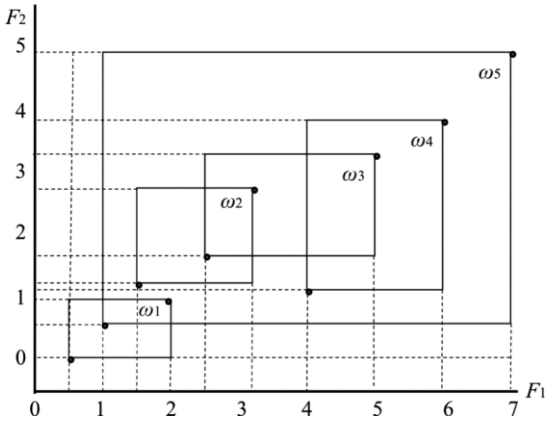

Fig. 3: An example of weakly monotone structure.

$\boldsymbol{R}^d$ and/or the series of the maximum sub-objects in $\boldsymbol{R}^d$ also become monotone. Therefore, we have the properties (1) and (2) whether the given series of objects is strongly monotone or weakly monotone.

In the above, we characterized monotone structures of N interval objects in the space $\boldsymbol{R}^d$ by the nesting property of $2N$ sub-objects in $\boldsymbol{R}^d$, i.e. the minimum sub object and the maximum sub-object, and obtained the correlation matrix $\boldsymbol{S}$ based on the Kendall or Spearman's rank correlation coefficient. As noted in the preceding, the monotone structures include any linear structure as a special case. On the other hand, a monotone structure may be approximated well by an appropriately selected linear structure. Therefore, we can use also the Pearson correlation coefficient to evaluate the degree of similarity between two features instead of the Kendall and Spearman's rank correlation coefficients.

2.3. The Object Splitting Method for SO-PCA

PROCEDURE 1: Object splitting method for SO-PCA. For a set of N objects $\omega_1, \omega_2, \ldots, \omega_N$ under d interval valued features, the object splitting method is executed by the following steps.

1. We split each object ω_k into the minimum sub object $\omega_{k(\min)}$ and the maximum sub-object $\omega_{k(\max)}$. As the result, we have a $(2N) \times d$ numerical data table.
2. We calculate the $d \times d$ correlation matrix $\boldsymbol{S}$ for the $(2N) \times d$ data table obtained in (1) based on the selected correlation coefficient, where we can use the Kendall or Spearman's rank correlation coefficient or the Pearson correlation coefficient.
3. We find the principal components based on the correlation matrix in (2).
4. We represent each symbolic object ω_k in the factor planes as the arrow line connecting from $\omega_{k(\min)}$ to $\omega_{k(\max)}$, or as the Cartesian join of $\omega_{k(\min)}$ and $\omega_{k(\max)}$, i.e. a rectangular region spanned by $\omega_{k(\min)}$ and $\omega_{k(\max)}$.

EXAMPLE 3: Fats' and oils' data (interval-valued data).

We applied the object splitting method to the Fats' and oils' data of Table 1. We used only four interval features. The contribution ratios of the first two principal components

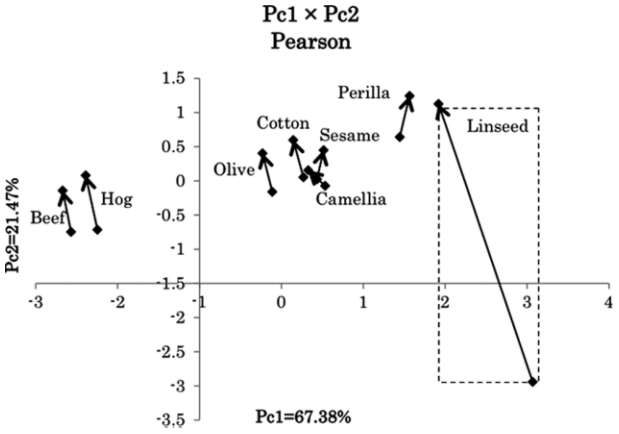

Fig. 4: The result of the SO-PCA for Fats' and oils' data (Pearson).

are 67.38% and 21.47% by the Pearson coefficient, and are 64.60% and 27.85% by the Spearman coefficient. Figs. 4 and 5 show the results of the SO-PCA based on the Pearson and the Spearman correlation matrices. These figures show the arrow line representation of eight objects in the factor planes. An arrow line connects from the minimum sub object to the maximum sub-object. The rectangle shown by dotted lines in Fig. 4 indicates the Cartesian join for *Linseed oil* that is the largest object. In Figs 4 and 5, the second principal component plays the size factor. On the other hand, in the first principal component, *Specific gravity* and *Iodine* value take positive weights, while *Freezing point* and *Saponification* value took negative weights. The mutual positions of the given eight symbolic objects are almost similar by the Pearson and the Spearman coefficients. However, we can see some differences in the lengths and the directions of some arrow lines by the Pearson and the Spearman coefficients. In this example, we should note that the arrow line representation supports a better

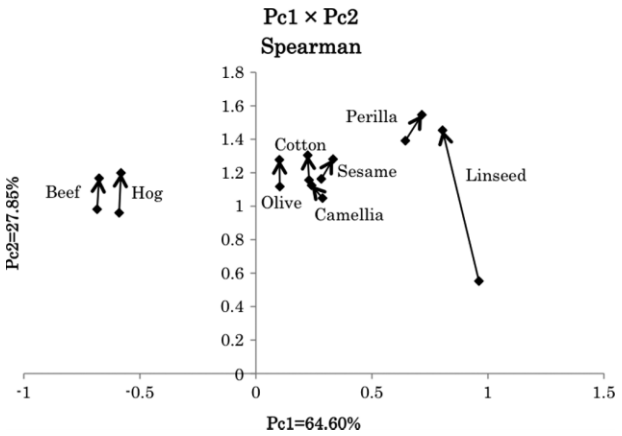

Fig. 5: The result of the SO-PCA for Fats' and oils' data (Spearman).

understanding for the descriptions of symbolic objects in the factor planes compared to the rectangular representation.

Chouakria *et al.* [6] presented a comparative study of the vertices method (V-PCA) and the centers method (C-PCA). The V-PCA is implemented on the numerical data table of the size $(N \times 2^d) \times d$, while the C-PCA is implemented on the size $N \times d$. Therefore, the C-PCA is stronger than the V-PCA in the computational complexity, when the number of descriptive features is large. The contribution ratios of the first two principal components for the fats' and oils' data of Table 1 are 68.29% and 20.23% by the V-PCA, and 75.23% and 15.09% by the C-PCA, respectively. The rectangular representations of objects for these two methods are similar, although their contribution ratios are different. Moreover, their results are also close to the arrow line representations in Figs 4 and 5.

