FOUNDATIONS OF QUANTUM THEORY: NON-RELATIVISTIC APPROACH

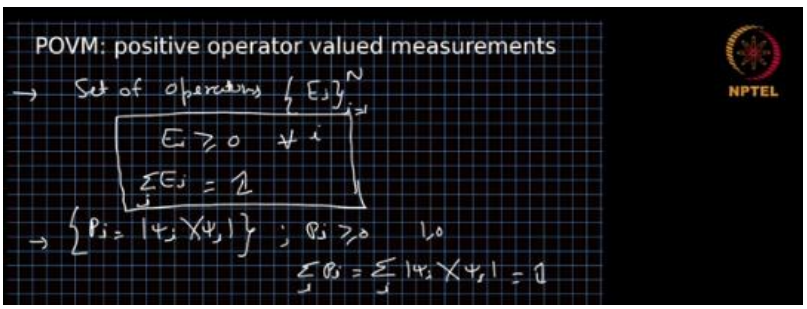
Dr. Sandeep K. Goyal
Department of Physical Sciences
IISER Mohali

Week-07
Lecture-19

Measurements: POVM

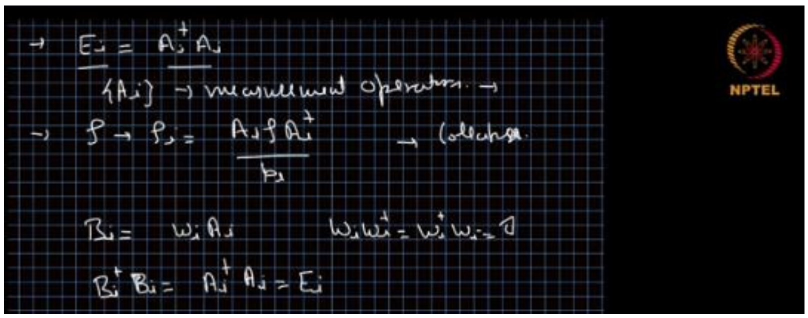
In this lecture, we will talk about the POVM that stands for Positive Operator Value Measurement. It's a type of generalized measurement and this contains every type of measurement allowed on a quantum system. So, this is the most general measurement we can have and any measurement can be written in terms of POVM. Mathematically, a POVM is a set of operators Ei where i is from 1 to n, so there are n number of operators such that Ei for every i is positive semi-definite for every i and Ei sum over i equals identity. So, these are the only two conditions this set of operators, the operators in this set needs to satisfy that every single one of them should be a positive semi definite operator and they all add up to identity.

(Refer slide time: 1:55)



So, if we take examples of projective measurements, so the projected P i, we have our Psi i. Psi i in a projective measurement where Psi i forms an orthonormal basis. And so, from here we can see that p i are all positive because their eigenvalues are one or zero and when we add p i that is sum over i, psi i outer product psi i, since i is an orthonormal and complete basis, this has to be identity. So, this becomes a valid measurement and this is one example of the POVM. So, a POVM is a set of operators which satisfy these two conditions. Now, this seems very general at the moment, but we will see how this can be used to perform measurement and how we extract information about the quantum system using these measurements. So here, of course, we need to talk about the Born rule of probability, how that generalizes for POVM and what is the state after the measurement. So, the probability pi of ith outcome is given by the expectation value of Ei and that is trace of rho times Ei. Since Ei is a positive operator we can always write it as Ai dagger Ai or some matrix Ai, we can appropriately choose Ai, then these set of Ai's become the measurement and the state after collapse goes to rho i and that is Ai rho Ai dagger over pi.

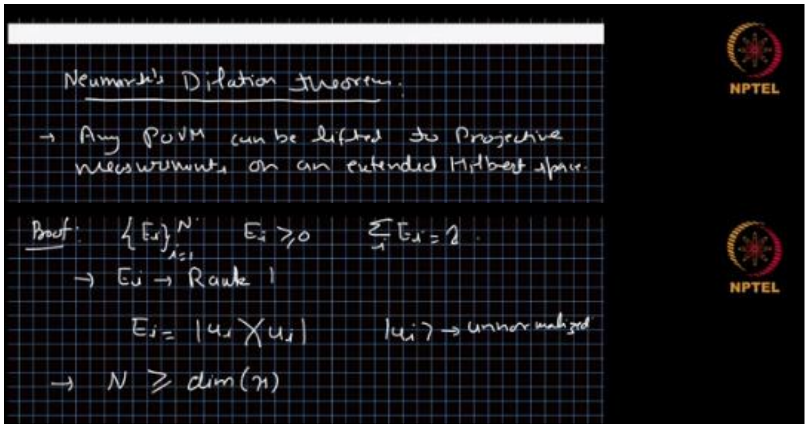
(Refer slide time: 4:16)