Lauro *et al.* [8] presented a comparative study of the V-PCA, the method called spaghetti PCA, and the method based on interval algebra and optimization theory. For the Fats' and oils' data of Table 1, their results of rectangular representations in the first factor planes are mutually similar. Among them, the spaghetti PCA is especially close to the result in Figs 4 and 5. The spaghetti PCA uses the main diagonals of the hyper-rectangles to represent multidimensional interval data. The contribution ratios of the first two principal components are 71.33% and 18.09%. In the representation of interval objects in the first factor plane, the lengths and the directions of the main diagonals of the rectangular regions are very similar to those of the arrow lines in Figs 4 and 5. The spaghetti PCA is a very different method from the object splitting method. However, we should point out the fact that the main diagonal of an object may be described by two end points: the minimum vertex and the maximum vertex.

In this section, we presented the object splitting method of PCA for interval objects. This method transforms the given $N \times d$ interval-valued data table into a $2N \times d$ standard numerical data table, then executes the PCA on the transformed data table. We should note that

1. The object splitting method is simple and works as well as other methods for interval objects. Especially, this method is easily applicable to large data tables.
2. The arrow line representation of objects in the factor planes is useful to provide insights about the mutual relationships of the given interval objects.

In the next section, we present the *quantile method*, which is an extension of the object splitting method and can manipulate not only interval-valued features but also other type features including histogram features and nominal multi-valued features.

III. COMMON REPRESENTATION BY QUANTILES

In the aggregation process of large data sets, the use of histograms is very natural and common to describe the reduced data sets. Billard and Diday [2,4] summarize empirical distribution functions and descriptive statistics for various feature types. Based on knowledge of distribution functions, the quantile method [12] provides a common framework to represent symbolic data described by features of different types. The basic idea is to express the observed feature values by some predefined quantiles of the underlying distribution. In the interval feature case, a distribution is assumed within each interval, e.g., uniform distribution (Bertrand and Goupil [15]). For a histogram feature, quantiles of any histogram may be obtained simply by interpolation, assuming the uniformity in each *bin* of the histogram [2,4,15]. Although the numbers of bins of the given histograms are mutually different in general, we can obtain the same number of quantiles for each histogram. For nominal multi-valued features, quantiles are determined from ranking defined on the categorical values based on their frequencies. Therefore, when we choose quartiles, for example, we can represent each feature value for different feature types in the same form of a 5-tuple (min,

Q_1, Q_2, Q_3, max). This common representation then allows for a unified approach to S-PCA. In the following subsections, we describe detail procedures to have quantile values for various feature types.

3.1. Quantiles for Interval-valued Feature

Let a feature F_j be an interval-valued feature and let each object $\omega_k \in \boldsymbol{U}$ be represented by an interval:

$$I_{kj} = [x_{kj(\min)}, x_{kj(\max)}], \; k = 1, 2,\ldots,N. \tag{18}$$

We assume that each interval has a uniform distribu tion [2,4,15]. Then, in the case of m quantiles, the resultant $(m - 1)$ quantile values become

$$Q_{kji} = x_{kj(\min)} + (x_{kj(\max)} - x_{kj(\min)}) \times i/m, \; i = 1, 2,\ldots,m - 1. \tag{19}$$

Therefore, each object $\omega_k \in \boldsymbol{U}$ for the feature F_j is described by an $(m + 1)$-tuple:

$$(x_{kj(\min)}, Q_{kj1}, Q_{kj\,2},\ldots,Q_{kj(m-1)}, x_{kj(\max)}), \; k = 1, 2,\ldots,N. \tag{20}$$

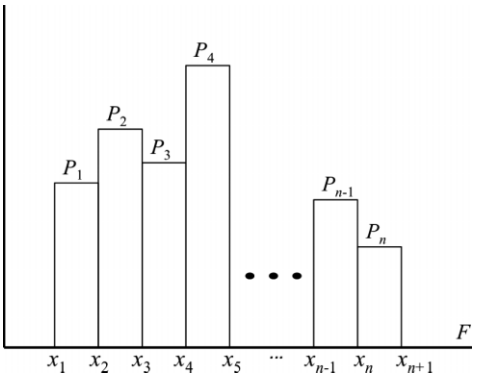

Fig. 6: A histogram-valued data.

3.2. Quantiles for Histogram-valued Feature

Let a feature F be a histogram feature and let an object $\omega \in \boldsymbol{U}$ be represented by a histogram in Fig. 6. Let the histogram be composed of n bins, and let p_i be the probability of the ith bin, where it is assumed that $p_1 + p_2 +\cdots+ p_n = 1$. Then, under the assumption that n bins (subintervals) have uniform distributions, we define the cumulative distribution function $F(x)$ of the histogram [2,4] as:

$$F(x) = 0 \text{ for } x \le x_1$$
$$F(x) = p_1(x - x_1)/(x_2 - x_1) \text{ for } x_1 \le x < x_2 \; F(x) = F(x_2) + p_2(x - x_2)/(x_3 - x_2) \text{ for } x_2 \le x < x_3 \cdots\cdots$$
$$F(x) = F(x_n) + p_n(x - x_n)/(x_{n+1} - x_n) \text{ for } x_n \le x < x_{n+1}$$
$$F(x) = 1 \text{ for } x_{n+1} \le x.$$

Then, in the case of m quantiles, we can find $(m + 1)$ values including $(m - 1)$ quantile values from the equations:

$$F(\min) = 0, \ (\text{i.e. } \min = x_1)$$
$$F(Q_2) = 1/m, \ F(Q_3) = 2/m \ldots, F(Q_m)$$
$$= (m - 1)/m, \text{ and}$$
$$F(\max) = 1, \ (\text{i.e. } \max = x_{n+1}).$$

Therefore, the object $\omega_k \in \boldsymbol{U}$ is described by an $(m + 1)$- tuple

$$(x_{\min}, Q_1, Q_2, \ldots, Q_{m-1}, x_{\max}). \tag{21}$$

In general, we can describe each object $\omega_k \in \boldsymbol{U}$ under a histogram-valued feature F_j by an $(m + 1)$-tuple:

$$(x_{kj(\min)}, Q_{kj1}, Q_{kj2}, \ldots, Q_{kj(m-1)}, x_{kj(\max)}),$$
$$k = 1, 2, \ldots, N. \tag{22}$$

It should be noted that the numbers of bins of the given histograms are mutually different in general. However, we can select an integer number m, and obtain $(m + 1)$-tuples as the common representation for all histograms.

3.3. Quantiles for Nominal (categorical) Multi-valued Feature

Let F_j be a multi-valued feature which takes n possible categorical values c_i, $i = 1, 2, \ldots, n$. For each i, let p_i be the relative frequency of categorical value c_i in terms of N objects [2,4,15]. Then, we sort the relative frequency values. For simplicity, we assume that:

$$p_1 \leq p_2 \leq \cdots \leq p_n. \tag{23}$$

According to this order, we suppose rank values 1, 2, ... , n for the categorical values $c_1, c_2, \ldots, c_n$, respectively. We define the cumulative distribution function for each object $\omega_k \in \boldsymbol{U}$ based on the rank values.

Let n_k be the number of possible categorical values taken by object $\omega_k \in \boldsymbol{U}$ under F_j. Let q_{ki} be the frequency value associated with the category c_i and given by

$$q_{ki} = 1/n_k \text{ if } c_i \text{ is a possible value for}$$
$$\omega_k \in \mathbf{U} \text{ under } F_j, = 0 \text{ otherwise.}$$

Therefore, we define a piecewise linear cumulative distribution function for each object $\omega_k \in \boldsymbol{U}$ based on uniform densities attached to rank values (see Example 4). Then we find $(m + 1)$ values including quantile values for the selected integer number m. Therefore, we can obtain again the common $(m + 1)$-tuple representation:

$$(x_{kj(\min)}, Q_{kj1}, Q_{kj2}, \ldots, Q_{kj(m-1)}, x_{kj(\max)}),$$
$$k = 1, 2, \ldots, N. \tag{24}$$

EXAMPLE 4: The *fifth* feature (*Major acids*) of Table 1 is an example of nominal multi-valued feature. We suppose the *quartile* case, i.e. $m = 4$. For this purpose, we use basically the procedure given in the above. However, in order to prevent ties of rank values, we use the sums of frequency

values attached to the category values of each object. Table 2 summarizes the results. From this table, the object *Linseed*, for example, has the min*imum* value *four* and the *maximum* value *nine*. Then, we assume uniform probability densities for the intervals associated with the ranking values:

$$[1,4[:0;[4,5[:0.2;[5,6[:0.2;[6,7[:0.2;[7,8[:0.2; [8,9[:0.2;[9,10]:0.2, \tag{25}$$

where we should note that the interval [9,10] is attached to the maximum rank value *nine*. The corresponding cumulative distribution function is a piecewise linear function $F(x)$ characterized by:

$$\begin{aligned} F(x) &= 0,\ 1 \le x < 4; \\ F(x) &= 0.2 \times (x-4),\ 4 \le x < 5; \\ F(x) &= 0.2 + 0.2 \times (x-5),\ 5 \le x < 6; \\ F(x) &= 0.4,\ 6 \le x < 7; \\ F(x) &= 0.4 + 0.2 \times (x-7),\ 7 \le x < 8; \\ F(x) &= 0.6 + 0.2 \times (x-8),\ 8 \le x < 9; \\ F(x) &= 0.8 + 0.2 \times (x-9),\ 9 \le x \le 10. \end{aligned} \tag{26}$$

By solving the equations $F(x) = 0.25$, $F(x) = 0.5$, and $F(x) = 0.75$, we obtain the quartile values $Q_1 = 5.25$,

Table 2: The sums of frequency values and rank values.

Object	Lu	A	C	Ln	M	S	P	L	O
Linseed	0	0	0	0.2	0.2	0	0.2	0.2	0.2
Perilla	0	0	0	0.2	0	0.2	0.2	0.2	0.2
Cotton	0	0	0	0	0.2	0.2	0.2	0.2	0.2
Sesame	0	0.2	0	0	0	0.2	0.2	0.2	0.2
Camellia	0	0	0	0	0	0	0	0.5	0.5
Olive	0	0	0	0	0	0.25	0.25	0.25	0.25
Beef	0	0	0.2	0	0.2	0.2	0.2	0	0.2
Hog	0.167	0	0	0	0.167	0.167	0.167	0.167	0.167
$\sum q_{ij}$	0.167	0.2	0.2	0.4	0.767	1.217	1.417	1.717	1.917
Rank	1	2	2	4	5	6	7	8	9

$Q_2 = 7.5$, and $Q_3 = 8.75$, respectively. Finally, we have the desired 5-tuple:

$$(4,\ 5.25,\ 7.5,\ 8.75,\ 10). \tag{27}$$

It should be noted that we are also able to treat other feature types such as discrete multi-valued features, binary features, etc., by assuming appropriate distribution functions.

IV. THE QUANTILE METHOD FOR S-PCA

Let $\boldsymbol{U} = \{\omega_1, \omega_2, \dots, \omega_N\}$ be given objects. Let each object ω_k be described by d features. In general, d features are a mixture of interval features, histogram features, nominal multi-valued features and other types.

DEFINITION 9: Quantile sub-objects.

Let each object $\omega_k \in \boldsymbol{U}$ be described with the given d features by $(m+1)$-tuples:

$$(x_{kj(\min)}, Q_{kj1}, Q_{kj2},\ldots,Q_{kj(m-1)}, x_{kj(\max)}),$$
$$j = 1, 2,\ldots,d;\ k = 1, 2,\ldots,N. \tag{28}$$

Then, we define the *quantile* sub-object ω_{kQi} as:

$$\boldsymbol{x}_{kQi} = (Q_{k1i}, Q_{k2i},\ldots,Q_{kdi}),\ i = 1.2,\ldots,m-1;\ \mathrm{k} = 1, 2,\ldots,N. \tag{29}$$

PROPOSITION 7: For each object $\omega_k \in \boldsymbol{U}$, the *min imum* sub-object $\omega_{k(\min)}$, $(m-1)$ *quantile* sub-objects $(\omega_{kQ1}, \omega_{kQ2},\ldots, \omega_{kQ(m-1)})$, and the *maximum* sub-object $\omega_{k(\max)}$ organize a monotone structure in the space $\boldsymbol{R}^d$.

Proof: From the definition of $(m+1)$ sub-objects, we can obtain the following nesting relations of the Cartesian join regions:

$$\boldsymbol{J}(\omega_{k(\min)}, \omega_{kQ1}) \subseteq \boldsymbol{J}(\omega_{k(\min)}, \omega_{kQ2}) \subseteq \cdots \subseteq \boldsymbol{J}(\omega_{k(\min)}, \omega_{kQ(m-1)}) \subseteq \boldsymbol{J}(\omega_{k(\min)}, \omega_{k(\max)}). \tag{30}$$

Thus, Definition 7 leads the conclusion.

PROPOSITION 8: Property of correlation matrix $\boldsymbol{S}$ by the *quantile method*

Let a series of objects $\omega_k \in \boldsymbol{U}$, $k = 1, 2,\ldots, N$, is mono tone in the space $\boldsymbol{R}^d$ and let the $d \times d$ correlation matrix $\boldsymbol{S}$ be obtained by applying the Kendall or Spearman's rank correlation coefficients to the $N \times (m+1)$ sub-objects of Definition 9. Then, the absolute value of each off-diagonal element of $\boldsymbol{S}$ is large.

Proof: From Proposition 7, $(m+1)$ sub-objects for each of N objects organize always a monotone structure in any subspace of $\boldsymbol{R}^d$. Therefore, if the given series of objects is monotone, their nesting property restrict the order of $N \times (m+1)$ sub-objects to be similar in any subspace of $\boldsymbol{R}^d$. This leads to the conclusion.

Now, the quantile method for general S-PCA is summa rized as follows.

PROCEDURE 2: The quantile method for S-PCA Let the set of N objects $\omega_1, \omega_2, \ldots, \omega_N$ be described by d features, which are a mixture of interval features, histogram features, nominal multi-valued features, and other types. Then, we execute the quantile method by the following steps.