This is a normalization, so these are the two rules this is the generalized Born rule of probability and this is the collapse of the wave, of the state of the quantum system. Now, one thing to understand here is in POVM we are only given ei is not Ai and we can find another set Bi which is omega i times Ai where Wi Wi dagger equals Wi dagger Wi equals identity. We can have a unitary matrix or many unitary matrices Wi, even then we see that Bi dagger Bi is same as Ai dagger Ai which is Ei. So, for the same set of POVM operators we can have many measurement operators, so the choice of the measurement operator depends on the experiment we are trying to perform. But the outcome, which is in terms of probabilities, they are independent of the measurement operators. As long as we have the same POVM element, Ei, we are okay and we will get the probabilities, the correct probabilities.

Eis are called effect. Just the name that Eis are called effects. So POVMs have N effects and each effect will give you a probability of measurement. And from that measurement outcome, you can estimate the states and things like that. Now we will see, we will elaborate over this POVM, but before that, we will talk about a nice theorem. It's called Neumark's dilation theorem. So, what is Neumark's dilation theorem? It says that any POVM can be lifted to projective measurements on an extended Hilbert space.

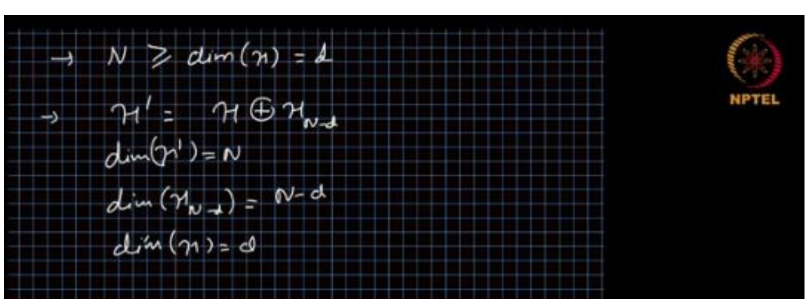
(Refer slide time: 8:03)



So, what this Neumark's dilation theorem says is that if we are given any POVM, it means we are given a set of observables Ei, a set of effects Ei, then we can always extend the Hilbert space, we can have a bigger Hilbert space, in which this POVM will look like a projective measurement. And since in the axioms of quantum mechanics, only projective measurements are defined and everything else should be derived from the projective measurement, in that way, this theorem becomes a very important theorem. Because here we are saying that ultimately everything is projective, but if we are trying to look at it from a restricted Hilbert space, then it will look like generalised measurements. So how do we prove this thing? So, we are given a set of effects, Eis. This is with our POVM, where Eis are positive, semi-definite, and sum over i, Ei is identity.

For the sake of simplicity, in this lecture, we will assume that Eis are rank one effects. So, what does that mean? That means Ei, we can write as Ui, outer product Ui. Where Ui is unnormalized. So, we are assuming that the effects are rank one. And they are positive operator. So, rank one positive operator means one eigenvalue will be non-zero and all the other eigenvalues will be zero.

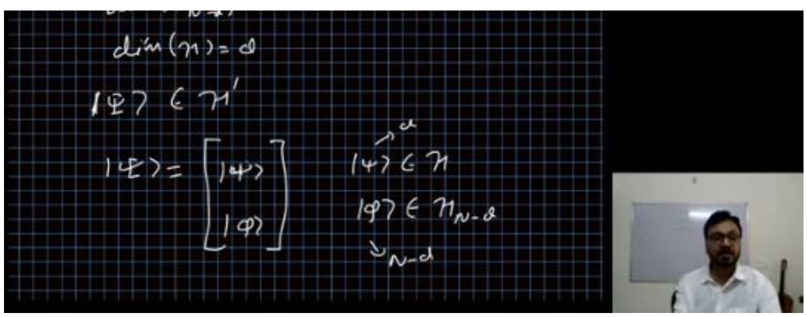
(Refer slide time: 9:15)



So, the corresponding eigenvector will be ui. We are choosing for every Ei and others we don't care. And the eigenvalue has been absorbed in the normalization of the vector ui. That's why ui is unnormalized. So, if we are assuming the Ei is R rank 1, then the number of the effects N is from 1 to N. So, there are N number of effects.