1. We select an integer value m $(1 \le m < N)$.
2. For each feature F_j, we find the common represen tation of N objects by the $(m+1)$-tuples:

$$(x_{kj(\min)}, Q_{kj1}, Q_{kj2},\ldots,Q_{kj(m-1)},$$
$$x_{kj(\max)}),\ k = 1, 2,\ldots,N.$$

3. For each object ω_k, we find $(m+1)$ d-dimensional sub-objects: the *minimum* sub-object $\omega_{k(\min)}$, $(m-1)$ *quantile* sub-objects, $\omega_{kQ1}, \omega_{kQ2}, \ldots, \omega_{kQ(m-1)}$, and the maximum sub-object $\omega_{k(\max)}$. Then we split each object into $(m+1)$ sub-objects. As the result, we have an $\{N \times (m+1)\} \times d$ numerical data table.
4. We calculate the $d \times d$ correlation matrix $\boldsymbol{S}$ for the $\{N \times (m+1)\} \times d$ data table obtained in 3) based on the selected correlation coefficient, where we can use the Kendall or Spearman's rank correlation coefficient, or the Pearson correlation coefficient.
5. We find the principal components based on the correlation matrix in 4).

In the factor planes, we can reproduce each object ω_k, $k = 1, 2, \ldots, N$, as a series of m arrow lines:

$$\omega_{k(\min)} \longrightarrow \omega_{kQ1} \longrightarrow \omega_{kQ2} \longrightarrow \cdots \longrightarrow \omega_{kQ(m-1)} \longrightarrow \omega_{k(\max)}. \tag{31}$$

As a different representation, we can use also a series of m rectangles.

In this procedure, if we select as $m = 1$, the quantile method is reduced to the original "object splitting method".

V. EXAMPLES OF THE QUANTILE METHOD FOR S-PCA

EXAMPLE 5: Fats' and oils' data

We illustrate the *quartile* case, i.e. $m = 4$. In this case, the common representation of each object under a feature is 5-tuple, i.e. (min, Q_1, Q_2, Q_3, max). For the fifth feature *Major acids*, we used the quantification in Example 4. For the data in Table 1, we obtain the necessary 5-tuples for each of the *eight* objects with respect to *five* features. Then, we split each object into *five* sub-objects, i.e. the *minimum* sub-object, three *quantile* sub-objects, and the *maximum* sub-object. Therefore, we have 40 sub-objects for the given *eight* objects. Table 3 shows a part of our data, where *five* sub-objects are presented only for *Linseed* and *Perilla*. Table 4 shows the correlation matrices by the Pearson and Spearman coefficients. The elements of the upper triangular matrix show the Pearson correlation coefficients and those of the lower triangular matrix show the Spearman correlation coefficients. For the Spearman coefficient, *Specific gravity* and *Iodine* values show the highest correlation, while, for the Pearson coefficient, *Specific gravity* and *Freezing point* values show the negatively highest coefficient. We can find another difference between the Spearman and the Pearson coefficients for the correlation between *Saponification* and *Major acids* values.

Figs. 7 and 8 show the results of the S-PCA based on the Pearson and Spearman correlation matrices. The contribution ratios of the first two principal components are 56.28% and 26.68% by the Pearson coefficient, and are 53.71% and 34.54% by the Spearman coefficient. Each object is described by a series of *four* arrow lines. In this example, the second principal component plays the role of the size factor, and *Major acids* and *Saponification* value take especially large positive weights. On the other hand, in the first principal component, *Specific gravity* and *Iodine* values take large positive weights, while *Freezing point* and *Saponification* values took large negative values.

Table 3: A part of the Fats' and oils' data (quartile data). $F_1 F_2 F_3 F_4 F_5$

	F_1	F_2	F_3	F_4	F_5
Linseed					
1	0.93000	−27	170	118	4
2	0.93125	−24.75	178.5	137.5	5.25
3	0.93250	−22.5	187	157	7.5
4	0.93375	−20.25	195.5	176.5	8.75
5	0.93500	−18	204	196	10
Perilla					
1	0.93000	−5	192	188	4
2	0.93175	−4.75	196	190.25	6.25
3	0.93350	−4.5	200	192.5	7.5
4	0.93525	−4.25	204	194.75	8.75
5	0.93700	−4	208	197	10

Table 4: The Spearman and the Pearson correlation matrices.

S	Spec.	Freez.	Iodine	Sapon.	M. acids
Spec.	1.0000	−0.8923	0.7682	−0.3187	0.2432
Freez.	−0.6309	1.0000	−0.6368	0.4968	−0.1138
Iodine	0.9582	−0.6142	1.0000	−0.3834	0.1107
Saponi.	−0.2044	0.6437	−0.1980	1.0000	0.3634
M. acids	0.2558	0.0398	0.1805	0.6428	1.0000

By the addition of *Major acids*, the lengths of objects became larger than those in Figs 4 and 5. The position of *Camellia* moved slightly towards the upper side. The mutual positions of eight objects are almost the same in Figs 7 and 8. However, some differences exist in the sizes and the directions of arrow lines.

EXAMPLE 6: Histogram data (Hardwood data)

The histogram data used here are selected from the US Geological Survey (Climate-Vegetation Atlas of North America [16]). The number of objects is 16 and the number of features is *eight*. Table 5 shows histogram values for 16 *hardwoods* under the feature: *Annual temperature* (ANNT). In this table, *N* is the number of preselected regions in which the hardwood exists. For example, *Acer East* is observed in 6869 regions. The ANNT of these 6869 regions is in the range from −2.3 to 23.8°C. In other words, 100% of 6869 *Acer East* samples exist in the range from −2.3 to 23.8°C, while 50% of 6869 samples exist in the range between −2.3 and 9.2°C, and so on. We selected the following eight features to describe objects (hardwoods). The data formats for other features F_2 –F_8 are the same with Table 5, viz.,

F_1: Annual temperature (ANNT) (°C);

F_2: January temperature (JANT) (°C);

F_3: July temperature (JULT) (°C);

F_4: Annual precipitation (ANNP) (mm);

F_5: January precipitation (JANP) (mm);

F_6: July precipitation (JULP) (mm);

F_7: Growing degree days on 5°C base × 1000 (GDC5); and

F_8: Moisture index (MITM).

In this example, *deciles* and *quartiles* describe each object, where the preselected number *m* is 6, and the 7-tuple is used as a common representation for the given
Ichino: The Quantile Method for Symbolic PCA 195

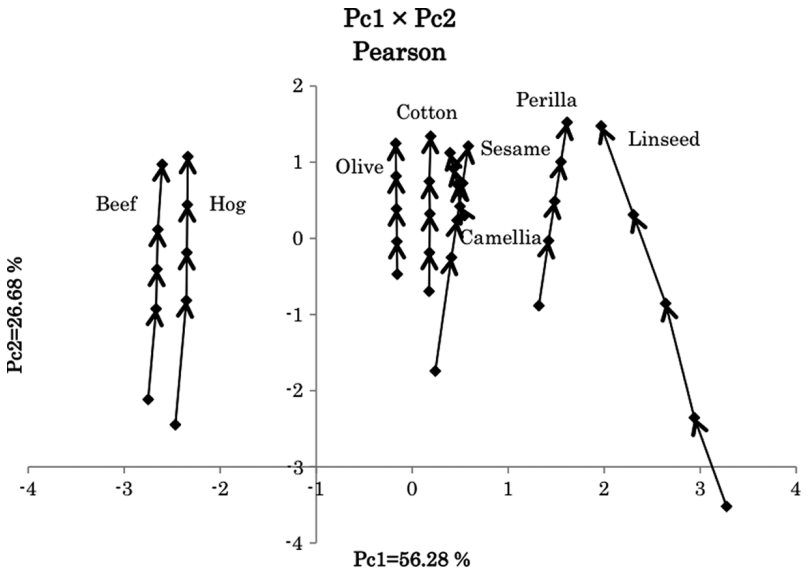

Fig. 7: The result of the S-PCA for Fats' and oils' data (Pearson).

16 hardwoods. According to the Procedure 2 for S-PCA in Section 4, we transform the given (16 objects) × (8 features) symbolic data table to a (16 × 7 sub-objects) × (8 features) standard numerical data table. Table 6 shows a part of the transformed data table.

Table 7 shows the 8 × 8 correlation matrices, where the upper triangular matrix shows the elements of the Pear son correlation matrix, and the lower triangular matrix shows the elements of the Spearman's rank correlation matrix. The Pearson and the Spearman correlation matrices are similar in many elements. However, some differences should be pointed out. Features F_1 (ANNT), F_2 (JANT), F_3 (JULT), and F_7 (GDC5) are highly correlated mutually for the Spearman coefficient. Feature F_4 (ANNP) is strongly correlated with features F_5 (JANP) and F_8 (MITH) for the Spearman coefficient, while F_4 (ANNP) is largely cor related with features F_5 (JANP) and F_6 (JULP) for the Pearson coefficient. We see also a difference between the Pearson and Spearman correlation coefficients concerning feature F_7 (GDC5).

The contribution ratios of the first two principal compo nents are 77.01% and 11.64% for the Pearson correlation matrix, and are 87.41% and 8.38% for the Spear man correlation matrix. Figs 9 and 10 show the arrow

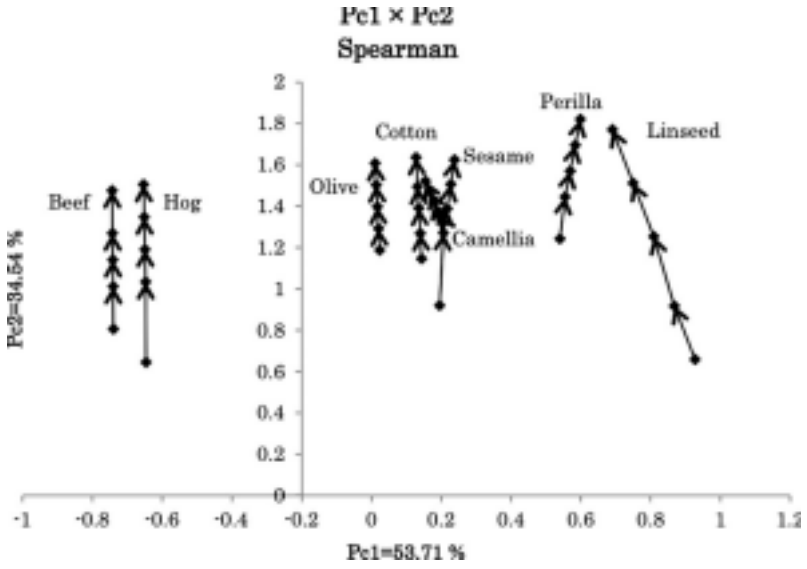

Fig. 8: The result of the S-PCA for Fats' and oils' data (Spearman).

Table 5: Histogram data (annual temperature).

	Annual temperature (°C)							
Taxon name	*N*	0%	10%	25%	50%	75%	90%	100%
Acer East	6 865	−2.3	0.6	3.8	9.2	14.4	17.9	23.8
Acer West	1 954	−3.9	0.2	1.9	4.2	7.5	10.3	20.6
Alnus East	10 144	−10.2	−4.4	−2.3	0.6	6.1	15.0	20.9
Alnus West	4 761	−12.2	−4.6	−3.0	0.3	3.2	7.6	18.7
Betula	16 815	−13.4	−8.4	−5.1	−1.0	3.9	12.6	20.3
Carya	4 638	3.6	7.5	10.0	13.6	17.2	19.4	23.5
Castanea	2 216	4.4	8.6	11.3	14.9	17.5	19.2	21.5
Fraxinus East	8 565	−2.3	1.4	4.3	8.6	14.1	17.9	23.2
Fraxinus West	1 095	2.6	9.4	11.5	17.2	21.2	22.7	24.4
Juglans East	4 138	1.3	6.9	9.1	12.4	15.5	17.6	21.4
Juglans West	526	7.3	12.6	14.1	16.3	19.4	22.7	26.6
Ostrya/Carpinus	5 348	1.2	4.4	7.0	11.4	16.0	19.2	28.0
Quercus East	7 360	−1.5	3.4	6.3	11.2	16.4	19.1	24.2
Quercus west	1 942	−1.5	6.0	9.5	14.6	17.9	19.9	27.2
Tilia	3 792	1.1	3.8	5.8	8.8	12.0	14.4	19.9
Ulmus	8 028	−2.3	1.7	4.9	9.7	15.3	18.6	23.8

Table 6: A part of quantile data.

Taxon name	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8
Acer East								
1	−2.3	−24.6	11.5	415	10	56	0.5	0.62
2	.6 0	−18.3	16.6	720	23	77	1.2	0.89
3	.8 3	−12.3	18.2	835	40	89	1.5	0.94
4	.2 9	−5.1	22.2	1010	69	100	2.5	0.97
5	.4 14	2.3	25.8	1200	96	113	3.6	0.99
6	.9 17	7.9	27.3	1355	127	135	4.8	0.99
7	.8 23	18.9	28.8	1630	166	222	6.8	1.00
Acer West								
1	−3.9	−23.8	7.1	105	5	0	0.1	0.14
2	.2 0	−11.8	11.3	380	28	8	0.5	0.49
3	.9 1	−10.1	12.8	505	54	23	0.7	0.61
4	.2 4	−6.9	14.9	750	92	38	1.1	0.75
5	.5 7	−1.3	17.6	1175	176	52	1.6	0.91
6	.3 10	3.3	19.9	1860	267	71	2.2	0.98
7	.6 20	11.0	29.2	4370	616	160	5.6	1.00
Alnus East								
1	−10.2	−30.9	7.1	220	9	28	0.1	0.22
2	−4.4	−26.5	13.2	380	19	58	0.6	0.53
3	−2.3	−22.7	14.8	475	23	74	0.8	0.69
4	.6 0	−18.1	16.5	770	46	91	1.1	0.93
5	.1 6	−8.0	19.8	1060	80	108	1.9	0.99
6	.0 15	3.7	25.7	1235	106	126	3.7	0.99
7	.9 20	14.1	29.1	1650	166	212	5.9	1.00
Alnus West								
1	−12.2	−30.5	7.1	170	4	0	0.1	0.22
2	−4.6	−25.7	11.5	335	18	21	0.5	0.49
3	−3.0	−21.6	12.8	410	23	41	0.7	0.59
4	.3 0	−15.1	14.4	510	37	57	0.9	0.72
5	.2 3	−7.6	15.6	790	93	74	1.1	0.87
6	.6 7	−0.8	17.5	1385	199	87	1.6	0.97
7	.7 18	10.8	28.3	4685	667	452	4.8	1.00

Table 7: The Spearman and the Pearson correlation matrices.