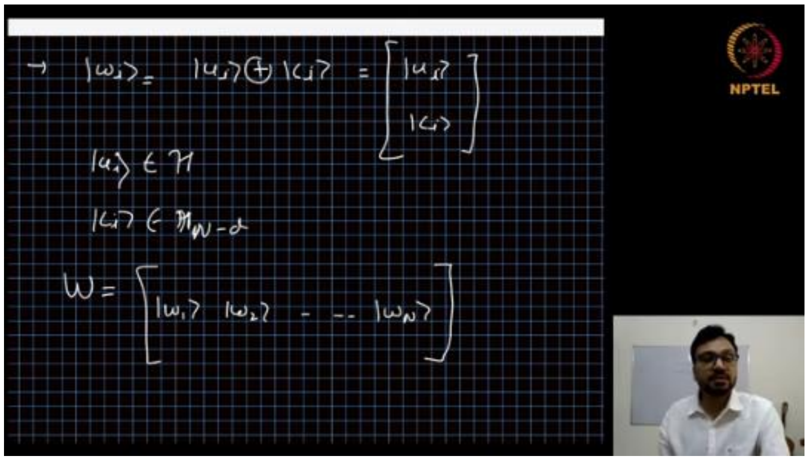
So, N must be greater than or equal to the dimension of the Hilbert space. Because, otherwise they will not add up to identity. The condition, this condition will not be satisfied if they have Eis to be rank one and they are less than the dimension of the Hilbert space and the equality only holds when all of uis are going to each other, only then the equality can hold, not will hold, but it can hold when all of them are also going to each other. Now we consider a Hilbert space H prime okay which is the Hilbert space of the system which we are calling H and then we add another Hilbert space H of n minus d so the dimension of the Hilbert space H was d and we add another Hilbert space of n minus d dimension. So the dimension of the Hilbert space H prime is n. And the dimension of the Hilbert space H n minus d is n minus d. And the dimension of H is d. So we can see that dimension of H plus dimension of n minus d. H n minus d total is n. So that's the dimension of H prime. So, what we are trying to say is if there was a vector capital Psi from the Hilbert space H prime, then this can be written as a vector of Psi, small psi and a small phi, where Psi belongs to H and phi belongs to H of n minus d.

(Refer slide time: 10:01)



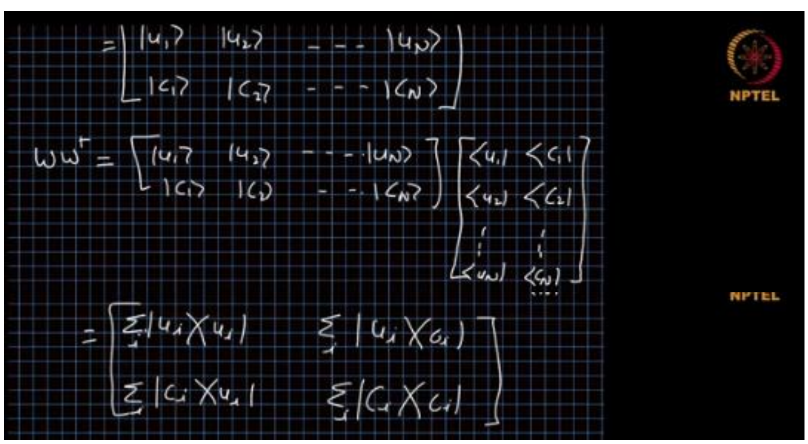
It means the dimension of psi is d and the dimension of phi is n minus d. This is what we mean by this symbol. This is what happens to the vector and we will see what happened to the operator also and this is called direct sum. So, what we have done is we have considered a Hilbert space H prime which is an extended Hilbert space in which a part of that is the original Hilbert space H which belongs to the quantum system and another part is the extra Hilbert space we have considered. Now we define operators or vectors wi which is ui direct sum ci. So, what we are saying is w i is defined in a way that which is u i and c i where u i belongs to H and c i belongs to H of n minus d. So, if we define a matrix W which is w1, w2, wn.

(Refer slide time: 11:06)



Then this is the environment. And if we write it further, it will become u1, u2, un, c1, c2, cN. What is WW dagger? It is u1 u2 uN, c1 c2 cN, u1 u2 uN c1 c2 cN and which will become we multiply this row with this column we get sum over i ui ui and we multiply this with this sum over i ui ci and we have ci ui and ci ci. For a moment, just consider this element.