S	ANNT	JANT	JULT	ANNP	JANP	JULP	GDC5	MITM
ANNT	1.0000	0.9836	0.9666	0.6524	0.5633	0.7573	0.9584	0.6381
JANT	0.9851	1.0000	0.9190	0.6354	0.5635	0.7438	0.9424	0.5754
JULT	0.9706	0.9242	1.0000	0.6856	0.5888	0.7673	0.9329	0.7477
ANNP	0.8154	0.7746	0.8660	1.0000	0.9349	0.8237	0.6695	0.6154
JANP	0.7828	0.7689	0.8139	0.9548	1.0000	0.7062	0.5854	0.4704
JULP	0.8356	0.7834	0.8859	0.8999	0.8149	1.0000	0.8302	0.6360
GDC5	0.9934	0.9650	0.9837	0.8211	0.7744	0.8635	1.0000	0.5792
MITM	0.7268	0.6678	0.8079	0.9569	0.8798	0.8899	0.7458	1.0000

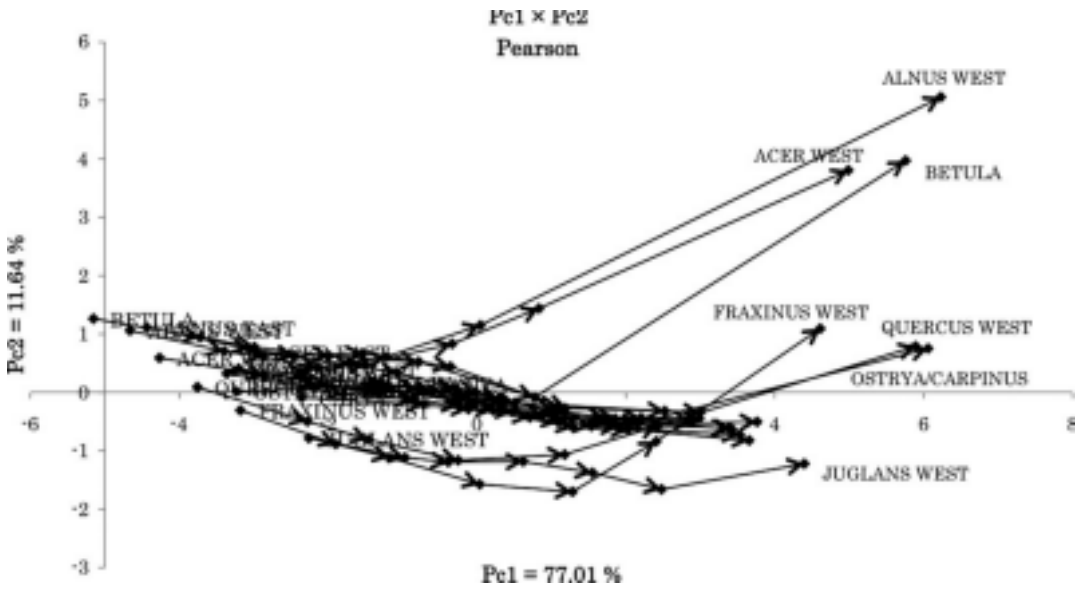

Fig. 9: The result of the S-PCA for hard woods (Pearson).

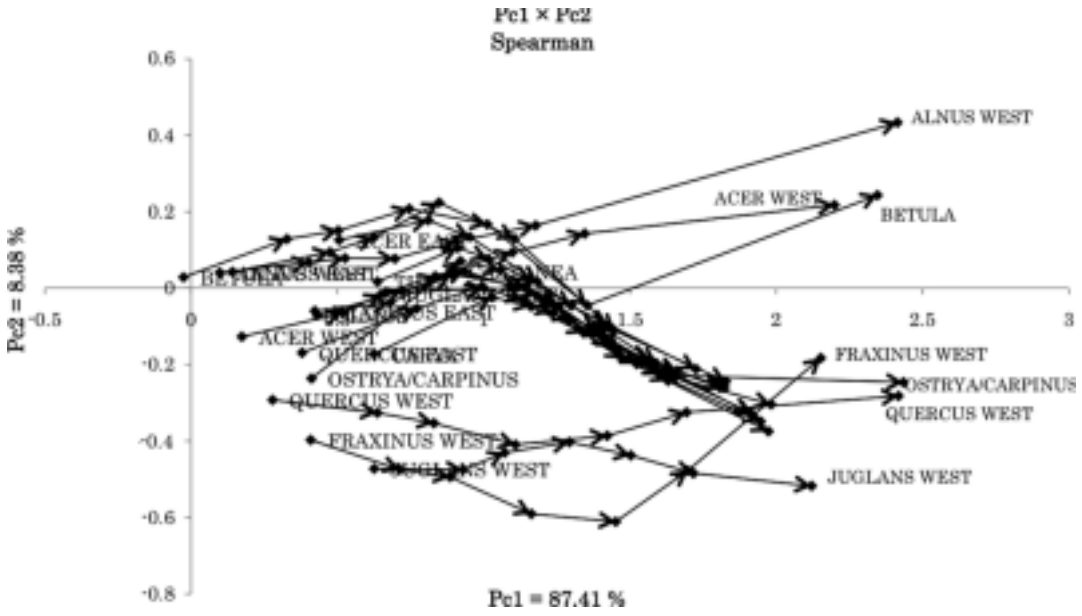

Fig. 10: The result of the S-PCA for hard woods (Spearman).

line representations of sixteen hardwoods in the factor planes by the Pearson and Spearman correlation matrices, respectively. In the two factor planes, the first principal component plays the role of the size factor, and the given eight features take similar positive weights. In the second principal component, four features concerning *precipita tion* and *moisture*, i.e. ANNP, JANP, JULP, and MITH, take positive weights, while other features for *temperature* and *growing degree*, i.e. ANNT, JANT, JULT, and GDC5, took negative weights. For the Spearman correlation matrix, *moisture* (MITH) takes an especially large positive weight for the second principal component. However, for the Pear son correlation matrix, the corresponding weight is very small.

In Fig. 9, many series of arrow lines tend to be slightly right down. Almost all kinds of hardwood in the eastern area of the US organize a large stream of arrow lines. This tendency of the main stream depends on temperature and precipitation. On the other hand, largely fluctuating and mutually separate streams are mainly composed of the hardwoods in the western area. For example, *Acer West*, *Alnus West*, *Betula*, and *Fraxinus West* most drastically change toward the upper right with the last decile. This change is heavily dependent on precipitation and moisture. In Fig. 10, the main stream of arrow lines has two branches. Each branch initially grows toward the upper right, and then changes direction toward right down. This property is not clear in Fig. 9. Generally, mutual arrow lines are clearly represented in Fig. 10. Therefore, in this example, the Spearman correlation matrix may be better than the Pearson correlation matrix. Since the quantile method is based on the monotonic property of the given set of objects, the use of the Spearman correlation matrix may be natural.