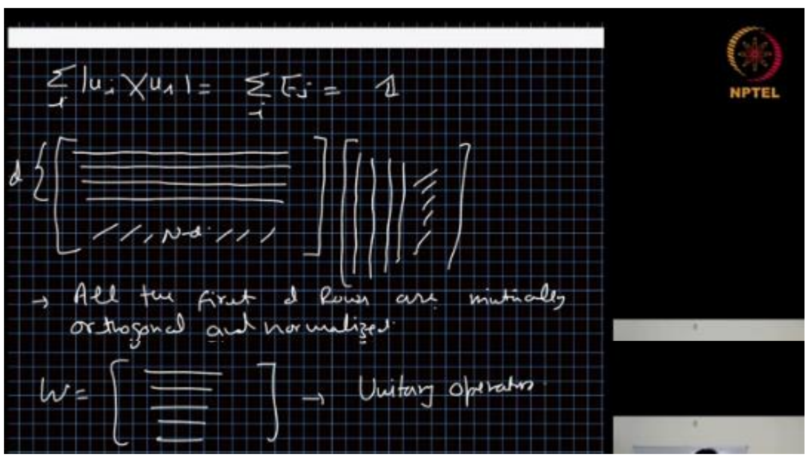
(Refer slide time: 12:30)



And this is nothing but sum over i Ei, which we know is identity. The condition that the sum over i ui ui can be seen from the matrix point of view. We have first d rows of the matrix W. Forget about this last n minus d rows. We are only interested in the first d rows. So, they are represented by the uis and then we multiply them with themselves.

And we get identity. It means all the rows are orthogonal to each other, orthogonal and normalized. All the first d rows are orthogonal and normalized or mutually orthogonal, let me write here mutually orthogonal and normalized. This is what it means that when we multiply the first row with first column then we get one and you multiply this first row with the second column we get zero and with the third we get zero four we get zero second row with second column we get one but everything else is zero that's how we get identity that this is this is what it shows that all the first d rows are usually orthogonal to each other they are not well the inner product is coming out to be one. So, we have a matrix W, where first d rows are orthogonal and rest n minus d rows, we have not decided what they are. They can be anything till now, c i vectors are unknown. So, the next d minus n minus d rows are unknown.

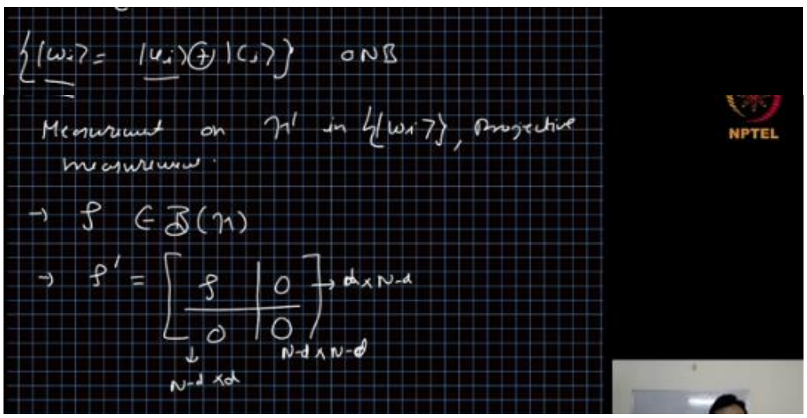
(Refer slide time: 15:37)



Now, if we have n dimensional vectors, n of n minus, n dimensional vectors and d of them are orthogonal. Then by using Gram Schmidt orthogonalization process or by something else we can make the rest of them also orthogonal. So, it's always possible to extend a set of d orthogonal vectors, d n dimensional orthogonal vectors to n n dimensional orthogonal vectors so we can always choose the n minus d rows which are mutually orthogonal and also orthogonal to the rest d. In that way, by choosing c appropriately, we can have W in which all the rows are orthogonal to each other and normalized. And that is the definition or one of the properties of a unitary operator. So if you have a matrix in which all the rows are normalized and orthogonal to each other, then it's a unitary operator.

,If you have a matrix which has all the columns normalized and orthogonal to each other, then it's a unitary operator. And they can be independently. So, by default, if all the rows are normalized and orthogonal to each other, then the columns will also come out to be normalized and orthogonal to each other. So, in that way, by choosing c's properly, we have defined wi, which are ui plus appropriately chosen ci such that the matrix W is a unitary matrix. Then the set wi represents an orthonormal basis because wis are the columns of the matrix w and w is a unitary.

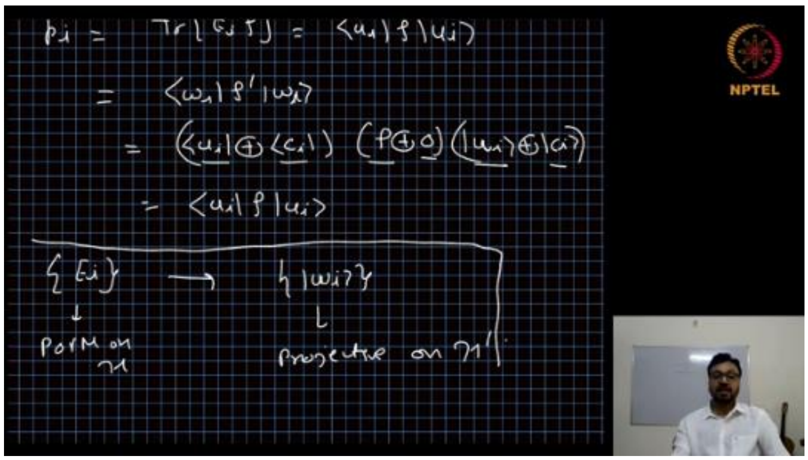
(Refer slide time: 18:00)



So, all the columns must be orthogonal to each other. In that way by extending the ui vector in a to a larger Hilbert space H prime and getting n dimensional vectors we have converted it into a part of an orthonormal basis. So now, if we perform measurement on the extended Hilbert space H prime in the basis omega i then it's the orthonormal basis. Then it is a projective measurement. Now, if we have a state rho, which is from the set of, which is from the Hilbert space H, then the, the density matrix rho prime in the extended Hilbert space will look like the following. It will be rho which is d by d, d dimensional, 0, 0 and a 0 which is n minus d by n minus d. So, this 0 is n minus d by n minus d. This is d by n minus d. This is n minus d by d. So, the whole matrix is n by n matrix.

So, we have just taken the row and padded it up with zeros to make it an n by n square matrix. Now probability pi was trace of Ei times rho, traditionally, like in the POVM definition which is for our POVM will be ui, rho ui because Ei is just a rank 1 projector, rank 1 POVM. Then this is same as wi rho prime wi because wi is ui direct sum ci. Rho prime is rho, direct sum zero and we have ui direct sum ci. Here the product was u will multiply with rho with u and c will multiply with 0 with p, so we get ui rho ui. So the projective measurement in the basis wi on an extended state rho prime is same as performing POVM on rho in the restricted Hilbert state. So in that way, we have gotten a mapping from the POVM in H, this is POVM on H to projective measurement on H prime.

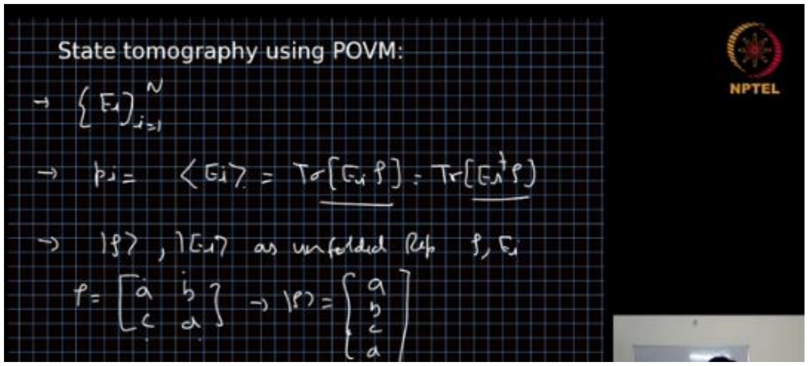
(Refer slide time: 19:4)



And this is the proof of the Neumark's dilation theorem. Any generalized measurement can be thought of as a projective measurement in an extended Hilbert space. Now, we move on to how we can use POVM to do the state tomography of a quantum system. So, again, let us consider we have a POVM Ei is from 1 to n. So, there are n elements, n effects in this POVM. Now, and the outcome of a measurement are the pis, which is the expectation value of Ei, which is trace of Ei rho.