VI. CONCLUDING REMARKS

We presented the quantile method for the S-PCA. The quantile method can treat not only histogram-valued data, but also nominal and ordinal multi-valued type data, and is simply based on the property of monotone structure of the given objects. By selecting a common integer number m, the quantile method transforms a given $N \times d$ complex symbolic data table to a simple $(N \times (m + 1)) \times d$ numerical data table. An important aspect is that we can select the integer m as a sufficiently small number compared to the number N of objects, and we can apply the traditional PCA simply to the $(N \times (m + 1)) \times d$ data table. We presented several experimental results in order to show the effectiveness of the quantile method. An arrow line representation of objects in the factor plane may be a useful tool to analyze complex symbolic data tables.

ACKNOWLEDGMENTS

This research was partly supported by Japan Society of the Promotion of Science (Grant-in Aid for Scientific Research (C) 19500130). The author wishes to thank referees and editors for suggestions leading improvements in this article. The author also acknowledges to Professor Paula Brito for her collaborations.

REFERENCES

1. H. H. Bock and E. Diday, eds., Analysis of Symbolic Data, Exploratory Methods for Extracting Statistical Information from Complex Data, Berlin, Springer-Verlag, 2000.
2. L. Billard and E. Diday, Symbolic Data Analysis: Concep tual Statistics and Data Mining, Chichester, Wiley, 2006.
3. E. Diday and M. Noirhomme-Fraiture, eds., Symbolic Data Analysis and the SODAS Software, Chichester, Wiley, 2008.
4. L. Billard and E. Diday, From the statistics of data to the statistics of knowledge: symbolic data analysis, J Am Stat Assoc 98(462) (2003), 470–487.

5. A. Chouakria, Extension de l'analyse en composantes principales a des donnees de type intervalle. Doctoral ´ Thesis; University of Paris IX Dauphine, 1998.
6. A. Chouakria, P. Cazes, and E. Diday, Symbolic princi pal component analysis, In Analysis of Symbolic Data, H.-H. Bock and E. Diday, eds. Berlin, Springer-Verlag, 2000.
7. C. Lauro and F. Palumbo, Principal component analysis of interval data: a symbolic data analysis approach, Comput Stat 15(1) (2000), 73–87.
8. C. Lauro, R. Verde, and A. Irpino, Principal component analysis of symbolic data described by intervals, In Sym bolic Data Analysis and the SODAS Software, E. Diday and M. Noirhomme-Fraiture, eds. Chichester, Wiley, 2008, 279–311.
9. M. Ichino, General metrics for mixed features—the Carte sian space theory for pattern recognition, In Proceedings on International Conference on Systems, Man, and Cybernet ics, Chaina, Beijin, 1988.
10. M. Ichino and H. Yaguchi, Generalized Minkowski metrics for mixed feature type data analysis, IEEE Trans Syst Man Cybern 24(4) (1994), 698–708.
11. M. Ichino, Symbolic principal component analysis based on the nested covering, In Proceedings ISI2007, Portugal, Lisbon, 2007.
12. M. Ichino, Symbolic PCA for histogram-valued data, In Proceedings of IASC 2008, Japan, Yokohama, 2008. [13] M. Ichino and H. Yaguchi, Symbolic pattern classifiers
13. based on the Cartesian system model, In Data Science, Classification, and Related Methods, C. Hayashi, *et al.*, eds. Tokyo, Springer-Verlag, 1998, 358–369.
14. C. Chatfield and A. J. Collins, Introduction to Multivariate Analysis, Chapter 4, Chapman and Hall, New York, 1984. [15] P. Bertrand and F. Goupil, Descriptive statistics for sym bolic data, In Analysis of Symbolic Data, H.-H. Bock and E. Diday, eds. Berlin, Springer-Verlag, 2000.
15. Histogram data by the U.S. Geological Survey, Climate Vegetation Atlas of North America, http://pubs.usgs.gov/pp/ p1650-b/. [Last accessed October 2, 2008].

This page is intentionally left blank

Tachyons are Messengers for Entangled Particles

Ardeshir Irani

ABSTRACT

Tachyons are Scalar Bosons that reside in our fourth dimensional spatial Universe. They are point particles with zero mass, zero charge, and zero spin. Based on the Lorentz Transformation for v > c, tachyons can travel backwards in time to enter our third dimensional spatial Universe and can travel faster than light speed, to transmit information instantaneously between two entangled particles, to change their spin values from a superposition state to a definitive up and down state. We will show that the laws of Quantum Mechanics can be retro-causal while the laws of Classical Mechanics must be causal.

Keywords: tachyon, electron, entanglement, superposition, qtd, lorentz transformation.

Classification: LCC Code: QC174.12

Language: English

Great Britain Journals Press

LJP Copyright ID: 925614
Print ISSN: 2631-8490
Online ISSN: 2631-8504

London Journal of Research in Science: Natural and Formal

Volume 23 | Issue 12 | Compilation 1.0

Tachyons are Messengers for Entangled Particles

Ardeshir Irani

ABSTRACT

Tachyons are Scalar Bosons that reside in our fourth dimensional spatial Universe. They are point particles with zero mass, zero charge, and zero spin. Based on the Lorentz Transformation for v > c, tachyons can travel backwards in time to enter our third dimensional spatial Universe and can travel faster than light speed, to transmit information instantaneously between two entangled particles, to change their spin values from a superposition state to a definitive up and down state. We will show that the laws of Quantum Mechanics can be retro-causal while the laws of Classical Mechanics must be causal.

Keywords: tachyon, electron, entanglement, superposition, qtd, lorentz transformation.

Author: The Dark Energy Research Institute, 12343 Chavers Avenue, Downey, Ca. 90242, USA.

I. INTRODUCTION

For QED (Quantum Electro Dynamics) between two charged particles the messenger particle is the photon, and for QCD (Quantum Chromo Dynamics) between three quarks or between quark anti-quark pairs the messenger particle is the gluon. We will show that for QTD (Quantum Tachyonic Dynamics) between two entangled particles the messenger particle is the tachyon. The tachyon is a point particle with zero mass and zero charge like the photon and the gluon. However, we will show that it has zero spin and hence it is a Scalar Boson like the Higgs Boson. Like the photon and the gluon, it does not interact with the Higgs field because if it did, it would receive mass to slow it down. Therefore, the tachyon is very different from the W+- and the Z^0 Bosons of the Weak interaction which get their mass from the Higgs field. So far in Physics literature tachyons have been postulated to travel faster than the speed of light, and in science fiction the use of tachyon beams is made to be able to consider sending messages to the past (1).

We will show this to be true mathematically using the Lorentz transformation $t = t_0/\sqrt{1 - v^2/c^2}$ which becomes $t = -i\,t_0/\sqrt{v^2/c^2 - 1}$ for v > c. The axis of the imaginary number "i" in the equation is perpendicular to the three real number axes x, y, z in our 3D space and therefore the tachyon lies in our fourth dimension (4D space), the negative sign indicating it can travel backwards in time from 4D to 3D since 3D space was created before 4D space. For v very close to c, v ≥ c, time does not exist because $\sqrt{v^2/c^2 - 1}$ becomes zero, and $t \to \infty$ as time becomes ageless. The photon with speed v = c does not age as it travels indefinitely through space. In the regime of v ≥ c, the tachyon behaves much like the photon. Because of it being massless, the problem of decaying into smaller mass particles does not exist. In that sense it is elementary like the photon, the gluon, and the electron which has the smallest mass of all the elementary particles. The situation changes dramatically as $v \to \infty$ (infinity), $\sqrt{v^2/c^2 - 1} \to \infty$, and the tachyonic motion is instantaneous as $t \to 0$. Traveling at infinite speed implies the time of travel for the tachyon is zero. Consider the case for $\gamma = -i/\sqrt{v^2/c^2 - 1}$, when v > c, we get from $E = \gamma mc^2$ and $p = \gamma mv$, $E^2 = p^2c^2 - m^2c^4$. For particles with mass m in 4D space since the Energy E must

be positive, p > mc which must be the case since v > c. In the case of the massless tachyon E = pc and from the wave-particle duality of Quantum Mechanics p = h/λ and E = hc/λ. By comparison the highest energy Gamma ray photon in our 3D Universe would have energy of about 10^{-8} Joules and this is because the speed of the photon cannot exceed 3x 10^{8} m/s. For the case of tachyons E = hc/λ and as the wavelength λ of the tachyon approaches zero the Energy of the tachyon approaches infinity. The tachyon borrows this tremendous amount of energy for a very short period from the vacuum energy density of 4D space based on Heisenberg's Uncertainty Principle ΔExΔt ≥ h/2π. As ΔE increases, Δt simultaneously decreases. This spontaneous burst of energy allows the tachyon to enter 3D space to interact with the two entangled electrons. Having changed their quantum states to a classical state as will be further discussed below, the tachyon gives up its borrowed energy to the vacuum energy density of 3D space. The equation wants to pull the tachyon in its now photonic-like normal state of eternal agelessness back into 4D space since its equation contains "i", the fourth dimension of space, and that is where tachyons must reside before being called upon once again to untangle two entangled particles in our 3D space. The portal between 3D and 4D space is the Black Hole. Once the tachyon enters the Black Hole in 3D it returns to its home in 4D space. This phenomenon is called the tachyonic loop of space-time. *The existence of our 4D space has been confirmed both theoretically and experimentally because that is where Dark Matter of our 3D Universe resides (2).* The reason we cannot observe the existence of tachyons in our 3D space is because they travel faster than the speed of light and therefore photons are unable to reach them to be able to observe them, and so they remain spooky (implying ghostly) particles as referred to by Einstein who called the interaction of tachyons with entangled particles "spooky action at a distance".

Matter entering 4D from our 3D space increases the energy of 4D space while decreasing the energy of 3D space. Tachyons borrow energy from 4D space energy density, returning the energy to 3D space energy density. Hence the reverse process takes place whereby the energy of 4D space is decreased while the energy of 3D space is increased. However, these two processes are not in equilibrium because in another 2.6 billion years our 3D Universe will become part of the 4D Universe (2), with all the energy existing in 4D space. This implies that the number of particles that are not entangled greatly exceeds the number of entangled particles in our 3D Universe. As an example, an electron and a positron can annihilate each other to produce two entangled photons but these entangled photons would be only a small fraction of the total number of photons that exist in our 3D Universe.

Using QTD we will explain the 2022 Nobel Prize in Physics which states that Quantum Mechanics is not a deterministic theory like Classical Mechanics with a speed limit that is the speed of light, since two entangled particles can exchange information instantaneously even when separated by huge distances in space. Consider two entangled electrons both with a superposition of spin up and spin down states. The tachyon travels back in time from 4D to 3D space to contact one of the entangled electrons to change its superposition state to a deterministic value of either spin up or spin down. Then it approaches the second entangled electron instantaneously to transmit the information to the second electron which changes its spin state to become the opposite spin state of the first electron. If the first electron has up spin, then the second electron will have down spin, and this happens instantaneously even if the two electrons are located at two opposite sides of our 3D Universe. Since the two electrons have equal and opposite spins (+1/2 and -1/2) the tachyon must carry a zero spin to be able to interact with both the electrons simultaneously, and hence it is classified as a Scalar Boson like the Higgs Boson. Only in Quantum Mechanics can the future change the past, and it is done to make Quantum Mechanics compatible with Classical Mechanics. Tachyons can change the indeterministic nature of Quantum Mechanics into the deterministic nature of Classical Mechanics.

Quantum theory is used to explain Schrodinger's cat in a box which can be in a superposition of being both alive and dead until one observes it at which time it will be found to be in a deterministic state of being either alive or dead. Along with the grandfather's paradox which implies going back in time to kill your own grandfather so that you cannot be born, Schrodinger's cat is a terribly wrong example widely

used in Physics because Quantum Mechanics and the superposition principle apply only to microscopic particles, but for the size of a human being or a cat we are in the Classical regime for which the rules of Quantum Mechanics cannot apply. The law of causality for Classical Mechanics will not allow large objects to move backwards in time. This notion of humans being able to travel back and forth in time has been popularized by science fiction movies such as Back to the Future, Spiderman: Across the Spider-Verse, Flash, etc. and by science fiction books (3). Since tachyons and anti-particles can move backwards in time this means that the future can change the past implying quantum mechanics does not obey causality, while classical mechanics obeys causality since you cannot go to the past to kill your own grandfather and then find yourself to be still alive magically in an alternate Universe as has been depicted in the movies. Movies are for entertainment while Physics is for reality.

II. CONCLUSION

Tachyons and the Higgs Boson should be included together in the Standard Model of Particle Physics since they are both the only known Scalar Bosons.

The speed of light no longer exists as a universal cosmic speed limit as in classical mechanics because besides tachyons, space too that is quantized can expand faster than light speed (4).

Tachyons only interact with entangled particles such as two entangled electrons, two entangled photons, or an entangled electron-positron pair. Their job is to end the entanglement by separating them to demystify these quantum elementary particles to the classical level. The basic difference between Quantum Mechanics and Classical Mechanics is the size of the object and the time for which it is being observed. While space itself may seem to be very large it can be broken up into small quantum bits to obey the laws of Quantum Mechanics. Hence space and time can become a space-time continuum only for much larger classical values when they blend to become Classical, but their true nature is Quantum Mechanical at the tiniest levels (4).

REFERENCES

1. Benford, G. (1980) Timescape, Simon and Schuler, USA. Winner of the 1980 Nebula Award.
2. Irani, A. (2021) Dark Energy, Dark Matter, and the Multiverse. Journal of High Energy Physics, Gravitation and Cosmology, 7, 172-190. https://doi.org/10.4236/jhepgc.2021.71009
3. Benford G. (2018) REWRITE Loops in the timescape, Saga Press, An imprint of Simon and Schuler, Inc. NY, NY 10020.
4. Irani, A. (2022) Faster than the speed of light is a Quantum Phenomenon. Journal of High Energy Physics, Gravitation and Cosmology, Vol.9, No.2, Apr. 2023.