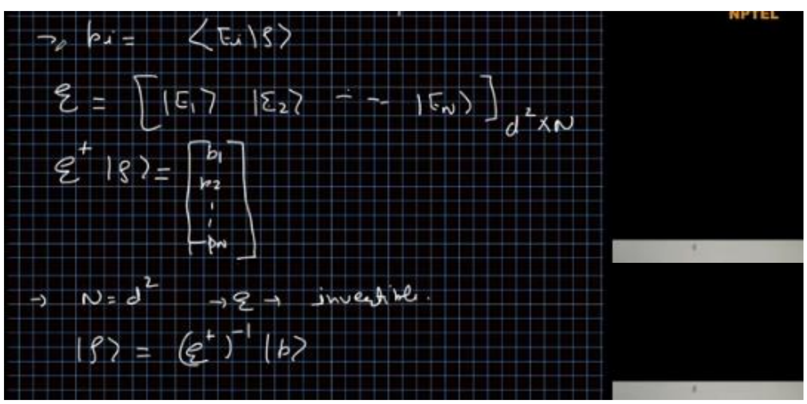
Since Ei is a Hermitian operator, it is a positive operator, so by default it is a Hermitian operator, so it can be written as Ei dagger rho. So, it is the same, Ei rho and Ei dagger rho are the same because Ei equals Ei rho. Now, if we have the, if we represent rho vector and Ei vector, the unfolded representation of the matrix rho and Ei. That is, we take the matrix rho, let us say rho is a 2 by 2 matrix with element a, b, c, d. Then the vector rho is a, b, c, d. Just be careful with the sequence, it is 1, 2, 3, 4, not the other way. So, this becomes the unfolded vector of the matrix rho, corresponding to the matrix rho.

(Refer slide time: 21:10)



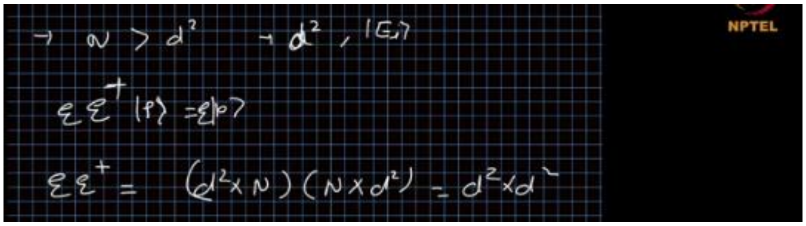
Similarly, we can have unfolded vector corresponding to Ei. Now, in that way, the probability Pi can be written as Ei vector and a rho inner product. Now, if we define a matrix xi, which is E1, E2, E3 and so on up to En, then we can write xi dagger acting on rho gives us a vector of probabilities p1, p2 up to pn. And let us see this xi matrix is the d square dimensional by n dimensional, d square by n dimensional matrix. Now probabilities are what we have as the outcome in the experiment and we want to do this is the only information given to us in an experiment and from here we want to see what is the density matrix and that is what we call the state tomography that from the experimental data we retrieve the information about the density matrix now take a simple case where n equals d square and xi is invertible.

(Refer slide time: 23:00)



If that is the case, then we can simply write rho equals xi dagger inverse acting on vector p. So, in that way we get the unfolded vector rho corresponding to the density matrix rho and from here we can calculate the we can reconstruct the density matrix rho. In case n is more than d square but still there are d square at least d square independent vectors ei, at least d square independent vector ei is, in that case we consider the equation again it was xi dagger rho equals p. We can multiply it with xi on both side now xi xi dagger will be n by d square and d square by xi xi dagger will be d square by n times n by d square matrix which is d square by d square matrix, just, let us see that in the beginning n was more so the xi was a rectangular matrix where the one dimension was bigger than the other. Now both the dimensions are same it's a square matrix and as we assume that there are at least d square independent Ei's in this xi matrix so xi xi dagger is invertible. And in that case we can write rho to be xi xi dagger inverse xi p and hence we can find the unfolded vector rho corresponding to the density matrix rho and from there we can reconstruct the density matrix rho.

(Refer slide time: 24:11)



And this is single short measurement unlike the measurement we did for qubits like photonic qubits where we had to perform first sigma x measurement then sigma y measurement then sigma z measurement and from there construct the density matrix back. Here if we can find appropriate POVM vectors Eis, from there in the single shot measurement, we will get enough data to construct the density matrix rho back. In that way, it can be very, very powerful. We will be considering more examples where the power of POVMs will be revealed over the projective measurements.

(Refer slide time: 24:50)